Anti-Powers in Primitive Uniform Substitutions

Mickaël Postic

Institut Camille Jordan, Université Claude Bernard Lyon 1, Lyon, France

`postic@math.univ-lyon1.fr`

Abstract

In a recent work, A. Berger and C. Defant showed that if x is a fixed point of a binary uniform and primitive morphism, then there exists a constant C such that for all positive integers i, k, beginning in position n in x is a k-anti-power with block length at most Ck. They ask whether this result extends to a broader class of morphic words. In this note we extend their results to fixed points of uniform primitive morphisms on arbitrary finite alphabets. Our methods make use of the recognisability of uniform primitive morphisms.

1 Introduction

Since its introduction by Fici et al. [FRSZ18], the notion of *anti-powers* has generated a lot of interest [BD19, FPS19, Def17, Gae18, Bur18, BFP18]. A k-anti-power is the concateation of k blocks of same lengths that are pairwise different. In one of the latest contributions, A. Berger and C. Defant [BD19] studied the block length of anti-powers arising in morphic words. Up to changing a little the definitions, Theorem 5 in [BD19] can be formulated the following way:

Theorem 1. *If w is aperiodic, fixed point of a primitive binary uniform morphism, then there is a constant $C = C(w)$ such that $\forall n,\ k \in \mathbb{N}$, w contains a k-anti-power with blocks of length at most Ck beginning at its nth position.*

They also asked to what extent are these results generalisable to a broader class of morphic words, and in particular if it was still true without the binary condition. Using the notion of recognisability first introduced by B. Mossé in [Mos96], we show that their results extend to fixed points of uniform primitive morphisms on arbitrary finite alphabets:

Theorem 2. *If σ is primitive and m-uniform, with an aperiodic fixed point x, there exist a constant $C = C(\sigma)$ such that : $\forall y \in X(\sigma)$, $\forall n, k \in \mathbb{N}$, y contains a k-antipower with block length at most Ck starting at position n.*

2 Preliminaries

We begin by recalling some basic notions pertaining to morphisms and words.

Let us fix, for the rest of this article, a set $\mathbb{A}$ called the *alphabet* whose elements are called *letters*. The set of *finite non empty words* $\mathbb{A}^+$ is $\bigcup_{n \in \mathbb{N}} \mathbb{A}^n$, while $\mathbb{A}^* = \mathbb{A}^+ \cup \{\varepsilon\}$ is the set of finite words (ε is the *empty* word). For a word $w \in \mathbb{A}^n$ we write $n = |w|$ the *length* of w. Let $\mathbb{A}^{\mathbb{N}} = \{a_1 a_2 \cdots | a_i \in \mathbb{A}\}$ be the set of *infinite words* over $\mathbb{A}$. For a finite or infinite word w, the notation $w_{[n,m]}$ will refer to the *factor* $w_n w_{n+1} \cdots w_m$. The set of factors of w will be denoted by Fact(w). A *morphism* σ over $\mathbb{A}$ is a map $\sigma :\ \mathbb{A}^* \to \mathbb{B}^*$ where $\mathbb{B}$ is another alphabet such that $\sigma(ww') = \sigma(w)\sigma(w')$. A *substitution* σ over $\mathbb{A}$ is a morphism where $\mathbb{A} = \mathbb{B}$. A *uniform* morphism σ is a morphism of constant length over letters: $\forall\ a, b \in \mathbb{A},\ |\sigma(a)| = |\sigma(b)|$.

Remark 1. In [BD19] the term morphism is used instead of substitution. We decided to stick to the term substitution, since it is commonly used in articles pertaining to recognizability.

Definition 3. Let $\mathbb{A} = \{a_1, \cdots, a_r\}$. A morphism σ over $\mathbb{A}$ is said to be *primitive* if:

$$\exists n, \ \forall i, \ \forall j, \ a_j \text{ occurs in } \sigma^n(a_i).$$

For a morphism σ, $x \in \mathbb{A}^{\mathbb{N}}$ is called a *fixed point* if $x = \sigma(x)$. The *shift orbit closure* $X(x)$ is the closure under the natural topology on $\mathbb{A}^{\mathbb{N}}$ of the orbit of x under the shift operator $\tau : a_1 a_2 \cdots \to a_2 a_3 \cdots$. If σ is primitive, it is easy to see that $X(x) = X(y)$ for any x and y fixed points of σ. Hence we can define $X(\sigma) = X(x)$ in this case.

Definition 4. A m-uniform primitive morphism σ is said to be *recognizable* if $\exists N \in \mathbb{N}$ such that $\forall y \in X(\sigma), \forall w \in \mathbb{A}^+, \sigma(y)_{[\alpha,\alpha+|w|-1]} = \sigma(y)_{[\beta,\beta+|w|-1]} = w$ with $|w| \geq N$ and $\alpha = 0 \pmod{m}$ then $\beta = 0 \pmod{m}$. N is refered to as *recognizability constant* of σ in this article.

Remark 2. Let σ be a m-uniform primitive recognizable morphism and N given by Definition 4. Then $\forall y \in X(\sigma), \forall w \in \mathbb{A}^+, \sigma(y)_{[\alpha,\alpha+|w|-1]} = \sigma(y)_{[\beta,\beta+|w|-1]} = w$ with $|w| \geq N + m$ gives $\alpha = \beta \pmod{m}$.

Proof. Denote $h = \beta - \alpha$ and $w' = \sigma(y)_{[m\lceil \frac{\alpha}{m} \rceil, \alpha+|w|-1]}$. Then $|w'| \geq N$ and $w' = \sigma(y)_{[h+m\lceil \frac{\alpha}{m} \rceil, h+\alpha+|w|-1]}$. By definition, this implies $h = 0 \pmod{m}$ hence $\alpha = \beta \pmod{m}$. $\square$

Remark 3. If $r = |\mathbb{A}| = 2$, an aperiodic word w that is a fixed point of a m-uniform morphism σ is uniformly recurrent if and only if σ is primitive.

Proof. Let us first suppose σ is primitive. Without loss of generality, let us fix $w = \sigma^{\infty}(0)$. Let us denote $\mathbb{A} = \{0; 1\}$ with $w = \sigma^{\infty}(0)$. It is easy to see that the n in Definition 3 is 2 or 1. Let then x be a factor of w. There exists k such that $x \in \text{Fact}(\sigma^k(0))$. Every factor of w of length at least $2m^{k+2}$ contains, for some $a \in \mathbb{A}$, $\sigma^{k+2}(a)$, so it contains $\sigma^k(0)$ hence x, and so w is uniformly recurrent.

Let us now suppose σ is not primitive. Since σ is m-uniform, w has to be eventually periodic. Indeed, if $\sigma(0) = 0^m$, then $w = 0^{\infty}$. If not and $\sigma(1) = 1^m$, then w contains arbitrary long plages of 1 ($1^{\infty} \in X(w)$), hence arbitrary long factors do not contain the factor 0. The only option left is $\sigma(0) = 1^m$ and $\sigma(1) = 0^m$. But this leads to no fixed point ($\sigma(0)$ must start with a 0). $\square$

We will now give some well-known results on substitutions and recognizability that we will need later:

Theorem 5 (Corollaire 3.2 in [Mos96]). *Let σ be a primitive m-uniform substitution and let x be aperiodic such that $\sigma(x) = x$. Then σ is recognizable.*

We will also use the following proposition from [HQZ99] (actually what is proved is somewhat stronger, but we only need this formulation):

Proposition 6. *If σ is a m-uniform morphism, and x is aperiodic with $x = \sigma(x)$ and $x_0 = a$, then $\exists N_1 \in \mathbb{N}$ such that $l \geq N_1$ implies that each occurrence $\sigma^l(a)$ in x is the image under σ of an occurrence of $\sigma^{l-1}(a)$ in x.*

3 Main Part

The goal of this part is to prove Therorem 2.

We will first give a lemma that is easily deduced from Proposition 6:

Lemma 7. *Let σ be a m-uniform morphism, and x aperiodic with $x = \sigma(x)$ and $x_0 = a$. Let $N_1 \in \mathbb{N}$ be given by Proposition 6. Then for every r, $l \in \mathbb{N}$, each occurrence $\sigma^{l+N_1+r}(a)$ in x is the image under σ^l of an occurrence of $\sigma^{N_1+r}(a)$ in x.*

Proof. By induction on l. Initialization is just the result of Proposition 6. Let $l \in \mathbb{N}$ be fixed. Suppose the result holds for l and let $x_{[\alpha,\alpha+m^{l+1+N_1+r}-1]} = \sigma^{l+1+N_1+r}(a)$. Then by the recurrence hypothesis, $\alpha = m^l \alpha'$ and $x_{[\alpha',\alpha'+m^{1+N_1+r}-1]} = \sigma^{1+N_1+r}(a)$. But now we can apply Proposition 6: $\alpha' = m\alpha''$ and $x_{[\alpha'',\alpha''+m^{N_1+r}-1]} = \sigma^{N_1+r}(a)$. So $x_{[\alpha,\alpha+m^{l+1+N_1+r}-1]} = \sigma^{l+1+N_1+r}(a)$ is the image under σ^{l+1} of an occurrence of $\sigma^{N_1+r}(a)$ in x. $\square$

We can now prove the following lemma:

Lemma 8. *If σ is a primitive m-uniform morphism, and x is aperiodic with $x = \sigma(x)$, $\exists N' \in \mathbb{N}$ such that $\forall i \in \mathbb{N}$, σ^i is recognizable with a recognizability constant less or equal to $m^i N'$.*

Proof. Let then σ be an aperiodic m-uniform morphism, and $x = \sigma(x)$ with $x_0 = a$. Let N_1 be given by Proposition 6, and let $y \in X(\sigma)$.

By Theorem 5, σ is recognizable. Let N be a recognizability constant of σ and let $r \in \mathbb{N}$ such that $m^{N_1+r} \geq N + m$. The prefix of x of length m^{N_1+r} is $p = \sigma^{N_1+r}(a)$. Since σ is primitive and uniformly recurrent, $\exists M \in \mathbb{N}$ such that every factor of x, hence of y, of length at least M contains p. I then claim that $N' = 2M$ has the required property.

Indeed, let w, $|w| \geq N'm^i$ be fixed. We show the following:

$$\sigma^i(y)_{[\alpha,\alpha+|w|-1]} = \sigma^i(y)_{[\beta,\beta+|w|-1]} = w \Rightarrow \beta = \alpha \pmod{m^i} \quad (1).$$

Let α, β be as in (1) and $h = \beta - \alpha$. So:

$$\sigma^i(y)_{[\alpha,\alpha+|w|-1]} = \sigma^i(y)_{[\alpha+h,\alpha+h+|w|-1]} = w.$$

By $y \in X(\sigma)$ we get an α' with

$$x_{[\alpha',\alpha'+|w|-1]} = x_{[\alpha'+h,\alpha'+h+|w|-1]} = w. \quad (2)$$

Since $|w| \geq N'm^i = 2Mm^i$, there exists $\gamma \geq 0$, $w' \in F(w)$ with
$w' = x_{[\gamma m^i,(\gamma+M)m^i-1]} = x_{[\gamma m^i+h,(\gamma+M)m^i+h-1]}$. Let $z = x_{[\gamma,\gamma+M-1]}$ so $w' = \sigma^i(z)$.
Since $|z| = M$, $\exists \gamma'$, $\gamma \leq \gamma' < \gamma' + m^{N_1+r} - 1 \leq \gamma + M - 1$ such that:

$$x_{[\gamma',\gamma'+m^{N_1+r}-1]} = \sigma^{N_1+r}(a).$$

Applying σ^i to x gives $x_{[\gamma'm^i,(\gamma'+m^{N_1+r})m^i-1]} = \sigma^{N_1+r+i}(a)$, and by (2),

$$x_{[\gamma'm^i,(\gamma'+m^{N_1+r})m^i-1]} = x_{[\gamma'm^i+h,(\gamma'+m^{N_1+r})m^i+h-1]} = \sigma^{N_1+r+i}(a).$$

Using Lemma 7, this implies $h = 0 \pmod{m^i}$: $x_{[\gamma'm^i+h,(\gamma'+m^{N_1+r})m^i+h-1]}$ is the σ^i image of $x_{[\delta,\delta+m^{N_1+r}]} = \sigma^{N_1+r}(a)$, hence

$$\gamma'm^i + h = \delta m^i \text{ and so } h = 0 \pmod{m^i}.$$

$\square$

Proof theorem 2. Let σ be aperiodic, primitive and m-uniform and let $C(\sigma) = (N'+1)m$ where N' is the constant given by Lemma 8. Let then $k, n \in \mathbb{N}$ and $y \in X(\sigma)$ be fixed.

Let $i \in \mathbb{N}$ be such that $m^{i-1} \leq k < m^i$. Consider then the k consecutive blocks of length $N'm^i + 1$ starting at position n: the block number s is then $y_{[n+s(N'm^i+1),n+(s+1)(N'm^i+1)-1]}$. We have $kC(\sigma) \geq (N'+1)m^i \geq N'm^i + 1$. Moreover, using Lemma 8, we get that two of these blocks, say blocks s and t, are equal implies the difference between their starting indices is 0 modulo m^i:

$$y_{[n+s(N'm^i+1),n+(s+1)(N'm^i+1)-1]} = y_{[n+t(N'm^i+1),n+(t+1)(N'm^i+1)-1]} \quad (1)$$

$$\Rightarrow n + s(N'm^i + 1) = n + t(N'm^i + 1) \pmod{m^i} \quad (2)$$

$$\Rightarrow s = t \pmod{m^i}. \quad (3)$$

But since k is smaller than m^i this implies $s = t$, which completes the proof. $\square$

4   Conclusion

Using the theory of recognizability, we have been able to improve the previous result to the class of aperiodic fixed points of primitive and m-uniform morphisms. The two followig examples show that the conditions aperiodic and primitive are tight. Let $\sigma : \begin{cases} 0 \to 01 \\ 1 \to 01 \end{cases}$. Since $(01)^\infty$, which is not aperiodic, does only contain two factor of each length, it cannot contain k-anti-powers for k greater than 2. Let then $\sigma : \begin{cases} 0 \to 010 \\ 1 \to 111 \end{cases}$. This substitution is not primitive, and give rise to the celebrated Cantor word $c = \sigma^\infty(0)$. Since c contains arbitrary long plages of 1, it is clear that Theorem 2 doesn't apply here.

In the other hand, this result might be extendable to the class of recognizable substitutions, alas, I was not successful in finding an equivalent to Lemma 8; this seems to be the key to extend this result.

References

[BD19]   Aaron Berger and Colin Defant. On Anti-Powers in Aperiodic Recurrent Words. *arXiv e-prints*, page arXiv:1902.01291, Feb 2019.

[BFP18]   Golnaz Badkobeh, Gabriele Fici, and Simon J. Puglisi. Algorithms for Anti-Powers in Strings. *Inform. Process. Lett.*, 137:57–60, 2018.

[Bur18]   Amanda Burcroff. (k, λ)-Anti-Powers and Other Patterns in Words. *Electron. J. Comb.*, 25:4–41, 2018.

[Def17]   Colin Defant. Anti-Power Prefixes of the Thue-Morse Word. *Electron. J. Comb.*, 24, 2017.

[FPS19]   Gabriele Fici, Mickaël Postic, and Manuel Silva. Abelian Anti-Powers in Infinite Words. *Adv. in Appl. Math.*, 108:67 – 78, 2019.

[FRSZ18]   Gabriele Fici, Antonio Restivo, Manuel Silva, and Luca Q. Zamboni. Anti-Powers in Infinite Words. *J. Comb. Theory, Ser. A*, 157:109–119, 2018.

[Gae18]   Marisa Gaetz. Anti-Power j-fixes of the Thue-Morse Word. 2018.

[HQZ99]   Charles Holton and Luca Q. Zamboni. Descendants Of Primitive Substitutions. *Theory Comput. Syst.*, 32:133–157, 04 1999.

[Mos96]   Brigitte Mossé. Reconnaissabilité des Substitutions et Complexité des Auites Automatiques. *Bull. Soc. Math. France*, 124(2):329–346, 1996.