Fundamenta Informaticae 117 (2012) 199–213
DOI 10.3233/FI-2012-696
IOS Press

A Representation Theorem for Holonomic Sequences Based on Counting Lattice Paths

Tomer Kotek* **and Johann A. Makowsky**† ‡
Department of Computer Science
Technion–Israel Institute of Technology
32000 Haifa, Israel
{*tkotek,janos*}*@cs.technion.ac.il*

Abstract. Using a theorem of N. Chomsky and M. Schützenberger one can characterize sequences of integers which satisfy linear recurrence relations with constant coefficients (C-finite sequences) as differences of two sequences counting words in regular languages. We prove an analog for P-recursive (holonomic) sequences in terms of counting certain lattice paths.

1. Introduction

A sequence $a(n)$ of integers is C-finite if there exists $n_0 \in \mathbb{N}$ such that $a(n)$ satisfies for every $n > n_0$ a linear recurrence relation with constant coefficients of the form

$$a(n) = \sum_{j=1}^{r} c_j a(n-j)$$

where $c_1, \ldots, c_r$ are integers[1] and $n_0 \geq r$. A well-known theorem by Chomsky and Schützenberger states:

*Partially supported by a grant of the Graduate School of the Technion–Israel Institute of Technology
†Partially supported by a grant of the Fund for Promotion of Research of the Technion–Israel Institute of Technology and grant ISF 1392/07 of the Israel Science Foundation (2007-2011)
‡Address for correspondence: Department of Computer Science, Technion–Israel Institute of Technology, 32000 Haifa, Israel
[1]We may assume without loss of generality that at least one of the c_i is non-zero.

Theorem 1. (N. Chomsky and P. Schützenberger [5])
For every regular language L, the sequence $a_L(n)$, which counts words of length n in L, is C-finite.

In [13] the authors rediscovered[2]:

Theorem 2. Every C-finite sequence is a difference $a_{L_1}(n) - a_{L_2}(n)$ for two regular languages L_1 and L_2. Alternatively, $a_{L_2}(n)$ can be replaced by a geometric sequence c^n.

Definition 3. Let $a(n)$ be a sequence of integers. We say that $a(n)$ is *holonomic* or *P-recursive* if there exist $n_0 \in \mathbb{N}$ such that $a(n)$ satisfies for every $n > n_0$ a linear recurrence relation with polynomial coefficients of the form

$$p_0(n) \cdot a(n) = \sum_{j=1}^{r} p_j(n) a(n-j) \tag{1}$$

where $p_0(x), \ldots, p_r(x) \in \mathbb{Z}[x]$ are polynomials over $\mathbb{Z}$, $p_0(n) \neq 0$ for every $n > n_0$ and it holds that $n_0 \geq r$. Additionally, if $p_0(x)$ is identically 1, $a(n)$ is said to be *simply P-recursive* or *SP-recursive*.

The purpose of this paper is to prove an analog to Theorem 2 for holonomic sequences. The precise statement of the Theorem is given in Section 2.3. In the sequel we refer to P-recursive and SP-recursive sequences, rather than to holonomic sequences. Holonomic sequences are known in the literature by yet another name, as they are the coefficients of D-finite generating functions.

2. Motivation and statement of the main theorem

2.1. Properties of P-recursive sequences

Lemma 4.

1. Both the class of SP-recursive sequences and the class of P-recursive sequences are closed under finite addition and multiplication.
2. If $a(n)$ is P-recursive, then there exists $c \in \mathbb{N}$ such that $a(n) \leq (n!)^c$.

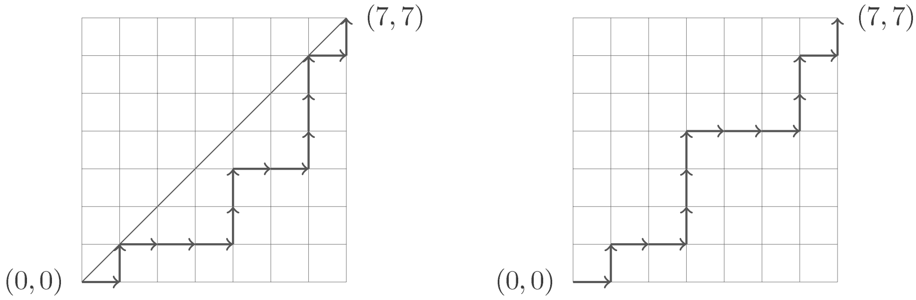
The upper bound in Lemma 4(2) can be reached easily, for every c, by some P-recursive sequence [9].

Remark 5. Lemma 4 also holds for P-recursive sequences over $\mathbb{Q}$, i.e. sequences of rational numbers which satisfy Equation (1). Notice that it makes no difference whether $p_0(x), \ldots, p_r(x)$ are polynomials over $\mathbb{Z}$ or $\mathbb{Q}$, since in the latter case one can multiply all the polynomials by their common denominator and obtain polynomials over $\mathbb{Z}$. We will use the closure properties of rational P-recursive sequences in the proof of our main theorem.

2.2. Counting sequences of lattice paths and their recurrence relations

The Catalan numbers C_{n+1} count the number of Dyck paths in the grid $[0,n]^2 = \{0,\ldots,n\} \times \{0,\ldots,n\}$, paths from $(0,0)$ to (n,n) which consist of steps $\{\uparrow, \rightarrow\}$ and do not go above the diagonal line $y = x$. Such a path is depicted in Figure 1(a).

[2] It turned out that this is implicitly already contained in [18].



(a) Catalan numbers (b) The central binomial coefficient

Figure 1. Paths counted by the Catalan numbers and the central binomial coefficient $\binom{2n}{n}$.

The Catalan numbers are P-recursive with the following P-recurrence:

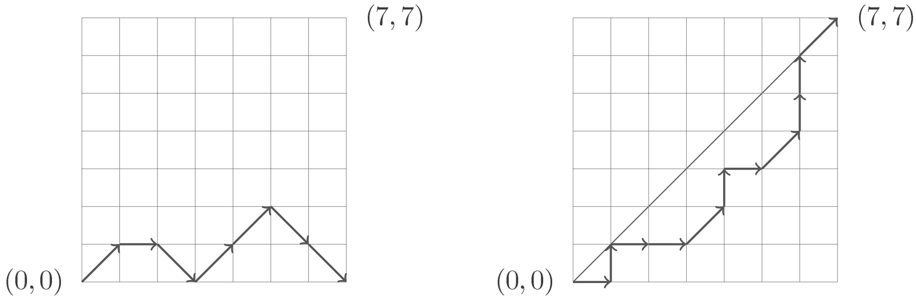
$$(n+2)C_{n+1} = 2(2n+1)C_n.$$

The central binomial coefficients $\binom{2n}{n}$ count paths similar to Dyck paths, except that they are no longer restricted to stay below the diagonal. Figure 1(b) shows such a path. They satisfy the following P-recurrence:

$$n \cdot \binom{2n}{n} = 2 \cdot (2n-1) \cdot \binom{2(n-1)}{n-1}.$$

Both C_{n+1} and $\binom{2n}{n}$ due to their modular properties. See [6, Theorems 2.1 and 4.3] for modular properties of C_{n+1} and $\binom{2n}{n}$, and [8, Proposition 2.2(iii)] for a discussion of modular properties of SP-recursive sequences.

Some other P-recursive sequences have natural interpretations as counting various lattice paths, such as Motzkin numbers and Schröder numbers. Motzkin numbers count paths in $[0,n]^2$ from $(0,0)$ to $(n,0)$ with steps $\nearrow, \rightarrow, \searrow$. Schröder numbers count paths in $[0,n]^2$ from $(0,0)$ to (n,n) with steps $\rightarrow, \uparrow, \nearrow$ which do not go above the diagonal. See examples in Figure 2.



(a) Motzkin numbers (b) Schröder numbers

Figure 2. Paths counted by Motzkin and Schröder numbers.

The factorial $n!$ is one of the simplest SP-recursive sequences with the recurrence $n! = n \cdot (n-1)!$. The factorial can be interpreted as the counting sequence of a class of lattice paths on the grid $[0,n]^2$.

Let $d_1(n)$ be the number of paths from $(0,0)$ to (n,n) with step set $\{\uparrow,\downarrow,\rightarrow\}$ which do not cross the diagonal line and are self-avoiding. An example of such a path is given in Figure 3(a).

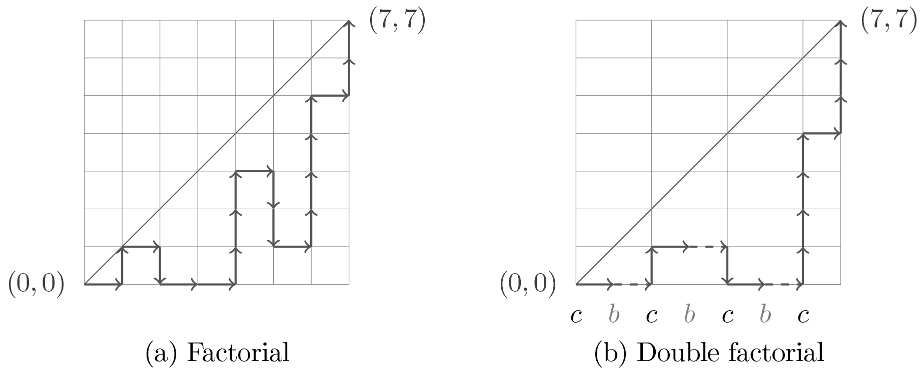
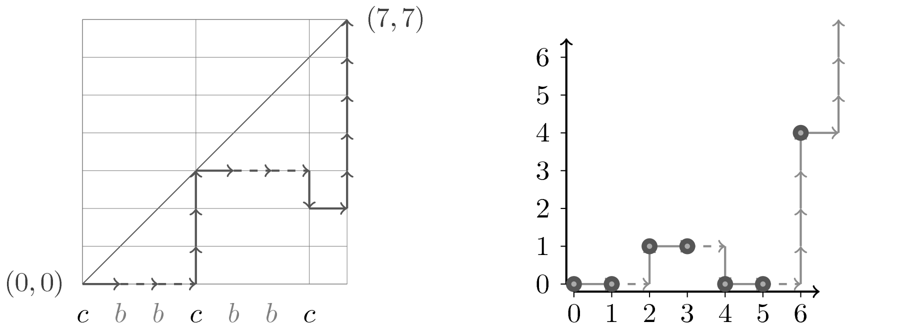


Figure 3. Paths count by the factorial and double factorial.

Consider a path in $[0,n]^2$ which is counted by $d_1(n)$. At any point (i,j), the only step which the path can take and which changes the column j is $\rightarrow$. Since the path starts at $(0,0)$ and ends at (n,n), it must take exactly one $\rightarrow$ step at each of the columns $0,\ldots,n-1$. Since the paths enumerated by $d_1(n)$ are self-avoiding, one can see that the paths can be determined uniquely by the points on the grid where $\rightarrow$ occurred.

Let $(0,0),(i_1,1),\ldots,(i_{n-1},n-1)$ be the points in which $\rightarrow$ occurs. We may think of these points as a function $f:\{0,\ldots,n-1\}\rightarrow\{0,\ldots,n-1\}$ which is defined as $f(j)=i_j$. The function f assigns each column j with the row i_j in which $\rightarrow$ occurs. It holds that $f(j)\leq j$ for every j. On the other hand, every function $f:\{0,\ldots,n-1\}\rightarrow\{0,\ldots,n-1\}$ for which $f(j)\leq j$ for every j is obtained by some legal lattice path of the paths counted by $d_1(n)$. Therefore, $d_1(n)$ is the number of functions $f:\{0,\ldots,n-1\}\rightarrow\{0,\ldots,n-1\}$ for which $f(j)\leq j$ for every j. Consequently, $d_1(n)=n!$.

Simple variants of $n!$ can be interpreted similarly. The double factorial $n!!$ is defined to be $n!!=n(n-2)(n-4)\cdots$. An interpretation in terms of lattice paths for $n!!$ can be obtained by augmenting the interpretation for $n!$ as follows. We label each column $j\in[0,n-1]=\{0,\ldots,n-1\}$ of the grid $[0,n]^2$ by a letter $w_{j+1}\in\{b,c\}$. The labeling of the columns induces a word $w=w_1\cdots w_n$ over the alphabet $\{b,c\}$ of length n. The last column is left unlabeled. Let $d_2(n,w)$ be the number of paths from $(0,0)$ to (n,n) with step set $\{\uparrow,\downarrow,\rightarrow\}$ which do not cross the diagonal, are self-avoiding, and in addition, the steps $\uparrow$ and $\downarrow$ may occur only in columns labeled with c. Let $w_{n!!}$ be a word of length n which belongs to the language $L_{!!}=L\left((bc)^\star+c(bc)^\star\right)$. The word $w_{n!!}$ is of form either $bcbc\cdots bc$ or $cbcbc\cdots bc$. The paths enumerated by $d_2(n,w_{n!!})$ can be distinguished uniquely by the points on the grid where $\rightarrow$ occurs in columns $j\in[0,n-1]$ which have the same parity as $n-1$, or equivalently, those columns $j\in[0,n-1]$ labeled c. Notice that on columns labeled b the only possible step is $\rightarrow$. The column n must consist of $\uparrow$ steps only in order to reach the upper right point of the grid (n,n). Let $d_2(n)$ be the sum over $w_{n!!}\in L_{!!}$ of length n of $d_2(n,w_{n!!})$. Then $d_2(n)=n!!$. Figure 3(b) illustrates a path counted by $d_2(n)$. The dashed right arrows $\dashrightarrow$ are used to indicate that no other step starting at the same column is possible, since the column is labeled b. As in the case of the factorial, the double factorial can be considered as counting functions, see Figure 4(d). The triple factorial $n!!!$ can be treated similarly, see Figure 4(c).



(c) Triple factorial

(d) The path in (b) as a function

Figure 4. Paths counted by the double and triple factorials.

Another simple variant of $n!$ which is SP-recursive is $n!^2$, which satisfies the recurrence $n!^2 = n^2 \cdot (n-1)!^2$. The sequence $n!^2$ can be interpreted as counting pairs of those paths counted by $d_1(n)$.

2.3. Main result

In this subsection we present our main theorem. We begin by defining the type of lattice paths which we will use to characterize P-recursive sequences. Let Σ be an alphabet.

Definition 6. Let $w = w_1 w_2 \cdots w_n$ be a word of length n over Σ and let $\sigma \in \Sigma$. A (w, σ)-*path* p is a lattice path in the grid $[0, n]^2$ satisfying the following conditions:

1. p starts at $(0, 0)$ and ends at (n, n),
2. every point (i, j) in p satisfies $i \leq j$, i.e. p does not cross the diagonal,
3. p consists of steps from $\{\rightarrow, \uparrow, \downarrow\}$,
4. p is self-avoiding, and
5. $w_{j+1} \neq \sigma$ implies the only step that occurs in column j is $\rightarrow$.

Definition 7. Let L be a regular language over Σ, k be a natural number, and $\bar{s} = (s_1, \ldots, s_k)$ be a tuple of Σ letters. The sequence $m_{L,\bar{s}}(n)$ counts the number of tuples $(w, p_1, \ldots, p_k)$ where $w \in L$, $|w| = n$ and each p_i is a (w, s_i)-path.

We can apply Definition 7 to $n!$, $n!!$ and $(n!)^2$. Let $L_c = L(c^\star)$. Then

$$n! = m_{L_c,c}(n), \;\; n!! = m_{L_{!!},c}(n), \;\; (n!)^2 = m_{L_c,(c,c)}(n)\,.$$

We are now ready to state our main theorem.

Theorem 8. Let $a(n)$ be a sequence of integers.

1. $a(n)$ is SP-recursive iff there exist regular languages L_1 and L_2 and a tuple $\bar{s}$ of letters of Σ such that $a(n) = m_{L_1,\bar{s}}(n) - m_{L_2,\bar{s}}(n)$.

2. $a(n)$ is P-recursive with leading polynomial $p_0(n)$ iff there exist regular languages L_1 and L_2 and a tuple $\bar{s}$ of letters of Σ such that

$$a(n) = \frac{m_{L_1,\bar{s}}(n) - m_{L_2,\bar{s}}(n)}{\prod_{j=1}^{n} p_0(j)} .$$

The proof uses elementary concepts from the theory of regular languages, see for example [12].

3. Proof of main theorem

In this section we present the proof to the main theorem, Theorem 8. The proof follows from the following two lemmas, which we will prove in Subsections 3.1 and 3.2 respectively.

Lemma 9. Let $a(n)$ be P-recursive with leading polynomial $p_0(n)$. There exist regular languages L_1 and L_2 and a tuple $\bar{s}$ of Σ letters such that

$$a(n) = \frac{m_{L_1,\bar{s}}(n) - m_{L_2,\bar{s}}(n)}{\prod_{j=1}^{n} p_0(j)} .$$

Lemma 10. Let $a(n) = m_{L,\bar{s}}(n)$ with L a regular language and $\bar{s}$ a tuple of letters of Σ. The sequence $a(n)$ is SP-recursive.

Using Lemma 9, Lemma 10 and the closure properties of P-recursive sequences given in Lemma 4 and Remark 5, the proof of Theorem 8 follows.

Proof:
[Proof of Theorem 8] Lemma 9 gives one direction of the proof in both cases, since $p_0(x) = 1$ if $a(n)$ is SP-recursive. For the other direction, Lemma 10 and the closure of P-recursive sequences under addition and multiplication imply $m_{L_1,\bar{s}}(n) - m_{L_2,\bar{s}}(n)$ is SP-recursive, which gives the second direction in the case of SP-recursive $a(n)$. The sequence of rational numbers $e(n) = \frac{1}{\prod_{j=1}^{n} p_0(j)}$ satisfies the P-recurrence $p_0(n) \cdot e(n) = e(n-1)$. By the closure of rational P-recursive sequences to multiplication, $(m_{L_1,\bar{s}}(n) - m_{L_2,\bar{s}}(n)) \cdot e(n)$ is P-recursive. □

For the proof of the Lemmas we look at the (w, σ)-paths as functions.

Definition 11. Let $w = w_1 w_2 \cdots w_n$ be a word of length n. Let $B_{w,\sigma}$ be the set of all functions $f : [0, n-1] \to [0, n-1]$ which satisfy the following conditions:

- $f(j) \leq j$ for every $j \in [0, n-1]$, and
- $f(j-1) \neq f(j)$ implies $w_{j+1} = \sigma$.

Proposition 12. For every w over Σ of length n and $\sigma \in \Sigma$, $|B_{w,\sigma}| = \prod_{j: w_j = \sigma} j$.

By $\prod_{j:w_j=\sigma} j$ we mean a product of all the indices j of the word w for which $w_j = \sigma$. Proposition 12 follows from the definition of $B_{w,\sigma}$.

Proposition 13. Let L be a regular language and let $\bar{s} = (s_1, \ldots, s_k) \in \Sigma^k$. It holds that

$$\sum_{w \in L:|w|=n} |B_{w,s_1}| \cdots |B_{w,s_k}| = m_{L,\bar{s}}(n)\,.$$

Proof:
For every $\sigma \in \Sigma$ and $w \in L$, let $B'_{w,\sigma}$ be the set of (w, σ)-paths. We will construct a bijection between $B_{w,\sigma}$ and $B'_{w,\sigma}$. The proposition then follows since by definition of $m_{L,\bar{s}}(n)$,

$$m_{L,\bar{s}}(n) = \sum_{w \in L:|w|=n} |B'_{w,s_1}| \cdots |B'_{w,s_k}|\,.$$

Let $\sigma \in \Sigma$ and $w \in L$. Let p be a (w, σ)-path. The path p contains exactly $n \rightarrow$ steps. There is exactly one $\rightarrow$ step in each column $j \in [0, n-1]$ of the grid $[0, n]^2$. Let $(i_0, 0), \ldots, (i_{n-1}, n-1)$ be the positions of occurrences of $\rightarrow$ in p. For each $j \in [0, n-2]$, the steps between (i_j, j) and $(i_{j+1}, j+1)$ must all be $\uparrow$ steps or all $\downarrow$ steps, since the path is self-avoiding. All the steps in the last column n must be $\uparrow$ steps in order to reach (n, n). Hence, the path p is determined uniquely by $(i_0, 0), \ldots, (i_{n-1}, n-1)$. It holds that $i_j \leq j$, since the path does not cross the diagonal. Let $f_p : [0, n-1] \rightarrow [0, n-1]$ be given as $f_p(j) = i_j$. The function f_p belongs to $B_{w,\sigma}$. Let $h : B'_{w,\sigma} \rightarrow B_{w,\sigma}$ be given by $h(p) = f_p$. The function h is injective, since p is determined uniquely by $(i_0, 0), \ldots, (i_{n-1}, n-1)$. It is not hard to see that h is also surjective, since it holds, for every $f \in B_{w,\sigma}$, that for every $j \in [0, n-1]$, $f(j) \leq j$. $\square$

3.1. Proof of Lemma 9

Let $n_0 \in \mathbb{N}$. Let $a(n)$ be a P-recursive sequence of integers satisfying, for $n > n_0$,

$$p_0(n) \cdot a(n) = \sum_{i=1}^{r} p_i(n) a(n-i) \tag{2}$$

where $p_0(x), \ldots, p_r(x) \in \mathbb{Z}[x]$ are polynomials over $\mathbb{Z}$, $p_0(n) \neq 0$ for every $n \in \mathbb{N}$ and it holds that $n_0 \geq r$. We assume for the sake of simplicity that $r = n_0$. The initial conditions of $a(n)$ are then $a(1), \ldots, a(r)$.

Since the p_i are polynomials over $\mathbb{Z}$, we may write the recurrence as follows:

$$a(n) = \frac{1}{p_0(n)} \left(\sum_{i=1}^{r_1} n^{\alpha_i} a(n - l_i) - \sum_{i=1}^{r_2} n^{\beta_i} a(n - g_i) \right)\,, \tag{3}$$

where $\alpha_1, \ldots, \alpha_{r_1}, \beta_1, \ldots, \beta_{r_2} \in \mathbb{N}$ and $l_1, \ldots, l_{r_1}, g_1, \ldots, g_{r_2} \in [r] = \{1, \ldots, r\}$. Let

$$sg(k) = \begin{cases} +, & a(k) \geq 0 \\ -, & a(k) < 0 \end{cases}\,.$$

Let

$$\Gamma_1 = \left\{\rho_1^-, \ldots, \rho_{r_1}^-, \rho_1^+, \ldots, \rho_{r_2}^+, c\right\}$$

and

$$\Gamma_2 = \left\{\pi_{k,t}^{sg(k)} \mid k \in [r], 1 \leq t \leq |a(k)|\right\} .$$

The alphabets Γ_1 and Γ_2 consist of previously unused letters. A word $w = w_1 w_2 \cdots w_n$ over alphabet $\Gamma = \Gamma_1 \cup \Gamma_2$ is called a *recurrence path* with respect to Equation (3) if the following conditions holds:

Cond 1 $w_1 \cdots w_r = c^{k-1}\pi_{k,t}^{sg(k)}c^{r-k}$, where $k \in [r]$ and $1 \leq t \leq |a(k)|$,

Cond 2 $w_{r+1} \cdots w_n$ belongs to $\{\rho_1^-, \ldots, \rho_{r_1}^-, \rho_1^+, \ldots, \rho_{r_2}^+, c\}^\star$,

Cond 3 $w_n = \rho_i^-$ or $w_n = \rho_i^+$ for some i,

Cond 4 for every m, if $w_m = \rho_i^-$ for some i then $w_{m-(l_i-1)} \cdots w_{m-1} = c^{l_i-1}$ and $w_{m-l_i} \neq c$, and

Cond 5 for every m, if $w_m = \rho_i^+$ for some i then $w_{n-(g_i-1)} \cdots w_{m-1} = c^{g_i-1}$ and $w_{m-g_i} \neq c$.

If $l_i = 1$ in Cond 4 then the requirement that $w_{m-(l_i-1)} \cdots w_m - 1 = c^{l_i-1}$ is void, and similarly for Cond 5. A word w which is a recurrence path with respect to Equation (3) describes the recursive choices in one branch of the recurrence tree, from the root $a(n)$ to a leaf, which represents an initial condition. The letters ρ_i^+ respectively ρ_i^- are used to indicate the choice of the i-th element of the first respectively second sum as the next recursive step. The letter c is a placeholder for the indices that are skipped over by the recurrence. The initial conditions are dealt with similarly using the $\pi_{k,t}^{sg(k)}$ letters. We consider every initial condition $a(k)$ as a sum of 1's or -1's, depending on whether $a(k)$ is positive or negative. To indicate that the initial condition $a(k)$ is chosen we set $w_k = \pi_{k,t}^{sg(k)}$ for some $t \leq |a(k)|$. Notice there are exactly $|a(k)|$ many ways to choose a word of the form $c^{k-1}\pi_{k,t}c^{r-k}$, one for each $t \leq |a(k)|$.

We may now rewrite Equation (3) as a sum of the values of the leafs of the recurrence tree. In other words, we rewrite Equation (3) as the sum over all words w of length n which are recurrence paths with respect to Equation (3). Let L_{rec} be the language of recurrence paths. Let $sign(w) \in \{1, -1\}$ be 1 if the parity of the number of letters in w of the form ρ_i^- or $\pi_{k,t}^-$ is even, and -1 otherwise. We get

$$a(n) = \frac{\displaystyle\sum_{\substack{w \in L_{rec}: \\ |w|=n}} \prod_{j:w_j=\rho_1^+} j^{\alpha_1} \cdots \prod_{j:w_j=\rho_{r_1}^+} j^{\alpha_{r_1}} \prod_{j:w_j=\rho_1^-} j^{\beta_1} \cdots \prod_{j:w_j=\rho_{r_2}^-} j^{\beta_{r_2}} \cdot sign(w)}{\displaystyle\prod_{j=1}^{n} p_0(j)} ,$$

where $\prod_{j:w_j=\sigma} j^m$ stands for the product of terms j^m over all indices $j \in [n]$ in w for which $w_j = \sigma$. By Proposition 12, $\prod_{j:w_j=\sigma} j$ is the number of functions in $B_{w,\sigma}$. Thus,

$$a(n) = \frac{\displaystyle\sum_{w \in L_{rec}:|w|=n} |B_{w,\rho_1^+}|^{\alpha_1} \cdots |B_{w,\rho_{r_1}^+}|^{\alpha_{r_1}} \cdot |B_{w,\rho_1^-}|^{\beta_1} \cdots |B_{w,\rho_{r_2}^-}|^{\beta_{r_2}} \cdot sign(w)}{\displaystyle\prod_{j=1}^{n} p_0(j)} .$$

It is not hard to see that L_{rec} is a regular language. Let

$$e_i = \sum_{\substack{1\le k\le r \\ t\le |a(k)|}} c^{k-1}\pi_{k,t}^{sg(k)}c^{r-k}\ , \text{ and}$$

$$e_r = \sum_{i=1}^{r_1} c^{l_i-1}\rho_i^- + \sum_{i=1}^{r_2} c^{g_i-1}\rho_i^+ \ .$$

L_{rec} is given by the regular expression $e_i \cdot e_r \cdot e_r{}^\star$. Let L_{even} and L_{odd} be the languages which consist of all words over Γ such that the number of letters of the form ρ_i^- or $\pi_{k,t}^-$ in the word is even respectively odd. L_{even} and L_{odd} are regular languages. By the closure of regular languages with respect to intersection, $L_{rec} \cap L_{even}$ and $L_{rec} \cap L_{odd}$ are regular as well, and by Proposition 13,

$$a(n) = \frac{m_{L_{rec}\cap L_{even},\bar{s}}(n) - m_{L_{rec}\cap L_{odd},\bar{s}}(n)}{\prod_{j=1}^n p_0(j)}$$

where $\bar{s} = \left(\rho_1^+,\dots,\rho_{r_1}^+,\rho_1^-,\dots,\rho_{r_2}^-\right)$.

3.2. Proof of Lemma 10

Let L be a regular language, let $\bar{s} = (s_1,\dots,s_k)$ be a tuple of Σ letters and let $a(n) = m_{L,\bar{s}}(n)$. The language L is accepted by a deterministic finite state automaton A with state set Q, an initial state q_0, a set of final states F and a state-transition function $\delta : Q\times\Sigma \to Q$. By abuse of notation we write $\delta(q,w)$ for the state reached by A on input w. For each $q\in Q$, let L_q be the language of the automaton A_q which is obtained from A by replacing the set F of final states with $\{q\}$. We will show that each $m_{L_q,\bar{s}}(n)$ is SP-recursive for every q. The lemma follows from this, since $m_{L,\bar{s}}(n) = \sum_{q\in F} m_{L_q,\bar{s}}(n)$, and the set of SP-recursive sequences is closed under finite addition.

We start with $m_{L_q,s}(n)$, where s is not a tuple, but rather a single letter, or in other words, $k=1$. By Proposition 13, $m_{L_q,s}(n)$ counts pairs (w,f) where $w\in L_q$ and $f\in B_{w,s}$. A word w of length n has a unique decomposition into a prefix of length $n-1$ and a suffix of length 1. Let $w = u\tau$, where $\tau\in\Sigma$ and u is a word of length $n-1$. The state reached by A_q on w, $\delta(q_0,w)$, depends only on τ and on the state reached by A_q on u, $\delta(q_0,u)$. More precisely, $\delta(q_0,w) = \delta(\delta(q_0,u),\sigma)$. Moreover, f has a unique restriction to $[0,n-2]$, $f' : [0,n-2] \to [0,n-2]$. The function f is defined uniquely by f' and the value $f(n-1)$. On the other hand, f' can be extended to $[0,n-1]$ in either n ways or exactly one way, depending on whether $w_n = s$. So, it holds that

$$m_{L_q,s}(n) = \overbrace{\sum_{\tau\in\Sigma-\{s\}}\ \sum_{u\tau\in L_q:|u|=n-1} |B_{u.s}|}^{(\bigcirc)} + \overbrace{\sum_{u\in L_p:|u|=n-1} n|B_{u,s}|}^{(\triangle)}\ . \tag{4}$$

The part $(\bigcirc)$ in Equation (4) counts pairs (w,f) where $w = u\cdot\tau \in L_q$, $\tau \ne s$, and, therefore, $f(n-1) = f(n-2)$. The part $(\triangle)$ in Equation (4) counts pairs (w,f) where $w = u\cdot s\in L_q$ and $f(n-1)\in\{0,\dots,n-1\}$.

Let $\mathbb{I}^{s}_{\delta(p,\tau)=q}$ be defined as follows,

$$\mathbb{I}^{s}_{\delta(p,\tau)=q} = \begin{cases} 0, & \delta(p,\tau) \neq q \\ n, & \delta(p,\tau) = q, \tau = s \\ 1, & \delta(p,\tau) = q, \tau \neq s \end{cases} .$$

We can rewrite equation (4) as follows:

$$m_{L_q,s}(n) = \sum_{p \in Q} \sum_{\tau \in \Sigma} \mathbb{I}^{s}_{\delta(p,\tau)=q} \cdot \left(\sum_{u \in L_p : |u| = n-1} |B_{u,s}| \right) .$$

Notice

$$m_{L_p,s}(n-1) = \sum_{u \in L_p : |u| = n-1} |B_{u,s}| ,$$

so

$$m_{L_q,s}(n) = \sum_{(p,\tau) \in Q \times \Sigma} \mathbb{I}^{s}_{\delta(p,\tau)=q} \cdot m_{L_p,s}(n-1) .$$

In a similar way we get for a tuple $\bar{s} = (s_1, \ldots, s_k)$,

$$m_{L_q,\bar{s}}(n) = \sum_{(p,\tau) \in Q \times \Sigma} \mathbb{I}^{\bar{s}}_{\delta(p,\tau)=q} \cdot m_{L_p,\bar{s}}(n-1) ,$$

where

$$\mathbb{I}^{\bar{s}}_{\delta(p,\tau)=q} = \begin{cases} 0, & \delta(p,\tau) \neq q \\ n^{|\{s_i = \tau : i \in [k]\}|}, & \delta(p,\tau) = q \end{cases} .$$

We get that $m_{L_q,\bar{s}}(n)$ satisfies a recurrence with polynomial coefficients in $m_{L_p,s}(n-1)$, $p \in Q$. To prove that $m_{L_q,\bar{s}}(n)$ is SP-recursive for every $q \in Q$, we first write the recurrences in matrix form. Let $\mathbf{v_n}$ be the column vector of length n given by $\mathbf{v_n} = \left(m_{L_q,\bar{s}}(n) : q \in Q \right)^{tr}$. Then there is a matrix M of size $|Q| \times |Q|$ of polynomials in n over $\mathbb{Z}$ such that $\mathbf{v_n} = M\mathbf{v_{n-1}}$. The matrix M is given by $M = (m_{q,p})$ where $m_{q,p} = \sum_{\tau \in \Sigma} \mathbb{I}^{s}_{\delta(p,\tau)=q}$. Let $\chi_M(\lambda)$ be the characteristic polynomial of M,

$$\chi_M(\lambda) = \det(\lambda \cdot \mathbf{1} - M) = \sum_{i=0}^{|Q|} c_i(n) \lambda^i .$$

Notice $c_i \in \mathbb{Z}[n]$ for every i, and $c_{|Q|} = 1$. The Cayley-Hamilton theorem for commutative rings states that, in the ring of matrices over $\mathbb{Z}[n]$, it holds that

$$\sum_{i=0}^{|Q|} c_i(n) M^i = 0 ,$$

or equivalently,

$$M^{|Q|+1} = \sum_{i=0}^{|Q|-1} -c_i(n) M^{i+1} .$$

Using that for every m, $\mathbf{v_m} = M^i \mathbf{v_{m-i}}$, we have

$$\mathbf{v_n} = M^{|Q|+1}\mathbf{v_{n-|Q|-1}} = \sum_{i=0}^{|Q|-1} -c_i(n) M^{i+1}\mathbf{v_{n-|Q|-1}} = \sum_{i=0}^{|Q|-1} -c_i(n)\mathbf{v_{n-|Q|+i}}\,. \tag{5}$$

Equation (5) gives $|Q|$ many SP-recurrences, one for each element $m_{L_q,\bar{s}}(n)$ of the vector $\mathbf{v_n}$.

4. Applications to counting permutations

Interesting classes of P-recursive functions arise from counting permutations. Gessel [10] and Noonan-Zeilberger [17] initiated the study of P-recursiveness and its relationship to patterns in permutations. M. Bona [2], and independently [14], proved the following theorem.

Theorem 14. The number $S_r(n)$ of permutations of length n containing exactly r subsequences of type 132 is a P-recursive function of n.

In [17] it is conjectured that for any given subsequence q, rather than just 132, and for any given r, the number of n-permutations containing exactly r subsequences of type q is a P-recursive function of n. However, later evidence has caused Zeilberger to change his mind [7] and conjecture that it is not P-recursive for $q = 1324$.

In this section we give two examples of interpretations of SP-recursive sequences which arise from counting permutations as counting lattice paths, and where the lattice paths interpretation is intuitive. The first example counts permutations with a fixed number of cycles. The purpose of this example is to show how counting permutations can be naturally interpreted in terms of counting (w, σ)-paths. The second example counts derangements. The purpose of this example is to illustrate how to derive from the explicitly given recurrence relation a lattice paths interpretation.

Unwinding the proof of Theorem 14 one can get an explicit recurrence relation for $S_r(n)$, and use our Theorem 8. In subsection 4.3 we give a simplifying assumption on the recurrence relation which allows us to give natural interpretations in terms of counting (w, σ)-paths, without using difference and division.

4.1. Counting permutations with a fixed number of cycles

The Stirling numbers of the first kind are denoted by $\left[{n \atop k}\right]$. They count the number of permutations of $[n]$ with exactly k cycles. The Stirling numbers of the first kind satisfy the recurrence

$$\left[{n+1 \atop k}\right] = n \cdot \left[{n \atop k}\right] + \left[{n \atop k-1}\right]\,. \tag{6}$$

Using the Cayley-Hamilton theorem one can show that $\left[{n+1 \atop k}\right]$ is SP-recursive.

$\left[{n+1 \atop k}\right]$ can be interpreted naturally as counting (w, σ)-paths. Every arrangement of $[0, n]$ into k cycles can be obtained uniquely by the following process, which is used in [11] to prove Equation (6). The process augments permutations of $[0, t]$ to create permutations of $[0, t+1]$ by adding the element $t+1$. We start with the only permutation of $\{0\}$. Every subsequent element $1, 2, \ldots, n$ either forms a new cycle or is added to an existing cycle. We require that exactly $k-1$ elements form new cycles during the process, which ends with the element n. This implies the permutation obtained at the end of the process has k cycles, since we begin the process with a permutation which consists of a single cycle.

Let L_k be the language over alphabet $\Sigma = \{A, B\}$ which consists of all words with exactly $k-1$ occurrences of the letter A. The letter A is used to indicate which elements form a new cycle during the process. Equivalently, the A's in a word w of L_k indicate which elements of $[n]$ are minimal in their cycles in the resulting permutation. In addition, 0 is also minimal in its cycle. The language L_k is a regular language given by the regular expression $(B^\star A)^{k-1} B^\star$.

We would like to show that $m_{L_k,B}(n) = \left[{n+1 \atop k}\right]$. There are j ways to add the element j to an existing cycle, by putting j to the right of any of the elements already in the permutation, $[0, j-1]$. The addition of j into one of the existing cycles corresponds to the letter B. The sequence $m_{L_k,B}(n)$ counts paths which allow any step of $\{\rightarrow, \uparrow, \downarrow\}$ in columns labeled B. In a column $j \in \{0, \ldots, n-1\}$ labeled B there are $j+1$ possibilities of movement for the lattice path before it leaves column j and does not return. On the other hand, there is only one way of forming a new cycle. The creation of a new cycle corresponds to the letter A, which occurs $k-1$ times in words of L_k. In a column j labeled A there is only one possibility of movement, since the only allowed step in this case is $\rightarrow$, which implies that upon arriving at the j-th column, the path leaves the column immediately (and does not return).

4.2. Counting derangements

The derangement numbers $D(n)$ count permutations of $[n]$ which have no fixed points. They satisfy the SP-recurrence

$$D(n) = (n-1) \cdot D(n-1) + (n-1) \cdot D(n-2) \tag{7}$$

with initial conditions $D(0) = 1$ and $D(1) = 0$. Let $D'(n)$ be the sequence such that $D(n+1) = D'(n)$, i.e. $D'(n)$ is the number of permutations of $[n+1]$ without fixed-points. The sequence $D'(n)$ satisfies the SP-recurrence

$$D'(n) = n \cdot D'(n-1) + n \cdot D'(n-2) \tag{8}$$

with initial conditions $D'(0) = D(1) = 0$ and $D'(1) = D(2) = 1$. Let the language L_{der} consist of all words $w = w_1 \cdots w_n$ over $\{a, b, c, d\}$ such that

(i) $w_1 = d$ and $w_2, \ldots, w_n \in \{a, b, c\}$, and

(ii) $w_i = c$ iff $w_{i+1} = b$.

One can show that

$$D'(n) = \sum_{w \in L_{der} : |w| = n} \; \prod_{i : w[i] \in \{a,b\}} i \tag{9}$$

where the summation is over words of length n in L_{der}.

We can think of Equation (9) as the sum over all paths in the recurrence tree of equation (8) from the root to a leaf $D'(1)$ (notice a path in the recurrence tree that ends in $D'(0) = 0$ has value 0). Such a path can be described by $1 = t_1 \leq \ldots \leq t_r = n$ such that for each i, $1 < i \leq r$, the difference of subsequent elements $t_i - t_{i-1}$ is either 1 or 2. The elements t_i in $[n]$ for which a recurrence step of the form $i \cdot D'(i-1)$ was chosen (i.e., those for which $t_i - t_{i-1} = 1$) are assigned letter a, whereas b is assigned to those elements t_i of $[n]$ which correspond to a choice of the form $i \cdot D'(i-2)$ (i.e., those t_i for which $t_i - t_{i-1} = 2$). We assign c to all the elements $i \in [n] - \{t_1, \ldots, t_r\}$, which are skipped by a recursive choice $j \cdot D'(j-2)$, where $j = i+1$. The letter d is assigned to the leafs $D'(1)$. By condition (i) the letter d is a place-holder for the leaf of the path in the recurrence tree, since it always occurs in w_1 and never for any other w_i. Condition (ii) requires that $i-1$ is skipped iff $i \cdot D'(i-2)$ is chosen for i. Notice this is a regular language, given by the regular expression

$$d(cb + a)^{\star} .$$

Equation (9) can be interpreted as counting the number of tuples (w, p_0, p_1) where $w \in L_{der}$ is of length $|w| = n$, p_0 is a (w, a)-path and p_1 is a (w, b)-path. That is, $D(n+1) = m_{L_{der},(a,b)}(n)$.

4.3. Interpreting the recurrence relation directly

By Theorem 8 every P-recursive sequence $a(n)$ with leading polynomial $p_0(n)$ can be represented as

$$a(n) = \frac{m_{L_1,\bar{s}}(n) - m_{L_2,\bar{s}}(n)}{\prod_{j=1}^{n} p_0(j)} .$$

Theorem 15. Let $a(n)$ be an SP-recursive sequence of integers

$$a(n) = \sum_{j=1}^{r} p_j(n) a(n-j)$$

such that for each $p_j(n)$, either $p_j(n)$ is a non-negative constant $p_j(n) = c \in \mathbb{N}$ or $p_j(n)$ tends to $+\infty$ as n tends to $+\infty$. Then $a(n)$ can be represented as

$$a(n+\delta) = m_{L,\bar{s}}(n) ,$$

where L is a regular language and $\delta \in \mathbb{N}$.

Proof:
For every $\delta \in \mathbb{N}$, $b_\delta(n) = a(n+\delta)$ satisfies an SP-recurrence

$$b_\delta(n) = \sum_{j=1}^{r} p_j(n+\delta) b_\delta(n-j) . \tag{10}$$

There exists $\delta \in \mathbb{N}$ such that for each j, $p_j(n+\delta)$ is a polynomial in n over the natural numbers $\mathbb{N}$. Therefore, there exist a regular language L and a tuple $\bar{s}$ of letters such that $a(n+\delta) = m_{L,\bar{s}}(n)$. □

5. Discussion and conclusions

Lattice walks on the grid $Grid_{n\times n}$ are given by a (usually) finite set of steps and are counted according to their lengths n. Lattice paths are self-avoiding walks on the grid $Grid_{n\times n}$ with given starting and end points $(0,0)$ and (n,k) for $k \leq n$. In most cases in the literature paths are counted as a function of n, which may or may not coincide with their length.

Counting lattice walks by length n cannot give a characterization of holonomic sequences, as their number is bounded by d^n, where d is the number of allowed steps. Nevertheless, studying the holonomicity of counting functions of lattice walks has attracted considerable attention in the literature. A unified treatment of counting lattices path is given in [1]. Holonomicity criteria were given in [3, 15], and non-holonomicity was studied in [4, 16, 15].

In order to get our characterization of holonomic sequences we modified the definition of lattice paths by introducing restrictions on when a certain step is allowed. For this purpose we marked the x-coordinates of the grid with letters from a finite alphabet, which emitted a labeling of the x-axis itself by a word. Our restrictions on the allowed steps were expressed in terms of words from a regular language. On the other hand, we only allowed the three types of steps $\{\uparrow, \downarrow, \rightarrow\}$. It remains open whether by a richer choice of steps the use of regular languages can be avoided.

Acknowledgments

We would like to thank Ira Gessel, Toufik Mansour, Daniel Marx and the anonymous referees for their helpful suggestions and comments.

References

[1] Banderier, C., Flajolet, P.: Basic analytic combinatorics of directed lattice paths, *Theoretical Computer Science*, **281 (1-2)**, 2002, 37–80.

[2] Bóna, M.: The number of permutations with exactly r 132-Subsequences is P-Recursive in the size!, *Advances in Applied Mathematics*, **18**, 1997, 510–522.

[3] Bousquet-Mélou, M.: Counting walks in the quarter plane, in: *Mathematics and Computer Science: Algorithms, trees, combinatorics and probabilities*, Trends in Mathematics, Birkhäuser, 2002, 49–67.

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

[5] Chomsky, N., Schützenberger, M.: The algebraic theory of context-free languages, *Comp. Prog. and Formal Systems*, North-Holland, Amsterdam, 1963.

[6] Deutsch, E., Sagan, B.: Congruences for Catalan and Motzkin numbers and related sequences, *Journal of Number Theory*, **117**(1), 2006, 191 – 215, ISSN 0022-314X.

[7] Elder, M., Vatter, V.: Problems and conjectures presented at the Third International Conference on Permutation Patterns, 2005, University of Florida.

[8] Fischer, E., Kotek, T., Makowsky, J.: Application of logic to combinatorial sequences and their recurrence relations, *Model Theoretic Methods in Finite Combinatorics*, 558, Amer. Math. Soc., Providence, RI, 2011, To appear.

[9] Gerhold, S.: On some non-holonomic sequences, *Electronic Journal of Combinatorics*, **11**, 2004, 1–7.

[10] Gessel, I.: Symmetric functions and P-recursiveness, *J. Comb. Theory, Ser. A*, **53**(2), 1990, 257–285.

[11] Graham, R. L., Knuth, D. E., Patashnik, O.: *Concrete mathematics: a foundation for computer science*, Addison-Wesley Pub., 1994.

[12] Hopcroft, J., Ullman, J.: *Introduction to automata theory, languages, and computation*, Addison Wesley, 1979.

[13] Kotek, T., Makowsky, J.: Definability of combinatorial functions and their linear recurrence relations, *Fields of Logic and Computation*, 2010.

[14] Mansour, T., Vainsthein, A.: Counting occurrences of 132 in a permutation, *Advances in Applied Mathematics*, **28**, 2002, 185–195.

[15] Mishna, M.: Classifying lattice walks restricted to the quarter plane, *Journal of Combinatorial Theory, Series A*, **116**, 2009, 460–477.

[16] Mishna, M., Rechnitzer, A.: Two non-holonomic lattice walks in the quarter plane, *Theor. Comput. Sci.*, **410**(38-40), 2009, 3616–3630.

[17] Noonan, J., Zeilberger, D.: The Goulden-Jackson cluster method: extensions, applications and implementations, *J. Differ. Equations Appl.*, **5 (4-5)**, 1999, 355–377.

[18] Salomaa, A., Soittola, M.: *Automata-theoretic aspects of formal power series*, Texts and Monographs in Computer Science, Springer, 1978.