Journal of Integer Sequences, Vol. 18 (2015),
Article 15.9.3

Factorizations of the Fibonacci Infinite Word

Gabriele Fici
Dipartimento di Matematica e Informatica
Università di Palermo
Via Archirafi 34
90123 Palermo
Italy
Gabriele.Fici@unipa.it

Abstract

The aim of this note is to survey the factorizations of the Fibonacci infinite word that make use of the Fibonacci words and other related words, and to show that all these factorizations can be easily derived in sequence starting from elementary properties of the Fibonacci numbers.

1 Preliminaries

The well-known sequence of Fibonacci numbers (sequence A000045 in the *On-Line Encyclopedia of Integer Sequences*) is defined by $F_1 = 1$, $F_2 = 1$ and for every $n > 2$, $F_n = F_{n+1} + F_{n+2}$. The first few values of the sequence F_n are reported in Table 1 for reference.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
F_n	1	1	2	3	5	8	13	21	34	55	89	144	233	377	610	987	1597	2584	4181	6765

Table 1: The first few values of the sequence of Fibonacci numbers.

A basic property of Fibonacci numbers (that can be easily proved by induction) is that 1 plus the sum of the first n Fibonacci numbers is equal to the $(n+2)$-th Fibonacci number:

$$1 + \sum_{i=1}^{n} F_i = F_{n+2}. \tag{1}$$

We recall here a famous result, usually attributed to Zeckendorf [13], but published earlier by Lekkerkerker [8] and which, in fact, is a special case of an older and more general result due to Ostrowski [10]. It permits us to use Fibonacci numbers as a basis for representing integers:

Theorem 1. *Every positive integer can be expressed uniquely as the sum of one or more distinct non-consecutive Fibonacci numbers F_n, $n > 1$.*

For example, $17 = 13+3+1 = F_7+F_4+F_2$, and there is no other way to write 17 as the sum of non-consecutive Fibonacci numbers (assuming the convention that F_1 is not used in the representation). Thus, one can represent natural numbers as strings of 0-1 bits, where the i-th bit (from the right) encodes the presence/absence of the $(i+1)$-th Fibonacci number in the representation given by Theorem 1. So for example the number 17 is represented by 100101. We call this representation of natural numbers the *Zeckendorf representation*.

The first few natural numbers and their Zeckendorf representations are displayed in Table 2, where we padded to the left with 0s in order to have strings of the same length. Note that with 6 bits one can represent the first 21 natural numbers. In fact, for every $n > 0$, there are exactly F_n integers whose leftmost 1 in the Zeckendorf representation is in position n (starting from the right). From (1), we derive that one needs n bits to represent the first F_{n+2} natural numbers.

The strings of length n forming the Zeckendorf representations of the first F_{n+2} natural numbers are precisely all the 0-1 strings of length n not containing two consecutive 1s. These strings are in lexicographic order if the natural numbers are in increasing order from 0 to $F_{n+2}-1$.

Zeck.	decimal	Zeck.	decimal	Zeck.	decimal
000000	0	010000	8	100100	16
000001	1	010001	9	100101	17
000010	2	010010	10	101000	18
000100	3	010100	11	101001	19
000101	4	010101	12	101010	20
001000	5	100000	13		
001001	6	100001	14		
001010	7	100010	15		

Table 2: The Zeckendorf representations of the first few natural numbers coded with 6 bits.

Let us define $f(n)$, for every $n \geq 0$, as the rightmost digit of the Zeckendorf representation of n. For every $n > 1$ we define the *n-th Fibonacci word* as the string

$$f_n = f(0)f(1)\cdots f(F_n - 1)$$

of length $|f_n| = F_n$. By convention, we set $f_1 = 1$. The first few Fibonacci words are shown in Table 3.

$$
\begin{aligned}
f_1 &= 1\\
f_2 &= 0\\
f_3 &= 01\\
f_4 &= 010\\
f_5 &= 01001\\
f_6 &= 01001010\\
f_7 &= 0100101001001\\
f_8 &= 010010100100101001010\\
f_9 &= 0100101001001010010100100101001001
\end{aligned}
$$

Table 3: The first few Fibonacci words.

We also define the *Fibonacci infinite word* f as the limit of f_n as n goes to infinity. That is, f is the infinite word whose n-th letter is the "parity" of the Zeckendorf representation of n:

$$f = f(0)f(1)f(2)f(3)\cdots = 0100101001001010010\cdots$$

In the Zeckendorf representation of an integer, when the n-th digit from the right is a 1, the $(n-1)$-th digit from the right is a 0. Hence, the rightmost $n-2$ digits of the Zeckendorf representations of the natural numbers from F_{n+1} to $F_{n+2}-1$ are the same rightmost $n-2$ digits of the Zeckendorf representations of the first F_n natural numbers. For example, the 2 rightmost digits of the Zeckendorf representations of 5, 6 and 7 are, respectively, 00, 01, 10, as well as the 2 rightmost digits of the Zeckendorf representations of the 0, 1 and 2. We deduce that for every $n > 2$, one has

$$f_n = f_{n-1}f_{n-2}. \tag{2}$$

For more details on Fibonacci words the reader can see, for instance, [1].

Recall that a *factorization* of an infinite word w is a sequence $(x_n)_{n\geq 1}$ of finite words such that w can be expressed as the concatenation of the elements of the sequence, i.e., $w = \prod_{n\geq 1} x_n$.

In general, exhibiting a factorization $(x_n)_{n\geq 1}$ of an infinite word w can be useful to better understand the combinatorics of w, provided the sequence $(x_n)_{n\geq 1}$ has non-trivial combinatorial properties—for example, all the words in the sequence are palindromes, squares, or prefixes of w.

Another point of view consists in defining a factorization by some general rule that can be applied to any infinite word. The sequence $(x_n)_{n\geq 1}$ is therefore determined by the particular instance of the infinite word w (as is the case, for example, in the Lempel-Ziv factorization or in the Lyndon factorization, that we will see below). In this case, the word w can have particular properties that make it a limit example for that particular factorization.

In next sections, we will show a number of factorizations of the Fibonacci infinite word that make use of the Fibonacci finite words and other related words. These factorizations have been introduced over the time in different papers, and we think it can be useful to collect them all together for reference. We also add some (at least to the best of our knowledge) novel factorizations. Moreover, we present these factorizations in an order that allows us to provide a short and elementary proof for each of them, despite the original proofs being sometimes more involved or more technical.

2 Fibonacci words and co-Fibonacci words

The first factorization of the Fibonacci infinite word we exhibit is the following.

Proposition 2. *The Fibonacci infinite word can be obtained by concatenating* 0 *and the Fibonacci words:*

$$\begin{aligned} f &= 0\prod_{n\geq 1} f_n \\ &= 0\cdot 1\cdot 0\cdot 01\cdot 010\cdot 01001\cdot 01001010\cdots \end{aligned} \tag{3}$$

Proof. Since for every $i \geq 1$, $|f_i| = F_i$, it is sufficient to prove that, for every $n \geq 1$, f_n occurs in f starting at position $1 + \sum_{i=1}^{n-1} F_i = F_{n+1}$. From (2), we have $f_{n+2} = f_{n+1}f_n$, so that f_n has an occurrence in f starting at position $|f_{n+1}| = F_{n+1}$. □

Let us consider the sequence p_n of the palindromic prefixes of f, also called *central words*. The first few values of the sequence p_n are displayed in Table 4, where ε denotes the empty word, i.e., the word of length 0.

$$\begin{aligned} &p_3 = \varepsilon \\ &p_4 = 0 \\ &p_5 = 010 \\ &p_6 = 010010 \\ &p_7 = 01001010010 \\ &p_8 = 0100101001001010010 \\ &p_9 = 01001010010010100101001001010010 \end{aligned}$$

Table 4: The first few central words.

As it is well-known, for every $n \geq 3$, p_n is obtained from f_n by removing the last two letters. More precisely, we have for every $n \geq 1$,

$$f_{2n+1} = p_{2n+1}01, \qquad f_{2n+2} = p_{2n+2}10. \tag{4}$$

The fundamental property of the central words is the following:

Lemma 3. *For every $n \geq 2$ one has*

$$p_{2n+1} = p_{2n-1}01p_{2n} = p_{2n}10p_{2n+1}, \qquad p_{2n+2} = p_{2n}10p_{2n+1} = p_{2n+1}01p_{2n}.$$

Proof. Follows immediately from (2) and (4). □

Remark 4. It is easy to see from (2) that for every $n \geq 4$, one has $f_n = f_{n-2}f_{n-3}f_{n-2}$. We have therefore from (3):

$$\begin{aligned} f &= 01001 \prod_{n\geq 2} f_n f_{n-1} f_n & (5)\\ &= 01001 \cdot (0 \cdot 1 \cdot 0)(01 \cdot 0 \cdot 01)(010 \cdot 01 \cdot 010)(01001 \cdot 010 \cdot 01001) \cdots \end{aligned}$$

Analogously, since $1 = f_3$, we can write

$$\begin{aligned} f &= 0100 \prod_{n\geq 2} f_{n-1} f_n f_{n-1} & (6)\\ &= 0100 \cdot (1 \cdot 0 \cdot 1)(0 \cdot 01 \cdot 0)(01 \cdot 010 \cdot 01)(010 \cdot 01001 \cdot 010) \cdots \end{aligned}$$

We now introduce a class of words that we call the *co-Fibonacci words*. Although this class has appeared previously in the literature [2], to the best of our knowledge no one has yet given a name to them.

The co-Fibonacci words f'_n are defined by complementing the last two letters in the Fibonacci words f_n, that is, $f'_n = p_n yx$, where x and y are the letters such that $f_n = p_n xy$. Equivalently, co-Fibonacci words can be defined by $f'_n = f_{n-2}f_{n-1}$ for every $n \geq 3$. The first few co-Fibonacci words are displayed in Table 5.

The following lemma is a direct consequence of Lemma 3.

Lemma 5. *For every $n \geq 2$ one has*

$$f'_{2n+1} = f_{2n}f'_{2n-1}, \qquad f_{2n+2} = f_{2n}f'_{2n+1}.$$

Proposition 6. *The Fibonacci word can be obtained by concatenating* 0 *and the odd co-Fibonacci words:*

$$\begin{aligned} f &= 0 \prod_{n\geq 1} f'_{2n+1} & (7)\\ &= 0 \cdot 10 \cdot 01010 \cdot 0100101001010 \cdots \end{aligned}$$

$$
\begin{aligned}
f'_3 &= 10 \\
f'_4 &= 001 \\
f'_5 &= 01010 \\
f'_6 &= 01001001 \\
f'_7 &= 0100101001010 \\
f'_8 &= 010010100100101001010
\end{aligned}
$$

Table 5: The first few co-Fibonacci words.

Proof. Follows directly from (3) replacing $f_{2n-1}f_{2n}$ with f'_{2n+1}. $\square$

Analogously, we have the following:

Proposition 7. *The Fibonacci word can be obtained by concatenating* 01 *and the even co-Fibonacci words:*

$$
\begin{aligned}
f &= 01\prod_{n\geq 1} f'_{2n+2} \qquad (8)\\
&= 01 \cdot 001 \cdot 01001001 \cdot 010010100100101001010 \cdots
\end{aligned}
$$

Proof. Follows directly from (3) replacing $f_{2n}f_{2n+1}$ with f'_{2n+2}. $\square$

3 Singular words

Let us define the *left rotation* of a non-empty word $w = w_1w_2\cdots w_n$, w_i letters, as the word $w^{\lambda} = w_nw_1\cdots w_{n-1}$. Analogously, the *right rotation* of w is defined as the word $w^{\rho} = w_2\cdots w_nw_1$.

The *singular words* $\hat{f}_n$ are defined by complementing the first letter in the left rotations of the Fibonacci words f_n. The first few singular words are displayed in Table 6. Note that for every $n \geq 1$, one has $\hat{f}_{2n+1} = 0p_{2n+1}0$ and $\hat{f}_{2n+2} = 1p_{2n+2}1$.

The singular words are palindromic factors of f but do not appear as prefixes of f (by the way, f also contains other palindromic factors besides the central words p_n and the singular words $\hat{f}_n$, e.g., 1001 or 01010, see [6] for more details). Their name comes from the fact that among the $F_n + 1$ factors of f of length F_n, there are F_n of them that can be obtained one from each other by iteratively applying left (or equivalently right) rotation and one, the singular word, whose left (or equivalently right) rotation is not a factor of f.

Wen and Wen [12] proved that the Fibonacci infinite word can be obtained by concatenating the singular words:

$$
\begin{aligned}
\hat{f}_1 &= 0\\
\hat{f}_2 &= 1\\
\hat{f}_3 &= 00\\
\hat{f}_4 &= 101\\
\hat{f}_5 &= 00100\\
\hat{f}_6 &= 10100101
\end{aligned}
$$

Table 6: The first few singular words.

Proposition 8. *The Fibonacci infinite word is the concatenation of the singular words:*

$$
\begin{aligned}
f &= \prod_{n\geq 1} \hat{f}_n \qquad (9)\\
&= 0 \cdot 1 \cdot 00 \cdot 101 \cdot 00100 \cdot 10100101 \cdots
\end{aligned}
$$

Proof. Indeed, (9) follows directly from (3) and the definition of singular words, observing that the Fibonacci words end by letter 0 and 1 alternatingly. □

The factorization (9) is in fact the Lempel-Ziv factorization of f. The Lempel-Ziv factorization is a factorization widely used in computer science for compressing strings [14]. The Lempel-Ziv factorization of a word w is $w = w_1w_2\cdots$ where w_1 is the first letter of w and for every $i \geq 2$, w_i is the shortest prefix of $w_iw_{i+1}\cdots$ that occurs only once in the word $w_1w_2\cdots w_i$. Roughly speaking, at each step one searches for the shortest factor that did not appear before.

Remark 9. It is easy to see, using for example (2), (4) and the definition of singular words, that for every $n \geq 4$, $\hat{f}_n = \hat{f}_{n-2}\hat{f}_{n-3}\hat{f}_{n-2}$. Therefore, from (9), we have

$$
\begin{aligned}
f &= 0100 \prod_{n\geq 2} \hat{f}_n\hat{f}_{n-1}\hat{f}_n \qquad (10)\\
&= 0100 \cdot (1 \cdot 0 \cdot 1)(00 \cdot 1 \cdot 00)(101 \cdot 00 \cdot 101)(00100 \cdot 101 \cdot 00100) \cdots
\end{aligned}
$$

Since $0 = \hat{f}_1$, we hence obtain

$$
\begin{aligned}
f &= 010 \prod_{n\geq 2} \hat{f}_{n-1}\hat{f}_n\hat{f}_{n-1} \qquad (11)\\
&= 010 \cdot (0 \cdot 1 \cdot 0)(1 \cdot 00 \cdot 1)(00 \cdot 101 \cdot 00)(101 \cdot 00100 \cdot 101) \cdots
\end{aligned}
$$

The factorization (11) is a sort of dual with Lucas numbers of the factorization in singular words (9). Indeed, the sequence of factor lengths in (9) is the sequence of Fibonacci numbers, while if in (11) one decomposes the first term as $01 \cdot 0$, then the sequence of factor lengths is the sequence of Lucas numbers A000032: 2,1,3,4,7,11, etc.

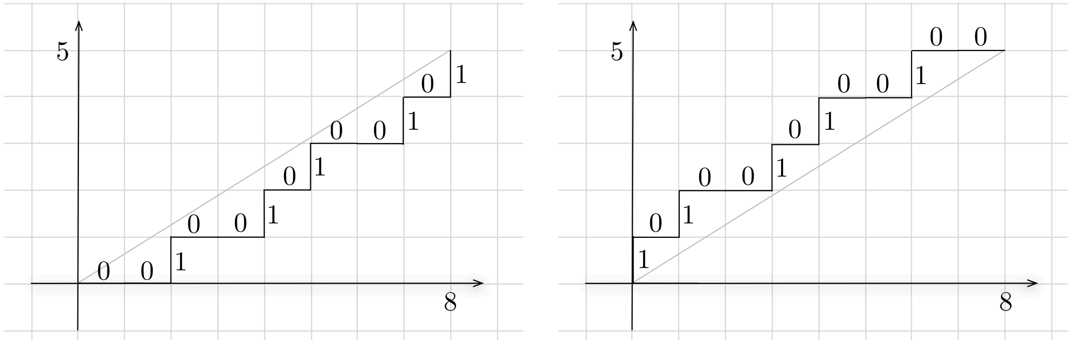

Figure 1: The lower Christoffel word $c_7 = 0010010100101$ (left) and the upper Christoffel word $\widetilde{c_7} = 1010010100100$ (right) are the best grid approximations, respectively from above and from below, of the Euclidean segment joining the points $(0,0)$ and $(8,5) = (F_6, F_5)$.

4 Christoffel words

The *lower Christoffel words* are defined by $c_n = 0p_n1$, for every $n \geq 3$. The lower Christoffel words are the Lyndon factors of f, i.e., they are lexicographically smaller than any of their proper suffixes (with respect to the order induced by $0 < 1$).

$$
\begin{aligned}
c_3 &= 01 \\
c_4 &= 001 \\
c_5 &= 00101 \\
c_6 &= 00100101 \\
c_7 &= 0010010100101 \\
c_8 &= 001001010010010100101
\end{aligned}
$$

Table 7: The first few lower Christoffel words.

If in the Euclidean plane one interprets each 0 by a horizontal unitary step and each 1 with a vertical unitary step, the lower Christoffel word c_n is the best grid approximation from below of the segment joining the point $(0,0)$ to the point (F_{n-1}, F_{n-2}) (see Figure 1).

Analogously, the *upper Christoffel words* are defined by $\widetilde{c_n} = 1p_n0$, for every $n \geq 3$. Therefore, the upper Christoffel words are the reversals of the lower Christoffel words (we use the notation $\widetilde{w}$ for the reversal, a.k.a. mirror image, of the word w). The upper Christoffel words are the anti-Lyndon factors of f, i.e., they are lexicographically greater than any of their proper suffixes (with respect to the order induced by $0 < 1$). Moreover, $\widetilde{c_n}$ is the best grid approximation from above of the segment joining the point $(0,0)$ to the point (F_{n-1}, F_{n-2}).

Remark 10. For every $n \geq 3$, $c_n = f_n^\lambda$ if n is even, $c_n = \widetilde{f_n^\lambda}$ if n is odd. Therefore, $\widetilde{c_n} = \widetilde{f_n^\lambda}$ if n is even, $\widetilde{c_n} = f_n^\lambda$ if n is odd.

$$\begin{aligned}
\widetilde{c_3} &= 10 \\
\widetilde{c_4} &= 100 \\
\widetilde{c_5} &= 10100 \\
\widetilde{c_6} &= 10100100 \\
\widetilde{c_7} &= 1010010100100 \\
\widetilde{c_8} &= 101001010010010100100
\end{aligned}$$

Table 8: The first few upper Christoffel words.

Lemma 11. *For every $n \geq 2$ one has*

$$c_{2n+1} = c_{2n}c_{2n-1}, \qquad c_{2n+2} = c_{2n}c_{2n+1},$$

and therefore

$$\widetilde{c_{2n+1}} = \widetilde{c_{2n-1}}\widetilde{c_{2n}}, \qquad \widetilde{c_{2n+2}} = \widetilde{c_{2n+1}}\widetilde{c_{2n}}.$$

Proof. The first part follows from Lemma 5 by applying the right rotation to each side of the equalities. The second part follows from the first by applying the reversal. ☐

The following result states that every Christoffel word is the product of two singular words.

Lemma 12. *For every $n \geq 1$ one has*

- $c_{2n+1} = \hat{f}_{2n-1}\hat{f}_{2n}$
- $c_{2n+2} = \hat{f}_{2n+1}\hat{f}_{2n}$

and therefore

- $\widetilde{c_{2n+1}} = \hat{f}_{2n}\hat{f}_{2n-1}$
- $\widetilde{c_{2n+2}} = \hat{f}_{2n}\hat{f}_{2n+1}$

Proof. Follows directly from Lemma 3 and the definitions of Christoffel and singular words. ☐

Melançon [9] proved that the Fibonacci word is the concatenation of the odd lower Christoffel words:

Proposition 13. *The Fibonacci word is the concatenation of the odd lower Christoffel words:*

$$\begin{aligned} f &= \prod_{n\geq 1} c_{2n+1} \qquad (12) \\ &= 01\cdot 00101\cdot 0010010100101\cdots \end{aligned}$$

Proof. Follows directly from (9) and Lemma 12. □

Actually, Melançon proved that (12) is precisely the Lyndon factorization of f. Recall that the Lyndon factorization of a word w is $w = \ell_1\ell_2\cdots$, where each ℓ_i is a Lyndon word and is lexicographically greater than or equal to ℓ_{i+1}. The uniqueness of such a factorization for finite words is a well-known theorem of Chen, Fox and Lyndon [4]. Siromoney et al. extended this factorization to infinite words [11].

Symmetrically, we have the following:

Proposition 14. *The Fibonacci word is the concatenation of* 0 *and the even upper Christoffel words:*

$$\begin{aligned} f &= 0\prod_{n\geq 2} \widetilde{c_{2n}} \qquad (13) \\ &= 0\cdot 100\cdot 10100100\cdot 101001010010010100100\cdots \end{aligned}$$

Proof. Follows directly from (9) and Lemma 12. □

In fact, it is easy to see that (13) is the Lyndon factorization of f if one takes the order induced by $1 < 0$.

We now present two other factorizations based on Christoffel words. To the best of our knowledge, these factorizations did not appear before in literature.

Proposition 15. *The Fibonacci word is the concatenation of* 010 *and the lower Christoffel words where each odd lower Christoffel word is squared:*

$$\begin{aligned} f &= 010\prod_{n\geq 1} c_{2n+1}^2 c_{2n+2} \qquad (14) \\ &= 010\cdot(01\cdot 01\cdot 001)(00101\cdot 00101\cdot 00100101)\cdots \end{aligned}$$

Proof. Follows directly from (6) and Lemma 12. Indeed, by Lemma 12, we have

$$c_{2n+1}c_{2n+1}c_{2n+2} = \hat{f}_{2n-1}\hat{f}_{2n}\cdot\hat{f}_{2n-1}\hat{f}_{2n}\cdot\hat{f}_{2n+1}\hat{f}_{2n} = (\hat{f}_{2n-1}\hat{f}_{2n}\hat{f}_{2n-1})(\hat{f}_{2n}\hat{f}_{2n+1}\hat{f}_{2n}).$$

□

Analogously, we have the following:

Proposition 16. *The Fibonacci word is the concatenation of* 0100 *and the upper Christoffel words where each even upper Christoffel word is squared:*

$$\begin{aligned} f &= 0100 \prod_{n\geq 1} \widetilde{c_{2n+1}}\,\widetilde{c_{2n+2}}^{\,2} && (15)\\ &= 0100 \cdot (10 \cdot 100 \cdot 100)(10100 \cdot 10100100 \cdot 10100100) \cdots \end{aligned}$$

Proof. Follows directly from (5) and Lemma 12. Indeed, by Lemma 12, we have

$$\widetilde{c_{2n+1}}\,\widetilde{c_{2n+2}}\,\widetilde{c_{2n+2}} = \hat{f}_{2n}\hat{f}_{2n-1} \cdot \hat{f}_{2n}\hat{f}_{2n+1} \cdot \hat{f}_{2n}\hat{f}_{2n+1} = (\hat{f}_{2n}\hat{f}_{2n-1}\hat{f}_{2n})(\hat{f}_{2n+1}\hat{f}_{2n}\hat{f}_{2n+1}).$$

□

5 Reversals of Fibonacci words

One of the most known factorizations of the Fibonacci infinite word, and perhaps the most surprising, is the following.

Proposition 17. *The Fibonacci word can be obtained also by concatenating the reversals of the Fibonacci words:*

$$\begin{aligned} f &= \prod_{n\geq 2} \widetilde{f_n} && (16)\\ &= 0 \cdot 10 \cdot 010 \cdot 10010 \cdot 01010010 \cdots \end{aligned}$$

Proof. It follows from the definitions that taking the right rotation of $\widetilde{f_n}$ and complementing the last letter produces the n-th singular word $\hat{f}_n$. Therefore, (16) follows directly from (9) observing that the reversals of the Fibonacci words start with 0 and 1 alternatingly. □

The factorization (16) is basically the Crochemore factorization of f—the only difference is that the Crochemore factorization starts with 0, 1, 0 and then coincides with the one above (see [3]). Recall that the Crochemore factorization of w is $w = c_1c_2\cdots$ where c_1 is the first letter of w and for every $i > 1$, c_i is either a fresh letter or the longest prefix of $c_ic_{i+1}\cdots$ occurring twice in $f_1f_2\cdots f_i$. For example, the Crochemore factorization of the word $w = 0101001$ is $0 \cdot 1 \cdot 010 \cdot 01$, since 010 occurs twice in 01010.

In 1995, de Luca [5] considered the following factorization:

Proposition 18. *The Fibonacci word can be obtained by concatenating the reversals of the even Fibonacci words.*

$$\begin{aligned} f &= \prod_{n\geq 2} \widetilde{f_{2n}} && (17)\\ &= 010 \cdot 01010010 \cdot 0101001010010010100 10 \cdots \end{aligned}$$

Proof. Applying the reversal to (2), we have that $\widetilde{f_n} = \widetilde{f_{n-2}}\widetilde{f_{n-1}}$, for every $n > 2$. So (17) follows directly from (16) by replacing $\widetilde{f_{2n-2}}\widetilde{f_{2n-1}}$ with $\widetilde{f_{2n}}$. □

In [5] de Luca proved that the factorization (17) has the following minimal property with respect to the lexicographical order: any non-trivial permutation of a finite number of the factors will produce an infinite word that is lexicographically greater than f.

Concatenating the reversals of the odd Fibonacci words instead of even ones still produces the Fibonacci word, if one prepends a 0:

Proposition 19. *The Fibonacci word can be obtained by concatenating* 0 *and the reversals of the odd Fibonacci words:*

$$\begin{aligned} f &= 0\prod_{n\geq 2}\widetilde{f_{2n+1}} \\ &= 0\cdot 10010\cdot 1001001010010\cdots \end{aligned} \tag{18}$$

Proof. Follows directly from (16) by replacing $\widetilde{f_{2n-1}}\widetilde{f_{2n}}$ with $\widetilde{f_{2n+1}}$. □

Recently [7], studying the so-called *open* and *closed* words, the following factorization has been proved:

Proposition 20. *The Fibonacci word can be obtained by concatenating* 01 *and the squares of the reversals of the Fibonacci words:*

$$\begin{aligned} f &= 01\prod_{n\geq 2}(\widetilde{f_n})^2 \\ &= 01\cdot(0\cdot 0)(10\cdot 10)(010\cdot 010)(10010\cdot 10010)\cdots \end{aligned} \tag{19}$$

Proof. Recalling that for every $n \geq 3$, one has $\widetilde{f_n} = \widetilde{f_{n-2}}\widetilde{f_{n-1}}$, we have, from (16), that $f = \widetilde{f_2}\widetilde{f_3}\widetilde{f_4}\cdots = 0\cdot\widetilde{f_1}\widetilde{f_2}\widetilde{f_2}\widetilde{f_3}\cdots = 01\cdot(\widetilde{f_2}\widetilde{f_2})(\widetilde{f_3}\widetilde{f_3})\cdots = 01\prod_{n\geq 2}(\widetilde{f_n})^2$. □

6 Generalization to standard Sturmian words

The Fibonacci word is the most prominent example of a *standard Sturmian word*. Let α be an irrational number such that $0 < \alpha < 1$, and let $[0; d_1 + 1, d_2, d_3, \ldots]$ be the continued fraction expansion of α. The sequence of words defined by $s_1 = 1$, $s_2 = 0$ and $s_n = s_{n-1}^{d_{n-2}}s_{n-2}$ for $n \geq 3$, converges to the infinite word w_α, called the standard Sturmian word of slope α. The sequence of words s_n is called the standard sequence of w_α. The Fibonacci word is the standard Sturmian word of slope $\alpha = (3 - \sqrt{5})/2$ and its standard sequence is the sequence of Fibonacci finite words, since one has $d_i = 1$ for every $i \geq 1$.

Most of the factorizations we described in this note can be generalized to any standard Sturmian word. However, the proofs become more technical and less easy to present.

References

[1] J. Berstel, Fibonacci words, a survey. In G. Rozenberg and A. Salomaa, editors, *The Book of L*, Springer-Verlang, 1985, pp. 11–25.

[2] J. Berstel, On the index of Sturmian words. In Juhani Karhumäki, Hermann Maurer, Gheorghe Păun, and Grzegorz Rozenberg, editors, *Jewels are Forever*, Springer, 1999, pp. 287–294.

[3] J. Berstel and A. Savelli, Crochemore factorization of Sturmian and other infinite words, In *Mathematical Foundations of Computer Science*, Vol. 4162 of *Lecture Notes in Computer Science*, Springer, 2006, pp. 157–166.

[4] K. T. Chen, R. H. Fox, and R. C. Lyndon, Free differential calculus, IV — the quotient groups of the lower central series, *Ann. Math.* **68** (1958), 81–95.

[5] A. de Luca, A division property of the Fibonacci word, *Inform. Process. Letters* **54** (1995), 307–312.

[6] A. de Luca and A. De Luca, Combinatorial properties of Sturmian palindromes, *Internat. J. Found. Comput. Sci.* **17** (2006), 557–573.

[7] A. De Luca and G. Fici, Open and closed prefixes of Sturmian words, In *Proceedings of the 9th International Conference on Words*, Vol. 8079 of *Lecture Notes in Computer Science*, Springer, 2013, pp. 132–142.

[8] C. G. Lekkerkerker, Voorstelling van natuurlijke getallen door een som van Fibonacci, *Simon Stevin* **29** (1951–52), 190–195.

[9] G. Melançon, Lyndon words and singular factors of Sturmian words, *Theoret. Comput. Sci.* **218** (1999), 41–59.

[10] A. Ostrowski, Bemerkungen zur theorie der diophantischen approximationen, *Abh. Math. Sem. Hamburg* **1** (1922), 250–251. Reprinted in Collected Mathematical Papers, Vol. 3, pp. 57–80.

[11] R. Siromoney, L. Matthew, V. R. Dare, and K. G. Subramanian, Infinite Lyndon words, *Inform. Process. Letters* **50** (1994), 101–104.

[12] Z.-X. Wen and Z.-Y. Wen, Some properties of the singular words of the Fibonacci word, *European J. Combin.* **15** (1994), 587–598.

[13] É. Zeckendorf, Reprsentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas, *Bull. Soc. Roy. Sci. Liège* **41** (1972), 179–182.

[14] J. Ziv and A. Lempel, A universal algorithm for sequential data compression, *IEEE Trans. Inf. Theor.* **23**(3) (1977), 337–343.

2010 *Mathematics Subject Classification*: Primary 68R15.
Keywords: Fibonacci word; Zeckendorf representation; Lyndon factorization; Lempel-Ziv factorization; Crochemore factorization.

(Concerned with sequences A000032 and A000045.)

Received July 30 2015; revised versions received August 3 2015; August 6 2015. Published in *Journal of Integer Sequences*, August 7 2015.

Return to Journal of Integer Sequences home page.