



Area-width scaling in generalised Motzkin paths



Nils Haug^{a,*}, Thomas Prellberg^a, Grzegorz Siudem^b

^a School of Mathematical Sciences, Queen Mary University of London, London, E1 4NS, United Kingdom

^b Faculty of Physics, Warsaw University of Technology, Koszykowa 75, PL-00-662 Warsaw, Poland

HIGHLIGHTS

- A family 1D lattice walks called ℓ -Motzkin paths is introduced.
- Motzkin, Schröder and Dyck paths correspond to $\ell = 1, 2, \infty$.
- The area-width scaling is analysed heuristically for all ℓ .
- The known Schröder path generating function is rederived.
- The scaling result for Schröder paths is verified rigorously.

ARTICLE INFO

Article history:

Received 25 January 2017

Received in revised form 5 April 2017

Available online 29 April 2017

Keywords:

Exact solutions

Scaling functions

Basic hypergeometric series

Schroeder paths

Saddle point method

Dominant balance

ABSTRACT

We consider a generalised version of Motzkin paths, where horizontal steps have length ℓ , with ℓ being a fixed positive integer. We first give the general functional equation for the area-width generating function of this model. Using a heuristic ansatz, we then derive the area-width scaling behaviour in terms of a scaling function in one variable for the special cases of Dyck, (standard) Motzkin and Schröder paths, before generalising our approach to arbitrary ℓ . We then rigorously derive the tricritical scaling of Schröder paths by applying the generalised method of steepest descents to the known exact solution for their area-width generating function. Our results show that for Dyck and Schröder paths, the heuristic scaling ansatz reproduces the rigorous results.

© 2017 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Vesicles consist of a closed lipid membrane enclosing a fluid and act as containers for molecules inside of biological cells [1]. Depending on parameters such as temperature and the osmotic pressure acting on the outside of the membrane, they favour different conformations. In the case of a high external pressure, the vesicles tend to minimise their volume, whereas more spatially extended configurations are typical in the case of low pressure.

Two-dimensional vesicles can be modelled as self-avoiding polygons (SAP), either in the continuum or on the lattice [2–4]. In this case, the volume of the vesicle becomes the area of the polygon and the perimeter is the distance covered when travelling around the polygon once. The partition function of vesicles, modelled as SAP on \mathbb{Z}^2 with fixed perimeter m , zero bending rigidity and subject to an external pressure ϵ , is given by

$$Z_m(q) = \sum_{n=0}^{\infty} p_{m,n} q^n, \quad (1)$$

where $p_{m,n}$ is the number of SAP of perimeter m and area n and $q = \exp(-\epsilon)$.

* Corresponding author.

E-mail address: n.a.haug@qmul.ac.uk (N. Haug).

The corresponding area–perimeter generating function is defined as

$$P(t, q) = \sum_{m=0}^{\infty} Z_m(q) t^m = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} p_{m,n} t^m q^n, \quad (2)$$

and can be interpreted physically as the grand-canonical partition function of the model, where both area and perimeter of the vesicle can fluctuate. The shape of the phase diagram of this ensemble has been discussed in [4]. In particular, it was shown that the radius of convergence $t_c(q)$ of $P(t, q)$, seen as a series in t , is positive for $q \leq 1$ and zero for $q > 1$. The point $(t, q) = (t_c(1), 1)$ is called a tri-critical point [5].

Based on exact enumeration data, Richard, Guttman and Jensen conjectured in [6] that the generating function of *rooted* SAP, given by $P_r(t, q) = t \frac{d}{dt} P(t, q)$ satisfies a q -functional equation of finite degree with polynomial coefficients. Subject to this conjecture, they argued that in a region around the tri-critical point, the singular part of $P_r(t, q)$ should obey the scaling relation

$$P_r^{(\text{sing})}(t_c(1) - z\epsilon^\phi, 1 - \epsilon) = \epsilon^\theta F(z) + o(\epsilon^\theta), \quad (3)$$

where the critical exponents are $\theta = 1/3$ and $\phi = 2/3$, and

$$F(z) = b_0 \frac{\text{Ai}'(b_1 z)}{\text{Ai}(b_1 z)}. \quad (4)$$

Here, $\text{Ai}(z)$ is the **Airy function** [7], defined for $z \in \mathbb{C}$ and $c_{\pm} = e^{\pm i\pi/3}$ as

$$\text{Ai}(z) = \frac{1}{2\pi i} \int_{c_-}^{c_+} \exp\left(\frac{u^3}{3} - zu\right) du, \quad (5)$$

and b_0 and b_1 are constants.

From the scaling function, interesting statistical properties of the model, such as the distribution of areas in the infinite perimeter ensemble, can then be deduced [8].

Despite extensive research however, there are almost no rigorous results concerning the enumeration of SAP [2]. Therefore, a mathematical validation of the conjecture made in [6] appears to be currently out of reach.

Progress in understanding the problem can be made by considering directed subclasses of SAP, for which a functional equation for the area–perimeter generating function is known. Examples for such models are Dyck paths, staircase polygons and directed column-convex polygons. From the functional equation, it is possible to extract the tri-critical scaling behaviour of the model, either by using a heuristic scaling ansatz [9] or by carrying out rigorous saddle point analysis on the exact solution for the generating function. The latter method has been applied in [10,11] to the exact solutions for the generating functions of staircase polygons and Dyck paths, which are known from [12,13]. For both models, it was shown that the scaling function is given by Eq. (4), with model-dependent values of b_0 and b_1 .

It is natural to try to generalise the scaling result for Dyck paths to other one-dimensional directed lattice walks, such as Motzkin and Schröder paths [14]. In this case, it is easier to consider the width, i.e. the distance between the start and the end point of the path rather than the perimeter. For Motzkin paths, which are closely related to RSOS configurations, the solution for the area-width generating function has been derived in [15], and for Schröder paths, the area-width generating function was derived in [16]. However, no corresponding scaling forms have been extracted yet.

Out of this motivation, we here analyse the scaling behaviour of a model called ℓ -Motzkin paths, with steps $(1, 1)$, $(1, -1)$ and $(\ell, 0)$, where ℓ is a fixed positive integer. This model has been studied previously in the combinatorics literature with a focus on bijections [17]. Motzkin and Schröder paths are included in this family of walks and correspond to the cases $\ell = 1$ and 2, respectively, whereas Dyck paths can be identified with the limiting case $\ell = \infty$ [18].

After defining the model of ℓ -Motzkin paths, we will re-derive the scaling behaviour for Dyck paths known from [11] by heuristically inserting a single-variable scaling ansatz into the functional equation for the generating function. This approach will then be extended to the cases of (standard) Motzkin and Schröder paths, the scaling behaviour of which has not been analysed yet. Then we will generalise our results to ℓ -Motzkin paths with arbitrary ℓ .

As a further result, we give an alternative derivation of the exact solution for the area-width generating function of Schröder paths, which can be expressed in terms of a quotient of basic hypergeometric series, similar to the well-known expression for Dyck paths [13]. From this we obtain the associated scaling form by rigorous saddle point analysis.

Our results show that the heuristic scaling ansatz reproduces the rigorous results for Dyck and Schröder paths. Moreover, we obtain the same scaling form for all values of ℓ , and therefore in particular for Motzkin paths. We will begin by precisely defining the model we consider.

2. The model of ℓ -Motzkin paths

Given $\ell \in \mathbb{N}$ and $s \in \mathbb{Z}_{\geq 0}$, we define an ℓ -Motzkin path of s steps to be a lattice walk $(x_i, y_i)_{i=0}^s$ on $\mathbb{Z}_{\geq 0}^2$ such that $(x_0, y_0) = (0, 0)$ and from any point (x, y) on the path, the walker can either step towards $(x + 1, y + 1)$, $(x + 1, y - 1)$ or towards $(x + \ell, y)$, corresponding to an up-, down- or horizontal step, respectively. Moreover, the path needs to end on the

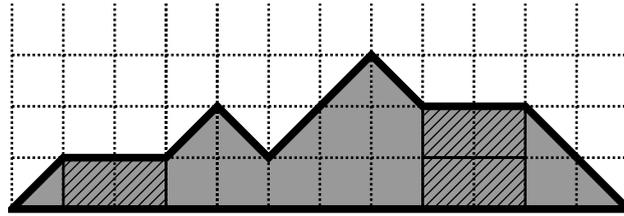


Fig. 1. A Schröder path of width 12, with two horizontal steps of length two, three (2×1) -rectangles below these steps (hatched), four pairs of up/down steps, area 12 below these steps, and thus total area 18.



Fig. 2. Graphical interpretation of Eq. (7). An ℓ -path either consists of zero steps, or it starts with a horizontal step, followed by an ℓ -path, or it starts with an up-step, followed by an ℓ -path, followed by a down-step, followed by another ℓ -path.

horizontal line, i.e. $(x_s, y_s) = (m, 0)$, where m is the width of the ℓ -Motzkin path. Fig. 1 shows an example trajectory for the case $\ell = 2$. Since we will only consider ℓ -Motzkin paths in this paper, we will shortly refer to them as ℓ -paths from now on. For given $\ell \in \mathbb{N}$, we define the generating function

$$G(s, u, p, q) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{v=0}^{\infty} \sum_{w=0}^{\infty} c_{k,l,v,w} s^k u^l p^v q^w, \tag{6}$$

where $c_{k,l,v,w}$ is the number of paths with k horizontal steps, $\frac{l}{2}$ pairs of up- and down-steps, v $(\ell \times 1)$ -rectangles under all the horizontal steps, and w unit squares under all the up- and down-steps (including the half unit squares directly underneath these steps). Thus the weight u is associated to the total number of up- and down-steps, q corresponds to the area under these steps, measured in unit squares of the lattice, and s and p weight the number of horizontal steps and the number of $(\ell \times 1)$ -rectangles underneath these steps, respectively. For example, the weight of the trajectory shown in Fig. 1 is $s^2 u^4 p^3 q^{12}$. Note that there is no explicit ℓ -dependence in $G(s, u, p, q)$. Instead, the area-width generating functions for ℓ -paths of different ℓ are obtained by choosing appropriate values for s and p , as will be explained below.

A functional equation for $G(s, u, p, q)$ can be obtained by noting that for a given ℓ , the set of all ℓ -paths can be divided into the following three subsets. In Fig. 2 we graphically illustrate this decomposition. The first subset only contains the trajectory of zero steps, which has weight 1. The second subset consists of all paths which start with a horizontal step, followed by a path (possibly of zero steps). The weight of a path in this set is thus the weight s of the horizontal step at the beginning, times the weight of the path attached to this initial step. Finally, the third subset contains all the ℓ -paths which start with an up-step. Their weight is given by the weight u of this initial up-step times the weight u of its complimentary down-step, times the weight of the path in between these two steps, times the weight of the path following the down-step. Moreover, the two triangular regions below the initial up-step and the corresponding down-step together contribute one unit of area to the total area underneath the diagonal steps, which is accounted for by a factor of q . Since the path between the initial up-step and its complimentary down-step is elevated by one, each up- or down-step in between generates one further area of unit size, weighted by q , and each horizontal step generates an $(\ell \times 1)$ -rectangle, weighted by p . Summing over the weights of the paths in all three subsets, this leads to the functional equation

$$G(s, u, p, q) = 1 + sG(s, u, p, q) + qu^2G(ps, qu, p, q)G(s, u, p, q). \tag{7}$$

Note that by iteration of Eq. (7), we obtain the continued fraction representation

$$G(s, u, p, q) = \frac{1}{1 - s - \frac{qu^2}{1 - ps - \frac{q^3u^2}{1 - p^2s - \frac{q^5u^2}{1 - p^3s - \dots}}}}, \tag{8}$$

which can be used to approximate $G(s, u, p, q)$ numerically. In the following, we will consider the generating function

$$G^{(\ell)}(a, t, q) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} p_{k,m,n} a^k t^m q^n, \tag{9}$$

where $p_{k,m,n}$ is the number of paths with $2k$ diagonal steps, width m and total area n enclosed between the trajectory and the bottom line, counted in units of lattice cells. We will refer to $G^{(\ell)}(a, t, q)$ as the area-width generating function. Since

each horizontal step of an ℓ -path increases the width of the path by ℓ and each $(\ell \times 1)$ -rectangle increases the total area, measured in units of lattice cells, by ℓ , we have the identity

$$G^{(\ell)}(a, t, q) = G(t^\ell, \sqrt{at}, q^\ell, q). \quad (10)$$

Substituting Eq. (10) into Eq. (7), we obtain the functional equation

$$G^{(\ell)}(a, t, q) = 1 + t^\ell G^{(\ell)}(a, t, q) + aqt^2 G^{(\ell)}(a, qt, q) G^{(\ell)}(a, t, q). \quad (11)$$

For $q = 1$, Eq. (11) is solved by

$$G^{(\ell)}(a, t, 1) = \frac{1 - t^\ell - \sqrt{(1 - t^\ell)^2 - 4at^2}}{2at^2}, \quad (12)$$

and setting $a = 1$ in Eq. (12), we obtain the generating functions of the Motzkin numbers for $\ell = 1$ and the large Schröder numbers for $\ell = 2$ ([19,20] and A001006 and A006318 in [21]).

For given $\ell \in \mathbb{N}$ and real $a > 0$, we denote the smallest positive value for which the discriminant $(1 - t^\ell)^2 - 4at^2$ vanishes by t_c and define $G_c = G^{(\ell)}(a, t_c, 1)$. From Eq. (12) it follows that

$$G_c = \frac{1 - t_c^\ell}{2at_c^2} = \frac{1}{\sqrt{at_c}}. \quad (13)$$

If $|t| < 1$ and we let ℓ tend to infinity, then the weight t^ℓ associated to horizontal steps becomes zero, thus $G^{(\infty)}(a, t, q) = G(0, \sqrt{at}, 0, q)$ satisfies the functional equation

$$G^{(\infty)}(a, t, q) = 1 + aqt^2 G^{(\infty)}(a, qt, q) G^{(\infty)}(a, t, q). \quad (14)$$

In this case, the parameter t only appears in powers of the product at^2 and therefore a can be set equal to one without loss of generality. We write $G^{(\infty)}(t, q) \equiv G^{(\infty)}(1, t, q)$. Eq. (14) is then readily identified as the functional equation for the area-width generating function of Dyck paths [13]. If $q = 1$, it is solved by the generating function of the Catalan numbers ([22] and A000108 in [21]), and for general q , the solution was given in [13].

In the next section, we are now going to analyse the scaling behaviour of $G^{(\ell)}(a, t, q)$ around the point $(a, t, q) = (a, t_c, 1)$ by using a heuristic ansatz. We will begin by treating the case of Dyck paths ($\ell = \infty$), for which the scaling function has been extracted rigorously via the method of steepest descents in [11]. Then we will apply the same approach to Motzkin and Schröder paths ($\ell = 1$ and $\ell = 2$), for which no scaling form has been derived yet in the literature, before generalising our approach to arbitrary ℓ .

3. Heuristic scaling ansatz

Given the known scaling behaviour of Dyck paths, the heuristic approach consists of assuming that in the vicinity of the point $(a, t, q) = (a, t_c, 1)$, also the area-width generating functions of other ℓ -paths satisfy a similar scaling relation. More precisely, we expect that there is a value $z_- < 0$ such that for $z \in (z_-, \infty)$ and $\epsilon \rightarrow 0^+$,

$$G^{(\ell)}(a, t(z, \epsilon), 1 - \epsilon) = G_c + \epsilon^\theta F_0(a, z) + \epsilon^{2\theta} F_1(a, z) + \mathcal{O}(\epsilon^{3\theta}) \quad (15)$$

where $t(z, \epsilon) = t_c - z\epsilon^\phi$; θ and ϕ are positive, non-integer critical exponents, and $F_0(a, z)$ and $F_1(a, z)$ are analytic functions for $z \in (z_-, \infty)$. The function $F_0(a, z)$ is the scaling function.

The heuristic approach is non-rigorous since it makes the assumption that $G^{(\ell)}(a, t(z, \epsilon), 1 - \epsilon)$ admits an expansion of the form (15). Note, however that the method of dominant balance used below forms part of a rigorous method for deriving the area limit distribution of two-dimensional polygon models [8].

For better readability, we omit the dependence of a from now on and write $F_0(a, z) \equiv F(z)$. We define

$$G_{sc}^{(\ell)}(a, z, \epsilon) = G_c + \epsilon^\theta F(z) + \epsilon^{2\theta} F_1(z). \quad (16)$$

By setting $z = \tau\epsilon^{-\phi}$ with $0 < \tau < t_c$, it follows from the positivity of the coefficients of the generating function $G^{(\ell)}(a, t, z)$ that for $a > 0$,

$$\lim_{z \rightarrow \infty} F(z) = -\infty. \quad (17)$$

Following [9], we now insert the RHS of Eq. (15) into the functional equation (11). Using a dominant balance argument, this uniquely determines the values of θ and ϕ and leads to an ODE for the function $F(z)$.

We will begin by validating the heuristic approach by reproducing the result for the scaling function of Dyck paths, known from [11].

3.1. Dyck paths ($\ell = \infty$)

The area-width generating function of Dyck paths satisfies Eq. (14), where, as explained above, a can be set to one without loss of generality. Substituting $a = 1$ and $t^\ell = 0$ into the solution for $q = 1$ given in Eq. (12), we obtain the critical values

$$t_c = \frac{1}{2} \quad \text{and} \quad G_c = 2. \tag{18}$$

Now we define the function

$$\Phi_\infty(z, \epsilon) = 1 - G_{sc}^{(\infty)}(z, \epsilon) + (1 - \epsilon) t(z, \epsilon)^2 G_{sc}^{(\infty)}(z + t_c \epsilon^{1-\phi} - z\epsilon, \epsilon) G_{sc}^{(\infty)}(z, \epsilon),$$

where $G_{sc}^{(\infty)}(z, \epsilon) \equiv G_{sc}^{(\infty)}(1, z, \epsilon)$ is given by Eq. (16), with an unknown function $F(z)$. Under the assumption that Eq. (15) holds, it follows from Eq. (14) that

$$\Phi_\infty(z, \epsilon) = \mathcal{O}(\epsilon^{3\theta}) \quad (\epsilon \rightarrow 0^+). \tag{19}$$

Expanding $\Phi_\infty(z, \epsilon)$ into a series in ϵ , we obtain

$$\begin{aligned} \Phi_\infty(z, \epsilon) = & (1 - G_c + t_c^2 G_c^2) + \epsilon^\theta ((2t_c^2 G_c - 1)F(z)) + \epsilon^{2\theta} ((2t_c^2 G_c - 1)F_1(z) + t_c^2 F(z)^2) + \epsilon^\phi (-2t_c G_c^2 z) \\ & + \epsilon^{1-\phi+\theta} (t_c^3 G_c F'(z)) + \epsilon^{1-\phi+2\theta} (t_c^3 (F(z)F'(z) + G_c F_1'(z))) + \epsilon^{\phi+\theta} (-4zt_c G_c F(z)) + \mathcal{O}(\epsilon^{3\theta}). \end{aligned}$$

The constant coefficient and the one of order ϵ^θ are zero by virtue of Eq. (18). For Eq. (19) to hold, the coefficient of the order of $\epsilon^{2\theta}$ in the above equation needs to be cancelled by another coefficient, hence one of the other exponents needs to equal 2θ . If $2\theta = \phi + \theta$, thus $\theta = \phi$, then the term of order ϵ^ϕ in the above equation could not be cancelled by any other term unless $\theta = 1$, which is impossible by the assumption that θ is not an integer. Likewise, it is impossible that $2\theta = 1 - \phi + 2\theta$, since ϕ is assumed to be non-integer. The third possibility is that $2\theta = \phi$, in which case the only way to obtain a solution $F(z)$ analytic at zero is to also have $2\theta = 1 - \phi + \theta$. The critical exponents hence necessarily satisfy the equations

$$2\theta - \phi = 0 \quad \text{and} \quad \theta + \phi = 1,$$

and thus $\theta = 1/3$ and $\phi = 2/3$. Inserting these exponents, the above equation simplifies to

$$\Phi_\infty(z, \epsilon) = \left[\frac{1}{4}F'(z) + \frac{1}{4}F(z) - 4z \right] \epsilon^{2/3} + \mathcal{O}(\epsilon).$$

From Eq. (19) we thus get the Riccati type ODE

$$F'(z) = Az - BF(z)^2, \tag{20}$$

where $A = 16$ and $B = 1$. In order to solve Eq. (20), we linearise it by using the ansatz

$$F(z) = b_0 \frac{f'(b_1 z)}{f(b_1 z)},$$

where

$$b_0 = \left(\frac{A}{B^2} \right)^{1/3} \quad \text{and} \quad b_1 = (AB)^{1/3}. \tag{21}$$

This leads to the second order ODE

$$f''(z) - zf(z) = 0, \tag{22}$$

the general solution of which is given by

$$f(z) = \lambda_1 \text{Ai}(z) + \lambda_2 \text{Bi}(z), \tag{23}$$

where $\lambda_1, \lambda_2 \in \mathbb{R}$, the Airy function $\text{Ai}(z)$ is defined in Eq. (5) and

$$\text{Bi}(z) = e^{-i\pi/6} \text{Ai}(ze^{-2i\pi/3}) + e^{i\pi/6} \text{Ai}(ze^{2i\pi/3}). \tag{24}$$

Inserting the solution (23) into Eq. (21), we obtain the general solution of Eq. (20) as

$$F(z) = b_0 \frac{(\lambda + 1) \text{Ai}'(b_1 z) + (\lambda - 1) \text{Bi}'(b_1 z)}{(\lambda + 1) \text{Ai}(b_1 z) + (\lambda - 1) \text{Bi}(b_1 z)}, \tag{25}$$

where $\lambda \in \mathbb{R}$. It now follows from the asymptotic behaviour of $\text{Ai}(z)$, $\text{Bi}(z)$ and their derivatives (Section 9.7 in [7]) that the only possibility to satisfy condition (17) is to set $\lambda = 1$. Thus, $F(z)$ has the form given in Eq. (4), and inserting the values $A = 16$ and $B = 1$ into Eq. (21), we obtain

$$b_0 = b_1 = 2^{4/3}. \tag{26}$$

Note that this result is not given explicitly in [11], since a different parametrisation was used in this reference. More precisely, the generating function $G(t, q)$ analysed there is related to the one discussed here by the relation $G(t, q) = G^{(\infty)}(\sqrt{t/q}, \sqrt{q})$, and therefore in [11], $b_0 = 2$ and $b_1 = 4$. However, one verifies that both expressions are equivalent by substituting $t \rightarrow qt^2$ and $q \rightarrow q^2$ into the above result.

We now repeat the same analysis for Motzkin and Schröder paths, for which the scaling behaviour has not yet been studied in the literature.

3.2. Motzkin paths ($\ell = 1$)

Setting $\ell = 1$ and $q = 1$ in Eq. (12), we get the critical values for standard Motzkin paths as

$$G_c = \frac{1 + 2\sqrt{a}}{\sqrt{a}} \quad \text{and} \quad t_c = \frac{1}{1 + 2\sqrt{a}}. \quad (27)$$

Analogous to the case of Dyck paths, we define $\Phi_1(z, \epsilon)$ from Eq. (11) as

$$\begin{aligned} \Phi_1(a, z, \epsilon) = & 1 - G_{sc}^{(1)}(a, z, \epsilon) + t(z, \epsilon)G_{sc}^{(1)}(a, z, \epsilon) \\ & + a(1 - \epsilon)t(z, \epsilon)^2 G_{sc}^{(1)}(a, z + t_c\epsilon^{1-\phi} - z\epsilon, \epsilon) G_{sc}^{(1)}(a, z, \epsilon). \end{aligned}$$

Again, assumption (15) implies that $\Phi_1(a, z, \epsilon) = \mathcal{O}(\epsilon^{3\theta})$ and requires the critical exponents to be $\theta = 1/3$ and $\phi = 2/3$. From the expansion

$$\Phi_1(a, z, \epsilon) = [aG_c t_c^4 F'(z) + a t_c^3 F(z)^2 - z(2aG_c^2 t_c^2 + G_c)] \epsilon^{2/3} + \mathcal{O}(\epsilon),$$

we are then lead to the same ODE (20) as for Dyck paths, with the coefficients now being

$$A = \frac{2G_c}{t_c^2} + \frac{1}{a t_c^3} \quad \text{and} \quad B = \sqrt{a}. \quad (28)$$

The final form of the scaling function is given by Eq. (4) with

$$b_0 = \left(\frac{2\sqrt{a} + 1}{a^2 t_c^3} \right)^{1/3} \quad \text{and} \quad b_1 = \sqrt{a} b_0. \quad (29)$$

3.3. Schröder paths ($\ell = 2$)

For Schröder paths, the critical values are given by

$$t_c = \sqrt{1 + 2a - 2\sqrt{a(a+1)}} \quad \text{and} \quad G_c = \frac{1}{\sqrt{a t_c}}. \quad (30)$$

As for Dyck and Motzkin paths, we define

$$\begin{aligned} \Phi_2(a, z, \epsilon) = & 1 - G_{sc}^{(2)}(a, z, \epsilon) + t(z, \epsilon)^2 G_{sc}^{(2)}(a, z, \epsilon) \\ & + a(1 - \epsilon)t(z, \epsilon)^2 G_{sc}^{(2)}(a, z + t_c\epsilon^{1-\phi} - z\epsilon, \epsilon) G_{sc}^{(2)}(a, z, \epsilon), \end{aligned}$$

and assumption (15) determines $\theta = 1/3$ and $\phi = 2/3$. Expanding $\Phi_2(a, z, \epsilon)$ in ϵ gives with these critical exponents and the above values for t_c and G_c ,

$$\Phi_2(a, z, \epsilon) = [aG_c t_c^4 F'(z) + a t_c^3 F^2(z) - z(2aG_c^2 t_c^2 + 2G_c t_c)] \epsilon^{2/3} + \mathcal{O}(\epsilon),$$

which again leads to Eq. (20), where the coefficients are now

$$A = \frac{2G_c}{t_c^2} + \frac{2}{a t_c^2} \quad \text{and} \quad B = \sqrt{a}. \quad (31)$$

Thus, also for Schröder paths, the scaling function is given by Eq. (4), with

$$b_0 = \left(\frac{2\sqrt{a} + 2t_c}{a^2 t_c^3} \right)^{1/3} \quad \text{and} \quad b_1 = \sqrt{a} b_0. \quad (32)$$

In the next section we now generalise the results obtained so far to general ℓ .

3.4. The case of general ℓ

Now we assume ℓ to be any positive integer. In this general case, it is not possible to give an expression for the critical value t_c as a function of a .

As in the special cases, we define

$$\begin{aligned} \Phi_\ell(a, z, \epsilon) = & 1 - G_{sc}^{(\ell)}(a, z, \epsilon) + t(z, \epsilon)^\ell G_{sc}^{(\ell)}(a, z, \epsilon) \\ & + a(1 - \epsilon)t(z, \epsilon)^2 G_{sc}^{(\ell)}(a, z + t_c\epsilon^{1-\phi} - z\epsilon, \epsilon) G_{sc}^{(\ell)}(a, z, \epsilon), \end{aligned} \tag{33}$$

and from the assumption that $\Phi_\ell(a, z, \epsilon) = \mathcal{O}(\epsilon^{3\theta})$ one obtains $\theta = 1/3$ and $\phi = 2/3$. Expanding the RHS of Eq. (33) in ϵ we get

$$\Phi_\ell(s, z, \epsilon) = [aG_c t_c^4 F'(z) + at_c^3 F(z)^2 - (2aG_c^2 t_c^2 + \ell G_c t_c^{\ell-1})z] \epsilon^{2/3} + \mathcal{O}(\epsilon),$$

which leads to Eq. (20) with

$$A = \frac{2G_c}{t_c^2} + \frac{\ell t_c^{\ell-4}}{a} \quad \text{and} \quad B = \sqrt{a}. \tag{34}$$

The solution of this equation is given in Eq. (4) with parameters

$$b_0 = \left(\frac{2\sqrt{a} + \ell t_c^{\ell-1}}{a^2 t_c^3} \right)^{1/3} \quad \text{and} \quad b_1 = \sqrt{a} b_0. \tag{35}$$

In the next section we are going give a quick derivation of the solution for the generating function of Schröder paths which is alternative to the one given in [16]. This will make it possible to analyse the scaling behaviour of $G^{(2)}(a, t, q)$ rigorously by means of the method of steepest descents, and compare the rigorous result with the one obtained heuristically in Section 3.3.

4. Exact solution for Schröder paths ($\ell = 2$)

The exact solution for the area-width generating function of Schröder paths ($\ell = 2$) has been derived in Eq.(4.41) of [16] by using the Enumerating Combinatorial Objects (ECO) method, and in this reference also a refined model with additional weights corresponding to the number of contacts of the walker with the line $y = 0$ was considered. Here we give an alternative derivation of this result.

Substituting $\ell = 2$ into Eq. (11), we obtain the functional equation

$$1 - t^2 G^{(2)}(a, t, q) + aqt^2 G^{(2)}(a, qt, q)G^{(2)}(a, t, q) = 0 \tag{36}$$

for the area-width generating function of Schröder paths. Inserting the ansatz

$$G^{(2)}(a, t, q) = \frac{H(a, qt^2, q)}{H(a, t^2, q)} \tag{37}$$

into Eq. (36) for $\ell = 2$, we get the linearised functional equation

$$aqt^2 H(q^2 t^2) + (t^2 - 1)H(qt^2) + H(t^2) = 0, \tag{38}$$

where we have abbreviated $H(t^2) \equiv H(a, t^2, q)$ for convenience. The fact that t only appears quadratic in Eq. (38) makes this equation easier to solve than the linearised functional equation for standard Motzkin paths. Namely, Eq. (38) is solved by a basic hypergeometric series [23], defined as

$$H(t) = \sum_{n=0}^{\infty} \frac{(-aq; q)_n}{(q; q)_n} q^{\binom{n}{2}} (-t)^n = {}_1\phi_1 \left(\begin{matrix} -aq \\ 0 \end{matrix}; q, t \right), \tag{39}$$

where the q -Pochhammer symbol is given by (Section 17.2 in [7])

$$(z; q)_n = (1 - z) \cdot (1 - qz) \cdots (1 - q^{n-1}z). \tag{40}$$

This result can be verified by straightforward substitution. For $a = 1$, $G^{(2)}(a, t, q)$ generates a q -deformation of the large Schröder numbers as defined in [24].

As an aside, we note that

$$G^{(2)} \left(\frac{1}{(h-1)q}, t\sqrt{q(h-1)}, q \right) = G_p(h, t, q), \tag{41}$$

where $G_p(h, t, q)$ is the generating function of Dyck paths, with weights h, t and q associated to the number of peaks, width and area, respectively. Here, a peak is called any up-step followed by a down-step. To verify Eq. (41), it suffices to see that via

a factorisation argument similar to the one used in the derivation of Eq. (7), it follows that $G_p(h, t, q)$ satisfies the functional equation

$$1 + ((h - 1)qt^2 - 1)G_p(t) + qt^2G_p(qt)G_p(t) = 0, \quad (42)$$

where $G_p(t) \equiv G_p(h, t, q)$. From Eq. (7) it follows that the LHS of Eq. (41) satisfies the same functional equation, and therefore both functions are identical. Since for a Dyck path, the number of valleys equals the number of peaks minus one, the weight of peaks corresponds physically to the bending rigidity of the vesicle membrane, as it is incorporated in more realistic models [25].

In order to validate our results from Section 3.3, we will now analyse the scaling behaviour of $G^{(2)}(a, t, q)$ by carrying out rigorous saddle point analysis. The same technique has been applied before to area–perimeter weighted staircase polygons and area-width-weighted Dyck paths [10,11].

5. Saddle point asymptotics for Schröder paths ($\ell = 2$)

We are now going to analyse the asymptotic behaviour of the area-width generating function of Schröder paths $G^{(2)}(a, t, q)$ around the tri-critical point $(a, t, q) = (a, t_c, 1)$ of the model by using the method of steepest descents, generalised to the case of two coalescing saddle points. To this purpose, we use the exact solution for $G^{(2)}(a, t, q)$, which has been derived in the last section. We will first derive the asymptotics of the functions $H(t^2)$ and $H(qt^2)$ defined in Eq. (39), and then use Eq. (37) to obtain the leading asymptotic behaviour of $G^{(2)}(a, t, q)$. The calculation is analogous to the one carried out for staircase polygons and Dyck paths in [10,11], therefore we will only outline the essential steps here. We will further assume that $a > 0$.

The first step is to write the series $H(t^2)$ as a contour integral. Using the formula

$$\frac{(-1)^{n+1}q^{\binom{n}{2}}}{(q; q)_n(q; q)_\infty} = \text{Res} \left[(z; q)_\infty^{-1}; z = q^{-n} \right], \quad (43)$$

valid for $n \in \mathbb{N}$, one obtains the expression

$$H(t^2) = \lim_{N \rightarrow \infty} \frac{A}{2\pi i} \int_{C_N} \frac{z^{-2 \log_q(t)}}{(-aq/z; q)_\infty (z; q)_\infty} dz \quad (44)$$

where \log_q denotes the logarithm with base q , $A = (q; q)_\infty (-aq; q)_\infty$ and the contour C_N surrounds exactly the N leftmost zeros of $(z; q)_\infty$. By estimating the integral in Eq. (44), one then shows that in the limit $N \rightarrow \infty$, the contour C_N can be exchanged by a straight line running from $\rho - i\infty$ to $\rho + i\infty$, where $0 < \rho < 1$. The two q -products in the denominator of the integrand in Eq. (44) can be estimated by using the formula [10]

$$\ln(z; q)_\infty = \frac{1}{\ln(q)} \text{Li}_2(z) + \frac{1}{2} \ln(1 - z) + \ln(q)R(z, q), \quad (45)$$

where $\text{Li}_2(z)$ is the Euler dilogarithm (Eq. 25.12.1 in [7]) and the remainder is sufficiently bounded in the relevant region of the \mathbb{C} plane if $q \rightarrow 1^-$. Inserting Eq. (45) into Eq. (44), we obtain the asymptotic expression

$$H(t^2) = \frac{A}{2\pi i} \int_{\rho - i\infty}^{\rho + i\infty} \exp\left(\frac{1}{\epsilon} f(z)\right) g(z) (1 + \mathcal{O}(\epsilon)) dz, \quad (46)$$

as $q \rightarrow 1^-$, where $\epsilon = -\ln(q)$,

$$f(z) = 2 \ln(t) \ln(z) + \text{Li}_2(z) + \text{Li}_2(-a/z), \text{ and} \quad (47)$$

$$g(z) = \sqrt{\frac{z}{(1-z)(z+a)}}. \quad (48)$$

The function $f(z)$ has the two saddle points

$$z_1 = \frac{1}{2} (1 - t^2 - \sqrt{d}); \quad z_2 = \frac{1}{2} (1 - t^2 + \sqrt{d}) \quad (49)$$

where $d = (1 - t^2)^2 - 4at^2$. For $d = 0$, the saddles coalesce in $z_m = (1 - t^2)/2$. It is now possible to apply the method of steepest descents to the integral on the RHS of Eq. (46) (see e.g. [26] for a general introduction).

From Theorem 1 in [27] it follows that there exists a transformation $u \mapsto z(u)$ which is regular in a domain containing z_1 and z_2 if d is sufficiently close to zero, such that

$$f(z) = \frac{1}{3} u^3 - \alpha u + \beta = p(u). \quad (50)$$

The polynomial $p(u)$ has saddle points $u_{1,2} = \pm\sqrt{\alpha}$. Since the transformation is regular, it necessarily maps these saddle points onto the two saddle points of $f(z)$, given in Eq. (49). From Eq. (50) we therefore obtain that

$$\alpha = \left(\frac{3}{4}(f(z_1) - f(z_2))\right)^{2/3} \quad \text{and} \quad \beta = \frac{1}{2}(f(z_1) + f(z_2)). \tag{51}$$

Applying the transformation defined by Eq. (50), Eq. (46) can be rewritten as

$$H(t^2) = \frac{A}{2\pi i} \int_{c-\infty}^{c+\infty} \exp\left(\frac{1}{\epsilon}p(u)\right) S(u)(1 + \mathcal{O}(\epsilon)) du, \tag{52}$$

with $S(u) = g(z(u))z'(u)$ and $c_{\pm} = \exp(\pm i\pi/3)$. In order to obtain the leading term of the asymptotic expansion of $H(t^2)$, we write

$$S(u) = p^{(0)} + u q^{(0)} + (u^2 - \alpha)S_1(u), \tag{53}$$

where $S_1(u)$ is a regular function of u , and the coefficients $p^{(0)}$ and $q^{(0)}$ can be determined by using that $S(u_{1,2}) = p^{(0)} \pm \sqrt{\alpha} q^{(0)}$ and inserting the saddle point values into the second derivative of Eq. (50) to obtain $z'(u_{1,2})$. This gives the expressions

$$p^{(0)} = \sqrt{\frac{\sqrt{\alpha}}{2}} \left(\frac{g(z_2)}{\sqrt{f''(z_2)}} + \frac{g(z_1)}{\sqrt{-f''(z_1)}} \right), \tag{54}$$

$$q^{(0)} = \sqrt{\frac{1}{2\sqrt{\alpha}}} \left(\frac{g(z_2)}{\sqrt{f''(z_2)}} - \frac{g(z_1)}{\sqrt{-f''(z_1)}} \right). \tag{55}$$

Inserting Eq. (53) into Eq. (52) and using the coefficients given in Eqs. (54)–(55), we arrive for $q \rightarrow 1^-$ at the asymptotic expression

$$H(t^2) \sim A \exp\left(\frac{\beta}{\epsilon}\right) \left(\epsilon^{1/3} p^{(0)} \text{Ai}\left(\frac{\alpha}{\epsilon^{2/3}}\right) - \epsilon^{2/3} q^{(0)} \text{Ai}'\left(\frac{\alpha}{\epsilon^{2/3}}\right) \right). \tag{56}$$

The asymptotic expression of $H(q^2 t)$ has the same form as Eq. (56), with $p^{(0)}$ and $q^{(0)}$ replaced by the coefficients $p^{(1)}$ and $q^{(1)}$, which are obtained by replacing $g(z)$ by $h(z) = g(z)/z$ in Eqs. (54)–(55). With this we arrive at the following result. For $q \rightarrow 1^-$,

$$G^{(2)}(a, t, q) \sim \frac{p^{(1)} \text{Ai}(\alpha\epsilon^{-2/3}) - q^{(1)} \text{Ai}'(\alpha\epsilon^{-2/3})\epsilon^{1/3}}{p^{(0)} \text{Ai}(\alpha\epsilon^{-2/3}) - q^{(0)} \text{Ai}'(\alpha\epsilon^{-2/3})\epsilon^{1/3}}, \tag{57}$$

where $\epsilon = -\ln(q)$. Note that this expression is uniform for a range of values of t and a including the critical point $d = 0$. In particular, setting $t = t_c - z\epsilon^{2/3}$, Eq. (57) gives for $\epsilon \rightarrow 0^+$,

$$G^{(2)}(a, t, q) = \frac{1}{z_m} \left(1 + \left(\frac{q^{(0)}}{p^{(0)}} - \frac{q^{(1)}}{p^{(1)}} \right) \frac{\text{Ai}'(\alpha\epsilon^{-2/3})}{\text{Ai}(\alpha\epsilon^{-2/3})} \epsilon^{1/3} + \mathcal{O}(\epsilon^{2/3}) \right). \tag{58}$$

Expanding the coefficients $(q^{(0)}/p^{(0)} - q^{(1)}/p^{(1)})$ and α up to linear order around the critical point, we obtain the coefficients given in Eq. (32), thereby confirming the validity of the heuristic scaling ansatz.

Fig. 3 shows the remarkable agreement of scaling function and partition function asymptotics for q close to one.

6. Summary and outlook

We analysed the scaling behaviour of the generating function of area-weighted Dyck, Motzkin and Schröder paths around the tri-critical point by using a heuristic ansatz and generalised this approach to ℓ -Motzkin paths with arbitrary ℓ . The scaling behaviour of Dyck paths had already been analysed in [11]. The result obtained is in agreement with the one derived in that reference. For Motzkin and Schröder paths, no scaling function had been given yet in the literature.

We gave a derivation of the area-width generating function of Schröder paths alternative to the one in [16]. The exact solution enabled us to obtain a rigorous result for the scaling behaviour of area-width weighted Schröder paths by applying the saddle point method, thereby confirming the result obtained via the heuristic approach.

The solution for the area-width generating functions of Schröder paths and Dyck paths is possible because in both cases, the parameter t only appears in one power in both corresponding functional equations. The generating function for Motzkin paths has been obtained in [15], but the derivation was more involved and led to a much more complicated expression. Exact solutions for the area-width generating functions for ℓ -Motzkin with $2 < \ell < \infty$ are not known yet. This therefore presents an open problem.

Another open question is whether there exists a rigorous method to obtain the scaling behaviour of a generating function directly from the functional equation it satisfies, without having to know the exact solution. So far, the only rigorous method known is by applying the method of steepest descents, which requires knowledge of the exact solution.

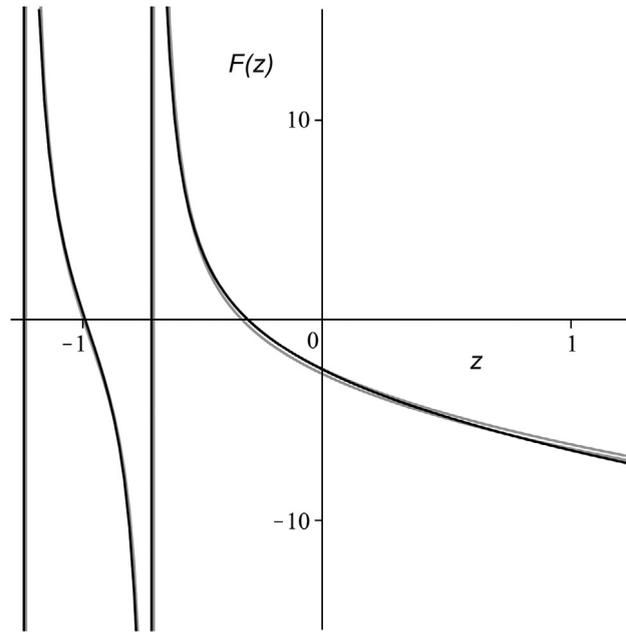


Fig. 3. Plot of the scaling function $F(z)$ given by Eq. (4) with coefficients (32) for $a = 1$ (black) against the approximation of the scaling function obtained directly from the generating function $G^{(2)}(1, t, q)$ and fixed values $\epsilon = 10^{-3}$ and 10^{-4} (grey).

Acknowledgements

GS thanks Queen Mary University of London for hospitality and acknowledges the financial support from the doctoral scholarship from the National Science Centre in Poland (grant no. UMO-2015/16/T/ST1/00528).

References

- [1] B. Alberts, Intracellular vesicular traffic, in: *Molecular Biology of the Cell*, Garland Science, Taylor and Francis Group, 2007, (Chapter 13).
- [2] A.J. Guttmann, Self-avoiding walks and polygons - an overview, *Asia Pac. Math. Newsl.* 2 (2012).
- [3] S. Leibler, R.R.P. Singh, M.E. Fisher, Thermodynamic behavior of two-dimensional vesicles, *Phys. Rev. Lett.* 59 (1987) 1989–1992.
- [4] M.E. Fisher, A.J. Guttmann, S.G. Whittington, Two-dimensional lattice vesicles and polygons, *J. Phys. A: Math. Gen.* 24 (1991) 3095–3106.
- [5] I.D. Lawrie, S. Sarbach, Theory of tricritical points, in: C. Domb, J.L. Lebowitz (Eds.), *Phase Transitions and Critical Phenomena*, Vol. 9, Academic Press, 1984, pp. 2–161. (Chapter 1).
- [6] C. Richard, A.J. Guttmann, I. Jensen, Scaling function and universal amplitude combinations for self-avoiding polygons, *J. Phys. A: Math. Gen.* 34 (2001) L495–L501.
- [7] NIST Digital Library of Mathematical Functions. <http://dlmf.nist.gov/>, Release 1.0.13 of 2016-09-16. F.W.J. Olver, A.B. Olde Daalhuis, D.W. Lozier, B.I. Schneider, R.F. Boisvert, C.W. Clark, B.R. Miller and B.V. Saunders, eds.
- [8] C. Richard, Limit distributions and scaling functions, in: A.J. Guttmann (Ed.), *Polygons, Polyominoes and Polycubes*, in: *Lecture Notes in Physics*, vol. 775, Springer, Netherlands, 2009, pp. 247–299.
- [9] C. Richard, Scaling behaviour of two-dimensional polygon models, *J. Stat. Phys.* 108 (3) (2002) 459–493.
- [10] T. Prellberg, Uniform q -series asymptotics for staircase polygons, *J. Phys. A: Math. Gen.* 28 (1995) 1289–1304.
- [11] N. Haug, T. Prellberg, Uniform asymptotics of area-weighted Dyck paths, *J. Math. Phys.* 56 (2015) 043301.
- [12] R. Brak, A.J. Guttmann, Exact solution of the staircase and row-convex polygon perimeter and area generating function, *J. Phys. A: Math. Gen.* 23 (20) (1990) 4581.
- [13] P. Flajolet, Combinatorial aspects of continued fractions, *Discrete Math.* 32 (1980) 125–161.
- [14] C. Krattenthaler, Lattice path enumeration, in: M. Bóna (Ed.), *Handbook of Enumerative Combinatorics*, CRC Press, Boca Raton, 2015, pp. 589–678. (Chapter 10).
- [15] A.L. Owczarek, T. Prellberg, Exact solution of the discrete $(1 + 1)$ -dimensional RSOS model with field and surface interactions, *J. Phys. A* 42 (49) (2009) 495003.
- [16] A.C. Oppenheim, R. Brak, A.L. Owczarek, Anisotropic step, surface contact, and area weighted directed walks on the triangular lattice, *Internat. J. Modern Phys. B* 16 (9) (2002) 1269–1299.
- [17] E. Pergola, R. Pinzani, S. Rinaldi, R.A. Sulanke, A bijective approach to the area of generalized Motzkin paths, *Adv. in Appl. Math.* 28 (3) (2002) 580–591.
- [18] J. Haglund, Catalan paths and q , t -enumeration, in: M. Bóna (Ed.), *Handbook of Enumerative Combinatorics*, CRC Press, Boca Raton, 2015, pp. 679–752. (Chapter 11).
- [19] E.W. Weisstein, Schröder Number. <http://mathworld.wolfram.com/SchroederNumber.html>, 2016. From MathWorld—A Wolfram Web Resource.
- [20] E.W. Weisstein, Motzkin Number. <http://mathworld.wolfram.com/MotzkinNumber.html>. From MathWorld—A Wolfram Web Resource.
- [21] OEIS Foundation Inc., The On-Line Encyclopedia of Integer Sequences. <http://oeis.org>, 2016.
- [22] M. Aigner, *A Course in Enumeration*, Springer-Verlag, Berlin, Heidelberg, 2007.
- [23] G. Gasper, M. Rahman, *Basic Hypergeometric Series*, in: *Encyclopedia of Mathematics and its Applications*, vol. 96, Cambridge University Press, 1990.
- [24] J. Bonin, L. Shapiro, R. Simion, Some q -analogues of the Schröder numbers arising from combinatorial statistics on lattice paths, *J. Statist. Plann. Inference* 34 (1993) 235–274.
- [25] U. Seifert, Configurations of fluid membranes and vesicles, *Adv. Phys.* 46 (1997) 13–137.
- [26] P. Flajolet, R. Sedgewick, Saddle point asymptotics, in: *Analytic Combinatorics*, Cambridge University Press, 2009, (Chapter 8).
- [27] C. Chester, B. Friedman, F. Ursell, An extension of the method of steepest descents, *Math. Proc. Cambridge Philos. Soc.* 53 (1957) 599–611.