UDC 511.32

ON DENSITY AND ERGODIC PROPERTIES OF THE INFINITE FIBONACCI WORD

Hamoud J., Abdullah D.
Moscow Institute of Physics and Technology (National Research University),
Dolgoprudny, Moscow region

Received 17.04.2025, revised 11.10.2025.

In this paper we explore combinatorial properties of Fibonacci words and their generalizations within the framework of combinatorics on words. The paper measures the diversity of subwords in Fibonacci words, showing non-decreasing growth for infinite sequences. We extend factor analysis to arithmetic progressions of symbols, highlighting generalized pattern distributions. Recent results link Sturmian sequences (including Fibonacci words) to unbounded binomial complexity and gap inequivalence, with implications for formal language theory and automata. In this work, the infinite word $\mathfrak{F} = \mathfrak{F}_b := ({}_bf_n)_{n\geqslant 0}$ is defined by concatenating non-negative base- $b \geqslant 2$ representation of the recursive $n!$.

Keywords: density, Fibonacci word, ergodic theory, sequence.

Vestnik TvGU. Seriya: Prikladnaya Matematika [Herald of Tver State University. Series: Applied Mathematics], 2025, № 4, 25–42, https://doi.org/10.26456/vtpmk742

1. Introduction

Throughout this paper, we examine a field of mathematics known as combinatorics on words, which studies the structures and properties of sequences of symbols — referred to as “words” — formed from a finite alphabet. This area, promoted by researchers such as [2, 7], is essential for understanding the complexity and patterns present in strings of symbols. It also interacts with disciplines such as formal language theory, automata theory, and number theory, among others.

In [16] (see also [14,23]), M. Lothaire considered an infinite word $a = y_0y_1 \dots y_n \dots$, where each block y_n belongs to the set $\{a, ab, abb\} = \delta(B)$. Since the substring bbb is forbidden, every occurrence of the letter a in the sequence can be followed by at most two b’s before the next a appears. The *Fibonacci sequence*, recursively defined by $F_0 = 0, F_1 = 1$ and $F_n = F_{n-1}+F_{n-2}$ for all $n \geq 3$, has been generalised in several ways. Consequently, there exists a corresponding special infinite word b over the alphabet

B such that $\delta(b) = a$. Separately, Grimm and Uwe in 2001 [18] introduced several important definitions concerning infinite square-free words and established bounds related to them.

The concept of *natural density* for a set of positive integers was introduced in [12] and is presented in Definition 1 for a set of positive integers.

Definition 1 (Natural Density [12])**.** *Let A be a set of positive integers. The natural density of A, denoted by $\delta(A)$, is*

$$\delta(A) := \lim_{x\to\infty} \frac{\#A(x)}{x}, \tag{1}$$

where $A(x) := A \cap [1, x]$ and $\#X$ is the number of elements in X.

A nonempty finite set Σ is called an alphabet. The elements of the set Σ are called letters. The alphabet consisting of b symbols from 0 to $b-1$ will then be denoted by $\Sigma_b = \{0, \dots, b-1\}$. A word $\mathbf{w}$ is a sequence of letters. The finite word $\mathbf{w}$ can be considered as a function of $\mathbf{w} : \{1, \cdots, |\mathbf{w}|\} \to \Sigma$, where $\mathbf{w}[i]$ is the letter in the i^{th} position. The length of the word $|\mathbf{w}|$ is the number of letters contained in it. The empty word is denoted by ε. Then we introduce infinite words as functions $\mathbf{w} : \mathbb{N} \to \Sigma$. The set of all finite words over Σ is denoted by Σ^*, and $\Sigma^+ = \Sigma^* \backslash \{\varepsilon\}$; the set of all infinite words is denoted by $\Sigma^{\mathbb{N}}$.

The concatenation of the finite words $\mathbf{U} = \mathbf{U}[1] \cdots \mathbf{U}[n], |\mathbf{U}| = n$ and $\mathbf{w} = \mathbf{w}[1] \cdots \mathbf{w}[m]$, $|\mathbf{w}| = m$ is the word

$$\mathbf{s} = \mathbf{U}\mathbf{w} = \mathbf{U} = \mathbf{U}[1] \cdots \mathbf{U}[n]\mathbf{w}[1] \cdots \mathbf{w}[m], \quad |\mathbf{s}| = |\mathbf{u}| + |\mathbf{w}| = n + m. \tag{2}$$

Let $\mathbf{U}$ and $\mathbf{w}$ be two words. If there are words $\mathbf{S}$ and $\mathbf{v}$ such that $\mathbf{w} = \mathbf{S}\mathbf{U}\mathbf{v}$, then the word $\mathbf{U}$ is called a factor of the word $\mathbf{w}$. The set of all factors of $\mathbf{w}$ is denoted by $\mathcal{L}_{\mathbf{w}}$ and it is called language generated by $\mathbf{W}$. If $\mathbf{s} = \varepsilon$, then $\mathbf{U}$ is called a prefix of the word $\mathbf{w}$, if $\mathbf{V} = \varepsilon$, it is named a suffix. Duaa and Meisami [6] emphasize that the factor $\mathbf{w}[i]\mathbf{w}[i+1] \cdots \mathbf{w}[j]$ where $i \leqslant j$ is denoted by $\mathbf{w}[i \cdots j]$.

Let $\mathcal{A} = \{a_0, \dots, a_m\}$ be a finite alphabet, whose elements are called in the following *digits*, and let $\mathcal{A}^*$ be the set of finite words over $\mathcal{A}$ as provided by Lai and Loreti [4]. A *substitution* is a function $\sigma : \mathcal{A}^* \to \mathcal{A}^*$ such that $\sigma(uv) = \sigma(u)\sigma(v)$ for all words $u, v \in \mathcal{A}^*$ and such that $\sigma(a)$ is not the empty word for every letter $a \in \mathcal{A}$. The domain $\mathcal{A}^*$ of a substitution σ can naturally be extended to the set of infinite sequences $\mathcal{A}^{\mathbb{N}}$ and to set of bi-infinite sequences $\mathcal{A}^{\mathbb{Z}}$ by concatenation. In particular, according to (2) if $\mathbf{w} = w_1 w_2 \cdots \in \mathcal{A}^{\mathbb{N}}$, then, $\sigma(\mathbf{w}) = \sigma(w_1)\sigma(w_2) \cdots$ and if $\mathbf{w} = \cdots w_{-1}.\, w_0 w_1 \cdots \in \mathcal{A}^{\mathbb{Z}}$ then $\sigma(\mathbf{w}) = \cdots \sigma(w_{-1}).\, \sigma(w_0)\sigma(w_1) \cdots$.

According to [2,4,6], in this paper we shall consider the alphabets $\mathcal{A}_m := \{1, \dots, m\}$ with $m \geq 2$. In 2013 Ramirez et al. [22] introduced k-Fibonacci word as an extension of the classical Fibonacci word, generalizing its properties to higher dimensions. These words were studied for their distinctive curves and patterns.

Example 1. The sequence of finite Fibonacci words is constructed using the substitution (or morphism) $\sigma : \{a, b\} \to \{a, b\}^*$ where $\sigma(a) = ab$ and $\sigma(b) = a$. Starting with $F_0 = b$ and $F_1 = a$, we generate the sequence by applying the rule $F_n = F_{n-1}F_{n-2}$ for $n \geq 2$.

The first few Fibonacci words are:

– $F_0 = b$

- $F_1 = a$
- $F_2 = F_1 F_0 = ab$
- $F_3 = F_2 F_1 = aba$
- $F_4 = F_3 F_2 = abaab$
- $F_5 = F_4 F_3 = abaababa$
- $F_6 = F_5 F_4 = abaababaabaab$

The infinite Fibonacci word, denoted $\mathcal{F}$, is the limit of this sequence. Let's analyze the density of the symbol 'a' in these finite prefixes. The density is the ratio of the number of occurrences of 'a' to the total length of the word.

- In F_4 = abaab: length is 5, number of 'a's is 3, density of 'a' $= 3/5 = 0.6$.
- In F_5 = abaababa: length is 8, number of 'a's is 5, density of 'a' $= 5/8 = 0.625$.
- In F_6 = abaababaabaab: length is 13, number of 'a's is 8, density of 'a' $= 8/13 \approx 0.615$.

As $n \to \infty$, the length of F_n is the Fibonacci number f_n, and the number of 'a's is f_{n-1}. The density of 'a's converges to:

$$\lim_{n\to\infty} \frac{|F_n|_a}{|F_n|} = \lim_{n\to\infty} \frac{f_{n-1}}{f_n} = \frac{1}{\varphi} = \varphi - 1 \approx 0.618,$$

where $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden ratio. This demonstrates how the density of symbols in the Fibonacci word is intrinsically linked to fundamental mathematical constants.

Recently, M. Rigo, et al. [15] presented the *Thue-Morse sequence*, which is the fixed point of the substitution $0 \to 01, 1 \to 10$, and has unbounded 1-gap k-binomial complexity for $k \geq 2$. Also, we note that for a Sturmian sequence and any integer $g \geq 1$, all sufficiently long factors of the sequence are pairwise g-gap k-binomially inequivalent for any $k \geq 2$.

This paper is organized as follows. In section 2 we introduce the basic concepts, definitions and properties necessary to understand the paper structure, section 3 deals with an introduction of the main topic of the research, and section 4 presents the main results of this paper.

2. Preliminaries

In this section we consider the Fibonacci word, which is generated by the substitution rules $0 \to 01$ and $1 \to 0$. This infinite sequence is a well-known object in combinatorics on words. Its density properties — especially the frequency of symbols — have been studied using various mathematical approaches, as formalized in Definition 6. Furthermore, studies [9–13] can be viewed to contribute to strengthening the ideas through this section.

The *period length* of the Fibonacci sequence modulo p is presented in Definition 2 and its use is shown in Example 2.

Definition 2. *The period length of the Fibonacci sequence modulo p, denoted $\pi(p)$, is the smallest integer $m \geq 1$ such that $F(n+m) \equiv F(n) \mod p$ for all $n \geq 0$. The restricted period length of the Fibonacci sequence modulo p, denoted $\alpha(p)$, is the smallest integer $m \geq 1$ such that $F(m) \equiv 0 \mod p$.*

Example 2. The Fibonacci sequence modulo 7 is

$$0, 1, 1, 2, 3, 5, 1, 6, 0, 6, 6, 5, 4, 2, 6, 1, 0, 1, 1, 2, 3, 5, 1, 6, \ldots$$

with period length 16.

In next Definition 3 we provide the definition of *Lucas numbers* due to their close connection with Fibonacci numbers, which is relevant to the analysis of density.

Definition 3 (Lucas numbers)**.** *The sequence of Lucas numbers $L(n)$ where $n \geq 0$ is defined as $L(0) = 2$, $L(1) = 1$, and $L(n+2) = L(n+1) + L(n)$ for $n \geq 0$.*

According to Definitions 2 and 3 we present the following Definition 4 to emphasize its properties.

Definition 4. *Let p be a prime number, and let $i \in \{0, 1, \ldots, \pi(p) - 1\}$. We say that i is a Lucas zero (with respect to p) if $L(i) \equiv 0 \mod p$ and a Lucas non-zero if $L(i) \not\equiv 0$* mod p.

Definition 5 (Fibonacci Sequence [24])**.** *Let F_k be the Fibonacci sequence given by $F_0 = 0, F_1 = 1, \ldots, F_{k+1} = F_k + F_{k-1}$ $(k \geq 1)$. Then we define the following function of its index:*

$$F_k = \frac{1}{\sqrt{5}} \left(\left(\frac{\sqrt{5}+1}{2} \right)^{n+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{n+1} \right), \tag{3}$$

where $F_n = \frac{1}{\sqrt{5}} \left(\varphi^{n+1} - \bar{\varphi}^{n+1} \right), \quad \varphi = \frac{\sqrt{5}+1}{2}, \quad \bar{\varphi} = 1 - \varphi.$

Our empirical analysis of the Fibonacci word f_{10} (a finite prefix of the infinite Fibonacci word) revealed a consistent pattern in the number of unique palindromic subwords for lengths less than thirty. Specifically, we observed that for every odd length L, there are exactly two unique palindromic subwords, and for every even length L, there is exactly one unique palindromic subword. This pattern is summarized in the table 1 below where we refer by SL to Subword Length and by NPS to Number of Palindromic Subwords.

Table 1: *Number of Palindromic Subwords*

SL	NPS	SL	NPS
1	2	11	2
2	1	12	1
3	2	13	2
4	1	14	1
5	2	15	2
6	1	16	1

SL	NPS	SL	NPS
7	2	17	2
8	1	18	1
9	2	19	2
10	1	20	1
21	2	26	1
22	1	27	2
23	2	28	1
24	1	29	2
25	2	30	1

Definition 6 (Density of Symbols). *[9] The Fibonacci word exhibits a precise asymptotic density of* 1*'s equal to* $\frac{1}{\phi^2}$*, where* $\phi = \frac{1+\sqrt{5}}{2}$ *is the golden ratio.*

This arises because the ratio of 0's to 1's converges to ϕ. Typically, this density is uniform according to Definition 8 [10]: for any position c and interval length m, the proportion of 1's in the substring ω_{c-m}^{c+m} converges uniformly to $\frac{1}{\phi^2}$ as $m \to \infty$. Paper [10] refers to any substring of length n, ensuring the density becomes increasingly homogeneous for large n. Then, we have:

1. **Critical Factorizations and Structural Properties** While not directly about symbol density, the Fibonacci word's structural regularity influences its density behavior. For example, Fibonacci words longer than five characters have **exactly one critical factorization** — position where the local period equals the global period [3]. This uniqueness contrasts with palindromes, which have at least two critical points (see [9,11]).
2. **Generalized Fibonacci Sequences and Density** Studies on generalized Fibonacci sequences (e.g., (r, a)-Fibonacci numbers) explore natural density in number-theoretic contexts. While these sequences have zero natural density in $\mathbb{Z}_{\geq 1}$, such results differ from symbol density in the Fibonacci word, highlighting distinct applications of "density" across mathematical domains (see [9,12]).

Definition 7. *For* $n \geq 0$*, the Fibonacci sequence* $F(n)_{n\geq 0}$ *is defined by the initial conditions* $F(0) = 0$ *and* $F(1) = 1$ *and the recurrence*

$$F(n+2) = F(n+1) + F(n). \tag{4}$$

It is well known that $F(n)_{n\geq 0}$ is periodic modulo m.

In next Definition 8, we establish from (1) the concept of *limiting density* of the Fibonacci sequence.

Definition 8 (Limiting density [10]). *Let* p *be a prime number. The limiting density of the Fibonacci sequence modulo powers of* p *is*

$$\operatorname{dens}(p) =: \lim_{\lambda\to\infty} \frac{|\{F(n) \bmod p^\lambda : n \geq 0\}|}{p^\lambda}. \tag{5}$$

Definition 9 (Arithmetic closure [6]). *Let* $\mathbf{w} = (a_n)_{n\geqslant 0} \in \Sigma^{\mathbb{N}}$*. The arithmetic closure of* $\mathbf{w}$ *is the set*

$$A(\mathbf{w}) = \{a_i a_{i+d} a_{i+2d} \cdots a_{i+kd} \mid d \geqslant 1, k \geqslant 0\}. \tag{6}$$

Actually, based on equations (4), (5), and (6), we introduce the following subsection to examine the topic of *Infinite Square-Free Words* in more detail.

2.1 Infinite Square-Free Words

Glen et al. state in [3] that any infinite word having exactly $n + 1$ distinct subwords of length n for every $n \in \mathbb{N}$ is a Sturmian word. In other words, an infinite word is Sturmian if and only if $p(w, n) = n + 1$, n $\geq$ 0. Note that this condition obviously implies $p(w, 1) = 2$, meaning the word uses exactly two different letters. The Fibonacci word sequence is generated as follows: $S(0) = a, S(1) = ab, S(2) = aba, S(3) = abaab, S(4) = abaababa, \ldots$

Figure 1 below shows representation of these subwords.

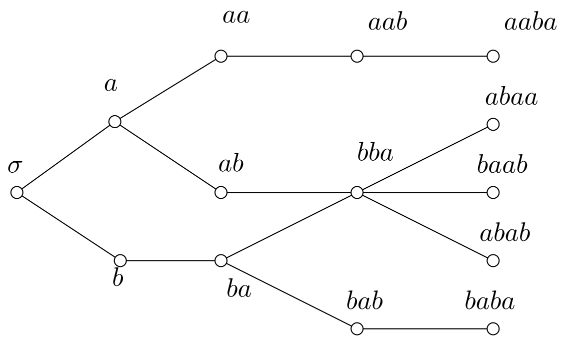

Fig. 1: *The subwords of the Fibonacci word.*

Proposition 1. *The infinite word of Thue-Morse has square factors.*

In fact, the only forecourt-free words over two letters a and b are:$\{a, b, ab, aba, bab\}$. On the opposite, there will be infinite square-free words over three letters, as we show that in Definition 10 and in Proposition 2.

Definition 10 (Morphism)**.** *Let $A = \{a, b\}$, $B = \{a, b, c\}$ be two alphabets. We define a morphism $\delta : B^* \to A^*$ by $\delta(c) = a, \delta(b) = ab, \delta(a) = abb$.*

Proposition 2. *For any infinite word $b = b_0b_1b_2 \ldots$ over the alphabet B, the image $\delta(b) = \delta(b_0)\delta(b_1) \ldots \delta(b_n)$ is a well-defined infinite word over A, and it begins with the letter a.*

Square-free words are defined in Definition 11.

Definition 11 (Square-Free Words)**.** *A word is called square-free if it contains no factor of the form xx (square), where x is a nonempty word.*

Definition 12 (Binary Square-Free Words)**.** *Binary square-free words are words over a binary alphabet. They are limited to short sequences.*

For example, the only square-free words over the alphabet a, b are a, b, ab, ba, aba, and bab. Any longer binary word necessarily contains a square.

Definition 13 (Ternary Square-Free Words [17, 19])**.** *Over a three-letter alphabet, square-free words can be infinitely long, and the number of such words grows exponentially with their length. The language of ternary square-free words is*

$$\mathcal{A} = \bigcup_{n \geq 0} \mathcal{A}(n) \subset \Sigma^{\mathbb{N}_0}.$$

Building on Definition 13, we now state Proposition 3, which provides bounds on the number of ternary square-free words.

Proposition 3. *[19–21] The number $s(n)$ of ternary square-free words of length n is bounded by*

$$6 \cdot 1.032^n \leq s(n) \leq 6 \cdot 1.379^n.$$

Brandenburg [19] noted in 1983 that for any $n \geq 3$, this upper bound can be systematically improved by calculating $a(n)$ for values of n that are as large as possible. Using the count $a(90) = 258, 615, 015, 792$, the upper bound becomes

$$s \leq 43\,102\,502\,632^{\frac{1}{88}} = 1.320\,829\,\ldots$$

The value $a(110)$ yields an improved upper bound of

$$s \leq 8\,416\,550\,317\,984^{\frac{1}{108}} = 1.317\,277\,\ldots$$

2.2 Sturmian Morphisms

A morphism is a function that maps words (finite sequences) to other words while preserving the structure of the sequences.

Definition 14 (Number of Occurrences)**.** *Let A be an alphabet and let $a \in A$ be a symbol. For a word $w \in A^*$ (where A^* denotes the set of all words over A), we denote by $|w|_a$ the number of occurrences of the symbol a in w. Similarly, $|w|_b$ denotes the number of occurrences of the symbol b in w.*

Sturmian morphisms specifically refer to the transformations that generate Sturmian sequences from simpler words.

Observation 1. A morphism $\psi = A^* \to A^*$ is Sturmian if the infinite word $\psi(x)$ is Sturmian for any Sturmian word x.

F. Mignosi et al. [1] define a morphism ψ to be *Sturmian* if and only if it belongs to the monoid generated by the morphisms E, φ, and $\tilde{\varphi}$ in any number and order; that is, $\psi \in \{E, \varphi, \tilde{\varphi}\}^*$. Furthermore, a morphism is called *standard* if, for every characteristic Sturmian word x, the image $\psi(x)$ is also a characteristic Sturmian word. Sturmian morphisms can be employed to generate infinite families of graphs. By applying these morphisms to base graphs, one can create complex structures that exhibit properties similar to those found in Sturmian sequences, such as balance and aperiodicity.

Definition 15. *Strings or sequences that contain subsets of characters that can create palindromic structures are referred to be scattered palindromic. Sporadic palindromic structures are more flexible than standard palindromes, which have tight symmetry requirements.*

For scattered palindromes, we can define them in terms of their indices. In Lemma 1 we present scattered palindromic conditions.

Lemma 1. *Let w be a word of order n. A sequence $s_1, s_2, \ldots, s_k$, $1 \leq k \leq n$ is scattered palindromic if there exists some mapping such that:*

$$s_i = s_j \quad \text{where } j = k + 1 - i \quad \text{for some } i, j.$$

These results are best consolidated in Theorem 1. We will, in turn, apply this theorem to obtain a refined method for calculating the density of the Fibonacci word.

Theorem 1. *Let w be a word over an alphabet A. Then every palindrome $v \in A^*$ is a subword of some word $x \in A^*$ if and only if the following holds: whenever $v = a_1 a_2 \ldots a_n, a_i \in A$ and $n \geq 0$, there exist letters $y_0, y_1, \ldots, y_n \in \mathrm{A}$ such that $\mathrm{x} = y_0 a_1 y_1 a_2 \ldots y_n a_n$. Moreover, the following inequality is satisfied:*

$$|p(w)| \leq |sp(w)|$$

for all $w \in \Sigma$ where $sp(w) = \sum_{t=1}^{|w|} sp_t(w)$ and t is the length of the word w.

3. Ergodic Theory

In this section, a sequence $(s_n)_{n \geqslant 0}$ of real numbers is said to be equidistributed or uniformly distributed on a non-degenerate interval $[a, b]$, if the proportion of terms that fall into a sub-interval is proportional to the length of this interval, i.e., if for any sub-interval $[c, d]$ of $[a, b]$ we have

$$\lim_{n \to \infty} \frac{\#\left(\{s_1, \ldots, s_n\} \cap [c, d]\right)}{n} = \frac{d - c}{b - a}. \tag{7}$$

The theory of uniform distribution modulo 1 deals with the distribution behavior of sequences of real numbers. The factor complexity of a finite or infinite word $\mathbf{W}$ is the function $k \mapsto \mathrm{p}_{\mathbf{w}}(k)$, which, for each integer k, gives the number $\mathrm{p}_{\mathbf{w}}(k)$ of distinct factors of length n in that word.

It is clear that the factor complexity is between zero and $(\#\Sigma)^k$. If $\mathrm{p}_{\mathbf{w}}(k) = (\#\Sigma)^k$, then the word $\mathbf{W}$ is said to have full factor complexity. This kind of words are called disjunctive words.

It is also easy to see that the factor complexity of any infinite word is a non-decreasing function, and the complexity of a finite word first increases, then decreases to zero.

Definition 16. *[6] The arithmetic complexity of an infinite word is the function that counts the number of words of a specific length composed of letters in arithmetic progression (and not only consecutive). In fact, it's a generalization of the complexity function.*

Proposition 4. *[4] For every $m \geq 2$ the substitution σ_2 is called Fibonacci substitution and the substitution σ_3 is called Tribonacci substitution.*

We use the symbol $|w|$ to denote the length of a finite word and we define the *weight* $|w|_i$ of a word w with respect to the i-th letter of $\mathcal{A}$, namely the number of occurrences of the letter a_i in the word w. According to (7) we present Definition 17 for bi-infinite word.

Definition 17 (Uniform frequency). *[4] Let* $\mathbf{w}$ *be an infinite (bi-infinite) word and let* $i = 1, \dots, m$. *If for all* $k \geq 0$ *(*$k \in \mathbb{Z}$*) the limit*

$$\lim_{n\to\infty} \frac{|w_{k+1}\cdots w_{k+n}|_j}{n} \tag{8}$$

exists uniformly with respect to k*, then it is called the uniform frequency* $f_j(\mathbf{w})$ *of the digit* a_j *in* $\mathbf{w}$*. Equivalently, if above limit exists, we can define*

$$f_j(\mathbf{w}) := \lim_{n\to\infty} \frac{\sup\{|w|_j \mid w \textit{ is a subword of length } n \textit{ of } \mathbf{w}\}}{n}. \tag{9}$$

The Perron-Frobenious eigenvalue of a substitution is defined as the largest eigenvalue of its adjacency matrix. We denote by ρ_m the Perron-Frobenious eigenvalue of σ_m. A substitution σ is a Pisot substitution if its Perron-Frobenious eigenvalue is a Pisot number, namely an algebraic integer greater than 1 whose conjugates are less than 1 in modulus.

Proposition 5. *[4] For all* $m \geq 2$, σ_m *is a Pisot substitution.*

In particular, the Perron-Frobenious eigevanlue ρ_m of σ_m is the Pisot number whose minimal polynomial is $\lambda^m - \lambda^{m-1} - \cdots - \lambda - 1$. Moreover the vector

$$d_m := (\rho_m^{-1}, \dots, \rho_m^{-m})$$

is a left eigenvector associated to ρ_m whose ℓ_0 norm is equal to 1.

Definition 18 (Equidistributed modulo 1). *A sequence* $(a_n)_{n\geqslant 0}$ *of real numbers is said to be equidistributed modulo 1 or uniformly distributed modulo 1 if the sequence of fractional parts of* $(a_n)_{n\geqslant 0}$ *is equidistributed in the interval* $[0, 1]$.

Actually, Theorem 2 states Weyl's criterion, a fundamental result in number theory and combinatorics. It is a key tool for analyzing distribution properties of sequences, especially for establishing equidistribution modulo 1.

Theorem 2 (Weyl's criterion [5]). *A sequence* $(a_n)_{n\geqslant 0}$ *is equidistributed modulo 1 if and only if for all non-zero integers* N,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \mathrm{e}^{2\pi \mathrm{i} N a_j} = 0. \tag{10}$$

Lemma 2. *[8] The fractional part of the sequence* $(\log(n!))_{n\geqslant 0}$ *is dense in* $[0, 1]$.

Proposition 6. *If* k *is any positive integer having* m *digits, there exists a positive integer* n *such that the first* m *digits of* $n!$ *constitute the integer* k.

Also we can state this result for any arbitrary base. Let $d_m \dots d_n$ be a word over Σ_b with $d_m \neq 0$. There exists n such that the base $_b$ expansion of $n!$ starts with $d_m \dots d_n$.

3.1 Statement of the problem

The infinite word $\mathfrak{F} = \mathfrak{F}_b := ({}_bf_n)_{n\geqslant 0}$ is defined by concatenating non-negative base-$b \geqslant 2$ representation of the recursive $n!$.

For example by concatenating base-10 representation of the recursive $n!$ we have:

$$\mathfrak{F} := (f_n)_{n\geqslant 0} = 11262412072050404032 0\cdots .$$

Problem 1. *What is the factor complexity of the* $\mathfrak{F}$*, i.e.* $\mathrm{p}_{\mathfrak{F}}(k)$*?*

Problem 2. *What is the arithmetic complexity, i.e.,* $\mathrm{a}_{\mathfrak{F}}(k)$*?*

Problem 3. *What is the density of arithmetic complexity, i.e.,* $\mathrm{a}_{\mathfrak{F}}(k)$*?*

4. Main Results

The concept of Fibonacci density, based on the fundamental principles of density, was first presented by one of the authors at the 65th Conference of the Moscow Institute of Physics and Technology in 2021. In Theorem 4 we state our main result. Since Sturmian words admit several equivalent characterizations and exhibit a wide range of combinatorial properties, their structure — in particular, their palindromic or return word properties — can be used to distinguish them. Moreover, according to Definition 9, both Proposition 7, Proposition 8, and Observation 2 follow directly as its immediate consequences.

Proposition 7. *The arithmetic complexity of* $\mathbf{w}$ *is the function* $a_{\mathbf{w}}$ *mapping* n *to the number* $\mathrm{a}_{\mathbf{w}}(n)$ *of words with length* n *in* $A(\mathbf{w})$*.*

Observation 2. If $\mathrm{a}_{\mathbf{w}}(k) = (\#\Sigma)^k$, then the word $\mathbf{w}$ is said to have full arithmetic complexity.

Proposition 8. *Let* r *denote the number of* 1 *'s. For any subword of length* n*, its density satisfies:*

$$|\frac{r\phi^2 - n}{n\phi^2}| \leq \frac{1}{n}. \tag{11}$$

Proposition 9. *Let* $\mathbf{W} \in \Sigma^{\mathbb{N}}$ *and* $\#\Sigma = k$*. Then for all* $n \in \mathbb{N}$*, according to Definition 9 we have*

$$1 \leqslant \mathrm{p}_{\mathbf{w}}(n) \leqslant \mathrm{a}_{\mathbf{w}}(n) \leqslant k^n.$$

Theorem 3. *[13] Let* $p \neq 2$ *be a prime number, and define* $e = \nu_p(F(p-\epsilon))$*. According to Definition 8 let define* $N(p) = |\{F(i) \bmod p^e\}|$ *where* i *is a Lucas non-zero, and let* $Z(p)$ *be the number of Lucas zeros* i *such that* $F(i) \not\equiv F(j) \mod p^e$ *for all Lucas non-zeros* j*. Then*

$$\mathrm{dens}(p) = \frac{N(p)}{p^e} + \frac{Z(p)}{2p^{2e-1}(p+1)}. \tag{12}$$

Proof. To clarify this, we denote the density by dens(p). Consider a prime $p \neq 2$, and let $e = \nu_p(F(p-\epsilon))$ is the coefficient, where $\nu_p(a)$ denotes the p-adic assessment of a, and F is a Lucas sequence evaluated at $p - \epsilon$ (with ϵ a small integer, here $\epsilon = 1$). Returning to Definition 8, we define $N(p) = |\{F(i) \bmod p^e\}|$, where i runs over the set of "Lucas non-zero" indices. If $F(i) \not\equiv F(j) \mod p^e$ for all Lucas non-zeros j, then in particular $F(i) \not\equiv 0 \pmod{p^e}$. Let $N(p)$ represents the cardinality of distinct residue classes modulo p^e and $Z(p)$ is the number of "Lucas zero" indices i.

For all Lucas positive number j, if $F(i) \equiv 0 \pmod{p^e}$ then $F(i) \not\equiv F(j) \pmod{p^e}$. Consequently, the density is given by

$$\text{dens}(p) = \frac{N(p)}{p^e} + \frac{Z(p)}{2p^{2e-1}(p+1)}.$$

As desire. □

Theorem 4. *Let $\Sigma_b = \{0, \cdots, b-1\}$ be the alphabet for base- b, then we have*

(i) $\mathfrak{F}b$ is disjunctive, meaning its factor complexity is full: $p\mathfrak{F}_b(k) = b^k$ for all $k \geq 1$.

(ii) The arithmetic complexity of $\mathfrak{F}b$ is also full: $a\mathfrak{F}_b(k) = b^k$ for all $k \geq 1$.

Proof. The infinite word $\mathfrak{F}_b$ is constructed by concatenating the base-b representations of the sequence $n!$ for $n = 0, 1, 2, \ldots$. We will prove each part of the theorem based on the properties of this specific construction.

(i) Proof of Full Factor Complexity:

A word is said to have full factor complexity if its set of factors of length k, denoted $\mathcal{L}_k(\mathfrak{F}_b)$, has size b^k for every $k \geq 1$. This is equivalent to saying the word is **disjunctive**, meaning every possible finite word over the alphabet Σ_b appears as a factor in $\mathfrak{F}_b$.

The fact that $\mathfrak{F}_b$ is disjunctive is a well-established result in combinatorics on words. It follows from a more general principle regarding sequences formed by concatenating function values. Specifically, for any given word $W \in \Sigma_b^*$ of length k, we need to show that W appears as a factor in $\mathfrak{F}_b$.

This property is a direct consequence of the result by H. G. Senge and E. G. Straus on the digits of numbers in different bases, and was later extended and popularized. A key result, as stated by authors like M. Dekking, states that for any word $W \in \Sigma_b^*$, there exists an integer n such that the base-b representation of $n!$ begins with the word W.

Let W be an arbitrary word of length k. Since there exists an n such that the base-b expansion of $n!$ starts with W, the word W is a prefix of the block corresponding to $n!$. As $\mathfrak{F}_b$ is the concatenation of all such blocks $(\ldots((n-1)!)_b(n!)_b\ldots)$, W is necessarily a factor of $\mathfrak{F}_b$. Since this holds for any arbitrary word $W \in \Sigma_b^*$, it means that the set of factors of $\mathfrak{F}_b$ is Σ_b^*. Therefore, for any length k, all b^k possible words of that length are present, and the factor complexity is full:

$$p_{\mathfrak{F}_b}(k) = b^k.$$

(ii) Proof of Full Arithmetic Complexity:

The arithmetic complexity, $a_{\mathbf{w}}(k)$, counts the number of factors of length k that can be extracted from a word $\mathbf{w}$ by taking letters at indices forming an arithmetic progression $i, i+d, i+2d, \ldots, i+(k-1)d$.

By definition, the set of factors (which corresponds to $d = 1$) is a subset of the set of arithmetically extracted words. This gives the fundamental inequality:

$$p_{\mathbf{w}}(k) \leq a_{\mathbf{w}}(k).$$

Additionally, the number of possible words of length k over an alphabet of size b is b^k, so we have:

$$p_{\mathbf{w}}(k) \leq a_{\mathbf{w}}(k) \leq b^k.$$

From part (i) of this proof, we have established that the factor complexity of $\mathfrak{F}_b$ is full, i.e., $p_{\mathfrak{F}_b}(k) = b^k$.

Substituting this result into the inequality, we get:

$$b^k \leq a_{\mathfrak{F}_b}(k) \leq b^k.$$

This inequality can only be satisfied if both sides are equal. Therefore, the arithmetic complexity must also be full:

$$a_{\mathfrak{F}_b}(k) = b^k.$$

This completes the proof. □

Actually, Lemma 3 concerns a semigroup $\mathbf{A}$ constructed within the infinite word $\mathfrak{F}_b$. Theorem 5 then gives a result about density that follows from the framework established in Theorem 4.

Lemma 3. *Let $(a_1, a_2, \ldots, a_n)$ be a sequence of words where $a_i \in \mathbf{A}$, among all the words $\mathbf{A} \in \Sigma_b^+$ that we formed in $\mathfrak{F}_b$, i.e., $\mathfrak{F}_b[i \cdots j] = \mathbf{A}$ for semi-group $\mathbf{A}$.*

Theorem 5 provides us with a practical application for calculating the density of a Fibonacci word through Fibonacci numbers, establishing a strong and close connection between these equivalent relations.

Theorem 5. *Let $\mathcal{F}$ be the infinite Fibonacci word generated by the substitution $\sigma(a) = ab$ and $\sigma(b) = a$. The asymptotic density of the letter 'a' in $\mathcal{F}$, denoted* dens(a), *exists and is given by:*

$$\operatorname{dens}(a) = \lim_{n\to\infty} \frac{|F_n|_a}{|F_n|} = \frac{1}{\varphi} = \varphi - 1$$

where F_n is the n-th finite Fibonacci word, $|F_n|$ is its length, $|F_n|_a$ is the number of occurrences of the letter 'a' in F_n, and $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden ratio.

Conclusion

The study establishes foundational results on the density and complexity properties of Fibonacci words, integrating combinatorial, algebraic, and dynamical perspectives. The Fibonacci word is confirmed as a *Sturmian word*, characterized by its minimal factor complexity $p(n) = n + 1$. This aligns with its generation via substitutions and hierarchical subword structure. The asymptotic density of a specific letter in the

Fibonacci word converges to $\varphi - 1 \approx 0.618$, where $\varphi = \frac{1+\sqrt{5}}{2}$. This arises from the recursive ratio

$$\lim_{n\to\infty} \frac{F(n)}{F(n+1)},$$

directly tied to the golden ratio. Both factor complexity and arithmetic complexity of the infinite Fibonacci word $\mathfrak{F}_b$ are proven to be full (b^k for base b), demonstrating that all possible subwords and arithmetic progressions appear infinitely often. For primes $p \neq 2$, the density $\text{dens}(p)$ is derived via counts of Fibonacci residues modulo p^e and Lucas zeros, combining combinatorial enumeration with modular arithmetic.

These results deepen the understanding of Fibonacci words as prototypical Sturmian systems, with applications in symbolic dynamics, number theory, and coding. The explicit density formulas and complexity proofs provide tools for analyzing pattern distributions in substitutive sequences.

Acknowledgements

The authors would like to express their sincere gratitude to the anonymous reviewers for their insightful comments and constructive suggestions, which have substantially enhanced the clarity and quality of this manuscript.

References

[1] Mignosi F., Seebold P., “Morphismes sturmiens et regles de Rauzy”, *Journal de Theorie des Nombres de Bordeaux*, **5**:2 (1993), 221–233.

[2] Rigo M., *Formal Languages, Automata and Numeration Systems 1: Introduction to Combinatorics on Words*, ISTE Ltd., London; John Wiley and Sons, Inc., Hoboken, NJ, 2014.

[3] Glen A., Wolff A., Clarke R., *On Sturmian and Episturmian Words, and Related Topics*, School of mathematical sciences discipline of pure mathematics, 2006.

[4] Lai A.C., Loreti P., *Trigonometric inequalities for Fibonacci chains*, 2025, https://doi.org/10.48550/arXiv.2503.11322.

[5] Kuipers L., Niderreiter H., *Uniform Distribution of Sequences*, A Wiley-Interscience publication, New York, 1974.

[6] Duaa A., Meisami M., “Factor and arithmetic complexity of concatenating the n!”, *Chebyshevskii Sbornik*, **24**:4 (2023), 341–344.

[7] Yong A., “Critique of Hirsch’s citation index: A combinatorial Fermi problem”, *Notices of the AMS*, **61**:9 (2014), 1040–1050, https://doi.org/10.48550/arXiv.1402.4357.

[8] Maxfield J.E., “A Note on n!”, *Mathematics Magazine*, **43** (1997), 64–67.

[9] Kirisci M, Karaisa A., “Fibonacci statistical convergence and Korovkin type approximation theorems”, *Journal of Inequalities and Applications*, 2017, 229, https://doi.org/10.1186/s13660-017-1503-z.

[10] *Is the density of 1s in the Fibonacci word uniform?*, MathOverflow, https://mathoverflow.net/questions/323614/is-the-density-of-1s-in-the-fibonacci-word-uniform.

[11] Harju T., Nowotka D., "Density of Critical Factorizations", *RAIRO - Theoretical Informatics and Applications*, **36**:3 (2002), 315–327, https://doi.org/10.1051/ita:2002016.

[12] Trojovsky P., "On the Natural Density of Sets Related to Generalized Fibonacci Numbers of Order r", *Axioms*, **10**:3 (2021), 144, https://doi.org/10.3390/axioms10030144.

[13] Bragman N., Rowland E., *Limiting density of the Fibonacci sequence modulo powers of p*, https://arxiv.org/abs/2202.00704, 2022.

[14] Bonizzoni P., Brattka V., Lowe B., "The Nature of Computation: Logic, Algorithms, Applications", *9th Conference on Computability in Europe, CiE 2013*, 2013.

[15] Rigo M., Stipulanti M., Whiteland M.A., "Gapped Binomial Complexities in Sequences", *IEEE International Symposium on Information Theory (ISIT)*, 2023, 1294–1299.

[16] Lothaire M., *Combinatorics on words*, Cambridge Mathematical Library, 2003.

[17] Sun X., "New lower bound on the number of ternary square-free words", *Journal of Integer Sequences*, **6**:3 (2003), 03.3.2.

[18] Grimm U., "Improved bounds on the number of ternary square-free words", *Journal of Integer Sequences*, **4**:2 (2001), 01.2.7.

[19] Grimm U., *Improved bounds on the number of ternary square-free words*, arXiv preprint math/0105245, 2001.

[20] Brinkhuis J., "Nonrepetitive sequences on three symbols", *Quarterly Journal of Mathematics*, **34** (1983), 145–149.

[21] Baake M., Elser V., Grimm U., "The entropy of square-free words", *Mathematical and Computer Modelling*, **26** (1997), 13–26.

[22] Ramirez J.L., "On the k-Fibonacci words", *Acta Universitatis Sapientiae, Informatica*, **5**:2 (2013), 212–226, https://doi.org/10.2478/ausi-2014-0011.

[23] Zolotov B., *Another Solution to the Thue Problem of Non-Repeating Words*, arXiv preprint arXiv:1505.00019, 2015.

[24] Zeckendorf E., "Representation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas", *Bulletin de la Societe Royale des Sciences de Liege*, **41** (1972), 179–182.

Citation

Hamoud J., Abdullah D., "On Density and Ergodic Properties of the Infinite Fibonacci Word", *Vestnik TvGU. Seriya: Prikladnaya Matematika [Herald of Tver State University. Series: Applied Mathematics]*, 2025, № 4, 25–42. https://doi.org/10.26456/vtpmk742

Authors Info

1. **Hamoud Jasem**
   PhD Student at Moscow Institute of Physics and Technology.

   Russia, 141701, Moscow region, Dolgoprudny, Institutskiy Pereulok, 9, MIPT.
   E-mail: hamoud.math@gmail.com

2. **Abdullah Duaa**
   PhD Student at Moscow Institute of Physics and Technology.

   Russia, 141701, Moscow region, Dolgoprudny, Institutskiy Pereulok, 9, MIPT.
   E-mail: duaa1992abdullah@gmail.com

О ПЛОТНОСТИ И ЭРГОДИЧЕСКИХ СВОЙСТВАХ БЕСКОНЕЧНОГО СЛОВА ФИБОНАЧЧИ

Хаммуд Д., Абдулла Д.
Московский физико-технический институт (национальный исследовательский университет), г. Долгопрудный, Московская обл.

Поступила в редакцию 17.04.2025, после переработки 11.10.2025.

В этой статье исследуются комбинаторные свойства слов Фибоначчи и их обобщения в рамках комбинаторики слов. В работе рассматриваются бесконечные последовательности, измеряется разнообразие подслов в словах Фибоначчи, демонстрируется неубывающий рост для бесконечных последовательностей. Мы расширяем факторный анализ на арифметические прогрессии символов, выделяя обобщенные распределения паттернов. Последние результаты связывают Штурмовские последовательности (включая слова Фибоначчи) с неограниченной биномиальной сложностью и неэквивалентностью промежутков, что имеет значение для теории формальных языков и автоматов. В данной работе бесконечное слово $\mathfrak{F} = \mathfrak{F}_b := ({}_b f_n)_{n \geqslant 0}$ определяется путем конкатенации неотрицательных базисных $b \geqslant 2$ представлений рекурсивного $n!$

Ключевые слова: плотность, слово Фибоначчи, эргодическая теория, последовательность.

Образец цитирования

Hamoud J., Abdullah D. On Density and Ergodic Properties of the Infinite Fibonacci Word // Вестник ТвГУ. Серия: Прикладная математика. 2025. №4. С. 25–42. https://doi.org/10.26456/vtpmk742

Список литературы

[1] Mignosi F., Seebold P. Morphismes sturmiens et regles de Rauzy // Journal de Theorie des Nombres de Bordeaux. 1993. Vol. 5, № 2. Pp. 221–233.

[2] Rigo M. Formal Languages, Automata and Numeration Systems 1: Introduction to Combinatorics on Words. London, Hoboken, NJ: ISTE Ltd., John Wiley and Sons, Inc., 2014.

[3] Glen A., Wolff A., Clarke R. On Sturmian and Episturmian Words, and Related Topics. School of mathematical sciences discipline of pure mathematics. 2006.

[4] Lai A.C., Loreti P. Trigonometric inequalities for Fibonacci chains. 2025. https://doi.org/10.48550/arXiv.2503.11322

[5] Kuipers L., Niderreiter H. Uniform Distribution of Sequences. New York: A Wiley-Interscience publication, 1974.

[6] Duaa A., Meisami M. Factor and arithmetic complexity of concatenating the n! // Chebyshevskii Sbornik. 2023. Vol. 24, № 4. Pp. 341–344.

[7] Yong A. Critique of Hirsch's citation index: A combinatorial Fermi problem // Notices of the AMS. 2014. Vol. 61, № 9. Pp. 1040–1050. https://doi.org/10.48550/arXiv.1402.4357

[8] Maxfield J.E. A Note on n! // Mathematics Magazine. 1997. Vol. 43. Pp. 64–67.

[9] Kirisci M, Karaisa A. Fibonacci statistical convergence and Korovkin type approximation theorems // Journal of Inequalities and Applications. 2017. ID 229. https://doi.org/10.1186/s13660-017-1503-z

[10] Is the density of 1s in the Fibonacci word uniform? [Electronic resource] // MathOverflow. URL: https://mathoverflow.net/questions/323614/is-the-density-of-1s-in-the-fibonacci-word-uniform.

[11] Harju T., Nowotka D. Density of Critical Factorizations // RAIRO - Theoretical Informatics and Applications. 2002. Vol. 36, № 3. Pp. 315–327. https://doi.org/10.1051/ita:2002016

[12] Trojovsky P. On the Natural Density of Sets Related to Generalized Fibonacci Numbers of Order r // Axioms. 2021. Vol. 10, № 3. ID 144. https://doi.org/10.3390/axioms10030144

[13] Bragman N., Rowland E. Limiting density of the Fibonacci sequence modulo powers of p. https://arxiv.org/abs/2202.00704. 2022.

[14] Bonizzoni P., Brattka V., Lowe B. The Nature of Computation: Logic, Algorithms, Applications // 9th Conference on Computability in Europe, CiE 2013. 2013.

[15] Rigo M., Stipulanti M., Whiteland M.A. Gapped Binomial Complexities in Sequences // IEEE International Symposium on Information Theory (ISIT). 2023. Pp. 1294–1299.

[16] Lothaire M. Combinatorics on words. Cambridge Mathematical Library, 2003.

[17] Sun X. New lower bound on the number of ternary square-free words // Journal of Integer Sequences. 2003. Vol. 6, № 3. ID 03.3.2.

[18] Grimm U. Improved bounds on the number of ternary square-free words // Journal of Integer Sequences. 2001. Vol. 4, № 2. ID 01.2.7.

[19] Grimm U. Improved bounds on the number of ternary square-free words. arXiv preprint math/0105245. 2001.

[20] Brinkhuis J. Nonrepetitive sequences on three symbols // Quarterly Journal of Mathematics. 1983. Vol. 34. Pp. 145–149.

[21] Baake M., Elser V., Grimm U. The entropy of square-free words // Mathematical and Computer Modelling. 1997. Vol. 26. Pp. 13–26.

[22] Ramirez J.L. On the k-Fibonacci words // Acta Universitatis Sapientiae, Informatica. 2013. Vol. 5, № 2. Pp. 212–226. https://doi.org/10.2478/ausi-2014-0011

[23] Zolotov B. Another Solution to the Thue Problem of Non-Repeating Words. arXiv preprint arXiv:1505.00019. 2015.

[24] Zeckendorf E. Representation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas // Bulletin de la Societe Royale des Sciences de Liege. 1972. Vol. 41. Pp. 179–182.

Сведения об авторах

1. **Хаммуд Джасем**
аспирант Московского физико-технического института.

Россия, 141700, Московская область, г. Долгопрудный, Институтский переулок, д. 9, МФТИ.
E-mail: *hamoud.math@gmail.com*

2. **Абдулла Дуа**
аспирант Московского физико-технического института.

Россия, 141700, Московская область, г. Долгопрудный, Институтский переулок, д. 9, МФТИ.
E-mail: *duaa1992abdullah@gmail.com*