Determining Sets of Quasiperiods of Infinite Words

Guilhem Gamard, Gwenaël Richomme

HAL Id: lirmm-01369373
https://hal-lirmm.ccsd.cnrs.fr/lirmm-01369373v1

Submitted on 20 Sep 2016

Determining Sets of Quasiperiods of Infinite Words

Guilhem Gamard[1] and Gwenaël Richomme[2]

1   LIRMM (CNRS, Univ. Montpellier)
    UMR 5506 - CC 477, 161 rue Ada, 34095, Montpellier Cedex 5, France
2   LIRMM (CNRS, Univ. Montpellier)
    UMR 5506 - CC 477, 161 rue Ada, 34095, Montpellier Cedex 5, France and
    Univ. Paul-Valéry Montpellier 3,
    Dpt MIAp, Route de Mende, 34199 Montpellier Cedex 5, France

Abstract

A word is quasiperiodic if it can be obtained by concatenations and overlaps of a smaller word, called a quasiperiod. Based on links between quasiperiods, right special factors and square factors, we introduce a method to determine the set of quasiperiods of a given right infinite word. Then we study the structure of the sets of quasiperiods of right infinite words and, using our method, we provide examples of right infinite words with extremal sets of quasiperiods (no quasiperiod is quasiperiodic, all quasiperiods except one are quasiperiodic, ...). Our method is also used to provide a short proof of a recent characterization of quasiperiods of the Fibonacci word. Finally we extend this result to a new characterization of standard Sturmian words using a property of their sets of quasiperiods.

1 Introduction

Many studies around words focus on the task of measuring regularities of strings. Various notions were introduced to that end, the strongest one being *periodicity*. Recall that an infinite word is *periodic* if it is obtained by infinite concatenation of occurrences of a word u, called a *period*.

In the context of text algorithms, Apostolico and Ehrenfeucht introduced [1] the notion of *quasiperiodicity*, which is a generalization of periodicity for finite words. A word $w = w_1 w_2 \ldots w_n$ is *quasiperiodic* if there exists another word $q \neq w$ (called a *cover* or a *quasiperiod* of w) such that w is covered with occurrences of q. More precisely, for all $i \in \{1, \ldots, n\}$, there exists $k \in \{0, \ldots, |q| - 1\}$ such that $w_{i-k} w_{i-k+1} \ldots w_{i-k+|q|-1}$ is an occurrence of q. For instance, the string "*ababa abababa ababababa ababa*" has quasiperiods *aba* and *ababa*, but it is not periodic.

This definition generalizes immediately to right infinite words (see [18]). As finite words may have several quasiperiods, infinite words may have infinitely many quasiperiods. Words with infinitely many quasiperiods are called *multi-scale quasiperiodic* (see [19]). In Section 2 we state a characterization of the set of quasiperiods of an infinite aperiodic word showing links between some extremal quasiperiods and some square factors and right special factors. An important consequence of this result is to provide a general method to determine the

set of quasiperiods of any infinite word. Next we describe several examples of uses of this method.

In [6], Christou, Crochemore and Iliopoulos provide characterizations of quasiperiods of Fibonacci strings. One of their motivations was that "Fibonacci strings are important in many concepts [3] and are often cited as a worst case example for many string algorithms." However, Fibonacci strings are not always the best words for this purpose. For example, Groult and Richomme [11] proved that the algorithms provided by Brodal and Pedersen [5] and by Iliopoulos and Mouchard [12] to compute all the quasiperiods of a word do not reach their worst case on Fibonacci strings. They proved that those algorithms were optimal, and provided a family of strings reaching the worst case. Nevertheless, the study of finite Fibonacci strings is indeed of great interest: see, e.g., references in [6].

Some of the results from [6] were recently reformulated by Mousavi, Schaeffer and Shallit as a new characterization of quasiperiods of the infinite Fibonacci word [21] (another one was given in [14]). They use this result, among many others, to show how to build automated proofs of some results about the Fibonacci word. Using the method to determine the set of quasiperiods of any infinite word described in Section 2, we provide a short proof of the above mentioned characterization (Section 4.1).

The infinite Fibonacci word is a special case of Sturmian words (and therefore Fibonacci strings are special cases of factors of Sturmian words). A natural question is whether the previous characterization of quasiperiods of the Fibonacci word can be extended to other Sturmian words. Unfortunately, some Sturmian words are not quasiperiodic [15]; more precisely, a Sturmian word is quasiperiodic if and only if it is not a Lyndon word. However, we can still extend our characterization to *standard* Sturmian words, i.e. Sturmian words having all their left special factors as prefixes (Section 4.3).

Sturmian words are not necessarily quasiperiodic, but their bi-infinite counterparts are always multi-scale quasiperiodic. This result can be extended to subshifts, i.e. topological spaces generated from languages by the shift operation. A subshift is *quasiperiodic* (resp. *multi-scale quasiperiodic*) if and only if it is generated by a word which is quasiperiodic (resp. multi-scale quasiperiodic). Monteil and Marcus proved [19] that all Sturmian subshifts are multi-scale quasiperiodic. They also proved that multi-scale quasiperiodic shifts have zero topological entropy, are minimal (their words are uniformly recurrent), and that all of their factors have frequencies (see [9] for a generalization to two-dimensional words).

The main tool of [19] is a so-called *derivation* operation, which takes the inverse image of a word by a well-chosen morphism. The derivative of a quasiperiodic word **w** is another word which describes the lengths of the overlaps in **w** between each two consecutive occurrences of its quasiperiods. While reading [19], one naturally asks whether the derivation operation preserves multi-scale quasiperiodicity. In other terms, given a multi-scale quasiperiodic word, does its derivative still has infinitely many quasiperiods? In Section 3, we show this is not the case. We provide a right infinite multi-scale quasiperiodic word whose derivative is non-quasiperiodic. While discussing properties of the derivation operation, we also provide a word such that each quasiperiod has the previous (in terms of length) one as a quasiperiod. This nested effect can be avoided; we provide a multi-scale quasiperiodic word with only non-quasiperiodic quasiperiods. The proof of properties of our examples all involve our general method.

Let us summarize the main parts of our paper. In Section 2, we present general properties of quasiperiods of right infinite words and a general method to determine them. In Section 3, we present our results around the derivation operation. In Section 4, we provide our proof of the characterization of the quasiperiods of the Fibonacci word and its generalization to the new characterization of standard Sturmian words.

We assume readers are aware of general results and definitions in Combinatorics on Words (see for instance [16, 17]). Let us recall basic definitions. All infinite words we consider are right infinite words over finite alphabet. Given an alphabet A, A^* (resp. A^ω) denotes the set of finite (resp. infinite) words. The empty word is denoted by ε. The length of a finite word u is denoted $|u|$. A word u is a *factor* of a word v if $v = pus$ for some words p and s. When p is empty (resp. s is empty), u is a *prefix* (resp. a *suffix*) of v. When p and s are not empty, u is an *internal factor* of v. If there exists at least two different letters α and β such that $u\alpha$ and $u\beta$ are factors of v, u is a *right special factor* of v. When αu and βu are factors of v, u is a *left special factor* of v. A *bispecial factor* is a factor which is both left and right special.

2 Basic properties of sets of quasiperiods

Here we provide materials that help to determine quasiperiods of an infinite word poiting first the two main results of this section. Our first result characterizes periodic infinite words using their sets of quasiperiods. We give its proof later, as a consequence of Proposition 2.3, an intermediate result to show Theorem 2.2.

▶ **Proposition 2.1.** *An infinite word is periodic if and only if all its sufficiently long prefixes belong to its set of quasiperiods.*

Consider an infinite word $\mathbf{w}$. As any quasiperiod of $\mathbf{w}$ is one of its prefixes, $\mathbf{w}$ has at most one quasiperiod of length n for each integer n. Thus the knowledge of the lengths of quasiperiods is equivalent to the knowledge of the quasiperiods themselves. After Proposition 2.1, it is clear that lengths of quasiperiods of an aperiodic infinite word, if there are any, are distributed into intervals of integers which are finite and disjoint (possibly singletons). The next theorem characterizes these intervals.

▶ **Theorem 2.2.** *Let $\mathbf{w}$ be an infinite aperiodic word and, for any integer i, let p_i denote the prefix of length i of $\mathbf{w}$. The set of lengths of quasiperiods of $\mathbf{w}$ is an union (possibly empty) of disjoint intervals of integers. If $[i, j]$ is such an interval, then there are no quasiperiods of lengths $i - 1$ nor $j + 1$. Moreover,*
- *$p_i p_i$ is a factor of $\mathbf{w}$;*
- *p_{i-1} is not an internal factor of $p_i p_{i-1}$;*
- *p_j is a right special factor of $\mathbf{w}$;*
- *for all k such that $i \le k < j$, the word p_k is not a right special factor of $\mathbf{w}$.*

This theorem induces a method to determine the quasiperiods of an aperiodic infinite word $\mathbf{w}$. This method consists in determining first among all prefixes p such that pp is a factor of $\mathbf{w}$ or such that p is right special, those such that $\mathbf{w}$ is p-quasiperiodic. With this information, using Theorem 2.2, we can deduce that:
- for any prefix p of $\mathbf{w}$ such that $\mathbf{w}$ is p-quasiperiodic and pp is a factor of $\mathbf{w}$, if q is the smallest right special prefix of $\mathbf{w}$ longer than p, then all prefixes π such that $|p| \le |\pi| \le |q|$ are quasiperiods of $\mathbf{w}$;
- for any prefix q of $\mathbf{w}$ which is a quasiperiod and right special, if p is the smallest prefix such that $|p| > |q|$ and pp is a factor of $\mathbf{w}$, then all prefixes π such that $|q| < |\pi| < |p|$ are *not* quasiperiods of $\mathbf{w}$ (observe that when $\mathbf{w}$ has finitely many quasiperiods, p does not exist).

Moreover, any multi-scale quasiperiodic word which is not periodic has infinitely many quasiperiods which are right special factors. If an infinite word has finitely many quasiperiods, then the longest one is right special. The converse is not true: some multi-scale quasiperiodic

words have infinitely many right special factors which are not quasiperiods. Construction of such counter-examples is left to the readers.

In Theorem 2.2, the existence of disjoint intervals is ensured by aperiodicity and Proposition 2.1. The first two items are consequence of Proposition 2.4 below, and the last two items are a consequence of Proposition 2.3. The end of this section is dedicated to the proof of these propositions.

▶ **Proposition 2.3.** *Let* **w** *be an infinite word with a quasiperiod* q. *Let* a *be the letter such that* qa *is a prefix of* **w**. *The word* qa *is a quasiperiod of* **w** *if and only if* q *is not right special.*

Proof. If q is not a right special factor of **w** then each of its occurrences is followed by the letter a and so, since q is a quasiperiod of **w**, qa is also a quasiperiod of **w**.

Conversely assume that qa is a quasiperiod of **w** and let b be a letter such that qb is a factor of **w**. Then qb is a factor of a qa-quasiperiodic factor u of **w**. For a length reason, we can assume that u is the overlap of two occurrences of qa. Therefore there exist words p_1, p_2, s_1 and s_2 such that $u = p_1\,qb\,s_1 = qa\,s_2 = p_2\,qa$ with $0 < |p_2| \leq |qa|$ and $0 < |s_2| \leq |qa|$. If either p_1 or s_1 is the empty word then $a = b$, so assume these words are not empty. By a classical result (see for instance [16, Prop. 1.3.4]), the equation $p_2\,qa = qa\,s_2$ implies that there exist words x and y and an integer k such that $p_2 = xy$ and $qa = (xy)^k x$ and $s_2 = yx$. In particular, qa is a factor of the periodic word $(xy)^\omega$. Equation $p_1\,qb\,s_1 = p_2\,qa$ implies that qb is also a factor of $(xy)^\omega$. As $|qa| = |qb| \geq |xy|$ (recall that $xy = p_2$ and $|qa| \geq |p_2|$), we conclude that $a = b$, so the word q is not right special factor of **w**. ◀

As shown below, Proposition 2.1 is a corollary of Proposition 2.3.

Proof of Proposition 2.1. Let **w** be an infinite word. If **w** is periodic with period of length n, then any prefix of **w** with length at least n is a quasiperiod of **w**.

Conversely assume that, for an integer n, all prefixes with length at least n are quasiperiods of **w**. For $i \geq 0$, let p_i be the prefix of length i. As **w** is p_n-quasiperiodic, **w** has at least two occurrences. There exists a word u such that both p_n and up_n are prefixes of **w**. By hypothesis and Proposition 2.3, p_n is not right special. Each of its occurrences extend to p_{n+1}. Hence both p_{n+1} and up_{n+1} are prefixes of **w**. Iterating this argument, for all $i \geq n$, both p_i and up_i are prefixes of **w**. Thus **w** $= u$**w**: **w** is periodic with period u. ◀

Proposition 2.3 provides a first piece of information on some extremal quasiperiods of a word. The next result provides further information. It generalizes an observation made in [14] for the smallest quasiperiod of a word.

▶ **Proposition 2.4.** *Assume that* qa *is a quasiperiod of an infinite word* **w** *for some word* q *and some letter* a. *The word* q *is not a quasiperiod of* **w** *if and only if the word* qaq *is a factor of* **w** *and* q *is not an internal factor of* qaq.

Proof. If qaq is not a factor of **w**, then each occurrence of qa is properly overlapped by the next occurrence of qa. This implies that each occurrence of q is overlapped by or concatenated to the next occurrence of q, that is **w** is q-quasiperiodic.

The converse is immediately true. If a word **w** contains qaq as a factor and if q is not an internal factor of qaq, then q cannot be a quasiperiod of **w**. ◀

It should be observed for understanding Theorem 2.2 that under the hypotheses "qa is a quasiperiod of **w**" and "q is not an internal factor of qaq", the word qaq is a factor of **w** if and only if $qaqa$ is a factor of **w**.

3 On multiscale properties

As explained in the introduction, the goal of this section is threefold. First we provide examples to illustrate the usage of Theorem 2.2. Second we want to show that the derivation operation introduced in [19] does not preserve multiscale quasiperiodicity. Third we aim to study the structure of the relation "is a quasiperiod of" for multiscale quasiperidic words.

A finite word u is said to be *superprimitive* if it is not quasiperiodic. It can be seen (for instance as a consequence of the study made in [14]) that the Fibonacci word has both infinitely many superprimitive quasiperiods and infinitely many non-superprimitive quasiperiods. Moreover, due to its morphic structure, the Fibonacci word has an infinite sequence of *nested quasiperiods*. In other terms, it has a sequence of quasiperiods $(q_n)_{n \geq 0}$ such that, for each $n \geq 0$, q_{n+1} is q_n-quasiperiodic. There are many possible structures for the relation "is a quasiperiod of" inside multiscale quasiperiodic words. We provide several extremal examples throughout this section.

3.1 A multiscale quasiperiodic word with only one superprimitive quasiperiod

▶ **Proposition 3.1.** *There exists a multiscale quasiperiodic word having only one superprimitive quasiperiod.*

This proposition is a direct corollary of the next lemma.

Recall that a morphism between two sets of words A^* and B^* (with A, B finite alphabets) is an application which commutes with concatenation. A morphism is entirely defined by the images of the letters. A morphism is called *non-erasing* if no image of letters is the empty word. As usual, for a word u and a morphism h, we denote $h^\omega(u)$ the word $\lim_{n \to \infty} h^n(u)$ when it exists.

▶ **Lemma 3.2.** *Let h be the morphism from $\{a, b\}^*$ to $\{a, b\}^*$ defined by $h(a) = abaaba$ and $h(b) = bababa$. The quasiperiods of $h^\omega(a)$ are exactly the words $h^n(aba)$ for $n \geq 0$, and all these words are aba-quasiperiodic.*

Proof. This lemma is a direct consequence of Theorem 2.2 and the following four steps. Indeed Steps 1 to 3 determine the prefixes p of $h^\omega(a)$ such that pp is a factor of $h^\omega(a)$ and $h^\omega(a)$ is p-quasiperiodic. Step 4 and Theorem 2.2 allow to conclude that words $h^n(aba)$ are the only quasiperiods of $h^\omega(a)$.

Step 1: The prefixes p such that pp is a factor of $h^\omega(a)$ are $h^n(a)$, $h^n(ab)$ and $h^n(aba)$.

Let p be such a word. If $|p| \leq 6$, it can be checked that $p \in \{a, ab, aba, h(a)\}$. Assume $|p| > 6$. The word p has $abaabab$ as a prefix. Observe that if $\pi abaabab$ is a prefix of $h^\omega(a)$ for a word π, then necessarily $\pi \in h(A^*)$. As pp is a factor of $h^\omega(a)$, there exist words π and p' such that $h(\pi)pp$ is a prefix of $h^\omega(a)$ and $p = h(p')$. Hence $h(\pi p' p')$ is a prefix of $h^\omega(a)$. As $h(a)$ and $h(b)$ are not prefixes of one another, $\pi p' p'$ is a prefix of $h^\omega(a)$. As $|\pi p' p'| < |h(\pi p' p')|$, the proof of this step ends by induction.

Step 2: Words $h^n(a)$ and $h^n(ab)$ are not quasiperiods of $h^\omega(a)$.

Assume $h^n(a)$ is a quasiperiod of $h^\omega(a)$ for a smallest integer n. Observe $n \neq 0$ and $n \neq 1$. Let $(\pi_k)_{k \geq 0}$ be a sequence of words such that $\pi_k h^n(a)$ is a prefix of $h^\omega(a)$ and, for $k \geq 0$, $|\pi_{k+1}| - |\pi_k| \leq |h^n(a)|$. Since $abaabab$ is a prefix of $h^n(a)$ (as $n \geq 2$), for each $k \geq 0$, $\pi_k = h(\pi'_k)$ for a word π'_k. Observe that $(\pi'_k h^{n-1}(a))_{k \geq 0}$ is a sequence of prefixes of $h^\omega(a)$ and $|\pi'_{k+1}| - |\pi'_k| = \frac{|\pi_{k+1}| - |\pi_k|}{6} \leq \frac{|h^n(a)|}{6} = |h^{n-1}(a)|$. Hence $h^{n-1}(a)$ is a quasiperiod of $h^\omega(a)$. This contradicts the choice of n.

Similarly one can prove that words $h^n(ab)$ are not quasiperiods of $h^\omega(a)$.

Step 3: Words $h^n(aba)$ are quasiperiods of $h^\omega(a)$.

This is a direct consequence of the fact that aba is a quasiperiod of any word in $h(a\{a,b\}^\omega)$ and so of $h^\omega(a)$.

Step 4: Words $h^n(aba)$ are right special factors of $h^\omega(a)$.

This is direct consequence of the fact that aba is a right special factor of $h^\omega(a)$ and $h(a)$ and $h(b)$ begin with different letters.

By the previous steps and Theorem 2.2, factors $h^n(aba)$ are both beginnings and endings of intervals of quasiperiods. Therefore, they are the only quasiperiods of $h^\omega(a)$. ◀

Let us observe that, as aba is a quasiperiod of $h(aba)$, for any $n \geq 1$, $h^n(aba)$ is $h^{n-1}(aba)$-quasiperiodic. Thus not only $h^\omega(a)$ has a unique superprimitive quasiperiod but its sequence of quasiperiods (sorted by increasing length) is a sequence of nested quasiperiods.

3.2   About normal form and derivation

In [20], Mouchard introduced two normal forms to decompose a quasiperiodic (finite) word. A *border* of a nonempty word u is a factor different from u which is both a prefix and a suffix of u. Let $\mathcal{B}(q)$ be the set of borders of q; let $\mathcal{L}(q)$ be the set of words u such that $q = uv$ with $v \in \mathcal{B}(q)$; and let $\mathcal{R}(q)$ be the set of words u such that $q = vu$ with $v \in \mathcal{B}(q)$. Note that the empty word belongs to $\mathcal{B}(q)$ and q belongs to $\mathcal{L}(q) \cap \mathcal{R}(q)$. Any q-quasiperiodic finite word can be decomposed as a concatenation of elements of $\mathcal{L}(q)$ (or as a concatenation of elements of $\mathcal{R}(q)$). Mouchard proved that if q is superprimitive then the decomposition over $\mathcal{L}(q)$ (resp. over $\mathcal{R}(q)$) is unique. This decomposition is called the *left normal form* (resp. *right normal form*) of the word.

As observed by Marcus and Monteil [19] this result extends naturally to infinite words. They introduced a derivation operation. Observe that any word has at most one element of $\mathcal{L}(q)$ of each length. If $\mathbf{w}$ is decomposed over $\mathcal{L}(q)$ and if $(\ell_n)_{n \geq 0}$ is the decomposition, then the *left derivated word* is the word $(|q| - |\ell_n|)_{n \geq 0}$ (written over the alphabet $\{0, \ldots, |q| - 1\}$). Marcus and Monteil showed that this derivation operation is a desubstitution operation, that is, the inverse operation of taking the image of an infinite word under a morphism. This morphism, that we called (following the idea of [19]) the *left integrating morphism*, is defined from $\{0, \ldots, |q| - 1\}^*$ to A^* by mapping i on the prefix of length $|q| - i$ of q.

For instance, consider the *Fibonacci word* $\mathbf{F}$, that is the fixed point of the *Fibonacci morphism* φ defined by $\varphi(a) = ab$ and $\varphi(b) = a$. We know it is aba-quasiperiodic (as $\mathbf{F}$ does not contain aaa and bb as factors and starts with ab - see also [14]). With this quasiperiod $q = aba$, the morphism used to derivate any q-quasiperiodic word is the morphism defined by $0 \mapsto aba$, $1 \mapsto ab$ which, up to a renaming of letters, is φ^2. Hence $\mathbf{F}$ is its own derivative word, and therefore can be derivated arbitrarily many times.

Because of this terminology of "derivation", one could expect that, if a word is multiscale quasiperiodic, then it could be derivated infinitely many times. The next result, combined with Lemma 3.2, disproves this intuition.

▶ **Proposition 3.3.** *Let $\mathbf{w}$ be a quasiperiodic word such that for all quasiperiods q there is no quasiperiod of length $|q| + 1$. Given any quasiperiod q of $\mathbf{w}$, the corresponding left derivated word is not quasiperiodic.*

Proof. Let q be any quasiperiod of $\mathbf{w}$ and let $\mathbf{x}$ be the corresponding left derivated word. We denote by ν the morphism underlying the derivation: $\mathbf{w} = \nu(\mathbf{x})$. By construction of ν, for all letters α in $\{0, \ldots, |q| - 1\}$, $\nu(\alpha)$ begins with the first letter, say a, of $\mathbf{w}$. Assume that Q is

a quasiperiod of $\mathbf{x}$. Then $\nu(Q)$ is a quasiperiod of $\mathbf{w}$ (as ν is a non-erasing morphism - basic fact mentioned in [15]). As $\nu(Q)$ is always followed by the letter a: $\nu(Q)a$ is a quasiperiod of $\mathbf{w}$, a contradiction with the hypothesis. ◄

As the derivation operation is associated to the left normal form, one can ask whether a similar definition associated to the right normal form could give a better behavior. To any q-quasiperiodic word, we call *right derived word*, the word $(|q| - |r_n|)_{n \geq 0}$ where $(r_n)_{n \geq 0}$ is the decomposition of $\mathbf{w}$ over $\mathcal{R}(q)$. We call *right integrating morphism*, the morphism defined by mapping i on the suffix of length $|q| - i$ of q. The example of the Fibonacci word, developed below, shows that right derivation does not preserve multiscale quasiperiodicity.

The smallest quasiperiod of the Fibonacci word is aba. The corresponding right integrating morphism is the morphism μ defined by $\mu(a) = aba$; $\mu(b) = ba$. One can observe that $\varphi^2(a)aba = aba\mu(a)$ and $\varphi^2(b)aba = aba\mu(b)$. Thus $\varphi^2(u)aba = aba\mu(u)$ for any word u. Applying the previous formula for arbitrary large prefixes of $\mathbf{F}$, we get $\varphi^2(\mathbf{F}) = aba\mu(\mathbf{F})$, that is $\mathbf{F} = \mu(a\mathbf{F})$. The right derived word of $\mathbf{F}$ is the word $a\mathbf{F}$. This word is a Lyndon infinite word and consequently it is not quasiperiodic (see [15]).

We end this section with an example of word for which both left and right derivation does not provide a multiscale quasiperiodic word. The proof of these properties is omitted (but still can be done using our general method). Let us consider the following four morphisms f, g, λ, χ, and the word $\mathbf{w}_{fg}$ defined by $\mathbf{w}_{fg} = f(g^\omega(a))$ and

$$f : \begin{cases} a \mapsto aba \\ b \mapsto ba \end{cases} \qquad g : \begin{cases} a \mapsto aba \\ b \mapsto bba \end{cases} \qquad \lambda : \begin{cases} a \mapsto aba \\ b \mapsto ab \end{cases} \qquad \chi : \begin{cases} a \mapsto baa \\ b \mapsto bab \end{cases}$$

► **Lemma 3.4.** *The word $\mathbf{w}_{fg}$ is equal to $\lambda(\chi^\omega(b))$. It is multiscale quasiperiodic and its quasiperiods are the words $f(g^n(a)) = \lambda(\chi^n(a))$. For any quasiperiod q of $\mathbf{w}_{fg}$, the right derived word of $\mathbf{w}_{fg}$ is $g^\omega(a)$ and its left derived word is $\chi^\omega(b)$. Both words $g^\omega(a)$ and $\chi^\omega(b)$ are not quasiperiodic.*

By lack of place, the proof of this lemma is omitted.

One can verify that $\mathbf{w}_{fg}$ is a fixed point of the morphism defined by $h(a) = a$ and $h(b) = babaab$ (this property is a consequence of $h(f(u)) = f(g(u))$ for all words u). This opens a new question: can any multiscale quasiperiodic word $\mathbf{w}$ be desubstituted into another multiscale quasiperiodic word?

3.3 A multiscale quasiperiodic word with all quasiperiods superprimitive

Let $q = abbababba$ and consider morphism ψ defined by:

$$\begin{cases} \psi(a) = (abbab)^7 = abbababbababbababbababbababbababbab \\ \psi(b) = bababba(q)^2(abbab)^2 = bababbaabbababbaabbababbaabbababbab \end{cases}$$

► **Proposition 3.5.** *The quasiperiods of the infinite word $\psi^\omega(a)$ are the words $\psi^n(q)$ with $n \geq 0$. Moreover each of these quasiperiods is superprimitive.*

This proposition is a synthesis of the next three lemmas.

► **Lemma 3.6.** *The word $\psi^\omega(a)$ is $\psi^n(q)$-quasiperiodic for each $n \geq 0$.*

Proof. As already recalled in the proof of Proposition 3.3, for any non-erasing morphism f and any infinite word $\mathbf{w}$, if $\mathbf{w}$ is q-quasiperiodic then $f(\mathbf{w})$ is $f(q)$-quasiperiodic. Hence to prove the lemma, we just need to prove that $\psi^\omega(a)$ is q-quasiperiodic.

As both words obtained from $\psi(a)$ and $ab\psi(b)$ removing their last b are q-quasiperiodic, for any infinite word $\mathbf{w}$, $\psi(a\mathbf{w})$ is q-quasiperiodic. In particular $\psi^\omega(a)$ is q-quasiperiodic. ◄

▶ **Lemma 3.7.** *For any $n \geq 0$, the word $\psi^n(q)$ is superprimitive.*

Proof. Assume by contradiction that n is the least integer such that $\psi^n(q)$ is quasiperiodic, and let Q be one of its quasiperiods. Necessarily $n \geq 1$. The word $\psi(a)ba$ is a prefix of $\psi^n(q)$. An exhaustive verification shows that a prefix of $\psi(a)ba$ is a border of $\psi^n(q)$ if and only if this prefix is of the form $(abbab)^\ell$ with $\ell \in [1; 7]$ when $n = 1$ and $\ell \in \{1, 2\}$ when $n \geq 2$. As any $abbab$-quasiperiodic word cannot contain the word aa as a factor, no prefix of $\psi(a)ba$ can be a quasiperiod of $\psi^n(q)$. It follows that $\psi(a)ba$ must be a prefix of Q.

Observe that if $\psi(a)ba$ is a factor of the image by ψ of a word (finite or infinite) u, then any occurrence of $\psi(a)ba$ in $\psi(u)$ corresponds to a prefix of the image of a suffix of u. Consequently, considering the last occurrence of Q in $\psi^n(q)$, we then deduce that $Q = \psi(Q')$ for some word Q'. That Q is a quasiperiod of $\psi^n(q)$ means there exists a double sequence of words $(p_i, s_i)_{1 \leq i \leq k}$ such that $\psi^n(q) = p_i Q s_i$ for each i in $[1; k]$, $p_1 = \varepsilon = s_k$ and, for each i in $[1; k-1]$, $|p_i Q| \geq |p_{i+1}| > |p_i|$. The observation at the beginning of the paragraph implies that, for each i in $[1; k]$, $p_i = \psi(p_i')$ for some word p_i'. As $Q = \psi(Q')$ and as images of letters by ψ have all the same length, for each i in $[1; k]$ $s_i = \psi(s_i')$ for some word s_i'. Injectivity of ψ implies that for each i in $[1; k]$, $\psi^{n-1}(q) = p_i' Q' s_i'$. Moreover $p_1' = \varepsilon = s_k'$. Observe for each i in $[1; k]$, $|p_i| = 35|p_i'|$ and $|q| = 35|Q'|$. Hence for each i in $[1; k-1]$ $|p_i'Q| \geq |p_{i+1}'| > |p_i'|$. Hence $\psi^{n-1}(a)$ is Q'-quasiperiodic. This contradicts the minimality in the choice of n. ◀

▶ **Lemma 3.8.** *If Q is a quasiperiod of $\psi^\omega(a)$ then $Q = \psi^n(q)$ for some integer $n \geq 0$.*

Proof. Observe that q is right special in $\psi^\omega(a)$, and so, as $\psi(a)$ and $\psi(b)$ begin with different letters, for all $n \geq 0$, the word $\psi^n(q)$ is right special. Thus by Theorem 2.2, we just have to prove that, if Q is a quasiperiod of $\psi^\omega(a)$ and QQ is a factor of $\psi^\omega(a)$, then $Q = \psi^n(q)$ for some integer $n \geq 0$. We use arguments similar to those used in the proof of Lemma 3.7.

The word $\psi(a)ba$ is a prefix of $\psi^\omega(a)$. An exhaustive verification shows that among all prefixes of this word, only q is a quasiperiod of $\psi^\omega(a)$. Let us assume that Q is a quasiperiod of $\psi^\omega(a)$ with $|Q| \geq |\psi(a)ba|$ and QQ a factor of $\psi^\omega(a)$. As in the proof of Lemma 3.7, we observe that if $u\psi(a)ba$ is a prefix of $\psi^\omega(a)$ then $u = \psi(v)$ for some word v. Thus this also holds if uQ is a prefix of $\psi^\omega(a)$. From the fact that QQ is a factor of $\psi^\omega(a)$, we deduce $Q = \psi(Q')$ for a word Q'. Moreover, possibly acting more precisely as in the proof of Lemma 3.7, we can see that Q' must be a quasiperiod of $\psi^\omega(a)$ with $Q'Q'$ a factor of $\psi^\omega(a)$. Hence by induction on $|Q|$, we can deduce that $Q = \psi^n(q)$ for some integer $n \geq 0$. ◀

To end this section, we emphasize the interest of the previous examples by mentioning that when there are arbitrarily large intervals of lengths of quasiperiods, then there exists arbitrarily large quasiperiods that are not superprimitive.

▶ **Lemma 3.9.** *Let $\mathbf{w}$ be an aperiodic multiscale quasiperiodic word for which there exist arbitrary large intervals $[i, j]$ of lengths of quasiperiods. This word $\mathbf{w}$ admits an infinite sequence of nested quasiperiods.*

Proof. Let q_0 be any quasiperiod of $\mathbf{w}$. By hypothesis, there exists an interval $[i, j]$ with $j - i \geq |q_0|$ such that for all integers k in $[i, j]$, the prefix of length k of $\mathbf{w}$ is one of its quasiperiods. As $j - i \geq |q_0|$, there exists an integer k in $[i, j]$ such that the prefix of length k of $\mathbf{w}$ is q_0-quasiperiodic. By iterating that reasoning, we can construct a sequence of nested quasiperiods of $\mathbf{w}$. ◀

4   Quasiperiods of standard Sturmian words

The starting result of this section is the recent characterization of the Fibonacci word of [21] mentioned in the introduction. We provide a short proof using the general method of Section 2. We also reformulate this result in such a way it could be generalized to all standard Sturmian words. This is done in Section 4.2 before showing in Section 4.3 this is a characteristic property of this family of words.

4.1   Fibonacci example

We denote by $(F_n)_{n\geq 0}$ the sequence of Fibonacci integers ($F_0 = 1$, $F_1 = 1$, $F_{n+2} = F_{n+1} + F_n$ for $n \geq 0$) and by $(f_n)_{n\geq 1}$ the sequence of finite Fibonacci words ($f_1 = a$, $f_2 = ab$, $f_{n+2} = f_{n+1}f_n$ for $n \geq 1$). It is well-known that the infinite Fibonacci word $\mathbf{F} = \lim_{n\to\infty} f_n$ is also the fixed point of the morphism φ defined by $\varphi(a) = ab$ and $\varphi(b) = a$.

▶ **Lemma 4.1** (see [21]). *For all $n \geq 0$, the prefix of length n of $\mathbf{F}$ is a quasiperiod of $\mathbf{F}$ if and only if $n \notin \{F_p - 1 \mid p \geq 0\}$.*

Proof. It is well-known that left special factors of $\mathbf{F}$ coincide with its prefixes (see for instance [4, Prop. 4.10.3]). Thus determining prefixes of $\mathbf{F}$ that are right special is equivalent to determining the bispecial factors of $\mathbf{F}$. Let us denote by $(g_n)_{n\geq 2}$ the sequence of prefixes of $\mathbf{F}$ of length $(F_{n+1} - 2)_{n\geq 2}$. These words are exactly the bispecial factors of $\mathbf{F}$ ($\mathbf{F}$ is a standard Sturmian word; by [7] palindromic prefixes of standard Sturmian word are its bispecial factors; lengths of palindromic prefixes of $\mathbf{F}$ are computed in [8]).

For any $n \geq 3$, $f_n f_n$ is a prefix of $\mathbf{F}$ (indeed $abaaba = (\varphi^2(a))^2$ is a prefix of $\mathbf{F}$ which is the fixed point of φ). Moreover as $\mathbf{F}$ is $\varphi^2(a)$-quasiperiodic, $\mathbf{F}$ is f_n-quasiperiodic.

By Theorem 2.2, for all $n \geq 3$, for each prefix π of $\mathbf{F}$ with $F_n \leq |\pi| \leq |g_n| = F_{n+1} - 2$, π is a quasiperiod of $\mathbf{F}$. Moreover the prefix of length $|g_n| + 1 = F_{n+1} - 1$ is not a quasiperiod of $\mathbf{w}$. Finally prefixes of $\mathbf{F}$ of length $F_0 - 1 = F_1 - 1 = 0$, $F_2 - 1 = 1$ or $F_3 - 1 = 2$ are not quasiperiods of $\mathbf{F}$.   ◀

As mentioned in the previous proof, bispecial factors of the Fibonacci word are its prefixes of length $F_{n+1} - 2$ for $n \geq 2$.

▶ **Corollary 4.2.** *For all $n \geq 0$, the prefix of length $n + 1$ of $\mathbf{F}$ is a quasiperiod of $\mathbf{F}$ if and only if the prefix of length n of $\mathbf{F}$ is not bispecial.*

4.2   Quasiperiods of standard Sturmian words

The study of quasiperiods in Sturmian words dates back to an original question of Marcus [18]: "Is every Sturmian word quasiperiodic?" This question was completely answered in [15]: a Sturmian word is quasiperiodic if and only if it is not a Lyndon word. In other words, in any Sturmian shift, all but two words are quasiperiodic. Episturmian words, a family of words that include Sturmian words, were also considered and a characterization of all quasiperiods of any episturmian word was provided (see [10, Th. 4.19]). This characterization is quite elaborate and uses the so-called directive word of the studied episturmian word. With a bit of work, Lemma 4.1 could be deduced from this characterization. This is also the case of the next result, which generalizes Corollary 4.2.

Let us recall that an infinite word is *Sturmian* if and only if it has exactly $n + 1$ factors of length n for all n. By the well-known Morse-Hedlund theorem, Sturmian words are the aperiodic words with the least possible number of factors. These words have exactly one left

special factor and one right special factor of each length. A Sturmian word is called *standard Sturmian* if its left special factors coincide with its prefixes. By [15], they are multiscale quasiperiodic.

▶ **Proposition 4.3.** *Let* $\mathbf{w}$ *be a standard Sturmian word and* n *a positive integer. Then the prefix of length* n *of* $\mathbf{w}$ *is a quasiperiod if and only if its prefix of length* $n - 1$ *is* not *bispecial.*

This proposition could be proved using [10, Th. 4.19]. We rather provide another argument, which may be reused in other contexts. We work with *graphs of words* and *return words*, for which we recall the definitions.

Let n be an integer and $\mathbf{w}$ be an infinite word. The *n-th order graph of words of* $\mathbf{w}$, denoted by $\mathcal{G}_{\mathbf{w}}(n)$, is the directed graph whose vertices are the factors of length n of $\mathbf{w}$, such that there is an edge between two vertices x and y if and only if $\mathbf{w}$ has a factor of length $n + 1$ which has x as a prefix and y as a suffix. Observe that a factor v of $\mathbf{w}$ is right special if and only if its vertex in $\mathcal{G}_{\mathbf{w}}(|v|)$ has at least two outgoing edges. Therefore the graph of words allows to visualize right special factors, so it can help searching for quasiperiods using Proposition 2.3.

Let u be a factor of $\mathbf{w}$. A word v is a *return word* for u in $\mathbf{w}$ if and only if uv is a factor of $\mathbf{w}$ which has exactly two occurrences of u, one as a prefix and one as a suffix. A factor of $\mathbf{w}$ is *recurrent* if and only if it occurs infinitely many times in $\mathbf{w}$. Each return word v of u in $\mathbf{w}$ corresponds to a path of length $|v|$ starting from u in $\mathcal{G}_{\mathbf{w}}(|u|)$ (but not all such paths induce return words). The introduction of return words to study quasiperiodicity stems from the following lemma.

▶ **Lemma 4.4.** *[10, Lem. 4.3] A finite word* v *is a quasiperiod of an infinite word* $\mathbf{w}$ *if and only if* v *is a recurrent prefix of* $\mathbf{w}$ *such that any return to* v *in* $\mathbf{w}$ *has length at most* $|v|$.

The graphs of words of Sturmian words are well-known since works from Arnoux and Rauzy [2]. We exploit this information to characterize quasiperiods of Sturmian words.

Let $\mathbf{w}$ be a Sturmian word. It has exactly one left special factor and one right special factor of each length. Let $\ell_n(\mathbf{w})$ and $r_n(\mathbf{w})$ denote respectively the left and right special factors of length n of $\mathbf{w}$. Since $\mathbf{w}$ is on a binary alphabet, $\ell_n(\mathbf{w})$ has exactly two incoming edges and all other vertices have only one incoming edge. Likewise, $r_n(\mathbf{w})$ has exactly two outgoing edges and all other vertices have only one outgoing edge. There are only two possible shapes for such a graph. If $r_n(\mathbf{w}) = \ell_n(\mathbf{w})$ then $G_n(\mathbf{w})$ is the union of two edge-disjoint paths who only share one vertex, $r_n(\mathbf{w})$. Otherwise, $G_n(\mathbf{w})$ is the union of three edge-disjoint paths, one from $\ell_n(\mathbf{w})$ to $r_n(\mathbf{w})$ and two from $r_n(\mathbf{w})$ to $\ell_n(\mathbf{w})$. These paths do not share vertices other than $\ell_n(\mathbf{w})$ and $r_n(\mathbf{w})$.

The path going from $\ell_n(\mathbf{w})$ to $r_n(\mathbf{w})$, and which might be empty if these two vertices are equal, is called the *special path*. The other two paths are called the *short path* and the *long path*, according to their respective lengths. If both are of the same length, we arbitrarily choose which one is the short path (this does not matter). Although the special path might be of length 0, the short and the long path have always at least 1 edge.

The length of a return word to $\ell_n(\mathbf{w})$ is the sum of the length of the special path and of one of the short or long paths. As a Sturmian word has $n + 2$ factors of length $n + 1$, the graph $G_n(\mathbf{w})$ has $n + 2$ edges (recall that each edge corresponds to a factor of length $n + 1$). Lemma 4.4 implies that $\ell_n(\mathbf{w})$ is a quasiperiod of $\mathbf{w}$ if and only if the short path of $G_n(\mathbf{w})$ is not of length 1 and the special path is not empty.

This situation is well-known; see for instance the description of evolution of graphs of words in [2]. It occurs exactly for integers n such that $r_{n-1}(\mathbf{w}) = \ell_{n-1}(\mathbf{w})$. This ends the proof of Proposition 4.3.

4.3   Standard Sturmian words : a new characterization

The converse of Proposition 4.3 holds and allows to provide the following new characterization of standard Sturmian words.

▶ **Theorem 4.5.** *Let* **w** *be an aperiodic word. The word* **w** *is standard Sturmian if and only if it is multiscale quasiperiodic and satisfies the following condition: for each positive integer n, the prefix of length n of* **w** *is a quasiperiod if and only if the prefix of length $n-1$ is not right special.*

Proof. The "only if" part corresponds to Proposition 4.3. Let us prove the "if" part. Let **w** be an aperiodic word such that, for each $n > 0$, the prefix of length n of **w** is a quasiperiod if and only if the prefix of length $n-1$ is not right special. Let a be the first letter of **w** and let $B = \text{alph}(\mathbf{w}) \setminus \{a\}$ with $\text{alph}(\mathbf{w})$ the set of letters occurring in **w**. The size of $\text{alph}(\mathbf{w})$ may be arbitrary, but is at least two (so B is not empty) since **w** is not periodic.

Step 1: The word **w** has no factor in B^* of length at least 2.

First, observe that factors of **w** belonging to B^* have bounded length. Indeed, **w** is quasiperiodic and any quasiperiod contains occurrences of the letter a and of letters from B. Hence lengths of factors belonging to B^* are bounded by the length of the smallest quasiperiod. Let x be a factor of **w** of maximal length among all factors belonging to B^*. As B is not empty, $|x| \geq 1$. Let p be the smallest prefix of **w** ending with x. By maximality in the definition of x, p is not right special: it is always followed by the letter a. Thus by hypothesis on **w**, pa is a quasiperiod of **w**.

By definition, p begins with a and ends with x. By maximality in the definition of x, p ends with ax and, by construction, does not contain any other occurrence of x. It follows that borders of pa are the words ε and a. Thus $\mathbf{w} \in p\{p, ap\}^\omega$.

Let π be the prefix of **w** of length $|p| - 1$ and let b be the last letter of p. We have $p = \pi b$ and $\mathbf{w} \in \pi b\{\pi b, a\pi b\}^\omega$. As x has only one occurrence in p as a suffix, it has no occurrence in π. Moreover a is the first letter of π. Assume that π is not right special. By hypothesis, it follows that p is a quasiperiod of **w**. By the choice on x and definition of p, p cannot be an internal factor of pap. Thus $\mathbf{w} = p^\omega$: a contradiction with aperiodicity of **w**. Thus π is right special. There exists a letter c different from b such that πc is a factor of **w**. This word πc occurs in a factor $a\pi$. Hence $a\pi = \pi c$ which implies $a = c$ and π is a power of a. Thus $|x| = 1$.

First corollary of Step 1: There exists an aperiodic word $\mathbf{w}'$ such that $\mathbf{w} = L_a(\mathbf{w}')$, where L_a is the morphism defined by $L_a(a) = a$, $L_a(x) = ax$ for any letter $x \neq a$.

The existence of $\mathbf{w}'$ is a reformulation of the result of Step 1. Aperiodicity of $\mathbf{w}'$ is a consequence of aperiodicity of **w**.

Second corollary of Step 1: **w** is a binary word.

Indeed assume that **w** contains at least three different letters a (its first letter), b and c, with the first occurrence of b occurring before the first occurrence of c. Any quasiperiod of **w** must contain b and c. By Step 1, b is always followed by the letter a. Let π be smallest prefix of **w** ending with b. The word πa is a prefix of **w**. As it does not contains c, πa cannot be a quasiperiod of **w**. By the properties of **w**, π is right special. This contradicts the fact that b is not right special.

Step 2: The smallest quasiperiod of **w** is its prefix $a^k ba$ ($k \geq 1$).

Indeed by Step 1 (and its second corollary), each occurrence of b is followed by the letter a. In particular the prefix $a^k b$ is not right special. By the properties of **w**, $a^k ba$ is a quasiperiod of **w**. As **w** is aperiodic, $a^k b$ is not a quasiperiod of **w**. Clearly **w** has no quasiperiod that are powers of the letter a. Hence $a^k ba$ is the smallest quasiperiod of **w**.

Step 3: For each integer n, the prefix of $\mathbf{w}'$ of length $n + 1$ is a quasiperiod of $\mathbf{w}'$ if and only if the prefix of $\mathbf{w}'$ of length n is not right special.

Let p be a prefix of $\mathbf{w}'$ and c be the letter such that pc is a prefix of $\mathbf{w}'$. We have to prove that pc is a quasiperiod of $\mathbf{w}'$ if and only if p is not right special in $\mathbf{w}'$. This is a direct consequence of the next four properties (the third one is an hypothesis on $\mathbf{w}$, the proof of the others are omitted by lack of place):

1. Let $\mathbf{x} \in \{a, b\}^\omega$. A word q is a quasiperiod of $\mathbf{x}$ if and only if both words $L_a(q)$ and $L_a(q)a$ are quasiperiods of $L_a(\mathbf{x})$.
2. $L_a(p)ac$ is a quasiperiod of $\mathbf{w}$ if and only if $L_a(pc)$ and $L_a(pc)a$ are quasiperiods of $\mathbf{w}$.
3. $L_a(p)ac$ is a quasiperiod of $\mathbf{w}$ if and only if $L_a(p)a$ is not right special in $\mathbf{w}$.
4. Let $\mathbf{x} \in \{a, b\}^\omega$. A word u is right special in $\mathbf{x}$ if and only if $L_a(u)a$ is right special in $L_a(\mathbf{x})$.

Step 4: $\mathbf{w}'$ is multi-scale quasiperiodic.

By Step 2, $a^k ba$ is a quasiperiod of $\mathbf{w}$. Hence $\mathbf{w} \in a^k b \{a^k b, a^{k+1} b\}^\omega$. As $\mathbf{w}$ is aperiodic, there exists an integer $i \geq 1$ such that $(a^k b)^i a^{k+1} b$ is a prefix of $\mathbf{w}$. Let j be the greatest integer such that of $((a^k b)^i a^{k+1} b)^j$ is a factor of $\mathbf{w}$ (aperiodicity of $\mathbf{w}$ implies the existence of j). Let p be any prefix of $\mathbf{w}$ beginning with $(a^k b)^i a^{k+1} b$ and ending with $((a^k b)^i a^{k+1} b)^{j-1} (a^k b)^i a^{k+1}$ (multiscale quasiperiodicity of $\mathbf{w}$ implies there exist infinitely many such p). As a^{k+2} is not a factor of $\mathbf{w}$, each occurrence of p in $\mathbf{w}$ is always followed by the letter b. Hence p is not right special and pb is a quasiperiod of $\mathbf{w}$. By maximality of j, $pbpb$ is not a factor of $\mathbf{w}$. Thus two consecutive occurrences of pb must overlap by a factor at least as long as $(a^k b)^i a^{k+1} b$. Let p' be the unique (by the properties of L_a) word such that $L_a(p') = pb$. The word p' is a quasiperiod of $\mathbf{w}'$. As there are infinitely many possible words p, and so p', $\mathbf{w}'$ is multiscale quasiperiodic.

Conclusion. We have proven that:

- $\mathbf{w}$ is a binary word (let $\{a, b\}$ be the alphabet of $\mathbf{w}$);
- for an aperiodic multi-scale quasiperiodic word $\mathbf{w}'$ and a letter x, $\mathbf{w} = L_x(\mathbf{w}')$;
- the word $\mathbf{w}'$ satisfies the condition which links its quasiperiods and its right special factors, like $\mathbf{w}$.

Hence we can iterate this argument on $\mathbf{w}'$ and so on. Thus $\mathbf{w}$ is $\{L_a, L_b\}$-adic, that is, there exists an infinite sequence $(s_i)_{i \geq 0}$ and an infinite sequence of letters $(\alpha_i)_{i \geq 1}$ such that $s_0 = \mathbf{w}$ and, for $i \geq 1$, $s_{i-1} = L_{\alpha_i}(s_i)$. By [13, Cor. 2.7], $\mathbf{w}$ is standard episturmian. As it is a binary word, $\mathbf{w}$ is Sturmian.   ◀

5 Conclusion

Proposition 2.1 and Theorem 4.5 show that multiscale quasiperiodicity is an interesting combinatorial notion as it allows to characterize some families of right infinite words. These characterizations can be extended to biinfinite words. For instance a biinfinite word is Sturmian if and only if it is multiscale quasiperiodic and satisfies: for each positive integer n, $\mathbf{w}$ has a quasiperiod of length n if and only if $\mathbf{w}$ has no bispecial factor of length $n - 1$. Nevertheless, the structure of sets of quasiperiods of biinfinite words still needs to be studied, because it is more complex as there may exist several quasiperiods having the same length.

Another important problem is left open at the end of Section 3.2. What is the exact link between desubstitution and multiscale quasiperiodicity? Can any multiscale quasiperiodic word $\mathbf{w}$ be desubstituted into another multiscale quasiperiodic word? If the answer is negative, what additional conditions does this property imply?

―――― **References** ――――

1 A. Apostolico and A. Ehrenfeucht. Efficient detection of quasiperiodicities in strings. *Theor. Comput. Sci.*, 119:247–265, 1993.

2 P. Arnoux and G. Rauzy. Représentation géométrique de suites de complexité $2n+1$. *Bull. Soc. Math. France*, 119:199–215, 1991.

3 J. Berstel. Fibonacci words - a survey. In *The book of L.* Springer-Verlag, 1986.

4 V. Berthé and M. Rigo, editors. *Combinatorics, Automata and Number Theory.* Number 135 in Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2010.

5 G. S. Brodal and C. N. S. Pedersen. Finding maximal quasiperiodicities in strings. In *Combinatorial Pattern Matching (CPM'2000), 11th Annual Symposium, CPM 2000, Montreal, Canada, June 21-23, 2000*, volume 1848 of *Lect. Notes in Comput. Science*, pages 397–411, 2000.

6 M. Christou, M. Crochemore, and I. Costas S. Quasiperiodicities in Fibonacci strings. Technical Report 1201.6162, ArXiv, 2002 (To appear in Ars Combinatoria).

7 A. de Luca. Sturmian words: structure, combinatorics, and their arithmetics. *Theor. Comput. Sci.*, 183:45–82, 1997.

8 S. Fischler. Palindromic prefixes and episturmian words. *J. Combin. Theory Ser. A*, 113(7):1281–1304, 2006.

9 G. Gamard and G. Richomme. Coverability in two dimensions. In A. Horia Dediu, E. Formenti, C. Martín-Vide, and B. Truthe, editors, *Language and Automata Theory and Applications - 9th International Conference, LATA 2015, Nice, France, March 2-6, 2015, Proceedings*, volume 8977 of *Lect. Notes in Comput. Science*, pages 402–413. Springer, 2015.

10 A. Glen, F. Levé, and G. Richomme. Quasiperiodic and Lyndon episturmian words. *Theor. Comput. Sci.*, 409(3):578–600, 2008.

11 R. Groult and G. Richomme. Optimality of some algorithms to detect quasiperiodicities. *Theoretical Computer Science*, 411:3110 – 3122, 2010.

12 C. S. Iliopoulos and L. Mouchard. An $o(n \log n)$ algorithm for computing all maximal quasiperiodicities in strings. In C. S. Calude and M. J. Dinneen, editors, *Combinatorics, Computation and Logic. Proceedings of DMTCS'99 and CATS'99*, Lect. Notes in Comput. Science, pages 262–272, Auckland, New-Zealand, 1999. Springer.

13 J. Justin and G. Pirillo. Episturmian words and episturmian morphisms. *Theor. Comput. Sci.*, 276(1-2):281–313, 2002.

14 F. Levé and G. Richomme. Quasiperiodic infinite words: some answers. *Bull. Europ. Assoc. Theoret. Comput. Sci.*, 84:128–238, 2004.

15 F. Levé and G. Richomme. Quasiperiodic Sturmian words and morphisms. *Theor. Comput. Sci.*, 372(1):15–25, 2007.

16 M. Lothaire. *Combinatorics on Words*, volume 17 of *Encyclopedia of Mathematics and its Applications.* Addison-Wesley, 1983. Reprinted in the *Cambridge Mathematical Library*, Cambridge University Press, UK, 1997.

17 M. Lothaire. *Algebraic Combinatorics on Words*, volume 90 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, 2002.

18 S. Marcus. Quasiperiodic infinite words. *Bull. Eur. Assoc. Theor. Comput. Sci.*, 82:170–174, 2004.

19 T. Monteil and S. Marcus. Quasiperiodic infinite words: multi-scale case and dynamical properties. arXiv:math/0603354v1, 2006.

20 L. Mouchard. Normal forms of quasiperiodic strings. *Theor. Comput. Sci.*, 249:313–324, 2000.

21 H. Mousavi, L. Schaeffer, and J. Shallit. Decision algorithms for Fibonacci-automatic words, I: basic results. *RAIRO Theor. Inform. Appl.*, To appear.