THE PARITY OF THE SUM-OF-DIGITS-FUNCTION OF GENERALIZED ZECKENDORF REPRESENTATIONS*

Michael Drmota and Johannes Gajdosik
Department of Discrete Mathematics, Technical University of Vienna
Wiedner Hauptstraße 8-10, A-1040, Vienna
(Submitted June 1996)

1. INTRODUCTION

Let $G=(G_n)$ be a strictly increasing sequence of positive integers with $G_1=1$. Then every nonnegative integer n has a digital expansion

$$n=\sum_{i\geq 1}\varepsilon_i G_i$$

with respect to basis G, where the digits $\varepsilon_i=\varepsilon_i(n)\geq 0$ are integers. This digital expansion is unique, when one assumes that the digits ε_i are chosen in such a way that the digital sum $\sum_{i\geq 1}\varepsilon_i$ is as small as possible; in this case, we will call the digital expansion a *proper digital expansion*. It is easy to see that the following algorithm provides this expansion.

1. For $n=0$, we have $\varepsilon_i(n)=0$ for every $i\geq 1$.
2. If $G_j\leq n<G_{j+1}$ and $n'=n-G_j$ has the proper expansion $n'=\sum_{i\geq 1}\varepsilon_i' G_i$, then the expansion of $n=\sum_{i\geq 1}\varepsilon_i G_i$ is given by $\varepsilon_i=\varepsilon_i'$ for $i\neq j$ and by $\varepsilon_j=\varepsilon_j'+1$.

The most prominent digital expansions are related to linear recurring sequences $G=(G_n)$, e.g., the binary (resp. the q-ary) expansion relies on $G_n=2^{n-1}$ (resp. on $G_n=q^{n-1}$). If G_n are the Fibonacci numbers, i.e., $G_n=F_{n+1}$, then we obtain the Zeckendorf expansion.

For each digital expansion with respect to a basis G, we can define a partial order in a quite natural way. We will say $a\leq_G b$ if and only if $\varepsilon_i(a)\leq\varepsilon_i(b)$ for every $i\geq 1$. It is well known that for every partial order there is a Möbius function (see [10], [13]). Let $s_G(n)$ denote the sum of digits of n. Then it will turn out that the Möbius function μ_G of a digital expansion to a basis G is given by $\mu_G(n)=(-1)^{s_G(n)}$ if $\max_{i\geq 1}\varepsilon_i(n)\leq 1$ and by $\mu_G(n)=0$ otherwise.

If G is a proper linear recurring sequence and if the initial conditions of G are properly chosen (see Section 3), then

$$M_G(N):=\sum_{n=0}^{N-1}\mu_G(n)$$

is either bounded or

$$M_G(N)=S_G(N):=\sum_{n=0}^{N-1}(-1)^{s_G(n)},$$

which we will see from calculating the Möbius function in Section 2. (We always define empty sums to be zero, i.e., $M_G(N)=S_G(N):=0$ for $N\leq 0$.)

---

* This work was supported by the Austrian Science Foundation, grant P10187-PHY. This paper, presented at the Seventh International Research Conference held in Graz, Austria, in July 1996, was scheduled to appear in the Conference Proceedings. However, due to limitations placed by the publisher on the number of pages allowed for the Proceedings, we are publishing the article in this issue of *The Fibonacci Quarterly* to assure its presentation to the widest possible number of readers in the mathematics community.

In Section 3 we will formulate conditions for G, under which we will be able to derive formulas for $S_G(N)$. We will also obtain a recursive formula for the generating function of $S_G(G_n)$, which we will analyze in Section 4 in order to obtain asymptotic information about $S_G(N)$.

Our main interest lies in the distribution of the $S_G(N)$ (resp. $M_G(N)$) when $0 \le N < m$ for large m. This means that we count the number of times $S_G(N)$ takes a certain value k when $0 \le N < m$: let $d_m(k) := |\{0 \le N < m : S_G(N) = k\}|$ be this number and let X_m be a random variable with probability distribution $\mathbf{P}(X_m = k) = d_m(k)/m$. Then we are interested in the asymptotic distribution of X_m for $m \to \infty$. Depending on the nature of the recurrence relation for G, we will observe significantly different behavior of X_m. First, we distinguish two cases:

1. either $S_G(G_n)$ is bounded for all initial conditions of G (Section 4.1), or
2. there are initial conditions of G such that $S_G(G_n)$ is unbounded (Section 4.2).

Since we can establish a linear recurrence relation for the $S_G(G_n)$, the first case is equivalent to the assumption that the characteristic polynomial of this recursion is a product of some z^{r-v} $(r - v \ge 0)$ and certain different cyclotomic polynomials. In this case, we can derive asymptotic formulas for $\mathbf{E}X_m$ and $\mathbf{V}X_m$, provided that the sequence G satisfies a certain technical condition. Our main result (Theorem 2) says that, in the case of unbounded variance, X_m satisfies a central limit theorem. (Note that there are sequences G for which $\mathbf{V}X_m$ is bounded, e.g., $G_n = 2^{n-1}$.)

2. THE MÖBIUS FUNCTION OF A DIGITAL EXPANSION

Let $G = (G_n)$ be a strictly increasing sequence of integers with $G_1 = 1$. As mentioned above, every nonnegative integer n has a digital expansion $n = \sum_{i \ge 1} \varepsilon_i G_i$ with nonnegative integral digits ε_i. It is called *proper digital expansion for n* if the digital sum $\sum_{i \ge 1} \varepsilon_i$ is as small as possible.

Lemma 1: Let $n = \sum_{i \ge 1} \varepsilon_i G_i$ be a proper digital expansion for n. Then any sum of the form $\sum_{i \ge 1} \varepsilon_i' G_i$ with integral digits ε_i', $i \ge 1$, satisfying $0 \le \varepsilon_i' \le \varepsilon_i$ is a proper digital representation for some $n' \le n$.

Proof: First, note that it follows from the algorithm stated in the Introduction that any digital expansion of the form $n_j = \sum_{i=1}^{j} \varepsilon_i G_i \le n$ is a proper one.

Next, we will use induction on the digital sum $s' = \sum_{i \ge 1} \varepsilon_i'$, where $0 \le \varepsilon_i' \le \varepsilon_i$. Obviously, there is nothing to show if $s' = 0$.

Now suppose that $n' = \sum_{i \ge 1} \varepsilon_i' G_i$ has digital sum s'. There exists $j \ge 1$ such that $\varepsilon_j' > 0$ and $\varepsilon_i' = 0$ for $i > j$. Then $G_j \le n' \le n_j < G_{j+1}$. Therefore, $n'' = n' - G_j$ can be represented by $n'' = \sum_{i=1}^{j} \varepsilon_i'' G_i$ with $\varepsilon_j'' = \varepsilon_j' - 1$ and $\varepsilon_i'' = \varepsilon_i'$ for $i \ne j$. Since $0 \le \varepsilon_i'' \le \varepsilon_i$ and its digital sum satisfies $\sum_{i \ge 1} \varepsilon_i'' = s' - 1 < s'$, this expansion for n'' is proper. Consequently, $\sum_{i \ge 1} \varepsilon_i' G_i$ is a proper expansion for n'. □

Now we introduce the Möbius functions $\mu(x, y)$ of a locally finite partial order $\le$ on a set X, i.e., all intervals $[x, y] = \{u \in X : x \le u \le y\}$ are finite (see [10], [13]). Any function $f : X^2 \to \mathbf{C}$ that satisfies $f(x, y) = 0$ for $x \not\le y$ will be called an *arithmetical function*. The convolution $f * g$ of two arithmetical functions f, g is given by

$$(f * g)(x, y) = \sum_{x \le u \le y} f(x, u) g(u, y).$$

Obviously δ, defined by $\delta(x,y)=1$ for $x=y$ and $\delta(x,y)=0$ otherwise, is the unit element of $*$. Furthermore, if $f(x,x)\neq 0$ for every $x\in X$, then there always exists an inverse arithmetical function f^{-1} satisfying $f^{-1}*f=\delta$. The Möbius function μ is defined as the inverse function of ζ given by $\zeta(x,y)=1$ if $x\le y$ and by $\zeta(x,y)=0$ otherwise. Especially, if $g=\zeta*f$, then f can be recovered by $f=\mu*g$. (We intend to use this Möbius function in future work for sieve methods in connection with specific problems of digital expansions.)

Theorem 1: Let $\le_G$ be the partial order on the nonnegative integers induced by the digital expansion with respect to a strictly increasing sequence of integers $G=(G_n)$ and suppose $m=\sum_{i\ge1}\varepsilon_i'G_i$ and $n=\sum_{i\ge1}\varepsilon_i''G_i$ are proper digital expansions of nonnegative integers m, n with $m\le_G n$, i.e., $\varepsilon_i'\le\varepsilon_i''$ for all i. Then

$$\mu(m,n)=\begin{cases}0 & \text{if there is an } i \text{ with } \varepsilon_i''-\varepsilon_i'>1,\\ (-1)^{\sum_{i\ge1}(\varepsilon_i''-\varepsilon_i')} & \text{otherwise.}\end{cases}$$

Proof: Since there is a natural bijection between $[m,n]=\{d\in\mathbb{N}_0 \mid m\le_G d\le_G n\}$ and $[0,n-m]$, we have $\mu(m,n)=\mu(0,n-m)$ if $m\le_G n$. (For $m\not\le_G n$, we have $\mu(m,n)=0$.)

Therefore, we will calculate only $\mu(0,n)$. From the definition of $\mu(x,y)$, it is clear that $\mu(0,0)=1$ and that

$$\sum_{0\le_G d\le_G n}\mu(0,d)=0 \quad \text{for } n>0.$$

Assume for a moment that $\varepsilon_i''\le 1$ for all i. We show that $\mu(0,\sum_{j=0}^{k-1}G_{i_j})=(-1)^k$ by induction on the digital sum $s=k$. Clearly, for $s=0$, we have $\mu(0,0)=1=(-1)^0$. Now assume that $s\ge1$ and that $\mu(0,\sum_{j=0}^{k-1}G_{i_j})=(-1)^k$ for all $k<s$. Then

$$\begin{aligned}0&=\sum_{0\le_G d\le_G\sum_{j=0}^{s-1}G_{i_j}}\mu(0,d)\\ &=(\mu(0,0))+(\mu(0,G_{i_0})+\mu(0,G_{i_1})+\cdots+\mu(0,G_{i_{s-1}}))\\ &\quad+(\mu(0,G_{i_0}+G_{i_1})+\mu(0,G_{i_0}+G_{i_2})+\cdots+\mu(0,G_{i_{s-2}}+G_{i_{s-1}}))+\cdots+\left(\mu\left(0,\sum_{j=0}^{s-1}G_{i_j}\right)\right)\\ &=1+\binom{s}{1}(-1)^1+\binom{s}{2}(-1)^2+\cdots+\binom{s}{s-1}(-1)^{s-1}+\mu\left(0,\sum_{j=0}^{s-1}G_{i_j}\right).\end{aligned}$$

Because of $\sum_{j=0}^{s}\binom{s}{j}(-1)^j=0$, it follows that $\mu(0,\sum_{j=0}^{s-1}G_{i_j})=(-1)^s$, which proves the theorem in this special case.

Now suppose that kG_i with $i\ge1$ and $k>1$ is a proper digital expansion. Then $0=\mu(0,0)+\mu(0,G_i)+\cdots+\mu(0,kG_i)$. Notice that $\mu(0,0)+\mu(0,G_i)=0$. Hence, it follows that $\mu(0,2G_i)=\mu(0,3G_i)=\cdots=\mu(0,kG_i)=0$.

Next, we show by induction on the digital sum $s(n)=\sum_{i\ge1}\varepsilon_i''$ that $\mu(0,n)=0$ whenever there is an i with $\varepsilon_i''>1$. We must start with $s(n)=2$ because $\varepsilon_i''>1$ cannot be satisfied when $s(n)<2$. Suppose that $s(n)=2$ and that there is some i with $\varepsilon_i''>1$. Then $m=2G_i$ and $\mu(0,m)=0$. Now assume the assertion holds for all natural numbers l with $s(l)<s(n)$ and assume there is a j with $\varepsilon_j''>1$. Then

$$-\mu(0,n) = \sum_{0 \le_G d <_G n} \mu(0,d) = \sum_{0 \le_G d <_G n,\, \forall i: \varepsilon_i(d) \le 1} \mu(0,d) + \sum_{0 \le_G d <_G n,\, \exists i: \varepsilon_i(d) > 1} \mu(0,d)$$
$$= \sum_{0 \le_G d <_G n,\, \forall i: \varepsilon_i(d) \le 1} \mu(0,d).$$

Define $n_1 := \sum_{i \ge 1} \min(\varepsilon_i'', 1) G_i$. Because of the existence of j with $\varepsilon_j'' > 1$, we have $0 < n_1 < n$ and

$$\sum_{0 \le_G d <_G n,\, \forall i: \varepsilon_i(d) \le 1} \mu(0,d) = \sum_{0 \le_G d \le_G n_1} \mu(0,d).$$

The right-hand side is, of course, zero, due to (2), which completes our proof. □

Since $\mu_G(m,n) = \mu_G(0, n-m)$ (if $m \le_G n$), it is sufficient to consider the restricted Möbius function $\mu_G(n) = \mu_G(0,n)$. As mentioned above, the main topic of this paper is to discuss the partial sums

$$M_G(N) = \sum_{n=0}^{N-1} \mu_G(n).$$

Nevertheless, we will rather discuss the partial sums $S_G(N)$, see (1), which will be motivated by the following proposition.

Proposition 1: Suppose that $G_n \ge 2G_{n-1}$ for all $n > 1$. Then $M_G(N)$ is bounded by 1. On the other hand, if $G_n \le 2G_{n-1}$ for all $n > 1$, then

$$M_G(N) = S_G(N) := \sum_{n=0}^{N-1} (-1)^{s_G(n)},$$

where $s_G(n)$ denotes the digital sum $s_G(n) = \sum_{i \ge 1} \varepsilon_i$ of the proper digital representation

$$n = \sum_{i \ge 1} \varepsilon_i G_i .$$

Proof: Due to Theorem 1, only those n with expansion coefficients 0 or 1 enter the sum. If $G_n \ge 2G_{n-1}$ for all $n > 1$, then all the digital expansions $\sum_{i \ge 1} \varepsilon_i G_i$ with $\varepsilon_i \in \{0, 1\}$ are proper ones. Hence, $M_G(N)$ attains only the same values as in the binary case in which the corresponding sum is 0 or ± 1.

If $G_n \le 2G_{n-1}$ for all $n > 1$, then in all the proper digital expansions only the digits 0 and 1 can occur, and the assertion follows from Theorem 1 with $m = 0$. □

Remark 1: We will see later that for all G considered here, $(a_1+1)G_{n-1} \ge G_n \ge a_1 G_{n-1}$ holds for $n > r$; therefore, $G_n \le 2G_{n-1}$ for all $n > 1$ is equivalent to $a_1 = 2$ and $r = 1$ or $a_1 = 1$ when the initial conditions of G are properly chosen. But if $a_1 > 2$ or $a_1 = 2$ and $r > 1$, and if $G_n \ge 2G_{n-1}$ holds for the initial values, then Proposition 1 applies and $M_G(N)$ is bounded. Because of this, we will investigate the function $S_G(N)$ rather than $M_G(N)$, keeping in mind that, in most cases, when $M_G(N)$ is of interest, both are the same.

Remark 2: If $G_n = 2^{n-1}$, then $t_n = (-1)^{s_G(n)}$ is the Thue-Morse sequence [11]. Since $t_{2n} + t_{2n+1} = 0$, we have $S_G(2n+1) = t_{2n} = t_n$, and we also have $S_G(2n) = 0$. Thus, it is not really interesting to study $S_G(N)$ in this case.

3. DIGITAL EXPANSIONS AND GENERATING FUNCTIONS

From this point on we will consider only integral linear recurring sequences $G = (G_n)_{n\geq 1}$ that satisfy assumptions 1-5 below (in Section 4.1 we will also need assumption 6):

1. $G_1 = 1$ and $G_{n+1} > G_n$ for $n \geq 1$.
2. $G_n = \sum_{i=1}^r a_i G_{n-i}$ for $n > r$ with some integers $a_i \geq 0$.
3. $G_{n-j} \geq \sum_{i=j+1}^r a_i G_{n-i}$ for $n > r$ and $1 \leq j < r$.
4. G satisfies no linear recursion with constant integer coefficients with a smaller degree.
5. The characteristic polynomial $z^r - \sum_{i=1}^r a_i z^{r-1} = \prod_{i=1}^r (z - \alpha_i)$ (of the above recursion) has only one real, positive, and simple root α_1 of maximal modulus.
6. Let $b_i = (a_i \bmod 2)(-1)^{a_1+\cdots+a_{i-1}}$ ($a_i \bmod 2 = 0$ if a_i is even and $a_i \bmod 2 = 1$ otherwise). Then
$$z^r - \sum_{i=1}^r b_i z^{r-i} = z^{r-v} \prod_{h=1}^{k'} \Phi_{k_h}(z) \tag{1}$$
is a product of z^{r-v} $(r - v \geq 0)$ and different cyclotomic polynomials $\Phi_{k_h}(z)$ $(k_1 < k_2 < \cdots < k_{k'})$, all of them dividing $z^p - 1$ with some fixed $p > r$. Furthermore, none of the α_i and no quotient α_i / α_j $(i \neq j)$ is a p^{th} root of unity.

Assumptions 1, 2, and 4 are natural. Therefore, only conditions 3, 5, and 6 need to be motivated.

Assumption 3 is necessary to show that $S(G_n)$ satisfies a linear recurrence; especially, it implies (6) in Proposition 2.

From assumption 5, we obtain $G_n = \beta\alpha_1^{n-1} + O((\alpha_1\gamma)^n)$ with some $\beta > 0$ and $0 \leq \gamma < 1$. Note that assumptions 2 and 3 imply $(a_1 + 1)G_{n-1} \geq G_n \geq a_1 G_{n-1}$ for $n > r$, which gives $a_1 \leq \alpha_1 \leq a_1 + 1$. Similarly, we get $a_1 \geq a_i$ for all i.

The first part of assumption 6 (concerning the cyclotomic factors) ensures that $S(G_n)$ is bounded. The assumption that α_i and α_i / α_j are not p^{th} roots of unity is frequently used in problems concerning digital expansions with respect to linear recurring sequences and avoids technical difficulties (see Lemma 2).

Usually, assumptions 3 and 5 are replaced by the stronger condition $a_1 \geq a_2 \geq \cdots \geq a_r$ and certain assumptions on the initial values of G (see, e.g., [8]; in this case, the second part of assumption 6 is also satisfied). However, there are other interesting examples, e.g., $a_1 = a_r = 1$, $a_2 = \cdots = a_{r-1} = 0$, that satisfy the above assumptions and are not of the form $a_1 \geq a_2 \geq \cdots \geq a_r$.

From here on, let $G = (G_n)$ be a fixed linear recurring sequence with assumptions 1-5. For notational convenience, we will omit the index G in the sequel.

Proposition 2: Let $b_i = (a_i \bmod 2)(-1)^{a_1+\cdots+a_{i-1}}$ ($a_i \bmod 2 = 0$ if a_i is even and $a_i \bmod 2 = 1$ otherwise). Then $S(G_n) = S_G(G_n)$ satisfies the linear recurrence
$$S(G_n) = \sum_{i=1}^r b_i S(G_{n-i}) \quad \text{for } n > r. \tag{2}$$
Furthermore, if n has the proper digital expansion $n = \sum_{j=1}^l \varepsilon_j G_j$, then
$$S\left(\sum_{j=1}^l \varepsilon_j G_j\right) = \sum_{j=1}^l (\varepsilon_j \bmod 2)(-1)^{\varepsilon_{j+1}+\cdots+\varepsilon_l} S(G_j). \tag{3}$$

Proof: We will first establish a set identity that holds for all nonnegative integers ε_j, regardless of whether $\sum_{j=1}^{l}\varepsilon_j G_j$ is a proper digital expansion or not:

$$\begin{aligned}\left\{a \,\middle|\, 0\le a<\sum_{j=1}^{l}\varepsilon_j G_j\right\} &= \bigcup_{j=1}^{l}\left\{a \,\middle|\, \sum_{h=j+1}^{l}\varepsilon_h G_h \le a < \sum_{h=j}^{l}\varepsilon_h G_h\right\}\\ &= \bigcup_{j=1}^{l}\left\{\sum_{h=j+1}^{l}\varepsilon_h G_h + a \,\middle|\, 0\le a<\varepsilon_j G_j\right\} = \bigcup_{j=1}^{l}\bigcup_{i=0}^{\varepsilon_j-1}\left\{\sum_{h=j+1}^{l}\varepsilon_h G_h + a \,\middle|\, iG_j\le a<(i+1)G_j\right\} \quad (4)\\ &= \bigcup_{j=1}^{l}\bigcup_{i=0}^{\varepsilon_j-1}\left\{\left(\sum_{h=j+1}^{l}\varepsilon_h G_h\right)+iG_j+a \,\middle|\, 0\le a<G_j\right\},\end{aligned}$$

where each union is disjoint. (Again, empty sums are set at zero.)

Now set $l=n-1$, $\varepsilon_j=a_{n-j}$ for $n-r\le j<n$ and $\varepsilon_j=0$ otherwise. Then one obtains for $n>r$, after interchanging i and j and shifting $i\to n-i$, $h\to n-h$,

$$\left\{a \,\middle|\, 0\le a<\sum_{i=n-r}^{n-1}a_{n-i}G_i\right\}=\bigcup_{i=1}^{r}\bigcup_{j=0}^{a_i-1}\left\{\left(\sum_{h=1}^{i-1}a_h G_{n-h}\right)+jG_{n-i}+a \,\middle|\, 0\le a<G_{n-i}\right\}. \quad (5)$$

From this we see that, for $n>r$,

$$\begin{aligned}S(G_n) &= \sum_{a=0}^{G_n-1}(-1)^{s(a)}=\sum_{i=1}^{r}\sum_{j=0}^{a_i-1}\sum_{a=0}^{G_{n-i}-1}(-1)^{s(\sum_{h=1}^{i-1}a_h G_{n-h}+jG_{n-i}+a)}\\ &= \sum_{i=1}^{r}\sum_{j=0}^{a_i-1}\sum_{a=0}^{G_{n-i}-1}(-1)^{(\sum_{h=1}^{i-1}a_h+j+s(a))}=\sum_{i=1}^{r}(-1)^{(\sum_{h=1}^{i-1}a_h)}\cdot S(G_{n-i})\sum_{j=0}^{a_i-1}(-1)^j\\ &= \sum_{i=1}^{r}(a_i \bmod 2)(-1)^{(\sum_{h=1}^{i-1}a_h)}S(G_{n-i})=\sum_{i=1}^{r}b_i S(G_{n-1})\end{aligned}$$

with $b_i:=(a_i \bmod 2)(-1)^{a_1+\cdots+a_{i-1}}$. Note that assumption 3 from above ensures that

$$s\left(\sum_{h=1}^{i-1}a_h G_{n-h}+jG_{n-i}+a\right)=\sum_{h=1}^{i-1}a_h+j+s(a). \quad (6)$$

You only have to start with $m=\sum_{h=1}^{i-1}a_h G_{n-h}+jG_{n-i}+a$ and apply the algorithms stated in the Introduction to deduce that $\varepsilon_{n-h}(m)=a_h$, $1\le h<i$ and $\varepsilon_{n-i}(m)=j$. (Of course, this procedure is standard in the study of such digital sequences (cf. [8], [9]). This proves equation (2).

The proof of (3) is quite similar. If we set $\sum_{j=1}^{l}\varepsilon_j G_j =: m+\varepsilon_l G_l$ in (4), we get

$$\{a \mid 0\le a<m+\varepsilon_l G_l\}=\bigcup_{i=0}^{\varepsilon_l-1}\{iG_l+a \mid 0\le a<G_l\}\cup\{\varepsilon_l G_l+a \mid 0\le a<m\}.$$

Let $\varepsilon_l G_l+m=\sum_{j=1}^{l}\varepsilon_j G_j$ be a proper digital expansion. Then it follows that

$$\begin{aligned}S(\varepsilon_l G_l+m) &= \sum_{a=0}^{\varepsilon_l G_l+m-1}(-1)^{s(a)}=\sum_{i=0}^{\varepsilon_l-1}\sum_{a=0}^{G_l-1}(-1)^{s(iG_l+a)}+\sum_{a=0}^{m-1}(-1)^{s(\varepsilon_l G_l+a)}\\ &= \sum_{i=0}^{\varepsilon_l-1}(-1)^i\sum_{a=0}^{G_l-1}(-1)^{s(a)}+(-1)^{\varepsilon_l}\sum_{a=0}^{m-1}(-1)^{s(a)}=(\varepsilon_l \bmod 2)S(G_l)+(-1)^{\varepsilon_l}S(m).\end{aligned} \quad (7)$$

Iterated use of equation (7) gives (3). □

Corollary: Let $d_m(k) := |\{0 \le a < m | S(a) = k\}|$ and $D_m(z)$ the corresponding generating function

$$D_m(z) = \sum_{k \in Z} d_m(k) z^k = \sum_{a=0}^{m-1} z^{S(a)}. \tag{8}$$

Then $D_{G_n}(z)$ (and $D_{G_n}(z^{-1})$) satisfy, for $n > r$, the relation

$$D_{G_n}(z) = \sum_{i=1}^{r} \sum_{j=0}^{a_i-1} z^{\left(\sum_{h=1}^{i-1} b_h S(G_{n-h}) + (-1)^{a_1+\cdots a_{i-1}} (j \bmod 2) S(G_{n-i})\right)} D_{G_{n-i}}\left(z^{(-1)^{a_1+\cdots a_{i-1}+j}}\right). \tag{9}$$

Proof: Suppose $n > r$. An iterated use of (7) gives, for $1 \le i \le r$, $j < a_i$, and $m < G_{n-i}$,

$$\begin{aligned} & S(a_1 G_{n-1} + \cdots + a_{i-1} G_{n-i+1} + j G_{n-i} + m) \\ & = (a_1 \bmod 2) S(G_{n-1}) + (-1)^{a_1} (a_2 \bmod 2) S(G_{n-2}) + \cdots \\ & \quad + (-1)^{a_1 + \cdots + a_{i-2}} (a_{i-1} \bmod 2) S(G_{n-i+1}) + (-1)^{a_1 + \cdots + a_{i-1}} (j \bmod 2) S(G_{n-i}) \\ & \quad + (-1)^{a_1 + \cdots + a_{i-1} + j} S(m) \\ & = \sum_{h=1}^{i-1} b_h S(G_{n-h}) + (-1)^{a_1 + \cdots + a_{i-1}} (j \bmod 2) S(G_{n-i}) + (-1)^{a_1 + \cdots + a_{i-1} + j} S(m). \end{aligned}$$

Note that, for $i = 1$, we just obtain $S(jG_{n-1} + m) = (j \bmod 2) S(G_{n-1}) + (-1)^j S(m)$. Hence, by using (5) and (8), we get

$$\begin{aligned} D_{G_n}(z) &= \sum_{m=0}^{G_n - 1} z^{S(m)} = \sum_{i=1}^{r} \sum_{j=0}^{a_i - 1} \sum_{m=0}^{G_{n-i}-1} z^{S(a_1 G_{n-1} + \cdots + a_{i-1} G_{n-i+1} + j G_{n-i} + m)} \\ &= \sum_{i=1}^{r} \sum_{j=0}^{a_i-1} z^{\left(\sum_{h=1}^{i-1} b_h S(G_{n-h}) + (-1)^{a_1 + \cdots + a_{i-1}} (j \bmod 2) S(G_{n-i})\right)} \sum_{m=0}^{G_{n-i}-1} z^{(-1)^{a_1 + \cdots + a_{i-1} + j} S(m)} \\ &= \sum_{i=1}^{r} \sum_{j=0}^{a_i-1} z^{\left(\sum_{h=1}^{i-1} b_h S(G_{n-h}) + (-1)^{a_1 + \cdots + a_{i-1}} (j \bmod 2) S(G_{n-i})\right)} D_{G_{n-i}}\left(z^{(-1)^{a_1 + \cdots + a_{i-1} + j}}\right). \quad \square \end{aligned}$$

4. ASYMPTOTIC ANALYSIS

We distinguish two cases: either $S(G_n)$ is bounded for all suitable initial conditions of G or it is not. The first case will be of special interest. It turns out that in this case the distribution of the values of $S(N)$ approximates a normal distribution for all suitable initial conditions of G (see Theorem 2).

4.1 Bounded $S(G_n)$

Proposition 3: Suppose that $S(G_n)$ is bounded. Then $S(G_n)$ satisfies a linear recursion for $n > N$ with some N, whose characteristic polynomial is a product of different cyclotomic polynomials.

Remark: This motivates the first part of assumption 6 in Section 3.

Proof: We know that every $S(m)$ is an integer and, therefore, can only attain a finite number of distinct values. So we see from (2) that $S(G_n)$ must be periodic (in n) for $n > N$. Let $p > r$ be some period of $S(G_n)$ and assume $n > N$. Then $S(G_{n+p}) - S(G_n) = 0$, which implies that $S(G_n)$ is a linear combination of powers of p^{th} roots of unity. Let $m(z)$ be the product of all cyclotomic polynomials corresponding to those roots of unity which appear in the representation of $S(G_n)$. Then $S(G_n)$ satisfies the linear recurrence related to $m(z)$. □

Proposition 4: Suppose that $S(G_n)$ is bounded. Then $D_{G_n}(z)$ (defined in (8)) and $D_{G_n}(z^{-1})$ satisfy, for $n > N$, a homogeneous linear recurrence with (in n) constant coefficients $a_i(z)$ that are analytic around $z = 1$ and satisfy $a_i(z) = a_i(z^{-1})$.

Proof: Let $p > r$ be a period of $S(G_n)$. Then, by splitting (9) into four parts, we get

$$\begin{aligned} D_{G_{k+sp}}(z) = & \sum_{i=\max(0,k-r)}^{k-1} \gamma_{k,i}(z) D_{G_{i+sp}}(z) + \sum_{i=k+p-r}^{p-1} \zeta_{k,i}(z) D_{G_{i+(s-1)p}}(z) \\ & + \sum_{i=\max(0,k-r)}^{k-1} \gamma_{k,p+i}(z) D_{G_{i+sp}}(z^{-1}) + \sum_{i=k+p-r}^{p-1} \zeta_{k,p+i}(z) D_{G_{i+(s-1)p}}(z^{-1}), \end{aligned} \tag{10}$$

with

$$\begin{aligned} \gamma_{k,i}(z) &= h_{k-i} z^{\left(\sum_{h=1}^{k-i-1} b_h m_{k-h} - (a_1 + \cdots + a_{k-i-1} \bmod 2) m_i\right)} \\ \gamma_{k,p+i}(z) &= \bar{h}_{k-i} z^{\left(\sum_{h=1}^{k-i-1} b_h m_{k-h} + (a_1 + \cdots + a_{k-i-1} \bmod 2) m_i\right)} \\ \zeta_{k,i}(z) &= h_{k+p-i} z^{\left(\sum_{h=1}^{k+p-i-1} b_h m_{k-h} - (a_1 + \cdots + a_{k+p-i-1} \bmod 2) m_i\right)} \\ \zeta_{k,p+i}(z) &= \bar{h}_{k+p-i} z^{\left(\sum_{h=1}^{k+p-i-1} b_h m_{k-h} - (a_1 + \cdots + a_{k+p-i-1} \bmod 2) m_i\right)}, \end{aligned}$$

where $m_i := S(G_i)$, $0 \le k < p$ and $0 \le i < p$ and

$$h_i = \begin{cases} |\{0 \le j < a_i \mid j \equiv a_1 + \cdots + a_{i-1} (2)\}| & \text{for } 1 \le i < r, \\ 0 & \text{otherwise,} \end{cases}$$

$$\bar{h}_i = \begin{cases} |\{0 \le j < a_i \mid j \equiv a_1 + \cdots + a_{i-1} + 1 (2)\}| & \text{for } 1 \le i \le r, \\ 0 & \text{otherwise.} \end{cases}$$

In the case $1 \le i \le r$, we can calculate

$$\begin{aligned} h_i &= \left\lfloor \frac{a_i + 1}{2} \right\rfloor - \begin{cases} 1 & \text{for } a_i \equiv a_1 + \cdots + a_{i-1} \ (2) \equiv 1 \ (2), \\ 0 & \text{otherwise,} \end{cases} \\ h_i + \bar{h}_i &= a_i, \\ h_i - \bar{h}_i &= b_i. \end{aligned} \tag{11}$$

Furthermore, we define $\gamma_{p+k,p+i}(z^{-1}) = \gamma_{k,i}(z)$, $\gamma_{p+k,i}(z^{-1}) = \gamma_{k,p+i}(z)$, $\zeta_{p+k,p+i}(z^{-1}) = \zeta_{k,i}(z)$, $\zeta_{p+k,i}(z^{-1}) = \zeta_{k,p+i}(z)$, and

$$\mathbf{d}_s(z) = \begin{pmatrix} \mathbf{d}_{1,s}(z) \\ \mathbf{d}_{2,s}(z) \end{pmatrix} = \begin{pmatrix} (D_{G_{0+sp}}(z), D_{G_{1+sp}}(z), \ldots, D_{G_{p-1+sp}}(z))^T \\ (D_{G_{0+sp}}(z^{-1}), D_{G_{1+sp}}(z^{-1}), \ldots, D_{G_{p-1+sp}}(z^{-1}))^T \end{pmatrix},$$

$$\mathbf{\Gamma}(z)=\begin{pmatrix}\mathbf{\Gamma}_{1,1}(z) & \mathbf{\Gamma}_{1,2}(z)\\ \mathbf{\Gamma}_{2,1}(z) & \mathbf{\Gamma}_{2,2}(z)\end{pmatrix}=\begin{pmatrix}(\gamma_{k,i}(z))_{0\le k,i<p} & (\gamma_{k,p+i}(z))_{0\le k,i<p}\\ (\gamma_{p+k,i}(z))_{0\le k,i<p} & (\gamma_{p+k,p+i}(z))_{0\le k,i<p}\end{pmatrix},$$

$$\mathbf{Z}(z)=\begin{pmatrix}\mathbf{Z}_{1,1}(z) & \mathbf{Z}_{1,2}(z)\\ \mathbf{Z}_{2,1}(z) & \mathbf{Z}_{2,2}(z)\end{pmatrix}=\begin{pmatrix}(\zeta_{k,i}(z))_{0\le k,i<p} & (\zeta_{k,p+i}(z))_{0\le k,i<p}\\ (\zeta_{p+k,i}(z))_{0\le k,i<p} & (\zeta_{p+k,p+i}(z))_{0\le k,i<p}\end{pmatrix}.$$

Then the identities $\mathbf{d}_{2,s}(z)=\mathbf{d}_{1,s}(z^{-1})$, $\mathbf{\Gamma}_{2,2}(z)=\mathbf{\Gamma}_{1,1}(z^{-1})$, $\mathbf{\Gamma}_{2,1}(z)=\mathbf{\Gamma}_{1,2}(z^{-1})$, $\mathbf{Z}_{2,2}(z)=\mathbf{Z}_{1,1}(z^{-1})$, and $\mathbf{Z}_{2,1}(z)=\mathbf{Z}_{1,2}(z^{-1})$ hold and (10) becomes

$$\mathbf{d}_s(z)=\mathbf{\Gamma}(z)\,\mathbf{d}_s(z)+\mathbf{Z}(z)\,\mathbf{d}_{s-1}(z),$$

or, formally,

$$\mathbf{d}_s(z)=((\mathbf{I}-\mathbf{\Gamma}(z))^{-1}\mathbf{Z}(z))\,\mathbf{d}_{s-1}(z).$$

Since the quadratic matrix $\mathbf{\Gamma}(1)$ consists of four quadratic $p\times p$-blocks that are lower triangle matrices with zero diagonal, it is an easy exercise to show that $\mathbf{I}-\mathbf{\Gamma}(1)$ is invertible. Hence, $(\mathbf{I}-\mathbf{\Gamma}(z))$ is invertible in a neighborhood of $z=1$.

Call $\mathbf{P}_{(z)}(l):=\det(l\mathbf{I}-\mathbf{\Theta}(z))$ the characteristic polynomial of the matrix

$$\mathbf{\Theta}(z):=(\mathbf{I}-\mathbf{\Gamma}(z))^{-1}\mathbf{Z}(z).$$

Then, by the theorem of Cayley-Hamilton, $\mathbf{P}_{(z)}(\mathbf{\Theta}(z))=\mathbf{0}$. From this, we see that the sequence $(D_{G_{i+sp}}(z))_{s\ge 0}$ satisfies a linear homogeneous recursion.

Finally, it follows from the definition of $\mathbf{\Gamma}$ and $\mathbf{Z}$ that $\mathbf{P}_{(z)}(l)=\mathbf{P}_{(z^{-1})}(l)$, from which we see that $a_i(z)=a_i(z^{-1})$. □

Let $A_i(z)$, $1\le i\le 2p$, denote the roots of the polynomial $\mathbf{P}_{(z)}(l)$, where z varies in a sufficiently small neighborhood of $z=1$. Since $a_i(z^{-1})=a_i(z)$, they satisfy $A_i(z^{-1})=A_i(z)$. Furthermore, there exist functions $B_{k,i}(z,s)$ that are polynomials in s such that

$$D_{G_{k+sp}}(z)=\sum_i B_{k,i}(z,s)A_i(z)^s. \tag{12}$$

Since $D_{G_{k+sp}}(1)=G_{k+sp}\sim\beta_1\alpha_1^{k-1}(\alpha_1^p)^s$, it might be expected that (locally around $z=1$) there exists a unique root $A_1(z)$ (satisfying $A_1(1)=\alpha_1^p$) of maximal modulus which is simple. The following lemma shows that this is true if assumption 6 in Section 3 holds.

Lemma 2: Suppose that assumptions 1-6 in Section 3 hold and let $v:=\max\{1\le i\le r\,|\,b_i\ne 0\}$. Then, with the above notation, the $2p$ roots of $\mathbf{P}_{(1)}(l)$ are α_i^p, $1\le i\le r$, where α_i, $1\le i\le r$, denote the roots of $z^r-\sum_{j=1}^r a_jz^{r-j}$, 0 with multiplicity $2p-r-v$, and 1 with multiplicity v.

Proof: From $D_{G_{k+sp}}(1)=G_{k+sp}=\sum_i\beta_i(k+sp)\alpha_i^{k+sp-1}\sim\beta_1\alpha_1^{k-1}(\alpha_1^p)^s$, we see that α_i^p surely are roots of $\mathbf{P}_{(1)}(l)$.

Since $\mathbf{I}-\mathbf{\Gamma}(1)$ is invertible, the multiplicity of 0 is $2p$ minus the rank of $\mathbf{Z}(1)$. $\mathbf{Z}(1)$ has a simple block structure. It is an easy exercise to show that its rank equals $r+v$. (Recall that $h_i+\overline{h}_i=a_i$ and $h_i-\overline{h}_i=b_i$.)

Similarly, the multiplicity of 1 is $2p$ minus the rank of

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_{1,1} & \mathbf{K}_{1,2} \\ \mathbf{K}_{1,2} & \mathbf{K}_{1,1} \end{pmatrix} = \mathbf{I} - \mathbf{\Gamma}(1) - \mathbf{Z}(1).$$

Observe that

$$\text{rk}\begin{pmatrix} \mathbf{K}_{1,1} & \mathbf{K}_{1,2} \\ \mathbf{K}_{1,2} & \mathbf{K}_{1,1} \end{pmatrix} = \text{rk}\begin{pmatrix} \mathbf{K}_{1,1} + \mathbf{K}_{1,2} & \mathbf{0} \\ \mathbf{K}_{1,2} & \mathbf{K}_{1,1} - \mathbf{K}_{1,2} \end{pmatrix}$$

and that $\mathbf{K}_{1,1} + \mathbf{K}_{1,2}$ (resp. $\mathbf{K}_{1,1} - \mathbf{K}_{1,2}$) are cyclic matrices with entries $1, -a_1, \ldots, -a_r, 0, \ldots, 0$ (resp. $1, -b_1, \ldots, -b_r, 0, \ldots, 0$). By [3, Lemma 3], the rank of $\mathbf{K}_{1,1} + \mathbf{K}_{1,2}$ is p (resp. the rank of $\mathbf{K}_{1,1} - \mathbf{K}_{1,2}$ is $p - \nu$), ν being equal to the number of different p^{th} roots of unity that are roots of $z^r - \sum_{j=1}^{r} b_j z^{r-j}$. Thus, rk $\mathbf{K} = 2p - \nu$, which completes the proof of the lemma. □

Let us define discrete random variables X_m by

$$\mathbf{P}(X_m = k) = \frac{d_m(k)}{m}. \tag{13}$$

(Recall that $d_m(k) := |\{0 \le a < m | S(a) = k\}|$.) It is well known that one can calculate mean and variance using the generating function:

$$\mu_m = \mathbf{E}X_m = \frac{1}{m} D_m'(1),$$

$$\sigma_m^2 = \mathbf{V}X_m = \frac{1}{m}\left(D_m''(1) + D_m'(1) - \frac{1}{m} D_m'(1)^2 \right).$$

From here on, we will assume 1-6 in Section 3.

Lemma 3: Let $A_1(z)$ be the unique root of maximal modulus of $\mathbf{P}_{(l)}(z)$. Then we have $A_1''(1) \ge 0$,

$$\mu_{G_{k+sp}} := \mathbf{E}X_{G_{k+sp}} = O(1) \quad \text{and} \quad \sigma^2_{G_{k+sp}} := \mathbf{V}X_{G_{k+sp}} = s\frac{A_1''(1)}{A_1(1)} + O(1)$$

as $s \to \infty$. Furthermore, if $A_1''(1) \ne 0$, then

$$\mathbf{E} \exp\left(it \frac{X_{G_{k+sp}} - \mu_{G_{k+sp}}}{\sigma_{G_{k+sp}}} \right) = \exp\left(-\frac{t^2}{2} \right)\left(1 + O\left(\frac{1}{\sqrt{s}} \right) \right)$$

as $s \to \infty$. This means that X_{G_m} is asymptotically Gaussian with mean μ_{G_m} and variance $\sigma^2_{G_m}$.

Proof: Let $A(z) = A_1(z)$ and $B_k(z) = B_{k,1}(z, s)$ in (12) (where the s-degree of the polynomial $B_{k,1}(z, s)$ is zero). Since $A'(1) = 0$, we obtain from (12) by differentiation,

$$D_{G_{k+sp}}(1) = B_k(1)A(1)^s + O((A(1)\gamma)^s),$$

$$D'_{G_{k+sp}}(1) = B_k(1)A(1)^s \frac{B_k'(1)}{B_k(1)} + O((A(1)\gamma)^s),$$

$$D''_{G_{k+sp}}(1) = B_k(1)A(1)^s \left(s\frac{A''(1)}{A(1)} + \frac{B_k''(1)}{B_k(1)} \right) + O((A(1)\gamma)^s),$$

with some $0 \le \gamma < 1$ properly chosen. From $D_{G_{k+sp}}(1) = G_{k+sp}$, we get

$$D'_{G_{k+sp}}(1) = G_{k+sp}\frac{B'_k(1)}{B_k(1)}(1+O(\gamma^s)),$$
$$D''_{G_{k+sp}}(1) = G_{k+sp}\left(s\frac{A''(1)}{A(1)}+\frac{B''_k(1)}{B_k(1)}\right)(1+O(\gamma^s)).$$

Both $D'_{G_{k+sp}}(1)$ and $D''_{G_{k+sp}}(1)$ are real, and because of $B_k(1)=\beta_1\alpha_1^{k-1}\in\mathbf{R}^+$, $B'_k(1)$ is real. Furthermore, $A''(1)$ and $B''_k(1)$ are real, too. From this, we obtain that

$$\mathbf{E}X_{G_{k+sp}} = \frac{B'_k(1)}{B_k(1)}(1+O(\gamma^s)) = O(1),$$
$$\mathbf{V}X_{G_{k+sp}} = \left(s\frac{A''(1)}{A(1)}+\frac{B''_k(1)}{B_k(1)}-\left(\frac{B'_k(1)}{B_k(1)}\right)^2\right)(1+O(\gamma^s)) = s\frac{A''(1)}{A(1)}+O(1),$$

from which it is clear that $A''(1)\geq 0$. Using $A'(1)=A'''(1)=0$, we get

$$A(e^t)^s = A(1)^s\exp\left(\frac{st^2}{2}\frac{A''(1)}{A(1)}\right)(1+O(st^4)).$$

Now suppose $A''(1)>0$, then we have

$$D_{G_{k+sp}}\left(e^{it/\sigma_{G_{k+sp}}}\right) = G_{k+sp}\exp\left(-\frac{t^2}{2}\right)\left(1+O\left(\frac{t}{\sqrt{s}}\right)+O\left(\frac{t^4+1}{s}\right)\right),$$

where the O-constants are independent of k. For any fixed t, we get

$$\mathbf{E}\exp\left(it\frac{X_{G_{k+sp}}-\mu_{G_{k+sp}}}{\sigma_{G_{k+sp}}}\right) = \frac{D_{G_{k+sp}}\left(e^{it/\sigma_{G_{k+sp}}}\right)}{G_{k+sp}}\exp\left(-it\frac{\mu_{G_{k+sp}}}{\sigma_{G_{k+sp}}}\right)$$
$$= \exp\left(-\frac{t^2}{2}\right)\left(1+O\left(\frac{1}{\sqrt{s}}\right)\right).$$

Thus, by Levi's theorem (see [7]), the normalized random variables $(X_{G_m}-\mu_{G_m})/\sigma_{G_m}$ converge weakly to normal distribution.

Remark: The use of generating functions for the proof of asymptotic normality probably started with Bender's paper [2]. Further references can be found in [5].

Now we will turn our attention to X_m, where m need not be an element of the basis G.

Theorem 2: Suppose that $G=(G_n)$ satisfies a linear recursion with restrictions 1-6 of Section 3. Then, with the above notation, we have

$$\mathbf{E}X_m = O(1) \quad\text{and}\quad \mathbf{V}X_m = \frac{l}{p}\frac{A''(1)}{A(1)}+O(1),$$

X_m being defined as in (13) and l being the length of the digital expansion of m. If $A''(1)>0$, then X_m is asymptotically Gaussian with mean value $\mathbf{E}X_m$ and variance $\mathbf{V}X_m\sim c\log m$ for some constant $c>0$, i.e.,

$$\lim_{m\to\infty}\frac{1}{m}\left|\left\{N<m: S(N)\leq \mathbf{E}X_m + x\sqrt{\mathbf{V}X_m}\right\}\right| = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\,dt.$$

Remark: The special case of $G_n = F_{n+1}$ (which leads to the original Zeckendorf representation) was discussed in [4]. There are also recent contributions to similar questions, e.g., Dumont and Thomas [6] prove asymptotic normality for substitution sequences by a different method, and Barat and Grabner [1] show the existence of a limiting distribution of G-additive functions.

Proof: Let $m = \sum_{i=1}^{l} \varepsilon_i G_i$ be the digital expansion of m. Iterated use of equation (7) yields, for $1 \le j \le l$, $i < \varepsilon_j$, and $a < G_j$,

$$S\left(\sum_{h=j+1}^{l} \varepsilon_h G_h + iG_j + a\right) = (\varepsilon_l \bmod 2)S(G_l) + (-1)^{\varepsilon_l} S\left(\sum_{h=j+1}^{l-1} \varepsilon_h G_h + iG_j + a\right)$$

$$= (\varepsilon_l \bmod 2)S(G_l) + (-1)^{\varepsilon_l}(\varepsilon_{l-1} \bmod 2)S(G_{l-1}) + \cdots + (-1)^{\varepsilon_l + \cdots + \varepsilon_{j+2}}(\varepsilon_{j+1} \bmod 2)S(G_{j+1})$$

$$+ (-1)^{\varepsilon_l + \cdots + \varepsilon_{j+1}}(i \bmod 2)S(G_j) + (-1)^{\varepsilon_l + \cdots + \varepsilon_{j+1} + i} S(a)$$

$$= \sum_{p=j+1}^{l} (-1)^{\varepsilon_l + \cdots + \varepsilon_{p+1}}(\varepsilon_p \bmod 2)S(G_p) + (-1)^{\varepsilon_l + \cdots + \varepsilon_{j+1}}(i \bmod 2)S(G_j) + (-1)^{\varepsilon_l + \cdots + \varepsilon_{j+1} + i} S(a),$$

and from (4) we see that

$$d_m(k) = \left|\left\{0 \le a < \sum_{i=1}^{l} \varepsilon_i G_k \,\middle|\, S(a) = k\right\}\right| = \sum_{j=1}^{l}\sum_{i=0}^{\varepsilon_j - 1}\left|\left\{0 \le a < G_j \,\middle|\, S\left(\sum_{h=j+1}^{l} \varepsilon_h G_h + iG_j + a\right) = k\right\}\right|$$

$$= \sum_{j=1}^{l}\sum_{i=0}^{\varepsilon_j - 1}\left|\left\{0 \le a < G_j \,\middle|\, S(a) = (-1)^{\varepsilon_l + \cdots + \varepsilon_{j+1} + i} \times \left(k - \sum_{\substack{p=j+1 \\ \varepsilon_p \equiv 1(2)}}^{l} (-1)^{\varepsilon_l + \cdots + \varepsilon_{p+1}} S(G_p) - (-1)^{\varepsilon_l + \cdots + \varepsilon_{j+1}}(i \bmod 2)S(G_j)\right)\right\}\right|$$

$$= \sum_{j=1}^{l}\sum_{i=0}^{\varepsilon_j - 1} d_{G_j}\left((-1)^{\varepsilon_l + \cdots + \varepsilon_{j+1} + i}\left(k - \sum_{\substack{p=j+1 \\ \varepsilon_p \equiv 1(2)}}^{l} (-1)^{\varepsilon_l + \cdots + \varepsilon_{p+1}} m_p - (-1)^{\varepsilon_l + \cdots + \varepsilon_{j+1}}(i \bmod 2)m_j\right)\right)$$

and

$$D_m(z) = \sum_{k \in \mathbf{Z}} d_m(k) z^k$$

$$= \sum_{j=1}^{l}\sum_{i=0}^{\varepsilon_j - 1}\sum_{k \in \mathbf{Z}} z^k d_{G_j}\left((-1)^{\varepsilon_l + \cdots + \varepsilon_{j+1} + i}\left(k - \sum_{\substack{p=j+1 \\ \varepsilon_p \equiv 1(2)}}^{l} (-1)^{\varepsilon_l + \cdots + \varepsilon_{p+1}} S(G_p) - (-1)^{\varepsilon_l + \cdots + \varepsilon_{j+1}}(i \bmod 2)S(G_j)\right)\right)$$

$$= \sum_{j=1}^{l}\sum_{i=0}^{\varepsilon_j - 1}\sum_{k \in \mathbf{Z}} d_{G_j}(k) z^{\left((-1)^{\varepsilon_l + \cdots + \varepsilon_{j+1} + i} k + \sum_{\substack{p=j+1 \\ \varepsilon_p \equiv 1(2)}}^{l} (-1)^{\varepsilon_l + \cdots + \varepsilon_{p+1}} m_p + (-1)^{\varepsilon_l + \cdots \varepsilon_{j+1}}(i \bmod 2)m_j\right)}$$

$$= \sum_{j=1}^{l} z^{\left(\sum_{\substack{p=j+1 \\ \varepsilon_p \equiv 1(2)}}^{l} (-1)^{\varepsilon_l+\cdots+\varepsilon_{p+1}} m_p\right)} \sum_{i=0}^{\varepsilon_j - 1} z^{\left((-1)^{\varepsilon_l+\cdots+\varepsilon_{j+1}} (i \bmod 2) m_j\right)} D_{G_j}\left(z^{\left((-1)^{\varepsilon_l+\cdots+\varepsilon_{j+1}+i}\right)}\right)$$

$$= \sum_{j=1}^{l} \sum_{i=0}^{\varepsilon_j - 1} z^{b(j,i)} D_{G_j}(z^{c(j,i)}), \tag{14}$$

in which

$$b(j,i) = \sum_{\substack{p=j+1 \\ \varepsilon_p \equiv 1(2)}}^{l} (-1)^{\varepsilon_l+\cdots+\varepsilon_{p+1}} m_p + (-1)^{\varepsilon_l+\cdots+\varepsilon_{j+1}} (i \bmod 2) m_j,$$

$$c(j,i) = (-1)^{\varepsilon_l+\cdots+\varepsilon_{j+1}+i}.$$

Differentiation of (14) yields

$$zD'_m(z) = \sum_{j=1}^{l} \sum_{i=0}^{\varepsilon_j - 1} \left(b(j,i) z^{b(j,i)} D_{G_j}(z^{c(j,i)}) + z^{b(j,i)} D'_{G_j}(z^{c(j,i)}) c(j,i) z^{c(j,i)} \right),$$

$$z\frac{\partial}{\partial z}(zD'_m(z)) = \sum_{j=1}^{l} \sum_{i=0}^{\varepsilon_j - 1} \Big(b(j,i)^2 z^{b(j,i)} D_{G_j}(z^{c(j,i)}) + 2b(j,i) z^{b(j,i)} D'_{G_j}(z^{c(j,i)}) c(j,i) z^{c(j,i)}$$
$$+ z^{b(j,i)} \left(z^{c(j,i)} D'_{G_j}(z^{c(j,i)}) + z^{2c(j,i)} D''_{G_j}(z^{c(j,i)}) \right) \Big).$$

It is an easy exercise to show $\sum_{j=1}^{l} (l-j+1)^k G_j \le C_k G_l$. Because the m_j are bounded, we get $b(j,i) = O(l-j+1)$ (uniformly in i) and

$$D'_m(1) = \sum_{j=1}^{l} \sum_{i=0}^{\varepsilon_j - 1} \left(b(j,i) D_{G_j}(1) + c(j,i) D'_{G_j}(1) \right)$$
$$= O\left(\sum_{j=1}^{l} (l-j+1) G_j \right) = O(G_l) = O(m)$$

and

$$\left.\frac{\partial}{\partial z}(zD'_m(z))\right|_{z=1} = \sum_{j=1}^{l} \sum_{i=0}^{\varepsilon_j - 1} \left(b(j,i)^2 D_{G_j}(1) + 2b(j,i) c(j,i) D'_{G_j}(1) + D'_{G_j}(1) + D''_{G_j}(1) \right)$$
$$= \sum_{j=1}^{l} \varepsilon_j D''_{G_j}(1) + O\left(\sum_{j=1}^{l} (l-j+1)^2 G_j \right) + O\left(\sum_{j=1}^{l} (l-j+1) G_j \right) + O\left(\sum_{j=1}^{l} G_j \right)$$
$$= \sum_{j=1}^{l} \varepsilon_j G_j \frac{j}{p} \frac{A''(1)}{A(1)} \left(1 + O\left(\frac{1}{j} \right) \right) + O(m)$$
$$= \frac{1}{p} \frac{A''(1)}{A(1)} \left(l \sum_{j=1}^{l} \varepsilon_j G_j - \sum_{j=1}^{l} \varepsilon_j G_j (l-j) \right) + O\left(\sum_{j=1}^{l} \varepsilon_j G_j \frac{1}{p} \frac{A''(1)}{A(1)} \right) + O(m)$$
$$= m \frac{l}{p} \frac{A''(1)}{A(1)} + O(m).$$

Thus, we have

$$\mathbf{E}X_m = O(1) \quad \text{and} \quad \mathbf{V}X_m = \frac{l}{p}\frac{A''(1)}{A(1)} + O(1). \tag{15}$$

Furthermore, by using (14), we obtain

$$D_m(e^{it/\sigma_m}) = \sum_{j=1}^{l} \exp\left(\frac{it}{\sigma_m} \sum_{\substack{p=j+1 \\ \varepsilon_p \equiv 1(2)}}^{l} (-1)^{\varepsilon_l+\cdots+\varepsilon_{p+1}} m_p \right)$$

$$\times \sum_{i=0}^{\varepsilon_j - 1} \exp\left(\frac{it}{\sigma_m} (-1)^{\varepsilon_l+\cdots+\varepsilon_{j+1}} (i \bmod 2) m_j \right) D_{G_j}\left(\exp\left(\frac{it}{\sigma_m} (-1)^{\varepsilon_l+\cdots+\varepsilon_{j+1}+i} \right)\right),$$

and for any fixed t,

$$D_{G_j}\left(\exp\left(\frac{it}{\sigma_m} (-1)^{\varepsilon_l+\cdots+\varepsilon_{j+1}+i} \right)\right) = D_{G_j}\left(\exp\left(\frac{it\frac{\sigma_{G_j}}{\sigma_m}}{\sigma_{G_j}} (-1)^{\varepsilon_l+\cdots+\varepsilon_{j+1}+i} \right)\right)$$

$$= G_j \exp\left(-\frac{t^2 \frac{j}{l}\left(1+O\left(\frac{1}{j}\right)\right)}{2} \right)\left(1+O\left(\frac{1}{\sqrt{j}}\right)\right) = G_j e^{-t^2/2} \exp\left(\frac{t^2}{2}\frac{l-j}{l} + O\left(\frac{1}{\sqrt{j}}\right)\right)$$

and

$$\sum_{i=0}^{\varepsilon_j-1} \exp\left(\frac{it}{\sigma_m} (-1)^{\varepsilon_l+\cdots+\varepsilon_{j+1}} (i \bmod 2) m_j \right) = \sum_{i=0}^{\varepsilon_j-1} \left(1 + O\left(\frac{1}{\sqrt{l}.} \right)\right)$$

$$= \varepsilon_j \left(1 + O\left(\frac{1}{\sqrt{l}} \right)\right) = \varepsilon_j \exp\left(O\left(\frac{1}{\sqrt{l}} \right)\right),$$

where the O-constants do not depend on l or j. Thus we get, for $0 < \vartheta < \frac{1}{2}$,

$$D_m(e^{it/\sigma_m}) e^{t^2/2} = \sum_{j=1}^{l} \varepsilon_j G_j \exp\left(\frac{it}{\sigma_m} \sum_{\substack{p=j+1 \\ \varepsilon_p \equiv 1(2)}}^{l} (-1)^{\varepsilon_l+\cdots+\varepsilon_{p+1}} m_p + O\left(\frac{l-j}{l}\right) + O\left(\frac{1}{\sqrt{j}}\right)\right)$$

$$= \sum_{l-l^\vartheta \le j \le l} \varepsilon_j G_j \exp\left(O\left(\frac{l-j}{\sqrt{l}}\right) + O\left(\frac{1}{\sqrt{j}}\right)\right) + \sum_{1 \le j < l-l^\vartheta} \varepsilon_j G_j O(1)$$

$$= \sum_{l-l^\vartheta \le j \le l} \varepsilon_j G_j \exp\left(O\left(l^{\vartheta-\frac{1}{2}}\right) + O\left(\frac{1}{\sqrt{l/2}}\right)\right) + O\left(G_{\lfloor l-l^\vartheta \rfloor}\right)$$

$$= \sum_{l-l^\vartheta \le j \le l} \varepsilon_j G_j \left(1 + O\left(l^{\vartheta-\frac{1}{2}}\right)\right) + O\left(\alpha_1^{l-l^\vartheta}\right) = \sum_{l-l^\vartheta \le j \le l} \varepsilon_j G_j + O\left(ml^{\vartheta-\frac{1}{2}}\right) + O\left(\alpha_l^{l-l^\vartheta}\right)$$

$$= m + O\left(ml^{\vartheta-\frac{1}{2}}\right) + O\left(\alpha_1^{l-l^\vartheta}\right) = m + O\left(ml^{\vartheta-\frac{1}{2}}\right)$$

and, finally (for any fixed t),

$$\mathbf{E}\exp\left(it\frac{X_m-\mu_m}{\sigma_m}\right)=\frac{D_m(e^{it/\sigma_m})}{m}\exp\left(-it\frac{\mu_m}{\sigma_m}\right)$$
$$=\exp\left(-\frac{t^2}{2}\right)\left(1+O(l^{9-\frac{1}{2}})\right)\exp\left(O\left(\frac{1}{\sqrt{l}}\right)\right)$$
$$=\exp\left(-\frac{t^2}{2}\right)\left(1+O(l^{9-\frac{1}{2}})\right)$$

and X_m is asymptotically Gaussian with mean μ_m and variance σ_m^2. □

The condition that $z^r-\sum_{i=1}^r b_i z^{r-i}$ (where $v=\max\{1\le i\le r\,|\,b_i\ne 0\}$) is a product of z^{r-v} and different cyclotomic polynomials is rather restrictive in the case in which $G_n\le 2G_{n-1}$ for $n>1$.

Proposition 5: Suppose that $G=(G_n)$ satisfies a linear recursion with restrictions 1-5 of Section 3 such that $G_n\le 2G_{n-1}$ for $n>1$. Then $z^r-\sum_{i=1}^r b_i z^{r-i}$ is a product of z^{r-v} and different cyclotomic polynomials, where $v=\max\{1\le i\le r\,|\,b_i\ne 0\}$, if and only if one of the following conditions holds:

1. $r=1$ and $a_1=2$: the binary system, or
2. $a_1=a_2=\cdots+a_r=1$: a generalized Zeckendorf representation.

Proof: First, let $B(z)=z^r-\sum_{i=1}^r b_i z^{r-i}$ be of the above type, then if $a_1>1$ we are in the first case. So let us assume $a_1=1$, then it follows that $a_i\in\{0,1\}$, $a_r=1$, and therefore $v=r$. From this, we see that $z^r-\sum_{i=1}^r b_i z^{r-i}$ must be a symmetric polynomial that yields $a_i=a_{r-i}$ for all $1\le i<r$. Now suppose $a_1=\cdots=a_{i-1}=1=a_r=\cdots=a_{r-i+1}$ and $a_i=0=a_{r-i}$ for some $1<i\le r-i$. Then, by assumption 3 in Section 3, we have that $G_{n-r+i}\ge\sum_{j=r-i+1}^r a_j G_{n-j}=\sum_{j=r-i+1}^r G_{n-j}$ for $n>r$ or, equivalently, that $G_n\ge\sum_{j=1}^i G_{n-j}$ for $n>i$. Because $G_n=\sum_{j=1}^r a_j G_{n-j}$ for $n>r$, it follows that $\sum_{j=i+1}^r a_j G_{n-j}\ge G_{n-i}$ for $n>r$. On the other hand we have, again by assumption 3, that $G_{n-i}\ge\sum_{j=i+1}^r a_j G_{n-j}$ for $n>r$, from which we see that $G_n=\sum_{j=1}^{r-i} a_{i+j}G_{n-j}$ for $n>r-i$, a contradiction to assumption 4.

Now let $r=1$ and $a_1=2$, then $v=0$ and $B(z)=z$. Finally, suppose $a_1=a_2=\cdots=a_r=1$. Then $b_i=(-1)^{i+1}$ and

$$B(z)=\sum_{i=0}^r(-1)^i z^{r-i}=\frac{z^{r+1}+(-1)^r}{z+1}$$

is of the desired type. □

4.2 Unbounded $S(G_n)$

Proposition 6: If $S(G_n)$ is unbounded, then there exists some α with $1<\alpha<\alpha_1$ (α_1 defined as in Section 3), $k\ge 1$, real numbers $\varphi_1,\dots,\varphi_k$, and polynomials $\beta_1(n),\dots,\beta_k(n)$, $\overline{\beta}_1(n),\dots,\overline{\beta}_k(n)$ not all of them zero, such that

$$S(G_n)=\alpha^n\sum_{i=1}^k(\beta_i(n)\cos(n\varphi_i)+\overline{\beta}_i(n)\sin(n\varphi_i))+O((\gamma\alpha)^n)$$

for some $\gamma\in(0,1)$.

Proof: Since $S(G_n)$ satisfies the linear recurrence of Proposition 2, this representation follows immediately. □

Theorem 3: Suppose that $G=(G_n)$ satisfies a linear recurrence as above such that $S(G_n)$ is unbounded. Then

$$\limsup_{m\to\infty}\frac{\log(|S(m)|)}{\log m}=\frac{\log\alpha}{\log\alpha_1}.$$

Proof: First, it follows from Proposition 6 that

$$\limsup_{m\to\infty}\frac{\log(|S(m)|)}{\log m}\geq\limsup_{m\to\infty}\frac{\log(|S(G_n)|)}{\log G_n}=\frac{\log\alpha}{\log\alpha_1}.$$

The upper bound follows from the second part of Proposition 2 and again by an application of Proposition 6: Let $m=\sum_{j=1}^{l}\varepsilon_j G_j$ be the proper digital expansion of m and let $C, K>0$ be large enough so that $|\beta_i(n)+\overline{\beta}_i(n)|<Cn^D$ for all n, i. Then we have, for $l\to\infty$,

$$\frac{\log(|S(m)|)}{\log m}\leq\frac{\log\left(\sum_{j=1}^{l}|S(G_j)|\right)}{\log(\varepsilon_l G_l)}\leq\frac{\log(l\alpha^l(Cl^D+C'\gamma^l))}{\log\varepsilon_l+\log G_l}$$
$$\leq\frac{l\log\alpha+(D+1)\log l+C''}{l\log\alpha_1+C'''}\to\frac{\log\alpha}{\log\alpha_1},$$

which completes our proof. □

Remark: It is also possible to discuss the function $F(m)=S(m)m^{-(\log\alpha)/(\log\alpha_1)}$ in more detail. It turns out that $F(m)$ is an almost periodic function, i.e., $S(m)$ has an almost fractal structure. You just have to adapt the methods used in [8] and [9].

5. CONCLUSIONS

Our starting point was the Möbius function $\mu_G(n)$ of the partial order which is induced by proper digital expansions with respect to a basis $G=(G_n)$. It turned out that $\mu_G(n)\in\{-1,0,1\}$, so it is a natural question to determine the distribution of these three values $-1, 0, 1$. If $G_{n+1}\geq 2G_n$ for all $n>1$, then the answer is very easy (see Proposition 1). Therefore, we restricted ourselves to the case $G_{n+1}\leq 2G_n$ for all $n>1$. Here $\mu_G(n)=(-1)^{s_G(n)}$. Thus, $\mu_G(n)\neq 0$ for all $n\geq 0$ and $M_G(N)=S_G(N)$ is exactly the difference between the number of $n<N$ with $\mu_G(n)=1$ and the number of $n<N$ with $\mu_G(n)=-1$. In the case of linear recurring sequences $G=(G_n)$ (satisfying certain natural conditions), we proved that in any case $M_G(N)=o(N)$, i.e., $-1, +1$ are asymptotically equidistributed.

More precisely, we discussed the distribution of values of $S_G(N)$ (which can also be considered in the case $G_{n+1}\geq 2G_n$). It turns out that there are two essentially different cases, the case of bounded $S_G(G_n)$ and the case of unbounded $S_G(G_n)$. If $S_G(G_n)$ is unbounded, then $S_G(N)$ has an almost fractal structure (see Theorem 3 and the Remark following it). However, if $S_G(G_n)$ is bounded for all suitable initial conditions of G, then the values $S_G(N)$ admit a Gaussian limit law in the following sense: If X_n is a random variable defined by

$$\mathbf{P}(X_N=k)=\frac{1}{N}|\{n<N\,|\,S_G(n)=k\}|$$

then X_N is asymptotically Gaussian with bounded mean value and variance $\mathbf{V}X_N\sim c\log N$, provided that $c\neq 0$ (Theorem 2).

Since $S_G(G_n)$ satisfies the linear recurrence (2), it follows that $S_G(G_n)$ is periodic (for sufficiently large n) if it is bounded. This can only occur for all suitable initial conditions of G if and only if the roots of the characteristic polynomial $B(z) = z^r - \sum_{j=1}^r b_j z^{r-j}$ of (2) are 0 or roots of unity. Therefore, the assumption on $B(z)$ in Theorem 2, this is assumption 6 in Section 3, is quite natural.

Finally, we want to recall that the only recurring sequences $G = G(n)$ satisfying assumptions 1-5 such that $a_1 = 1$ (i.e., $G_{n+1} < 2G_n$) and that $B(z)$ is the product of $z^{r-\nu}$ and cyclotomic polynomials are generalized Fibonacci numbers (Proposition 5). They satisfy a recursion of the form $G_n = G_{n-1} + \cdots + G_{n-r}$. Here Theorem 2 applies. Hence, the values of $M_G(N)$ with respect to generalized Zeckendorf representations satisfy a central limit law.

REFERENCES

1. G. Barat & P. J. Grabner. "Distribution Properties of G-Additive Functions." *J. Number Th.* **60** (1996):103-23.
2. E. A. Bender. "Central and Local Limit Theorems Applied to Asymptotic Enumeration." *J. Combin. Theory Ser. A* **15** (1973):91-111.
3. M. Drmota & M. Skalba. "Relations between Polynomial Roots." *Acta Arith.* **71** (1995):65-77.
4. M. Drmota & M. Skalba. "The Parity of the Zeckendorf Sum-of-Digits-Function." Manuscript, 1995.
5. M. Drmota & M. Soria. "Marking in Combinatorial Constructions: Generating Functions and Limiting Distributions." *Theor. Comput. Sci.* **144** (1995):67-99.
6. J. M. Dumont & A. Thomas. "Gaussian Asymptotic Properties of the Sum-of-Digits Function." *J. Number Th.* **62** (1997):19-38.
7. W. Feller. *An Introduction to Probability Theory and Its Applications.* Vols. I and II. New York: Wiley, 1966.
8. P. J. Grabner & R. F. Tichy. "Contributions to Digit Expansions with Respect to Linear Recurrences." *J. Number Th.* **36** (1990):160-69.
9. P. J. Grabner & R. F. Tichy. "α-Expansions, Linear Recurrences, and the Sum-of-Digits Function." *Manuscripta Math.* **70** (1991):311-24.
10. J. H. Van Lint & R. M. Wilson. *A Course in Combinatorics.* Cambridge: Cambridge University Press, 1992.
11. H. M. Morse. "Recurrent Geodesics on a Surface of Negative Curvature." *Trans. Amer. Math. Soc.* **22** (1921):84-100.
12. A. Pethö & R. F. Tichy. "On Digit Expansions with Respect to Linear Recurrences." *J. Number Th.* **33** (1989):243-56.
13. R. P. Stanley. *Enumerative Combinatorics.* Vol. I. Monterey, CA: Wadsworth & Brooks/-Cole Advanced Books & Software, 1986.

AMS Classification Numbers: 11A63, 11B39

❖❖❖