Random trees between two walls: exact partition function

2003 J. Phys. A: Math. Gen. 36 12349

(http://iopscience.iop.org/0305-4470/36/50/001)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **36** (2003) 12349–12366 PII: S0305-4470(03)65568-X

Random trees between two walls: exact partition function

J Bouttier, P Di Francesco and E Guitter

Service de Physique Théorique, CEA/DSM/SPhT, Unité de Recherche Associée au CNRS, CEA/Saclay, 91191 Gif sur Yvette Cedex, France

E-mail: bouttier@spht.saclay.cea.fr, philippe@spht.saclay.cea.fr and guitter@spht.saclay.cea.fr

Received 26 June 2003, in final form 8 September 2003
Published 2 December 2003
Online at stacks.iop.org/JPhysA/36/12349

Abstract

We derive the exact partition function for a discrete model of random trees embedded in a one-dimensional space. These trees have vertices labelled by integers representing their position in the target space, with the solid-on-solid constraint that adjacent vertices have labels differing by ± 1. A non-trivial partition function is obtained whenever the target space is bounded by walls. We concentrate on the two cases where the target space is (i) the half-line bounded by a wall at the origin or (ii) a segment bounded by two walls at a finite distance. The general solution has a soliton-like structure involving elliptic functions. We derive the corresponding continuum scaling limit which takes the remarkable form of the Weierstrass $\wp$ function with constrained periods. These results are used to analyse the probability for an evolving population spreading in one dimension to attain the boundary of a given domain with the geometry of the target (i) or (ii). They also translate, via suitable bijections, into generating functions for bounded planar graphs.

PACS numbers: 02.50.−r, 05.40.−a, 05.50.+q

1. Introduction

In this paper, we consider a simple model describing the embedding of a random tree in one dimension. More precisely, we consider random *rooted* trees whose vertices are labelled by integers representing their possible discrete positions in a one-dimensional target space. We moreover impose that two neighbouring vertices on the tree have labels differing by +1 or −1, which allows us to view the edges of the tree as rigid segments of unit length embedded in the real line (see figure 1 for an illustration). We choose to consider the case of so-called *planar* trees, i.e. we count as *distinct* all trees obtained by permuting any two descendent subtrees at a given vertex. This model is nothing but a discrete version of the so called one-dimensional Brownian snake [1] which is used to describe branching processes, for instance the spreading

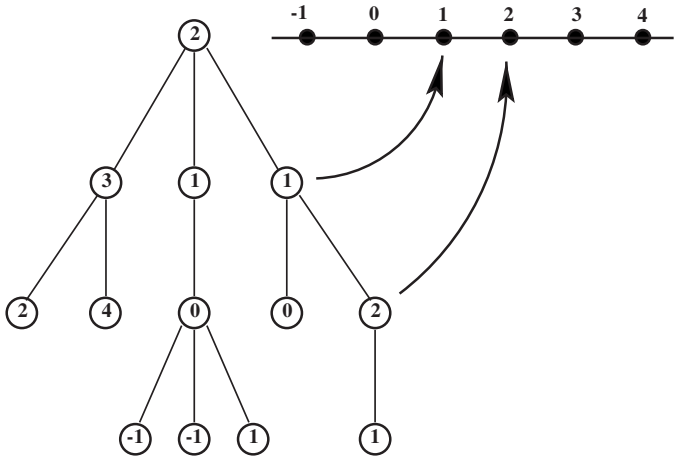

Figure 1. A sample rooted labelled planar tree, with root (top) vertex labelled by 2. Neighbouring vertices have labels differing by ± 1. These labels may be viewed as positions on a target integer line as indicated.

of a population in a one-dimensional target space. In this language, we may think of our rooted trees as representing 'genealogical trees' for the lineage of an initial individual (materialized by the root vertex), while the labels represent the positions in space of all the descendent individuals. The spreading process is modelled here by demanding that each individual lives at distance 1 from its parent[1]. We will discuss this interpretation in detail in section 5.

Alternatively, we may view this model as a statistical solid-on-solid (SOS) model with heights given by the labels, and whose base space is a random tree. The SOS rule of having neighbouring heights differing by ± 1 is responsible for the 'integrability' of the model. However, as opposed to the two-dimensional lattice case where a roughening transition takes place, we expect here that the discrete nature of the heights is eventually irrelevant when working with large tree-like base spaces.

We shall consider a statistical sum over all such trees with a weight g per edge, and with a fixed position, say n, of the root vertex. With no bounds on the labels, the partition function is trivial as it amounts to counting rooted planar trees with N edges (in number c_N where $c_N = \binom{2N}{N}/(N+1)$ is the Nth Catalan number), each of which gives rise to 2^N possible embeddings. Similarly, the width $w_N = \sqrt{\langle n^2 \rangle_N}$ for the fluctuations of the labels n in a random tree of size N and with root label 0 is easily obtained as

$$w_N^2 = \frac{1}{2}\left(\frac{4^N}{\binom{2N}{N}} - 1\right) \tag{1.1}$$

hence $w_N \sim (\pi N/4)^{1/4}$ for large N.

The problem becomes more interesting in the presence of a wall, say at position -1, which amounts to imposing that all labels be non-negative. Similar trees were introduced under the name of 'well-labelled trees' in [2] in connection with the enumeration of rooted planar tetravalent graphs, with vertex labels representing the (necessarily non-negative) geodesic distance on the graph to the root vertex[2]. The explicit form of the corresponding partition

[1] The particular discrete spreading rule chosen here should not affect the universality of the continuum answer.

[2] In [2], a slightly different constraint is imposed on the labels, demanding that neighbouring vertices have labels differing by $0, \pm 1$. Our results will be easily extended to this modified case in section 6, with no fundamental difference.

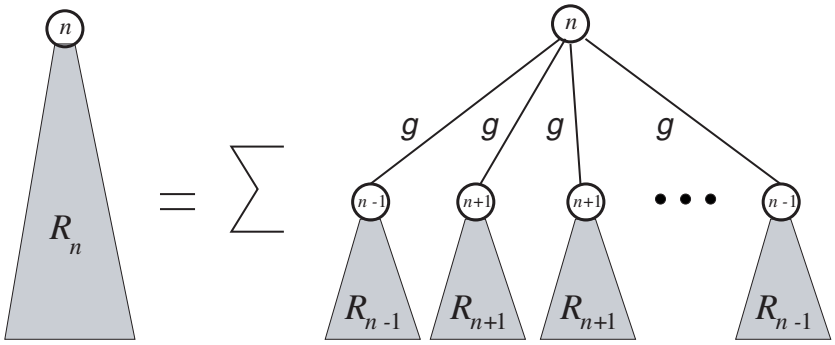

Figure 2. A sketch of the master equation (2.1). A tree contributing to R_n is decomposed according to the sequence of descendents of its root (labelled by n). Each descendent vertex is arbitrarily labelled by $n \pm 1$ henceforth the descendent subtrees are generated by $R_{n\pm1}$ accordingly. Each edge connected to the root is given a weight g. The summation over all possible configurations produces the rhs of equation 2.1.

function as a function of g and n was given in [3] in the context of planar graph enumeration and will be recalled in the next section. A remarkable outcome of this solution is the emergence of discrete soliton-like expressions, suggesting an underlying integrable structure.

The purpose of this paper is to study the statistics of trees with labels now belonging to a *finite set*, say $\{0, 1, \ldots, L\}$, which amounts to having two walls at positions -1 and $L + 1$ in the embedding target space. In the language of evolving populations, introducing walls gives access to the probability for the population to attain predefined boundaries or to remain confined within a predefined connected domain. The two-wall situation corresponds to the generic case where this domain is compact, i.e. is a segment (with two boundaries). Such boundary conditions correspond to the so-called restricted solid-on-solid (RSOS) version of the problem, in which heights are restricted to belong to a finite segment.

The paper is organized as follows. In section 2, we introduce a master equation for the partition function of labelled rooted trees and recall its solutions both without walls and with one wall. Section 3 is devoted to the derivation of the two-wall solution, involving elliptic functions. The corresponding continuum limit is derived in section 4 and expressed in terms of the Weierstrass $\wp$ function with constrained periods. As an application of these results we study in section 5 a particular stochastic process describing the evolution of a population which spreads in one dimension. We discuss in section 6 the solution of a slightly different problem corresponding to a dilute SOS version in which neighbouring vertices of the tree have labels differing by ±1 *or* 0. We gather a few concluding remarks in section 7, where we discuss the integrability of our models in particular in connection with planar graph enumeration and matrix models. The precise connection between labelled trees and planar graphs is further detailed in the appendix.

2. Enumeration of labelled rooted trees

Let R_n denote the generating function for labelled rooted planar trees with a root at position n. By a decomposition according to the possible local environments of the root vertex, characterized by the sequence of labels $n - 1$ or $n + 1$ of its adjacent vertices (see figure 2), we immediately get the equation

$$R_n = \frac{1}{1 - g(R_{n+1} + R_{n-1})}. \tag{2.1}$$

Note also that from their combinatorial definition as counting functions, the R_n are required to have a series expansion in powers of g starting as $R_n = 1 + O(g)$, a condition sufficient to determine all the R_n from equation (2.1). Relation (2.1) is valid for all accessible values of n and it may be supplemented by boundary conditions to account for the possible presence of walls. We choose these conditions among three categories: no wall, one wall and two walls, corresponding to restrictions on the allowed labels n, respectively, no restriction, $n \geqslant 0$ and $0 \leqslant n \leqslant L$.

In the absence of a wall, all the R_n are equal due to translational invariance, to a function R satisfying $R = 1/(1 - 2gR)$ and $R = 1 + O(g)$, namely

$$R = R(g) \equiv \frac{1 - \sqrt{1 - 8g}}{4g} = \sum_{N=0}^{\infty} g^N 2^N c_N. \tag{2.2}$$

This formula displays the critical value $g_c = 1/8$ of g, while the coefficient of g^N clearly counts rooted planar trees with N edges (c_N) with 2^N possible embeddings.

In the presence of one wall at position -1, we must write $R_{-1} = 0$ and consider equation (2.1) only for $n \geqslant 0$. This system may be solved order by order in g using as a seed the vanishing of all the coefficients of the series for R_{-1} and the order zero values $R_n = 1 + O(g)$ for all $n \geqslant 0$. In a more global way, the solution was worked out in [3] by replacing the condition of existence of a power series expansion for each R_n by the condition that $R_n \to R$ at large n with R as above. Indeed, it is clear from the combinatorial definition that $R - R_n = O(g^{n+1})$ as the wall at position -1 may not be reached with less than $(n + 1)$ edges from position n. The solution of equation (2.1) with these boundary conditions reads

$$R_n = R\frac{u_n u_{n+4}}{u_{n+1}u_{n+3}} \qquad u_n = x^{(n+1)/2} - x^{-(n+1)/2} \tag{2.3}$$

with $R = R(g)$ as in equation (2.2) and where x is the solution of

$$x + \frac{1}{x} + 2 = \frac{1}{gR} \tag{2.4}$$

with, say, modulus less than 1, namely $x = x(g) \equiv (1 - (1 - 8g)^{1/4})/(1 + (1 - 8g)^{1/4})$. Note that x is real for all $g \leqslant 1/8$ and admits a convergent series expansion in g with *positive integer* coefficients. The small g behaviour $x(g) = g + O(g^2)$ ensures the above property that $R_n = R + O(g^{n+1})$. Solution (2.3) is readily checked by noting that equation (2.1) translates into the following trilinear equation for the u:

$$u_n u_{n+2} u_{n+4} = \frac{1}{R} u_{n+1} u_{n+2} u_{n+3} + gR(u_{n-1}u_{n+3}u_{n+4} + u_n u_{n+1} u_{n+5}) \tag{2.5}$$

easily verified upon substituting $gR = x/(1+x)^2$ and $1/R = (1+x^2)/(1+x)^2$. The particular form of solution (2.3) was identified in [3] as a stationary one-soliton solution to the KP equation [4].

In the presence of two walls, we must write $R_{-1} = R_{L+1} = 0$ and consider equation (2.1) for $0 \leqslant n \leqslant L$. This system may again be solved order by order in g from the boundary conditions that all coefficients of R_{-1} and R_{L+1} vanish and $R_n = 1 + O(g)$ for all n, $0 \leqslant n \leqslant L$. For a finite value of L, one may also eliminate all but one R_n from the finite set of algebraic equations (2.1) so as to obtain an algebraic equation of degree $[(L + 3)/2]$ for each R_n, with a unique solution such that $R_n = 1 + O(g)$. In a more global way, we intend to generalize solution (2.3) to this two-wall case. This is done in the next section.

3. Two-wall solution

3.1. General solution via elliptic functions

From the one-soliton structure of solution (2.3), it is natural to look for a more general soliton-like solution to equation (2.1) in the 'elliptic' form

$$\begin{aligned} R_n &= R\frac{u_n u_{n+4}}{u_{n+1}u_{n+3}} \\ u_n &= (x^{(n+1)/2} - x^{-(n+1)/2})\prod_{j=1}^{\infty}(1-q^j x^{n+1})\left(1-\frac{q^j}{x^{n+1}}\right) \end{aligned} \tag{3.1}$$

with an additional free real parameter q such that $|q| < 1$. For $q = 0$, we recover solution (2.3) provided R and x are given by equations (2.2) and (2.4), depending on g only. As we shall now see, for a general q, we may tune R and x in equation (3.1) as functions of *both* g *and* q so as to satisfy equation (2.1). The above solution clearly satisfies $R_{-1} = 0$ and the value of the free parameter q may finally be adjusted so as to ensure $R_{L+1} = 0$.

To fix the functions R and x in terms of g and q, we again write equation (2.1) in the form of equation (2.5), and note that from the definition of u_n, we have $u_{-1} = 0$ and $u_{-2-k} = -u_k$. Hence, taking $n = -1$ in equation (2.5) we get $gR^2 = u_1/u_3$, while taking $n = -2$, we obtain $gR = (u_0/u_1)^2$, which generalizes equation (2.4). This leads to

$$g = \frac{u_0^4 u_3}{u_1^5} \tag{3.2}$$

which implicitly determines $x(g, q)$ as a function of g and q. More precisely, for a fixed q with $0 \leqslant q < 1$, this equation has four solutions for $0 \leqslant g \leqslant 1/8$, a real positive one $x_0(g, q)$ with, say, modulus less than 1 together with its inverse $x_2(g, q) = 1/x_0(g, q)$, and a solution on the unit circle $x_1(g, q)$ with, say, positive imaginary part together with its inverse $x_3(g, q) = 1/x_1(g, q)$. For a fixed q with $-1 < q \leqslant 0$, we have the same pattern of solutions provided $g \geqslant 1/8$ and does not exceed some upper bound. When $g = 1/8$, all solutions coalesce to $x = 1$ independently of q. For a definite value of the position L of the second wall, the proper choice of solution will be discussed below. In particular, the presence of two walls at a finite distance increases the radius of convergence of the R_n as series of g as it reduces the entropy of configurations. This requires exploring values of $g > 1/8$. Note that for $q = 0$, we have $x_0(g, 0) = x(g)$, corresponding to the solution of section 2, while $x_1(g, 0) = (1 + \mathrm{i}(1 - 8g)^{1/4})/(1 - \mathrm{i}(1 - 8g)^{1/4})$.

The function R is now given by

$$R = \frac{u_1^3}{u_0^2 u_3} \tag{3.3}$$

with two determinations $R^{(0)}(g, q)$ and $R^{(1)}(g, q)$ corresponding, respectively, to the substitutions $x = x_0(g, q)$ and $x = x_1(g, q)$ in the u_n, or equivalently $x_2(g, q)$ and $x_3(g, q)$, respectively, as the change $x \to 1/x$ amounts to $u_n \to -u_n$, leaving R and all R_n invariant. It remains to verify that, for these particular choices of x and R, the u_n of equation (3.1) actually satisfy the identity (2.5) for all n. Let us introduce the notation $q = \exp(2\mathrm{i}\pi\tau)$, $x^{n+1} = \exp(2\mathrm{i}\pi z)$ so that $u_n = \theta_1(z)$ where θ_1 is the (unnormalized) Jacobi theta function with nome q and argument z:

$$\theta_1(z) \equiv \theta(z|\tau) = 2\mathrm{i}\sin(\pi z)\prod_{j\geqslant 1}(1 - 2q^j\cos(2\pi z) + q^{2j}). \tag{3.4}$$

We also introduce the notation $x = \exp(2\mathrm{i}\pi\alpha)$, so that equation (2.5), together with the particular choices (3.2) and (3.3), translates into a theta function identity

$$\left(\frac{\theta_1(2\alpha)}{\theta_1(\alpha)}\right)^2 \theta_1(z)\theta_1(z+2\alpha)\theta_1(z+4\alpha) = \frac{\theta_1(4\alpha)}{\theta_1(2\alpha)}\theta_1(z+\alpha)\theta_1(z+2\alpha)\theta_1(z+3\alpha) + \theta_1(z-\alpha)\theta_1(z+3\alpha)\theta_1(z+4\alpha) + \theta_1(z)\theta_1(z+\alpha)\theta_1(z+5\alpha). \tag{3.5}$$

To prove it, we note that both sides have the same transformations when $z \to z+1$ and $z \to z+\tau$, due to the properties $\theta_1(z+1) = -\theta_1(z)$ and $\theta_1(z+\tau) = -q^{-\frac{1}{2}}\exp(-2\mathrm{i}\pi z)\theta_1(z)$. Taking the ratio of rhs/lhs of equation (3.5), we get an elliptic function, with poles possibly at $z = 0, -2\alpha, -4\alpha$ in a fundamental cell. One easily checks by examining the rhs at these values that the corresponding residues all vanish and therefore the ratio is a constant, easily shown to be 1 by taking for instance its value at $z = -\alpha$, which completes the proof of the identity. With this notation, we have the relations

$$g = \frac{\theta_1^4(\alpha)\theta_1(4\alpha)}{\theta_1^5(2\alpha)} \tag{3.6}$$

and

$$R = \frac{\theta_1(2\alpha)^3}{\theta_1(\alpha)^2\theta_1(4\alpha)} \tag{3.7}$$

while

$$R_n = \frac{\theta_1(2\alpha)^3}{\theta_1(\alpha)^2\theta_1(4\alpha)}\frac{\theta_1((n+1)\alpha)\theta_1((n+5)\alpha)}{\theta_1((n+2)\alpha)\theta_1((n+4)\alpha)} \tag{3.8}$$

with $\theta_1(z)$ as in equation (3.4). As this solution involves elliptic functions, it is natural to study its transformation under the modular transformation $\tau \to -1/\tau$. From the transformation $\theta_1(z/\tau|-1/\tau) = -\mathrm{i}(-\mathrm{i}\tau)^{1/2}\exp(\mathrm{i}\pi z^2/\tau)\theta_1(z|\tau)$, we find that the *physical quantities* such as g as given by equation (3.6) and R_n as given by equation (3.8) are invariant under $(\alpha,\tau) \to (\tilde\alpha,\tilde\tau) \equiv (\alpha/\tau, -1/\tau)$ or equivalently $(x,q) \to (\tilde x,\tilde q)$ with $\tilde x = x^{1/\tau}$ and $\tilde q = \exp(-2\mathrm{i}\pi/\tau)$. This is not the case for intermediate factors such as R or any of the u_n taken independently. We deduce from these properties that the *same* solution is reached by the parameters (x,q) and by their modular transforms $(\tilde x,\tilde q)$.

It is interesting to study the modular transformation of the solutions $x_i(g,q)$, $i = 0,1,2,3$ of equation (3.6) as introduced above. Clearly, for a fixed g, the modular invariance of the rhs of equation (3.6) implies that $\tilde x_i(g,q) = x_i(g,q)^{1/\tau}$ is a valid solution when $q \to \tilde q$, hence we may write $x_i(g,q)^{1/\tau} = x_{\sigma(i)}(g,\tilde q)$ for some permutation $\sigma \in S_4$. Iterating this transformation, and using $\tilde{\tilde q} = q$ while $\tau\tilde\tau = -1$, we deduce that $x_{\sigma^2(i)}(g,q) = x_i(g,q)^{-1}$, which brings us back to the same solution we started from, but with the determination $x_i(g,q)^{-1}$ instead of $x_i(g,q)$. Therefore σ is a *circular permutation* of the four indices.

For $q < 0$, the modular transform $\tilde q$ becomes complex, and we will make no use of the above remarks. On the other hand, for $0 < q < 1$, the modular parameter $\tau = \mathrm{i}t$ may be taken purely imaginary, and the modular transformation reduces to $t \to 1/t$ hence $\tilde q$ is also real and in the range $(0,1)$. It is then easy to see that the modular transformation sends the solution $x_0(g,q)$ to $x_1(g,\tilde q)$ and $x_1(g,q)$ to $x_2(g,\tilde q)$ which leads to the same R_n as $x_0(g,\tilde q)$. In other words, the same physical solution may be described by either some q and the real determination $x_0(g,q)$ or by its modular transform $\tilde q$ and the determination on the unit circle $x_1(g,\tilde q)$.

3.2. Boundary condition at $n = L + 1$

We now implement the two-wall boundary condition described above, in which we require $R_{L+1} = 0$, or equivalently $u_{L+5} = 0$. This is achieved by demanding that $x^{L+6} = q^m$ for some $m \in \mathbb{Z}$. From the above study, we have at our disposal x-solutions either real positive and smaller than 1, or on the unit circle with positive imaginary part (we discard the equivalent $1/x$-solution). This leaves us for *positive* q with two possibilities: (i) $x = \exp(2i\pi k/(L+6))$, $k = 1, 2, \ldots$, corresponding to $m = 0$ and (ii) $x = q^{m/(L+6)}$, $m = 1, 2, \ldots$ while at *negative* q, only solution (i) survives. As the solution R_n must be positive by definition for $n = 0, 1, \ldots, L$, it is easily verified that only $k = 1$ in case (i) and $m = 1$ in case (ii) are admissible to prevent sign changes for the u_n and the R_n. In a more physical language, higher values of k or m correspond to higher modes with oscillations in the range $[0, L]$. Taking for instance $m = 2$ corresponds to a first vanishing of R_n at the coordinate $n = L/2 - 2$.

For any given L, we may always pick solution (i) and define $q_L(g)$ as the unique real solution of

$$x_1(g, q_L(g)) = \exp(2i\pi/(L+6)). \tag{3.9}$$

Introducing the notation $L(g)$ for the value of L such that $x_1(g, q = 0) = \exp(2i\pi/(L+6))$, namely

$$L(g) = \frac{\pi}{\arctan((1-8g)^{1/4})} - 6 \tag{3.10}$$

we have that $q_L(g) > 0$ for $L > L(g)$ and $q_L(g) < 0$ for $L < L(g)$, also valid for $g > 1/8$ with the convention that $L(g) = \infty$ in this range. The length $L(g)$ may be taken as a measure of the typical extent in the embedding space of the random trees at a fixed value of g, in the absence of walls.

On physical grounds, we expect a qualitative change of behaviour to occur at wall distances L of the order of $L(g)$. For $L \gg L(g)$, the tree behaves as a compact object of typical extent $L(g)$ 'diffusing' between the walls and feeling them only when approaching at distances of the order of $L(g)$ and smaller. For $L \ll L(g)$, the walls strongly squeeze the tree and L is the only relevant scale of the problem. In the first regime, the *profile* $\{R_n\}_{0 \leqslant n \leqslant L}$, always maximal in the middle, is mainly constant and decreases significantly only at distances of order $L(g)$ from the walls. In the second regime, the profile will vary over the whole range between the walls. Alternatively, for a fixed value of L, we may invert these conditions into $q_L(g) > 0$ for $g < g_L$ and $q_L(g) < 0$ for $g > g_L$, where

$$g_L = \frac{1}{8}\left(1 - \tan^4\left(\frac{\pi}{L+6}\right)\right). \tag{3.11}$$

Note that $0 < g_L < 1/8$ for all $L = 0, 1, 2, \ldots$.

Besides solution (3.9), we have another possibility for the choice of q, namely, $q = q'_L(g)$, which solves condition (ii)

$$x_0(g, q'_L(g)) = q'_L(g)^{1/(L+6)}. \tag{3.12}$$

This alternative solution exists for all $L > L(g)$ and is *a priori* distinct from the previous solution (3.9). In practice, the *physical solution* for fixed g and L is *unique* as one easily checks that solution (3.9) is the modular transform of (3.12), namely $q_L(g) = \tilde{q'}_L(g)$ and $x_1(g, q_L(g)) = x_0(g, q'_L(g))^{1/\tau'_L(g)}$ with $q'_L(g) = \exp(2i\pi\,\tau'_L(g))$. The two choices (3.9) or (3.12) therefore lead to the same physical quantities R_n.

To obtain the combinatorial series expansions for the R_n in g it is simpler to work with solution (3.12). This is always possible as g is small (hence we may work in the regime $g < g_L$). Moreover, we have $x_0(g, q'_L(g)) = g + O(g^2)$ while $q'_L(g) = g^{L+6}(1 + O(g))$. This

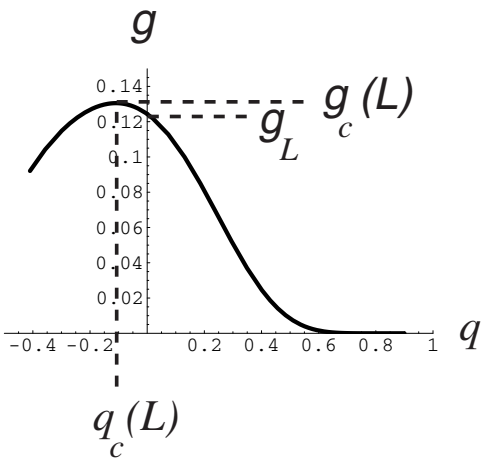

Figure 3. Plot of $g(q)$ as given by equation (3.13) for $L = 6$. The function decreases from a maximum value $g_c(L)$ at some negative value $q_c(L)$ down to 0 at $q = 1$. At $q = 0$, we have $g = g_L$ as in equation (3.11). We have $g_L < \frac{1}{8} < g_c(L)$.

shows in particular that the present R_n and the former $R(g)$ of equation (2.2) have the same expansion up to order $\min(n+1, L-n+1)$ in g, as expected.

On the other hand, to have a more global approach it is best to work with solution (3.9), which is valid for any g. We may then use relations (3.6)–(3.8) with $\alpha = 1/(L+6)$ to parametrize the solution with q as follows:

$$\begin{aligned} g(q) &= \frac{\theta_1^4\left(\frac{1}{L+6}\right)\theta_1\left(\frac{4}{L+6}\right)}{\theta_1^5\left(\frac{2}{L+6}\right)} \\ R(q) &= \frac{\theta_1^3\left(\frac{2}{L+6}\right)}{\theta_1^2\left(\frac{1}{L+6}\right)\theta_1\left(\frac{4}{L+6}\right)} \\ R_n(q) &= \frac{\theta_1^3\left(\frac{2}{L+6}\right)}{\theta_1^2\left(\frac{1}{L+6}\right)\theta_1\left(\frac{4}{L+6}\right)}\frac{\theta_1\left(\frac{n+1}{L+6}\right)\theta_1\left(\frac{n+5}{L+6}\right)}{\theta_1\left(\frac{n+2}{L+6}\right)\theta_1\left(\frac{n+4}{L+6}\right)} \end{aligned} \tag{3.13}$$

with θ_1 as in equation (3.4) . This form displays clearly the symmetry $R_n = R_{L-n}$ expected from the symmetry of the problem, as a consequence of the relation $\theta_1(1-z) = \theta_1(z)$.

For fixed L, the function $g(q)$ starts from $g = 0$ at $q = 1$ and increases as q decreases, passing through g_L of equation (3.11) at $q = 0$, and reaching a maximum value $g_c(L) > 1/8$ attained at some negative value of $q = q_c(L)$ (see figure 3 for illustration). The quantity $g_c(L)$ is nothing but the radius of convergence of all the series in g appearing in the problem, and governs the leading growth of the number of configurations as a function of the number N of edges in the tree as $g_c(L)^{-N}$. We have, for instance, $g_c(1) = 1/4$, $g_c(2) = 3 - 2\sqrt{2}$, $g_c(3) = 4/27$. We also see that $g_c(L) \to 1/8$ when $L \to \infty$. The branch of the solution for $q < q_c(L)$ is discarded as unphysical.

To conclude this section, let us use solution (3.13) to display for fixed L the exact profile $\{R_n\}_{0 \leqslant n \leqslant L}$ for some particular values of q, namely

(a) a positive value of q realizing $g < g_L$, in which case the profile is flat except for a region distant by typically $L(g)$ from the walls;
(b) the value $q = 0$ where $g = g_L$ and all theta functions degenerate into trigonometric functions;

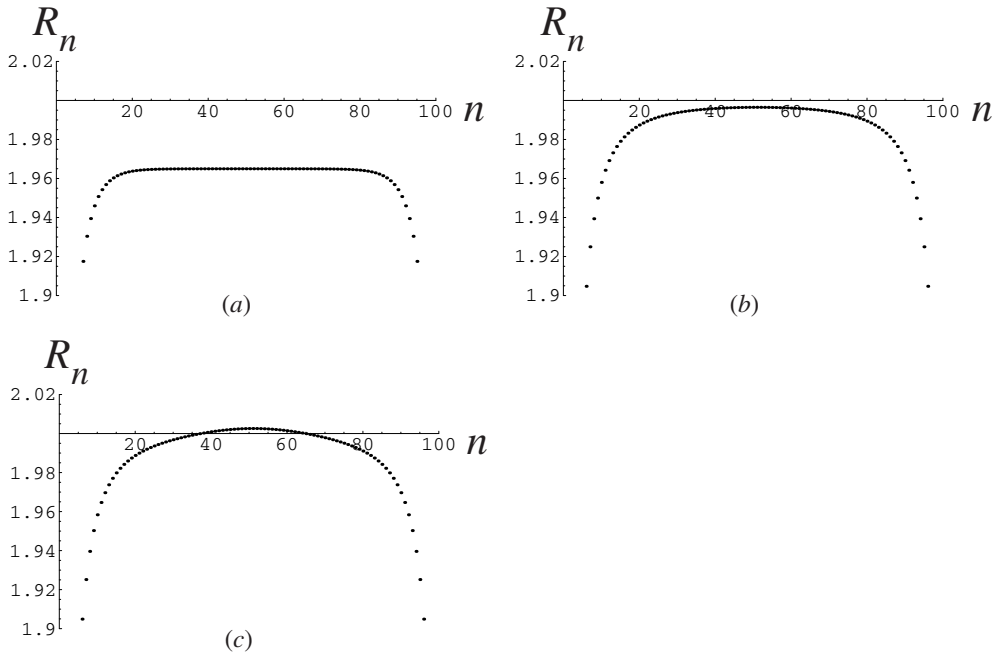

Figure 4. Typical profiles $\{R_n\}_{0\leqslant n\leqslant L}$ for $L = 100$ as obtained from equation (3.13) for three particular values of q: (*a*) a positive value realizing $g \leqslant g_L$ where the profile is flat except for a region of extent $L(g)$ from the walls (here $q = 0.25$ and $L(g) \simeq 18$); (*b*) $q = 0\,(g = g_L)$ and (*c*) $q = q_c(L)\,(g = g_c(L))$ where the profile is slightly peaked in the middle.

(c) the negative value $q = q_c(L)$ corresponding to $g = g_c(L)$, where the profile varies over the whole range $[0, L]$ and is slightly peaked in the middle.

These profiles are represented in figure 4 for $L = 100$.

4. Continuum limit

To begin with, let us derive the continuum limit of the one-wall solution (2.3). It is reached by letting $g \to 1/8$ and $n \to \infty$ simultaneously as

$$g = \frac{1}{8}(1 - \epsilon^4) \qquad n = \frac{r}{\epsilon} \tag{4.1}$$

with $\epsilon \to 0$ playing the role of the inverse of a correlation length. This leads to

$$x(g) = \frac{1-\epsilon}{1+\epsilon} \qquad R(g) = \frac{2}{1+\epsilon^2} \tag{4.2}$$

and finally, expanding R_n up to order 2 in ϵ:

$$R_n = 2(1 - \epsilon^2 \mathcal{U}(r)) \tag{4.3}$$

where $\mathcal{U}$ is the scaling function

$$\mathcal{U}(r) = 1 + \frac{3}{\sinh^2(r)} \tag{4.4}$$

describing the 'repulsive potential' felt by the tree as a function of the rescaled distance r from the wall at $r = 0$.

In the presence of the second wall, we let in addition $L \to \infty$ by keeping the quantity $\omega = (L+6)\epsilon/2$ fixed. This ratio of the two characteristic lengths of the problem, namely L and the typical extent of the unconstrained tree $L(g) \sim \pi/\epsilon$, is the only physical scale surviving the continuum limit. Expanding the rhs of the first line of equation (3.13) at small ϵ up to order 4, we get

$$g = \frac{1}{8}\left(1 - \frac{\epsilon^4}{\omega^4}\left(\frac{5}{32}\left(\frac{\theta_1'''(0)}{\theta_1'(0)}\right)^2 - \frac{3}{32}\frac{\theta_1^{(5)}(0)}{\theta_1'(0)}\right)\right) \tag{4.5}$$

to be identified with $g = 1/8(1-\epsilon^4)$ as in equation (4.1). This fixes implicitly the value of τ as a function of ω by

$$\omega^4 = \frac{5}{32}\left(\frac{\theta_1'''(0)}{\theta_1'(0)}\right)^2 - \frac{3}{32}\frac{\theta_1^{(5)}(0)}{\theta_1'(0)}. \tag{4.6}$$

We may now expand R and R_n as given by equation (3.13) up to order 2 in ϵ to obtain the relevant scaling function

$$\begin{aligned} R &= 2\left(1 - \frac{\epsilon^2}{4\omega^2}\frac{\theta_1'''(0)}{\theta_1'(0)}\right) \\ R_n &= 2(1-\epsilon^2\mathcal{U}(r)) \end{aligned} \tag{4.7}$$

with r as in equation (4.1) and where the scaling function $\mathcal{U}(r)$ is now given by

$$\mathcal{U}(r) = \frac{\theta_1'''(0)}{4\omega^2\theta_1'(0)} - 3\frac{\mathrm{d}^2}{\mathrm{d}r^2}\log\theta_1\left(\frac{r}{2\omega}\right) = 3\wp(r) \tag{4.8}$$

where we have identified the Weierstrass $\wp$ function [5] with half-periods ω and $\omega' = \tau\omega$ where τ is fixed by equation (4.6), which amounts to

$$g_2(\omega, \omega') = \tfrac{4}{3} \tag{4.9}$$

where g_2 is the first elliptic invariant of the Weierstrass function. The scaling function $\mathcal{U}(r)$ may be viewed as the potential felt by the random tree in the presence of the two walls.

When sending the second wall to infinity, i.e. taking $\omega \to \infty$, we immediately recover the above result (4.4) as $g_2 = 4/3$ fixes the second half-period $\omega' = \mathrm{i}\pi/2$, in which case the scaling function degenerates into (4.4). Taking $L = L(g)$, i.e. $\omega = \pi/2$, fixes $\omega' = \mathrm{i}\infty$ which corresponds to $q = 0$, in which case the scaling function degenerates into a trigonometric function

$$\mathcal{U}(r) = \frac{3}{\sin^2(r)} - 1. \tag{4.10}$$

The values of $\omega \in [\pi/2, \infty](L \geqslant L(g))$ correspond to taking $q \in [0,1]$, while those in $[0, \pi/2](L \leqslant L(g))$ are obtained for $q \in [-q_*, 0]$, where $q_* = -\lim_{L\to\infty} q_c(L) = \exp(-\pi\sqrt{3})$.

Note that we could have derived the scaling limit of R_n without using the explicit solution (3.13) by plugging the ansatz $R_n = 2(1-\epsilon^2\mathcal{U}(r))$ into the original equation (2.1) and expanding up to order 4 in ϵ to obtain a differential equation for $\mathcal{U}$, namely

$$\mathcal{U}'' = 2(\mathcal{U}^2 - 1) \tag{4.11}$$

and require that $\mathcal{U}(r)$ diverge at $r = 0$ and $r = 2\omega$, with no divergence in-between. This equation is to be compared with that satisfied by $\wp$, namely $\wp'' = 6\wp^2 - g_2/2$, which allows us to identify $\mathcal{U}(r) = 3\wp(r)$ provided $g_2 = 4/3$.

5. A simple application: escape probability from a fixed domain for a spreading population

As mentioned in the introduction, rooted planar trees may be used to model discrete branching evolution processes. Let us assume, for instance, that an initial parent individual (materialized by the root vertex) gives rise in one generation to a number $k \geqslant 0$ of children individuals with probability p_k, each child itself independently giving rise to subsequent generations with the same probabilities. The resulting genealogical tree is nothing but a *planar* tree (without labels). In the following, we concentrate on the most natural choice $p_k = (1-p)p^k$, with $0 \leqslant p \leqslant 1$, and where the prefactor $(1-p)$ ensures the correct normalization and may be interpreted as the probability of death without descendents. This choice is expected to capture all possible physics of the problem, as the parameter p allows us to explore all possible values of the average number of children $p/(1-p)$, known to be the only relevant quantity in the problem [6].

We may now turn the branching process into a *spatial* branching process by allowing the individuals to spread in a one-dimensional target space. More precisely, we consider a discrete version of the problem in which the individuals may occupy integer-valued positions, with the diffusion rule that each child lives at a position differing by ± 1 from that of its parent, with an equal probability of $1/2$. Using these positions as vertex labels, we may write the following master equation for the extinction probability of a family:

$$E_n(T) = \frac{1-p}{1-\frac{1}{2}p(E_{n+1}(T-1)+E_{n-1}(T-1))} \tag{5.1}$$

where $E_n(T)$ stands for the probability that an individual at position n has no more descendents at generation T, with the initial condition $E_n(0) = 0$. Equation (5.1) is obtained by enumerating all possible configurations of the first generation children, and noting that the joint probability of extinction of all their descendents is the product of individual extinction probabilities before generation $T-1$. Comparing equation (5.1) to equation (2.1), we see that $E_n \equiv \lim_{T\to\infty} E_n(T)$ obeys the same equation (2.1) as $(1-p)R_n$, provided we set

$$g = \frac{p(1-p)}{2}. \tag{5.2}$$

This allows us to identify $E_n = (1-p)R_n$ by noting that this choice corresponds precisely to the stable fixed point of the recursion relation (5.1). We may therefore interpret $(1-p)R_n$ as the probability of extinction of a family whose first generation's parent is at position n. Accordingly, we interpret $1-(1-p)R_n$ as the survival probability for such a family.

For unconstrained positions (no wall in the former language), we simply have a translation invariant probability of survival

$$S(p) \equiv \frac{2p-1+|2p-1|}{2p} = \begin{cases} 0 & p \in \left[0, \frac{1}{2}\right] \\ \dfrac{2p-1}{p} & p \in \left[\frac{1}{2}, 1\right] \end{cases} \tag{5.3}$$

obtained by substituting $g = p(1-p)/2$ into equation (2.2). This result is totally insensitive to the diffusion process, and is the same as that for unlabelled trees. It displays a first-order singularity (discontinuous derivative) at $p = p_c = \frac{1}{2}$. This classical result in the theory of branching processes [6] is a particular case of a more general statement for so-called Galton–Watson processes that the genealogical tree is almost surely finite (here $S(p) = 0$) if and only if the average number of children is less than or equal to one (here corresponding to $\sum j p_j = p/(1-p) \leqslant 1$, namely $p \leqslant 1/2$).

Apart from the extinction probability $E_n(T)$, one natural quantity to study is the probability $C_n(T)$ for the population spreading from a position n to remain *confined* within a given

connected domain $\mathcal{D}$ until generation T. This quantity is readily seen to obey the same equation (5.1) as $E_n(T)$ for all n in $\mathcal{D}$, but with a different initial value at $T = 0$, namely $C_n(0) = 1$, and also the condition that $C_n(T) = 0$ for all n outside $\mathcal{D}$. We distinguish the two possible cases where (i) $\mathcal{D}$ is a half-line, say $[0, \infty)$ or (ii) $\mathcal{D}$ is a segment, say $[0, L]$. The only relevant conditions are at the boundary of the domain and read respectively (i) $C_{-1}(T) = 0$ and (ii) $C_{-1}(T) = C_{L+1}(T) = 0$. In the limit when $T \to \infty$, we may therefore identify $C_n(\infty) = (1-p)R_n$ with R_n given by our (i) one-wall and (ii) two-wall solutions of sections 2 and 3. The results will be best expressed in terms of the quantity $S_n = 1 - C_n(T = \infty)$ which is nothing but the probability for the population to escape from the domain.

In the one-wall case, solution (2.3) leads to the population's escape probability

$$S_n = 1 - \frac{1-|2p-1|}{2p}\frac{(1-x^{n+1})(1-x^{n+5})}{(1-x^{n+2})(1-x^{n+4})} \tag{5.4}$$

where

$$x = x(p) \equiv \frac{1-|2p-1|^{1/2}}{1+|2p-1|^{1/2}}. \tag{5.5}$$

Note that for any fixed $n \geqslant 0$ the escape probability S_n is strictly positive as soon as $p > 0$, and moreover that it still displays a singularity at $p = p_c = 1/2$ but of weaker *third-order* type (discontinuity of the third derivative) as is readily seen by expanding S_n up to order 3 in powers of $|2p-1|$. Note also that $\lim_{n\to\infty} S_n = S(p)$ as in equation (5.3), expressing the equivalence in probability between surviving forever and reaching infinitely distant points. Note finally the following simple expression for the escape probability S_n at the transition point $p = \frac{1}{2}$:

$$S_n\left(p = \frac{1}{2}\right) = \frac{3}{(n+2)(n+4)} \qquad n \geqslant 0. \tag{5.6}$$

The scaling limit (4.1) may be used to study the vicinity of the transition point by setting $p = p_c(1+\eta\epsilon^2)$ with $\eta = \pm 1$ according to whether we approach the transition from above or below. Equation (4.3) allows us to interpret the scaling function $\mathcal{U}(r)$ as describing the scaling behaviour of the escape probability $S_n \sim \mathcal{S}_n$ around $p = \frac{1}{2}$, with

$$\mathcal{S}_n = \epsilon^2(\mathcal{U}(n\epsilon) + \eta) = |2p-1|\left(\frac{3}{\sinh^2(n|2p-1|^{1/2})} + 1\right) + (2p-1) \tag{5.7}$$

valid in the scaling region of large n and $n\epsilon = O(1)$. This scaling function displays clearly the above-mentioned third-order transition with

$$\mathcal{S}_n = \frac{3}{n^2} + (2p-1) + \frac{n^2}{5}(2p-1)^2 - \frac{2n^4}{63}|2p-1|^3 + O((2p-1)^4 n^6). \tag{5.8}$$

We have represented in figure 5 the escape probabilities $S_n(p)$ and their limit $S(p)$ as functions of $p \in [0, 1]$. We have also blown up the critical region around $p = 1/2$ to compare the exact solution with its scaling limit.

In the case of two walls, solution (3.13) leads directly to the probability $S_n = 1-(1-p)R_n$ of escaping from the interval $[0, L]$.

For illustration, we have displayed in figure 6 the increase in the escape probability $S_n^{(L)}(p) - S_n(p)$ from the one-wall situation to that with two walls for $L = 5$ and $n = 0, 1, 2$, as functions of $p \in [0, 1]$. The increase is maximal at $p = \frac{1}{2}$.

For finite L, the critical value $g_c(L) > 1/8$ is never attained as $g = p(1-p)/2 \in [0, 1/8]$ for $p \in [0, 1]$, hence the singularity at $p = \frac{1}{2}$ is suppressed. However it is restored in the scaling limit where $L \to \infty$ with $L|2p-1|^{1/2}$ fixed, as $g_c(L) \to 1/8$ in this case. The

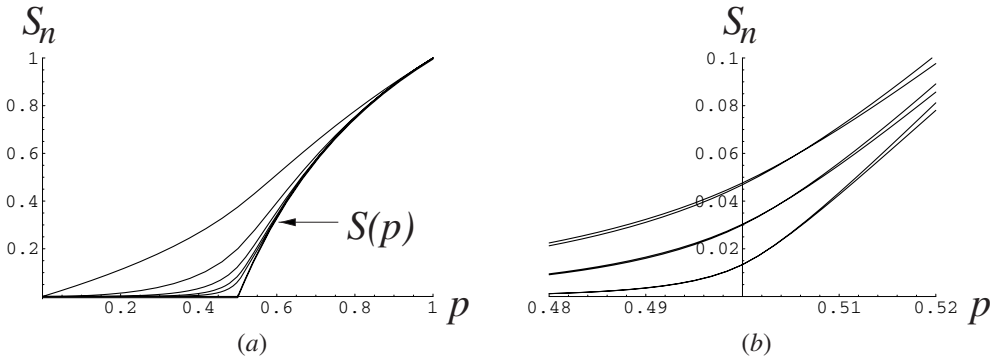

Figure 5. Escape probabilities S_n of equation (5.4) as functions of p in the presence of one wall: (*a*) for $p \in [0, 1]$ and from top to bottom $n = 0, 1, 2, 3, 4$ as well as $n = \infty$ in which case we recover the no-wall solution $S(p)$ of equation (5.3); (*b*) for p in the critical region $p \sim \frac{1}{2}$ and from top to bottom $n = 5, 7, 12$, together with the expected scaling limits $\mathcal{S}_n$ as given by equation (5.7) with a proper shift of $n \to n + 3$ ensuring the perfect matching of the curves.

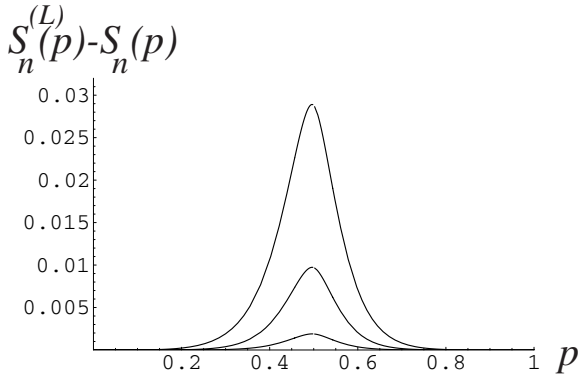

Figure 6. Plot of the increase in the escape probability $S_n^{(L)}(p) - S_n(p)$ from the one-wall situation to that with two walls for $L = 5$ and $n = 0, 1, 2$ (bottom to top).

scaling function $\mathcal{U}(r)$ given by equation (4.8) again describes the scaling behaviour of the escape probability $S_n \sim \mathcal{S}_n$ in the vicinity of $p = \frac{1}{2}$, with the result

$$\mathcal{S}_n = 3|2p-1|\wp(n|2p-1|^{1/2}) + (2p-1) \tag{5.9}$$

where the Weierstrass $\wp$ function must be taken with fixed half-periods $\omega = L|2p-1|^{1/2}/2$ and $\omega' = \tau\omega$, such that $g_2(\omega, \omega') = \frac{4}{3}$. We again note that $\mathcal{S}_n$ displays a third-order singularity at $p = \frac{1}{2}$ by expanding

$$\mathcal{S}_n = \frac{3}{n^2} + (2p-1) + \frac{3}{20}g_2(\omega,\omega')n^2(2p-1)^2 + \frac{3}{28}g_3(\omega,\omega')n^4|2p-1|^3 + O((2p-1)^4 n^6) \tag{5.10}$$

where g_2 and g_3 stand for the elliptic invariants of the Weierstrass function and with the constraint that $g_2(\omega, \omega') = 4/3$ which fixes ω' and consequently g_3 as functions of $\omega = L|2p-1|^{1/2}/2$. Note that the singularity of $\mathcal{S}_n$ at $p = 1/2$ disappears at the modular invariant point $\tau = \mathrm{i}$, i.e. $q = \exp -2\pi$, where $g_3(\omega, \mathrm{i}\omega) = 0$, which causes all odd powers of $|2p-1|$ to vanish in the series expansion (5.10).

Of particular simplicity is the case when $q = 0$ in equation (3.13), corresponding to $g = g_L$ as in equation (3.11), i.e. $p = (1 - \tan^2(\pi/(L+6)))/2$ or $p = (1 + \tan^2(\pi/(L+6)))/2$. The formula for the associated escape probability S_n reads

$$S_n = \begin{cases} \dfrac{1}{\cos\left(\frac{2\pi}{L+6}\right)} \left(\dfrac{\sin\left(\frac{\pi}{L+6}\right)\sin\left(\frac{3\pi}{L+6}\right)}{\sin\left(\pi\frac{n+2}{L+6}\right)\sin\left(\pi\frac{n+4}{L+6}\right)} - 2\sin^2\left(\frac{\pi}{L+6}\right) \right) & p = \frac{1}{2}\left(1 - \tan^2\left(\frac{\pi}{L+6}\right)\right) \\ \dfrac{\sin\left(\frac{\pi}{L+6}\right)\sin\left(\frac{3\pi}{L+6}\right)}{\sin\left(\pi\frac{n+2}{L+6}\right)\sin\left(\pi\frac{n+4}{L+6}\right)} & p = \frac{1}{2}\left(1 + \tan^2\left(\frac{\pi}{L+6}\right)\right). \end{cases} \tag{5.11}$$

This allows for framing the exact value of S_n at $p = 1/2$ between these two values for all L. For large L, we may identify $\omega = \lim_{L\to\infty} L|2p-1|^{1/2}/2 = \pi/2$ for both values of $p = (1 \pm \tan^2(\pi/(L+6)))/2$, and $\tau = \mathrm{i}\infty$ to ensure $q = 0$, in which case the scaling function reduces to (4.10) and therefore equation (5.9) turns into

$$\mathcal{S}_n = |2p-1| \left(\frac{3}{\sin^2(n|2p-1|^{1/2})} - 1 \right) + (2p-1) \tag{5.12}$$

in agreement with the large L limit of equation (5.11) .

6. Another solvable case of tree embedding: dilute SOS model on a random tree

We may consider a slightly different version of labelled trees, in which we impose the weaker constraint that any two adjacent vertices of the tree must have labels differing by ± 1 *or* 0. As shown in [2], this version of labelled trees is that involved in the enumeration of tetravalent planar graphs. In the language of spreading of a population, a child may now stay at the same position as its parent. This corresponds to a dilute SOS version of the case studied in this paper.

The main recursion relation is now replaced by

$$R_n = \frac{1}{1 - g(R_{n+1} + R_n + R_{n-1})} \tag{6.1}$$

where R_n is the generating function for rooted trees with root vertex labelled by $n \in \mathbb{Z}$, and a weight g per edge. Again, we may consider three types of boundaries: no wall, one wall and two walls. The no-wall case is easily solved, with all the R_n equal to the solution of $R = 1/(1 - 3gR)$ with $R = 1 + O(g)$, namely

$$R = R(g) \equiv \frac{1 - \sqrt{1 - 12g}}{6g}. \tag{6.2}$$

The one-wall case corresponds to setting $R_{-1} = 0$ and only considering the R_n for $n \geqslant 0$. It was solved in [3], with the result

$$R_n = R\frac{u_n u_{n+3}}{u_{n+1}u_{n+2}} \qquad u_n = x^{(n+1)/2} - x^{-(n+1)/2} \tag{6.3}$$

with $R = R(g)$ of equation (6.2), and where x is the solution of $x + 1/x + 4 = 1/(gR)$ with, say, modulus less than 1. Note the slight difference in the index shifts when compared with equation (2.3). The main recursion relation reduces this time to a quartic equation for the u_n:

$$u_n u_{n+1} u_{n+2} u_{n+3} = \frac{1}{R} u_{n+1}^2 u_{n+2}^2 + gR\left(u_{n-1}u_{n+2}^2 u_{n+3} + u_n^2 u_{n+3}^2 + u_n u_{n+1}^2 u_{n+4}\right) \tag{6.4}$$

supplemented by the initial condition $u_{-1} = 0$. Equation (6.4) is easily checked for all k by setting $1/R = (x^2 + x + 1)/(x^2 + 4x + 1)$ and $gR = x/(x^2 + 4x + 1)$. Finally, in the two-wall

case where we require $R_{-1} = R_{L+1} = 0$ and only consider $n = 0, 1, 2, \ldots, L$, we have found the elliptic solution

$$\begin{aligned} R_n &= R\frac{u_n u_{n+3}}{u_{n+1}u_{n+2}} \\ u_n &= \theta_1((n+1)\alpha) \end{aligned} \tag{6.5}$$

with $x = \exp(2\mathrm{i}\pi\alpha)$ and θ_1 as in equation (3.4), which solves equation (6.4) provided we take

$$\begin{aligned} R &= 4\frac{\theta_1(\alpha)\theta_1(2\alpha)}{\theta_1'(0)\theta_1(3\alpha)}\left(\frac{\theta_1'(\alpha)}{\theta_1(\alpha)} - \frac{1}{2}\frac{\theta_1'(2\alpha)}{\theta_1(2\alpha)}\right) \\ g &= \frac{\theta_1'(0)^2\theta_1(3\alpha)}{16\theta_1(\alpha)^2\theta_1(2\alpha)\left(\frac{\theta_1'(\alpha)}{\theta_1(\alpha)} - \frac{1}{2}\frac{\theta_1'(2\alpha)}{\theta_1(2\alpha)}\right)^2}. \end{aligned} \tag{6.6}$$

The boundary conditions are again satisfied for two choices of the parameter x: (i) $x = \exp(2\mathrm{i}\pi/(L+5))$ and (ii) $x = q^{1/(L+5)}$ (when $q \geqslant 0$), the latter leading to the same physical solution as the former by modular invariance. Picking again the first solution, we must take $\alpha = 1/(L+5)$, and may view the equations for the solution as parametrized by q. This leads to the continuum limit, upon taking $g = (1-\epsilon^4)/12$, $n = r/\epsilon$, $2\omega = (L+5)\epsilon$ and $R_n = 2(1-\epsilon^2\mathcal{U}(r))$. We end up with a scaling function $\mathcal{U}(r) = 2\wp(r)$ in terms of the Weierstrass $\wp$ function with half-periods ω as above and ω' fixed by now requiring that $g_2(\omega, \omega') = 3$. We also recover the one-wall case by taking the limit $L \to \infty$, namely $\omega = \infty$ while $\omega' = \mathrm{i}\pi/\sqrt{6}$, in which case $\mathcal{U}(r) = 1 + 3/\sinh^2(\sqrt{3/2}r)$, a result already obtained in [3].

Note finally that all scaling functions coincide with those of section 4 up to a global rescaling $r \to \sqrt{\frac{3}{2}}r$ and $\omega \to \sqrt{\frac{3}{2}}\omega$. This confirms the expected universality of the continuum limit.

7. Conclusion

In this paper, we have extensively studied a model of random rooted planar trees embedded in a discrete one-dimensional target space. In particular, we have derived explicit expressions for the partition function of the model with various target spaces, namely the whole integer line, a half-line and a segment. To obtain these, we have shown that the partition functions actually obey recursion relations and that the particular target at hand translates into various boundary conditions. We have also derived the corresponding scaling functions in the continuum limit for which the recursion relations turn into differential equations. A different approach, popular among probabilists, consists in studying directly the continuum limit of embedded random trees in the form of continuum spatial branching processes [1], giving rise to partial differential equations. It should be possible in this context to solve these equations with wall-type boundary conditions, and to recover our continuum results. Some work in this direction appeared recently [10], where an analogue of our one-wall case can be found.

The striking simplicity of solutions (2.3), (3.1) as well as (6.3), (6.5) is directly linked to the 'integrability' of the corresponding nonlinear recursion relations (2.1) and (6.1), respectively. One possible explanation for the integrability uses the interpretation of labelled trees in the context of planar graph enumeration. More precisely, as shown in detail in the appendix, the SOS and dilute SOS models on trees, respectively, occur in the enumeration of rooted planar Eulerian triangulations (i.e. triangulations with bicoloured faces) and of rooted planar quadrangulations. In both cases, the labels n correspond to geodesic distances along the graph from the root. This reformulation suggests a possible connection with matrix models,

known to be integrable. As already hinted in [3] in the context of graph enumeration, equations (2.1) and (6.1) are very similar to those arising in the context of the matrix models used for generating Eulerian triangulations and quadrangulations, respectively, namely $R_n = (n/N)/(1 - g(R_{n+1} + R_{n-1}))$ and $R_n = (n/N)/(1 - g(R_{n+1} + R_n + R_{n-1}))$, where N is the size of the matrices. The remarkable point is that the index n is no longer related to the geodesic distance along the graphs, but to their *genus*. Soliton theory seems to indicate that integrability survives when changing the recursion into $R_n = (\alpha+\beta n)/(1-g(R_{n+1}+R_{n-1}))$ or $R_n = (\alpha+\beta n)/(1-g(R_{n+1}+R_n+R_{n-1}))$ for some constants α and β, which could interpolate between the two problems.

Beyond the two (quadratic) examples of this paper, we have at our disposal a host of nonlinear recursion relations all used for enumerating possibly decorated planar graphs while keeping track of some geodesic distances, and which were found to be integrable as well (see [3] for details). The solutions display some multicritical behaviour corresponding to higher order critical points, with non-trivial hierarchies of scaling functions. It would be interesting to find some proper interpretation of these equations in the context of embedded trees or alternatively of population-spreading processes.

It would also be interesting to classify the target spaces leading to integrable models of embedded trees, or alternatively to spot among all possible discrete spatial branching processes those with integrability properties. We may then hope, by changing the nature of the discrete target space, to reach new critical points.

Appendix. Graph interpretation

As mentioned above, well-labelled trees, i.e. trees with non-negative labels corresponding to the one-wall situation in the dilute SOS version of section 6, were introduced in [2] in the context of graph enumeration. More precisely, it was shown that there exists a bijection between these well-labelled rooted planar trees and rooted planar quadrangulations. This allows us to interpret the quantity R_n of equation (6.3) as the generating function for planar quadrangulations with both a marked (origin) vertex and a marked oriented edge linking a vertex with geodesic distance m from the origin to a vertex with geodesic distance $m+1$ from the origin with $m \leqslant n$, and with an activity g per vertex. Besides this equivalence, there exists yet another bijection, now in the dual language, between rooted planar tetravalent graphs (dual to the above quadrangulations) and decorated (so-called blossom) binary trees [7]. This bijection was extended so as to keep track of geodesic distances between faces in [3]. In this language, the generating function Z_n for two-leg planar tetravalent graphs with the two legs distant by at most n was shown to obey the recursion relation $Z_n = 1+gZ_n(Z_{n+1}+Z_n+Z_{n-1})$, with $Z_{-1} = 0$, a direct consequence of the above bijection with blossom binary trees. This equation is nothing but yet another form of equation (6.1) and allows us to identify $Z_n = R_n$ of equation (6.3).

Remarkably, our slightly simpler equation (2.1) also admits two analogous interpretations in terms of graphs, now related to the enumeration of rooted planar Eulerian triangulations (i.e. triangulations with bicoloured faces, say in black and white) or dually to that of rooted trivalent bipartite planar graphs (say with black and white vertices). In this dual language, we may rely on a bijection [8] between rooted trivalent bipartite planar graphs and properly decorated binary trees. Keeping track of the graph-geodesic distances in these binary trees leads to the equation $Z_n = 1 + gZ_n(Z_{n+1} + Z_{n-1})$ for the generating function Z_n of two-leg trivalent bicoloured planar graphs with the two legs attached to vertices of opposite colours and at a geodesic distance of at most n. In this approach, the proper definition of the geodesic distance on the graphs makes use of oriented paths linking faces, with the constraint that

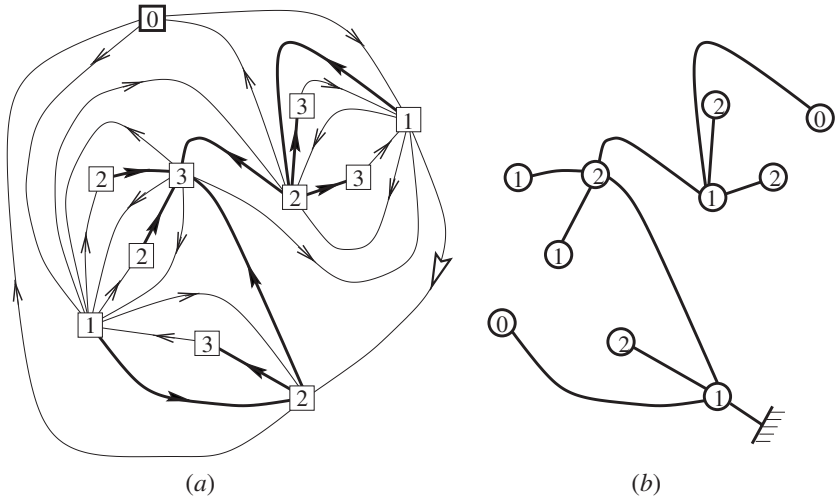

Figure 7. A sample Eulerian triangulation (*a*) with a marked origin vertex (labelled 0) and a marked oriented edge (big empty arrow). We have indicated for each vertex its geodesic distance from the origin (respecting the edge orientations). A labelled rooted tree (*b*) is obtained by retaining for each clockwise-oriented triangle the edge connecting the two farthest vertices from the origin, and picking as the root the end of the previously marked edge. The vertex labels on the tree are simply the distances of the graph vertices from the origin minus one.

a step across an edge between two faces always leaves the black vertex on the left. The equation for Z_n is yet another form of equation (2.1) which allows us to identify $Z_n = R_n$ of equation (2.3).

In terms of triangulations, a bijection similar to that of [2] may be established between rooted planar Eulerian triangulations and the well-labelled (SOS) trees corresponding to the one-wall situation of section 2 as follows. First we replace the face bicoloration by the compatible orientation of all edges in such a way that each triangle is either clockwise- or anticlockwise oriented. This allows us to define the geodesic distance from a vertex to another as the length of a minimal path respecting the orientation of edges. Picking some vertex as the origin, a well-labelled tree is obtained by retaining for each clockwise-oriented triangle the edge linking the two farthest vertices from the origin, and labelling each vertex by its distance from the origin minus one, as illustrated in figure 7 (see [9] for details and proofs). This allows us to interpret R_n as the generating function for planar Eulerian triangulations with a marked (origin) vertex and a marked oriented edge linking a vertex with geodesic distance m from the origin to a vertex with geodesic distance $m + 1$ from the origin with $m \leqslant n$, and with an activity g per vertex.

In the language of graphs, we may use our two-wall solutions to enumerate *bounded* graphs as follows. The quantity

$$G_n^{(L)} = R_n^{(L)} - R_{n-1}^{(L-1)} \tag{A.1}$$

where $R_n^{(L)}$ is the solution of equation (2.1) (resp. (6.1)) with two walls at positions -1 and $L+1$, is the generating function for Eulerian triangulations (resp. for quadrangulations) with a marked (origin) vertex, a marked oriented edge linking a vertex with geodesic distance n from the origin to a vertex with geodesic distance $n + 1$ from the origin and which are 'bounded' in the sense that the geodesic distance of *all the vertices* from the origin is less than or equal to

$L+1$. As an example, the rooted Eulerian triangulation of figure 7 contributes to $G_1^{(L)}$ for all $L \geqslant 2$.

References

[1] See, for example, Le Gall J-F 1999 *Spatial Branching Processes, Random Snakes and Partial Differential Equations* (Boston, MA: Birkhauser)

[2] Chassaing P and Schaeffer G 2002 Random planar lattices and integrated superBrownian excursion *Preprint* math.CO/0205226 (*Proba. Theory Relat. Fields* at press)

[3] Bouttier J, Di Francesco P and Guitter E 2003 Geodesic distance in planar graphs *Nucl. Phys.* B **663** 535–67

[4] See, for example, Jimbo M and Miwa T 1983 Solitons and infinite dimensional Lie algebras *Publ. Inst. Math. Sci.* **19** 943–1001, equation (2.12)

[5] Bateman H 1953 *Higher Transcendental Functions* vol 2 (New York: McGraw-Hill)

[6] Karlin S and Taylor H 1975 *A First Course in Stochastic Processes* (New York: Academic)

[7] Schaeffer G 1997 Bijective census and random generation of Eulerian planar maps *Electron. J. Combinatorics* **4** R20

See also Schaeffer G 1998 Conjugaison d'arbres et cartes combinatoires aléatoires *PhD Thesis* Université Bordeaux I

[8] Bousquet-Mélou M and Schaeffer G 2000 Enumeration of planar constellations *Adv. Appl. Math.* **24** 337–68

See also Poulalhon D and Schaeffer G 2002 A note on bipartite Eulerian planar maps *Preprint* http://www.loria.fr/∼schaeffe

Bouttier J, Di Francesco P and Guitter E 2002 Counting colored random triangulations *Nucl. Phys.* B **641** 519–32

[9] Bouttier J, Di Francesco P and Guitter E 2003 Statistics of planar graphs viewed from a vertex: a study via labelled trees *Nucl. Phys.* B at press *Preprint* cond-mat/0307606 and SPhT/03-104

[10] Delmas J-F 2002 Computation of moments for the length of the one dimensional ISE support *Preprint* available at http://cermics.enpc.fr/∼delmas