On the Joint Path Lengths Distribution in Random Binary Trees*

January 5, 2005

Charles Knessl
Dept. Mathematics, Statistics & Computer Science
University of Illinois at Chicago
Chicago, Illinois 60607-7045
U.S.A
knessl@uic.edu

Wojciech Szpankowski
Department of Computer Science
Purdue University
W. Lafayette, IN 47907
U.S.A.
spa@cs.purdue.edu

Abstract

During the *10th Seminar on Analysis of Algorithms*, MSRI, Berkeley, June 2004, Knuth posed the problem of analyzing the left and the right path length in a random binary trees. In particular, Knuth asked about properties of the generating function of the joint distribution of the left and the right path lengths. In this paper, we mostly focus on the asymptotic properties of the distribution of the difference between the left and the right path lengths. Among other things, we show that the Laplace transform of the appropriately normalized moment generating function of the path difference satisfies the *first Painlevé transcendent*. This is a nonlinear differential equation that has appeared in many modern applications, from nonlinear waves to random matrices. Surprisingly, we find out that the difference between path lengths is of the order $n^{5/4}$ where n is the number of nodes in the binary tree. This was also recently observed by Marckert and Janson. We present precise asymptotics of the distribution's tails and moments. We shall also discuss the joint distribution of the left and right path lengths. Throughout, we use methods of analytic algorithmics such as generating functions and complex asymptotics, as well as methods of applied mathematics such as the WKB method.

Key Words: Binary trees, Catalan number, path length, Painlevé transcendent, WKB method.

*The work was supported by NSF Grants CCR-0208709 and DMS-02-02815, NIH Grant R01 GM068959-01, and NSA Grant MDA 904-03-1-0036

1 Introduction

Trees are the most important nonlinear structures that arise in computer science. Applications are in abundance; here we discuss binary unlabeled ordered trees (further called binary trees) and study their asymptotic properties when the number of nodes, n, becomes large. While various interesting questions concerning statistics of randomly generated binary trees were investigated since Euler and Cayley [6, 12, 13, 18, 20, 21], recently novel applications have been surfacing. In 2003 Seroussi [16], when studying universal types for sequences and Lempel-Ziv'78 parsings, asked for the number of binary trees of *given* path length (sum of all paths from the root to all nodes). This was an open problem; partial solutions are reported in [11, 17].

During the *10th Seminar on Analysis of Algorithms*, MSRI, Berkeley, June 2004, Knuth asked to analyze the joint distribution of the left and the right path lengths in random binary trees. This problem received a lot of attention in the community (cf. related papers [8, 15]) and leads to an interesting analysis, that encompasses several other problems studied recently [8, 11, 14, 15, 16, 17, 20]. Here, we mostly focus on the asymptotic properties of the distribution of the difference between the left and the right path lengths. However, we also obtain some results for the joint distribution of the left and the right path lengths in a random binary tree.

In the standard model, that we adopt here, one selects uniformly a tree among all binary unlabeled ordered trees built on n nodes, $\mathcal{T}_n$ (where $|\mathcal{T}_n| = \binom{2n}{n}\frac{1}{n+1}$ =Catalan number). Many deep and interesting results concerning the behavior of binary trees in the standard model were uncovered. For example, Flajolet and Odlyzko [4] and Takacs [20] established the average and the limiting distribution for the height (longest path), while Louchard [14] and Takacs [19, 20, 21] derive the limiting distribution for the path length. As we indicate below, these limiting distributions are expressible in terms of the Airy's function (cf. [1, 2]). Recently, Seroussi [16, 17], and Knessl and Szpankowski [11] analyzed properties of random binary trees when selected uniformly from the set $\mathcal{T}_t$ of all binary trees of given path length t. Among other results, they enumerated the number of trees in $\mathcal{T}_t$ and analyze the number of nodes in a randomly selected tree from $\mathcal{T}_t$.

We now summarize our main results and put them into a bigger perspective. Let $N_n(p, q)$ be the number of binary trees built on n nodes with the right path length equal to p and the left path length equal to q. It is easy to see that its generating function $G_n(w, v)$ satisfies

$$G_n(w, v) = \sum_{i=0}^{n} w^i v^{n-i} G_i(w, v) G_{n-i}(w, v) \tag{1.1}$$

with $G_0(w, v) = 1$. Summing over n we obtain the triple transform $C(w, v, z)$ (cf. also (2.12) below) that satisfies

$$C(w, v, z) = 1 + zC(w, v, wz)C(w, v, vz). \tag{1.2}$$

This is exactly the equation that Knuth asked to analyze.

The above functional equation encompasses many properties of binary trees. Let us first set $w = v$ and define $C(w,z) = C(w,w,z)$. Recurrence (1.2) then becomes

$$C(w,z) = 1 + zC^2(w,zw). \tag{1.3}$$

Observe that this equation is asymmetric with respect to z and w. When enumerating trees in $\mathcal{T}_n$, we set $w = 1$ to get the well known algebraic equation $C(1,z) = 1 + zC^2(1,z)$ that can be explicitly solved as $C(1,z) = \left(1 - \sqrt{1-4z}\right)/(2z)$, leading to the Catalan number. A randomly (uniformly) selected tree from $\mathcal{T}_n$ has path length T_n that is asymptotically distributed as Airy's distribution [19, 20], that is,

$$\Pr\{T_n/\sqrt{2n^3} \le x\} \to W(x)$$

where $W(x)$ is the Airy distribution function defined by its moments [5]. The Airy distribution arises in surprisingly many contexts, such as parking allocations, hashing tables, trees, discrete random walks, area under a Brownian bridge, etc. [5, 14, 19, 20, 21].

Setting $z = 1$ in (1.3) we arrive at

$$C(w,1) = 1 + C^2(w,w)$$

which is not *algebraically* solvable. Observe that the coefficient of w^t in $C(w,1)$ enumerates binary trees with a *given* path length t. In [11, 17] it was shown that

$$[w^t]C(w,1) = |\mathcal{T}_t| = \frac{1}{(\log_2 t)\sqrt{\pi t}} 2^{\frac{2t}{\log_2 t}\left(1 + c_1 \log^{-2/3} t + c_2 \log^{-1} t + O(\log^{-4/3} t)\right)}$$

for large t, where c_1 and c_2 are explicitly computable constants.

Let us now set $v = 1$ in (1.2). Then

$$C(w,1,z) = 1 + zC(w,1,wz)C(w,1,z)$$

while $G_n(w,1) = [z^n]C(w,1,z)$ satisfies

$$G_{n+1}(w,1) = \sum_{i=1}^{n} w^i G_i(w,1)G_{n-i}(w,1). \tag{1.4}$$

Observe that $G_n(w,1)$ is the generating function of the right path length. Actually, recurrence (1.4) was studied by Takacs in [19] when analyzing the area under a Bernoulli random walk. Also, it appears in the Kleitman-Winston conjecture [9, 22].

Finally, we turn attention to results presenting in this paper. We first analyze the limiting distribution of the difference $\mathcal{D}_n = \mathcal{R}_n - \mathcal{L}_n$ where $\mathcal{R}_n, \mathcal{L}_n$ are the right and left path lengths, respectively. We observe that the difference $\mathcal{D}_n$ is of order $n^{5/4}$. This was also recently observed by Marckert [15] and Janson [8]. Among other things, we show that the tail of the distribution is thicker than that of the Gaussian distribution. More precisely,

$$n^{-5/4} \cdot \Pr\{\mathcal{D}_n = \beta n^{5/4}\} \sim \sqrt{\frac{5}{6}}(5\beta)^{1/3} c_0 \exp\left(-\frac{3}{4}5^{1/3}\beta^{4/3}\right)[1 + O(\beta^{-4/3})]$$

as $\beta \to \infty$, where c_0 is a constant.

Next, we analyze moments of $\mathcal{D}_n$. We first observe that odd moments vanish, while the *normalized* even moments satisfy (asymptotically) a certain *nonlinear* recurrence that occurs in various forms in many other problems, that are described by nonlinear functional equations similar to (1.1) (e.g., quicksort, linear hashing, enumeration of trees in $\mathcal{T}_t$). In these cases, usually the limiting distribution can be characterized only by moments. We conjecture that these problems constitute a new class of distributions determined by moments. More precisely, let Z be a (normalized) limiting distribution of such a process. Then for some $a_n \to \infty$ we have $\mathbf{E}[Z^k]/a_k = c_k$ such that in general c_k satisfies

$$c_{k+1} = \alpha_k + \beta_k c_k + \gamma_k \sum_{i=0}^{k} c_i c_{k-i} \tag{1.5}$$

with some initial conditions, and given α_k, β_k and γ_k. In our case (cf. also [8]) the even normalized moments of $\mathcal{D}_n$ converge as $\mathbf{E}[\mathcal{D}_n^{2m+2}]/n^{5(m+1)/2} \to (2m+2)!\sqrt{\pi}\Delta_m$ for any integer $m \geq 0$, where Δ_m satisfy the recurrence similar to (1.5) (cf. (2.23) and (2.24) below). Similar recurrences appear in the quicksort [10], linear hashing [5], path length in binary trees [12, 14, 20, 21], area under Bernoulli walk [19], enumeration of trees with given path length [11], and many others [6, 13, 18].

Finally, we analyze the moment generating function of $\mathcal{D}_n$. We observe that the Laplace transform of an appropriately normalized moment generating function satisfies the *first Painlevé transcendent* nonlinear differential equation [7]

$$0 = U_1^2(\phi) + 2U_1''(\phi) - 4\phi.$$

This also appears in many modern applications, including nonlinear waves and random matrices. We shall also discuss the joint distribution of the left and the right path lengths.

Throughout, we use methods of analytic algorithmics such as generating functions and complex asymptotics, as well as methods of applied mathematics, such as the WKB method. We add that moments of $\mathcal{D}_n$ were recently analyzed by Janson [8] using a Galton-Watson branching process approach, and the limiting distribution is implicit in Marckert [15], who applied Brownian analysis. As pointed out by Janson [8] "many different methods are useful and valuable, even for the same types of problems, and should be employed without prejudice".

2 Problem Statement and Summary of Results

Let $N(p, q; n)$ be the number of binary trees with n nodes that have a total right path length p and a total left path length q. We also set

$$N(p, q; n) = \bar{N}(p + q, p - q; n) \tag{2.1}$$

and note the obvious symmetry relation

$$N(p,q;n) = N(q,p;n). \tag{2.2}$$

We shall mostly focus on analyzing the difference between the right and left path length, and this we denote by

$$J = p - q. \tag{2.3}$$

It is well known [6] that the total number of trees with n nodes is the Catalan number

$$C_n = \frac{1}{n+1}\binom{2n}{n}. \tag{2.4}$$

Then we define the probability distribution of the path length difference, $\mathcal{D}_n$, by

$$P_-(J;n) = \text{Prob}[\mathcal{D}_n = J] = \frac{1}{C_n}\sum_{i=0}^{\binom{n}{2}-|J|} N(i, i+|J|; n), \qquad J \in \left[-\binom{n}{2}, \binom{n}{2}\right]. \tag{2.5}$$

Here we used the fact that the left or right path in a tree with n nodes can be at most $\binom{n}{2}$. We can easily verify that

$$P_-\left(\binom{n}{2} - 1; n\right) = 0, \tag{2.6}$$

$$P_-\left(\binom{n}{2}; n\right) = P_-\left(\binom{n}{2} - 2; n\right) = P_-\left(\binom{n}{2} - 3; n\right) = 1/C_n, \tag{2.7}$$

since there are no trees where the path length difference is one below the maximum, and exactly one tree (out of C_n) has this difference either zero or two or three below the maximum value of $\binom{n}{2}$. In view of (2.2) we have $P_-(J;n) = P_-(-J;n)$ so it is sufficient to analyze (2.5) for $J \geq 0$.

The generating function of $N(p,q)$ in (2.1)

$$G_n(w,v) = \sum_p \sum_q N(p,q;n) w^p v^q \tag{2.8}$$

satisfies the recurrence

$$G_{n+1}(w,v) = \sum_{i=0}^{n} w^i v^{n-i} G_i(w,v) G_{n-i}(w,v), \;\; n \geq 0, \tag{2.9}$$

subject to the initial condition

$$G_0(w,v) = 1. \tag{2.10}$$

From (2.9) we also obtain the functional equation

$$C(w,v,z) = 1 + zC(w,v,wz)C(w,v,vz) \tag{2.11}$$

for the triple transform

$$C(w,v,z) = \sum_{n=0}^{\infty} G_n(w,v) z^n = \sum_n \sum_p \sum_q N(p,q;n) z^n w^p v^q. \tag{2.12}$$

We note that $G_n(1,1) = C_n$ and from (2.5) we obtain

$$P_-(J;n) = \frac{1}{C_n}[w^J] G_n\left(w, \frac{1}{w}\right). \tag{2.13}$$

We study the limit $n \to \infty$, with an appropriate scaling of p and q. First we consider the path length difference, with J scaled as

$$J = \beta n^{5/4} = O(n^{5/4}). \tag{2.14}$$

For a fixed β we shall obtain

$$P_-(J;n) \sim n^{-5/4} p_-(\beta) \tag{2.15}$$

where $p_-(\beta)$ is a probability density that can be represented as

$$\begin{aligned} p_-(\beta) &= \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} e^{-\beta b}[1+\sqrt{\pi}\bar{H}(b)]db \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\beta i x}[1+\sqrt{\pi}\bar{H}(ix)]dx \\ &= \frac{1}{\pi}\int_0^{\infty} \cos(\beta x) S(x) dx, \end{aligned} \tag{2.16}$$

where $S(x) = 1 + \sqrt{\pi}\bar{H}(ix)$ and

$$1 + \sqrt{\pi}\bar{H}(b) = \int_{-\infty}^{\infty} e^{\beta b} p_-(\beta) d\beta. \tag{2.17}$$

Thus the left side of (2.17) is the moment generating function of $p_-(\beta)$, which is an *entire* function of b.

While we do not have an explicit formula for $p_-(\beta)$, we have the following asymptotic and numerical values:

$$p_-(\beta) \sim \sqrt{\frac{5}{6}}(5\beta)^{1/3} c_0 \exp\left(-\frac{3}{4}5^{1/3}\beta^{4/3}\right)[1 + O(\beta^{-4/3})] \tag{2.18}$$

as $\beta \to \infty$, with

$$c_0 \approx .5513288; \tag{2.19}$$

$$p_-(0) \approx .45727 \ ; \ p_-''(0) \approx -.71462. \tag{2.20}$$

We also find that the density has two inflection points, at $\beta = \pm\beta_c$, with

$$\beta_c \approx .75898. \tag{2.21}$$

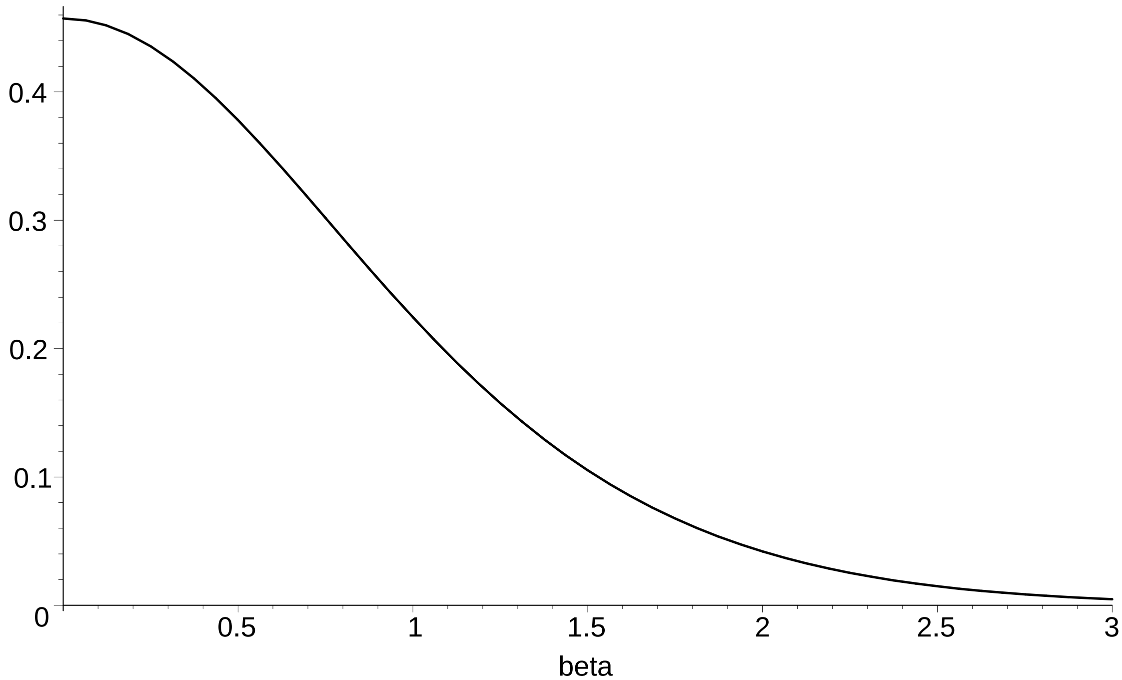

Figure 1: The density $p_-(\beta)$ for $\beta \in [0, 3]$.

This is our first main result.

The function $p_-(\beta)$ is graphed in Figure 1 over the range $\beta \in [0, 3]$, and the derivatives $p'_-(\beta)$ and $p''_-(\beta)$ are given over the same range in Figure 2. The graph of $p_-(\beta)$ somewhat resembles a Gaussian, but it differs from the Gaussian density in at least three important respects. First, the tail is clearly thicker in view of (2.18). For any Gaussian density, we would have $\beta_c p_-(0) = 1/\sqrt{2\pi} = .39894\ldots$ while for the present density (2.20) and (2.21) yield the value $\beta_c p_-(0) \approx .34705$. Finally, the Gaussian density would be an entire function of β, while we will show that if we view $p_-(\beta)$ as a function of the complex variable β, this function has an essential singularity at $\beta = 0$. While for $|\beta|$ small and β real, $p_-(\beta)$ is locally Gaussian, its behavior for $|\beta|$ small and β imaginary is quite different.

The function $\bar{H}$ in (2.16) is an entire function satisfying $\bar{H}(b) = \bar{H}(-b)$ and $\bar{H}(0) = 0$. Denoting its Taylor series as

$$\bar{H}(b) = \sum_{m=0}^{\infty} b^{2m+2} \Delta_m \tag{2.22}$$

and setting

$$\Delta_m = \frac{1}{\Gamma\left(\frac{5}{2}m + 2\right)} \tilde{\Delta}_m \tag{2.23}$$

we find that $\tilde{\Delta}_m$ satisfies the nonlinear recurrence

$$\tilde{\Delta}_{m+1} = \frac{(5m+6)(5m+4)}{8} \tilde{\Delta}_m + \frac{1}{4} \sum_{\ell=0}^{m} \tilde{\Delta}_\ell \tilde{\Delta}_{m-\ell}, \ m \geq 0 \tag{2.24}$$

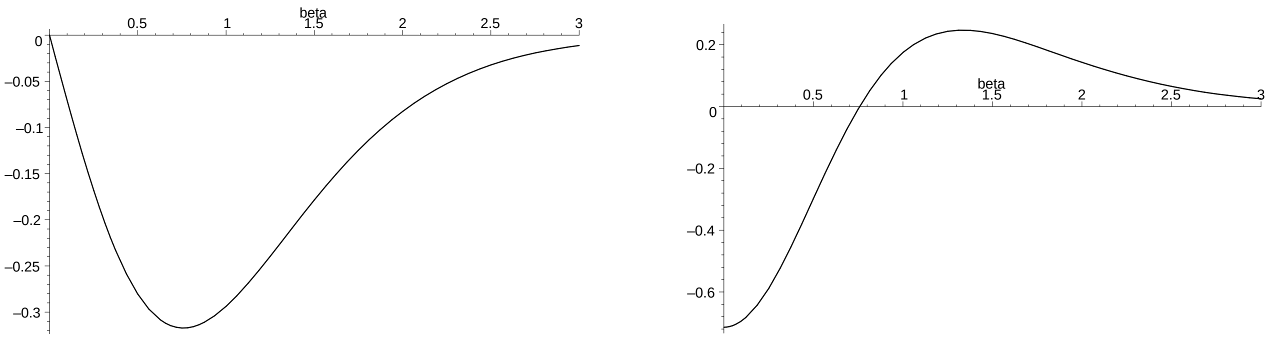

Figure 2: The first derivative $p'_-(\beta)$ and the second derivative $p''_-(\beta)$.

with

$$\Delta_0 = \tilde{\Delta}_0 = \frac{1}{4}. \tag{2.25}$$

In view of (2.22) and (2.17) the variance of the limiting density $p_-(\beta)$ in (2.15) is

$$\int_{-\infty}^{\infty} \beta^2 p_-(\beta) d\beta = 2\sqrt{\pi}\Delta_0 = \frac{\sqrt{\pi}}{2}. \tag{2.26}$$

We thus observe that the even moments of the difference converge as follows

$$\frac{\mathbf{E}[\mathcal{D}_n^{2m+2}]}{n^{5(m+1)/2}} \to (2m+2)!\sqrt{\pi}\Delta_m.$$

This should be compared with Janson [8].

Furthermore, setting

$$\bar{H}(b) = b^{6/5}\Delta(b^{4/5}) = B^{3/2}\Delta(B), \qquad b = B^{5/4} \tag{2.27}$$

we shall show that for b, $B > 0$ the function $\Delta(B)$ satisfies the nonlinear integral equation

$$0 = \int_0^B \Delta(\xi)\Delta(B-\xi)d\xi + 2B^2\Delta(B) + 2\frac{\sqrt{B}}{\sqrt{\pi}} - \frac{4}{\sqrt{\pi}}\int_0^B \frac{\Delta'(\xi)}{\sqrt{B-\xi}}d\xi. \tag{2.28}$$

We also have $\Delta(B) \sim B/4$ as $B \to 0^+$ and, in view of (2.22),

$$\Delta(B) = \sum_{m=0}^{\infty} B^{1+\frac{5}{2}m}\Delta_m. \tag{2.29}$$

The following asymptotic properties hold:

$$\Delta_m = k'\frac{e^{m/2}}{\sqrt{m}} m^{-m/2} 10^{-m/2}\left[1 - \frac{4}{15m} + O(m^{-2})\right], \; m \to \infty, \tag{2.30}$$

$$\Delta(B) = c_0 B e^{B^5/20}[1 - B^{-5} + O(B^{-10})], \; B \to \infty, \tag{2.31}$$

$$\bar{H}(b) = c_0 b^2 e^{b^4/20}[1 - b^{-4} + O(b^{-8})], \; b \to \infty. \tag{2.32}$$

Here k' and c_0 are related by

$$c_0 = 2\sqrt{\pi}k', \tag{2.33}$$

so that the numerical value of k' can be obtained from (2.19).

We can also infer the behavior of $\bar{H}(b)$ for purely imaginary values of b. Letting $b = ix$ with $x > 0$ and using (2.22) yields

$$1 + \sqrt{\pi}\bar{H}(ix) \equiv S(x) = \sum_{m=0}^{\infty} \Delta_{m-1}(-1)^m \sqrt{\pi} x^{2m} \tag{2.34}$$

where

$$\Delta_{-1} \equiv \frac{1}{\sqrt{\pi}}. \tag{2.35}$$

Then if

$$\bar{H}(ix) = -y^{3/2}\Lambda(y) = -x^{6/5}\Lambda(x^{4/5}), \qquad y = x^{4/5} \tag{2.36}$$

we find that $\Lambda(y)$ satisfies

$$0 = \int_0^y \Lambda(\xi)\Lambda(y-\xi)d\xi + 2y^2\Lambda(y) - 2\frac{\sqrt{y}}{\sqrt{\pi}} + \frac{4}{\sqrt{\pi}}\int_0^y \frac{\Delta'(\xi)}{\sqrt{y-\xi}}d\xi. \tag{2.37}$$

Note that this equation differs from (2.28) only slightly, by the signs of the last two terms. However, (2.37) can be analyzed by a Laplace transform whereas (2.28) cannot. Indeed, setting

$$U(\phi) = \int_0^\infty e^{-y\phi}\Lambda(y)dy \tag{2.38}$$

we obtain from (2.37)

$$0 = 2U''(\phi) + U^2(\phi) + 4\sqrt{\phi}U(\phi) - \phi^{-3/2}. \tag{2.39}$$

Also, since we know the behavior of $\Lambda(y)$ as $y \to 0^+$ we must have

$$U(\phi) \sim \frac{1}{4\phi^2}, \ \phi \to +\infty. \tag{2.40}$$

Setting

$$U(\phi) = -2\sqrt{\phi} + U_1(\phi) \tag{2.41}$$

we obtain from (2.39) and (2.40)

$$0 = U_1^2(\phi) + 2U_1''(\phi) - 4\phi, \tag{2.42}$$

with

$$U_1(\phi) = 2\sqrt{\phi} + \frac{1}{4}\phi^{-2}[1 + o(1)], \qquad \text{for} \quad \phi \to \infty.$$

The second order nonlinear ODE in (2.42) is (after a slight rescaling) the *first Painlerè transcendent* [7]. This classic problem has been studied for over 100 years, and modern applications in nonlinear waves and random matrices have been found in recent years. It is well known that each singularity of $U_1(\phi)$ is a double pole, and the Laurent expansion near a singularity at $\phi = -\nu_*$ has the form

$$U_1(\phi) = \frac{-12}{(\phi + \nu_*)^2} + O(1), \ \phi \to -\nu_*. \tag{2.43}$$

Let us denote by ν_* the singularity with the largest real part. Note that to uniquely fix this we need the second term in the expansion of $U_1(\phi)$ as $\phi \to \infty$, as given below (2.42). In view of (2.43), (2.38), and (2.41) we then obtain

$$\Lambda(y) - \frac{1}{\sqrt{\pi}} y^{-3/2} \sim -12 y e^{-\nu_* y}, \ y \to \infty \tag{2.44}$$

and (2.36) then yields

$$\sqrt{\pi}\bar{H}(ix) + 1 = S(x) \sim 12\sqrt{\pi} x^2 \exp(-\nu_* x^{4/5}), \ x \to \infty. \tag{2.45}$$

This yields the behavior of the moment generating function of the density $p_-(\beta)$ along the imaginary axis. The constant ν_* is found numerically as

$$\nu_* = 3.41167\ldots \ . \tag{2.46}$$

Finally, we discuss the joint distribution for the total left and right path lengths. This problem we formulate, but do not analyze. Introducing the scaling

$$p + q = \alpha n^{3/2}, \ p - q = \beta n^{5/4} \tag{2.47}$$

we obtain

$$\begin{aligned} P(p, q; n) &= \text{Prob [right path length} = p\text{, left path length} = q \mid \text{number of nodes} = n] \\ &= N(p, q; n)/C_n \\ &\sim 2n^{-11/4} \frac{1}{(2\pi i)^2} \int_{-i\infty}^{i\infty} \int_{-i\infty}^{i\infty} e^{-a\alpha} e^{-b\beta} [1 + \sqrt{\pi} H(a, b)] da db \\ &\equiv 2n^{-11/4} p(\alpha, \beta). \end{aligned} \tag{2.48}$$

Thus

$$1 + \sqrt{\pi} H(a, b) = \int_0^\infty e^{a\alpha} \int_{-\infty}^\infty e^{b\beta} p(\alpha, \beta) d\beta d\alpha \tag{2.49}$$

is the moment generating function of the two-dimensional density, which has support over the range $\alpha \geq 0$ and $\beta \in \mathbb{R}$. We have $H(0, 0) = 0$ and $p(\alpha, -\beta) = p(\alpha, \beta)$.

Setting $\alpha = 0$ with $\bar{H}(b) = H(0,b)$ we obtain the marginal distribution of the path length difference, with

$$p_-(\beta) = \int_0^\infty p(\alpha,\beta)d\alpha.$$

The marginal distribution of the total path length (without distinguishing between right and left paths) is given by

$$p_+(\alpha) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty}[1+\sqrt{\pi}H(a,0)]e^{-a\alpha}da$$

so that

$$1+\sqrt{\pi}H(a,0) = \int_0^\infty e^{a\alpha}p_+(\alpha)d\alpha.$$

This has been previously shown to follow an Airy distribution [11].

The function $H(a,b)$ satisfies the integral integration

$$\begin{aligned} 0 &= \int_0^1 \frac{H(x^{3/2}a, x^{5/4}b)H((1-x)^{3/2}a,(1-x)^{5/4}b)}{[x(1-x)]^{3/2}}dx \\ &+ \frac{2}{\sqrt{\pi}}\int_0^1 \left\{\frac{H((1-x)^{3/2}a,(1-x)^{5/4}b)}{(1-x)^{3/2}} - H(a,b)\right\}\frac{dx}{x^{3/2}} \\ &- \frac{4}{\sqrt{\pi}}H(a,b) + (4a+2b^2)\left[\frac{1}{\sqrt{\pi}} + H(a,b)\right]. \end{aligned} \tag{2.50}$$

In terms of the generating function in (2.8) the scaling (2.47) translates to

$$w = 1 + \frac{b}{n^{5/4}} + \frac{a}{n^{3/2}},\ \ v = 1 - \frac{b}{n^{5/4}} + \frac{a}{n^{3/2}} \tag{2.51}$$

and then, for fixed a and b and $n \to \infty$,

$$G_n(w,v) \sim \frac{4^n}{n^{3/2}\sqrt{\pi}}[1+\sqrt{\pi}H(a,b)]. \tag{2.52}$$

Here we used the asymptotic behavior of the Catalan numbers C_n. We have thus identified the scaling (2.47) and the problem (2.50) that must be analyzed to obtain the joint distribution of left and right paths in binary trees with large numbers of nodes n. While it does not seem feasible to solve (2.50) exactly, we believe that an asymptotic analysis for a and/or b large should be possible, and from this one can obtain asymptotic properties of the joint density $p(\alpha,\beta)$ for α and/or $|\beta|$ large.

3 The Basic Integral Equation

We shall derive (2.50) with the scaling in (2.51) and (2.52). In [11], we previously analyzed $G_n(w,w)$ for $n \to \infty$ and various ranges of w. This corresponds to computing the total path length in a tree with n nodes. The distribution of the path length is $P_+(I;n) = \sum_{p+q=I} N(p,q;n)/C_n$ and this is known to follow an Airy distribution in the limit $n \to \infty$, with the scaling $I = p + q = O(n^{3/2})$.

We begin by computing "moments" of $N(p,q)$, by expanding (2.9) about $(w,v) = (1,1)$. Since $G_n(1,1) = C_n$ and $G_n(w,v) = G_n(v,w)$, we write

$$\begin{aligned} G_n(w,v) &= C_n + A_n(w-1) + A_n(v-1) \\ &+ E_n(w-1)^2 + D_n(w-1)(v-1) + E_n(v-1)^2 + \cdots \end{aligned} \tag{3.1}$$

where

$$A_n = \frac{\partial}{\partial w} G_n(w,v)|_{(1,1)}, \quad E_n = \frac{1}{2}\frac{\partial^2}{\partial w^2} G_n(w,v)|_{(1,1)}, \quad D_n = \frac{\partial^2}{\partial w \partial v} G_n(w,v)|_{(1,1)}. \tag{3.2}$$

By iterating (2.9) for the first few values on n, we find that

$$G_1(w,v) = 1\ , \quad G_2(w,v) = w + v \tag{3.3}$$

$$G_3(w,v) = w^3 + w^2 v + w v^2 + v^3 + wv \tag{3.4}$$

$$G_4(w,v) = (w^3 + v^3)(w^3 + w^2 v + w v^2 + v^3 + wv) + (w^2 v + w v^2)(w + v). \tag{3.5}$$

In general, $G_n(w,v)$ is a polynomial in w and v of degree $\binom{n}{2}$.

By using (3.1) in (2.9), we are led to the recurrences

$$A_{n+1} = \sum_{i=0}^{n} i C_i C_{n-i} + 2\sum_{i=0}^{n} C_i A_{n-i}, \tag{3.6}$$

$$D_{n+1} = \sum_{i=0}^{n} i(n-i) C_i C_{n-i} + 2\sum_{i=0}^{n} D_i C_{n-i} + 2\sum_{i=0}^{n} A_i A_{n-i} + 2n \sum_{i=0}^{n} C_{n-i} A_i, \tag{3.7}$$

and

$$E_{n+1} = \sum_{i=0}^{n} \binom{i}{2} C_i C_{n-i} + 2\sum_{i=0}^{n} E_i C_{n-i} + \sum_{i=0}^{n} A_i A_{n-i} + n \sum_{i=0}^{n} C_{n-i} A_i. \tag{3.8}$$

By using

$$\sum_{i=0}^{n} \binom{i}{2} C_i C_{n-i} = \sum_{i=0}^{n} \frac{(n-i)^2 + i^2 - n}{4} C_i C_{n-i}$$

we can symmetrize (3.8). Then defining

$$F_n = D_n - 2E_n \tag{3.9}$$

we obtain from (3.7) and (3.8)

$$F_{n+1} = \sum_{i=0}^{n} \left[i(n-i) - \frac{(n-i)^2}{2} - \frac{i^2}{2} + \frac{n}{2} \right] C_i C_{n-i} + 2 \sum_{i=0}^{n} C_{n-i} F_i. \tag{3.10}$$

We can solve (3.6) – (3.10), e.g., by using generating functions and the fact that $\sum_{n=0}^{\infty} C_n z^n = \frac{1}{2z}(1 - \sqrt{1-4z})$. A lengthy but straightforward analysis yields

$$A_n = \frac{1}{2}4^n - \frac{3n+1}{2n+2}\binom{2n}{n} = \frac{1}{2}4^n - \frac{3n+1}{2}C_n, \tag{3.11}$$

$$F_n = -n4^{n-1} + \frac{n^2}{n+1}\binom{2n}{n} = -n4^{n-1} + n^2 C_n \tag{3.12}$$

and

$$D_n = \frac{1}{6}\frac{1}{n+1}\binom{2n}{n}(5n^3 + 30n^2 + 25n + 6) - (7n+4)4^{n-1}. \tag{3.13}$$

Then from (3.9), (3.12) and (3.13), we get

$$E_n = \frac{1}{12}C_n(5n^3 + 24n^2 + 25n + 6) - \frac{3n+2}{n}4^n. \tag{3.14}$$

Using Stirling's formula and the fact that

$$C_n = \frac{4^n}{n^{3/2}}\left[\frac{1}{\sqrt{\pi}} + O(n^{-1})\right], \ n \to \infty$$

we obtain the large n estimates

$$A_n = \frac{1}{2}4^n - \frac{3}{2}\frac{1}{\sqrt{\pi n}}4^n + O(4^n n^{-3/2}), \tag{3.15}$$

$$E_n = \left[\frac{5}{12}\frac{n^{3/2}}{\sqrt{\pi}} - \frac{3}{4}n + O(\sqrt{n})\right]4^n, \tag{3.16}$$

$$D_n = \left[\frac{5}{6}\frac{n^{3/2}}{\sqrt{\pi}} - \frac{7}{4}n + O(\sqrt{n})\right]4^n, \tag{3.17}$$

$$F_n = D_n - 2E_n \sim -\frac{1}{4}n4^n. \tag{3.18}$$

In view of (3.1) and (3.15) – (3.18), it would seem that a natural scaling would have p and q both $O(n^{3/2})$, say with

$$p = \alpha_1 n^{3/2}, \ q = \beta_1 n^{3/2}, \tag{3.19}$$

and then scale the generating function variables w and v as

$$w - 1 = a_1 n^{-3/2}, \;\; v - 1 = b_1 n^{-3/2}. \tag{3.20}$$

This is certainly correct if $w = v$, and then all the terms in (3.1), as well as the cubic and higher order terms, become of the same order ($O(4^n n^{-3/2})$) as $n \to \infty$. However, we show in Appendix A that with the scaling (3.19) and (3.20), we ultimately obtain an integral equation for the leading term in the expansion of $G_n(w, v)$, as follows

$$G_n(w, v) = C_n + \frac{4^n}{n^{3/2}} F(a_1, b_1; n) \sim \frac{4^n}{n^{3/2}} \left[\frac{1}{\sqrt{\pi}} + F(a_1, b_1) \right]. \tag{3.21}$$

The solution of the equation has the form $F(a_1, b_1) = F_0(a_1 + b_1)$, which depends on a_1 and b_1 only through their sum. This leads to a limiting joint density for the distribution of the left and right paths of the form $P(p, q; n) \sim n^{-3} p_0(\alpha_1, \beta_1)$, where $p_0(\alpha_1, \beta_1) = \delta(\alpha_1 - \beta_1) \times$ [function of $(\alpha_1 + \beta_1)$]. Here δ is the Dirac delta function, and in p_0 the latter function corresponds to the Airy density. But this suggests that the scaling (3.19) is too "thick" in the difference $p - q$, and that the correct scaling must have $p - q = o(n^{3/2})$.

To infer the right scaling needed to obtain a genuine two-dimensional density, we rewrite the quadratic terms in (3.1) as

$$E_n (w + v - 2)^2 + F_n (w - 1)(v - 1). \tag{3.22}$$

In view of (3.16) and (3.18) the two terms in (3.22) balance each other, and also the constant and linear terms in (3.1), if we scale w and v as

$$w + v - 2 = O(n^{-3/2}), \;\; (w - 1)(v - 1) = O(n^{-5/2}). \tag{3.23}$$

This is accomplished by setting

$$w = 1 + \frac{b}{n^{5/4}} + \frac{a}{n^{3/2}}, \;\; v = 1 - \frac{b}{n^{5/4}} + \frac{a}{n^{3/2}}. \tag{3.24}$$

Note that when $w = v$, $b = 0$ and (3.24) is consistent with the scaling $w - 1 = O(n^{-3/2})$ required for the limiting Airy distribution of the total pathlength. The cubic terms in (3.1) would take the form

$$\begin{aligned} &H_n[(w - 1)^3 + (v - 1)^3] + K_n (w - 1)(v - 1)(w + v - 2) \\ &= H_n (w + v - 2)^3 + (K_n - 3H_n)(w - 1)(v - 1)(w + v - 2). \end{aligned} \tag{3.25}$$

We can show that $H_n = O(n^3 4^n)$ as $n \to \infty$, and $K_n \sim 3H_n$ with $K_n - 3H_n = O(n^{5/2} 4^n)$. Hence the scaling (3.24) also leads to the cubic terms being $O(n^{-3/2} 4^n)$. It should be possible to show by induction that *all* terms in the Taylor expansion (3.1) balance. Instead we show that the scaling (3.24) leads to a non-trivial integral equation, whose solution leads to a two-dimensional density for the distribution of left and right paths in the tree, without the collapse that we see with (3.19) and (3.20).

We use (3.24) in (2.9) and get

$$G_n(w,v) = C_n + \frac{4^n}{n^{3/2}} H(a,b;n). \tag{3.26}$$

From (3.24) we have $a = \frac{1}{2}n^{3/2}(w+v-2)$ and $b = \frac{1}{2}n^{5/4}(w-v)$, so that

$$G_i(w,v) = C_i + \frac{4^i}{i^{3/2}} H\left(\left(\frac{i}{n}\right)^{3/2} a,\ \left(\frac{i}{n}\right)^{5/4} b; i\right), \tag{3.27}$$

with a similar expression for $G_{n-i}(w,v)$. Thus (2.9) becomes

$$C_{n+1} + \frac{4^{n+1}}{(n+1)^{3/2}} H\left(\left(1+\frac{1}{n}\right)^{3/2} a, \left(1+\frac{1}{n}\right)^{5/4} b; n+1\right) = \tilde{T}_1 + \tilde{T}_2 + \tilde{T}_3 + \tilde{T}_4 \tag{3.28}$$

where

$$\tilde{T}_1 = \sum_{i=0}^{n} \left(1 + \frac{b}{n^{5/4}} + \frac{a}{n^{3/2}}\right)^i \left(1 - \frac{b}{n^{5/4}} + \frac{a}{n^{3/2}}\right)^{n-i} C_i C_{n-i} \tag{3.29}$$

$$\begin{aligned} \tilde{T}_2 = \sum_{i=0}^{n} \left(1 + \frac{b}{n^{5/4}} + \frac{a}{n^{3/2}}\right)^i \left(1 - \frac{b}{n^{5/4}} + \frac{a}{n^{3/2}}\right)^{n-i} C_i \frac{4^{n-i}}{(n-i)^{3/2}} \\ \times H\left(\left(1-\frac{i}{n}\right)^{3/2} a, \left(1-\frac{i}{n}\right)^{5/4} b; n-i\right) \end{aligned} \tag{3.30}$$

and

$$\begin{aligned} \tilde{T}_3 = \sum_{i=0}^{n} \left(1 + \frac{b}{n^{5/4}} + \frac{a}{n^{3/2}}\right)^i \left(1 - \frac{b}{n^{5/4}} + \frac{a}{n^{3/2}}\right)^{n-i} C_{n-i} \frac{4^i}{i^{3/2}} \\ \times H\left(\left(\frac{i}{n}\right)^{3/2} a, \left(\frac{i}{n}\right)^{5/4} b; n-i\right). \end{aligned} \tag{3.31}$$

$$\begin{aligned} \tilde{T}_4 &= \sum_{i=0}^{n} \left(1 + \frac{b}{n^{5/4}} + \frac{a}{n^{3/2}}\right)^i \left(1 - \frac{b}{n^{5/4}} + \frac{a}{n^{3/2}}\right)^{n-i} \\ &\times \frac{4^n}{i^{3/2}(n-i)^{3/2}} H\left(\left(\frac{i}{n}\right)^{3/2} a, \left(\frac{i}{n}\right)^{5/4} b; i\right) H\left(\left(1-\frac{i}{n}\right)^{3/2} a, \left(1-\frac{i}{n}\right)^{5/4} b; n-i\right). \end{aligned} \tag{3.32}$$

We estimate individually the terms $\tilde{T}_j$ for $j = 1, 2, 3, 4$. We assume that $H(a,b;n)$ tends to the limit $H(a,b)$ as $n \to \infty$ so that (3.26) becomes

$$G_n(w,v) \sim \frac{4^n}{n^{3/2}}\left[\frac{1}{\sqrt{\pi}} + H(a,b)\right],\ n \to \infty. \tag{3.33}$$

Note also that $G_n(1,1) = C_n$ implies that $H(0,0) = 0$.

First, by using the Euler MacLaurin formula we approximate the sum in (3.32) by an integral and obtain

$$\tilde{T}_4 \sim \frac{4^n}{n^2}\int_0^1 \frac{H(x^{3/2}a, x^{5/4}b)H((1-x)^{3/2}a,(1-x)^{5/4}b)}{[x(1-x)]^{3/2}}dx. \tag{3.34}$$

The fact that $H(a,b)$ is analytic in (a,b) with $H(0,0) = 0$ insures that the integral in (3.34) is finite, in spite of the singularities at $x = 0$ and $x = 1$.

Next we estimate the difference between $\tilde{T}_1$ and C_{n+1}, using the fact that the Catalan numbers satisfy the recurrence relation $C_{n+1} = \sum_{i=0}^{n} C_i C_{n-i}$. We thus have

$$\begin{aligned}\tilde{T}_1 - C_{n+1} &= \sum_{i=0}^{n}\left[\left(1+\frac{b}{n^{5/4}}+\frac{a}{n^{3/2}}\right)^i\left(1-\frac{b}{n^{5/4}}+\frac{a}{n^{3/2}}\right)^{n-i} - 1\right]C_iC_{n-i} \qquad (3.35)\\ &= \sum_{i=0}^{n}\left\{(2i-n)\frac{b}{n^{5/4}} + \frac{na}{n^{3/2}} + \frac{b^2}{n^{5/2}}\left[\binom{i}{2}+\binom{n-i}{2} - i(n-i)\right] + \dots\right\}C_iC_{n-i}.\end{aligned}$$

Now we use the asymptotic relations

$$\begin{aligned}\sum_{i=0}^{n} i(n-i)C_iC_{n-i} &\sim \frac{4^n}{\sqrt{\pi}}\int_0^1 \frac{dx}{\sqrt{x}\sqrt{1-x}} = O(4^n)\\ \sum_{i=0}^{n}\binom{i}{2}C_iC_{n-i} &\sim \frac{n^2}{2}\sum_{i=0}^{n} C_i \frac{4^n}{\sqrt{\pi}n^{3/2}} \sim \frac{4^n}{\sqrt{\pi}}\sqrt{n}\\ \sum_{i=0}^{n} C_iC_{n-i} = C_{n+1} &\sim \frac{4^{n+1}}{\sqrt{\pi}}n^{-3/2}.\end{aligned}$$

Using the above in (3.35) yields

$$\tilde{T}_1 - C_{n+1} = \frac{4^n}{\sqrt{\pi}}\frac{4a+2b^2}{n^2} + O(4^n n^{-5/2}). \tag{3.36}$$

We next write $\tilde{T}_2$ in (3.30) as

$$\begin{aligned}
\tilde{T}_2 &= \sum_{i=0}^{n} C_i 4^{-i}\left(1+\frac{b}{n^{5/4}}+\frac{a}{n^{3/2}}\right)^i\left(1-\frac{b}{n^{5/4}}+\frac{a}{n^{3/2}}\right)^{n-i} \\
&\times \left[\frac{4^n}{(n-i)^{3/2}}H\left(\left(1-\frac{i}{n}\right)^{3/2}a,\left(1-\frac{i}{n}\right)^{5/4}b;n-i\right)-\frac{4^n}{n^{3/2}}H(a,b;n)\right] \\
&+ \frac{4^n}{n^{3/2}}H(a,b;n)\sum_{i=0}^{n}C_i4^{-i}\left[\left(1+\frac{b}{n^{5/4}}+\frac{a}{n^{3/2}}\right)^i\left(1-\frac{b}{n^{5/4}}+\frac{a}{n^{3/2}}\right)^{n-i}-1\right] \\
&+ \frac{4^n}{n^{3/2}}H(a,b;n)\sum_{i=0}^{n}C_i4^{-i}.
\end{aligned} \tag{3.37}$$

Using the generating function for the C_n we can easily show that

$$\sum_{i=0}^{n} C_i 4^{-i} = 2-\frac{2}{\sqrt{\pi n}}+O(n^{-3/2}),\ \ n\to\infty. \tag{3.38}$$

Expanding the next to last term in (3.37) as in (3.35) and using

$$\sum_{i=0}^{n}\frac{2i-n}{n^{5/4}}bC_i4^{-i} = -\frac{2b}{n^{1/4}}+O(n^{-3/4}),$$

$$\sum_{i=0}^{n}\binom{i}{2}C_i4^{-i}n^{-5/2} \sim \sum_{i=0}^{n}\frac{\sqrt{i}}{2\sqrt{\pi}}n^{-5/2} = O(n^{-1}),$$

$$\begin{aligned}
\sum_{i=0}^{n}\binom{n-i}{2}C_i4^{-i}n^{-5/2} &= \sum_{i=0}^{n}\frac{(n-i)(n-i-1)}{2}C_i4^{-i}n^{-5/2} \\
&= \frac{1}{2\sqrt{n}}\sum_{i=0}^{\infty}C_i4^{-i}+O\left(\frac{1}{n}\right)\sim\frac{1}{\sqrt{n}}, \\
\sum_{i=0}^{n} i(n-i)C_i4^{-i}n^{-5/2} &= O(n^{-1})
\end{aligned}$$

we obtain

$$\begin{aligned}
\sum_{i=0}^{n}C_i4^{-i}&\left[\left(1+\frac{b}{n^{5/4}}+\frac{a}{n^{3/2}}\right)^i\left(1-\frac{b}{n^{5/4}}+\frac{a}{n^{3/2}}\right)^{n-i}-1\right] \\
&= -\frac{2b}{n^{1/4}}+\frac{2a+b^2}{\sqrt{n}}+O(n^{-3/4}).
\end{aligned} \tag{3.39}$$

The first sum in the right side of (3.37) we approximate by the Euler-MacLaurin formula and use $H(a,b;n)\to H(a,b)$ as $n\to\infty$; thus we obtain

$$\frac{4^n}{n^2\sqrt{\pi}}\int_0^1 \frac{1}{x^{3/2}}\left[\frac{H(1-x)^{3/2}a,(1-x)^{5/4}b)}{(1-x)^{3/2}} - H(a,b)\right]dx. \tag{3.40}$$

Here we also used $C_i 4^{-i} \sim i^{-3/2}/\sqrt{\pi}$ for $i \to \infty$. By combining (3.38) – (3.40), we see that (3.37) becomes

$$\begin{aligned}
\tilde{T}_2 &= \frac{4^n}{n^{3/2}}H(a,b;n)\left[2 - \frac{2}{\sqrt{\pi n}} + O(n^{-3/2})\right] \\
&+ \frac{4^n}{n^{3/2}}H(a,b;n)\left[-\frac{2b}{n^{1/4}} + \frac{2a+b^2}{\sqrt{n}} + O(n^{-3/4})\right] \\
&+ \frac{4^n}{n^2}\frac{1}{\sqrt{\pi}}\int_0^1 \frac{1}{x^{3/2}}\left[\frac{H((1-x)^{3/2}a,(1-x)^{5/4}b)}{(1-x)^{3/2}} - H(a,b)\right]dx.
\end{aligned} \tag{3.41}$$

Finally, we consider $\tilde{T}_3$ in (3.31). By changing the index on the sum from i to $n-i$ and using the symmetry $H(a,b;n) = H(a,-b;n)$ we see that $\tilde{T}_3$ is the same as $\tilde{T}_2$ with b replaced by $-b$. Then adding $\tilde{T}_3$ to $\tilde{T}_2$ it follows that the terms of order $O(4^n n^{-7/4})$ cancel and hence

$$\begin{aligned}
\tilde{T}_2 + \tilde{T}_3 &= \frac{4^{n+1}}{n^{3/2}}H(a,b;n)\left[1 - \frac{1}{\sqrt{\pi n}} + O(n^{-3/2})\right] + \frac{4^n}{n^2}\left\{(4a+2b^2)H(a,b;n)\right. \\
&+ \left.\frac{2}{\sqrt{\pi}}\int_0^1 \frac{1}{x^{3/2}}\left[\frac{H((1-x)^{3/2}a,(1-x)^{5/4}b)}{(1-x)^{3/2}} - H(a,b)\right]dx\right\} + O(4^n n^{-5/2}).
\end{aligned} \tag{3.42}$$

We use (3.42), (3.36) and (3.34) in (3.28) and note that

$$\frac{4^{n+1}}{(n+1)^{3/2}}H\left(\left(1+\frac{1}{n}\right)^{3/2}a, \left(1+\frac{1}{n}\right)^{5/4}b; n+1\right) - \frac{4^{n+1}}{n^{3/2}}H(a,b;n) = O(4^n n^{-5/2}). \tag{3.43}$$

Thus at order $O(4^n n^{-2})$ we obtain from (2.9) and (3.33) the integral equation in (2.50). Using the above procedure we can refine (3.33) to an expansion of the form

$$G_n(w,v) = \frac{4^n}{n^{3/2}}\left[\frac{1}{\sqrt{\pi}} + H(a,b) + \frac{1}{\sqrt{n}}H^{(1)}(a,b) + O(n^{-1})\right] \tag{3.44}$$

where $H^{(1)}$ can be characterized as the solution of a *linear* integral equation, whose kernel involves the solution $H(a,b)$ to (2.50).

We can recast (2.50) as follows. We assume first that $a < 0$ and $b > 0$, and set

$$H(a,b) = (-a)T((-a)^{2/3}, (-a)^{2/3}b^{-4/5}) = (-a)T(A,\theta) \tag{3.45}$$

where

$$A = (-a)^{2/3}, \quad \theta = (-a)^{2/3}b^{-4/5}. \tag{3.46}$$

Using (3.45) and (3.46) in (2.50), dividing by a^2/A, and integrating by parts, with

$$\int_0^1 [T(A-Ax,\theta)-T(A,\theta)]\frac{1}{x^{3/2}}dx = 2T(A,\theta)-2T(0,\theta)+\int_0^1 \frac{2}{\sqrt{x}}\frac{d}{dx}T(A-Ax,\theta)dx,$$

we see that (2.50) becomes

$$\begin{aligned} 0 &= \int_0^A T(\xi,\theta)T(A-\xi,\theta)d\xi + \frac{2A^{1/2}}{\sqrt{\pi}}\theta^{-5/2} \\ &- \frac{4}{\sqrt{\pi}}\int_0^A \frac{1}{\sqrt{A-\xi}}\frac{d}{d\xi}T(\xi,\theta)d\xi + (2A^2\theta^{-5/2}-4A)T(A,\theta). \end{aligned} \tag{3.47}$$

Here we also used $T(0,\theta)=-1$, since (3.15) shows that $G_n(w,v)-C_n \sim A_n(w+v-2) = 2an^{-3/2}A_n \sim a4^n n^{-3/2}$, for $a\to 0$ and $n\to\infty$, and hence $H(a,b)\sim a$ as $a\to 0$. From (3.45) it then follows that $T(A,\theta)\to -1$ as $A\to 0$.

The integral operator in (3.47) involves only the A variable, and θ appears only as a parameter. Introducing the Laplace transform

$$\hat{T}(S,\theta) = \int_0^\infty T(A,\theta)e^{-AS}dA \tag{3.48}$$

in (3.47) leads to

$$0 = \hat{T}^2 - 4\sqrt{S}\hat{T} + 4\hat{T}_S + 2\theta^{-5/2}\hat{T}_{SS} + \frac{1}{S^{3/2}}\theta^{-5/2} - \frac{4}{\sqrt{S}}. \tag{3.49}$$

Furthermore, letting

$$\hat{T}(S,\theta) = 2\sqrt{S} + T_1(S,\theta) \tag{3.50}$$

we obtain from (3.49)

$$0 = T_1^2 + 4T_{1,S} + 2\theta^{-5/2}T_{1,SS} - 4S. \tag{3.51}$$

We will show in Appendix B that an asymptotic analysis of the functional equation in (2.11), with the scaling (3.24), leads to a limiting nonlinear ODE that is equivalent to (3.51).

In the sections that follow, we shall analyze (2.50) when $a=0$. We conclude this section by expressing the distribution in (2.48), in the limit $n\to\infty$, in terms of the solution $H(a,b)$ to (2.50). We have, inverting (2.8),

$$P(p,q;n) = \frac{1}{C_n}\frac{1}{(2\pi i)^2}\oint\oint w^{-p-1}v^{-q-1}G_n(w,v)dwdv \tag{3.52}$$

where the integrals are closed loops about the origins in the w- and v- planes. In view of the scaling (3.24) and (2.47), we have

$$
\begin{aligned}
w^{-p-1}v^{-q-1} &= (wv)^{-1}(wv)^{-\frac{p+q}{2}}\left(\frac{w}{v}\right)^{-\frac{p-q}{2}} \qquad (3.53)\\
&= \left[1+\frac{2a}{n^{3/2}}+O(n^{-2})\right]^{-\alpha n^{3/2}/2}\left[1+\frac{2b}{n^{5/4}}+O(n^{-3/2})\right]^{-\beta n^{5/4}/2} \sim e^{-a\alpha}e^{-b\beta}.
\end{aligned}
$$

Using (3.53), (3.33), and $C_n \sim 4^n n^{-3/2}/\sqrt{\pi}$ in (3.52), and approximating the loop integrals by ones over Bromwich contours, we obtain precisely (2.48). This shows that $n^{11/4}P(p,q;n)$ tends to a limit when p and q are scaled as in (2.47). We note that the factor of 2 in (2.48) arises from the Jacobian $dwdv = |\partial(w,v)/\partial(a,b)|dadb$.

4 Analysis of the Difference Between Right and Left Paths

We get $a = 0$ in (2.50) with

$$\bar{H}(b) = H(0,b), \qquad (4.1)$$

which yields

$$
\begin{aligned}
0 &= \int_0^1 \frac{\bar{H}(x^{5/4}b)\bar{H}((1-x)^{5/4}b)}{[x(1-x)]^{3/2}}dx + \frac{2b^2}{\sqrt{\pi}} \qquad (4.2)\\
&+ \frac{2}{\sqrt{\pi}}\int_0^1 \left[\frac{\bar{H}((1-x)^{5/4}b)}{(1-x)^{3/2}} - \bar{H}(b)\right]\frac{dx}{x^{3/2}} + \left(2b^2 - \frac{4}{\sqrt{\pi}}\right)\bar{H}(b).
\end{aligned}
$$

We first take b real and, since $\bar{H}(b) = \bar{H}(-b)$, we may take $b > 0$. Introducing $B = b^{4/5}$ and $\Delta(B) = \bar{H}(b)b^{-6/5}$ as in (2.27) we see that (4.2) becomes

$$
\begin{aligned}
0 &= \int_0^1 b^{12/5}\Delta(Bx)\Delta(B-Bx)dx + \frac{2}{\sqrt{\pi}}B^{5/2} \qquad (4.3)\\
&+ \frac{2}{\sqrt{\pi}}b^{6/5}\int_0^1 [\Delta(B-Bx) - \Delta(B)]\frac{dx}{x^{3/2}} + \left(2b^2 - \frac{4}{\sqrt{\pi}}\right)b^{6/5}\Delta(B).
\end{aligned}
$$

After dividing by $b^{8/5} = B^2$ and integrating by parts in the second integral, (4.3) becomes (2.28). By examining (2.28) as $B \to 0$ we see that $\Delta(B) \sim \Delta_0 B$ and the first two terms in the right side of (2.28) are $O(B^3)$, while the last two terms balance to give $\Delta_0 = 1/4$. Thus

$$\bar{H}(b) \sim \frac{1}{4}b^2, \ b \to 0 \qquad (4.4)$$

and this gives the variance of the limiting density $p_-(\beta)$ in (2.26), as the variance is $\sqrt{\pi}\bar{H}''(0)$.

While it seems difficult to solve (2.28) exactly, we can infer the behavior of $\Delta(B)$ as $B \to \infty$. Let us write

$$\Delta(B) = \exp[\Phi(B)] \tag{4.5}$$

and rewrite (2.28) as

$$0 = 2\int_0^{B/2} e^{\Phi(B-\xi)}\Delta(\xi)d\xi + 2B^2 e^{\Phi(B)} + O(\sqrt{B}) - \frac{4}{\sqrt{\pi}}\int_0^B \frac{1}{\sqrt{\xi}}\Phi'(B-\xi)e^{\Phi(B-\xi)}d\xi. \tag{4.6}$$

Now we assume that $\Phi \gg \Phi' \gg \Phi''$ as $B \to \infty$ and we view (4.5) as a WKB type expansion. Since $\Delta(B)$ will grow exponentially, the $O(\sqrt{B})$ term in (4.6) will not affect the asymptotic series. The first integral in (4.6) we treat as an implicit Laplace integral, where the factor $e^{\Phi(B-\xi)}$ will concentrate the integrand for ξ small (more precisely for $\xi = O(1/\Phi'(B))$). We thus write

$$\begin{aligned} 2\int_0^{B/2} e^{\Phi(B-\xi)}\Delta(\xi)d\xi &\sim 2\int_0^{\infty} e^{\Phi(B)}e^{-\xi\Phi'(B)}[1+O(\xi^2\Phi''(B))]\Delta'(0)\xi \;\; d\xi \\ &= \frac{1}{2}e^{\Phi(B)}\left(\frac{1}{\Phi'(B)}\right)^2\left[1+O\left(\frac{\Phi''(B)}{(\Phi'(B))^2}\right)\right]. \end{aligned} \tag{4.7}$$

We shall see that the leading term in (4.7) suffices to obtain the first three terms in the expansion of $\Phi(B)$ and the first two terms in that of $\Delta(B)$.

The second integral in (4.6) we expand as

$$\begin{aligned} \int_0^B \frac{1}{\sqrt{\xi}}\Phi'(B-\xi)e^{\Phi(B-\xi)}d\xi &= \int_0^{\infty}\frac{1}{\sqrt{\xi}}\left[\Phi' - \xi\Phi'' + \frac{\xi^2}{2}\Phi''' + O(\xi^3\Phi^{(iv)})\right]e^{\Phi}e^{-\xi\Phi'} \\ &\times \;\; e^{\frac{1}{2}\xi^2\Phi''}e^{-\frac{1}{6}\xi^3\Phi'''}\left[1+O(\xi^4\Phi^{(iv)})\right]d\xi, \end{aligned} \tag{4.8}$$

where Φ and all its derivatives are understood to be evaluated at B, in the right side of (4.8). Again the integral becomes concentrated where ξ is small. In (4.8) we introduce the scaling

$$\xi = \frac{\zeta}{\Phi'(B)} \tag{4.9}$$

to obtain

$$\begin{aligned} & e^{\Phi}\int_0^{\infty}\frac{e^{-\zeta}}{\sqrt{\zeta}}\sqrt{\Phi'}\left[1 - \zeta\frac{\Phi''}{(\Phi')^2} + \frac{\zeta^2}{2}\frac{\Phi'''}{(\Phi')^3} + \cdots\right] \\ \times & \left[1 + \frac{\zeta^2}{2}\frac{\Phi''}{(\Phi')^2} + \frac{\zeta^4}{8}\frac{(\Phi'')^2}{(\Phi')^4} + \cdots\right]\left[1 - \frac{\zeta^3}{6}\frac{\Phi'''}{(\Phi')^3} + \cdots\right]d\zeta, \end{aligned} \tag{4.10}$$

where again $\Phi = \Phi(B)$, $\Phi' = \Phi'(B)$, etc. Evaluating the integrals in (4.10) in terms of the Gamma function, using

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \qquad \frac{1}{2}\Gamma\left(\frac{1}{2}\right) - \Gamma\left(\frac{3}{2}\right) = -\frac{1}{8}\sqrt{\pi},$$
$$\frac{1}{2}\Gamma\left(\frac{5}{2}\right) - \frac{1}{6}\Gamma\left(\frac{7}{2}\right) = \frac{1}{16}\sqrt{\pi}, \qquad \frac{1}{8}\Gamma\left(\frac{9}{2}\right) - \frac{1}{2}\Gamma\left(\frac{7}{2}\right) = -\frac{15}{128}\sqrt{\pi}$$

and then using (4.10) and (4.7) in (4.6) we obtain, after cancelling the common factor $e^{\Phi(\beta)}$,

$$\frac{1}{2(\Phi')^2} + 2B^2 = 4\sqrt{\Phi'}\left[1 - \frac{1}{8}\frac{\Phi''}{(\Phi')^2} + \frac{1}{16}\frac{\Phi'''}{(\Phi')^3} - \frac{15}{128}\frac{(\Phi'')^2}{(\Phi')^4} + O(B^{-15})\right]. \tag{4.11}$$

Here we wrote the error term as $O(B^{-15})$, anticipating that $\sqrt{\Phi'} \sim B^2/2$ so that $\Phi' \sim B^4/4$ and $\Phi \sim B^5/20$ as $B \to \infty$.

From (4.11) it follows that $\Phi'(B)$ has an expansion of the form

$$\Phi'(B) = \nu_0 B^4 + \nu_1 B^{-1} + \nu_2 B^{-6} + O(B^{-11}). \tag{4.12}$$

Using (4.12) yields

$$\frac{\Phi'''}{(\Phi')^3} \sim \frac{12}{\nu_0^2}B^{-10}, \qquad \frac{(\Phi'')^2}{(\Phi')^4} \sim \frac{16}{\nu_0^2}B^{-10}$$

and

$$\frac{\Phi''}{(\Phi')^2} = \frac{4}{\nu_0}B^{-5}\left[1 - \frac{9}{4}\frac{\nu_1}{\nu_0}B^{-5} + O(B^{-10})\right].$$

Using the above in (4.11) and equating coefficients of B^2, B^{-3} and B^{-8} leads to the relations

$$\begin{aligned} 2 &= 4\sqrt{\nu_0}, \qquad 0 = \frac{\nu_1}{2\nu_0} - \frac{1}{2\nu_0}, \\ \frac{1}{2\nu_0^2} &= 4\sqrt{\nu_0}\left[\frac{9}{8}\frac{\nu_1}{\nu_0^2} - \frac{9}{8}\frac{1}{\nu_0^2} + \frac{\nu_2}{2\nu_0} - \frac{1}{8}\left(\frac{\nu_1}{\nu_0}\right)^2 - \frac{\nu_1}{4\nu_0^2}\right] \end{aligned}$$

so that $\nu_0 = 1/4$, $\nu_1 = 1$ and $\nu_2 = 5$. It follows that

$$\Phi'(B) = \frac{1}{4}B^4 + B^{-1} + 5B^{-6} + O(B^{-11})$$

and hence

$$\Phi(B) = \frac{1}{20}B^5 + \log B + \text{constant} - B^{-5} + O(B^{-10})$$

and (4.5) yields

$$\Delta(B) = c_0 B e^{B^5/20}[1 - B^{-5} + O(B^{-10})] \tag{4.13}$$

for $B\to\infty$ and some constant c_0. We shall later determine c_0 numerically. We have thus obtained (2.31) using a formal WKB expansion. Then (2.32) follows immediately from (2.27) and the fact that $B^5=b^4$.

An alternate approach to analyzing (4.2) is via a Taylor expansion in powers of b. Expanding $\bar{H}(b)$ as in (2.22) corresponds to expanding $\Delta(B)$ in the form in (2.29). Using (2.29) in (2.28) leads to

$$\begin{aligned} 0 &= \sum_{m=0}^{\infty}\sum_{\ell=0}^{m}\Delta_\ell\Delta_{m-\ell}\mathcal{B}\left(\frac{5}{2}\ell+2,\frac{5}{2}(m-\ell)+2\right)B^{\frac{5}{2}m+3}+2\sum_{m=0}^{\infty}B^{\frac{5}{2}m+3}\Delta_m+\frac{2}{\sqrt{\pi}}\sqrt{B} \\ &- \frac{4}{\sqrt{\pi}}\sum_{m=0}^{\infty}B^{\frac{5}{2}m+1}\left(1+\frac{5}{2}m\right)\Delta_m\mathcal{B}\left(\frac{5}{2}m+1,\frac{1}{2}\right). \end{aligned} \tag{4.14}$$

Here

$$\mathcal{B}(x,y)=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

is the Beta function. Since $\mathcal{B}\left(1,\frac{1}{2}\right)=2$ we obtain $\Delta_0=1/4$ and then for $m\geq 0$ we compare coefficients of $B^{\frac{5}{2}m+3}$ in (4.14), which leads to the recurrence

$$\frac{4}{\sqrt{\pi}}\Delta_{m+1}\frac{\Gamma\left(\frac{5}{2}m+\frac{7}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{5}{2}m+4\right)}\frac{5m+7}{2}=2\Delta_m+\sum_{\ell=0}^{m}\frac{\Gamma\left(\frac{5}{2}\ell+2\right)\Gamma\left(\frac{5}{2}(m-\ell)+2\right)}{\Gamma\left(\frac{5}{2}m+4\right)}\Delta_\ell\Delta_{m-\ell} \tag{4.15}$$

We can somewhat simplify this equation by getting

$$\tilde{\Delta}_m=\Gamma\left(\frac{5}{2}m+2\right)\Delta_m \tag{4.16}$$

which leads to

$$\tilde{\Delta}_{m+1}=\frac{(5m+6)(5m+4)}{8}\tilde{\Delta}_m+\frac{1}{4}\sum_{\ell=0}^{m}\tilde{\Delta}_\ell\tilde{\Delta}_{m-\ell}. \tag{4.17}$$

We have thus derived (2.24). Despite the nonlinearity in (4.17) being in the form of a convolution sum, we cannot solve (4.17) by using generating functions, due to the rapid growth of $\tilde{\Delta}_m$ as $m\to\infty$. Note that since $\tilde{\Delta}_m>0$ for all m we immediately obtain the lower bound

$$\tilde{\Delta}_m\geq k_0\left(\frac{25}{8}\right)^m\Gamma\left(m+\frac{6}{5}\right)\Gamma\left(m+\frac{4}{5}\right), \tag{4.18}$$

where k_0 is a positive constant. The right side of (4.18) follows by dropping the nonlinear term in (4.17) and replacing $=$ by $\geq$. We can take $k_0=1/\left[4\Gamma\left(\frac{6}{5}\right)\Gamma\left(\frac{4}{5}\right)\right]$, since $\tilde{\Delta}_0=1/4$.

We next analyze (4.17) for $m\to\infty$ by using a discrete WKB-type expansion. Anticipating that the nonlinear part of (4.17) becomes negligible compared to $\tilde{\Delta}_{m+1}$ as $m\to\infty$, we write

$$\tilde{\Delta}_m = k\left(\frac{25}{8}\right)^m \Gamma\left(m+\frac{6}{5}\right)\Gamma\left(m+\frac{4}{5}\right)[1+\varepsilon(m)] \tag{4.19}$$

where $\varepsilon(m) \to 0$ as $m \to \infty$.

Using (4.19) in (4.17) and retaining only the boundary terms in the sum (corresponding to $\ell = 0, 1$ and $\ell = m-1, m$) we obtain

$$k\left(\frac{25}{8}\right)^{m+1} \Gamma\left(m+1+\frac{6}{5}\right)\Gamma\left(m+1+\frac{4}{5}\right)[1+\varepsilon(m+1)] + \cdots \tag{4.20}$$

$$\begin{aligned} &= \frac{k}{8}(5m+6)(5m+4)\left(\frac{25}{8}\right)^m \Gamma\left(m+\frac{6}{5}\right)\Gamma\left(m+\frac{4}{5}\right)[1+\varepsilon(m)] \\ &+ \frac{1}{2}\tilde{\Delta}_0 k\left(\frac{25}{8}\right)^m \Gamma\left(m+\frac{6}{5}\right)\Gamma\left(m+\frac{4}{5}\right)[1+\varepsilon(m)] \\ &+ \frac{1}{2}\tilde{\Delta}_1 k\left(\frac{25}{8}\right)^{m-1} \Gamma\left(m+\frac{1}{5}\right)\Gamma\left(m-\frac{1}{5}\right)[1+\varepsilon(m-1)] + \cdots . \end{aligned}$$

From (4.15) and (4.17) we have

$$\tilde{\Delta}_1 = \frac{49}{64} = \left(\frac{7}{8}\right)^2, \quad \Delta_1 = \frac{7}{60}\frac{1}{\sqrt{\pi}} \tag{4.21}$$

and thus the fourth moment of $p_-(\beta)$ is

$$\int_{-\infty}^{\infty} \beta^4 p_-(\beta) d\beta = 24\sqrt{\pi}\Delta_1 = \frac{14}{5}, \tag{4.22}$$

where we used (2.17) and (2.27). Using $\Gamma(z+1) = z\Gamma(z)$ to simplify the left side of (4.20) we see that as $m \to \infty$ this equation becomes

$$\varepsilon(m+1) - \varepsilon(m) \sim \varepsilon'(m) = \frac{1+\varepsilon(m)}{(5m+6)(5m+4)} + O(m^{-4})$$

so that $\varepsilon'(m) \sim \frac{1}{25}m^{-2}$ and hence

$$\varepsilon(m) \sim -\frac{1}{25}\frac{1}{m}, \; m \to \infty. \tag{4.23}$$

From (4.23), (4.19) and (4.16) we obtain

$$\begin{aligned} \Delta_m &= k\left(\frac{25}{8}\right)^m \frac{\Gamma\left(m+\frac{6}{5}\right)\Gamma\left(m+\frac{4}{5}\right)}{\Gamma\left(\frac{5}{2}m+2\right)}\left[1-\frac{1}{25m}+O(m^{-2})\right] \\ &= k\sqrt{2\pi}\frac{2\sqrt{2}}{5\sqrt{5}}\frac{e^{m/2}}{\sqrt{m}} m^{-m/2} 10^{-m/2}\left[1-\frac{4}{15m}+O(m^{-2})\right], \end{aligned} \tag{4.24}$$

where the last equality follow from Stirling's approximation, in the form

Table 1:

m	Δ_m	$\Delta_m(10m)^{m/2}\sqrt{m}e^{-m/2}$	$\Delta_m(10m)^{m/2}\sqrt{m}e^{-m/2}/\left(1-\frac{4}{5m}\right)$
10	.71125(–9)	.15154	.15570
20	.73829(–20)	.15349	.15557
30	.64124(–32)	.15416	.15554
40	.10779(–44)	.15450	.15553
50	.52860(–58)	.15470	.1555341
60	.96628(–72)	.15484	.1555319
70	.77532(–86)	.15493	.1555306
80	.30690(–100)	.15501	.1555298
90	.65424(–115)	.15506	.1555292
100	.80422(–130)	.15511	.1555288
1000	.69013(–1785)	.15548	.1555270
2000	.63932(–3869)	.15550	.1555270

$$\Gamma(m+\alpha) = m^m e^{-m}\sqrt{\frac{2\pi}{m}}m^{\alpha}\left[1+\left(\frac{\alpha^2}{2}-\frac{\alpha}{2}+\frac{1}{12}\right)\frac{1}{m}+O(m^{-2})\right].$$

This analysis suggests that the lower bound (4.18) is correct asymptotically to leading order, albeit with a different constant replacing k_0. Note that (4.24) agrees with (2.30) if

$$k' = k\sqrt{2\pi}\frac{2\sqrt{2}}{5\sqrt{5}}. \tag{4.25}$$

Also, the terms retained in going from (4.17) to (4.20) are sufficient to obtain the $O(m^{-2})$ error term in (4.24), though we shall not calculate it.

The constant k cannot be determined from (4.20), and must be obtained numerically. In Table 4 we compute Δ_m by iterating (4.15).

For m in the range $[10, 2000]$ we then give $\Delta_m(10m)^{m/2}\sqrt{m}e^{-m/2}$ and $\Delta_m(10m)^{m/2}\sqrt{m}e^{-m/2}$ $/\left(1-\frac{4}{15m}\right)$. Both of these sequences should converge to k' as $m\to\infty$, with the latter converging more rapidly, as our analysis suggests it behaves as $k'[1+O(m^{-2})]$. The data in Table 1 confirm precisely our formal asymptotic analysis, and show that

$$k' \approx .1555270. \tag{4.26}$$

In Table 1 the notation .71125(−9) means $.71125\times 10^{-9}$, etc.

We next show that the asymptotic relations in (2.30) and (2.31) are consistent with (2.29). We argue that for $B\to\infty$ the dominant contribution to the sum in (2.29) comes from large values of m, where (4.24) applies. Thus we write

$$
\begin{aligned}
\Delta(B) &= \sum_{m=0}^{\infty} B^{1+\frac{5}{2}m}\Delta_m \qquad (4.27)\\
&= \int_1^{\infty} B^{1+\frac{5}{2}x} k' \left(\frac{e}{10x}\right)^{x/2} \frac{1}{\sqrt{x}}\left[1+\frac{4}{15x}+O(x^{-2})\right]dx.
\end{aligned}
$$

Here we approximated the sum by an integral. The correction terms from the Euler Maclaurin formula in this approximation are smaller than any term in the asymptotic series in (4.24).

Getting

$$
\phi = \phi(x, B) = -\frac{x}{2}\log x + \frac{5}{2}x\log B + \frac{x}{2}\log\left(\frac{e}{10}\right) \qquad (4.28)
$$

we view the integral in (4.27) as a Laplace type integral. Then the major contribution will come from where $\phi_x = 0$, and this occurs at

$$
x = x_0 = \frac{1}{10}B^5. \qquad (4.29)
$$

Noting that

$$
\phi(x_0, B) = \frac{x_0}{2}\log\left(\frac{B^5}{x_0}\right) + \frac{x_0}{2}\log\left(\frac{e}{10}\right) = \frac{1}{20}B^5 \qquad (4.30)
$$

we expand ϕ in Taylor series about $x = x_0$. Hence (4.27) becomes

$$
\begin{aligned}
\Delta(B) = k'Be^{B^5/20}\int_{-\infty}^{\infty} \exp\left[-\frac{1}{4x_0}(x-x_0)^2 + \frac{1}{12x_0^2}(x-x_0)^3 - \frac{1}{24x_0^3}(x-x_0)^4 + \cdots\right] \qquad (4.31)\\
\times\frac{1}{\sqrt{x_0}}\left[1-\frac{x-x_0}{2x_0}+\frac{3(x-x_0)^2}{8x_0^2}+\ldots\right]\left[1-\frac{4}{15x_0}+\cdots\right]dx.
\end{aligned}
$$

Scaling $x = x_0 + 2\sqrt{x_0}\xi$ leads to

$$
\begin{aligned}
\Delta(B) &= 2k'Be^{B^5/20}\int_{-\infty}^{\infty} e^{-\xi^2}\left[1+\frac{2}{3\sqrt{x_0}}\xi^3+\frac{2}{9x_0}\xi^6+O(x_0^{-3/2})\right] \qquad (4.32)\\
&\times \left[1-\frac{1}{\sqrt{x_0}}\xi+\frac{3}{2x_0}\xi^2\right]\left[1-\frac{4}{15x_0}+O(x_0^{-3/2})\right]dx\\
&= 2k'Be^{B^5/20}\sqrt{\pi}\left[1+\frac{1}{6x_0}\right]\left[1-\frac{4}{15x_0}+O(x_0^{-2})\right].
\end{aligned}
$$

In view of (4.29) we see that (4.32) agrees with the WKB expansion (4.13), provided that $c_0 = 2\sqrt{\pi}k'$, as in (2.33).

In Table 2, we calculate $\Delta(B)$ numerically for B in the range $[1, 7]$, and compare this to $\Delta(B)B^{-1}e^{-B^5/20}$ and also to $\Delta(B)B^{-1}e^{-B^5/20}/(1-B^{-5})$. Both of these functions should

Table 2:

B	$\Delta(B)$	$\Delta(B)B^{-1}e^{-B^5/20}$	$\Delta(B)B^{-1}e^{-B^5/20}/(1-B^{-5})$
1	.33176	.31558	–
2	5.0774	.51256	.5290967
3	.31140(6)	.54893	.5512023
4	.37924(23)	.55078	.5513226
5	.19895(69)	.55115	.5513282
6	.23615(170)	.55125	.5513287
7	.35143(366)	.55129	.5513288
8	.15579(713)	.55131	.5513288

Table 3:

$\eta = x^2$	$-\sqrt{\pi}\bar{H}(ix)$
5	.84204
10	.95962
15	.98659
20	.99477
25	.99772

converge to c_0, with the latter converging much more rapidly, as our analysis predicts that it behaves as $c_0[1 + O(B^{-10})]$ for $B \to \infty$. The numerics validate our asymptotic analysis, give the value in (2.19) for c_0, and also confirm (2.33) (with (4.26)).

We next consider $\bar{H}(b)$ and (2.22) for b purely imaginary. This analysis will facilitate the numerical calculation of $p_-(\beta)$, which we expressed in (2.16) as a Fourier integral involving $S(x) = 1 + \sqrt{\pi}\bar{H}(ix)$. We get

$$b = ix, \quad x > 0. \tag{4.33}$$

From (2.32) we see that as $b \to +\infty$, corresponding to $x \to -i\infty$, we have $1 + \sqrt{\pi}\bar{H}(b) \sim \sqrt{\pi}\bar{H}(b)$, with an exponentially small error. However, this ceases to be true for some complex ranges of b, i.e., there is a Stoke's phenomenon in the asymptotic behavior of $\bar{H}(b)$.

With (4.33) and (2.22) we have

$$\bar{H}(ix) = -x^2 \sum_{m=0}^{\infty} (-1)^m \Delta_m x^{2m}, \tag{4.34}$$

and we note that the estimate in (2.30) shows that $\bar{H}(\cdot)$ is an entire function. From (2.16) we would expect $S(x)$ to decay as $x \to \infty$ and thus $\bar{H}(ix) \to -1/\sqrt{\pi}$ as $x \to \infty$.

In Table 3 we give numerical values of $-\bar{H}(ix)\sqrt{\pi}$ with $\eta = x^2$, for η in the range $[5, 25]$. The data indeed indicates that the quantity approaches the value 1 quite rapidly, as $\eta, x \to \infty$. To obtain the rate of approach it becomes convenient to define $\Delta_{-1} = 1/\sqrt{\pi}$ and then

$$S(x) \equiv \sum_{m=-1}^{\infty} (-1)^{m+1} x^{2m+2} \sqrt{\pi} \Delta_m = 1 + \sqrt{\pi} \bar{H}(ix). \tag{4.35}$$

Setting

$$\bar{H}(b) = \tilde{H}(x) = \tilde{H}(-ib) \tag{4.36}$$

we see that (4.2) becomes

$$\begin{aligned} 0 &= \int_0^1 \frac{\tilde{H}(xu^{5/4})\tilde{H}(x(1-u)^{5/4})}{[u(1-u)]^{3/2}} du - \frac{2x^2}{\sqrt{\pi}} \\ &+ \frac{2}{\sqrt{\pi}} \int_0^1 \left[\frac{\tilde{H}((1-u)^{5/4}x)}{(1-u)^{3/2}} - \tilde{H}(x) \right] \frac{du}{u^{3/2}} - \left(2x^2 + \frac{4}{\sqrt{\pi}} \right) \tilde{H}(x). \end{aligned} \tag{4.37}$$

Then letting

$$\tilde{H}(x) = -x^{6/5} \Lambda(x^{4/5}), \quad y = x^{4/5}$$

we obtain from (4.37) the equation (2.37). Its analysis, using a Laplace transform, we already indicated in (2.38) – (2.43). Inverting the Laplace transform in (2.38) and using (2.41) leads to

$$\begin{aligned} \Lambda(y) &= \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{y\phi} [-2\sqrt{\phi} + U_1(\phi)] d\phi \\ &= \frac{1}{2\pi i} \frac{d}{dy} \int_{-i\infty+\varepsilon}^{i\infty+\varepsilon} e^{y\phi} \left[-\frac{2}{\sqrt{\phi}} + \frac{U_1(\phi)}{\phi} \right] d\phi \\ &= -\frac{2}{\sqrt{\pi}} \frac{d}{dy} (y^{-1/2}) + \frac{1}{2\pi i} \frac{d}{dy} \int_{-i\infty+\varepsilon}^{i\infty+\varepsilon} e^{y\phi} \frac{U_1(\phi)}{\phi} d\phi. \end{aligned} \tag{4.38}$$

The asymptotic behavior of the last integral as $y \to +\infty$ is obtained by shifting the contour to the left, until we encounter the first singularity of $U_1(\phi)$, which is the double pole at $\phi = -\nu_*$. Thus (4.38) becomes

$$\Lambda(y) - \frac{1}{\sqrt{\pi}} y^{-3/2} \sim -\frac{12}{2\pi i} \int_{-i\infty}^{i\infty} \frac{e^{y\zeta}}{\zeta^2} e^{-y\nu_*} d\zeta \tag{4.39}$$

which leads to (2.44). Then, since $y = x^{4/5}$, we have

$$
\begin{aligned}
S(x) &= 1+\sqrt{\pi}\bar{H}(ix) = 1+\sqrt{\pi}\tilde{H}(x) \\
&= 1-\sqrt{\pi}x^{6/5}\Lambda(x^{4/5}) \\
&= \left(\Lambda(y)-\frac{1}{\sqrt{\pi}}y^{-3/2}\right)(-\sqrt{\pi}y^{3/2}) \\
&\sim -12y^{5/2}e^{-\nu_* y}\sqrt{\pi}
\end{aligned}
$$

and this is the result in (2.45).

To determine ν_* numerically, we can solve the Painlevè equation (2.42) numerically and see where it blows up. However, it is somewhat difficult to impose accurately the condition as $\phi \to \infty$. Thus we instead calculate Δ_m from (2.23) and (2.24) for m sufficiently large so we can precisely estimate

$$
R(x) \equiv -\frac{1}{x^{4/5}} \log\left\{\frac{1}{12}\sum_{m=0}^{\infty}(-1)^m x^{2m-2}\Delta_{m-1}\right\}. \tag{4.40}
$$

Our analysis suggests that

$$
R(x) \to \nu_* \text{ as } x \to \infty. \tag{4.41}
$$

Note that the right side of (4.40) corresponds to solving the asymptotic relation (2.45) for ν_*, in terms of the sum in (4.35). In doing this we divided the sum by $12x^2$, which should improve the convergence, since it adds to $R(x)$ the terms of order $O(x^{-4/5}\log x)$ and $O(x^{-4/5})$, which correspond to the first two correction terms to the limit in (4.41).

In Table 4 we consider x in the range $[1,8]$ and compute the sum $\sum_{m=0}^{\infty}(-1)^m x^{2m}\Delta_{m-1}$ as well as $R(x)$. The data again confirm our asymptotic analysis and lead to the value of ν_* in (2.46). We note that there is a lot of cancelation in the alternating sum. To obtain the result for $x=8$ we needed to truncate the sum at about $m = 2,500$ and do the calculation with 175 digits of precision !

In the next section we use our results for the function $\bar{H}(b)$ to study the limiting density $p_-(\beta)$, both asymptotically and numerically.

5 Asymptotic and Numerical Studies of $p_-(\beta)$

We study $p_-(\beta)$ in (2.16) which is a proper density which satisfies $p_-(-\beta) = p_-(\beta)$,

$$
\int_{-\infty}^{\infty} p_-(\beta)d\beta = 1,\ \int_{-\infty}^{\infty}\beta^2 p_-(\beta)d\beta = \frac{1}{2}\sqrt{\pi} \ \text{ and } \ \int_{-\infty}^{\infty}\beta^4 p_-(\beta)d\beta = \frac{14}{5}.
$$

Higher order moments follow easily from (2.17) and (2.22) – (2.24).

We obtain the tail behavior as $\beta \to \infty$. We argue that since $\bar{H}(b)$ is entire, there must be a saddle point in the integral

Table 4:

x	$\sum_{m=0}^{\infty}(-1)^m x^{2m}\Delta_{m-1}$	$R(x)$
1	.36868	3.482732
2	.12273	3.428240
3	.29170(–1)	3.411948
4	.61976(–2)	3.411285
5	.12817(–2)	3.411601
6	.26478(–3)	3.411674
7	.55131(–4)	3.411675
8	.11616(–4)	3.411672

$$\frac{1}{2\pi i}\int_{-i\infty}^{i\infty} e^{-\beta b}[1+\sqrt{\pi}\bar{H}(b)]db \tag{5.1}$$

along the positive real b-axis, far from the origin. Then we shift the integration contour in (5.1) far to the right and use the estimate in (2.32), thus obtaining for $\beta \to \infty$

$$p_-(\beta) \sim \frac{1}{2\pi i}\int_{Br} \sqrt{\pi}c_0 e^{b^4/20} e^{-b\beta} b^2[1+O(b^{-4})]db. \tag{5.2}$$

The integrand in (5.2) has a saddle point where

$$\frac{d}{db}\left(-b\beta+\frac{b^4}{20}\right)=0 \Rightarrow b=b_0(\beta)=(5\beta)^{1/3}. \tag{5.3}$$

The directions of steepest descent are $\arg(b-b_0)=\pm\pi/2$ and the standard saddle point estimate gives

$$\begin{aligned} p_-(\beta) &\sim \frac{1}{2\pi i}\sqrt{\pi}c_0 b_0^2 \exp\left(\frac{b_0^4}{20}-b_0\beta\right)\int_{-i\infty}^{i\infty}\exp\left(\frac{3b_0^2}{10}\zeta^2\right)d\zeta \\ &= c_0\sqrt{\frac{5}{6}}b_0(\beta)\exp\left[-\frac{3}{4}\beta b_0(\beta)\right] \end{aligned} \tag{5.4}$$

which is precisely the result in (2.18). Note that using the $O(b^{-4})$ correction term in (2.32) will allow us to compute an $O(\beta^{-4/3})$ correction to (2.18). It appears that the asymptotic series for $\bar{H}(b)$ involves powers of b^{-4}, and that in $p_-(\beta)$ involves powers of $\beta^{-4/3}$.

Next we discuss the analyticity of $p_-(\beta)$ about $\beta=0$. If the density were analytic we could write the last equation in (2.16) as

$$p_-(\beta)=\frac{1}{\pi}\sum_{m=0}^{\infty} d_m\beta^{2m}\frac{(-1)^m}{(2m)!} \tag{5.5}$$

where

$$d_m = \int_0^\infty x^{2m} S(x) dx. \tag{5.6}$$

Using the estimate (2.45) for $S(x)$ we can use (5.6) to infer the behavior of d_m for $n \to \infty$:

$$\begin{aligned} d_m &\sim \int_0^\infty x^{2m} 12\sqrt{\pi} x^2 \exp(-\nu_* x^{4/5}) dx \\ &= \int_0^\infty y^{\frac{5}{2}(m+1)} e^{-\nu_* y} 12\sqrt{\pi} \frac{5}{4} y^{\frac{1}{4}} dy \\ &= 12\sqrt{\pi}\frac{5}{4} \nu_*^{-\frac{5}{2}m - \frac{15}{4}} \Gamma\left(\frac{5}{2}m + \frac{15}{4}\right). \end{aligned} \tag{5.7}$$

It follows that $d_m/(2m)!$ grows faster than exponentially (roughly like $m^{m/2}$) and thus the series (5.5) cannot define an analytic function. By appropriate contour rotation in (2.16) we believe that we can show that $p_-(\beta)$ has an essential singularity at $\beta = 0$, and infer its behavior as $\beta \to 0$ for all values of $\arg(\beta) \in [0, 2\pi)$. However, this would require knowing the behavior of $\bar{H}(b)$ for $b \to \infty$ and arbitrary $\arg(b)$, which we did not analyze. Note that this would lead also to a better understanding of how the relation (2.32) (presumably valid also in some sector (range of $\arg(b)$) containing the real axis) transitions to the relation $\bar{H}(b) \sim -1/\sqrt{\pi}$, which is certainly true for $\arg(b) = \pm\pi/2$.

We next calculate numerically $p_-(\beta)$ for moderate values of $\beta \geq 0$. First we discretize the integral(s) in (2.16) using a step size h, which yields the approximation

$$\begin{aligned} p_-(\beta) &\doteq \frac{h}{2\pi} \sum_{J=-\infty}^{\infty} e^{-J\beta h i} [1 + \sqrt{\pi} \bar{H}(Jhi)] \\ &= \frac{h}{2\pi} \sum_{J=-\infty}^{\infty} e^{-J\beta h i} \sum_{m=0}^{\infty} (-1)^m \sqrt{\pi} \Delta_{m-1} (Jh)^{2m} \\ &= \frac{h}{2\pi} \left[1 + \sum_{J=1}^{\infty} 2\cos(J\beta h) \sum_{m=0}^{\infty} (-1)^m \sqrt{\pi} \Delta_{m-1} (Jh)^{2m} \right]. \end{aligned} \tag{5.8}$$

Our basic procedure is as follows: choose some $\beta_{\max}$ (say $\beta_{\max} = 3$), choose a small step size h, truncate the limit on the outer sum at $J = N_{\max}$ (with $hN_{\max} \gg 1$), calculate the inner sum to some specified precision (to hold uniformly in the range $Jh \in [0, hN_{\max}]$), and plot the right side of (5.8) as a function of β over the range $[0, \beta_{\max}]$.

Our numerical studies show that it is desirable to estimate the truncated tail in the J-sum in (5.8) analytically. More precisely, we write (2.16) as

$$\begin{aligned} p_-(\beta) &= \frac{1}{2\pi}\left[\int_0^{hN_{\max}} 2\cos(\beta x)S(x)dx + \int_{hN_{\max}}^{\infty} 2\cos(\beta x)S(x)dx\right] \qquad (5.9)\\ &\doteq \frac{h}{2\pi}\left[1 + \sum_{J=1}^{N_{\max}} 2\cos(J\beta h)S(Jh) - \cos(\beta N_{\max}h)S(N_{\max}h)\right]\\ &+ \frac{1}{\pi}\int_{hN_{\max}}^{\infty}\cos(\beta x)S(x)dx. \end{aligned}$$

Here the term involving $S(N_{\max}h)$ comes from the Euler MacLaurin approximation to the first integral in (5.9). The last integral in (5.9) we approximate using (2.45), thus defining

$$\text{TAIL }(\beta; hN_{\max}) \equiv \frac{12}{\sqrt{\pi}}\int_{hN_{\max}}^{\infty} x^2\cos(\beta x)\exp(-\nu_* x^{4/5})dx. \qquad (5.10)$$

For $hN_{\max}\to\infty$ we can further approximate the integral in (5.10) using

$$\int_z^{\infty}\cos(\beta x)S(x)dx \sim -\frac{12\sqrt{\pi}}{\beta}\sin(\beta z)z^2\exp(-\nu_* z^{4/5}),\ \ z\to\infty. \qquad (5.11)$$

However, this approximation is not valid for small β and it is not hard to evaluate (5.10) numerically. We thus follow the procedure outlined below (5.8) by using the refined approximation in (5.9), with (5.10). Note that approximations for the derivatives of $p_-(\beta)$ can then be obtained by *analytically* differentiating (5.9). This method was used to obtain the graphs in Figure 1 and Figure 2. Here we used $h = .025$ and $hN_{\max} = 7.5$, so that $N_{\max} = 300$. Without the tail estimate in (5.10), much smaller values of h and $1/N_{\max}$ would be needed.

We illustrate our approach in Table 5. Here we compute $p_-(\beta)$ at the values $\beta = 0, 1, 2, 3$ and also give $p_-''(0)$ and the inflection point β_c. The table illustrates the convergence as $h \to 0$ and $hN_{\max}\to\infty$, and leads to the values in (2.20) and (2.21). We comment that in obtaining Table 5, care must be taken to accurately calculate $S(x)$ in the range $x \in [0, N_{\max}h]$. For example, for $N_{\max}h = 8$ we needed to truncate the m-sum in (5.8) (which defines $S(Jh)$) at $m = 1500$ and use 100 digits of precision. We can estimate the number of terms we need to retain in the sum, in terms of x, analytically. For example, take $x = 10$. Then if we truncate the sum at $m = M = 3000$ we have $x^{2M} = 10^{6000}$, while Δ_{3000} is of the order of 10^{-6066}. Thus this truncation should give $S(10)$ correctly to about 60 digits. We find that $S(10) \approx .95290(-6)$, and this required doing the calculation in 250 digits of precision.

Appendix A

Here we briefly discuss the scalings (3.19) and (3.20). By using (3.20) and (3.21) in (2.9) we obtain

Table 5:

h	$hN_{\max}$	$p_-(0)$	$p_-(1)$	$p_-(2)$	$p_-(3)$	$p''_-(0)$	β_c
0.1	5	.45727	.22450	.041958	.0047815	–.71460	.75901
0.05	5	.45727	.22450	.041957	.0047818	–.71461	.75899
0.025	5	.45727	.22450	.041957	.0047819	–.71462	.75899
0.01	5	.45727	.22450	.041957	.0047820	–.71462	.75898
0.1	7.5	.45727	.22450	.041957	.0047820	–.71462	.75898
0.05	7.5	.45727	.22450	.041957	.0047819	–.71462	.75898
0.025	7.5	.45727	.22450	.041957	.0047819	–.71462	.75898

$$C_{n+1} + \frac{4^{n+1}}{(n+1)^{3/2}} F\left(\left(1+\frac{1}{n}\right)^{3/2} a_1, \left(1+\frac{1}{n}\right)^{3/2} b_1; n+1\right) = T_1 + T_2 + T_3 + T_4 \quad \text{(A.1)}$$

where

$$\begin{aligned}
T_1 &= \sum_{i=0}^{n} \left(1+\frac{a_1}{n^{3/2}}\right)^i \left(1+\frac{b_1}{n^{3/2}}\right)^{n-i} C_i C_{n-i}, \\
T_2 &= \sum_{i=0}^{n} \left(1+\frac{a_1}{n^{3/2}}\right)^i \left(1+\frac{b_1}{n^{3/2}}\right)^{n-i} C_i \frac{4^{n-i}}{(n-i)^{3/2}} F\left(\left(1-\frac{i}{n}\right)^{3/2} a_1 \left(1-\frac{i}{n}\right)^{3/2} b_1; n-i\right), \\
T_3 &= \sum_{i=0}^{n} \left(1+\frac{a_1}{n^{3/2}}\right)^i \left(1+\frac{b_1}{n^{3/2}}\right)^{n-i} C_{n-i} \frac{4^i}{i^{3/2}} F\left(\left(\frac{i}{n}\right)^{3/2} a_1, \left(\frac{i}{n}\right)^{3/2} b_1; i\right),
\end{aligned}$$

and

$$\begin{aligned}
T_4 &= \sum_{i=0}^{n} \left(1+\frac{a_1}{n^{3/2}}\right)^i \left(1+\frac{b_1}{n^{3/2}}\right)^{n-i} \frac{4^n}{[i(n-i)]^{3/2}} F\left(\left(\frac{i}{n}\right)^{3/2} a_1, \left(\frac{i}{n}\right)^{3/2} b_1; i\right) \\
&\times\ F\left(\left(1-\frac{i}{n}\right)^{3/2} a_1, \left(1-\frac{i}{n}\right)^{3/2} b_1; n-i\right).
\end{aligned}$$

This is analogous to (3.28) - (3.32).

Omitting details we expand the terms in (A.1) as $n \to \infty$, assuming that $F(a_1, b_1; n) \to F(a_1, b_1)$ as $n \to \infty$. By Euler MacLaurin we have

$$T_4 \sim \frac{4^n}{n^2} \int_0^1 \frac{F(a_1 x^{3/2}, b_1 x^{3/2}) F(a_1(1-x)^{3/2}, b_1(1-x)^{3/2})}{[x(1-x)]^{3/2}} dx. \quad \text{(A.2)}$$

Then for T_1 we obtain

$$
\begin{aligned}
T_1 - C_{n+1} &\sim \sum_{i=0}^{n} \frac{ia_1 + (n-i)b_1}{n^{3/2}} C_i C_{n-i} &\text{(A.3)}\\
&= \frac{a_1+b_1}{n^{3/2}} \sum_{i=0}^{n} iC_iC_{n-i}\\
&= \frac{a_1+b_1}{n^{3/2}} \sum_{i=0}^{n} (n-i)C_{n-i}C_i\\
&\sim \frac{a_1+b_1}{\sqrt{n}} \sum_{i=0}^{\infty} \frac{4^n}{n^{3/2}} \frac{1}{\sqrt{\pi}} 4^{-i} C_i\\
&= \frac{4^n}{n^2} \frac{2(a_1+b_1)}{\sqrt{\pi}}.
\end{aligned}
$$

We next use $T_2(b_1, a_1) = T_3(a_1, b_1)$ and write the former as

$$
\begin{aligned}
T_2 &= \frac{4^n}{n^{3/2}} \sum_{i=0}^{n} 4^{-i} C_i \left(1 + \frac{a_1}{n^{3/2}}\right)^i \left(1 + \frac{b_1}{n^{3/2}}\right)^{n-i} F(a_1, b_1) &\text{(A.4)}\\
&+ \sum_{i=0}^{n} \left(1 + \frac{a_1}{n^{3/2}}\right)^i \left(1 + \frac{b_1}{n^{3/2}}\right)^{n-i} C_i \frac{4^{n-i}}{n^{3/2}} \left[\frac{F\left(\left(1-\frac{i}{n}\right)^{3/2} a_1, \left(1-\frac{i}{n}\right)^{3/2} b_1; n-i\right)}{\left(1-\frac{i}{n}\right)^{3/2}} - F(a_1, b_1)\right].
\end{aligned}
$$

The second term in (A.4) can be approximated for $i = O(n)$ by an integral, which leads to

$$
\frac{4^n}{n^2} \int_0^1 \frac{1}{\sqrt{\pi} x^{3/2}} \left[\frac{F((1-x)^{3/2} a_1, (1-x)^{3/2} b_1)}{(1-x)^{3/2}} - F(a_1, b_1)\right] dx. \qquad \text{(A.5)}
$$

Using (3.38) we write the first part of (A.4) as

$$
\begin{aligned}
&\frac{4^n}{n^{3/2}} \sum_{i=0}^{n} 4^{-i} C_i \left[\left(1 + \frac{a_1}{n^{3/2}}\right)^i \left(1 + \frac{b_1}{n^{3/2}}\right)^{n-i} - 1 \right] F(a_1, b_1) + F(a_1, b_1) \frac{4^n}{n^{3/2}} \left[2 - \frac{2}{\sqrt{\pi n}} + O(n^{-3/2}) \right] \quad \text{(A.6)}\\
&\sim \frac{4^n}{n^{3/2}} \left[\sum_{i=0}^{n} 4^{-i} C_i \frac{ia_1 + (n-i)b_1}{n^{3/2}} + 2 - \frac{2}{\sqrt{\pi n}} \right] F(a_1, b_1) = \frac{4^n}{n^{3/2}} \left[2 + \frac{2b_1}{\sqrt{n}} - \frac{2}{\sqrt{\pi n}} + O(n^{-1}) \right] F(a_1, b_1).
\end{aligned}
$$

Using (A.2) – (A.6) in (A.1) we obtain in the limit $n \to \infty$ the nonlinear integral equation

$$
\begin{aligned}
0 &= \int_0^1 \frac{F(x^{3/2} a_1, x^{3/2} b_1) F((1-x)^{3/2} a_1, (1-x)^{3/2} b_1)}{[x(1-x)]^{3/2}} dx &\text{(A.7)}\\
&+ \frac{2}{\sqrt{\pi}} \int_0^1 \left\{ \frac{F((1-x)^{3/2} a_1, (1-x)^{3/2} b_1)}{(1-x)^{3/2}} - F(a_1, b_1) \right\} \frac{dx}{x^{3/2}}\\
&- \frac{4}{\sqrt{\pi}} F(a_1, b_1) + 2(a_1 + b_1) \left[\frac{1}{\sqrt{\pi}} + F(a_1, b_1) \right].
\end{aligned}
$$

To analyze (A.7) we take $a_1, b_1 < 0$ and set

$$F(a_1, b_1) = (-a_1 - b_1)\bar{F}(A_1, B_1);\;\; A_1 = (-a_1)^{2/3}, B_1 = (-b_1)^{2/3}. \tag{A.8}$$

After some simplification (A.7) then becomes

$$\begin{aligned} 2\bar{F}(A_1, B_1) &= \int_0^1 \bar{F}(A_1(1-x), B_1(1-x))\bar{F}(A_1 x, B_1 x)dx \\ &- \frac{4}{\sqrt{\pi}} \frac{1}{A_1^{3/2} + B_1^{3/2}} \int_0^1 \frac{1}{\sqrt{1-x}} \frac{d}{dx}[\bar{F}(A_1 x, B_1 x)]dx. \end{aligned} \tag{A.9}$$

Now let $C = (A_1^{3/2} + B_1^{3/2})^{2/3} = (-a_1 - b_1)^{2/3}$ with

$$\bar{F}(A_1, B_1) = F_*(C), \tag{A.10}$$

so that (A.9) becomes

$$2CF_*(C) = \int_0^C F_*(y)F_*(C-y)dy - \frac{4}{\sqrt{\pi}} \int_0^C \frac{1}{\sqrt{C-y}} F_*'(y)dx. \tag{A.11}$$

Note that this integral equation, upon getting $F_*(C) = \frac{1}{2}\bar{D}(2^{-2/3}C)$, is that obtained in [11] for the Airy distribution of total path length.

We have thus shown that the scaling (3.20) leads to the approximation $G_n(w, v) \sim 4^n n^{-3/2}[\pi^{-1/2} + F_0(a_1 + b_1)]$, where $F_0(a_1 + b_1) = F_*((-a_1 - b_1)^{2/3})$ for $a_1, b_1 < 0$. The corresponding limiting density, with the scaling (3.19) is

$$p_0(\alpha_1, \beta_1) = \frac{1}{(2\pi i)^2} \int_{-i\infty}^{i\infty} \int_{-i\infty}^{i\infty} e^{-\alpha_1 a_1} e^{\beta_1 b_1} [1 + \sqrt{\pi} F_0(a_1 + b_1)] da_1 db_1. \tag{A.12}$$

This implies that $p_0(\alpha_1, \beta_1)$ has the form $\delta(\alpha_1 - \beta_1)$ times an Airy distribution. Thus seeing the fine structure of the distribution of the left and right path difference requires a different scaling, that has $p - q = o(n^{3/2})$.

Appendix B

Here we briefly discuss the functional equation (2.11) for the triple generating function of $N(p, q; n)$. We use the scaling (3.24) and also set

$$z = \frac{1}{4}\left(1 + \frac{\zeta}{n}\right) = \frac{1}{4} + O(n^{-1}) \tag{B.1}$$

with

$$C(w, v, z) = \bar{C}(\zeta, a, b) = \bar{C}\left((4z-1)n, \frac{1}{2}n^{3/2}(w + v - 2), \frac{1}{2}n^{5/4}(w - v)\right). \tag{B.2}$$

With (B.1) and (B.2), (2.11) becomes

$$\begin{aligned}\bar{C}(\zeta, a, b) &= 1 + \frac{1}{4}\left(1 + \frac{\zeta}{n}\right)\bar{C}\left(\zeta + \frac{b}{n^{1/4}} + \frac{a}{\sqrt{n}} + \frac{\zeta b}{n^{5/4}} + \frac{\zeta a}{n^{3/2}}, a, b\right) \\ &\times \ \bar{C}\left(\zeta - \frac{b}{n^{1/4}} + \frac{a}{\sqrt{n}} - \frac{\zeta b}{n^{5/4}} + \frac{\zeta a}{n^{3/2}}, a, b\right).\end{aligned} \tag{B.3}$$

Next we let

$$\bar{C}(\zeta, a, b) = 2\left[1 + \frac{1}{\sqrt{n}} D(\zeta, a, b; n)\right] \tag{B.4}$$

where $D = O(1)$ as $n \to \infty$. Then expanding D as

$$D(\zeta, a, b; n) = D^{(0)}(\zeta, a, b) + n^{-1/4} D^{(1)}(\zeta, a, b) + O(n^{-1/2}) \tag{B.5}$$

we obtain the limiting ODE

$$0 = \zeta + 2aD^{(0)}_{\zeta} + b^2 D^{(0)}_{\zeta\zeta} + [D^{(0)}]^2. \tag{B.6}$$

When $b = 0$ this becomes a Ricatti equation that can be solved explicitly in terms of Airy functions. When $a = 0$ it reduces to the Painlevè transcendent. We can also relate (B.6) to (3.51). By setting

$$\zeta = (-a)^{2/3}(-S), \ D^{(0)} = (-a)^{1/3}\bar{D} \tag{B.7}$$

and recalling that $\theta = (-a)^{2/3} b^{-4/5}$ (cf. (3.46)) we obtain from (B.6)

$$O = -S + 2\bar{D}_S + \bar{D}^2 + \theta^{-5/2}\bar{D}_{SS}. \tag{B.8}$$

Then $2\bar{D}$ satisfies the same equation (3.51) as the function T_1.

References

[1] M. Abramowitz, and I. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1964.

[2] G. Andrews, R. Askey and R. Roy, *Special Functions*, Cambridge University Press, 1999.

[3] C. Bender and S. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, Mc-Graw Hill, 1978.

[4] P. Flajolet, and A. Odlyzko, The Average Height of Binary Trees and Other Simple Trees, *J. Computer and System Sciences*, 25, 171–213, 1982.

[5] P. Flajolet, P. Poblete, and A. Viola, On the Analysis of Linear Probing Hashing, *Algorithmica*, 22, 490–515, 1998.

[6] P. Flajolet and R. Sedgewick, *Analytical Combinatorics*, in preparation.

[7] E. L. Ince, *Ordinary Differential Equations*, Dover, New York 1956.

[8] S. Janson, Left and Right Pathlengths in Random Binary Trees, preprint, 2004.

[9] J. Kim and B. Pittel, Confirming the Kleitman–Watson Conjecture on the Largest Coefficient in a q-Catalan Number, *J. Comb. Theory*, Ser. A, 92, 197-206, 2000.

[10] C. Knessl and W. Szpankowski, Quicksort algorithm again revisited, *Discrete Math. Theor. Comput. Sci.*, 3 43–64, 1999.

[11] C. Knessl and W. Szpankowski, Enumeration of Binary Trees, Lempel-Ziv'78 Parsings, and Universal Types, *Proc. of the Second Workshop on Analytic Algorithmics and Combinatorics*, (ANALCO04), Vancouver, 2005.

[12] D. E. Knuth, *Selected Papers on the Analysis of Algorithms*, Cambridge University Press, Cambridge, 2000.

[13] D. E. Knuth, *The Art of Computer Programming. Fundamental Algorithms. Combinatorial Algorithms*, Vol. 4, 2004; see `http://www-cs-faculty.stanford.edu/ knuth/news.html`.

[14] G. Louchard, The Brownian Excursion Area: A Numerical Analysis, *Comp. & Maths. with Appls.*, 10, 413-417, 1984.

[15] J-F. Marckert, The Rotation Correspondence is Asymptotically a Dilatation, *Random Structures & Algorithms*, 24, 118-132, 2004.

[16] G. Seroussi, On Universal Types, *Proc. ISIT 2004*, pp. 223, Chicago, 2004.

[17] G. Seroussi, "On the Number of t-ary Trees with a Given Pathlength", preprint.

[18] W. Szpankowski, *Average Case Analysis of Algorithms on Sequences*, John Wiley & Sons, New York, 2001.

[19] L. Takács, A Bernoulli Excursion and its Various Applications, *J. Appl. Probab.*, 23, 557-585, 1991.

[20] L. Takács, Conditional Limit Theorems for Branching Processes, *J. Applied Mathematics and Stochastic Analysis*, 4, 263-292, 1991.

[21] L. Takács, The Asymptotic Distribution of the Total Heights of Random Rooted Trees, *Acta Sci. Math. (Szegad)*, 57, 613-625, 1993.

[22] K. Watson and D. Kleitman, On the Asymptotic Number of Tournament Score Sequences, *J. Comb. Theory*, Ser. A, 35, 208-230, 1983.