Counting ordered patterns in words generated by morphisms

Sergey Kitaev[1], Toufik Mansour[2], and Patrice Séébold[3]

[1] Reykjavík University, Kringlan 1, 103 Reykjavík, ICELAND
sergey@ru.is
[2] Department of Mathematics, Haifa University, 31905 Haifa, ISRAEL
toufik@math.haifa.ac.il
[3] LIRMM, Univ. Montpellier 2, CNRS, 161 rue Ada, 34392 Montpellier, FRANCE
Patrice.Seebold@lirmm.fr

Abstract. We start a general study of counting the number of occurrences of ordered patterns in words generated by morphisms. We consider certain patterns with gaps (classical patterns) and that with no gaps (consecutive patterns). Occurrences of the patterns are known, in the literature, as rises, descents, (non-)inversions, squares and p-repetitions. We give recurrence formulas in the general case, then deducing exact formulas for particular families of morphisms. Many (classical or new) examples are given illustrating the techniques and showing their interest.

Keywords: morphisms, patterns, rises, descents, inversions, repetitions

1 Introduction

In algebraic combinatorics, an occurrence of a pattern p in a permutation π is a subsequence of π (of the same length as that of p) whose elements are in the same relative order as those in p. For example, the permutation $\pi = 536241$ contains an occurrence of the pattern $p = 2431$. Babson and Steingrímsson introduced generalized patterns where two adjacent elements of a pattern must also be adjacent in the permutation [2].

In combinatorics on words, an occurrence of a pattern p in a word u is a factor of u having the same shape as p. For example the word $u = abaabaaabab$ contains an occurrence of the pattern $p = \alpha\alpha\beta\alpha\alpha\beta$.

Burstein [4] realized a "mixing" of these two notions by considering ordered alphabets. An occurrence of an (ordered) pattern in a word is a factor or a subsequence having the same shape, and in which the relative order of the letters is the same as that in the pattern. In [5] one computed the number of occurrences of many of ordered patterns in the Peano words. In the present paper we start a general study of counting the number of occurrences of ordered patterns in words generated by morphisms.

2 Preliminaries

2.1 Definitions and notations

We refer to [7] for standard definitions in combinatorics on words.

Let n be a non-negative integer. The *incidence matrix* of f^n is the $k \times k$ matrix $M(f^n) = (m_{n,i,j})_{1 \le i,j \le k}$ where $m_{n,i,j}$ is the number of occurrences of the letter a_i in the word $f^n(a_j)$, i.e., $m_{n,i,j} = |f^n(a_j)|_{a_i}$.

Property 1 *For every $n \in \mathbb{N}$, $M(f)^n = M(f^n)$.*

2.2 Ordered patterns

Let A be a totally ordered alphabet and let $\aleph$ be the ordered alphabet whose letters are the first n positive integers in the usual order (thus $\aleph = \{1, 2, \ldots, n\}$).

An *ordered pattern* is any word over $\aleph \cup \{\#\}$, $\# \notin \aleph$, without two consecutive $\#$. If a pattern contains at least one $\#$, not at the very beginning or at the very end, it is an *ordered pattern with gaps*; otherwise it is an *ordered pattern with no gaps*. Moreover, in this paper the four ordered patterns u, $\#u$, $u\#$, and $\#u\#$ are considered to be the same (but of course $u\#u$ is not the same pattern as uu). In particular, if x is a word over $\aleph$, we will write $(x\#)^\ell$ or $(\#x)^\ell$ to represent the ordered pattern $x\#x\# \cdots \#x$ containing l occurrences of the word x.

A word v over A *contains an occurrence of the ordered pattern*

$$u = u_1 \# u_2 \# \cdots \# u_n,$$

$u_i \in \aleph^+$ and $n \ge 1$, if $v = w_0 v_1 w_1 v_2 w_2 \cdots w_{n-1} v_n w_n$ and there exists a literal morphism $f : \aleph^* \to A^*$ such that $f(u_i) = v_i$, $1 \le i \le n$, and if $x, y \in \aleph$, $x < y \Rightarrow f(x) < f(y)$. Thus the word v contains an occurrence of the ordered pattern u if v contains a subsequence v' which is equal to $f(u')$ where u' is obtained from u by deleting all the occurrences of $\#$, with the additional condition that two adjacent (not separated by $\#$) letters in u must be adjacent in v. The number of different occurrences of u as an ordered pattern in v is denoted by $|v|_u$.

Example. Let $A = \{a, b, c, d, e, f\}$ with $a < b < c < d < e < f$. The word $v = eafdbc$ contains one occurrence of the ordered pattern 2#31, namely the subsequence efd ($|e\,a\,f\,d\,b\,c|_{2\#31} = 1$). In v, the ordered pattern 2#3#1 occurs in three occurrences: efd, efb, and efc ($|e\,a\,f\,d\,b\,c|_{2\#3\#1} = 3$); the ordered pattern 231 does not occur in v ($|e\,a\,f\,d\,b\,c|_{231} = 0$).

3 Ordered patterns with gaps and morphisms

Let k be an integer ($k \ge 2$) and A the k-letter ordered alphabet $A = \{a_1 < a_2 < \cdots < a_k\}$. Let f be any morphism on A: for $1 \le i \le k$, $f(a_i) = a_{i_1} \ldots a_{i_{p_i}}$ with $p_i \ge 0$ ($p_i = 0$ if and only if $f(a_i) = \varepsilon$).

3.1 Inversions, non-inversions, and repetitions with gaps of f^n

In what follows we are interested in some particular forms of ordered patterns. In accordance with permutations theory, an *inversion* (resp. *non-inversion*) is an occurrence of the ordered pattern 2#1 (resp. 1#2). *Repetitions with gaps of one letter* are occurrences of the ordered patterns $(1\#)^p$, $p \geq 1$.

Inversions and non-inversions Let n be a non-negative integer.

The *vector* $RG(f^n)$ *of non-inversions* (resp. *vector* $DG(f^n)$ *of inversions*) of f^n is the k vector whose i-th entry is the number of occurrences of the ordered pattern 1#2 (resp. 2#1) in the word $f^n(a_i)$, i.e.,

$$RG(f^n) = (|f^n(a_i)|_{1\#2})_{1\leq i\leq k} \qquad DG(f^n) = (|f^n(a_i)|_{2\#1})_{1\leq i\leq k}.$$

Our goal is to obtain recurrence formulas giving the entries of $RG(f^{n+1})$ and $DG(f^{n+1})$. Since $f^{n+1} = f^n \circ f = f \circ f^n$, we have two different ways to compute $RG(f^{n+1})$ and $DG(f^{n+1})$.

Let ℓ be an integer, $1 \leq \ell \leq k$. Either $|f^{n+1}(a_\ell)|_{1\#2}$ (resp. $|f^{n+1}(a_\ell)|_{2\#1}$) will be obtained from the word $f(a_\ell)$ and the entries of $RG(f^n)$ (resp. $DG(f^n)$) (see 1. below), or it will be computed from the values of $RG(f)$ (resp. $DG(f)$) and $f^n(a_\ell)$ (see 2. below).

1. From $f^{n+1} = f^n \circ f$:

Since $f(a_\ell) = a_{\ell_1} \dots a_{\ell_{p_\ell}}$, the number of occurrences of the ordered pattern 1#2 in $f^{n+1}(a_\ell) = f^n(f(a_\ell)) = f^n(a_{\ell_1} \dots a_{\ell_{p_\ell}})$ is obtained by adding two values:

• The number of occurrences of the ordered pattern 1#2 in each $f^n(a_{\ell_i})$, $1 \leq i \leq p_\ell$. Since the ℓ-th column of the incidence matrix of f indicates which letters appear in $f(a_\ell)$ (and how many times), this number is obtained by multiplying the vector $RG(f^n)$ by the ℓ-th column of $M(f)$, i.e., it is equal to $\sum_{t=1}^{k} |f^n(a_t)|_{1\#2} \cdot m_{1,t,\ell}$.

• The number of occurrences of the ordered pattern 1#2 in each of the $f^n(a_{\ell_i} a_{\ell_j})$, $1 \leq i < j \leq p_\ell$, where the letter corresponding to 1 is in $f^n(a_{\ell_i})$ and the letter corresponding to 2 is in $f^n(a_{\ell_j})$. In what follows we will call such an occurrence of 1#2 in $f^n(a_{\ell_i} a_{\ell_j})$ an *external* occurrence of the ordered pattern 1#2 in $f^n(a_{\ell_i} a_{\ell_j})$, and denote it $|f^n(a_{\ell_i} a_{\ell_j})|^{ext}_{1\#2}$.

The value of $|f^n(a_{\ell_i} a_{\ell_j})|^{ext}_{1\#2}$ is obtained by adding, for all the integers r, $1 \leq r \leq k-1$, the product of the number of occurrences of the letter a_r in $f^n(a_{\ell_i})$ ($|f^n(a_{\ell_i})|_{a_r}$) by the number of occurrences of all the letters of $f^n(a_{\ell_j})$ greater than a_r ($|f^n(a_{\ell_j})|_{a_s}$, $r+1 \leq s \leq k$). This gives

$$\sum_{r=1}^{k-1} (m_{n,r,\ell_i} \cdot \sum_{s=r+1}^{k} m_{n,s,\ell_j}).$$

The number of external occurrences of 1#2 in all the $f^n(a_{\ell_i} a_{\ell_j})$, $1 \leq i < j \leq p_\ell$, is thus given by

$$\sum_{1\leq i<j\leq p_\ell} |f^n(a_{\ell_i} a_{\ell_j})|^{ext}_{1\#2} = \sum_{1\leq i<j\leq p_\ell} (\sum_{r=1}^{k-1} (m_{n,r,\ell_i} \cdot \sum_{s=r+1}^{k} m_{n,s,\ell_j})).$$

2. From $f^{n+1} = f \circ f^n$.
Let $q_\ell = |f^n(a_\ell)|$: $f^{n+1}(a_\ell) = f(f^n(a_\ell)) = f(a_{\ell'_1} \dots a_{\ell'_{q_\ell}})$. Here the number of occurrences of the ordered pattern 1#2 in $f^{n+1}(a_\ell)$ is obtained by adding:

• The number of occurrences of the ordered pattern 1#2 in each $f^n(a_{\ell'_i})$, $1 \leq i \leq q_\ell$. As above it is equal to $\sum_{t=1}^{k} |f(a_t)|_{1\#2} \cdot m_{n,t,\ell}$.

• The number of external occurrences of the ordered pattern 1#2 in each of the $f(a_{\ell'_i} a_{\ell'_j})$, $1 \leq i < j \leq q_\ell$. As above, this number is given by

$$\sum_{1 \leq i < j \leq q_\ell} |f(a_{\ell'_i} a_{\ell'_j})|^{ext}_{1\#2} = \sum_{1 \leq i < j \leq q_\ell} (\sum_{r=1}^{k-1} (m_{1,r,\ell'_i} \cdot \sum_{s=r+1}^{k} m_{1,s,\ell'_j})).$$

The same reasoning applies for calculating the entries of $DG(f^{n+1})$, replacing 1#2 by 2#1 and "greater" by "smaller". Thus we have the following.

Proposition 1. *For each letter $a_\ell \in A$, let p_ℓ and q_ℓ be such that $f(a_\ell) = a_{\ell_1} \dots a_{\ell_{p_\ell}}$ and $f^n(a_\ell) = a_{\ell'_1} \dots a_{\ell'_{q_\ell}}$. Then, for all $n \in \mathbb{N}$,*

$$|f^{n+1}(a_\ell)|_{1\#2} = \sum_{1 \leq i < j \leq q_\ell} (\sum_{r=1}^{k-1} (m_{1,r,\ell'_i} \cdot \sum_{s=r+1}^{k} m_{1,s,\ell'_j})) + \sum_{t=1}^{k} |f(a_t)|_{1\#2} \cdot m_{n,t,\ell}, \quad (1)$$

$$|f^{n+1}(a_\ell)|_{2\#1} = \sum_{1 \leq i < j \leq q_\ell} (\sum_{r=2}^{k} (m_{1,r,\ell'_i} \cdot \sum_{s=1}^{r-1} m_{1,s,\ell'_j})) + \sum_{t=1}^{k} |f(a_t)|_{2\#1} \cdot m_{n,t,\ell}. \quad (2)$$

Of course, the analysis we have done above could be realized to compute more complex ordered patterns with gaps, such as 1#23, 1#2#3, $\cdots$ The only difficulty is to adapt the computation of external inversions and non-inversions.

Repetitions of one letter Let n be a non-negative integer and p a positive integer. The *vector of p-repetitions with gaps of one letter* of f^n is the k vector whose i-th entry is the number of occurrences of the ordered pattern $(1\#)^p$ in the word $f^n(a_i)$, i.e., $R_pG(f^n) = (|f^n(a_i)|_{(1\#)^p})_{1 \leq i \leq k}$. The following is obvious.

Proposition 2. *For each letter $a_\ell \in A$ and for all $n \in \mathbb{N}$,*

$$|f^n(a_\ell)|_{(1\#)^p} = \sum_{t=1}^{k} \binom{m_{n,t,\ell}}{p}. \quad (3)$$

3.2 Some examples in the binary case

The Thue-Morse morphism The *Thue-Morse morphism* μ ([10],[9],[8]) is defined by $\mu(a_1) = a_1 a_2$, $\mu(a_2) = a_2 a_1$. It generates the famous *Thue-Morse sequence* $\mathbf{t} = \mu^\omega(a_1)$ which has been widely studied.

For every positive integers n, the incidence matrix of μ^n is

$$M(\mu^n) = \begin{bmatrix} 2^{n-1} & 2^{n-1} \\ 2^{n-1} & 2^{n-1} \end{bmatrix}.$$

Thus, from equations (1), (2), and (3) we obtain

$$\begin{aligned} |\mu^{n+1}(a_1)|_{1\#2} &= |\mu^{n+1}(a_2)|_{1\#2} = 2^{2(n-1)} + |\mu^n(a_1)|_{1\#2} + |\mu^n(a_2)|_{1\#2}, \\ |\mu^{n+1}(a_1)|_{2\#1} &= |\mu^{n+1}(a_2)|_{2\#1} = 2^{2(n-1)} + |\mu^n(a_1)|_{2\#1} + |\mu^n(a_2)|_{2\#1}, \\ |\mu^n(a_1)|_{(1\#)^p} &= |\mu^n(a_2)|_{(1\#)^p} = 2 \cdot \binom{2^{n-1}}{p}. \end{aligned}$$

Since $RG(\mu) = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $DG(\mu) = \begin{bmatrix} 0 & 1 \end{bmatrix}$, Proposition 1 gives the following well known result.

Corollary 1. *For any integer $n \geq 2$,*

$$RG(\mu^n) = DG(\mu^n) = \begin{bmatrix} 2^{2n-3} & 2^{2n-3} \end{bmatrix} \text{ and } R_pG(\mu^n) = \begin{bmatrix} 2 \cdot \binom{2^{n-1}}{p} & 2 \cdot \binom{2^{n-1}}{p} \end{bmatrix}.$$

The Fibonacci morphism The *Fibonacci morphism* φ is defined by $\varphi(a_1) = a_1a_2$, $\varphi(a_2) = a_1$. It generates the well known *Fibonacci sequence* $\mathbf{f} = \varphi^\omega(a_1)$ which is the prototype of a Sturmian word (see, e.g., [7]).

Let $(F_n)_{n\geq -1}$ be the sequence of Fibonacci numbers: $F_{-1} = 0$, $F_0 = 1$, $F_n = F_{n-1} + F_{n-2}$ for $n \geq 1$. The following property of Fibonacci numbers will be useful below.

Property 2 *For every positive integer n, $F_n \cdot F_{n-2} = F_{n-1}^2 + \begin{cases} 1 & \text{if } n \text{ is even,} \\ -1 & \text{if } n \text{ is odd.} \end{cases}$*

An easy computation gives that, for every positive integer n, the incidence matrix of φ^n is $M(\varphi^n) = \begin{bmatrix} F_n & F_{n-1} \\ F_{n-1} & F_{n-2} \end{bmatrix}$.

The vector of non-inversions of φ is $RG(\varphi) = \begin{bmatrix} 1 & 0 \end{bmatrix}$. Moreover equation (1) (see Property 2) gives, for $n \geq 1$

$$\begin{aligned} |\varphi^{n+1}(a_1)|_{1\#2} &= m_{n,1,1} \cdot m_{n,2,2} + |\varphi^n(a_1)|_{1\#2} + |\varphi^n(a_2)|_{1\#2} \\ &= F_n \cdot F_{n-2} + |\varphi^n(a_1)|_{1\#2} + |\varphi^n(a_2)|_{1\#2} \\ &= F_{n-1}^2 + |\varphi^n(a_1)|_{1\#2} + |\varphi^n(a_2)|_{1\#2} + \begin{cases} 1 & \text{if } n \text{ is even,} \\ -1 & \text{if } n \text{ is odd} \end{cases} \end{aligned}$$

The vector of inversions of φ is $DG(\varphi) = \begin{bmatrix} 0 & 0 \end{bmatrix}$. Moreover, equation (2) gives, for $n \geq 1$

$$\begin{aligned} |\varphi^{n+1}(a_1)|_{2\#1} &= m_{n,2,1} \cdot m_{n,1,2} + |\varphi^n(a_1)|_{2\#1} + |\varphi^n(a_2)|_{2\#1} \\ &= F_{n-1}^2 + |\varphi^n(a_1)|_{2\#1} + |\varphi^n(a_2)|_{2\#1}. \end{aligned}$$

Now, $|\varphi^{n+1}(a_2)|_{1\#2} = |\varphi^n(a_1)|_{1\#2}$ and $|\varphi^{n+1}(a_2)|_{2\#1} = |\varphi^n(a_1)|_{2\#1}$ because $\varphi(a_2) = a_1$.

From this we obtain formulas to compute, for every $n \geq 0$, $|\varphi^{n+2}(a_1)|_{1\#2}$ and $|\varphi^{n+2}(a_1)|_{2\#1}$ from the sequence of Fibonacci numbers.

Corollary 2. *For every integer $n \geq 0$,*

$$|\varphi^{n+2}(a_1)|_{2\#1} = \sum_{p=0}^{n} F_p F_{n-p}^2 \,,$$
$$|\varphi^{n+2}(a_1)|_{1\#2} = |\varphi^{n+2}(a_1)|_{2\#1} + F_n + \begin{cases} 1 \text{ if } n \text{ is odd,} \\ -1 \text{ if } n \text{ is even.} \end{cases}$$

Regarding repetitions of one letter, $R_pG(\varphi) = \left[\binom{1}{p} + \binom{1}{p} \; \binom{1}{p} \right]$ and, for $n \geq 0$, the vector $R_pG(\varphi^{n+2})$ is obtained from equation (3).

Corollary 3. *For any integer $n \geq 0$, $R_pG(\varphi^{n+2}) = \left[\binom{F_{n+2}}{p} + \binom{F_{n+1}}{p} \; \binom{F_{n+1}}{p} + \binom{F_n}{p} \right]$.*

4 A particular family of morphisms

Let k be an integer ($k \geq 2$) and A the k-letter ordered alphabet $A = \{a_1 < a_2 < \cdots < a_k\}$. In this section we are interested in morphisms f having the following particularities:

1. there exists a positive integer m such that $|f(a_1)|_{a_i} = m$, $1 \leq i \leq k$,
2. there exists a positive integer d such that $|f(a_2 \ldots a_k)|_{a_i} = d$, $1 \leq i \leq k$,
3. for every i, j, $1 \leq i, j \leq k$, $|f(a_i a_j)|_{1\#2}^{ext} = |f(a_j a_i)|_{1\#2}^{ext}$.

(Conditions 1. and 2. are particular cases of the more general situation, considered in Theorem 1 below, in which the alphabet A is partitioned in sets A_1, A_2, ..., A_n such that, for each A_i, the sum of the number of occurrences of each letter in the images of letters of A_i is the same.) In this case we are able to give direct formulas to compute $|f^{n+1}(a_1)|_{1\#2}$ and others from m, d, and n.

Proposition 3. *For every positive integer n,*

$$\begin{aligned}
&|f^{n+1}(a_1)|_{1\#2} \\
&\quad = m(d+m)^{n-1} \sum_{i=1}^{k} |f(a_i)|_{1\#2} + \tfrac{[m(d+m)^{n-1}-1]m(d+m)^{n-1}}{2} \sum_{j=1}^{k} |f(a_j a_j)|_{1\#2}^{ext} \\
&\quad +m^2(d+m)^{2n-2} \sum_{1 \leq i < j \leq k} |f(a_i a_j)|_{1\#2}^{ext} \,, \\
&|f^{n+1}(a_2 \ldots a_k)|_{1\#2} \\
&\quad = d(d+m)^{n-1} \sum_{i=1}^{k} |f(a_i)|_{1\#2} + \tfrac{[d(d+m)^{n-1}-1]d(d+m)^{n-1}}{2} \sum_{j=1}^{k} |f(a_j a_j)|_{1\#2}^{ext} \\
&\quad +d^2(d+m)^{2n-2} \sum_{1 \leq i < j \leq k} |f(a_i a_j)|_{1\#2}^{ext} \,.
\end{aligned}$$

Now the same reasoning can be applied for

$$|f^{n+1}(a_1)|_{2\#1} \text{ and } |f^{n+1}(a_2 \ldots a_k)|_{2\#1},$$

because of the following obvious property.

Property 3 *Let f be a morphism on A. For every non-negative integer n, and for every integers i, j, $1 \leq i, j \leq k$, $|f^n(a_i a_j)|_{1\#2}^{ext} = |f^n(a_j a_i)|_{2\#1}^{ext}$.*

Thus, using equation (2), we have the following.

Proposition 4. *For every positive integer n,*

$$\begin{aligned}
&|f^{n+1}(a_1)|_{2\#1} \\
&\quad = m(d+m)^{n-1}\sum_{i=1}^{k}|f(a_i)|_{2\#1} + \tfrac{[m(d+m)^{n-1}-1]m(d+m)^{n-1}}{2}\sum_{j=1}^{k}|f(a_ja_j)|_{2\#1}^{ext} \\
&\quad +m^2(d+m)^{2n-2}\sum_{1\le i<j\le k}|f(a_ia_j)|_{2\#1}^{ext}\ , \\
&|f^{n+1}(a_2\dots a_k)|_{2\#1} \\
&\quad = d(d+m)^{n-1}\sum_{i=1}^{k}|f(a_i)|_{2\#1} + \tfrac{[d(d+m)^{n-1}-1]d(d+m)^{n-1}}{2}\sum_{j=1}^{k}|f(a_ja_j)|_{2\#1}^{ext} \\
&\quad +d^2(d+m)^{2n-2}\sum_{1\le i<j\le k}|f(a_ia_j)|_{2\#1}^{ext}\ .
\end{aligned}$$

The previous reasoning can of course be applied if conditions 1. and 2. are verified for any partition of the alphabet (in Propositions 3 and 4 the partition is in two sets $A = \{a_1\} \cup \{a_2 \dots a_k\}$). Then we obtain the following general result.

Theorem 1. *Let k be an integer ($k \geq 2$), and A the k-letter ordered alphabet $A = \{a_1 < a_2 < \dots < a_k\}$. Let f be a morphism on A fulfilling the following conditions:*

- *there exist a positive integer p and a set of p positive integers $\{m_1, \dots, m_p\}$ such that A can be partitioned into p subsets $A_1, \dots, A_p$ with $\sum_{a\in A_\ell}|f(a)|_{a_i} = m_\ell$, $1 \leq i \leq k$,*
- *for every i, j, $1 \leq i, j \leq k$, $|f(a_ia_j)|_{1\#2}^{ext} = |f(a_ja_i)|_{1\#2}^{ext}$.*

Let $M = m_1 + \dots + m_p$ and let $u = 1\#2$ or $u = 2\#1$. For every integer $n \geq 1$ and for each A_ℓ, $1 \leq \ell \leq p$,

$$\begin{aligned}
\sum_{a\in A_\ell}|f^{n+1}(a)|_u = m_\ell M^{n-1}\sum_{i=1}^{k}|f(a_i)|_u + \tfrac{(m_\ell M^{n-1}-1)m_\ell M^{n-1}}{2}\sum_{j=1}^{k}|f(a_ja_j)|_u^{ext} \\
+m_\ell^2 M^{2n-2}\sum_{1\le i<j\le k}|f(a_ia_j)|_u^{ext}.
\end{aligned}$$

5 Examples

In this section we give a series of examples of application of Theorem 1.

5.1 The Thue-Morse morphism

The Thue-Morse morphism (see Section 3.2) is the simplest example of a morphism fulfilling conditions 1. to 3. above. Indeed $m = d = 1$, and

$$|\mu(a_1a_2)|_{1\#2}^{ext} = |a_1a_2a_2a_1|_{1\#2}^{ext} = 1 = |a_2a_1a_1a_2|_{1\#2}^{ext} = |\mu(a_2a_1)|_{1\#2}^{ext},$$

$|\mu(a_1a_1)|_{1\#2}^{ext} = |\mu(a_2a_2)|_{1\#2}^{ext} = 1$. Since $|\mu(a_1)|_{1\#2} = |\mu(a_2)|_{2\#1} = 1$, and $|\mu(a_1)|_{2\#1} = |\mu(a_2)|_{1\#2} = 0$, we obtain from Propositions 3 and 4 that

$$|\mu^{n+1}(a_1)|_{1\#2} = |\mu^{n+1}(a_1)|_{2\#1} = |\mu^{n+1}(a_2)|_{1\#2} = |\mu^{n+1}(a_2)|_{2\#1} = 2^{2n-1}.$$

5.2 The Prouhet morphisms

Let $k \geq 2$, and let A be the k-letter ordered alphabet $A = \{a_1 < \cdots < a_k\}$. The *Prouhet morphism* π_k ([9]) is defined on A by $\pi_k(a_i) = a_i a_{i+1} \ldots a_k a_1 \ldots a_{i-1}, 1 \leq i \leq k$. As above we obtain a corollary of Theorem 1.

Corollary 4. *For every i, $1 \leq i \leq k$, and for every positive integer n,*

$$|\pi_k^{n+1}(a_i)|_{1\#2} = \frac{(k-1)k^n}{12}\left(3k^{n+1} + k - 2\right),$$

$$|\pi_k^{n+1}(a_i)|_{2\#1} = \frac{(k-1)k^n}{12}\left(3k^{n+1} - k + 2\right).$$

5.3 The Arshon morphisms

Let k be any even positive integer. The morphism β_k ([1]) is defined, for every r, $1 \leq r \leq k/2$, by

$$\begin{array}{ll} a_{2r-1} & \mapsto a_{2r-1}a_{2r} \ldots a_{k-1}a_k a_1 a_2 \ldots a_{2r-3}a_{2r-2}, \\ a_{2r} & \mapsto a_{2r-1}a_{2r-2} \ldots a_2 a_1 a_k a_{k-1} \ldots a_{2r+1}a_{2r}. \end{array}$$

Corollary 5. *Let k be any even positive integer. For every i, $1 \leq i \leq k$, and for every positive integer n,*

$$\begin{array}{l} |\beta_k^{n+1}(a_i)|_{1\#2} = \frac{k^{n-1}}{4}\left[k^{n+2} \cdot (k-1) + 2k\right], \\ |\beta_k^{n+1}(a_i)|_{2\#1} = \frac{k^{n-1}}{4}\left[k^{n+2} \cdot (k-1) - 2k\right]. \end{array}$$

Example. For every i, $1 \leq i \leq k$, and for every $n \geq 1$,

$$|\beta_6^{n+1}(a_i)|_{1\#2} = 6^{n-1} \cdot (45 \cdot 6^n + 3), \quad |\beta_6^{n+1}(a_i)|_{2\#1} = 6^{n-1} \cdot (45 \cdot 6^n - 3).$$

5.4 Three other examples

1. Let A be the four-letter ordered alphabet $A = \{a_1 < a_2 < a_3 < a_4\}$. Define the morphism f on A by $f(a_1) = a_1a_3a_2a_4$, $f(a_2) = \varepsilon$, $f(a_3) = a_1a_4$, $f(a_4) = a_2a_3$.

The morphism f fulfills the conditions of Theorem 1. Here we choose $p = 3$, $A = A_1 \cup A_2 \cup A_3$ with $A_1 = \{a_1\}$, $A_2 = \{a_2\}$, $A_3 = \{a_3, a_4\}$, and $m_1 = m_3 = 1$, $m_2 = 0$, thus $M = 2$.

Corollary 6. *For every positive integer n,*

$$\begin{array}{l} |f^{n+1}(a_1)|_{1\#2} = |f^{n+1}(a_3a_4)|_{1\#2} = 3 \cdot 2^{n-1} \cdot (2^{n+1} + 1), \\ |f^{n+1}(a_1)|_{2\#1} = |f^{n+1}(a_3a_4)|_{2\#1} = 3 \cdot 2^{n-1} \cdot (2^{n+1} - 1), \\ |f^{n+1}(a_2)|_{1\#2} = |f^{n+1}(a_2)|_{2\#1} \quad = 0. \end{array}$$

2. Let A be the five-letter ordered alphabet $A = \{a_1 < a_2 < a_3 < a_4 < a_5\}$. Define the morphism g on A by $g(a_1) = a_1a_3a_5a_4a_2$, $g(a_2) = a_4a_2a_3$, $g(a_3) = a_5a_1$, $g(a_4) = a_1a_5$, $g(a_5) = a_2a_3a_4$.

The morphism g fulfills the conditions of Theorem 1. Here we choose $p = 3$, $A = A_1 \cup A_2 \cup A_3$ with $A_1 = \{a_1\}$, $A_2 = \{a_2, a_4\}$, $A_3 = \{a_3, a_5\}$, and $m_1 = m_2 = m_3 = 1$, thus $M = 3$.

Corollary 7. *For every positive integer n,*

$$|g^{n+1}(a_1)|_{1\#2} = |g^{n+1}(a_2a_4)|_{1\#2} = |g^{n+1}(a_3a_5)|_{1\#2} = 3^{n-1} \cdot (5 \cdot 3^{n+1} + 2),$$
$$|g^{n+1}(a_1)|_{2\#1} = |g^{n+1}(a_2a_4)|_{2\#1} = |g^{n+1}(a_3a_5)|_{2\#1} = 3^{n-1} \cdot (5 \cdot 3^{n+1} - 2).$$

3. Let A be the three-letter ordered alphabet $A = \{a < b < c\}$. Define the morphism h on A by $h(a) = aba\,cab\,cac\,bab\,cba\,cbc$, $h(b) = aba\,cab\,cac\,bca\,bcb\,abc$, $h(c) = aba\,cab\,cba\,cbc\,acb\,abc$.
This morphism was proved square-free by Brandenburg in [3]. It fulfills the conditions of Theorem 1 with $p = 3$, $A = A_1 \cup A_2 \cup A_3$ with $A_1 = \{a\}$, $A_2 = \{b\}$, $A_3 = \{c\}$, and $m_1 = m_2 = m_3 = 6$, thus $M = 18$.

Corollary 8. *For every $x \in A$ and for every positive integer n,*

$$|h^{n+1}(x)|_{1\#2} = 6 \cdot 18^{n-1} \cdot (9 \cdot 18^{n+1} + 40),$$
$$|h^{n+1}(x)|_{2\#1} = 6 \cdot 18^{n-1} \cdot (9 \cdot 18^{n+1} - 40).$$

6 Ordered patterns with no gaps and morphisms

6.1 Rises, descents, and squares of f^n

Let k be an integer ($k \geq 2$) and A the k-letter ordered alphabet $A = \{a_1 < a_2 < \cdots < a_k\}$. Let f be any morphism on A: for $1 \leq i \leq k$, $f(a_i) = a_{i_1} \ldots a_{i_{p_i}}$ with $p_i \geq 0$ ($p_i = 0$ if and only if $f(a_i) = \varepsilon$).

The *vector of rises* (resp. *vector of descents*, resp. *vector of squares of one letter*) of f^n is the k vector whose i-th entry is the number of occurrences of the ordered pattern 12 (resp. 21, resp. 11) in the word $f^n(a_i)$, i.e.,

$$R(f^n) = (|f^n(a_i)|_{12})_{1\leq i\leq k}, \quad D(f^n) = (|f^n(a_i)|_{21})_{1\leq i\leq k},$$
$$R_2(f^n) = (|f^n(a_i)|_{11})_{1\leq i\leq k}.$$

We define two sequences of k vectors, $(F(f^n))_{n\in\mathbb{N}}$ and $(L(f^n))_{n\in\mathbb{N}}$, where $F(f^n)[i]$ is the first letter of $f^n(a_i)$ and $L(f^n)[i]$ is the last letter of $f^n(a_i)$ if $f^n(a_i) \neq \varepsilon$, and $F(f^n)[i] = L(f^n)[i] = 0$ if $f^n(a_i) = \varepsilon$. Of course these two sequences take their values in a finite set: they are ultimately periodic. Thus they can be computed *a priori* from f.

Given a non-negative integer n, let $\aleph'$ be the subset of $\aleph$ such that, for each $i \in \aleph$, $f^n(a_i) \neq \varepsilon$ if and only if $i \in \aleph'$. We associate to the two vectors $F(f^n)$ and $L(f^n)$ an application $C_n^{12} : \aleph' \times \aleph' \to \{0, 1\}$ defined by

$$C_n^{12}(i,j) = \begin{cases} 1, \text{ if } L(f^n)[i] < F(f^n)[j] \\ 0, \text{ if } L(f^n)[i] \geq F(f^n)[j]. \end{cases}$$

Similarly we define

$$C_n^{21}(i,j) = \begin{cases} 1, \text{ if } L(f^n)[i] > F(f^n)[j] \\ 0, \text{ if } L(f^n)[i] \le F(f^n)[j], \end{cases} \quad C_n^{11}(i,j) = \begin{cases} 1, \text{ if } L(f^n)[i] = F(f^n)[j] \\ 0, \text{ if } L(f^n)[i] \ne F(f^n)[j]. \end{cases}$$

For any morphism f on A, there exists a least integer M_f ($M_f \le k$ and M_f depends only on f) such that, for every positive integer n and every $a \in A$, $f^n(a) = \varepsilon$ if and only if $f^{M_f}(a) = \varepsilon$. By convention, if f is a nonerasing morphism then $M_f = 0$. The integer M_f is known in the literature about L-systems as the *mortality exponent* of f ([6]).

Now let ℓ be an integer, $1 \le \ell \le k$. One has $f(a_\ell) = a_{\ell_1} \dots a_{\ell_{p_\ell}}$ and we denote by $\ell'_1 \dots \ell'_{p'_\ell}$ the subsequence of $\ell_1 \dots \ell_{p_\ell}$ such that $f^{n+1}(a_\ell) = f^n(a_{\ell'_1} \dots a_{\ell'_{p'_\ell}})$ for every $n \ge M_f$. This means that, for every $n \ge M_f$, a letter a_{ℓ_i} appears in $a_{\ell_1} \dots a_{\ell_{p_\ell}}$ but not in $a_{\ell'_1} \dots a_{\ell'_{p'_\ell}}$ if and only if $f^n(a_{\ell_i}) = \varepsilon$. Of course $p'_\ell \le p_\ell$, and if $M_f = 0$ then $p'_\ell = p_\ell$ for each $1 \le \ell \le k$.

Here also, as in Section 3, the number of occurrences of the ordered pattern 12 in $f^{n+1}(a_\ell) = f^n(a_{\ell_1} \dots a_{\ell_{p_\ell}}) = f^n(a_{\ell'_1} \dots a_{\ell'_{p'_\ell}})$ ($n \ge M_f$) is obtained by adding two values: (1) the number of occurrences of the ordered pattern 12 in each $f^n(a_{\ell_i})$, $1 \le i \le p_\ell$. As in the previous case, this number is equal to $\sum_{t=1}^{k} |f^n(a_t)|_{12} \cdot m_{1,t,\ell}$, and (2) the number of external occurrences of the ordered pattern 12 in $f^n(a_{\ell'_i} a_{\ell'_j})$ for each subsequence $a_{\ell'_i} a_{\ell'_j}$ of $f(a_\ell)$, $1 \le i < j \le p'_\ell$. But the only possibility for 12 to be an external occurrence in $f^n(a_{\ell'_i} a_{\ell'_j})$ is that $j = i+1$ and the last letter of $f^n(a_{\ell'_i})$ is smaller than the first letter of $f^n(a_{\ell'_j})$. Thus, the number of occurrences of such patterns is only the number of times $L(f^n)[i] < F(f^n)[i+1]$ with $i+1 \le p'_\ell$, i.e., the number of times $C_n^{12}(\ell'_i, \ell'_{i+1}) = 1$ for $1 \le i \le p'_\ell - 1$.

We proceed similarly with the patterns 21 and 11. Consequently we have the following proposition.

Proposition 5. *For each letter $a_\ell \in A$, $f(a_\ell) = a_{\ell_1} \dots a_{\ell_{p_\ell}}$, and for all $n \ge M_f$, let $\ell'_1 \dots \ell'_{p'_\ell}$ be the subsequence of $\ell_1 \dots \ell_{p_\ell}$ such that $f^{n+1}(a_\ell) = f^n(a_{\ell'_1} \dots a_{\ell'_{p'_\ell}})$ and $f^n(a_{\ell'_i}) \ne \varepsilon$, $1 \le i \le p'_\ell$. Then*

$$|f^{n+1}(a_\ell)|_{12} = \sum_{t=1}^{k} |f^n(a_t)|_{12} \cdot m_{1,t,\ell} + \sum_{i=1}^{p'_\ell - 1} C_n^{12}(\ell'_i, \ell'_{i+1}), \tag{4}$$

$$|f^{n+1}(a_\ell)|_{21} = \sum_{t=1}^{k} |f^n(a_t)|_{21} \cdot m_{1,t,\ell} + \sum_{i=1}^{p'_\ell - 1} C_n^{21}(\ell'_i, \ell'_{i+1}), \tag{5}$$

$$|f^{n+1}(a_\ell)|_{11} = \sum_{t=1}^{k} |f^n(a_t)|_{11} \cdot m_{1,t,\ell} + \sum_{i=1}^{p'_\ell - 1} C_n^{11}(\ell'_i, \ell'_{i+1}). \tag{6}$$

6.2 Some examples

No external rises, no external descents, no external squares Let us suppose that the morphism f is such that, for all i and j, $L(f)[i] \geq F(f)[j]$ (there are no external rises). According to equation (4), in this case, for each letter $a_\ell \in A$, $f(a_\ell) = a_{\ell_1} \dots a_{\ell_{p_\ell}}$, and for all $n \geq M_f$,

$$|f^{n+1}(a_\ell)|_{12} = \sum_{t=1}^{k} |f^n(a_t)|_{12} \cdot m_{1,t,\ell}.$$

Moreover, if the above inequality is strict then, according to equation (5),

$$|f^{n+1}(a_\ell)|_{21} = \sum_{t=1}^{k} |f^n(a_t)|_{21} \cdot m_{1,t,\ell} + p'_\ell - 1.$$

Now if we suppose that, conversely to the previous case, the morphism f is such that, for all i and j, $L(f)[i] \leq F(f)[j]$ (there are no external descents) then we obtain the same result by switching 12 and 21 in the above formulas.

To end, if we suppose that the morphism f is such that, for all i and j, $L(f)[i] \neq F(f)[j]$ then, according to equation (6), for each letter $a_\ell \in A$, $f(a_\ell) = a_{\ell_1} \dots a_{\ell_{p_\ell}}$, and for all $n \geq M_f$,

$$|f^{n+1}(a_\ell)|_{11} = \sum_{t=1}^{k} |f^n(a_t)|_{11} \cdot m_{1,t,\ell}.$$

The Thue-Morse morphism Since $R(\mu) = \begin{bmatrix} 1 & 0 \end{bmatrix}$, $D(\mu) = \begin{bmatrix} 0 & 1 \end{bmatrix}$ and $R_2(\mu) = \begin{bmatrix} 0 & 0 \end{bmatrix}$ we obtain again a well known result.

Corollary 9. *For any integer $n \geq 0$,*

$$\begin{aligned} R(\mu^{2n}) &= \begin{bmatrix} \frac{4^n-1}{3} & \frac{4^n-1}{3} \end{bmatrix} = D(\mu^{2n}) = R_2(\mu^{2n}) \\ R(\mu^{2n+1}) &= \begin{bmatrix} \frac{2(4^n-1)}{3}+1 & \frac{2(4^n-1)}{3} \end{bmatrix} \\ D(\mu^{2n+1}) &= \begin{bmatrix} \frac{2(4^n-1)}{3} & \frac{2(4^n-1)}{3}+1 \end{bmatrix} \\ R_2(\mu^{2n+1}) &= \begin{bmatrix} \frac{2(4^n-1)}{3} & \frac{2(4^n-1)}{3} \end{bmatrix}. \end{aligned}$$

The Fibonacci morphism Since $R(\varphi) = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $D(\varphi) = R_2(\varphi) = \begin{bmatrix} 0 & 0 \end{bmatrix}$ we have again a well known result.

Corollary 10. *For any integer $n \geq 1$,*

$$\begin{aligned} R(\varphi^n) &= \begin{bmatrix} F_{n-1} & F_{n-2} \end{bmatrix} \\ D(\varphi^{2n}) &= \begin{bmatrix} F_{2n-1} & F_{2n-2}-1 \end{bmatrix} = R_2(\varphi^{2n+1}) \\ R_2(\varphi^{2n}) &= \begin{bmatrix} F_{2n-2}-1 & F_{2n-3} \end{bmatrix} = D(\varphi^{2n-1}). \end{aligned}$$

Erasing morphisms Let A be the four-letter ordered alphabet $A = \{a_1 < a_2 < a_3 < a_4\}$.

1. Here we consider the erasing morphism f, given in Section 5.4, defined on A by $f(a_1) = a_1a_3a_2a_4$, $f(a_2) = \varepsilon$, $f(a_3) = a_1a_4$, $f(a_4) = a_2a_3$. One has $M_f = 1$.

Starting from $R(f) = \left[2\ 0\ 1\ 1\right]$, we obtain the following corollary of Proposition 5.

Corollary 11. *For any integer $n \geq 1$, $R_2(f^n) = \left[0\ 0\ 0\ 0\right]$ and*

$$\text{if } n \text{ is even} \begin{cases} R(f^n) = \left[2^n\ 0\ \frac{2^{n+1}+1}{3}\ \frac{2^n-1}{3}\right] \\ D(f^n) = \left[2^n - 1\ 0\ \frac{2^{n+1}-2}{3}\ \frac{2^n-4}{3}\right], \end{cases}$$

$$\text{if } n \text{ is odd} \begin{cases} R(f^n) = \left[2^n\ 0\ \frac{2^{n+1}-1}{3}\ \frac{2^n+1}{3}\right] \\ D(f^n) = \left[2^n - 1\ 0\ \frac{2^{n+1}-4}{3}\ \frac{2^n-2}{3}\right]. \end{cases}$$

2. Now we consider the erasing morphism g defined on A by $g(a_1) = a_1a_2a_4a_3$, $g(a_2) = a_3$, $g(a_3) = \varepsilon$, $g(a_4) = a_1a_2a_4$. Here we have $M_f = 2$.

Corollary 12. $R(g) = \left[2\ 0\ 0\ 2\right]$*,* $D(g) = \left[1\ 0\ 0\ 0\right]$*,* $R_2(g) = \left[0\ 0\ 0\ 0\right]$*, and, for any integer $n \geq 2$,*

$$R(g^n) = \left[2^n\ 0\ 0\ 2^n\right], \qquad D(g^n) = \left[2^{n-1} + 2^{n-2} - 1\ 0\ 0\ 2^{n-1} + 2^{n-2} - 1\right]$$
$$R_2(g^n) = \left[2^{n-2}\ 0\ 0\ 2^{n-2}\right].$$

References

1. S. Arshon, Démonstration de l'existence de suites asymétriques infinies, *Mat. Sb.* **44** (1937), 769–777 (in Russian), 777–779 (French summary).
2. E. Babson, E. Steingrímsson, Generalized permutation patterns and a classification of the Mahonian statistics, Sém. Lothar. Comb. **44** (2000), Art. B44b.
3. F.-J. Brandenburg, Uniformly growing k-th power-free homomorphisms, *Theoret. Comput. Sci.* **23** (1983), 69–82.
4. A. Burstein, Enumeration of words with forbidden patterns, Ph.D. thesis, University of Pennsylvania, 1998.
5. S. Kitaev, T. Mansour, P. Séébold, The Peano curve and counting occurrences of some patterns, *J. Autom., Lang. Combin.* **9** (2004) 4, 439–455.
6. F. Levé, G. Richomme, On a conjecture about finite fixed points of morphisms, *Theoret. Comput. Sci.* **339-1** (2005), 103–128.
7. M. Lothaire, *Algebraic Combinatorics on Words*, vol.90 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge, UK, 2002.
8. M. Morse, Recurrent geodesics on a surface of negative curvature, *Trans. Amer. Math. Soc.* **22** (1921), 84–100.
9. M. E. Prouhet, Mémoire sur quelques relations entre les puissances des nombres, *Comptes Rendus Acad. Sci. Paris* **33** (1851), 225.
10. A. Thue, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, *Vidensk.-Selsk. Skrifter. I. Mat. Nat. Kl.* **1** Kristiania (1912), 1–67. Reprinted in *Selected Mathematical Papers of Axel Thue*, T. Nagell, editor, Universitetsforlaget, Oslo, 1977, 413-478.