

Exact solution of two classes of prudent polygons

Uwe Schwerdtfeger

Fakultät für Mathematik

Universität Bielefeld

Postfach 10 01 31, 33501 Bielefeld, Germany

November 1, 2021

Abstract

Prudent walks are self-avoiding walks on a lattice which never step into the direction of an already occupied vertex. We study the closed version of these walks, called prudent polygons, where the last vertex of the walk is adjacent to its first one. More precisely, we give the half-perimeter generating functions of two subclasses of prudent polygons on the square lattice, which turn out to be algebraic and non-D-finite, respectively.

1 Introduction

The enumeration of self-avoiding walks (SAW) and polygons (SAP) on a lattice by their number of steps [11] is a long standing problem in combinatorics. Extrapolation of series data from exact enumeration has led to high precision estimates of the exponential growth rate and subexponential corrections but an exact solution of either problem (i.e. finding the generating function, see below) seems out of reach. Rechnitzer [16] has shown that the anisotropic generating function of SAPs on the square lattice is not D-finite. A (possibly multivariate) function $f(\mathbf{z})$ is *D-finite*, if the vector space over $\mathbb{C}(\mathbf{z})$ spanned by its derivatives is finite dimensional. In the univariate case this means that f is a solution of a homogeneous linear ordinary differential equation with polynomial coefficients. At present one tries to find solvable subclasses with large exponential growth rates. This approach is particularly successful in two dimensions. We will restrict to the square lattice in our paper.

One promising class are so-called *prudent walks* (PW) [5, 14], which are SAWs never stepping towards an already occupied vertex. Note that a general prudent walk is not *reversible*, i.e. the walk traversed backwards from its terminal vertex to its initial vertex may not be prudent. Since SAWs are counted modulo translation, we may choose

the initial vertex of a PW to be the origin $(0,0)$. The full problem of PW is unsolved, but recently Bousquet-Mélou [2] succeeded in enumerating a substantial subclass. We adopt the terminology of her paper and use the same methods to obtain the generating functions for the corresponding polygon models defined below. Every nearest neighbour walk on the square lattice has a minimal bounding rectangle containing it, referred to as the *box* of the walk. It is easy to see that each unit step of a prudent walk ends on the boundary of its current box. (This is *not* a characterisation of PWs, e.g. the walk $(0,0) \rightarrow (0,1) \rightarrow (1,1) \rightarrow (2,1) \rightarrow (2,0) \rightarrow (1,0)$ is not prudent.) This property allows the definition of the following subclasses. Call a PW *one-sided*, if every step ends on the *top* side and *two-sided*, if every step ends on *top* or on the *right* side of the current box. Similarly, a PW is referred to as *three-sided* if every step ends on the *left*, *top* or the *right* side of the current box and additionally each *left step* and each *right step* that ends on the *bottom side* of the current box *inflates* the box.

Remark. *i)* As soon as the width of the box of a PW is greater than one, the latter additional condition is redundant. It rules out certain configurations which can occur only if the box width is equal to one, namely “downward zig-zags” of width one, e.g. $\dots \rightarrow (1,0) \rightarrow (1,-1) \rightarrow (0,-1) \rightarrow (0,-2) \rightarrow (1,-2) \rightarrow (1,-3) \rightarrow \dots$ which needlessly complicate the computations below.

ii) Duchi [5] introduced two-sided and three-sided PWs as *type-1* and *type-2* PWs, respectively. In [4, 7] the authors also employ her notation.

Explicit expressions for the generating functions of one-, two- and three-sided PWs have been found so far confirming the data obtained in [4, 7] by computer enumeration. The first class consists of partially directed walks and has a rational generating function. The second class was shown to have an algebraic generating function by Duchi [5] and recently in [2] the third class was solved and the generating function was found not to be D-finite.

Guttmann [4, 7, 9] proposed to study the polygon version of the problem, meaning walks, whose last vertex is adjacent to the starting vertex. We exclude single edges from this definition. As above, the property of being prudent demands a starting vertex and a terminal vertex. So prudent polygons are *rooted* polygons with a directed root edge. Note further that a prudent polygon (PP) which ends, say, to the right of the origin (i.e. in the vertex $(1,0)$) may never step right of the line $x = 1$, and furthermore if the walk hits that line it has to head directly to the vertex $(1,0)$. So prudent polygons are *directed* in the sense that they contain a corner of their box. Moreover, a k -sided PP can be interpreted as a $(k - 1)$ -sided PW confined in a half-plane, see also Section 6. In this paper we deal with the polygon versions of the two-sided and three-sided walks, referred to as two-sided and three-sided PPs. Enumeration of one-sided PPs is trivial, since these are simply rows of unit cells. We give explicit expressions for the half-perimeter generating functions of two-sided PPs and three-sided PPs and show that the latter is not D-finite, as expected on numerical grounds [4, 7]. To our knowledge three-sided PPs are the first *exactly solved* polygon model with a non-D-finite *half-perimeter* generating function. Enumeration of the full class of PPs remains an open problem, as for the walk case.

Outline: In Section 2 we give functional equations for the generating functions which are

based on decompositions of the classes in question, in Section 3 we solve those by the kernel method [1, 2, 3, 12] and in Section 4 we study the analytic behaviour of the generating functions of two-sided and three-sided prudent polygons. Section 5 is dedicated to the random generation of PPs.

2 Functional equations

In combinatorial enumeration of objects from a class $\mathcal{P}$ (say PPs) with respect to counting parameters $c_1, \dots, c_n : \mathcal{P} \rightarrow \mathbb{Z}_{\geq 0}$ (say perimeter, area etc.) the (multivariate) power series

$$P(x_1, \dots, x_n) = \sum_{\alpha \in \mathcal{P}} x_1^{c_1(\alpha)} \dots x_n^{c_n(\alpha)}$$

is called the *generating function* of $\mathcal{P}$. In the following we will count prudent polygons by half-perimeter and other, so-called *catalytic counting parameters*. The variables in the generating function marking the latter are called *catalytic variables*. Their introduction allows us to translate certain combinatorial decompositions into non-trivial functional equations for the associated generating functions [8]. Furthermore, we identify a PP (a “closed” PW) with the collection of unit cells it encloses and build larger PPs from smaller ones by attaching unit cells in a prudent fashion, i.e. the new boundary walk with the same initial vertex remains prudent.

A two-sided prudent polygon either ends at the vertex above the origin or at the vertex to the right of it. This partitions two-sided PPs into two subsets, which can be transferred into each other by the reflection in the diagonal $x = y$. So it suffices to enumerate prudent polygons ending on the top of their box. Here two cases occur, namely the “degenerate” case when the first steps of the walk are left. The resulting PP is simply a row of unit cells pointing to the left. These have a half-perimeter generating function $t^2 + t^3 + \dots = t^2/(1-t)$. In the “generic case” such a PP is a bar graph turned upside down, i.e. a column convex polyomino containing the top side of its bounding box, cf. Figure 1. Denote by $B(t, u, w)$

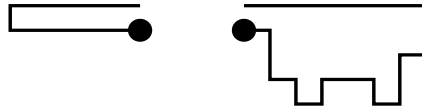


Figure 1: Degenerate (left) and generic 2-sided PPs ending on the top of the box

the generating function of bar graphs counted by half-perimeter, width and height of the rightmost column (catalytic parameter), marked by t , u and w respectively. Here w is the catalytic variable. The width parameter is not a catalytic parameter. However, it will be important in the study of three-sided PPs. We follow the lines of [1].

Lemma 2.1. *The generating function $B(t, u, w)$ of bar graphs satisfies the functional equation*

$$B(t, u, w) = u \left(\frac{t^2 w}{1 - wt} + \frac{wt (B(t, u, 1) - B(t, u, w))}{1 - w} + \frac{B(t, u, w) t^2 w}{1 - wt} \right). \quad (2.1)$$

Proof. A bar graph is either a single column, or it is obtained by attaching a new column to the right side of a bar graph. The decomposition is sketched in Figure 2.

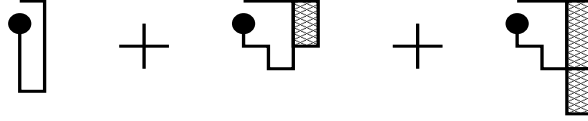


Figure 2: Illustration of the decomposition underlying functional equation (2.1)

Single columns of height ≥ 1 contribute $ut^2w/(1 - wt)$ to the generating function. The polygons obtained by adding a column which is shorter than or equal to the old rightmost column contribute the second summand. This is seen as follows. A polygon of half-perimeter n , width k and rightmost column height l contributing $t^n u^k w^l$ to $B(t, u, w)$ gives rise to l polygons whose rightmost column is shorter or equal. Their contribution sums up to

$$tu \sum_{j=1}^l t^n u^k w^j = tuw \frac{t^n u^k 1^l - t^n u^k w^l}{1 - w}. \quad (2.2)$$

Summing this over all polygons gives the second summand. The third summand corresponds to adding a larger column. To this end duplicate the rightmost column and attach a non-empty column below the so obtained new rightmost column. A so obtained bar graph can be viewed as an ordered pair of a bar graph and a column. The generating function of those pairs is the third summand of the rhs. This finishes the proof. $\square$

The walk constituting the boundary of a three-sided PP has $(0, 0)$ as its initial vertex and $(1, 0)$ or $(-1, 0)$ or $(0, 1)$ as its terminal vertex. Those walks with terminal vertex $(0, 1)$ may not step above the line $y = 1$ and they have to move directly to the vertex $(0, 1)$ as soon as they step upon that line. This leads to two sorts of bar graphs either rooted on their left or on their right side, see Figure 3.



Figure 3: Three-sided PPs with terminal vertex $(0, 1)$ are bar graphs

So only those three-sided PPs are of further interest, which end in $(1, 0)$ or $(-1, 0)$. Both classes are transformed into each other by a reflection in the line $x = 0$. We study those

ending to the right of the origin in the vertex $(1, 0)$. Again a degenerate and a generic case are distinguished, according to whether such a PP reaches its terminal vertex from below via the vertex $(1, -1)$ (“counterclockwise around the origin”) or from above, via the vertex $(1, 1)$ (“clockwise”). In the degenerate case we simply obtain a single column. In the generic case, a (possibly empty) sequence of initial down steps is followed by a left step. So denote by $R(t, u, w)$ the generating function of the generic three-sided PPs ending in the vertex $(1, 0)$ counted by half-perimeter, the length of the top row and the distance of the top left corner of the top row and the top left corner of the box, marked by t , u , and w , respectively, cf. Figure 4. Here both u and w are catalytic variables.

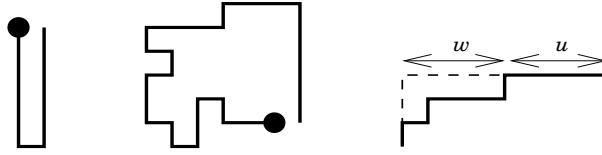


Figure 4: Degenerate and generic three-sided PPs, catalytic variables

Lemma 2.2. *The generating function $R(t, u, w)$ of generic three-sided PPs satisfies the functional equation*

$$R(t, u, w) = ut (B(t, u) + t) + \frac{ut (R(t, w, w) - R(t, u, w))}{w - u} + \frac{ut^2 (R(t, u, w) - R(t, u, ut))}{w - ut} + R(t, u, ut) ut (B(t, u) + t), \quad (2.3)$$

where $B(t, u) := B(t, u, 1)$ is the generating function of bar graphs counted by half-perimeter and width.

Proof. The decomposition we use is sketched in Figure 5.

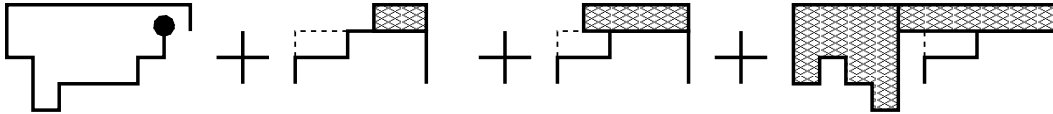


Figure 5: Illustration of the decomposition underlying functional equation (2.3)

The polygons in question contain the top right corner of their box. This corner is some point $(1, y)$. If $y = 1$, then the PP is either the unit square containing $(0, 0)$ and $(1, 1)$ or a bar graph as above with that unit square glued to the right. This yields the first summand. A PP with $y > 1$ is obtained in one of the following three ways from a PP with top right corner $(1, y - 1)$. The first is to add a new row on top, which is shorter than or equal to the original top row. A similar computation as in (2.2) (with some additional book keeping

on w) yields the second summand. The second way to obtain a larger PP from a smaller one is by adding a new row on top, which is longer than the original top row, but does not inflate the box to the left. Again a treatment similar to the computation in (2.2) yields the third summand. The third way to extend a PP is to add a row on top of length equal to the width of the box plus one and possibly an arbitrary bar graph. This finally yields the fourth summand and the functional equation is complete. $\square$

Remark. As in the case of general SAPs [10] we can define a concatenation of two three-sided PPs. Roughly speaking, the one PP can be enlarged by inserting the other one at the top corner of the leftmost column, see Figure 6.

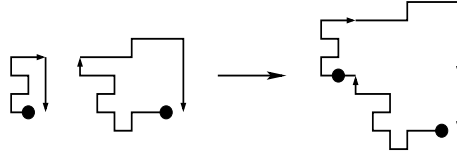


Figure 6: Concatenating two 3-sided PPs

The numbers $pp_3^{(m)}$ of three-sided PPs hence satisfy $pp_3^{(m+n)} \geq pp_3^{(m)} \cdot pp_3^{(n)}$. This implies the existence of a connective constant β , i.e. a representation $pp_3^{(m)} = \exp(\beta m + o(m))$. The precise value for β and the subexponential corrections are given in Section 4. The converse inequality holds for prudent walks, since breaking an $m + n$ step PW after m steps leaves one with a pair of prudent walks of respective lengths m and n .

We now turn to unrestricted PPs. They can be partitioned into eight subclasses according to their end point being $(1, 0)$, $(0, 1)$, $(-1, 0)$ or $(0, -1)$ and their orientation (clockwise or counterclockwise around the origin). All eight classes can be transformed into each other by symmetry operations of the square. Hence it suffices to enumerate those PPs ending in $(1, 0)$ which reach their endpoint via the vertex $(1, 1)$ (clockwise). We denote this class by $\mathcal{F}$ and by $F(u, w, x) := F(t, u, w, x)$ its generating function. $G(u, w, x) := G(t, u, w, x)$ and $H(u, w, x) := H(t, u, w, x)$ are defined as the generating functions of the two auxiliary subclasses $\mathcal{G} \supseteq \mathcal{H}$ of $\mathcal{F}$ specified below. We have the following functional equation.

Lemma 2.3. *The power series F , G and H satisfy a system of functional equations. For $X = F, G, H$ the single equations are of the form*

$$\begin{aligned} X(u, w, x) = & \frac{tux(X(w, w, x) - X(u, w, x))}{w - u} \\ & + \frac{t^2ux(X(u, w, x) - X(u, ut, x))}{w - ut} + I_X(u, w, x), \end{aligned} \quad (2.4)$$

where the formal power series $I_X(u, w, x) := I_X(t, u, w, x)$ is equal to

$$I_X(u, w, x) := \begin{cases} G(x, x, u) & \text{if } X=F, \\ t^2ux + t^2uxF(x, xt, w) + xH(x, x, u) & \text{if } X=G, \\ t^2uxw^{-1}G(x, xt, w) & \text{if } X=H. \end{cases}$$

Proof. The proof relies on a decomposition similar to that of three-sided PPs. The PPs of the class $\mathcal{F}$ all contain the top right corner of their box. In the generating function $F(u, w, x)$ of $\mathcal{F}$ the variable u marks the length of the top row, w marks the distance of the top left corner of the top row to that of the box and x marks the height of the box.

We define the classes $\mathcal{G}$ and $\mathcal{H}$. $\mathcal{G}$ consists of the unit square together with those PPs in $\mathcal{F}$ which are obtained by attaching a piece (a collection of unit cells) on top of a given PP in $\mathcal{F}$, such that the top side of the box is shifted by one unit, the left side by at least one unit and the bottom side by an arbitrary number of units, see Figure 7.

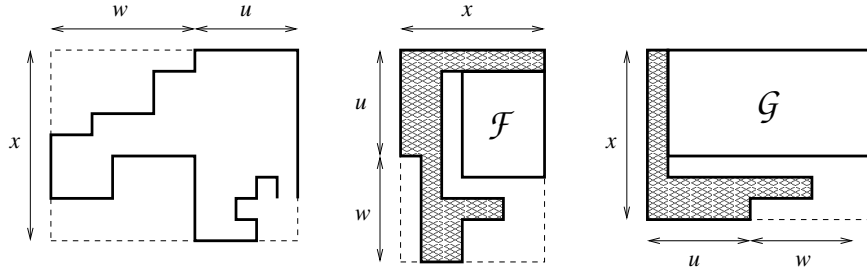


Figure 7: Illustrations of the classes $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{H}$, catalytic variables

Note, that the polygons in $\mathcal{G}$ contain the top left corner of their box. The catalytic variable u in the generating function $G(u, w, x)$ marks the length of the leftmost column, w marks the distance of the lower left corner of the leftmost column to the bottom left corner of the box and x marks the width of the box. The class $\mathcal{H}$ is the subclass of $\mathcal{G}$ obtained by glueing a piece to the left of the leftmost column and thereby shifting the left side of the box by exactly one unit and the bottom side by at least one unit, see Figure 7. The variable u in $H(u, w, x)$ marks the length of the bottom-most row, w marks the distance minus one of the lower right corner of that row to the bottom right corner of the box. The variable x marks the height of the box.

Now the functional equations are derived similarly to the three-sided case. Every PP in $\mathcal{F}$ can be extended by adding a new row on top which is shorter than or equal to the old top row or longer than the old top row, but does not inflate the box to the left. These two operations contribute the first and the second summand in the equation for $X = F$, as in the proof of equation (2.3). Inflating the box to the left yields a PP in $\mathcal{G}$, explaining the expression for I_F .

As for the functional equation $G(u, w, x)$ of the class $\mathcal{G}$ the first two summands on the rhs correspond to adding a new column to the left, the expression $xH(x, x, u)$ to adding a piece which shifts the bottom boundary of the box, in an analogous fashion as above.

The unit square contributes t^2ux , the term $t^2uxF(x, xt, w)$ corresponds to the “minimal” polygons in $\mathcal{G}$ obtained by adding a top row on an arbitrary PP of length equal to the width of the box plus one.

The minimal PPs in $\mathcal{H}$ are those obtained by extending a PP in $\mathcal{G}$ adding a column to the left of length equal to the height of the box plus one. This explains the term for I_H . The rest of the rhs corresponds to adding a new bottom row. $\square$

3 Solution by the kernel method

The following result has already been obtained in [15] as the solution of an algebraic equation which arises from a different decomposition of the class. We derive it here for completeness and to recall the “classic” kernel method as applied in [1].

Theorem 3.1. *The generating function $B(t, u) := B(t, u, 1)$ of bar graphs counted by half-perimeter and width is equal to*

$$B(t, u) = \frac{1 - t - u(1 + t)t - \sqrt{t^2(1 - t)^2u^2 - 2t(1 - t^2)u + (1 - t)^2}}{2tu}. \quad (3.1)$$

Proof. The functional equation 2.1 is equivalent to

$$0 = (t^2uw(1 - w) - uwt(1 - wt) - (1 - w)(1 - wt)) B(t, u, w) + tuw(1 - wt)B(t, u, 1) + t^2uw(1 - w) \quad (3.2)$$

The kernel equation

$$0 = (t^2uw(1 - w) - uwt(1 - wt) - (1 - w)(1 - wt))$$

is a quadratic equation in the catalytic variable w and has the following unique *power series* solution $q(t, u)$

$$q(t, u) = \frac{1 + (1 - u)t + ut^2 - \sqrt{t^2(1 - t)^2u^2 - 2t(1 - t^2)u + (1 - t)^2}}{2t}. \quad (3.3)$$

Upon substituting $w = q(t, u)$ into Eq. (3.2), the terms with $B(t, u, w)$ are cancelled and we can solve for $B(t, u, 1)$, which leads to (3.1). $\square$

Remark. In principle, $B(t, u, w)$ can also be computed, by substituting the result for $B(t, u, 1)$ into Eq. (3.2).

By setting $u = w = 1$ in the bar graph generating function, adding the contribution of the degenerate two-sided PPs and multiplication by 2, we obtain the following result so far conjectured by series extrapolation from exact enumeration data [4].

Corollary 3.2. *The generating function of two-sided prudent polygons is equal to*

$$\begin{aligned}
PP_2(t) &= \frac{1}{t} \left(\frac{1 - 3t + t^2 + 3t^3}{1 - t} - \sqrt{(1 - t)(1 - 3t - t^2 - t^3)} \right) \\
&= 4z^2 + 6z^3 + 12z^4 + 28z^5 + 72z^6 + 196z^7 + 552z^8 + 1590z^9 \\
&\quad + 4656z^{10} + 13812z^{11} + 41412z^{12} + 125286z^{13} + 381976z^{14} \\
&\quad + 1172440z^{15} + 3620024z^{16} + 11235830z^{17} + 35036928z^{18} \\
&\quad + 109715014z^{19} + 344863872z^{20} + \dots
\end{aligned} \tag{3.4}$$

Now we turn to the three-sided case. Note that the sum of the catalytic counting parameters, namely the length of the top row and the distance of its top left corner to the top left corner of the box, is equal to the width of the polygon. We have the following result for the generic three-sided PPs ending on the right. It is derived in a similar way as the corresponding result on PWs in [2].

Theorem 3.3. *The functional equation 2.3 has a unique power series solution. For the generating function $R(t, w, w)$ of generic three-sided prudent polygons ending on the right and counted by half-perimeter and width we have an explicit expression as a an infinite sum of formal power series*

$$R(t, w, w) = \sum_{k \geq 0} L\left((tq^2)^k w\right) \prod_{j=0}^{k-1} K\left((tq^2)^j w\right). \tag{3.5}$$

Here

$$q := q(t, 1) = \frac{t^2 + 1 - \sqrt{1 - 4t + 2t^2 + t^4}}{2t}, \tag{3.6}$$

with $q(t, u)$ as in (3.3) in the proof of Theorem 3.1. K and L are given by

$$K(w) = \frac{(1 - t)q - 1 - ((1 - t + t^2)q - 1)(B(t, qw) + t)w}{1 - t(1 + t)q - (t(1 - t - t^3)q + t^2)(B(t, qw) + t)w} \tag{3.7}$$

and

$$L(w) = \frac{(1 + t^2 - (1 - 2t + 2t^2 + t^4)q)(B(t, qw) + t)w}{1 - t(1 + t)q - (t(1 - t - t^3)q + t^2)(B(t, qw) + t)w}, \tag{3.8}$$

where $B(t, u)$ is the generating function of bar graphs as in (3.1).

Proof. The functional equation (2.3) is equivalent to

$$\begin{aligned}
0 &= \left(ut^2(w - u) - ut(w - ut) - (w - u)(w - ut) \right) R(t, u, w) \\
&\quad + \left(ut(B(t, u) + t)(w - u)(w - ut) - t^2u(w - u) \right) R(t, u, ut) \\
&\quad \quad \quad + ut(w - ut)R(t, w, w) \\
&\quad \quad \quad + ut(w - u)(w - ut)(B(t, u) + t).
\end{aligned} \tag{3.9}$$

We first solve the kernel equation

$$(ut^2(w - u) - ut(w - ut) - (w - u)(w - ut)) = 0$$

for u and w . The unique power series solutions are $U(t, w) = q(t)w$ resp. $W(t, u) = q(t)tu$, with $q(t)$ as in (3.6). We substitute $w = W(t, u)$ in Eq. (3.9) and obtain an expression for $R(t, u, ut)$ in terms of $R(t, qtu, qtu)$, namely

$$R(t, u, ut) = \frac{(q - 1) R(t, qtu, qtu) + (1 - q)(1 - tq)(B(t, u) + t)u}{1 - qt - (1 - q)(1 - tq)(B(t, u) + t)u}. \quad (3.10)$$

Substitute this into Eq. (3.9) and set $u = U(t, w)$. This relates $R(t, w, w)$ and $R(t, wtq^2, wtq^2)$ as follows:

$$R(t, w, w) = K(w) \cdot R(t, wtq^2, wtq^2) + L(w), \quad (3.11)$$

with $K(w)$ and $L(w)$ as in (3.7) and (3.8). $K(w)$ and $L(w)$ are a formal power series in t , which is seen as follows: $B(t, qw)$ is well-defined as a formal power series in t as $[t^N] B(t, u)$ is a polynomial in u of degree at most $N - 1$. Furthermore by the definition of B we see $B(t, u) = t^2u + O(t^3)$. The denominator is now easily seen to be $1 + O(t)$, so both $K(w)$ and $L(w)$ are well-defined as formal power series in t . Inspecting the first few coefficients we see $(1 - t)q - 1 = O(t^3)$ and $1 - (1 - t + t^2)q = O(t^2)$, so the numerator of $K(w)$ is $O(t^3)$. In a similar way the numerator of $L(w)$ is seen to be $w \cdot O(t^2)$. Moreover we have $tq^2 = t + O(t^2)$. So we can iterate Eq. (3.11) and obtain formula (3.5). $\square$

Remark. *i)* We have the following alternative expressions for $K(w)$ and $L(w)$:

$$K(w) = \frac{((1 - q)(1 - qt)qwt(B(t, qw) + t) + t^2q(q - 1))(q - 1)}{q(1 - qt)^2((1 - q)qwt(B(t, qw) + t) + t)} \quad (3.12)$$

and

$$L(w) = \frac{(1 - qt)(1 - q^2t)(q - 1)qtw(B(t, qw) + t)}{q(1 - qt)^2((1 - q)qwt(B(t, qw) + t) + t)}. \quad (3.13)$$

The expressions (3.7) and (3.8) were obtained by expressing powers of q in terms of t , e.g.

$$\begin{aligned} q^2 &= (t(t^2 + 1)q - t)/t^2, \\ q^3 &= (t(t^4 + 2t^2 - t + 1)q - t^3 - t)/t^3, \\ q^4 &= (tq(t^6 + 3t^4 - 2t^3 + 3t^2 - 2t + 1) - t + t^2 - 2t^3 - t^5)/t^4. \end{aligned}$$

ii) In principle one could also compute $R(t, u, w)$. To obtain the generating function of all three-sided PPs we sum up the contributions of the degenerate PPs and those ending on top, multiply by two and obtain

$$PP_3(t) = 2 \left(\frac{t^2}{1 - t} + B(t, 1) + R(t, 1, 1) \right).$$

The first few terms of the series $PP_3(t)$ are

$$\begin{aligned} PP_3(t) &= 6t^2 + 10t^3 + 24t^4 + 66t^5 + 198t^6 + 628t^7 + 2068t^8 + 7004t^9 + 24260t^{10} \\ &\quad + 85596t^{11} + 306692t^{12} + 1113204t^{13} + 4085120t^{14} + 15131436t^{15} + 56495170t^{16} \\ &\quad + 212377850t^{17} + 803094926t^{18} + 3052424080t^{19} + 11653580124t^{20} + \dots \end{aligned}$$

4 Analytic properties of the generating functions

So far we have considered the generating functions in question as *formal* power series. A crude estimate on the number of SAPs of half-perimeter n is 4^{2n} which is the total number of all nearest neighbour walks on the square lattice of length $2n$. So the series $PP_2(t)$ and $PP_3(t)$ converge at least in the open disc $\{|t| < 1/16\}$ and represent analytic functions there. This section deals with the analytic properties of these functions. We first discuss the analytic structure of the generating function of two-sided PPs.

Proposition 4.1. *The generating function $PP_2(t)$, cf. (3.4), is algebraic of degree 2, with its dominant singularity a square root singularity at $t = \rho$, where ρ is the unique real root of the equation*

$$\frac{1 - 4t + 2t^2 + t^4}{1 - t} = 1 - 3t - t^2 - t^3 = 0. \quad (4.1)$$

With $\theta = \sqrt[3]{26 + 6\sqrt{33}}$ the exact value for ρ can be written as

$$\rho = \frac{\theta^2 - \theta - 8}{3\theta} = 0.2955977\dots$$

The number $pp_2^{(m)}$ of two-sided PPs of half-perimeter m is asymptotically

$$pp_2^{(m)} \sim A \cdot \rho^{-m} \cdot m^{-3/2} \quad (m \rightarrow \infty),$$

where

$$A = \frac{\sqrt{(-37 + 11\sqrt{33})\theta^2 + (-152 + 8\sqrt{33})\theta + 32}}{4\sqrt{6\pi}\rho} = 0.8548166\dots$$

Remark: *i)* The generating function of two-sided prudent *walks* is algebraic with its dominant singularity a simple pole at $\bar{\sigma} = 0.403\dots$. Its coefficients are asymptotically equal to $\kappa \cdot \bar{\sigma}^{-m}$, where $\kappa = 2.51\dots$, cf. [2].

ii) The asymptotic number of bar graphs as well as staircase polygons (counted by half-perimeter) and Dyck paths (by half-length) is of the form $\kappa \cdot \mu^n \cdot n^{-3/2}$. Furthermore, the area random variables in the fixed-perimeter (fixed-length) ensembles of all three models are known to converge weakly to the *Airy distribution* [6, 17, 19].

The analytic structure of $PP_3(t)$ is far more complicated due to the analytic structure of $R(t, 1, 1)$, which is stated in the main result Theorem 4.4. In what follows we make frequent use of the following facts about the series q :

Lemma 4.2. *The series q , $(1 - t)q - 1$, q^2t , $t(1 + t)q$ and $t(1 - t - t^3)q + t^2$ have non-negative integer coefficients. For $|t| \leq \rho$ we have the estimates*

$$|q| \leq \frac{|t|^2 + 1}{2|t|}, \quad |q^2t| \leq 1, \quad |(1 - t)q - 1| \leq \rho, \quad |1 - t(1 + t)q| \geq \rho. \quad (4.2)$$

Equality holds if and only if $t = \rho$. Furthermore

$$q(\rho) = \frac{\rho^2 + 1}{2\rho} = \frac{1}{\sqrt{\rho}}. \quad (4.3)$$

The singular behaviour of $B\left(t, q(q^2t)^N\right)$ and $B(t, qw)$ plays an important role in the study of $R(t, 1, 1)$.

Lemma 4.3. *For $N \geq 0$ the dominant singularity of $B\left(t, q(q^2t)^N\right)$ is σ_N , which is the unique solution in the interval $[0, \rho)$ of the equation*

$$u(t) - q(q^2t)^N = \frac{1}{t} \cdot \frac{1 - \sqrt{t}}{1 + \sqrt{t}} - q(q^2t)^N = 0.$$

In particular, $\sigma := \sigma_0 = \tau^2 = 0.2441312\dots$, where τ is the unique real root of the polynomial $t^5 + 2t^2 + 3t - 2$. The sequence $\{\sigma_N, N \geq 0\}$ is monotonically increasing and converges to ρ . Furthermore $B(t, qw)$ is analytic in the polydisc $\{|t| < \rho\} \times \{|w| < \sqrt{\rho}\}$.

Proof. $B(t, u)$ is singular if and only if

$$t^2(1 - t)^2u^2 - 2t(1 - t^2)u + (1 - t)^2 = 0.$$

The relevant solution $u(t)$ with $u(\rho) = 1$ is

$$u(t) = \frac{1}{t} \cdot \frac{1 - \sqrt{t}}{1 + \sqrt{t}}. \quad (4.4)$$

$B\left(t, q(q^2t)^N\right)$ is singular if $q(q^2t)^N = u(t)$. This equation has a solution σ_N in the interval $(0, \rho)$, as $u(t) \rightarrow 1$ and $q(q^2t)^N \rightarrow (\rho^2 + 1)/2\rho = 1/\sqrt{\rho} > 1$, for $t \rightarrow \rho$. Here u is strictly decreasing and $q(q^2t)^N$ strictly increasing. We further see that σ_N converges to ρ , as for arbitrary fixed t with $0 < t < \rho$ we can choose N sufficiently large, such that $u(t) > q(q^2t)^N$, see Lemma 4.2. So $\sigma_N \geq t$, which shows the convergence. Monotonicity follows, as $q(q^2t)^{N+1} < q(q^2t)^N$ for $t \in (0, \rho)$. All these singularities are square root singularities, as the expressions under the root are analytic in $|t| < \rho$. $B(t, qw)$ is singular, if $w = u(t)/q$ and hence

$$|w| = \frac{|u(t)|}{|q|} \geq \sqrt{\rho}u(\rho) = \sqrt{\rho},$$

with equality if and only if $t = \rho$. So there is no singularity inside the polydisc. $\square$

Now we are ready to state the main result, which is proven in the subsequent lemmas.

Theorem 4.4. *The function $R(t, 1, 1)$ is analytic in the disc $\{|t| < \sigma\}$ with its unique dominant singularity a square root singularity at σ . Moreover it is meromorphic in the slit disc*

$$D_{\sigma, \rho} = \{|t| < \rho\} \setminus [\sigma, \rho),$$

and it has infinitely many square root singularities in the set $\{\sigma_N, N = 0, 1, 2, \dots\}$. In particular, $R(t, 1, 1)$ is not D -finite.

Remark. *i)* The number $pp_3^{(m)}$ of three-sided PPs of half perimeter m is asymptotically equal to $\kappa \cdot \sigma^{-m} \cdot m^{-3/2}$ for some positive constant κ . In particular, two-sided PPs are exponentially rare among three-sided PPs.

ii) The generating function of three-sided prudent *walks* has its dominant singularity a simple pole at $\bar{\sigma} = 0.403\dots$, as in the two-sided case. It is meromorphic in some larger disc of radius $\bar{\rho} = \sqrt{2} - 1$ with infinitely many simple poles in the interval $[\bar{\sigma}, \bar{\rho})$. Its coefficients grow like $\kappa \cdot \bar{\sigma}^{-m}$, for some $\kappa > 0$ [2].

Possible singularities of $R(t, 1, 1)$ in $D_{\sigma, \rho}$ are zeroes of the denominators of $K(w)$ and $L(w)$, places, where the representation (3.5) diverges, and square root singularities of $B(t, q(q^2t)^N)$. Now we investigate the analytic properties of the single summands in the representation (3.5).

Lemma 4.5. 1. $K((q^2t)^N)$ and $L((q^2t)^N)$ are analytic in $\{|t| < \sigma_N\}$.

2. $K((q^2t)^N w)$ and $L((q^2t)^N w)$ are analytic in $\{|t| < \rho\} \times \{|w| < \sqrt{\rho}\}$.

Proof. With the above definition of $u(t)$ and a short computation we obtain the estimate

$$\left| B(t, q(q^2t)^N) \right| < B(|t|, u(|t|)) = \sqrt{|t|}. \quad (4.5)$$

The denominator of $K(w)$ and $L(w)$ is

$$1 - T(t, w) = 1 - t(1+t)q - (t(1-t-t^3)q + t^2)(B(t, qw) + t)w.$$

$T(t, w)$ is a power series in t and w with non-negative coefficients and $T(0, w) = 0$. Hence we have the estimate

$$T(t, (q^2t)^N) \leq T(\sigma_N, (q(\sigma_N)^2 \sigma_N)^N) \leq T\left(\sigma_N, \frac{u(\sigma_N)}{q(\sigma_N)}\right).$$

A computation shows that the function $1 - T(t, u(t)/q(t))$

$$1 - T\left(t, \frac{u(t)}{q(t)}\right) = 1 - t(1+t)q - (t(1-t-t^3)q + t^2)(\sqrt{t} + t)\frac{u(t)}{q(t)}$$

has no zeroes in $[\sigma, \rho]$. This finishes the proof of the first assertion, as $K((q^2t)^N)$ and $L((q^2t)^N)$ do not have poles inside $\{|t| < \sigma_N\}$. Furthermore, the denominator $1 - T(t, (q^2t)^N w)$ is analytic in the polydisc $\{|t| < \rho\} \times \{|w| < \sqrt{\rho}\}$, with the only singular point $(t, w) = (\rho, \sqrt{\rho})$ on its boundary. As above we see

$$\left| T(t, (q^2t)^N w) \right| \leq T(|t|, |w|) \leq T\left(\rho, \frac{u(\rho)}{q(\rho)}\right) = T(\rho, \sqrt{\rho}),$$

and hence the denominator is non-zero in the domain in question and $K((q^2t)^N w)$ and $L((q^2t)^N w)$ are both analytic in the polydisc. $\square$

Lemma 4.6. 1. The series representation (3.5) of $R(t, 1, 1)$ is a series of algebraic functions, which converges compactly in the slit disc $D_{\sigma, \rho} = \{|t| < \rho\} \setminus [\sigma, \rho)$ to a meromorphic function.

2. Furthermore the corresponding representation of $R(t, w, w)$ converges compactly in the polydisc $\{|t| < \rho\} \times \{|w| < \sqrt{\rho}\}$ to an analytic function.

3. The Taylor expansion of $R(t, w, w)$ about $(t, w) = (0, 0)$ converges absolutely in $\{|t| < \rho\} \times \{|w| < \sqrt{\rho}\}$.

Proof. For the first assertion choose $0 < r < \rho$. We look at the disc $\{|t| \leq r\}$. The term independent of w in the numerator of $K(w)$ is strictly less than ρ for $|t| \leq r$ and the corresponding term in the denominator is strictly larger than ρ , see Lemma 4.2. So we can choose N large such that $\sigma_N > r$ and $|K((q^2 t)^N)| < 1$ for $|t| \leq r$. Split the series at N . The summands for $k = 0, \dots, N-1$ sum up to a function which is meromorphic in the slit disc $\{|t| \leq r\} \setminus [\sigma, r]$. In the rest of the series take out the common factors to obtain

$$\prod_{j=0}^{N-1} K((tq^2)^j) \sum_{k \geq 0} L((tq^2)^{N+k}) \prod_{j=0}^{k-1} K((tq^2)^{N+j}). \quad (4.6)$$

The first product is a meromorphic function in the slit disc. $L((tq^2)^{N+k})$ is easily seen to converge uniformly to 0 in $|t| \leq r$ as $k \rightarrow \infty$. In $|t| \leq r$ all summands are holomorphic (see the above discussion) and the sum can be estimated by a geometric series and hence converges uniformly in the compact disc $\{|t| \leq r\}$. By Montel's theorem the limit of the sum is again analytic. This finishes the proof for the first assertion. The second assertion is proven along the similar lines. By the multivariate version of Montel's theorem [18] the limit function is also analytic in the domain in question and thus the third assertion follows. $\square$

Lemma 4.7. $R(t, 1, 1)$ is singular at infinitely many of the σ_N . Furthermore, $R(t, 1, 1)$ is singular at σ .

Proof. Terms singular at σ_N only show up in the summands for $k \geq N$. The sum of these (4.6) is equal to

$$\prod_{j=0}^{N-1} K((tq^2)^j) \left[L((tq^2)^N) + K((tq^2)^N) R(t, (tq^2)^{N+1}, (tq^2)^{N+1}) \right]. \quad (4.7)$$

In order to show that the singularity σ_N does not cancel, only the term in square brackets is of interest. Singular terms show up in the numerators and the common denominator of $K((tq^2)^N)$ and $L((tq^2)^N)$. We now manipulate the expressions (3.12) and (3.13) for $K(w)$ and $L(w)$ in order to get rid of singular terms in the denominator, where the factor

$$(1 - q)qwt(B(t, qw) + t) + t$$

leads to a singularity at σ_N for $w = (tq^2)^N$. Write

$$qwt(B(t, qw) + t) = A(w) - \phi(w),$$

where

$$A(w) = \frac{1}{2}(1 + t - qw(1 + t)t)$$

$$\phi(w) = \frac{1}{2}\sqrt{t^2(1 - t)^2(qw)^2 - 2t(1 - t^2)qw + (1 - t)^2}.$$

Then $A((tq^2)^N)$ is analytic in $\{|t| < \rho\}$. After multiplication of the numerator and denominator with $(1 - q)A(w) + t + (1 - q)\phi(w)$ there is no more occurrence of ϕ in the denominator. We now have to collect the terms involving $\phi(w)$ in the numerators of $K(w)$ and $L(w)$. In the numerator of $K(w)$ the terms involving $\phi(w)$ sum up to

$$P_K(w)\phi(w) := t(q - 1)^2(1 - q^2t)\phi(w).$$

The terms involving $\phi(w)$ in the numerator of $L(w)$ sum up to

$$P_L(w)\phi(w) := (1 - qt)(1 - q^2t)(1 - q)t\phi(w).$$

So the singularity at σ_N can only cancel if

$$-\frac{P_L((\sigma_N q(\sigma_N)^2)^N)}{P_K((\sigma_N q(\sigma_N)^2)^N)} = R\left(\sigma_N, (\sigma_N q(\sigma_N)^2)^{N+1}, (\sigma_N q(\sigma_N)^2)^{N+1}\right). \quad (4.8)$$

In order to prove that this equation can hold for at most finitely many of the σ_N , we show that for σ_N sufficiently close to ρ the lhs of eq. (4.8) is strictly decreasing while the rhs is strictly increasing. Since (σ_N) is monotonically increasing and converges to ρ this will finish the proof. We first prove the assertion on the rhs. The Taylor expansion of $R(t, w, w)$ about $(0, 0)$ has non-negative coefficients and represents $R(t, w, w)$ in the polydisc $\{|t| < \rho\} \times \{|w| < \sqrt{\rho}\}$ by Lemma 4.6. By the definition of σ_N and $u(t)$ we have

$$(\sigma_N q(\sigma_N)^2)^{N+1} = u(\sigma_N) q(\sigma_N) \sigma_N$$

The rhs of the last equation is strictly increasing for sufficiently large N and converges to $\sqrt{\rho}$ as $N \rightarrow \infty$. The sequence σ_N is also strictly increasing by Lemma 4.3. So for large enough N the sequence $R\left(\sigma_N, (\sigma_N q(\sigma_N)^2)^{N+1}, (\sigma_N q(\sigma_N)^2)^{N+1}\right)$ is strictly increasing.

Now we turn to the lhs of eq. (4.8). A computation yields

$$-\frac{P_L((\sigma_N q(\sigma_N)^2)^N)}{P_K((\sigma_N q(\sigma_N)^2)^N)} = \frac{1 - \sigma_N q(\sigma_N)}{q(\sigma_N) - 1},$$

which easily seen to be ultimately strictly decreasing. This finishes the proof of Lemma 4.7. $\square$

The Lemmas 4.5, 4.6 and 4.7 together constitute a proof of Theorem 4.4.

5 Random three-sided prudent polygons

In [2] prudent walks of a given fixed length are generated uniformly at random with a refined version of a method proposed in [13]. We briefly describe a version of the method tailored to our particular needs. The main ingredient in the sampling procedure are *generating trees*. These are rooted trees with their nodes labelled in such a way that if two nodes bear the same label, then the multisets of the labels of their children are the same. In this section we present generating trees and sampling procedures for the various classes of prudent polygons.

The decompositions underlying the functional equations (2.1), (2.3) and (2.4) (see also Figures 2, 5 and 7) yield rules according to which a larger PP from the respective class can be constructed starting from a smaller one. We refine these building steps such that each step increases the half-perimeter by one. The result is a step-by-step procedure which allows to generate any PP of half-perimeter m in a *unique* way, starting from the unit square, such that after the k th step we have a PP of half-perimeter $k+2$, $k = 0, 1, \dots, m-2$. To put it differently, we can organise the polygons in a rooted tree $\mathcal{T}$, with the unit square as the root and the polygons of half-perimeter m as the nodes on level $m-2$. So a random PP of half-perimeter m corresponds to a random path of length $m-2$ in that tree starting from the root. A polygon α is a *child* of a polygon π , if it is obtained by one of the following six construction steps.

1. Attaching a new top row which is shorter than or equal to the current top row,
2. attaching a unit square to the left side of the current top row,
3. attaching a new leftmost column which is shorter than or equal to the current leftmost column,
4. attaching a unit square to the bottom side of the leftmost column,
5. attaching a new bottom-most row which is shorter than or equal to the current bottom-most row, and
6. attaching a unit square to the right side of the bottom-most row.

Any of these steps, if applicable, increases the half-perimeter by one, see Figure 8 below.



Figure 8: The types of steps used to obtain generating trees

Remark. *i)* Steps of types 2,4 and 6 are only admissible if the current top row (leftmost column, bottom row) is longer than or equal to the second row from top (second column

from the left, second row from the bottom). Additionally, a step of type 6 is forbidden, if the bottom row is only one unit shorter than the width of the box.

ii) In the proof of the functional equation (2.3) the steps of types 3 and 4 are encapsulated in the “attaching a bar graph to the left” operation. Hence any generic three-sided PP can be generated starting from the unit square by using only steps of the first four types.

iii) Building bar graphs only requires steps of type 1 and 2. Here we reflected the bar graphs discussed earlier in the line $x = y$.

In order to compute the appropriate probabilities according to which each step in the random path in the tree is chosen, we associate to each PP a label encoding the admissible steps which can be applied to enlarge it. We give a labelling for unrestricted PPs, since the labellings of the other classes are obtained as specialisations thereof. The labels are five-tuples $(a, e, k, l, p) \in \{B, L, T\} \times \{y, n\} \times \mathbb{Z}_{\geq 0}^3$. The letter a encodes the last building step. It is equal to T (top), if the last step was of type 1 (which inflated the box to the top), or of type 2 but *without* inflating the box to the left. a is equal to L (left) if the last step was of type 2 *and* thereby inflating the box to the left, of type 3 or of type 4 but *without* inflating the box. Finally, a is equal to B (bottom) if the last building step was of type 4 *with* an inflation of the box, or of types 5 or 6.

If $a = T$, the parameter $e \in \{y, n\}$ (yes/no) indicates if the current top row is longer than or equal to the second row from the top, and hence if a step of type 2 is applicable. Similarly, if $a = L$, e decides if a step of type 4 can be performed, i.e. if the leftmost column is shorter than or equal to the second but leftmost one. Finally, if $a = B$, e decides whether a step of type 6 can be performed.

The parameter k always denotes the length of the top row, and l is either the distance of the left end of the top row to the left side of the box or the length of the leftmost column or the length of the bottom row, depending on whether $a = T$ or $a = L$ or $a = B$ respectively. The unit square receives the label $(L, n, 1, 1, 0)$.

Labels for the generating tree of PPs				
a	e	k	l	p
T	top row extendable?	top row	dist. of box to top row	height
L	left col. extendable?	top row	length of leftmost col.	dist. of box to left col.
B	bottom row extendable?	top row	length of bottom row	height

Remark. In the proof of equation (2.4) we introduced two subclasses $\mathcal{G}$ and $\mathcal{H}$. The polygons with $a = L$ (resp. $a = B$) are precisely those in $\mathcal{G}$ (resp. $\mathcal{H}$).

The construction steps yield the following rewriting rules for the labels associated with general PPs.

$$(T, n, k, l, p) \rightarrow \begin{cases} (T, n, i, l + k - i, p + 1), & i = 1, \dots, k - 1 \\ (T, y, k, l, p + 1) \end{cases} \quad (5.1)$$

$$(T, y, k, l, p) \rightarrow \begin{cases} (T, n, i, l + k - i, p + 1), & i = 1, \dots, k - 1 \\ (T, y, k, l, p + 1) \\ (T, y, k + 1, l - 1, p), & \text{if } l \geq 1 \\ (L, n, k + 1, 1, p - 1), & \text{if } l = 0 \end{cases} \quad (5.2)$$

$$(L, n, k, l, p) \rightarrow \begin{cases} (T, n, i, k - i, l + p + 1), & i = 1, \dots, k - 1 \\ (T, y, k, 0, l + p + 1) \\ (L, n, k + 1, i, l + p - i), & i = 1, \dots, l - 1 \\ (L, y, k + 1, l, p) \end{cases} \quad (5.3)$$

$$(L, y, k, l, p) \rightarrow \begin{cases} (T, n, i, k - i, l + p + 1), & i = 1, \dots, k - 1 \\ (T, y, k, 0, l + p + 1) \\ (L, n, k + 1, i, l + p - i), & i = 1, \dots, l - 1 \\ (L, y, k + 1, l, p) \\ (L, y, k, l + 1, p - 1) & \text{if } p \geq 1 \\ (B, n, k, 1, l + 1) & \text{if } p = 0 \end{cases} \quad (5.4)$$

$$(B, n, k, l, p) \rightarrow \begin{cases} (T, n, i, k - i, p + 1), & i = 1, \dots, k - 1 \\ (T, y, k, 0, p + 1) \\ (L, n, k + 1, i, p - i), & i = 1, \dots, p - 1 \\ (L, y, k + 1, l, 0) \\ (B, n, k, i, p + 1), & i = 1, \dots, l - 1 \\ (B, y, k, l, p + 1) & \text{if } k - 1 > l \\ (B, n, k, l, p + 1) & \text{if } k - 1 = l \end{cases}$$

$$(B, y, k, l, p) \rightarrow \begin{cases} (T, n, i, k - i, p + 1), & i = 1, \dots, k - 1 \\ (T, y, k, 0, p + 1) \\ (L, n, k + 1, i, p - i), & i = 1, \dots, p - 1 \\ (L, y, k + 1, l, 0) \\ (B, n, k, i, p + 1), & i = 1, \dots, l - 1 \\ (B, y, k, l, p + 1) \\ (B, n, k, l + 1, p) & \text{if } k - 1 = l \\ (B, y, k, l + 1, p) & \text{if } k - 1 > l \end{cases}$$

The labelled rooted tree generated according to these rewriting rules with its root labelled $(L, n, 1, 1, 0)$ is a generating tree for unrestricted prudent polygons (more precisely, for the class $\mathcal{F}$, cf. Section 2). This tree is of course isomorphic to the tree $\mathcal{T}$ defined above having PPs as nodes, simply by replacing each PP by its label.

As mentioned above, choosing a PP of half-perimeter m uniformly at random is equivalent to choosing an $m - 2$ -step path starting from the root uniformly at random. This is

achieved by picking each step in the path according to an appropriate probability which in turn can be expressed in terms of *extension numbers*. If π is a polygon of half-perimeter $m - s$ (a path of length $m - s - 2$), then $EX(\pi, s)$ denotes the number of polygons of half-perimeter m which can be reached from π in s construction steps, or equivalently of extensions of length s of the path. Denote by $Ch(\pi)$ the set of polygons obtained from π in one step, i.e. the children of π in the generating tree. Now the right probability to choose $\alpha \in Ch(\pi)$ in the course of our random sampling procedure is equal to

$$\mathbb{P}(\alpha|\pi) = \frac{EX(\alpha, s-1)}{EX(\pi, s)}.$$

The numbers $EX(\pi, s)$ can be computed recursively, namely

$$EX(\pi, s) = \begin{cases} 1 & \text{if } s = 0, \\ \sum_{\alpha \in Ch(\pi)} EX(\alpha, s-1) & \text{otherwise.} \end{cases}$$

The crucial observation is that $EX(\pi, \cdot)$ *only depends on the label of π* , which allows an efficient computation.

For unrestricted PPs, in the first $m - 2$ levels of the tree $O(m^3)$ different labels occur since none of the parameters exceeds m . It hence takes $O(m^4)$ operations to compute the all required extension numbers. We have implemented the procedure and computed these numbers up to $m = 80$. See Figure 12 at the end of the paper for some samples.

Modifications for two-sided PPs. As remarked above, generating two-sided PPs only requires steps of types 1 and 2. The only required information for the building procedure is the length of the top row and if the top row is extendable. We hence only need labels (T, n, k) and (T, y, k) obtained from the labels $(T, \cdot, k, l, p)$ above by leaving the parameters l and p unconsidered. The rewriting rule (5.1) can be adapted unchanged (up to deleting the last two coordinates) and in (5.2) simply omit the last line (and the “if $l \geq 1$ ”-clause in the second but last line). The unit square receives the label $(T, y, 1)$. There are $O(m)$ different labels in the first $m - 2$ levels of the generating tree, and hence $O(m^2)$ extension numbers have to be computed. See Figure 10 for some samples of half-perimeter 250.

Modifications for three-sided PPs. For the generation of three-sided PPs steps 1, 2, 3 and 4 suffice. For an appropriate labelling we can hence dump down the $(B, \cdot, k, l, p)$ labels and use labels (T, n, k, l) , (T, y, k, l) , (L, n, k, l) and (L, y, k, l) obtained from the labels $(T, \cdot, k, l, p)$ and $(L, \cdot, k, l, p)$ by simply discarding the parameter p . The rewriting rules (5.1), (5.2) and (5.3) are adapted without change. In the rule (5.4) drop the last line. The unit square is labelled with $(L, n, 1, 1)$. We have $O(m^2)$ different labels on the first $m - 2$ levels of the tree and hence $O(m^3)$ extension numbers have to be computed. See Figure 11 for some samples of half-perimeter 250.

6 Conclusion

We have solved the class of two-sided and three-sided prudent polygons, the generating function being algebraic in the former and non-D-finite in the latter case. The analysis

shows that two-sided PPs are exponentially rare among three-sided PPs which is different from the corresponding walk models where the growth rates are equal.

It would be nice to solve the class of general prudent polygons. We expect that the involved functional equations require three or more catalytic variables, which is the case for the equation found for the walk model.

Since the exponential growth rates of SAWs and SAPs are known to be equal [10] it is also interesting to compare the exponential growth rates of k -sided PWs and PPs. To that end it suffices to study PPs ending in $(1, 0)$. As already mentioned in the introduction, a k -sided PP ending in $(1, 0)$ may never step right of the line $x = 1$ and it heads towards the vertex $(1, 0)$ as soon as it hits that line for the first time in a point $(1, y_0)$. Up to that step the boundary walk of that k -sided PP is genuinely $k - 1$ -sided. This yields an injective map sending a k -sided PP to a $k - 1$ -sided PW simply by reflecting the segment joining $(1, y_0)$ and $(1, 0)$ in the line $y = y_0$, see Figure 9. We denote the so-obtained subclass of

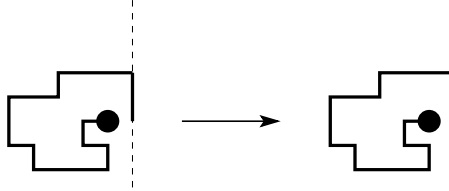


Figure 9: Embedding of k -sided PPs into $k - 1$ -sided PWs

$k - 1$ -sided PWs by “embedded k -sided PPs”. If we count PPs by full perimeter, their exponential growth rates become $1/\sqrt{\rho} = 1.83\dots$ for two-sided PPs and $1/\sqrt{\sigma} = 2.02\dots$ for three-sided PPs. It is known that the exponential growth rate of PWs is equal to $1 + \sqrt{2} = 2.41\dots$ in the one-sided case and equal to $2.48\dots$ in the two- and three-sided cases [2]. The latter rate is also expected for unrestricted PWs [4, 7]. Consequently, for $k = 2, 3$, our results imply that k -sided PPs are exponentially rare among k -sided PWs *and*, via embedding, among $k - 1$ -sided PWs. Furthermore, the rate of three-sided PPs is even smaller than that of one-sided PWs. This is not surprising looking at the pictures in Figure 11, as such a PP roughly consists of two “almost” one-sided PWs, one heading to the far left followed by one “almost directed” walk up and to the right (and the closing tail). We expect that this heuristic argument also applies in the general case, which is also supported by an estimated value of approximately $2.1 < 1 + \sqrt{2}$ for the growth rate of general PPs [4, 7].

Acknowledgements

Part of this work was done during a stay of the author at the University of Melbourne. He would like to thank MASCOS and the members of the Statistical Mechanics and Combinatorics group for their kind hospitality. In particular he would like to thank Tony Guttmann

for introducing the problem to him. He thanks Christoph Richard for helpful comments on the manuscript. He would also like to acknowledge financial support by the German Research Council (DFG) within CRC 701.

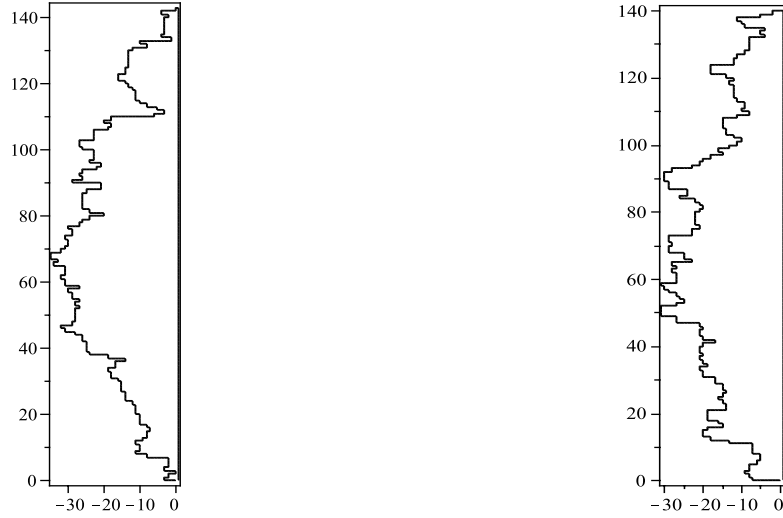


Figure 10: Random 2-sided PPs of half-perimeter 250

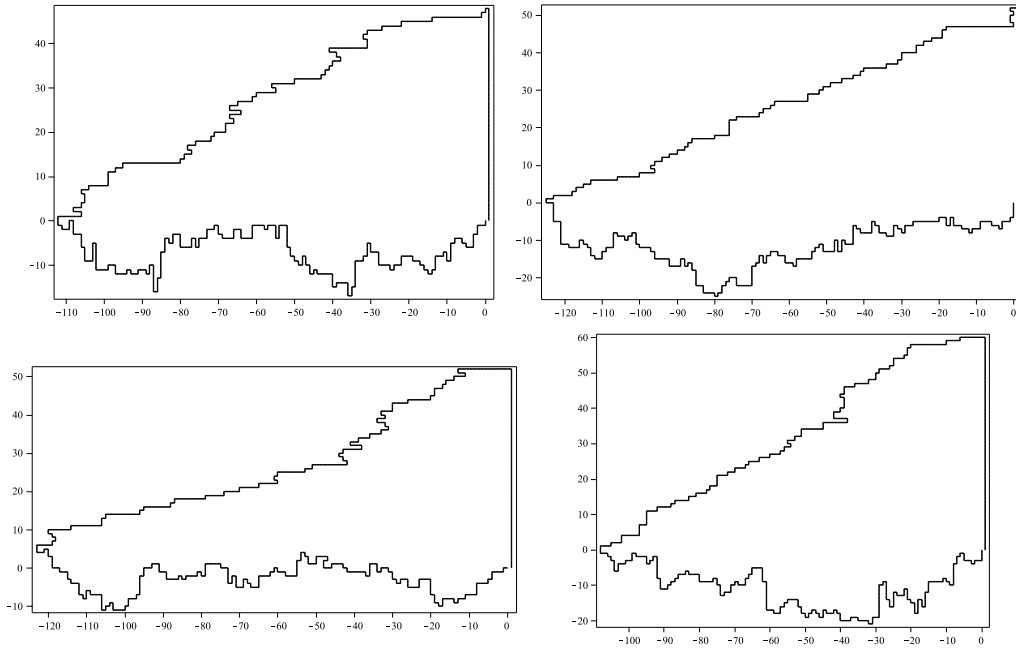


Figure 11: Random 3-sided PPs of half-perimeter 250

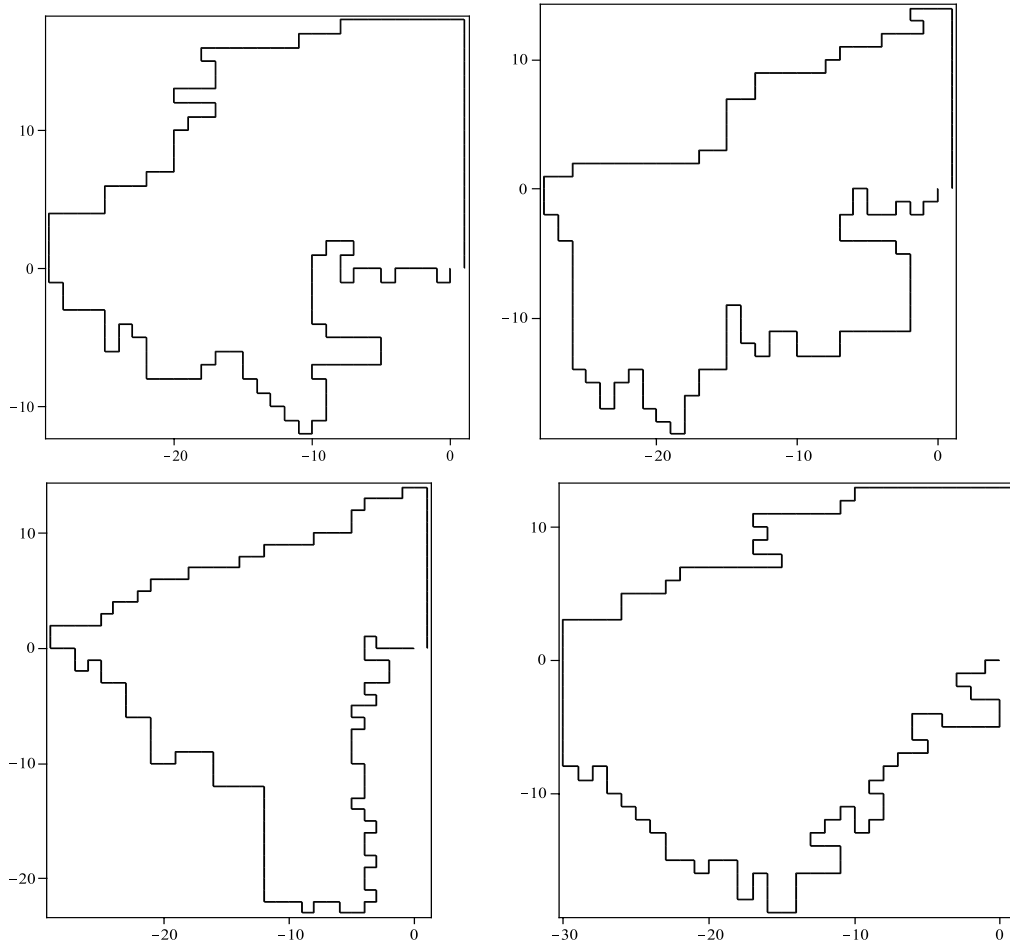


Figure 12: Random unrestricted PPs of half-perimeter 80

References

- [1] M. Bousquet-Mélou, A method for the enumeration of various classes of column-convex polygons, *Discrete Math.* **154**(1996), 1-25, doi:10.1016/0012-365X(95)00003-F
- [2] M. Bousquet-Mélou, Families of prudent self-avoiding walks (2008), arXiv: math.CO 0804.4843.
- [3] M. Bousquet-Mélou and A. Jehanne, Polynomial equations with one catalytic variable, algebraic series and map enumeration, *J. Combin. Theory, Ser. B* **96**(2006), No.5, 623-672, doi:10.1016/j.jctb.2005.12.003
- [4] J. Dethridge, A.J. Guttmann and I. Jensen, Prudent self-avoiding walks and polygons. In *Random Polymers*, EURANDOM, Eindhoven, The Netherlands (June 2007).
- [5] E. Duchi, On some classes of prudent walks. In *FPSAC'05*, Taormina, Italy (2005).

- [6] P. Duchon, Q -grammars and wall polyominoes, *Ann. Comb.* **3**(1999), No. 2-4, 311-321, doi:10.1007/BF01608790
- [7] T.M. Garoni, Timothy M., A.J. Guttmann, I. Jensen and J. Dethridge, Prudent walks and polygons, *J. Phys. A* **42**(2009), No. 9, doi: 10.1088/1751-8113/42/9/095205
- [8] I.P. Goulden and D.M. Jackson, *Combinatorial Enumeration*, New York, Wiley (1983).
- [9] A.J. Guttmann, Some solvable, and as yet unsolvable, polygon and walk models, *IOP, J. Phys.: Conference Series* **42**(2006), 98-110, doi: 10.1088/1742-6596/42/1/011
- [10] J.M. Hammersley, The number of polygons on a lattice, *Proc. Camb. Philos. Soc.* **57** (1961), 516-523.
- [11] N. Madras and G. Slade, *The self-avoiding walk*, Probability and Its Applications. Boston, MA: Birkhäuser. xiv, 425 p. (1993).
- [12] M. Mishna and A. Rechnitzer, Two non-holonomic lattice walks in the quarter plane, *Theoret. Comput. Sci.* **410**(2009), No. 38-40, 3616-3630, doi:10.1016/j.tcs.2009.04.008
- [13] A. Nijenhuis and H. Wilf, *Combinatorial algorithms for computers and calculators. 2nd ed.*, Computer Science and Applied Mathematics. New York -San Francisco - London: Academic Press. xv, 302 p. (1978)
- [14] P. Pr  a, Exterior self-avoiding walks on the square-lattice, unpublished manuscript (1997).
- [15] T. Prellberg and R. Brak, Critical exponents from nonlinear functional equations for partially directed cluster models, *J. Stat. Phys.* **78**(1995), No. 3-4, 701-730, doi:10.1007/BF02183685
- [16] A. Rechnitzer, Haruspicy 2: The anisotropic generating function of self-avoiding polygons is not D-finite, *J. Combin. Theory, Ser. A* **113**(2006), Issue 3, 520-546, doi:10.1016/j.jcta.2005.04.010
- [17] C. Richard, Limit distributions and scaling functions. In *Polygons, Polyominoes and Polycubes*, 247-299, Lecture Notes in Physics. Berlin/Heidelberg: Springer (2009) doi:10.1007/978-1-4020-9927-4_11
- [18] V. Scheidemann, *Introduction to complex analysis in several variables*, Birkh  user, Basel (2005).
- [19] L. Tak  cs, A Bernoulli excursion and its various applications. *Adv. Appl. Probab.* **23**(1991), Issue 3, 557-585, doi:10.2307/1427622