



Combinatorial constructions for the Zeckendorf sum of digits of polynomial values

Thomas Stoll

► To cite this version:

Thomas Stoll. Combinatorial constructions for the Zeckendorf sum of digits of polynomial values. Ramanujan Journal (The), 2013, 32 (2), pp.227-243. 10.1007/s11139-012-9422-6 . hal-01278548

HAL Id: hal-01278548

<https://hal.science/hal-01278548>

Submitted on 24 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

COMBINATORIAL CONSTRUCTIONS FOR THE ZECKENDORF SUM OF DIGITS OF POLYNOMIAL VALUES

THOMAS STOLL

ABSTRACT. Let $p(X) \in \mathbb{Z}[X]$ with $p(\mathbb{N}) \subset \mathbb{N}$ be of degree $h \geq 2$ and denote by $s_F(n)$ the sum of digits in the Zeckendorf representation of n . We study by combinatorial means three analogues of problems of Gelfond (1967/68), Stolarsky (1978) and Lindström (1997) concerning the distribution of s_F on polynomial sequences. First, we show that for $m \geq 2$ we have $\#\{n < N : s_F(p(n)) \equiv a \pmod{m}\} \gg_{p,m} N^{4/(6h+1)}$ (Gelfond). Secondly, we find the extremal minimal and maximal orders of magnitude of the ratio $s_F(p(n))/s_F(n)$ (Stolarsky). Third, we prove that $\limsup_{n \rightarrow \infty} s_F(p(n))/\log_\varphi(p(n)) = 1/2$, where φ denotes the golden ratio (Lindström).

1. INTRODUCTION

Denote by $(F_j)_{j \geq 0} = 0, 1, 1, 2, 3, 5, 8, 13, \dots$ the sequence of Fibonacci numbers defined by $F_0 = 0$, $F_1 = 1$ and

$$(1) \quad F_{j+2} = F_{j+1} + F_j, \quad j \geq 0.$$

Zeckendorf's theorem (see [23] or [1, p.105]) says that every positive integer n can be uniquely written as

$$(2) \quad n = \sum_{j \geq 0} \varepsilon_j F_{j+2},$$

with the digits $\varepsilon_j \in \{0, 1\}$ under the condition that $\varepsilon_j \varepsilon_{j+1} = 0$ for all $j \geq 0$. In what follows, we will say that the right hand side of (2) has *non-adjacent terms* or, more informally, that the ε_j 's are *non-adjacent* to refer to this condition. For convenience, we will say that the digit ε_j is at *position* $j+2$. Let s_F be the sum of digits function defined in this numeration system, *i.e.*,

$$s_F(n) = \sum_{j \geq 0} \varepsilon_j.$$

The Zeckendorf sum of digits function s_F shares many properties with the sum of digits function in the usual q -ary positional representation of integers¹ and many analogues of well-known results have been found so far. For instance, Coquet and Van Den Bosch [4] showed that s_F has fractal summatory behaviour, and Drmota and Skalba [8] proved a Newman parity phenomenon. Also, there are various results known about the distribution of s_F on polynomial sequences. Drmota and Steiner [9] showed that a proper renormalization of $s_F(n^h)$ is asymptotically normally distributed. Nowadays, several extensions of these results to numeration systems are known that use more general linear recurring sequences as digital base

¹In the q -ary expansion, one uses q^j in place of F_{j+2} in (2) with digits ε_j in the set $\{0, 1, \dots, q-1\}$.

sequences. We refer to the article of Wagner [22] and the references given therein for a rather extensive list.

The aim of the present paper is to study – by purely combinatorial and elementary means – three Zeckendorf analogues of problems concerning the distribution of the sum of digits on polynomial values, that have so far been considered for the q -ary (or binary) case.

In what follows, let $p(x) \in \mathbb{Z}[x]$ with $p(\mathbb{N}) \subset \mathbb{N}$ be a polynomial of degree $h \geq 2$ and set

$$p(x) = a_h x^h + a_{h-1} x^{h-1} + \cdots + a_0.$$

In 1967/68, Gelfond [11] proved equidistribution in arithmetic progressions for the q -ary sum of digits of linear polynomials. He then asked to find an asymptotic formula ($n \rightarrow \infty$) for the q -ary sum of digits of $p(n)$, $h \geq 2$, in a fixed arithmetic progression. We refer to the recent paper of Drmota, Mauduit and Rivat [6] and the references given therein for a detailed description and the most recent result to the problem. A complete answer to Gelfond's question is available only when either the base q is particularly large [6], *i.e.*, $q \geq \exp(67h^3(\log h)^2)$, or the degree h of $p(x)$ is particularly small [11], [18], *i.e.*, $h \leq 2$.

The Zeckendorf expansion of integers can be seen as an expansion with respect to $\varphi = (1 + \sqrt{5})/2 < 2$, the golden ratio, and it is the purpose of this paper to find a general lower bound for this expansion for polynomials of arbitrary degree $h \geq 2$. The equidistribution result for linear polynomials in expansions with respect to linear recurrences has been found in Drmota and Skalba [8] (see (59), (60) in that paper).

Our main result complements the investigation put forward by Dartyge and Tenenbaum [5], and by the author [21], and is as follows.²

Theorem 1.1. *Let $m \geq 2$. Then for all $a \in \mathbb{Z}$ with $0 \leq a < m$ we have*

$$\#\{n < N : s_F(p(n)) \equiv a \pmod{m}\} \gg_{p,m} N^{4/(6h+1)}.$$

In 1978, Stolarsky [20] studied the pointwise relationship between the binary sum of digits of n^h versus the binary sum of digits of n . Recently, Hare, Laishram and Stoll [15] found the maximal and minimal extremal orders of magnitude of the related ratio in the q -ary case and settled hereby an open problem posed by Stolarsky.

In the present paper, we are interested to study the same problem for the Zeckendorf sum of digits function. The function s_F is different from its q -ary counterpart in an important respect, namely, it is in general *not submultiplicative* (compare with [15, Proposition 2.2] or [19]). A simple example for this is

$$\begin{aligned} s_F(37 \cdot 42) &= s_F(1554) = s_F(F_{16} + F_{14} + F_{12} + F_9 + F_6 + F_4 + F_2) = 7, \\ s_F(37) \cdot s_F(42) &= s_F(F_9 + F_4) \cdot s_F(F_9 + F_6) = 2 \cdot 2 = 4. \end{aligned}$$

We therefore might expect that $s_F(n^h)/s_F(n)$ for fixed $h \geq 2$ can be (up to the multiplicative constants) as large as $\log n$ and as small as $(\log n)^{-1}$. This is a different

²We write $f \ll_\omega g$ resp. $f \gg_\omega g$ if there exists a positive constant c depending at most on ω such that $f \leq cg$ resp. $f \geq cg$.

matter for the q -ary sum of digits [15] where the submultiplicativity property yields an upper bound of the form $(\log n)^{1-1/h}$.

Theorem 1.2. *There exist constants $C_1 = C_1(p) > 0$ and $C_2 = C_2(p) > 0$ such that*

$$(3) \quad \frac{s_F(p(n))}{s_F(n)} > C_1 \log n,$$

respectively,

$$(4) \quad \frac{s_F(p(n))}{s_F(n)} < \frac{C_2}{\log n},$$

for an infinite number of integers n . In particular, we have

$$\limsup_{n \rightarrow \infty} \frac{s_F(p(n))}{s_F(n)} = +\infty, \quad \liminf_{n \rightarrow \infty} \frac{s_F(p(n))}{s_F(n)} = 0.$$

The function s_F is, like its q -ary counterpart, a *subadditive* function (see Proposition 2.1). This allows the following interpretation: $s_F(n)$ equals the minimal number of Fibonacci numbers needed to write n as a sum thereof. Finding polynomial values with small and large values of s_F in turn is related to some extent to the classical investigation of finding perfect powers among Fibonacci numbers and their sums. A deep result of Bugeaud, Mignotte and Siksek [3] says that the only powers n^h among Fibonacci numbers are 1, 8 and 144. In our language, this means that

$$\# \{(n, h) : s_F(n^h) = 1, \quad n \geq 2, \quad h \geq 2\} = 2.$$

One might ask, whether the number of pairs (n, h) remains still small (or finite) when $s_F(n^h) = 1$ is replaced by a less stringent condition. From the method of proof of Theorem 1.2 we obtain the following result. Recall that φ denotes the golden ratio.

Corollary 1.3. *There exist $C = C(p) > 0$ and $N_0 = N_0(p) > 0$, such that for all $N \geq N_0$,*

$$\# \{n : s_F(p(n)) \leq C, \quad s_F(n) = N, \quad n \geq 1\} \geq 1.$$

Moreover, if $p(x) = x^h$ then for all $N \geq 8h + 1$ we have

$$\# \{n < 56\varphi^{4N} : s_F(n^h) \leq (2h + 1)(8h + 2), \quad s_F(n) = N\} \geq 1.$$

In other words, for sufficiently large N there always exists an integer n such that n is the sum of exactly N distinct, non-adjacent Fibonacci numbers, and $p(n)$ can be expressed as the sum of at most $C(p)$ Fibonacci numbers.

Our third result concerns a Zeckendorf analogue of a theorem of Lindström [17] (see also Drmota and Rivat [7] for the special case of squares) who proved that for $\varepsilon > 0$ there exists an n such that the density of 1's in the binary digital expansion of $p(n)$ is at least $1 - \varepsilon$.

Theorem 1.4. *We have*

$$\liminf_{n \rightarrow \infty} \frac{s_F(p(n))}{\log_{\varphi} p(n)} = 0, \quad \limsup_{n \rightarrow \infty} \frac{s_F(p(n))}{\log_{\varphi} p(n)} = \frac{1}{2}.$$

Bugeaud, Luca, Mignotte and Siksek [2] also found all powers that are at most one away from a Fibonacci number. Rephrased in our context, this can be seen as finding powers with sum of digits values that are close to the maximal and minimal possible ones. Theorem 1.4 says that for any polynomial p and $n \rightarrow \infty$ we can get arbitrarily “close” to a Fibonacci number both from above and from below, where “close” here means that the values differ by adding or subtracting only “few” Fibonacci numbers.

The paper is structured as follows. In Section 2 we provide some necessary notation and tools to treat the three problems. Sections 3, 4 and 5 are then devoted to the proofs of Theorem 1.1, Theorem 1.2 and Corollary 1.3, and Theorem 1.4, respectively.

2. NOTATION AND AUXILIARY RESULTS

We first show that s_F is a subadditive function.

Proposition 2.1. *For all integers $n_1, n_2 \geq 0$ we have*

$$s_F(n_1 + n_2) \leq s_F(n_1) + s_F(n_2).$$

Proof. Write n_1 and n_2 in Zeckendorf expansion, *i.e.*,

$$n_1 = \sum_j \varepsilon_j^{(1)} F_{j+2} \quad \text{and} \quad n_2 = \sum_j \varepsilon_j^{(2)} F_{j+2},$$

and consider

$$(5) \quad n_1 + n_2 = \sum_j (\varepsilon_j^{(1)} + \varepsilon_j^{(2)}) F_{j+2}.$$

The right hand side of (5) is in general not the Zeckendorf expansion of $n_1 + n_2$ since on the one hand, $\varepsilon_j^{(1)} + \varepsilon_j^{(2)} \in \{0, 1, 2\}$ and, on the other hand, there is no condition that prohibits adjacent terms. However, we have carry propagation rules at our disposal to get back the Zeckendorf expansion of $n_1 + n_2$. First, if we encounter two adjacent 1’s, say, $(\varepsilon_{j'}^{(1)} + \varepsilon_{j'}^{(2)}) = (\varepsilon_{j'+1}^{(1)} + \varepsilon_{j'+1}^{(2)}) = 1$, we can replace these two 1’s by a single 1 at position $j' + 2$ because of

$$(6) \quad 1 \cdot F_{j'+1} + 1 \cdot F_{j'} = 1 \cdot F_{j'+2}.$$

(At first, we neglect the addition that has to be performed at position $j' + 2$.) When we apply this rule for the first time, we decrease $s_F(n_1) + s_F(n_2)$ by 1. Second, if we encounter a 2 at position $j' \geq 4$, say, $(\varepsilon_{j'}^{(1)} + \varepsilon_{j'}^{(2)}) = 2$, then we can replace this non-admissible number with the aid of the recurrence

$$(7) \quad 2 \cdot F_{j'} = 1 \cdot F_{j'+1} + 1 \cdot F_{j'-2}.$$

Note that here we have equality of coefficients $2 = 1 + 1$, so that $s_F(n_1) + s_F(n_2)$ remains unmodified when we apply this rule for the first time. For $j' \leq 3$ we use

$$(8) \quad 2 \cdot F_3 = 1 \cdot F_4 + 1 \cdot F_2 \quad \text{and} \quad 2 \cdot F_2 = 1 \cdot F_3,$$

in either case the value of $s_F(n_1) + s_F(n_2)$ is not increased. The replacement rules (6)–(8) shift a portion of the summands to the more significant digits. We now repeat the application of these rules. Note that we can use (7) also when we encounter a number ≥ 3 at some position, to shift the contribution to the more significant digits. After a finite number of applications of these rules (in any order)

we will end up with the Zeckendorf expansion of $n_1 + n_2$. This concludes the proof. \square

Consider the Zeckendorf expansions of the m integers n_1, n_2, \dots, n_m ,

$$n_i = \sum_{j=k_i}^{K_i} \varepsilon_j^{(i)} F_{j+2}, \quad i = 1, 2, \dots, m,$$

with $\varepsilon_{k_i}^{(i)} = \varepsilon_{K_i}^{(i)} = 1$. Suppose that $k_1 \leq k_2 \leq \dots \leq k_m$ (this can always be achieved by reordering of the n_i). We say that n_1, n_2, \dots, n_m (or their expansions) are *non-interfering*, if $k_{i+1} - K_i \geq 2$ for all $i = 1, 2, \dots, m-1$. More informally, non-interfering integers have digital blocks that do not overlap. For non-interfering n_1, n_2, \dots, n_m we have

$$s_F(n_1 + n_2 + \dots + n_m) = s_F(n_1) + s_F(n_2) + \dots + s_F(n_m),$$

a fact that will be central in our combinatorial constructions.

We will often use the following simple lemmas that follow easily from Binet's formula (see e.g. [14, 10.14]) or straightforward inductive arguments.

Denote by $(L_j)_{j \geq 0}$ the sequence of Lucas numbers defined by $L_0 = 2$, $L_1 = 1$ and

$$(9) \quad L_{j+2} = L_{j+1} + L_j, \quad j \geq 0.$$

Lemma 2.2. *For $k > l \geq 1$ we have*

$$(10) \quad F_k - F_l = \begin{cases} F_{k-1} + F_{k-3} + \dots + F_{l+1}, & \text{if } k \equiv l \pmod{2}, \\ F_{k-1} + F_{k-3} + \dots + F_{l+2} + F_{l-1}, & \text{if } k \not\equiv l \pmod{2}. \end{cases}$$

Lemma 2.3. *For $k \geq l \geq 0$ we have*

$$(11) \quad L_k L_l = L_{k+l} + (-1)^l L_{k-l},$$

$$(12) \quad L_k F_l = F_{k+l} - (-1)^l F_{k-l}.$$

Our constructions are based on formulas for powers of Lucas numbers. These formulas are proven in [10] and can be obtained by an easy inductive argument on h and (11).

Proposition 2.4. *For all $k \geq 0$ and $h \geq 1$ we have*

$$(13) \quad L_k^h = \sum_{i=0}^{(h-1)/2} \binom{h}{i} (-1)^{ik} L_{(h-2i)k}, \quad h \text{ odd},$$

$$L_k^h = \sum_{i=0}^{h/2-1} \binom{h}{i} (-1)^{ik} L_{(h-2i)k} + \binom{h}{h/2} (-1)^{kh/2}, \quad h \text{ even}.$$

Proof. For the convenience of the reader we here outline the induction step from h odd to $h + 1$ even, the other case is similar. We have

$$\begin{aligned}
L_k^{h+1} &= \sum_{i=0}^{(h-1)/2} \binom{h}{i} (-1)^{ik} L_{(h-2i)k} L_k \\
&= \sum_{i=0}^{(h-1)/2} \binom{h}{i} (-1)^{ik} (L_{(h-2i+1)k} + (-1)^k L_{(h-2i-1)k}) \\
&= \sum_{i=1}^{(h-1)/2} \left(\binom{h}{i} + \binom{h}{i-1} \right) (-1)^{ik} L_{(h+1-2i)k} \\
&\quad + L_{(h+1)k} + \binom{h}{\frac{h-1}{2}} (-1)^{(\frac{h+1}{2})k} L_0,
\end{aligned}$$

and the statement follows by observing that $\binom{h}{i} + \binom{h}{i-1} = \binom{h+1}{i}$, $2\binom{h}{(h-1)/2} = \binom{h+1}{(h+1)/2}$. \square \square

It is well-known that multiples of Fibonacci numbers mF_k have finite length as $k \rightarrow \infty$, a fact that is true more generally for any numeration system defined by a Pisot recurrence [12] [13]. Also, it is well-known that the digits in the finite block only get shifted when k increases. Multiples of Lucas numbers mL_k have finite length, too, with the similar behaviour about shifted digits. We make this explicit in the following proposition.

Here and throughout, for an integer $m \geq 1$ we denote by $u = u(m) \geq 0$ the unique integer u with $F_{2u} < m \leq F_{2u+2}$. We set $v = v(m) = 1$ if $F_{2u+1} \leq m \leq F_{2u+2}$ and 0 else. From Binet's formula we see that

$$(14) \quad u(m) \leq \frac{\log(m+1)}{2 \log \varphi} + \frac{\log \sqrt{5}}{2 \log \varphi} < \frac{\log m}{2 \log \varphi} + 2.$$

Proposition 2.5. *There exist non-adjacent*

$$\varepsilon_{-(2u+1)}(m), \varepsilon_{-2u}(m), \dots, \varepsilon_{2u+v}(m) \in \{0, 1\}, \quad \varepsilon_{-(2u+1)}(m) = \varepsilon_{2u+v}(m) = 1,$$

depending only on m , such that for all $k \geq 2u + 3$,

$$(15) \quad mL_k = \sum_{j=-(2u+1)}^{2u+v} \varepsilon_j(m) F_{k+j}.$$

Proof. We prove (15) by induction on $s_F(m)$. If $m = F_{2l+1}$ for $l \geq 1$ then $u = l$, $v = 1$ and by (12),

$$F_{2l+1} L_k = F_{k+2u+1} + F_{k-(2u+1)}.$$

If $m = F_{2l}$ for $l \geq 1$ then $u = l - 1$, $v = 1$ and by (10),

$$\begin{aligned}
F_{2l} L_k &= F_{k+2l} - F_{k-2l} \\
&= F_{k+2u+1} + F_{k+2u-1} + \dots + F_{k-(2u+1)}.
\end{aligned}$$

For the induction step, set $m' = m + F_l$ where $0 < m < F_{l-1}$ and assume (15) for $s_F(m)$. We can suppose that $l \geq 4$. Write $u' = u(m + F_l)$ and $v' = v(m + F_l)$. Since

m' is not a Fibonacci number we have $u' = \lfloor l/2 \rfloor$ and $v' = (l \bmod 2)$. Therefore, $k \geq 2u' + 3$ implies $k > l$ and by (12) (we write $u = u(m)$ and $v = v(m)$),

$$\begin{aligned} m'L_k &= (F_{k+2u+v} + w + F_{k-(2u+1)}) + (F_{k+l} - (-1)^l F_{k-l}) \\ (16) \quad &= (F_{k+l} + F_{k+2u+v}) + w + (F_{k-(2u+1)} - (-1)^l F_{k-l}), \end{aligned}$$

where w is either zero or a positive integer whose most significant digit is at a position $\leq k + 2u + v - 2$ and least significant digit at a position $\geq k - (2u + 1) + 2$. Note also that $l \geq (2u + v) + 2$. Thus

$$k + 2u' + v' = k + 2 \left\lfloor \frac{l}{2} \right\rfloor + (l \bmod 2) = k + l$$

indeed denotes the position of the most significant digit in (16). It remains to prove that

$$k - (2u' + 1) = k - 2 \left\lfloor \frac{l}{2} \right\rfloor - 1$$

is the position of the least significant digit of $m'L_k$. We distinguish two cases. If l is odd then $l \geq 2u + v + 2$ implies that $l \geq 2u + 3$ and $m'L_k = F_{k+l} + F_{k+2u+v} + w + F_{k-(2u+1)} + F_{k-l}$. Therefore, the least significant digit is at position

$$k - l = k - 2 \left\lfloor \frac{l}{2} \right\rfloor - 1$$

and we are done. Now, assume that l is even. Then by (10),

$$F_{k-(2u+1)} - F_{k-l} = F_{k-(2u+1)-1} + F_{k-(2u+1)-3} + \cdots + F_{k-l+2} + F_{k-l-1},$$

and the least significant digit of $m'L_k$ is at position

$$k - l - 1 = k - 2 \left\lfloor \frac{l}{2} \right\rfloor - 1.$$

This finishes the proof. \square

The next lemma shows that when we subtract a “small” positive integer from mL_k then we obtain a large sum of digits when k gets large.

Lemma 2.6. *For all $m \geq 1$, $k \geq 2u + 3$ and $1 \leq b < F_{k-2u-2}$ we have*

$$(17) \quad s_F(mL_k - b) \geq \frac{k}{2} - u - \frac{1}{2} - s_F(b - 1).$$

For integers $m_1, m_2, k_1, k_2 \geq 1$ with

$$2k_1 \geq 2u(m_1) + 3, \quad 2k_2 \geq 2u(m_2) + 3, \quad \text{and} \quad m_2 L_{2k_2} < F_{2k_1 - 2u(m_1) - 2}$$

we have

$$\begin{aligned} (18) \quad s_F(m_1 L_{2k_1} + m_2 L_{2k_2}) &= C_3(m_1, m_2), \\ s_F(m_1 L_{2k_1} - m_2 L_{2k_2}) &= k_1 - k_2 + C_4(m_1, m_2), \end{aligned}$$

where $C_3(m_1, m_2)$ and $C_4(m_1, m_2)$ do not depend on k_1 and k_2 .

Proof. Consider

$$s_F(mL_k - b) = s_F((mL_k - F_{k-2u-1}) + (F_{k-2u-1} - b)).$$

Let $l \geq 3$ be such that $F_{l-1} \leq b < F_l$ and take $l' \in \{l, l+1\}$ with $l' \equiv k - 2u - 1 \pmod{2}$. Then

$$(19) \quad F_{k-2u-1} - F_{l'} = F_{l'+1} + F_{l'+3} + \cdots + F_{k-2u-2}$$

and by subadditivity,

$$(20) \quad \frac{l'}{2} - 1 \leq s_F(F_{l'} - 1) = s_F((F_{l'} - b) + (b - 1)) \leq s_F(F_{l'} - b) + s_F(b - 1).$$

The least significant digit of mL_k is at position $k - 2u - 1$, thus the least significant digit of $mL_k - F_{k-2u-1}$ is at a position $\geq k - 2u + 1$. By (19), the most significant digit of $F_{k-2u-1} - F_{l'}$ can be found at position $k - 2u - 2$. Therefore, the Zeckendorf expansions of $mL_k - F_{k-2u-1}$ and of $F_{k-2u-1} - F_{l'}$ involve different, non-interfering blocks of digits and

$$s_F(mL_k - F_{l'}) = s_F(mL_k - F_{k-2u-1}) + s_F(F_{k-2u-1} - F_{l'}).$$

Furthermore, the least significant digit of $F_{k-2u-1} - F_{l'}$ is at position $l' + 1$ and $F_{l'} - b < F_{l'}$. By using (20) and Proposition 2.5 we conclude

$$(21) \quad \begin{aligned} s_F(mL_k - b) &= s_F(mL_k - F_{k-2u-1}) + s_F(F_{k-2u-1} - F_{l'}) + s_F(F_{l'} - b) \\ &= s_F(mL_k) - 1 + \frac{(k - 2u - 1) - l'}{2} + s_F(F_{l'} - b) \\ &\geq 1 + \frac{k - 2u - 1 - l'}{2} + \frac{l'}{2} - 1 - s_F(b - 1) \\ &\geq -u + \frac{k}{2} - \frac{1}{2} - s_F(b - 1). \end{aligned}$$

The second statement follows from Proposition 2.5 and (21). \square \square

3. GELFOND'S PROBLEM

For $k \geq 1$ consider

$$(22) \quad t(k) = m_3 L_{6k} - m_2 L_{4k} + m_1 L_{2k} + m_0 L_0,$$

with positive real parameters $m_0, m_1, m_2, m_3 \geq 1$ as a linear form in even-indexed Lucas numbers. For $l \geq 1$ set

$$(23) \quad T_l(k) = t(k)^l = \sum_{i=0}^{3l} c_i L_{2ik}$$

as the formal expansion of the l -th power of $t(k)$ in terms of even-indexed Lucas numbers. We obtain this expansion by using relation (11) a number of $(l - 1)$ times. The following lemma is the key result in the proof of Theorem 1.1 and is an analogue to Lemma 2.1 of [21].

Lemma 3.1. *Let $M > 1$. For all $m_3, m_2, m_1, m_0 \in \mathbb{R}$ with*

$$(24) \quad 1 \leq m_0, m_1, m_3 < M, \quad 0 < m_2 < \frac{1}{l^3 \cdot (32M)^l}$$

we have $c_{3l} > 0$, $c_{3l-1} < 0$ and

$$(25) \quad c_i > 0 \quad \text{for } i = 0, 1, 2, \dots, 3l - 2.$$

Moreover, for $i = 0, 1, \dots, 3l$,

$$(26) \quad |c_i| < \frac{1}{2}(8M)^l.$$

Proof. The coefficient c_i of L_{2ik} in (23) is bounded above in absolute value by the coefficient of L_{2ik} in the expansion of

$$(ML_{6k} + ML_{4k} + ML_{2k} + ML_0)^l = M^l(L_{6k} + L_{4k} + L_{2k} + L_0)^l.$$

Expanding $(L_{6k} + L_{4k} + L_{2k} + L_0)^l$ with the help of (11) gives a total of $2^{l-1} \cdot 4^l$ summands, so that each individual coefficient in the expansion of (23) is bounded by

$$(27) \quad M^l 2^{l-1} \cdot 4^l = \frac{1}{2}(8M)^l.$$

We further have $c_{3l} = m_3^l > 0$ and $c_{3l-1} = -lm_2m_3^{l-1} < 0$. Suppose now that $0 \leq i \leq 3l-2$ and consider the coefficient of L_{2ik} in

$$(28) \quad T_l(k) = (m_3L_{6k} + m_1L_{2k} + m_0L_0)^l + r(k),$$

where

$$(29) \quad r(k) = \sum_{j=1}^l \binom{l}{j} (-m_2L_{4k})^j (m_3L_{6k} + m_1L_{2k} + m_0L_0)^{l-j}.$$

Consider the first summand in (28). Since $m_3, m_1, m_0 \geq 1$ the coefficient of L_{2ik} in the expansion of $(m_3L_{6k} + m_1L_{2k} + m_0L_0)^l$ is ≥ 1 . Note also that by (11) all Lucas numbers with index $0, 2k, 4k, \dots, (6l-4)k$ appear in the expansion of the first term. Set

$$r(k) = \sum_{i=0}^{3l-1} d_i L_{2ik}.$$

We want to show that for sufficiently small $m_2 > 0$ the coefficient of L_{2ik} in the first summand in (28) is dominant. To this end, we assume $m_2 < 1$ so that $m_2 \geq m_2^j$ for $1 \leq j \leq l$. The term L_{4k}^j , $1 \leq j \leq l$, in (29) can be written as a linear form of at most $(l+1)$ even-indexed Lucas numbers where each individual coefficient is bounded above by 2^{l-1} . Similarly, the expansion of $(m_3L_{6k} + m_1L_{2k} + m_0L_0)^{l-j}$ involves at most $(l-1) \cdot 3 + 1 = 3l-2$ even-indexed Lucas numbers with each coefficient being bounded by $\frac{1}{2}(8M)^{l-1}$. Using $\binom{l}{j} < 2^l$ and $M > 1$ we hence find that

$$|d_i| < l 2^l m_2 \cdot 2^{l-1} \cdot \frac{1}{2}(8M)^{l-1} \cdot 2(l+1)(3l-2) < m_2 l^3 (32M)^l.$$

This means that if $m_2 < l^{-3} \cdot (32M)^{-l}$ then the terms

$$L_0, L_{2k}, L_{4k}, \dots, L_{2(3l-2)k}$$

in the expression of $T_l(k)$ have positive coefficients. □ □

We are now ready for the proof of Theorem 1.1.

Theorem 1.1. We consider first the case of monomials $p(x) = x^h$. Let $\mu \geq 1$ be an integer such that

$$(30) \quad \varphi^\mu > h^3 \varphi (32\varphi)^h$$

and take $M = \varphi > 1$ in Lemma 3.1. Then, for all integers m_0, m_1, m_2, m_3 with

$$(31) \quad \varphi^{\mu-1} \leq m_0, m_1, m_3 < \varphi^\mu, \quad 1 \leq m_2 < \varphi^\mu / (h^3 \varphi (32\varphi)^h),$$

the expansion of $T_h(k) = (t(k))^h$ has all positive integral coefficients with the only exception of the coefficient of $L_{2(3h-1)k}$ which is negative. Observe that for all μ

with (30) the intervals for m_0, m_1, m_2, m_3 in (31) are non-empty. Let $k \in \mathbb{Z}$ be such that

$$(32) \quad k > h\mu + 5h + 3.$$

The inequalities (32) and (26) yield that

$$(33) \quad \varphi^k > \varphi^{h\mu} \cdot \varphi^{5h} > (8\varphi^\mu)^h > \frac{1}{2}(8\varphi^\mu)^h > |c_i|, \quad \text{for all } i = 0, 1, \dots, 3h,$$

where c_i here denotes the coefficient of L_{2ik} in $T_h(k)$. Roughly speaking, if k is sufficiently large, then the digital structure of $T_h(k)$ will split according to Proposition 2.5. We will now make this process explicit. Consider

$$t(k)^h = c_{3h}L_{6hk} - |c_{3h-1}|L_{2(3h-1)k} + \sum_{i=0}^{3h-2} c_i L_{2ik}.$$

In order to apply Proposition 2.5 we assume that

$$(34) \quad 2ik \geq 2u(|c_i|) + 3 \quad \text{for } i = 1, 2, \dots, 3h.$$

From (14) and (33) it is easy to see that this condition is verified for all k with (32). Second, in order to have noninterfering blocks of digits we want to have

$$(35) \quad |c_i|L_{2ik} < F_{2(i+1)k-2u(|c_{i+1}|)-2} \quad \text{for } i = 0, 1, \dots, 3h-1.$$

It is again direct (but rather tedious) calculation that this condition is verified whenever k satisfies (32). We therefore get by Lemma 2.6 that

$$\begin{aligned} (36) \quad s_F(t(k)^h) &= s_F(c_{3h}L_{6hk} - |c_{3h-1}|L_{2(3h-1)k}) + \sum_{i=0}^{3h-2} s_F(c_i L_{2ik}) \\ &= 3hk - (3h-1)k + C_3(c_{3h}, |c_{3h-1}|) + C_4(c_{3h-2}, c_{3h-3}, \dots, c_0) \\ &= k + C_5(c_{3h}, |c_{3h-1}|, \dots, c_1, c_0), \end{aligned}$$

where C_3, C_4 and C_5 do not depend on k . If we fix a quadruple (m_0, m_1, m_2, m_3) in the ranges (31), the quantity C_5 does not depend on k and is constant whenever k satisfies (32). We therefore find that

$$(37) \quad s_F(t(k)^h), \quad k = h\mu + 5h + 4, h\mu + 5h + 5, \dots, h\mu + 5h + m + 3,$$

runs through a complete set of residues mod m . Hence, in any case, we hit a fixed arithmetic progression mod m (which might be altered by C_5) for some k with

$$h\mu + 5h + 4 \leq k \leq h\mu + 5h + m + 3.$$

Thus, for μ with (30) and by (31) we find $\gg_{h,m} \varphi^{4\mu}$ distinct integers that are all $\ll_{h,m} \varphi^{\mu(6h+1)}$ (note that the coefficient “6” comes from the use of L_{6k} instead of the factor “3” and the use of q^{3k} in [21]). Thus, we get the statement for monomials.

The general case of a polynomial $p(x) = a_h x^h + \dots + a_0$ of degree $h \geq 2$ with $p(\mathbb{N}) \subset \mathbb{N}$ follows easily from what we have already proven. We may assume that all coefficients a_i , $0 \leq i \leq h$, are positive since otherwise we can find $e = e(p) \in \mathbb{N}$ depending only on p such that $p(x + e)$ has positive coefficients only. Write

$$t(k)^l = \sum_{i=0}^{3l} c_i^{(l)} L_{2ik}, \quad l = 0, 1, \dots, h.$$

The coefficients $c_i^{(l)}$, $i = 0, 1, \dots, 3l - 2$ and $c_{3l}^{(l)}$ are positive by Lemma 3.1. Obviously, we have that $c_{3h-1}^{(h)} < 0$. It is therefore sufficient to ensure that for all $l = 1, \dots, h - 1$ one can achieve $|c_{3l-1}^{(l)}| < c_{3l-1}^{(l+1)}$. As a consequence, the negative Lucas term in the expansion of $t(k)^l$ is then compensated by the corresponding positive term in $t(k)^{l+1}$. By (31), we have that $|c_{3l-1}^{(l)}| \ll_l \varphi^{\mu l}$ and $|c_{3l-1}^{(l+1)}| \gg_l \varphi^{\mu(l+1)}$, as $\mu \rightarrow \infty$. This means that for a given polynomial p there is $\mu_0 = \mu_0(p)$ such that for all $\mu \geq \mu_0$ the polynomial $p(t(k))$ will have all but one negative coefficient in the expansion of even-indexed Lucas numbers. Now, exactly the same reasoning as for the monomial case works, and gives the desired general result. \square \square

4. STOLARSKY'S PROBLEM

Theorem 1.2. First, assume that $p(x) = x^h$. We define

$$(38) \quad t_m(k) = mL_{4k} + mL_{2k} - L_0,$$

where $m \geq 5$ is a fixed integer. With the help of (11) we calculate

$$(39) \quad \begin{aligned} t_m(k)^2 &= m^2 L_{8k} + 2m^2 L_{6k} + (m^2 - 4m)L_{4k} \\ &\quad + (2m^2 - 4m)L_{2k} + (2m^2 + 2)L_0, \end{aligned}$$

and

$$\begin{aligned} t_m(k)^3 &= m^3 L_{12k} + 3m^3 L_{10k} + 3m^2(m-2)L_{8k} \\ &\quad + 4m^2(m-3)L_{6k} + m(7m^2 - 6m + 12)L_{4k} \\ &\quad + 3m(3m^2 - 4m + 4)L_{2k} + (3m^3 - 12m^2 - 4)L_0. \end{aligned}$$

Since $m \geq 5$, all coefficients in these two linear forms in even-indexed Lucas numbers are positive. Since $t_m(x)^h = (t_m(x)^2)^{h_1} \cdot (t_m(x)^3)^{h_2}$ for some nonnegative integers h_1, h_2 , we get that

$$(40) \quad t_m(x)^h = \sum_{i=0}^{2h} c_i(m) L_{2ik}$$

has $c_i(m) > 0$ for all $i = 0, 1, \dots, 4h$. Moreover, each $c_i(m)$ is bounded from above by the corresponding coefficient of L_{2ik} in the expansion of $(mL_{4k} + mL_{2k} + mL_0)^h = m^h(L_{4k} + L_{2k} + L_0)^h$ which is bounded by $m^h \cdot 2^{h-1} \cdot 3^h = \frac{1}{2}(6m)^h$, thus

$$0 < c_i(m) < (6m)^h.$$

This means that for each fixed $m \geq 5$ the linear form (40) is made of summands that by Proposition 2.5 have finite length. Therefore, as in the previous section, for sufficiently large k these summands correspond to digital blocks that do not interfere, hence we can calculate the overall sum of digits by breaking down the calculation to the sum of digits of the summands. This gives, as $k \rightarrow \infty$,

$$(41) \quad s_F(t_m(k)^h) = \sum_{i=0}^{2h} s_F(c_i(m)L_{2ik}) \ll (2h+1) \log((6m)^h),$$

and, on the other hand, by Lemma 2.6,

$$(42) \quad \begin{aligned} s_F(t_m(k)) &= s_F(mL_{4k}) + s_F(mL_{2k} - L_0) \\ &\geq 1 + k - u(m) - \frac{1}{2} - 1 \gg k. \end{aligned}$$

This shows that

$$\lim_{k \rightarrow \infty} \frac{s_F(t_m(k)^h)}{s_F(t_m(k))} = 0.$$

More precisely, since $k = O(\log(t_m(k)))$ for fixed m , we get by (41) and (42) that there exists a constant $C_2 = C_2(h)$ such that

$$\frac{s_F(n^h)}{s_F(n)} < \frac{C_2}{\log n}$$

has infinitely many solutions in positive integers n .

The step to the consideration of a general polynomial $p(x)$ is very similar as before. Without loss of generality assume that $p(x)$ has only positive coefficients. The coefficients in the expansion of $t_m(k)^l$, for $l = 2, 3, \dots, h$ as linear forms in Lucas numbers are positive by construction, so the coefficients in $p(t_m(k))$ in the expansion as a linear form in Lucas numbers $L_0, L_{2k}, \dots, L_{4kh}$ are positive with the only possible exception of L_0 which possibly stems from the contributing term $t_m(k)^1 = mL_{4k} + mL_{2k} - L_0$. But this is taken care of by choosing $m \geq m_0(p)$ sufficiently large, in order to have enough compensation for this negative term, say, in the power $t_m(k)^2$ (note that in (39) the coefficient of L_0 is $2m^2 + 2$). We then end up with a linear form with only positive coefficients and we conclude as before.

For the first part of Theorem 1.2, we set $r(k) = L_{2k+1}$. Since $r(k) = F_{2k+2} + F_{2k}$ we have $s_F(r(k)) = 2$ for all $k \geq 1$. By (11) and an easy induction we see that $L_{2k+1}^h < L_{h(2k+1)}$ for all $h \geq 2$. Therefore, by (13) and a trivial estimate of the summands,

$$L_{2k+1}^h = L_{h(2k+1)} - b,$$

where

$$0 < b < 2^h L_{(h-2)(2k+1)} \left(\frac{h}{2} + 1 \right).$$

In particular, for $k \geq k_0(h)$ sufficiently large, we have $b < F_{(h-1)(2k+1)}$ and by (17),

$$\begin{aligned} s_F(L_{h(2k+1)} - b) &\geq \frac{(2k+1)h}{2} - 1 - \frac{1}{2} - s_F(b-1) \\ &\geq \frac{(2k+1)h}{2} - \frac{3}{2} - \frac{(h-1)(2k+1)}{2} = k-1. \end{aligned}$$

Therefore, we get $s_F(r(k)^h)/s_F(r(k)) \geq \frac{k-1}{2}$, and conclude that there is a constant $C_1 = C_1(h) > 0$ such that

$$\frac{s_F(n^h)}{s_F(n)} > C_1 \log n$$

has infinitely many solutions in positive integers n . A minor variation of the above argument gives the result for general polynomials $p(x)$. In fact, we see that $s_F(r(k) + e) = O_p(1)$ where $e = e(p)$ is such that $p(x + e)$ only has positive coefficients. \square \square

Corollary 1.3. Set $m = 5$ in (38). Then, for $k \geq 4$, we get

$$s_F(t_5(k)) = s_F(F_{4k+5} + F_{4k-5} + F_{2k+5} + F_{2k-5} - 2) = k-1.$$

Therefore, by (41) and the reasoning from above, there exists $C = C(p)$ such that $s_F(p(t_5(k))) \leq C$ for all $k \geq 4$. This gives the general case of Theorem 1.3. In the case of $p(x) = x^h$, we compute the quantities explicitly. We have $0 < c_i(5) < 30^h$. If we assume that k is such that $2k \geq (2u+1) + 1 + (2u+1)$ then by Proposition 2.5

the terms in (40) will not interfere. A calculation then shows that it suffices to take $k \geq 8h + 2 \geq 2u(30^h) + 2$ to get

$$s_F(t_5(k)^h) \leq (2h + 1) \cdot (2u + 2) \leq (2h + 1)(8h + 2).$$

Finally, one has $t_5(k) < F_{4k+6} < \frac{\varphi^{10}}{\sqrt{5}} \cdot \varphi^{4(k-1)} + 1$. This finishes the proof of Corollary 1.3. \square \square

5. LINDSTRÖM'S PROBLEM

Theorem 1.4. The liminf statement follows directly from inequality (4) of Theorem 1.2. In fact, we have

$$\frac{s_F(p(n))}{\log_\varphi p(n)} = \frac{s_F(p(n))}{s_F(n)} \cdot \frac{s_F(n)}{\log_\varphi p(n)},$$

and the second factor is $O(1)$ for any fixed $p(x)$.

For the limsup result we use a construction similar to the one used by Lindström [17] (we use convenient choices of the parameters to simplify the argument). Again, we first consider the case of $p(x) = x^h$. For $k \geq h$ put

$$t(k) = k^{h+2} L_{2k(k+1)} - \sum_{i=1}^k L_{2ik} + k^{h+2} L_0,$$

and consider

$$t(k)^h = \sum_{i=0}^{(k+1)h} c_i L_{2ik}.$$

We will show here that there is $k_0(h)$ such that for all $k \geq k_0(h)$ and $i = 0, 1, \dots, (k+1)h$,

$$(43) \quad \begin{aligned} c_i &> 0, & \text{if } i \equiv 0 \pmod{k+1}; \\ c_i &< 0, & \text{if } i \not\equiv 0 \pmod{k+1}. \end{aligned}$$

The argument is similar to the one used to prove Lemma 3.1. For convenience set $m = k^{h+2}$, $l = k + 1$ and write

$$(44) \quad t(k)^h = m^h (L_{2kl} + L_0)^h - hm^{h-1} \left(\sum_{i=1}^{l-1} L_{2ik} \right) (L_{2kl} + L_0)^{h-1}$$

$$(45) \quad + \sum_{j=2}^h \binom{h}{j} m^{h-j} \left(- \sum_{i=1}^{l-1} L_{2ik} \right)^j (L_{2kl} + L_0)^{h-j}.$$

The first term on the right hand side of (44) can be written as a linear combination of Lucas number with index $0, 2kl, 4kl, \dots, 2hkl$ having only positive coefficients. Moreover, each coefficient $\geq m^h$. The second term in (44) (with the minus sign in front) is a linear combination of Lucas number with index in

$$\{2jk : 0 \leq j \leq hk\} \setminus \{2klj : 0 \leq j \leq h\}.$$

having only negative coefficients. Each coefficient is $\geq hm^{h-1}$ in modulus. To prove (43) it is therefore sufficient to show that the expansion of (45) as a sum of

even-indexed Lucas numbers has coefficients that are $< \min(m^h, hm^{h-1})$ in modulus. A similar argument as in Lemma 3.1 shows that each coefficient in the expansion of (45) is bounded in modulus by

$$(h-1)2^h m^{h-2} \cdot 2^{h-1}(l-1)^h \cdot 2^{(h-3)+(h-2)} \cdot 2 \cdot ((l-1)h+1) \cdot ((h-2)+1).$$

Therefore, if we assume $k \geq h$ (thus, $\min(m^h, hm^{h-1}) = hm^{h-1}$), it is sufficient to have

$$(h-1)^2 2^{4h-5} m^{h-2} (l-1)^h (h(l-1)+1) < hm^{h-1}.$$

It is easily seen that this holds true for $k \geq k_1(h)$. Next, in order to split the digital structure and to have noninterfering terms, we need to ensure the condition in Lemma 2.6. We have $|c_i| < m^h \cdot 2^{h-1}(l+1)^h$ (compare with (33)) and by (14) it is therefore enough to show that for sufficiently large k ,

$$2k \geq 2 \left(\frac{\log(m^h \cdot 2^{h-1}(l+1)^h)}{\log 2\varphi} + 2 \right) + 3.$$

This is true for all sufficiently large $k \geq k_2(h)$. Lemma 2.6 yields that, as $k \rightarrow \infty$,

$$s_F(t(k)^h) = hk^2 + O_h(1),$$

$$\log_\varphi(t(k)^h) \leq h \cdot 2k(k+1) + O_h(\log k).$$

We therefore conclude that

$$\liminf_{k \rightarrow \infty} \frac{s_F(t(k)^h)}{\log_\varphi(t(k)^h)} \geq \liminf_{k \rightarrow \infty} \frac{hk^2 + O_h(1)}{2hk(k+1) + O_h(\log k)} = 1/2,$$

which proves the statement for monomials $p(x) = x^h$. The consideration of general polynomials is straightforward here, since the conditions (43) do not depend on h and therefore the signs of the c_i 's for different powers match up. We leave the details to the reader. \square \square

ACKNOWLEDGEMENT

The author is grateful to C. Dartyge for valuable remarks on an earlier version of the paper.

REFERENCES

- [1] J.-P. Allouche, J. Shallit, *Automatic sequences: Theory, applications, generalizations*, Cambridge University Press, Cambridge, 2003.
- [2] Y. Bugeaud, F. Luca, M. Mignotte, S. Siksek, Fibonacci numbers at most one away from a perfect power, *Elem. Math.* **63** (2008), no. 2, 65–75.
- [3] Y. Bugeaud, M. Mignotte, S. Siksek, Classical and modular approaches to exponential Diophantine equations. I. Fibonacci and Lucas perfect powers, *Ann. of Math. (2)* **163** (2006), no. 3, 969–1018.
- [4] J. Coquet, P. Van Den Bosch, A summation formula involving Fibonacci digits, *J. Number Theory* **22** (1986), no. 2, 139–146.
- [5] C. Dartyge, G. Tenenbaum, Congruences de sommes de chiffres de valeurs polynomiales, *Bull. London Math. Soc.* **38** (2006), no. 1, 61–69.
- [6] M. Drmota, C. Mauduit, J. Rivat, The sum-of-digits function of polynomial sequences, *J. London Math. Soc.* **84** (2011), 81–102.
- [7] M. Drmota, J. Rivat, The sum of digits function of squares, *J. London Math. Soc.* **72** (2005), 273–292.
- [8] M. Drmota, M. Skalba, The parity of the Zeckendorf sum-of-digits function, *Manuscripta Math.* **101** (2000), no. 3, 361–383.

- [9] M. Drmota, W. Steiner, The Zeckendorf expansion of polynomial sequences, *J. Théor. Nombres Bordeaux* **14** (2002), no. 2, 439–475.
- [10] H.H. Ferns, Products of Fibonacci and Lucas Numbers, *Fibonacci Quart.*, **7**, No. 1 (Feb. 1969), 1–13.
- [11] A. O. Gelfond, Sur les nombres qui ont des propriétés additives et multiplicatives données, *Acta Arith.* **13** (1967/1968), 259–265.
- [12] P. J. Grabner, R. F. Tichy, I. Nemes, A. Pető, On the least significant digit of Zeckendorf expansions, *Fibonacci Quart.* **34** (1996), 147–151.
- [13] P. J. Grabner, R. F. Tichy, I. Nemes, A. Pető, Generalized Zeckendorf expansions, *Appl. Math. Lett.* **7** (1994), 25–28.
- [14] G. H. Hardy, E. M. Wright, *An Introduction to the Theory of Numbers*, Sixth Edition, Oxford University Press, 2008.
- [15] K. G. Hare, S. Laishram, T. Stoll, Stolarsky’s conjecture and the sum of digits of polynomial values, *Proc. Amer. Math. Soc.* **139** (2011), 39–49.
- [16] K. G. Hare, S. Laishram, T. Stoll, The sum of digits of n and n^2 , *IJNT* **7** (2011), 1737–1759.
- [17] B. Lindström, On the binary digits of a power, *J. Number Theory* **65** (1997), 321–324.
- [18] C. Mauduit, J. Rivat, La somme des chiffres des carrés, *Acta Math.* **203** (2009), 107–148.
- [19] T. Rivoal, On the bits counting function of real numbers, *J. Aust. Math. Soc.* **85** (2008), no. 1, 95–111.
- [20] K. B. Stolarsky, The binary digits of a power, *Proc. Amer. Math. Soc.* **71** (1978), 1–5.
- [21] T. Stoll, The sum of digits of polynomial values in arithmetic progressions, *Functiones et Approximatio Commentarii Mathematici*, accepted (Oct 21, 2011); preprint available from the author’s webpage.
- [22] S. Wagner, Numbers with fixed sum of digits in linear recurrent number systems, *The Ramanujan Journal* **14** (2007), 43–68.
- [23] E. Zeckendorf, Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas, *Bull. Soc. Roy. Sci. Liège* **41** (1972), 179–182.

E-mail address: `stoll@iml.univ-mrs.fr`