BASE PHI REPRESENTATIONS AND GOLDEN MEAN BETA-EXPANSIONS

F. MICHEL DEKKING

Abstract. In the base phi representation, any natural number is written uniquely as a sum of powers of the golden mean with coefficients 0 and 1, where it is required that the product of two consecutive digits is always 0. In this paper, we give precise expressions for those natural numbers for which the kth digit is 1, proving two conjectures for $k = 0, 1$. The expressions are all in terms of generalized Beatty sequences.

1. Introduction

Base phi representations were introduced by George Bergman in 1957 ([2]). Base phi representations are also known as beta-expansions of the natural numbers, with $\beta = (1+\sqrt{5})/2 = \varphi$, the golden mean.
A natural number N is written in base phi 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. Similar to base 10 numbers, we write these representations as

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

The base phi representation of a number N is unique ([2]). Our main concern will be the distribution of the digit $d_k = d_k(N)$ over the natural numbers $N \in \mathbb{N}$, where $k \geq 0$. Several authors have interpreted this in the frequency sense. The following result was conjectured by Bergman, and proved in [7].

Theorem 1.1. *The frequency of* 1*'s in* $(d_0(N))$ *exists, and* $\lim_{N\to\infty} \frac{1}{N} \sum_{M=1}^{N} d_0(M) = \frac{1}{\varphi+2} = \frac{5-\sqrt{5}}{10}$.

A more detailed description, obviously implying the previous theorem, was conjectured by Baruchel in 2018 (see A214971 in [9]):

Conjecture 1.1. *Digit* $d_0(N) = 1$ *if and only if* $N = \lfloor n\varphi \rfloor + 2n + 1$ *for a natural number* n, *or* $N = 1$.

Here $\lfloor \cdot \rfloor$ denotes the floor function, and $(\lfloor n\varphi \rfloor)$ is the well-known lower Wythoff sequence. The corresponding result for digit d_1 was conjectured by Kimberling in 2012 (see A054770 in [9]):

Conjecture 1.2. *Digit* $d_1(N) = 1$ *if and only if* $N = \lfloor n\varphi \rfloor + 2n - 1$ *for a natural number* n.

Both conjectures will be proved in Section 5. In Sections 2, 3, and 4, we introduce some objects and tools used in the proof. Finally, Section 6 gives the result for any digit $d_k(N)$ with $k \geq 1$ of the base phi expansion.

In future work, we plan to extend our results to the metallic means, or more generally to arbitrary quadratic bases, as defined and analyzed in [3].

2. Generalized Beatty Sequences

The sequences occurring in the conjectures are sequences V of the type $V(n) = p\lfloor n\alpha \rfloor + q\,n + r$, $n \geq 1$, where α is a real number, and p, q, and r are integers. As in [1], we call them *generalized Beatty sequences.* If S is a sequence, we denote its sequence of first order differences as ΔS, i.e., ΔS is defined by

$$\Delta S(n) = S(n+1) - S(n), \quad \text{for } n = 1, 2 \ldots.$$

It is well-known ([8]) that the sequence $\Delta(\lfloor n\varphi \rfloor)$ is equal to the Fibonacci word $x_{1,2} = 1211212112\ldots$ on the alphabet $\{1, 2\}$. More generally, we have the following simple lemma.

Lemma 2.1. ([1]) *Let $V = (V(n))_{n\geq 1}$ 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.*

3. Morphisms

A morphism is a map from the set of infinite words over an alphabet to itself, respecting the concatenation operation. The canonical example is the Fibonacci morphism σ on the alphabet $\{0, 1\}$ given by

$$\sigma(0) = 01, \quad \sigma(1) = 0.$$

A central role in this paper is played by the morphism γ on the alphabet $\{\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}\}$ given by

$$\gamma(\mathrm{A}) = \mathrm{AB}, \quad \gamma(\mathrm{B}) = \mathrm{C}, \quad \gamma(\mathrm{C}) = \mathrm{D}, \quad \gamma(\mathrm{D}) = \mathrm{ABC}.$$

In the following, we write $|w|$ for the length of a finite word w. Here are some useful properties of the morphism γ.

Lemma 3.1. *The morphism γ has the following properties*

i) $|\gamma^n(\mathrm{A})| = L_n$, *for all $n \geq 2$, where L_n is the nth Lucas number (see next section).*

ii) $\gamma^n(\mathrm{A}) = \gamma^n(\mathrm{C})$ *and* $\gamma^n(\mathrm{A}) = \gamma^{n+1}(\mathrm{B})$ *for all $n \geq 2$.*

Proof. i) Starting at $n = 2$, it follows by induction from the recursion of the Lucas numbers that one has $|\gamma^n(\mathrm{A})| = L_n,\ |\gamma^n(\mathrm{B})| = L_{n-1},\ |\gamma^n(\mathrm{C})| = L_n,\ |\gamma^n(\mathrm{D})| = L_{n+1}$.

ii) This follows immediately from $\gamma^2(\mathrm{A}) = \gamma(\mathrm{AB}) = \mathrm{ABC} = \gamma(\mathrm{D}) = \gamma^2(\mathrm{C})$. □

It is notationally convenient to extend the semigroup of words to the free group of words. For example, one has $\mathrm{DC}^{-1}\mathrm{B}^{-1}\mathrm{BC} = \mathrm{D}$.

4. Lucas Numbers

The Lucas numbers $(L_n) = (2, 1, 3, 4, 7, 11, 18, 29, 47, 76, 123, 199, 322, \dots)$ are defined by

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

The Lucas numbers have a particularly simple base phi representation.
From the well-known formula $L_{2n} = \varphi^{2n} + \varphi^{-2n}$, and the recursion $L_{2n+1} = L_{2n} + L_{2n-1}$, we have for all $n \geq 1$

$$\beta(L_{2n}) = 10^{2n} \cdot 0^{2n-1}1, \quad \beta(L_{2n+1}) = 1(01)^n \cdot (01)^n.$$

Exercise. Show that the base phi representation of $L_{2n+1}+1$ equals $\beta(L_{2n+1}+1) = 10^{2n+1} \cdot (10)^n 01$ – see also Lemma 3.3 (2) in [7], but note that these authors write the digits in reverse order.

Since $\beta(L_{2n})$ consists of only 0's between the exterior 1's, the following lemma is obvious.

Lemma 4.1. *For all $n \geq 1$ and $k = 1, \dots, L_{2n-1}$ one has $\beta(L_{2n}+k) = \beta(L_{2n}) + \beta(k) = 10\dots0\,\beta(k)\,0\dots01$.*

As in [6], [7], and [10], the strategy will be to partition the natural numbers in intervals $[L_n + 1, L_{n+1}]$, and establish recursive relations for the β-expansions of the numbers in these intervals. However, an analogous formula as in Lemma 4.1 starting from an *odd* Lucas number does not exist. To obtain recursive relations, the interval $[L_{2n+1}+1, L_{2n+2}-1]$ has to be divided into three subintervals. These three intervals are

$$\begin{aligned} I_n &= [L_{2n+1}+1,\, L_{2n+1}+L_{2n-2}-1], \\ J_n &= [L_{2n+1}+L_{2n-2}, L_{2n+1}+L_{2n-1}], \\ K_n &= [L_{2n+1}+L_{2n-1}+1,\, L_{2n+2}-1]. \end{aligned}$$

Note that I_n and K_n have the same length $L_{2n-2}-1$, that J_n has length $L_{2n-3}+1$, and that the starting point $L_{2n+1}+L_{2n-2}$ of J_n can be written as $2L_{2n}$.

From parts b. and c. of Proposition 3.1 and part c. of Proposition 3.2 in the paper by Sanchis and Sanchis ([10]), we obtain[1] recursions for the beta-expansions of the natural numbers in the intervals I_n, K_n, and J_n.

Lemma 4.2. ([10]) *For all $n \geq 2$ and $k = 1, \dots, L_{2n-2}-1$*

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

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

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

As an illustration, we write out what Lemma 4.2 gives for $n = 2$. In the first part, k takes the values 1 and $L_2 - 1 = 2$, giving $(10)^{-1}\beta(5)(01)^{-1} = 00 \cdot 10$ and $(10)^{-1}\beta(6)(01)^{-1} = 10 \cdot 00$. So, the beta expansions of $L_5 + 1 = 12$, $L_5 + 2 = 13$, $L_5 + L_3 + 1 = 16$, and $L_5 + L_3 + 2 = 17$ are

$$\begin{aligned} &\beta(12) = 100000 \cdot 101001,\ \beta(13) = 100010 \cdot 001001, \\ &\beta(16) = 101000 \cdot 100001,\ \beta(17) = 101010 \cdot 000001. \end{aligned}$$

[1]N.B.: these authors write the beta-expansions in reverse order.

In the second part of Lemma 4.2, k takes the values 0 and $L_1 = 1$, giving $(10)^{-1}\beta(3)(01)^{-1} = 0\cdot$ and $(10)^{-1}\beta(4)(01)^{-1} = 1\cdot$. So, the beta expansions of $L_5 + L_2 + 1 = 14$ and $L_5 + L_2 + 1 = 15$ are

$$\beta(14) = 100100 \cdot 001001,\ \beta(15) = 100101 \cdot 001001.$$

5. A Proof of the Conjectures

The conjectures in the introduction will be part of the following more general result.

Theorem 5.1. *Let $\beta(N) = (d_i(N))$ be the base phi representation of a natural number N. Then:*

$d_0(N) = 1$ *if and only if $N = \lfloor n\varphi \rfloor + 2n + 1$ for some natural number n,*

$d_1d_0(N) = 10$ *if and only if $N = \lfloor n\varphi \rfloor + 2n - 1$ for some natural number n,*

$d_1d_0d_{-1}(N) = 000$ *if and only if $N = \lfloor n\varphi \rfloor + 2n$ for some natural number n,*

$d_1d_0d_{-1}(N) = 001$ *if and only if $N = 3\lfloor n\varphi \rfloor + n + 1$ for some natural number n.*

A proof of Theorem 5.1 will be given just after the proof of Theorem 5.3.

It is convenient to code the four possibilities for the digits of N by a map T to an alphabet of four letters $\{\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}\}$. We let

$$T(N) = \mathrm{A} \text{ if and only if } d_1d_0(N) = 10, \quad T(N) = \mathrm{B} \text{ if and only if } d_1d_0d_{-1}(N) = 000,$$
$$T(N) = \mathrm{C} \text{ if and only if } d_0(N) = 1, \quad T(N) = \mathrm{D} \text{ if and only if } d_1d_0d_{-1}(N) = 001.$$

Thus, we have the following scheme.

N	$\beta(N)$	$T(N)$
1	1	C
2	$10 \cdot 01$	A
3	$100 \cdot 01$	B
4	$101 \cdot 01$	C
5	$1000 \cdot 1001$	D
6	$1010 \cdot 0001$	A
7	$10000 \cdot 0001$	B
8	$10001 \cdot 0001$	C

N	$\beta(N)$	$T(N)$
9	$10010 \cdot 0101$	A
10	$10100 \cdot 0101$	B
11	$10101 \cdot 0101$	C
12	$100000 \cdot 101001$	D
13	$100010 \cdot 001001$	A
14	$100100 \cdot 001001$	B
15	$100101 \cdot 001001$	C
16	$101000 \cdot 100001$	D

N	$\beta(N)$	$T(N)$
17	$101010 \cdot 000001$	A
18	$1000000 \cdot 000001$	B
19	$1000001 \cdot 000001$	C
20	$1000010 \cdot 010001$	A
21	$1000100 \cdot 010001$	B
22	$1000101 \cdot 010001$	C
23	$1001000 \cdot 100101$	D
24	$1001010 \cdot 000101$	A

The reader may check the validity of the following T-values, which we use in the proof of Theorem 5.3:

$$T(L_{2n}) = \mathrm{B},\ T(L_{2n} + 1) = \mathrm{C},\ T(L_{2n+1} + 1) = \mathrm{D} \quad \text{for all } n \geq 1.$$

Theorem 5.2. *The sequence $(T(N))_{N\geq 2}$ is the unique fixed point of the morphism γ.*

Theorem 5.2 is an immediate consequence of Theorem 5.3.

Theorem 5.3. *Let γ be the morphism given by* A ↦ AB, B ↦ C, C ↦ D, D ↦ ABC. *Then,*

a) $T(2)\,T(3)\cdots T(L_n{+}1) = \gamma^n(\mathrm{A})$ *for $n \geq 2$,*

b) $T(L_n{+}2)\,T(L_n{+}3)\cdots T(L_{n+1}{+}1) = \gamma^{n-1}(\mathrm{A})$ *for $n \geq 3$.*

Proof. We prove a) and b) simultaneously by induction.

For $n = 2$ and $L_2 = 3$, one finds $T(2)T(3)T(4) = \mathrm{ABC}$, which equals $\gamma^2(\mathrm{A})$.

Also for $n = 3$, one has $T(2)T(3)T(4)T(5) = \mathrm{ABCD} = \gamma^3(\mathrm{A})$.

Part b) for $n = 3$ is checked by $T(6)T(7)T(8) = \mathrm{ABC} = \gamma^2(\mathrm{A})$.

In the following, we do not formally perform an induction step $n \to n + 1$, but show how T-images of intervals can be expressed in T-images of intervals with lower indices. We have for part a)

$$\begin{aligned} T(2)\cdots T(L_{n+1}+1) &= T(2)\cdots T(L_n+1)\,T(L_n+2)\cdots T(L_{n+1}+1) \\ &= \gamma^n(\mathrm{A})\,\gamma^{n-1}(\mathrm{A}) \\ &= \gamma^n(\mathrm{AB}) = \gamma^{n+1}(\mathrm{A}). \end{aligned}$$

Here, we used Lemma 3.1 part ii).

For part b), this formula follows for even indices directly from Lemma 4.1 and part a):

$$\begin{aligned} T(L_{2n}+2)\cdots T(L_{2n+1})\,T(L_{2n+1}+1) &= T(L_{2n}+2)\cdots T(L_{2n+1})\,\mathrm{D} \\ &= T(2)\dots T(L_{2n-1})\,\mathrm{D} \\ &= T(2)\dots T(L_{2n-1})\,T(L_{2n-1}+1) = \gamma^{2n-1}(\mathrm{A}). \end{aligned}$$

For odd indices, we use Lemma 4.2. We have

$$\begin{aligned} &T(L_{2n+1}+1)\cdots T(L_{2n+1}+L_{2n-2}-1) \\ &\qquad = T(L_{2n+1}+1)\,\gamma^{2n-2}(\mathrm{A})\,T(L_{2n}+1)^{-1}\,T(L_{2n})^{-1} = \mathrm{D}\,\gamma^{2n-2}(\mathrm{A})\,\mathrm{C}^{-1}\mathrm{B}^{-1}, \\ &T(L_{2n+1}+L_{2n-2})\cdots T(L_{2n+1}+L_{2n-1}) \\ &\qquad = T(L_{2n-2})\,T(L_{2n-2}+1)\,\cdots T(L_{2n-1}+1)\,T(L_{2n-1}+1)^{-1} = \mathrm{B}\,\mathrm{C}\,\gamma^{2n-3}(\mathrm{A})\,\mathrm{D}^{-1}, \\ &\quad T(L_{2n+1}+L_{2n-1}+1)\cdots T(L_{2n+2}-1) = \mathrm{D}\,\gamma^{2n-2}(\mathrm{A})\,\mathrm{C}^{-1}\mathrm{B}^{-1}. \end{aligned}$$

Concatenating the T-images of the intervals I_n, J_n, and K_n, we obtain, using Lemma 3.1 ii)

$$\begin{aligned} &T(L_{2n+1}+2)\cdots T(L_{2n+2}+1) \\ &\quad = T(L_{2n+1}+1)^{-1}\,\mathrm{D}\,\gamma^{2n-2}(\mathrm{A})\,\mathrm{C}^{-1}\mathrm{B}^{-1}\mathrm{B}\,\mathrm{C}\,\gamma^{2n-3}(\mathrm{A})\,\mathrm{D}^{-1}\mathrm{D}\,\gamma^{2n-2}(\mathrm{A})\,\mathrm{C}^{-1}\mathrm{B}^{-1}\mathrm{BC} \\ &\quad = \gamma^{2n-2}(\mathrm{A})\,\gamma^{2n-3}(\mathrm{A})\,\gamma^{2n-2}(\mathrm{A}) = \gamma^{2n-2}(\mathrm{ABC}) = \gamma^{2n-2}(\gamma^2(\mathrm{A})) = \gamma^{2n}(\mathrm{A}). \end{aligned}$$

□

Proof of Theorem 5.1. From Theorem 5.2, we know that the digit $d_0(N) = 1$ if and only if $T(N) = \mathrm{C}$, where (with some abuse of notation) $T = \mathrm{CABCABCD}\dots$ is the fixed point of γ, prefixed by C. We see from the form of γ^2 that (apart from the prefix C) T is a concatenation of the words ABC and D. Suppose we apply a code: $\psi(\mathrm{ABC}) = 0$, $\psi(\mathrm{D}) = 1$. Then γ induces a morphism σ on the alphabet $\{0, 1\}$:

$$\sigma: \quad 0 \mapsto \psi(\gamma(\mathrm{ABC})) = \psi(\mathrm{ABCD}) = 01, \quad 1 \mapsto \psi(\gamma(\mathrm{D})) = \psi(\mathrm{ABC}) = 0.$$

We see that σ is the Fibonacci morphism, with fixed point $x_{0,1}$, given as A003849 in [9]. But, the 0's in $x_{0,1}$ occur at positions $\lfloor n\varphi \rfloor$ for $n = 1, 2 \dots$ (see, e.g., [8]). Because the differences between the indices of the positions of C in T are expanded by two by the inverse of ψ, and because of the prefix C, this implies that the C's occur at positions $\lfloor n\varphi \rfloor + 2n + 1$ for $n = 0, 1, \dots$. But, A's always occur at two places before a C, implying that the positions of A are given by $\lfloor n\varphi \rfloor + 2n - 1$ for $n = 1, \dots$. Similarly, the positions of B are given by $\lfloor n\varphi \rfloor + 2n$.

Finding the positions of D is more involved. In the following display, we underline the locations of D in the images of A, B, C, and D under the morphism γ^4:

$$\gamma^4: \quad \mathrm{A} \mapsto \mathrm{ABC\underline{D}ABC},\ \mathrm{B} \mapsto \mathrm{ABC\underline{D}},\ \mathrm{C} \mapsto \mathrm{ABC\underline{D}ABC},\ \mathrm{D} \mapsto \mathrm{ABC\underline{D}ABCABC\underline{D}}.$$

We see from this that the difference between the indices of occurrence of D in $T = \gamma^4(T)$ is always 4 or 7. Moreover, these differences as generated by A, B, C, and D under γ^4 are respectively 7, 4, 7, and the pair 7,4. Mapping A $\mapsto$ 7, B $\mapsto$ 4, C $\mapsto$ 7, and D $\mapsto$ 74, the morphism γ induces for A, C, and B a morphism $7 \mapsto 74$ and $4 \mapsto 7$. Moreover, this morphism is compatible with the part induced by D: $74 \mapsto 747$. It follows that the sequence of differences of indices of occurrence of D is nothing else but the Fibonacci word $x_{7,4}$ on the alphabet $\{7, 4\}$. Lemma 2.1 then gives that this word, written as a sequence, equals $(3\lfloor n\varphi\rfloor + n + 1)_{n\geq 1}$. □

Remark 5.1. With induction, using Lemma 4.1 and 4.2, one proves that $d_1 d_0(N) = 10$ forces $d_{-1}(N) = 0$. It follows that Theorem 5.1 implies that

Digit $d_{-1}(N) = 1$ if and only if $N = 3\lfloor n\varphi\rfloor + n + 1$ for some natural number n.

6. A General Result

Here, we give an expression for the set of N with $d_k(N) = 1$ for any $k > 1$. Recall that we partitioned the natural numbers in Lucas intervals $\Lambda_{2n} = [L_{2n}, L_{2n+1}]$ and $\Lambda_{2n+1} = [L_{2n+1} + 1, L_{2n+2} - 1]$.

The basic idea behind this partition is that if

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

then the left most index $L = L(N)$ and the right most index $R = R(N)$ satisfy

$$L(N) = 2n = |R(N)| \text{ if and only if } N \in \Lambda_{2n},$$
$$L(N) = 2n+1,\ |R(N)| = 2n+2 \text{ if and only if } N \in \Lambda_{2n+1}.$$

This is not hard to see from the simple expressions we have for the β-expansions of the Lucas numbers, see also Theorem 1 in [5]. For the cardinality $|\Lambda_n|$ of Λ_n, we have

$$|\Lambda_n| = \lfloor \varphi^{n+1}\rfloor - \lfloor \varphi^n\rfloor.$$

Theorem 6.1. *Let $\beta(N) = (d_i(N))$ be the base phi representation of a natural number N, and let $k \geq 2$. Then $d_k(N) = 1$ if and only if N is a member of one of the generalized Beatty sequences $(\lfloor n\varphi\rfloor L_k + nL_{k-1} + r)$, where $r = r_1,\ r_1 + 1, \dots, r_1 + |\Lambda_k| - 1$, with $r_1 = -L_{k-1}$ if k is even, and $r_1 = -L_{k-1} + 1$ if k is odd.*

Proof. It turns out that the coding with the alphabet $\{\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}\}$ is still useful. We extend this alphabet to an alphabet $\{\mathrm{A}_0, \mathrm{A}_1, \mathrm{B}_0, \mathrm{B}_1, \mathrm{C}_0, \mathrm{C}_1, \mathrm{D}_0, \mathrm{D}_1\}$ via the extended coding T_+ defined for $j = 0, 1$ by

$$T_+(N) = \mathrm{A}_j \text{ if and only if } d_{-2}(N) = j,\ T(N) = \mathrm{A},\ \dots,$$
$$T_+(N) = \mathrm{D}_j \text{ if and only if } d_{-2}(N) = j,\ T(N) = \mathrm{D}_j.$$

We also want to extend the morphism γ to a morphism γ_+. Here, it turns out that one has to extend γ^{k+2} instead of γ. For simplicity in notation, we suppress the dependence on k in γ_+.

We obtain γ_+ by looking at $\gamma^{k+2}(\mathrm{A})\gamma^{k+2}(\mathrm{B})\gamma^{k+2}(\mathrm{C})\gamma^{k+2}(\mathrm{D})$ – note that this word is always a prefix of $(T(N))_{N\geq 2}$ as a consequence of Theorem 5.2. We define

$$\begin{aligned}
\gamma_+(\mathrm{A}_0) &= \gamma_+(\mathrm{A}_1) = T_+(2)\cdots T_+(L_{k+2}+1),\\
\gamma_+(\mathrm{B}_0) &= \gamma_+(\mathrm{B}_1) = T_+(L_{k+2}+2)\cdots T_+(L_{k+2}+L_{k+1}+1) = T_+(L_{k+2}+2)\cdots T_+(L_{k+3}+1),\\
\gamma_+(\mathrm{C}_0) &= \gamma_+(\mathrm{C}_1) = T_+(L_{k+3}+2)\cdots T_+(L_{k+3}+L_{k+2}+1) = T_+(L_{k+3}+2)\cdots T_+(L_{k+4}+1),\\
\gamma_+(\mathrm{D}_0) &= \gamma_+(\mathrm{D}_1) = T_+(L_{k+4}+2)\cdots T_+(L_{k+4}+L_{k+3}+1) = T_+(L_{k+4}+2)\cdots T_+(L_{k+5}+1).
\end{aligned}$$

In view of the complexity of the proof, we start with the case $k=2$, so $\gamma^{k+2}=\gamma^4$, and γ_+ has the form:

$$\begin{aligned}
\gamma_+(\mathrm{A}_0) &= \gamma_+(\mathrm{A}_1) = \mathrm{A}_0\mathrm{B}_1\mathrm{C}_1\mathrm{D}_0\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0,\\
\gamma_+(\mathrm{B}_0) &= \gamma_+(\mathrm{B}_1) = \mathrm{A}_0\mathrm{B}_1\mathrm{C}_1\mathrm{D}_0,\\
\gamma_+(\mathrm{C}_0) &= \gamma_+(\mathrm{C}_1) = \mathrm{A}_0\mathrm{B}_1\mathrm{C}_1\mathrm{D}_0\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0,\\
\gamma_+(\mathrm{D}_0) &= \gamma_+(\mathrm{D}_1) = \mathrm{A}_0\mathrm{B}_1\mathrm{C}_1\mathrm{D}_0\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0\mathrm{A}_0\mathrm{B}_1\mathrm{C}_1\mathrm{D}_0.
\end{aligned}$$

Here, the $\mathrm{B}_1\mathrm{C}_1$ in $\gamma_+(\mathrm{A}_j)$ is coming from the first couple of 1's in $d_2(N)$ occurring in the interval $\Lambda_2=[L_2,L_3]=[3,4]$.

We claim that $(T_+(N))_{N\geq 2}$ is the unique fixed point of γ_+. We will prove this in a manner similar to the proof of Theorem 5.3.

Claim.
⊞ a) $T_+(2)\cdots T_+(L_{4n}+1)=\gamma_+^n(\mathrm{A}_0)$ for $n\geq 1$.
⊞ b) $T_+(L_{4n}+2)\cdots T_+(L_{4n+1}+1)=\gamma_+^n(\mathrm{B}_0)$ for $n\geq 1$.
⊞ c) $T_+(L_{4n+1}+2)\cdots T_+(L_{4n+2}+1)=\gamma_+^n(\mathrm{C}_0)$ for $n\geq 1$.
⊞ d) $T_+(L_{4n+2}+2)\cdots T_+(L_{4n+3}+1)=\gamma_+^n(\mathrm{D}_0)$ for $n\geq 1$.
⊞ e) $T_+(L_{4n+3}+2)\cdots T_+(L_{4n+4}+1)=\gamma_+^n(\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0)$ for $n\geq 1$.

Proof of the Claim. This will be done by induction, with an unexpected twist.
First the case $n=1$.
By definition, one has ⊞ a) $T_+(2)\cdots T_+(L_4+1)=\gamma_+(\mathrm{A}_0)$, ⊞ b) $T_+(L_4+2)\cdots T_+(L_5+1)=\gamma_+(\mathrm{B}_0)$, ⊞ c) $T_+(L_5+2)\cdots T_+(L_6+1)=\gamma_+(\mathrm{C}_0)$, and ⊞ d) $T_+(L_6+2)\cdots T_+(L_7+1)=\gamma_+(\mathrm{D}_0)$. What remains is ⊞ e) $T_+(L_7+2)\cdots T_+(L_8+1)=\gamma_+(\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0)$, which can be proved using Lemma 4.2:
the central part of $\beta(L_7+k)$ equals $\beta(L_5+k)$ for $k=1,\dots L_4-1$, yielding $T_+(L_7+2)\cdots T_+(L_7+L_4-1)=\gamma_+(\mathrm{C}_0)\mathrm{C}_0^{-1}\mathrm{B}_0^{-1}$. Similarly, $T_+(L_7+L_5+1)\cdots T_+(L_8-1)=\mathrm{D}_0\gamma_+(\mathrm{C}_0)\mathrm{C}_0^{-1}\mathrm{B}_0^{-1}$. In between, we have $T_+(L_7+L_4)\cdots T_+(L_7+L_4+L_3)=\mathrm{B}_0\mathrm{C}_0\gamma_+(\mathrm{B}_0)\mathrm{D}_0^{-1}$. Pasting these three words together, and adding the two letters $T_+(L_8)=\mathrm{B}_0$ and $T_+(L_8+1)=\mathrm{C}_0$, we obtain the word $\gamma_+(\mathrm{C}_0\mathrm{B}_0\mathrm{C}_0)=\gamma_+(\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0)$.

Next, we make the induction step $n\to n+1$.

⊞ a) Here, one splits $T_+(2)\cdots T_+(L_{4(n+1)}+1)$ into five subwords $T_+(L_{4n+j}+2)\cdots T_+(L_{4n+j+1}+1)$, $j=0,\dots,4$. The induction hypothesis then gives

$$T_+(2)\cdots T_+(L_{4(n+1)}+1)=\gamma_+^n(\mathrm{A}_0)\gamma_+^n(\mathrm{B}_0)\gamma_+^n(\mathrm{C}_0)\gamma_+^n(\mathrm{D}_0)\gamma_+^n(\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0)=\gamma_+^{n+1}(\mathrm{A}_0).$$

⊞ b) From Lemma 4.1, one obtains, from the induction hypothesis again with a splitting,

$$T_+(L_{4(n+1)}+2)\cdots T_+(L_{4(n+1)+1}+1) = T_+(2)\cdots T_+(L_{4n+3}+1) =$$
$$\gamma_+^n(\mathrm{A}_0)\gamma_+^n(\mathrm{B}_0)\gamma_+^n(\mathrm{C}_0)\gamma_+^n(\mathrm{D}_0) = \gamma_+^{n+1}(\mathrm{B}_0).$$

⊞ c) This is more involved, as we have to use Lemma 4.2. This lemma yields

$$\begin{aligned}&T_+(L_{4(n+1)+1}+2)\cdots T_+(L_{4(n+1)+1}+L_{4n+2}-1)\\ &= T_+(L_{4(n+1)-1}+2)\cdots T_+(L_{4(n+1)-1}+L_{4n+2}-1)\\ &= T_+(L_{4n+3}+2)\cdots T_+(L_{4n+4}-1)\\ &= \gamma_+^n(\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0)\mathrm{C}_0^{-1}\mathrm{B}_0^{-1},\end{aligned}$$

where we used part e) of the induction hypothesis in the last step. For the 'middle part', Lemma 4.2 yields

$$\begin{aligned}&T_+(L_{4(n+1)+1}+L_{4n+2})\cdots T_+(L_{4(n+1)+1}+L_{4n+3})\\ &= T_+(L_{4n+2})\cdots T_+(L_{4n+3}) = \mathrm{B}_0\mathrm{C}_0\gamma_+^n(\mathrm{D}_0)\mathrm{D}_0^{-1}.\end{aligned}$$

The last part is similar to the first part. Pasting the three parts together, and adding $\mathrm{B}_0\mathrm{C}_0$ at the end, we obtain

$$\begin{aligned}&T_+(L_{4(n+1)+1}+2)\cdots T_+(L_{4(n+1)+2}+1)\\ &= \gamma_+^n(\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0)\mathrm{C}_0^{-1}\mathrm{B}_0^{-1}\mathrm{B}_0\mathrm{C}_0\gamma_+^n(\mathrm{D}_0)\mathrm{D}_0^{-1}\mathrm{D}_0\gamma_+^n(\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0)\mathrm{C}_0^{-1}\mathrm{B}_0^{-1}\mathrm{B}_0\mathrm{C}_0\\ &= \gamma_+^n(\mathrm{A}_0\mathrm{B}_1\mathrm{C}_1)\gamma_+^n(\mathrm{D}_0)\gamma_+^n(\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0)\\ &= \gamma_+^{n+1}(\mathrm{C}_0).\end{aligned}$$

⊞ d) From Lemma 4.1, one obtains

$$\begin{aligned}&T_+(L_{4(n+1)+2}+2)\cdots T_+(L_{4(n+1)+3}+1)\\ &= T_+(2)\cdots T_+(L_{4n+5}+1)\\ &= T_+(2)\cdots T_+(L_{4n+4}+1)T_+(L_{4n+4}+2)\cdots T_+(L_{4n+5}+1)\\ &= \gamma_+^{n+1}(\mathrm{A}_0)\,\gamma_+^{n+1}(\mathrm{B}_0) = \gamma_+^{n+1}(\mathrm{D}_0).\end{aligned}$$

Here, we could not use the induction hypothesis, but can apply part a) and b) that were already proved.

⊞ e) Again, we have to use Lemma 4.2. This lemma yields

$$\begin{aligned}&T_+(L_{4(n+1)+3}+2)\cdots T_+(L_{4(n+1)+3}+L_{4n+2}-1)\\ &= T_+(L_{4(n+1)+1}+2)\cdots T_+(L_{4(n+1)+1}+L_{4n+4}-1)\\ &= T_+(L_{4n+5}+2)\cdots T_+(L_{4n+6}-1) = \gamma_+^{n+1}(\mathrm{C}_0)\mathrm{C}_0^{-1}\mathrm{B}_0^{-1},\end{aligned}$$

where we used part c) that was already proved. For the 'middle part', Lemma 4.2 yields

$$\begin{aligned}&T_+(L_{4(n+1)+3}+L_{4n+4})\cdots T_+(L_{4(n+1)+3}+L_{4n+5})\\ &= T_+(L_{4n+4})\cdots T_+(L_{4n+5}) = \mathrm{B}_0\mathrm{C}_0\gamma_+^{n+1}(\mathrm{B}_0)\mathrm{D}_0^{-1},\end{aligned}$$

where we used part b), that was already proved above.
The last part is similar to the first part. Pasting the three parts together, we obtain

$$T_+(L_{4(n+1)+3}+2)\cdots T_+(L_{4(n+1)+4}+1) = \gamma_+^{n+1}(\mathrm{C}_0)\gamma_+^{n+1}(\mathrm{B}_0)\gamma_+^{n+1}(\mathrm{C}_0) = \gamma_+^{n+1}(\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0).$$

This finishes the proof of the claim. □

To finish the proof of the theorem for the case $k = 2$, we note that the situation is almost identical[2] to the appearance of D in $\gamma^4(\mathrm{A}), \ldots, \gamma^4(\mathrm{D})$ at the end of the proof of Theorem 5.2: the words $\mathrm{B}_1\mathrm{C}_1$ occur at indices that differ by 7 or 4, and these differences occur as $x_{7,4}$, the Fibonacci word on the alphabet $\{7, 4\}$. An application of Lemma 2.1 then gives that the numbers N with $d_2(N) = 1$ occur as $N = 3\lfloor n\varphi\rfloor + n + r$, with two possibilities for r, which are found to be $r = 0$ and $r = -1$.

Consider, in general, the case of an even integer $2k$, $k = 1, 2\ldots.$ One first proves that $(T_+(N))_{N\geq 2}$ is the unique fixed point of γ_+, following the same scheme as in the proof for the $k = 2$ case. Next, one has to sort out where the N with $d_{2k}(N) = 1$ appear with respect to the $\gamma_+(\mathrm{A}_0), \ldots, \gamma_+(\mathrm{D}_0)$ in the fixed point of γ_+. The first time $d_{2k}(N) = 1$ appears is for $N = L_{2k}$, the first number in Λ_{2k}, and all other N in Λ_{2k} also have $d_{2k}(N) = 1$. By Lemma 4.1, these trains of N's, with $d_{2k}(N) = 1$, also appear at the end of Λ_{2k+2} (excepting $N = L_{2k+3} + 1$). Because they cannot appear in Λ_{2k+1}, this *is* the second appearance of the train. Application of Lemma 4.2, and another time Lemma 4.1, then gives that the third appearance is in Λ_{2k+3}, and the fourth and fifth appearance are in Λ_{2k+4}. Moreover, these three Lucas intervals correspond – except for one or two symbols at the begin and at the end – to the intervals used to define $\gamma_+(\mathrm{B}_0)$, $\gamma_+(\mathrm{C}_0)$, and $\gamma_+(\mathrm{D}_0)$, and at the same time, it shows that $\gamma_+(\mathrm{C}_0) = \gamma_+(\mathrm{A}_0)$ and $\gamma_+(\mathrm{D}_0) = \gamma_+(\mathrm{A}_0)\gamma_+(\mathrm{B}_0)$.

This means that the situation is much like the appearance of $\mathrm{B}_1\mathrm{C}_1$ in the words $\gamma_+(\mathrm{A}_0), \ldots, \gamma_+(\mathrm{D}_0)$ in the $k = 2$ case treated above: the trains occur at indices that differ by L_{2k+2} or L_{2k+1}, and these differences occur as $x_{L_{2k+2},L_{2k+1}}$, the Fibonacci word on the alphabet $\{L_{2k+2}, L_{2k+1}\}$. An application of Lemma 2.1 then gives that the numbers N in the train occur as $\lfloor n\varphi\rfloor L_{2k} + nL_{2k-1} + r$ for some r, since

$$L_{2k+2} - L_{2k+1} = L_{2k}, \quad \text{and} \quad 2L_{2k+1} - L_{2k+2} = L_{2k-1}.$$

Substituting $n = 1$, corresponding to the first train, with first element $N = L_{2k}$, gives $r_1 = -L_{2k-1}$. The length of the train is $|\Lambda_{2k}|$.

The proof for odd integers $2k + 1$ follows the same steps, the sole difference being that r_1 turns out to be one larger, because Λ_{2k} starts at L_{2k}, but Λ_{2k+1} starts at $L_{2k+1}+1$. □

Remark 6.1. Note that we also have $|\Lambda_{2n}| = L_{2n-1}+1$, and $|\Lambda_{2n+1}| = L_{2n}-1$, the expressions used in [10]. It can therefore be checked easily that our Theorem 6.1 implies the main result of [10] (for positive k).

Remark 6.2. A result similar to Theorem 6.1 will hold for digits $d_N(k)$ with k negative, but the situation is somewhat more complex. One has, for example,

digit $d_{-2}(N) = 1$ if and only if $N = 4\lfloor n\varphi\rfloor + 3n + r$ for $r = 2$, 3, or 4 and some nonnegative integer n.

Here is a proof of this statement. We define the extended coding T_+ on $\{\mathrm{A}_0, \mathrm{A}_1, \mathrm{B}_0, \mathrm{B}_1, \mathrm{C}_0, \mathrm{C}_1, \mathrm{D}_1\}$ as before. For $j = 0, 1$:

$$T_+(N) = \mathrm{A}_j \text{ if and only if } d_{-2}(N) = j,\ T(N) = \mathrm{A},\ \ldots,\ T_+(N) = \mathrm{D}_j$$
$$\text{if and only if } d_{-2}(N) = j,\ T(N) = \mathrm{D}_j.$$

[2]This observation also leads to a more or less independent proof of Theorem 6.1 for $k = 2$: $\mathrm{B}_1\mathrm{C}_1$ occurs always immediately before D_0, so the positions of B_1, respectively C_1, are just those of D in Theorem 5.1, shifted by -1 and -2.

The morphism γ_+ is more complex now:

$$\begin{aligned}
\gamma_+(\mathrm{A}_0) &= \mathrm{A}_0\mathrm{B}_0\mathrm{C}_0\mathrm{D}_0\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0,\\
\gamma_+(\mathrm{A}_1) &= \mathrm{A}_1\mathrm{B}_1\mathrm{C}_1\mathrm{D}_0\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0,\\
\gamma_+(\mathrm{B}_0) &= \gamma_+(\mathrm{B}_1) = \mathrm{A}_1\mathrm{B}_1\mathrm{C}_1\mathrm{D}_0,\\
\gamma_+(\mathrm{C}_0) &= \gamma_+(\mathrm{C}_1) = \mathrm{A}_0\mathrm{B}_0\mathrm{C}_0\mathrm{D}_0\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0,\\
\gamma_+(\mathrm{D}_0) &= \mathrm{A}_1\mathrm{B}_1\mathrm{C}_1\mathrm{D}_0\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0\mathrm{A}_1\mathrm{B}_1\mathrm{C}_1\mathrm{D}_0.
\end{aligned}$$

Here, the $\mathrm{A}_1\mathrm{B}_1\mathrm{C}_1$ in $\gamma_+(\mathrm{A}_1)$ is coming from the first triple of 1's in $d_{-2}(N)$ occurring in the interval $\Lambda_1 \cup \Lambda_2 = [2, 3, 4]$.

We claim that $(T_+(N))_{N\geq 2}$ is the unique fixed point of γ_+. We will prove this in a manner similar to the proof of Theorem 6.1. Note that although $\gamma_+^n(\mathrm{A}_0) \neq \gamma_+^n(\mathrm{A}_1)$, one has that $\gamma_+^n(\mathrm{B}_0) = \gamma_+^n(\mathrm{B}_1)$ and $\gamma_+^n(\mathrm{C}_0) = \gamma_+^n(\mathrm{C}_1)$ for all n.

Claim.

⊞ a) $T_+(2)\cdots T_+(L_{4n}+1) = \gamma_+^n(\mathrm{A}_1)$ for $n \geq 1$.

⊞ b) $T_+(L_{4n}+2)\cdots T_+(L_{4n+1}+1) = \gamma_+^n(\mathrm{B}_1)$ for $n \geq 1$.

⊞ c) $T_+(L_{4n+1}+2)\cdots T_+(L_{4n+2}+1) = \gamma_+^n(\mathrm{C}_1)$ for $n \geq 1$.

⊞ d) $T_+(L_{4n+2}+2)\cdots T_+(L_{4n+3}+1) = \gamma_+^n(\mathrm{D}_0)$ for $n \geq 1$.

⊞ e) $T_+(L_{4n+3}+2)\cdots T_+(L_{4n+4}+1) = \gamma_+^n(\mathrm{A}_0\mathrm{B}_0\mathrm{C}_0)$ for $n \geq 1$.

Proof of the Claim. This will be done by induction. Except for the change $\mathrm{A}_0 \mapsto \mathrm{A}_1$ the case $n = 1$ is literally the same as in the proof of Theorem 6.1. The induction step $n \to n+1$ can also be performed in the same way as in proof of Theorem 6.1, making the substitutions $\gamma_+^n(\mathrm{A}_0) \mapsto \gamma_+^n(\mathrm{A}_1)$ and $\gamma_+^{n+1}(\mathrm{A}_0) \mapsto \gamma_+^{n+1}(\mathrm{A}_1)$ at the appropriate places. □

Obtaining the positions of the 1's for the case $k = -2$ is more involved. We can still compare the situation to the appearance of D in $\gamma^4(\mathrm{A}), \ldots, \gamma^4(\mathrm{D})$ at the end of the proof of Theorem 5.2. There, the differences of the indices of positions (P_i) of D's occur according to the following pattern:

$T(N)$:	A	B	C	D	A	B	C	A	B	C	D	A	B	C	D
ΔP :	7	4	7	7 4	7	4	7	7	4	7	7 4	7	4	7	7 4

Moreover, it was proved that $\Delta P = x_{7,4}$, the Fibonacci word on the alphabet $\{7, 4\}$. In $(T_+(N))$, the role of D is taken over by the letter A_1. Let ΔQ be the sequence of differences of the indices of positions A_1 in $(T_+(N))$. Inspection of the five words $\gamma_+(\mathrm{A}_0), \ldots, \gamma_+(\mathrm{D}_0)$ leads to the conclusion that the ΔQ will be 7 or 11. The difference will be 11 if and only if a B_j is followed by a $\mathrm{C}_{j'}$, or when a D_0 is followed by a A_0. Because B_1 is always followed by C_1, it is always the case that D_0 is followed by A_0. It follows that we obtain ΔQ from ΔP by substituting every 4 7 in ΔP by 11:

$T_+(N)$:	A_1	B_1	C_1	D_0	A_0	B_0	C_0	A_1	B_1	C_1	D_0	A_0	B_0	C_0	D_0	$\cdots$
ΔQ :	7	11		7	11	11		7	11		7	11	11		7	$\cdots$

What is ΔQ? From Proposition 3 in [4], one obtains that the word $bx_{a,b}$ is the fixed point of the morphism $\tau : a \mapsto baa, b \mapsto ba$. Consider the morphism $\psi : a \mapsto 10, b \mapsto 0$. Then τ induces

$$10 = \psi(a) \mapsto \psi\tau(a) = \psi(baa) = 01010, \quad 0 = \psi(b) \mapsto \psi\tau(b) = \psi(ba) = 010,$$

which is equivalent to the morphism $0 \mapsto 010$ and $1 \mapsto 01$, which happens to be σ^2, where σ is the Fibonacci morphism. It follows that $\psi(bx_{a,b}) = x_{0,1}$. Now take $a = 11$, $b = 7$, and replace $0, 1$ by $7, 4$. Then, ψ can be considered as an inverse of the map $4\,7 \mapsto 11$ and $7 \mapsto 7$ from ΔP to ΔQ. It follows that $\Delta Q = 7x_{11,7}$. An application of Lemma 2.1 then gives that the numbers N, with $d_{-2}(N) = 1$ and $T_+(N) = \mathrm{A}_1$, occur as $N = 4\lfloor n\varphi\rfloor + 3n + 2$ for $n \geq 1$, except that $N = 2$ is missing. We obtain all occurrences by letting the generalized Beatty sequence start at $N = 2$, by adding the index $n = 0$. This leads to the announced expression.

NOTE ADDED IN PROOF: Using the idea of return words, some of the proofs in this paper can be streamlined.

References

[1] J.-P. Allouche and F. M. Dekking, *Generalized Beatty sequences and complementary triples*, arXiv: 1809.03424v3 [math.NT]. To appear in Mosc. J. Comb. Number Theory (2019).

[2] G. Bergman, *A number system with an irrational base*, The Mathematics Magazine, **31** (1957), 98–110.

[3] E. B. Burger, D. C. Clyde, C. H. Colbert, G. H. Shin, and Z. Wang, *Canonical Diophantine representations of natural numbers with respect to quadratic "bases"*, J. Number Theory, **133** (2013), 1372–1388.

[4] M. Dekking, *Substitution invariant Sturmian words and binary trees*, Integers, **18A** (2018), #A7, 1–14.

[5] P. J. Grabner, I. Nemes, A. Pethö, and R. F. Tichy, *Generalized Zeckendorf decompositions*, Appl. Math. Lett., **7** (1994), 25–28.

[6] E. Hart, *On using patterns in the beta-expansions To study Fibonacci-Lucas products*, The Fibonacci Quarterly, **36.5** (1998), 396–406.

[7] E. Hart and L. Sanchis, *On the occurrence of F_n in the Zeckendorf decomposition of nF_n*, The Fibonacci Quarterly, **37.1** (1999), 21–33.

[8] M. Lothaire, *Algebraic combinatorics on words*, Cambridge University Press, 2002.

[9] OEIS Foundation Inc. (2011), The On-Line Encyclopedia of Integer Sequences, `https://oeis.org`.

[10] G. R. Sanchis and L. A. Sanchis, *On the frequency of occurrence of α^i in the α-expansions of the positive integers*, The Fibonacci Quarterly, **39.2** (2001), 123–173.

MSC2010: 11D85, 11A63, 68R15

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

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