Closed Ziv-Lempel factorization of the m-bonacci words

Marieh Jahannia
School of Mathematics, Statistics and Computer Science, College of Science
University of Tehran, Tehran, Iran
mjahannia@ut.ac.ir

Morteza Mohammad-noori
School of Mathematics, Statistics and Computer Science, College of Science
University of Tehran, Tehran, Iran

Narad Rampersad
Department of Mathematics and Statistics
University of Winnipeg, Winnipeg, Canada
n.rampersad@uwinnipeg.ca

Manon Stipulanti
Department of Mathematics
University of Liège, Liège, Belgium
m.stipulanti@uliege.be

November 30, 2021

Abstract

A word w is said to be closed if it has a proper factor x which occurs exactly twice in w, as a prefix and as a suffix of w. Based on the concept of Ziv-Lempel factorization, we define the closed z-factorization of finite and infinite words. Then we find the closed z-factorization of the infinite m-bonacci words for all $m \geq 2$. We also classify closed prefixes of the infinite m-bonacci words.

Keywords: Ziv–Lempel factorization; closed; Fibonacci word; m-bonacci words; episturmian words.

---

1 Introduction

Factorization of words is an important topic in combinatorics on words, which roughly consists in breaking a given word into concatenation of other words, called factors. Some specific factorizations require that those factors satisfy some special properties. Some various types of factorizations studied in the literature are the Ziv-Lempel factorization, the Crochemore factorization, the Lyndon factorization and the grammar-based factorization [6, 22, 23, 24]. The Ziv-Lempel factorization, or z-factorization for short, was introduced by Ziv and Lempel for finite words [20] and then was extended to infinite words [4]. This factorization has several applications in data compression [27] and text processing [19]. Ghareghani et al. [13] determined z-factorizations for standard episturmian words. We introduced the palindromic z-factorizations by requiring each factor to be a palindrome and computed this factorization for the m-bonacci words [14]. In this work, based on the notion of closed words, which appeared in [5], we introduce the closed z-factorization and apply it to the infinite Fibonacci word and then to all m-bonacci words, for $m > 2$. We also characterize closed prefixes of the m-bonacci word h_ω and obtain the binary word $x = oc(h_\omega)$ associated with closed prefixes of h_ω (defined by $x_n = 1$ if the prefix of length n of h_ω is closed; otherwise, $x_n = 0$). The connection of this word with the sequence of m-bonacci numbers then appears as a consequence.

The paper is organized as follows. In Section 2, we present some notation and definitions needed in the rest of the paper. In Section 3, we study the closed z-factorization of the Fibonacci word. In Section 4 we study the closed z-factorization of the m-bonacci words and consider numerous properties of this factorization. The main result of this section is reported as Theorem 42 which shows the closed z-factorization of the m-bonacci words. Section 5 is devoted to the link between the closed and palindromic z-factorizations of the m-bonacci words. In Section 6, we characterize closed prefixes of the m-bonacci words and we give the oc-sequence of the m-bonacci words. Finally in Section 7 we mention some open problems.

2 Preliminaries

Let A be a finite alphabet. The elements of A^* are called (finite) words over A. We denote the empty word by ε and we let $A^+ = A^* \setminus \{\varepsilon\}$. For every finite word w, we let $|w|$ denote its *length*. A word z is a *factor* of $w \in A^*$, and we write $z \prec w$, if $w = uzv$ for some $u, v \in A^*$. The factor z is said to be *proper* if $u, v \neq \varepsilon$. We say that z is a *prefix* (resp., *suffix*) of w, and we denote this by $z \lhd w$ (resp., $z \rhd w$), if $u = \varepsilon$ (resp., $v = \varepsilon$). The set of factors of a word w is denoted by $\boldsymbol{F}(w)$. For a factor z of a word w, we let $|w|_z$ denote the number of occurrences of z in w. We say that z is a *right special factor* of w if za and zb are factors of w for some distinct letters $a, b \in A$.

Let $w = w_1 w_2 \cdots w_n$ with $w_i \in A$ for all $i \in \{1, \ldots, n\}$. We let w^R denote the *reverse* of w, that is, $w^R = w_n \cdots w_2 w_1$. If $w = w^R$, w is called a *palindrome* or a *palindromic word*. We let $(w)^+$ denote the *palindromic closure* of w, that is, the shortest palindrome having w as a prefix. For example, $(race)^+ = racecar$. If v is the longest palindromic suffix of w, say

$w = uv$, then $(w)^+ = uvu^R$. For each word $u \in A^*$, we use the notation u^{-1} as below. If $w = uv$, then we let $u^{-1}w = v$ and $wv^{-1} = u$. This simply yields $(uv)^{-1} = v^{-1}u^{-1}$, consequently, $w^{-1} = w_n^{-1} \cdots w_2^{-1} w_1^{-1}$.

Let $\boldsymbol{t}$ be an infinite word and let w be a factor of $\boldsymbol{t}$. The word v is said to be a *return word* of w if v begins with an occurrence of w and ends exactly just before the next occurrence of w in $\boldsymbol{t}$. If v is a return word of w, then vw is said to be a *complete return word* of w. For example, $abacaba$ and $ccabacc$ are complete return words of aba and cc. The notion of return words was introduced to study primitive substitutive sequences [9]. In [16], Justin and Vuillon presented a new characterization of Sturmian words using return words. They also characterized the return words of factors of standard episturmian words.

A non-empty word x is called a *border* of w if x is both a prefix and a suffix of w. A word w is said to be *closed* if it is a single letter or has a border x such that it does not have any other occurrence in w, in other words, $|w|_x = 2$. In this case, we call w the *frontier* of x. As an example, the word $w = mamma$ is closed, because ma appears only as a prefix and a suffix of w. The notion of closed words appeared in the study of trapezoidal words [5]. If w is not closed then it is said to be *open*. In [11], Fici illustrates several aspects of open and closed words and factors.

In [12], Fici et al. studied words with the smallest number of closed factors. In [3], Badkobeh et al. showed that a length-n word contains at least $n+1$ distinct closed factors and characterized those words having exactly $n+1$ closed factors. Badkobeh et al. described an efficient solution to the shortest and longest closed factorizations [2]. The shortest (resp., longest) closed factorization of a string is obtained by factorizing it into shortest (resp., longest) closed factors. In [7], A. De Luca et al. studied closed prefixes of Sturmian words and introduced the oc-sequence of a word w, as $oc(w)$, which is the binary sequence whose n-th term is 1 if the length-n prefix of w is closed, or 0 if it is open. They showed that this sequence is deeply related to the combinatorial and periodic structure of a word. In [5], Bucci et al. studied closed prefixes of Fibonacci words and investigated the oc-sequence of the Fibonacci word, $oc(F)$. Note that, for a given infinite word $\boldsymbol{t}$ and a nonempty factor w of $\boldsymbol{t}$, every complete return word of w is closed because it contains the factor w exactly twice, once as a prefix and once as a suffix.

Let A be a finite alphabet. A mapping $\psi : A^* \to B^*$ is called a *morphism* if $\psi(uv) = \psi(u)\psi(v)$ for all $u, v \in A^*$. A morphism ψ is said to be *prolongable* if there exists a letter $a \in A$ and a word $x \in A^*$ such that $\psi(a) = ax$ and $\psi^i(x) \neq \varepsilon$ for all $i \geq 0$. In this case, the word $\psi^n(a)$ is a proper prefix of $\psi^{n+1}(a)$ for all $n \geq 0$. Therefore, the infinite word $\psi^\omega(a) = \lim_{n\to\infty} \psi^n(a)$ is a fixed point of ψ. For every morphism $\psi : A^* \to B^*$ and each word $u \in A^*$, we define $\psi(u^{-1}) = (\psi(u))^{-1}$. This is justified by applying ψ on $uu^{-1} = u^{-1}u = \varepsilon$.

A *factorization* of a word consists in decomposing it into consecutive factors, which satisfy some special properties. Given an infinite word $\boldsymbol{w}$, the *Ziv-Lempel factorization* or the *z-factorization* of $\boldsymbol{w}$ is $z(\boldsymbol{w}) = (z_1, z_2, \ldots)$ where z_i is the shortest prefix of $z_i z_{i+1} \cdots$ that occurs exactly once in $z_1 z_2 \cdots z_i$. We introduce the closed z-factorization $cz(\boldsymbol{w}) = (z_1, z_2, \ldots)$ of $\boldsymbol{w}$ by requiring that each factor z_i is closed. Notice that self-referencing is allowed, i.e., the previous occurrence of z_i may overlap with itself.

3 Closed z-factorization of the Fibonacci word

Recall that the sequence of *finite Fibonacci words* is given by $f_{-1} = 1$, $f_0 = 0$ and $f_n = f_{n-1} f_{n-2}$ for all $n \geq 1$. Furthermore, for all $n \geq 0$, f_n is the n-th iteration of the morphism $\sigma: A^* \to A^*$ on the letter 0 defined by $\sigma(0) = 01$, $\sigma(1) = 0$, that is, $f_n = \sigma^n(0)$. The first few Fibonacci words are given in Table 1. The infinite Fibonacci word f_ω is given by $f_\omega = \lim_{n\to\infty} f_n$. Equivalently we have, $f_\omega = \sigma^\omega(0)$.

n	-1	0	1	2	3	4	5
f_n	1	0	01	010	01001	01001010	0100101001001

Table 1: The first few Fibonacci words $(f_n)_{n\geq -1}$.

The sequence of Fibonacci numbers is given by the recurrence relation $F_n = F_{n-1} + F_{n-2}$ for all $n \geq 1$ where $F_{-1} = 1$, $F_0 = 1$. The sequence of Fibonacci words is related to the latter sequence of numbers since $F_n = |f_n|$ for all $n \geq -1$.

We note that the infinite Fibonacci word belongs to the class of *Sturmian words*, that is, infinite aperiodic binary words with minimal factor complexity. These words were presented in [21] and are widely studied in the literature because they have several equivalent definitions and many various optimal properties, see for instance [18, Chapter 2].

In [26], Wen and Wen defined the n-th *singular word* w_n of f_ω by $w_{-2} = \varepsilon$, $w_{-1} = 0$, $w_0 = 1$ and for $n \geq 1$, $w_n = a f_n b^{-1}$, where $ab \in \{01, 10\}$ is the length-2 suffix of f_n. It is easy to see that $|w_n| = F_n$ for all $n \geq -1$. The first few singular words of the Fibonacci word f_ω are displayed in Table 2.

n	-2	-1	0	1	2	3	4	5
w_n	ε	0	1	00	101	00100	10100101	0010010100100

Table 2: The first few singular words $(w_n)_{n\geq -2}$ of the Fibonacci word f_ω.

The following lemma summarizes some properties of the singular words of the Fibonacci word f_ω that are useful in the following.

Lemma 1. [*26*, Property 2] *The singular words $(w_n)_n$ of the Fibonacci word have the following properties.*

1. *For all $n \geq -1$, $w_n \not\prec w_{n+1}$.*
2. *For all $n \geq 1$, $w_n = w_{n-2} w_{n-3} w_{n-2}$.*
3. *For all $n \geq -2$, w_n is a palindrome.*
4. *For all $n \geq -1$, $w_n \not\prec \prod_{j=-1}^{n-1} w_j$.*

It is known that the infinite Fibonacci word f_ω can be written as the concatenation of the singular words $(w_n)_{n\geq -1}$ [26], which turns out to be the z-factorization of f_ω [10].

Lemma 2. [*10*, Proposition 8] *The infinite Fibonacci word f_ω is the concatenation of the singular words, that is,*

$$f_\omega = \prod_{n\geq -1} w_n.$$

Our main goal in this section is to prove Theorem 5, which gives the closed z-factorization of the infinite Fibonacci word f_ω. We will make use of the following two lemmas.

Lemma 3. [*26*, Lemma 3] *Let $n \geq 1$ and write $w_n w_{n+1} = u_1 u_2 u_3$ (or $w_{n+1} w_n = u_1 u_2 u_3$) with $0 < |u_1| < F_n$ and $0 < |u_3| < F_{n+1}$. Then u_2 is not a singular word.*

Lemma 4. *For all $n \geq -1$, w_n is closed.*

Proof. From Table 2, the assertion can be easily verified for the values $-1 \leq n \leq 3$. Assume that $n \geq 4$. It follows from Lemma 1 that $w_n = w_{n-2} w_{n-3} w_{n-2}$. So w_{n-2} is a border of w_n. It suffices to show that w_{n-2} is neither a proper factor of $w_{n-2} w_{n-3}$ nor that of $w_{n-3} w_{n-2}$. We proceed by contradiction and suppose that w_{n-2} is a proper factor of $w_{n-2} w_{n-3}$ (the other case is similar). There exist non-empty words u_1 and u_3 over $\{0,1\}$ such that $w_{n-2} w_{n-3} = u_1 w_{n-2} u_3$. Using the fact that $|w_{n-2}| = F_{n-2}$ and $|w_{n-3}| = F_{n-3}$, we have $|u_1 u_3| = F_{n-3}$. So $0 < |u_1| < F_{n-3} < F_{n-2}$ and $0 < |u_3| < F_{n-3}$. This contradicts Lemma 3. □

Theorem 5. *The closed z-factorization of the Fibonacci infinite word is*

$$cz(f_\omega) = (w_{-1}, w_0, w_1, \ldots).$$

Proof. Based on Lemma 2, $z(f_\omega) = (w_{-1}, w_0, w_1, \ldots)$. By Lemma 4, the factors w_n are closed, which shows that this factorization is also $cz(f_\omega)$. □

4 Closed z-factorization of the m-bonacci word

The *Tribonacci word* is the most natural extension of the Fibonacci word to a three-letter alphabet and has been studied in many papers, see for instance [1, 25]. To describe such an extension to a finite alphabet of arbitrary size greater than 1, for every integer $m \geq 2$, we define the *m-bonacci word* as the fixed point of the morphism φ_m given in the following definition.

Definition 6. For $m \geq 2$, let $A_m = \{0, 1, \ldots, m-1\}$ and let φ_m be the morphism defined by

$$\varphi_m : A_m^* \to A_m^*, 0 \mapsto 01, \ldots, (m-2) \mapsto 0(m-1), (m-1) \mapsto 0.$$

The sequence of *finite m-bonacci words* denoted as $(h_n^{(m)})_{n\geq 0}$, or briefly as $(h_n)_{n\geq 0}$, is given by $h_n = \varphi_m^n(0)$ for all $n \geq 0$. The *infinite m-bonacci word h_ω* is the fixed point of the morphism φ_m which starts with 0.

Example 7. Suppose that $m = 2$. Then $A_2 = \{0, 1\}$ and $\varphi_2 : 0 \mapsto 01, 1 \mapsto 0$. Also the infinite 2-bonacci word $h_\omega^{(2)}$ is exactly the infinite Fibonacci word. Furthermore, for all $n \geq 0$, $f_n = h_n$.

The first few finite m-bonacci words are given in Table 3 for some values of the parameter m.

n	0	1	2	3	4	5
$h_n^{(2)}$	0	01	010	01001	01001010	0100101001001
$h_n^{(3)}$	0	01	0102	0102010	0102010010201	010201001020101020100102
$h_n^{(4)}$	0	01	0102	01020103	010201030102010	01020103010201001020103010201
$h_n^{(5)}$	0	01	0102	01020103	0102010301020104	0102010301020104010201030102010

Table 3: The first few words of the sequence $\left(h_n^{(m)}\right)_{n\geq 0}$ for $m \in \{2,3,4,5\}$.

Remark 8. In addition to the morphism φ_m, we define several objects related to m in this section, where the parameter m is deleted in the notation of most of them for the sake of clarity. These objects are h_n, u_n, μ_n, ψ_n, z_n, $\underline{n}$, $\underline{\underline{n}}$ and $\hat{n}$. This is justifiable firstly because fixing the parameter m at the beginning of the statement of each of the upcoming results removes the danger of confusion and secondly because deleting m from the notation proposes a considerable simplicity and convenience in presenting formulas.

Notation 9. For every integer $n \geq 0$, let $\underline{n} := n \bmod m$. Note that $\underline{n} \in \{0,1,\ldots,m-1\}$.

Remark 10. The words h_n, represented in Definition 6, can be defined in a recursive way as follows:

$$h_n = \begin{cases} 0, & \text{if } n = 0; \\ h_{n-1}\cdots h_0 n, & \text{if } 1 \leq n \leq m-1; \\ h_{n-1}\cdots h_{n-m}, & \text{if } n \geq m. \end{cases} \tag{1}$$

Notation 11. Let m and n be integers. We use the notation $m \mid n$ as m divides n and $m \nmid n$ as m does not divide n.

Lemma 12. *Let $n \geq 1$. The word h_n^R starts with $\underline{n}0$ if $m \nmid n$; 01 otherwise.*

Proof. We proceed by induction on $n \geq 1$. The result holds true for $n = 1$ as $h_1^R = 10$. Assume that $n \geq 2$. There are two cases to consider according to the value of n.

Case 1. Suppose that $n \leq m-1$. Using Equation (1), we obtain $h_n^R = nh_0^R \cdots h_{n-1}^R$. As $h_0^R = 0$, we get the result.

Case 2. Suppose that $n \geq m$. From Equation (1), we deduce that $h_n^R = h_{n-m}^R \cdots h_{n-1}^R$. Using the induction hypothesis, h_{n-m}^R starts with $\underline{n-m}0 = \underline{n}0$ if $m \nmid n$; 01 otherwise. This ends the proof. □

Similarly to the Fibonacci word, which is a typical example of Sturmian words, m-bonacci words are typical examples of episturmian words over A_m. Episturmian words extend Sturmian words to larger alphabets. Their construction and properties can be found in [8, 15, 17] for instance. Inspired by the construction of standard episturmian words presented in [17], we need a restricted version of that construction and prior to this, some definitions are required as well.

Definition 13. For each letter $a \in A_m$, we define the morphism $\psi_a^{(m)}: A_m^* \to A_m^*$, ψ_a for short, by $\psi_a(a) = a$, and $\psi_a(b) = ab$ for all $b \in A_m \setminus \{a\}$. Furthermore, we define the sequence of morphisms $\mu_n^{(m)}: A_m^* \to A_m^*$, μ_n for short, by $\mu_0 = \text{id}$ and $\mu_n = \psi_{\underline{0}} \circ \psi_{\underline{1}} \cdots \circ \psi_{\underline{n-1}}$ for $n > 0$, where $id : A_m^* \to A_m^*$ is the identity morphism.

In Table 4, we compute the images of letters in A_3 under the morphisms $\mu_0, \mu_1, \dots, \mu_4$. Observe that $h_0 = \mu_0(0)$, $h_1 = \mu_1(1)$, $h_2 = \mu_2(2)$, $h_3 = \mu_3(0)$ and $h_4 = \mu_4(1)$.

$a \in A_3$	0	1	2
$\mu_0^{(3)}(a) = id(a)$	0	1	2
$\mu_1^{(3)}(a) = \psi_0(a)$	0	01	02
$\mu_2^{(3)}(a) = \psi_0 \circ \psi_1(a)$	010	01	0102
$\mu_3^{(3)}(a) = \psi_0 \circ \psi_1 \circ \psi_2(a)$	0102010	010201	0102
$\mu_4^{(3)}(a) = \psi_0 \circ \psi_1 \circ \psi_2 \circ \psi_0(a)$	0102010	0102010010201	01020100102

Table 4: The images of letters in A_3 under morphisms $\mu_i^{(3)}$, $i \in \{0, \dots, 4\}$.

Definition 14. We let $\left(u_n^{(m)}\right)_{n\geq 1}$, or briefly $(u_n)_{n\geq 1}$, denote the sequence of palindromic prefixes of h_ω starting with $u_1 = \varepsilon$ and sorted by increasing length.

Notice that $u_1^{(m)} = \varepsilon$ and $u_2^{(m)} = 0$ for all $m \geq 2$. It is clear that $h_\omega = \lim_{n\to\infty} u_n$. In Table 5, we show the first few elements of the sequence $\left(u_n^{(m)}\right)_{n\geq 1}$ for $m \in \{2,3,4,5\}$.

n	1	2	3	4	5	6
$u_n^{(2)}$	ε	0	010	010010	01001010010	0100101001001010010
$u_n^{(3)}$	ε	0	010	0102010	01020100102010	010201001020101020100102010
$u_n^{(4)}$	ε	0	010	0102010	010201030102010	010201030102010010201030102010
$u_n^{(5)}$	ε	0	010	0102010	010201030102010	0102010301020104010201030102010

Table 5: The first few words of the sequence $\left(u_n^{(m)}\right)_{n\geq 1}$ for $m \in \{2,3,4,5\}$.

Lemma 15. [*17*, Section 2.1] *The following identities hold.*

1. *For all $n \geq 0$, $h_n = \mu_n(\underline{n})$.*
2. *For all $n \geq 1$, $u_{n+1} = (u_n\ \underline{n-1})^{(+)}$.*
3. *For all $n \geq 1$, we have $u_{n+1} = h_{n-1}u_n$.*
4. *For all $n \geq 2$, we have $u_n = h_0^R\, h_1^R \cdots h_{n-2}^R$.*
5. *For all $n \geq 1$,*

$$h_n = \begin{cases} u_{n+1}\, n, & \text{if } 1 \leq n \leq m-1; \\ u_n u_{n-m}^{-1}, & \text{if } n \geq m. \end{cases}$$

Definition 16. Let $(x_n)_{n\geq 0}$ be a sequence of words and let $(\nu_n)_{(n\geq 0)}$ be a sequence of morphisms, where $x_n \in A_m^*$ and $\nu_n : A_m^* \to A_m^*$ for $n \geq 0$. If a and b be integers where $a \leq b$, then we let $\mathop{\circ}_{i=a}^{b} \nu_i = \nu_a \circ \nu_{a+1} \circ \cdots \circ \nu_b$. Moreover, if $a > b$, then we let $\mathop{\circ}_{i=a}^{b} \nu_i = id$ and $\prod_{i=a}^{b} x_i = \varepsilon$.

In the two following lemmas, we start counting the positions of letters at 1.

Lemma 17. *Let $m \geq 2$ and let k and k' be distinct letters of A_m. Suppose that the word $x = x_1 \cdots x_n$ with $x_1, \ldots, x_{n-1} \in A_m \backslash \{k\}$ contains exactly p occurrences of k' that occur at positions $a_1, \ldots, a_p$. Moreover, let $y = \psi_k(x)$.*

1. *If $x_n = k$, then $|y| = 2n-1$; otherwise $|y| = 2n$.*
2. *The word y contains n occurrences of the letter k that appear in all odd positions and p occurrences of the letter k' that appear at positions $2a_1, \ldots, 2a_p$.*

Proof. Let $x = x_1 \cdots x_{n-1} x_n$, where $x_1, \ldots, x_{n-1} \in A_m \setminus \{k\}$ and $x_n \in A_m$. Since $y = \psi_k(x)$, we obtain

$$y = kx_1kx_2 \cdots kx_{n-1}ky_n, \tag{2}$$

where $y_n = \varepsilon$ if $x_n = k$; $y_n = x_n$ otherwise. Therefore, $|y|$ equals either $2n-1$ (when $x_n = k$) or $2n$ (when $x_n \neq k$). Furthermore, by Equation (2), for each $1 \leq j \leq \lfloor \frac{y}{2} \rfloor$ the equation $y_{2j} = x_j$ holds. This proves the second part of the lemma. □

Lemma 18. *Let $n \geq m$. We have*

$$\Big(\mathop{\circ}_{i=n-m+1}^{n-1} \psi_{\underline{i}}\Big)(\underline{n-m+1}) = \prod_{j=2}^{m} \Big(\mathop{\circ}_{i=n-m+1}^{n-j} \psi_{\underline{i}}\Big)(\underline{n-j+1}). \tag{3}$$

Proof. First, observe that the length of the composition in the left-hand side of Equation (3) is $m-1$. BâĂŇy applying Lemma 17 $(m-1)$ times, both sides of Equation (3) equal the word y of length $2^{m-1}-1$ that consists of letters $\underline{n-1}, \underline{n-2}, \ldots, \underline{n-m+1}$ and for each $j \in \{1, \ldots, m-1\}$, the letter $\underline{n-j}$ appears exactly in positions $(2t+1)2^{m-j-1}$ with $t \in \{0, 1, \ldots, 2^{j-1}-1\}$. The result now follows. □

Lemma 19. *The following identities hold.*

1. *For all $n \geq m$, $\mu_n(\underline{n-m+1}) = h_{n-1} h_{n-2} \cdots h_{n-m+1}$.*
2. *For all $1 \leq n \leq m-1$, $\mu_{n-1}(n) = u_n\, n$.*
3. *For all $n \geq 1$, $h_n^R = 0^{-1} \varphi_m(h_{n-1}^R) 0$.*
4. *For all $n \geq 2$, $u_n = \varphi_m(u_{n-1}) 0$.*

Proof.

1. The result is obtained by applying $\mathop{\circ}_{i=0}^{n-m} \psi_{\underline{i}}$ on both sides of Equation (3) and using the identities $h_j = \mu_j(\underline{j}) = \psi_{\underline{0}} \circ \psi_{\underline{1}} \circ \cdots \circ \psi_{\underline{j-1}}(\underline{j})$ which hold for each $j \in \{n-1, \ldots, n-m+1\}$ by Part (1) of Lemma 15.
2. By Part (1) of Lemma 15, considering $1 \leq n \leq m-1$, we obtain $h_{n-1} = \mu_{n-1}(n-1)$ and $h_n = \mu_n(n)$. Now replacing μ_n by $\mu_{n-1} \circ \psi_{n-1}$ in the last equation easily yields

$$h_n = h_{n-1}\, \mu_{n-1}(n). \tag{4}$$

On the other hand, by Parts (3) and (5) of Lemma 15, we get $h_n = h_{n-1} u_n\, n$. Comparing this with Equality (4) yields the result.

3. We proceed by induction on $n \geq 1$. The result holds true for the case $n = 1$ because $h_1^R = 10 = 0^{-1}\varphi_m(h_0^R)0$ as $h_0^R = 0$. We divide the proof into three cases according to the value of n.

   Case 1. Suppose that $n \leq m-1$. By (1), we have $h_n^R = n h_0^R \cdots h_{n-1}^R$. Using the induction hypothesis, we get $h_n^R = n00^{-1}\varphi_m(h_0^R)0 \cdots 0^{-1}\varphi_m(h_{n-1}^R)0 = 0^{-1}\varphi_m((n-1)h_0^R \cdots h_{n-2}^R)0 = 0^{-1}\varphi_m(h_{n-1}^R)0$.

   Case 2. Suppose that $n = m-1$. Using (1), we have $h_n^R = h_0^R \cdots h_{n-1}^R$. By the induction hypothesis, we find that $h_n^R = 00^{-1}\varphi_m(h_0^R)0 \cdots 0^{-1}\varphi_m(h_{n-1}^R)0 = 0^{-1}\varphi_m((n-1)h_0^R \cdots h_{n-2}^R)0 = 0^{-1}\varphi_m(h_{n-1}^R)0$.

   Case 3. Suppose that $n \geq m$. By (1), we deduce that $h_n^R = h_{n-m}^R \cdots h_{n-1}^R$. By the induction hypothesis, we get $h_n^R = 0^{-1}\varphi_m(h_{n-m-1}^R)0 \cdots 0^{-1}\varphi_m(h_{n-1}^R)0 = 0^{-1}\varphi_m(h_{n-m-1}^R \cdots h_{n-2}^R)0 = 0^{-1}\varphi_m(h_{n-1}^R)0$.

4. For $n = 2$, we have $u_2 = 0 = \varphi_m(u_1)0$ since $u_1 = \varepsilon$. Now suppose that $n \geq 3$. Using Part (4) of Lemma 15 and $h_0 = 0$, we obtain $u_n = 0 h_1^R \cdots h_{n-2}^R$. Replacing the words h_j^R in the right-hand side from Part (3) of Lemma 19 yields $u_n = \varphi_m(h_0^R h_1^R \cdots h_{n-3}^R)0$ whence the result is obtained by Part (4) of Lemma 15. □

We define a sequence of words $(z_n^{(m)})_{n\geq 0}$ in terms of $(h_n^{(m)})_{n\geq 0}$ that will be useful in the sequel to obtain the closed z-factorization of the m-bonacci words.

Definition 20. We define the sequence $(z_n^{(m)})_{n\geq 0}$, denoted briefly $(z_n)_{n\geq 0}$, by $z_0 = 0$, $z_1 = 1$, $z_2 = 020$ and

1. If $m = 2$, then for all $n \geq 3$, $z_n = (\underline{n-3})^{-1}\, h_{n-3}^R\, h_{n-2}^R\, \underline{n-2}$.
2. If $m \geq 3$, then

$$z_n = \begin{cases} (n-3)^{-1}\, h_{n-3}^R\, h_{n-2}^R\, n\, h_0^R\, h_1^R \cdots h_{n-3}^R\, (n-2), & \text{if } 3 \leq n \leq m-1; \\ (\underline{n-3})^{-1}\, h_{n-3}^R\, h_{n-2}^R\, h_{n-m}^R\, h_{n-m+1}^R\, h_{n-m+2}^R \cdots h_{n-3}^R\, \underline{n-2}, & \text{if } n \geq m. \end{cases}$$

The recursive equation satisfied by the sequence $(|z_n|)_n$ is given in Part (1) of Lemma 25. We show in Lemma 22 that for $m = 2$, the words z_n are exactly the singular words of the Fibonacci word. The case $m = 3$ is studied in the following example.

Example 21. The case $m = 3$ corresponds to the Tribonacci word. For all $n \geq 3$, we have

$$z_n^{(3)} = (\underline{n-3})^{-1}\, h_{n-3}^R\, h_{n-2}^R\, h_{n-3}^R\, \underline{n-2}.$$

Table 6 shows the first few words of the sequence $(z_n^{(m)})_{n\geq 0}$ for $2 \leq m \leq 5$.

Lemma 22. *For all $n \geq 0$, $w_{n-1} = z_n^{(2)}$.*

n	0	1	2	3	4	5
$z_n^{(2)}$	0	1	00	101	00100	10100101
$z_n^{(3)}$	0	1	020	1001	02010102	010010201020100
$z_n^{(4)}$	0	1	020	10301	020100102	010301020101020103
$z_n^{(5)}$	0	1	020	10301	0201040102	0103010201001020103

Table 6: The first few words of the sequence $(z_n^{(m)})_{n\geq 0}$ for $m = 2,3,4,5$.

Proof. The cases $n \in \{0,1,2\}$ are easily handled using Tables 2 and 6. Assume that $n \geq 3$. By definition, we get $w_{n-1} = a f_{n-1} b^{-1}$, where ab is the length-2 suffix of f_{n-1}. It can be verified by induction on n that for all $n \geq 1$, f_n ends with 01 (resp., 10) if n is odd (resp., even). Therefore, we can write $w_{n-1} = \underline{n-2}\, f_{n-1} (\underline{n-1})^{-1}$. Moreover, using the recursive definition of f_{n-1}, we get $w_{n-1} = \underline{n-2}\, f_{n-2} f_{n-3} (\underline{n-1})^{-1}$. Using Part (3) of Lemma 1, we have $w_{n-1} = (\underline{n-1})^{-1} f_{n-3}^R f_{n-2}^R\, \underline{n-2}$. On the other hand, Definition 20 gives $z_n^{(2)} = (\underline{n-3})^{-1} (h_{n-3}^{(2)})^R (h_{n-2}^{(2)})^R\, \underline{n-2}$. By Definition 6, we have $h_n^{(2)} = f_n$ for all n, which completes the proof. □

Definition 23. For every integer $n \geq 0$, let

$$\underline{\underline{n}} = \begin{cases} \varepsilon, & \text{if } m | n; \\ \underline{n}, & \text{otherwise.} \end{cases} \tag{5}$$

Moreover, let $\hat{n} = (\underline{\underline{n}})^{-1}\underline{n}$. Consequently

$$\hat{n} = \begin{cases} 0, & \text{if } m | n; \\ \varepsilon, & \text{otherwise.} \end{cases}$$

and $\hat{n} = \underline{n}(\underline{\underline{n}})^{-1}$.

In Table 7, the first few values of $\underline{\underline{n}}$ are displayed for $m \in \{2,3,4,5\}$.

	$\underline{\underline{0}}$	$\underline{\underline{1}}$	$\underline{\underline{2}}$	$\underline{\underline{3}}$	$\underline{\underline{4}}$	$\underline{\underline{5}}$	$\underline{\underline{6}}$	$\underline{\underline{7}}$	$\underline{\underline{8}}$	$\underline{\underline{9}}$
$m = 2$	ε	1	ε	1	ε	1	ε	1	ε	1
$m = 3$	ε	1	2	ε	1	2	ε	1	2	ε
$m = 4$	ε	1	2	3	ε	1	2	3	ε	1
$m = 5$	ε	1	2	3	4	ε	1	2	3	4

Table 7: The first few values of $\underline{\underline{n}}$ for $m \in \{2,3,4,5\}$.

Lemma 24. *For all $n \geq 0$, $\varphi_m(\underline{n}) = 0\,\underline{\underline{n+1}}$.*

Proof. The proof follows immediately from the definition of the morphism φ_m and Definition 23. □

In the first part of the following lemma, we study the length of the word z_n and in the second part, we find the first and the last letters of z_n.

Lemma 25. *1. For all $n \geq m+1$, $|z_n| = |z_{n-1}| + |z_{n-2}| + \cdots + |z_{n-m}|$.*

2. Let $n \geq 2$. The word z_n ends with $\underline{n-2}$. Moreover, z_n starts with 1 if $m|n-3$; 0 otherwise.

Proof. 1. The case $m = 2$ follows easily by induction on n and Definition 20. Suppose that $m \geq 3$. We proceed by induction on n. Using Definition 20 and Equation 1, we get

$$|z_n| = \begin{cases} |h_{n-3}| + 2|h_{n-2}|, & \text{if } 3 \leq n \leq m-1; \\ |h_{n-3}| + |h_{n-2}| + |h_{n-m}| + |h_{n-m+1}| + \cdots + |h_{n-3}|, & \text{if } n \geq m. \end{cases} \tag{6}$$

For the base case $n = m+1$ we have, $|z_{m+1}| = |h_{m-2}| + |h_{m-1}| + |h_1| + |h_2| + \cdots + |h_{m-2}|$. From Equation (1), we know that $|h_{m-2}| = |h_0| + |h_1| + \cdots + |h_{m-3}| + 1$. So, we get $|z_{m+1}| = |h_{m-1}| + 3|h_{m-2}| - 2$ since $|h_0^R| = 1$. By using (6) several times, we obtain $|z_m| + |z_{m-1}| + \cdots + |z_1| = |h_{m-1}| + 3|h_{m-2}| - 2 = |z_{m+1}|$, as desired.
Now, suppose that $n \geq m+2$. By using (6) several times, we have

$$|z_{n-1}| + |z_{n-2}| + \cdots + |z_{n-m}| = |h_{n-3}| + |h_{n-2}| + |h_{n-m}| + |h_{n-m+1}| + \cdots + |h_{n-3}| = |z_n|.$$

2. Using Definition 20, the result holds true for the case $n = 2$ as $z_2 = 010$. Assume that $n \geq 3$. By Definition 20 again, z_n starts with $(\underline{n-3})^{-1} h_{n-3}^R$ and ends with $\underline{n-2}$. If $m|n-3$, then Lemma 12 says that h_{n-3}^R starts with $01 = \underline{n-3}1$ which ends the proof. □

In the following remark, we rewrite the definition of the sequence z_n in a slightly different form. This helps us to prove Lemma 28 which gives a characterization of z_n in terms of the morphism φ_m and the word z_{n-1}.

Remark 26. For all $n \geq 3$, the word z_n can be written as $z_n = \alpha_n\, h_{n-3}^R\, h_{n-2}^R\, \beta_n\, z_n'\, \gamma_n$, where

$$z_n' = \prod_{i=(n-m)_*}^{(n-3)} h_i^R, \quad (j)_* := \max\{0, j\}, \quad (\alpha_n, \beta_n, \gamma_n) = \begin{cases} \big((n-3)^{-1}, n, n-2\big), & \text{if } n \leq m-1; \\ ((\underline{n-3})^{-1}, \varepsilon, \underline{n-2}), & \text{if } n \geq m. \end{cases}$$

Lemma 27. *For any sequence $(i_1, \ldots, i_k)$ of positive integers, we have*

$$\varphi_m\Big(\prod_{j=1}^{k} h_{i_j}^R\Big) = 0\Big(\prod_{j=1}^{k} h_{i_j+1}^R\Big)0^{-1}.$$

Consequently, for all $n \geq m+1$,

$$\varphi_m\Big(\prod_{i=n-m-1}^{n} h_i^R\Big) = 0\Big(\prod_{i=n-m}^{n+1} h_i^R\Big)0^{-1}.$$

Proof. To prove the first part, from Part (3) of Lemma 19, we find that

$$\varphi_m(h_{i_1}^R h_{i_2}^R \cdots h_{i_k}^R) = 0h_{i_1+1}^R 0^{-1} 0 h_{i_2+1}^R 0^{-1} \cdots 0 h_{i_k+1}^R 0^{-1}.$$

The second part is obtained from the first one. □

Lemma 28. *For all* $n \geq 2$, $z_n = (\widehat{n-3})^{-1}\varphi_m(z_{n-1})\widehat{n-2}$.

Proof. The proof is by induction on n. Assume that the result holds true up to $n-1$. If $n=2$, then $z_2 = \varphi_m(z_1)0 = (\widehat{2-3})^{-1}\varphi_m(z_1)\widehat{2-2}$. If $n=3$, then $z_3 = 0^{-1}\varphi_m(z_2) = (\widehat{3-3})^{-1}\varphi_m(z_2)\widehat{3-2}$. Now suppose that $n \geq 4$. Let $B_n = (\widehat{n-3})^{-1}\varphi_m(z_{n-1})\widehat{n-2}$. We divide the proof into two cases according to the value of n.

Case 1. Suppose that $n \leq m$. Using the induction hypothesis and Remark 26, we have

$$B_n = (\widehat{n-3})^{-1}\varphi_m\big((n-4)^{-1}h_{n-4}^R h_{n-3}^R (n-1)\prod_{i=0}^{n-4} h_i^R\,(n-3)\big)\,\widehat{n-2}.$$

From Part (3) of Lemma 19 and Lemmas 24 and 27, we get

$$B_n = (\widehat{n-3})^{-1}(0(n-3))^{-1}0h_{n-3}^R\,h_{n-2}^R\,a\prod_{i=0}^{n-3} h_i^R 0^{-1}0(n-2)\,\widehat{n-2},$$

where $a = \varepsilon$ if $n = m$; $a = n$ otherwise. As a consequence, we find that

$$B_n = (\underline{n-3})^{-1}h_{n-3}^R\,h_{n-2}^R\,a\prod_{i=0}^{n-3} h_i^R\,\underline{n-2}.$$

The result can be easily deduced using Remark 26.

Case 2. Assume that $n \geq m+1$. From the induction hypothesis and then Remark 26, we get

$$B_n = (\widehat{n-3})^{-1}\varphi_m\big((\underline{n-4})^{-1}h_{n-4}^R h_{n-3}^R \prod_{i=n-m-1}^{n-4} h_i^R(\underline{n-3})\big)\,\widehat{n-2}.$$

Part (3) of Lemma 19 and Lemmas 24 and 27 implies that

$$B_n = (\widehat{n-3})^{-1}(0\,(\underline{\underline{n-3}}))^{-1}\,0\,h_{n-3}^R\,h_{n-2}^R\prod_{i=n-m}^{n-3} h_i^R\,0^{-1}\,0\,(\underline{\underline{n-2}})\,\widehat{n-2}.$$

Using Definition 23, we obtain

$$B_n = (\underline{n-3})^{-1}h_{n-3}^R\,h_{n-2}^R\prod_{i=n-m}^{n-3} h_i^R\,\underline{n-2}.$$

Consequently, Remark 26 gives that $B_n = z_n$. □

Definition 29. For $n \geq 0$, let $P_n = \prod_{k=0}^{n-1} z_k$.

Lemma 30. *We have* $P_0 = \varepsilon$, $P_1 = 0$, $P_2 = 01$ *and for all* $n \geq 3$,

$$P_n = \varphi_m(P_{n-1})\,\widehat{n-3}.$$

Proof. The cases $n \in \{0,1,2,3\}$ are trivially verified. Assume that $n \geq 4$. By Lemma 28, we get

$$\begin{aligned} P_n = z_0\,z_1\cdots z_{n-1} &= 0\cdot 1\cdot(\widehat{-1})^{-1}\varphi_m(z_1)\widehat{0}\cdot(\widehat{0})^{-1}\varphi_m(z_2)\widehat{1}\cdot\cdots\cdot(\widehat{n-4})^{-1}\varphi_m(z_{n-2})\,\widehat{n-3} \\ &= \varphi_m(z_0\,z_1\cdots z_{n-2})\,\widehat{n-3} = \varphi_m(P_{n-1})\,\widehat{n-3}. \end{aligned}$$

□

Lemma 31. *We have the following factorization for the infinite m-bonacci word.*

$$h_\omega = \prod_{n\geq 0} z_n.$$

Proof. We need to prove the following two items.

1. For all $n \geq 1$, $|P_n| > |P_{n-1}|$,

2. $(P_n)_{n\geq 0}$ is a sequence of prefixes of h_ω.

To prove (1), note that $|P_n| = |P_{n-1}| + |z_{n-1}|$ for all $n \geq 1$. This yields $|P_n| > |P_{n-1}|$ since $|z_{n-1}| > 0$. Let us prove (2). We proceed by induction on n. The m-bonacci word h_ω starts with 01. Therefore, it is clear that P_n is a prefix of h_ω for $n \in \{0,1,2\}$. Now suppose that $n \geq 3$ and P_{n-1} is a prefix of h_ω. Using Lemma 30, we have $P_n = \varphi_m(P_{n-1})\widehat{n-3}$. The proof is divided into two cases according to whether m divides $n-3$ or not.

Case 1. Suppose that $m|n-3$. Thus, $\widehat{n-3} = 0$ and then $P_n = \varphi_m(P_{n-1})0$. By the induction hypothesis, there exists a letter $a \in A_m$ and an infinite word $\boldsymbol{z}$ over A_m such that $h_\omega = P_{n-1}a\boldsymbol{z}$. Since h_ω is the fixed point of φ_m, we know that $h_\omega = \varphi_m(h_\omega) = \varphi_m(P_{n-1}a\boldsymbol{z})$. Since $\varphi_m(a) = 0b$ with $b \in \{\varepsilon, 0, 1, \ldots, m-1\}$, we get $h_\omega = \varphi_m(P_{n-1})0\,b\,\varphi_m(\boldsymbol{z}) = P_n\,b\,\varphi_m(\boldsymbol{z})$, showing that P_n is also a prefix of h_ω.

Case 2. Suppose that $m \nmid n-3$. Therefore, $\widehat{n-3} = \varepsilon$ and then, $P_n = \varphi_m(P_{n-1})$. Thus $\varphi_m(P_{n-1}) = P_n$ is a prefix of $\varphi_m(h_\omega) = h_\omega$, as required. □

Justin and Vuillon [16] studied the return words of factors of standard episturmian words and their occurrences. In order to prove the main theorem of this section, we need to mention some useful lemmas. Recall from Definition 14 that u_n is the sequence of palindromic prefixes of h_ω.

Lemma 32. [*16*, Corollary 4.1] *Let $v \in A_m^*$ be any finite factor of h_ω. Let $j(v) \geq 1$ be such that $u_{j(v)}$ is the shortest palindromic prefix of h_ω which contains v as a factor, say $u_{j(v)} = fvg$ with $f, g \in A_m^*$. Then, y is a return word of v if and only if fyf^{-1} is a return word of $u_{j(v)}$. Moreover, the return words of the palindromic prefix $u_{j(v)+1}$ are $\mu_{j(v)}(i)$ for all $i \in A_m$.*

The following definition, which will be useful in this paper, is mentioned in [18].

Definition 33. Given an alphabet A, a set $X \subset A^+$ of non-empty words is a code on A if every word $w \in A^*$ has at most one factorization using words of X.

Lemma 34. [*14*, Lemma 15] *The set $\{0\underline{\underline{i}} \mid 1 \leq i \leq m\} = \{01, 02, \ldots, 0(m-1), 0\}$ of non-empty words is a code on the alphabet A_m.*

Lemma 35. [*14*, Lemma 17] *Let $x, y \in A_m^*$ be two finite words.*

1. *If $\varphi_m(x)0$ is a factor of $\varphi_m(y)0$, then x is a factor of y.*

2. *If $\varphi_m(x)$ is a factor of $\varphi_m(y)$ and x does not end with the letter $m-1$, then x is a factor of y.*

Lemma 36. *For $n \geq 3$, $(\underline{n-3})^{-1} h_{n-3}^R \underline{n-2}$ is not a factor of u_{n-1}.*

Proof. The proof is by induction on n. The statement can be readily verified for $n = 3$. Now assume that $n \geq 4$ and that the assertion holds for values less than n. We proceed by contradiction. Suppose that $(\underline{n-3})^{-1} h_{n-3}^{R}\, \underline{n-2} \prec u_{n-1}$. Using Parts (3) and (4) of Lemma 19, we have $(\underline{n-3})^{-1} 0^{-1} \varphi_m(h_{n-4}^{R}) 0\, \underline{n-2} \prec \varphi_m(u_{n-2})0$. So we get

$$(\underline{n-3})^{-1} 0^{-1} \varphi_m\big((\underline{n-4})(\underline{n-4})^{-1} h_{n-4}^{R} (\underline{n-3})(\underline{n-3})^{-1}\big)\, 0\, \underline{n-2} \prec \varphi_m(u_{n-2})0.$$

From Lemma 24 and then Definition 23, we find that

$$(\widehat{n-3})^{-1} \varphi_m\big((\underline{n-4})^{-1} h_{n-4}^{R} (\underline{n-3})\big)\, \widehat{n-2} \prec \varphi_m(u_{n-2})0. \tag{7}$$

We consider three cases according to whether m divides $n-3$ or $n-2$ or none of them.

Case 1. Suppose that $m \nmid n-3$ and $m \nmid n-2$. Replacing $\widehat{n-3} = \widehat{n-2} = \varepsilon$ into (7), we get $\varphi_m\big((\underline{n-4})^{-1} h_{n-4}^{R}\, \underline{n-3}\big) \prec \varphi_m(u_{n-2})0$. Since $\varphi_m\big((\underline{n-4})^{-1} h_{n-4}^{R}\, \underline{n-3}\big)$ ends with $\underline{\underline{n-2}} \neq 0$, we find that $\varphi_m\big((\underline{n-4})^{-1} h_{n-4}^{R}\, \underline{n-3}\big) \prec \varphi_m(u_{n-2})$. Since $(\underline{n-4})^{-1} h_{n-4}^{R}\, \underline{n-3}$ ends with $\underline{n-3} \neq m-1$, we deduce from Part (2) of Lemma 35 that $(\underline{n-4})^{-1} h_{n-4}^{R}\, \underline{n-3} \prec u_{n-2}$, contradicting the induction hypothesis.

Case 2. Suppose that $m \mid n-3$ and $m \nmid n-2$. Substituting $\widehat{n-3} = 0$ and $\widehat{n-2} = \varepsilon$ into (7), we have $0^{-1} \varphi_m\big((\underline{n-4})^{-1} h_{n-4}^{R}\, \underline{n-3}\big) \prec \varphi_m(u_{n-2})0$. By Lemma 12, $0^{-1} \varphi_m\big((\underline{n-4})^{-1} h_{n-4}^{R}\, \underline{n-3}\big)$ starts with 1. From Lemma 12 and Part (4) of Lemma 15, $\varphi_m(u_{n-2})$ starts with 0. Hence, $\varphi_m\big((\underline{n-4})^{-1} h_{n-4}^{R}\, \underline{n-3}\big)$ is a factor of $\varphi_m(u_{n-2})0$. As in Case (1), we reach a contradiction since $\underline{\underline{n-2}} \neq 0$ and $\underline{n-3} \neq m-1$.

Case 3. Suppose that $m \nmid n-3$ and $m \mid n-2$. By replacing $\widehat{n-3} = \varepsilon$ and $\widehat{n-2} = 0$ into (7), we find that $\varphi_m\big((\underline{n-4})^{-1} h_{n-4}^{R}\, \underline{n-3}\big)\, 0 \prec \varphi_m(u_{n-2})\, 0$. From Part (1) of Lemma 35, $(\underline{n-4})^{-1} h_{n-4}^{R}\, \underline{n-3} \prec u_{n-2}$, contradicting the induction hypothesis. □

Lemma 37. *For all $n \geq 0$, z_n is not a factor of z_{n+1}.*

Proof. If $m = 2$, then the result follows from Part (1) of Lemma 1 and Lemma 22. We prove it for $m \geq 3$. The result is obviously true for $n \in \{0,1,2\}$, see Table 6. Let $n \geq 3$. By induction on n, we assume that the result holds true up to $n-1$ and we show that it is still true for n. By contradiction, suppose that $z_n \prec z_{n+1}$. We obtain from Lemma 28 that $(\widehat{n-3})^{-1} \varphi_m(z_{n-1})\, \widehat{n-2} \prec (\widehat{n-2})^{-1} \varphi_m(z_n)\, \widehat{n-1}$. So, we get

$$(\widehat{n-3})^{-1} \varphi_m(z_{n-1})\, \widehat{n-2} \prec \varphi_m(z_n)\, \widehat{n-1}. \tag{8}$$

We divide the proof into two cases according to whether m divides $n-3$ or not.

Case 1. Suppose that $m \nmid n-3$. Substituting $\widehat{n-3} = \varepsilon$ into (8), we find that

$$\varphi_m(z_{n-1})\, \widehat{n-2} \prec \varphi_m(z_n)\, \widehat{n-1}. \tag{9}$$

Now we consider three cases according to whether m divides $n-1$ or $n-2$ or none of them.

Case 1-1. Assume that $m \nmid n-1$ and $m \nmid n-2$. Plugging $\widehat{n-1} = \widehat{n-2} = \varepsilon$ into (9), we have $\varphi_m(z_{n-1}) \prec \varphi_m(z_n)$. Part (2) of Lemma 25 implies that z_{n-1} ends with $\underline{n-3}$. Since $m \nmid n-2$ implies $\underline{n-2} \neq m$ and thus $\underline{n-3} \neq m-1$, using Part (2) of Lemma 35, z_{n-1} is a factor of z_n, contradicting the induction hypothesis.

Case 1-2. Suppose that $m|n-1$ and $m \nmid n-2$. Replacing $\widehat{n-1}=0$ and $\widehat{n-2}=\varepsilon$ into (9), we find that $\varphi_m(z_{n-1}) \prec \varphi_m(z_n)0$. Using Part (2) of Lemma 25, z_{n-1} ends with $\underline{n-3}$. It follows that $\varphi_m(z_{n-1})$ ends with $0\underline{\underline{n-2}}$. In conclusion, $\varphi_m(z_{n-1}) \prec \varphi_m(z_n)$. As in Case (1-1), we reach a contradiction.

Case 1-3. Suppose that $m \nmid n-1$ and $m \mid n-2$. Plugging $\widehat{n-1}=\varepsilon$ and $\widehat{n-2}=0$ into (9), we get $\varphi_m(z_{n-1})0 \prec \varphi_m(z_n)$. We then have, $\varphi_m(z_{n-1})\,0 \prec \varphi_m(z_n)\,0$. Part (1) of Lemma 35 gives, $z_{n-1} \prec z_n$ which contradicts the induction hypothesis.

Case 2. Suppose that $m|n-3$. By replacing $\widehat{n-3}=0$ into (8), we get $0^{-1}\varphi_m(z_{n-1})\widehat{n-2} \prec \varphi_m(z_n)$. We deduce from Part (2) of Lemma 25 that z_n (resp., z_{n-1}) starts with 1 (resp., 0). So $0^{-1}\varphi_m(z_{n-1})$ (resp., $\varphi_m(z_n)$) starts with 1 (resp., 0). Since each occurrence of 1 in $\varphi_m(z_n)$ is preceded by a 0, we conclude that $\varphi_m(z_{n-1})\,\widehat{n-2} \prec \varphi_m(z_n)$. and thus, $\varphi_m(z_{n-1}) \prec \varphi_m(z_n)$. As in Case (1-1), we reach a contradiction since $\underline{n-3} \neq m-1$. □

Lemma 38. *Let $n \geq 1$ and γ_n be the last letter of z_n. The word z_n is not a factor of $z_{n-1}z_n\gamma_n^{-1}$.*

Proof. To prove the result, we proceed by induction on n. If $n=1$, then $z_1=1$ is not a factor of $z_0z_1\gamma_1^{-1}=z_0=0$. If $n=2$, then the word $z_1z_2\gamma_2^{-1}$ does not contain z_2. Now assume that $n \geq 3$ and that the result is true for all values less than n. We proceed by contradiction and suppose that $z_n \prec z_{n-1}z_n\gamma_n^{-1}$. By Part (2) of Lemma 25, $\gamma_n=\underline{n-2}$. Lemma 28 implies that

$$(\widehat{n-3})^{-1}\varphi_m(z_{n-1})\,\widehat{n-2} \prec (\widehat{n-4})^{-1}\varphi_m(z_{n-2})\,\widehat{n-3}\,(\widehat{n-3})^{-1}\varphi_m(z_{n-1})\,\widehat{n-2}\,(\underline{n-2})^{-1}.$$

So, we have $(\widehat{n-3})^{-1}\varphi_m(z_{n-1})\,\widehat{n-2} \prec \varphi_m(z_{n-2}z_{n-1})\,(\underline{\underline{n-2}})^{-1}$. Therefore,

$$(\widehat{n-3})^{-1}\varphi_m(z_{n-1})\,\widehat{n-2} \prec \varphi_m(z_{n-2}z_{n-1}(\underline{n-3})^{-1}\,\underline{n-3})\,(\underline{\underline{n-2}})^{-1}.$$

Using Definition 23, we find that

$$(\widehat{n-3})^{-1}\varphi_m(z_{n-1})\,\widehat{n-2} \prec \varphi_m(z_{n-2}z_{n-1}(\underline{n-3})^{-1})\,0. \tag{10}$$

We divide the proof into three cases according to whether m divides $n-3$ or $n-2$ or none of them.

Case 1. Assume that $m \nmid n-3$ and $m \nmid n-2$. Plugging $\widehat{n-3}=\widehat{n-2}=\varepsilon$ into (10), we find that $\varphi_m(z_{n-1}) \prec \varphi_m\big(z_{n-2}z_{n-1}(\underline{n-3})^{-1}\big)\,0$. From Part (2) of Lemma 25, z_{n-1} ends with $\underline{n-3}$. Thus, $\varphi_m(z_{n-1})$ ends with $0\underline{\underline{n-2}}$. Using Definition 23, $\underline{\underline{n-2}} \neq 0$. It follows that $\varphi_m(z_{n-1}) \prec \varphi_m\big(z_{n-2}z_{n-1}(\underline{n-3})^{-1}\big)$. As z_{n-1} ends with $\underline{n-3}$, Part (2) of Lemma 35 tells us that $z_{n-1} \prec z_{n-2}z_{n-1}(\underline{n-3})^{-1}$, since $m \nmid n-2$ implies $\underline{n-2} \neq m$ and thus $\underline{n-3} \neq m-1$. This contradicts the induction hypothesis.

Case 2. Assume that $m \mid n-3$ and $m \nmid n-2$. By replacing $\widehat{n-3}=0$ and $\widehat{n-2}=\varepsilon$ into (10), we have $0^{-1}\varphi_m(z_{n-1}) \prec \varphi_m\big(z_{n-2}z_{n-1}(\underline{n-3})^{-1}\big)\,0$. By Part (2) of Lemma 25, z_{n-1} and $z_{n-2}z_{n-1}$ start with 1. So $0^{-1}\varphi_m(z_{n-1})$ starts with 2 and $\varphi_m(z_{n-2}z_{n-1})$ starts with 0. We conclude that $\varphi_m(z_{n-1}) \prec \varphi_m\big(z_{n-2}z_{n-1}(\underline{n-3})^{-1}\big)\,0$. By Part (2) of Lemma 25, $\varphi_m(z_{n-1})$ ends with 01 since $\underline{n-3}=0$. So, $\varphi_m(z_{n-1}) \prec \varphi_m\big(z_{n-2}z_{n-1}(\underline{n-3})^{-1}\big)$. As in Case (1), we reach a contradiction.

Case 3. Assume that $m \nmid n-3$ and $m \mid n-2$. Plugging $\widehat{n-3}=\varepsilon$ and $\widehat{n-2}=0$ into (10), we have $\varphi_m(z_{n-1})\,0 \prec \varphi_m\big(z_{n-2}z_{n-1}(\underline{n-3})^{-1}\big)\,0$. Part (1) of Lemma 35 implies that $z_{n-1} \prec z_{n-2}z_{n-1}(\underline{n-3})^{-1}$, contradicting the induction hypothesis. □

Lemma 39. *For all $n \geq 1$, z_n is not a factor of P_n.*

Proof. Using Part (4) of Lemma 1 and Lemma 22, the assertion is true for the case $m = 2$. Let us suppose that $m \geq 3$. The proof is by induction on n. It can be easily checked for $n \in \{1,2\}$ using Table 6. Now suppose that z_j is not a factor of P_j for all $3 \leq j \leq n-1$. We show it is still true for $j = n$. Arguing by contradiction, assume that $z_n \prec P_n$. From Lemmas 28 and 30,

$$(\widehat{n-3})^{-1}\,\varphi_m(z_{n-1})\,\widehat{n-2} \prec \varphi_m(P_{n-1})\,\widehat{n-3}. \tag{11}$$

The proof is divided into three cases according to whether m divides $n-3$ or $n-2$ or none of them.

Case 1. Assume that $m \nmid n-3$ and $m \nmid n-2$. Substituting $\widehat{n-3} = \widehat{n-2} = \varepsilon$ into (11), we get $\varphi_m(z_{n-1}) \prec \varphi_m(P_{n-1})$. Part (2) of Lemma 25 implies that z_{n-1} ends with $\underline{n-3}$. Thus, using Part (2) of Lemma 35, z_{n-1} is a factor of P_{n-1}, since $m \nmid n-2$ implies $\underline{n-2} \neq m$ and thus $\underline{n-3} \neq m-1$, which contradicts the induction hypothesis.

Case 2. Assume that $m \mid n-3$ and $m \nmid n-2$. Replacing $\widehat{n-3} = 0$ and $\widehat{n-2} = \varepsilon$ into (11), we find that $0^{-1}\,\varphi_m(z_{n-1}) \prec \varphi_m(P_{n-1})0$. By Part (2) of Lemma 25, z_{n-1} begins and ends with 0 and P_{n-1} begins with $z_0 = 0$. Thus $0^{-1}\varphi_m(z_{n-1})$ (resp., $\varphi_m(P_{n-1})0$) begins and ends with 1 (resp., 0). It follows that $\varphi_m(z_{n-1})$ is a factor of $\varphi_m(P_{n-1})$. As in Case (1), we reach a contradiction since $\underline{n-3} \neq m-1$.

Case 3. Assume that $m \nmid n-3$ and $m \mid n-2$. Plugging $\widehat{n-3} = \varepsilon$ and $\widehat{n-2} = 0$ into (11), we have $\varphi_m(z_{n-1})0 \prec \varphi_m(P_{n-1})$. So, we have $\varphi_m(z_{n-1})0 \prec \varphi_m(P_{n-1})0$. From Part (1) of Lemma 35, $z_{n-1} \prec P_{n-1}$ contradicting the induction hypothesis. □

Lemma 40. *For all $n \geq 0$, z_n is closed.*

Proof. The result follows from Lemma 4 for the case $m = 2$. Now suppose that $m \geq 3$. Using Table 6, the result is clearly true for the cases $n \in \{0,1,2\}$. We assume that $n \geq 3$. Let us show, equivalently, that there exists a border v of z_n which $|z_n|_v = 2$, that is, we prove that z_n is a complete return to v. Set $v = (\underline{n-3})^{-1}\,h^R_{n-3}\,\underline{n-2}$. Clearly, v is a border of z_n. There are two cases to consider according to the value of n.

Case 1. Suppose that $n \leq m-1$. Observe that using Remark 26,

$$z_n = (n-3)^{-1}\,h^R_{n-3}\,h^R_{n-2}\,n\prod_{i=0}^{n-3} h^R_i\,(n-2).$$

In order to show that z_n is a complete return to v, it suffices to prove that y is a return word of v, where

$$y = (n-3)^{-1}\,h^R_{n-3}\,h^R_{n-2}\,\underline{n}\prod_{i=0}^{n-4} h^R_i\,(n-3). \tag{12}$$

For this purpose, we prove that fyf^{-1} is a return word of $u_{j(v)}$, where f and $j(v)$ are those of Lemma 32. First, we find the minimal integer $j(v)$ such that the word v is a factor of $u_{j(v)}$. We obtain from Lemma 36 that v does not occur in u_{n-1}. From Part (4) of Lemma 15, we have $h^R_{n-3}\,h^R_{n-2} \triangleright u_n$. Using Lemma 12, h^R_{n-3} (resp., h^R_{n-2}) starts with $n-3$ (resp., $n-2$). Thus, $(n-3)^{-1}\,h^R_{n-3}\,(n-2) \prec u_n$. We conclude that $j(v) = n$. Now, let

$$f = \prod_{i=0}^{n-4} h^R_i\,(n-3). \tag{13}$$

Plugging (13) and (12) into fyf^{-1} and then using Part (4) of Lemma 15, we obtain

$$fyf^{-1} = \prod_{i=0}^{n-2} h^R_i\ n\ = u_n\,n.$$

Using Part (2) of Lemma 19, we know that $fyf^{-1} = \mu_{n-1}(n)$. Then, by Lemma 32, y is a return word of v. So, the desired conclusion is obtained in this case.

Case 2. Suppose that $n \geq m$. The proof is obtained in the same manner as the first case. By Remark 26, we find that

$$z_n = (\underline{n-3})^{-1}\,h^R_{n-3}\,h^R_{n-2} \prod_{i=(n-m)_*}^{n-3} h^R_i\ \underline{n-2}.$$

Now let

$$y = (\underline{n-3})^{-1}\,h^R_{n-3}\,h^R_{n-2} \prod_{i=(n-m)_*}^{n-4} h^R_i\ \underline{n-3}. \tag{14}$$

Similarly to the first case, the minimal integer $j(v)$ such that v is a factor of $u_{j(v)}$ equals $j(v) = n$. Substituting (13) and (14) into fyf^{-1}, we get

$$fyf^{-1} = \prod_{i=0}^{n-4} h^R_i\ h^R_{n-3}\,h^R_{n-2} \prod_{i=(n-m)_*}^{n-4} h^R_i\,(\prod_{i=0}^{n-4} h^R_i)^{-1} = \prod_{i=0}^{n-2} h^R_i\,(\prod_{i=0}^{n-m-1} h^R_i)^{-1}.$$

By Part (4) of Lemma 15, we have $fyf^{-1} = u_n\,u^{-1}_{n-m+1}\ = h_{n-2}\cdots h_{n-m}$. Now using Part (1) of Lemma 19, we find that $fyf^{-1} = \mu_{n-1}(\underline{n-m})$. We conclude from Lemma 32 that y is a return word of v. So, we obtain the desired conclusion in this case. □

Lemma 41. *For all $n \geq 4$, the word $(\underline{n-3})^{-1}h^R_{n-3}$ is not a factor of u_{n-3}.*

Proof. The proof is by induction on n. The case $n = 4$ can be easily checked by hand. Assume that $n \geq 5$ and that the claim holds true for all values less than n and consider the case n. Suppose to the contrary that $(\underline{n-3})^{-1}\,h^R_{n-3} \prec u_{n-3}$. Using Parts (3) and (4) of Lemma 19,

$$(\underline{n-3})^{-1}\,0^{-1}\,\varphi_m(h^R_{n-4})0 \prec \varphi_m(u_{n-4})0.$$

So, $(\underline{n-3})^{-1}\,0^{-1}\,\varphi_m\big((\underline{n-4})(\underline{n-4})^{-1}h^R_{n-4}\big)\,0 \prec \varphi_m(u_{n-4})0$. Using Lemma 24, we get

$$(\widehat{n-3})^{-1}\,\varphi_m\big((\underline{n-4})^{-1}h^R_{n-4}\big)\,0 \prec \varphi_m(u_{n-4})0. \tag{15}$$

Now we divide the proof into two cases according to whether m divides $n-3$ or not.

Case 1. Suppose that $m \nmid (n-3)$. We obtain that $\varphi_m((\underline{n-4})^{-1}h^R_{n-4})0 \prec \varphi_m(u_{n-4})0$ by replacing $\widehat{n-3} = \varepsilon$ into (15). Part (1) of Lemma 35 implies that $(\underline{n-4})^{-1}h^R_{n-4} \prec u_{n-4}$ which contradicts the induction hypothesis.

Case 2. Suppose that $m \mid (n-3)$. we find that $0^{-1}\varphi_m\big((\underline{n-4})^{-1}h^R_{n-4}\big)0 \prec \varphi_m(u_{n-4})0$ by plugging $\widehat{n-3} = 0$ into (15). Using Lemma 12, $0^{-1}\varphi_m\big((\underline{n-4})^{-1}h^R_{n-4}\big)$ begins with 1. Also, from Lemma 12 and Part (4) of Lemma 15, $\varphi_m(u_{n-4})$ begins with 0. Therefore, $\varphi_m\big((\underline{n-4})^{-1}h^R_{n-4}\big)0 \prec \varphi_m(u_{n-4})0$. As in Case (1), we reach a contradiction. □

In the following theorem, we obtain the closed z-factorization of the m-bonacci word based on the sequence of words z_n.

Theorem 42. *The closed z-factorization of the m-bonacci word is*

$$cz(h_\omega) = (z_0, z_1, z_2, z_3, \ldots).$$

Proof. First note that using Theorem 5, the case $m = 2$ is covered. Suppose that $m \geq 3$. From Lemma 31, $h_\omega = \prod_{n\geq 0} z_n$. So, the first few factors of h_ω are $\{z_0, \ldots z_6\}$. It can be easily checked that z_i, $1 \leq i \leq 6$, are closed z-factors of h_ω. Now assume that $n \geq 7$. In order to prove the statement, we need to show the following three claims.

1. The word z_n is closed.

2. The word z_n does not appear in $P_{n+1}\gamma_n^{-1}$ where γ_n is the last letter of z_n.

3. Every closed prefix of z_n has already appeared in P_n.

Claim (1) is true by Lemma 40 and Claim (2) is true respectively by Lemmas 37, 38 and 39. To prove Claim (3), we find the largest closed prefix of z_n and prove that this prefix has already appeared in P_n. Lemma 40 implies that a frontier of z_n is $(\underline{n-3})^{-1}h^R_{n-3}\,\underline{n-4}$. Now we set $v = (\underline{n-3})^{-1}h^R_{n-3}$ and find the closed prefix of z_n with border v. We divide the proof into two cases according to the value of n.

Case 1. Assume that $n \leq m-1$. In order to prove the statement, we prove that the word y is a return word of v, where

$$y = (\underline{n-3})^{-1}h^R_{n-3}\,\underline{n-2}. \tag{16}$$

To show this, we need to prove that fyf^{-1} is a return word of $u_{j(v)}$ where f and $j(v)$ are those of Lemma 32. Using Part (5) of Lemma 15, $u_{n-2} = (\underline{n-3})^{-1}h^R_{n-3}$. Therefore, the minimal integer $j(v)$ such that v is a factor of $u_{j(v)}$ equals $j(v) = n-2$ and thus, set $f = \varepsilon$. Replacing $f = \varepsilon$ and Equation (16) into fyf^{-1}, we get $fyf^{-1} = (\underline{n-3})^{-1}h^R_{n-3}(n-2)$. By Part (5) of Lemma 15 and then Part (2) of Lemma 19, we find that

$$fyf^{-1} = u_{n-2}(n-2) = \mu_{n-3}(n-2).$$

Lemma 32 implies that y is a return word of v. So, $(\underline{n-3})^{-1}h^R_{n-3}(\underline{n-2})(\underline{n-3})^{-1}h^R_{n-3}$ is a closed prefix of z_n. Using Definition 20 and replacing z_{n-1}, z_{n-2} and z_{n-3} into P_n, we find that $(\underline{n-3})^{-1}h^R_{n-3}(\underline{n-2})(\underline{n-3})^{-1}h^R_{n-3}$ is a factor of P_n which ends the proof in this case.

Case 2. Suppose that $n \geq m$. This case is similar to the first case. Set

$$y = (\underline{n-3})^{-1}h^R_{n-3}(h^R_{n-m-3})^{-1}\,\underline{n-3}. \tag{17}$$

Lemma 41 tells us that v is not a factor of u_{n-3}. On the other hand, using Part (4) of Lemma 15, we get

$$u_{n-2} = h_0^R \cdots h_{n-4}^R = h_0^R \cdots h_{n-m-4}^R h_{n-3}^R = h_0{}^R \cdots h_{n-m-4}^R \underline{n-3}(\underline{n-3})^{-1} h_{n-3}^R.$$

Therefore, the minimal integer $j(v)$ such that v occurs in $u_{j(v)}$ equals $j(v) = n-2$. Now set

$$f = h_0{}^R h_1{}^R \cdots h_{n-m-4}^R \underline{n-3}. \tag{18}$$

Substituting Equations (17) and (18) into fyf^{-1}, we find that

$$fyf^{-1} = h_0^R h_1^R \cdots h_{n-4}^R (h_0^R h_1^R \cdots h_{n-m-3}^R)^{-1}.$$

Using Part (4) of Lemma 15 and then Part (1) of Lemma 19, we have

$$fyf^{-1} = u_{n-2} u_{n-m-1}^{-1} = h_{n-4} h_{n-5} \cdots h_{n-m-2} = \mu_{n-3}(\underline{n-m-2}).$$

Lemma 32 implies that fyf^{-1} is a return word of u_{n-2}. Thus, y is a return word of v, that is, the word $(\underline{n-3})^{-1} h_{n-3}^R (h_{n-m-3}^R)^{-1} h_{n-3}^R$ is a closed prefix of z_n. The result is obtained by using Definition 20 and substituting z_{n-1}, z_{n-2} and z_{n-3} into P_n. □

5 Relation between the palindromic and closed z-factorizations of the m-bonacci words

In this section, we link two kinds of factorizations of the m-bonacci words, namely the palindromic and closed z-factorizations. In [14], we introduced a variation of the z-factorization, the palindromic z-factorization, in which each factor is palindromic. Also, we computed this factorization for the Fibonacci word and more generally for the m-bonacci words. The palindromic z-factorization of a word $\boldsymbol{w}$ is $pz(\boldsymbol{w}) = (p_1, p_2, \ldots)$ such that p_i is the shortest palindromic prefix of $p_i p_{i+1} \cdots$ which occurs exactly once in $p_1 p_2 \cdots p_i$.

In the following lemma, the length of the n-th palindromic z-factor of the m-bonacci word is expressed by the previous m palindromic z-factors.

Lemma 43. [*14*, Corollary 26] *Let $pz(h_\omega) = (p_1, p_2, \ldots)$ be the palindromic z-factorization of the m-bonacci word. If m is even, then, for all $n \geq m-1$, we have*

$$|p_n| = |p_{n-1}| + |p_{n-2}| + \cdots + |p_{n-m}|.$$

If m is odd, then, for all $n \geq m-1$, we have

$$|p_n| = |p_{n-1}| + |p_{n-2}| + \cdots + |p_{n-m}| + (-1)^n.$$

We compare these two types of factorizations in the following corollary.

Corollary 44. *Let $pz(h_\omega) = (p_1, p_2, \ldots)$ and $cz(h_\omega) = (z_1, z_2, \ldots)$ be respectively the palindromic z-factorization and the closed z-factorization of h_ω. If $m = 2$, then for all $n \geq m-1$, $|z_n| = |p_n|$. If $m \geq 3$, then for every even integer m and for all $n \geq m-1$, $|z_n| = |p_n|$.*

Proof. The case $m = 2$ follows from Theorem 5 and the case $m \geq 3$ follows from Part (1) of Lemma 25 and Lemma 43. □

6 The oc-sequence of the m-bonacci words

The notion of the oc-sequence of a word w is introduced in [7]. It is a binary sequence whose n-th element is 1 if the length-n prefix of w is closed; otherwise, it is 0. In this section, our aim is to show that the sequence of the lengths of the maximum consecutive 1's in $oc(h_\omega^{(m)})$ is exactly the m-bonacci word.

In the following lemma, we prove that the palindromic prefixes of the m-bonacci words are closed.

Lemma 45. *For all $n \geq 2$, u_n is closed.*

Proof. To prove this statement, it suffices to find a border v of u_n such that $|u_n|_v = 2$. Using Lemma 15, we have $u_n = h_{n-2}u_{n-1}$. As u_n is a palindrome by its definition, we find that u_{n-1} is a border of u_n. We show that $v = u_{n-1}$. We proceed by induction on n. The result is clear for $n \in \{2,3\}$ by Table 5. Now suppose that $n \geq 4$ and the result holds true up to $n-1$. We will show that it is still true for n. We proceed by contradiction and suppose that u_{n-1} is a proper factor of u_n. Therefore, there exist non-empty words s and t such that $u_n = s\,u_{n-1}\,t$. From Part (4) of Lemma 19, we obtain that $\varphi_m(u_{n-1})0 = s\,\varphi_m(u_{n-2})0t$. From Lemma 34, we deduce that $\varphi_m(u_{n-1})$ and $\varphi_m(u_{n-2})$ have a unique factorization using words of the set $C = \{01, 02, \ldots, 0(m-1), 0\}$. Thus, we have $\varphi_m(u_{n-1})0 = y_1 \cdots y_k 0$ and $\varphi_m(u_{n-2}) = x_1 \cdots x_\ell$, with $x_1, \ldots, x_\ell, y_1, \ldots, y_k \in C$ and $\ell, k \geq 1$. By uniqueness of the factorization, there exists $1 \leq i \leq k$ such that for all $1 \leq j \leq l$, we have $x_j = y_{i+j-1}$. Also, $s = y_1 \cdots y_{i-1}$ and $0t = y_{i+l} \cdots y_k 0$. Thus, there exist words $s', t' \in A_m^*$ such that $\varphi_m(s') = s$ and $\varphi_m(t') = t$. Finally, we deduce that

$$\varphi_m(u_{n-1})0 = s\varphi_m(u_{n-2})0t = \varphi_m(s')\varphi_m(u_{n-2})\varphi_m(t') = \varphi_m(s'u_{n-2}t').$$

By injectivity of φ_m, u_{n-2} is a proper factor of u_{n-1}, contradicting the induction hypothesis. □

Definition 46. Let $n \geq 2$. We define t_n by

$$t_n = \begin{cases} (n-1)h_0^R \cdots h_{n-3}^R, & \text{if } 2 \leq n \leq m-1; \\ h_{n-m-1}^R h_{n-m}^R \cdots h_{n-3}^R, & \text{if } n \geq m. \end{cases}$$

Then we have, $h_{n-1}^R = t_n h_{n-2}^R$. Hence, $u_n t_n h_{n-2}^R = u_{n+1}$.

Definition 47. Let w be a prefix of the m-bonacci word h_ω and let $n(w)$ be the unique positive integer satisfying $|u_{n(w)}| < |w| \leq |u_{n(w)+1}|$. Then, w is a prefix of *type-1* if it satisfies

$$|u_n(w)| + |t_{n(w)}| < |w| \leq |u_{n(w)+1}| \tag{19}$$

and of *type-2* if it satisfies

$$|u_{n(w)}| < |w| \leq |u_{n(w)}| + |t_{n(w)}|. \tag{20}$$

Lemma 48. *Let w be a prefix of the m-bonacci word h_ω satisfying $|u_{n(w)}| < |w| \leq |u_{n(w)+1}|$ for some positive integer $n(w)$ and let $v = u_{n(w)}^{-1}w$.*

1. *If w is a prefix of type-1, then $v = t_{n(w)}x$, where x is a non-empty prefix of $h^R_{n(w)-2}$.*

2. *If w is a prefix of type-2, then v is a non-empty prefix of $t_{n(w)}$ (possibly $v = t_{n(w)}$).*

Proof. The result directly follows from Inequalities (19) and (20) since $u_{n(w)+1} = u_{n(w)}t_{n(w)}h^R_{n(w)-2}$ and $w = u_{n(w)}v$. □

In the following theorem, we characterize closed (and open) prefixes of the m-bonacci word.

Theorem 49. *We have the following properties.*

1. *The prefixes of type-1 of the m-bonacci word are closed.*

2. *The prefixes of type-2 of the m-bonacci word are open.*

Proof. First observe that $0 = u_2$ is the type-1 prefix of h_ω. Now we examine the prefixes of the m-bonacci word whose length is greater than 1. To prove the first item in the statement, consider each prefix w of h_ω with $w = u_{n(w)}v$ and $v = t_{n(w)}x$ such that $\varepsilon \neq x \lhd h^R_{n(w)-2}$. In order to prove the assertion, we need to find a frontier u' of w. There are two cases to consider according to the value of $n(w)$.

Case 1. Assume that $n(w) \le m-1$. Therefore, $t_{n(w)} = (n(w)-1)h^R_0 \cdots h^R_{n(w)-3}$. As $n(w)-1$ does not occur in $u_{n(w)-1}$, the longest border of w is $h^R_0 h^R_1 \cdots h^R_{n(w)-3}x = u_{n(w)-1}x$. Using the proof of Lemma 45, $u_{n(w)-1}$ has no internal occurrence in $u_{n(w)}$. Hence, $u_{n(w)-1}x$ has no internal occurrence in w and we are done.

Case 2. Assume that $n(w) \ge m$. Thus, $t_{n(w)} = h^R_{n(w)-m-1}h^R_{n(w)-m} \cdots h^R_{n(w)-3}$. Set $u' = u_{n(w)-m}t_{n(w)}x = u_{n(w)-1}x$. It is obvious that u' is a suffix of w. Also by Part (3) of Lemma 15, u' is a prefix of w. As in Case (1), $u_{n(w)-1}x$ has no internal occurrence in w and we get the result. To prove the second item in the statement, let $w = u_{n(w)}v$, where $v \lhd t_{n(w)}$. We consider two cases according to the value of $n(w)$.

Case 1. Suppose that $n(w) \le m$. Then, we have $t_{n(w)} = (n(w)-1)h^R_0 \cdots h^R_{n(w)-3}$. Since $n(w)-1$ does not occur in $u_{n(w)}$, the longest border of w is $(n(w)-1)^{-1}v$ which is also a border of $u_{n(w)}$. So, it appears three times in w. In other words, w is open.

Case 2. Suppose that $n(w) \ge m$. We prove that the longest border of w is $u_{n(w)-m}v$. Suppose that the word xv is a border of w. As $w = u_{n(w)}v$, x is a border of $u_{n(w)}$. Since $u_{n(w)}$ is palindromic, x is a palindromic prefix of $u_{n(w)}$. Thus it is of the form $x = u_i$ for some i. As v is a prefix of $t_{n(w)} = h^R_{n(w)-m-1}h^R_{n(w)-m} \cdots h^R_{n(w)-3}$, the longest u_i, $i \ge 2$, such that u_iv is a border of w is $u_{n(w)-m}$. On the other hand, $u_{n(w)-m}v$ is a border of $u_{n(w)}$. Therefore, it occurs three times in w and thus w is open. □

Corollary 50. *The sequence of lengths of the maximum consecutive 1's in the oc-sequence of the m-bonacci word is exactly $|h_i|$, for all $i \ge 0$.*

Proof. Using Inequality (19), we deduce that the number of consecutive closed prefixes of the m-bonacci words h_ω is equal to $|t_n| + |h^R_{n-2}| = |h^R_{n-1}|$ and we get the result. □

Example 51. Let t_n be the sequence of Tribonacci numbers where $T_{-1} = 1$, $T_0 = 1$, $T_1 = 2$ and for all $n \ge 2$, $T_n = T_{n-1} + T_{n-2} + T_{n-3}$. Then, we have $oc(h^{(3)}_\omega) = 10\prod_{i\ge0} 1^{T_i}0^{T_{i-1}+T_i}$. Table 8 shows the first few values of the oc-sequence for the infinite Tribonacci word.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
$h_\omega^{(3)}$	0	1	0	2	0	1	0	0	1	0	2	0	1	0	1	0	2	0	1	0	0	1	0	2
$oc(h_\omega^{(3)})$	1	0	1	0	0	1	1	0	0	0	1	1	1	1	0	0	0	0	0	0	1	1	1	1

Table 8: The first few values of the oc-sequence of the Tribonacci word

7 Open problems

It is interesting to find the closed z-factorization of other infinite words such as episturmian words and automatic words. Another interesting problem is to obtain the closed c-factorizations of infinite words and find a relation between the closed z-factorization and the closed c-factorization of infinite words. We leave it as an open problem to characterize the closed c-factorization of the m-bonacci word.

Problem 52. Let $c(h_\omega) = (c_0, c_1, c_2, \ldots)$ be the closed c-factorization of the m-bonacci word. For all $m \geq 3$ and $n \geq 2m-1$, we conjecture that $|c_i| = |h^m_{i-m+1}|$.

As a further research, it would be interesting to find the closed non self-referencing z-factorization of the m-bonacci words, i.e, no previous occurrence of z_i can overlap with itself.

Acknowledgements

Manon Stipulanti is supported by the FNRS Research Grant 1.B.397.20.

References

[1] P. Arnoux, G. Rauzy, “Représentation géometrique de suites de complexité $2n+1$”, *Bull. Soc. Math.* France 119 (2) (1991) pp. 199–215.

[2] G. Badkobeh, H. Bannai, K. Goto, Tomohiro I, C. S. Iliopoulos, S. Inenaga, S. J. Puglisi, S. Sugimoto, “Closed factorization”, *Disc. App. Math.*, 212 (2016) pp. 23–29.

[3] G. Badkobeh, G. Fici, Z. Lipták, “On the number of closed factors in a word”. *LATA 2015, 9th International Conference on Language and Automata Theory and Applications, LNCS*, 8977 (2015) pp. 381–390.

[4] J. Berstel, A. Savelli, “Crochemore factorization of Sturmian and other infinite words”. *In Proc. MFCSâĂŹ06, LNCS 4162, Springer*, 2006, pp. 157–166.

[5] M. Bucci, A. De Luca, G. Fici, “Enumeration and structure of trapezoidal words”, *Theoret. Comp. Sci.*, 468 (2013) pp. 12–22.

[6] M. Crochemore and W. Rytter, “Text algorithms”, *The Clarendon Press Oxford University Press*, 1994.

[7] A. De Luca, G. Fici, L. Q. Zamboni, “The sequence of open and closed prefixes of a Sturmian word”, *Adv. in App. Math.* 90 (2017), pp. 27–45

[8] X. Droubay, J. Justin, G. Pirillo, “Episturmian words and some constructions of de Luca and Rauzy”, *Theoret. Comp. Sci.*, 255 (2001) pp. 539–553.

[9] F. Durand, “A characterization of substitutive sequences using return words”, *Disc. Math.*, 179 (1998) pp. 89–101.

[10] G. Fici, “Factorizations of the Fibonacci infinite word”, *J. Int. Seq.* 18, no. 9, Article 15.9.3, (2015).

[11] G. Fici, “Open and closed words”, *âĂŇBull. of the europ. ass. for Theoret. Comp. Sci.*, 123, (2017), pp. 140–149.

[12] G. Fici, Z. Lipták, “Words with the smallest number of closed factors”, *In 14th Mons Days of Theoret. Comp. Sci.* (2012).

[13] N. Ghareghani, M. MohammadâĂŞNoori, P. Sharifani, “On z-factorization and c-factorization of standard episturmian words”, *Theoret. Comp. Sci.* 412 (2011), pp. 5232–5238.

[14] M. Jahannia, M. Mohammad-Noori, N. Rampersad, M. Stipulanti, “Palindromic Ziv-Lempel and Crochemore factorizations of m-Bonacci infinite words”, *Theoret. Comp. Sci.*, 790 (2019), pp. 16–40.

[15] J. Justin, L. Vuillon, “Episturmian word: shifts, morphisms and numeration systems”, *Inter. Journal of Found. of Comp. Sci.*, 15(2) (2004), pp. 329–348.

[16] J. Justin, L. Vuillon, “Return words in Sturmian and episturmian words”, *Theoret. Inf. and App.*, 34 (2000), pp. 343–356.

[17] J. Justin, G. Pirillo, “Episturmian words and episturmian morphisms”, *Theoret. Comp. Sci.*, 276 (2002), pp. 281–313.

[18] M. Lothaire, “Algebraic combinatorics on words”, *Encyclopedia of Mathematics and Its Applications, Cambridge University Press*, vol. 90, (2002).

[19] R. Kolpakov, G. Kucherov, “On maximal repetitions in words”, *J. Disc. Alg.* 1(1) (2000), pp. 159–186.

[20] A. Lempel, J. Ziv, “On the complexity of finite sequences”, *IEEE Transac. in Inf. Theo.* IT-22(1976), pp. 75–81.

[21] M. Morse, G. A.Hedlund, “Symbolic dynamics II. Sturmian trajectories”, *American J. Math.*, 62(1940), pp. 1–42.

[22] W. Rytter, “Compressed and fully compressed pattern-matching in one and two dimensions”, *Proc. IEEE88* (11) (2000), pp. 1769–1778.

[23] W. Rytter, “Application of LempelâĂŞZiv factorization to the approximation of grammar-based compression”. *Theoret. Comp. Sci.* 302 (2003), pp. 211–222.

[24] R. Siromoney, L. Mathew, D.R. Dare, K.G.Subramanian, “Infinite Lyndon words”, *Inform. Process. Lett.* 50(1994), pp. 101–104.

[25] B. Tan, Z-Y. Wen, “Some properties of the Tribonacci sequence”, *European J. Combin.*, 28(2007), pp. 1703–1719.

[26] Z-X. Wen,Z-Y. Wen, “Some properties of the singular words of the Fibonacci word”, *European J. Combin.*, Volume 15, Issue 6, (1994), pp. 587–598.

[27] J. Ziv, A. Lempel, “A universal algorithm for sequential data compression”, *IEEE Trans. Inform. Theory* 23(3) (1977), pp. 337–343.