The Number of Valid Factorizations of Fibonacci Prefixes

P. Bonardo and A. E. Frid
Aix-Marseille Université
CNRS, Centrale Marseille
I2M
Marseille, France
anna.e.frid@gmail.com

Jeffrey Shallit
School of Computer Science
University of Waterloo
Waterloo, ON N2L 3G1
Canada
shallit@uwaterloo.ca

July 13, 2018

Abstract

We establish several recurrence relations and an explicit formula for $V(n)$, the number of factorizations of the length-n prefix of the Fibonacci word into a (not necessarily strictly) decreasing sequence of standard Fibonacci words. In particular, we show that the sequence $V(n)$ is the shuffle of the ceilings of two linear functions of n.

1 Introduction

In the classical Fibonacci, or Zeckendorf, numeration system [5, 10], a positive integer is represented as a sum of Fibonacci numbers:

$$n = F_{m_k} + F_{m_{k-1}} + \cdots + F_{m_0},$$

where $m_k > m_{k-1} > \cdots > m_0 \geq 2$ and, as usual, $F_0 = 0$, $F_1 = 1$, and $F_{m+2} = F_{m+1} + F_m$ for all $m \geq 0$. For example, $16 = 13 + 3 = F_7 + F_4 = [100100]_F$, where a digit in brackets is 1 if the respective Fibonacci number appears in the sum, and 0 otherwise. Here a representation ends by the digits corresponding to $F_4 = 3$, $F_3 = 2$ and $F_2 = 1$.

Under the condition that m_i and m_{i+1} are never consecutive, that is, $m_{i+1} - m_i \geq 2$, or, equivalently, that the Fibonacci numbers F_i are chosen greedily, such a *canonical* representation is unique, and the language L_V of all canonical representations is given by the regular expression $\epsilon + 1(0 + 01)^*$. At the same time, if consecutive Fibonacci numbers are allowed, but at most once each, the number of such *legal* representations of n is the well-known integer sequence A000119 from the Online Encyclopedia of Integer Sequences (OEIS) [7]. Its values oscillate between 1 on numbers of the form $F_i - 1$ and $\sqrt{n+1}$ on numbers of the form $n = F_i^2 - 1$ [9].

For example, since

$$16 = 13+3 = 8+5+3 = 8+5+2+1 = 13+2+1$$
$$= [100100]_F = [11100]_F = [11011]_F = [100011]_F,$$

the number of legal representations of 16 is 4. Each legal representation of n can be obtained from a canonical one by a series of replacements

$$\cdots 100 \cdots \longleftrightarrow \cdots 011 \cdots ,$$

corresponding to the replacement of a Fibonacci number F_{m+2} by $F_{m+1}+F_m$.

In this paper, we allow even more freedom in Fibonacci representations of n, allowing the transformations

$$\cdots k0l \cdots \longleftrightarrow \cdots (k-1)\ 1 (l+1) \cdots \tag{1}$$

for all $k > 0$, $l \geq 0$. Note that the introduced transformation corresponds to passing from a sum of the form $kF_{m+1} + lF_{m-1}$ to the sum $(k-1)F_{m+1} + F_m + (l+1)F_{m-1}$, and, in particular, does not change the represented number.

The representations that can be obtained from the canonical one by a series of transformations as in (1) are called *valid* and were introduced in [4] in a more general setting because of their link to the Fibonacci word and factorizations of its prefixes, explained below. Clearly, each legal representation is valid, but the opposite is not true. For example, starting from the legal representation $16 = [11011]_F$, we can find two more valid representations

$$16 = [10121]_F = [1221]_F,$$

and starting from the legal representation $16 = [11100]_F$, we find a new representation

$$16 = [20000]_F,$$

so that the total number of valid representations of 16 is 7.

Let $V(n)$ denote the number of valid representations of n. The goal of this paper is to prove a precise formula for $V(n)$, given below in Theorem 1. Our formula demonstrates that the values of $V(n)$ are determined by the shuffle of two straight lines of irrational slope; see Fig. 1.

2 Notation and Sturmian representations

We use notation common in combinatorics on words; the reader is referred, for example, to [3] for an introduction. Given a finite word u, we denote its length by $|u|$. The power u^k just means the concatenation $u^k = \underbrace{u \cdots u}_{k}$. The i'th symbol of a finite or infinite word u is denoted by $u[i]$, so that $u = u[1]u[2]\cdots$. A factor $w[i+1]w[i+2]\cdots w[j]$ of a finite or infinite word w, or, more precisely, its occurrence starting from position $i+1$ of w, is denoted by $w(i..j]$. In particular, for $j \geq 0$, the word $w(0..j]$ is the prefix of w of length j.

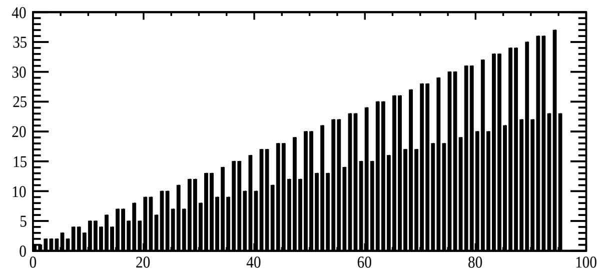

Figure 1: First 100 values of $V(n)$

The *standard Fibonacci sequence* (f_n) of words over the binary alphabet $\{a, b\}$ is defined as follows:

$$f_{-1} = b, \qquad f_0 = a, \qquad f_{n+1} = f_n f_{n-1} \text{ for all } n \geq 0. \tag{2}$$

The word f_n is called also the *standard word of order* n. In particular, $f_1 = ab$, $f_2 = aba$, $f_3 = abaab$, $f_4 = abaababa$, and so on. From the definition, we easily see that the length of f_n is the Fibonacci number F_{n+2}.

The infinite word

$$\mathbf{f} = \lim_{n\to\infty} f_n = abaababaabaababaababa\cdots$$

is called the Fibonacci infinite word. Here we index it starting with $\mathbf{f}[1] = a$.

In the *Fibonacci*, or *Zeckendorf numeration system*, a non-negative integer $N < F_{n+3}$ is represented as a sum of Fibonacci numbers

$$N = \sum_{0\leq i\leq n} k_i F_{i+2}, \tag{3}$$

where $k_i \in \{0, 1\}$ for $i \geq 0$. In the canonical version of the definition, the following condition holds:

$$\text{for } i \geq 1, \text{ if } k_i = 1, \text{ then } k_{i-1} = 0. \tag{4}$$

Under this nonadjacency condition, the representation of N is unique up to leading zeros. However, by removing the nonadjacency condition, we get multiple representations: for example, $14 = F_7 + F_2 = F_6 + F_5 + F_2 = F_6 + F_4 + F_3 + F_2$. We call such representations *legal* and denote a representation $N = \sum_{0\leq i\leq n} k_i F_{i+2}$ by $N = [k_n \cdots k_0]_F$. If the condition (4) holds, we call the representation *canonical*.

Let $L(n)$ denote the number of legal representations of n. The sequence $(L(n))$ is well-studied (see, e.g., [2]) and listed in the OEIS as sequence A000119. In particular, $1 \leq L(n) \leq \sqrt{n+1}$, and both bounds are precise [9].

The following lemma is a particular case of [4, Prop. 2].

Lemma 1. *For all $k_0, \ldots, k_n$ such that $k_i \in \{0,1\}$, the word $f_n^{k_n} f_{n-1}^{k_{n-1}} \cdots f_0^{k_0}$ is a prefix of the Fibonacci word $\mathbf{f}$.*

So $L(n)$ is also the number of ways to factor the prefix $\mathbf{f}(0..n]$ of the Fibonacci word as a sequence of standard words in strictly decreasing order.

To expand this definition, in this note we consider all factorizations of Fibonacci prefixes $\mathbf{f}(0..n]$ as a concatenation of standard words in (non-strictly) decreasing order. We write $N = [k_n \cdots k_0]_F$ and call this representation of N *valid* if $k_i \geq 0$ for all i and $\mathbf{f}(0..N] = f_n^{k_n} f_{n-1}^{k_{n-1}} \cdots f_0^{k_0}$. Note that according to the previous lemma, every legal representation is valid, but not the other way around. For example, $\mathbf{f}(0..14] = (abaab)(aba)(aba)(aba)$, making the representation $14 = [1300]_F$ valid. Theorem 1 of [4] says, in particular, that valid representations are exactly those that can be obtained from the canonical one by a series of transformations (1).

The number of valid representations of N is denoted by $V(N)$, and this note is devoted to the study of the sequence $(V(n))$, recently listed in the OEIS as A300066. Clearly, $V(n) \geq L(n)$, and moreover, we prove an explicit formula for $V(n)$ that implies its linear growth.

3 Result

As is well-known, the Fibonacci infinite word

$$\mathbf{f} = abaababa \cdots$$

is the fixed point of the Fibonacci morphism $\mu : a \to ab, b \to a$; moreover, for each $n \geq 1$, we have $f_n = \mu(f_{n-1})$. Consequently, if $N = \sum_{0 \leq i \leq n} k_i F_{i+2}$ for $k_i \in \{0,1\}$, that is, $N = [k_n \cdots k_0]_F$, then Lemma 1 implies that

$$\mu(\mathbf{f}(0..N]) = \mu(\mathbf{f}(0..[k_n \cdots k_0]_F]) = \mu(f_n^{k_n} \cdots f_0^{k_0}) = f_{n+1}^{k_n} \cdots f_1^{k_0} = \mathbf{f}(0..[k_n \cdots k_0 0]_F].$$

Let φ denote the golden ratio: $\varphi = \frac{1+\sqrt{5}}{2}$. It is important that the Fibonacci word is a Sturmian word of slope $1/(\varphi+1) = 1/\varphi^2$ and zero intercept, that is, for all n, we have

$$\mathbf{f}[n] = \begin{cases} a, & \text{if } \{n/\varphi^2\} < 1 - 1/\varphi^2; \\ b, & \text{otherwise.} \end{cases} \tag{5}$$

Here $\{x\} = x - \lfloor x \rfloor$ denotes the fractional part of x.

Proposition 1. *If $\mathbf{f}[n] = a$, all valid representations of n end with an even number of 0s. If $\mathbf{f}[n] = b$, all of them end with an odd number of 0s.*

Proof. It suffices to consult the definition of a valid representation and notice that f_i ends with a if and only if i is even. □

We now state our main result.

Theorem 1. *If* $\mathbf{f}[n] = a$*, then* $V(n) = \lceil n/\varphi^3 \rceil$*, or, equivalently,* $V(n)$ *is equal to the number of occurrences of* b *in* $\mathbf{f}(0..n]$*, plus one. If* $\mathbf{f}[n] = b$*, then* $V(n) = \lceil n/\varphi^3 \rceil$*, or, equivalently,* $V(n)$ *is equal to the number of occurrences of* aa *in* $\mathbf{f}(0..n]$*, plus one.*

To prove the theorem, we will need several more propositions.

Proposition 2.

(a) $V([r0]_F) \geq V([r]_F)$ *for all* $r \in \{0,1\}^*$*.*

(b) If $r = r'10^{2k}$ *for some* $k \geq 0$*, then* $V([r0]_F) = V([r]_F)$*.*

Proof. (a): Consider a factorization $\mathbf{f}(0..[r]_F] = f_n^{k_n} f_{n-1}^{k_{n-1}} \cdots f_0^{k_0}$. Applying the Fibonacci morphism μ to both sides, we get the factorization $\mathbf{f}(0..[r0]_F] = f_{n+1}^{k_n} f_n^{k_{n-1}} \cdots f_1^{k_0}$. So the number of factorizations of $\mathbf{f}(0..[r0]_F]$ (which is equal to $V([r0]_F)$) is at least as large as the number of factorizations of $\mathbf{f}(0..[r]_F]$ (which is equal to $V([r]_F)$).

(b) If, in addition $r = r'10^{2k}$ for some $k \geq 0$, we see that $\mathbf{f}(0..[r]_F]$ ends with f_{2k} and $\mathbf{f}(0..[r0]_F]$ ends with f_{2k+1}, which in turn ends with b. From Proposition 1, no factorization of $\mathbf{f}(0..[r0]_F]$ ends with f_0; that is, such a factorization must be of the form $\mathbf{f}(0..[r0]_F] = f_{n+1}^{k_n} f_n^{k_{n-1}} \cdots f_1^{k_0}$. Taking the μ-preimage, we get the factorization $\mathbf{f}(0..[r]_F] = f_n^{k_n} f_{n-1}^{k_{n-1}} \cdots f_0^{k_0}$, thus establishing a bijection and the equality $V([r0]_F) = V([r]_F)$. □

Proposition 3. *We have*

$$V([z10^{2k}]_F) = V([z10^{2k-2}]_F) + V([z(01)^k]_F).$$

for all $z \in \{0,1\}^*$ *and all* $k \geq 1$*.*

Proof. Proposition 1 tells us that $\mathbf{f}[[z10^{2k}]_F] = a$, and moreover, since $k > 0$, the prefix of length $[z10^{2k}]_F$ of s ends with aba, which is a suffix of f_{2k}. Consider a valid factorization $[z10^{2k}]_F = f_n^{k_n} f_{n-1}^{k_{n-1}} \cdots f_0^{k_0}$. If $k_0 = 0$, then $k_1 = 0$ since f_1 ends with b, so the factorization is of the form $f_n^{k_n} f_{n-1}^{k_{n-1}} \cdots f_2^{k_2}$. Taking the μ^2-preimage, we get a factorization $f_{n-2}^{k_n} f_{n-3}^{k_{n-1}} \cdots f_0^{k_2}$ of $[z10^{2k-2}]_F$. Moreover, μ^2 is a bijection: every factorization $[z10^{2k-2}]_F$ corresponds to a factorization of $[z10^{2k}]_F$ with $k_0 = k_1 = 0$.

On the other hand, if $k_0 \neq 0$, then $k_0 = 1$ since the word that we factor ends with aba. Removing this last occurrence of $f_0 = a$, we get the prefix of s of length $[z10^{2k}]_F - 1 = [z(01)^k 0]_F$. From Proposition 2, the number of valid factorizations of $[z(01)^k 0]_F$ is equal to that of $[z(01)^k]_F$. Combining the two possibilities, we get the statement of the proposition. □

Proposition 4. *For all* $z \in \{0,1\}^*$ *and for all* $k \geq 1$*, we have*

$$V([z10^k1]_F) = \begin{cases} V([z10^{k+1}]_F), & \text{if } k \text{ is odd;} \\ V([z10^k]_F) + V([z(01)^{k/2}]_F), & \text{if } k \text{ is even.} \end{cases}$$

Proof. If k is odd, then $[z10^k1]_F = [z10^{k+1}]_F + 1$, and the prefix $\mathbf{f}(0..[z10^{k+1}]_F]$ was considered in the previous proposition. It ends with aba, and the symbol added to get $\mathbf{f}(0..[z10^k1]_F]$ is also a. So $\mathbf{f}(0..[z10^k1]_F]$ ends with $abaa$, and all valid factorizations end with f_0. This means that the number of valid factorizations of $\mathbf{f}(0..[z10^k1]_F]$ is equal to that of $\mathbf{f}[0..[z10^{k+1}]_F]$; that is, $V([z10^k1]_F) = V([z10^{k+1}]_F)$.

If k is even, $k > 0$, then $\mathbf{f}(0..[z10^k1]_F]$ ends with $f_3f_0 = abaaba$. In particular, the last factor of any valid factorization of $\mathbf{f}(0..[z10^k1]_F]$ is either $f_0 = a$, or $f_2 = aba$. Indeed, $f_4 = abaababa$ and thus for all $l > 2$ the f_{2l} do not have a common suffix with $\mathbf{f}(0..[z10^k1]_F]$. So, letting $V_2(n)$ denote the number of factorizations of $\mathbf{f}(0..n]$ of the form $f_n^{k_n} f_{n-1}^{k_{n-1}} \cdots f_2^{k_2}$, we get

$$\begin{aligned} V([z10^k1]_F) &= V([z10^k1]_F - 1) + V_2([z10^k1]_F - 3) \\ &= V([z10^{k+1}]_F) + V_2([z(01)^{k/2}00]_F) \\ &= V([z10^k]_F) + V([z(01)^{k/2}]_F). \end{aligned}$$

Here the last equality follows from Proposition 2 (for the first addend) and by taking μ^{-2} of each factorization (for the second one). □

Propositions 2 to 4 give a full list of recurrence relations sufficient to compute $V(n)$ for every $n > 1$, starting from $V(1) = 1$. Before using them to prove the main theorem, we consider two particular cases.

Corollary 1. *For all $k \geq 1$ we have*

$$V(F_{2k+1} - 1) = V(F_{2k+1} - 2) = F_{2k-1}$$

and

$$V(F_{2k+2} - 2) = F_{2k}$$

Proof. For $k = 1$, the equalities can be easily checked: $V(F_3 - 1) = V(1) = V(F_3 - 2) = V(0) = 1 = F_1$, and $V(F_4 - 2) = V(1) = 1 = F_2$. We also observe that $F_{2k+1} - 1 = [(10)^{k-1}1]_F$, $F_{2k+1} - 2 = [(10)^{k-1}0]_F$, and $F_{2k+2} - 2 = [(10)^{k-1}01]_F$. Now we assume that the equalities hold for k, and use Propositions 3 and 4 to prove they hold for $k + 1$:

$$\begin{aligned} V(F_{2k+3} - 2) &= V([(10)^k0]_F) = V([(10)^{k-1}1]_F) + V([(10)^{k-1}01]_F) \\ &= V(F_{2k+1} - 1) + V(F_{2k+2} - 2) = F_{2k-1} + F_{2k} = F_{2k+1}, \\ V(F_{2k+3} - 1) &= V([(10)^k1]_F) = V([(10)^k0]_F) = V(F_{2k+3} - 2) = F_{2k+1}, \\ V(F_{2k+4} - 2) &= V([(10)^k01]_F) = V([(10)^k0]_F) + V([(10)^{k-1}01]_F) \\ &= V(F_{2k+3} - 2) + V(F_{2k+2} - 2) = F_{2k+1} + F_{2k} = F_{2k+2}. \end{aligned}$$

□

Corollary 2. *For all $k \geq 1$, we have*

$$V(F_{2k}) = V(F_{2k+1}) = F_{2k-2} + 1.$$

Proof. For $k = 1$, the equalities can be easily checked: $V(F_2) = V(1) = V(F_3) = V(2) = 1 = F_0 + 1$. Suppose the equalities hold for k; let us prove them for $k+1$. With Proposition 3, we have

$$V(F_{2k+2}) = V([10^{2k}]_F) = V([10^{2k-2}]_F) + V([(10)^{k-1}1]_F) = F_{2k-2}+1+F_{2k-1} = F_{2k}+1,$$

and with Proposition 2, we have

$$V(F_{2k+3}) = V([10^{2k+1}]_F) = V([10^{2k}]_F) = V(F_{2k+2}) = F_{2k} + 1.$$

□

Proposition 5. *Let $n = [z]_F$ and $n' = [z0]_F$ be such that $\mathbf{f}[n] = a$. Then $\lceil n/\varphi^2 \rceil = \lceil n'/\varphi^3 \rceil$.*

Proof. Let us write the canonical Fibonacci representation of n as $\sum_{0 \le i \le l} F_{m_i}$. Since $\mathbf{f}[n] = a$, from Proposition 1 we get that m_1 is even.

Now $F_k = \frac{1}{\sqrt{5}}(\varphi^k - \psi^k)$, where $\psi = \frac{1-\sqrt{5}}{2}$, $-1 < \psi < 0$. So

$$n = \sum_{0 \le i \le l} F_{m_i} = \frac{1}{\sqrt{5}} \left(\sum_{0 \le i \le l} \varphi^{m_i} - \sum_{0 \le i \le l} \psi^{m_i} \right)$$

and

$$\frac{n}{\varphi^2} = \frac{1}{\sqrt{5}} \left(\sum_{0 \le i \le l} \varphi^{m_i - 2} - \frac{1}{\varphi^2} \sum_{0 \le i \le l} \psi^{m_i} \right).$$

At the same time,

$$n' = \sum_{0 \le i \le l} F_{m_i+1} = \frac{1}{\sqrt{5}}(\sum_{0 \le i \le l} \varphi^{m_i+1} - \sum_{0 \le i \le l} \psi^{m_i+1})$$

and thus

$$\frac{n'}{\varphi^3} = \frac{1}{\sqrt{5}} \left(\sum_{0 \le i \le l} \varphi^{m_i - 2} - \frac{1}{\varphi^3} \sum_{0 \le i \le l} \psi^{m_i+1} \right).$$

The difference between the two ratios is

$$\frac{n'}{\varphi^3} - \frac{n}{\varphi^2} = \frac{\psi^2}{\sqrt{5}\varphi^2} \left(1 - \frac{\psi}{\varphi} \right) S,$$

where

$$S = \sum_{0 \le i \le l} \psi^{m_i} = \psi^{m_1} \sum_{0 \le i \le l} \psi^{m_i - m_1}.$$

Let us estimate S. Since m_1 is even and $\psi^{m_1} > 0$, an upper bound for S is

$$S < \psi^{m_1} \sum_{k=0}^{\infty} \psi^{2k} = \frac{\psi^{m_1}}{1 - \psi^2} \le \frac{1}{1 - \psi^2},$$

whereas a lower bound is

$$S > \psi^{m_1}\left(1+\sum_{k=1}^{\infty}\psi^{2k+1}\right) > \psi^{m_1}\left(1+\sum_{k=0}^{\infty}\psi^{2k+1}\right) = \psi^{m_1}\left(1+\frac{\psi}{1-\psi^2}\right) = 0.$$

So

$$0 < \frac{n'}{\varphi^3} - \frac{n}{\varphi^2} < \frac{\psi^2}{\sqrt{5}\varphi^2}\left(1-\frac{\psi}{\varphi}\right)\frac{1}{1-\psi^2} = \frac{1}{\varphi^2}.$$

Together with (5), meaning that $\{n/\varphi^2\} < 1 - 1/\varphi^2$, this implies the statement of the Proposition. □

Proof of Theorem 1. Let us start with the case of $\mathbf{f}[n] = a$ and proceed by induction starting with $V(1) = 1$. There are three subcases:

(a) $n = [z10^{2k}]_F$, $k > 0$;

(b) $n = [z10^k1]_F$, k odd;

(c) $n = [z10^k1]_F$, k even.

(a) Since $n = [z10^{2k}]_F$ and $k > 0$, Proposition 3 gives $V(n) = V([z10^{2k}]_F) = V([z10^{2k-2}]_F) + V([z(01)^k]_F)$. Write $[z10^{2k-2}]_F = n'$ and $[z(01)^k]_F = n''$. Note that Proposition 1 gives $\mathbf{f}[n'] = \mathbf{f}[n''] = a$. At the same time, $n''+1 = [z10^{2k-1}]_F$ and thus $\mathbf{f}[n''+1] = b$. Now (5) implies that $\{n'/\varphi^2\} \in (0, 1-1/\varphi^2)$ and $\{(n''+1)/\varphi^2\} \in (1-1/\varphi^2, 1)$. Also, the Fibonacci representation of n' is obtained from that of $n''+1$ by a one-symbol shift to the left, so, due to the Fibonacci recurrence relation, $n' + n'' + 1 = n$.

Let us consider the sum $t = \{n'/\varphi^2\} + \{(n''+1)/\varphi^2\}$. From the inclusions above, we see that t belongs to the interval $(1-1/\varphi^2, 2-1/\varphi^2)$. But we also know that $\{n/\varphi^2\} = \{(n'+n''+1)/\varphi^2\} \in (0, 1-1/\varphi^2)$, since $\mathbf{f}[n] = a$. So

$$\{n/\varphi^2\} = \{n'/\varphi^2\} + \{(n''+1)/\varphi^2\} - 1,$$

which is equivalent to $\lfloor n/\varphi^2 \rfloor = \lfloor n'/\varphi^2 \rfloor + \lfloor (n''+1)/\varphi^2 \rfloor + 1$ and to $\lfloor n/\varphi^2 \rfloor = \lfloor n'/\varphi^2 \rfloor + \lfloor n''/\varphi^2 \rfloor + 1$ (since $\lfloor n''/\varphi^2 \rfloor = \lfloor (n''+1)/\varphi^2 \rfloor$). Since all the numbers under consideration are irrational, and thus every ceiling is just the floor plus 1, we get

$$\lceil n/\varphi^2 \rceil = \lceil n'/\varphi^2 \rceil + \lceil n''/\varphi^2 \rceil.$$

To establish the statement of the theorem for this subcase, it is sufficient to use Proposition 3 and the induction hypothesis: $V(n') = \lceil n'/\varphi^2 \rceil$ and $V(n'') = \lceil n''/\varphi^2 \rceil$.

(b): Here $n = [z10^{2k-1}1]_F$ and $k > 0$. It suffices to refer to the previous subcase and to Proposition 4: $V(n) = V(n-1) = V([z10^{2k}]_F) = \lceil (n-1)/\varphi^2 \rceil$. It remains to notice that $\lceil (n-1)/\varphi^2 \rceil = \lceil n/\varphi^2 \rceil$, since $\mathbf{f}[n-1] = a$.

(c): Here $n = [z10^{2k}1]_F$ and $k > 0$. We use Proposition 4: $V([z10^{2k}1]_F) = V([z10^{2k}]_F) + V([z(01)^k]_F)$. As above, write $n' = [z10^{2k}]_F$ and $n'' = [z(01)^k]_F$;

then $n = n' + n'' + 2$, whereas $V(n) = V(n') + V(n'')$. By the induction hypothesis, $V(n') = \lceil n'/\varphi^2 \rceil$ and $V(n'') = \lceil n''/\varphi^2 \rceil$.

We have $\mathbf{f}[n] = a$ and $\mathbf{f}[n-1] = b$, implying from (5) that $\{(n-1)/\varphi^2\} \in (1 - 1/\varphi^2, 1)$ and thus $\{n/\varphi^2\} \in (0, 1/\varphi^2)$. At the same time, $\mathbf{f}[n'] = \mathbf{f}[n''] = a$ implies $\{n'/\varphi^2\}, \{n''/\varphi^2\} \in (0, 1 - 1/\varphi^2)$ and thus

$$\{n'/\varphi^2\} + \{n''/\varphi^2\} + \{2/\varphi^2\} \in (2/\varphi^2, 2).$$

Comparing it to $\{n/\varphi^2\} = \{(n' + n'' + 2)/\varphi^2\} \in (0, 1/\varphi^2)$, we see that

$$\{n/\varphi^2\} = \{n'/\varphi^2\} + \{n''/\varphi^2\} + \{2/\varphi^2\} - 1.$$

But since $n = n' + n'' + 2$ and $x = \lfloor x \rfloor + \{x\}$ for every x, it also means that

$$\lfloor n/\varphi^2 \rfloor = \lfloor n'/\varphi^2 \rfloor + \lfloor n''/\varphi^2 \rfloor + 1.$$

Finally, since k/φ^2 is not an integer for any integer $k > 0$, we have $\lceil k/\varphi^2 \rceil = \lfloor n/\varphi^2 \rfloor + 1$, so that

$$\lceil n/\varphi^2 \rceil = \lceil n'/\varphi^2 \rceil + \lceil n''/\varphi^2 \rceil.$$

It remains to use the induction hypothesis to establish

$$V(n) = V(n') + V(n'') = \lceil n'/\varphi^2 \rceil + \lceil n''/\varphi^2 \rceil = \lceil n/\varphi^2 \rceil,$$

which was to be proved.

To complete the part of the proof concerning $\mathbf{f}[n] = a$, it remains to notice that $\lceil n/\varphi^2 \rceil$ is equal to the number of bs in $\mathbf{f}(0..n]$ plus one due to (5).

Now for $\mathbf{f}[n] = b$, it is sufficient to combine Propositions 1, 2 and 5: if $\mathbf{f}[n] = b$, then $n = [r0]_F$, where $m = [r]_F$ and $\mathbf{f}[m] = a$. Then

$$V(n) = V(m) = \lceil m/\varphi^2 \rceil = \lceil n/\varphi^3 \rceil.$$

Here $\mathbf{f}(0..n] = \mu(\mathbf{f}(0..m])$, and so the occurrences of aa in $\mathbf{f}(0..n]$ correspond exactly to occurrences of b in $\mu(\mathbf{f}(0..m])$. The theorem is proved. □

The theorem ensures that the sequence $(V(n))$ grows as shown in Fig. 1. The two visible straight lines correspond to the symbols of the Fibonacci word equal to a (the upper line) or b (the lower line).

4 Fibonacci-regular representation

A sequence $(s(n))_{n \geq 0}$ is said to be *Fibonacci-regular* if there exist an integer k, a row vector v of dimension k, a column vector w of dimension k, and a $k \times k$ matrix-valued morphism ρ on $\{0, 1\}^*$ such that

$$s([z]_F) = v\rho(z)w$$

for all canonical Fibonacci representations $z \in L_V$. The triple (v, ρ, w) is called a *linear representation*; see, for example, [6].

Berstel [2] gave the following linear representation for the function $L(n)$ we mentioned previously in Section 2:

$$v = [1\,0\,0\,0], \quad \rho(0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad \rho(1) = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad w = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Hence $L(n)$ is Fibonacci-regular.

We can find a similar representation for the function $V(n)$. For technical reasons it is easier to deal with the reversed Fibonacci representation; one can then obtain the ordinary linear representation by interchanging the roles of the vectors and taking the transposes of the matrices.

Theorem 2. *$V(n)$ has the reversed linear representation (t, γ, u), where*

$$t = [1\ 0\ 0\ 0\ 0\ 0\ 0\ 0], \qquad u = [1\ 1\ 1\ 1\ 1\ 2\ 1\ 4]^T$$

$$\gamma(0) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 2 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & -3 & 3 & 0 & 1 & 0 \\ -1 & -1 & 0 & 2 & 3 & 0 & 1 & 0 \end{bmatrix}, \quad \gamma(1) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Proof. Define $g(x) = V([x]_F)$ if x is a valid canonical representation (that is, containing no leading zeroes, and no two consecutive 1's), and 0 otherwise. It suffices to show, for all $x \in \{0,1\}^*$, that

$$\begin{bmatrix} g(xi) \\ g(xi0) \\ g(xi1) \\ g(xi00) \\ g(xi000) \\ g(xi100) \\ g(xi0000) \\ g(xi10000) \end{bmatrix} = \gamma(i) \begin{bmatrix} g(x) \\ g(x0) \\ g(x1) \\ g(x00) \\ g(x000) \\ g(x100) \\ g(x0000) \\ g(x10000) \end{bmatrix}. \tag{6}$$

Once we prove this, it is then easy to see (using induction on $|z|$) that, if z is the Fibonacci representation of n, then $t\gamma(z^R)u = V(n)$.

Thus it suffices to verify Eq. (6). This is equivalent to proving the following identities for x.

$$g(x01) = -g(x) + g(x0) + g(x00) \tag{7}$$
$$g(x0100) = -g(x) + 2g(x00) + g(x000) \tag{8}$$
$$g(x00000) = g(x) - g(x0) - 3g(x00) + 3g(x000) + g(x0000) \tag{9}$$
$$g(x010000) = -g(x) - g(x0) + 2g(x00) + 3g(x000) + g(x0000) \tag{10}$$
$$g(x10) = g(x1) \tag{11}$$
$$g(x1000) = g(x100). \tag{12}$$

Identities (11) and (12) are particular cases of Proposition 2 (b).

To prove (7), consider separately two cases: if x ends with an even number of zeros, then $g(x) = g(x0)$ due to Proposition 2 (b) and $g(x00) = g(x01)$ due to Proposition 4, so the identity holds. If x ends with an odd number of zeros, $x = z10^{2k+1}$, $k \geq 0$, then due to Proposition 4,

$$g(x01) = g(z10^{2k+2}1) = g(z10^{2k+2}) + g(z(01)^{k+1}) = g(x0) + g(z(01)^{k+1}).$$

On the other hand, due to Propositions 2 and 3,

$$g(x00) = g(x0) = g(z10^{2k+2}) = g(z10^{2k}) + g(z(01)^{k+1}) = g(x) + g(z(01)^{k+1}).$$

Comparing these equalities, we get (7).

To prove (8), it is sufficient to use Proposition 3 to get

$$g(x0100) = g(x01) + g(x001),$$

and then to use (7) twice, for $g(x01)$ and for $g(x001)$.

To prove (9), we again have to consider two cases. If $x = z10^{2k}$, $k \geq 0$, then due to Proposition 2, $g(x0^5) = g(x0000)$, $g(x000) = g(x00)$, $g(x0) = g(x)$, and the equality holds. If now $x = z10^{2k+1}$, $k \geq 0$, then (9) immediately simplifies with Proposition 2 as

$$g(z10^{2k+6}) - g(z10^{2k+4}) = 3[g(z10^{2k+4} - g(z10^{2k+2})] - [g(z10^{2k+2}) - g(z10^{2k})].$$

Applying Proposition 3, we reduce it to

$$g(z(01)^{k+3}) = 3g(z(01)^{k+2}) - g(z(01)^{k+1}),$$

or, writing $y = z(01)^{k+1}$ and applying Proposition 2 again,

$$g(y0100) = 3g(y00) - g(y).$$

But this is exactly (8) since $g(y00) = g(y000)$.

Finally, to prove (10), it is sufficient to use Propositions 3 and 2 to get

$$g(x010000) = g(x0100) + g(x00101) = g(x0100) + g(x00100).$$

Now (10) is obtained immediately by summing up (8) applied to x and to $x0$:

$$g(x0100) = -g(x) + 2g(x00) + g(x000)$$

and

$$g(x00100) = -g(x0) + 2g(x000) + g(x0000).$$

□

References

[1] J.-P. Allouche and J. Shallit. *Automatic Sequences: Theory, Applications, Generalizations*. Cambridge University Press, 2003.

[2] J. Berstel. An exercise on Fibonacci representations. *Theor. Inform. Appl.* **35** (2001) 491–498.

[3] J. Berstel and P. Séébold. Sturmian words. In: M. Lothaire, *Algebraic Combinatorics on Words*. Cambridge University Press, 2002. Chapter 2, pp. 45–110.

[4] A. E. Frid, Sturmian numeration systems and decompositions to palindromes. *European J. Combin.* **71** (2018) 202–212.

[5] C. G. Lekkerkerker. Voorstelling van natuurlijke getallen door een som van getallen van Fibonacci. *Simon Stevin* **29** (1952) 190–195.

[6] Hamoon Mousavi, Luke Schaeffer, and Jeffrey Shallit. Decision algorithms for Fibonacci-automatic words, I: Basic results. *RAIRO Inform. Théorique* **51** (2016), 39–66.

[7] N. J. A. Sloane et al. The On-Line Encyclopedia of Integer Sequences. Available at `https://oeis.org`.

[8] A. Ostrowski. Bemerkungen zur Theorie der diophantischen Approximationen. *Hamb. Abh.* **1** (1921) 77–98.

[9] P. K. Stockmeyer, A smooth tight upper bound for the Fibonacci representation function $R(N)$, *Fibonacci Quart.* **46/47** (2) (2009), 103–106.

[10] E. Zeckendorf. Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas. *Bull. Soc. Roy. Sci. Liège* **41** (1972) 179–182.