The sum of digits functions of the Zeckendorf and the base phi expansions

Michel Dekking, F.

DOI
10.1016/j.tcs.2021.01.011

Publication date
2021

Document Version
Final published version

Published in
Theoretical Computer Science

Theoretical Computer Science 859 (2021) 70–79

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

The sum of digits functions of the Zeckendorf and the base phi expansions

F. Michel Dekking

DIAM, Delft University of Technology, Faculty EEMCS, P.O. Box 5031, 2600 GA Delft, the Netherlands

ARTICLE INFO

Article history:
Received 6 August 2020
Received in revised form 23 November 2020
Accepted 5 January 2021
Available online 8 January 2021
Communicated by M. Sciortino

Keywords:
Zeckendorf expansion
Base phi
Wythoff sequence
Fibonacci word
Generalized Beatty sequence

ABSTRACT

We consider the sum of digits functions for both base phi, and for the Zeckendorf expansion of the natural numbers. For both sum of digits functions we present morphisms on infinite alphabets such that these functions viewed as infinite words are letter-to-letter projections of fixed points of these morphisms. We characterize the first differences of both functions a) with generalized Beatty sequences, or unions of generalized Beatty sequences, and b) with morphic sequences.

1. Introduction

Perhaps the most famous word in language theory is the Thue-Morse word

$$s_2 := 0110100110010110\ldots,$$

fixed point of the morphism $\mu: 0 \to 01, 1 \to 10$. Here a morphism is a map from infinite words to infinite words that preserves the concatenation operation on the set of words. The remarkable property of s_2 is that it can also be obtained from binary expansions: $s_2(N)$ gives the parity of the sum of digits in the binary expansion of the natural number N. The sum of digits in base 2 written as an infinite word equals

$$s_{\rm TM} := 0112122312232334,\ldots,$$

and $s_{\rm TM}$ is fixed point of the morphism $j \to j, j+1$ on the infinite alphabet $\{0, 1, 2, \ldots\}$. The reason for this is simple: the number $2N$ has the same number of digits as N, and $2N+1$ has one more digit than N.

The question arises: do similar connections to language theory hold for expansions in other bases?
For expansions in integer bases b, b a natural number, it is not hard to establish that the answer is positive.
In this paper we consider the case where the powers of 2 are replaced by the Fibonacci numbers (the Zeckendorf expansion), respectively the powers of the golden mean $\varphi = (1+\sqrt{5})/2$ (the base phi expansion).

For both expansions we give in Theorem 3, respectively Theorem 11 a morphism on an infinite alphabet, such that the sum of digits functions of these expansions considered as infinite words are letter-to-letter projections of fixed points of

E-mail address: F.M.Dekking@TUDelft.nl.

https://doi.org/10.1016/j.tcs.2021.01.011

these morphisms. This means that the sum of digits functions are morphic words, defined in general as letter-to-letter projections of fixed points of morphisms.

We then will show how these results permit to give precise information on the first differences of the sum of digits functions. The *first differences* of a function $f : \mathbb{N}_0 \to \mathbb{N}_0$ are given by the function Δf defined by

$$\Delta f(N) = f(N+1) - f(N), \quad \text{for } N = 0, 1, 2 \ldots.$$

We shall focus on the signs of Δf. A number N is called a *point of increase* of a function $f : \mathbb{N}_0 \to \mathbb{N}_0$ if $\Delta f(N) > 0$. It is called a *point of constancy* if $\Delta f(N) = 0$, and a *point of decrease* if $\Delta f(N) < 0$.

For base two it is simple to see that the points of increase of the sum of digits function $s_{\rm TM}$ are given by the even numbers, and that the points of constancy and decrease are given by the numbers 1 mod 4, respectively 3 mod 4.

We will prove (Theorem 4 and Theorem 12) for both the Zeckendorf representation and the base phi expansion that the points of increase, constancy and decrease are all given by unions of generalized Beatty sequences, as studied in [1]. These are sequences V of the type

$$V(n) = p\lfloor n\alpha\rfloor + qn + r, \quad n \ge 1,$$

where α is a real number, and p, q, and r are integers. Here we denoted the floor function by $\lfloor\cdot\rfloor$.

We will also prove that the first differences of the sequences of points of increase, constancy and decrease are all morphic sequences. See Theorem 5 for the Zeckendorf representation, and Theorem 13 for the base phi expansion.

A prominent role in this paper, both for base phi and the Zeckendorf expansion, is played by $(\lfloor n\varphi\rfloor)$, the well known lower Wythoff sequence.

A standard result (see, e.g., [14]) is that the sequence $\Delta(\lfloor n\varphi\rfloor)$ is equal to the Fibonacci word $x_{1,2} = 1211212112\ldots$ on the alphabet $\{1, 2\}$, i.e., the unique fixed point of the morphism $1 \to 12, 2 \to 1$. More generally, we have the following simple lemma.

Lemma 1. ([1]) *Let $V = (V(n))_{n\ge1}$ be the generalized Beatty sequence defined by $V(n) = p\lfloor n\varphi\rfloor + qn + r$, and let ΔV be the sequence of its first differences. Then ΔV is the Fibonacci word on the alphabet $\{2p+q,\ p+q\}$. Conversely, if $x_{a,b}$ is the Fibonacci word on the alphabet $\{a, b\}$, then any V with $\Delta V = x_{a,b}$ is a generalized Beatty sequence $V = ((a-b)\lfloor n\varphi\rfloor) + (2b-a)n + r)$ for some integer r.*

Let $A(n) = \lfloor n\varphi\rfloor$, and $B(n) = \lfloor n\varphi^2\rfloor$. It is well known that A and B form a pair of Beatty sequences, i.e., they are disjoint with union $\mathbb{N}$. In the next lemma, VA is the composition given by $VA(n) = V(A(n))$.

Lemma 2. ([1]) *Let V be a generalized Beatty sequence given by $V(n) = p\lfloor n\varphi\rfloor + qn + r$, $n \ge 1$. Then VA and VB are generalized Beatty sequences with parameters $(p_{VA}, q_{VA}, r_{VA}) = (p+q, p, r-p)$ and $(p_{VB}, q_{VB}, r_{VB}) = (2p+q, p+q, r)$.*

2. The Zeckendorf sum of digits function

Let $F_0 = 0, F_1 = 1, F_2 = 1, \ldots$ be the Fibonacci numbers. Ignoring leading and trailing zeros, any natural number N can be written uniquely with digits $d_i = 0$ or 1, as

$$N = \sum_{i\ge0} d_i F_{i+2},$$

where $d_i d_{i+1} = 11$ is not allowed. We denote the Zeckendorf expansion of N as $Z(N)$, with digits $d_i(N)$.

Let $s_{\rm Z}$ be the sum of digits of such an expansion: for $N \ge 0$

$$s_{\rm Z}(N) = \sum_{i\ge0} d_i(N).$$

We have

$$(s_{\rm Z}(N)) = (0, 1, 1, 1, 2, 1, 2, 2, 1, 2, 2, 2, 3, 1, 2, 2, 2, 3, 2, 3, 3, 1, 2, 2, 2, \ldots)$$

Our first result is that $s_{\rm Z}$ is a morphic sequence. The alphabet will consist of symbols $\binom{j}{0}$ and $\binom{j}{1}$. Note that in the theorem $\binom{j}{0}\binom{j+1}{1}$ is a word of length 2 over this alphabet.

Theorem 3. *The function $s_{\rm Z}$, as a sequence, is a morphic sequence on an infinite alphabet, i.e., $(s_{\rm Z}(N))$ is a letter to letter projection of a fixed point of a morphism τ. The alphabet is $\{0, 1, \ldots, j, \ldots\} \times \{0, 1\}$, and τ is the morphism given by*

$$\tau(\binom{j}{0}) = \binom{j}{0}\binom{j+1}{1},$$
$$\tau(\binom{j}{1}) = \binom{j}{0}.$$

The letter-to-letter map is given by the projection on the first coordinate: $\binom{j}{i} \to j$ *for* $i = 0, 1$. *The fixed point* x_τ *of* τ *with initial symbol* $\binom{0}{0}$ *projected on the first coordinate equals* $(s_Z(N))$.

Proof. See the Comments of sequence A007895 in [15] for a proof of this. □

Let I_Z, C_Z and D_Z be the functions listing the points of increase, constancy, and decrease of the function s_Z. We have[1]

$$I_Z = (0, 3, 5, 8, 11, 13, 16, \dots),\ C_Z = (1, 2, 6, 9, 10, 14, \dots),\ D_Z = (4, 7, 12, 17, 20, 25, \dots).$$

To state our results it is actually convenient to define $D_Z = (-1, 4, 7, 12, 17, 20, 25, \dots)$.

When (a_n) and (b_n) are two increasing sequences, indexed by $\mathbb{N}$, then we mean by the union of (a_n) and (b_n) the increasing sequence whose terms go through the set $\{a_n, b_n : n \in \mathbb{N}\}$.

Theorem 4. *The function* I_Z*, the points of increase of the function* s_Z*, is given for* $n = 1, 2, \dots$ *by*

$$I_Z(n) = \lfloor n\varphi \rfloor + n - 2.$$

The function C_Z*, the points of constancy of the function* s_Z*, is given for* $n = 1, 2, \dots$ *by the union of the two generalized Beatty sequences with terms*

$$2\lfloor n\varphi \rfloor + n - 2 \quad \text{and} \quad 3\lfloor n\varphi \rfloor + 2n - 3.$$

The function D_Z*, the points of decrease of the function* s_Z*, is given for* $n = 1, 2 \dots$ *by*

$$D_Z(n) = 2\lfloor n\varphi \rfloor + n - 4.$$

Proof. Let I_Z be the sequence of the points of increase of the function s_Z.

Projection on the second coordinate of τ yields the Fibonacci morphism σ_F given by

$$\sigma_F(0) = 01,\ \sigma_F(1) = 0.$$

Thus the second coordinates of the fixed point of τ equal the infinite Fibonacci word $x_{0,1}$ = 0100101001001.... Obviously, the increase points of s_Z occur if and only if the word $(j, 0)(j+1, 1)$ occurs in the fixed point x_τ of τ if and only if the word 01 occurs in $x_{0,1}$. Since 11 does not occur in $x_{0,1}$, this means that we have to shift the positions of 1's in $x_{0,1}$ by 1. It is well known that the positions of 1 are given by the upper Wythoff sequence $(\lfloor n\varphi^2 \rfloor) = (\lfloor n\varphi \rfloor + n)$. Since the first coordinate of the fixed point of τ starts from index 0, and the second from index 1, we have to replace n by $n+1$, and this yields the first result of Theorem 4.

The points of constancy are more difficult to characterize with the fixed point x_τ than the points of increase. We therefore take another approach. Write $Z(N) = \dots w$, where w is a word of length 4. Then w can be any word of the 0-1-words of length 4 containing no 11. Obviously, the three words $w = 0000$, $w = 0100$ and $w = 1000$ give points of increase.

Furthermore the numbers N with $Z(N)$ ending in $w = 0001, 1001$ and $w = 0010$ give

$$Z(N) = \dots 001 \Rightarrow Z(N+1) \doteq \dots 002 \doteq \dots 010, \quad Z(N) = \dots 0010 \Rightarrow Z(N+1) \doteq \dots 0011 \doteq \dots 0100.$$

We see that these give points of constancy.

Finally, we show that the N with $Z(N)$ having suffix $w = 0101$ or $w = 1010$ give points of decrease. In the following two computations the $\doteq$-sign indicates that we use also non-admissible Zeckendorf representations.

$$Z(N) = \dots 0101 \quad \Rightarrow \quad Z(N+1) \doteq \dots 0102 \doteq \dots 0110 \doteq \dots 1000,$$
$$Z(N) = \dots 01010 \quad \Rightarrow \quad Z(N+1) \doteq \dots 01011 \doteq \dots 01100 \doteq \dots 10000.$$

In both cases at least one digit 1 is lost, so these N are the points of decrease.

With this knowledge we can apply Theorem 2.3 and Proposition 2.8 in the paper [9], obtaining that one part of I_Z is given by the generalized Beatty sequence $(2\lfloor n\varphi \rfloor + n - 2)$ and the other part is given by $(3\lfloor n\varphi \rfloor + 2n - 3)$.

[1] I_Z is the sequence A026274 in [15].

Again from Theorem 2.3 and Proposition 2.8 in the paper [9], we obtain that $(D_Z(n+1))$ is the union of the two generalized Beatty sequences $(3\lfloor n\varphi\rfloor + 2n - 1)$ and $(5\lfloor n\varphi\rfloor + 3n - 1)$.

It is not a simple matter to see that this union is given by the single generalized Beatty sequence $(2\lfloor n\varphi\rfloor + n - 4)$, where the index starts at $n = 2$.

Let us write $V(p,q,r) = (p\lfloor n\varphi\rfloor + qn + r)_{n\ge 1}$. We have proved so far that $I_Z = V(1,1,-2)$, and C_Z is the union of $V(2,1,-2)$ and $V(3,2,-3)$. If we add 2 to all terms of these sequences, we obtain the three sequences $V(1,1,0)$, $V(2,1,0)$, and $V(3,2,-1)$.

The triple of sequences

$$\{V(1,1,0), V(2,1,0), V(1,1,-1)\}$$

is known as the 'first classical complementary triple', i.e., these are three disjoint sequences with union $\mathbb{N}$. See page 334 in [1]. The third sequence of this triple, $V(1,1,-1)$, can be written as a disjoint union of the two sequences $V(3,2,-1)$ and $V(2,1,-2)$, by Lemma 2. Thus

$$\{V(1,1,0), V(2,1,0), V(3,2,-1), V(2,1,-2)\}$$

forms a complementary quadruple. If we subtract 2 from all terms of these four sequences, the first gives I_Z, the second and the third together, C_Z. Since $\{I_Z, C_Z, D_Z\}$ is a complementary triple, with union $\{-1,0,1,2,\dots\}$ this implies that $(D_Z(n+1))$ has to be equal to $V(2,1,-4)$. □

Next, we give a characterization of I_Z, C_Z and D_Z in terms of morphisms.

Theorem 5. *The points of increase of the function s_Z are given by the sequence I_Z, which has $I_Z(1) = 0$, and ΔI_Z is the fixed point of the Fibonacci morphism $3 \to 32,\ 2 \to 3$.*

The points of constancy of the function s_Z are given by the sequence C_Z, which has $C_Z(1) = 1$, and ΔC_Z is the fixed point of the 2-block Fibonacci morphism on the alphabet $\{1,4,3\}$ given by $1 \to 14,\ 3 \to 14,\ 4 \to 3$.

The points of decrease of the function s_Z are given by the sequence D_Z, which has $D_Z(1) = -1$, and ΔD_Z is the fixed point of the Fibonacci morphism $5 \to 53,\ 3 \to 5$.

For the proof of Theorem 5 we have to make some preparations. Let $\Lambda_3 := \{2\}$, $\Psi_3 := \{0,1\} =: [0,1]$, and define for $n \ge 4$ the intervals of integers Λ_n and Ψ_n by

$$\Lambda_n := [F_n, F_{n+1} - 1],\ \ \Psi_n := [0, F_n - 1].$$

The (Λ_n) form a partition of $\mathbb{N}_0 \setminus \{0,1\}$, and the (Ψ_n) satisfy

$$\Psi_{n+1} = \Psi_n \cup \Lambda_n. \tag{1}$$

For an interval I, let $C_Z(I)$ denote the points of increase lying in the interval I. Also, let $\Delta C_Z(I)$ denote the first differences of the points of increase lying in the interval I, considered as a word on the alphabet $\{1,2,3,4\}$. At first sight, the latter definition is problematic, as one has to know the first point of increase after the last element of $C_Z(I)$. However, we shall only consider intervals $I = \Lambda_n$ and $I = \Psi_{n+1}$, which both are followed by Λ_{n+1}, and one verifies easily that the first point of increase in Λ_{n+1} is always the second point. Actually, this follows directly from the following lemma.

Lemma 6. *For all $n \ge 3$ one has $C_Z(\Lambda_{n+1}) = C_Z(\Psi_n) + F_{n+1}$.*

Proof. We used the notation $A + y = \{x + y : x \in A\}$ for a set A, and a number y. The lemma follows from the basic Zeckendorf recursion: the numbers N in Λ_{n+1} all have a digit 1 added to the expansion of the number $N - F_{n+1}$. □

Let h be the morphism on the alphabet $\{1,3,4\}$ given by

$$h(1) = 14,\ h(3) = 14,\ h(4) = 3.$$

Proposition 7. *For all $n \ge 5$ one has* (i) $\Delta C_Z(\Psi_n) = h^{n-4}(3)$ (ii) $\Delta C_Z(\Lambda_n) = h^{n-5}(3)$.

Proof. The proof is by induction. For $n = 5$, we have $\Psi_5 = [0,4]$, which has two points of constancy: $N = 1$ and $N = 2$. Therefore $C_Z(\Psi_5) = 14 = h(3)$. Here the difference 4 is coming from $N = 6$, the second point of the interval Λ_5. We further have $\Lambda_5 = [5,7]$, which has one point of constancy $N = 6$. Therefore $C_Z(\Lambda_5) = 3$.

Suppose the result has been proved till n.

(i) By equation (1),

$$\Delta C_Z(\Psi_{n+1}) = \Delta C_Z(\Psi_n)\Delta C_Z(\Lambda_n) = h^{n-4}(3)h^{n-5}(3) = h^{n-5}(h(3)3) = h^{n-5}(143) = h^{n-5}(h^2(3)) = h^{n-3}(3).$$

(ii) Directly from Lemma 6: $\Delta C_Z(\Lambda_{n+1}) = \Delta C_Z(\Psi_n) = h^{n-4}(3)$. □

Proof of Theorem 5. The statements on I_Z and D_Z follow immediately from Lemma 1.

The statement on C_Z follows from Proposition 7, part (i), since $h^n(3) = h^n(1)$ for all $n > 0$. □

3. The base phi expansion

A natural number N is written in base phi ([2]) if N has the form

$$N = \sum_{i=-\infty}^{\infty} d_i \varphi^i,$$

with digits $d_i = 0$ or 1, and where $d_i d_{i+1} = 11$ is not allowed. Ignoring leading and trailing 0's, the sum is actually finite, and the base phi representation of a number N is unique ([2]).

We write these expansions as

$$\beta(N) = d_L d_{L-1} \dots d_1 d_0 \cdot d_{-1} d_{-2} \dots d_{R+1} d_R.$$

Let for $N \geq 0$

$$s_\beta(N) := \sum_{k=L}^{k=R} d_k(N)$$

be the sum of digits function of the base phi expansions. We have

$$(s_\beta(N)) = (0, 1, 2, 2, 3, 3, 3, 2, 3, 4, 4, 5, 4, 4, 4, 5, 4, 4, 2, 3, 4, 4, 5, 5, 5, 4, 5, 6, 6, 7, 5, 5, 5, 6, \dots).$$

The case of base phi is considerably more complicated than the Zeckendorf case. We need several preparations, before we can prove Theorem 11 in Section 3.2, Theorem 12 in Section 3.3 and Theorem 13 in Section 3.4.

3.1. The recursive structure theorem

The result of this section was anticipated in [10], [11], and [16], and proved in [8].

The Lucas numbers $(L_n) = (2, 1, 3, 4, 7, 11, 18, 29, 47, 76, 123, \dots)$ are defined by

$$L_0 = 2, \quad L_1 = 1, \quad L_n = L_{n-1} + L_{n-2} \quad \text{for } n \geq 2.$$

For $n \geq 2$ we are interested in three consecutive intervals given by

$$I_n := [L_{2n+1} + 1,\ L_{2n+1} + L_{2n-2} - 1],$$
$$J_n := [L_{2n+1} + L_{2n-2},\ L_{2n+1} + L_{2n-1}],$$
$$K_n := [L_{2n+1} + L_{2n-1} + 1,\ L_{2n+2} - 1].$$

To formulate the next theorem, it is notationally convenient to extend the semigroup of words to the free group of words. For example, one has $110^{-1}01^{-1}00 = 100$.

Theorem 8. ***[Recursive Structure Theorem] I*** *For all $n \geq 1$ and $k = 1, \dots, L_{2n-1}$ one has $\beta(L_{2n} + k) = \beta(L_{2n}) + \beta(k) = 10\dots0\,\beta(k)\,0\dots01$.*

II *For all $n \geq 2$ and $k = 1, \dots, L_{2n-2} - 1$*

$$I_n: \quad \beta(L_{2n+1} + k) = 1000(10)^{-1}\beta(L_{2n-1} + k)(01)^{-1}1001,$$
$$K_n: \quad \beta(L_{2n+1} + L_{2n-1} + k) = 1010(10)^{-1}\beta(L_{2n-1} + k)(01)^{-1}0001.$$

Moreover, for all $n \geq 2$ and $k = 0, \dots, L_{2n-3}$

$$J_n: \quad \beta(L_{2n+1} + L_{2n-2} + k) = 10010(10)^{-1}\beta(L_{2n-2} + k)(01)^{-1}001001.$$

It is crucial to our analysis to partition the natural numbers in what we call the Lucas intervals, given by $\Lambda_0 := [0, 1]$, and for $n = 1, 2 \dots$ by

$$\Lambda_{2n} := [L_{2n}, L_{2n+1}], \quad \Lambda_{2n+1} := [L_{2n+1} + 1, L_{2n+2} - 1].$$

If $I = [k, \ell]$ and $J = [\ell + 1, m]$ are two adjacent intervals of integers, then we write $IJ = [k, m]$.

We code the Lucas intervals with four symbols ⓪, ①, ② and ③ (for extra readability these symbols are in color in the web version of this article), by a code Ψ in the following way:

$$\Psi(\Lambda_0) = ⓪,\ \Psi(\Lambda_1) = ①,\ \Psi(\Lambda_2) = ②,\ \Psi(\Lambda_3) = ③.$$

We then code $\Psi(\Lambda_4) = \Psi(\Lambda_0)\Psi(\Lambda_1)\Psi(\Lambda_2) = ⓪①②$, $\Psi(\Lambda_5) = \Psi(\Lambda_3)\Psi(\Lambda_2)\Psi(\Lambda_3) = ③②③$, and in general by induction, suggested by Theorem 8:

$$\Psi(\Lambda_{2n+2}) = \Psi(\Lambda_0)\Psi(\Lambda_1)\Psi(\Lambda_2)\dots\Psi(\Lambda_{2n}),$$
$$\Psi(\Lambda_{2n+1}) = \Psi(\Lambda_{2n-1})\Psi(\Lambda_{2n-2})\Psi(\Lambda_{2n-1}).$$

Let σ be the morphism on the alphabet {⓪, ①, ②, ③} defined by

$$\sigma(⓪) = ⓪①, \quad \sigma(①) = ②③, \quad \sigma(②) = ⓪①②, \quad \sigma(③) = ③②③.$$

Lemma 9. *For each $n \ge 0$ we have $\Psi(\Lambda_{2n+2}) = \sigma^n(②)$, $\Psi(\Lambda_{2n+3}) = \sigma^n(③)$.*

Proof. By induction. For $n = 0$: $\Psi(\Lambda_2) = ②$, $\Psi(\Lambda_3) = ③$. The induction step:

$$\Psi(\Lambda_{2n+5}) = \Psi(\Lambda_{2n+3})\ \Psi(\Lambda_{2n+2})\ \Psi(\Lambda_{2n+3}) = \sigma^n(③)\sigma^n(②)\sigma^n(③) = \sigma^{n+1}(③).$$

Also, using the simple identity $\sigma(②)③\sigma(②) = \sigma^2(②)$ in the last step:

$$\Psi(\Lambda_{2n+4}) = \Psi(\Lambda_0)\Psi(\Lambda_1)\Psi(\Lambda_2)\dots\Psi(\Lambda_{2n})\Psi(\Lambda_{2n+1})\Psi(\Lambda_{2n+2}) = \Psi(\Lambda_{2n+2})\Psi(\Lambda_{2n+1})\Psi(\Lambda_{2n+2}) =$$
$$\sigma^n(②)\sigma^{n-1}(③)\sigma^n(②) = \sigma^{n+1}(②) \quad \square$$

We will now show that the fixed point x_σ of the morphism σ is quasi-Sturmian, and determine its complexity function p_σ, i.e., $p_\sigma(n)$ is the number of words of length n that occurs in x_σ. Let $g_{a,b}$ the morphism on the alphabet $\{a, b\}$ given by

$$g_{a,b}(a) = baa,\ \ g_{a,b}(b) = ba. \tag{2}$$

The morphism $g_{a,b}$ is well-known, and closely related to the Fibonacci morphism. In fact, $x_g = bx_{a,b}$, if x_g is the fixed point of $g_{a,b}$, and $x_{a,b}$ is the fixed point of the Fibonacci morphism $a \to ab$, $b \to a$ (see [3]).

Proposition 10. *The fixed point x_σ of σ is equal to the decoration $\delta(x_g)$ of the fixed point x_g of $g = g_{a,b}$. The decoration morphism δ is given by $\delta(a) = ②③$, $\delta(b) = ⓪①$. For all $n \ge 1$ one has $p_\sigma(n) = n + 3$.*

Proof. For the two words ⓪① and ②③ occurring in x_σ we find

$$\sigma(⓪①) = ⓪①②③, \quad \sigma(②③) = ⓪①②③②③.$$

In other words,

$$\sigma(\delta(a)) = \delta(baa) = \delta(g(a)), \quad \sigma(\delta(b)) = \delta(ba) = \delta(g(b)).$$

Thus $\sigma\,\delta = \delta\,g$, which implies $\sigma^n\,\delta = \delta\,g^n$ for all n. Since x_σ has prefix ⓪①$= \delta(b)$, with b the prefix of x_g, this implies the first part of the proposition.

For the second part, Proposition 8 in [4] is not conclusive, as we do not know a priori the constant n_0. But there is a direct computation possible. The complexity function of the Sturmian word x_g is given by $p(n) = n + 1$. We have, distinguishing between words of even and odd length, and then splitting according to words occurring at even or odd positions in x_σ,

$$p_\sigma(2n) = p(n) + p(n+1) = n + 1 + n + 2 = 2n + 3, \quad p_\sigma(2n+1) = p(n+1) + p(n+1) = 2n + 4. \quad \square$$

Proposition 10 in combination with the main result of the paper [13], explains why the factors of x_σ have a simple return word structure. This lies at the basis of Theorem 12 in Section 3.3.

3.2. A morphic sequence representation of s_β

The image under a morphism δ of the fixed point x of a morphism, will be called a *decoration* of x. It is well known that such a $\delta(x)$ is a morphic sequence, i.e., the letter to letter projection of the fixed point of a morphism. This is the way we formulate the morphic sequence result in the next theorem.

Theorem 11. *The function s_β, as a sequence, is a decoration of a morphic sequence on an infinite alphabet, i.e., $(s_\beta(N))$ is an image under a morphism δ of a fixed point of a morphism γ. The alphabet is $\{0, 1, \ldots, j, \ldots\} \times \{\textcircled{0}, \textcircled{1}, \textcircled{2}, \textcircled{3}\}$, and γ is the morphism given for $j \geq 0$ by*

$$\gamma(\binom{j}{\textcircled{0}}) = \binom{j}{\textcircled{0}}\binom{j}{\textcircled{1}},$$
$$\gamma(\binom{j}{\textcircled{1}}) = \binom{j}{\textcircled{2}}\binom{j}{\textcircled{3}},$$
$$\gamma(\binom{j}{\textcircled{2}}) = \binom{j+2}{\textcircled{0}}\binom{j+2}{\textcircled{1}}\binom{j+2}{\textcircled{2}},$$
$$\gamma(\binom{j}{\textcircled{3}}) = \binom{j+1}{\textcircled{3}}\binom{j+2}{\textcircled{2}}\binom{j+1}{\textcircled{3}}.$$

The decoration map is given by the morphism δ:

$$\delta(\binom{j}{\textcircled{0}}) = 0 + j, 1 + j, \quad \delta(\binom{j}{\textcircled{1}}) = 2 + j, \quad \delta(\binom{j}{\textcircled{2}}) = 2 + j, 3 + j, \quad \delta(\binom{j}{\textcircled{3}}) = 3 + j, 3 + j.$$

The image $\delta(x_\gamma)$ of the fixed point x_γ of γ with initial symbol $\binom{0}{\textcircled{0}}$ equals $(s_\beta(N))$.

Proof. One combines Theorem 8 with Lemma 9. We see from part **I** of Theorem 8, that the number of 1's in the expansion of N from Λ_{2n+2} is 2 more than the number of 1's in the corresponding N' in $\Lambda_0\Lambda_1 \ldots \Lambda_{2n}$. This gives the three upper indices $j + 2$ in $\gamma(\binom{j}{\textcircled{2}})$. Similarly, part **II** gives that the number of 1's in the three intervals Λ_{2n-1}, Λ_{2n-2}, and Λ_{2n-1} is increased by 1, by 2, and respectively 1 for the corresponding N' in the interval Λ_{2n+1}. This gives the three upper indices in $\gamma(\binom{j}{\textcircled{3}})$. The lower indices are given by the morphism σ. This all happens at the level of the shifted versions of the four intervals $\Lambda_0, \Lambda_1, \Lambda_2$ and Λ_3. Here $\Lambda_0 = [0, 1]$ with $s_\beta(0) = 0$ and $s_\beta(1) = 1$; $\Lambda_1 = \{2\}$ with $s_\beta(2) = 2$; $\Lambda_2 = [3, 4]$ with $s_\beta(3) = 2$ and $s_\beta(4) = 3$; $\Lambda_3 = [5, 6]$ with $s_\beta(5) = 3$ and $s_\beta(6) = 3$. This yields the decorations δ, taking in to account the corresponding increments of the sum of digits. $\square$

We illustrate Theorem 11 with the following table.

N	0 1	2	3 4	5 6	7 8 9 10 11	12 13 14 15 16 17
$s_\beta(N)$	0 1	2	2 3	3 3	2 3 4 4 5	4 4 4 5 4 4
Lucas interval	Λ_0	Λ_1	Λ_2	Λ_3	Λ_4	Λ_5
shifted Lucas intervals	Λ_0	Λ_1	Λ_2	Λ_3	Λ_0 Λ_1 Λ_2	Λ_3 Λ_2 Λ_3
⓪,①,②,③-coding	⓪	①	②	③	⓪①②	③②③

Remark. In the paper [7] the base phi analogue of the Thue-Morse sequence, i.e., the sequence $(s_\beta(N) \mod 2)$, is shown to be a morphic sequence. This result follows also from Theorem 11, by mapping $2j$ to 0, and $2j + 1$ to 1. The morphisms found in this way are on a larger alphabet than the morphism in [7].

3.3. Generalized Beatty sequences for s_β

Let I_β be the sequence listing the points of increase of $s_\beta(N)$. We see that the first six points of increase are $I_\beta(1) = 0$, $I_\beta(2) = 1$, $I_\beta(3) = 3$, $I_\beta(4) = 7$, $I_\beta(5) = 8$, $I_\beta(6) = 10$. Similarly we define C_β and D_β.

Theorem 12. *The sequence I_β, the points of increase of the function s_β, is given by the union of the two generalized Beatty sequences*

$$(\lfloor n\varphi \rfloor + 2n)_{n\geq 0}, \text{ and } (4\lfloor n\varphi \rfloor + 3n + 1)_{n\geq 0}.$$

The sequence C_β, the points of constancy of the function s_β, is given by the union of the four generalized Beatty sequences

$$(3\lfloor n\varphi \rfloor + n + 1)_{n\geq 1}, \ (4\lfloor n\varphi \rfloor + 3n + 2)_{n\geq 0}, \ (7\lfloor n\varphi \rfloor + 4n + 2)_{n\geq 0}, \text{ and } (11\lfloor n\varphi \rfloor + 7n + 4)_{n\geq 1}.$$

The sequence D_β, the points of decrease of the function s_β, is given by the union of the three generalized Beatty sequences

$$(4\lfloor n\varphi \rfloor + 3n - 1)_{n\geq 1}, \ (7\lfloor n\varphi \rfloor + 4n)_{n\geq 1}, \text{ and } (7\lfloor n\varphi \rfloor + 4n + 4)_{n\geq 1}.$$

Proof. I: Points of increase

Any occurrence of a ⓪ gives two points of increase, namely the pair $0+j, 1+j$, and the pair $1+j, 2+j$. Here we use that ⓪ is always followed by ①. Similarly, any occurrence of a ② gives a point of increase $2+j, 3+j$.

As a consequence we obtain the numbers N which are point of increase by the sequences of occurrences of ⓪, and those of ②. How do we obtain these sequences? We have to study the return words to ⓪, and ②. The sets of these return words are respectively

{⓪①②③, ⓪①②③②③}, and {②③, ②③⓪①}.

Both ⓪, and ② induce the descendant morphism $g_{a,b}$ (the descendant morphism is a generalization of the derived morphism, see [12]). Here we coded $b :=$ ⓪①②③, $a :=$ ⓪①②③②③, respectively $b :=$ ②③, $a :=$ ②③⓪①.

The occurrences of ⓪ in the fixed point of σ occur at distances given by the lengths of $\delta($⓪①②③$)$ and $\delta($⓪①②③②③$)$. These are $|\delta($⓪①②③$)| = 7$, and $|\delta($⓪①②③②③$)| = 11$. It then follows from Lemma 1 that the increase points are given by the union of the two generalized Beatty sequences $V'(4,3,0)$ and $V'(4,3,1)$, where the $'$ indicates that these start from index 0. Similarly, the occurrences of ② have first differences 7 and 4, giving the generalized Beatty sequence $V(3,1,-1)$.

This is not yet the first result in Theorem 12, but by Lemma 2 the sequence $V(1,2,0)$ splits into the two sequences $V(3,1,-1)$ and $V(4,3,0)$. Adding $N=0$ to $V(1,2,0)$ and to $V(4,3,0)$ then yields the result on I_β in Theorem 12.

II: Points of constancy

Any occurrence of a ① gives a point of constancy, namely the pair $2+j, 2+j$. Here we use that ① is always followed by ②. Similarly, any occurrence of a ③ gives a point of constancy $3+j, 3+j$.

But there are more points of constancy. At the inner boundary of $\Lambda_2\Lambda_3$ in the quadruple $\Lambda_0\Lambda_1\Lambda_2\Lambda_3$ occurs $3, 3$. However, this is not the case at the inner boundary of the interval $\Lambda_2\Lambda_3$ in the triple $\Lambda_3\Lambda_2\Lambda_3$ in Λ_5. Since $\Psi(\Lambda_0\Lambda_2\Lambda_3\Lambda_4)$ = ⓪①$\sigma($①$)$, and $\Psi(\Lambda_3\Lambda_2\Lambda_3) = \sigma($③$)$ these points of constancy occur if and only if $\sigma($①$)$ occurs in the fixed point of σ.

This still does not yet exhaust all possibilities: there is the point $N = 14$ with $s_\beta(N) = s_\beta(N+1) = 4$ in Λ_5, not yet covered by the previous sequences. This induces points of constancy occurring at all shifted Λ_5, which occur if and only if $\sigma($③$)$ occurs in the fixed point of σ. Since any Λ_k for $k > 5$ can be written as a union of shifted versions of the three intervals $\Lambda_0\Lambda_1\Lambda_2\Lambda_3$, Λ_4, and Λ_5, we have covered all possibilities.

As a consequence we obtain the numbers N which are point of increase by the sequences of occurrences of ①, ③, $\sigma($①$)$, and $\sigma($③$)$. As before, all four have a set of two return words, and a descendant morphism that is equal to g. For ① the δ-images have lengths 11 and 7, for ③ the δ-images have lengths 7 and 4, for $\sigma($①$)$, the δ-images have lengths 29 and 18, and for $\sigma($③$)$ the δ-images have lengths 18 and 11. Application of Lemma 1 then gives the four generalized Beatty sequences of C_β in Theorem 12.

III: Points of decrease

The first point of decrease is $N = 6$, which occurs at the end of Λ_3, so $N+1 = 7$ occurs at the beginning of $\Lambda_4 = \Lambda_0\Lambda_1\Lambda_2$. This gives occurrences of points of decrease at every occurrence of ③⓪. This word has two return words: $b :=$ ③⓪①②, and $a :=$ ③⓪①②③②. These induce as descendant morphism the morphism g, once more. As $|\delta(a)| = 11$, and $|\delta(b)| = 7$, this leads to the sequence $V'(4,3,-1)$.

The next point of decrease is at $N = 11$, occurring at the inner boundary of the adjacent $\Lambda_4\Lambda_5$. The third point of decrease is at $N = 15$, which lies inside Λ_5. The coding of Λ_5 is $\Psi(\Lambda_5)$ = ③②③ = $\sigma($③$)$. As in the previous section, this gives the sequence $V(7,4,0)$ for the occurrences of the decrease points $N = 11$, and later shifts. Then $V(7,4,4)$ gives the occurrences of the decrease points $N = 15 = 11+4$, and later shifts. Again, since any Λ_k for $k > 5$ can be written as a union of intervals $\Lambda_0\Lambda_1\Lambda_2\Lambda_3$, Λ_4, and Λ_5, we have covered all possibilities. This finishes the D_β part of Theorem 12. □

3.4. Morphisms for the first differences

As for the Zeckendorf expansion, we have seen in the previous section that the points of constancy have a more complicated structure than the points of increase or the points of decrease. This phenomenon expresses itself also in the 'morphic versions' of the characterization.

Theorem 13. *The points of increase of the function s_β are given by the sequence I_β, which has $I_\beta(1) = 0$, and ΔI_β is the fixed point of the morphism on the alphabet $\{1, 2, 4\}$ given by*

$$1 \to 12,\ 2 \to 4,\ 4 \to 1244.$$

The points of constancy of the function s_β are given by the sequence C_β, which has $C_\beta(1) = 2$, and ΔC_β is a morphic sequence, given by the letter-to-letter projection $1 \to 1, 2 \to 2, 3 \to 3, 3' \to 3, 4 \to 4$ of the fixed point of the morphism on the alphabet $\{1, 2, 3, 3', 4\}$ given by

$$1 \to 43,\ 2 \to 21,\ 3 \to 21,\ 3' \to 13'43,\ 4 \to 13'4.$$

The points of decrease of the function s_β are given by the sequence D_β, which has $D_\beta(1) = 6$, and ΔD_β is the shift by one of the fixed points of the morphism on the alphabet $\{2, 4, 5, 7\}$ *given by*

$$2 \to 542,\ 4 \to 542,\ 5 \to 7,\ 7 \to 7542.$$

Proof. We use in all three cases the return words to ⓪ which are b := ⓪①②③ and a := ⓪①②③②③ to follow the occurrences of the points of increase, constancy and decrease. The important property of these return words is that the *first* occurrence of the points of increase is at the same position in the decorated a and b, and the same holds for the points of constancy and decrease.

Proof. I: Points of increase

We take in to account the increase in the differences of the occurrences of the increase points in the decorations

$$\delta(\binom{j}{⓪}\binom{j}{①}\binom{j}{②}\binom{j}{③}) = 0+j, 1+j, 2+j, 2+j, 3+j, 3+j, 3+j,$$
$$\delta(\binom{j}{⓪}\binom{j}{①}\binom{j}{②}\binom{j}{③}\binom{j}{②}\binom{j}{③}) = 0+j, 1+j, 2+j, 2+j, 3+j, 3+j, 3+j, 2+j, 3+j, 3+j, 3+j,$$

of the extended return words a and b. For a these differences are $1, 2, 4$ and 4. For b the differences between the occurrences of the increase points are $1, 2$, and 4. Recall here, that the last 4 comes from the first increase point of the next word. It follows that we can obtain ΔI_β by decorating the fixed point of the morphism g given by $a \to baa, b \to ba$ with the two words 124 and 1244. To turn this decorated fixed point in to a fixed point, we apply the natural algorithm (cf. the proof of Corollary 9 in [5]). In this case this gives the following block map on the alphabet $\{a_1, a_2, a_3, a_4, b_1, b_2, b_3\}$:

$$a_1a_2a_3a_4 \to b_1b_2b_3a_1a_2a_3a_4a_1a_2a_3a_4$$
$$b_1b_2b_3 \to b_1b_2b_3a_1a_2a_3a_4.$$

The most efficient way to turn this in to a morphism:

$$a_1 \to b_1b_2,\ a_2 \to b_3,\ a_3 \to a_1a_2a_3a_4,\ a_4 \to a_1a_2a_3a_4$$
$$b_1 \to b_1b_2,\ b_2 \to b_3,\ b_3 \to a_1a_2a_3a_4.$$

The associated letter-to-letter map λ is given by $\lambda(a_1a_2a_3a_4) = 1244$, $\lambda(b_1b_2b_3) = 124$. We see that we can consistently merge a_1 and b_1 to the letter 1, a_2 and b_2 to the letter 2, and a_3 and b_3 to the letter 4. Renaming a_4 by 4, this then yields the morphism $1 \to 12, 2 \to 4, 4 \to 1244$ as generating morphism for ΔI_β.

II: Points of constancy

We follow the same strategy as in part **I**. The differences of the occurrences of points of constancy in the decorated versions of a and b are now $2, 1, 4$ and $3, 1, 3, 4$. Decorating the fixed point of the morphism g on $\{a, b\}$ by $a \to 214$, and $b \to 3134$ this time leads to a morphism on the alphabet $\{1, 2, 3, 3', 4\}$ given by

$$1 \to 43,\ 2 \to 21,\ 3 \to 21,\ 3' \to 13'43,\ 4 \to 13'4.$$

The letter-to-letter projection $1 \to 1, 2 \to 2, 3 \to 3, 3' \to 3, 4 \to 4$ of the fixed point of this morphism on the alphabet $\{1, 2, 3, 3', 4\}$ yields the sequence ΔC_β (where $C_\beta(1) = 2$).

III: Points of decrease

The differences of the occurrences of points of decrease in the decorated versions of a and b are 7 and $5, 4, 2$. Decorating the fixed point of the morphism $g_{a,b}$ by $a \to 7$, and $b \to 542$ this time leads to a morphism on the alphabet $\{2, 4, 5, 7\}$ given by

$$2 \to 542,\ 4 \to 542,\ 5 \to 7,\ 7 \to 7542.$$

The unique fixed point of this morphism on the alphabet $\{2, 4, 5, 7\}$ yields the sequence ΔD_β, when we put $D_\beta(1) = -1$. □

4. Alternative proofs of Theorem 12 and 13

The proofs of Theorem 12 and 13 have been based entirely on the properties of the infinite morphism γ of Theorem 11. The question rises whether there is also a more local approach based on the digit blocks of the expansion as was used for the points of constancy, and the points of decrease of the Zeckendorf sum of digits function. Here we give a sketch of how this might be achieved for the points of increase of the base phi expansion. We say a number N is of type B if $d_1d_0d_{-1}(N) = 000$, and of type E if $d_2d_1d_0(N) = 001$. One can then prove the following.

Proposition 14. *A number N is a point of increase of* $(s_\beta(N))$ *if and only if N is of type* B *or of type* E.

Next, Theorem 5.1 from the paper [6] gives that type B occurs along the generalized Beatty sequence $(\lfloor n\varphi\rfloor + 2n)_{n\ge 0}$, and one can deduce from Remark 6.3 in the same paper that type E occurs along the generalized Beatty sequence $(4\lfloor n\varphi\rfloor + 3n + 1)_{n\ge 0}$. This gives the alternative proof of the I_B-part of Theorem 12, based on Proposition 14.

We next give a proof of the $\Delta I_{\rm B}$ part of Theorem 13, directly from Theorem 12 by a purely combinatorial argument.

Alternative proof of Theorem 13. Let

$$I_{\rm B} := (\lfloor n\varphi\rfloor + 2n)_{n\ge 0}, \quad I_{\rm E} := (4\lfloor n\varphi\rfloor + 3n + 1)_{n\ge 0}.$$

By Lemma 1, the difference sequence of the sequence $(\lfloor n\varphi\rfloor + 2n, n \ge 1)$ is equal to the Fibonacci word $x_{4,3} = 4344344344\ldots$ on the alphabet $\{4, 3\}$, and the difference sequence of the sequence $(4\lfloor n\varphi\rfloor + 3n + 1, n \ge 1)$ is the Fibonacci word $x_{11,7} = 11, 7, 11, 11, 7, \ldots$. However, in Theorem 12 the sequences start at $n = 0$, yielding the two difference sequences

$$\Delta I_{\rm B} = 3x_{4,3} = 34344344344\ldots, \quad \Delta I_{\rm E} = 7x_{11,7} = 7, 11, 7, 11, 11, 7, \ldots.$$

Recall that the sequences $bx_{a,b}$ are fixed points of the morphisms $g_{a,b}$ from Equation (2) given by $g_{a,b}(a) = baa$, $g_{a,b}(b) = ba$. The return words of 3 in $\Delta I_{\rm B}$ are 34 and 344. We code these words by the differences that they yield between successive occurrences of 3's, i.e., by the letters 7 and 11. Then, since

$$g_{4,3}(34) = 34\,344, \quad g_{4,3}(344) = 34\,344\,344,$$

the return words induce a derived morphism

$$7 \to 7, 11, \quad 11 \to 7, 11, 11.$$

This derived morphism happens to be equal to $g_{11,7}$, the morphism giving the sequence $\Delta I_{\rm E}$. This implies that to merge the two sequences $I_{\rm B}$ and $I_{\rm E}$ to obtain I, one has to replace the 3's in $\Delta I_{\rm B}$ by 1, 2. This decoration of $\Delta I_{\rm B}$, induces a morphism μ on the alphabet $\{1, 2, 4\}$ in the usual way, given by

$$\mu(1) = 12, \ \mu(2) = 4, \ \mu(4) = 1244.$$

This proves the theorem. □

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J.-P. Allouche, F.M. Dekking, Generalized Beatty sequences and complementary triples, Mosc. J. Comb. Number Theory 8 (2019) 325–342, https://doi.org/10.2140/moscow.2019.8.325.
[2] G. Bergman, A number system with an irrational base, Math. Mag. 31 (1957) 98–110.
[3] J. Berstel, P. Séébold, A remark on morphic Sturmian words, RAIRO Theor. Inform. Appl. 28 (1994) 255–263.
[4] J. Cassaigne, Sequences with grouped factors, in: DLT'97, Developments in Language Theory III, Thessaloniki, Aristotle University of Thessaloniki, 1998, pp. 211–222.
[5] F.M. Dekking, Morphic words, Beatty sequences and integer images of the Fibonacci language, Theor. Comput. Sci. 809 (2019) 407–417, https://doi.org/10.1016/j.tcs.2019.12.036.
[6] F.M. Dekking, Base phi representations and golden mean beta-expansions, Fibonacci Q. 58 (2020) 38–48.
[7] M. Dekking, The sum of digits function of the base phi expansion of the natural numbers, Integers 20 (#A45) (2020) 1–6.
[8] F. Michel Dekking, How to add two natural numbers in base phi, Fibonacci Q. (2021), in press.
[9] F. Michel Dekking, The structure of Zeckendorf expansions, arXiv:2006.06970v1, 2020, to appear in Integers (2021).
[10] E. Hart, On using patterns in the beta-expansions to study Fibonacci-Lucas products, Fibonacci Q. 36 (1998) 396–406.
[11] E. Hart, L. Sanchis, On the occurrence of F_n in the Zeckendorf decomposition of nF_n, Fibonacci Q. 37 (1999) 21–33.
[12] C. Holton, L.Q. Zamboni, Descendants of primitive substitutions, Theory Comput. Syst. 32 (1999) 133–157.
[13] Y. Huang, Z.Y. Wen, The sequence of return words of the Fibonacci sequence, Theor. Comput. Sci. 593 (2015) 106–116.
[14] M. Lothaire, Algebraic Combinatorics on Words, Cambridge University Press, 2002.
[15] On-line encyclopedia of integer sequences, founded by N.J.A. Sloane, electronically available at http://oeis.org.
[16] G.R. Sanchis, L.A. Sanchis, On the frequency of occurrence of α^i in the α-expansions of the positive integers, Fibonacci Q. 39 (2001) 123–137.