THE DISTRIBUTION OF THE SUM-OF-DIGITS FUNCTION

MICHAEL DRMOTA AND JOHANNES GAJDOSIK

Abstract. By using a generating function approach it is shown that the sum-of-digits function (related to specific finite and infinite linear recurrences) satisfies a central limit theorem. Additionally a local limit theorem is derived.

1. Introduction

Let $G = (G_j)_{j\ge 0}$ be a strictly increasing sequence of integers with $G_0 = 1$. Then every non-negative integer n has a (unique) proper G-ary digital expansion

$$n = \sum_{j\ge 0} \epsilon_j(n) G_j$$

with integer digits $\epsilon_j(n) \ge 0$ provided that

$$\sum_{j<k} \epsilon_j(n) G_j < G_k$$

for all $k \ge 0$. The sum-of-digits functions $s_G(n)$ is given by

$$s_G(n) = \sum_{j\ge 0} \epsilon_j(n)$$

and the aim of this paper is to get an insight to the distribution of $s_G(n)$, i.e. to the behaviour of the numbers

$$a_{Nk} = |\{n < N : s_G(n) = k\}|.$$

It is also very convenient to consider a related sequence of (discrete) random variables X_N, $N \ge 0$, defined by

$$\mathbf{Pr}[X_N = k] = \frac{a_{Nk}}{N}.$$

Expected value and variance of X_N are given by

$$\mathbf{E}X_N = \frac{1}{N}\sum_{n<N} s_G(n) \quad \text{and by} \quad \mathbf{V}X_N = \frac{1}{N}\sum_{n<N} (s_G(n) - \mathbf{E}X_N)^2 \tag{1.1}$$

There is a vast literature concerning asymptotic properties of $\mathbf{E}X_N$ and $\mathbf{V}X_N$ and on the distribution of X_N.

Asymptotic and exact formulas for $\mathbf{E}X_N$ are due to Bush [3], Bellman and Shapiro [2], Delange [5], and Trollope [20] for q-ary digital expansions and due to Pethö and Tichy [17] and Grabner and Tichy [12, 13] for G-ary digital expansions with respect to linear recurrences. Corresponding formulas for higher moments $\mathbf{E}X_N^d$ and for the variance $\mathbf{V}X_N$ can be found in Coquet [4], Kirschenhofer [15], Kennedy and Cooper [14], Grabner, Kirschenhofer, Prodinger, and Tichy [11], and in Dumont and Thomas [7].

The asymptotic distribution of X_N and related problems are discussed in Schmidt [19], Schmid [18], Bassily and Katai [1], and in Dumont and Thomas [8].

Date: January 16, 1998.

Key words and phrases. digital expansions, central limit theorem.

Institut für Algebra und Diskrete Mathematik, TU Wien, Wiedner Hauptstrasse 8-10/118, A-1040 Wien, Austria, email:drmota@tuwien.ac.at.

This research was supported by the Austrian Science Foundation FWF, grant P10187-PHY.

The main purpose of this paper is to prove asymptotic normality (of the distribution of X_N) by the use of generating functions, where it is also possible to derive a local limit law. (A similar approach was used in [6].)

2. Results

In the present paper we will deal with basis sequences $G = (G_j)_{j\ge 0}$ which satisfy specific finite or infinite linear recurrences.

2.1. **Finite Recurrences.**

In the first case we will make the following assumptions:

1. There exist non-negative integers a_i, $1 \le i \le r$, such that (for $j \ge r$)
$$G_j = \sum_{i=1}^{r} a_i G_{j-i}.$$
2. $\gcd\{\mathrm{i} \ge 1 : \mathrm{a_i} \ne 0\} = 1$.
3. For all $j > r$ and $1 \le k < r$ we have
$$G_{j-k} \ge \sum_{i=k+1}^{r} a_i G_{j-i}.$$

In section 3 we will show that the above assumptions imply that the characterictic polynomial
$$P(u) = u^r - \sum_{i=1}^{r} a_i u^{r-i}$$
has a unique root α of maximal modulus which is real and positive (i.e. all other roots α' of $P(u)$ satisfy $|\alpha'| < \alpha$) and that
$$G_j \sim C\alpha^j \quad (j \to \infty)$$
for some constant $C > 0$.

Remark . Usually (e.g. see [12]) it is assumed that $a_1 \ge a_2 \ge \cdots \ge a_r > 0$ and that $G_j = \sum\limits_{i=1}^{j} a_i G_{j-i} + 1$ for $j < r$. In this case all assumptions are satisfied. (Furthermore, $P(u)$ is irreducible and α is a Pisot number.)
If the sequence of a_i, $1 \le i \le r$, is not decreasing then the situation is more complicated, e.g. if $a_1 = a_r = 1$ and $a_i = 0$ for $i \ne 1, r$ then condition 3. is satisfied, too. However, if $r = 4$, $a_1 = a_3 = a_4 = 1$, and $a_2 = 0$ then 3. is violated.

2.2. **Infinite Recurrences.**

In the second case our starting point is Parry's α-expansion of 1 (see [16])
$$1 = \frac{a_1}{\alpha} + \frac{a_2}{\alpha^2} + \frac{a_3}{\alpha^3} + \cdots,$$
where $\alpha > 1$ is a real number and a_i, $i \ge 1$ are positive integers. (In the case of ambiguity we take the infinite representation of 1.) The sequence $G = (G_j)_{j\ge 0}$ is now defined by
$$G_0 = 1,\ G_j = \sum_{i=1}^{j} a_i G_{j-i} + 1 \quad (j > 0). \tag{2.1}$$
If we further set
$$A(u) = \sum_{i\ge 1} a_i u^i \quad \text{and} \quad G(u) = \sum_{j\ge 0} G_j u^j$$
then
$$G(u) = \frac{1}{(1-u)(1-A(u))}$$
and it follows that $z_0 = 1/\alpha > 0$ is the only singularity on the circle of convergence $|z| = z_0$, which is a simple pole. Hence
$$G_j \sim C'\alpha^j \quad (j \to \infty)$$

and so we are in similar situation as above.

2.3. **Asymptotic Properties.** First we state a theorem concerning expected value and variance of X_N (defined) in (1.1). Actually this statement is more or less a collection of well known facts (see [5, 17, 13, 15, 7, 10]). More precisely, much more is known about the following $\mathcal{O}(1)$-terms. Therefore we will not present a proof.

Theorem 2.1. *Suppose that $G=(G_j)_{j\ge 0}$ satisfies a finite or infinite linear recurrence of the above types. Set*

$$G(z,u)=\sum_{j\ge 1}\left(\sum_{l=0}^{a_j-1} z^l\right) z^{a_1+\cdots+a_{j-1}} u^j$$

and let $1/\alpha(z)$ denote the (analytic) solution $u=1/\alpha(z)$ of the equation

$$G(u,z)=1$$

for z in a sufficiently small (complex) neighbourhood of $z_0=1$ such that $\alpha(1)=\alpha$. Then

$$\mathbf{E}X_N=\frac{1}{N}\sum_{n<N} s_G(n)=\mu\frac{\log N}{\log\alpha}+\mathcal{O}(1)$$

and

$$\mathbf{V}X_N=\frac{1}{N}\sum_{n<N}(s_G(n)-\mathbf{E}X_N)^2=\sigma^2\frac{\log N}{\log\alpha}+\mathcal{O}(1),$$

where

$$\mu=\frac{\alpha'(1)}{\alpha}\quad \text{and}\quad \sigma^2=\frac{\alpha''(1)}{\alpha}+\mu-\mu^2.$$

Our main result concerns the distribution properties of X_N. We prove asymptotic normality in the weak sense and provide a local limit law.

Theorem 2.2. *Suppose that $G=(G_j)_{j\ge 0}$ satisfies a finite or infinite linear recurrence of the above types. If $\sigma^2\ne 0$ then for every $\varepsilon>0$*

$$\frac{1}{N}|\{n<N: s_G(n)<\mathbf{E}X_N+x\mathbf{V}X_N\}|=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}t^2}dt+\mathcal{O}((\log N)^{-1/2+\varepsilon}) \tag{2.2}$$

uniformly for all real x as $N\to\infty$.
Furthermore

$$|\{n<N: s_G(n)=k\}|=\frac{N}{\sqrt{2\pi\mathbf{V}X_N}}\left(\exp\left(-\frac{(k-\mathbf{E}X_N)^2}{2\mathbf{V}X_N}\right)+\mathcal{O}((\log N)^{-1/2+\varepsilon})\right) \tag{2.3}$$

uniformly for all non-negative integers k as $N\to\infty$.

In section 3. we collect some preliminaries which will be used in sections 4. and 5. for the proofs of (2.2) and (2.3).

3. Preliminaries

3.1. **Finite Recurrences.** We will first collect some basis facts which will be needed in the sequel.

Lemma 3.1. *Suppose that $G_0,G_1,\dots,G_{r-1}$ are positive, that $G_j=\sum_{i=1}^r a_iG_{j-i}$ for $j\ge r$, where $a_i\ge 0$, $1\le i\le r$, and that $\gcd\{i\ge 0: a_i\ne 0\}=1$. Then the characterictic polynomial $P(u)=u^r-\sum\limits_{i=1}^{r} a_iu^{r-i}$ has a unique root α of maximal modulus which is real and >1. Furthermore,*

$$G_j=C\alpha^j+\mathcal{O}(\alpha^{(1-\delta)j}) \tag{3.1}$$

for a real constant $C>0$ and some $\delta>0$.

Proof. First we show that $P(u)$ has a unique positive real root $\alpha > 1$ of maximal modulus. Set $G(u) = 1 - u^r P(u^{-1}) = \sum_{j=1}^{r} a_j u^j$. Then $G(u)$ is strictly increasing for real $u \ge 0$. Since $G(0) = 0$ and $\lim_{u\to\infty} G(u) = \infty$ there uniquely exists $u_0 > 0$ with $G(u_0) = 1$. Since G_n is strictly increasing we have $\sum_{j=1}^{r} a_j = G(1) > 1$ and consequently $u_0 < 1$. Furthermore, $G'(u_0) = \sum_{j=1}^{r} j a_j u_0^{j-1} > 0$. Thus, $\alpha = 1/u_0 > 1$ is a simple root of $P(u)$.

If $|u| < u_0$ then
$$|G(u)| \le G(|u|) < G(u_0) = 1.$$
Furthermore, if $|u| = u_0$ but $u \ne u_0$ then the gcd-condition $\gcd\{i \ge 1 : a_i \ne 0\} = 1$ implies that
$$|G(u)| < G(|u|) = 1.$$
Consequently, there are no roots of $P(u)$ other than α with modulus $\ge \alpha$.

Next it is clear that G_j has a representation of the form (3.1) for some real C. We only have to show that $C > 0$. For this purpose define $F_j(x_0, \dots, x_{r-1})$ by $F_j(x_0, \dots, x_{r-1}) = x_j$ if $0 \le j < r$ and by
$$F_j(x_0, \dots, x_{r-1}) = \sum_{i=1}^{r} a_i F_{j-1}(x_0, \dots, x_{r-1})$$
for $j \ge r$. Then $F_j(x_0, \dots, x_{r-1})$ is multilinear and monotonic in all variables. Furthermore, $F_j(G_0, \dots, G_{r-1}) = G_j$ and $F_j(1, \alpha, \dots, \alpha^{r-1}) = \alpha^j$. Hence, by setting $c_1 = \min_{0\le j<r} G_j \alpha^{-j}$ we obtain
$$c_1\alpha^j = F_j(c_1, c_1\alpha, \dots, c_1\alpha^{r-1}) \le F_j(G_0, \dots, G_{r-1}) = G_j = C\alpha^j + \mathcal{O}(\alpha^{(1-\delta)j}).$$
Thus $C > 0$. □

Lemma 3.2. *Suppose that $G = (G_j)_{j\ge0}$ satisfies the above conditions 1.–3.*

If $0 \le k < a_i$ (for $1 \le i \le r$), $0 \le m < G_{j-i}$, and $j \ge r$ then
$$n = a_1 G_{j-1} + \cdots + a_{i-1} G_{j-i+1} + k G_{j-i} + m$$
has digits
$$\begin{aligned} \epsilon_l(n) &= \epsilon_l(m) \quad (0 \le l < j - i) \\ \epsilon_{j-i}(n) &= k \\ \epsilon_l(n) &= a_{j-l} \quad (j - i < l < j). \end{aligned}$$

Suppose that n has the digital expansion $n = \epsilon_L(n)G_L + \epsilon_{L-1}(n)G_{L-1} + \cdots + \epsilon_0(n)G_0$. If $j \le L$, $k < \epsilon_j(n)$, and $m < G_j$ then
$$n' = \epsilon_L(n)G_L + \epsilon_{L-1}(n)G_{L-1} + \cdots + \epsilon_{j+1}(n)G_{j+1} + kG_j + m$$
has digits
$$\begin{aligned} \epsilon_l(n') &= \epsilon_l(m) \quad (0 \le l < j) \\ \epsilon_j(n') &= k \\ \epsilon_l(n') &= \epsilon_l(n) \quad (j < l \le L). \end{aligned}$$

Proof. Since $kG_{j-i} + m < a_i G_i$ we obtain for $1 \le i' < i$ by condition 3.
$$n - \sum_{i''=1}^{i'} a_{i''} G_{j-i''} < G_j - \sum_{i''=1}^{i'} a_{i''} G_{j-i''} = \sum_{i''=i'+1}^{r} a_{i''} G_{j-i''} \le G_{j-i'}.$$
Thus, $\epsilon_l(n) = a_{j-l}$ for $j - i < l < j$. Similarly,
$$n - \sum_{i''=1}^{i-1} a_{i''} G_{j-i''} - kG_{j-i} = m < G_{j-i}$$

implies $\epsilon_{j-i}(n) = k$ and consequently $\epsilon_l(n) = \epsilon_l(m)$ for $0 \le l < j-i$. This completes the proof of the first part.

Since $kG_j + m < \epsilon_j(n)G_j$ we have for $0 \le i < L-j$

$$n' - \sum_{i'=0}^{i} \epsilon_{L-i'}G_{L-i'} < G_{L-i+1}$$

which gives $\epsilon_l(n') = \epsilon_l(n)$ for $j < l \le L$. Finally

$$n' - \sum_{i'=0}^{L-j-1} \epsilon_{L-i'}G_{L-i'} - kG_j < G_j$$

provides $\epsilon_j(n') = k$ and $\epsilon_l(n') = \epsilon_l(m)$ for $0 \le l < j$. □

Next let

$$b_{j,k} = a_{G_j k} = |\{n < G_j : s_G(n) = k\}|$$

and set

$$b_j(z) = \sum_{k \ge 0} b_{j,k} z^k$$

and

$$B(z,u) = \sum_{j \ge 0} b_j(z) u^j.$$

Lemma 3.3. *For $j \ge r$ we have*

$$b_j(z) = \sum_{i=1}^{r} \sum_{l=0}^{a_i - 1} z^{a_1 + \cdots + a_{i-1} + l} b_{j-i}(z). \tag{3.2}$$

Proof. First observe that the set $\{n \in \mathbf{Z} : 0 \le n < G_j\}$ $(j \ge r)$ can be represented as a disjoint union of the form

$$\bigcup_{i=1}^{r} \bigcup_{l=0}^{a_i - 1} \left\{ \sum_{h=1}^{i-1} a_h G_{h-j} + lG_{j-i} + m : 0 \le m < G_{j-i} \right\}.$$

Thus by Lemma 3.2

$$b_{j,k} = \sum_{i=1}^{r} \sum_{l=0}^{a_i - 1} b_{j-i, k-a_1-\cdots-a_{i-1}-l}.$$

which provides (3.2). □

Corollary . *We have*

$$B(z,u) = \frac{P(z,u)}{1 - G(z,u)},$$

in which $G(z,u)$ is defined in Theorem 2.1 and

$$P(z,u) = \sum_{j=1}^{r} b_j(z) \left(1 - \sum_{i=1}^{r-j} \left(\sum_{l=0}^{a_i - 1} z^l \right) z^{a_1 + \cdots + a_{i-1}} u^i \right) u^j.$$

3.2. **Infinite Recurrences.** In the case of digital expansions which are related to Parry's α-expansion we have similar properties.

Lemma 3.4. ([13]) *Let $G = (G_j)_{j \ge 0}$ be given by (2.1). Then a finite sequence $(\epsilon_0, \epsilon_1, \dots, \epsilon_L)$ of nonnegative integers constitute the G-ary digits $\epsilon_j = \epsilon_j(n)$ of $n = \sum_{j=0}^{L} \epsilon_j G_j$ if and only if*

$$(\epsilon_k, \epsilon_{k-1}, \dots, \epsilon_0, 0, 0, \dots) < (a_1, a_2, \dots) \tag{3.3}$$

for $k = 0, 1, \dots, L$, where "$<$" denotes the lexicographic order.

Remark . Note that (3.3) implies that Lemma 3.2 holds for the infinite case, too.

As above let

$$b_j(z) = \sum_{k\ge 0} b_{j,k} z^k,$$

where $b_{j,k} = a_{G_j k} = |\{n < G_j : s_G(n) = k\}|$, and

$$B(z,u) = \sum_{j\ge 0} b_j(z) u^j.$$

By using (3.3) (or the corresponding version of Lemma 3.2 for the infinite case) we obtain a similar representation of $B(z,u)$ as in Corollary of Lemma 3.3 in the case of finite recurrences.

Lemma 3.5. ([13]) *We have*

$$B(z,u) = \frac{P(z,u)}{1-G(z,u)},$$

in which $G(z,u)$ is defined in Theorem 2.1 and

$$P(z,u) = 1 + \sum_{j\ge 1} z^{a_1+\cdots+a_j} u^j.$$

3.3. **Composition.** A similar procedure which led us to recurrences for $b_{j,k}$ (resp. to generating function identities in Corollary of Lemma 3.3 and to Lemma 3.5) can also be used to extract a_{Nk} from $b_{j,k} = a_{G_j k}$.

Lemma 3.6. *Suppose that*

$$N = \sum_{l=0}^{L-1} e_l G_{j_l}$$

is the G-ary digital expansion of N, where $j_0 > j_1 > \cdots > j_{L-1}$ and $e_j > 0$. Then

$$\sum_{k\ge 0} a_{Nk} z^k = \sum_{l=0}^{L-1} b_{j_l}(z) z^{\sum_{h=0}^{l-1} e_h} \sum_{i=0}^{e_l-1} z^i. \tag{3.4}$$

Proof. By Lemma 3.2 (which is also valid in the case of infinite recurrences) we have

$$\begin{aligned} a_{Nk} &= \sum_{l=0}^{L-1} \sum_{i=0}^{e_l-1} \left| \left\{ m < G_{j_l} : s_G(m) = k - \sum_{h=0}^{l-1} e_h - i \right\} \right| \\ &= \sum_{l=0}^{L-1} \sum_{i=0}^{e_l-1} b_{j_l, k-\sum_{h=0}^{l-1} e_h - i}. \end{aligned} \tag{3.5}$$

Thus (3.4) follows. □

4. Global Limit Law

The first step is to obtain proper information of $b_j(z)$.

Proposition 4.1. *Suppose that $G = (G_j)_{j\ge 0}$ satisfies a finite or infinite recurrence of the above types. Then*

$$b_j(z) = C(z)\alpha(z)^j + \mathcal{O}\left(\alpha^{(1-\delta)j}\right) \tag{4.1}$$

uniformly for z contained in a sufficiently small complex neighbourhood of $z_0 = 1$ as $j \to \infty$, where $\alpha(z)$ is defined in Theorem 2.1 and $C(z)$ is an analytic function with $C(1) = C$ resp. $C(1) = C'$.

Proof. Firstly, let G_n satisfy a finite linear recurrence of the above type. Then $1 - G(1,u) = u^r P(u^{-1})$, where $P(u)$ is the characteristic polynomial. Thus $G(1,\alpha^{-1}) = 0$. Furthermore, since $\frac{\partial}{\partial u} G(z,u) < 0$ for real and non-negative z, u there exists an analytic function $\alpha(z)$ (for z in a sufficiently small complex neighbourhood of $z_0 = 1$) with $G(z, 1/\alpha(z)) = 0$ and $\alpha(1) = \alpha$. Similarly we have $G(1,\alpha^{-1}) = 0$ and $\frac{\partial}{\partial u} G(z,u) < 0$ in the case of infinite linear recurrence. Thus, there also exists an analytic function $\alpha(z)$ with $G(z,1/\alpha(z)) = 0$ and $\alpha(1) = \alpha$.

In both cases there exist analytic functions $G_1(z,u)$, $R(z,u)$, $R_1(z,u)$ such that

$$
\begin{aligned}
1 - G(z,u) &= \left(u - \frac{1}{\alpha(z)}\right) G_1(z,u),\\
G_1(z,1/\alpha(z)) &\neq 0,\\
\frac{P(z,u)}{G_1(z,u)} &= \frac{P(z,1/\alpha(z))}{G_1(z,1/\alpha(z))} + \left(u - \frac{1}{\alpha(z)}\right) R_1(z,u)
\end{aligned}
$$

for z, u in a complex neighbourhood of $z_0 = 1, u_0 = \alpha$. Since $1/\alpha = 1/\alpha(1)$ is the unique (and simple) zero of $G(1,u) = 0$ on the circle $|u| = 1/\alpha$ and since there are no zeroes for $|u| < 1/\alpha$ the function $G_1(z,u)$ can be analytically continued to $|u| < 1/\alpha + \varepsilon$ (for some sufficiently small $\varepsilon > 0$) if z varies in a (sufficiently small) neighbourhood of $z_0 = 1$. Withoug loss of generality we may assume that $|1/\alpha(z)| \le 1/\alpha + \varepsilon/4$. Hence, $R_1(z,u)$ can be analytically continued to the same region and we obtain for $|u| < 1/\alpha + \varepsilon$

$$
\frac{P(z,u)}{1 - G(z,u)} = \frac{C(z)}{1 - u\alpha(z)} + R_1(z,u),
$$

where $C(z) = -\alpha(z)P(z,1/\alpha(z))/G_1(z,1/\alpha(z))$. Finally, by Cauchy's formula we get

$$
\begin{aligned}
b_j(z) &= \frac{1}{2\pi i} \int\limits_{|u|=1/\alpha+\varepsilon/2} \frac{P(z,u)}{1-G(z,u)} \frac{du}{u^{j+1}}\\
&= C(z)\alpha(z)^j + \frac{1}{2\pi i} \int\limits_{|u|=1/\alpha+\varepsilon/2} R_1(z,u) \frac{du}{u^{j+1}}\\
&= C(z)\alpha(z)^j + \mathcal{O}\left(\alpha^{(1-\delta)j}\right),
\end{aligned}
$$

with some $\delta > 0$. This completes the proof of Proposition 4.1. □

With help of Proposition 4.1 and Lemma 3.6 we can prove asymptotic normality of X_N. Observe that

$$
\frac{1}{N} \sum_{k \ge 0} a_{Nk} e^{ikt} = \mathbf{E} e^{itX_N}
$$

is the characteristic function of X_N.

Proposition 4.2. *Suppose that $\sigma^2 \neq 0$ and set $\mu_N = \mathbf{E}X_N$ and $\sigma_N^2 = \mathbf{V}X_N$. Then for every $\varepsilon > 0$ we have uniformly for $|t| \le (\log N)^{1/2-\varepsilon}$*

$$
e^{-it\mu_N/\sigma_N} \frac{1}{N} \sum_{k \ge 0} a_{Nk} e^{ikt/\sigma_N} = e^{-t^2/2} + \mathcal{O}((\log N)^{-1/2+\varepsilon}). \tag{4.2}
$$

Proof. Set $f(z) = \log \alpha(e^z)$ in an open neighbourhood of $z = 0$. Then we have

$$
\alpha(e^{it}) = \alpha e^{i\mu t - \sigma^2 t^2/2 + \mathcal{O}(t^3)},
$$

with $\mu = f'(0) = \alpha'(1)/\alpha$ and $\sigma^2 = f''(0) = \alpha''(1)/\alpha + \mu - \mu^2$ (see Theorem 2.2). Hence, by using Proposition 4.1

$$
b_j(e^{it}) = G_j e^{j(i\mu t - \sigma^2 t^2/2)} e^{\mathcal{O}(t + jt^3 + (1-\delta)^j)} + \mathcal{O}\left(\alpha^{(1-\delta)j}\right)
$$

in an open neighbourhood of $t=0$ in $\mathbf{R}$. Now suppose that $N=\sum\limits_{l=0}^{L-1} e_l G_{j_l}$ with $j_0>j_1>\cdots>j_{L-1}$ and $e_l>0$ is the G-ary expansion of N. Then by Lemma 3.6 and the trivial estimate $\sum\limits_{s=0}^{e_l-1} e^{its}=e_l(1+\mathcal{O}(t))$ we have

$$\begin{aligned}\sum_{k\ge 0} a_{Nk}e^{ikt} &= \sum_{l=0}^{L-1} b_{j_l}(e^{it})e^{it\sum_{h=0}^{l-1}e_h}\sum_{s=0}^{e_l-1}e^{its}\\ &= \sum_{l=0}^{L-1} e_l G_{j_l} e^{ij_l\mu t-j_l\sigma^2t^2/2+it\sum_{h=0}^{l-1}e_h} e^{\mathcal{O}(t+j_lt^3+(1-\delta)^{j_l})}+\mathcal{O}(N^{(1-\delta)}).\end{aligned}$$

Now observe that

$$\frac{it}{\sigma_N}=\frac{it}{\sigma\, j_0^{1/2}}\left(1+\mathcal{O}(j_0^{-1})\right),$$

and that

$$e^{-it\mu_N/\sigma_N}=e^{-it(\mu/\sigma)j_0^{1/2}(1+\mathcal{O}(j_0^{-1}))}.$$

Hence

$$\begin{aligned}\mathbf{E}e^{it(X_N-\mu_N)/\sigma_N} &= e^{-it\mu_N/\sigma_N}\frac{1}{N}\sum_{k\ge 0}a_{Nk}e^{ikt/\sigma_N}\\ &= e^{-t^2/2}\sum_{l=0}^{L-1}\frac{e_lG_{j_l}}{N}e^{it/(\sigma j_0^{1/2})\left(\sum_{h=0}^{l-1}e_h-\mu(j_0-j_l)\right)-t^2(j_l-j_0)/(2j_0)}\\ &\qquad\times e^{\mathcal{O}\left(j_0^{-1/2}(|t|+|t|^3)+(1-\delta)^{j_l}\right)}+\mathcal{O}(N^{-\delta}).\end{aligned}$$

Let $\varepsilon>0$ be a (small) real number and let κ be defined by $j_{\kappa-1}>j_0-j_0^{\varepsilon}\ge j_\kappa$. Then $\kappa\le j_0-j_\kappa<j_0^{\varepsilon}$ and consequently

$$\begin{aligned}\mathbf{E}e^{it(X_N-\mu_N)/\sigma_N} &= e^{-t^2/2}\sum_{l=0}^{\kappa-1}\frac{e_lG_{j_l}}{N}e^{\mathcal{O}\left(|t|\sigma j_0^{\varepsilon-1/2}+t^2j_0^{\varepsilon-1}+j_0^{-1/2}(|t|+|t|^3)+(1-\delta)^{j_0-j_0^{\varepsilon}}\right)}\\ &\quad+\mathcal{O}\left(\sum_{l=\kappa}^{L-1}\frac{G_{j_l}}{N}\right)+\mathcal{O}(N^{-\delta})\\ &= e^{-t^2/2}e^{\mathcal{O}\left(|t|\sigma j_0^{\varepsilon-1/2}+t^2j_0^{\varepsilon-1}+|t|^3j_0^{-1/2}+(1-\delta)^{j_0-j_0^{\varepsilon}}\right)}+\mathcal{O}(\alpha^{-j_0^{\varepsilon}}).\end{aligned}$$

Since $j_0=(\log N)/(\log\alpha)+\mathcal{O}(1)$ this implies (4.2) directly for $|t|\le(\log N)^{\varepsilon/3}$. Furthermore, since

$$e^{-t^2/2+\mathcal{O}(|t^3|j_0^{-1/2})}\le e^{-cj_0^{2\varepsilon/3}}=\mathcal{O}(j_0^{-1})$$

for $(\log N)^{\varepsilon/3}\le|t|\le(\log N)^{1/2-\varepsilon}$ and a sufficiently small $c>0$ we finally obtain the full version of (4.2). □

We can now use Proposition 4.2 to prove the first part of Theorem 2.2.

Proof. Set

$$\Delta_N(t)=e^{-t^2/2}-\mathbf{E}e^{it(X_N-\mu_N)/\sigma_N}.$$

Then by Esseen's [9, p. 32] inequality we have

$$\frac{1}{N}|\{n<N : s_G(n)<\mathbf{E}X_N+x\mathbf{V}X_N\}|=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac12t^2}dt+\mathcal{O}\left(\frac{1}{T}+\int_{-T}^{T}\left|\frac{\Delta_N(t)}{t}\right|dt\right).$$

Choosing $T=(\log N)^{1/2-\varepsilon}$ we directly obtain from Proposition 4.2 and by applying the estimate

$$e^{-it\mu_N/\sigma_N}\frac{1}{N}\sum_{k\ge 0}a_{Nk}e^{ikt/\sigma_N}=1+\mathcal{O}(t^2)$$

for $|t| \le (\log N)^{-2}$ that

$$\int_{-T}^{T} \left| \frac{\Delta_N(t)}{t} \right| dt = \mathcal{O}\left((\log N)^{-1/2+\varepsilon} (\log\log N) \right).$$

Hence, (2.2) follows. □

5. Local Limit Law

In order to prove a local limit law for X_N, i.e. the second part or Theorem 2.2, we need more precise information on the behaviour of $b_j(z)$.

Proposition 5.1. *Suppose that $G = (G_j)_{j\ge 0}$ satisfies a finite or infinite recurrence of the above types. Then there exist $\eta > 0$ and $\delta > 0$ such that*

$$b_j(e^{it}) = C(e^{it})\alpha(e^{it})^j + \mathcal{O}\left(\alpha^{(1-\delta)j}\right) \tag{5.1}$$

uniformly for $|t| \le \eta$, where $C(z)$ and $\alpha(z)$ are as in Proposition 4.1, and

$$b_j(e^{it}) = \mathcal{O}\left(\alpha^{(1-\delta)j}\right) \tag{5.2}$$

uniformly for $\eta \le |t| \le \pi$.

Proof. Obviously, (5.1) follows from (4.1) for some $\eta > 0$.

For the proof of (5.2) we just have to observe that $|G(z,u)| < G(|z|,|u|) \le 1$ if $|z| \le 1$, $z \ne 1$, and $|u| \le 1/\alpha$. Hence, by continuity there exist $\varepsilon > 0$ and $\tau > 0$ such that $|1 - G(u, e^{it})| \ge \tau$ uniformly for (real) t with $\eta \le |t| \le \pi$ and (complex) u with $|u| \le 1 + \varepsilon$. Thus, $B(e^{it}, u)$ is analytic (and therefore bounded) in this range and we obtain

$$\begin{aligned} b_j(e^{it}) &= \frac{1}{2\pi i} \int\limits_{|u|=1/\alpha+\varepsilon} B(e^{it}, u) \frac{du}{u^{j+1}} \\ &= \mathcal{O}\left(\alpha^{(1-\delta)j}\right), \end{aligned}$$

with some $\delta > 0$. □

With help of Proposition 5.1 it is possible to derive asymptotic expansions for $b_{j,k}$ via saddle point approximations.

Proposition 5.2. *We have*

$$b_{j,k} = \frac{G_j}{\sqrt{2\pi j\sigma^2}} \left(\exp\left(-\frac{(k - j\mu)^2}{2j\sigma^2} \right) + \mathcal{O}(j^{-1/2}) \right)$$

uniformly for all $j, k \ge 0$.

Proof. We again use Cauchy's formula

$$b_{j,k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} b_j(e^{it}) e^{-ikt}\, dt.$$

Since

$$\int_{\eta \le |t| \le \pi} |b_j(e^{it})|\, dt = \mathcal{O}\left(\alpha^{(1-\delta)j}\right) = \mathcal{O}(G_j/j)$$

we just have to evaluate

$$I = \frac{1}{2\pi} \int_{|t| \le j^{-\varepsilon}} b_j(e^{it}) e^{-ikt}\, dt + \frac{1}{2\pi} \int_{j^{-\varepsilon} \le |t| \le \eta} b_j(e^{it}) e^{-ikt}\, dt = I_1 + I_2,$$

where $0 < \varepsilon < \frac{1}{6}$. From $\alpha(e^{it}) = \alpha e^{i\mu t - \sigma^2 t^2/2 + \mathcal{O}(t^3)}$ it follows that there exists a constant $c > 0$ such that $|\alpha(e^{it})| \le e^{-ct^2}$ for $|t| \le \eta$. Hence,

$$I_2 \le \frac{1}{\pi} \int_{j^{-\varepsilon}}^{\infty} e^{-cjt^2}\, dt + \mathcal{O}\left(\alpha^{(1-\delta)j}\right) = \mathcal{O}\left(e^{-cj^{1-2\varepsilon}}\right) + \mathcal{O}\left(\alpha^{(1-\delta)j}\right) = \mathcal{O}(G_j/j)$$

Finally,

$$
\begin{aligned}
I_1 &= \frac{1}{2\pi}\int_{|t|\le j^{-\varepsilon}} C\alpha^j e^{it(j\mu-k)-j\sigma^2t^2/2}\left(1+\mathcal{O}(j|t^3|+|t|)\right)\,dt + \mathcal{O}\left(\alpha^{(1-\delta)j}\right)\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} C\alpha^j e^{it(j\mu-k)-j\sigma^2t^2/2}\,dt + \mathcal{O}\left(\int_{|t|>j^{-\varepsilon}} C\alpha^j e^{-j\sigma^2t^2/2}\,dt\right)\\
&\quad + \mathcal{O}\left(\int_{|t|\le j^{-\varepsilon}} C\alpha^j e^{-j\sigma^2t^2/2}(j|t^3|+|t|)\,dt\right) + \mathcal{O}\left(\alpha^{(1-\delta)j}\right)\\
&= \frac{C\alpha^j}{\sqrt{2\pi j\sigma^2}}\exp\left(-\frac{(k-j\mu)^2}{2j\sigma^2}\right) + \mathcal{O}(\alpha^j/j)\\
&= \frac{G_j}{\sqrt{2\pi j\sigma^2}}\exp\left(-\frac{(k-j\mu)^2}{2j\sigma^2}\right) + \mathcal{O}(G_j/j).
\end{aligned}
$$

This completes the proof of Proposition 5.2. □

Finally Proposition 5.2 and Lemma 3.6 can be used to complete the proof of Theorem 2.2.

Proof. As in the proof of Proposition 4.2 we suppose that $N = \sum_{l=0}^{L-1} e_l G_{j_l}$ (with $j_0 > j_1 > \cdots > j_{L-1}$ and $e_l > 0$) is the G-ary expansion of N. Furthermore, let $\varepsilon > 0$ be a (small) real number and let κ be defined by $j_{\kappa-1} > j_0 - j_0^{\varepsilon} \ge j_\kappa$. Then by (3.5)

$$
\begin{aligned}
a_{Nk} &= \sum_{l=0}^{L-1}\sum_{i=0}^{e_l-1} b_{j_l,k-\sum_{h=0}^{l-1} e_h - i}\\
&= \sum_{l=0}^{\kappa-1}\sum_{i=0}^{e_l-1} b_{j_l,k-\sum_{h=0}^{l-1} e_h - i} + \mathcal{O}\left(\sum_{l=\kappa}^{L-1}\frac{G_{j_l}}{j_l^{1/2}}\right)\\
&= \sum_{l=0}^{\kappa-1}\sum_{i=0}^{e_l-1} \frac{G_{j_l}}{\sqrt{2\pi j_l\sigma^2}}\exp\left(-\frac{\left(k-\sum_{h=0}^{l-1} e_h - ij_l\mu\right)^2}{2j_l\sigma^2}\right) + \mathcal{O}(G_{j_0}/j_0).
\end{aligned}
$$

If $l < \kappa$ and $|k-\mu_N| = \mathcal{O}(j_0^{1/2}\log j_0)$ then

$$
\begin{aligned}
&\frac{(k-\mu_N)^2}{2\sigma_N^2} - \frac{\left(k-\sum_{h=0}^{l-1} e_h - ij_l\mu\right)^2}{2j_l\sigma^2}\\
&\qquad = \frac{(k-\mu_N)^2 - \left(k-\sum_{h=0}^{l-1} e_h - ij_l\mu\right)^2}{2\sigma_N^2} + \left(k-\sum_{h=0}^{l-1} e_h - ij_l\mu\right)^2\left(\frac{1}{2\sigma_N^2}-\frac{1}{2j_l\sigma^2}\right)\\
&\qquad = \mathcal{O}\left(j_0^{\varepsilon-1/2}\log j_0\right) + \mathcal{O}\left(j_0^{-1}(\log j_0)^2\right),
\end{aligned}
$$

where we have used $\mu_N = j_0\mu + \mathcal{O}(1)$ and $\sigma_N^2 = j_0\sigma^2 + \mathcal{O}(1)$. Hence, from

$$
\sqrt{\frac{j_0+\mathcal{O}(1)}{j_l}}\exp\left(\frac{(k-\mu_N)^2}{2\sigma_N^2} - \frac{\left(k-\sum_{h=0}^{l-1} e_h - ij_l\mu\right)^2}{2j_l\sigma^2}\right) = 1 + \mathcal{O}\left(j_0^{\varepsilon-1/2}\log j_0\right)
$$

we obtain

$$\begin{aligned}a_{Nk} &= \frac{N}{\sqrt{2\pi\sigma_N^2}} \exp\left(-\frac{(k-\mu_N)^2}{2\sigma_N^2}\right) \sum_{l=0}^{k-1} \frac{e_l G_{j_l}}{N} \left(1 + \mathcal{O}\left(j_0^{\varepsilon-1/2} \log j_0\right)\right) + \mathcal{O}(G_{j_0}/j_0) \\ &= \frac{N}{\sqrt{2\pi\sigma_N^2}} \left(\exp\left(-\frac{(k-\mu_N)^2}{2\sigma_N^2}\right) + \mathcal{O}\left(j_0^{\varepsilon-1/2} \log j_0\right)\right).\end{aligned}$$

If $|k-\mu_N| \geq j_0^{1/2} \log j_0$ then we have for $l < \kappa$

$$\begin{aligned}b_{j_l, k-\sum_{h=0}^{l-1} e_h - i} &= \mathcal{O}\left(\alpha^{j_l} j_0^{-1/2} \exp\left(-\frac{(\log j_0)^2}{4\sigma^2}\right)\right) \\ &= \mathcal{O}\left(\alpha^{j_l} j_0^{-1}\right)\end{aligned}$$

which finally gives

$$\begin{aligned}a_{Nk} &= \mathcal{O}((\alpha^{j_0} j_0^{-1}) + \mathcal{O}\left(\sum_{l=\kappa}^{L-1} \frac{G_{j_l}}{j_l^{1/2}}\right) \\ &= \mathcal{O}(G_{j_0}/j_0).\end{aligned}$$

This completes the proof of Theorem 2.2. □

References

[1] N. L. Bassily and I. Kátai, Distribution of the values of q-additive functions on polynomial sequences, *Acta Math. Hung.* **68** (1995), 353–361.

[2] R. Bellman and H. N. Shapiro, On a problem in additive number theory, *Ann. Math.* **49** (1948), 333–340.

[3] L. E. Bush, An asymptotic formula for the average sum of the digits of integers, *Am. Math. Monthly* **47** (1940), 154–156.

[4] J. Coquet, Power sums of digital sums, *J. Number Th.* **22** (1986), 161–176.

[5] H. Delange Sur la fonction sommatoire de la fonction "Somme de Chiffres", *L'Enseignement math.* **21** (1975), 31–77.

[6] M. Drmota and M. Skałba, The parity of the Zeckendorf sum-of-digits-function, preprint.

[7] J. M. Dumont and A. Thomas, Digital sum moments and substitutions, *Acta Arith.* **64** (1993), 205–225.

[8] J. M. Dumont and A. Thomas, Gaussian asymptotic properties of the sum-of-digits functions, *J. Number Th.* **62** (1997), 19–38.

[9] C.-G. Esseen, Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law, *Acta Math.* **77** (1945), 1–125.

[10] J. Gajdosik, Kombinatorische Faktorisierungen und Ziffernentwicklungen, *thesis*, TU Wien, 1996.

[11] P. J. Grabner, P. Kirschenhofer, H. Prodinger, and R. F. Tichy, On the moments of the sum-of-digits function, in: *Applications of Fibonacci Numbers* **5** (1993), 263–271

[12] P. J. Grabner and R. F. Tichy, Contributions to digit expansions with respect to linear recurrences, *J. Number Th.* **36** (1990), 160–169.

[13] P. Grabner and R. F. Tichy, α-Expansions, linear recurrences, and the sum-of-digits function, *manuscripta math.* **70** (1991), 311–324.

[14] R. E. Kennedy and C. N. Cooper, An extension of a theorem by Cheo and Yien concerning digital sums, *Fibonacci Q.* **29** (1991), 145–149.

[15] P. Kirschenhofer, On the variance of the sum of digits function, *Lecture Notes Math.* **1452** (1990), 112–116.

[16] W. Parry, On the β-expansion of real numbers, *Acta Math. Acad. Sci. Hung.*, **12** (1961), 401–416.

[17] A. Pethö and R. F. Tichy, On digit expansions with respect to linear recurrences, *J. Number Th.* **33** (1989), 243–256.

[18] J. Schmid, The joint distribution of the binary digits of integer multiples, *Acta Arith.* **43** (1984), 391–415.

[19] W. M. Schmidt, The joint distribution the digits of certain integer s-tuples, *Studies in pure mathematics, Mem. of P. Turan* (1983), 605–622.

[20] H. Trollope, An explicit expression for binary digital sums, *Meth. Mag.* **41** (1968), 21–25.