Entropy **2013**, *15*, 3312-3324; doi:10.3390/e15093312

OPEN ACCESS

entropy

ISSN 1099-4300

www.mdpi.com/journal/entropy

Article

On the Entropy of a Two Step Random Fibonacci Substitution

Johan Nilsson

Department of Mathematics, Universität Bielefeld, Postfach 100131, D–33501 Bielefeld, Germany; E-Mail: jnilsson@math.uni-bielefeld.de; Tel.: +49-521-106-4778; Fax: +49-521-106-6481

Received: 8 July 2013; in revised form: 30 July 2013 / Accepted: 19 August 2013 / Published: 23 August 2013

Abstract: We consider a random generalization of the classical Fibonacci substitution. The substitution we consider is defined as the rule mapping, $\mathtt{a} \mapsto \mathtt{baa}$ and $\mathtt{b} \mapsto \mathtt{ab}$, with probability p, and $\mathtt{b} \mapsto \mathtt{ba}$, with probability $1-p$ for $0 < p < 1$, and where the random rule is applied each time it acts on a $\mathtt{b}$. We show that the topological entropy of this object is given by the growth rate of the set of inflated random Fibonacci words, and we exactly calculate its value.

Keywords: combinatorics on words; asymptotic enumeration; symbolic dynamics

---

1. Introduction

In [1], Godrèche and Luck define the random Fibonacci chain by the generalized substitution:

$$\theta : \begin{cases} \mathtt{a} \mapsto \mathtt{b} \\ \mathtt{b} \mapsto \begin{cases} \mathtt{ab} & \text{with probability } p \\ \mathtt{ba} & \text{with probability } 1-p \end{cases} \end{cases}$$

for $0 < p < 1$ and where the random rule is applied each time θ acts on a $\mathtt{b}$. They introduce the random Fibonacci chain when studying quasi-crystalline structures and tilings in the plane. In their paper, it is claimed (without proof) that the topological entropy of the random Fibonacci chain is given by the growth rate of the set of inflated random Fibonacci words. This was later, with a combinatorial argument, proven in a more general context in [2].

The renewed interest in this system, and in possible generalizations, stems from the observation that the natural geometric generalization of the symbolic sequences by tilings of the line had to be Meyer sets with entropy and interesting spectra [3]. There is now a fair understanding of systems that emerge from the local mixture of inflation rules that each define the same hull. However, little is known so far about

more general mixtures. Here, we place our attention to one such generalization. It is still derived from the Fibonacci rule, but mixes inflations that define distinct hulls.

In this paper, we consider the randomized substitution, ϕ, defined by:

$$\phi = \begin{cases} \mathtt{a} \mapsto \mathtt{baa} \\ \mathtt{b} \mapsto \begin{cases} \mathtt{ab} & \text{with probability } p \\ \mathtt{ba} & \text{with probability } 1-p \end{cases} \end{cases}$$

for $0 < p < 1$ and where the random rule is applied each time ϕ acts on a $\mathtt{b}$. The substitution, ϕ, is a mixture of two substitutions, whose hulls are different. This is true, since the hull of the substitution, $(\mathtt{a}, \mathtt{b}) \mapsto (\mathtt{baa}, \mathtt{ab})$, contains words with the sub-words, $\mathtt{aaa}$ and $\mathtt{bb}$, but neither of these sub-words are to be found in any word of the hull of $(\mathtt{a}, \mathtt{b}) \mapsto (\mathtt{baa}, \mathtt{ba})$. For a more detailed survey of the differences and similarities of the generated hulls of these two substitutions, see [4].

Before we can state our main theorem in detail, we need to introduce some notation. A word, w, over an alphabet, Σ, is a finite sequence, $w_1 w_2 \dots w_n$, of symbols from Σ. We let, here, $\Sigma = \{\mathtt{a}, \mathtt{b}\}$. We denote a sub-word of w by $w[a, b] = w_a w_{a+1} w_{a+2} \dots w_{b-1} w_b$, and similarly, we let $W[a, b] = \{w[a, b] : w \in W\}$. By $|\cdot|$, we mean the length of a word and the cardinality of a set. Note that $|w[a, b]| = b - a + 1$. When indexing the brackets with a letter, α, from the alphabet, $|\cdot|_\alpha$, we shall mean the numbers of occurrences of α in the enclosed word.

For two words, $u = u_1 u_2 u_3 \dots u_n$ and $v = v_1 v_2 v_3 \dots v_m$, we denote by uv the concatenation of the two words, that is, $uv = u_1 u_2 u_3 \dots u_n v_1 v_2 \dots v_m$. Similarly, we let, for two sets of words, U and V, their product be the set, $UV = \{uv : u \in U, v \in V\}$, containing all possible concatenations.

Letting ϕ act on the word, $\mathtt{a}$, repeatedly yields an infinite sequence of words, $r_n = \phi^{n-1}(\mathtt{a})$. We know that $r_1 = \mathtt{a}$ and $r_2 = \mathtt{baa}$. However, r_3 is one of the words, $\mathtt{abbaabaa}$ or $\mathtt{babaabaa}$, with probability p or $1 - p$. The sequence, $\{r_n\}_{n=1}^{\infty}$, converges in distribution to an infinite random word, r. We say that r_n is an inflated word (under ϕ) in generation n, and we introduce, here, sets that correspond to all inflated words in generation n;

Definition 1. *Let* $A_1 = \{\mathtt{a}\}$, $B_1 = \{\mathtt{b}\}$, *and for* $n \geq 2$, *we define recursively:*

$$A_n = B_{n-1} A_{n-1} A_{n-1}$$
$$B_n = A_{n-1} B_{n-1} \cup B_{n-1} A_{n-1}$$

and we let $A := \lim_{n\to\infty} A_n$ *and* $B := \lim_{n\to\infty} B_n$.

The sets, A and B, are indeed well defined. This is a direct consequence of Corollary 6. It is clear from the definition of A_n and B_n that all their elements have the same length, that is, for all $x, y \in A_n$ (or $x, y \in B_n$), we have $|x| = |y|$. By induction, it easily follows that for $a \in A_n$, we have $|a| = f_{2n}$ and for $b \in B_n$, we have $|b| = f_{2n-1}$, where f_m is the mth Fibonacci number, defined by $f_{n+1} = f_n + f_{n-1}$ with $f_0 = 0$ and $f_1 = 1$.

For a word, w, we say that x is a sub-word of w if there are two words, u, v, such that $w = uxv$. The sub-word set, $F(S, n)$, is the set of all sub-words of length n of words in S. The *combinatorial entropy* of the random Fibonacci chain is defined as the limit, $\lim_{n\to\infty} \frac{1}{n} \log |F(A, n)|$. The combinatorial entropy

is known to equal the topological entropy for our type of systems; see [5]. The existence of this limit is direct by Fekete's lemma [6], since we have sub-additivity, $\log|F(S,n+m)| \leq \log|F(S,n)| + \log|F(S,m)|$. We can now state the main result in this paper.

Theorem 2. *The logarithm of the growth rate of the size of the set of inflated random Fibonacci words equals the topological entropy of the random Fibonacci chain, that is:*

$$\lim_{n\to\infty}\frac{\log|A_n|}{f_{2n}} = \lim_{n\to\infty}\frac{\log|B_n|}{f_{2n-1}} = \lim_{n\to\infty}\frac{\log|F(C,n)|}{n} = \frac{1}{\tau^3}\log 2 \tag{1}$$

where τ is the golden mean, $\tau = \frac{1+\sqrt{5}}{2}$ and $C \in \{A,B\}$.

The outline of the paper is that we start by studying the sets, A_n and B_n. Next, we give a finite method for finding the sub-word set, $F(A,n)$, (which, we will see, is the same as $F(B,n)$). Thereafter, we derive some Diophantine properties of the Fibonacci number that will play a central part when we look at the distribution of the letters in words from $F(A,n)$. Finally, we present an estimate of $|F(A,n)|$, leading up to the proof of Theorem 2.

2. Inflated Words

In this section, we present the sets of inflated words and give an insight to their structure. The results presented here will also play an important role for the results in the coming sections.

Proposition 3. *Let $u,v \in A_n$ (or both in B_n). Then, $u \neq v$, if and only if $\{\phi(u)\} \cap \{\phi(v)\} = \emptyset$, where, here, $\{\phi(z)\}$ denotes the set of all possible words that can be obtained by applying ϕ on z.*

Proof. Let $u \neq v$, and assume that $w \in \{\phi(u)\} \cap \{\phi(v)\}$. Denote by ϕ_u and ϕ_v the special choices of ϕ, such that $w = \phi_u(u) = \phi_v(v)$. Let k be the first position, such that $u_k \neq v_k$, where $u = u_1u_2\dots u_m$ and $v = v_1v_2\dots v_m$. Then, we may assume $u_k = \mathtt{a}$ and $v_k = \mathtt{b}$; otherwise, just swap the names of u and v. Since we have $\phi(\mathtt{a}) = \mathtt{baa}$, we see that we must have $\phi_v(v_k) = \phi_v(\mathtt{b}) = \mathtt{ba}$. However, then, also, $\phi_v(v_kv_{k+1}) = \phi_v(\mathtt{bb}) = \mathtt{baab}$. This then implies $u_{k+1} = \mathtt{b}$, since, if we have $u_{k+1} = \mathtt{a}$, then there must be two consecutive $\mathtt{a}$s in w, and we could not find a continuation in v. Hence, we have $\phi_u(u_ku_{k+1}) = \phi_u(\mathtt{ab}) = \mathtt{baaba}$. As previously, v must continue with a $\mathtt{b}$. We now see that we are in a cycle, where $|\phi_u(u_ku_{k+1}\dots u_{k+s})| = 3+2s$ and $|\phi_v(v_kv_{k+1}\dots v_{k+s})| = 2(s+1)$. Since there is no $s \in \mathbb{N}$, such that $3+2s = 2(s+1)$, we conclude that there can be no such w. □

We can now turn to the question of counting the elements in the sets, A_n and B_n.

Proposition 4. *For $n \geq 2$, we have:*

$$|A_n| = 2^{f_{2n-3}-1} \quad \textit{and} \quad |B_n| = 2^{f_{2n-4}+1}$$

Proof. Let us start with the proof of the the size of A_n. From the Definition 1 of A_n and B_n, it follows by induction that $|x|_{\mathtt{b}} = f_{2n-2}$ for $x \in A_n$. Combining this with Proposition 3, we find the recursion:

$$|A_n| = |A_{n-1}| \cdot 2^{|x|_{\mathtt{b}}} = |A_{n-1}| \cdot 2^{f_{2n-4}} \tag{2}$$

The size of A_n now follows from Equation (2) by induction. For the size of B_n, we have, by the definition of B_n and that we already know the size of A_n,

$$|B_n| = \frac{|A_{n+1}|}{|A_n||A_n|} = \frac{2^{f_{2n-1}-1}}{2^{f_{2n-3}-1} \cdot 2^{f_{2n-3}-1}} = 2^{f_{2n-4}+1}$$

which completes the proof. □

From Proposition 4, the statements of the logarithmic limits of the sets, A_n and B_n, in Theorem 2 follows directly. Our next step is to give some result on sets of prefixes of A_n and B_n. These results will play a central role when we later look at sets of sub-words.

Proposition 5. *For $n \geq 2$, we have:*

$$A_n[1, f_{2n} - 1] \subset A_{n+1}[1, f_{2n} - 1] \tag{3}$$

$$A_n[1, f_{2n} - 1] \subset \big(B_n A_n\big)[1, f_{2n} - 1] \tag{4}$$

Proof. Let us first consider (3). We give a proof by induction on n. For the basis case, $n = 2$, we have:

$$A_2[1, f_{2\cdot 2} - 1] = A_2[1, 2] = \{\mathtt{ab}\} \subset \{\mathtt{ab}, \mathtt{ba}\} = A_3[1, f_4 - 1]$$

Now, assume for induction that Equation (3) holds for $2 \leq n \leq p$. Then, for $n = p + 1$, we have by the induction assumption:

$$\begin{aligned}
A_{p+1}[1, f_{2(p+1)} - 1] &= \big(B_p A_p A_p\big)[1, f_{2(p+1)} - 1] \\
&\subseteq \big((A_p B_p \cup B_p A_p) A_p\big)[1, f_{2(p+1)} - 1] \\
&= \big(B_{p+1} A_p\big)[1, f_{2(p+1)} - 1] \\
&= B_{p+1}\big(A_p[1, f_{2p} - 1]\big) \\
&\subset B_{p+1}\big(A_{p+1}[1, f_{2p} - 1]\big) \\
&= \big(B_{p+1} A_{p+1}\big)[1, f_{2(p+1)} - 1] \\
&= \big(B_{p+1} A_{p+1} A_{p+1}\big)[1, f_{2(p+1)} - 1] \\
&= A_{p+2}[1, f_{2(p+1)} - 1]
\end{aligned}$$

which completes the induction and the proof of Equation (3). Let us turn to the proof of Equation (4). By the help of Equation (3), we have:

$$\begin{aligned}
A_n[1, f_{2n} - 1] &= \big(B_{n-1} A_{n-1} A_{n-1}\big)[1, f_{2n} - 1] \\
&= B_{n-1} A_{n-1}\big(A_{n-1}[1, f_{2(n-1)} - 1]\big) \\
&\subset B_{n-1} A_{n-1}\big(A_n[1, f_{2(n-1)} - 1]\big) \\
&= \big(B_{n-1} A_{n-1} A_n\big)[1, f_{2n} - 1] \\
&\subseteq \big(B_n A_n\big)[1, f_{2n} - 1]
\end{aligned}$$

which concludes the proof. □

From Proposition 5, it is straight forward, by recalling the recursive definition of A_n and B_n, to derive the following equalities on prefix-sets.

Corollary 6. *For $n \geq 3$, we have:*

$$\begin{aligned} A_n[1, f_{2(n-1)} - 1] &= A_{n+1}[1, f_{2(n-1)} - 1] \\ B_n[1, f_{2(n-1)} - 1] &= A_n[1, f_{2(n-1)} - 1] \\ B_n &= B_{n+1}[1, f_{2n-1}] \end{aligned}$$

We end the section by proving a result on suffixes of the sets, A_n and B_n, that we shall make use of in the next sections.

Proposition 7. *For $n \geq 2$, we have:*

$$A_n[f_{2n-2} + 2, f_{2n}] \subseteq B_n[2, f_{2n-1}] \tag{5}$$

$$B_n[2, f_{2n-1}] = B_{n+1}[f_{2n} + 2, f_{2n+1}] \tag{6}$$

Proof. We give a proof by induction on n. For the basis case, $n = 2$, we have:

$$A_2[f_2 + 2, f_4] = A_2[2, 3] = \{\mathtt{a}\} \subseteq \{\mathtt{a}, \mathtt{b}\} = B_2[2, 2]$$

Now, assume for induction that Equation (5) holds for $2 \leq n \leq p$. Then, for the induction step, $n = p+1$, we have by the induction assumption:

$$\begin{aligned} A_{p+1}[f_{2(p+1)-2} + 2, f_{2(p+1)}] &= \big(B_p A_p A_p\big)[f_{2(p+1)-2} + 2, f_{2(p+1)}] \\ &= \big(A_p A_p\big)[f_{2p-2} + 2, 2f_{2p}] \\ &= \big(A_p[f_{2p-2} + 2, f_{2p}]\big) A_p \\ &\subseteq \big(B_p[2, f_{2p-1}]\big) A_p \\ &= \big(B_p A_p[2, f_{2p+1}]\big) \\ &\subseteq B_{p+1}[2, f_{2(p+1)-1}] \end{aligned}$$

which completes the induction and the proof of Equation (5). For the proof of Equation (6), we have:

$$B_n[2, f_{2n-1}] = \big(A_n B_n\big)[f_{2n} + 2, f_{2n} + f_{2n-1}] \subseteq B_{n+1}[f_{2n} + 2, f_{2n} + f_{2n-1}]$$

and for the converse inclusion, we have by Equation (5):

$$\begin{aligned} B_{n+1}[f_{2n} + 2, f_{2n} + f_{2n-1}] &= \big(A_n B_n \cup B_n A_n\big)[f_{2n}2, f_{2n} + f_{2n-1}] \\ &= \big(B_n[2, f_{2n-1}]\big) \cup \big(A_n[f_{2n-2} + 2, f_{2n}]\big) \\ &\subseteq B_n[2, f_{2n-1}] \end{aligned}$$

which proves the equality (6). □

3. Sets of Sub-Words

Here, we investigate properties of the sets of sub-words, $F(A, m)$ and $F(B, m)$. We will prove that they coincide, and moreover, we show how to find them by considering finite sets, which will be central when estimating their size, depending on m.

First, we turn our attention to proving that it is indifferent if we consider sub-words of A_n or of B_n.

Proposition 8. *For $n \geq 1$, we have:*

$$F(A_{n+1}, f_{2n} - 1) = F(B_{n+1}, f_{2n} - 1)$$

Proof. Let us first turn to the proof of the inclusion:

$$F(A_{n+1}, f_{2n} - 1) \subseteq F(B_{n+1}, f_{2n} - 1) \tag{7}$$

Let $x_{(k)} \in A_{n+1}[k, k - 1 + f_{2n} - 1]$ for $1 \leq k \leq f_{2n+1} + 2$. It is clear that $x_{(k)} \in F(A_{n+1}, f_{2n} - 1)$ for any k. We have to prove that also $x_{(k)} \in F(B_{n+1}, f_{2n} - 1)$.

For $1 \leq k \leq f_{2n-1} + 2$, we have:

$$x_{(k)} \in F(B_nA_n, f_{2n} - 1) \subseteq F(B_{n+1}, f_{2n} - 1)$$

For $f_{2n-1} + 3 \leq k \leq f_{2n} + 1$, we have by Corollary 6, which $x_{(k)}$ must be a sub-word of:

$$\begin{aligned}\big(A_nA_n\big)[3, f_{2n} + f_{2n-2} - 1] &= \big(A_nB_n\big)[3, f_{2n} + f_{2n-2} - 1] \\ &= B_{k+1}[3, f_{2n} + f_{2n-2} - 1]\end{aligned}$$

For $f_{2n} + 2 \leq k \leq f_{2n+1} + 2$, we have by Proposition 7:

$$\begin{aligned}\big(B_nA_nA_n\big)[f_{2n} + 2, f_{2n+2}] &= \big(A_n[f_{2n-2} + 2, f_{2n}]\big)A_n \\ &\subseteq \big(B_n[2, f_{2n-1}]\big)A_n \\ &\subseteq B_{n+1}[2, f_{2n+1}]\end{aligned}$$

which concludes the proof of the inclusion (7). For the converse inclusion, it is enough to consider sub-words of A_nB_n, since any sub-word of B_nA_n clearly is a sub-word of A_{n+1}. Therefore, let $y_{(k)} \in (A_nB_n)[k, k - 1 + f_{2n} - 1]$ for $1 \leq k \leq f_{2n-1} + 1$. We now proceed as in the case above.

For $1 \leq k \leq f_{2n-2} + 1$, we have:

$$\begin{aligned}(A_nB_n)[1, f_{2n} + f_{2n-2} - 1] &= A_n\big(B_n[1, f_{2n-2} - 1]\big) \\ &= A_n\big(A_n[1, f_{2n-2} - 1]\big) \\ &= A_{n+1}[f_{2n+1} + 1, f_{2n+1} + f_{2n-2} - 1]\end{aligned}$$

For $f_{2n-2} + 2 \leq k \leq f_{2n-1} + 2$, we have:

$$\begin{aligned}(A_nB_n)[f_{2n-2} + 2, f_{2n-1} + 2] &= \big(A_n[f_{2n-2} + 2, f_{2n}]\big)A_n \\ &= \big(B_n[2, f_{2n-1}]\big)A_n \\ &= A_{n+1}[2, f_{2n+1}]\end{aligned}$$

which completes the proof. □

The above result shows that the set of sub-words from A_n and B_n coincide if the sub-words are not chosen too long. If we consider the limit sets, A and B, their sets of sub-words turn out to be the same. We have the following:

Proposition 9. *For $m \geq 1$, we have $F(A, m) = F(B, m)$.*

Proof. Let $x \in F(A, m)$. Then, there is an n, such that:

$$x \in F(A_n, m) \subseteq F(A_nB_n \cup B_nA_n, m) = F(B_{n+1}, m) \subseteq F(B, m)$$

Similarly, if $x \in F(B, m)$. Then, there is an n, such that:

$$x \in F(B_n, m) \subseteq F(B_nA_nA_n, m) = F(A_{n+1}, m) \subseteq F(A, m)$$

which completes the proof. □

The direct consequence of Proposition 9 is that we find the topological entropy in Equation (1) independent if we look at sub-words from A or B.

Now, let us turn to the question of finding $F(A, m)$ from a finite set, A_n, and not having to consider the infinite set, A.

Proposition 10. *For $n \geq 2$, we have:*

$$F(A_{n+1}, f_{2n} - f_{2n-3}) = F(A_{n+2}, f_{2n} - f_{2n-3})$$

Proof. It is clear that $F(A_{n+1}, f_{2n} - f_{2n-3}) \subseteq F(A_{n+2}, f_{2n} - f_{2n-3})$ holds for all $n \geq 2$. For the reverse inclusion, assume that $x \in F(A_{n+2}, f_{2n} - f_{2n-3})$. Note that we can write A_{n+1} and A_{n+2} on the form:

$$\begin{aligned} A_{n+1} &= B_nA_nA_n, \\ A_{n+2} &= B_nA_nB_nA_nA_nB_nA_nA_n \cup A_nB_nB_nA_nA_nB_nA_nA_n \end{aligned} \tag{8}$$

From Equation (8), we see that any x is a sub-word of any element in some of the seven sets:

$$\begin{aligned} &A_nA_n, \quad B_nA_n, \quad A_nB_n, \quad A_nB_nA_n, \\ &B_nB_n, \quad A_nB_nB_n, \quad B_nB_nA_n \end{aligned} \tag{9}$$

in such a way that the first letter in x is in the first factor (that is, A_n or B_n) of the sets. If x is a sub-word of A_nA_n or B_nA_n or completely contained in A_n, it is clear that we have $x \in F(A_{n+1}, f_{2n} - f_{2n-3})$. For the case when x is a sub-word of A_nB_n, it follows from Proposition 8 that we have $x \in F(A_{n+1}, f_{2n} - f_{2n-3})$.

If x is a sub-word of a word in $A_nB_nA_n$, such that x begins in the first A_n factor and ends in the second, then we have that x is a sub-word of a word in the set:

$$\begin{aligned} &\big(A_n[f_{2n-3} + f_{2n-1} + 2, f_{2n}]\big)B_{n-1}A_{n-1}\big(A_n[1, f_{2n-4} - 1]\big) \\ &\quad = \big(A_n[f_{2n-3} + f_{2n-1} + 2, f_{2n}]\big)B_{n-1}A_{n-1}\big(A_{n-1}[1, f_{2n-4} - 1]\big) \\ &\quad = \big(A_n[f_{2n-3} + f_{2n-1} + 2, f_{2n}]\big)\big(A_n[1, f_{2n-1} + f_{2n-4} - 1]\big) \\ &\quad = \big(A_nA_n\big)[f_{2n-3} + f_{2n-1} + 2, f_{2n} + f_{2n-1} + f_{2n-4} - 1] \end{aligned}$$

and we see that we have $x \in F(A_{n+1}, f_{2n} - f_{2n-3})$.

If x is a sub-word of a word in B_nB_n, let us first consider the case when it is a sub-word of $B_nB_{n-1}A_{n-1}$. Then, it follows that:

$$B_nB_{n-1}A_{n-1} \subseteq B_n\big(A_n[1, f_{2n-1}]\big) = \big(B_nA_n\big)[1, 2f_{2n-1}]$$

so x is a sub-word of a word in A_{n+1}. For the the second case, $B_nA_{n-1}B_{n-1}$, we have:

$$\begin{aligned} B_nA_{n-1}B_{n-1} &= A_{n-1}\big(B_{n-1}A_{n-1}\big)B_{n-1} \cup \big(B_{n-1}A_{n-1}A_{n-1}\big)B_{n-1} \\ &\subseteq \big(A_nB_nA_n\big)[f_{2n-1} + 1, 3f_{2n-1}] \cup \big(A_nB_n\big)[1, 2f_{2n-1}] \end{aligned}$$

and again, x is a sub-word of a word in A_{n+1}, by what we just proved above.

If x is a sub-word of a word in $A_nB_nB_n$, we have by Corollary 6:

$$\begin{aligned} &\big(A_nB_nB_n\big)[f_{2n-1} + f_{2n-3} + 1, 2f_{2n} - f_{2n-3} - 1] \\ &\quad = \big(A_n[f_{2n-1} + f_{2n-3} + 1, f_{2n}]\big)B_n\big(B_n[1, f_{2n-4} - 1]\big) \\ &\quad = \big(A_n[f_{2n-1} + f_{2n-3} + 1, f_{2n}]\big)B_n\big(A_n[1, f_{2n-4} - 1]\big) \end{aligned}$$

which shows that x is a sub-word of a word in A_{n+1} by what we previously have shown.

Finally, if x is a sub-word of a word in $B_nB_nA_n$, we first consider the case when x is a sub-word of a word in $B_nB_{n-1}A_{n-1}A_n$. By Corollary 6, we have:

$$\begin{aligned} &\big(B_nB_nA_n\big)[2f_{2n-3} + 1, f_{2n+1} - f_{2n-3} - 1] \\ &\quad = \big(B_n[2f_{2n-3} + 1, f_{2n-1}]\big)B_{n-1}A_{n-1}\big(A_n[1, f_{2n-4} - 1]\big) \\ &\quad = \big(B_n[2f_{2n-3} + 1, f_{2n-1}]\big)B_{n-1}A_{n-1}\big(A_{n-1}[1, f_{2n-4} - 1]\big) \\ &\quad = \big(B_n[2f_{2n-3} + 1, f_{2n-1}]\big)\big(A_n[1, f_{2n-1} + f_{2n-4} - 1]\big) \end{aligned}$$

which, by the help of the previous case, shows that x is a sub-word of a word in A_{n+1}. For the last case, $B_nA_{n-1}B_{n-1}A_n$, we have by Corollary 6 and Proposition 7:

$$\begin{aligned} &\big(B_nB_nA_n\big)[2f_{2n-3} + 1, f_{2n+1} - f_{2n-3} - 1] \\ &\quad = \big(B_n[2f_{2n-3} + 1, f_{2n-1}]\big)A_{n-1}B_{n-1}\big(A_n[1, f_{2n-4} - 1]\big) \\ &\quad = \big(B_{n-1}[2f_{2n-3} - f_{2n-2} + 1, f_{2n-3}]\big)A_{n-1}B_{n-1}\big(A_{n-1}[1, f_{2n-4} - 1]\big) \\ &\quad = \big(B_n[2f_{2n-3} - f_{2n-2} + 1, f_{2n-1}]\big)\big(A_n[1, f_{2n-2} - 1]\big) \end{aligned}$$

and again, we see that x is a sub-word of a word in A_{n+1} by what we have proven above. □

The result of Proposition 10 can be extended to hold for sub-words from elements A_n and A_{n+k}, where $k \geq 1$. A straight forward argument via induction gives:

$$F(A_{n+1}, f_{2n} - f_{2n-3}) = F(A_{n+k}, f_{2n} - f_{2n-3}) \tag{10}$$

for $k \geq 1$. By combining Proposition 10 and Equation (10), we can now prove that to find the factors set, it is sufficient to only consider a finite set.

Proposition 11. *For $n \geq 2$, we have:*

$$F(A_{n+1}, f_{2n} - f_{2n-3}) = F(A, f_{2n} - f_{2n-3}) \tag{11}$$

Proof. It is clear that we have $F(A_{n+1}, f_{2n} - f_{2n-3}) \subseteq F(A, f_{2n} - f_{2n-3})$. For the reversed inclusion, let $x \in F(A, f_{2n} - f_{2n-3})$. Then, there is a smallest $m \geq n+1$, such that x is a sub-word of an element of A_m. Then, Equation (10) gives:

$$x \in F(A_m, f_{2n} - f_{2n-3}) = F(A_{n+1}, f_{2n} - f_{2n-3})$$

which shows the desired inclusion. □

4. Fibonacci Numbers Revisited

In this section, we shall restate, and adopt for our purposes, some of the Diophantine properties of the Fibonacci numbers and use them to derive results on the distribution of the letters in the words in the sets, A_n and B_n. Let us introduce the notation:

$$\tau = \frac{1+\sqrt{5}}{2} \qquad \text{and} \qquad \widehat{\tau} = \frac{1-\sqrt{5}}{2}$$

for the roots of $x^2 - x - 1 = 0$. It is well known that τ and $\widehat{\tau}$ appear in Binet's formula, the Fibonacci numbers; see [7]:

$$f_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n = \frac{1}{\sqrt{5}}\left(\tau^n - \widehat{\tau}^{\,n}\right) \tag{12}$$

From Equation (12), it is with induction straight forward to derive:

$$f_n = \tau f_{n-1} + \widehat{\tau}^{\,n-1} = \tau^2 f_{n-2} + \widehat{\tau}^{\,n-2} \tag{13}$$

Definition 12. *Let $\|\cdot\|$ denote the smallest distance to an integer.*

By using the special property, $\tau^2 = \tau + 1$, we have for an integer, k, the following line of equalities:

$$\left\|\frac{1}{\tau^2}\,k\right\| = \left\|\frac{\tau-1}{\tau}\,k\right\| = \left\|k - \frac{1}{\tau}\,k\right\| = \left\|\frac{1}{\tau}\,k\right\| = \left\|(\tau-1)k\right\| = \left\|\tau k\right\|$$

From Equation (13), it follows that:

$$\|\tau f_n\| = \left\|f_{n+1} - \widehat{\tau}^{\,n}\right\| = \frac{1}{\tau^n}$$

since $\widehat{\tau} = -\frac{1}{\tau}$. For an integer, k, which is not a Fibonacci number, we have the following estimate of how far away from an integer τk is.

Proposition 13. *For a positive integer, k, such that $f_{n-1} < k < f_n$, we have:*

$$\|\tau k\| > \frac{1}{\tau^{n-2}} \tag{14}$$

Proof. We give a proof by induction on n. For the basis case, $n = 5$, the statement of the proposition follows by an easy calculation. Now, assume for induction that Equation (14) holds for $5 \leq n \leq p$. For the induction step, $n = p + 1$, let $f_p < k < f_{p+1}$. Then, if $k - f_{p-1}$ is not a Fibonacci number, we have:

$$\|\tau k\| = \big\|\tau(k - f_p) + \tau f_p\big\| \geq \big\|\tau \underbrace{(k - f_p)}_{<f_{p-1}}\big\| - \big\|\tau f_p\big\| > \frac{1}{\tau^{p-3}} - \frac{1}{\tau^p} > \frac{1}{\tau^{p-2}}$$

If $k - f_{p-1} = f_m$ for some $m < p - 1$, then:

$$\|\tau k\| \geq \|\tau f_m\| - \|\tau f_p\| = \frac{1}{\tau^m} - \frac{1}{\tau^p} \geq \frac{1}{\tau^{p-2}} - \frac{1}{\tau^p} = \frac{1}{\tau^{p-1}} \qquad \square$$

Proposition 14. *Let $x \in A_n[1, k]$ for $1 \leq k \leq f_{2n}$ (or $x \in B_n[1, k]$ for $1 \leq k \leq f_{2n-1}$) and $n \geq 2$. Then:*

$$|x|_{\mathtt{b}} \in \left\{ \left\lfloor \tfrac{1}{\tau^2} k \right\rfloor, \left\lceil \tfrac{1}{\tau^2} k \right\rceil \right\} \tag{15}$$

Proof. We give a proof by induction on n. The basis case, $n = 2$, follows by considering each of the words contained in A_2 and B_2. To be able to use Proposition 13 in the induction step, we have to consider the basis step, $n = 3$, as well, but only for the set, B_3 (since the words in A_2 are of length ≥ 3). This is, however, seen to hold by a straight forward enumeration of the elements of B_3.

Now, assume for induction that Equation (15) holds for $2 \leq n \leq p$, for words both from A_n and B_n. For the induction step, $n = p + 1$, let us first derive an identity of which we shall later make use. Let q and m be positive integers, such that $f_{m-1} < q < f_m$. Then, by the help of Proposition 13, we have:

$$\begin{aligned} \left\lfloor \frac{1}{\tau^2}(q - f_{m-1}) \right\rfloor &= \left\lfloor \frac{1}{\tau^2} q - f_{m-3} - \widehat{\tau}^{\,m-1} \right\rfloor \\ &= \left\lfloor \frac{1}{\tau^2} q + \frac{(-1)^m}{\tau^{m-1}} \right\rfloor - f_{m-3} \\ &= \left\lfloor \frac{1}{\tau^2} q \right\rfloor - f_{m-3} \end{aligned} \tag{16}$$

With the same argumentation, we can derive a similar result for $\lceil \cdot \rceil$. For the induction step, we consider first the number of $\mathtt{b}$s in prefixes of words in $A_{p+1} = B_p A_p A_p$. It is clear from the induction assumption that Equation (15) holds for $1 \leq k \leq f_{2p-1}$. For $f_{2p-1} < k < f_{2p}$ or $f_{2p} < k < f_{2p+1}$, let $x = uv \in A_{p+1}[1, k]$, where $u \in B_p$. By the induction assumption, we may assume that $|v|_{\mathtt{b}}$ is given by rounding downwards, (the result is obtained in a similar way for the case with $\lceil \cdot \rceil$). By Equation (16), it now follows that:

$$|uv|_{\mathtt{b}} = |u|_{\mathtt{b}} + |v|_{\mathtt{b}} = f_{2p-3} + \left\lfloor \frac{1}{\tau^2}(k - f_{2p-1}) \right\rfloor = \left\lfloor \frac{1}{\tau^2} k \right\rfloor$$

For $k = f_{2p}$, we have:

$$|uv|_{\mathtt{b}} = f_{2p-3} + \left\lfloor \frac{1}{\tau^2}(f_{2p} - f_{2p-1}) \right\rfloor = \left\lfloor \frac{1}{\tau^2} f_{2p} + \frac{1}{\tau^{2p-1}} \right\rfloor = \left\lceil \frac{1}{\tau^2} f_{2p} \right\rceil$$

For $f_{2p+1} < k < f_{2p+2}$, let $x = uvw \in A_{p+1}[1, k]$, where $u \in B_p$ and $v \in A_p$. Then, the induction assumption and Equation (16) gives:

$$\begin{aligned} |uvw|_{\mathtt{b}} &= |u|_{\mathtt{b}} + |v|_{\mathtt{b}} + |w|_{\mathtt{b}} \\ &= f_{2p-3} + f_{2p-2} + \left\lfloor \frac{1}{\tau^2}(k - f_{2p-1} - f_{2p}) \right\rfloor \\ &= f_{2p-1} + \left\lfloor \frac{1}{\tau^2}(k - f_{2p+1}) \right\rfloor \\ &= \left\lfloor \frac{1}{\tau^2} k \right\rfloor \end{aligned}$$

For the last case, $k = f_{2p+2}$, we have:

$$|x|_{\mathtt{b}} = \left\lfloor \frac{1}{\tau^2}(f_{2p+2}) \right\rfloor = \left\lfloor f_{2p} + \frac{1}{\tau^{2p}} \right\rfloor = f_{2p}$$

The case when we consider words from B_{p+1} is treated in the same way, but where we do not need to do the induction step for the case $n = 3$. This completes the induction and the proof. □

Proposition 15. *Let $x \in F(A_{n+2}, f_{2n})$ for $n \geq 2$. Then:*

$$f_{2n-2} - 1 \leq |x|_{\mathtt{b}} \leq f_{2n-2} + 1 \tag{17}$$

Proof. Let us first turn our attention to the upper bound in Equation (17). In the same way as in the proof of Proposition 10, we consider sub-words of the seven sets, given in Equation (9).

If x is a sub-word, beginning at position $2 < k \leq f_{2n}$, in an element in A_nA_n or A_nB_n, then:

$$|x|_{\mathtt{b}} \leq f_{2n-2} + \left\lceil \frac{1}{\tau^2}\big((k - f_{2n}) + f_{2n}\big) \right\rceil - \left\lfloor \frac{1}{\tau^2} k \right\rfloor \leq f_{2n-2} + 1$$

since the number of $\mathtt{b}$s in a word in A_n is f_{2n-2}, and a word in A_n is of length f_{2n}. The proof of the upper bound in Equation (17) and for the other sets in Equation (9) is obtained in the same way.

For the lower bound, we have:

$$\begin{aligned} |x|_{\mathtt{b}} &\geq \left\lfloor \frac{1}{\tau^2}(k + f_{2n}) \right\rfloor - \left\lceil \frac{1}{\tau^2} k \right\rceil \\ &= f_{2n-2} + \left\lfloor \frac{1}{\tau^2} k + \frac{1}{\tau^{2n-2}} \right\rfloor - \left\lceil \frac{1}{\tau^2} k \right\rceil \\ &\geq f_{2n-2} - 1 \end{aligned}$$

for any $x \in F(A_{n+2}, f_{2n})$. □

5. Estimating the Size of the Sub-Word Set

We shall in this section give an estimate of the sub-word set, $F(A, f_{2n})$, and give the final part of the proof of Theorem 2. Let us introduce the set:

$$C_n = \phi\big(F(A, f_{2n-2} + 1)\big)$$

By Proposition 15, we can estimate the number of bs in words in $F(A, f_{2n-2}+1)$. This estimate then gives that we have bounds on the length of words in C_n. That is, for $x \in C_n$, we have:

$$|x| = |x|_{\mathtt{a}} + |x|_{\mathtt{b}} \geq 3(f_{2n-3}-1) + 2(f_{2n-4}+2) = f_{2n}+1 \tag{18}$$

and:

$$|x| = |x|_{\mathtt{a}} + |x|_{\mathtt{b}} \leq 3(f_{2n-3}+2) + 2(f_{2n-4}-1) = f_{2n}+4 \tag{19}$$

Proposition 16. *For $n \geq 2$, we have:*

$$F(A, f_{2n}) = F\big(C_n, f_{2n}\big)$$

Proof. The set, $F\big(C_n, f_{2n}\big)$, is created by inflating words from $F(A, f_{2n-2}+1)$, which are then cut into suitable lengths. This implies that $F(A, f_{2n}) \supseteq F\big(C_n, f_{2n}\big)$.

For the converse inclusion, let $x \in F(A, f_{2n})$. Then, there is a word, $w \in A_{n+1}$, and words, u, v, such that $uxv \in A_{n+2}$ and $uxv \in \phi(w)$. For any word, $z \in F\big(\{w\}, f_{2n-2}+1\big)$, we have from Equation (18) that any $s \in \phi(z)$ fulfills $f_{2n}+1 \leq |s|$. This gives that there is a word, $z_x \in F\big(\{w\}, f_{2n-2}+1\big)$, such that x is a sub-word of a word in $\phi(z_x)$, which implies $x \in F\big(C_n, f_{2n}\big)$. □

Proposition 17. *For $n \geq 2$, we have:*

$$|F(A, f_{2n})| \leq 2^{f_{2n-3}+2n} \cdot 5^{n-1} \tag{20}$$

Proof. We give a proof by induction on n. For the basis case, $n = 2$, we have:

$$|F(A, f_4)| = 7 \leq 160 = 2^{f_1+4} \cdot 5$$

Assume for induction that Equation (20) holds for $2 \leq n \leq p$. For the induction step, $n = p+1$, note that from Equations (18) and (19), it follows that $|F(\{x\}, f_{2p+2})| \leq 5$ for $x \in C_{p+1}$. By Proposition 15, we have that the number of bs in $u \in F(A, f_{2p}+1)$ is at most $f_{2p-2}+2$. This gives, then, with the help of the induction assumption:

$$\begin{aligned}
\big|F\big(A, f_{2p+2}\big)\big| &\leq |C_{p+1}| \cdot 5 \\
&\leq \big|F\big(A, f_{2p}\big)\big| \cdot 2^{f_{2p-2}+2} \cdot 5 \\
&\leq 2^{f_{2p-3}+2p} \cdot 5^{p-1} \cdot 2^{f_{2p-2}+2} \cdot 5 \\
&= 2^{f_{2(p+1)-3}+2(p+1)} \cdot 5^{p}
\end{aligned}$$

which completes the proof. □

We can now turn to proving the last equality in Equation (1) and, thereby, completing the proof of Theorem 2. By Proposition 17, we have:

$$\begin{aligned}
\lim_{n\to\infty} \frac{\log|F(A, f_{2n})|}{f_{2n}} &\leq \lim_{n\to\infty} \frac{\log\left(2^{f_{2n-3}+2n} \cdot 5^{n-1}\right)}{f_{2n}} \\
&= \lim_{n\to\infty} \frac{f_{2n-3}+2n}{f_{2n}} \log 2 + \frac{n-1}{f_{2n}} \log 5 \\
&= \frac{1}{\tau^3} \log 2
\end{aligned}$$

which implies the equality in Equation (1).

A further generalization of the random Fibonacci substitutions would be to study the structure occurring when mixing two substitutions with different inflation multipliers. This, however, seems to be a far more complex question.

Acknowledgments

The author wishes to thank M. Baake and M. Moll at Bielefeld University, Germany, for our discussions and for reading drafts of the manuscript. This work was supported by the German Research Council (DFG), via CRC 701.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Godrèche, C.; Luck, J.M. Quasiperiodicity and randomness in tilings of the plane. *J. Stat. Phys.* **1989**, *55*, 1–28.
2. Nilsson, J. On the entropy of a family of random substitutions, *Monatsh. Math.* **2012**, *168*, 563–577.
3. Baake, M.; Moll, M. Random Noble Means Substitutions; In *Aperiodic Crystals*; Schmid, S., Withers, R.L., Lifshitz, R., Eds.; Springer: Dordrecht, The Netherlands, 2013; In press.
4. Luck, J.M.; Godrèche, C.; Janner, A.; Janssen, T. The nature of the atomic surfaces of quasiperiodic self-similar structures. *J. Phys. A: Math. Gen.* **1993**, *26*, 1951–1999.
5. Lind, D.; Marcus, B. *An Introduction to Symbolic Dynamics and Coding*; Cambridge University Press: Cambridge, UK, 1995.
6. Fekete, M. Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit. ganzzahligen Koeffizienten. *Math. Zeitschrift* **1923**, *17*, 228–249.
7. Graham, R.L.; Knuth, D.E.; Patashnik, O. *Concrete Mathematics*; Addison-Wesley Publishing Company: Reading, MA, USA, 1994.