WALKS IN THE QUARTER PLANE
A FUNCTIONAL EQUATION APPROACH

MIREILLE BOUSQUET-MÉLOU

ABSTRACT. We study planar walks that start from a given point (i_0, j_0), take their steps in a finite set $\mathfrak{S}$, and are confined in the first quadrant $x \geq 0, y \geq 0$. Their enumeration can be attacked in a systematic way: the generating function $Q(x, y; t)$ that counts them by their length (variable t) and the coordinates of their endpoint (variables x, y) satisfies a linear functional equation encoding the step-by-step description of walks. For instance, for the square lattice walks starting from the origin, this equation reads

$$\left(xy - t(x + y + x^2y + xy^2)\right) Q(x, y; t) = xy - xtQ(x, 0; t) - ytQ(0, y; t).$$

The central question addressed in this paper is the *nature* of the series $Q(x, y; t)$. When is it algebraic? When is it D-finite (or holonomic)? Can these properties be derived from the functional equation itself?

Our first result is a new proof of an old theorem due to Kreweras, according to which one of these walk models has, for mysterious reasons, an algebraic generating function. Then, we provide a new proof of a holonomy criterion recently proved by M. Petkovšek and the author. In both cases, we work directly from the functional equation.

RÉSUMÉ. Considérons les chemins du plan qui partent d'un point (i_0, j_0) donné, choisissent leurs pas dans un ensemble fini $\mathfrak{S}$, et restent confinés dans le premier quadrant $x \geq 0, y \geq 0$. On peut attaquer leur énumération de façon systématique : la série génératrice $Q(x, y; t)$ qui les énumère selon la longueur (variable t) et le point d'arrivée (variables x, y) est toujours solution d'une équation fonctionnelle linéaire qui code la construction pas à pas de ces chemins. Par exemple, l'équation régissant les chemins de la grille carrée qui, issus de l'origine, restent dans le premier quadrant, est donnée ci-dessus.

La question centrale qui motive cet article porte sur le *nature* de la série $Q(x, y; t)$. Quand est-elle algébrique ? holonome ? Comment déduire ces propriétés de l'équation fonctionnelle qu'elle satisfait ?

Notre premier résultat est une nouvelle preuve d'un vieux théorème dû à Kreweras, qui affirme qu'un de ces modèles de chemins a, pour des raisons mystérieuses, une série génératrice algébrique. Ensuite, nous donnons une nouvelle preuve d'un critère d'holonomie récemment démontré par M. Petkovšek et l'auteur. Dans les deux cas, le point de départ est l'équation fonctionnelle qui rgit la srie $Q(x, y; t)$.

1. Walks in the quarter plane

The enumeration of lattice walks is one of the most venerable topics in enumerative combinatorics, which has numerous applications in probabilities [15, 27, 36]. These walks take their steps in a finite subset $\mathfrak{S}$ of $\mathbb{Z}^d$, and might be constrained in various ways. One can only cite a small percentage of the relevant litterature, which dates back at least to the next-to-last century [1, 18, 24, 30, 31]. Many recent publications show that the topic is still active [4, 6, 11, 20, 22, 32, 33].

After the solution of many explicit problems, certain patterns have emerged, and a more recent trend consists in developping methods that are valid for generic sets of steps. A special attention is being paid to the *nature* of the generating function of the walks under consideration. For instance, the generating function for unconstrained walks on the line $\mathbb{Z}$ is rational, while the generating function for walks constrained to stay in the half-line $\mathbb{N}$ is always algebraic [3]. This result has often been described in terms of *partially directed*

2-dimensional walks confined in a quadrant (or *generalized Dyck walks* [13, 19, 25, 26]), but is, essentially, of a 1-dimensional nature.

Similar questions can be addressed for *real* 2-dimensional walks. Again, the generating function for unconstrained walks starting from a given point is clearly rational. Moreover, the argument used for 1-dimensional walks confined in $\mathbb{N}$ can be recycled to prove that the generating function for the walks that stay in the half-plane $x \geq 0$ is always algebraic. What about doubly-restricted walks, that is, walks that are confined in the quadrant $x \geq 0, y \geq 0$?

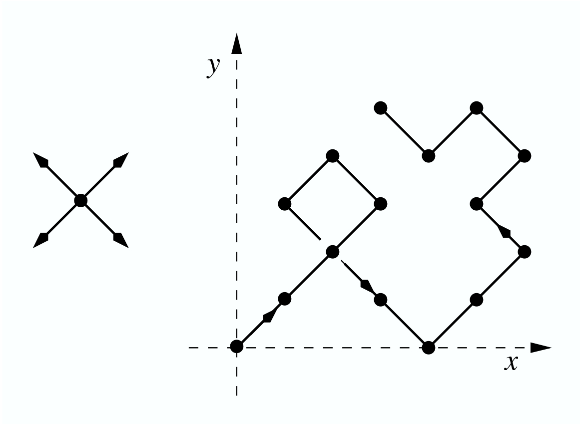


FIGURE 1. A walk on the diagonal square lattice confined in the first quadrant.

A rapid inspection of the most standard cases suggests that these walks might have always a D-finite generating function[1]. The simplest example is probably that of the diagonal square lattice, where the steps are North-East, South-East, North-West and South-West (Figure 1): by projecting the walks on the x- and y-axes, we obtain two decoupled prefixes of Dyck paths, so that the length generating function for walks that start from the origin and stay in the first quadrant is

$$\sum_{n\geq 0} \binom{n}{\lfloor n/2 \rfloor}^2 t^n,$$

a D-finite series. For the ordinary square lattice (with North, East, South and West steps), the generating function is

$$\sum_{m,n\geq 0} \binom{m+n}{m} \binom{m}{\lfloor m/2 \rfloor} \binom{n}{\lfloor n/2 \rfloor} t^{m+n} = \sum_{n\geq 0} \binom{n}{\lfloor n/2 \rfloor} \binom{n+1}{\lceil n/2 \rceil} t^n,$$

another D-finite series. The first expression comes from the fact that these walks are shuffles of two prefixes of Dyck walks, and the Chu-Vandermonde identity transforms it into the second simpler expression.

In both cases, the number of n-step walks grows asymptotically like $4^n/n$, which prevents the generating function from being algebraic (see [16] for the possible asymptotic behaviours of coefficients of algebraic series).

The two above results can be refined by taking into account the coordinates of the endpoint: if $a_{i,j}(n)$ denotes the number of n-step walks of length n ending at (i,j), then

[1]A series $F(t)$ is D-finite (or *holonomic*) if it satisfies a linear differential equation with polynomial coefficients in t. Any algebraic series is D-finite.

we have, for the diagonal square lattice:

$$\sum_{i,j,n\geq 0} a_{i,j}(n)x^i y^j t^n = \sum_{i,j,n\geq 0} \frac{(i+1)(j+1)}{(n+1)^2}\binom{n+1}{\frac{n-i}{2}}\binom{n+1}{\frac{n-j}{2}}x^i y^j t^n,$$

where the binomial coefficient $\binom{n}{(n-i)/2}$ is zero unless $0 \leq i \leq n$ and $i \equiv n \bmod 2$. Similarly, for the ordinary square lattice,

$$\sum_{i,j,n\geq 0} a_{i,j}(n)x^i y^j t^n = \sum_{i,j,n\geq 0} \frac{(i+1)(j+1)}{(n+1)(n+2)}\binom{n+2}{\frac{n+i-j+2}{2}}\binom{n+2}{\frac{n-i-j}{2}}x^i y^j t^n. \quad (1)$$

These two series can be seen to be D-finite in their three variables.

This holonomy, however, is not the rule: as proved in [10], walks that start from $(1,1)$, take their steps in $\mathfrak{S} = \{(2,-1),(-1,2)\}$ and always stay in the first quadrant have a non-D-finite length generating function. The same holds for the subclass of walks ending on the x-axis.

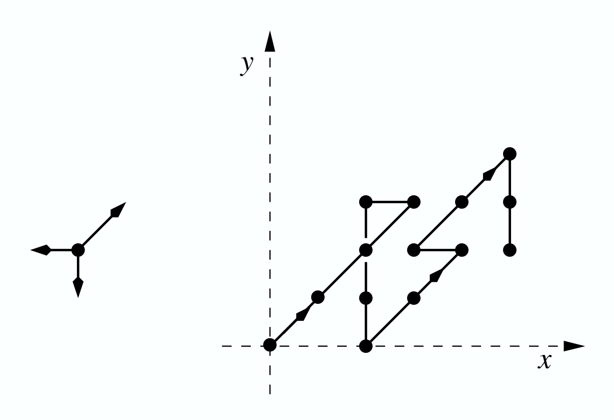


FIGURE 2. Kreweras' walks in a quadrant.

At the other end of the hierarchy, another walk model displays a mysteriously simple algebraic generating function: when the starting point is $(0,0)$, and the allowed steps South, West and North-East (Figure 2), the number of walks of length $3n+2i$ ending at the point $(i,0)$ is

$$\frac{4^n(2i+1)}{(n+i+1)(2n+2i+1)}\binom{2i}{i}\binom{3n+2i}{n}. \quad (2)$$

This result was first proved by Kreweras in 1965 [24, Chap. 3], and then rederived by Niederhausen [32] and Gessel [18]. It is, however, not well-understood, for two reasons:

– no direct proof of (2) is known, even when $i = 0$. The number of walks ending at the origin is closely related to the number of non-separable planar maps, to the number of *cubic* non-separable maps [34, 35, 38, 39], and to the number of two-stack sortable permutations [5, 40, 41]. All available proofs of (2) are rather long and complicated. Moreover, in all of them, the result is *checked* rather than *derived*.

– most importantly, the three-variate generating function for these walks can be shown to be algebraic [18], but none of the proofs explain combinatorially this algebraicity.

All problems of walks confined in a quadrant can be attacked by writing a functional equation for their three-variate generating function, and it is this uniform approach that we discuss here. This functional equation simply encodes the step-by-step construction of the walks. For instance, for square lattice walks, we can write

$$\begin{aligned} Q(x,y;t) &:= \sum_{i,j,n\geq 0} a_{i,j}(n)x^i y^j t^n \\ &= 1+t(x+y)Q(x,y;t)+t\,\frac{Q(x,y;t)-Q(0,y;t)}{x}+t\,\frac{Q(x,y;t)-Q(x,0;t)}{y}, \end{aligned}$$

that is,

$$\left(xy - t(x+y+x^2y+xy^2)\right) Q(x,y;t) = xy - xtQ(x,0;t) - ytQ(0,y;t), \tag{3}$$

and the solution of this equation, given by (1), is D-finite (but transcendental). Similarly, for the diagonal square lattice, we have

$$\left(xy - t(1+x^2)(1+y^2)\right) Q(x,y;t) = xy - t(1+x^2)Q(x,0;t) - t(1+y^2)Q(0,y;t) + tQ(0,0;t),$$

with again a D-finite transcendental solution, while for Kreweras' algebraic model, we obtain

$$\left(xy - t(x+y+x^2y^2)\right) Q(x,y;t) = xy - xtQ(x,0;t) - ytQ(0,y;t). \tag{4}$$

Finally, the equation that rules the non-holonomic model of [10] is

$$\left(xy - t(x^3+y^3)\right) Q(x,y;t) = x^2y^2 - tx^3Q(x,0;t) - ty^3Q(0,y;t).$$

The general theme of this paper is the following: the above equations completely solve, in some sense, the problem of enumerating the walks. But they are not the kind of solution one likes, especially if the numbers are simple, or if the generating function is actually algebraic! How can one derive these simple solutions from the functional equations? And what is the essential difference between, say, Eqs. (3) and (4), that makes one series transcendental, and the other algebraic?

We shall answer some of these questions. Our main result is a new proof of (2), which we believe to be simpler that the three previous ones. It has, at least, one nice feature: we *derive the algebraicity from the equation* without having to guess the formula first. Then, we give a new proof of a (refinement of) a holonomy criterion that was proved combinatorially in [10]: if the set of steps $\mathfrak{S}$ is symmetric with respect to the y-axis and satisfies a *small horizontal variations* condition, then the generating function for walks with steps in $\mathfrak{S}$, starting from any given point (i_0, j_0), is D-finite. This result covers the two above D-finite transcendental cases, but not Kreweras' model... We finally survey some perspectives of this work.

Let us conclude this section with a few more formal definitions on walks and power series.

Let $\mathfrak{S}$ be a finite subset of $\mathbb{Z}^2$. A walk with steps in $\mathfrak{S}$ is a finite sequence $w = (w_0, w_1, \ldots, w_n)$ of vertices of $\mathbb{Z}^2$ such that $w_i - w_{i-1} \in \mathfrak{S}$ for $1 \leq i \leq n$. The number of steps, n, is the *length* of w. The starting point of w is w_0, and its endpoint is w_n. The *complete generating function* for a set $\mathfrak{A}$ of walks starting from a given point $w_0 = (i_0, j_0)$ is the series

$$A(x,y;t) = \sum_{n\geq 0} t^n \sum_{i,j\in\mathbb{Z}} a_{i,j}(n)x^i y^j,$$

where $a_{i,j}(n)$ is the number of walks of $\mathfrak{A}$ that have length n and end at (i,j). This series is a formal power series in t whose coefficients are polynomials in $x, y, 1/x, 1/y$. We shall often denote $\bar{x} = 1/x$ and $\bar{y} = 1/y$.

Given a ring $\mathbb{L}$ and k indeterminates $x_1, \ldots, x_k$, we denote by $\mathbb{L}[x_1, \ldots, x_k]$ the ring of polynomials in $x_1, \ldots, x_k$ with coefficients in $\mathbb{L}$, and by $\mathbb{L}[[x_1, \ldots, x_k]]$ the ring of formal power series in $x_1, \ldots, x_k$ with coefficients in $\mathbb{L}$. If $\mathbb{L}$ is a field, we denote by $\mathbb{L}(x_1, \ldots, x_k)$ the field of rational functions in $x_1, \ldots, x_k$ with coefficients in $\mathbb{L}$.

Assume $\mathbb{L}$ is a field. A series F in $\mathbb{L}[[x_1,\dots,x_k]]$ is *rational* if there exist polynomials P and Q in $\mathbb{L}[x_1,\dots,x_k]$, with $Q\neq 0$, such that $QF=P$. It is *algebraic* (over the field $\mathbb{L}(x_1,\dots,x_k)$) if there exists a non-trivial polynomial P with coefficients in $\mathbb{L}$ such that $P(F,x_1,\dots,x_k)=0$. The sum and product of algebraic series is algebraic.

The series F is *D-finite* (or *holonomic*) if the partial derivatives of F span a finite dimensional vector space over the field $\mathbb{L}(x_1,\dots,x_k)$ (this vector space is a subspace of the fraction field of $\mathbb{L}[[x_1,\dots,x_k]]$); see [37] for the one-variable case, and [28, 29] otherwise. In other words, for $1\le i\le k$, the series F satisfies a non-trivial partial differential equation of the form

$$\sum_{\ell=0}^{d_i} P_{\ell,i}\frac{\partial^\ell F}{\partial x_i^\ell}=0,$$

where $P_{\ell,i}$ is a polynomial in the x_j. Any algebraic series is holonomic. The sum and product of two holonomic series is still holonomic. The specializations of an holonomic series (obtained by giving values from $\mathbb{L}$ to some of the variables) are holonomic, if well-defined. Moreover, if F is an *algebraic* series and $G(t)$ is a holonomic series of one variable, then the substitution $G(F)$ (if well-defined) is holonomic [29, Prop. 2.3].

2. A new proof of Kreweras' result

Consider walks that start from $(0,0)$, are made of South, West and North-East steps, and always stay in the first quadrant (Figure 2). Let $a_{i,j}(n)$ be the number of n-step walks of this type ending at (i,j). We denote by $Q(x,y;t)$ the complete generating function of these walks:

$$Q(x,y;t):=\sum_{i,j,n\ge 0} a_{i,j}(n)x^iy^jt^n.$$

Constructing the walks step by step yields the following equation:

$$\big(xy-t(x+y+x^2y^2)\big)\,Q(x,y;t)=xy-xtQ(x,0;t)-ytQ(0,y;t). \tag{5}$$

We shall often denote, for short, $Q(x,y;t)$ by $Q(x,y)$. Let us also denote the series $xtQ(x,0;t)$ by $R(x;t)$ or even $R(x)$. Using the symmetry of the problem in x and y, the above equation becomes:

$$\big(xy-t(x+y+x^2y^2)\big)\,Q(x,y)=xy-R(x)-R(y). \tag{6}$$

This equation is equivalent to a recurrence relation defining the numbers $a_{i,j}(n)$ by induction on n. Hence, it defines completely the series $Q(x,y;t)$. Still, the characterization of this series we have in mind is of a different nature:

Theorem 1. *Let $X\equiv X(t)$ be the power series in t defined by*

$$X=t(2+X^3).$$

Then the generating function for Kreweras' walks ending on the x-axis is

$$Q(x,0;t)=\frac{1}{tx}\left(\frac{1}{2t}-\frac{1}{x}-\left(\frac{1}{X}-\frac{1}{x}\right)\sqrt{1-xX^2}\right).$$

Consequently, the length generating function for walks ending at $(i,0)$ is

$$[x^i]Q(x,0;t)=\frac{X^{2i+1}}{2.4^i\ t}\left(C_i-\frac{C_{i+1}X^3}{4}\right),$$

where $C_i = \binom{2i}{i}/(i+1)$ *is the* i*-th Catalan number. The Lagrange inversion formula gives the number of such walks of lenth* $3n+2i$ *as*

$$a_{i,0}(3n+2i) = \frac{4^n(2i+1)}{(n+i+1)(2n+2i+1)}\binom{2i}{i}\binom{3n+2i}{n}.$$

The aim of this section is to derive Theorem 1 from the functional equation (5).

2.1. **The obstinate kernel method.** The kernel method is basically the only tool we have to attack Equation (6). This method had been around since, at least, the 70's, and is currently the subject of a certain rebirth (see the references in [2, 3, 9]). It consists in coupling the variables x and y so as to cancel the *kernel* $K(x,y) = xy - t(x+y+x^2y^2)$. This should give the "missing" information about the series $R(x)$.

As a polynomial in y, this kernel has two roots

$$\begin{aligned} Y_0(x) &= \frac{1-t\bar{x}-\sqrt{(1-t\bar{x})^2-4t^2x}}{2tx} &&= t+\bar{x}t^2+O(t^3), \\ Y_1(x) &= \frac{1-t\bar{x}+\sqrt{(1-t\bar{x})^2-4t^2x}}{2tx} &&= \frac{\bar{x}}{t}-\bar{x}^2 - t-\bar{x}t^2+O(t^3). \end{aligned}$$

The elementary symmetric functions of the Y_i are

$$Y_0+Y_1 = \frac{\bar{x}}{t}-\bar{x}^2 \quad \text{and} \quad Y_0Y_1 = \bar{x}. \tag{7}$$

The fact that they are polynomials in $\bar{x} = 1/x$ will play a very important role below.

Only the first root can be substituted for y in (6) (the term $Q(x, Y_1; t)$ is not a well-defined power series in t). We thus obtain a functional equation for $R(x)$:

$$R(x) + R(Y_0) = xY_0. \tag{8}$$

It can be shown that this equation – once restated in terms of $Q(x,0)$ – defines uniquely $Q(x,0;t)$ as a formal power series in t with polynomial coefficients in x. Equation (8) is the standard result of the kernel method.

Still, we want to apply here the *obstinate* kernel method. That is, we shall not content ourselves with Eq. (8), but we shall go on producing pairs (X,Y) that cancel the kernel and use the information they provide on the series $R(x)$.

Let $(X,Y) \neq (0,0)$ be a pair of Laurent series in t with coefficients in a field $\mathbb{K}$ such that $K(X,Y) = 0$. We define $\Phi(X,Y) = (X',Y)$, where $X' = (XY)^{-1}$ is *the other solution* of $K(x,Y) = 0$, seen as a polynomial in x. Similarly, we define $\Psi(X,Y) = (X,Y')$, where $Y' = (XY)^{-1}$ is the other solution of $K(X,y) = 0$. Note that Φ and Ψ are involutions. Let us examine their action on the pair (x, Y_0). We obtain the following diagram[2]

All these pairs of power series cancel the kernel, and we have framed the ones that can be legally substituted[3] in the main functional equation (6). We thus obtain *two* equations for the unknown series $R(x)$:

$$R(x) + R(Y_0) = xY_0, \tag{9}$$

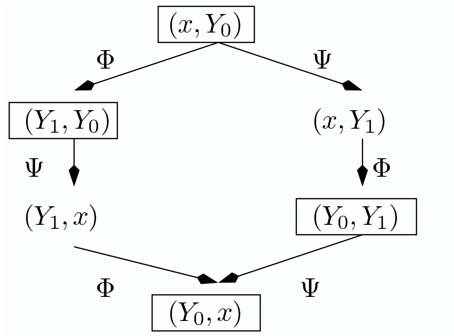
$$R(Y_0) + R(Y_1) = Y_0Y_1 = \bar{x}. \tag{10}$$

This obstinate kernel method is the first original ingredient of this proof.

Remark. Let p,q,r be three nonnegative numbers such that $p+q+r=1$. Take $x = (pr)^{1/3}q^{-2/3}$, $y = (qr)^{1/3}p^{-2/3}$, and $t = (pqr)^{1/3}$. Then $K(x,y;t) = 0$, so that $R(x)+R(y) =$

[2]The transformations Φ and Ψ are related to, but distinct from, the transformations ξ and η of [14, Sect. 2.4]. The underlying idea – finding "all" pairs of roots of the kernel – is the same.

[3]For this choice of steps, the series $Q(Y_0, Y_1; t)$ and $Q(0, Y_1; t)$ *are* well-defined!



xy. This equation can be given a probabilistic interpretation by considering random walks that make a North-East step with (small) probability r and a South (resp. West) step with probability p (resp. q). This probabilistic argument, and the equation it implies, is the starting point of Gessel's solution of Kreweras problem [18, Eq. (21)].

2.2. **Symmetric functions of Y_0 and Y_1.** After the kernel method, the next tool of our approach is the extraction of the positive part of power series. More precisely, let $S(x;t)$ be a power series in t whose coefficients are Laurent polynomials in x:

$$S(x;t) = \sum_{n\geq 0} t^n \sum_{i\in\mathbb{Z}} s_i(n) x^i t^n,$$

where for each $n \geq 0$, only finitely many coefficients $s_i(n)$ are non-zero. We define the positive part of this series by

$$S^+(x;t) := \sum_{n\geq 0} t^n \sum_{i\in\mathbb{N}} s_i(n) x^i t^n.$$

This is where the values of the symmetric functions of Y_0 and Y_1 become crucial: the fact that they only involve negative powers of x (see (7)) will simplify the extraction of the positive part of certain equations. This observation is the second original ingredient of this proof.

Lemma 2. *Let $F(u, v;t)$ be a power series in t with coefficients in $\mathbb{C}[u, v]$, such that $F(u, v;t) = F(v, u;t)$. Then the series $F(Y_0, Y_1;t)$, if well-defined, is a power series in t with polynomial coefficients in $\bar{x}$. Moreover, the constant term of this series, taken with respect to $\bar{x}$, is $F(0, 0;t)$.*

Proof. All symmetric polynomials of u and v are polynomials in $u+v$ and uv with complex coefficients.

■

We now want to form a symmetric function of Y_0 and Y_1, starting from the equations (9–10). The first one reads

$$R(Y_0) - xY_0 = -R(x).$$

By combining both equations, we then obtain the companion expression:

$$R(Y_1) - xY_1 = R(x) + 2\bar{x} - 1/t.$$

Taking the product[4] of these two equations gives

$$(R(Y_0) - xY_0)(R(Y_1) - xY_1) = -R(x)(R(x) + 2\bar{x} - 1/t).$$

[4] An alternative derivation of Kreweras' result, obtained by considering the divided difference $(R(Y_0) - xY_0 - R(Y_1) + xY_1)/(Y_0 - Y_1)$, will be discussed on the complete version of this paper.

The extraction of the positive part of this identity is made possible by Lemma 2. Given that $R(x;t) = xtQ(x,0;t)$, one obtains:

$$x = -t^2x^2Q(x,0)^2 + (x-2t)Q(x,0) + 2tQ(0,0),$$

that is,

$$t^2x^2Q(x,0)^2 + (2t-x)Q(x,0) - 2tQ(0,0) + x = 0. \tag{11}$$

2.3. **The quadratic method.** Equation (11) – which begs for a combinatorial explanation – is typical of the equations obtained when enumerating planar maps, and the rest of the proof will be routine to all maps lovers. This equation can be solved using the so-called *quadratic method*, which was first invented by Brown [12]. The formulation we use here is different both from Brown's original presentation and from the one in Goulden and Jackson's book [21]. This new formulation is convenient for generalizing the method to equations of higher degree with more unknowns [8].

Equation (11) can be written as

$$P(Q(x), Q(0), t, x) = 0, \tag{12}$$

where $P(u,v,t,x) = t^2x^2u^2 + (2t-x)u - 2tv + x$, and $Q(x,0)$ has been abbreviated in $Q(x)$. Differentiating this equation with respect to x, we find

$$\frac{\partial P}{\partial u}(Q(x), Q(0), t, x)\frac{\partial Q}{\partial x}(x) + \frac{\partial P}{\partial x}(Q(x), Q(0), t, x) = 0.$$

Hence, if there exists a power series in t, denoted $X(t) \equiv X$, such that

$$\frac{\partial P}{\partial u}(Q(X), Q(0), t, X) = 0, \tag{13}$$

then one also has

$$\frac{\partial P}{\partial x}(Q(X), Q(0), t, X) = 0, \tag{14}$$

and we thus obtain a system of three polynomial equations, namely Eq. (12) written for $x = X$, Eqs. (13) and (14), that relate the three unknown series $Q(X)$, $Q(0)$ and X. This puts us in a good position to write an algebraic equation defining $Q(0) = Q(0,0;t)$.

Let us now work out the details of this program: Eq. (13) reads $X = 2t^2X^2Q(X) + 2t$, and since the right-hand side is a multiple of t, it should be clear that this equation defines a unique power series $X(t)$. The system of three equations reads

$$\begin{cases} t^2X^2Q(X)^2 + (2t-X)Q(X) - 2tQ(0) + X = 0, \\ 2t^2X^2Q(X) + 2t - X = 0, \\ 2t^2XQ(X)^2 - Q(X) + 1 = 0. \end{cases}$$

Eliminating $Q(X)$ between the last two equations yields $X = t(2 + X^3)$, so that the series X is the parameter introduced in Theorem 1. Going on with the elimination, we finally obtain

$$Q(0,0;t) = \frac{X}{2t}\left(1 - \frac{X^3}{4}\right),$$

and the expression of $Q(x,0;t)$ follows from (11).

∎

3. A holonomy criterion

Using functional equations, we can recover, and actually refine, a holonomy criterion that was recently proved combinatorially [10]. Let $\mathfrak{S}$ be a finite subset of $\mathbb{Z}^2$. We say that $\mathfrak{S}$ is symmetric with respect to the y-axis if

$$(i,j) \in \mathfrak{S} \Rightarrow (-i,j) \in \mathfrak{S}.$$

We say that $\mathfrak{S}$ has small horizontal variations if

$$(i,j) \in \mathfrak{S} \Rightarrow |i| \leq 1.$$

The usual square lattice steps satisfy these two conditions. So do the steps of the diagonal square lattice (Figure 1).

Theorem 3. *Let $\mathfrak{S}$ be a finite subset of $\mathbb{Z}^2$ that is symmetric with respect to the y-axis and has small horizontal variations. Let $(i_0, j_0) \in \mathbb{N}^2$. Then the complete generating function $Q(x,y;t)$ for walks that start from (i_0, j_0), take their steps in $\mathfrak{S}$ and stay in the first quadrant is D-finite.*

A combinatorial argument proving the holonomy of $Q(1,1;t)$ is presented in [10].

3.1. **Example.** Before we embark on the proof of this theorem, let us see the principle of the proof at work on a simple example: square lattice walks confined in a quadrant. The functional equation satisfied by their complete generating function is

$$(15)\quad \left(xy - t(x+y+x^2y+xy^2)\right) Q(x,y) = xy - xtQ(x,0) - ytQ(0,y) = xy - R(x) - R(y),$$

where, as in Kreweras' example, we denote by $R(x)$ the series $txQ(x,0)$. The kernel $K(x,y) = xy - t(x+y+x^2y+xy^2)$, considered as a polynomial in y, has two roots:

$$Y_0(x) = \frac{1 - t(x+\bar{x}) - \sqrt{(1-t(x+\bar{x}))^2 - 4t^2}}{2t} = t + (x+\bar{x})t^2 + O(t^3),$$

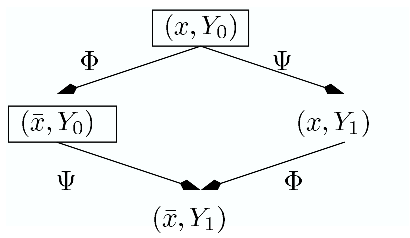
$$Y_1(x) = \frac{1 - t(x+\bar{x}) + \sqrt{(1-t(x+\bar{x}))^2 - 4t^2}}{2t} = \frac{1}{t} - x - \bar{x} - t - (x+\bar{x})t^2 + O(t^3).$$

The elementary symmetric functions of the Y_i are

$$Y_0 + Y_1 = \frac{1}{t} - x - \bar{x} \quad \text{and} \quad Y_0 Y_1 = 1.$$

Observe that they are no longer polynomials in $\bar{x} = 1/x$.

If, as above, we apply to the pair (x, Y_0) the transformations Φ and Ψ, we obtain a very simple diagram:



From the two pairs that can be substituted for (x,y) in Equation (15), we derive the following system:

$$\begin{aligned} R(x) + R(Y_0) &= xY_0, \\ R(\bar{x}) + R(Y_0) &= \bar{x}Y_0. \end{aligned}$$

From here, the method has to diverge from what we did in Kreweras' case. Eliminating $R(Y_0)$ between the two equations gives

$$R(x) - R(\bar{x}) = (x - \bar{x})Y_0. \tag{16}$$

Since $R(0) = 0$, extracting the positive part of this identity gives $R(x)$ as *the positive part of an algebraic series*. It is known that the positive part of a D-finite series is always D-finite [28]. In particular, the series $R(x)$ is D-finite. The same holds for $Q(x, 0)$, and, by (15), for $Q(x, y)$.

This argument is enough for proving the holonomy of the series, but, given the simplicity of this model, we can proceed with explicit calculations. Given the polynomial equation defining Y_0,

$$Y_0 = t(1 + \bar{x}Y_0 + xY_0 + Y_0^2) = t(1 + \bar{x}Y_0)(1 + xY_0),$$

the Lagrange inversion formula yields the following expression for Y_0:

$$Y_0 = \sum_{m\ge 0}\sum_{i\in\mathbb{Z}} \frac{x^i\, t^{2m+|i|+1}}{2m+|i|+1}\binom{2m+|i|+1}{m+|i|}\binom{2m+|i|+1}{m}.$$

Since $R(0) = 0$, extracting the positive part in the identity (16) now gives, after some reductions,

$$R(x) = txQ(x, 0) = \sum_{m\ge 0}\sum_{i\ge 0} \frac{x^{i+1}t^{2m+i+1}(i + 1)}{(2m + i + 1)(2m + i + 2)}\binom{2m+i+2}{m+i+1}\binom{2m+i+2}{m}.$$

This naturally fits with the general expression (1).

3.2. **Proof of Theorem 3.** We define two Laurent polynomials in y by

$$P_0(y) := \sum_{(0,j)\in\mathfrak{S}} y^j \quad\text{and}\quad P_1(y) := \sum_{(1,j)\in\mathfrak{S}} y^j.$$

Let $-p$ be the largest down move; more precisely,

$$p = \max(0, \{-j : (i, j) \in \mathfrak{S} \text{ for some } i\}).$$

The functional equation obtained by constructing walks step-by-step reads:

$$K(x, y)Q(x, y) = x^{1+i_0}y^{p+j_0} - ty^pP_1(y)Q(0, y) - t\sum_{(i,-j)\in\mathfrak{S}}\sum_{m=0}^{j-1}\left(Q_m(x) - \delta_{i,1}Q_m(0)\right)x^{1-i}y^{p+m-j} \tag{17}$$

where

$$K(x, y) = xy^p\left(1 - tP_0(y) - t(x + \bar{x})P_1(y)\right)$$

is the kernel of the equation, and $Q_m(x)$ stand for the coefficient of y^m in $Q(x, y)$. All the series involved in this equation also depend on the variable t, but it is omitted for the sake of briefness. For instance, $K(x, y)$ stands for $K(x, y; t)$.

As above, we shall use the kernel method – plus another argument – to solve the above functional equation. The polynomial $K(x, y)$, seen as a polynomial in y, admits a number of roots, which are Puiseux series in t with coefficients in an algebraic closure of $\mathbb{Q}(x)$. Moreover, all these roots are distinct. As $K(x, y; 0) = xy^p$, exactly p of these roots, say $Y_1, \ldots, Y_p$, vanish at $t = 0$. This property guarantees that these p series can be substituted for y in (17), which yields

$$x^{1+i_0}Y^{p+j_0} = tY^pP_1(Y)Q(0, Y) + t\sum_{(i,-j)\in\mathfrak{S}}\sum_{m=0}^{j-1}\left(Q_m(x) - \delta_{i,1}Q_m(0)\right)x^{1-i}\,Y^{p+m-j} \tag{18}$$

for any $Y = Y_1, \ldots, Y_p$.

Given the symmetry of K in x and $\bar{x}$, each of the Y_i is invariant by the transformation $x \to 1/x$. Replacing x by $\bar{x}$ in the above equation gives, for any $Y = Y_1, \ldots, Y_p$,

$$\bar{x}^{1+i_0} Y^{p+j_0} = tY^p P_1(Y) Q(0,Y) + t \sum_{(i,-j)\in\mathfrak{S}} \sum_{m=0}^{j-1} \left(Q_m(\bar{x}) - \delta_{i,1} Q_m(0)\right) x^{i-1} Y^{p+m-j}. \tag{19}$$

We now combine (18) and (19) to eliminate $Q(0,Y)$:

$$(x^{1+i_0} - \bar{x}^{1+i_0}) Y^{p+j_0} = t \sum_{(i,-j)\in\mathfrak{S}} \sum_{m=0}^{j-1} \left(x^{1-i} Q_m(x) - x^{i-1} Q_m(\bar{x})\right) Y^{p+m-j}$$

for any $Y = Y_1, \ldots, Y_p$. This is the generalization of Eq. (16). The right-hand side of the above equation is a polynomial P in Y, of degree at most $p-1$. We know its value at p points, namely $Y_1, \ldots, Y_p$. The Lagrange interpolation formula implies that these p values completely determine the polynomial. As the left-hand side of the equation is algebraic, then each of the coefficients of P is also algebraic. That is,

$$t \sum_{(i,-j)\in\mathfrak{S}} \sum_{m=0}^{j-1} \left(x^{1-i} Q_m(x) - x^{i-1} Q_m(\bar{x})\right) y^{p+m-j} = \sum_{m=0}^{p-1} A_m(x) y^m,$$

where each of the A_m is an algebraic series. Let us extract the positive part of this identity. Given that i can only be 0, 1 or -1, we obtain

$$t \sum_{(i,-j)\in\mathfrak{S}} \sum_{m=0}^{j-1} \left(x^{1-i} Q_m(x) - \delta_{i,1} Q_m(0)\right) y^{p+m-j} = \sum_{m=0}^{p-1} H_m(x) y^m$$

where $H_m(x) := A_m^+(x)$ is the positive part of $A_m(x)$. Again, this series can be shown to be D-finite. Going back to the original functional equation (17), this gives

$$K(x,y)Q(x,y) = x^{1+i_0} y^{p+j_0} - ty^p P_1(y) Q(0,y) - \sum_{m=0}^{p-1} H_m(x) y^m.$$

Let us finally[5] consider the kernel $K(x,y)$ as a polynomial in x. One of its roots, denoted below X, is a formal power series in t that vanishes at $t=0$. Replacing x by this root allows us to express $Q(0,y)$ as a D-finite series:

$$ty^p P_1(y) Q(0,y) = X^{1+i_0} y^{p+j_0} - \sum_{m=0}^{p-1} H_m(X) y^m.$$

The functional equation finally reads

$$K(x,y)Q(x,y) = (x^{1+i_0} - X^{1+i_0}) y^{p+j_0} - \sum_{m=0}^{p-1} (H_m(x) - H_m(X)) y^m.$$

Since the substitution of an algebraic series into a D-finite one gives another D-finite series, this equation shows that $Q(x,y)$ is D-finite.

■

[5]In the square lattice case, the symmetry of the model in x and y makes this step unnecessary: once the holonomy of $Q(x,0)$ is proved, the holonomy of $Q(x,y)$ follows.

4. **Perspectives**

This paper was written in a rush, right after the material of Section 2 was found. This new proof of Kreweras' formula suggests numerous questions and research directions, which I would like to explore in the coming weeks (or months...). Here are some of these questions.

4.1. **Other starting points.** It was observed by Gessel in [18] that the method he used to prove Kreweras' result was hard to implement for a starting point different from the origin. The reason of this difficulty is that, unlike the method presented here, Gessel's approach *checks* the known expression of the generating function, but does not *construct* it. I am confident that the new approach of Section 2 can be used to solve such questions. If the starting point does not lie on the main diagonal, the $x-y$ symmetry is lost; the diagram of Section 2.1 also loses its symmetry, and gives *four* different equations between the *two* unknown functions $Q(x,0)$ and $Q(0,y)$.

4.2. **Other algebraic walk models.** A close examination of the ingredients that make the proof of Section 2 work might help to construct other walk models which, for non-obvious reasons, would have an algebraic generating function. Note that for some degenerate sets of steps, like those of Figure 3, the quadrant condition is equivalent to a half-plane condition and thus yields an algebraic series.

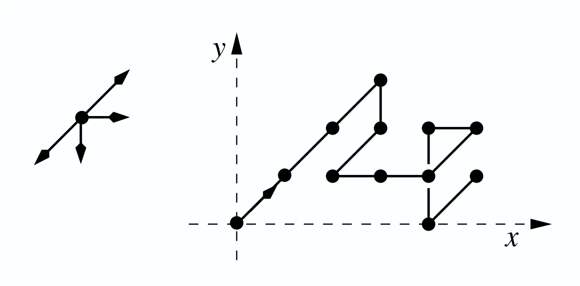


FIGURE 3. A degenerate set of steps.

I have started a systematic exploration of walks with few steps and only one up step: the non-trivial algebraic cases do not seem to be legion! However, I met in this exploration one model that seems to yield nice numbers (with a D-finite generating function) and for which the method of Section 2 "almost" works. I then realized that the same problem had been communicated to me, under a slightly different form, by Ira Gessel, a few months ago. I plan to explore this model further.

4.3. **Other equations.** Any combinatorial problem that seems to have an algebraic generating function and for which a linear functional equation with two "catalytic" variables (in the terminology of Zeilberger [42]) is available is now likely to be attacked via the method of Section 2. These conditions might seem very restrictive, but there is at least one such problem! The *vexillary* involutions, which were conjectured in 1995 to be counted by Motzkin numbers, satisfy the following equation:

$$\left(1+\frac{t^2x^2y}{1-x}+\frac{t^2y}{1-y}\right)F(x,y;t)=\frac{t^2x^2y^2}{(1-ty)(1-txy)}+t\left(1+\frac{ty}{1-y}\right)F(xy,1;t)+\frac{t^2x^2y}{1-x}F(1,y;t).$$

The conjecture was recently proved via a difficult combinatorial construction [23]. I have been able to apply successfully the method of Section 2 to this equation [7].

4.4. **Random walks in the quarter plane.** Random walks in the quarter plane are naturally studied in probabilities. Given a Markov chain on the first quadrant, a central question is the determination of an/the invariant measure $(p_{i,j})_{i,j\geq 0}$. The invariance is equivalent to a linear equation satisfied by the series $P(x,y) = \sum p_{i,j}x^iy^j$, in which the variables x and y are "catalytic". A whole recent book is devoted to the solution of this equation in the case where the walk has small horizontal and vertical variations [14]. This book contains *one* example for which the series $P(x,y)$ is algebraic: no suprise, the steps of the corresponding walk are exactly Kreweras' steps... This result is actually due to Flatto and Hahn [17]. The equation satisfied by the series $P(x,y)$ does not work exactly like the equations for complete generating functions like $Q(x,y;t)$: roughly speaking, the third variable t is replaced by the additional constraint $P(x,y) = 1$. However, I hope that either a refinement of the enumeration problem that would contain the invariant distribution as a limit distribution, or a direct adaptation of the method of Section 2 to the context of $P(x,y)$, will give a new, simpler proof of Flatto and Hahn's result. Their proof is based on non-trivial complex analysis, and uses a parametrisation of the roots of the kernel by elliptic functions, which are *not* algebraic. It seems that a large detour is done to end up with an algebraic series. I hope my new approach will offer a significant shortcut, by staying in the field of algebraic series.

Acknowledgements. To my shame, I must recall that, in the lecture that I gave at FPSAC'01 in Phoenix, I mentioned (part of) Kreweras' result as a conjecture. I am very grateful to Ira Gessel who enlighted my ignorance by giving me the right references.

References

[1] D. André, Solution directe du problème résolu par M. Bertrand, *C. R. Acad. Sci. Paris* **105** (1887) 436–437.

[2] C. Banderier, M. Bousquet-Mélou, A. Denise, D. Gardy, D. Gouyou-Beauchamps and P. Flajolet, Generating functions for generating trees, to appear in *Discrete Math.*

[3] C. Banderier and P. Flajolet, Basic analytic combinatorics of directed lattice paths, to appear in *Theoret. Comput. Sci.*

[4] E. Barcucci, E. Pergola, R. Pinzani and S. Rinaldi, A bijection for some paths on the slit plane, *Adv. in Applied Math.* **26** no. 2 (2001) 89–96.

[5] M. Bousquet-Mélou, Multi-statistic enumeration of two-stack sortable permutations, *Electronic J. Combinatorics* **5** (1998) R21.

[6] M. Bousquet-Mélou, Walks on the slit plane: other approaches, *Advances in Applied Math.* **27** No. 2-3 (2001) 243–288.

[7] M. Bousquet-Mélou, The number of vexillary involutions, in preparation.

[8] M. Bousquet-Mélou, On (some) functional equations arising in enumerative combinatorics, in *Proceedings of FPSAC'01*, Arizona State University, pp. 83–89.

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

[10] M. Bousquet-Mélou and M. Petkovšek, Walks confined in a quadrant are not always D-finite, submitted.

[11] M. Bousquet-Mélou and G. Schaeffer, Walks on the slit plane, preprint 2000, `arXiv:math.CO/0012230`. To appear in *Proba. Theory Related Fields.*

[12] W. G. Brown, On the existence of square roots in certain rings of power series, *Math. Annalen* **158** (1965) 82-89.

[13] P. Duchon, On the enumeration and generation of generalized Dyck words, *Discrete Math.* **225** (2000) 121–135.

[14] G. Fayolle, R. Iasnogorodski and V. Malyshev, *Random walks in the quarter plane*, Vol. 40 of *Applications of Mathematics: Stochastic Modelling and Applied Probabilities*, Springer-Verlag, Berlin Heidelberg, 1999.

[15] W. Feller, *An introduction to probability theory and its applications*, Vol.1, John Wiley and Sons, New York, 1950.

[16] P. Flajolet, Analytic models and ambiguity of context-free languages, *Theoret. Comput. Sci.* **49** (1987) 283–309.

[17] L. Flatto and S. Hahn, Two parallel queues created by arrivals with two demands I. *SIAM Journal of Applied Math.* **44** (1984) 1041–1053.

[18] I. M. Gessel, A probabilistic method for lattice path enumeration, *J. Stat. Planning Infer.* **14** (1986) 49–58.

[19] I. M. Gessel, A factorization for formal Laurent series and lattice path enumeration, *J. Combin. Theory. Ser. A* **28** (1980) 321–337.

[20] I. M. Gessel and D. Zeilberger, Random walk in a Weyl chamber, *Proc. Amer. Math. Soc.* **115** (1992) 27-31.

[21] I. P. Goulden and D. M. Jackson, *Combinatorial enumeration,* John Wiley and Sons, 1983.

[22] D. J. Grabiner, Random walk in an alcove of an affine Weyl group, and non-colliding random walks on an interval, in *Proceedings of FPSAC 2001*, Tempe, Arizona, pp. 217–225.

[23] O. Guibert, E. Pergola and R. Pinzani, Vexillary involutions are enumerated by Motzkin numbers, *Annals of Combinatorics* **5** (2001) 1–22.

[24] G. Kreweras, Sur une classe de problèmes liés au treillis des partitions d'entiers, *Cahiers du B.U.R.O.* **6** (1965) 5–105.

[25] J. Labelle, Langages de Dyck généralisés, *Ann. Sci. Math. Québec* **17** (1993) 53–64.

[26] J. Labelle and Y.-N. Yeh, Generalized Dyck paths, *Discrete Math.* **82** (1990) 1–6.

[27] G. F. Lawler, *Intersections of random walks*, Probabilities and its applications, Birkhäuser Boston, 1991.

[28] L. Lipshitz, The diagonal of a D-finite power series is D-finite, *J. Algebra* **113** (1988) 373–378.

[29] L. Lipshitz, D-finite power series, *J. Algebra* **122** (1989) 353–373.

[30] S. G. Mohanty, *Lattice path counting and applications*, Academic Press, 1979.

[31] T. V. Narayana, A partial order and its applications to probability theory, *Sankhya* **21** (1959) 91–98.

[32] H. Niederhausen, The ballot problem with three candidates, *Europ. J. of Combinatorics* **4** (1983) 161-167.

[33] H. Niederhausen, Lattice paths between diagonal boundaries, *Electronic J. Combinatorics* **5** (1998) R30.
[34] D. Poulalhon and G. Schaeffer, A bijection for loopless triangulations of a polygon with interior points, FPSAC 2002.
[35] G. Schaeffer, Random sampling of large planar maps and convex polyhedra, in *Proceedings of the 31th annual ACM Symposium on the Theory of Computing* (STOC'99), Atlanta, ACM press, 1999.
[36] F. Spitzer, *Principles of random walk*, The University Series in Higher Mathematics, Van Nostrand Company, Princeton, 1964.
[37] R. P. Stanley, Differentiably finite power series, *European J. Combin.* **1** (1980) 175–188.
[38] W. T. Tutte, A census of planar triangulations, *Canad. J. Math.* **14** (1962) 21–38.
[39] W. T. Tutte, A census of planar maps, *Canad. J. Math.* **15** (1963) 249–271.
[40] J. West, Sorting twice through a stack, *Theoret. Comput. Sci.* **117** (1993) 303–313.
[41] D. Zeilberger, A proof of Julian West's conjecture that the number of two-stack-sortable permutations of length n is $2(3n)!/((n+1)!(2n+1)!)$, *Discrete Math.* **102** (1992) 85–93.
[42] D. Zeilberger, The umbral transfer-matrix method: I. Foundations. *J. Comb. Theory, Ser. A*, **91** (2000) 451–463.

E-mail address: `bousquet@labri.fr`

CNRS, LaBRI, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence Cedex, France