The sum of digits function of the base phi expansion of the natural numbers

Dekking, Michel

Publication date
2020

Document Version
Final published version

Published in
Integers

THE SUM OF DIGITS FUNCTION OF THE BASE PHI EXPANSION OF THE NATURAL NUMBERS

Michel Dekking
Dept. of Mathematics, Delft University of Technology, Delft, The Netherlands
F.M.Dekking@math.tudelft.nl

Received: 11/25/19, Accepted: 5/23/20, Published: 6/18/20

Abstract

In the base phi expansion any natural number is written uniquely as a sum of powers of the golden mean with digits 0 and 1, where one requires that the product of two consecutive digits is always 0. In this paper we show that the sum of digits function modulo 2 of these expansions is a morphic sequence. In particular we prove that — like for the Thue-Morse sequence — the frequency of 0's and 1's in this sequence is equal to 1/2.

1. Introduction

Base phi representations were introduced by George Bergman in 1957 [1]. 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 is represented as

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

with digits $d_i = 0$ or 1, and where $d_i d_{i+1} = 11$ is not allowed. We write these expansions as

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

Ignoring leading and trailing 0's, the base phi representation of a number N is unique, as shown by Bergman.

Let for $N \geq 0$

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

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

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

In this paper we study the base phi analogue of the Thue-Morse sequence (where the base equals 2), i.e., the sequence

$$(x_\beta(N)) := (s_\beta(N) \mod 2) = 0, 1, 0, 0, 1, 1, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, \dots$$

Recall that a morphism is a map from the set of infinite words over an alphabet to itself, respecting the concatenation operation. The Thue Morse sequence is the fixed point starting with 0 of the morphism $0 \to 01,\ 1 \to 10$.

Theorem 1. *The sequence x_β is a morphic sequence, i.e., the letter-to-letter image of the fixed point of a morphism.*

This theorem permits us to answer a number of natural questions one may ask about x_β, for example: will a word 00000 ever occur? What are the frequencies of 0 and 1?

We end this introduction by mentioning some related work. In [2] asymptotic expressions for $\sum_{N<x} s_\beta(N)$ as $x \to \infty$ were obtained. In [7], so-called α-irreducibles were introduced, which might serve as building blocks for $s_\beta(N)$. An α-irreducible is a natural number N, such that if $\beta(N) = \beta(N') + \beta(N'')$ with $N' < N''$, then $N' = 0$ and $N'' = N$. The first twelve α-irreducibles are 1,2,3,5,6,7,12,13,14,16,17,18. Grabner and Prodinger give a detailed asymptotic description of the counting function A, where $A(n)$ is the number of α-irreducibles among $1, 2, \dots, n$. From their Theorem 1, and Lemma 1 and Lemma 2 in the next section, one can obtain new insights in A. Let (L_n) be the Lucas numbers. The even Lucas intervals $[L_{2n}, L_{2n+1}]$ will contain no α-irreducibles, with exception of $N = L_{2n}$. The odd Lucas intervals $[L_{2n+1}+1, L_{2n+2}-1]$, with $N = L_{2n+2}$ added, will contain two shifted copies of the α-irreducibles in the previous (extended) odd Lucas interval. Since $L_{2n+1} \sim \varphi^{2n+1}$, this directly implies the crude asymptotics of the counting function: $A(n) \asymp n^\rho$, with $\rho = \log 2/\log\varphi^2$.

2. Properties of the Base Phi Representation

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

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

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 $n \geq 1$

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

The properties of base phi expansion of the natural numbers are intrinsically determined by the *Lucas intervals*:

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

When we add $\Lambda_0 := [0, 1]$, these intervals partition the natural numbers as $n = 0, 1, 2 \dots$. The partition elements correspond to the lengths of the expansions:

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

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

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 [6].

For two expansions $\beta(N)$ and $\beta(N')$, we write $\beta(N) + \beta(N')$ for the digit-wise addition of these expansions, tacitly assuming that 0's have been added to the left and/or right of these expansions to make this well-defined. Since $\beta(L_{2n})$ consists of only 0's between the exterior 1's, the following lemma is obvious.

Lemma 1. ([3]) *For all $n \geq 1$ and $k = 0, \dots, L_{2n-1}$ one has $\beta(L_{2n}+k) = \beta(L_{2n}) + \beta(k)$.*

This gives a recursive relation for the expansions in the Lucas interval Λ_{2n}. To obtain recursive relations for the interval Λ_{2n+1}, this interval has to be divided into three subintervals. These three intervals are

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

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

Lemma 2. ([11], [3])[1] *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 $k = 0, \dots, L_{2n-3}$,

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

3. The Sequence x_β is Morphic

If $V = [K, K+1, \dots, L]$ is an interval of natural numbers, then we write

$$x_\beta(V) := [x_\beta(K), x_\beta(K+1), \dots, x_\beta(L)]$$

[1]See [4] for a comprehensive proof of Lemma 2

for the consecutive sums of digits modulo 2 of these numbers.

Since $x_\beta(L_{2n}) = 0$ and $x_\beta(0) = 0$, Lemma 1 implies directly the following lemma.

Lemma 3. (EVEN) *For* $n \geq 1$ *one has* $x_\beta(\Lambda_{2n}) = x_\beta([0, L_{2n-1}])$.

The mirror morphism on $\{0,1\}$ is defined by $\overline{0} = 1, \overline{1} = 0$.

We obtain from Lemma 2 with $x_\beta(I_n) = x_\beta(K_n) = \overline{x_\beta(\Lambda_{2n-1})}$, and $x_\beta(J_n) = x_\beta(\Lambda_{2n-2})$ the following.

Lemma 4. (ODD) *For* $n \geq 1$ *one has* $x_\beta(\Lambda_{2n+1}) = \overline{x_\beta(\Lambda_{2n-1})}x_\beta(\Lambda_{2n-2})\overline{x_\beta(\Lambda_{2n-1})}$.

We illustrate the base phi expansions with the following table.

N	$\beta(N)$	$x_\beta(N)$	Lucas interval
0	0	0	Λ_0
1	1	1	Λ_0
2	$10 \cdot 01$	0	Λ_1
3	$100 \cdot 01$	0	Λ_2
4	$101 \cdot 01$	1	Λ_2
5	$1000 \cdot 1001$	1	Λ_3
6	$1010 \cdot 0001$	1	Λ_3
7	$10000 \cdot 0001$	0	Λ_4
8	$10001 \cdot 0001$	1	Λ_4
9	$10010 \cdot 0101$	0	Λ_4
10	$10100 \cdot 0101$	0	Λ_4
11	$10101 \cdot 0101$	1	Λ_4
12	$100000 \cdot 101001$	0	Λ_5

Let τ be the morphism on the alphabet $A := \{1, \ldots, 8\}$ defined by

$$\begin{array}{llll} \tau(1) = 12, & \tau(2) = 312, & \tau(3) = 47, & \tau(4) = 8312, \\ \tau(5) = 56, & \tau(6) = 756, & \tau(7) = 83, & \tau(8) = 4756. \end{array}$$

Define the mirroring morphism μ on A by

$$\mu: \ 1 \to 5, \ 2 \to 6, \ 3 \to 7, \ 4 \to 8, \ 5 \to 1, \ 6 \to 2, \ 7 \to 3, \ 8 \to 4.$$

Then τ is mirror invariant: $\tau\mu = \mu\tau$.

Theorem 2. *Let* x_β *be the sum of digits function of the base phi expansions of the natural numbers. Let* $\lambda : A^* \to \{0,1\}$ *be the letter-to-letter morphism given by*

$$\lambda(1) = \lambda(3) = \lambda(6) = \lambda(8) = 0, \text{ and } \lambda(2) = \lambda(4) = \lambda(5) = \lambda(7) = 1.$$

Then $x_\beta = \lambda(t)$, *where* $t = 1231247123\ldots$ *is the fixed point of* τ *starting with 1.*

Theorem 2 is a direct consequence of the following result. Note that $\overline{\lambda\tau} = \lambda\tau\mu$.

Proposition 1. *For $n = 1, 2\dots$ one has $x_\beta(\Lambda_{2n}) = \lambda(\tau^n(1))$, and $x_\beta(\Lambda_{2n+1}) = \lambda(\tau^n(3))$.*

Proof. By induction. For $n = 1$ one has $x_\beta(\Lambda_2) = 01 = \lambda(12) = \lambda(\tau(1))$, and $x_\beta(\Lambda_3) = 11 = \lambda(47) = \lambda(\tau(3))$. From Lemma 3 and the induction hypothesis we have

$$\begin{aligned} x_\beta(\Lambda_{2n+2}) &= x_\beta([0, L_{2n-1}])x_\beta([L_{2n-1}+1, L_{2n}-1])x_\beta([L_{2n}, L_{2n+2}]) \\ &= \lambda(\tau^n(1))\lambda(\tau^{n-1}(3))\lambda(\tau^n(1)) \\ &= \lambda(\tau^{n-1}(12312)) = \lambda(\tau^{n+1}(1)). \end{aligned}$$

From Lemma 4 and the induction hypothesis we have

$$\begin{aligned} x_\beta(\Lambda_{2n+3}) &= \overline{x_\beta(\Lambda_{2n+1})}x_\beta(\Lambda_{2n})\overline{x_\beta(\Lambda_{2n+1})} \\ &= \overline{\lambda(\tau^n(3)}\lambda(\tau^n(1))\overline{(\lambda(\tau^n(3))} \\ &= \lambda(\tau^n(7))\lambda(\tau^n(1))\lambda(\tau^n(7)) \\ &= \lambda(\tau^n(717)) = \lambda(\tau^n(47)) = \lambda(\tau^{n+1}(3)). \end{aligned}$$ □

Since τ is mirror invariant, the letters a and $\mu(a)$ have the same frequency for $a \in A$. As $\overline{\lambda} = \lambda\mu$, this implies the following.

Proposition 2. *The letters 0 and 1 have frequency $\frac{1}{2}$ in x_β.*

It is well-known that the words of length 2 in the Thue-Morse sequence have frequencies $\frac{1}{6}$ for 00 and 11, and $\frac{1}{3}$ for 01 and 10. Here is the corresponding result for the golden mean sum of digits function.

Proposition 3. *In x_β the words 00 and 11 have frequency $\frac{1}{10}\sqrt{5}$, and the words 01 and 10 have frequency $\frac{1}{2} - \frac{1}{10}\sqrt{5}$.*

Proof. As in [10] we compute the frequencies $\nu[ab]$ of the words ab of length 2 occurring in the fixed point t of the morphism τ by using the 2-block substitution $\tau^{[2]}$. The words of length 2 occurring in the fixed point t of the morphism τ are

12, 23, 24, 28, 31, 35, 47, 56, 64, 67, 68, 71, 75, 83.

When we code the 14 words of length 2 by $\ell_1, \dots, \ell_{14}$, in the order given above, then $\tau^{[2]}$ is given for the letters $\ell_1, \dots \ell_7$ by

$$\ell_1 \to \ell_1\ell_2,\, \ell_2 \to \ell_5\ell_{13},\, \ell_3 \to \ell_5\ell_{14},\, \ell_4 \to \ell_5\ell_{13},\, \ell_5 \to \ell_7\ell_{12},\, \ell_6 \to \ell_7\ell_{13},\, \ell_7 \to \ell_{14}\ell_5\ell_{14}.$$

The $\tau^{[2]}$-images of $\ell_8, \dots, \ell_{14}$ follow from this by mirror-symmetry. The first 7 components of the normalized eigenvector of the incidence matrix of the morphism $\tau^{[2]}$ are given by

$$\left[\tfrac{1}{4} - \tfrac{1}{20}\sqrt{5},\ \tfrac{1}{2} - \tfrac{1}{5}\sqrt{5},\ \tfrac{3}{20} - \tfrac{1}{20}\sqrt{5},\ \tfrac{1}{5}\sqrt{5} - \tfrac{2}{5},\ \tfrac{1}{10},\ \tfrac{3}{20} - \tfrac{1}{20}\sqrt{5},\ \tfrac{3}{20}\sqrt{5} - \tfrac{1}{4}\right].$$

This means that, e.g., $\nu[12] = \frac{1}{4} - \frac{1}{20}\sqrt{5}$, and $\nu[31] = \frac{1}{10}$. The frequency of 00 equals $\mu[00] = \nu[13] + \nu[68] + \nu[83] = \frac{1}{10}\sqrt{5}$. □

Remark. Christian Mauduit with Michael Drmota and Joël Rivat proved that the Thue-Morse sequence is normal along squares (see [5]). We conjecture that this also holds for the sum of digits function modulo 2 of the basis phi expansion of the natural numbers, i.e., for $(x_\beta(n^2))$.

Acknowledgement. I am grateful to the organizers of the CIRM conference "Prime Numbers, Determinism and Pseudorandomness" for creating an excellent environment for remembering Christian Mauduit and his work. I also thank Peter Grabner for providing a relevant reference.

References

[1] G. Bergman, A number system with an irrational base, *Math. Mag.* **31** (1957), 98–110.

[2] C. Cooper and R. E. Kennedy, The first moment of the number of 1's function in the beta-expansion of the positive integers, *Journal of Institute of Mathematics & Computer Sciences* **14** (2001), 69–77.

[3] M. Dekking, Base phi representations and golden mean beta-expansions, *Fibonacci Quart.* **58** (2020), 38–48.

[4] M. Dekking, How to add two natural numbers in base phi. To appear in *Fibonacci Quart.* (2020).

[5] M. Drmota, C. Mauduit, and J. Rivat, Normality along squares, *J. Eur. Math. Soc.* **21** (2019), 507–548.

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

[7] P. J. Grabner and H. Prodinger, Additive irreducibles in α-expansions, *Publ. Math. Debrecen* **80** (2012), 405–415.

[8] E. Hart, On using patterns in the beta-expansions to study Fibonacci-Lucas products, *Fibonacci Quart.* **36** (1998), 396–406.

[9] E. Hart and L. Sanchis, On the occurrence of F_n in the Zeckendorf decomposition of nF_n, *Fibonacci Quart.* **37** (1999), 21–33.

[10] M. Queffélec, *Substitution Dynamical Systems – Spectral Analysis*, Lecture Notes in Mathematics 1294, 2nd ed., Springer, Berlin, 2010.

[11] G.R. Sanchis and L.A. Sanchis, On the frequency of occurrence of α^i in the α-expansions of the positive integers, *Fibonacci Quart.* **39** (2001), 123–173.