ELSEVIER Theoretical Computer Science 158 (1996) 35–51

Theoretical Computer Science

About the p-paperfolding words

Michel Koskas

Université Bordeaux I, Algorithmique Arithmétique eXpérimentale. UMR n° 9936 CNRS. 351, cours de la Libération F-33405 Talence Cedex, France

Received April 1994; revised December 1994
Communicated by D. Perrin

Abstract

Let p be an integer greater than or equal to 2. The aim of this paper is to study the language associated to a p-paperfolding sequence. It is known that the number of factors of length n of a 2-paperfolding sequence (i.e. its complexity function) is $P(n) = 4n$ for $n \geqslant 7$. It is also known that the language of all the factors of all 2-paperfolding sequences is not context-free and that its generating function is transcendental.

We show that the complexity function of a p-paperfolding sequence is either strictly subaffine or ultimately linear. The first case never happens if $p = 2$ or 3. In the second case, the complexity function is either $P(n) = 2n$ or $P(n) = 4n$ for n large enough. We give a simple necessary and sufficient condition for the number of special factors to be p-automatic. We finally show that, for any given p, the language of all factors of all p-paperfolding sequences is not context-free, and that the associated generating series is not algebraic.

0. Introduction

Let u be a sequence taking its values in a finite alphabet. How "complicated" is it? One possible answer to this question is the following definition.

Definition 1. The complexity (or factor-complexity) of a sequence is the function $n \mapsto P(n)$ where $P(n)$ is the number of factors (blocks) of length n of the sequence.

For a survey on the complexity of sequences one can read [3]. A family of sequences for which the complexity function has been computed consists of the *paperfolding sequences*. These sequences are obtained by repeatedly folding a piece of paper onto itself. At each step, one can fold the paper in two different ways, thus generating uncountably many sequences.

It is known that *all* the paperfolding sequences have the same complexity $P(n)$. Furthermore, one has $P(n) = 4n$ for $n \geqslant 7$ (see [1, 4]).

SSDI 0304-3975(95)00008-9

In [4], the authors also study the language $\mathscr{E}$ of *all* the factors of *all* paperfolding sequences. They compute the *complexity function* of this language (i.e. the number of words of given length in $\mathscr{E}$) and they show that $\mathscr{E}$ is not context-free and has a transcendental generating function.

In this work we study the p-paperfolding sequences which are constructed by folding a piece of paper onto itself in p parts and iterating this process, ([11], see also [14] for related sequences). These sequences are Toeplitz sequences (see [3] for a survey) and some of their complexity functions are in $O(n)$ (see [8]). We prove here the following results:

- The complexity function of a 3-paperfolding sequence is either $P(n) = 2n$ for $n \geqslant 1$ or $P(n) = 4n$ for n large enough.
- If $p \geqslant 4$, the complexity function of a p-paperfolding sequence is either strictly subaffine or linear. In the first case we give a necessary and sufficient condition for the number of special factors to be p-automatic. In the second case we show that, for n large enough, $P(n) = 2n$ or $P(n) = 4n$.
- For any given p the language of all factors of all p-paperfolding sequences is not context-free and its generating function is transcendental.

1. p-paperfolding sequences

1.1. Construction

Fold a piece of paper in p parts, (there are 2^{p-1} possibilities, each of them is called a *folding instruction*). Repeat this operation an infinite number of times, then unfold the paper. One obtains on its edge a sequence of "mountains" and "valleys" which is by definition a p-paperfolding sequence.

If at each step one chooses the same folding instruction, then the sequence is called a *regular* p-paperfolding sequence. Otherwise it is called a *generalized* p-paperfolding sequence.

For instance, one obtains a regular 3-paperfolding sequence by choosing at each step the folding instruction $\vee\wedge$ (this means that the paper is folded in three parts, the left side being folded towards up, and the right side being folded towards down):

1 $\vee\ \wedge$
2 $\vee\ \wedge\ \vee\ \vee\ \wedge\ \wedge\ \vee\ \wedge$
3 $\vee\ \wedge\ \vee\ \vee\ \wedge\ \wedge\ \vee\ \wedge\ \vee\ \vee\ \wedge\ \vee\ \vee\ \wedge\ \wedge\ \vee\ \wedge\ \wedge\ \vee\ \wedge\ \vee\ \vee\ \wedge\ \wedge\ \vee\ \wedge$
4 $\vee\ \wedge\ \vee\ \vee\ \wedge\ \wedge\ \vee\ \wedge\ \vee\ \vee\ \wedge\ \vee\ \vee\ \wedge\ \wedge\ \vee\ \wedge\ \wedge\ \vee\ \wedge\ \vee\ \vee\ \wedge\ \wedge\ \vee\ \wedge \ldots$

One can also obtain a *generalized* 3-paperfolding sequence, choosing to fold randomly at each step the paper according to the folding instructions $\vee\vee$, $\vee\wedge$, $\wedge\vee$ or $\wedge\wedge$. Take for instance for the *last* three folding instructions equal to $\vee\wedge$, $\vee\vee$

and ∧ ∧. This gives:

1 ∨ ∧
2 ∨ ∧ ∨ ∨ ∧ ∨ ∨ ∧
3 ∨ ∧ ∨ ∨ ∧ ∨ ∨ ∧ ∧ ∨ ∧ ∧ ∨ ∧ ∧ ∨ ∧ ∧ ∨ ∧ ∨ ∨ ∧ ∨ ∨ ∧
4 ∨ ∧ ∨ ∨ ∧ ∨ ∨ ∧ ∧ ∨ ∧ ∧ ∨ ∧ ∧ ∨ ∧ ∧ ∨ ∧ ∨ ∨ ∧ ∨ ∨ ∧ ...

A first algorithm to construct a p-paperfolding sequence is the following. Suppose that we have folded the paper $n+1$ times, if we unfold it n times, we have a sequence obtained after n folds and, to obtain the sequence issued from our $n+1$ folds, we have to unfold it a last time. Let w_n be the word on the alphabet $\{\vee, \wedge\}$, obtained after folding n times, *but following the last n folding instructions.* Let $\alpha_1 \ldots \alpha_{p-1}$ be the first folding instruction. One has

$$w_{n+1} = w_n \alpha_1 f(w_n) \alpha_2 f^2(w_n) \ldots \alpha_{p-1} f^{p-1}(w_n),$$

where $f(\omega_n)$ is obtained by reading w_n backwards and interchanging the symbols $\vee$ and $\wedge$, (of course $f^2(w_n) = w_n$). (From now on, the word $f(w)$ will be denoted by $\hat{w}$).

Taking again the generalized 3-paperfolding example: (the symbol ↓ denotes the place of the third, second and first folding instructions)

1 ∨ ∧ ↓ ↓
2 ∨ ∧ ∨ ∨ ∧ ∨ ∨ ∧ ↓ ↓
3 ∨ ∧ ∨ ∨ ∧ ∨ ∨ ∧ ∧ ∨ ∧ ∧ ∨ ∧ ∧ ∨ ∧ ∧ ∨ ∧ ∨ ∨ ∧ ∨ ∨ ∧ ↓...
4 ∨ ∧ ∨ ∨ ∧ ∨ ∨ ∧ ∧ ∨ ∧ ∧ ∨ ∧ ∧ ∨ ∧ ∧ ∨ ∧ ∨ ∨ ∧ ∨ ∨ ∧ ...

A particular case is the 2-paperfolding.

Example (*general 2-paperfolding sequences*). One can obtain all 2-paperfolding sequences by randomly folding a paper in two (at each step the left side on or under the right one). The preceding algorithm gives, (denoting $\vee$ by 0 and $\wedge$ by 1), choosing 0 at the first step, 1 at the second, 1 at the third, and 0 at the fourth: (↓ denotes again the place of the third, the second and the first folding instructions)

1 0 ↓
2 0 1 1 ↓
3 0 1 1 1 0 0 1 ↓
4 0 1 1 1 0 0 1 0 0 1 1 0 0 0 1

It has been shown in [1] that for any 2-paperfolding sequence, the complexity function is $P(n) = 4n$ for $n \geqslant 7$.

1.2. p-paperfolding sequences defined as Toeplitz sequences

Definitions. Let us consider the alphabet $\mathscr{A} = \{0, 1\}$. Let a_n ($n \geqslant 0$) be a sequence of words of length $p - 1$ over $\mathscr{A}$, ($a_n = a_n[1] \dots a_n[p-1]$).

One defines ${}^t0 = 1$, ${}^t1 = 0$, and ${}^ta_n = {}^ta_n[1]\,{}^ta_n[2] \dots {}^ta_n[p-1]$. One also defines $\bar{a}_n = a^n[p-1] \dots a_n[1]$, and denotes by $\hat{a}_n$ the word ${}^t\overline{a_n}$.

Let $\bullet$ be a new symbol. A periodic word over $\mathscr{A} \cup \{\bullet\}(\zeta_1 \dots \zeta_T \zeta_1 \dots \zeta_T \zeta_1 \dots \zeta_T \dots)$ will be denoted $(\zeta_1 \zeta_2 \dots \zeta_T)^\infty$.

Let B_n the sequence defined over $\mathscr{A} \cup \{\bullet\}$ by

$$B_n = (a_n \bullet \hat{a}_n \bullet)^\infty.$$

Let $A_0^1 = B_0$ and for $n \in \mathbb{N}$, define the periodic sequence A_{n+1}^1 by replacing the letters $\bullet$ of A_n^1 by the letters of B_{n+1}. When n goes to infinity, the sequence A_n^1 tends to a limit over $\mathscr{A}$. This limit, denoted by A^1 is called the p-paperfolding sequence according to the sequence of instructions $(B_n)_{n \geqslant 1}$.

From now on, A^1 will denote a sequence obtained as above. One writes $A^1 = A^1[k]_{k \in \mathbb{N}^*}$.

In the same way, for $i \geqslant 1$, let $A_0^i = B_{i-1}$ and for $n \in \mathbb{N}$ define the periodic sequence A_{n+1}^i obtained by replacing the letters $\bullet$ in A_n^i by the letters of B_{n-1+i}. This sequence also tends to a limit A^i. From now on, A^i will denote such a p-paperfolding sequence.

A finite word occurring at least one time in a p-paperfolding sequence will be called a *p-paperfolding word*.

Finally a finite word $m = m[1] \dots m[n]$ is said to be 2-periodic if $\forall j$, $1 \leqslant j \leqslant n-2$, $m[j] = m[j+2]$.

Proposition 2. *This construction yields exactly the p-paperfolding sequences.*

The proof is left to the reader.

1.3. Example

With $p = 2$, a_n is reduced to a single symbol: $\forall n$, $a_n = 0$ or $a_n = 1$. Let us construct a prefix of such a Toeplitz sequence, with $a_0 = 0$, $a_1 = 1$, $a_2 = 1$ and $a_3 = 0$:

$$B_0 = (0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet \dots),$$

$$B_1 = (1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet \dots),$$

$$B_2 = (1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet \dots),$$

$$B_3 = (0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet \dots).$$

This gives us, for the A_n^0 sequences ($0 \leqslant n \leqslant 3$):

$$A_0^0 = 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet \dots,$$

then:

$$B_1 = \quad 1 \quad \bullet \quad 0 \quad \bullet \quad 1 \quad \bullet \quad 0 \quad \bullet \quad 1 \quad \bullet \ldots$$

$$A_0^1 = 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet 0 \bullet 1 \bullet \ldots$$

$$A_1^1 = 0\ 1\ 1 \bullet 0\ 0\ 1 \bullet 0\ 1\ 1 \bullet 0\ 0\ 1 \bullet 0\ 1\ 1 \bullet \ldots$$

hence

$$B_2 = \quad 1 \quad \bullet \quad 0 \quad \bullet \quad 1 \ldots$$

$$A_1^1 = 0\ 1\ 1 \bullet 0\ 0\ 1 \bullet 0\ 1\ 1 \bullet 0\ 0\ 1 \bullet 0\ 1\ 1 \bullet \ldots$$

$$A_2^1 = 0\ 1\ 1\ 1\ 0\ 0\ 1 \bullet 0\ 1\ 1\ 0\ 0\ 0\ 1 \bullet 0\ 1\ 1\ 1 \ldots$$

and

$$B_3 = \quad 0 \quad \bullet$$

$$A_2^1 = 0\ 1\ 1\ 1\ 0\ 0\ 1 \bullet 0\ 1\ 1\ 0\ 0\ 0\ 1 \bullet 0\ 1\ 1\ 1 \ldots$$

$$A_3^1 = 0\ 1\ 1\ 1\ 0\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 1 \bullet 0\ 1\ 1\ 1 \ldots$$

We recognize here the beginning of the example given in the preceding paragraph.

2. Induction formulas satisfied by the complexity of p-paperfolding sequences

2.1. ν-factors

Let $\nu \in [0, p-1]$ be an integer, and $i \geqslant 1$. A ν-factor of A^i is a factor of A^i beginning at least once at an index $j \equiv \nu (\mathrm{mod}\, p)$. We define $\varphi_\nu^i(m)$ as being the number of ν-factors of A^i of length m.

We also define

$$\varphi^i(m) = \sum_{\nu=0}^{\nu=p-1} \varphi_\nu^i(m).$$

We finally define $P_{\text{even}}^i(m)$ (resp. $P_{\text{odd}}^i(m)$) as being the number of factors of A^i of length m beginning at least once at an even (resp. odd) index $P^i(m)$ as being the number of factors of A^i of length m. In general,

$$P^i(m) \leqslant P_{\text{even}}^i(m) + P_{\text{odd}}^i(m)$$

and

$$P^i(m) \leqslant \varphi^i(m).$$

Lemma 3. *No factor of length greater than or equal to $p+2$ of a p-paperfolding sequence M is 2-periodic.*

Proof. From the construction of these Toeplitz sequences, it is clear that any letter of index $(kp-1)$ is different from the letter of index $kp+1$. Indeed, $A_0^i = (m \bullet \hat{m} \bullet m \bullet \hat{m} \bullet \dots)$ and the last letter of m is different from the first letter of $\hat{m}$. As any factor of M of length greater than or equal to $p+2$ contains a letter of index multiple of p which is neither its first nor its last letter, the result holds. □

Lemma 4. *Let k be an integer not multiple of p. Then no factor of length greater than or equal to $p+2k$ of a p-paperfolding sequence is 2k-periodic.*

The proof is exactly as above.

We now give a first lemma which will allow us to compute the factors of a p-paperfolding sequence by counting separately the factors which begin at a given index modulo p.

Synchronization lemma. *For any distinct numbers v and $v' \in [0, p-1]$, any factor M of length greater than or equal to $p(p+2)-1$ of a p-paperfolding sequence cannot be simultaneously v-factor and v'-factor.*

Proof. Let u be a factor of M of length $n \geqslant p(p+2)-1$ occurring simultaneously at an index i_1 and an index i_2, with $i_1 \not\equiv i_2 \pmod p$.

As $n \geqslant p(p+2)-1$ and $i_1 \not\equiv i_2 \pmod p$, there are at least $p+2$ letters of index multiple of p in the first or in the second occurrence of u (possibly in both). Let us suppose they are in the first one. These letters form a word w of length greater than or equal to $p+2$, which is a factor of the p-paperfolding sequence obtained by "forgetting" the first folding instruction.

This word w is also the subword of the second occurrence of u formed with the letters congruent to $i_2 - i_1 \pmod p$. As $i_1 - i_2 \not\equiv 0 \pmod p$, the word w is 2-period. But $|w| \geqslant p+2$. This contradicts the Lemma 3. □

We now give two induction lemmas which will help us to compute the complexity functions of the p-paperfolding sequences.

Lemma 5. *If $\hat{a}_1 = a_1$, then $\forall i$, $0 \leqslant i \leqslant p-1$, $\forall j \geqslant p$, one has*

$$\varphi_i^1(j) = P^2(l_{i,j}),$$

where $l_{i,j}$ is the number of integers multiple of p in $[i, i+j-1]$.

Lemma 6. *One has*

$$l_{i,j} = \left\lfloor \frac{j}{p} \right\rfloor + \begin{cases} 0 & \text{if} \begin{cases} r_i \neq 0 \text{ and } r_i + r_j - 1 \leqslant p, \\ \text{or} \\ r_i = r_j = 0, \end{cases} \\ 1 & \text{otherwise.} \end{cases}$$

with r_i, r_j the remainders of the euclidian division of i and j by p.

Proof. Define

$$k_{i,j} = \left\lfloor \frac{j}{p} \right\rfloor + \begin{cases} 0 & \text{if} \begin{cases} r_i \neq 0 \text{ and } r_i + r_j - 1 \leqslant p, \\ or \\ r_i = r_j = 0, \end{cases} \\ 1 & \text{otherwise.} \end{cases}$$

By definition,

$$l_{i,j} = \left\lfloor \frac{i+j-1}{p} \right\rfloor - \left\lfloor \frac{i-1}{p} \right\rfloor.$$

Case 1: If $i \equiv 0 \pmod p$:

$$\left\lfloor \frac{i-1}{p} \right\rfloor = \left\lfloor \frac{i}{p} \right\rfloor - 1,$$

$$\left\lfloor \frac{i+j-1}{p} \right\rfloor = \left\lfloor \frac{i}{p} \right\rfloor + \left\lfloor \frac{j}{p} \right\rfloor + \left\lfloor \frac{r_j - 1}{p} \right\rfloor.$$

If $j \equiv 0 \pmod p$ then $\left\lfloor \frac{i+j-1}{p} \right\rfloor = \left\lfloor \frac{i}{p} \right\rfloor + \left\lfloor \frac{j}{p} \right\rfloor - 1$ and $l_{i,j} = \left\lfloor \frac{j}{p} \right\rfloor = k_{i,j}$.

If $j \not\equiv 0 \pmod p$ then $\left\lfloor \frac{i+j-1}{p} \right\rfloor = \left\lfloor \frac{i}{p} \right\rfloor + \left\lfloor \frac{j}{p} \right\rfloor$ and $l_{i,j} = \left\lfloor \frac{j}{p} \right\rfloor + 1 = k_{i,j}$.

Case 2: If $i \not\equiv 0 \pmod p$:

$$\left\lfloor \frac{i-1}{p} \right\rfloor = \left\lfloor \frac{i}{p} \right\rfloor,$$

$$\left\lfloor \frac{i+j-1}{p} \right\rfloor = \left\lfloor \frac{i}{p} \right\rfloor + \left\lfloor \frac{j}{p} \right\rfloor + \left\lfloor \frac{r_i + r_j - 1}{p} \right\rfloor;$$

hence

$$\left\lfloor \frac{i+j-1}{p} \right\rfloor = \left\lfloor \frac{i}{p} \right\rfloor + \left\lfloor \frac{j}{p} \right\rfloor \quad \text{if } r_i + r_j \leqslant p,$$

which may be written:

$$l_{i,j} = \left\lfloor \frac{j}{p} \right\rfloor = k_{i,j}$$

and

$$\left\lfloor \frac{i+j-1}{p} \right\rfloor = \left\lfloor \frac{i}{p} \right\rfloor + \left\lfloor \frac{j}{p} \right\rfloor + 1 \quad \text{if } r_i + r_j \geqslant p + 1$$

and then

$$l_{i,j} = \left\lfloor \frac{j}{p} \right\rfloor + 1 = k_{i,j}. \quad \square$$

Proof of Lemma 5. Let

$$Z = \{\text{factors of } A^1 \text{ of length } j \text{ beginning at an index} \equiv i \pmod p\}.$$

Let $Y = \{\text{factors of } A^2 \text{ of length } l_{i,j}\}$. The map which associates to a word m of Z the subword m' of Y which is composed with the letters of m of indices multiple of p is clearly bijective, hence the result holds. $\square$

We now give the second induction lemma.

Lemma 7. *If* $\hat{a}_1 \neq a_1$, *then* $\forall j \geqslant p$

$$\varphi_j^1(j) = P^2_{\text{even}}(l_{i,j}) + P^2_{\text{odd}}(l_{i,j}),$$

with

$$l_{i,j} = \left\lfloor \frac{j}{p} \right\rfloor + \begin{cases} 0 & \textit{if} \begin{cases} & r_i \neq 0 \textit{ and } r_i + r_j - 1 \leqslant p, \\ \textit{or} & \\ & r_i = r_j = 0, \end{cases} \\ 1 & \textit{otherwise.} \end{cases}$$

Proof. we recall that $l_{i,j}$ is the number of integers multiple of p in $[i, i+j-1]$. Let $\Gamma = \{\text{factors of } A^1 \text{ of length } j \text{ beginning at least once at an index} \equiv i (\text{mod}\, p)\}$. Define Θ_0 and Θ_1 by

$$\Theta_0 = \{\text{factrors of } A^2 \text{ of length } l_{i,j} \text{ beginning at least once at an even index}\},$$

and

$$\Theta_1 = \{\text{factors of } A^2 \text{ of length } l_{i,j} \text{ beginning at least once at an odd index}\}.$$

Let

$$\Theta = (\Theta_0 \times \{0\}) \cup (\Theta_1 \times \{1\}).$$

Define the map Φ as follows:

$$\Phi \to \Theta,$$

$$m \mapsto \left(m_{(p)}, \left\lfloor \frac{i}{p} \right\rfloor + 1 (\text{mod}\, 2) \right),$$

where $m_{(p)}$ is the subword of m formed of the letters of m of indices multiple of p.

This map is clearly bijective, hence the result. $\square$

We now give a general induction lemma.

General induction lemma. (i) *If* $\hat{a}_1 \neq a_1$ *then* $\forall j \geqslant p$,

$$\varphi^1(j) = (p - r_j)\left(P^2_{\text{even}}\left(\left\lfloor \frac{j}{p} \right\rfloor\right) + P^2_{\text{odd}}\left(\left\lfloor \frac{j}{p} \right\rfloor\right)\right) + r_j\left(P^2_{\text{even}}\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right) + P^2_{\text{odd}}\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right)\right),$$

and

$$\varphi^1(j+1) - \varphi^1(j) = \left(P^2_{\text{even}}\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right) - P^2_{\text{even}}\left(\left\lfloor \frac{j}{p} \right\rfloor\right)\right) + \left(P^2_{\text{odd}}\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right) - P^2_{\text{odd}}\left(\left\lfloor \frac{j}{p} \right\rfloor\right)\right),$$

r_j *being the remainder in the euclidian division of* j *by* p.

(ii) *If* $\hat{a}_1 = a_1$ *then*

$$\forall j \geqslant p,\ \varphi^1(j) = (p - r_j)P^2\left(\left\lfloor \frac{j}{p} \right\rfloor\right) + r_j P^2\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right)$$

and

$$\varphi^1(j+1) - \varphi^1(j) = P^2\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right) - P^2\left(\left\lfloor \frac{j}{p} \right\rfloor\right),$$

where r_j *is the remainder in the euclidian division of* j *by* p.

(iii) *If* p *is odd then* $\forall a_1 \in A^{p-1}$, *one has* $\forall j \geqslant p$:

$$P^1_{\text{odd}}(j+1) + P^1_{\text{even}}(j+1) - P^1_{\text{odd}}(j) - P^1_{\text{even}}(j) = \left(P^2_{\text{odd}}\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right) - P^2_{\text{odd}}\left(\left\lfloor \frac{j}{p} \right\rfloor\right)\right) + \left(P^2_{\text{even}}\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right) - P^2_{\text{even}}\left(\left\lfloor \frac{j}{p} \right\rfloor\right)\right).$$

Proof. (i) Using Lemma 7, the proof is the same as for the following point.

(ii) One has, using Lemma 5, $\varphi^1(j) = \sum_{i=1}^{i=p} \varphi_i^2(l_{i,j})$. By counting separately the values of i giving the same number $l_{i,j}$, one has:

$$\varphi^1(j) = (p - r_j)\varphi^2\left(\left\lfloor \frac{j}{p} \right\rfloor\right) + r_j\varphi^2\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right).$$

This last equality implies the result.

(iii) One has, for odd p and for $j \geqslant p$, the following.

If $j \equiv 0 \pmod p$ and $j \equiv 0 \pmod 2$, then

$$P^1_{\text{odd}}(j) + P^1_{\text{even}}(j) = (p - r_j + 3)\left(P^2_{\text{odd}}\left(\left\lfloor \frac{j}{p} \right\rfloor\right) + P^2_{\text{even}}\left(\left\lfloor \frac{j}{p} \right\rfloor\right)\right) + (r_j - 3)\left(P^2_{\text{odd}}\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right) + P^2_{\text{even}}\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right)\right).$$

If $j \not\equiv 0 \pmod{p}$ and $j \equiv 0 \pmod{2}$, then

$$P^1_{\text{odd}}(j) + P^1_{\text{even}}(j) = (p - r_j + 2)\left(p^2_{\text{odd}}\left(\left\lfloor \frac{j}{p} \right\rfloor\right) + P^2_{\text{even}}\left(\left\lfloor \frac{j}{p} \right\rfloor\right)\right)$$
$$+ (r_j - 2)\left(P^2_{\text{odd}}\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right) + P^2_{\text{even}}\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right)\right).$$

If $j \equiv 0 \pmod{p}$ and $j \not\equiv 0 \pmod{2}$, then

$$P^1_{\text{odd}}(j) + P^1_{\text{even}}(j) = (p - r_j + 4)\left(P^2_{\text{odd}}\left(\left\lfloor \frac{j}{p} \right\rfloor\right) + P^2_{\text{even}}\left(\left\lfloor \frac{j}{p} \right\rfloor\right)\right)$$
$$+ (r_j - 4)\left(P^2_{\text{odd}}\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right) + P^2_{\text{even}}\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right)\right).$$

If $j \not\equiv 0 \pmod{p}$ and $j \not\equiv 0 \pmod{2}$, then

$$P^1_{\text{odd}}(j) + P^1_{\text{even}}(j) = \, = (p - r_j + 3)\left(P^2_{\text{odd}}\left(\left\lfloor \frac{j}{p} \right\rfloor\right) + P^2_{\text{even}}\left(\left\lfloor \frac{j}{p} \right\rfloor\right)\right)$$
$$+ (r_j - 3)\left(P^2_{\text{odd}}\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right) + P^2_{\text{even}}\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right)\right).$$

In each case the point (iii) holds. This concludes the proof of the lemma. □

3. A necessary and sufficient condition for the number of special factors of a p-paperfolding sequence to be automatic

In this part we study the number of *special factors*: a factor w of a p-paperfolding sequence is called special if and only if $w0$ and $w1$ are both factors of this sequence. The number of special factors of length n is given by $P(n + 1) - P(n)$, if P is the complexity function of this paperfolding sequence (this holds for any binary sequence).

Proposition 8. *If there exists an integer j_0 such that $\hat{a}_{j_0} \neq a_{j_0}$, then for any $i \geqslant j_0$ and for any $j \geqslant p^i$ one has*

$$P^1(j+1) - P^1(j) = P^{i+1}_{\text{odd}}\left(\left\lfloor \frac{j}{p^i} \right\rfloor + 1\right) - P^{i+1}_{\text{odd}}\left(\left\lfloor \frac{j}{p^i} \right\rfloor\right)$$
$$+ P^{i+1}_{\text{even}}\left(\left\lfloor \frac{j}{p^i} \right\rfloor + 1\right) - P^{i+1}_{\text{even}}\left(\left\lfloor \frac{j}{p^i} \right\rfloor\right).$$

If $\forall j_0 \in \mathbb{N}^*$, $\hat{a}_{j_0} = a_{j_0}$, then for any integer i and any integer j such that $j \geqslant p^i$ one has

$$P^1(j+1) - P^1(j) = P^{i+1}\left(\left\lfloor \frac{j}{p^i} \right\rfloor + 1\right) - P^{i+1}\left(\left\lfloor \frac{j}{p^i} \right\rfloor\right).$$

Proof. This proposition is an easy consequence of the Synchronization Lemma and of the General Induction Lemma. □

Theorem 9. *Let m be a word of length $p-1$. Let C^1_m be the complexity function of the word: $(m0\hat{m}0m1\hat{m}1)^\infty$, C^2_m be the sum $P_{\text{odd}}+P_{\text{even}}$ of the same word. Let $\zeta^1(m)=(C^1_m(2)-C^1_m(1),\dots,C^1_m(p')-C^1_m(p-1))$ and $\zeta^2(m)=(C^2_m(2)-C^2_m(1),\dots,C^2_m(p)-C^2_m(p-1))$.*

Let a_n be a sequence of folding instructions. Let $n_0\in\mathbb{N}\cup\{+\infty\}$, such that $\forall j\geqslant n_0$, $a_i=\hat{a}_i$. Let ζ be the map $\zeta(n)=\zeta^1(a_n)$ if $n<n_0$ and $\zeta(n)=\zeta^2(n)$ otherwise. Then the sequence $(P^1(n+1)-P^1(n))$ is p-automatic if and only if the sequence (ζ_n) is ultimately periodic.

Proof. The complexity function P of any p-paperfolding sequence verifies $P(1)=2$ and $P(2)=4$, $P_{\text{even}}(1)=2$, $P_{\text{odd}}(1)=2$, $P_{\text{even}}(2)=4$ and $P_{\text{odd}}(2)=4$. Hence the sequence (A^n) and $(a_n0\hat{a}_n0a_n1\hat{a}_n1)^\infty$ have the same complexity function for arguments smaller than or equal to p.

The p-kernel of a sequence $u=u[n]_{n\in\mathbb{N}^*}$ is the set of sequences of the form $u[p^kn+r]_{n\in\mathbb{N}^*}$ with $k\in\mathbb{N}^*$ and $0\leqslant r\leqslant p^k-1$ (this terminology has been introduced by Salon in the multi-index case in [12, 13]). A sequence u is p-automatic if and only if its p-kernel is finite (see [6]). Proposition 8 helps us to compute easily the p-kernel of this sequence and to realize that it is finite if and only if the sequence ζ_n is ultimately periodic. □

4. Computation of the complexity function

Theorem 10. *If the complexity function of a p-paperfolding sequence is ultimately affine, then it is ultimately linear.*

Proof. If the complexity function P^1 is ultimately affine, then the functions $P^i(i\geqslant 1)$ are all ultimately affine, say $P^i(n)=a_in+b_i$. Furthermore, the a_i's are eventually equal.

Let $N_0\in\mathbb{N}\cup\{\infty\}$ such that for any $n\geqslant N_0$ $\hat{a}_n=a_n$.

If $N_0=+\infty$ then all the a_i's are the number 4. One has $\varphi^1(n+1)-\varphi^1(n)=(P^i_{\text{odd}}+p^i_{\text{even}})(\lfloor n/p^i\rfloor+1)-(P^i_{\text{odd}}+p^i_{\text{even}})(\lfloor n/p^i\rfloor)$ for i as large as we wish and $n\geqslant p^i$. (i is such that $\hat{a}_i\neq a_i$). For $n=p^i$ one has $\varphi^1(n+1)-\varphi^1(n)=(P^i_{\text{odd}}+p^i_{\text{even}})(2)-(P^i_{\text{odd}}+p^i_{\text{even}})(1)$ (General Induction Lemma). Hence $\varphi^1(n+1)-\varphi^1(n)=4$. But $\varphi(n)=P^1(n)$ if $n\geqslant p(p+2)-1$, and $P^1(n)=a_1n+b_1$ if n is large enough. So $P^1(n)=4n+b_1$ for n large enough. We have then $\varphi^1(n)-\varphi(p^i)=4(n-p^i)$. But $\varphi^1(p^i)=4p^i$. (Lemma 7). Hence $P^1(n)=4n$.

If $N_0<+\infty$ then let $i\geqslant N_0$ be an integer. We have $a_i=2$. One has (General Induction Lemma) $\varphi^i(n+1)-\varphi^i(n)=p^{i+j}(\lfloor n/p^j\rfloor+1)-P^{i+j}(\lfloor n/p^j\rfloor)$ for $j\geqslant 0$. Taking $n=p^j$ with j large enough, we have $\varphi^i(n+1)-\varphi^i(n)=2$ for n large enough.

But we have $\varphi^i(n) = P^i(n)$ for $n \geqslant p(p+2) - 1$ (Synchronization Lemma) and $P^i(n) = a_i n + b_i$ for n large enough. We conclude that $a_i = 2$. In this case $\varphi^1(n) - \varphi^1(p^j) = 2(n - p^j)$ and $\varphi^1(p^j) = 2p^j$ (Lemma 5). □

Theorem 11. *There exist words a in* $\mathscr{A}^{p-1}$ *such that* $\zeta^1(a) = (2, 2, \ldots, 2)$ *and* $\zeta^2(a) = (4, 4, \ldots, 4)$. *The complexity function of* A^1 *is ultimately affine (hence ultimately linear) if and only if* $\zeta(a_n) = (2, 2, \ldots, 2)$ *for n large enough or* $\zeta(a_n) = (4, 4, \ldots, 4)$ *for n large enough.*

Proof. For any number p and any sequence a_n of folding instructions, $(n \in \mathbb{N}^*)$, the complexity function verifies $P^1(1) = 2$, $P^1(2) = 4$, $P^1_{\text{odd}}(1) = 2$, $P^1_2(2) = 4$, $P^1_{\text{even}}(1) = 2$ and $P^2_{\text{even}}(2) = 4$. The word $a = 0^{p-1}$ verifies $\zeta^1(a) = (2, 2, \ldots, 2)$ and $\zeta^2(a) = (4, 4, \ldots, 4)$. The complexity function of A^1 is ultimately affine if and only if $(P^1(n+1) - P^1(n))$ is constant for n large enough. If one wants this sequence to be ultimately constant, one must have $\zeta(a_n) = (2, 2, \ldots, 2)$ for n large enough or, $\zeta(a_n) = (4, 4, \ldots, 4)$ for n large enough (Proposition 8). On the other hand, this condition is sufficient using the same proposition. □

Theorem 12. *If* $p = 2$, *then* $P^1(n) = 4n$ *for n large enough.*

If $p = 3$ *and if* $\forall i \in \mathbb{N}^*$ $\hat{a}_i = a_i$, *then* $P^1(n) = 2n$ *for* $n \geqslant 1$.

If $p = 3$ *and if there exists i such that* $\hat{a}_i \neq a_i$, *then* $P^1(n) = 4n$ *for n large enough.*

If $p \geqslant 4$, *then there exist folding instructions sequences leading to strictly sub-affine complexity functions, and others leading to affine (hence linear) complexity functions. In this last case, if* $\forall i \in \mathbb{N}^*$ $\hat{a}_i = a_i$, *then* $P^1(n) = 2n$ *for n large enough, and if there exists i such that* $\hat{a}_i \neq a_i$, *then* $P^1(n) = 4n$ *for n large enough.*

Proof. If $p = 2$ (resp. $p = 3$) then $\forall a \in \mathscr{A}^{p-1}, \zeta(a) = (2)$ (resp. $(2, 2)$). Hence the generalized 2-paperfolding and the generalized 3-paperfolding sequences have ultimately linear complexities (Theorem 11).

If $p = 2$, $\forall a \in \mathscr{A}^{p-1}, \hat{a} \neq a$, $P^1_{\text{odd}}(1) = 2$, $P^1_{\text{even}}(1) = 2$, $P^1_{\text{odd}}(2) = 4$, $P^1_{\text{even}}(2) = 4$ and hence $P^1(n+1) - P^1(n) = 4$ for n large enough, which implies $P^1(n) = 4n$ for n large enough (Proposition 8 and Theorem 10).

If $p = 3$ and $\forall i \in \mathbb{N}^*$ $\hat{a}_i = a_i$, since $P^1(1) = 2$, $P^1(2) = 4$ and $P^1(3) = 6$ one has $P^1(n) = 2n$ for n large enough, (Proposition 8 and Theorem 10). If there exists i such that $\hat{a}_i \neq a_i$, since $P^1_{\text{odd}}(1) = 2$, $P^1_{\text{even}}(1) = 2$, $P^1_{\text{odd}}(2) = 4$, $P^1_{\text{even}}(2) = 4$, $P^1_{\text{odd}}(3) = 6$, $P^1_{\text{even}}(3) = 6$, one has $P^1(n) = 4n$ for n large enough (Proposition 8 and Theorem 10).

If $p \geqslant 4$, the cardinality of the image of ζ is strictly greater than 1, and there are folding instructions sequences leading to strictly subaffine complexity functions, (because for instance $a = 0^{p-1}$ and $a = 01^{p-2}$ have distinct images by ζ. Indeed, $(0^p 1^{p-1} 0^p 1^{p+1})^\infty$ verifies $P(1) = 2$, $P(2) = 4$, $P(3) = 6$, hence its image by ζ begins by the numbers 2, 2 while $(01^{p-2} 0^{p-1} 1001^{p-1} 0^{p-2} 11)^\infty$ verifies $P(1) = 2$, $P(2) = 4$, $P(3) = 8$, and its image by ζ begins with the numbers 2, 4).

The complexity function is ultimately affine if and only if

$$\zeta(a_n) = (2, 2, \ldots, 2)$$

for n large enough or

$$\zeta(a_n) = (4, 4, \ldots, 4)$$

for n large enough. In this case, we conclude as in the case $p = 3$. □

5. The language generated by the p-paperfolding sequences for a fixed p is not context-free

In [4], the authors considered the language of all factors of all 2-paperfolding sequences. We will consider in the same way the language of all the p-paperfolding sequences for a fixed p.

Lemma 13. *Let w be a factor of A^1. Then w^n is not a factor of A^1 as soon as $n \geqslant p(p+2)$.*

Lemma 14. *If w^n is a factor of A^1 with $n \geqslant p(p+2)$ then the length of w, denoted by $|w|$, is a power of p.*

Proof. Let

$$w^n = A^1[j] \ldots A^1[j + n|w| - 1].$$

One has

$$w_1 = A^1[j] \ldots A^1[j + (n-1)|w| - 1],$$

which is equal to the factor

$$w_2 = A^1[j + |w|] \ldots A^1[j + n|w| - 1],$$

these two factors being equal to w^{n-1}. But $|w^{n-1}| \geqslant p(p+2) - 1$. These two factors must hence begin at a same index modulo p, (Synchronization Lemma), hence $|w|$ is multiple of p.

Now the subword w_2 formed with the letters of w of index multiple of p is a factor of (the p-paperfolding sequence) A^2, so is w_2^n. If w_2 is not the empty word, its length is a multiple of p, and the length of w is a multiple of p^2. By an immediate induction, one shows that $|w|$ is a power of p.

Proof of Lemma 13. Let us suppose that $|w| = p^k$. The word w^n is a factor of A^1. Let w_1^n be the subword formed of the letters of w appearing at an index mutiple of p. This word w_1^n is a factor of A^2. Repeating this operation $k-1$ times, we finally find a factor of A^{k+1} of the form α^n, α belonging to $\mathscr{A}$, and $n \geqslant p(p+2)$. This is in contradiction with the synchronization lemma. □

Theorem 15. *For any p, the language $\mathscr{L}$ of all factors of all p-paperfolding sequences is not context-free.*

Proof. we will use here the same idea as in [4, 5] (see also [9]). The pumping lemma for context-free languages (see [7]) easily implies that in any infinite context-free language, there are arbitrarily large powers. Lemma 13 hence implies that the language $\mathscr{L}$ of all factors of all p-paperfolding sequences is not context-free. □

Theorem 16. *The generating series of the language $\mathscr{L}$ is transcendental.*

Remark 17. This last proposition implies that, if the complementary language $\{0,1\}^*\backslash\mathscr{L}$ were context-free, then it would be ambiguously context-free (Chomsky–Schützenberger's Theorem).

To prove this theorem, we need the following lemma.

Lemma 18. *The number $\mathscr{N}_p(j)$ of factors of length j appearing in at least one p-paperfolding sequence verifies:*

$$\mathscr{N}_p(j) = 2^j \quad \text{if } j \leqslant p+1;$$

$$\mathscr{N}_p(j) = 2^{p-1}\left((p-r_j)\,\mathscr{N}_p\left(\left\lfloor \frac{j}{p} \right\rfloor\right) + r_j \mathscr{N}_p\left(\left\lfloor \frac{j}{p} \right\rfloor + 1\right)\right) \text{for } j \geqslant p(p+2)-1.$$

Proof. A factor $u[1]u[2]\ldots u[j]$ is a p-paperfolding word if and only if there exists $j_0 \in [1,p]$, such that

(a) the words $m(k) = u[j_0+kp+1]u[j_0+kp+2]\ldots u[\alpha_k]$ (with $j_0+kp+1 \leqslant j$ and $\alpha_k = \inf(j_0+kp+p-1, j)$ verify $m(k) = \hat{m}(k+1)$,

(b) the word $u[j_0]u[j_0+p]\ldots u[j_0+\beta p]$ is a p-paperfolding factor with β the largest integer such that $j_0+\beta p \leqslant j$.

If $j \leqslant p+1$ it is then clear that any factor of length j is a p-paperfolding factor.

If $j \geqslant p(p+2)-1$ then the synchronization lemma remains true for all factors of all p-paperfolding sequences and $\mathscr{N}_p(j)$ may be obtained by summing $\mathscr{N}_p^{\nu}(j)$ on the p-paperfolding factors beginning at indices congruent to ν modulo p, for $1 \leqslant \nu \leqslant p$.

But, $\mathscr{N}_p^{\nu}(j) = 2^{p-1}\mathscr{N}_p(l_{\nu,j})$ (same proof as for the generalized induction lemma). The multiplicative factor comes from the choice that we have at each step for the folding instruction. Summing on ν, one finds the result. □

Let us prove now Theorem 16. Let $\mathscr{F}(X) = \sum_{k=0}^{k=+\infty} \mathscr{N}_p(k)X^k$ the generating series of the language $\mathscr{L}$. Let us consider the following polynomials:

$$P_1(X) = \sum_{k=0}^{k=p^2+2p-2} \mathscr{N}_p(k)X^k,$$

$$P_2(X) = \sum_{k=0}^{k=p+1} \mathscr{N}_p(k)X^k,$$

and

$$P_3(X) = \sum_{k=0}^{k=p+2} \mathcal{N}_p(k)X^k.$$

One has

$$\mathcal{F}(X) = P_1(X) + \sum_{r=0}^{r=p-1} \sum_{k=p+2}^{k=+\infty} 2^{p-1}((p-r)\mathcal{N}_p(k) + r\mathcal{N}_p(k+1))X^r(X^p)^k$$
$$+ 2^{p-1}(\mathcal{N}_p(p+1) + (p-1)\mathcal{N}_p(p+2))X^{p-1}(X^p)^{p+1}.$$

Let us consider

$$\mathcal{F}(X) = P_1(X) + \sum_{r=0}^{r=p+1} 2^{p-1}((p-r)X^r(\mathcal{F}(X^p) - P_2(X^p)))$$
$$+ \frac{r2^{p-1}X^r}{X^p}(\mathcal{F}(X^p) - P_3(X^p)).$$

The degree of the polynomial

$$P_1(X) - 2^{p-1}\sum_{r=1}^{r=p-1}(p-r)X^rP_2(X^p) - 2^{p-1}\frac{rX^r}{X^p}P_3(X^p)$$
$$+ 2^{p-1}(\mathcal{N}_p(p+1) + (p-1)\mathcal{N}_p(p+2))X^{p-1}(X^p)^{p+1}$$

is smaller than or equal to $p(p+2)-1$. Let us compute its leading coefficient. The polynomial P_1 does not contribute to this coefficient because it degree is exactly p^2+2p-2. The coefficient of the term of degree p^2+2p-1 is hence

$$-2^{p-1}\mathcal{N}_p(p+1) - 2^{p-1}(p-1)\mathcal{N}_p(p+2) + 2^{p-1}(\mathcal{N}_p(p+1)$$
$$+ (p-1)\mathcal{N}_p(p+2)) = 0.$$

Hence the degree of the polynomial

$$P_1(X) - 2^{p-1}\sum_{r=1}^{r=p-1}(p-r)X^rP_2(X^p) - 2^{p-1}\frac{X^r}{X^p}P_3(X^p)$$
$$+ 2^{p-1}(\mathcal{N}_p(p+1) + (p-1)\mathcal{N}_p(p+2))X^{p-1}(X^p)^{p+1}$$

is smaller than or equal to $p(p+2)-2$.

Let us suppose firstly that this degree is greater than or equal to $p(p+1)$. Let us suppose moreover that $\mathcal{F}$ is a rational function of degree α. Since $d \not\equiv -1 \pmod p$, $(p(p+1) \leqslant d \leqslant p(p+2)-2)$, it is impossible to have equality $\alpha p + p - 1 = d$, and one has hence $\alpha = \max(\alpha p + p - 1, d)$.

If $d > \alpha p + p - 1$, then $\alpha = d$, which means that $\alpha > p\alpha + p - 1$, which is impossible.

Hence, $\alpha p + p - 1 > d$ and $\alpha = \alpha p + p - 1$, which means that $\alpha = -1$ and $\alpha > d$, which is clearly impossible.

Finally, if $d \geqslant p(p+1)$ then $\mathcal{F}$ cannot be a rational function.

Let us prove now that the degree of the polynomial

$$P_1(X) - 2^{p-1}\sum_{r=1}^{r=p-1}(p-r)X^r P_2(X^p) - 2^{p-1}\frac{rX^r}{X^p}P_3(X^p)$$
$$+\, 2^{p-1}(\mathcal{N}_p(p+1) + (p-1)\mathcal{N}_p(p+2))X^{p-1}(X^p)^{p+1}$$

is actually greater than or equal to $p(p+1)$.

For that, it is sufficient to exhibit a p-paperfolding factor beginning at two different (modulo p) indices (not necessarily in the same p-paperfolding sequence).

We can do that the following way.

- If p is even: let u be the factor

$$u = \underbrace{0^p 1^p \ldots 0^p}_{p+1 \text{ paquets}}.$$

It can be obtained either as

$$\underbrace{m_1 0 \hat{m}_1 1 \ldots m_1 0}_{m_1 \text{ repeated } \lfloor p/2 \rfloor + 1 \text{ times}}$$

or as

$$\underbrace{0 m_1 1 \hat{m}_1 \ldots 0 m_1}_{m_1 \text{ repeated } \lfloor p/2 \rfloor + 1 \text{ times}}$$

with $m_1 = 0^{p-1}$.

- If p is odd: let u be the factor

$$u = \underbrace{010101 \ldots 01}_{p(p+1) \text{ letters}}$$

It can be obtained either as

$$\underbrace{m_2 0 m_2 0 \ldots m_2 0.}_{m_2 \text{ repeated } p+1 \text{ times}}$$

with

$$m_2 = \underbrace{0101 \ldots 01}_{p-1 \text{ letters}} \, (m_2 = \hat{m}_2)$$

or as

$$\underbrace{0 m_3 0 m_3 \ldots 0 m_3}_{m_3 \text{ repeated } p+1 \text{ times}}$$

with

$$m_3 = \underbrace{1010\ldots10}_{p-1 \text{ letters}}$$

Finally, $\mathscr{F}(X)$ is not a rational function.

Now $\mathscr{F}$ is an entire power series with convergence radius 1 and integer coefficients. The theorem of Polya–Carlson (see [10]) asserts that $\mathscr{F}$ is either a rational function or a transcendental function.

Since it is not rational, it is trancendental.

Acknowledgements

The author would like to thank warmly Jean-Paul Allouche and Thomas Downarrowicz for comments on an earlier version and Désiré Razafy Adriamampianina for interesting discussions.

References

[1] J.-P Allouche, The number of factors in a paperfolding sequence, *Bull. Austral. Math. Soc.* **46** (1992) 23–32.

[2] J.-P. Allouche and R. Bacher, Toeplitz sequences, paperfolding, Towers of Hanoi, and progression-free sequences of integers, *Enseign. Math.* **38** (1992) 315–327.

[3] J.-P. Allouche, Sur la complexité des suites infines, *Bull. Belg. Math. Soc.* **1** (1994) 133–143.

[4] J.-P. Allouche and M. Bousquet-Mélou, Canonical positions for the factors in the paperfolding sequences, *Theoret. Comput. Sci.* **129** (1994) 263–278.

[5] J.-P. Allouche and M.' Bousquet-Mélou, Facteurs de suites de Rudin-Shapiro généralisées, *Bull. Belg. Math. Soc.* (1994) 145–164.

[6] Cobham, Uniform Tag Sequences, *Math. Systems Theory* **6** (1972) 164–192.

[7] J.E. Hopcroft and J.D. Ullman, *Introduction to Automata Theory, Languages and Computation* (*Addison-Wesley*, Reading, MA 1979).

[8] M. Koskas, Complexité de suites de Toeplitz, Discrete Math., to appear

[9] S. Lehr, A result about languages concerning paperfolding sequences, *Math. Systems Theory* **25** (1992) 309–313.

[10] G. Polya, Collected Papers, Vol. 1, in: R.P. Boas, ed., Singularities of analytic functions, (*M.I.T. Press, Cambridge*, 1974).

[11] D. Razafy Andriamampianina, Le p-pliage de papier, *Ann. Fac. Sci. Toulouse Math.* **5** (1989) 410–414.

[12] O. Salon, Suites automatiques à multi-indices, *Séminaire de Théorie des Nombres de Bordeaux*, Exposé 4 (1986–1987) 4-01–4-27 (Followed by an appendix of J. Shallit).

[13] O. Salon, Suites automatiques à multi-indices at algébricité, *C.R. Acad. Sci. Paris Sér I. Math.* **305** (1987) 501–504.

[14] Z.X. Wen and Z.Y. Wen, Somes studies on the (p, q)-type sequences, *Theoret. Comput. Sci.* **94** (1992) 373–393.