THE JOINT DISTRIBUTION OF GREEDY AND LAZY FIBONACCI EXPANSIONS

Wolfgang Steiner
Vienna University of Technology, Department of Geometry,
Wiedner Hauptstraße 8–10/113, A–1040 Vienna, Austria

July 29, 2002

1. Introduction

Every non-negative integer n has at least one digital expansion

$$n = \sum_{k \geq 2} \epsilon_k F_k,$$

with digits $\epsilon_k \in \{0,1\}$. The maximal expansion with respect to the lexicographic order on $(\ldots, \epsilon_4, \epsilon_3, \epsilon_2)$ is the *Zeckendorf expansion* or, more generally, the *greedy expansion*, which has been studied by Zeckendorf [7] and many others. (Lexicographic order means $(\ldots, \epsilon_3, \epsilon_2) < (\ldots, \epsilon'_3, \epsilon'_2)$ if $\epsilon_k < \epsilon'_k$ for some $k \geq 2$ and $\epsilon_j \leq \epsilon'_j$ for all $j \geq k$.) The minimal expansion with respect to this order is the less known *lazy expansion*, which was introduced by Erdős and Joó [4] (for q-ary expansions of 1, $1 < q < 2$). For example, 100 has greedy expansion $100 = 89+8+3 = F_{11}+F_6+F_4$ and lazy expansion $100 = 55+21+13+5+3+2+1 = F_{10}+F_8+F_7+F_5+F_4+F_3+F_2$. Denote the digits of the greedy expansion by $\epsilon^g_k(n)$ and those of the lazy expansion by $\epsilon^\ell_k(n)$.

The aim of this work is to study the structure of the possible digit sequences in order to obtain distributional results for the *sum-of-digits functions*

$$s_g(n) = \sum_{k \geq 2} \epsilon^g_k(n) \quad \text{and} \quad s_\ell(n) = \sum_{k \geq 2} \epsilon^\ell_k(n).$$

This work was supported by the Austrian Science Foundation, grant S8302-MAT.

2. Results

It is well known that Zeckendorf expansions have no two subsequent ones (because the pattern $(0,1,1)$ could be replaced by $(1,0,0)$) and that every finite sequence with no two subsequent ones is a Zeckendorf expansion of some integer (see Zeckendorf [7]). Symmetrically, lazy expansions have no two subsequent zeros preceeded by a one, because $(1,0,0)$ could be replaced by $(0,1,1)$, and it is not difficult to see that every such sequence is the lazy expansion of some integer (see Lemma 1).

For $s_g(n)$, Grabner and Tichy [5] proved (in the context of digital expansion related to linear recurrences) that its mean value is given by

$$\frac{1}{N}\sum_{n<N} s_g(n) = \frac{1}{\alpha^2+1}\log_\alpha N + f_1(\log_\alpha N) + \mathcal{O}\left(\frac{\log N}{N}\right),$$

where f_1 is periodic with period 1, continuous and nowhere differentiable and α denotes the golden number $\frac{1+\sqrt{5}}{2}$. For the variance, Dumont and Thomas [2] obtained (in the more general context of numeration systems associated with primitive substitutions on finite alphabets)

$$\frac{1}{N}\sum_{n<N}\left(s_g(n) - \frac{1}{\alpha^2+1}\log_\alpha N\right)^2 = \frac{1}{5\sqrt{5}}\log_\alpha N + f_2(\log_\alpha N)\log_\alpha N + o(1),$$

where f_2 is again periodic with period 1, continuous and nowhere differentiable. In [3], they showed that the distribution is asymptotically normal, i.e.

$$\frac{1}{N}\#\left\{n<N \,\middle|\, \frac{s_g(n) - \frac{1}{\alpha^2+1}\log_\alpha N}{5^{-3/4}\sqrt{\log_\alpha N}} < x\right\} \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt.$$

This is also a special case of a result of Drmota and Steiner [1], where generalizations of the sum-of-digits functions are studied.

The distribution of $s_\ell(n)$ has not been studied yet, but it is easy to replace the greedy expansions in [1] by lazy expansions and to obtain similar asymptotics (with expected value $\frac{\alpha^2}{\alpha^2+1}\log_\alpha N$). Instead of doing this, we will directly prove the following central limit theorem for the joint distribution of $s_g(n)$ and $s_\ell(n)$.

Theorem 1. *We have, as $N\to\infty$,*

$$\frac{1}{N}\#\left\{n<N \,\middle|\, \frac{s_g(n)-\mu_g\log_\alpha N}{\sigma\sqrt{\log_\alpha N}} < x_g, \frac{s_\ell(n)-\mu_\ell\log_\alpha N}{\sigma\sqrt{\log_\alpha N}} < x_\ell\right\}$$
$$\to \frac{1}{2\pi\sqrt{1-C^2}}\int_{-\infty}^{x_\ell}\int_{-\infty}^{x_g} e^{-\frac{1}{2(1-C^2)}(t_g^2+t_\ell^2-2Ct_gt_\ell)}dt_g dt_\ell$$

with $\alpha = \frac{1+\sqrt{5}}{2}$, $\mu_g = \frac{1}{\alpha^2+1}$, $\mu_\ell = \frac{\alpha^2}{\alpha^2+1}$, $\sigma = 5^{-3/4}$ *and* $C = 9 - 5\alpha \approx 0.90983$.

This means that the two sum-of-digits functions are strongly correlated. If one of them is large for some n, the probability of the other one to be large is very high. (The distribution is the Gaussian distribution with covariance matrix $\left(\begin{smallmatrix} 1 & C \\ C & 1 \end{smallmatrix}\right)$.)

Similarly to [1], corresponding results can be proved for F-additive functions, for sequences of primes and for polynomial sequences $P(n)$, $n \in \mathbb{N}$, or $P(p)$, $p \in \mathbb{P}$.

3. Proofs

First we prove the characterization of lazy expansions given in Section 2.

Lemma 1. *The lazy expansions are exactly those sequences* $(\epsilon_k)_{k\geq 2} \in \{0,1\}^{\mathbb{N}}$ *with* $(\epsilon_k, \epsilon_{k-1}, \epsilon_{k-2}) \neq (1,0,0)$ *for all* $k \geq 4$ *and only a finite number of* $\epsilon_k = 1$.

Proof. As already noted, the pattern $(1,0,0)$ does not occur because it could be replaced by $(0,1,1)$ and it suffices therefore to show that no two such sequences represent the same number. For an integer $n \in \{F_k - 1, F_k, \ldots, F_{k+1} - 2\}$, we must have $\epsilon_j^\ell(n) = 0$ for all $j \geq k$ since $\epsilon_j^\ell(n) = 1$ implies

$$\sum_{i=2}^{j} \epsilon_i^\ell(n) F_i \geq F_j + F_{j-2} + F_{j-4} + \cdots = F_{j+1} - 1.$$

On the other hand, we have $\epsilon_{k-1}^\ell(n) = 1$ since the sum over all F_j, $2 \leq j \leq k-2$, is

$$\sum_{j=2}^{k-2} F_j = (F_{k-2} + F_{k-4} + \ldots) + (F_{k-3} + F_{k-5} + \ldots) = F_{k-1} - 1 + F_{k-2} - 1 = F_k - 2$$

and hence too small. The number of possible expansions with these properties is easily seen to be F_{k-1} (by induction on k), thus equal to $\#\{F_k - 1, F_k, \ldots, F_{k+1} - 2\}$, and the lemma is proved.

In order to study the joint structure of the greedy and lazy digits, we show that

$$D_k(n) = \sum_{j=2}^{k-1} (\epsilon_j^\ell(n) - \epsilon_j^g(n)) F_j = \sum_{j=k}^{\infty} (\epsilon_j^g(n) - \epsilon_j^\ell(n)) F_j$$

can only take three values.

Lemma 2. $D_k(n)$, $k \geq 3$, *can only take the values 0,* F_k *and* F_{k-1}.

Proof. We show that

$$\sum_{j\geq 3} (\epsilon_j' - \epsilon_j'') F_k = \sum_{j\geq 2} \epsilon_j F_j \tag{1}$$

with $\epsilon_j, \epsilon'_j, \epsilon''_j \in \{0,1\}$ implies

$$\sum_{j\geq 3}(\epsilon'_j - \epsilon''_j)F_{j+i} = \sum_{j\geq 2}\epsilon_j F_{j+i} - \delta F_i \tag{2}$$

for all $i > 0$ with $\delta \in \{0,1\}$. It suffices to prove (2) for $i = 1$. Then the general equation follows then by induction on i with $F_{j+i} = F_{j+i-1} + F_{j+i-2}$.

Since F_j is given by $F_j = \frac{1}{\sqrt{5}}\alpha^j - \frac{1}{\sqrt{5}}\left(-\frac{1}{\alpha}\right)^j$, we obtain

$$F_{j+1} - \alpha F_j = \frac{1}{\alpha\sqrt{5}}\left(-\frac{1}{\alpha}\right)^j + \frac{\alpha}{\sqrt{5}}\left(-\frac{1}{\alpha}\right)^j = \left(-\frac{1}{\alpha}\right)^j.$$

Hence "(2) $-\ \alpha \times$ (1)" with $i = 1$ yields

$$-\delta = \sum_{j\geq 3}(\epsilon'_j - \epsilon''_j - \epsilon_j)\left(-\frac{1}{\alpha}\right)^j - \epsilon_2\frac{1}{\alpha^2}$$

and δ is bounded by

$$-\delta < \frac{2}{\alpha^3} + \frac{1}{\alpha^4} + \frac{2}{\alpha^5} + \frac{1}{\alpha^6} + \cdots = \left(\frac{2}{\alpha^3} + \frac{1}{\alpha^4}\right)\frac{1}{1-\alpha^{-2}} = \frac{1}{\alpha}\alpha = 1.$$

Since δ is an integer, we have thus $\delta \geq 0$. For the lower bound, we get

$$-\delta > -\frac{1}{\alpha^2} - \frac{1}{\alpha^3} - \frac{2}{\alpha^4} - \frac{1}{\alpha^5} - \frac{2}{\alpha^6} - \cdots = -\left(\frac{2}{\alpha^2} + \frac{1}{\alpha^3}\right)\alpha + \frac{1}{\alpha^2} = -\alpha + \frac{1}{\alpha^2} = -1 - \frac{1}{\alpha^3}.$$

Hence $\delta \in \{0,1\}$ and (2) is proved. If either $\epsilon_2 = 0$ or $\epsilon_j = 0$ for all $j \geq 4$, then we obtain $-\delta > -1$ and thus $\delta = 0$.

Clearly we have

$$\left|\sum_{j\geq k}(\epsilon^g_j(n) - \epsilon^\ell_j(n))F_{j-k+3}\right| = \sum_{j\geq 2}\epsilon_j F_j$$

for some $\epsilon_j \in \{0,1\}$, since the term on the left side is a non-negative integer. By (2), we get

$$|D_k(n)| = \left|\sum_{j\geq k}(\epsilon^g_j(n) - \epsilon^\ell_j(n))F_j\right| = \sum_{j\geq 2}\epsilon_j F_{j+k-3} - \delta F_{k-3}$$

for all $k \geq 4$. Since $D_k(n)$ is bounded by

$$|D_k(n)| \leq \sum_{j=2}^{k-1} F_j = F_{k+1} - 2,$$

ϵ_j must be zero for all $j \geq 5$ and, if $\epsilon_2 = 1$, for $j \geq 4$. Hence we have $\delta = 0$, ϵ_j must be zero for all $j \geq 4$ and the only possible values for $|D_k(n)|$ are 0, F_k and F_{k-1}.

Since greedy expansions have no two subsequent ones and lazy expansions have no two subsequent zeros (in the range of its ones), we have, for $k \geq 4$,

$$D_k(n) \geq (F_{k-2}+F_{k-4}+\ldots)-(F_{k-1}+F_{k-3}+\ldots) = (F_{k-1}-1)-(F_k-1) = -F_{k-2}$$

and thus $D_k(n) \geq 0$ if $\epsilon_j^\ell(n) = 1$ for some $j \geq k$. Otherwise we have $\sum_{j=2}^{k-1} \epsilon_j^\ell(n) F_j = n$. Hence $D_k(n)$ is non-negative for $k \geq 4$. Clearly $|D_3(n)| \in \{0,1\}$ and $D_3(n) = D_4(n) - 2(\epsilon_3^\ell(n) - \epsilon_3^g(n))$. Because of $D_4(n) \in \{0,2,3\}$, $D_3(n)$ is non-negative and the lemma is proved.

Remark. δ in (2) can be 1, e.g. $F_3+F_5-F_4 = F_4+F_2$ and $F_4+F_6-F_5 = F_5+F_3-1$. This is due to $2F_k = F_{k+1} + F_{k-2}$, but for $k = 3$ we also have $2F_3 = F_4 + F_2$.

Lemma 3. *For $F_K \leq n \leq F_{K+1} - 2$, the digits $\epsilon_k^g(n), \epsilon_k^\ell(n)$ have the following properties:*

1. $\epsilon_k^g = 0$ *for all* $k > K$, $\epsilon_K^g = 1$, $\epsilon_{K-1}^g = 0$
2. $\epsilon_k^\ell = 0$ *for all* $k \geq K$, $\epsilon_{K-1}^\ell = 1$
3. $(\epsilon_k^g, \epsilon_k^\ell) = (1,0)$ *implies* $(\epsilon_{k-1}^g, \epsilon_{k-1}^\ell) = (0,1)$.
4. *If* $(\epsilon_{k+1}^g, \epsilon_{k+1}^\ell) \neq (0,1)$, *then* $(\epsilon_k^g, \epsilon_k^\ell) = (0,1)$ *implies* $(\epsilon_{k-1}^g, \epsilon_{k-1}^\ell) = (0,1)$ *with probability* $\frac{F_{k-3}+1}{F_k-1}$ *and* $(\epsilon_{k-1}^g, \epsilon_{k-1}^\ell) = (0,0)$, $(\epsilon_{k-1}^g, \epsilon_{k-1}^\ell) = (1,1)$ *with probabilities* $\frac{F_{k-2}-1}{F_k-1}$.
5. *If* $(\epsilon_{k+1}^g, \epsilon_{k+1}^\ell) = (0,1)$, *then* $(\epsilon_k^g, \epsilon_k^\ell) = (0,1)$ *implies* $(\epsilon_{k-1}^g, \epsilon_{k-1}^\ell) = (1,0)$ *with probability* $\frac{F_{k-2}-1}{F_{k-2}+1}$ *and* $(\epsilon_{k-1}^g, \epsilon_{k-1}^\ell) = (0,0)$, $(\epsilon_{k-1}^g, \epsilon_{k-1}^\ell) = (1,1)$ *with probabilities* $\frac{1}{F_{k-2}+1}$. *In the latter cases, the* $(\epsilon_j^g, \epsilon_j^\ell)$ *are alternately* $(0,0)$ *and* $(1,1)$ *for* $j < k$.
6. $(\epsilon_k^g, \epsilon_k^\ell) = (1,1)$ *resp.* $(\epsilon_k^g, \epsilon_k^\ell) = (0,0)$ *imply* $(\epsilon_{k-1}^g, \epsilon_{k-1}^\ell) = (0,1)$, *if the digits are not determined by 4. and* $k < K$.

Proof. 1. is obvious and 2. follows from the proof of Lemma 1. Furthermore, these n are the only integers with these properties (and their number is $F_{K-1} - 1$). 3. follows directly from the properties of greedy and lazy expansions. For the other properties, we use Lemma 2 and $D_{k-1} = D_k + (\epsilon_k^g - \epsilon_k^\ell) F_k$.

In 5., we must have $D_{k+2} = F_{k+2}$, $D_{k+1} = F_k$ and $D_k = 0$. Hence $(\epsilon_{k-1}^g, \epsilon_{k-1}^\ell)$ cannot be $(0,1)$. Furthermore, $(\epsilon_{k-1}^g, \epsilon_{k-1}^\ell) = (0,0)$ implies $D_{k-1} = 0$ and $\epsilon_{k-2}^\ell = 1$.

Thus $\epsilon^g_{k-2} = 1$. Similarly $(\epsilon^g_{k-1}, \epsilon^\ell_{k-1}) = (1,1)$ implies $(\epsilon^g_{k-2}, \epsilon^\ell_{k-2}) = (0,0)$. Inductively, we get the alternating sequence, i.e. only one possibility for the last digits. For $(\epsilon^g_{k-1}, \epsilon^\ell_{k-1}) = (1,0)$, the situation is similar to that of $k-1 = K$ and we have therefore $F_{k-2} - 1$ possibilities. This gives the stated probabilities.

In 4., we must have $D_{k+1} = F_{k+1}$ and $D_k = F_{k-1}$. Then we have $F_{k-3}+1$ possibilities for $(\epsilon^g_{k-1}, \epsilon^\ell_{k-1}) = (0,1)$ (see 5.). $(\epsilon^g_{k-1}, \epsilon^\ell_{k-1}) = (1,1)$ and $(\epsilon^g_{k-1}, \epsilon^\ell_{k-1}) = (0,0)$ imply, with $D_{k-1} = F_{k-1}$, $(\epsilon^g_{k-2}, \epsilon^\ell_{k-2}) = (0,1)$ and hence $F_{k-2} - 1$ possibilities. This also proves 6.

Remark. For $n = F_{K+1} - 1$, the unique digital expansion is given by $\epsilon_{K-2j} = 1$ for all $j \le K/2 - 1$ and $\epsilon_{K-1-2j} = 0$ for all $j < K/2 - 1$. Note that for these n, $s_g(n)$ is as large as possible whereas $s_\ell(n)$ is as small as possible (in the "neighbourhood" of n) while, for "typical" n, large $s_g(n)$ entails large $s_\ell(n)$.

Lemma 3 shows that we get simple transition probabilities from ϵ_k to ϵ_{k-1} if we exclude those n whose digital expansions terminate by alternating $(1,1)$'s and $(0,0)$'s. Thus define the sets

$$\mathcal{S}_{J,K} = \big\{n \in \{F_K, \ldots, F_{K+1}-1\} \mid (\epsilon^g_k(n), \epsilon^\ell_k(n)) \notin \{(0,0),(1,1)\} \text{ for some } k \le J\big\}$$

for $K \ge J+3$. The number of excluded n is

$$\#\big(\{F_K, F_K+1, \ldots, F_{K+1}-2\} \setminus \mathcal{S}_{J,K}\big) = F_{K-J+1}.$$

(In case $(\epsilon^g_J, \epsilon^\ell_J) = (0,0)$, we have F_{K-J} possibilities for $\epsilon^g_{J+1}, \ldots, \epsilon^g_{K-2}$, and in case $(\epsilon^g_J, \epsilon^\ell_J) = (1,1)$, we have F_{K-J-1} possibilities for $\epsilon^g_{J+2}, \ldots, \epsilon^g_{K-2}$.)

Define a sequence of random vectors $(X_{k,J,K})_{k\ge 2}$ by

$$\mathbf{Pr}[X_{k,J,K} = (b^g, b^\ell)] = \frac{1}{\#\mathcal{S}_{J,K}} \#\{n \in \mathcal{S}_{J,K} \mid \epsilon^g_k(n) = b^g, \epsilon^\ell_k(n) = b^\ell\}.$$

Lemma 3 shows that this is a Markov chain, i.e.

$$\begin{aligned}\mathbf{Pr}[X_{k-1,J,K} = (b^g_{k-1}, b^\ell_{k-1}) | X_{k,J,K} = (b^g_k, b^\ell_k), X_{k+1,J,K} = (b^g_{k+1}, b^\ell_{k+1}), \ldots] \\ = \mathbf{Pr}[X_{k-1,J,K} = (b^g_{k-1}, b^\ell_{k-1}) | X_{k,J,K} = (b^g_k, b^\ell_k)],\end{aligned}$$

if we make a distinction between $X_{k+1,J,K} = (0,1)$ and $X_{k+1,J,K} \ne (0,1)$ in case $X_{k,J,K} = (0,1)$ (otherwise we had a Markov chain of order 2), say $X_{k,J,K} = (0,1)^1$ if $X_{k,J,K} = (0,1) \ne X_{k+1,J,K}$ and $X_{k,J,K} = (0,1)^2$ if $X_{k,J,K} = (0,1) = X_{k+1,J,K}$.

The transition matrix $P_{k,J}$ defined by

$$\begin{pmatrix} \mathbf{Pr}[X_{k-1,J,K} = (0,0)] \\ \mathbf{Pr}[X_{k-1,J,K} = (0,1)^1] \\ \mathbf{Pr}[X_{k-1,J,K} = (0,1)^2] \\ \mathbf{Pr}[X_{k-1,J,K} = (1,0)] \\ \mathbf{Pr}[X_{k-1,J,K} = (1,1)] \end{pmatrix} = P_{k,J} \begin{pmatrix} \mathbf{Pr}[X_{k,J,K} = (0,0)] \\ \mathbf{Pr}[X_{k,J,K} = (0,1)^1] \\ \mathbf{Pr}[X_{k,J,K} = (0,1)^2] \\ \mathbf{Pr}[X_{k,J,K} = (1,0)] \\ \mathbf{Pr}[X_{k,J,K} = (1,1)] \end{pmatrix}$$

is, for $k \geq J+3$,

$$P_{k,J} = \begin{pmatrix} 0 & \frac{F_{k-2}-F_{k-J}}{F_k - F_{k-J+2}} & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & \frac{F_{k-3}-F_{k-J-1}}{F_k - F_{k-J+2}} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{F_{k-2}-F_{k-J}}{F_k - F_{k-J+2}} & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{\alpha^2} + \mathcal{O}\left(\alpha^{-k}\right) & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & \frac{1}{\alpha^3} + \mathcal{O}\left(\alpha^{-k}\right) & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{1}{\alpha^2} + \mathcal{O}\left(\alpha^{-k}\right) & 0 & 0 & 0 \end{pmatrix},$$

i.e. the Markov chain is almost homogeneous. Denote the limit of this matrix for $k \to \infty$ by P. Its eigenvalues are $1, -\frac{1}{\alpha}, -\frac{1}{\alpha^2}$ and 0. Thus the probability distribution is almost stationary with

$$\begin{aligned} \mathbf{Pr}[X_{k,J,K} = (0,0)] &= \frac{1}{\alpha(\alpha^2+1)} + \mathcal{O}\left(\alpha^{-\min(k,K-k)}\right) \\ \mathbf{Pr}[X_{k,J,K} = (0,1)^1] &= \frac{\alpha}{\alpha^2+1} + \mathcal{O}\left(\alpha^{-\min(k,K-k)}\right) \\ \mathbf{Pr}[X_{k,J,K} = (0,1)^2] &= \frac{1}{\alpha^2(\alpha^2+1)} + \mathcal{O}\left(\alpha^{-\min(k,K-k)}\right) \\ \mathbf{Pr}[X_{k,J,K} = (1,0)] &= \frac{1}{\alpha^2(\alpha^2+1)} + \mathcal{O}\left(\alpha^{-\min(k,K-k)}\right) \\ \mathbf{Pr}[X_{k,J,K} = (1,1)] &= \frac{1}{\alpha(\alpha^2+1)} + \mathcal{O}\left(\alpha^{-\min(k,K-k)}\right) \end{aligned}$$

for $J < k < K$.

For a given $N = \sum_{k=2}^{L} \epsilon_k^g(N) F_k$ with $\epsilon_L^g(N) = 1$ (i.e. $L \approx \log_\alpha N$), define

$$\mathcal{S}_N = \bigcup_{k=L-[L^\eta]}^{L} \bigcup_{K=L-[L^\eta]}^{k-1} \left(\mathcal{S}_{[L^\eta],K} + \sum_{j=k+1}^{L} \epsilon_j^g(N) F_j \right)$$

and a sequence of random vectors $(Y_{k,N})_{k\geq 2}$ by

$$\mathbf{Pr}[Y_{k,N} = (b^g, b^\ell)] = \frac{1}{\#\mathcal{S}_N} \#\{n \in \mathcal{S}_N \mid \epsilon_k^g(n) = b^g, \epsilon_k^\ell(n) = b^\ell\}.$$

This sequence is close to what we need because of

$$\#\big(\{0,\ldots,N-1\} \setminus \mathcal{S}_N\big) = \mathcal{O}\left(L^\eta F_{L-[L^\eta]} + L^{2\eta} F_{L-2[L^\eta]}\right) = \mathcal{O}\left(\frac{(\log N)^\eta N}{\alpha^{(\log_\alpha N)^\eta}}\right) \quad (3)$$

and, for $[L^\eta] \leq k \leq [L - L^\eta]$, the $Y_{k,N}$ are a Markov chain with transition matrices $P_{k,[L^\eta]}$. For $[L^\eta] \leq k \leq L - 2[L^\eta]$, the distribution of $Y_{k,N}$ is thus almost stationary with the probabilities of $X_{k,J,K}$ and error terms $\mathcal{O}\left(\alpha^{-L^\eta}\right)$.

Lemma 4. *The $Y_{k,N} = (Y^g_{k,N}, Y^\ell_{k,N})$ satisfy a central limit theorem for $L^\eta \le k \le L - 2L^\eta$. More precisely, we have, for all $a_g, a_\ell \in \mathbb{R}$, as $N \to \infty$,*

$$\sum_{k=[L^\eta]}^{L-2[L^\eta]} \frac{a_g(Y^g_{k,N} - \mu_g) + a_\ell(Y^\ell_{k,N} - \mu_\ell)}{\sigma\sqrt{L}} \Rightarrow \mathcal{N}(0, a_g^2 + a_\ell^2 + 2a_g a_\ell C),$$

where $\mathcal{N}(M,V)$ denotes the normal law with mean value M and variance V.

Proof. For the mean value, we have

$$\begin{aligned}\mathbf{E}\, Y^g_{k,N} &= \mathbf{Pr}[Y^g_{k,N} = (1,0)] + \mathbf{Pr}[Y^g_{k,N} = (1,1)] \\ &= \frac{1}{\alpha^2(\alpha^2+1)} + \frac{1}{\alpha(\alpha^2+1)} + \mathcal{O}\left(\alpha^{-L^\eta}\right) = \mu_g + \mathcal{O}\left(\alpha^{-L^\eta}\right)\end{aligned}$$

and

$$\mathbf{E}\, Y^\ell_{k,N} = \mu_\ell + \mathcal{O}\left(\alpha^{-L^\eta}\right).$$

Hence the mean value of the sum converges to zero. The variance is given by

$$\begin{aligned}&\mathbf{Var}\left(\sum_{k=[L^\eta]}^{L-2[L^\eta]} a_g(Y^g_{k,N} - \mu_g) + a_\ell(Y^\ell_{k,N} - \mu_\ell)\right) \\ &= \mathbf{Var} \sum_{k=[L^\eta]}^{L-2[L^\eta]} a_g Y^g_{k,N} + \mathbf{Var} \sum_{k=[L^\eta]}^{L-2[L^\eta]} a_\ell Y^\ell_{k,N} + 2\,\mathbf{Cov}\left(\sum_{k=[L^\eta]}^{L-2[L^\eta]} a_g Y^g_{k,N}, \sum_{k=[L^\eta]}^{L-2[L^\eta]} a_\ell Y^\ell_{k,N}\right) \\ &= L\sigma^2(a_g^2 + a_\ell^2) + \mathcal{O}\left(L^\eta\right) + 2a_g a_\ell \sum_{k=[L^\eta]}^{L-2[L^\eta]} \sum_{j=[L^\eta]-k}^{L-2[L^\eta]-k} \mathbf{Cov}\left(Y^g_{k,N}, Y^\ell_{k+j,N}\right).\end{aligned}$$

(The calculation of the variance of $\sum Y^g_{k,N}$ and $\sum Y^\ell_{k,N}$ is similar to that in [1] and to that of the covariance hereafter.) The covariance is given by

$$\mathbf{Cov}(Y^g_{k,N}, Y^\ell_{k+j,N}) = \mathbf{Pr}[Y^g_{k,N} = 1, Y^\ell_{k+j,N} = 1] - \mathbf{Pr}[Y^g_{k,N} = 1]\mathbf{Pr}[Y^\ell_{k+j,N} = 1].$$

For $j = 0$, we obtain, with $(\alpha^2+1)^2 = 5\alpha^2$,

$$\mathbf{Cov}(Y^g_{k,N}, Y^\ell_{k,N}) = \frac{1}{\alpha(\alpha^2+1)} - \frac{\alpha^2}{(\alpha^2+1)^2} + \mathcal{O}\left(\alpha^{-L^\eta}\right) = -\frac{1}{5\alpha^4} + \mathcal{O}\left(\alpha^{-L^\eta}\right).$$

The approximated transition matrix has the form $P = QDQ^{-1}$

$$\begin{pmatrix} \frac{1}{\alpha(\alpha^2+1)} & 1 & -\frac{\alpha^3}{\alpha^2+1} & 1 & 1 \\ \frac{\alpha}{\alpha^2+1} & -\alpha & \frac{\alpha^3}{\alpha^2+1} & 0 & 0 \\ \frac{1}{\alpha^2(\alpha^2+1)} & \frac{1}{\alpha} & -\frac{\alpha^2}{\alpha^2+1} & 0 & 0 \\ \frac{1}{\alpha^2(\alpha^2+1)} & -1 & \frac{\alpha^4}{\alpha^2+1} & -1 & -2 \\ \frac{1}{\alpha(\alpha^2+1)} & 1 & -\frac{\alpha^3}{\alpha^2+1} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{\alpha} & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{\alpha^2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & -\frac{1}{\alpha} & -\alpha & 1 & 1 \\ 1 & -\frac{1}{\alpha^2} & -\alpha^2 & 1 & 1 \\ 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & -\alpha & 0 & 1 \end{pmatrix}$$

and the transition matrix of order j ($P^j = QD^jQ^{-1}$) is given by

$$P^j = \frac{1}{\alpha^2+1}\begin{pmatrix} \frac{1}{\alpha} & \frac{1}{\alpha} & \frac{1}{\alpha} & \frac{1}{\alpha} & \frac{1}{\alpha} \\ \alpha & \alpha & \alpha & \alpha & \alpha \\ \frac{1}{\alpha^2} & \frac{1}{\alpha^2} & \frac{1}{\alpha^2} & \frac{1}{\alpha^2} & \frac{1}{\alpha^2} \\ \frac{1}{\alpha^2} & \frac{1}{\alpha^2} & \frac{1}{\alpha^2} & \frac{1}{\alpha^2} & \frac{1}{\alpha^2} \\ \frac{1}{\alpha} & \frac{1}{\alpha} & \frac{1}{\alpha} & \frac{1}{\alpha} & \frac{1}{\alpha} \end{pmatrix} + \left(-\frac{1}{\alpha}\right)^j \begin{pmatrix} 1 & -\frac{1}{\alpha} & 1 & -\alpha & 1 \\ -\alpha & 1 & -\alpha & \alpha^2 & -\alpha \\ \frac{1}{\alpha} & -\frac{1}{\alpha^2} & \frac{1}{\alpha} & -1 & \frac{1}{\alpha} \\ -1 & \frac{1}{\alpha} & -1 & \alpha & -1 \\ 1 & -\frac{1}{\alpha} & 1 & -\alpha & 1 \end{pmatrix}$$
$$+\frac{1}{\alpha^2+1}\left(-\frac{1}{\alpha^2}\right)^j \begin{pmatrix} -\alpha^3 & \alpha & -\alpha^3 & \alpha^5 & -\alpha^3 \\ \alpha^4 & -\alpha^2 & \alpha^4 & -\alpha^6 & \alpha^4 \\ \alpha^3 & -\alpha & \alpha^3 & -\alpha^5 & \alpha^3 \\ -\alpha^2 & 1 & -\alpha^2 & \alpha^4 & -\alpha^2 \\ -\alpha^3 & \alpha & -\alpha^3 & \alpha^5 & -\alpha^3 \end{pmatrix}$$

Clearly

$$\begin{aligned}
&\mathbf{Pr}[Y^g_{k,N} = 1, Y^\ell_{k+j,N} = 1] \\
&= \mathbf{Pr}[Y_{k+j,N} = (0,1)^1]\Big(\mathbf{Pr}[Y_{k,N} = (1,0)|Y_{k+j,N} = (0,1)^1] + \mathbf{Pr}[Y_{k,N} = (1,1)|Y_{k+j,N} = (0,1)^1]\Big) \\
&\quad + \mathbf{Pr}[Y_{k+j,N} = (0,1)^2]\Big(\mathbf{Pr}[Y_{k,N} = (1,0)|Y_{k+j,N} = (0,1)^2] + \mathbf{Pr}[Y_{k,N} = (1,1)|Y_{k+j,N} = (0,1)^2]\Big) \\
&\quad + \mathbf{Pr}[Y_{k+j,N} = (1,1)]\Big(\mathbf{Pr}[Y_{k,N} = (1,0)|Y_{k+j,N} = (1,1)] + \mathbf{Pr}[Y_{k,N} = (1,1)|Y_{k+j,N} = (1,1)]\Big).
\end{aligned}$$

Note that the contribution of the first matrix of P^j to this probability is just $\mu_g\mu_\ell$ and that of the second matrix is zero. Hence we have, for $j > 0$,

$$\begin{aligned}
\mathbf{Cov}(Y^g_{k,N}, Y^\ell_{k+j,N}) &= \frac{1}{\alpha^2+1}\left(-\frac{1}{\alpha^2}\right)^j \left(\frac{\alpha(1+\alpha)}{\alpha^2+1} + \frac{-\alpha^2-\alpha^3}{\alpha^2(\alpha^2+1)} + \frac{-\alpha^2-\alpha^3}{\alpha(\alpha^2+1)}\right) + \mathcal{O}\left(\alpha^{-L^\eta}\right) \\
&= -\frac{1}{5}\left(-\frac{1}{\alpha^2}\right)^j + \mathcal{O}\left(\alpha^{-L^\eta}\right).
\end{aligned}$$

For $j < 0$, we get similarly

$$\mathbf{Cov}(Y^g_{k,N}, Y^\ell_{k-|j|,N}) = -\frac{1}{5}\left(-\frac{1}{\alpha^2}\right)^{|j|} + \mathcal{O}\left(\alpha^{-L^\eta}\right).$$

Therefore we have

$$\begin{aligned}
&\sum_{j=[L^\eta]-k}^{L-2[L^\eta]-k} \mathbf{Cov}\left(Y^g_{k,N}, Y^\ell_{k+j,N}\right) \\
&= -\frac{1}{5}\left(\frac{1}{\alpha^4} + 2\sum_{j=1}^{\infty}\left(-\frac{1}{\alpha^2}\right)^j\right) + \mathcal{O}\left(L\alpha^{-L^\eta}\right) + \mathcal{O}\left(\alpha^{-2\min(k-[L^\eta],L-2[L^\eta]-k)}\right)
\end{aligned}$$

With

$$C = -\frac{1}{5\sigma^2}\left(\frac{1}{\alpha^4} + 2\sum_{j=1}^{\infty}\left(-\frac{1}{\alpha^2}\right)^j\right) = -\frac{\alpha^2+1}{\alpha}\left(\frac{1}{\alpha^4} - \frac{2}{\alpha^2+1}\right) = 9 - 5\alpha,$$

we obtain

$$\mathbf{Var}\left(\sum_{k=[L^\eta]}^{L-2[L^\eta]} a_g(Y_{k,N}^g - \mu_g) + a_\ell(Y_{k,N}^\ell - \mu_\ell)\right) = L\sigma^2(a_g^2 + a_\ell^2 + 2a_g a_\ell C) + \mathcal{O}\left(L^\eta\right).$$

We apply the central limit theorem for Markov chains or mixing sequences (e.g. Theorem 2.1 of Peligrad [6]) and the lemma is proved.

Because of (3), we have

$$\begin{aligned}
&\frac{1}{N}\#\left\{n < N \,\middle|\, \frac{s_g(n) - \mu_g \log_\alpha N}{\sigma\sqrt{\log_\alpha N}} < x_g, \frac{s_\ell(n) - \mu_\ell \log_\alpha N}{\sigma\sqrt{\log_\alpha N}} < x_\ell\right\} \\
&\quad \to \frac{1}{\#\mathcal{S}_N}\#\left\{n \in \mathcal{S}_N \,\middle|\, \frac{s_g(n) - \mu_g \log_\alpha N}{\sigma\sqrt{\log_\alpha N}} < x_g, \frac{s_\ell(n) - \mu_\ell \log_\alpha N}{\sigma\sqrt{\log_\alpha N}} < x_\ell\right\} \\
&\quad \to \frac{1}{\#\mathcal{S}_N}\#\left\{n \in \mathcal{S}_N \,\middle|\, \sum_{k=[L^\eta]}^{L-2[L^\eta]} \frac{\epsilon_k^g(n) - \mu_g}{\sigma\sqrt{L}} < x_g, \sum_{k=[L^\eta]}^{L-2[L^\eta]} \frac{\epsilon_k^\ell(n) - \mu_\ell}{\sigma\sqrt{L}} < x_\ell\right\}.
\end{aligned}$$

With Lemma 4, the theorem is proved.

References

1. M. Drmota and W. Steiner, *The Zeckendorf Expansion of Polynomial Sequences*, to appear in J. Théor. Nombres Bordx., http://www.geometrie.tuwien.ac.at/steiner/ze.ps.gz.
2. J. M. Dumont and A. Thomas, *Digital sum moments and substitutions*, Acta Arith. **64** (1993), no. 3, 205–225.
3. J. M. Dumont and A. Thomas, *Gaussian asymptotic properties of the sum-of-digits functions*, J. Number Th. **62** (1997), 19–38.
4. P. Erdős and I. Joó, *On the Expansion* $1 = \sum q^{-n_i}$, Period. Math. Hung. **23** (1991), no. 1, 25–28.
5. P. J. Grabner and R. F. Tichy, *Contributions to digit expansions with respect to linear recurrences*, J. Number Th. **36** (1990), 160–169.
6. M. Peligrad, *On the Asymptotic Normality of Sequences of Weak Dependent Random Variables*, J. Theor. Probab. **9** (1996), no. 3, 703–715.
7. 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.

AMS Classification Numbers: 11A63, 11N60