Abelian Anti-Powers in Infinite Words

Gabriele Fici
Dipartimento di Matematica e Informatica, Università di Palermo, Italy
gabriele.fici@unipa.it

Mickael Postic
Institut Camille Jordan, Université Claude Bernard Lyon 1, Lyon, France
postic@math.univ-lyon1.fr

Manuel Silva
Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Lisbon, Portugal
mnasilva@gmail.com

Abstract

We introduce and study the notion of an abelian anti-power in the context of combinatorics on words. An abelian anti-power of order k (or simply an abelian k–anti-power) is a concatenation of k consecutive words of the same length having pairwise distinct Parikh vectors. This definition therefore generalizes to the abelian setting the notion of a k–anti-power, as introduced in [G. Fici et al., *Anti-powers in infinite words*, J. Comb. Theory, Ser. A, 2018], that is a concatenation of k pairwise distinct words of the same length. In particular, we deal with the question to determine whether a word contains abelian k–anti-powers for arbitrarily large k. A word with bounded abelian complexity clearly cannot contain abelian anti-powers of every order. We show that the Sierpìnski word (whose abelian complexity grows logarithmically) does not contain abelian 11–anti-powers. Another question is to find words with low factor complexity that contain both abelian powers and abelian anti-powers of every order. We show that all paperfolding words have this property.

1 Introduction

Many of the classical definitions in combinatorics on words (e.g., period, power, factor complexity, etc.) have a counterpart in the abelian setting, though they may not enjoy the same properties.

Recall that the Parikh vector $P(w)$ of a word w over a finite ordered alphabet $\mathbb{A} = \{a_1, a_2, \ldots, a_{|\mathbb{A}|}\}$ is the vector whose i-th component is equal to the number of occurrences of the letter a_i in w, $1 \leq i \leq |\mathbb{A}|$. For example, the Parikh vector of $w = abbca$ over $\mathbb{A} = \{a, b, c\}$ is $P(w) = (2, 2, 1)$. This notion is at the basis of the abelian combinatorics on words, where two words are considered equivalent if and only if they have the same Parikh vector.

The fundamental result of Morse and Hedlund [7] (an infinite word is aperiodic if and only if its factor complexity is unbounded) does not hold anymore in the case of the abelian complexity (the function that counts the number of distinct Parikh vectors of factors of length n for each n), as there exist aperiodic words with bounded abelian complexity. In fact, Richomme et al. [8] observed that if a word has bounded abelian complexity, then it contains abelian powers of any order — an abelian power of order k is a concatenation of k words having the same Parikh vector. However, this is not a characterization of words with bounded abelian complexity. Madill and Rampersad proved that the regular paperfolding word has unbounded abelian complexity [6], and Štěpán Holub proved that it contains abelian powers of every order [5]. Actually, Holub proved that all paperfolding words have this property.

In a recent paper [4], the first and the third author, together with Antonio Restivo and Luca Zamboni, introduced the notion of an anti-power. An *anti-power of order k*, or simply a *k–anti-power*, is a concatenation of k consecutive pairwise distinct words of the same length. E.g., *aabaaabbbaba* is a 4–anti-power.

In [4], it is proved that the existence of powers of every order or anti-powers of every order is an unavoidable regularity for infinite words:

Theorem 1. *[4] Every infinite word contains powers of every order or anti-powers of every order.*

Note that in the previous statement there is no hypothesis on the alphabet size.

In this paper, we extend the notion of an anti-power to the abelian setting.

Definition 2. An *abelian anti-power of order k*, or simply an *abelian k–anti-power*, is a concatenation of k consecutive words of the same length having pairwise distinct Parikh vectors.

For example, $aabaaabbbabb$ is an abelian 4–anti-power. Notice that an abelian k–anti-power is a k–anti-power but the converse does not necessarily holds (which is dual to the fact that a k–power is an abelian k–power but the converse does not necessarily holds).

We think that an analogous of Theorem 1 may still hold in the case of abelian anti-powers, but unfortunately the proof of Theorem 1 does not seem to be generalizable to the abelian setting.

Problem 1. Does every infinite word contain abelian powers of every order or abelian anti-powers of every order?

Clearly, if a word has bounded abelian complexity, then it cannot contain abelian anti-powers of every order. However, we show in this paper that the converse is not true. Indeed, we prove that the Sierpiǹski word, which is the fixed point starting with a of the substitution $(a \to aba, b \to bbb)$, does not contain abelian 11–anti-powers. The Sierpiǹski word has logarithmic abelian complexity (by construction) and contains abelian powers of every order (as it contains arbitrarily long blocks of b's).

An infinite word can contain both abelian powers of every order and abelian anti-powers of every order. This is the case, for example, of any word with full factor complexity. However, finding a class of words with low factor complexity satisfying this property seems a more difficult task. Indeed, most of the well-known examples of aperiodic words (Thue-Morse, Sturmian words, etc.) have bounded abelian complexity, hence they cannot contain abelian anti-powers of every order — whereas, by the aforementioned remark of Richomme et al. [8], they contain abelian powers of every order. Building upon the theory that Štěpán Holub developed to prove that all paperfolding words contain abelian powers of every order [5], we prove that all paperfolding words contains also abelian anti-powers of every order.

2 Sierpiǹski Word

Recall that the Sierpiǹski word s is the fixed point starting with a of the substitution

$$\begin{aligned} \sigma : a &\to aba \\ b &\to bbb \end{aligned}$$

so that the word s begins as follows:

$$ababbbababbbbbbbbbababbbabab^{27}a\cdots$$

Therefore, s can be obtained as the limit, for $n \to \infty$, of the sequence of words $(s_n)_{n\geq 0}$ defined by: $s_0 = a$, $s_{n+1} = s_n b^{3^n} s_n$ for $n \geq 1$. Notice that for every n one has $|s_n| = 3^n$.

Theorem 3. *The Sierpiǹski word s does not contain 11–anti-powers, hence it does not contain abelian 11–anti-powers.*

Proof. Suppose that s contains an 11–anti-power $u = u_1u_2\cdots u_{11}$, of length $11m$. Let us then consider the first occurrence of u in s. Let n be the smallest integer such that u occurs in $s_{n+1}b^{3^{n+1}}$ but not in $s_nb^{3^n}$.

Let us first suppose that no u_i is equal to b^m for some i. Then $u_1\cdots u_{10}$ is a factor of $s_{n+1} = s_nb^{3^n}s_n$, so $10m < 3^{n+1}$ hence $m < 3^{n-1}$. Then, by minimality of n, there are only two possible cases: either u_1 starts before the block b^{3^n}, or u_1 starts in the block b^{3^n} and ends in s_n.

In the first case, by minimality of n, u ends after the block b^{3^n}, and since no u_i equals b^m, we get $2m > 3^n$, which is in contradiction with $m < 3^{n-1}$.

If u_1 starts in the block b^{3^n} and ends in s_n, $u_2 \cdots u_{10}$ is a factor of $s_n = s_{n-1} b^{3^{n-1}} s_{n-1}$ and so $9m < 3^n$ hence $m < 3^{n-2}$. By minimality of n, u_{11} ends after the block $b^{3^{n-1}}$. Again, since no u_i equals b^m, we get $2m > 3^{n-1}$, which is in contradiction with $m < 3^{n-2}$.

Let us then suppose that $u_{11} = b^m$, so that $u_1 \cdots u_9$ is a factor of s_{n+1}. The same reasoning as before holds, since $(9m < 3^{n+1}) \Rightarrow (m < 3^{n-1})$ and $(8m < 3^n) \Rightarrow (2m < 3^{n-1})$. If $u_1 = b^m$, $u_2 \cdots u_{10}$ is a factor of s_n with no $u_i = b^m$ and we can again apply the same reasoning.

Finally, suppose that $u_i = b^m$ with $i \neq 1$ and $i \neq 11$. Hence, $u_1 \cdots u_{10}$ is a factor of $s_{n+1} = s_n b^{3^n} s_n$, and $10m < 3^{n+1}$. If u_1 starts before the block b^{3^n} (and ends after by minimality of n), we get $3m > 3^n$ since otherwise u would contain two blocks b^m. If u_1 does not start before the block b^{3^n}, then by minimality of n it starts in this block, so $u_2 \cdots u_{10}$ is a factor of $s_n = s_{n-1} b^{3^{n-1}} s_{n-1}$ which ends after the block $b^{3^{n-1}}$, again by minimality of n. This shows that $9m < 3^n$, and at the same time $3m > 3^{n-1}$, which produces a contradiction. $\square$

3 Paperfolding Words

In what follows, we recall the combinatorial framework for dealing with paperfolding words introduced in [5], although we use the alphabet $\{0, 1\}$ instead of $\{1, -1\}$.

A paperfolding word is the sequence of ridges and valleys obtained by unfolding a sheet of paper which has been folded infinitely many times. 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 factor complexity $c(n)$, and that $c(n) = 4n$ for $n \geq 7$ [1].

The *regular paperfolding word*

$$\mathbf{p} = 00100110001101100010011100110110 \cdots$$

is obtained by folding at each step in the same way. It can be defined as a Toeplitz word (see [2] for a definition of Toeplitz words) as follows: Consider the infinite periodic word $\gamma = (0?1?)^\omega$, defined over the alphabet $\{0, 1\} \cup \{?\}$. Then define $p_0 = \gamma$ and, for every $n > 0$, p_n as the word obtained from p_{n-1} by replacing the symbols ? with the letters of γ. So,

$$\begin{aligned}
p_0 &= 0?1?0?1?0?1?0?1?0?1?0?1?0?1?0?1? \cdots, \\
p_1 &= 001?011?001?011?001?011?001? \cdots, \\
p_2 &= 0010011?0011011?0010011?0011 \cdots, \\
p_3 &= 001001100011011?001001110011 \cdots,
\end{aligned}$$

etc. Thus, $\mathbf{p} = \lim_{n\to\infty} p_n$, and hence $\mathbf{p}$ does not contain occurrences of the symbol ?.

More generally, one can define a paperfolding word $\mathbf{f}$ by considering the two infinite periodic words $\gamma = (0?1?)^\omega$ and $\bar{\gamma} = (1?0?)^\omega$. Then, let $\boldsymbol{b} = b_0 b_1 \cdots$ be an infinite word over $\{0, 1\}$, called *the sequence of instructions*. Define $(\gamma_n)_{n\geq 0}$ where, for every n, $\gamma_n = \gamma$ if $b_n = 0$ or $\gamma_n = \bar{\gamma}$ if $b_n = 1$. The paperfolding word $\mathbf{f}$ *associated with* $\boldsymbol{b}$ is the limit of the sequence of words f_n defined by $f_0 = \gamma_0$ and, for every $n > 0$, f_n is obtained from f_{n-1} by replacing the symbols ? with the letters of γ_n.

Recall that every positive integer i can be uniquely written as $i = 2^k(2j + 1)$, where k is called the *order* of i (a.k.a. the 2-adic valuation of i), and $(2j + 1)$ is called the *odd part* of i. One can verify that the previous definition of $\mathbf{f}$ is equivalent to the following: for every $i = 1, 2, \ldots$ define $w_i = (-1)^j b_k$, where $i = 2^k(2j + 1)$. Then $f_i = 0$ if $w_i = 1$ and $f_i = 1$ if $w_i = -1$. This is equivalent to

$$f_i = 1 \quad \text{iff} \quad i \equiv 2^k(2 + b_k) \mod 2^{k+2}.$$

Remark 1. The regular paperfolding word corresponds to the sequence of instructions $\boldsymbol{b} = 1^\omega$.

Definition 4. Let $\mathbf{f}$ be a paperfolding word. An occurrence of 1 in $\mathbf{f}$ at position i is said to be *of order* k if the letter at position i is ? in f_{k-1} and 1 in f_k. We consider the 1's occurring in f_0 as of order 0.

Hence, in the paperfolding word $\mathbf{f}$, the 1's of order 0 appear at positions $2 + b_0 + 4t$, $t \geq 0$, the 1's of order 1 appear at positions $2(2 + b_1 + 4t)$, $t \geq 0$, and, in general, the 1's of order k appear at positions $2^k(2 + b_k + 4t)$, $t \geq 0$.

Let $\mathbf{f} = f_1 f_2 \cdots$ be a paperfolding word associated with the sequence $\boldsymbol{b} = b_0 b_1 \cdots$. A factor of $\mathbf{f}$ of length n starting at position $\ell+1$, denoted by $\mathbf{f}[\ell+1,\ldots,\ell+n]$, contains a number of 1's that is given by the sum, for all $k \geq 0$, of the 1's of order k in the interval $[\ell+1, \ell+n]$. For each k, since the 1's of order k are at distance 2^{k+2} one from another, the number of occurrences of 1's of order k in $\mathbf{f}[\ell+1,\ldots,\ell+n]$ is given by

$$\left\lfloor \frac{n-\ell}{2^{k+2}} \right\rfloor + \varepsilon_{k,b_k}(\ell, n),$$

where $\varepsilon_{k,b_k}(\ell, n) \in \{0,1\}$ depends on the sequence $\boldsymbol{b}$ (in fact, b_k determines the positions of the occurrences of the 1's of order k in $\mathbf{f}$). We set

$$\Delta(\ell, n) = \sum_{k \geq 0} \varepsilon_{k,b_k}(\ell, n)$$

the number of "extra" 1's in $\mathbf{f}[\ell+1,\ldots,\ell+n]$.

For example, in the prefix $\mathbf{p}[1,14]$ of length 14 of the regular paperfolding word, we know that there are at least $3 = \lfloor \frac{14}{4} \rfloor$ 1's of rank 0, $1 = \lfloor \frac{14}{8} \rfloor$ of rank 1 and $0 = \lfloor \frac{14}{16} \rfloor$ of rank 2. In the interval $[1,14]$ there are three 1's of rank 0 (at positions 3, 7 and 11), two 1's of rank 1 (at positions 6 and 14), and one 1 of rank 2 (at position 12), so we have in $\mathbf{p}[1,14]$ no extra 1 of rank 0, i.e., $\varepsilon_{0,1}(0,14) = 0$, one extra 1 of rank 1, i.e., $\varepsilon_{1,1}(0,14) = 1$ and one extra 1 of rank 2, i.e., $\varepsilon_{2,1}(0,14) = 1$, so that $\Delta(0,14) = 2$.

We set

$$\mathcal{E}_{k,b_k}(\ell, d, m) = (\varepsilon_{k,b_k}(\ell, \ell+d), \ldots, \varepsilon_{k,b_k}(\ell+(m-1)d, \ell+md))$$

and

$$\Delta(\ell, d, m) = \sum_{k \geq 0} \mathcal{E}_{k,b_k}(\ell, d, m) = (\Delta(\ell, \ell+d), \ldots, \Delta(\ell+(m-1)d, \ell+md)).$$

The factor of $\mathbf{f}$ of length dm starting at position $\ell+1$ is an abelian k-power if and only if the components of the vector $\Delta(\ell, d, m)$ are all equal, while it is an abelian k-anti-power if and only if the components of the vector $\Delta(\ell, d, m)$ are pairwise distinct.

The next result (Lemma 4 of [5]) will be the fundamental ingredient for the construction of abelian anti-powers in paperfolding words.

Lemma 5 (Additivity Lemma)**.** *Let $\ell, \ell' \geq 0$, and $d, d' \geq 1$ be positive integers with ℓ' and d' both even. Let r be such that $2^r > \ell + md$, and for each $k \geq 0$ the following implication holds: if $\mathcal{E}_{k,1}(\ell', d', m) \neq \mathcal{E}_{k,-1}(\ell', d', m)$ then $b_k = b_{k+r}$.*

Then

$$\Delta(\ell, d, m) + \Delta(\ell', d', m) = \Delta(\ell + 2^r \ell', d + 2^r d', m).$$

Using the Additivity Lemma, Holub [5] proved that all paperfolding words contain abelian powers of every order. We are now using the Additivity Lemma to prove that all paperfolding words contain abelian anti-powers of every order. We start with the regular paperfolding word, then we extend the proof to all paperfolding words.

3.1 Regular paperfolding word

Let

$$\begin{array}{cccc} \Phi & : & \{0,1\}^2 & \to & \{x,y,z\} \\ & & 00 & \mapsto & x \\ & & 01 & \mapsto & y \\ & & 10 & \mapsto & y \\ & & 11 & \mapsto & z \end{array}$$

be the morphism that identifies words of length 2 over the alphabet $\{0,1\}$ that are abelian equivalent. We have the following lemma:

Lemma 6. *Let $n \geq 3$ be an integer. Let $p = \boldsymbol{p}[\ell+1,\ldots,\ell+2^n] = u_1 v_1 \cdots u_{2^{n-1}} v_{2^{n-1}}$ be a factor of $\boldsymbol{p}$ of length 2^n. Then, no $q < 2^{n-1}$ exists such that*

$$\Phi(p) = \Phi(u_1 v_1) \cdots \Phi(u_{2^{n-1}} v_{2^{n-1}}) = \Phi(u_{q+1} v_{q+1}) \cdots \Phi(u_{2^{n-1}} v_{2^{n-1}}) \Phi(u_1 v_1) \cdots \Phi(u_q v_q). \qquad (1)$$

Proof. First, notice that if q' is the smallest solution of (1), then $q'|2^{n-1}$. Indeed, writing $w_i = \Phi(u_i v_i)$, we have

$$\begin{aligned} w_1 \cdots w_{2^{n-1}} &= w_1 \cdots w_{q'} w_{q'+1} \cdots w_{2^{n-1}} \\ &= w_{q'+1} \cdots w_{2^{n-1}} w_1 \cdots w_{q'}, \end{aligned}$$

and since two words commute if and only if they are powers of the same word, there exists a word z and integers a and b such that

$$w_1 \cdots w_{q'} = z^a \text{ and } w_{q'+1} \cdots w_{2^{n-1}} = z^b.$$

This gives $|z| \cdot (a+b) = 2^{n-1}$ and $|z| \cdot a = q'$. By the minimality of q', we have that $a = 1$ and so $|z| = q'|2^{n-1}$. Thus, $q' = 2^j$ for some integer $j < n$.

By the Toeplitz construction of $\mathbf{p}$, we immediately have that

$$u_1 v_1 \cdots u_{2^{n-1}} v_{2^{n-1}} = a v_1 \overline{a} v_2 a v_3 \overline{a} \cdots \overline{a} v_{2^{n-1}}$$

or

$$u_1 v_1 \cdots u_{2^{n-1}} v_{2^{n-1}} = u_1 a u_2 \overline{a} u_3 a u_4 \overline{a} \cdots u_{2^{n-1}} \overline{a}$$

with $a \in \{0,1\}$ and $\overline{a} = 1 - a$.

Suppose $q' \neq 1$ and $q' \neq 2^{n-1}$. Since q' is even, we have that $\Phi(u_i v_i) = \Phi(u_{i+q'} v_{i+q'})$ implies $u_i v_i = u_{i+q'} v_{i+q'}$. But this cannot be the case, since two consecutive letters of order j occur in $\mathbf{p}$ at distance 2^{j+1}. Since $j \leq n-2$, we have $2^{j+2} \leq 2^n$, so the factor p contains at least two consecutive letters of order j. Suppose that the first of such letters is u_i; then $u_{i+q'}$ is at distance $2q' = 2^{j+1}$, so $u_{i+q'} \neq u_i$, against the hypothesis that q' is a solution of (1).

Thus, we must have $q' = 1$ or $q' = 2^{n-1}$. Since $n \geq 3$, $\mathbf{p}[\ell+1, \ldots, \ell+2^n]$ contains two consecutive letters of order 1. Let us first suppose that v_i is a 1 of order 1 and v_{i+2} is a 0 of order 1. Then, $\Phi(u_i v_i) = \Phi(11) \neq \Phi(10) = \Phi(u_{i+2} v_{i+2})$. The other cases would give $10u_{i+1}v_{i+1}11$ with v_i a 0 of order 1 and v_{i+2} a 1 of order 1, $10u_{i+1}v_{i+1}00$ and $00u_{i+1}v_{i+1}10$ if u_i is a 1 of order 1 and u_{i+2} a 0 of order 1 or vice versa. Every case leads to $\Phi(u_i v_i) \neq \Phi(u_{i+2} v_{i+2})$. This implies $q' \neq 1$ and so $q' = 2^{n-1}$. By minimality of q', the only solution of (1) is $q = 2^{n-1}$. ☐

Theorem 7. *The regular paperfolding word contains abelian m-anti-powers for every $m \geq 2$.*

Proof. The proof is mainly based on the Additivity Lemma. Fix m. To prove the result it is sufficient to find a vector $\Delta(s, d, m)$ filled with pairwise distinct components. Let k be an integer such that $2^k \geq m$. Consider the first factor of length $2^{k+2} - 1$ containing a 1 of order k in the middle; our factor is then of the form

$$w1w'$$

with $|w| = |w'| = 2^{k+1} - 1$. Since for every positive integers i, j

$$p_j \text{ of order } i \Rightarrow p_{j+2^{i+2}} = p_j \neq p_{j+2^{i+1}} \tag{2}$$

then, up to applying a translation, we can suppose $w = w'$. In fact, the equality is true for every letter of order smaller than k by (2). Now, taking the smallest order $r > k$ of a letter in w or w' such that this letter differs from 1, if we consider the factor translated of 2^{r+1}, by (2) the letters of order smaller than r are the same and the letter we considered becomes a 1. Since the length of $w1w'$ is $2^{k+2} - 1$ and the distance between two letters of order higher than k is at least 2^{k+2}, we have that in less than 2 steps we get $w1w$ with every letter of order greater than k being a 1. Writing $\ell + 1$ the starting position of an occurrence in $\mathbf{p}$ of the factor $w1w$, we set $\ell' = \ell$ if ℓ is even or $\ell' = \ell + 1$ otherwise. Consider the vectors

$$\Delta(\ell', 2, 2^k), \Delta(\ell'+2, 2, 2^k), \Delta(\ell'+4, 2, 2^k), \Delta(\ell'+6, 2, 2^k), \ldots, \Delta(\ell'+2^{k+1}-2, 2, 2^k).$$

We claim that these vectors are pairwise distinct. By contradiction, if $\Delta(\ell'+2p, 2, 2^k) = \Delta(\ell'+2q, 2, 2^k)$ for some p, q with $p \leq q$, then we have that $\Phi(p_{\ell'+2p+1} \cdots p_{\ell'+2p+2^{k+1}}) = \Phi(p_{\ell'+2q+1} \cdots p_{\ell'+2q+2^{k+1}})$. Since the factor we are considering is $w1w$, we have $p_{l'+2p+1} \cdots p_{l'+2q-1} = p_{l'+2p+1+2^{k+1}} \cdots p_{l'+2q-1+2^{k+1}}$ and so

$$\Phi(u_{\ell'+2p+1} \cdots u_{\ell'+2p+2^{k+1}}) = \Phi(u_{\ell'+2q+1} \cdots u_{\ell'+2p+2^{k+1}} u_{\ell'+2p+1} \cdots u_{\ell'+2q+2^{k+1}})$$

but this contradicts Lemma 7.

Finally, as the vectors are different, we use the Additivity Lemma to obtain a vector whose components are pairwise distinct: Applying n times the Additivity Lemma on $\Delta(\ell'+2p,2,2^k)$ one can obtain $n\Delta(\ell'+2p,2,2^k)$. It then suffices to take a sequence of integers $\alpha_0,\dots,\alpha_{2^k-1}$ increasing enough to have

$$\Sigma_{i=0}^{2^k-1}\alpha_i\Delta(s'+2i,2,2^k),$$

a vector whose components are pairwise distinct. Indeed, labelling a_j the j-th component of this vector and $x_{i,j}$ the j-th component of $\Delta(s'+2i,2,2^k)$, we have

$$a_j = a_{j'} \Leftrightarrow \Sigma_{i=0}^{2^k-1}\alpha_i x_{i,j} = \Sigma_{i=0}^{2^k-1}\alpha_i x_{i,j'} \Leftrightarrow \Sigma_{i=0}^{2^k-1}\alpha_i(x_{i,j}-x_{i,j'}) = 0.$$

By "increasing enough", we precisely mean $\alpha_r > \Sigma_{i=0}^{r-1}\alpha_i \sup_{0\le q,q'\le 2^k-1}(x_{i,q}-x_{i,q'})$, so that by decreasing induction we have that for every i, with $0\le i\le 2^k-1$, one has $x_{i,j}=x_{i,j'}$. In particular, this gives $\Delta(\ell'+2j,2,2^k)=\Delta(\ell'+2j',2,2^k)$, which implies $j=j'$. Hence, all the components are pairwise distinct and the proof is complete. ☐

3.2 All paperfolding words

To generalize the result above to all paperfolding words, one has to take care of the condition $b_i = b_{i+r}$ in the Additivity Lemma.

Lemma 7 can be modified so that the translation is not by 2 but by 2^u, for any $u>1$. Let

$$\begin{array}{cccc}\phi & : & \{0,1\}^{2^u} & \to & \mathbb{N}\\ & & (a_1\cdots a_{2^u}) & \mapsto & |\{i \mid a_i = 1\}|\end{array}$$

be the morphism that identifies words of length 2^u over $\{0,1\}$ that are abelian equivalent. Then we have the following lemma, analogous to Lemma 7:

Lemma 8. *Let $n\ge u+3$ be an integer and let $\boldsymbol{f}$ be a paperfolding word. Every factor $f=\boldsymbol{f}[\ell+1,\ell+2^n]=a_{1,1}a_{1,2}\cdots a_{2^{n-1},2^u-1}a_{2^{n-1},2^u}$ of $\boldsymbol{f}$ of length 2^n satisfies the following property: If q is such that*

$$\begin{aligned}&\phi(f)=\phi(a_{1,1}\cdots a_{1,2^u})\cdots\phi(a_{2^{n-1},1}\cdots a_{2^{n-1},2^u})=\\&\phi(a_{q+1,1}\cdots a_{q+1,2^u})\cdots\phi(a_{2^{n-1},1}\cdots a_{2^{n-1},2^u})\phi(a_{1,1}\cdots a_{1,2^u})\cdots\phi(a_{q,1}\cdots a_{q,2^u}),\end{aligned}$$

then $q=2^{n-1}$.

Proof. The proof of Lemma 7 mainly applies here; all we have to change is the part where we are using the Toeplitz construction to justify $j=n-1$. Here, in each 2^u-tuple one can find one letter of order $u-1$ and one letter of higher order. Using (2), we then see that $\phi(a_{i,1}\cdots a_{i,2^u})$ is totally determined by the letter of order $u-1$ and the letter of higher order in $(a_{i,1}\cdots a_{i,2^u})$. Applying again (2) to the letter of order $u-1$ we see that we can apply exactly the same reasoning as in the proof of Lemma 7 (in a sense, our new ϕ is the previous one modulo the letters of order smaller than $u-1$). Thus, we can follow the same proof than in Lemma 7. ☐

Now we can prove the main theorem:

Theorem 9. *Every paperfolding word $\boldsymbol{f}$ contains abelian m-anti-powers for every $m\ge 2$.*

Proof. Let k be an integer such that $2^k\ge m$. As before, we will prove that $\mathbf{f}$ contains abelian 2^k-anti-powers, hence it will contain abelian m-anti-powers. Since the alphabet $\{0,1\}$ is finite, there must exist a factor $b_{u-1}\cdots b_{u+k+4}$ of $\boldsymbol{b}$ that occurs infinitely often. As before, let us start with the first block of length $2^{u+k+2}-1$ containing a 1 of order $u+k$ in the middle; our block is then

$$w1w'$$

with $|w| = |w'| = 2^{u+k+1} - 1$. As before, in at most two steps, we can have $w = w'$, and the maximum order of a letter appearing in this factor is $u + k + 4$. Again, writing ℓ the starting position of an occurrence of this factor, we set $\ell' = \ell$ if ℓ is even or $\ell' = \ell + 1$ otherwise. Consider the vectors

$$\Delta(\ell', 2^u, 2^k), \Delta(\ell' + 2^u, 2^u, 2^k), \Delta(\ell' + 2^{u+1}, 2^u, 2^k), \ldots, \Delta(\ell' + 2^{u+k+1} - 2^u, 2^u, 2^k).$$

Here again, these vectors are pairwise distinct: if $\Delta(\ell' + 2^u p, 2^u, 2^k) = \Delta(\ell' + 2^u q, 2^u, 2^k)$, we have that

$$\phi(a_{1,1}, \cdots, a_{1,2^u}) \cdots \phi(a_{2^{n-1},1} \cdots a_{2^{n-1},2^u}) = \phi() CHECK - THE - INDEXES$$

and this contradicts Lemma 9.

Moreover, $\varepsilon_{i,0}(l' + 2^u p, 2^u, 2^k) \neq \varepsilon_{i,1}(l' + 2^u p, 2^u, 2^k) \Rightarrow u - 1 \leq i \leq u + k + 4$, using (2) and the fact that no letter of order higher than $u + k + 4$ appears in the factor $w1w$. So, choosing r such that $2^r > l' + 2^{u+k+1} - 2^u + 2^{u+k}$ and $b_{u-1} \cdots b_{u+k+4} = b_{r+u-1} \cdots b_{r+u+k+4}$, we can apply the Additivity Lemma and, as for the regular paperfolding word, construct an abelian 2^k-anti-power that occurs as a factor in $\mathbf{f}$. $\square$

References

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

[2] J. Cassaigne and J. Karhumäki. Toeplitz words, generalized periodicity and periodically iterated morphisms. *Eur. J. Comb.*, 18(5):497–510, 1997.

[3] A. de Luca and S. Varricchio. *Finiteness and Regularity in Semigroups and Formal Languages.* Monographs in Theoretical Computer Science (An EATCS Series). Springer, 1st edition, 1999.

[4] G. Fici, A. Restivo, M. Silva, and L. Q. Zamboni. Anti-powers in infinite words. *J. Comb. Theory, Ser. A*, 157:109–119, 2018.

[5] S. Holub. Abelian powers in paper-folding words. *J. Comb. Theory, Ser. A*, 120(4):872–881, 2013.

[6] B. Madill and N. Rampersad. The abelian complexity of the paperfolding word. *Discr. Math.*, 313(7):831–838, 2013.

[7] M. Morse and G. A. Hedlund. Symbolic dynamics. *Amer. J. Math.*, 60:1–42, 1938.

[8] G. Richomme, K. Saari, and L. Zamboni. Abelian complexity of minimal subshifts. *J. Lond. Math. Soc.*, 83(1):79–95, 2011.