HAL
open science

Families of prudent self-avoiding walks

Mireille Bousquet-Mélou

To cite this version:

Mireille Bousquet-Mélou. Families of prudent self-avoiding walks. Journal of Combinatorial Theory, Series A, 2010, 117 (3), pp.313–344. hal-00276681v3

HAL Id: hal-00276681

https://hal.science/hal-00276681v3

Submitted on 9 Oct 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

FAMILIES OF PRUDENT SELF-AVOIDING WALKS

MIREILLE BOUSQUET-MÉLOU

Abstract. A self-avoiding walk (SAW) on the square lattice is *prudent* if it never takes a step towards a vertex it has already visited. Prudent walks differ from most classes of SAW that have been counted so far in that they can wind around their starting point.

Their enumeration was first addressed by Préa in 1997. He defined 4 classes of prudent walks, of increasing generality, and wrote a system of recurrence relations for each of them. However, these relations involve more and more parameters as the generality of the class increases.

The first class actually consists of *partially directed walks*, and its generating function is well-known to be rational. The second class was proved to have an algebraic (quadratic) generating function by Duchi (2005). Here, we solve exactly the third class, which turns out to be much more complex: its generating function is not algebraic, nor even D-finite.

The fourth class — general prudent walks — is the only isotropic one, and still defeats us. However, we design an isotropic family of prudent walks on the triangular lattice, which we count exactly. Again, the generating function is proved to be non-D-finite.

We also study the asymptotic properties of these classes of walks, with the (somewhat disappointing) conclusion that their endpoint moves away from the origin at a *positive speed*. This is confirmed visually by the random generation procedures we have designed.

1. Introduction

1.1. Families of self-avoiding walks

The study of self-avoiding walks is a famous "elementary" problem in combinatorics, which is also of interest in probability theory and in statistical physics [21]. Recall that, given a lattice with some origin O, a self-avoiding walk (SAW) is a lattice path starting from O that does not visit the same vertex twice (Fig. 1).

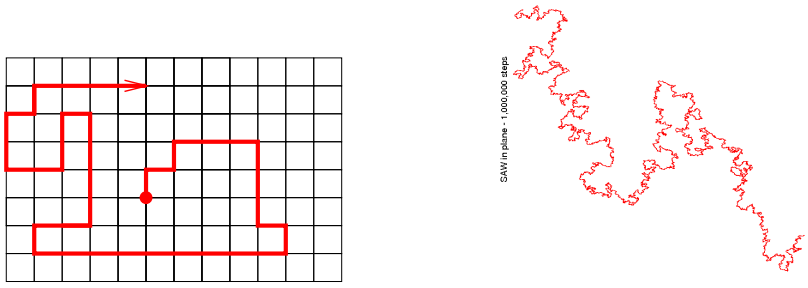

Figure 1. A self-avoiding walk on the square lattice, and a (quasi-)random SAW of length 1,000,000, constructed by Kennedy using a pivot algorithm [19].

It is strongly believed that, for two-dimensional lattices, the number $c(n)$ of n-step SAW and the average end-to-end distance of these walks satisfy

$$c(n) \sim \alpha \mu^n n^\gamma \quad \text{and} \quad \mathbb{E}(D_n) \sim \kappa n^\nu$$

Key words and phrases. Enumeration, self-avoiding walks, D-finite series.

MBM was partially supported by the French "Agence Nationale de la Recherche", project SADA ANR-05-BLAN-0372.

where $\gamma = 11/32$ and $\nu = 3/4$. The growth constant μ is lattice-dependent. Several independent but so far non-rigorous methods predict that $\mu = \sqrt{2+\sqrt{2}}$ on the honeycomb lattice. Moreover, numerical studies suggest that μ may also be a bi-quadratic number for the square lattice [18]. On the probability side, it has been proved that, if the scaling limit of SAW exists and has some conformal invariance property, it must be described by the process SLE(8/3) (stochastic Loewner evolution) [20]. This would imply that the predicted values of γ and ν are correct.

The fact that all these conjectures only deal with *asymptotic* properties of SAW tells us how far the problem is from the reach of *exact* enumeration. The followers of this discipline thus focus on the study of sub-classes of SAW. A simple family consists of *partially directed* walks, that is, self-avoiding walks formed of North, East and West steps. It is easy to see that their generating function is rational [28, Example 4.1.2],

$$\sum_n c(n)t^n = \frac{1+t}{1-2t-t^2}, \tag{1}$$

which gives $c(n) \sim \alpha\mu^n$, with $\mu = 1+\sqrt{2} = 2.41...$ The above series can be refined by taking into account the coordinates (X_n, Y_n) of the endpoint, and the analysis of the result gives:

$$\mathbb{E}(X_n) = 0, \quad \mathbb{E}(X_n^2) \simeq n \quad \text{and} \quad \mathbb{E}(Y_n) \simeq n.$$

We use the notation $a_n \simeq b_n$ as a shorthand for $a_n \sim \alpha b_n$ for some positive constant α.

The prudent walks studied in this paper form a more general class of SAW which have a natural kinetic description: a walk is *prudent* if it never takes a step pointing towards a vertex it has already visited. In other words, the walk is so cautious that it only takes steps in directions where the road is perfectly clear. Various examples are shown on Fig. 2. In particular, partially directed walks are prudent. Note that it is equivalent to require that the new step avoids the past convex hull of the walk: hence prudent walks are the discrete counterpart of the walks evolving in discrete time but continuous space studied by Angel *et al.* [2].

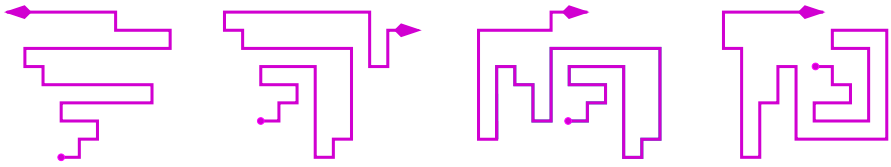

FIGURE 2. Four prudent walks: the first one is partially directed (or: 1-sided), the other ones are respectively 2-sided, 3-sided and 4-sided.

These walks have already been studied in the past under different names: self-directed walks [30], outwardly directed SAW [27], exterior SAW [25], and finally prudent walks [11, 10]. We refrain from the temptation of inventing one more name and stick to the latter terminology. The first two papers above deal with Monte-Carlo simulations. Préa [25] was, to our knowledge, the first to address enumerative questions. He wrote recurrence relations defining an array of numbers $c(n; i, j, h)$ that count prudent walks according to their length (n) and to three additional *catalytic* parameters (i, j, h). By this, we mean that these parameters are essential to the existence of these recurrence relations, and that it is far from obvious how to derive from them a recursion for, say, some numbers $c(n; i, j)$ that would only take into account two of the catalytic parameters (or for $c(n)$). Préa also defined four natural families of prudent walks of increasing generality, called *k-sided*, for k ranging from 1 to 4. In particular, 1-sided walks coincide with partially directed walks, and 4-sided walks coincide with general prudent walks. (Precise definitions will be given below.) He wrote recurrence relations for each of these classes: three catalytic parameters are needed for general (4-sided) prudent walks, but two suffice for 3-sided walks, while one is enough for 2-sided walks. No catalytic parameter is needed for 1-sided walks. This reflects the increasing generality of these four classes of walks.

Recall that the generating function of 1-sided walks (partially directed walks) is rational (1). Duchi [11] proved that 2-sided walks have an algebraic (quadratic) generating function. She also found an algebraic generating function for 3-sided walks, but there was a subtle flaw in her derivation, which was detected by Guttmann [16]. He and Dethridge performed a numerical study of prudent walks, in order to get an idea of their asymptotic enumeration, and of the properties of the associated generating functions [10]. In particular, they conjectured that the length generating function of general prudent walks is not *D-finite*, that is, does not satisfy any linear differential equation with polynomial coefficients. This implies that it is not algebraic.

1.2. **Contents**

In Section 2 of this paper, we collect functional equations that define the generating functions of the four classes of prudent walks introduced by Préa [25]. This is not really original, as these equations are basically equivalent to Préa's recurrence relations. Moreover, similar equations were written by Duchi [11]. Our presentation may be a bit more systematic.

In Sections 3 to 5 we address the solution of these equations. The case of 1-sided walks is immediate and leads the above rational generating function. We then recall how the *kernel method* solves linear equations with *one* catalytic variable. In particular, it provides the generating function of 2-sided walks (Section 3). The extension of this method to linear equations with *two* catalytic variables is not yet completely understood, although a number of papers have been devoted to instances of such equations recently [8, 5, 6, 9, 17, 23, 22]. Underlying the *bi-variate kernel method* is a certain group, which depends on the equation. Roughly speaking, the instances that have been studied suggest that the solution is "nice" if the group is finite. This belief is confirmed by the example of 3-sided walks. The corresponding equation is associated with an infinite group, but we can still solve it, and prove that the generating function of these walks is *not D-finite*, having a rather complex singularity structure (Section 4). We also prove that the growth constants of 3-sided walks and 2-sided walks are the same. It is actually predicted from numerical experiments that the growth constant of *general* prudent walks is also the same [10].

The final equation, which deals with general prudent walks and involves *three* catalytic variables, still defeats us. This is a bit annoying, as the other classes are by definition anisotropic. However, our understanding of the role of catalytic variables leads us to introduce a new isotropic class of prudent walks, on the triangular lattice, which are described by (only) two catalytic variables (Fig. 3, left). Again, the associated equation corresponds to an infinite group. We solve it, and prove that the generating function of triangular prudent walks is not D-finite, having a natural boundary (Section 5).

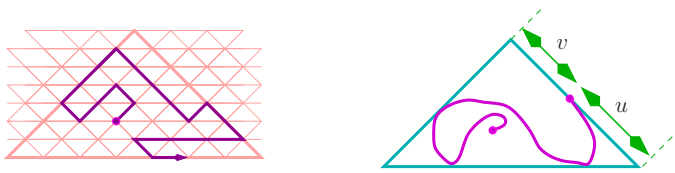

FIGURE 3. Left: A triangular prudent walk in a box of size 7. Right: The two catalytic parameters involved in the enumeration.

We also refine our equations to take into account parameters related to the end-to-end distance of a prudent walk: for instance the coordinates of the endpoint, or the size of the smallest rectangle containing the walk. Extending our solutions to these refined equations is harmless, but the conclusion we draw from these results is somewhat disappointing: the prudent walks we can solve drift away from the origin at a *positive speed*. In other words, the end-to-end distance grows linearly with the length of the walk. We do not know what happens for general (4-sided) prudent walks.

Finally, in Section 6, we address the uniform random generation of n-step prudent walks. Their step-by-step recursive structure allows for a standard recursive approach, in the spirit of [24]. This approach first requires a precomputation stage, followed by a generation stage, which is usually linear. We emphasize that the (costly) precomputation may require less information than the generation itself, and we use this to optimize the precomputation stage. Our final procedures involve the precomputation and storage of up to $O(n^4)$ numbers (for general prudent walks), so that the typical length we can reach is a few hundred. This still provides interesting pictures (Figs. 4, 6, 7, 8).

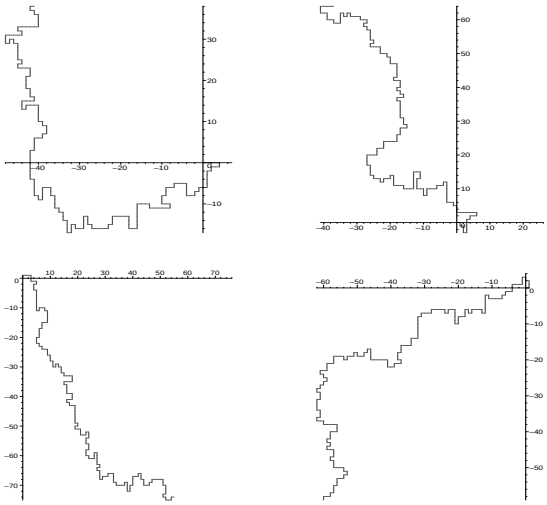

FIGURE 4. Random prudent walks.

1.3. **Families of prudent walks**

Let us conclude this long introduction with some definitions and notations. The *box* of a square lattice walk is the smallest rectangle that contains it. It is not hard to see that the endpoint of a prudent walk is always on the border of the box. This means that every new step either walks on the border of the box, or moves one of its four edges.

Using this box, we can give a kinetic description of partially directed walks: a prudent walk is partially directed if its endpoint, as the walk grows, always lies on the top of the box. This is why partially directed walks are also called 1-*sided*. The generalisation of this terminology is natural: a prudent walk is 2-*sided* if its endpoint always lies either on the top, or on the right edge of the box. It is 3-*sided* if its endpoint is always on the top, right or left edge of the box[1]. Of course, 4-*sided* walks coincide with general prudent walks (Fig. 2).

Consider now a walk on the triangular lattice. Define the (triangular) *box* of the walk as the smallest triangle pointing North that contains the walk. The walk is a *triangular prudent walk* if each new step either inflates the box, or walks along one edge of the box in a prudent way (that is, not pointing to an already visited vertex). An example is shown in Fig. 3. Note that this is *not* the natural counterpart of a square lattice prudent walk: this counterpart would just require the walk to avoid pointing to an already visited vertex. But then the natural "box" would be a hexagon: every new step would either inflate the hexagonal box, or walk along its border. As we will see below, the number of edge lengths of the box (3 for a hexagon, 1 for a triangle) is directly related to the number of catalytic parameters we have to introduce, and this is what makes triangular prudent walks relatively easy to handle.

[1]As noticed by Uwe Schwerdtfeger, we actually require this to hold as well for non-integral points, that is, when the walk is considered as a continuous process. The aim of this is to avoid considering the walk ESW, among others, as a 3-sided walk.

Given a class of walks $\mathcal{C}$, the *generating function* of walks of $\mathcal{C}$, counted by their length, is

$$C(t) = \sum_{w \in \mathcal{C}} t^{|w|},$$

where $|w|$ denotes the length of the walk w. The generalisation of this definition to the series $C(t; u_1, \ldots, u_k)$ counting walks according to their length and to k additional parameters is immediate. We will often drop the length variable t, denoting this series $C(u_1, \ldots, u_k)$. Recall that a one-variable series $C(t)$ is *algebraic* if it satisfies a polynomial equation $P(t, C(t)) = 0$, and *D-finite* if it satisfies a linear ODE with polynomial coefficients, $P_k(t)C^{(k)}(t) + \cdots + P_1(t)C'(t) + P_0(t)C(t) = 0$. Every algebraic series is D-finite. We refer to [29] for generalities about these classes of power series.

2. **Functional equations**

The construction of functional equations for all the families of prudent walks we study rely on the same principle, which we first describe on 1-sided (partially directed) walks.

Consider a 1-sided walk. If it ends with a horizontal step, we can extend it in two different ways: either we repeat the last step, or we change direction and add a North (N) step. Otherwise, the walk is either empty or ends with a North step, and we have three ways (N, E and W) to extend it. This shows that North steps, which move the top edge of the box, play a special role in these walks. Our functional equation is obtained by answering the following question: where is the last North step, and what has happened since then?

More specifically, let $P(t)$ denote the length generating function of 1-sided prudent walks. The contribution to $P(t)$ of walks that contain no North step (horizontal walks) is

$$1 + 2\sum_{n \geq 1} t^n = \frac{1+t}{1-t}.$$

The other walks are obtained by concatenating a 1-sided walk, a North step, and then a horizontal walk. Their contribution is thus

$$P(t)\ t\ \frac{1+t}{1-t}.$$

Adding these two contributions gives the equation

$$P(t) = \frac{1+t}{1-t} + t\ \frac{1+t}{1-t} P(t),$$

from which we readily derive the rational expression (1).

The principle of this recursive description extends to k-sided walks for each k. We say that a step of a k-sided walk is *inflating* if, at the time it was added to the walk, it shifted one of the k edges of the box that are relevant in the definition of k-sided walks. For instance, when $k = 2$, an inflating step moves the top or right edge of the box. We write our equations by answering the question: where is the last inflating step, and what has happened since then?

Since then, the walk has grown *without creating a new inflating step*. What does that mean? Assume $k \geq 2$, that the last inflating step was North, and that, since then, the walk has taken m East steps. Then m cannot be arbitrarily large, otherwise one or several of these East steps would be inflating, having moved the right edge of the box. This observation explains why we have to take into account other "catalytic" parameters in our enumeration of prudent walks. For instance, for a 2-sided walk, we keep track of the distance between the endpoint and the NE corner of the box, using a new variable u (Fig. 5). For 3-sided walks ending on the top of the box, we keep track of the distances between the endpoint and the NE and NW corners of the box (variables u and v). For 3-sided walks ending on the right edge of the box, we keep track of the distance between the endpoint and the NE corner (variable u) and of the width of the box (variable w). For 4-sided walks ending on the top of the box, we keep track of the distances between the endpoint and the NE and NW corners (variables u and v), and of the height of the box (variable w). These parameters, and the names of the corresponding variables,

are schematized in Fig. 5. They give rise to series with one, two or three catalytic variables. For instance, for 4-sided walks ending on the top of their box, we will use the series

$$T(t;u,v,w) \equiv T(u,v,w) = \sum_{i,j,h} T_{i,j,h} u^i v^j w^h,$$

where $T_{i,j,h} \equiv T_{i,j,h}(t)$ counts 4-sided walks ending on the top of their box, at a distance i (resp. j) from the NE (resp. NW) corner, such that the height of the box is h. Similar notation will be used for the other classes of walks.

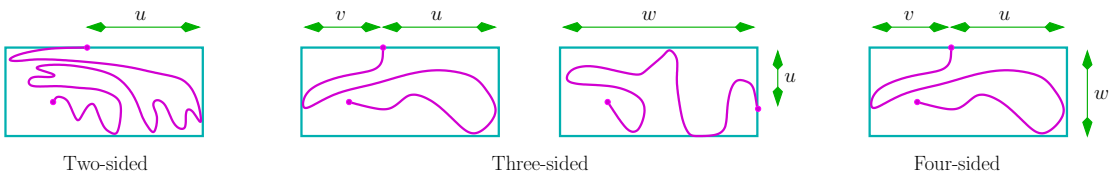

FIGURE 5. Catalytic variables for k-sided walks, $k \geq 2$.

Finally, for triangular prudent walks ending on the right edge of their (triangular) box, we keep track of the distances between the endpoint and the SE and N corners of the box (variables u and v, see Fig. 3).

2.1. **Two-sided prudent walks**

Lemma 1. *The generating function $T(t;u) \equiv T(u)$ of 2-sided walks ending on the top of their box satisfies*

$$\left(1 - \frac{tu(1-t^2)}{(1-tu)(u-t)}\right) T(u) = \frac{1}{1-tu} + t\,\frac{u-2t}{u-t}\,T(t).$$

The generating function of 2-sided walks, counted by their length and the distance of the endpoint to the NE corner of the box, is

$$P(t;u) = 2T(t;u) - T(t;0).$$

Proof. We partition the set of 2-sided walks ending on the top of their box into 3 classes, depending on the existence and direction of the last inflating step (LIS). This step, if it exists, has moved the right or top edge of the box.

1. Neither the top nor the right edge has ever moved: the walk is a sequence of West steps. The generating function for this class is
$$\frac{1}{1-tu}.$$
2. The LIS goes East. This implies that the endpoint of the walk was on the right edge of the box before that step. After that East step, the walk has made a sequence of N steps to reach the top of the box. Observe that, by symmetry, the series $T(t;u)$ also counts walks ending on the right edge of the box by the length and the distance between the endpoint and the NE corner. These two observations give the generating function for this class as
$$t \sum_{i \geq 0} T_i t^i = tT(t).$$
3. The LIS goes North. After this step, there is either an (unbounded) sequence of West steps, or a bounded sequence of East steps. This gives the generating function for this class as
$$\frac{t^2 u}{1-tu} T(u) + t \sum_{i \geq 0} T_i \sum_{k=0}^{i} t^k u^{i-k} = \frac{t^2 u}{1-tu} T(u) + \frac{t}{u-t}\left(uT(u) - tT(t)\right).$$

Adding the 3 terms gives the functional equation satisfied by $T(u)$.

The expression of $P(t;u)$ relies on an inclusion-exclusion argument: we first double the contribution of $T(u)$ to take into account walks ending on the right edge of the box, and then subtract the series $T(0)$ counting those that end at the NE corner. ■

2.2. **Three-sided prudent walks**

Lemma 2. *The generating functions $T(t;u,v) \equiv T(u,v)$ and $R(t;u,w) \equiv R(u,w)$ that count respectively 3-sided walks ending on the top and on the right edge of their box are related by*

$$\left(1 - \frac{tuv(1-t^2)}{(u-tv)(v-tu)}\right) T(u,v) = 1 + tuR(t,u) + tvR(t,v) - \frac{t^2 v}{u-tv} T(tv,v) - \frac{t^2 u}{v-tu} T(u,tu) \quad (2)$$

$$\left(1 - \frac{tuw(1-t^2)}{(u-t)(1-tu)}\right) R(u,w) = \frac{1}{1-tu} + tT(tw,w) - \frac{t^2 w}{u-t} R(t,w). \quad (3)$$

The generating function of 3-sided walks, counted by the length and by the width of the box, is

$$P(t;u) = T(t;u,u) + 2R(t;1,u) - 2T(t;u,0) - \frac{t}{1-t}.$$

Proof. We partition the set of 3-sided walks ending on the top of their box into 4 subsets, depending on the existence and direction of the LIS, which has moved the right, left or top edge of the box.

1. There is no inflating step at all: the walk is empty and contributes 1 to the generating function.
2. The LIS goes East. This case is analogous to Case (2) of 2-sided walks, with generating function
$$tv \sum_{i,j} R_{i,j} t^i v^j = tvR(t,v).$$
3. Symmetrically, the case where the LIS goes West is counted by
$$tuR(t,u).$$
4. If the LIS is a North step, it is followed by a bounded number of West steps, or by a bounded number of East steps. This case is counted by:

$$t \sum_{i,j \geq 0} T_{i,j} \left(\sum_{k=0}^{i} t^k u^{i-k} v^{j+k} + \sum_{k=0}^{j} t^k u^{i+k} v^{j-k} - u^i v^j \right)$$
$$= \frac{t}{u-tv} \left(uT(u,v) - tvT(tv,v) \right) + \frac{t}{v-tu} \left(vT(u,v) - tuT(u,tu) \right) - tT(u,v).$$

Adding the 4 terms gives the first equation of the lemma.

For 3-sided walks ending on the right edge of their box, the last inflating step cannot go West. Three cases remain:

1. There is no inflating step at all: the walk consists of South steps. The generating function for this class is
$$\frac{1}{1-tu}.$$
2. The LIS goes East. This case is analogous to Case (3) of 2-sided walks: the LIS is followed by an unbounded number of South steps, or by a bounded number of North steps. The generating function is
$$\frac{t^2 uw}{1-tu} R(u,w) + \frac{tw}{u-t} \left(uR(u,w) - tR(t,w) \right).$$
3. The LIS goes North. This case is analogous to Case (2) of 2-sided walks. The generating function is found to be
$$tT(tw,w).$$

Adding the 3 terms gives the functional equation for $R(u,w)$.

The expression of $P(t;u)$ again relies on an inclusion-exclusion argument, based on the enumeration of walks ending on a prescribed set of edges of their box. ■

2.3. General prudent walks on the square lattice

Lemma 3. *The generating function $T(t;u,v,w) \equiv T(u,v,w)$ of prudent walks ending on the top of their box satisfies*

$$\left(1-\frac{tuvw(1-t^2)}{(u-tv)(v-tu)}\right)T(u,v,w) = 1+\mathcal{G}(w,u)+\mathcal{G}(w,v)-\frac{tv}{u-tv}\mathcal{G}(v,w)-\frac{tu}{v-tu}\mathcal{G}(u,w)$$

with $\mathcal{G}(u,v) \equiv \mathcal{G}(t;u,v) = tvT(t;u,tu,v)$.

The generating function of prudent walks, counted by the length and the half-perimeter of the box, is

$$P(t;u) = 1+4T(t;u,u,u)-4T(t;0,u,u).$$

We have learnt from [10] that Andrew Rechnitzer has independently obtained the functional equation satisfied by $T(u,v,w)$.

Proof. Again, we partition the set of prudent walk ending on the top of their box into 4 subsets, depending on the existence and direction of the last inflating step. Note that the LIS cannot be a South step.

1. There is no inflating step: the walk is empty, and contributes 1 to the generating function.
2. The LIS goes East. This case is analogous to Case (2) of 2-sided walks. Using the obvious symmetry between prudent walks ending on the top and on the right edge of their box, we find that the generating function for this class is
$$tv\sum_{i,j,h} T_{i,j,h}t^j w^{i+j} v^h = tvT(w,tw,v).$$
3. Symmetrically, the generating function of prudent walks in which the LIS goes West is
$$tuT(w,tw,u).$$
4. Finally, the generating function of prudent walks in which the LIS goes North is analogous to Case (4) of 3-sided walks ending on the top, with generating function:

$$\frac{wt}{u-tv}\left(uT(u,v,w)-tvT(tv,v,w)\right) + \frac{wt}{v-tu}\left(vT(u,v,w)-tuT(u,tu,w)\right) - twT(u,v,w).$$

Adding the 4 terms provides the functional equation for $T(u,v,w)$, given the obvious symmetry $T(u,v,w) = T(v,u,w)$. The expression for $P(t;u)$ again relies on an inclusion-exclusion argument, based on the enumeration of walks ending on a prescribed set of edges of their box. ■

2.4. Triangular prudent walks

Lemma 4. *The generating function $R(t;u,v) \equiv R(u,v)$ of triangular prudent walks ending on the right edge of their box satisfies*

$$\left(1-\frac{tuv(1-t^2)(u+v)}{(u-tv)(v-tu)}\right)R(u,v) = 1+tu(1+t)\frac{v-2tu}{v-tu}R(u,tu)+tv(1+t)\frac{u-2tv}{u-tv}R(tv,v). \quad (4)$$

The generating function of triangular prudent walks, counted by the length and the size of the box, is

$$P(t;u) = 1+3R(t;u,u)-3R(t;u,0). \quad (5)$$

Proof. We partition the set of prudent walks ending on the right edge of their box into 7 subsets, depending on the existence and direction of the last inflating step.

1. There is no inflating step: the walk is empty, and contributes 1 to the generating function.

2. The LIS is a NW step. This implies that the endpoint of the walk was on the left edge of the box before that step. Thanks to the obvious symmetry between walks ending on the left edge and right edge of their box, we obtain the generating function of walks of this type as
$$tu\sum_{i,j} R_{i,j}t^i u^{i+j} = tuR(tu,u).$$
3. The LIS is a W step. Again, the endpoint of the walk was on the left edge of the box before that step. The generating function of walks of this type is
$$tu\sum_{i,j} R_{i,j}t^{i+1} u^{i+j} = t^2uR(tu,u).$$

The cases where the LIS goes SE or SW are very similar to the two previous cases. The endpoint of the walk was on the bottom edge of the box before the last inflating step.

4. The generating function of walks such that the LIS is a SE step is
$$tv\sum_{i,j} R_{i,j}t^j v^{i+j} = tvR(v,tv).$$
5. The generating function of walks such that the LIS is a SW step is
$$tv\sum_{i,j} R_{i,j}t^{j+1} v^{i+j} = t^2vR(v,tv).$$

We are left with the two richer cases where the LIS goes East or North-East. The endpoint of the walk was already on the right edge of the box before the last inflating step, and the LIS is followed by a bounded sequence of SE or NW steps.

6. The generating function of walks such that the LIS is an E step is
$$tv\sum_{i,j} R_{i,j}\left(\sum_{k=0}^{i} t^k u^{i-k} v^{j+k} + \sum_{k=1}^{j+1} t^k u^{i+k} v^{j-k}\right)$$
$$= \frac{tv}{u-tv}\left(uR(u,v) - tvR(tv,v)\right) + \frac{t^2u}{v-tu}\left(vR(u,v) - tuR(u,tu)\right).$$
7. The generating function of walks such that the LIS goes NE is
$$tu\sum_{i,j} R_{i,j}\left(\sum_{k=0}^{j} t^k v^{j-k} u^{i+k} + \sum_{k=1}^{i+1} t^k u^{i-k} v^{j+k}\right)$$
$$= \frac{tu}{v-tu}\left(vR(u,v) - tuR(u,tu)\right) + \frac{t^2v}{u-tv}\left(uR(u,v) - tvR(tv,v)\right).$$

We add the 7 terms above and note that $R(u,v) = R(v,u)$ to obtain the functional equation for $R(u,v)$. The expression of $P(t;u)$ again relies on an inclusion-exclusion argument. ■

3. Enumeration and asymptotic properties of 2-sided prudent walks

In this section, we recall how the *kernel method* works on linear equations with one catalytic variable [7, 3, 26], using the example of 2-sided walks. We first recover Duchi's algebraic generating function, and then refine our enumeration to keep track of other statistics like the end-to-end distance of the walk. This section is a sort of warm-up before solving the more difficult equations of Sections 4 and 5.

Proposition 5. *The generating function $P(t;u)$ of 2-sided walks, counted by their length and by the distance between the endpoint and the NE corner of the box, is*
$$P(t;u) = \frac{2(1-t^2)(1-t)U}{(1-uU)(1-tU)(2t-U)} - 1$$

where

$$U \equiv U(t) = \frac{1-t+t^2+t^3-\sqrt{(1-t^4)(1-2t-t^2)}}{2t}.$$

In particular, the length generating function is

$$P(t;1) = \frac{1}{1-2t-2t^2+2t^3}\left(1+t-t^3+t(1-t)\sqrt{\frac{1-t^4}{1-2t-t^2}}\right).$$

Proof. We start from the functional equation of Lemma 1, written as

$$\big((1-tu)(u-t)-tu(1-t^2)\big)\,T(u) = u-t+t(1-tu)(u-2t)T(t). \tag{6}$$

The series $U \equiv U(t)$ given in the proposition is the only power series in t that cancels the *kernel* of this equation, that is, the polynomial $\big((1-tu)(u-t)-tu(1-t^2)\big)$. The series $T(U) \equiv T(t;U)$ is well-defined. Replacing u by U in the equation cancels the left-hand side, and hence the right-hand side, giving

$$T(t) = \frac{U-t}{t(1-tU)(2t-U)}.$$

Then (6) gives $T(u)$, and the second equation of Lemma 1 provides $P(t;u)$. ∎

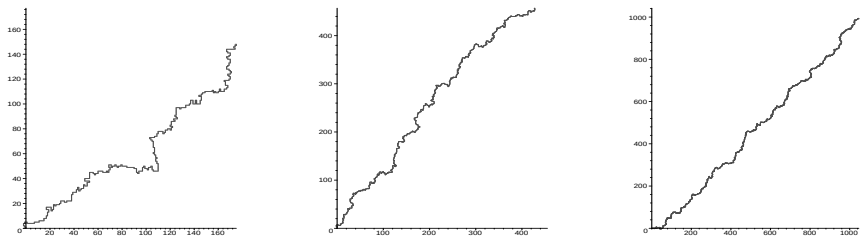

FIGURE 6. Random 2-sided walks of length 500, 1354 and 3148.

Proposition 6 (Asymptotic properties of 2-sided walks). *The length generating function of 2-sided walks has a unique singularity of minimal modulus, $\rho \simeq 0.403$, which is a simple pole and satisfies $1-2\rho-2\rho^2+2\rho^3=0$. The number of n-step 2-sided walks satisfies*

$$p_n \sim \kappa\mu^n \quad \textit{with} \quad \mu = \frac{1}{\rho} \simeq 2.48 \quad \textit{and} \quad \kappa = \frac{\rho\,(3\rho-1)}{(3\rho+1)\,(5\rho-2)} \simeq 2.51.$$

The distance between the endpoint and the NE corner of the box in a random n-step 2-sided walk follows asymptotically a geometric law of parameter 2ρ. In particular, the average value of this distance is asymptotically constant, equal to $\frac{2\rho}{1-2\rho} \simeq 4.15$.

Let X_n (resp. Y_n) denote the abscissa (resp. ordinate) of the endpoint of a random n-step 2-sided walk. Then the mean and variance of X_n+Y_n satisfy

$$\mathbb{E}(X_n+Y_n) \sim \mathfrak{m}\,n, \qquad \mathbb{V}(X_n+Y_n) \sim \mathfrak{s}^2\,n,$$

where

$$\mathfrak{m} = \frac{\rho+1}{3\,\rho+1} \simeq 0.63 \quad \textit{and} \quad \mathfrak{s}^2 = \frac{4\,(\rho+1)^2\,\rho}{(3\,\rho+1)^3\,(1-\rho)} \simeq 0.49,$$

and the variable $\frac{X_n+Y_n-\mathfrak{m}\,n}{\mathfrak{s}\sqrt{n}}$ converges in law to a standard normal distribution.

Finally,

$$\mathbb{E}(X_n-Y_n) = 0, \qquad \mathbb{V}(X_n-Y_n) \sim \mathfrak{s}^2\,n,$$

where

$$\mathfrak{s}^2 = \frac{\rho\left(\rho^2-2\right)(1+\rho)}{(\rho^2+\rho-1)\,(3\,\rho-1)\,(1+3\,\rho)} \simeq 5.17,$$

and the variable $\frac{X_n-Y_n}{\mathfrak{s}\sqrt{n}}$ converges in law to a standard normal distribution.

These asymptotic properties are in good agreement with the random 2-sided walks of Fig. 6.

Proof. We start from the expression of $P(t;1)$ given in Proposition 5. The singularities of $P(t;1)$ are found among the roots of the denominator $(1-2t-2t^2+2t^3)$ and those of the discriminant $(1-t^4)(1-2t-t^2)$. It is not hard to see that the smallest one (in modulus) is ρ, corresponding to a simple pole of the series. In the expression of $P(t;u)$ in terms of U given in Proposition 5, this pole is found when $U=2t$. This implies $\sqrt{(1-\rho^4)(1-2\rho-\rho^2)}=(1-4\rho^2)/2$. The asymptotic behaviour $p_n \sim \kappa\mu^n$ easily follows, with $\mu=\rho^{-1}$.

To study the distance to the NE corner, consider $P(t;u)$ (given in Proposition 5). For u in a neighborhood of 1, $P(t;u)$ still admits ρ as its unique dominant singularity, which remains a simple pole. This gives

$$[t^n]P(t;u) \sim \frac{1-U(\rho)}{1-uU(\rho)}\,[t^n]P(t;1),$$

with $U(\rho)=2\rho$, and the result follows using a continuity theorem for probability generating functions [14, Thm. IX.1].

Then, we enrich our enumeration by taking into account the sum of the coordinates of the endpoint, using a new variable z. The functional equations of Lemma 1 become:

$$\begin{aligned}\left(1-\frac{tuz^2(1-t^2)}{(z-tu)(u-tz)}\right)T(t,z;u) &= \frac{z}{z-tu}+tz\,\frac{u-2tz}{u-tz}T(t,z;tz),\\ P(t,z;u) &= 2T(t,z;u)-T(t,z;0).\end{aligned}$$

We solve them in the same way that we solved the case $z=1$, and obtain:

$$P(t,z;u)=\frac{2z^3(1-t^2)(1-tz)U}{(z^2-uU)(z-tU)(2tz-U)}-1$$

with

$$U\equiv U(t,z)=z\,\frac{1-tz+t^2+t^3z-\sqrt{(1-t^2)(1+t-tz+t^2z)(1-t-tz-t^2z)}}{2t}.$$

For z in a neighborhood of 1, the radius of convergence ρ_z of $P(t,z;1)$ is reached when $U(\rho_z,z)=2z\rho_z$, or $1-2z\rho_z-2\rho_z^2+2z\rho_z^3=0$. One easily checks that $P(t,z;1)$ satisfies the *meromorphic schema* of [14, Thm. IX.9], and the limit behaviour of X_n+Y_n follows.

The study of the distance between the endpoint and the first diagonal is similar. We first refine the enumeration by taking into account the difference X_n-Y_n, using a new variable z. When establishing the functional equation satisfied by $T(t,z;u)\equiv T(z;u)$, one must note that walks ending on the right edge of the box have generating function $T(t,1/z;u)$. The equations finally read

$$\begin{aligned}\left(1+\frac{tu\left(1-t^2\right)}{(z-tu)\,(u-tz)}\right)T(z;u) &= \frac{z}{z-tu}+tzT(\bar z;t\bar z)-\frac{t^2}{u-tz}T(z;tz),\\ P(t,z;u) &= T(t,z;u)+T(t,\bar z;u)-T(t,z;0)\end{aligned}\tag{7}$$

with $\bar z=1/z$. The kernel method provides a linear equation between $T(z;tz)$, $T(\bar z;t\bar z)$, involving the (quadratic) series $U(z)\equiv U(t,z)$ that cancels the kernel of (7). Replacing z by $1/z$ gives a second linear equation, now involving $T(z;tz)$, $T(\bar z;t\bar z)$ and $U(\bar z)$. Thus both series $T(z;tz)$ and $T(\bar z;t\bar z)$ can be expressed in terms of $U(z)$ and $U(\bar z)$. It is then straightforward to obtain expressions for $T(t,z;u)$ and $P(t,z;u)$ in terms of $U(z)$ and $U(\bar z)$ (we recommend using a formal algebra system). In particular, each of these series is algebraic of degree 4.

By an obvious symmetry argument, $\mathbb{E}(X_n-Y_n)=0$. As most continuity theorems involve *non-negative* random variables, we consider $n+X_n-Y_n$, which is clearly non-negative, with mean n. The associated generating function is $P(tz,z;1)$. For z in a neighborhood of 1, the radius of convergence $\rho_z\equiv\rho$ of this series is reached when

$$\rho^7z^8+\rho^5\left(\rho^2+2\,\rho-4\right)z^6-\rho^3\left(4\,\rho^2+5\,\rho-5\right)z^4+\rho\,\left(\rho-2+5\,\rho^2\right)z^2-2\,\rho+1=0,$$

and is again a simple pole of $P(tz,z;1)$. One easily checks that we are again in the meromorphic schema of [14, Thm. IX.9], and the limit behaviour of $n+X_n-Y_n$ (and consequently, of X_n-Y_n) follows. ■

4. Enumeration and asymptotic properties of 3-sided prudent walks

Proposition 7. *The generating function of 3-sided walks ending on the top of their box satisfies*

$$T(t;u,tu)=\sum_{k\geq 0}(-1)^k\frac{\prod_{i=0}^{k-1}\left(\frac{t}{1-tq}-U(uq^{i+1})\right)}{\prod_{i=0}^{k}\left(\frac{tq}{q-t}-U(uq^{i})\right)}\left(1+\frac{U(uq^k)-t}{t(1-tU(uq^k))}+\frac{U(uq^{k+1})-t}{t(1-tU(uq^{k+1}))}\right)$$

where

$$U(w)\equiv U(t;w)=\frac{1-tw+t^2+t^3w-\sqrt{(1-t^2)(1+t-tw+t^2w)(1-t-tw-t^2w)}}{2t}$$

is the only power series in t satisfying $(U-t)(1-tU)=twU(1-t^2)$, and

$$q\equiv q(t)=U(t;1)=\frac{1-t+t^2+t^3-\sqrt{(1-t^4)(1-2t-t^2)}}{2t}.$$

The generating function $P(t;u)$ of 3-sided walks, counted by their length and by the width of the box, can be expressed rationally in terms of $U(u)$ and $T(t;u,tu)$ (see (10)). When $u=1$, this gives the length generating function as

$$P(t;1)=\frac{1}{1-2t-t^2}\left(2\,t^2q\,T(t;1,t)+\frac{(1+t)\left(2-t-t^2q\right)}{1-tq}\right)-\frac{1}{1-t}.$$

Note that the series $U(t;1)$ is the algebraic series that occurs in the solution of 2-sided walks (Proposition 5).

Proof. In the equation (3) satisfied by $R(u,w)$, the only catalytic variable is u (there is no occurrence of $R(\cdot,w')$ with $w'\neq w$). Thus we can apply the standard kernel method: setting $u=U(w)$ cancels the coefficient of $R(u,v)$, and we are left with

$$twR(t,w)=\frac{U(w)-t}{t}\left(\frac{1}{1-tU(w)}+tT(tw,w)\right). \tag{8}$$

Recall that T is symmetric in its two (catalytic) variables. In particular, $T(tw,w)=T(w,tw)$. The equation (2) satisfied by $T(u,v)$ involves the series $R(t,u)$ and $R(t,v)$. We use (8) to express them in terms of $U(u)$, $U(v)$, $T(u,tu)$ and $T(v,tv)$, and obtain an equation that involves only the series T:

$$\left(1-\frac{tuv(1-t^2)}{(u-tv)(v-tu)}\right)T(u,v)=$$
$$1+\frac{U(u)-t}{t(1-tU(u))}+\frac{U(v)-t}{t(1-tU(v))}-\left(\frac{tv}{v-tu}-U(u)\right)T(u,tu)-\left(\frac{tu}{u-tv}-U(v)\right)T(v,tv). \tag{9}$$

Now both u and v play catalytic roles. We want to cancel the kernel of this new equation, namely the polynomial $K(u,v)=(u-tv)(v-tu)-tuv(1-t^2)$, by an appropriate choice of v. Thus v will be a function of u and t. As $K(u,v)$ is *homogeneous* in u and v, the dependency of v in u is extremely simple: $K(u,v)$ vanishes for $v=qu$, where $q=U(t;1)$ only depends on t (of course, $K(u,v)$ also vanishes for $v=u/q$, but this is not a power series in t). This simplicity is crucial to writing the solution in an explicit form. Replacing v by qu in (9) gives

$$T(u,tu)=-\frac{\frac{t}{1-tq}-U(uq)}{\frac{tq}{q-t}-U(u)}T(uq,tuq)+\frac{1}{\frac{tq}{q-t}-U(u)}\left(1+\frac{U(u)-t}{t(1-tU(u))}+\frac{U(uq)-t}{t(1-tU(uq))}\right).$$

Observe that $tq/(q-t)=\frac{1-tq}{1-t^2}=1+O(t)$, while $U(u)=O(t)$. Hence $\frac{tq}{q-t}-U(u)$ is invertible in the ring of power series in t with coefficients in $\mathbb{Q}[u]$. Moreover, $\frac{t}{1-tq}-U(uq)=O(t^3)$, so that

we can iterate the above equation indefinitely, replacing u by uq, then by uq^2, and so on. The net result is the expression of $T(u,tu)$ given in the proposition.

We now seek an expression for $P(t;u)$, which was given in terms of $T(u,u)$, $T(u,0)$ and $R(1,u)$ in the third equation of Lemma 2. We wish to express each of these series in terms of $U(u)$ (which is known explicitly) and $T(u,tu)$, which we have just determined. To express $R(1,u)$ (or, equivalently, $R(1,w)$), we combine the case $u=1$ of (3) with (8). For the other two series, namely $T(u,u)$ and $T(u,0)$, we specialize (9) to $v=u$, and then to $v=0$ (using $U(0)=t$). Putting together the three pieces gives an expression for $P(t;u)$ in terms of $U(u)$ and $T(u,tu)$, which can (for instance) be written as follows:

$$P(t;u) = \frac{1}{1-2t-t^2}\left(2\,t^2U(u)T(u,tu) + \frac{(1+t)\left(2-t-t^2U(u)\right)}{1-tU(u)}\right) \\ - \frac{2(1-U(u))(1+t)(1-u)}{(1-2t-t^2)(1-t-tu-t^2u)}\left(t^2T(u,tu) + \frac{(t+1)\,t}{1-tU(u)}\right) - \frac{1}{1-t}. \quad (10)$$

Note that the second term has a factor $(u-1)$: this makes the specialization $u=1$ obvious, and gives the announced expression of $P(t;1)$. ■

Where is the group? In the introduction (Section 1.2), we wrote that a group is associated with every linear equation with two catalytic variables, and that the group associated with 3-sided walks is infinite. What is this group? It is generated by two transformations Φ and Ψ that act on ordered pairs (u,v) and leave $K(u,v)/u/v$ unchanged:

$$\Phi(u,v) = \left(\frac{v^2}{u}, v\right), \quad \Psi(u,v) = \left(u, \frac{u^2}{v}\right).$$

It is easy to see that they generate an infinite group. Indeed, if one repeatedly applies $\Psi\circ\Phi$ to the pair (u,v), one obtains all pairs $(v^{2i}/u^{2i-1}, v^{2i+1}/u^{2i})$ for $i\geq 0$. The role of this group was first recognized in the study of Markov chains in the quarter plane [13]. See also the more recent and combinatorial papers [8, 5, 6, 9, 17, 23, 22]. As announced in the introduction, we now proceed to show that the series $P(t;1)$ is not D-finite.

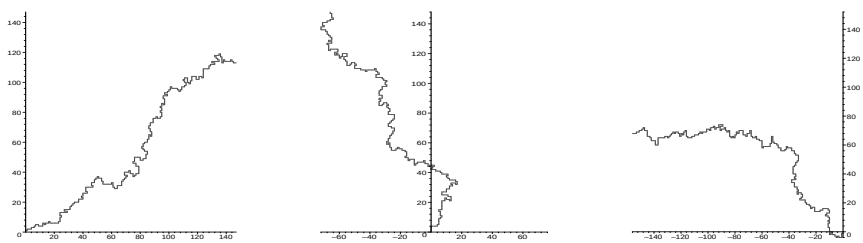

FIGURE 7. Random 3-sided walks of length 400.

Proposition 8 (Nature of the g.f. and asymptotic properties of 3-sided walks). *The length generating function $P(t;1)$ of 3-sided walks is meromorphic in the disk $\mathcal{D}=\{t:|t|<t_c\}$, with $t_c=\sqrt{2}-1$. It has infinitely many poles in this disk, and thus cannot be D-finite. On the segment $(0,t_c)$, the poles are simple and form an increasing sequence $(t_i)_{i\geq 0}$ that tends to t_c.*

The pole t_0 is the unique dominant singularity of $P(t;1)$ and coincides with the radius of convergence ρ of the series counting 2-sided walks (Proposition 6). Thus $t_0=\rho$ satisfies $1-2\rho-2\rho^2+2\rho^3=0$, and the number of n-step 3-sided walks is

$$p_n \sim \kappa\mu^n \quad \textit{with} \quad \mu = \frac{1}{\rho} \simeq 2.48$$

for some $\kappa>0$.

Let W_n denote the width of the box of a random n-step 3-sided walk. Then

$$\mathbb{E}(W_n) \sim \mathfrak{m}\, n, \qquad \mathbb{V}(W_n) \sim \mathfrak{s}^2\, n,$$

where

$$\mathfrak{m} = \frac{1+\rho}{2(1+3\rho)} \simeq 0.31 \quad \textit{and} \quad \mathfrak{s}^2 = \frac{3\rho\,(1+\rho)\left(385 - 1148\,\rho^2 - 494\,\rho\right)}{16\,(\rho^2+\rho-1)\,(3\,\rho-1)^3\,(1+3\,\rho)^3} \simeq 1.41,$$

and the variable $\frac{W_n - \mathfrak{m}\, n}{\mathfrak{s}\sqrt{n}}$ converges in law to a standard normal distribution.

Proof. We start from the expression of $P(t;1)$ given in Proposition 7. The polynomial $1-2t-t^2$ does not vanish in $\mathcal{D}$ (although it vanishes at t_c), and we will show below that $q(t)$ is analytic in $\mathcal{D}$, and that $1-tq(t)$ does not vanish in this domain. Hence most of our analysis will focus on the series $T(t;1,t)$. We will prove that it is meromorphic in $\mathcal{D}$, with a sequence of real positive simple poles $\rho = t_0 < t_1 < t_2 < \cdots < t_c$ satisfying

$$\frac{t_i q(t_i)}{q(t_i) - t_i} = U(q(t_i)^i).$$

Clearly, our first concern will be the series $U(u)$.

• **The series $U(u)$.** The quadratic equation that defines $U(u)$ gives

$$U(u) = t + \tilde{U}(u) \quad \text{with} \quad \tilde{U}(u) = \frac{tuU(u)}{1 - \frac{t\tilde{U}(u)}{1-t^2}}, \tag{11}$$

which shows that both $U(u)$ and $\tilde{U}(u)$ have coefficients in $\mathbb{N}[u]$. In particular, for all t and u such that $U(|t|;|u|)$ converges, $U(t;u)$ also converges and satisfies $|U(t;u)| \le U(|t|;|u|)$.

For $u=1$, we find that the radius of convergence of $U(t;1) = q(t)$ is at t_c. Thus $U(t;1)$ is analytic in $\mathcal{D}$. Moreover, we note that $U(t_c;1) = 1$. The non-negativity of the coefficients of $U(t;1) = q(t)$ implies that $|q(t)| < 1$ for $t \in \mathcal{D}$.

Consequently, for $t \in \mathcal{D}$ and $i \ge 0$, the series $U(t;q(t)^i) \equiv U(q^i)$ is (absolutely) convergent, and thus analytic in $\mathcal{D}$. Moreover, $\left|tU(q^i)\right| < t_c\, U(t_c;1) = t_c < 1$, so that $1 - tU(q^i)$ does not vanish in $\mathcal{D}$.

• **The numerator and denominator of $T(t;1,t)$.** Recall the expression of $T(t;u,tu)$ given in Proposition 7. Note that $q(t)^i \to 0$ as $i \to \infty$, both as a power series in t (because $q(t) = O(t)$) and for every $t \in \mathcal{D}$ (because $|q(t)| < 1$). In particular, the term $tq/(q-t) - U(q^i) = t^2/(q-t) - \tilde{U}(q^i)$ converges to $t^2/(q-t)$. Let

$$D(t) = \prod_{i \ge 0} \left(1 - \frac{q-t}{t^2}\, \tilde{U}(q^i)\right). \tag{12}$$

As $\frac{q-t}{t^2} = 1 + O(t)$ and $\tilde{U}(q^i) = O(t^i)$, this is a well-defined series in t. We write

$$T(t;1,t) = \frac{N(t)}{D(t)} \tag{13}$$

with

$$N(t) = \sum_{k \ge 0} (-1)^k N_k(t)$$

and

$$N_k(t) = \left(\frac{q-t}{t^2}\right)^{k+1} \prod_{i=0}^{k-1} \left(\frac{t}{1-tq} - U(q^{i+1})\right) \prod_{i>k} \left(1 - \frac{q-t}{t^2} \tilde{U}(q^i)\right) T_k(t) \tag{14}$$

with

$$T_k(t) = 1 + \frac{U(q^k) - t}{t(1 - tU(q^k))} + \frac{U(q^{k+1}) - t}{t(1 - tU(q^{k+1}))}.$$

We will now prove that $D(t)$ and $N(t)$ are analytic in $\mathcal{D}$, so that $T(t;1,t)$ is meromorphic.

• **The series $D(t)$ is analytic in $\mathcal{D}$.** As discussed above, every term of the product $D(t)$ is analytic in $\mathcal{D}$. We still need to prove that the product converges in $\mathcal{D}$. For $t \in \mathcal{D}$ and $|u| \leq 1$, the equations (11) imply that $|\tilde{U}(t;u)| \leq |u|\tilde{U}(|t|;1)$. Consequently,

$$|\tilde{U}(q^i)| \leq |q(t)|^i \, \tilde{U}(|t|;1), \tag{15}$$

with $|q(t)| < 1$, so that the series $\sum_i |\tilde{U}(q^i)|$ is convergent for $t \in \mathcal{D}$. The same holds for the product $D(t)$.

• **The series $N(t)$ is analytic in $\mathcal{D}$.** It follows from the properties of q, U and $\tilde{U}$ that every summand N_k (given by (14)) is analytic in $\mathcal{D}$. We will prove the convergence of the series $\sum_{k\geq 0} |N_k(t)|$ by bounding $N_k(t)$. Let us begin with $T_k(t)$. First, we note that

$$\frac{U(t;u)-t}{t(1-tU(t;u))} = \frac{u(1-t^2)U(t;u)}{1-tU(t;u)}$$

is uniformly bounded (by a constant) for $t \in \mathcal{D}$ and $|u| \leq 1$: hence $T_k(t)$ is uniformly bounded by a constant for $t \in \mathcal{D}$.

Let us now bound the infinite product occurring in $N_k(t)$. We write

$$\begin{aligned}\left|\prod_{i>k}\left(1-\frac{q-t}{t^2}\tilde{U}(t;q^i)\right)\right| &\leq \exp\left(\sum_{i>k}\frac{|q-t|}{|t|^2}|\tilde{U}(t;q^i)|\right)\\ &\leq \exp\left(\sum_{i>k}\frac{|q-t|}{|t|^2}\tilde{U}(|t|;1)|q(t)|^i\right) \leq \exp\left(\frac{|q-t|}{|t|^2}\frac{\tilde{U}(|t|;1)|}{1-|q(t)|}\right) < \infty.\end{aligned}$$

The second inequality follows from (15).

Hence it suffices to prove the convergence of

$$\sum_k |M_k(t)| \quad \text{with} \quad M_k(t) = \left(\frac{q-t}{t^2}\right)^{k+1}\prod_{i=0}^{k-1}\left(\frac{t}{1-tq} - U(q^{i+1})\right).$$

Recall that $U(t;q(t)^k) \to U(t;0) = t$ as $k \to \infty$. Hence

$$\frac{M_k(t)}{M_{k-1}(t)} = \frac{q-t}{t^2}\left(\frac{t}{1-tq} - U(q^k)\right) \to \frac{q(q-t)}{1-tq}.$$

Due to the positivity of the coefficients of the series $q(t) = U(t;1)$ and $q(t) - t = \tilde{U}(t;1)$, the modulus of this ratio is strictly bounded in $\mathcal{D}$ by the value it takes at t_c, namely 1. Thus $|M_k(t)/M_{k-1}(t)|$ converges, as k grows, to a limit that is less than 1: the convergence of $\sum_k |M_k(t)|$ follows, and implies that $N(t)$ is analytic in $\mathcal{D}$.

Consequently, $T(t;1,t) = N(t)/D(t)$ is meromorphic in $\mathcal{D}$. In this disk, all its singularities are poles, found among the zeroes of $D(t)$. The product form (12) of $D(t)$ leads us to study the zeroes of each factor, or, equivalently, the zeroes of $tq/(q-t) - U(q^i)$.

• **Zeroes of $tq/(q-t) - U(q^i)$.** We fix $i \geq 0$ and focus on the real interval $(0, t_c)$. On this interval, $U(t;q(t)^i)$ increases from 0 to 1. The identity

$$\frac{tq}{q-t} = \frac{1-tq}{1-t^2} = 1 - \frac{t\tilde{U}(t;1)}{1-t^2}$$

shows that $tq/(q-t)$ decreases from 1 to $1/\sqrt{2} < 1$. Thus there exists a unique t_i in $(0, t_c)$ such that $t_i q(t_i)/(q(t_i) - t_i) = U(q(t_i)^i)$. Moreover, as $U(q^i) \geq U(q^{i+1})$ for $t \in (0, t_c)$, one has $t_0 < t_1 < t_2 < \cdots < t_c$ and

$$\frac{t_i q(t_i)}{q(t_i) - t_i} - U(q(t_i)^k) > 0 \quad \text{for } k > i. \tag{16}$$

Each of these zeroes is simple, as the derivative of $tq/(q-t)$ is negative. As $T(t;1,t)$ is meromorphic in $\mathcal{D}$, its poles are isolated, so that the increasing sequence $(t_i)_i$ can only converge to t_c.

The equation satisfied by $\rho := t_0$ reads $q(\rho) = 2\rho$, which yields the cubic equation of the proposition.

• **Existence of infinitely many poles.** It remains to prove that each t_j, for $j \geq 0$, is actually a pole of $T(t;1,t) = N(t)/D(t)$, that is, that $N(t_j) \neq 0$. Recall the expression (14) of N_k. For $t = t_j$, one has $U(q^j) = tq/(q-t)$, or, equivalently, $\tilde{U}(q^j) = t^2/(q-t)$. Hence

$$N(t) = \sum_{k \geq j} (-1)^k N_k(t).$$

We will show the following properties: for $t = t_j$,

(A) $(-1)^j N_j(t) > 0$, and for $k > j$, the signs of $(-1)^k N_k(t)$ alternate, starting from a positive sign: $(-1)^{k-j-1} N_k(t) > 0$,

(B) $|N_{j+1}(t)| > |N_{j+2}(t)| > |N_{j+3}(t)| > \cdots$

The combination of these two properties implies that $N(t_j) > 0$, and in particular that t_j is indeed a pole of $T(t;1,t)$.

Let us write $t = t_j$, and study the sign of $N_k(t)$. We begin with the signs of $T_k(t)$ and $q-t$. We have already seen that $U(u) - t = \tilde{U}(u) > 0$ and $tU(u) < 1$ for $u \in (0,1)$. In particular,

$$T_k(t) \geq 1 \quad \text{and} \quad q - t > 0.$$

Let us move to the infinite product occurring in $N_k(t)$. Now for $i > k \geq j$,

$$1 - \frac{q-t}{t^2}\tilde{U}(q^i) = 1 - \frac{\tilde{U}(q^i)}{\tilde{U}(q^j)} > 0$$

as $q \in (0,1)$ and $\tilde{U}(u)$ is an increasing function of u, for $0 < u < 1$.

We are left with the sign of $t/(1-tq) - U(q^{i+1})$, for $i \geq 0$. We will prove that, still denoting $t = t_j$,

$$U(q^{j+2}) < \frac{t}{1-tq} < U(q^{j+1}). \tag{17}$$

Given that the sequence $U(q^i)$ decreases as i grows, this gives

$$\begin{aligned} \frac{t}{1-tq} - U(q^{i+1}) \quad &< 0 \quad \text{for } 0 \leq i \leq j, \\ &> 0 \quad \text{for } i \geq j+1, \end{aligned} \tag{18}$$

and concludes the proof of Property (A).

The key to prove (17) is to observe that the function $u \mapsto U(t;u)$ is increasing for $u \in (0,1)$, with an explicit inverse, easily derived from the quadratic equation defining $U(u)$:

$$u = \frac{(U(u)-t)(1-tU(u))}{tU(u)(1-t^2)}.$$

Thus (17) is equivalent to

$$q^{j+2} < \frac{q(1-tq-t^2)}{(1-t^2)(1-tq)} < q^{j+1}.$$

Recall that we write $t = t_j$. Hence $U(q^j) = tq/(q-t)$, so that

$$q^j = \frac{(q-t-t^2q)}{q(1-t^2)(q-t)}.$$

Hence we are left with proving that

$$\frac{q(q-t-t^2q)}{q-t} < \frac{q(1-tq-t^2)}{1-tq} < \frac{(q-t-t^2q)}{q-t},$$

which is easily seen to hold: the first inequality boils down to $q < 1$, and the second one to $q - t - t^2q - t^3q > 0$. But the quadratic equation defining q gives

$$q - t - t^2q - t^3q = tq(1+t)(1-2t) > 0.$$

It remains to prove Property (B), that is, for $k \geq j+2$,

$$\frac{|N_k(t)|}{|N_{k-1}(t)|} = \frac{\left|\frac{t}{1-tq} - U(q^k)\right|}{\left|\frac{tq}{q-t} - U(q^k)\right|} \frac{|T_k(t)|}{|T_{k-1}(t)|} < 1.$$

First, note that the sequence $(U(q^k))_k$ is decreasing. Hence the same holds for $(T_k(t))_k$. Thus it suffices to show that for $k \geq j+2$,

$$\left|\frac{t}{1-tq} - U(q^k)\right| < \left|\frac{tq}{q-t} - U(q^k)\right|.$$

Given (16) and (18), we are left with proving

$$\frac{t}{1-tq} < \frac{tq}{q-t},$$

which boils down to $q^2 < 1$ and is immediate.

This concludes the proof of the properties of $T(t;1,t)$ stated at the beginning of the proof. These properties also hold for $P(t;1)$, as shown by the last equation of Proposition 7.

• **The asymptotic number of 3-sided walks.** We have proved that $T(t;1,t)$ and $P(t;1)$ are meromorphic in $\mathcal{D}$, with a smallest real positive pole at $t_0 = \rho$. By Pringsheim's theorem, ρ is the radius of convergence of $P(t;1)$. In order to prove that the coefficients of $P(t;1)$ behave asymptotically as $\kappa\rho^{-n}$, we need to prove that there exists $\varepsilon > 0$ such that $P(t;1)$ has no singularity other than ρ in the disk of radius $\rho + \varepsilon$.

As $P(t;1)$ is meromorphic in $\mathcal{D}$, its poles are isolated. Thus we only have to prove that there is no other pole of modulus ρ. Recall that all poles t of $P(t;1)$ are roots of $tq/(q-t) = U(q^i)$ for some $i \geq 0$. Let t be a complex number of modulus ρ, with $t \neq \rho$. The positivity of the coefficients of $U(t;u)$ imply that for $i \geq 0$,

$$|U(t;q^i)| < U(\rho;q(\rho)^i) \leq U(\rho;1)$$

while

$$\left|\frac{tq}{q-t}\right| = \left|1 - \frac{t\tilde{U}(t;1)}{1-t^2}\right| > 1 - \frac{\rho\tilde{U}(\rho;1)}{1-\rho^2} = U(\rho;1).$$

Hence $|U(q^i)| < |tq/(q-t)|$ for all i and t cannot be a pole.

• **The width of the box.** The series $P(t;u)$ that counts 3-sided prudent walks by their length and the width of the box is given by (10). A perturbation of the singularity analysis that we have performed for $P(t;1)$ reveals that, for u in some neighborhood of 1, $P(t;u)$ admits a unique dominant singularity, which is a simple isolated pole obtained when $U(u) = tq/(q-t)$, or

$$t\,(t-1)^3\,(t+1)^3\,u^2 + \left(t^3 + 2\,t^2 - 2\,t + 1\right)(t-1)^2\,(t+1)^2\,u - t\left(1 - 2\,t^2 - t + t^4 + 2\,t^3\right).$$

We are again in the meromorphic schema of [14, Thm. IX.9], and the limit behaviour of W_n follows. ■

5. Enumeration and asymptotic properties of triangular prudent walks

We now turn our attention to the triangular prudent walks of Fig. 3. Recall that the box of such a walk is a triangle that points North. If this box has size k, we say that the walk *spans a box* of size k.

Proposition 9. *The number of triangular prudent walks that span a box of size k is*

$$\tilde{p}_k = 2^{k-1}(k+1)(k+2)!.$$

More precisely, the number of triangular prudent walks that span a box of size k, and end on the right edge of their box at distance i from the North corner (and thus at distance $j = k - i$ from the South-West corner) is

$$\tilde{r}_{i,j} = \begin{cases} \dfrac{2^k(k+2)!}{3} & \text{if } i = 0 \text{ or } j = 0, \\ \dfrac{2^k(k+2)!}{6} & \text{otherwise.} \end{cases}$$

Proof. We specialize the functional equations of Lemma 4 to $t = 1$. Remarkably, the kernel of (4) reduces to 1. Denoting $\tilde{R}(u,v) = R(1;u,v)$, we have

$$\tilde{R}(u,v) = 1 + 2u\frac{v-2u}{v-u}\tilde{R}(u,u) + 2v\frac{u-2v}{u-v}\tilde{R}(v,v).$$

This deprives us of our favourite tool (how could we cancel a kernel that reduces to 1?), but still yields a simple solution. Let $\tilde{R}_k(u,v)$ be the homogeneous component of degree k in $\tilde{R}(u,v)$. With the notation $\tilde{r}_{i,j}$ used in the proposition,

$$\tilde{R}_k(u,v) = \sum_{i+j=k} \tilde{r}_{i,j} u^i v^j.$$

The functional equation satisfied by $\tilde{R}(u,v)$ is equivalent to the following recurrence of order 1:

$$\tilde{R}_k(u,v) = 2\,\frac{u(v-2u)\tilde{R}_{k-1}(u,u) - v(u-2v)\tilde{R}_{k-1}(v,v)}{v-u},$$

with initial condition $\tilde{R}_0(u,v) = 1$. One readily checks that the polynomial

$$\tilde{R}_k(u,v) = \frac{2^k(k+2)!}{6}\left(u^k + v^k + \frac{u^{k+1} - v^{k+1}}{u-v}\right)$$

satisfies this recursion, and yields the values $\tilde{r}_{i,j}$ given in the proposition. The expression of $\tilde{p}_k$ then follows using the second equation of Lemma 4. ∎

Remark. The fact that the numbers $\tilde{p}_k$ grow faster than exponentially is not unexpected. In particular, given a square of size k, the number of partially directed (1-sided) walks of height k that fit in this square is $(k+1)^{k+2}$. Indeed, such walks are completely determined by choosing the abscissas of the k vertical steps, of the starting point and of the endpoint.

We now move to the length enumeration of triangular prudent walks.

Proposition 10. *Set $u = \frac{x(1-t)}{(1+tx)(1+t^2x)}$, where x is a new variable. The generating function of triangular prudent walks ending on the right edge of their box satisfies*

$$R(t;u,tu) = (1+xt)(1+xt^2)\sum_{k\geq 0} \frac{t^{\binom{k+1}{2}}\left(xt(1-2t^2)\right)^k}{(xt(1-2t^2);t)_{k+1}}\left(\frac{xt^3}{1-2t^2};t\right)_k \tag{19}$$

where we have used the standard notation

$$(a;q)_n = (1-a)(1-aq)\cdots(1-aq^{n-1}).$$

The generating function of triangular prudent walks, counted by their length and the size of the box, is

$$P(t;u) = 1 + \frac{6tu(1+t)}{1-t-2tu(1+t)}\big(1 + t\,(2u(1+t)-1)\,R(t;u,tu)\big).$$

Comments

1. The parametrisation of u in terms of the length variable t and another variable x is just a convenient way to write $R(t;u,tu)$. Equivalently, we have

$$R(t;u,tu) = (1+Xt)(1+Xt^2)\sum_{k\geq 0}\frac{t^{\binom{k+1}{2}}\left(Xt(1-2t^2)\right)^k}{(Xt(1-2t^2);t)_{k+1}}\left(\frac{Xt^3}{1-2t^2};t\right)_k$$

where

$$X \equiv X(u) = \frac{1-t-ut-ut^2-\sqrt{(1-t)(1-t-2ut-2ut^2+u^2t^2-u^2t^3)}}{2ut^3}$$

is the only power series in t satisfying $u(1+tX)(1+t^2X) = X(1-t)$. In particular, the length generating function of triangular prudent walks is

$$P(t;1) = \frac{6t(1+t)}{1-3t-2t^2}\big(1+t\,(1+2t)\,R(t;1,t)\big) \tag{20}$$

where

$$R(t;1,t) = (1+Y)(1+tY)\sum_{k\geq 0}\frac{t^{\binom{k+1}{2}}\left(Y(1-2t^2)\right)^k}{(Y(1-2t^2);t)_{k+1}}\left(\frac{Yt^2}{1-2t^2};t\right)_k \tag{21}$$

and

$$Y = tX(1) = \frac{1-2t-t^2-\sqrt{(1-t)(1-3t-t^2-t^3)}}{2t^2}. \tag{22}$$

Note that

$$Y = \frac{t}{1-t}(1+Y)(1+tY). \tag{23}$$

2. The value of $R(u,tu)$ is especially simple when $x = 1/(1-2t^2)$, that is, for $u = \frac{1-2t^2}{(1-t^2)(1+2t)}$. Indeed, for this value of u,

$$\begin{aligned} R(t;u,tu) &= \frac{(1+t-2t^2)(1-t^2)}{(1-2t^2)^2}\sum_{k\geq 0}\frac{t^{k+\binom{k+1}{2}}}{(t;t)_{k+1}}\left(\frac{t^3}{(1-2t^2)^2};t\right)_k \\ &= \frac{1-t}{t(1-2t)}\left(-1+\prod_{m\geq 1}(1+t^m)\left(1-\frac{t^{2m+1}}{(1-2t^2)^2}\right)\right). \end{aligned}$$

The product form follows from the following identity [1, Corollary 2.7]:

$$\sum_{n\geq 0}\frac{t^{\binom{n+1}{2}}(a;t)_n}{(t;t)_n} = \prod_{m\geq 1}(1+t^m)(1-at^{2m-1}). \tag{24}$$

This can be used to prove that neither $R(t;u,tu)$ nor $P(t;u)$ are D-finite. However, we will derive from (20–21) the following finer result on the length generating function $P(t;1)$.

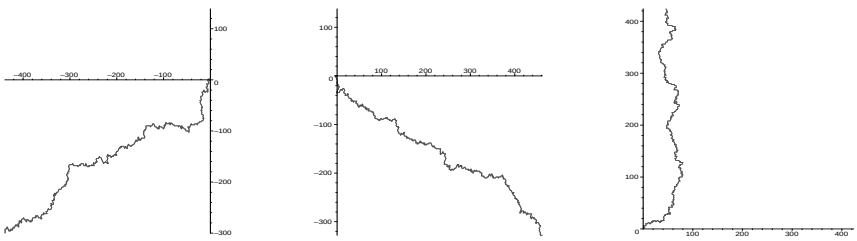

FIGURE 8. Random triangular prudent walks of length 500.

Proposition 11 (Nature of the g.f. and asymptotic properties of triangular prudent walks). *The length generating function $P(t;1)$ of triangular prudent walks is meromorphic in the domain $\mathcal{D} = \{t : |t| < 1\} \setminus [t_c, 1)$, where $t_c \simeq 0.295...$ is the real root of $1 - 3t - t^2 - t^3$. In this domain, it has infinitely many poles, so that it cannot be D-finite. These poles accumulate on a portion of the unit circle, so that $P(t;1)$ has a natural boundary.*

The series $P(t;1)$ has a unique dominant singularity, which is a simple pole $\rho = (\sqrt{17}-3)/4 \simeq 0.280$. Hence the number of triangular prudent walks of length n satisfies

$$p_n \sim \kappa \left(\frac{3+\sqrt{17}}{2} \right)^n$$

for some positive constant κ.

Let S_n denote the size of the box of a random n-step prudent walk. Then the mean and variance of S_n satisfy

$$\mathbb{E}(S_n) \sim \mathfrak{m}\, n, \qquad \mathbb{V}(S_n) \sim \mathfrak{s}^2\, n,$$

where

$$\mathfrak{m} = \frac{1}{2}\left(1 + \frac{1}{\sqrt{17}}\right) \quad \text{and} \quad \mathfrak{s}^2 = \frac{12}{17\sqrt{17}},$$

and the variable $\frac{S_n - \mathfrak{m}\, n}{\mathfrak{s}\sqrt{n}}$ converges in law to a standard normal distribution.

Proof of Proposition 10. The kernel of the functional equation (4) reads:

$$K(u,v) = (u - tv)(v - tu) - tuv(1 - t^2)(u + v).$$

Again we want to cancel it by an appropriate choice of u and v. The kernel is not homogeneous in u and v, as it was in the case of 3-sided walks, but it has another interesting property: the curve $K(u,v) = 0$ has genus 0, and thus admits a rational parametrisation, namely

$$K(U(x), U(tx)) = 0 \quad \text{for} \quad U(x) = \frac{x(1-t)}{(1+tx)(1+t^2x)}.$$

Now take an indeterminate x, and set $u = U(x)$ and $v = U(tx)$. The series $R(u,v)$ is well-defined (as a series in t with coefficients in $\mathbb{R}[x]$). Since the kernel vanishes, it follows from (4) that

$$1 + tu(1+t)\frac{v - 2tu}{v - tu} R(u, tu) + tv(1+t)\frac{u - 2tv}{u - tv} R(tv, v) = 0.$$

Recall that $R(u,v) = R(v,u)$ for all u and v. Denoting $\Phi(x) = R(u(x), tu(x))$, this equation can be rewritten as

$$\Phi(x) = \frac{(1+xt)(1+xt^2)}{1 - xt(1-2t^2)} + \frac{xt^2(1+xt)(1-2t^2-xt^3)}{(1+xt^3)(1-xt(1-2t^2))}\Phi(xt).$$

Iterating it gives the value (19) of $R(u, tu)$.

The expression of $P(t;u)$ in terms of $R(u,tu)$ follows from (5), after specializing (4) to $v = u$ and then $v = 0$. ■

Where is the group? We have solved a new linear equation with two catalytic variables u and v. Again, two transformations Φ and Ψ leave $K(u,v)/u/v$ unchanged:

$$\Phi(u,v) = \left(\frac{v^2}{u(1+v-t^2v)}, v \right), \quad \Psi(u,v) = \left(u, \frac{u^2}{v(1+u-t^2u)} \right).$$

They generate an infinite group. Indeed, if one repeatedly applies $\Psi \circ \Phi$ to the pair $(U(x), U(tx))$, one obtains all pairs $(U(t^{2i}x), U(t^{2i+1}x))$ for $i \geq 0$. We now proceed to show that the series $P(t;1)$ is not D-finite.

Proof of Proposition 11. Recall the expression (20) of $P(t;1)$. The denominator $1 - 3t - 2t^2$ vanishes at ρ and will be responsible for the simple dominant pole of $P(t;1)$. We will see that the series $R(t;1,t) \equiv R(1,t)$, given by (21), has a larger radius of convergence than $P(t;1)$. More precisely, we will prove that $R(1,t)$ is meromorphic in $\mathcal{D}$, with a unique real pole at

$t_0 \simeq 0.288 \in (\rho, t_c)$, where $4\,{t_0}^4 + 2\,{t_0}^3 - 6\,{t_0}^2 - 2\,t_0 + 1 = 0$, and infinitely many non-real poles that accumulate on a portion of the unit circle.

• **The series $Y(t)$.** The square root occurring in the series Y given by (22) vanishes at $t = 1$, at $t = t_c \simeq 0.295$ and at two other points of modulus $1.83... > 1$. Hence Y has radius t_c, but can be analytically continued in the domain $\mathcal{D}$. In this domain, $Y(t)$ is bounded. Moreover, it is easily seen to be increasing on the segment $(-1, t_c)$.

• **The series $R(1,t)$ is meromorphic in $\mathcal{D}$.** Let us write

$$R(1,t) = \frac{N(t)}{D(t)}$$

where

$$D(t) = (Y(1-2t^2); t)_\infty = \prod_{i \geq 0} (1 - Yt^i(1-2t^2))$$

and

$$N(t) = (1+Y)(1+tY) \sum_{k \geq 0} t^{\binom{k+1}{2}} \left(Y(1-2t^2)\right)^k \left(\frac{Yt^2}{1-2t^2}; t\right)_k (Yt^{k+1}(1-2t^2); t)_\infty.$$

Clearly $N(t)$ and $D(t)$ are analytic in $\mathcal{D}$ (as $|t| < 1$). The cancellation of $(1-2t^2)$ does not create any singularity, as

$$(1-2t^2)^k \left(\frac{Yt^2}{1-2t^2}; t\right)_k = (1-2t^2-Yt^2)(1-2t^2-Yt^3)\cdots(1-2t^2-Yt^{k+1}).$$

Hence $R(1,t)$ is meromorphic in $\mathcal{D}$, and its poles are found among the values of t such that $Y(t)t^\ell(1-2t^2) = 1$ for some $\ell \in \mathbb{N}$. All poles are simple, as $Y(t)t^\ell(1-2t^2) = 1$ implies $Y(t)t^m(1-2t^2) = t^{m-\ell} \neq 1$ for $m \neq \ell$.

• **Real poles of $R(1,t)$.** A standard study of the functions $t \mapsto Y(t)$ and $t \mapsto 1/t^\ell/(1-2t^2)$ on the interval $(-1, t_c)$ reveals that the only possible real pole of $R(1,t)$ is at $t_0 \simeq 0.288 < t_c$, where $Y(t_0) = 1/(1-2t_0^2)$. This implies $4\,{t_0}^4 + 2\,{t_0}^3 - 6\,{t_0}^2 - 2\,t_0 + 1 = 0$.

• **Existence of infinitely many poles.** Let us call a *critical value* any $t \in \mathcal{D}$ satisfying $Y(t)t^\ell(1-2t^2) = 1$ for some ℓ. We have just exhibited a real critical value t_0. Fig. 9 shows the critical values obtained for various values of ℓ. We will first prove that $\mathcal{D}$ contains infinitely many critical values. Then, we will show that almost of them, and in particular t_0, are poles of $R(1,t)$.

To prove the former point, we prove that there exist $0 < \theta_0 < \theta_1 < \pi$ such that every point $e^{i\theta}$ with $\theta_0 < |\theta| < \theta_1$ is an accumulation point of critical values. As $Y(\bar{z}) = \overline{Y(z)}$, it suffices to consider the case $\theta > 0$. Let $\theta \in (0, \pi)$ be of the form $p\pi/q$ for two integers p and q. Let $\ell = 2kq$ with $k \geq 0$. Take t in the vicinity of $e^{i\theta}$, that is $t = e^{i\theta}(1+s)$ with s small. As Y is analytic in the neighbourhood of $e^{i\theta}$,

$$Y(t)t^\ell(1-2t^2) = Y(e^{i\theta})(1-2e^{2i\theta}) \exp(\ell s + O(\ell s^2))(1 + O(s)).$$

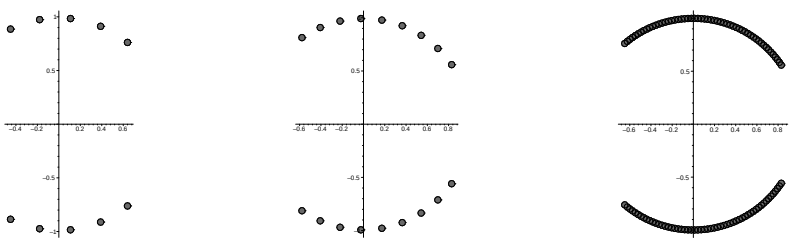

FIGURE 9. The roots of $Y(t)t^\ell(1-2t^2) = 1$ lying inside the domain $\mathcal{D}$ for $\ell = 20$, $\ell = 30$, $\ell = 200$.

This shows that there exists, in the neighbourhood of $e^{i\theta}$, a solution t of $Y(t)t^\ell(1-2t^2)=1$ for every (large) $\ell = 2kq$. This solution reads $t = e^{i\theta}(1+s)$ with

$$s \sim -\frac{1}{\ell}\log\left((1-2e^{2i\theta})Y(e^{i\theta})\right).$$

However, t will only be critical if it lies in $\mathcal{D}$, that is, if $|t|<1$. But

$$|t|^2 = 1 - \frac{2}{\ell}\log\left(|1-2e^{2i\theta}||Y(e^{i\theta})|\right) + o(\ell^{-1}),$$

so that t is critical (for ℓ large) if and only if $f(\theta) := |1-2e^{2i\theta}||Y(e^{i\theta})| > 1$. The function f is continuous on $(0,\pi)$. Moreover, at $t = e^{i\pi/2}$, one has $(1-2t^2)=3$ and $Y(t) = -1+i\pm e^{-i\pi/4}$, so that $f(\pi/2)>1$. This proves the existence of an interval (θ_0,θ_1) where every point θ satisfies $f(\theta)>1$ and is thus an accumulation point of critical values. Fig. 10 shows a plot of $f(\theta)$, and gives estimates for the best possible values $\theta_0 \simeq 0.59$ and $\theta_1 \simeq 2.28$, in good agreement with Fig. 9.

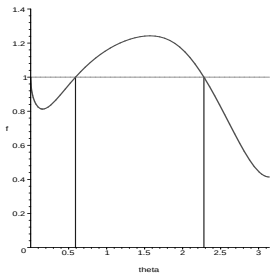

FIGURE 10. The function $f(\theta) := |1-2e^{2i\theta}||Y(e^{i\theta})|$.

Let $t\in\mathcal{D}$ satisfy $Y(t)t^\ell(1-2t^2)=1$. Then t is a pole of $R(1,t)$ if and only if $N(t)\neq 0$. But

$$N(t) = (1+Y)(1+tY)\sum_{k\geq 0} t^{\binom{k+1}{2}} t^{-k\ell}\left(\frac{t^{2-\ell}}{(1-2t^2)^2};t\right)_k (t^{k+1-\ell};t)_\infty.$$

If $k<\ell$, the kth summand is zero because of the term $(t^{k+1-\ell};t)_\infty$. We thus take $k=\ell+j$, with $j\geq 0$. We also rewrite $(1+Y)(1+tY)$ using (23). This gives

$$\begin{aligned} N(t) &= t^{-1-\binom{\ell+1}{2}}\frac{1-t}{1-2t^2}\left(\frac{t^{2-\ell}}{(1-2t^2)^2};y\right)_\ell (t;t)_\infty \sum_{j\geq 0}\frac{t^{\binom{j+1}{2}}}{(t;t)_j}\left(\frac{t^2}{(1-2t^2)^2};t\right)_j \\ &= t^{-1-\binom{\ell+1}{2}}\frac{1-t}{1-2t^2}\left(\frac{t^{2-\ell}}{(1-2t^2)^2};t\right)_\ell (t;t)_\infty \prod_{m\geq 1}(1+t^m)\left(1-\frac{t^{2m+1}}{(1-2t^2)^2}\right). \end{aligned}$$

The final product expression relies on (24).

Could this product vanish for our critical value $t\in\mathcal{D}$? As $|t|<1$, the only factors that might be zero are of the form $1-t^j/(1-2t^2)^2$, for $j\geq 2-\ell$. Each such factor has only a finite number of zeroes in $\mathcal{D}$. In particular, those for which $j<0$, taken together, will only vanish for a *finite* number of critical values, while there are infinitely many such values. For $j\geq 0$, the roots of $1-t^j/(1-2t^2)^2$ that lie in the unit disk satisfy $|1-2t^2|<1$. That is, they are inside the 'glasses' curve $|1-2t^2|=1$ plotted on Fig. 11. In particular, they remain away from the accumulation points we have exhibited on the unit circle, so that only finitely many critical values are non-poles.

Could t_0, the unique critical value found on the segment $(-1,t_c)$, not be a pole? For $t=t_0$, one has

$$\begin{array}{ll} \frac{t^j}{(1-2t^2)^2} \geq \frac{1}{(1-2t^2)^2} > 1 & \text{for } j\leq 0, \\ \frac{t^j}{(1-2t^2)^2} \leq \frac{t}{(1-2t^2)^2} \simeq 0.4 < 1 & \text{for } j\geq 1. \end{array}$$

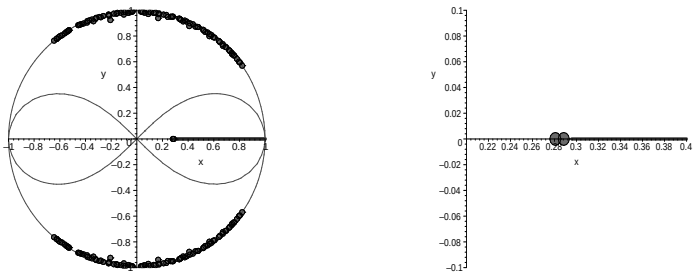

FIGURE 11. The singular landscape of $P(t;1)$: this series is meromorphic inside the unit disk except on a cut $[t_c, 1)$, with $t_c \simeq 0.295$. There are two simple poles on the real axis before t_c, namely $\rho \simeq 0.280$ and $t_0 \simeq 0.288$. Almost all critical values, which accumulate on a portion of the unit circle, are simple poles. The 'glasses' curve is $|1-2t^2| = 1$. To the right, a zoom in the neighbourhood of 0.28.

Hence $N(t_0) \neq 0$, so that t_0 is indeed a pole of $R(1,t)$, and, by Pringsheim's theorem, its radius of convergence.

• **The singular landscape of $R(1,t)$ and $P(t;1)$.** We can now conclude as to the nature of the series $R(1,t)$: this series is meromorphic in the domain $\mathcal{D}$, with a dominant pole at t_0 and infinitely many poles that accumulate on a portion of the unit circle, $\{e^{i\theta} : \theta_0 < |\theta| < \theta_1\}$.

In view of (20), the same holds for the series $P(t;1)$, with t_0 replaced by ρ as the denominator $1-3t-2t^2$ induces a new pole $\rho < t_0$. This is illustrated in Fig. 11. Clearly, $P(t;1)$ has no other pole of modulus less that t_0. Hence the number of n-step triangular prudent walks is, as announced, asymptotic to $\kappa\rho^{-n}$.

• **The size of the box.** Consider the expression of $P(t;u)$ given in Proposition 10. The singular pattern we have found for $u=1$ persists when u is in a neighborhood of 1. The radius of $R(u,tu)$ is reached when $Xt(1-2t^2) = 1$, but the pole of $P(t;u)$ given by the cancellation of $1-t-2tu(1+t)$ is smaller in modulus. We are again in the meromorphic schema of [14, Thm. IX.9], and the limit behaviour of S_n follows.

■

6. Random generation

6.1. Generating trees and random generation

A *generating tree* is a rooted tree with labelled nodes satisfying the following property: if two nodes have the same label, the multisets of labels of their children are the same. Given a label ℓ, the multiset of labels of the children of a node labelled ℓ is denoted $\mathrm{Ch}(\ell)$. The rule that gives $\mathrm{Ch}(\ell)$ as a function of ℓ is called the *rewriting rule* of the tree.

Consider a class $\mathcal{C}$ of walks closed by taking prefixes: if w belongs to $\mathcal{C}$, then so do all prefixes of w (this is the case for our classes of prudent walks). One can then display the elements of $\mathcal{C}$ as the nodes of a generating tree: the empty walk sits at the root of the tree, and the children of the node labelled by w are the walks of the form ws belonging to $\mathcal{C}$, where s is a step. Fig. 12 (top) shows the first few levels of the generating tree of 2-sided walks. An even simpler example is the tree of *directed walks* (walks with N and E steps), which has the following rewriting rule:

$$w \to \begin{cases} wN \\ wS \end{cases}.$$

In this context, the (uniform) random generation of a walk of $\mathcal{C}$ of size n is equivalent to the random generation of a path of length n in the tree, starting from the root. This path can be built recursively as follows: one starts from the empty walk (the root of the tree), and builds

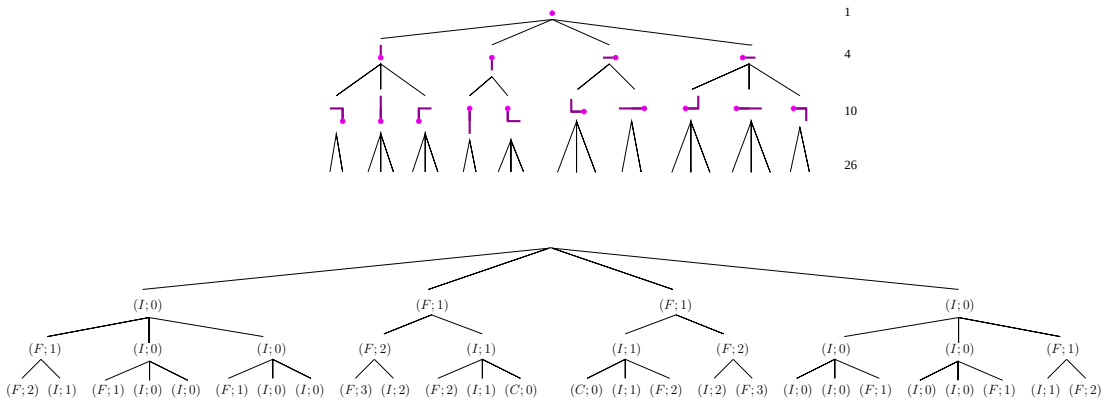

FIGURE 12. The generating tree of 2-sided walks (top) and one of its good labellings (bottom).

the walk step by step, choosing each new step *with the right probability*. If at some point one has obtained a walk w of length $n-m$, with $m>0$, then the probability of choosing s as the next step must be

$$\mathbb{P}(s|w) = \frac{\mathrm{Ex}(ws, m-1)}{\mathrm{Ex}(w,m)}$$

where $\mathrm{Ex}(w,m)$ is the number of *extensions of* w of length m, that is, of walks w' of length m such that the concatenation ww' belongs to $\mathcal{C}$. Clearly, $\mathrm{Ex}(w,m)=0$ if $w \notin \mathcal{C}$. For $w \in \mathcal{C}$, the numbers $\mathrm{Ex}(w,m)$ can be computed inductively:

$$\mathrm{Ex}(w,m) = \left\{ \begin{array}{ll} 1 & \text{if } m=0, \\ \sum_{ws \in \mathrm{Ch}(w)} \mathrm{Ex}(ws, m-1) & \text{otherwise.} \end{array} \right.$$

This is the basic principle of recursive random generation, initiated in [24]. The (time and space) complexity of the procedure that generates an n-step walk of $\mathcal{C}$ depends on how efficiently one solves the following two problems:

(1) compute the numbers $\mathrm{Ex}(w,m)$, for all walks w of length $n-m$,
(2) given a walk w of length less than n, decide which steps s are *admissible*, that is, are such that ws belongs to $\mathrm{Ch}(w)$.

We will explain how both tasks can be achieved by introducing *good labellings* of the class $\mathcal{C}$. A function L defined on $\mathcal{C}$ is a good labelling if for all $w \in \mathcal{C}$, the multiset $\{L(w') : w' \in \mathrm{Ch}(w)\}$ depends only on $L(w)$. This means that the tree obtained by replacing each node w of the generating tree by $L(w)$ is itself a generating tree. We will illustrate this discussion with the class of 1-sided walks.

(1) The numbers $\mathrm{Ex}(w,m)$**.** When $\mathcal{C}$ is the class of 1-sided walks, these numbers only depend on the *direction* (horizontal or vertical, H or V for short) of the final step of w (and of course on m). Denote this direction by $L(w)$. Then it is easy to see that L is a good labelling: the tree obtained by replacing each node w of the generating tree by $L(w)$ is itself a generating tree, with rewriting rule

$$V \to \left\{ \begin{array}{l} V \\ H \\ H \end{array} \right. , \quad H \to \left\{ \begin{array}{l} V \\ H \end{array} \right. . \tag{25}$$

(By convention, the label of the empty walk is set to V.)

For a general class $\mathcal{C}$, any good labelling L allows us to compute the numbers $\mathrm{Ex}(w,m)$. Indeed, they can be rewritten as $\mathrm{Ex}(L(w),m)$, with

$$\mathrm{Ex}(\ell,m) = \left\{ \begin{array}{ll} 1 & \text{if } m=0, \\ \sum_{\ell' \in \mathrm{Ch}(\ell)} \mathrm{Ex}(\ell', m-1) & \text{otherwise.} \end{array} \right. \tag{26}$$

In the case of partially directed walks, we have

$$\mathrm{Ex}(V,m) = \mathrm{Ex}(V,m-1) + 2\,\mathrm{Ex}(H,m-1), \quad \mathrm{Ex}(H,m) = \mathrm{Ex}(V,m-1) + \mathrm{Ex}(H,m-1).$$

A good labelling is *compact* if it takes its values in $\mathbb{Z}^k$, with k fixed. In this case, it is usually easy to determine $\mathrm{Ch}(\ell)$, given ℓ. In order to fasten the computation of the extension numbers, one tries to minimize the number of distinct labels $L(w)$ occurring in the first n levels of the generating tree. This solves Problem (1) above.

(2) Admissible steps. We are left with the second problem: determine the admissible steps. Moreover, we also need to decide, for every admissible step s, which of the elements of $\mathrm{Ch}(L(w))$ is the label of ws. That is, we need a way to *correlate* new steps with new values of the label. In our random generation procedures, this will be achieved thanks to a second good labelling P having the following properties:

- it refines the first one; that is, $P(w) = (L(w)|M(w))$, and the rewriting rule of P extends that of L,
- the last step of w is $S(P(w))$, for a simple function S.

For instance, in the case of 1-sided walks, we keep the label V for walks ending with a vertical step, but introduce two versions of H, namely $(H \mid 1)$ (for a final E step) and $(H \mid -1)$ (for a final W step), and refine the rule (25) as follows:

$$V \to \begin{cases} V \\ (H \mid 1) \\ (H \mid -1) \end{cases}, \quad (H \mid 1) \to \begin{cases} V \\ (H \mid 1) \end{cases}, \quad (H \mid -1) \to \begin{cases} V \\ (H \mid -1) \end{cases}.$$

The last step function is given by $S(V) = N$, $S(H \mid 1) = E$, $S(H \mid -1) = W$. This new rewriting rule is less compact than (25), and not needed for the calculation of the extension numbers (why compute and store separately $\mathrm{Ex}((H \mid 1), m)$ and $\mathrm{Ex}((H \mid -1), m)$ while they are equal?). However, it *is* needed for the generation stage.

In our random generation procedures below we use the L-labels to compute the extension numbers, and the P-labels to determine the admissible steps and associate with them the correct new label. Accordingly, the algorithm involves two procedures:

- The procedure $\mathrm{Ex}(\ell, m)$ computes the extension numbers using (26),
- The generation procedure reads (denoting $p[1] = \ell$ for a P-label $p = (\ell|m)$):

```
gen:=proc(n)
Initialisation w:=[]:  p:=P(w):  ℓ:=p[1]:  m:=n:
while m>0 do
[p1,..., pk]:= Ch(p):
for i from 1 to k do ℓi:=pi[1] od:
U:=Uniform(0,1):  d:=U*Ex(ℓ,m):
i:=1:
while d> Ex(ℓ1,m-1) +··· +Ex(ℓi,m-1) do i:=i+1:  od:
p:=pi:  s:=S(p):  w:=ws:
od:
```

Since all our procedures have this form, the following sections only describe the labellings L and P that we use for each class of walks. In all cases, the number of children of a P-label is bounded, so that the generation procedure is achieved in linear time. We first discuss the simple case of 2-sided walks, and then the more complex case of general prudent walks (on the square lattice). Once the principles underlying the corresponding labellings are understood, it is not difficult to adapt them to 3-sided prudent walks and to triangular prudent walks.

6.2. **Two-sided prudent walks**

A non-empty 2-sided walk w can be of three different types, I, C or F, depending on the nature of its final step s:

(I) either s is an *inflating* step, which moves the North or East edge of the box;

(C) or s walks along the North or East edge of the box, moving the endpoint *closer* to the NE corner,
(F) or s walks along the North or East edge, moving the endpoint *further* away from the NE corner.

We label w by $L(w) = (T; i)$, where T is the type of w, and i is the distance of the endpoint of w to the NE corner of its box. It is not hard to realize that L is a good labelling: the labels of the children of a node w of the generating tree only depend on $L(w)$. These labels are described by the following rewriting rule:

$$(I;i) \to \left\{ \begin{array}{ll} (I;i) & \\ (F;i+1) & \\ (C;i-1) & \text{if } i>0 \\ (I;0) & \text{if } i=0 \end{array} \right. , \quad (C;i) \to \left\{ \begin{array}{ll} (I;i) & \\ (C;i-1) & \text{if } i>0 \\ (I;0) & \text{if } i=0 \end{array} \right. , \quad (F;i) \to \left\{ \begin{array}{l} (I;i) \\ (F;i+1) \end{array} \right. .$$

We do not define the label of the empty walk. Two of its four children have label $(I; 0)$ (corresponding to North and East steps), while the other two (corresponding to South and West steps) have label $(F; 1)$. The top of the label based tree is shown in Fig. 12 (bottom).

The number of extensions of length m of a walk labelled ℓ can now be computed using (26). For a walk of length m and label $(T; i)$, one has $i \le m$. Hence $O(n)$ different labels occur in the first n levels of the tree, and the calculation of the extension numbers $\mathrm{Ex}(\ell, m)$ needed to generate a random walk of length n requires $O(n^2)$ operations.

The L-label $L(w)$ does not determine the last step s of w. For the generation stage, we refine it by defining $P(w) = (L(w) \mid s)$. The last step function is of course $S(P(w)) = s$. The generating tree based on the labelling P has rewriting rule:

$$(I;i \mid s) \to \left\{ \begin{array}{lll} (I;i & \mid & s) & \\ (F;i+1 & \mid & 3-s) & \\ (C;i-1 & \mid & 1-s) & \text{if } i>0 \\ (I;0 & \mid & 1-s) & \text{if } i=0 \end{array} \right. ,$$

$$(C;i \mid s) \to \left\{ \begin{array}{lll} (I;i & \mid & 1-s) & \\ (C;i-1 & \mid & s) & \text{if } i>0 \\ (I;0 & \mid & s) & \text{if } i=0 \end{array} \right. , \quad (F;i \mid s) \to \left\{ \begin{array}{lll} (I;i & \mid & 3-s) \\ (F;i+1 & \mid & s) \end{array} \right. .$$

We have written $s = 0$ (resp. 1, 2, 3) for a North (resp. East, South, West) step. The four children of the root have labels $(I; 0 \mid 0)$, $(I; 0 \mid 1)$, $(F; 1 \mid 2)$, $(F; 1 \mid 3)$.

Remark. As shown in [11], the language on the alphabet $\{N, S, E, W\}$ that naturally encodes 2-sided walks is *non-ambiguous context-free*. This implies that one can generate these walks using another type of recursive approach, based on factorisations of the walks [15]. This approach achieves a better complexity than our naive step-by-step construction: only $O(n)$ operations are needed in the precalculation stage. Moreover, it can be optimized further using the principles of *Boltzmann approximate-size sampling* [12]. As the other classes of prudent walks we have studied cannot be described by context-free grammars, we have focussed above on the step-by-step recursive approach. Independently, we have also implemented the Boltzmann approach, which produced the second and third walks of Fig. 6.

6.3. **General prudent walks on the square lattice**

A non-empty prudent walk w can be of two different types, I or A, depending on the nature of its final step s:

(I) either s is an *inflating* step, which moves one of the edges of the box,
(A) or s walks *along* one of the edges of the box.

Consider the last edge of the box that has moved while constructing w. Say it is the North edge (the other cases are treated similarly after a rotation). Then w ends on this edge. If the final step s is inflating (North), or walks along the top edge in *clockwise* direction (East), let i

(resp. j) be the distance between the endpoint of w and the NE (resp. NW) corner of the box. Otherwise, s walks along the top edge in a counterclockwise direction (West): exchange the roles of i and j. The top edge has length $i+j$. Let h the other dimension of the box. We label w by $L(w) = (I; \{i,j\}, h)$ if w is inflating, by $L(w) = (A; i, j, h)$ otherwise. If $i = j$, the set $\{i,j\}$ has to be understood as the multiset where i is repeated twice.

It is not hard to realize that the labels of the children of a node w only depend on $L(w)$. These labels are described by the left part of the rewriting rules given below (ignoring for the moment what follows the vertical bar). As i, j and h are bounded by n, the calculation of the extension numbers needed to generate a random walk of length n will require $O(n^4)$ operations.

The label $L(w)$ does not determine the last step s of w. For the generating stage, we refine it into a P-label as follows:

- for a walk of type I, we keep track of the *ordered* pair (i,j) and of the last step s; hence the P-label reads $(I; \{i,j\}, h \,|\, i, j, s)$,
- for a walk of type A, we keep track of the last step s and of its direction d around the box ($d = 1$ if s goes in a clockwise direction around the box, $d = -1$ otherwise). The refined label thus reads $(A; i, j, h \,|\, s, d)$.

The combined rewriting rule reads:

$$(I; \{i,j\}, h \mid i, j, s) \to \begin{cases} (I; \{i,j\}, h+1 & | \quad i, j, s) & \\ (A; i-1, j+1, h & | \quad s+1, 1) & \text{if } i > 0 \\ (A; j-1, i+1, h & | \quad s-1, -1) & \text{if } j > 0 \\ (I; \{0,h\}, j+1 & | \quad h, 0, s+1) & \text{if } i = 0 \\ (I; \{0,h\}, i+1 & | \quad 0, h, s-1) & \text{if } j = 0 \end{cases},$$

$$(A; i, j, h \mid s, d) \to \begin{cases} (I; \{i,j\}, h+1 & | \quad i \overset{d}{-} j, j \overset{d}{-} i, s-d) & \\ (A; i-1, j+1, h & | \quad s, d) & \text{if } i > 0 \\ (I; \{0,h\}, j+1 & | \quad h \overset{d}{-} 0, 0 \overset{d}{-} h, s) & \text{if } i = 0 \end{cases},$$

with

$$i \overset{d}{-} j = \begin{cases} i & \text{if } d = 1, \\ j & \text{if } d = -1. \end{cases} \tag{27}$$

Again, we have written $s = 0$ (resp. 1, 2, 3) if s goes North (resp. East, South, West), and these values are taken modulo 4. The four children of the root have labels $(I; \{0,0\}, 1 \mid 0, 0, s)$ for $s = 0, 1, 2, 3$.

6.4. **Three-sided prudent walks**

A non-empty 3-sided walk can be of five different types, I_v, I_h, A, C, or F, depending on the nature of its final step s:

- (I_v) either s is a *vertical* inflating step (it moves the North edge of the box),
- (I_h) or s is a *horizontal* inflating step (it moves the East or West edge of the box),
- (A) or s walks *along* the top edge,
- (C) or s walks upwards on a vertical edge, moving the endpoint of the walk *closer* to the top end of the edge,
- (F) or s walks downwards on a vertical edge, moving the endpoint *further* away from the top end of the edge.

For a walk w of type I_v or A, we denote by i and j the distances between the endpoint of the walk and the NE and NW corners, with the same convention as for general prudent walks (Section 6.3). Otherwise, i denotes the distance between the endpoint of w and the top end of the vertical edge where w ends, and j denotes the width of the box. The label of a walk w of type T is $L(w) = (I_v; \{i,j\})$ if $T = I_v$ and $(T; i, j)$ otherwise. Then L is a good labelling, described by the first part of the rewriting rules given below. The calculation of the extension numbers needed to generate a random walk of length n requires $O(n^3)$ operations.

The L-labelling does not determine the last step s of w. For the generating stage, we refine it as follows:

- for a walk of type I_v, we keep track of the ordered pair (i, j),
- for a walk of type A or I_h, we keep track of the last step s,
- for a walk of type C or F, we keep track of the direction d of the last step around the box: $d = 1$ if the last step goes in a clockwise direction around the box or consists only of South steps, $d = -1$ otherwise.

The last step is explicitly given by the refined label for types A and I_h, and is otherwise obtained via the last step function S defined by $S(I_v; \ldots) = S(C; \ldots) = 0$, $S(F; \ldots) = 2$. The generating rules read, with the notation (27):

$$(I_v; \{i,j\} \mid i,j\ ) \quad \to \quad \left\{ \begin{array}{lll} (I_v; \{i,j\} & \mid & i,j) \\ (A; i-1, j+1 & \mid & 1) \quad \text{if } i > 0 \\ (A; j-1, i+1 & \mid & 3) \quad \text{if } j > 0 \\ (I_h; 0, j+1 & \mid & 1) \quad \text{if } i = 0 \\ (I_h; 0, i+1 & \mid & 3) \quad \text{if } j = 0 \end{array} \right. ,$$

$$(I_h; i,j \mid s) \quad \to \quad \left\{ \begin{array}{lll} (I_h; i, j+1 & \mid & s) \\ (F; i+1, j & \mid & 2-s) \\ (C; i-1, j & \mid & s-2) \quad \text{if } i > 0 \\ (I_v; \{0,j\} & \mid & j \xrightarrow{s-2} 0, 0 \xrightarrow{s-2} j) \quad \text{if } i = 0 \end{array} \right. ,$$

$$(A; i,j \mid s) \quad \to \quad \left\{ \begin{array}{lll} (I_v; \{i,j\} & \mid & j \xrightarrow{s-2} i, i \xrightarrow{s-2} j) \\ (A; i-1, j+1 & \mid & s) \quad \text{if } i > 0 \\ (I_h; 0, j+1 & \mid & s) \quad \text{if } i = 0 \end{array} \right. ,$$

$$(C; i,j \mid d) \quad \to \quad \left\{ \begin{array}{lll} (I_h; i, j+1 & \mid & 2+d) \\ (C; i-1, j & \mid & d) \quad \text{if } i > 0 \\ (I_v; \{0,j\} & \mid & j \xrightarrow{d} 0, 0 \xrightarrow{d} j) \quad \text{if } i = 0 \end{array} \right. ,$$

$$(F; i,j \mid d) \quad \to \quad \left\{ \begin{array}{lll} (F; i+1, j & \mid & d) \\ (I_h; i, j+1 & \mid & 2-d) \\ (I_h; i, 1 & \mid & 3) \quad \text{if } j = 0 \end{array} \right. .$$

The four children of the root have labels $(I_h; 0, 1|1)$, $(I_h; 0, 1|3)$, $(I_v; \{0, 0\}|0, 0)$ and $(F; 1, 0|1)$.

6.5. **Triangular prudent walks**

The principles that underlie the generation of triangular prudent walks are the same as for prudent walks on the square lattice (Section 6.3). The type of a walk w depends on whether its last step s is *inflating* (type I) or walks *along* the box (type A). Consider the last edge of the box that has changed while constructing w. Assume it is the right edge (the other cases are treated similarly after a rotation). If s points up (NW or NE), then i (resp. j) denotes the distance between the endpoint of w and the top vertex (resp. bottom right vertex) of the box. Otherwise, the roles of i and j are exchanged. Define $L(w) = (T; i, j)$, where T is the type of w. Then L is a good labelling. From this, one computes the extension numbers needed to generate a walk of length n in time $O(n^3)$.

During the generation stage, we also keep track of the last step s of w, and of its direction d around the box: if the last edge that has changed is the right one, then $d = 1$ if s goes East of South-East, $d = -1$ otherwise. The other cases are treated similarly after a rotation. The

combined rewriting rules read:

$$(I; i, j \mid s, d) \quad \rightarrow \quad \begin{cases} (I; i, j+1 & \mid \quad s, d) & \\ (I; j, i+1 & \mid \quad s-d, -d) & \\ (A; j-1, i+1 & \mid \quad s-2d, -d) & \\ (A; i-1, j+1 & \mid \quad s+d, d) & \text{if } i > 0 \\ (I; 0, j+1 & \mid \quad s+d, -d) & \text{if } i = 0 \\ (I; j, 1 & \mid \quad s+2d, d) & \text{if } i = 0 \end{cases},$$

$$(A; i, j \mid s, d) \quad \rightarrow \quad \begin{cases} (I; i, j+1 & \mid \quad s-d, d) & \\ (I; j, i+1 & \mid \quad s-2d, -d) & \\ (A; i-1, j+1 & \mid \quad s, d) & \text{if } i > 0 \\ (I; 0, j+1 & \mid \quad s, -d) & \text{if } i = 0 \\ (I; j, 1 & \mid \quad s+d, d) & \text{if } i = 0 \end{cases}.$$

Again, the steps are encoded by $0, 1, \ldots, 5$ in clockwise order, and considered modulo 6. The NW step is encoded by 0. The six children of the root have labels $(I; 0, 1 \mid s, d)$, with $(s, d) = (0, -1), (1, 1), (2, -1), (3, 1), (4, -1), (5, 1)$.

7. Some questions and perspectives

The enumeration of general prudent walks remains an open problem, although we have established a functional equation defining their generating function (Lemma 3). Recall that this g.f. is expected to be non-D-finite, and to have the same radius of convergence as the g.f.s of 2-sided and 3-sided walks (Propositions 6 and 8) [10].

It may be worth investigating the number of prudent walks that span a given box. This number is indeed simple in the case of triangular prudent walks (Proposition 9), even though the length generating function is not D-finite.

Regarding random generation, we think that we have optimized the step-by-step recursive approach. However, it only provides walks with a few hundred steps. We believe that a better understanding of our results, and of the combinatorics of prudent walks, should lead to more efficient sampling procedures.

All the walks we have been able to count move away from the origin at a positive speed. Would it be possible to introduce a non-uniform distribution on these walks that would make them significantly more compact? Recall that the end-to-end distance in an n-step self-avoiding walk is expected to grow like $n^{3/4}$ only.

In the same spirit, it should be possible to study *kinetic* versions of prudent walks, where at each tick of the clock, the walk takes, with equal probability, one of the available prudent steps. The resulting distribution on n-step walks is no longer uniform. This version is actually the one that was studied in [30] and [27]. Clearly, the random generation of these kinetic walks can be performed in linear time, and does not require any precomputation (in terms of generating trees, one chooses uniformly one of the children). Random trajectories differ significantly from the uniform case (Fig. 13). Their behaviour has very recently been explained by Beffara *et al.* [4].

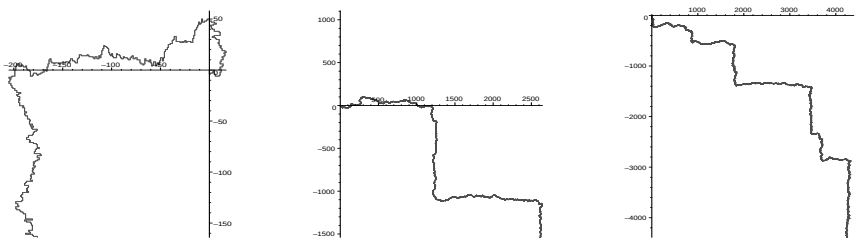

FIGURE 13. Kinetic prudent walks with 1000, 10000 and 20000 steps.

Acknowledgements. I am indebted to Tony Guttmann for discussions on prudent walks, which gave a new impulse to my interest in these objects.

References

[1] G. E. Andrews. *The theory of partitions*. Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1976. Encyclopedia of Mathematics and its Applications, Vol. 2.

[2] O. Angel, I. Benjamini, and B. Virág. Random walks that avoid their past convex hull. *Electron. Commun. Probab.*, 8:6–16, 2003.

[3] C. Banderier, M. Bousquet-Mélou, A. Denise, P. Flajolet, D. Gardy, and D. Gouyou-Beauchamps. Generating functions for generating trees. *Discrete Math.*, 246(1-3):29–55, 2002.

[4] V. Beffara, S. Friedli, and Y. Velenik. Scaling limit of the prudent walk. ArXiv:0902.4312.

[5] M. Bousquet-Mélou. Walks in the quarter plane: Kreweras' algebraic model. *Ann. Appl. Probab.*, 15(2):1451–1491, 2005.

[6] M. Bousquet-Mélou and M. Mishna. Walks with small steps in the quarter plane. ArXiv:0810.4387.

[7] M. Bousquet-Mélou and M. Petkovšek. Linear recurrences with constant coefficients: the multivariate case. *Discrete Math.*, 225(1-3):51–75, 2000.

[8] M. Bousquet-Mélou and M. Petkovšek. Walks confined in a quadrant are not always D-finite. *Theoret. Comput. Sci.*, 307(2):257–276, 2003.

[9] M. Bousquet-Mélou and G. Xin. On partitions avoiding 3-crossings. *Sém. Lothar. Combin.*, 54:Art. B54e, 21 pp. (electronic), 2005/06.

[10] J. C. Dethridge and A. J. Guttmann. Prudent self-avoiding walks. *Entropy*, 8:283–294, 2008.

[11] E. Duchi. On some classes of prudent walks. In *FPSAC'05*, Taormina, Italy, 2005.

[12] P. Duchon, P. Flajolet, G. Louchard, and G. Schaeffer. Boltzmann samplers for the random generation of combinatorial structures. *Combin. Probab. Comput.*, 13(4-5):577–625, 2004.

[13] G. Fayolle, R. Iasnogorodski, and V. Malyshev. *Random walks in the quarter-plane: Algebraic methods, boundary value problems and applications*, volume 40 of *Applications of Mathematics*. Springer-Verlag, Berlin, 1999.

[14] P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.

[15] P. Flajolet, P. Zimmerman, and B. Van Cutsem. A calculus for the random generation of labelled combinatorial structures. *Theoret. Comput. Sci.*, 132(1-2):1–35, 1994.

[16] A. J. Guttmann. Personal communication. 2005.

[17] E. J. Janse van Rensburg, T. Prellberg, and A. Rechnitzer. Partially directed paths in a wedge. *J. Combin. Theory Ser. A*, 115(4):623–650, 2008. ArXiv math/0609834.

[18] I. Jensen and A. J. Guttmann. Self-avoiding polygons on the square lattice. *J. Phys. A*, 32(26):4867–4876, 1999.

[19] T. Kennedy. A faster implementation of the pivot algorithm for self-avoiding walks. *J. Statist. Phys.*, 106(3-4):407–429, 2002.

[20] G. F. Lawler, O. Schramm, and W. Werner. On the scaling limit of planar self-avoiding walk. In *Fractal geometry and applications: a jubilee of Benoît Mandelbrot, Part 2*, volume 72 of *Proc. Sympos. Pure Math.*, pages 339–364. Amer. Math. Soc., Providence, RI, 2004.

[21] N. Madras and G. Slade. *The self-avoiding walk*. Probability and its Applications. Birkhäuser Boston Inc., Boston, MA, 1993.

[22] M. Mishna. Classifying lattice walks restricted to the quarter plane. *J. Combin. Theory Ser. A*, 116(2):460–477, 2009. Arxiv:math/0611651.

[23] M. Mishna and A. Rechnitzer. Two non-holonomic lattice walks in the quarter plane. *Theoret. Comput. Sci.*, 410(38-40):3616–3630, 2009. Arxiv:math/0701800.

[24] A. Nijenhuis and H. S. Wilf. *Combinatorial algorithms*. Academic Press Inc., New York, second edition, 1978.

[25] P. Préa. Exterior self-avoiding walks on the square lattice. Manuscript, 1997.

[26] H. Prodinger. The kernel method: a collection of examples. *Sém. Lothar. Combin.*, 50:Art. B50f, 19 pp. (electronic), 2003/04.

[27] S. B. Santra, W. A. Seitz, and D. J. Klein. Directed self-avoiding walks in random media. *Phys. Rev. E*, 63(6):067101–1–4, 2001.

[28] R. P. Stanley. *Enumerative combinatorics. Vol. 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997.

[29] R. P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

[30] L. Turban and J.-M. Debierre. Self-directed walk: a Monte Carlo study in two dimensions. *J. Phys. A*, 20(3):679–686, 1987.

CNRS, LaBRI, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence Cedex, France
E-mail address: mireille.bousquet@labri.fr