QUASIPERIODIC INFINITE WORDS: SOME ANSWERS

F. Levé, G. Richomme
LaRIA, UPJV,
33 rue Saint Leu,
80039 Amiens Cedex 01, France
{Florence.Leve,Gwenael.Richomme}@u-picardie.fr

Abstract

We answer some questions about infinite quasiperiodic words asked by Marcus in Bulletin 82 of the European Assocation of Theoretical Computer Science.

1 Introduction

The notion of repetition in Strings is central in a lot of researches (see for instance [7],[8]). In this vein, Apostolico and Ehrenfeucht introduced the notion of quasiperiodic finite words [2] in the following way: "a string z is quasiperiodic if there is a second string $w \neq z$ such that every position of z falls within some occurrence of w in z". The reader can consult [1] for a short survey of studies concerning quasiperiodicity. In [10], Marcus extends this notion to infinite words and he leaves open six questions. We bring answers to the fourth first questions. The questions will be recalled while treating them.

After some generalities, in Section 3, we recall the notion of quasiperiodic (finite or infinite) words. Section 4 answers the first question providing an example of Sturmian words which is not quasiperiodic. In answer to the second question, Section 5 shows that quasiperiodic words can have an exponential complexity. Answers to the third and fourth questions are given in Section 7. They are based on a characterization of the set of quasiperiods of the Fibonacci word stated in Section 6. In conclusion, we briefly consider the two last questions.

2 Generalities

We assume the reader is familiar with combinatorics on words and morphisms (see, e.g., [8]). We precise our notations.

Given an alphabet A (a non-empty set of letters), A^* is the set of finite words over A including the empty word ε. The length of a word w is denoted by $|w|$. A word u is a *factor* of w if there exist words p and s such that $w = pus$. If $p = \varepsilon$ (resp. $s = \varepsilon$), u is a *prefix* (resp. *suffix*) of w. A word u is a *border* of a word w if u is both a prefix and a suffix of w. A factor u of a word w is said *proper* if $w \neq u$.

Given an alphabet A, a(n endo)*morphism* f on A is an application from A^* to A^* such that $f(uv) = f(u)f(v)$ for any words u, v over A. A morphism on A is entirely defined by the images of elements of A. Given a morphism f, *powers* of f are defined inductively by $f^0 = Id$, $f^i = ff^{i-1}$ for integers $i \geq 1$ (composition of applications is denoted just by juxtaposition). When for a letter a, $f(a) = ax$ with $x \neq \varepsilon$, for all $n \geq 0$, $f^n(a)$ is a prefix of $f^{n+1}(a)$. If moreover, for all $n \geq 0$, $|f^n(a)| < |f^{n+1}(a)|$, the limit $\lim_{n\to\infty} f^n(a)$ is the infinite word denoted $f^\omega(a)$ having all the $f^n(a)$ as prefixes. This limit is also a fixed point of f.

3 Quasiperiodicity

We need both definitions of finite and infinite quasiperiodic words.

Let us take from [3] definitions in the finite case. A string w *covers* another string z if for every $i \in \{1, \ldots, |z|\}$, there exists $j \in \{1, \ldots, |w|\}$ such that there is an occurrence of w starting at position $i - j + 1$ in string z. Alternatively we say that w is a *quasiperiod* of z. If z is covered by $w \neq z$, then z is *quasiperiodic*. A string z is *superprimitive* if it is not quasiperiodic (Marcus [10] calls minimal such words). One can observe that any word of length 1 is not quasiperiodic. The string

$$z = abaababaabaababaaba$$

has *aba*, *abaaba*, *abaababaaba* as quasiperiods. Only *aba* is superprimitive. More generally in [3], it is proved that any quasiperiodic word has exactly one superprimitive quasiperiod. This is a consequence of the fact that any quasiperiod of a finite word w is a proper border of w.

When defining infinite quasiperiodic words, instead of considering the starting indices of the occurrences of a quasiperiod, for convenience, we choose to consider the words preceding the occurrences of a quasiperiod. An infinite word $\underline{w}$ is *quasiperiodic* if there exist a finite word x and words $(p_n)_{n\geq 0}$ such that $p_0 = \varepsilon$ and, for $n \geq 0$, $0 < |p_{n+1}| - |p_n| \leq |x|$ and $p_n x$ is a prefix of $\underline{w}$. We say that x *covers* $\underline{w}$. The word x is also called a *quasiperiod* and we say that the sequence $(p_n x)_{n\geq 0}$ is *a covering sequence of prefixes of the word* $\underline{w}$. In [9], Marcus proves that any infinite word having all finite words as factors is not quasiperiodic. In [10], several examples of quasiperiodic words are given. They have all the form

$$(rs)^{i_1} r (rs)^{i_2} r \ldots$$

for two words $r, s \neq \varepsilon$ and non-zero integers $(i_n)_{n\geq 1}$. Let us give another example of quasiperiodic word which does not follow this form. For this, for $k \geq 2$, we consider the endomorphism φ_k defined on $\Sigma_k = \{a_1, \ldots, a_k\}$ by $\varphi_k(a_i) = a_1 a_{i+1}$ if $i \neq k$ and $\varphi_k(a_k) = a_1$. This morphism extends the well-known Fibonacci morphism defined on $\{a, b\}^*$ by $\varphi(a) = ab$ and $\varphi(b) = a$. Marcus mentioned that the Fibonacci word, the fixed point of φ, is quasiperiodic (with quasiperiod $aba = \varphi^2(a)$). More generally, we have:

Lemma 3.1. *For $k \geq 2$, the fixed word $\varphi_k^\omega(a_1)$ is quasiperiodic with quasiperiod $\varphi_k^k(a_1)$.*

Proof. Let $(u_i)_{i\geq 0}$ be the sequence of words defined by $u_0 = \varepsilon$, $u_i = \varphi_k(u_{i-1})a_1$ for $i \geq 1$. The reader can verify the following properties for $1 \leq i \leq k$:

- $u_i = u_{i-1}a_i u_{i-1}$. In particular, u_j is a prefix of u_i when $0 \leq i \leq j \leq k$. (We can also note that u_i is a palindrom.)
- $\varphi_k^i(a_j) = u_i a_{i+j}$ for $1 \leq j \leq k - i$, $\varphi_k^i(a_j) = u_i u_{i-1+j-k}^{-1}$ for $k - i < j \leq k$, where uv^{-1} denotes the word w such that $u = wv$ (this notation can be used only if v is a suffix of u).

Consequently we can observe that $\varphi_k^k(a_1) = u_k u_0^{-1} = u_k$ covers each word $\varphi_k^k(a_j a_1) = u_k u_{j-1}^{-1} u_k$. The word $\varphi_k^\omega(a_1)$ can be decomposed over $\{\varphi_k^k(a_1), \varphi_k^k(a_2a_1), \ldots, \varphi_k^k(a_k a_1)\}$. So $\varphi_k^k(a_1)$ covers $\varphi_k^\omega(a_1)$. □

We end this section with another definitions. We will need to consider infinite words covered by two words and not only one. We say that the set $\{xa, xb\}$ *covers* $\underline{w}$ if xa and xb are factors of w and there exist words $(p_n x a_n)_{n\geq 0}$ with $a_n \in \{a, b\}$ such that $p_0 = \varepsilon$, and, for $n \geq 0$, $0 < |p_{n+1}| - |p_n| \leq |x| + 1$ and $p_n x a_n$ is a prefix of $\underline{w}$. Once again the sequence $(p_n x a_n)_{n\geq 0}$ is called *a covering sequence of prefixes of the word $\underline{w}$*.

4 About Sturmian words

The first question in [10] is: "Is every Sturmian word quasiperiodic?". Proposition 4.1 below provides a negative answer.

Proposition 4.1. *Not all Sturmian words are quasiperiodic.*

Let us recall that there are several equivalent definitions of Sturmian words (see [4] for instance). A convenient tool to deal with Sturmian words is the set of Sturmian endomorphisms $\{\varphi, \tilde{\varphi}, E\}$ where $\tilde{\varphi}$ and E are defined on $\{a, b\}$ by

$\tilde{\varphi}(a) = ba$, $\tilde{\varphi}(b) = a$ and $E(a) = b$, $E(b) = a$. The set $\{\varphi, \tilde{\varphi}, E\}^*$ is exactly the set of morphisms that preserves Sturmian words (the image of a Sturmian word is Sturmian) [12]. It is also well-known that if a Sturmian morphism generates an infinite word then this fixed point is a Sturmian word. Let us consider the Sturmian morphism $E\tilde{\varphi}\varphi E$. Proposition 4.1 is a corollary of the following result:

Lemma 4.2. *The infinite word* $(E\tilde{\varphi}\varphi E)^\omega(a)$ *is not quasiperiodic.*

Proof. Let $f = E\tilde{\varphi}\varphi E$: $f(a) = ab$, $f(b) = abb$. The proof of this lemma holds by contradiction. Assume that x is a quasiperiod of $f^\omega(a)$ of minimal length. Note that x is a prefix of $f^\omega(a) = ababbababbabb\ldots$ We observe that $|x| \geq 5$. By construction of $f^\omega(a)$, x ends with abb, ab or a. If x ends with abb, for each word p such that px is a prefix of $f^\omega(a)$, there exist words p', x' such that $p = f(p')$, $x = f(x')$ and $|x'| < |x|$. In fact x' does not depend on p. Consequently we can verify that x' is a quasiperiod of $f^\omega(a)$. This contradicts the choice of x. If x ends with a then xb is also a quasiperiod of $f^\omega(a)$. So we can assume x ends with ab and $f^\omega(a)$ has no quasiperiod of length less than or equal to $|x| - 2$. Let $(p_n)_{n\geq 0}$ be a covering sequence of prefixes of $f^\omega(a)$. Since x starts with a, there exist words $(p'_n)_{n\geq 0}$ and a unique word x' such that $p_n = f(p'_n)$ and $x = f(x')ab$. Moreover from $|x| \geq 5$, we deduce $|x'a| < |x| - 2$. Consequently $x'a$ cannot be a quasiperiod of $f^\omega(a)$. It follows that xa and xb are factors of $f^\omega(a)$ (otherwise $xb = f(x'b)$ or $x = f(x'a)$ and we can deduce that $x'b$ or $x'a$ is a quasiperiod of $f^\omega(a)$). So $\{xa, xb\}$ covers $f^\omega(a)$.

Let y be a non-empty word such that $\{ya, yb\}$ covers $f^\omega(a)$. Note that y must end with ab. Let $(p_n y a_n)_{n\geq 0}$ (with $a_n \in \{a, b\}$ for all $n \geq 0$) be a covering sequence of prefixes of $f^\omega(a)$. Since $|y| \geq 2$, y starts with ab. Consequently there exist words $(p'_n)_{n\geq 0}$ and y' such that $p_n = f(p'_n)$ and $y = f(y')ab$. Moreover $(p'_n y' a_n)_{n\geq 0}$ is a covering sequence of prefixes of $f^\omega(a)$. From what precedes, it follows that the word y is one of the words x_n defined by $x_0 = \varepsilon$ and $x_n = f(x_{n-1})$ for $n \geq 1$. Note that we can see by induction that for all $n \geq 1$, $x_n b$ has no proper suffix which is a prefix of x_n.

Let us consider again the quasiperiod x. The word xb is a factor of $f^\omega(a)$. Since x covers $f^\omega(a)$ and since x starts with a, the word xb has a proper suffix which is a prefix of x. Since $x \neq \varepsilon$, this contradicts what was said about the x_n's.

So $(E\tilde{\varphi}\varphi E)^\omega(a)$ has no quasiperiod. □

Let us observe that the word $\underline{w}$ such that $(E\tilde{\varphi}\varphi E)^\omega(a) = a\underline{w}$ starts with ba and can be decomposed over $\{ba, bba\}$. So it is quasiperiodic with quasiperiod bab.

5 Complexity

The second question of [10] is “What about the complexity function of a quasiperiodic infinite word?”. Let recall that the complexity function $p_{\underline{w}}(n)$ of an infinite word $\underline{w}$ is the function which associates to each integer $n \geq 0$ the number of factors of length n of $\underline{w}$. Our aim is to show that there exists no relation between quasiperiodic words and complexities.

First we consider words with lowest complexity. It is well known (see [8] for instance) that a word $\underline{w}$ has a bounded complexity if and only if $\underline{w} = uv^\omega$ for words $u, v \neq \varepsilon$. When $u = \varepsilon$, v is a quasiperiod of $\underline{w}$. When $u \neq \varepsilon$, $\underline{w}$ can be quasiperiodic as for instance $ab(aba)^\omega$ or it can be non-quasiperiodic as for instance ab^ω.

In [11], it is shown that Sturmian words are the words with lowest unbounded complexity. We know there exist quasiperiodic (the Fibonacci word [10]) and non-quasiperiodic Sturmian words (Section 4).

In [5], Cassaigne characterizes couples of integers (α, β) for which there exists an integer $n_0 \geq 0$ and an infinite word over $\{a, b\}$ having complexity $\alpha n + \beta$ for all $n \geq n_0$. They are the couples in $\{0, 1\} \times (\mathbb{N} \setminus \{0\}) \cup (\mathbb{N} \setminus \{0, 1\}) \times \mathbb{Z}$. When $\alpha \geq 1$, the word given for example by Cassaigne is a quasiperiodic word. More precisely, it is the word $g_{l,j}(\varphi_k^\omega(a_1))$ where j, l, k are suitable integers, $g_{l,j}$ is the morphism defined by $g_{l,j}(a_i) = a^l b^{i+j}$ and φ_k is the morphism defined in Section 3. We leave open the question to find non-quasiperiodic words with these complexities when $\alpha \notin \{0, 1\}$.

We end this section showing that there exist quasiperiodic words with exponential complexity. As already said, in [9], it is shown that all words having all finite words as factors are not quasiperiodic. These words are those with complexity $p(n) = 2^n$ for all $n \geq 0$.

Let $\underline{w}$ be such a word over $\{a, b\}$. Since $\varphi^2(a) = aba$ and $\varphi^2(b) = ab$, the word $\varphi^2(\underline{w})$ is quasiperiodic with quasiperiod aba.

We now evaluate the complexity of the word $\varphi^2(\underline{w})$. For this, let a_n (resp. b_n, c_n) be the number of factors of $\varphi^2(\underline{w})$ ending with b (resp. ba, aa). We have $p(0) = 1$, $p(1) = 2$, $p_{\underline{w}}(n) = a_n + b_n + c_n$ for $n \geq 2$. Since a^3 and b^2 are not factors of $\varphi^2(\underline{w})$, we have $a_2 = b_2 = c_2 = 1$ and for $n \geq 2$, $a_{n+1} = b_n + c_n$, $b_{n+1} = a_n$, $c_{n+1} = b_n$. Consequently for $n \geq 3$, $a_{n+1} = a_{n-1} + a_{n-2}$, $b_{n+1} = b_{n-1} + b_{n-2}$, $c_{n+1} = c_{n-1} + c_{n-2}$. So $p(2) = 3$ and for $n \geq 3$, $p(n + 1) = p(n - 1) + p(n - 2)$.

The first values of the sequence $(p(n))_{n \geq 1}$ are:

$$2, 3, 4, 5, 7, 9, 12, 16, 21, 28, 37, 49, \ldots$$

This sequence is part of the Padovan sequence (see sequence A000931 in [13]) defined by $a_0 = 1$, $a_1 = 0$, $a_2 = 0$ and for $n \geq 3$ $a_n = a_{n-2} + a_{n-3}$ (more precisely $p(n) = a_{n+8}$ for $n \geq 1$). It is known (see [13] for instance) that a_n is asymptotic to

$r^n/(2 * r + 3)$ where $r = 1.3247179572447\ldots$, is the real root of $x^3 = x + 1$ (r is called the plastic constant [14]). So $p(n) = \theta(r^n)$.

To end with complexity, let us quote a new question: what is the maximal complexity of a quasiperiodic infinite word?

6 Quasiperiods of the Fibonacci word

In order to answer other questions of [10], we characterize the quasiperiods of the Fibonacci word (Proposition 6.5). In particular, we show that this word has an infinite number of superprimitive quasiperiods (Proposition 6.6). We start with a useful lemma:

Lemma 6.1. *Let $\underline{w}$ be an infinite word over $\{a, b\}$ and let $x \in \{a, b\}^*$.*

If x is a quasiperiod of $\varphi(\underline{w})$ then x verifies one of the three following properties:

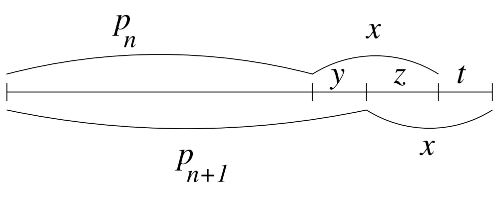
a) $x = \varphi(x')$ where x' is a quasiperiod of $\underline{w}$.

b) $x = \varphi(x')a$ where x' is a quasiperiod of $\underline{w}$.

c) $x = \varphi(x')a$ for a word x' such that $\{x'a, x'b\}$ covers $\underline{w}$.

Proof. Assume that x is a quasiperiod of $\varphi(\underline{w})$. Since $\varphi(\underline{w})$ starts with the letter a and since x is a prefix of $\varphi(\underline{w})$, x starts with a.

First we consider the case where x ends with the letter b. Since x is a factor of $\varphi(\underline{w})$, x ends with ab. Let $(p_n x)_{n\geq 0}$ be a covering sequence of prefixes of $\varphi(\underline{w})$. Let $n \geq 0$. We have $|p_n| < |p_{n+1}| \leq |p_n x|$. So $|p_{n+1}| \leq |p_n x| < |p_{n+1} x|$ and the following situation holds.



There exist words y, z, t such that $x = yz = zt$. Note $|y| = |p_{n+1}| - |p_n| \neq 0$. Since x starts with a and ends with b, there exist words y', z', t', p'_n such that $y = \varphi(y')$, $z = \varphi(z')$, $t = \varphi(t')$, $p_n = \varphi(p'_n)$ and $p_{n+1} = \varphi(p'_n y')$. Since p_n is a prefix of p_{n+1} which is itself a prefix of $p_n x$, we deduce $|p'_n| < |p'_{n+1}| \leq |p'_n x'|$ where $x' = y'z'$. Since φ is injective, we also have $x' = z't'$.

What precedes is valid for all $n \geq 0$ and due to injectivity of φ, the words $(p'_n)_{n\geq 0}$ and x' are defined uniquely. So $(p'_n x')_{n\geq 0}$ is a covering sequence of $\underline{w}$, that is, x' is a quasiperiod of $\underline{w}$. Moreover $x = \varphi(x')$.

(Note that since x ends with b, x' ends with a.)

Now we consider the case where x ends with the letter a.

If xa is a quasiperiod of $\varphi(\underline{w})$, one can see as previously that $x = \varphi(x')$ for a quasiperiod x' of $\underline{w}$.

If each occurrence of x is followed by an occurrence of the letter b, then x ends with the letter a. Since it also starts with a, there exists a word x' such that $x = \varphi(x')a$. Moreover in this current case, xx cannot be a factor of $\varphi(\underline{w})$. So $\varphi(x')$ is a quasiperiod of $\varphi(\underline{w})$. We can deduce that x' is a quasiperiod of $\underline{w}$.

Finally we have to consider the case where x ends with the letter a and some occurrences of x are followed by a and others by b. Thus we cannot say that xa or xb is a quasiperiod of $\underline{w}$. Let $(p_n x)_{n\geq 0}$ be a covering sequence of prefixes of $\varphi(\underline{w})$. For $n \geq 0$, let a_n be the letter such that $p_n x a_n$ is a prefix of $\varphi(\underline{w})$. Let b_n be the letter in $\{a, b\} \setminus \{a_n\}$. There exist a word x' and prefixes p'_n of $\underline{w}$ such that $p_n = \varphi(p'_n)$, $x = \varphi(x')a$. Moreover $xa_n = \varphi(x'b_n)$ when $a_n = b$ and $x = \varphi(x'b_n)$ when $a_n = a$. We can deduce that $(p'_n x' b'_n)_{n\geq 0}$ is a covering sequence of prefixes of w, that is, $\{x'a, x'b\}$ covers $\underline{w}$. ◻

The converse of Lemma 6.1 partially holds. We let the reader verify that:

Lemma 6.2. *If a word x verifies Case a or b in Lemma 6.1 then x is a quasiperiod of $\underline{w}$.*

This does not hold if x fulfils Case c. We can only deduce that $\{xa, xb\}$ covers $\varphi(\underline{w})$. Indeed there exist words $\underline{w}$, x' such that $x = \varphi(x'a)$ does not cover $\varphi(\underline{w})$. For instance if $x' = ab$ and $\underline{w} = abaaba(bba)^\omega$ then $\{x'a, x'b\} = \{aba, abb\}$ covers $\underline{w}$. But $\varphi(x')a = abaa$ does not cover $\varphi(\underline{w}) = abaabaabaaba(baaa)^\omega$.

In order to characterize the quasiperiods of the Fibonacci word $\underline{F}$, we need to study what happens when Case c of Lemma 6.1 occurs.

Lemma 6.3. *The words y such that $\{ya, yb\}$ covers $\underline{F}$ are the words $(u_n)_{n\geq 0}$ defined by $u_0 = \varepsilon$, $u_{n+1} = \varphi(u_n)a$ for $n \geq 1$.*

Moreover for $n \geq 2$, u_n is a quasiperiod of $\underline{F}$.

The proof of this lemma is a direct consequence of the following result and of the fact that $u_2 = aba$ is a quasiperiod of $\underline{F}$.

Lemma 6.4. *Let $\underline{w}$ be an infinite word over $\{a, b\}$ that does not contain bb as factor. Let $y \in \{a, b\}^*$ such that $|y| \geq 1$.*

The set $\{ya, yb\}$ covers $\varphi(\underline{w})$ if and only if $y = \varphi(z)a$ and $\{za, zb\}$ covers w.

Proof. First assume that $\{za, zb\}$ covers w. Since each occurrence of $\varphi(zb)$ in $\varphi(\underline{w})$ is followed by the letter a, the set $\{\varphi(za), \varphi(zb)a\} = \{\varphi(z)ab, \varphi(z)aa\}$ covers $\varphi(\underline{w})$.

Assume now that $\{ya, yb\}$ covers $\varphi(w)$ and $|y| \geq 1$. The word y starts and ends with the letter a. Let $(p_n ya_n)_{n\geq 0}$ be a sequence of prefixes covering $\varphi(\underline{w})$ with $a_n \in \{a, b\}$. There exist words p'_n and z such that $p_n = \varphi(p'_n)$ and $y = \varphi(z)a$. For $n \geq 0$, let $b_n = a$ if $a_n = b$ and $b_n = b$ if $a_n = a$. Let $n \geq 0$. By definition of a covering sequence of prefixes, $|p_n| < |p_{n+1}| \leq |p_n ya_n|$. If $a_n = b$, $|\varphi(p'_n)| < |\varphi(p'_{n+1})| \leq |\varphi(p'_n zb_n)|$. Since p'_n is a prefix of p_{n+1} itself a prefix of $p'_n zb_n$, $|p'_n| < |p'_{n+1}| \leq |p'_n zb_n|$. Observe now that $yaya$ is not a factor of $\varphi(\underline{w})$ (Indeed since y starts and ends with a, this would imply that aaa is a factor of $\varphi(\underline{w})$ but this is not possible since bb is not a factor of $\underline{w}$). Thus when $a_n = a$, $|p_n| < |p_{n+1}| < |p_n ya_n|$, that is, $|p_n| < |p_{n+1}| \leq |p_n y| = |\varphi(p'_n zb_n)|$. Once again $|p'_n| < |p'_{n+1}| \leq |p'_n zb_n|$. So the sequence $(p'_n zb_n)_{n\geq 0}$ is a covering sequence of prefixes of $\underline{w}$, that is, $\{za, zb\}$ covers $\underline{w}$. □

Now we can describe the set of quasiperiods of the Fibonacci word. Let $Q_0 = \{aba\}$ and for $n \geq 1$ $Q_n = \{\varphi(u), \varphi(u)a \mid u \in Q_{n-1}\}$.

Proposition 6.5. *The set of quasiperiods of the Fibonacci word is* $\bigcup_{n\geq 0} Q_n$.

The proof of this proposition is a consequence of Lemmas 6.1 and 6.3.

To illustrate the previous proposition, let us note that the first quasiperiods of $\underline{F}$ are: *aba*, *abaab*, *abaaba*, *abaababa*, *abaababaa*, *abaababaab*, *abaababaaba*.

Let $f_n = \varphi^n(a)$ for $n \geq 0$. We can see by induction that, for any integer $n \geq 2$, $Q_n = \{f_{n+2} f_{i_1} f_{i_2} \ldots f_{i_k} \mid 0 \leq k \leq n, n-1 \geq i_1 > i_2 > \ldots i_k = 0\}$. So Q_n contains 2^n distinct elements. Moreover for $x \in Q_n$, $|f_{n+2}| \leq |x| < |f_{n+3}|$. It follows that $\bigcup_{i=0}^{n} Q_i$ has $2^{n+1} - 1$ distinct elements each of length less than $|f_{n+3}|$.

The reader interested by similar results can consult [6] that provides a description of the quasiperiods of the words f_k considered as circular.

Proposition 6.5 shows in particular that the Fibonacci word has an infinite number of quasiperiods. This could have been obtained observing the particular quasiperiods $\varphi^n(aba)$ for $n \geq 0$.

We now want to state a great difference between quasiperiodic infinite words and quasiperiodic finite words. We have already recalled that any quasiperiodic finite word has a unique superprimitive quasiperiod. We show that the Fibonacci word has an infinite number of superprimitive quasiperiods.

Proposition 6.6. *The set of quasiperiods of $\underline{F}$ are the words $(q_n)_{n\geq 0}$ defined (for $n \geq 0$) by:*

$$q_{2n} = f_{2n+1} \prod_{i=0}^{n} f_{2(n-i)}$$

$$q_{2n+1} = f_{2n+2} \prod_{i=0}^{n} f_{2(n-i)+1}$$

Before proving this proposition let us give the first superprimitive quasiperiods: $q_0 = f_1 f_0 = aba$, $q_1 = f_2 f_1 = abaab$, $q_2 = f_3 f_2 f_0 = abaababaa$, $q_3 = f_4 f_3 f_1 = abaababaabaabab$.

Proof of Proposition 6.6. First we note that all the words q_n are quasiperiods of $\underline{F}$. Indeed $q_0 = aba \in Q_0$ and for $n \geq 0$, $q_{2n+1} = \varphi(q_{2n}) \in Q_{2n+1}$ and $q_{2n+2} = \varphi(q_{2n+1})a \in Q_{2n+2}$. Observe that the sequence $(q_n)_{n\geq 0}$ is a sequence of length increasing words.

We now want to prove that q_n does not cover q_m for any $n < m$. For this note that if n is odd and m is even, q_n ends with b and q_m ends with a. So q_n does not cover q_m. Case n even and m odd is similar. Assume now $n = 2p$ and $m = 2q$. The word q_{2n} ends with $ab \prod_{i=0}^{n} f_{2(n-i)}$ whereas the word q_{2m} ends with $\prod_{i=0}^{n+1} f_{2(n+1-i)}$ and so with $ba \prod_{i=0}^{n} f_{2(n-i)}$. Once again q_{2n} cannot cover q_{2m}. Case n and m both odd is similar.

To end the proof of Proposition 6.6, we need to see that $\underline{F}$ does not have a quasiperiod x that covers q_n for an integer $n \geq 0$ with $|x| < |q_n|$. This can be stated showing by induction that the set of superprimitive quasiperiods of $\underline{F}$ that belong to Q_n is $\{q_i \mid 0 \leq i \leq n\}$. This is a consequence of Lemma 6.2. $\square$

7 About set of quasiperiods

In this section, we consider the third and four questions in [10]. The third one is an open question: "What about the set of quasiperiods of an infinite word?".

As seen in previous section, there exists at least one word (the Fibonacci word) which has an infinite number of quasiperiods. It is easy to construct other such examples. Indeed taking two infinite words u and v over $\{a, b\}$ having the same quasiperiod z starting with the letter a, considering the morphism f defined by $f(a) = u$, $f(b) = v$, the fixed point $f^\omega(a)$ has the quasiperiods $f^n(z)$ for any $n \geq 0$. Indeed for any word w having a quasiperiod x, $f(w)$ has $f(x)$ as quasiperiod.

We now want to show that there are a lot of intermediate cases between infinite words having no quasiperiod [10] and infinite words having an infinite number of superprimitive quasiperiods (as the Fibonacci word).

Lemma 7.1. *Let* $\underline{w} = abaaba(bba)^\omega$*. For* $n \geq 0$*, the word* $\varphi^{n+1}(\underline{w})$ *has* $\bigcup_{i=0}^{n} Q_i$ *as set of quasiperiods and* $q_0, \ldots, q_n$ *as superprimitive quasiperiods.*

We let the reader prove this result using lemmas of the previous section.

In [10], Marcus defines "the quasiperiodicity of order p that where the intersection of two different occurrences of a superprimitive quasiperiod is never larger than p, but sometimes it is equal to p". He provides examples of words for each positive order. He asks: "Does there exist a quasiperiodic infinite word which is of no order p ($p = 1, 2, 3, \ldots$)? The Fibonacci word provides a positive answer.

Lemma 7.2. *The Fibonacci word has no order.*

Proof. The covering sequences of prefixes of $\underline{F}$ associated with the quasiperiod $q_0 = aba$ starts with $\varepsilon, aba, abaab$. In particular the second and third occurrences of aba overlap and the overlap has length $1 = |f_0|$. We have seen in the proof of Proposition 6.6 that $q_{2n} = \varphi(q_{2n-1})$ and $q_{2n+1} = \varphi(q_{2n})a$. Thus by induction on $n \geq 0$, one can see that q_n has two occurrences overlapping with an overlap of length at least $|f_n|$. So the Fibonacci word has no order. ◻

8 Conclusion

The two last questions of [10] concerns infinite words such that all their factors are also quasiperiodic. These questions should certainly be precised or transformed since any prefix of length 1 of a non-empty word is not quasiperiodic. One possible way to transform Question 6 (Does there exist a non-quasiperiodic infinite word such that all its factors are quasiperiodic?) is to search for a non-quasiperiodic infinite word having an infinity of quasiperiodic prefixes. We provide such an example.

Let $(u_n)_{n\geq 0}$, $(v_n)_{n\geq 0}$ be the sequences of words defined by $u_0 = aa$, and for $n \geq 0$, $v_n = u_n b$, $u_{n+1} = v_n^2$. So $u_0 = aa$, $v_0 = aab$, $u_1 = aabaab$, $v_1 = aabaabb$, $u_2 = aabaabbaabaabb$, $v_2 = aabaabbaabaabbb$. Of course, since each u_n is a square it is a quasiperiodic word. By induction one can see that for any $n \geq 0$, the word b^{n+1} occurs only once in v_n as a suffix. So v_n is superprimitive. Consequently the word $\lim_{n\to\infty} u_n = \lim_{n\to\infty} v_n$ is a non-quasiperiodic infinite word having infinitely many quasiperiodic prefixes.

Acknowledgments. The authors thank Francis Wlazinski for fruitful discussions.

References

[1] A. Apostolico and M. Crochemore. String pattern matching for a deluge survival kit. In J. Abello, P.M. Pardalos, and M.G.C. Resende, editors, *Handbook of Massive Data Sets*. Kluwer Academic Publishers, 2001.

[2] A. Apostolico and A. Ehrenfeucht. Efficient detection of quasiperiodicities in strings. *Theoretical Computer Science*, 119:247–265, 1993.

[3] A. Apostolico, M. Farach, and C. S. Iliopoulos. Optimal superprimitivity testing for strings. *Information Processing Letters*, 39(1):17–20, 1991.

[4] J. Berstel and P. Séébold. *Algebraic Combinatorics on Words (*M. Lothaire), volume 90, chapter 2. Sturmian words. Cambridge Mathematical Library, 2002.

[5] J. Cassaigne. Complexité et facteurs spéciaux. *Bulletin of the Belgian Mathematical Society*, 4:67–88, 1997.

[6] C. S. Iliopolos, D. Moore, and W. F. Smyth. The covers of a circular Fibonacci string. *J. Combinatorial Math. and Combinatorial Computing*, 26:227–236, 1998.

[7] M. Lothaire. *Algebraic Combinatorics on words*, volume 90 of *Encyclopedia of Mathematics*. Cambridge University Press, Cambridge, UK, 2002.

[8] M. Lothaire. *Applied Combinatorics on Words*. To appear. (see http://www-igm.univ-mlv.fr/˜berstel).

[9] S. Marcus. Bridging two hierarchies of infinite words. *Journal of Universal Computer Science*, 8:292–296, 2002.

[10] S. Marcus. Quasiperiodic infinite words. *Bulletin of the European Association for Theoretical Computer Science (EATCS)*, 82:170–174, 2004.

[11] M. Morse and G.A. Hedlund. Symbolic Dynamics II: Sturmian trajectories. *Amer. J. Math.*, 61:1–42, 1940.

[12] P. Séébold. Fibonacci morphisms and Sturmian words. *Theoretical Computer Science*, 88:365–384, 1991.

[13] N. J. A. Sloane. The on-line encyclopedia of integer sequences. http://www.research.att.com/˜njas/sequences/index.html.

[14] E. W Weisstein. Plastic constant. From *MathWorld* – A Wolfram Web resource. http://mathworld.wolfram.com/PlasticConstant.html.