Poster submission to FPSAC'02 (revised March 13, 2002).

LATTICE PATHS WITH AN INFINITE SET OF JUMPS

CYRIL BANDERIER AND DONATELLA MERLINI

Abstract. Whereas walks on $\mathbb{N}$ with a finite set of jumps were the subject of numerous studies, walks with an infinite number of jumps remain quite rarely studied. Even for relatively well structured models, the classical approach with context-free grammars fails as we deal with rewriting rules over an infinite alphabet. However, several classes of such walks offer a surprising structure: we make here explicit the associated bivariate functions, and give several theorems on their nature (rational, algebraic) via the kernel method or Riordan arrays theory. We give some examples of recent problems in combinatorics or theoretical computer science which lead to such rules.
Résumé. Tandis que les propriétés énumératives et asymptotiques des marches sur $\mathbb{N}$ avec un nombre fini de sauts ont fait l'objet de nombreuses études, les marches avec un nombre infini de sauts demeurent assez peu étudiées. Même pour des modèles relativement structurés, on ne peut utiliser les approches classiques par grammaires algébriques, puisqu'il s'agit de règles de récriture sur un alphabet infini. Toutefois, diverses classes de telles marches offrent une surprenante structure : nous explicitons ici la nature (algébrique, rationnelle) de la série génératrice bivariée associée (via la méthode du noyau ou la théorie des tableaux de Riordan). Nous illustrons l'intêret de telles marches en combinatoire et informatique théorique par quelques exemples.

Introduction

A considerable number of problems from computer science deals with a sum of independent identical distributed random variables $\Sigma_n = X_1 + X_2 + \cdots + X_n$ (where each of the X_i's assumes integer values). We will consider here the following model of random walks: the walk starts (at time 0) from a point Σ_0 of $\mathbb{Z}$ and at time n, one makes a jump $X_n \in \mathbb{Z}$; so the new position is given by the recurrence $\Sigma_n = \Sigma_{n-1} + X_n$ where, when $\Sigma_{n-1} = k$, the X_n's are constrained to belong to a fixed set $\mathcal{P}_k$ (that is, the possible jumps depend on the position of the walk).

These "walks on $\mathbb{Z}$" are homogeneous in time (that is to say, the set of jumps when one is at position k is independent from the time). When the positions Σ_n's are constrained to be nonnegative, we talk about "walks on $\mathbb{N}$". The probabilistic model under consideration here is the uniform distribution on all paths of length n.

When the sets $\mathcal{P}_k$'s are equal to a fixed set $\mathcal{P}$ (the simplest interesting case being $\mathcal{P} = \{-1, +1\}$), the corresponding walks have been deeply studied both in combinatorics and in probability theory. We refer to [3] for asymptotic properties of such "walks on $\mathbb{N}$ with a finite set of jumps". When the sets $\mathcal{P}_k$'s are unbounded, both enumeration and asymptotics become cumbersome: contrary to the previous case, the walks are not space-homogeneous (the set of available jumps depends on the position) and it is not possible to generate them by context-free grammars. However, if the sets $\mathcal{P}_k$'s have a "combinatorial" shape, it is reasonable to hope that the generating function associated to the corresponding walk would have some nice properties. We show here that this hope is legitimate and we present several classes of such walks, for which we are able to give the nature of their generating function.

Our results have potential impacts on the theory of generating trees (generation of combinatorial objects), the enumeration of general classes of lattice paths, and on the study of rewriting rules on an infinite alphabet.

A definition of the generating function associated to the walk is given in Section 1. In this first section, we also present the generating tree and Riordan array viewpoints. In Section 2, we give several theorems related to the nature of the generating functions associated to some walks and then we give some asymptotic results. In Section 3, we give some examples of problems in which some of the new classes of walks that we study in this article appeared.

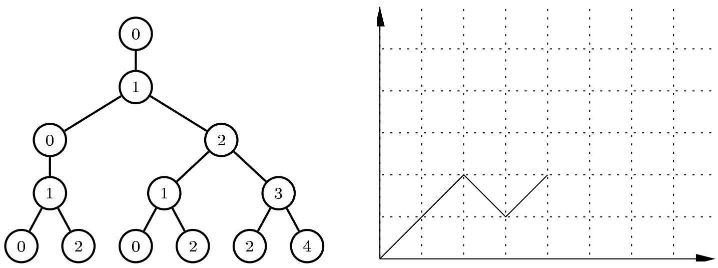

FIGURE 1. The generating tree of the walk on $\mathbb{N}$ with jumps $\mathcal{P} = \{+1, -1\}$ starting in 0 (and up to length $n = 4$). Each branch corresponds to a path. The branch $(0, 1, 2, 1, 2)$ corresponds to the path drawn on the lattice.

1. Lattice paths and generating trees

In combinatorics, it is classical to represent a particular walk as a path in a two dimensional lattice. Thus the drawing corresponds to the walk of length n linking the points $\big((0, \Sigma_0), (1, \Sigma_1), \ldots, (n, \Sigma_n)\big)$. It is also convenient to represent all the walks of length $\leq n$ as a tree of height n, where the root (at level 0 by convention) is labeled with the starting point of the walks and where the label of each node at level n encodes a possible position of the walk (see Figure 1).

We note $w_{n,k}$ the number of walks on $\mathbb{N}$ of length n going from the starting point to k (or, equivalently, the number of nodes with label k at level n in the tree) and we want to find the bivariate generating function

$$W(z, u) = \sum_{n \geq 0} w_n(u) z^n = \sum_{k \in \mathbb{Z}} W_k(z) u^k = \sum_{k \in \mathbb{Z}, n \geq 0} w_{n,k} u^k z^n \,,$$

where u encodes the final altitude of the walk (the label in the tree), z the length of the walk (the level in the tree), and where $w_n(u)$ is a Laurent polynomial (that is, a polynomial with finitely many monomials of negative and positive degree). When the walk is constrained to remain nonnegative (or equivalently when negative labels in the tree are not allowed), we consider similarly the bivariate generating function

$$F(z, u) = \sum_{n \geq 0} f_n(u) z^n = \sum_{k \in \mathbb{N}} F_k(z) u^k = \sum_{k \in \mathbb{N}, n \geq 0} f_{n,k} u^k z^n \,. \tag{1}$$

Generating trees and rewriting rules. The concept of generating trees has been used from various points of view and was introduced in the literature by Chung, Graham, Hoggatt and Kleiman [6] to examine the reduced Baxter permutations. This technique has been successively applied to other classes of permutations. A generating tree is a rooted labeled tree with the property that if v_1 and v_2 are any two nodes with the same label then, for each label ℓ, v_1 and v_2 have exactly the same number of children with label ℓ. To specify a generating tree it therefore suffices to specify: 1) the label of the root; 2) a set of rules explaining how to derive from the label of a parent the labels of all of its children. Points 1) and 2) define what we call a *rewriting rule*. For example, Figure 1 illustrates the upper part of the generating tree which corresponds to the rewriting rule $[(0), \{(k) \rightsquigarrow (k-1)(k+1)\}]$.

Riordan arrays We introduce now the concept of *matrix associated to a generating tree*: this is an infinite matrix $\{d_{n,k}\}_{n,k \in \mathbb{N}}$ where $d_{n,k}$ is the number of nodes at level n with label $k + r$, r being the label of the root. For example, the matrix associated to the generating tree of the Figure 1 is the following:

n/k	0	1	2	3	4
0	1				
1	0	1			
2	1	0	1		
3	0	2	0	1	
4	2	0	3	0	1

Many such matrices can be studied from a *Riordan array* viewpoint. In fact, the concept of a Riordan array provides a remarkable characterization of many lower triangular arrays that arise in combinatorics and algorithm analysis. The theory has been introduced in the literature in 1991 by Shapiro, Getu, Woan and Woodson [11]. Riordan arrays are a powerful tool in the study of many counting problems [7].

A Riordan array is an infinite lower triangular array $\{d_{n,k}\}_{n,k\in\mathbb{N}}$, defined by a pair of formal power series $D = (d(z), h(z))$, such that the k-th column is given by $d(z)(zh(z))^k$, i.e.:

$$d_{n,k} = [z^n]d(z)(zh(z))^k, \qquad n,k \geq 0.$$

From this definition we have $d_{n,k} = 0$ for $k > n$. The bivariate generating function for D is:

$$\sum_{n,k\geq 0} d_{n,k}u^k z^n = \frac{d(z)}{1 - uzh(z)}.$$

In what follows, we always assume that $d(0) \neq 0$; if we also have $h(0) \neq 0$ then the Riordan array is said to be *proper*; in the proper-case the diagonal elements $d_{n,n}$ are different from zero for all $n \in \mathbb{N}$. The most simple example is the Pascal triangle for which we have

$$\binom{n}{k} = [z^n]\frac{1}{1-z}\left(\frac{z}{1-z}\right)^k,$$

where we recognize the proper Riordan array with $d(z) = h(z) = 1/(1-z)$. Proper Riordan arrays are characterized by the existence of a sequence $A = \{a_i\}_{i\in\mathbb{N}}$ with $a_0 \neq 0$, called the *A-sequence*, such that every element $d_{n+1,k+1}$ can be expressed as a linear combination, with coefficients in A, of the elements in the preceding row, starting from the preceding column:

$$d_{n+1,k+1} = a_0 d_{n,k} + a_1 d_{n,k+1} + a_2 d_{n,k+2} + \cdots$$

It can be proved that $h(z) = A(zh(z))$, $A(z)$ being the generating function for A. For example, for the Pascal triangle we have: $A(z) = 1 + z$ and the previous relation reduces to the well-known recurrence relation for binomial coefficients. The A-sequence doesn't characterize completely $(d(z), h(z))$ because $d(z)$ is independent of $A(z)$. But it can be proved that there exists a unique sequence $Z = \{z_0, z_1, z_2, \ldots\}$, such that every element in column 0 can be expressed as a linear combination of all the elements of the preceding row:

$$d_{n+1,0} = z_0 d_{n,0} + z_1 d_{n,1} + z_2 d_{n,2} + \cdots$$

This property has been recently studied in [7], where it is proved that $d(z) = d(0)/(1 - zZ(zh(z)))$, $Z(z)$ being the generating function for Z. Thus the triple $(d(0), Z(z), A(z))$ characterizes every proper Riordan array.

2. Walks on $\mathbb{Z}$ with an infinite set of negative jumps

2.1. **Lattice paths and generating trees.** Consider a sequence $(e_i(k))_{i\geq -a}$ (for a given integer $a > 0$) of polynomials assuming nonnegative integers values then the walk with an infinite set of jumps under consideration here are of the following kind:

$$[(r), \{(k) \rightsquigarrow (0)^{e_k(k)}(1)^{e_{k-1}(k)}(2)^{e_{k-2}(k)} \ldots (k-1)^{e_1(k)}(k)^{e_0(k)} \ldots (k+a)^{e_{-a}(k)}\}], \tag{2}$$

where the exponent $e_i(k)$ is the multiplicity of the jumps $-i$ when one is at position k and where r is the starting position of the walk (or equivalently, the root of the associated generating tree).

If the sequence of polynomials $(e_i(k))_{i\geq -a}$ is ultimately $e_i(k) = 0$, then the situation covers the case of walks with a finite set of jumps. If the sequence is ultimately $e_i(k) = 1$, then this covers the case of "factorial rules" which are of great interests for the generation of combinatorial objects [4] and for which it was proven in [2] that the associated generating functions are algebraic.

We still note $f_{n,k}$ the number of walks on $\mathbb{N}$ of length n going from the starting point to k and we want to find the bivariate generating function $F(z,u) = \sum_{n,k\geq 0} f_{n,k}u^k z^n$. These random walks on $\mathbb{N}$ can equivalently be seen as lattice paths, generating trees and also as Riordan arrays (when $a = 1$).

Rule	EIS approximate description	Generating Function $F(z,u)$
	Rational OGF	**OGF**
$(0),\{(k)\rightsquigarrow(0)^k(k+1)\}$	F_0, $F(z,1)$: powers of 2	$\frac{1-2z-z^2}{1-(u+2)z-2uz^2}$
$(0),\{(k)\rightsquigarrow(0)^{2k}(k+1)\}$	$F(z,1)$:A001333 continued fraction convergents to $\sqrt{2}$ F_0: A052542 (ECS)	$\frac{1-2z+z^2}{1-(u+2)z+(2u-1)z^2+uz^3}$
$(0),\{(k)\rightsquigarrow(0)^{3k}(k+1)\}$	$F(z,1)$: A026150 (ECS)	$\frac{1-2z+z^2}{1-(u+2)z+(2u-2)z^2+2uz^3}$
$(0),\{(k)\rightsquigarrow(0)^{4k}(k+1)\}$	$F(z,1)$: A046717 half of 3^n	$\frac{1-2z+z^2}{1-(u+2)z+(2u-3)z^2+3uz^3}$
$(0),\{(k)\rightsquigarrow(0)^k(k+1)(k+2)\}$	$F(z,1)$: A001075 and F_0: A005320 Pell's equation	$\frac{1-4z+4z^2}{1-(4+u+u^2)z+(4u^2+u-1)z^2-\dots}$
$(1),\{(k)\rightsquigarrow(0)(1)^2(k)(k+2)^2(k+3)^5\}$	6^n and A003464 $(6^n-1)/5$	$\frac{(4u-1)z-u}{(1-6z)((2u^2+1)z-1)}$
$(0),\{(k)\rightsquigarrow(0)^{k^2}(2)^{3k-1}(3)(k)(k+1)^2(k+3)^5\}$		see Theorem 1
	Algebraic OGF	**OGF**
$(1),\{(k)\rightsquigarrow(1)\dots(k+s-2)(k+s-1)\}$	$F(z,1)$: s-ary trees	
$(1),\{(k)\rightsquigarrow(1)^2\dots(k)^2(k+1)\}$	$F(z,1)$: A001003 Schröder's second problem	$\frac{u}{2}\frac{1-(2u+1)z-\sqrt{1-6z-z^2}}{(1-u)z+(u^2+u)z^2}$
$(0),\{(k)\rightsquigarrow(0)^{k^2}(2)^{3k-1}(3)(k)(k+1)^2(k+3)^5\}$		
$(0),\{(k)\rightsquigarrow(0)^k(1)^{k-1}\dots(k-1)^1(k)^0(k+1)\}$	A036765 $F(z,1)$: rooted trees with a degree constraint	equation of degree 3
$(0),\{(k)\rightsquigarrow(0)^{k+2}(1)^{k+1}\dots(k-1)^3(k)^2(k+1)\}$	F_0: A006013 A046648 noncrossing trees on a circle $F(z,1)$: A001764 ternary trees	equation of degree 3
$(0),\{(k)\rightsquigarrow(0)^{k+3}\dots(k-1)^4(k)^3(k+1)^2(k+2)\}$	$F(z,1)$: A066357 planar trees with root parity constraint	equation of degree 4
$(0),\{(k)\rightsquigarrow(0)^{C_k}\dots(k-1)^{C_1}(k)^{C_0}(k+1)\}$ (where C_k is the k-th Catalan number)	F_0: A006318 large Schröder numbers	$\frac{1}{2}\frac{3-(4u+1)z-\sqrt{1-6z-z^2}}{1-3uz+(2u^2+u)z^2}$
$(0),\{(k)\rightsquigarrow(0)^{C_k}\dots(k-1)^{C_1}(k+1)\}$	F_0: A052705 (ECS)	$\frac{1}{2}\frac{3-(4u+2)z-\sqrt{1-4z-4z^2}}{1-(3u+2)z+(2u^2-2u+1)z^2}$
$(0),\{(k)\rightsquigarrow(0)^{T_k}\dots(k-1)^{T_1}(k)^{T_0}(k+1)\}$ (where T_k is the k-th tri-Catalan number)	F_0: A054727 noncrossing forests of rooted trees	equation of degree 3

TABLE 1. Some rewriting rules with simple combinatorial patterns. The ordinary generating functions $F(z,1)$ and $F_0(z)$ are defined as in Equation 1.

In Table 1, we give a list of rewriting rules with simple combinatorial patterns, the reference to famous numbers or combinatorial problems they refer to, the generating function $F(z,1)$, and the numbers identifying the corresponding sequences in the On-Line Encyclopedia of Integer Sequences `http://www.research.att.com/~njas/sequences/`; ECS stands for the Encyclopedia of Combinatorial Structures `http://algo.inria.fr/encyclopedia/`.

2.2. Rationality and algebraicity of classes of rewriting rules.

Theorem 1. *For any constant $B\geq 0$, the rule*

$$[(r),\{(k)\rightsquigarrow(0)^{e_k(k)}\dots(B)^{e_{k-B}(k)}\quad(k)^{e_0}\dots(k+a)^{e_{-a}}\}]$$

(where $e_k(k),\dots,e_{k-B}(k)$ are polynomial in k, $e_i(k)=0$ for $0<i<k-B$ and $e_i(k)=e_i$, some fixed constants, for $i\leq 0$) has a rational generating function $F(z,u)$.

Proof. First, we illustrate the general case by the following example:

$$[(0),\{(k)\rightsquigarrow(0)^{k^2}(2)^{3k-1}(3)(k)(k+1)^2(k+3)^5\}],$$

for which $B=3$, the polynomials in k are $e_k(k)=k^2, e_{k-1}=0, e_{k-2}=3k-1, e_{k-3}=1$, and the fixed constants are $e_0=1$, $e_{-1}=2$, $e_{-2}=0$, $e_{-3}=5$.

The part $(k) \rightsquigarrow (0)^{k^2}$ implies a transformation $u^k \rightsquigarrow k^2u^0$. The part $(k) \rightsquigarrow (2)^{3k-1}$ implies a transformation $u^k \rightsquigarrow (3k-1)u^2$. The part $(k) \rightsquigarrow (3)$ implies a transformation $u^k \rightsquigarrow u^3$. It is possible to perform all these transformations using the derivation, evaluation in $u = 1$ and multiplication by a monomial: in the first case, the multiplicity k^2 is obtained by $\partial(u\partial(u^k))$ and then evaluating in $u = 1$; for the second case, the multiplicity $3k-1$ is obtained by taking $\partial_u(u^{3k})/u$ and then evaluating in $u = 1$; for the third case simply evaluate in $u = 1$ and multiply by u^3. The part $(k) \rightsquigarrow (k)(k+1)^2(k+3)^5$ gives $u^k \rightsquigarrow P(u)u^k$ where $P(u) = 1 + 2u + 5u^3$. All these transformations are in fact linear, so to act on u^k or a polynomial in u (like $f_n(u)$) is the same. Finally, evaluating $\partial(u\partial f_n(u))$ in $u = 1$ gives $f_n''(1) + f_n'(1)$ and evaluating $u^2\partial_u f_n(u^3)/u$ in $u = 1$ gives $u^2(3f_n'(1) - f_n(1))$, so these trivial simplifications gives the following recurrence:

$$f_{n+1}(u) = P(u)f_n(u) + u^0(f_n''(1) + f_n'(1)) + u^2(3f_n'(1) - f_n(1)) + u^3 f_n(1)\,.$$

Multiplying by z^{n+1} and summing for $n \geq 0$ leads to the functional equation

$$(1 - zP(u))F(z,u) = 1 + z(u^3-1)F(z,1) + z(3u^2+1)F'(z,1) + zF''(z,1)\,.$$

Taking the first 2 derivatives and instantiating in $u = 1$ gives a rational system of full rank, hence $F(z,u)$ is rational:

$$F(z,u) = \frac{u^3(22z^2 - 112z^3 - z) + u^2(480z^3 - 60z^2) + 528z^3 - 250z^2 + 31z - 1}{(1 - zP(u))(872z^3 - 212z^2 + 30z - 1)}\,.$$

For the general case, one has the following functional equation

$$(1 - zP(u))F(z,u) = u^r + z\sum_{i=0}^{d} t_i(u)\partial_u^i F(z,1)$$

(d is the largest degree of the polymonials $e_i(k)$, and the t_i's are some Laurent polynomials which can be made explicit). Taking the first d derivatives and instantiating in $u = 1$ gives a system (for $m = 0,\ldots,d$):

$$\begin{aligned}\partial_u^m u^r \quad &+ \quad \left(\sum_{i=0}^{m-1}\left(z\partial_u^m t_i(1) + z\binom{m}{i}\partial_u^{m-i}P(1)\right)\partial_u^i F(z,1)\right)\\ &+ \quad \left(z\partial_u^m t_i(1) - (1 - zP(1))\right)\partial_u^m F(z,1) + z\sum_{i=m+1}^{d}\partial_u^m t_i(1)\partial_u^i F(z,1) = 0\,.\end{aligned}$$

This gives a matricial equation $M.\overrightarrow{F} = \overrightarrow{v}$ where $\overrightarrow{F} = (\partial_u^0 F(z,1),\ldots,\partial_u^d F(z,1))^T$ and $\overrightarrow{v} = (u^r,0,\ldots,0)^T$. The coefficients of the main diagonal of M are $-1+z\ldots$ (as they are the coefficients of the $\partial_u^m F(z,1)$ summand) and all the other coefficient of M are monomials in z of degree 1. Thus, one has $[z^0]\det M = \pm 1$ and then $\det M \neq 0$. Consequently, this system is of full rank. Solving it gives rational expressions for the $\partial_u^i F(z,1)$ and for $F(z,u)$. □

We now give a generalization of a result of [2] which was giving the algebraicity of "factorial rules": we allow here initial multiplicities which are not space-homogeneous.

Theorem 2. *For a constant $B \geq 0$, the rule*

$$[(r),\{(k) \rightsquigarrow (0)^{e_k(k)}\ldots(B)^{e_{k-B}(k)}(B+1)\ldots(k-b-1)(k-b)^{e_b}\ldots(k+a)^{e_{-a}}\}]$$

(where $e_k(k),\ldots,e_{k-B}(k)$ are polynomial in k, $e_i(k) = 1$ for $b < i < k-B$ and $e_i(k) = e_i$, some fixed constants, for $i \leq b$) has an algebraic generating function $F(z,u)$.

Proof. We illustrate the general case by the following example:

$$[(0),\{(k) \rightsquigarrow (0)^{k^2}(2)^{3k^5-2}(6)(7)\ldots(k-5)(k-4)^2(k-2)^3(k)(k+3)^2(k+23)\}]$$

for which $B = 5, b = 4, a = 23$, the polynomials in k are $e_k(k) = k^2$, $e_{k-2}(k) = 3k^5 - 2$, $e_{k-1}(k) = e_{k-3}(k) = e_{k-4}(k) = e_{k-5}(k) = 0$ and the fixed constants are $e_4 = 2$, $e_2 = 3$, $e_0 = 1$, $e_{-3} = 2$,

$e_{-23}=1$. One sets $P(u)=2u^{-4}+3u^{-2}+1+2u^3+u^{23}$, the recurrence is

$$f_{n+1}(u)=P(u)f_n(u)-\{u^{<0}\}P(u)f_n(u)+\sum_{i=0}^{5} t_i(u)\partial_u^i f_n(1)\,,$$

where $\{u^{<0}\}$ stands for the sum of the monomials in u with a negative degree. Multiplying by z^{n+1} and summing for $n\geq 0$ leads to the functional equation

$$(1-zP(u))F(z,u)=1-z\sum_{k=0}^{4-1} r_k(u)F_k(z)+z\sum_{i=0}^{5} t_i(u)\partial_u^i F(z,1)\,, \tag{3}$$

where $r_k(u):=\{u^{<0}\}P(u)u^k$ and $t_i(u)$ are (Laurent) polynomials which can be made explicit.

One can use the kernel method (we refer to [3, 5] for recent applications of this method) to solve this equation. We call $1-zP(u)$ the *kernel* of the equation. Solving $1-zP(u)=0$ with respect to u gives 4 roots $u_1(z)$, $u_2(z)$, $u_3(z)$ and $u_4(z)$ which are Puiseux series in $z^{1/4}$ and which tend to zero in 0. There are also 23 others roots which behave like $z^{-1/23}$ around 0, so we call $u_1,\ldots,u_4$ the *small* roots of the kernel. Plugging the 4 small roots of the kernel in Equation 3 and considering the 6 other equations obtained by taking the first 5 derivatives of Equation 3 (and then setting $u=1$) gives a system of full rank with 10 equations with 10 unknown univariate generating functions, which are thus all algebraic, and then one has a formula for $F(z,u)$, involving the u_i, which implies its algebraicity. For the general case, simply replace 4 by b and 5 by d in Equation 3 and then one can argue as in Theorem 1 above, with a new matricial equation $M.\overrightarrow{F}=\overrightarrow{v}$; looking at the valuation in z of each entries in M (some of them involves the small roots u_i's, but at most a product of b of them) gives $\det M\neq 0$ and thus a system of full rank, so $F(z,u)$ can be expressed as a rational function in z, u and the small roots u_i's. As these roots are algebraic, $F(z,u)$ is algebraic. □

Consider now the case where, for each i, the exponent $e_i(k)$ of the rule (2) is a constant (that is, the polynomial in $e_i(k)$ does not depend on k, so one simply writes e_i). How far can we relate the behavior of the walk

$$[(0),\{(k)\rightsquigarrow (0)^{e_k}(1)^{e_{k-1}}\ldots(k-2)^{e_2}(k-1)^{e_1}(k)^{e_0}(k+1)^{e_{-1}}\ldots(k+a)^{e_{-a}}\}] \tag{4}$$

to the generating function of the exponents $E(u)=\sum_{i\geq -a} e_i u^i$? We give here a first element of answer:

Theorem 3. *Consider the rule*

$$[(0),\{(k)\rightsquigarrow (0)^{e_k}(1)^{e_{k-1}}\ldots(k-1)^{e_1}(k)^{e_0}\ldots(k+a)^{e_{-a}}\}]\,. \tag{5}$$

If the generating function of the exponents $E(u)$ is algebraic then the bivariate generating function of the walk $F(z,u)$ is algebraic. For $a=1$, one has

$$F(z,u)=\frac{F_0(z)}{1-u\,e_{-1}z\,F_0(z)}\qquad \textit{with}\qquad F_0(z)=\frac{1}{e_{-1}z}\,E^{<-1>}(\frac{1}{z})$$

where $E^{<-1>}$ is the compositional inverse of $E(u)$ and where e_{-1} is the multiplicity of the $+1$ jump. More generally, for $a\geq 1$, the generating function $F(z,u)$ is expressed in terms of the a solutions $u_1(z),\ldots,u_a(z)$ of $1-zE(u)=0$ which satisfy $u_i(z)\sim 0$ for $z\sim 0$:

$$F(z,u)=F_0(z)\prod_{i=1}^{a}\frac{1}{1-uu_i(z)}=\sum_{k\geq 0}F_0(z)\left(\sum_{i_1+\cdots+i_a=k}u_1^{i_1}\ldots u_a^{i_a}\right)u^k\,.$$

One has

$$F_0(z)=\frac{(-1)^{a+1}}{ze_{-a}}\prod_{i=1}^{a}u_i(z)\qquad \textit{and}\qquad F(z,1)=\frac{-1}{ze_{-a}}\prod_{i=1}^{a}\frac{1}{1-\frac{1}{u_i(z)}}\,.$$

Proof. For $a=1$, the first identity reflects the combinatorial decomposition (one to one correspondence, in fact) "a walk from 0 to $k+1$" is "a walk from 0 to k" then followed by a jump $+1$ then followed by "a walk from $k+1$ to $k+1$ never going below $k+1$". The generating function of these last walks is clearly $F_0(z)$, thus one has $F_{k+1}(z)=F_k(z)e_{-1}zF_0(z)=F_0(z)(ze_{-1}F_0(z))^{k+1}$.

For the walks corresponding to the rule (5), the set of jumps is given by $E(1/u)$; if one reverses the time direction, one gets a new walk where the set of available jumps is given by $E(u)$. Define $\widetilde{F}(z,u)$ as the corresponding generating function (one starts at altitude 0), one has:

$$\widetilde{f}_{n+1}(u) = \{u^{\geq 0}\}E(u)\widetilde{f}_n(u), \qquad \widetilde{f}_0(u) = 1$$

where $\{u^{\geq 0}\}$ stands for the sum of all monomials in u with a nonnegative degree. Multiplying by z^{n+1} and summing for $n \geq 0$ gives

$$\widetilde{F}(z,u) = \widetilde{f}_0(u) + zE(u)\widetilde{F}(z,u) - z\{u^{-1}\}\frac{e_{-1}}{u}\widetilde{F}(z,u)\,,$$

that one rewrites as the following functional equation

$$(1 - zE(u))\widetilde{F}(z,u) = 1 - z\frac{e_{-1}}{u}\widetilde{F}_0(z)\,.$$

Then solving the "kernel" $1 - zE(u) = 0$ with respect to u gives a series $u_1(z) = E^{<-1>}(1/z)$, which is algebraic as the compositional inverse of an invertible algebraic function is algebraic (simply plug the inverse in the polynomial equation $\Phi(E(u), u) = 0$ satisfied by $E(u)$ to check this fact).

If one then evaluates the above functional equation at $u = u_1(z)$, one gets $0 = 1 - z\frac{e_{-1}}{u_1}\widetilde{F}_0(z)$ and thus $\widetilde{F}_0(z) = \frac{u_1}{e_{-1}z}$. As one has $\widetilde{F}_0(z) = F_0(z)$ (a walk from 0 to 0 from left to right is still a walk from 0 to 0 from right to left), one gets the result from the theorem. Note that if one sets $\widetilde{f}_0(u) = \frac{1}{1-u}$, $\widetilde{F}_0$ enumerates walks from anywhere to 0, so $\widetilde{F}_0(z) = \frac{u_1/(ze_{-1})}{1-u_1} = F(z,1)$, which is coherent with the theorem (case $a = 1$).

For $a \geq 1$, one sets $P(u) := \sum_{i=-a}^{-1} e_i u^i$; one has

$$(1 - zE(u))\widetilde{F}(z,u) = \widetilde{f}_0(u) - z\{u^{<0}\}P(u)\widetilde{F}(z,u)\,.$$

This is rewritten as

$$(1 - zE(u))\widetilde{F}(z,u) = \widetilde{f}_0(u) - z\sum_{k=0}^{a-1} r_k(u)\widetilde{F}_k(z)\,. \tag{6}$$

where $r_k(u) := \{u^{<0}\}P(u)u^k$ is a Laurent polynomial with monomials of degree going from -1 down to $k - a$.

The kernel equation $1 - zE(u) = 0$ has a roots $u_1(z), \ldots, u_a(z)$ which are Puiseux series in $z^{1/a}$ and which tend to 0 when z tends to 0. When $\widetilde{f}_0(u) = 1$, plugging these roots in the functional equation shows that they correspond to the a roots of the polynomial $u^a - zu^a \sum_{k=0}^{a-1} r_k(u) F_k(z)$, whose leading term is z^a and whose constant term is so $-ze_{-a}\widetilde{F}_0(z)$. This gives $\widetilde{F}_0(z) = \frac{-\prod_{i=1}^{a} u_i}{-ze_{-a}}$. When $\widetilde{f}_0(u) = \frac{1}{1-u}$ this gives a system of a equations for a unknowns (the $\widetilde{F}_k$'s). Solving it for $\widetilde{F}_0$ gives $F(z,1)$. Solving the $\widetilde{F}_0$ for $\widetilde{f}_0(u) = u^k$ gives the $F_k(z)$.

For $a = 1$, the Riordan arrays approach that we presented in Section 1 also gives the algebraicity of $F(z,u)$. In fact, a theorem from [10] gives $F(z,u) = \frac{d(z)}{1-uzh(z)}$ where $h(z) = A(zh(z))$ and $d(z) = 1/(1 - zZ(zh(z)))$ for the rule $[(0), \{(k) \rightsquigarrow (0)^{z_k}(1)^{a_k}(2)^{a_{k-1}} \ldots (k)^{a_1}(k+1)^{a_0}\}]$. For $a > 1$, the matrix associated (see Section 1) to the rule (4) is called a *horizontally stretched Riordan array*. With this concept, it can be shown, like with the kernel method, that the algebraicity of the corresponding generating function $F(z,u)$ depends on the algebraicity of $A(z) = \sum_{k\geq 0} a_k z^k$ and $F_0(z), \ldots, F_{a-1}(z)$ (the generating functions of the first a columns of the matrix). While the theory of Riordan arrays has been intensively studied, the theory of stretched Riordan arrays, from a generating function point of view, is still in progress. □

Remark: as D-finite functions are not necessarily closed under compositional inverse, it is not true that if $E(u)$ is D-finite, then $F(z,1)$ or $F_0(z)$ (and a fortiori $F(z,u)$) are D-finite, even in the case $a = 1$.

We end with a last application of the kernel method.

Theorem 4. *Consider the rewriting rule (4) when the e_i's are ultimately constants (say, equal to a constant C after rang b): $[(0), \{(k) \rightsquigarrow (0)^C \dots (k-b-1)^C (k-b)^{e_b} \dots (k)^{e_0} \dots (k+a)^{e_{-a}}\}]$. Then $F(z,u)$ is algebraic and satisfies*

$$F(z,u) = \frac{\prod_{i=0}^{b} u - u_i(z)}{K(z,u)}\,,$$

where the u_i's and K are defined as below.

Proof. One has the recurrence $f_{n+1}(u) = C\frac{f_n(u)-f_n(1)}{u-1} + P(u)f_n(u)$ this leads to the functional equation

$$\left(1 - zP(u) - z\frac{C}{u-1}\right)F(z,u) = 1 - \frac{zC}{u-1}F(z,1) - z\sum_{k=0}^{b-1}\{u^{<0}\}P(u)u^k F_k(z) \tag{7}$$

where $P(u) = \sum_{i=1}^{b}(e_i - C)\frac{1}{u^i} + \sum_{i=0}^{a} e_{-i}u^i$. Define the kernel K as $K(u,z) = u^b(1-u)(1-zP(u) - \frac{zC}{u-1})$. It has b roots $u_1(z), \dots, u_b(z)$ which are Puiseux series in $z^{1/b}$ and which tend to 0 in 0 and one root $u_0(z)$ which tends to 1 in 0. These are exactly the $b+1$ roots of the right hand part of (7) (once multiplied by $(1-u)u^b$). So $F(z,u) = \frac{\prod_{i=0}^{b} u-u_i(z)}{K(z,u)}$, where the u_i's are the $b+1$ small roots of the kernel. □

2.3. **Asymptotics.** Given a peculiar rule for Theorem 1, 2, 3 or 4, it is possible to find an asymptotic expansion for the number of walks. It is not really possible to merge all these results in a single one, as the rules are too unconstrained. However, for the algebraic case, a kind of universality holds for the behavior of the roots of the kernel. This leads to following theorem, which has to be adapted case by case for rules of Theorems 2 and 3 (and is easily applied to rules of Theorem 4).

Theorem 5. *The number of walks of length n for the "factorial" rule*

$$[(0), \{(k) \rightsquigarrow (0)(1)\dots(k-b-1)(k-b)^{e_b} \dots (k)^{e_0} \dots (k+a)^{e_{-a}}\}]$$

(where $e_i(k) = 1$ for $b < i \leq k$ and $e_i(k) = e_i$, some fixed constants, for $i \leq b$) has the following asymptotics $A\frac{\rho^{-n}}{\sqrt{2\pi n^3}}$, where A and ρ are algebraic constants depending on the finite set of jumps $\mathcal{P}$.

Proof. See [1] for a proof and applications to the limit laws of final altitude and number of factors. The approach is similar to the one used for walks with a finite number of jumps but there are some complications due to the fact that the kernel is now of the kind $1 - z\phi(u)$ where $\phi(u)$ is not unimodal. One can however establish that the real positive root u_0 now dominates and has a square-root behavior. □

3. Examples

We now give a series of examples from combinatorics or computer science in which rewriting rules studied in Section 2 appear.

Example 1. *Two families of rules leading to an algebraic generating function.*
For the rule $[(0), \{(k) \rightsquigarrow (0)^{e_k}(1)^{e_{k-1}} \dots (k-1)^{e_1}(k)^{e_0}(k+1)\}]$, where e_k for $i \geq 0$ is the number of t-ary trees with k nodes, $F(z,u)$ satisfies a algebraic equation of degree t. E.g., for $t = 3$, one has: $1 - \big(3 + (4-3u)z\big)F(z,u) - \big(-3 + (6u-7)z + (-3u^2+8u-3)z^2\big)F(z,u)^2 - \big(1 + (3-3u)z + (3u^2 - 7u+3)z^2 + (-u^3+4u^2-3u+1)z^3\big)F(z,u)^3 = 0$.

For the rule $[(0), \{(k) \rightsquigarrow (0)^{c+k}(1)^{c+k-1} \dots (k-2)^{c+2}(k-1)^{c+1}(k)^c(k+1)\}]$, $F(z,u)$ satisfies an algebraic equation of degree 3:
$\big((1-2u)z^2 + (c-(c+1)+2u^2)\big)F^3 + \big((u-2)z + (-c-2+4u-2u^2)z^2\big)F^2 + \big(1+(2-2u)z\big)F = 1$.
□

Example 2. *Tennis ball problem.* Let $s \geq 2$ be an integer and consider the following problem known as *the s-tennis ball problem.* At the first turn one is given balls numbered one through s. One throws one of them out of the window onto the lawn. At the second turn balls numbered $s+1$

through $2s$ are brought in and now one throws out on the lawn any of the $2s-1$ remained. Then balls $2s+1$ through $3s$ are brought in and one throws out one of the $3s-2$ available balls. The game continues for n turns. At this point, one picks up the n balls in the lawn and consider the ordered sequence $B = (b_1, b_2, \dots, b_n)$ with $b_1 < b_2 < \cdots < b_n$. This sequence will be called a *tennis ball s-sequence* and the first question is: how many tennis ball s-sequences of length n exist? The second question is: what is the sum of all the balls in all the possible s-sequences of length n ? Obviously, if we answer to both these questions, we also know the average sum of the balls in an s-sequence of length n. The general case $s \geq 1$ has been studied in [8] from a generating function viewpoint. In fact, the authors consider an infinite tree with root 0 and with s children. Each $(n+1)$-length path in this tree corresponds to an s-sequence of length n. This infinite tree is isomorphic to the generating tree with specification $[(1), \{(k) \rightsquigarrow (1)\dots(k+s-2)(k+s-1)\}]$.

By using this result the authors find that the number of tennis ball s-sequences of length n are counted by T_{n+1}, where $T_n = \frac{1}{1+(s-1)n}\binom{sn}{n}$ (the number of s-ary trees with n-nodes) and the cumulative sum of all the balls thrown onto the lawn in n turn is

$$\Sigma_n = \frac{1}{2}(sn^2 + (3s-1)n + 2s)T_{n+1} - \frac{1}{2}\sum_{k=0}^{n+1}\binom{sk}{k}\binom{s(n+1-k)}{n+1-k}.$$

□

Example 3. A new rewriting rule for $(4,2)$-tennis ball problem.

The problem of balls on the lawn admits many other variants. For example, one could be supplied with s balls at each turn but now throw out t balls at a time with $t < s$. The general (s,t) case is an open problem while the $(4,2)$ case has been treated in [8], where the authors study the problem by introducing a bilabeled generating tree technique. Anyway, recently Merlini and Sprugnoli found that the problem can be expressed by the rule (4) with $e_i = i+3$ and $a = 2$, namely:

$$[(0), \{(k) \rightsquigarrow (0)^{k+3}(1)^{k+2}(2)^{k+1}\dots(k+2)\}] \tag{8}$$

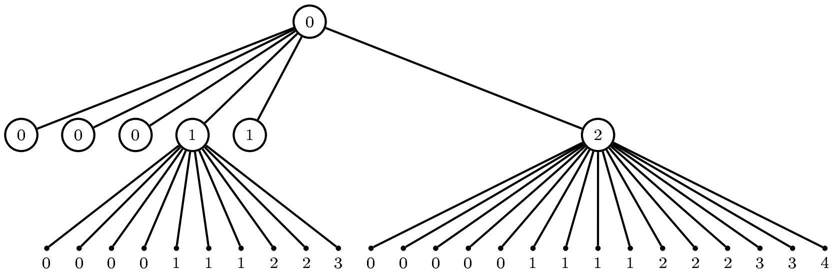

Figure 2. The partial generating tree for the specification (8)

In fact, if we don't care of the order of the balls thrown away, so that the configuration $(1,4)$, $(5,8)$, $(2,10)$ is considered to be the same as $(1,2)$, $(4,5)$, $(8,10)$, it can be proved that the number of $(4,2)$-sequences of length $2n$ in which the last-but-one element is $2n+k-1$ corresponds to the number of nodes with label k at level n in the generating tree of Figure 2 (for example, the possible sequences of length 2 are $(1,2)$, $(1,3)$, $(1,4)$, $(2,3)$, $(2,4)$ and $(3,4)$). □

Example 4. *Printers.*

In [9] the authors present a combinatorial model for studying the characteristics of job scheduling in a slow device, for example a printer in a local network. The policy usually adopted by spooling systems is called *First Come First Served* (FCFS) and can be realized by queuing the processes according to their arrival time and by using a FIFO algorithm. A job (printing a file) consists in a finite number of *actions* (printing-out a single page). Each action takes constant time to be performed (a *time slot*). If we fix n time slots, and suppose that at the end of the period the queue becomes empty, while it was never empty before, the successive states of the jobs queue can be

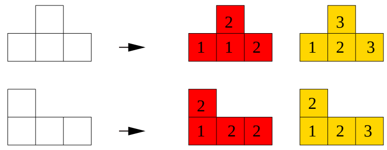

FIGURE 3. The schedules corresponding to two particular 1-histograms.

described by a combinatorial structure called *labeled 1-histograms*. A 1-*histogram* of length n is a histogram whose last column only contains 1 cell and, whenever a column is composed by k cells, then the next column contains at least $k-1$ cells. It is at all obvious that a 1-histogram corresponds to a path in the generating tree produced by the specification $[(1), (k) \rightsquigarrow (1) \dots (k+1)]$. A *labeled* 1-*histograms* of length n is a 1-histogram in which we label each cell according to some rules (see [9] for the details). Figure 3 illustrates the possible schedules for two particular 1-histograms of length 3: the first one, for example, corresponds to i) a first job which consists in printing two pages and a second job, which starts at time slot 2, and corresponds to printing a page at time slot 3, and ii) three different jobs which consists in printing a single page, the first at time slot 1, the second at time slot 2 and the third at time slot 3, after queuing at time slot 2. It can be proved that the number of schedules of length n with k jobs request at the first time slot corresponds to the number of nodes at level n having label $k+1$ in the generating tree with specification:

$$[(1), \{(k) \rightsquigarrow (1)^2 \dots (k)^2 (k+1)\}] .$$

This gives that the number S_n of possible schedules corresponds to the n^{th} small Schröder number, that is, the generating function for S_n is $(1-3z-\sqrt{1-6z+z^2})/(4z)$.

□

Acknowledgements. This work was partially supported by the Future and Emerging Technologies programme of the EU under contract number IST-1999-14186 (ALCOM-FT), by the INRIA postdoctoral programme and by the Max-Planck Institut. The first author also benefited of pasta and of an invitation in Florence in June and November 2001. Finally, the authors are grateful to an anonymous referee for her constructive comments.

REFERENCES

[1] Cyril Banderier. Limit laws for basic parameters of lattice paths with unbounded jumps. *Submitted*, 2002.
[2] Cyril Banderier, Mireille Bousquet-Mélou, Alain Denise, Philippe Flajolet, Danièle Gardy, and Dominique Gouyou-Beauchamps. Generating functions for generating trees. *Discrete Mathematics*, 246(1-3):29–55, 2002.
[3] Cyril Banderier and Philippe Flajolet. Basic analytic combinatorics of directed lattice paths. *To appear in Theoretical Computer Science*, June 2001.
[4] Elena Barcucci, Alberto Del Lungo, Elisa Pergola, and Renzo Pinzani. From Motzkin to Catalan permutations. *Discrete Mathematics*, 217(1-3):33–49, 2000. Formal power series and algebraic combinatorics (Vienna, 1997).
[5] Mireille Bousquet-Mélou. Walks on the slit plane: other approaches. *Applied Math.*, 27(2-3):243–288, 2001.
[6] F. R. K. Chung, R. L. Graham, V. E. Hoggatt, and M. Kleiman. The number of Baxter permutations. *Journal of Combinatorial Theory, Series A*, 24:382–394, 1978.
[7] D. Merlini, D. G. Rogers, R. Sprugnoli, and M. C. Verri. On some alternative characterizations of Riordan arrays. *Canadian Journal of Mathematics*, 49(2):301–320, 1997.
[8] D. Merlini, D. G. Rogers, R. Sprugnoli, and M. C. Verri. The tennis ball problem, 2002. To appear in Journal of Combinatorial Theory-A.
[9] D. Merlini, R. Sprugnoli, and M. C. Verri. Waiting patterns for a printer. In *FUN with algorithms 2, E. Lodi, L. Pagli, N. Santoro, Editors*, Carleton Scientific, pages 183–198, 2001.
[10] D. Merlini and M. C. Verri. Generating trees and proper Riordan Arrays. *Discrete Mathematics*, 218:167–183, 2000.
[11] L. W. Shapiro, S. Getu, W.-J. Woan, and L. Woodson. The Riordan group. *Discrete Applied Mathematics*, 34:229–239, 1991.

E-mail address: `Cyril.Banderier,at,inria.fr`, `http://algo.inria.fr/banderier`

INRIA-Rocquencourt, 78150 Le Chesnay (France) & Max-Planck-Institut, 66123 Saarbrücken (Germany)

E-mail address: `merlini@dsi.unif.it`, `http://www.dsi.unifi.it/~merlini/`

Dipartimento di Sistemi e Informatica, Universit degli Studi di Firenze, Via Lombroso 6/17, 50134 Firenze (Italy)