JOURNAL DE THÉORIE DES NOMBRES DE BORDEAUX

MICHAEL DRMOTA
WOLFGANG STEINER

The Zeckendorf expansion of polynomial sequences

Journal de Théorie des Nombres de Bordeaux, tome 14, n° 2 (2002), p. 439-475
<http://www.numdam.org/item?id=JTNB_2002__14_2_439_0>

© Université Bordeaux 1, 2002, tous droits réservés.

L'accès aux archives de la revue « Journal de Théorie des Nombres de Bordeaux » (http://jtnb.cedram.org/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
Article numérisé dans le cadre du programme Numérisation de documents anciens mathématiques
http://www.numdam.org/

Journal de Théorie des Nombres
de Bordeaux **14** (2002), 439–475

The Zeckendorf expansion of polynomial sequences

par Michael DRMOTA et Wolfgang STEINER

dédié à Michel Mendès France à l'occasion de son 65ème anniversaire

RÉSUMÉ. Nous montrons que la fonction 'somme de chiffres' de Zeckendorf $s_Z(n)$ lorsque n parcourt l'ensemble des nombres premiers ou bien une suite polynomiale d'entiers satisfait un théorème central limite. Nous obtenons aussi des résultats analogues pour d'autres fonctions du même type. Nous montrons également que le développement de Zeckendorf et le développement standard en base q des entiers sont asymptotiquement indépendants.

ABSTRACT. In the first part of the paper we prove that the Zeckendorf sum-of-digits function $s_Z(n)$ and similarly defined functions evaluated on polynomial sequences of positive integers or primes satisfy a central limit theorem. We also prove that the Zeckendorf expansion and the q-ary expansions of integers are asymptotically independent.

1. Introduction

Let $q \geq 2$ be an integer. Then a real-valued function f defined on the non-negative integers is called q-additive if f satisfies

$$f(0) = 0 \quad \text{and} \quad f(n) = \sum_{k\geq 0} f(\epsilon_{q,k}(n)q^k),$$

where $\epsilon_{q,k}(n) \in \{0, 1, \ldots, q-1\}$ are the digits in the q-ary expansion

$$n = \sum_{k\geq 0} \epsilon_{q,k}(n)q^k$$

of the integer $n \geq 0$. For example, the sum-of-digits function

$$s_q(n) = \sum_{j\geq 0} \epsilon_{q,k}(n)$$

is a q-additive function. The distribution behaviour of q-additive functions has been discussed by several authors (starting most probably with M.

Manuscrit reçu le 20 avril 2001.
This research was supported by the Austrian Science Foundation FWF, grant S8302-MAT.

Mendès France [18] and H. Delange [3], see also Coquet [2], Dumont and Thomas [10, 11], Manstavičius [16], and [6] for a list of further references). Most papers deal with the average value or the distribution of q-additive function. There are, however, also laws of the iterated logarithm and more generally a Strassen law for the sum of digits function due to Manstavičius [17]. (It seems to be difficult to generalize such a law to the Zeckendorf sum-of-digits function since a corresponding *Fundamental Lemma* seems to be out of reach at the moment, even the generalization to a joint law of two q-ary sum-of-digits function is not obvious, see [8].)

The most general central limit theorem for q-additive functions f is due to Manstavičius [16], where the distribution of the values $f(n)$ $(0 \le n < N)$ is considered. In this paper we are interested in the distribution of $f(P(n))$ $(0 \le n < N)$, where $P(x)$ is an integer polynomial. Here the best known result is due to Bassily and Kátai [1].[1] (Here and in the sequel $\Phi(x)$ denotes the distribution function of the standard normal law.)

Theorem 1. *Let f be a q-additive function such that $f(bq^k) = \mathcal{O}(1)$ as $k \to \infty$ and $b \in \{0, \dots, q-1\}$. Assume that $\frac{D_q(N)}{(\log N)^\eta} \to \infty$ as $N \to \infty$ for some $\eta > 0$ and let $P(n)$ be a polynomial with integer coefficients, degree r and positive leading term. Then, as $N \to \infty$,*

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \frac{f(P(n)) - M_q(N^r)}{D_q(N^r)} < x\right\} \to \Phi(x)$$

and

$$\frac{1}{\pi(N)}\#\left\{p \in \mathbb{P}, p < N \,\middle|\, \frac{f(P(p)) - M_q(N^r)}{D_q(N^r)} < x\right\} \to \Phi(x),$$

where

$$M_q(N) := \sum_{k=0}^{[\log_q N]} \mu_{k,q}, \qquad D_q(N)^2 = \sum_{k=0}^{[\log_q N]} \sigma_{k,q}^2$$

and

$$\mu_{k,q} := \frac{1}{q}\sum_{b=0}^{q-1} f(bq^k), \qquad \sigma_{k,q}^2 := \frac{1}{q}\sum_{b=0}^{q-1} f^2(bq^k) - \mu_{k,q}^2.$$

This result relies on the fact that suitably modified centralized moments converge.

The main purpose of this paper is to extend this result to certain G-ary digital expansions. Let $a \ge 1$ be an integer and the sequence $G = (G_k)_{k \ge 0}$ be defined by the linear recurrence

$$G_k = aG_{k-1} + G_{k-2}, \qquad G_0 = 1,\ G_1 = a+1.$$

[1] This theorem was only stated (and proved) for $\eta = \frac{1}{3}$. However, a short inspection of the proof shows that $\eta > 0$ is sufficient.

Now every integer $n \geq 0$ has a unique *digital expansion*

$$n = \sum_{k \geq 0} \epsilon_{G,k}(n) G_k$$

with integer digits $0 \leq \epsilon_{G,k}(n) \leq a$ provided that

$$\sum_{k=0}^{j} \epsilon_{G,k}(n) G_k < G_{j+1}$$

for all $j \geq 0$ (which means that $\epsilon_{G,k-1}(n) = 0$ if $\epsilon_{G,k}(n) = a$). A special case of these expansions is the Zeckendorf expansion where $a = 1$ and the G_k are the Fibonacci numbers.

A function f is said to be G-additive, if

$$f(0) = 0 \quad \text{and} \quad f(n) = \sum_{k \geq 0} f(\epsilon_{G,k}(n) G_k).$$

Alternatively we have

$$f(n) = \sum_{k \geq 0} f_k(\epsilon_{G,k}(n)),$$

where $f_k(b) := f(bG_k)$.

First we will prove the following theorem concerning the distribution of the sequence $f(n)$, $0 \leq n < N$. The proof essentially relies on the fact that the possible G-ary digital expansions can be represented by a Markov chain. Note that the sequence G_k is also given by

$$G_k = \frac{\alpha(\alpha+1)}{\alpha^2+1}\alpha^k - \frac{\alpha-1}{\alpha^2+1}\left(-\frac{1}{\alpha}\right)^k, \tag{1.1}$$

where α is the positive root of the characteristic polynomial of the linear recurrence

$$\chi(x) = x^2 - ax - 1.$$

Theorem 2. *Let G be as above, f a G-additive function such that $f_k(b) = \mathcal{O}(1)$ as $k \to \infty$ for $b \in \{0, \dots, a\}$. Then, for all $\eta > 0$, the expected value of $f(n)$, $0 \leq n < N$, is given by*

$$E_N := \frac{1}{N} \sum_{n<N} f(n) = M(N) + \mathcal{O}\left((\log N)^\eta\right), \tag{1.2}$$

where

$$M(N) = M_G(N) = \sum_{k=0}^{[\log_\alpha N]} \mu_k \quad \textit{with} \quad \mu_k = \frac{\alpha}{\alpha^2+1} \sum_{b=1}^{a-1} f_k(b) + \frac{1}{\alpha^2+1} f_k(a).$$

Furthermore, set

$$D(N)^2 = D_G(N)^2 = \sum_{j,k=0}^{[\log_\alpha N]} \sigma_{j,k}^{(2)}$$

with

$$\sigma_{j,k}^{(2)} = \begin{cases} \frac{\alpha}{\alpha^2+1} \sum_{b=1}^{a-1} f_k(b)^2 + \frac{1}{\alpha^2+1} f_k(a)^2 - \mu_k^2 & \text{if } j = k \\ \left(-\frac{1}{\alpha^2}\right)^{|j-k|} \mu_{\min(j,k)} \overline{\mu}_{\max(j,k)} & \text{if } j \neq k, \end{cases}$$

where

$$\overline{\mu}_k = -\frac{\alpha}{\alpha^2+1} \sum_{b=1}^{a-1} f_k(b) + \frac{\alpha^2}{\alpha^2+1} f_k(a).$$

Assume further that there exists a constant $c > 0$ such that $\sigma_{k,k}^{(2)} \geq c$ for all $k \geq 0$. Then, as $N \to \infty$,

$$\frac{1}{N} \# \left\{ n < N \,\middle|\, \frac{f(n) - M(N)}{D(N)} < x \right\} \to \Phi(x) \tag{1.3}$$

and

$$\frac{1}{N} \sum_{n<N} \left(\frac{f(n) - M(N)}{D(N)} \right)^h \to \int_{-\infty}^{\infty} x^h \, d\Phi(x) \tag{1.4}$$

for all positive integers h.

(1.3) has been shown by Drmota [5] for strongly G-additive functions f, i.e.

$$f(n) = \sum_{k \geq 0} f(\epsilon_k(n)).$$

Furthermore, it should be noted that (1.4) provides an asymptotic relation for the variance, too, however, without an error term:

$$V_N := \frac{1}{N} \sum_{n<N} (f(n) - E_N)^2 \sim D(N)^2. \tag{1.5}$$

We will use Theorem 2 and a method similar to Bassily and Kátai's to prove Theorem 3.

Theorem 3. *Let G, f be as in Theorem 2 and $P(n)$ a polynomial with integer coefficients, degree r and positive leading term. Then, as $N \to \infty$,*

$$\frac{1}{N} \# \left\{ n < N \,\middle|\, \frac{f(P(n)) - M(N^r)}{D(N^r)} < x \right\} \to \Phi(x), \tag{1.6}$$

$$\frac{1}{\pi(N)} \# \left\{ p < N \,\middle|\, \frac{f(P(p)) - M(N^r)}{D(N^r)} < x \right\} \to \Phi(x), \tag{1.7}$$

and

$$\frac{1}{N}\sum_{n<N}\left(\frac{f(P(n))-M(N^r)}{D(N^r)}\right)^h \to \int_{-\infty}^{\infty} x^h\,d\Phi(x),$$

$$\frac{1}{\pi(N)}\sum_{p<N}\left(\frac{f(P(p))-M(N^r)}{D(N^r)}\right)^h \to \int_{-\infty}^{\infty} x^h\,d\Phi(x)$$

for all positive integers h, if we set $f(P(n)) = -f(-P(n))$ for $P(n) < 0$.

Note that definition of $f(P(n))$ for $P(n) < 0$ has no influence on the result, because the number of non-negative integers n with $P(n) < 0$ is negligible.

Our next results concern the indepence of different digital expansions. For example, in [6] the following property is shown. Suppose that q_1, q_2 are two coprime integers and f_1, f_2 q_1- resp. q_2-additive functions satisfying the assumptions of Theorem 1. Then we have, as $N \to \infty$,

$$\frac{1}{\pi(N)}\#\left\{n < N \,\middle|\, \frac{f_i(n) - M_{q_i}(N)}{D_{q_i}(N)} < x_i \ (i = 1,2)\right\} \to \Phi(x_1)\Phi(x_2),$$

i.e. the distribution of the pairs $(f_1(n), f_2(n))$, $0 \le n < N$, can be considered as independent.

We will extend this property to our more general situation.

Theorem 4. *Suppose that f_1, f_2 are two functions satisfying one of the following conditions.*

(i) *$q_1, q_2 \ge 2$ are two positive coprime integers and f_1, f_2 q_1- resp. q_2-additive functions satisfying the assumptions of Theorem 1. Furthermore set $M_i(N) := M_{q_i}(N)$ and $D_i(N) := D_{q_i}(N)$ $(i = 1,2)$.*

(ii) *$q \ge 2$ is an integer and $f_1(n)$ a q-additive function satisfying the assumptions of Theorem 1. $a \ge 1$ is an integer and $f_2(n)$ is a G-additive function satisfying the assumptions of Theorem 2. Furthermore set $M_1(N) := M_q(N)$, $D_1(N) := D_q(N)$ and $M_2(N) := M_G(N)$, $D_2(N) := D_G(N)$.*

(iii) *$a_1, a_2 \ge 1$ are two different integers such that $\sqrt{\frac{a_1^2+4}{a_2^2+4}}$ is irrational, $G = (G_j)_{j\ge 0}$ and $H = (H_j)_{j\ge 0}$ the corresponding linear recurrent sequences, and f_1, f_2 G- resp. H-additive functions satisfying the assumptions of Theorem 2. Furthermore set $M_1(N) := M_G(N)$, $D_1(N) := D_G(N)$ and $M_2(N) := M_H(N)$, $D_2(N) := D_H(N)$.*

Let $P_1(x), P_2(x)$ be two polynomials with integer coefficients, degrees r_1, r_2 and positive leading term. Then, as $N \to \infty$,

$$(1.8)\quad \frac{1}{N}\#\left\{n < N \,\middle|\, \frac{f_i(P_i(n)) - M_i(N^{r_i})}{D_i(N^{r_i})} < x_i \ (i = 1,2)\right\} \to \Phi(x_1)\Phi(x_2)$$

and

$$(1.9)\quad \frac{1}{\pi(N)}\#\left\{p < N \,\middle|\, \frac{f_i(P_i(p)) - M_i(N^{r_i})}{D_i(N^{r_i})} < x_i \ (i = 1,2)\right\} \to \Phi(x_1)\Phi(x_2).$$

The paper is organized in the following way. Section 2 is devoted to the proof of Theorem 2. Section 3 provides a plan of the proof of Theorem 3. Sections 4–6 collect some preliminaries which are needed for the proof of Theorem 3 in Section 7. Finally, the proof of Theorem 4 is presented in Section 8.

2. Proof of Theorem 2

Our aim is to study the distribution behaviour of $f(n)$, $0 \le n < N$, i.e. the random variable Y_N defined by

$$\mathbf{Pr}[Y_N \le x] := \frac{1}{N}\#\{n < N : f(n) \le x\}.$$

If we define $\zeta_{k,N}$ by

$$\mathbf{Pr}[\zeta_{k,N} \le x] := \frac{1}{N}\#\{n < N : f_k(\epsilon_k(n)) \le x\}$$

and $\xi_{k,N}$ by

$$\mathbf{Pr}[\xi_{k,N} = b] := \frac{1}{N}\#\{n < N : \epsilon_k(n) = b\} \qquad (b \in \{0, \dots, a\}),$$

then we obviously have

$$Y_N = \sum_{k \ge 0} \zeta_{k,N} = \sum_{k \ge 0} f_k(\xi_{k,N}).$$

i.e. Y_N is a (weighted) sum of $\xi_{k,N}$. Therefore, we will first have a detailed look at $\xi_{k,N}$. It turns out that ξ_{k,G_j} constitutes an almost stationary Markov chain, as the next lemma shows. We want to mention that this fact is also a consequence of results from Dumont and Thomas [10, 11]. In our case this is a quite simple observation. Therefore we decided to present a short proof of this fact, too. This procedure is simpler and shorter than introducing the notation of [10, 11] and to specialize afterwards.

Lemma 1. *For fixed j, the random variables $(\xi_{k,G_j})_{0 \le k \le j-1}$ form a Markov chain with*

$$(2.1)\qquad \mathbf{Pr}[\xi_{k,G_j} = 1] = \mathbf{Pr}[\xi_{k,G_j} = 2] = \cdots = \mathbf{Pr}[\xi_{k,G_j} = a-1],$$

$$(2.2)\qquad \mathbf{Pr}[\xi_{k+1,G_j} = 1 | \xi_{k,G_j} = b] = \cdots = \mathbf{Pr}[\xi_{k+1,G_j} = a-1 | \xi_{k,G_j} = b],$$

$$(2.3)\qquad \mathbf{Pr}[\xi_{k+1,G_j} = b | \xi_{k,G_j} = 1] = \cdots = \mathbf{Pr}[\xi_{k+1,G_j} = b | \xi_{k,G_j} = a-1],$$

(for all j, k, b) and

$$\begin{pmatrix} \mathbf{Pr}[\xi_{k+1,G_j} = 0] \\ \mathbf{Pr}[\xi_{k+1,G_j} = 1] \\ \mathbf{Pr}[\xi_{k+1,G_j} = a] \end{pmatrix} = P_{k,j} \begin{pmatrix} \mathbf{Pr}[\xi_{k,G_j} = 0] \\ \mathbf{Pr}[\xi_{k,G_j} = 1] \\ \mathbf{Pr}[\xi_{k,G_j} = a] \end{pmatrix}, \tag{2.4}$$

where

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

with initial states

$$\mathbf{Pr}[\xi_{0,G_j} = 0] = \frac{G_{j-1}}{G_j} = \frac{1}{\alpha} + \mathcal{O}\left(\frac{1}{\alpha^{2j}}\right)$$

and

$$\mathbf{Pr}[\xi_{0,G_j} = 1] = \mathbf{Pr}[\xi_{0,G_j} = 2] = \cdots = \mathbf{Pr}[\xi_{0,G_j} = a] = \frac{1}{\alpha+1} + \mathcal{O}\left(\frac{1}{\alpha^{2j}}\right).$$

Remark. The matrices $P_{k,j}$ are no transition matrices of a Markov process, but they describe transition matrices in view of the relations (2.1)–(2.3). However, it turned out to be easier to work with 3×3-matrices instead of $(a+1) \times (a+1)$-matrices.

Proof. A sequence $(\epsilon_i)_{i \geq 0}$ of non-negative integers is a G-ary digital expansion of an integer n, if and only if $\epsilon_i \leq a$ for all $i \geq 0$, $\epsilon_{i-1} = 0$ if $\epsilon_i = a$ and $\epsilon_i \neq 0$ only for a finite number of i (cf. e.g. Grabner and Tichy [13]). Let

$$\mathcal{B}_j = \{(\epsilon_0, \ldots, \epsilon_{j-1}) : \epsilon_i \leq a,\ \epsilon_{i-1} = 0 \text{ if } \epsilon_i = a\}$$

be the set of G-ary digital expansions for $n < G_j$. Then

$$\mathbf{Pr}[\xi_{k,G_j} = b] = \frac{1}{G_j} \#\{(\epsilon_0, \ldots, \epsilon_{j-1}) \in \mathcal{B}_j : \epsilon_k = b\}$$

and it can be easily seen that (2.1) holds. For $k = 0$, even $\mathbf{Pr}[\xi_{0,G_j} = a]$ is equal to $\mathbf{Pr}[\xi_{0,G_j} = 1]$.

We have

$$\#\{(\epsilon_0, \ldots, \epsilon_{j-1}) \in \mathcal{B}_j : \epsilon_0 = 0\} = \#\{(\epsilon_0, \ldots, \epsilon_{j-1}) \in \mathcal{B}_j : \epsilon_{j-1} = 0\},$$

because we can take a block $(0, \epsilon_1, \ldots, \epsilon_{j-1})$ of the set on the left side of the equation, shift it to the left, set $\epsilon_{j-1} = 0$ and get a one-one correspondence to the blocks on the right side. Therefore

$$\mathbf{Pr}[\xi_{0,G_j} = 0] = \frac{G_{j-1}}{G_j} = \frac{\frac{\alpha+1}{D}\alpha^{j-1} - \frac{-\alpha^{-1}+1}{D}(-\alpha)^{-j+1}}{\frac{\alpha+1}{D}\alpha^{j} - \frac{-\alpha^{-1}+1}{D}(-\alpha)^{-j}} = \frac{1}{\alpha} + \mathcal{O}\left(\frac{1}{\alpha^{2j}}\right).$$

Since the other probabilities $\mathbf{Pr}[\xi_{0,G_j} = b]$, $1 \le b \le a$, are equal, we have

$$\begin{aligned}\mathbf{Pr}[\xi_{0,G_j} = 1] = \cdots = \mathbf{Pr}[\xi_{0,G_j} = a] &= \frac{1}{a}(1 - \mathbf{Pr}[\xi_{0,G_j} = 0]) \\ &= \frac{1}{\alpha+1} + \mathcal{O}\left(\frac{1}{\alpha^{2j}}\right).\end{aligned}$$

Now we show that we have a Markov chain.

$$\begin{aligned}&\mathbf{Pr}[\xi_{k+1,G_j} = b_{k+1} | \xi_{k,G_j} = b_k, \ldots, \xi_{0,G_j} = b_0] \\ &= \frac{\mathbf{Pr}[\xi_{k+1,G_j} = b_{k+1}, \xi_{k,G_j} = b_k, \ldots, \xi_{0,G_j} = b_0]}{\mathbf{Pr}[\xi_{k,G_j} = b_k, \ldots, \xi_{0,G_j} = b_0]} \\ &= \frac{\#\{(\epsilon_0, \ldots, \epsilon_{j-1}) \in \mathcal{B}_j : (\epsilon_0, \ldots, \epsilon_{k+1}) = (b_0, \ldots, b_{k+1})\}}{\#\{(\epsilon_0, \ldots, \epsilon_{j-1}) \in \mathcal{B}_j : (\epsilon_0, \ldots, \epsilon_k) = (b_0, \ldots, b_k)\}} \\ &= \frac{\#\{(\epsilon_{k+1}, \ldots, \epsilon_{j-1}) \in \mathcal{B}_{j-k-1} : \epsilon_{k+1} = b_{k+1}\}}{\#\{(\epsilon_k, \ldots, \epsilon_{j-1}) \in \mathcal{B}_{j-k} : \epsilon_k = b_k\}} \\ &= \frac{\mathbf{Pr}[\xi_{0,G_{j-k-1}} = b_{k+1}]\, G_{j-k-1}}{\mathbf{Pr}[\xi_{0,G_{j-k}} = b_k]\, G_{j-k}},\end{aligned}$$

where the third equation is valid only if $(b_0, \ldots, b_{k+1}) \in \mathcal{B}_{k+2}$. Otherwise the probability is 0 (for $b_{k+1} = a$, $b_k \neq 0$, $(b_0, \ldots, b_k) \in \mathcal{B}_{k+1}$) or undefined (for $(b_0, \ldots, b_k) \notin \mathcal{B}_{k+1}$). If the probability is defined, we thus have

$$\begin{aligned}&\mathbf{Pr}[\xi_{k+1,G_j} = b_{k+1} | \xi_{k,G_j} = b_k, \ldots, \xi_{0,G_j} = b_0] \\ &\qquad = \mathbf{Pr}[\xi_{k+1,G_j} = b_{k+1} | \xi_{k,G_j} = b_k]\end{aligned}$$

with the probabilities

$$\begin{aligned}&\mathbf{Pr}[\xi_{k+1,G_j} = 0 | \xi_{k,G_j} = 0] = \frac{1}{\alpha} + \mathcal{O}\left(\frac{1}{\alpha^{2(j-k)}}\right), \\ &\mathbf{Pr}[\xi_{k+1,G_j} = 0 | \xi_{k,G_j} = c] = \frac{\alpha+1}{\alpha^2} + \mathcal{O}\left(\frac{1}{\alpha^{2(j-k)}}\right) \qquad (1 \le c \le a), \\ &\mathbf{Pr}[\xi_{k+1,G_j} = b | \xi_{k,G_j} = 0] = \frac{1}{\alpha+1} + \mathcal{O}\left(\frac{1}{\alpha^{2(j-k)}}\right) \qquad (1 \le b \le a), \\ &\mathbf{Pr}[\xi_{k+1,G_j} = b | \xi_{k,G_j} = c] = \frac{1}{\alpha} + \mathcal{O}\left(\frac{1}{\alpha^{2(j-k)}}\right) \qquad \left(\begin{array}{l} 1 \le b \le a-1, \\ 1 \le c \le a \end{array}\right), \\ &\mathbf{Pr}[\xi_{k+1,G_j} = a | \xi_{k,G_j} = c] = 0 \qquad (1 \le c \le a).\end{aligned}$$

Similarly to (2.1), (2.2) and (2.3) are easy to see. Hence

$$\begin{aligned}\mathbf{Pr}[\xi_{k+1,G_j} = b] &= \sum_{c=0}^{a} \mathbf{Pr}[\xi_{k+1,G_j} = b|\xi_{k,G_j} = c]\mathbf{Pr}[\xi_{k,G_j} = c] \\ &= \mathbf{Pr}[\xi_{k+1,G_j} = b|\xi_{k,G_j} = 0]\mathbf{Pr}[\xi_{k,G_j} = 0] \\ &\quad + \mathbf{Pr}[\xi_{k+1,G_j} = b|\xi_{k,G_j} = a]\mathbf{Pr}[\xi_{k,G_j} = a] \\ &\quad + (a-1)\mathbf{Pr}[\xi_{k+1,G_j} = b|\xi_{k,G_j} = 1]\mathbf{Pr}[\xi_{k,G_j} = 1]\end{aligned}$$

and the transition from ξ_{k,G_j} to ξ_{k+1,G_j} is entirely determined by (2.4). □

Corollary 1. *The probability distribution of ξ_{k,G_j} is given by*

$$\mathbf{Pr}[\xi_{k,G_j} = b] = p_b + \mathcal{O}\left(\frac{1}{\alpha^{2\min(k,j-k)}}\right)$$

with

$$p_b = \begin{cases} \frac{\alpha+1}{\alpha^2+1} & \text{if } b = 0 \\ \frac{\alpha}{\alpha^2+1} & \text{if } 1 \le b \le a-1 \\ \frac{1}{\alpha^2+1} & \text{if } b = a. \end{cases}$$

Proof. Let P be the matrix obtained by neglecting the $\mathcal{O}\left(\frac{1}{\alpha^{2(j-k)}}\right)$ terms in the matrix $P_{k,j}$. The eigenvalues of P are 1, 0 and $-\frac{1}{\alpha^2}$ and the eigenvector to the eigenvalue 1 with $p_0 + (a-1)p_1 + p_a = 1$ is $\left(\frac{\alpha+1}{\alpha^2+1}, \frac{\alpha}{\alpha^2+1}, \frac{1}{\alpha^2+1}\right)^t$. □

Lemma 1 suggests to approximate the digital distribution by a stationary Markov chain $(X_k, k \ge 0)$, with (stationary) probability distribution $\mathbf{Pr}[X_k = b] = p_b$, $0 \le b \le a$, and transition matrix P, i.e.

$$\begin{aligned}&\mathbf{Pr}[X_{k+1} = 0|X_k = 0] = \frac{1}{\alpha}, \\ &\mathbf{Pr}[X_{k+1} = 0|X_k = c] = \frac{\alpha+1}{\alpha^2} \qquad (1 \le c \le a), \\ &\mathbf{Pr}[X_{k+1} = b|X_k = 0] = \frac{1}{\alpha+1} \qquad (1 \le b \le a), \\ &\mathbf{Pr}[X_{k+1} = b|X_k = c] = \frac{1}{\alpha} \qquad (1 \le b \le a-1,\ 1 \le c \le a), \\ &\mathbf{Pr}[X_{k+1} = a|X_k = c] = 0 \qquad (1 \le c \le a).\end{aligned} \tag{2.5}$$

The next lemma shows how we can quantify this approximation for finite dimensional distributions.

Lemma 2. *For every $h \ge 1$ and integers $0 \le k_1 < k_2 < \cdots < k_h < j$ we have*

$$\mathbf{Pr}[\xi_{k_1,G_j} = b_1, \dots, \xi_{k_h,G_j} = b_h] = q_{k_1,\dots,k_h,b_1,\dots,b_h} + \mathcal{O}\left(\frac{1}{\alpha^{2\min(k_1,j-k_h)}}\right)$$

for all $b_1, \dots, b_h \in \{0, \dots, a\}$, *where*

$$q_{k_1,\dots,k_h,b_1,\dots,b_h} = \mathbf{Pr}[X_{k_1} = b_1, \dots, X_{k_h} = b_h].$$

Proof. For $0 \le k < l < j$ we have

$$P_{k,j}P_{k+1,j} \cdots P_{l-1,j} = P^{l-k} + \mathcal{O}\left(\alpha^{-2(j-l)}\right)$$

and consequently

$$\mathbf{Pr}[\xi_{l,G_j} = b_2 | \xi_{k,G_j} = b_1] = \mathbf{Pr}[X_l = b_2 | X_k = b_1] + \mathcal{O}\left(\alpha^{-2(j-l)}\right). \tag{2.6}$$

Since

$$\begin{aligned}&\mathbf{Pr}[\xi_{k_1,G_j} = b_1, \dots, \xi_{k_h,G_j} = b_h] \\ &= \mathbf{Pr}[\xi_{k_h,G_j} = b_h | \xi_{k_{h-1},G_j} = b_{h-1}]\mathbf{Pr}[\xi_{k_{h-1},G_j} = b_{h-1} | \xi_{k_{h-2},G_j} = b_{h-2}] \cdots \\ &\qquad \cdots \mathbf{Pr}[\xi_{k_2,G_j} = b_2 | \xi_{k_1,G_j} = b_1]\mathbf{Pr}[\xi_{k_1,G_j} = b_1],\end{aligned}$$

we just have to apply (2.6) and Corollary 1 and the lemma follows. □

The case of general N is very similar.

Lemma 3. *The probability distribution of* $\xi_{k,N}$ *for* $G_j \le N < G_{j+1}$ *with* $j > k$ *is given by*

$$\mathbf{Pr}[\xi_{k,N} = b] = \mathbf{Pr}[\xi_{k,G_j} = b] + \mathcal{O}\left(\frac{1}{\alpha^{j-k}}\right) \tag{2.7}$$

for all $b \in \{0, \dots, a\}$.

Furthermore, the joint distribution for $0 \le k_1 < k_2 < \cdots < k_h < j$ *is given by*

$$\mathbf{Pr}[\xi_{k_1,N} = b_1, \dots, \xi_{k_h,N} = b_h] = \mathbf{Pr}[\xi_{k_1,G_j} = b_1, \dots, \xi_{k_h,G_j} = b_h] + \mathcal{O}\left(\frac{1}{\alpha^{j-k_h}}\right)$$

for all $b_1, \dots, b_h \in \{0, \dots, a\}$.

Proof. For $N = \sum_{i=0}^{j} \epsilon_i G_i$, we have

$$\{n < N\} = \{n < \epsilon_j G_j\} \cup \left(\{n < \epsilon_{j-1}G_{j-1}\} + \epsilon_j G_j\right) \cup \dots \cup \left(\{n < \epsilon_0 G_0\} + \sum_{i=1}^{j} \epsilon_i G_i\right).$$

Therefore

$$\mathbf{Pr}[\xi_{k,N} = b] = \frac{1}{N}\Big(\#\{n < \epsilon_j G_j \mid \epsilon_k = b\} + \#\{n < \epsilon_{j-1}G_{j-1} \mid \epsilon_k = b\} + \cdots$$

$$+ \#\{n < \epsilon_{k+1}G_{k+1} \mid \epsilon_k = b\} + \begin{cases} \sum_{i=0}^{k-1} \epsilon_i G_i & \text{if } \epsilon_k = b \\ 0 & \text{otherwise} \end{cases} \Big)$$

$$= \frac{1}{N}\left(\epsilon_j G_j \mathbf{Pr}[\xi_{k,G_j} = b] + \cdots + \epsilon_{[\frac{k+j}{2}]} G_{[\frac{k+j}{2}]} \mathbf{Pr}[\xi_{k,G_{[\frac{k+j}{2}]}} = b]\right)$$

$$+ \mathcal{O}\left(\frac{1}{N} G_{[\frac{k+j}{2}]}\right)$$

$$= \mathbf{Pr}[\xi_{k,G_j} = b] + \mathcal{O}\left(\frac{1}{\alpha^{j-k}}\right),$$

where we have used

$$\mathbf{Pr}[\xi_{k,G_j} = b] = \mathbf{Pr}[\xi_{k,G_{j-l}} = b] + \mathcal{O}\left(\frac{1}{\alpha^{j-l-k}}\right) \quad \text{for } k \le j - l.$$

A similar reasoning can be done for the joint distribution, e.g. we have for $l < k < j$:

$$\text{(2.8)} \quad \mathbf{Pr}[\xi_{k,N} = b, \xi_{l,N} = c] = \frac{1}{N} \sum_{i=k+1}^{j} \epsilon_i G_i \mathbf{Pr}[\xi_{k,G_i} = b, \xi_{l,G_i} = c]$$

$$+ \frac{1}{N} \begin{cases} \sum_{i=l+1}^{k-1} \epsilon_i G_i \mathbf{Pr}[\xi_{l,G_i} = c] + \begin{Bmatrix} \sum_{i=0}^{l-1} \epsilon_i G_i & \text{if } \epsilon_l = c \\ 0 & \text{otherwise} \end{Bmatrix} & \text{if } \epsilon_k = b \\ 0 & \text{otherwise} \end{cases}$$

Thus, we can proceed in the same way. □

We now turn to the derivation of $E_N = \mathbf{E} Y_N$, i.e. to the proof of (1.2), the first part of Theorem 2. Since

$$Y_N = \sum_{k=0}^{j} \zeta_{k,N}$$

for $N < G_{j+1}$, the expected value of Y_N is given by

$$\mathbf{E}\, Y_N = \sum_{k=0}^{j} \mathbf{E}\, \zeta_{k,N} = \sum_{k=A}^{B} \mathbf{E}\, \zeta_{k,N} + \mathcal{O}\left((\log N)^{\eta}\right),$$

where

$$\text{(2.9)} \qquad A = [(\log N)^{\eta}] \ \text{ and } \ B = [\log_\alpha N] - [(\log N)^{\eta}]$$

and $\eta > 0$ is a sufficiently small number (to be chosen in the sequel). Furthermore, we have

$$\mathbf{E}\,\zeta_{k,N} = \sum_{b=0}^{a} \mathbf{Pr}[\xi_{k,N} = b] f_k(b) = \mu_k + \mathcal{O}\left(\frac{1}{\alpha^{2\min(k,j-k)}}\right),$$

which implies

$$\mathbf{E}\,Y_N = \frac{1}{N}\sum_{n<N} f(n) = M(N) + \mathcal{O}\left((\log N)^{\eta}\right).$$

It seems that the variance $\mathbf{Var}\,Y_N$ cannot be treated in a similar (easy) way. Therefore, we use some additional assumptions and present a proof of (1.4) together with the distributional result (1.3).

The above calculation indicates that we just have to concentrate on digits $\epsilon_k(n)$ with $A \le k \le B$ (defined in (2.9)). The reason is that we obtain uniform estimates for this range. The following lemma is a direct consequence of Lemmata 2 and 3. Note that it is not necessary to assume that $k_1, \ldots, k_h$ are ordered and that they are distinct.

Lemma 4. *For every $h \ge 1$ and for every $\lambda > 0$ we have*

$$\frac{1}{N}\#\{n < N \mid \epsilon_{k_1}(n) = b_1, \ldots, \epsilon_{k_h}(n) = b_h\} = q_{k_1,\ldots,k_h,b_1,\ldots,b_h} + \mathcal{O}\left(\frac{1}{(\log N)^{\lambda}}\right)$$

uniformly for all integers

$$A \le k_1, k_2, \ldots, k_h \le B$$

(where A, B are defined in (2.9) with an arbitrary $\eta > 0$) and $b_1, b_2, \ldots, b_h \in \{0, 1, \ldots, a\}$, where

$$q_{k_1,\ldots,k_h,b_1,\ldots,b_h} = \mathbf{Pr}[X_{k_1} = b_1, \ldots, X_{k_h} = b_h].$$

This observation causes that we have to truncate the given function $f(n)$ and have to consider

$$\overline{f}(n) = \sum_{k=A}^{B} f_k(\epsilon_k) = f(n) + \mathcal{O}\left((\log N)^{\eta}\right).$$

In order to finish the proof of Theorem 2 it is (luckily) enough to prove

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \frac{\overline{f}(n) - \overline{M}(N)}{\overline{D}(N)} < x\right\} \to \Phi(x), \tag{2.10}$$

where

$$\overline{M}(N) = \sum_{k=A}^{B} \mu_k, \qquad \overline{D}(N)^2 = \sum_{j,k=A}^{B} \sigma_{j,k}^{(2)}.$$

This is due to the following lemma and (2.11).

Lemma 5. *Suppose that* $D(N)/(\log N)^{\overline{\eta}} \to \infty$ *for some* $\overline{\eta} > \eta/2$. *Then we have*

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \frac{f(n) - M(N)}{D(N)} < x\right\} \to \Phi(x)$$

for all $x \in \mathbb{R}$ *if and only if*

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \frac{\overline{f}(n) - \overline{M}(N)}{\overline{D}(N)} < x\right\} \to \Phi(x)$$

for all $x \in \mathbb{R}$.

Furthermore, if for all $h \geq 0$

$$\frac{1}{N}\sum_{n<N}\left(\frac{\overline{f}(n) - \overline{M}(N)}{\overline{D}(N)}\right)^h \to \int_{-\infty}^{\infty} x^h \, d\Phi(x)$$

then we also have

$$\frac{1}{N}\sum_{n<N}\left(\frac{f(n) - M(N)}{D(N)}\right)^h \to \int_{-\infty}^{\infty} x^h \, d\Phi(x)$$

and conversely.

Proof. We consider the three (sequences of) random variables

$$X_N = \frac{f(\cdot) - M(N)}{D(N)}, \quad Y_N = \frac{f(\cdot) - M(N)}{\overline{D}(N)}, \quad Z_N = \frac{\overline{f}(\cdot) - \overline{M}(N)}{\overline{D}(N)}.$$

Suppose first that the limiting distribution of X_N is Gaussian and that all moments converge. Since

$$\lim_{N\to\infty} \frac{D(N)}{\overline{D}(N)} = 1$$

and $Y_N = X_N \frac{D(N)}{\overline{D}(N)}$ the same is true for Y_N.

Further, we know that

$$\lim_{N\to\infty} \|Y_N - Z_N\|_\infty = 0.$$

Thus, it immediately follows that the limiting distribution of Z_N is the same as that of Y_N and that all moments of Z_N converge to the same limits as the moments of Y_N.

It is also clear that the converse implications are valid. This completes the proof of Lemma 5. □

Therefore it is sufficient to show that the moments

$$A_h(N) = \frac{1}{N}\sum_{n<N}\left(\frac{\overline{f}(n) - \overline{M}(N)}{\overline{D}(N)}\right)^h$$

converge to the corresponding moments of the normal law. We will do this in two steps. First we prove a central limit theorem (with convergence of moments) for the exact Markov process and then we compare these moments to those of $\overline{f}(n)$, i.e. (1.4). Obviously the proof (1.3) of Theorem 2 is completed then.

The next lemma provides a central limit theorem for $\sum f_k(X_k)$, where X_k is the stationary Markov process defined by (2.5).

Lemma 6. *Suppose that there exists a constant $c > 0$ such that $\sigma_{j,j}^{(2)} \geq c$ for all $j \geq 0$. Then we have*

$$D(N)^2 \gg \log N, \qquad \overline{D}(N)^2 \gg \log N \tag{2.11}$$

and the sums of the random variables $f_k(X_k)$ satisfy a central limit theorem. More precisely

$$\frac{\sum_{k=A}^{B} f_k(X_k) - \overline{M}(N)}{\overline{D}(N)} \Rightarrow \mathcal{N}(0,1)$$

and for all $h \geq 0$ we have, as $N \to \infty$,

$$\mathbf{E}\left(\frac{\sum_{k=A}^{B} f_k(X_k) - \overline{M}(N)}{\overline{D}(N)}\right)^h \to \int_{-\infty}^{\infty} x^h \, d\Phi(x).$$

Proof. Let

$$P(x,A) := \mathbf{Pr}[X_{k+1} \in A | X_k = x]$$

(which does not depend on k) denote the transition function of the Markov chain $(X_k, k \geq 0)$ and

$$\beta := 1 - \sup_{x_1,x_2,A} |P(x_1,A) - P(x_2,A)|$$

its ergodicity coefficient. If the f_k are injective on $\{0,\ldots,a\}$, then $(f_k(X_k), k \geq 0)$ is a Markov chain with ergodicity coefficient β and we get, by Lemma 2 of Dobrušin [4] and with $\mathbf{Var} f_k(X_k) = \sigma_{k,k}^{(2)} \geq c$,

$$\mathbf{Var} \sum_{k=s}^{s'} f_k(X_k) \geq \frac{c}{100}(s' - s + 1)\beta.$$

If some of the f_k are not injective, we get the same result by considering injective functions $\tilde{f}_k$ which tend to f_k. Since $D(N)^2 = \mathbf{Var} \sum_{k=0}^{[\log_\alpha N]} f_k(X_k)$ and $\overline{D}(N)^2 = \mathbf{Var} \sum_{k=A}^{B} f_k(X_k)$, this proves (2.11) if β is positive.

Suppose $\beta = 0$. Then there exist $x_1, x_2 \in \{0,\ldots,a\}$ and a set A such that $P(x_1,A) = 0$ and $P(x_2,A) = 1$, because $P(x,A)$ attains just finitely many values. We have $P(x,\{0\}) > 0$ for all x. Hence, if $0 \in A$, we get a contradiction to $P(x_1,A) = 0$ and, if $0 \notin A$, we get a contradiction to $P(x_2,A) = 1$. Therefore we have $\beta > 0$.

For each $h \geq 2$, the moments $\mathbf{E}\,|f_k(X_k)|^h$ are jointly bounded because of $f_k(b) = \mathcal{O}(1)$. Hence, if the f_k are injective, all conditions of Theorem 4 of Lifšic [15] are satisfied and we have convergence of (absolute) moments to those of the normal distribution. An inspection of Lifšic' proof shows that, as above, this is valid for non-injective f_k too. □

Now we are able to compare the moments of $\overline{f}(n)$ and $\sum f_k(X_k)$.

Lemma 7. *For every $h \geq 1$ and every $\lambda > 0$ we have*

$$\frac{1}{N}\sum_{n<N}\left(\frac{\overline{f}(n)-\overline{M}(N)}{\overline{D}(N)}\right)^h = \mathbf{E}\left(\frac{\sum_{k=A}^{B} f_k(X_k)-\overline{M}(N)}{\overline{D}(N)}\right)^h + \mathcal{O}\left(\frac{1}{(\log N)^\lambda}\right).$$

Proof. We have

$$\begin{aligned}\frac{1}{N}\sum_{n<N}\left(\frac{\overline{f}(n)-\overline{M}(N)}{\overline{D}(N)}\right)^h &= \frac{1}{N}\sum_{n<N}\left(\frac{\sum_{k=A}^{B}(f_k(\epsilon_k(n))-\mu_k)}{\overline{D}(N)}\right)^h \\ &= \frac{1}{N}\sum_{n<N}\sum_{k_1=A}^{B}\cdots\sum_{k_h=A}^{B}\prod_{i=1}^{h}\frac{f_{k_i}(\epsilon_{k_i}(n))-\mu_{k_i}}{\overline{D}(N)}\end{aligned}$$

$$= \sum_{A\leq k_1,\ldots,k_h\leq B}\ \sum_{0\leq b_1,\ldots,b_h\leq a}\frac{1}{N}\#\{n<N \mid \epsilon_{k_1}(n)=b_1,\ldots,\epsilon_{k_h}(n)=b_h\} \times\prod_{i=1}^{h}\frac{f_{k_i}(b_i)-\mu_{k_i}}{\overline{D}(N)}$$

and

$$\mathbf{E}\left(\frac{\sum_{k=A}^{B} f_k(X_k)-\overline{M}(N)}{\overline{D}(N)}\right)^h = \sum_{A\leq k_1,\ldots,k_h\leq B}\ \sum_{0\leq b_1,\ldots,b_h\leq a}\mathbf{Pr}[X_{k_1}=b_1,\ldots,X_{k_h}=b_h]\prod_{i=1}^{h}\frac{f_{k_i}(b_i)-\mu_{k_i}}{\overline{D}(N)}.$$

By Lemmata 4 and 6, these expressions are equal up to an error term $\mathcal{O}\left((\log N)^{h/2-\lambda}\right)$. Since λ can be chosen arbitrarily, the lemma is proved. □

3. Plan of the Proof of Theorem 3

We set M, D and $\overline{f}$ as in Theorem 2 with the only difference $B := [r\log_\alpha N] - A$ $(A = [(\log N)^\eta])$. Then an argument similar to

Lemma 5 shows that it is enough to prove

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \frac{\overline{f}(P(n)) - \overline{M}(N^r)}{\overline{D}(N^r)} < x\right\} \to \Phi(x)$$

and

$$\frac{1}{\pi(N)}\#\left\{p < N \,\middle|\, \frac{\overline{f}(P(p)) - \overline{M}(N^r)}{\overline{D}(N^r)} < x\right\} \to \Phi(x).$$

In fact, we prove that the centralized moments

$$B_h(N) = \frac{1}{N}\sum_{n<N}\left(\frac{\overline{f}(P(n)) - \overline{M}(N^r)}{\overline{D}(N^r)}\right)^h$$

and

$$C_h(N) = \frac{1}{\pi(N)}\sum_{p<N}\left(\frac{\overline{f}(P(p)) - \overline{M}(N^r)}{\overline{D}(N^r)}\right)^h$$

converge (for $N \to \infty$) by comparing them to $A_h(N^r)$. By proceeding as in the proof of Lemma 7 and by using the following lemma, it follows that for each fixed integer $h \geq 0$, $B_h(N) - A_h(N^r) \to 0$ and $C_h(N) - A_h(N^r) \to 0$ as $N \to \infty$. (Of course, this proves Theorem 3. We just have to replace Lemma 4 by the following property.)

Lemma 8 (Main Lemma). *Let $P(n)$ be an integer polynomial of degree $r \geq 1$ and positive leading term. Then for every $h \geq 1$ and for every $\lambda > 0$ we have*

$$\frac{1}{N}\#\{n < N \mid \epsilon_{k_1}(P(n)) = b_1, \ldots, \epsilon_{k_h}(P(n)) = b_h\}$$
$$= q_{k_1,\ldots,k_h,b_1,\ldots,b_h} + \mathcal{O}\left(\frac{1}{(\log N)^\lambda}\right)$$

and

$$\frac{1}{\pi(N)}\#\{p < N \mid \epsilon_{k_1}(P(p)) = b_1, \ldots, \epsilon_{k_h}(P(p)) = b_h\}$$
$$= q_{k_1,\ldots,k_h,b_1,\ldots,b_h} + \mathcal{O}\left(\frac{1}{(\log N)^\lambda}\right)$$

uniformly for all integers

$$(\log N)^\eta \leq k_1, k_2, \ldots, k_h \leq \log_\alpha N^r - (\log N)^\eta$$

and $b_1, b_2, \ldots, b_h \in \{0, 1, \ldots, a\}$. (The $q_{k_1,\ldots,k_h,b_1,\ldots,b_h}$ are as in Lemma 4.)

It turns out that this lemma can be proved similarly to that of Bassily and Kátai [1], i.e. with help of exponential sums. The only difficulty is to get a nice condition for extracting the digits $\epsilon_k(n)$ without using greedy algorithms. This problem is solved in the next section with help of a proper tiling of the unit square. Section 5 provides proper estimates for exponential sums. These are the two main ingredients of the proof which is then completed in Sections 6 and 7.

4. Tilings

The aim of this section is to provide proper tilings of the plane corresponding to our digital expansions in order to get an analogue to q-ary expansions where we have

$$\epsilon_{q,k}(n) = b \iff \left\langle \frac{n}{q^{k+1}} \right\rangle \in \left[\frac{b}{q}, \frac{b+1}{q}\right), \tag{4.1}$$

if $\langle x \rangle$ denotes the fractional part of x.

For our expansions, we will have to take into account the values of $\left\langle \frac{n}{\alpha^k(\alpha+1)} \right\rangle$ and $\left\langle \frac{n}{\alpha^{k+1}(\alpha+1)} \right\rangle$. By taking just one value into account, there are overlaps and we cannot get something like (4.1) or (4.2).

Proposition 1. *Let A_b, $0 \le b \le a$, denote rectangles in the plane $\mathbb{R}^2$ defined as the convex hull of the following corners:*

$$A_0 : \left(-\frac{\alpha}{\alpha^2+1}, \frac{\alpha^2}{\alpha^2+1}\right), (0,1), \left(\frac{\alpha+1}{\alpha^2+1}, -\frac{\alpha-1}{\alpha^2+1}\right), \left(\frac{1}{\alpha^2+1}, -\frac{\alpha}{\alpha^2+1}\right),$$

$$A_b : \left(\frac{(b-1)\alpha+1}{\alpha^2+1}, \frac{\alpha^2-\alpha+b}{\alpha^2+1}\right), \left(\frac{b\alpha+1}{\alpha^2+1}, \frac{\alpha^2-\alpha+b+1}{\alpha^2+1}\right),$$

$$\left(\frac{(b+1)\alpha+1}{\alpha^2+1}, -\frac{\alpha-b-1}{\alpha^2+1}\right), \left(\frac{b\alpha+1}{\alpha^2+1}, -\frac{\alpha-b}{\alpha^2+1}\right) \text{ for } b \in \{1, \dots, a-1\},$$

$$A_a : \left(\frac{\alpha^2-\alpha}{\alpha^2+1}, \frac{a\alpha}{\alpha^2+1}\right), \left(\frac{\alpha^2-\alpha+1}{\alpha^2+1}, \frac{\alpha^2}{\alpha^2+1}\right), (1,0), \left(\frac{\alpha^2}{\alpha^2+1}, -\frac{1}{\alpha^3+\alpha^2}\right).$$

Then these rectangles induce a periodic tiling of the plane with periods $\mathbb{Z}\times\mathbb{Z}$, i.e. they constitute a partition of the unit square modulo 1. Their slopes are $(\alpha, 1)$, $(-1, \alpha)$ and their areas are $\lambda_2(A_b) = p_b$, $b = 0, \dots, a$, with p_b as in Corollary 1. Furthermore, if $\epsilon_k(n) = b$ then

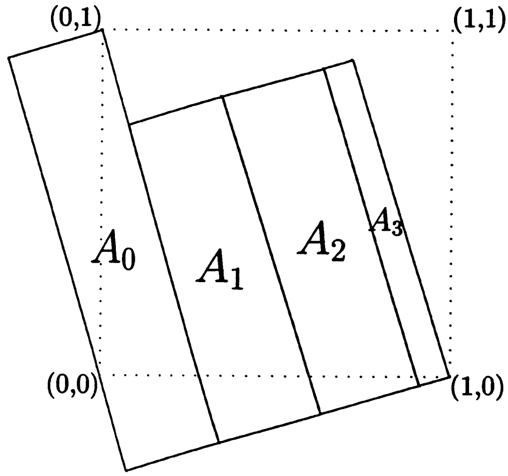
$$\left(\left\langle \frac{n}{\alpha^k(\alpha+1)} \right\rangle, \left\langle \frac{n}{\alpha^{k+1}(\alpha+1)} \right\rangle\right) \in (A_b \bmod 1) + \mathcal{O}\left(\alpha^{-k}\right). \tag{4.2}$$

Essentially, this proposition says that there is an analogue to (4.1) for G-ary expansions with a small error of order $\mathcal{O}\left(\alpha^{-k}\right)$ for the k-th digit. We want to remark that Farinole [12] considered a very similar question.

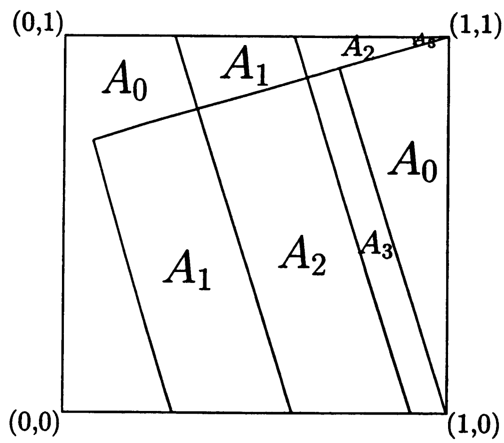
Remark. The rectangles A_b modulo 1 constitute a Markov partition of the toral automorphism with matrix

$$\begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix}.$$

Example. Before proving the proposition, we illustrate the example $a = 3$:



which looks like follows in $\mathbb{R}^2/\mathbb{Z}^2$:



Proof of Proposition 1. Suppose that n is given by $n = \sum \epsilon_j G_j$. Then we have

$$\begin{aligned}
&\left\langle \frac{n}{\alpha^k(\alpha+1)} \right\rangle \\
&\quad = \left\langle \cdots + \epsilon_{k+1}\frac{\alpha^2}{\alpha^2+1} + \epsilon_k \frac{\alpha}{\alpha^2+1} + \epsilon_{k-1}\frac{1}{\alpha^2+1} + \cdots \right\rangle + \mathcal{O}\left(\frac{1}{\alpha^k}\right) \\
&\quad = \left\langle \cdots + \epsilon_{k+1}\frac{(-\alpha)^{-1}\alpha}{\alpha^2+1} + \epsilon_k \frac{\alpha}{\alpha^2+1} + \epsilon_{k-1}\frac{1}{\alpha^2+1} + \cdots \right\rangle + \mathcal{O}\left(\frac{1}{\alpha^k}\right) \\
&\quad = \left\langle \frac{(x+\epsilon_k+y)\alpha}{\alpha^2+1} \right\rangle + \mathcal{O}\left(\frac{1}{\alpha^k}\right)
\end{aligned}$$

with the abbreviations

$$x = \cdots + \epsilon_{k+2}(-\alpha)^{-2} + \epsilon_{k+1}(-\alpha)^{-1},$$
$$y = \epsilon_{k-1}\alpha^{-1} + \epsilon_{k-2}\alpha^{-2} + \cdots,$$

where we have used (1.1) and that $\frac{\alpha}{\alpha^2+1}\alpha^j - \frac{\alpha}{\alpha^2+1}(-\alpha)^{-j}$ is an integer for all $j \geq 0$ (see (5.2)). Similarly we get

$$\left\langle \frac{n}{\alpha^{k+1}(\alpha+1)} \right\rangle = \left\langle \frac{-\alpha^2 x + \epsilon_k + y}{\alpha^2+1} \right\rangle + \mathcal{O}\left(\frac{1}{\alpha^k}\right).$$

By Rényi [19], we know that $(\epsilon_{k-1}, \epsilon_{k-2}, \ldots) < (\epsilon'_{k-1}, \epsilon'_{k-2}, \ldots)$ (lexicographically) implies

$$\epsilon_{k-1}\alpha^{-1} + \epsilon_{k-2}\alpha^{-2} + \cdots < \epsilon'_{k-1}\alpha^{-1} + \epsilon'_{k-2}\alpha^{-2} + \cdots$$

Hence, if $\epsilon_k < a$, then y is bounded by

$$0 \leq y < a\alpha^{-1} + a\alpha^{-3} + a\alpha^{-5} + \cdots = 1$$

and by

$$0 \leq y < a\alpha^{-2} + a\alpha^{-4} + a\alpha^{-6} + \cdots = \alpha^{-1}$$

if $\epsilon_k = a$. Similarly, x is bounded by

$$x < a\alpha^{-2} + a\alpha^{-4} + a\alpha^{-6} + \cdots = \alpha^{-1}$$

for all ϵ_k, by

$$x > -a\alpha^{-1} - a\alpha^{-3} - a\alpha^{-5} - \cdots = -1$$

for $\epsilon_k = 0$ and by

$$x > -(a-1)\alpha^{-1} - a\alpha^{-3} - a\alpha^{-5} - \cdots = \alpha^{-1} - 1$$

for $\epsilon_k > 0$.

If we put these limits into $\left(\frac{(x+\epsilon_k+y)\alpha}{\alpha^2+1}, \frac{-\alpha^2 x+\epsilon_k+y}{\alpha^2+1}\right)$, we obtain the given corners for A_b. It is now an easy exercise that (the interiors of) these rectangles are pairwisely disjoint (and situated as in the example) and that they induce a periodic tiling in $\mathbb{R}^2$ with periods $\mathbb{Z}^2$. □

5. Exponential Sums

In order to prove the Main Lemma we have to study exponential sums of the form

$$\frac{1}{N}\sum_{n<N} e\left(\frac{S}{\alpha+1}P(n)\right)$$

and

$$\frac{1}{\pi(N)}\sum_{p<N} e\left(\frac{S}{\alpha+1}P(p)\right),$$

where

$$S = \frac{m_{1,1}}{\alpha^{k_1}} + \frac{m_{1,2}}{\alpha^{k_1+1}} + \cdots + \frac{m_{h,1}}{\alpha^{k_h}} + \frac{m_{h,2}}{\alpha^{k_h+1}}$$

with integers $m_{i,j}$ $(1 \le i \le h, 1 \le j \le 2)$ and $e(x) := e^{2\pi i x}$, as usual.

Lemma 9. *Let $m_{i,j}$, $i \in \{1, \ldots, h\}, j \in \{1,2\}$ be integers with $|m_{i,j}| \le (\log N)^\delta$ for all i, j and*

$$(\log N)^\eta \le k_1 < k_2 < \cdots < k_h \le \log_\alpha N^r - (\log N)^\eta$$

for arbitrary constants $\delta > 0$, $\eta > 0$. Then, if $S \ne 0$,

$$\frac{\alpha^{(\log N)^{\eta'}}}{N^r} \ll |S| \ll \alpha^{-(\log N)^{\eta'}}$$

for all $\eta' < \eta$.

Proof. Clearly we have

$$S \ll \frac{(\log N)^\delta}{\alpha^{k_1}} \le \frac{(\log N)^\delta}{\alpha^{(\log N)^\eta}} \ll \alpha^{-(\log N)^{\eta'}}.$$

For the lower bound, we first remark that α^k is given by

$$\alpha^k = G'_k \alpha + G'_{k-1}, \tag{5.1}$$

where the sequence $(G'_j)_{j \ge 0}$ is defined by $G'_0 = 0$, $G'_1 = 1$ and $G'_j = aG'_{j-1} + G'_{j-2}$ for $j \ge 2$. Therefore we have

$$S = \frac{m_{1,1}\alpha^{k_h - k_1 + 1} + m_{1,2}\alpha^{k_h - k_1} + \cdots + m_{h,1}\alpha + m_{h,2}}{\alpha^{k_h+1}} = \frac{A\alpha + B}{\alpha^{k_h+1}},$$

with

$$A = m_{1,1}G'_{k_h - k_1 + 1} + m_{1,2}G'_{k_h - k_1} + \cdots + m_{h,1}$$

and

$$B = m_{1,1}G'_{k_h - k_1} + m_{1,2}G'_{k_h - k_1 - 1} + \cdots + m_{h,2}.$$

We have

$$|(A\alpha + B)(A\alpha^{-1} - B)| = |A^2 - aAB - B^2| \ge 1$$

if $A \ne 0$ or $B \ne 0$ and

$$A\alpha^{-1} - B \ll (\log N)^\delta$$

because G'_j is given by

$$G'_j = \frac{\alpha}{\alpha^2 + 1}\alpha^j - \frac{\alpha}{\alpha^2 + 1}(-\alpha)^{-j} \tag{5.2}$$

(cf. (1.1)). Hence

$$|S| \gg \frac{1}{(\log N)^\delta \alpha^{k_h}} \gg \frac{\alpha^{(\log N)^{\eta'}}}{N^r}.$$

□

The next two lemmata are adapted from Lemma 6.2 and Theorem 10 of Hua [14].

Lemma 10. *Let $P(n)$ be a polynomial of degree r with leading coefficient β. For every $\tau_0 > 0$, we have a $\tau > 0$ such that*

$$N^{-r}(\log N)^{\tau} < \beta < (\log N)^{-\tau}$$

implies

$$\frac{1}{N}\sum_{n<N} e(P(n)) = \mathcal{O}\left((\log N)^{-\tau_0}\right)$$

as $N \to \infty$.

Lemma 11. *Let $P(n)$ be as in Lemma 10. For every $\tau_0 > 0$, we have a $\tau > 0$ such that*

$$N^{-r}(\log N)^{\tau} < \beta < (\log N)^{-\tau}$$

implies

$$\frac{1}{\pi(N)}\sum_{p<N} e(P(p)) = \mathcal{O}\left((\log N)^{-\tau_0}\right).$$

as $N \to \infty$.

Note that we can apply these two lemmas for $\beta = S/(\alpha+1)$ with $S \neq 0$ for any choice of $\tau > 0$ since

$$\alpha^{-(\log N)^{\eta}} \ll (\log N)^{-\tau}.$$

Lemma 10 can be deduced for $r \geq 12$ from Theorem I in Chapter VI of Vinogradov [20] because of

$$\beta = \frac{1}{q} + \frac{\theta}{q^2} \text{ with } \theta \leq 1,\ (\log N)^{\tau} < q < N^{r}(\log N)^{-\tau}.$$

if $\beta \in [\frac{1}{q}, \frac{1}{q+1}]$. For general r, the two lemmata can be proved by replacing q by $\frac{1}{\beta}$ in the proofs of Lemma 6.2 and Theorem 10 of Hua and using the following lemma.[2]

Lemma 12.

$$\sum_{n=F+1}^{F+[\frac{1}{\beta}]} \min\left(U, \frac{1}{2\|n\beta\|}\right) \ll U + \frac{1}{\beta}\log\frac{1}{\beta},$$

where $\|x\| = \min(\langle x\rangle, 1-\langle x\rangle)$.

Proof. In each of the intervals $[m\beta, (m+1)\beta)$ and $(1-(m+1)\beta, 1-m\beta]$, $0 \leq m \leq \frac{1}{2}[\frac{1}{\beta}]$, we have at most one $\{n\beta\}$. Therefore

$$\sum_{n=F+1}^{F+[\frac{1}{\beta}]} \min\left(U, \frac{1}{2\|n\beta\|}\right) \leq 2\sum_{m=0}^{\frac{1}{2}[\frac{1}{\beta}]} \min\left(U, \frac{1}{2m\beta}\right) \ll U + \frac{1}{\beta}\log\frac{1}{\beta}$$

□

[2]Unfortunately we could not find a direct reference for Lemmata 10 and 11.

6. The Boundary of the Tilings

Lemma 13. *Let $P(x)$ be an arbitrary polynomial of degree r and $\Delta > 0$. Set*

$$E_{k,b}(\Delta) := \#\left\{n \le N \,\middle|\, \left(\left\langle \frac{P(n)}{\alpha^k(\alpha+1)}\right\rangle, \left\langle \frac{P(n)}{\alpha^{k+1}(\alpha+1)}\right\rangle\right) \in U_b(\Delta)\right\},$$

$$F_{k,b}(\Delta) := \#\left\{p \le N \,\middle|\, \left(\left\langle \frac{P(p)}{\alpha^k(\alpha+1)}\right\rangle, \left\langle \frac{P(p)}{\alpha^{k+1}(\alpha+1)}\right\rangle\right) \in U_b(\Delta)\right\},$$

where

$$U_b(\Delta) = \left\{\left(x_1 + y_1 - \frac{1}{\alpha}y_2, x_2 + \frac{1}{\alpha}y_1 + y_2\right) \,\middle|\, (x_1, x_2) \in \partial A_b, |y_i| \le \frac{\Delta}{2}, i = 1,2\right\}.$$

(∂A_b denotes the boundary of A_b.) Let $(\log N)^\eta < k < \log_\alpha N^r - (\log N)^\eta$ for some (fixed) $\eta > 0$ and λ an arbitrary positive constant. Then, uniformly in k, we have

$$E_{k,b}(\Delta) \ll \Delta N + N(\log N)^{-\lambda}, \qquad F_{k,b}(\Delta) \ll \Delta\pi(N) + N(\log N)^{-\lambda}.$$

Proof. We use discrepancies to prove this lemma. The isotropic discrepancy J_N of the points $(x_{1,1}, x_{1,2}), \ldots, (x_{N,1}, x_{N,2})$ in $\mathbb{R}^2$ is defined by

$$J_N = \sup_{C \subseteq \mathbb{T}^2} \left| \frac{1}{N} \sum_{n=1}^{N} \chi_C(\{x_{n,1}\}, \{x_{n,2}\}) - \lambda_2(C) \right|,$$

where the supremum is taken over all convex subsets C of $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. It can be estimated by the normal discrepancy D_N which is defined by

$$D_N = \sup_{I \subseteq \mathbb{T}^2} \left| \frac{1}{N} \sum_{n=1}^{N} \chi_I(\{x_{n,1}\}, \{x_{n,2}\}) - \lambda_2(I) \right|,$$

where the supremum is taken over all 2-dimensional intervals I of $\mathbb{T}^2$:

$$D_N \le J_N \le (8\sqrt{2}+1)\sqrt{D_N}$$

(see Theorem 1.12 of Drmota and Tichy [9]).

To get an estimate for D_N we use the following version of Erdős-Turán-Koksma's inequality:

$$D_N \ll \frac{1}{M} + \sum_{\substack{(m_1,m_2)\in\mathbb{Z}^2\setminus(0,0):\\ |m_1|,|m_2|\le M}} \min\left(\frac{1}{|m_1|}, \frac{1}{|m_2|}, \frac{1}{|m_1 m_2|}\right) \times \left| \frac{1}{N} \sum_{n=1}^{N} e(m_1 x_{n,1} + m_2 x_{n,2}) \right|,$$

where M is an arbitrary positive integer (and $\frac{1}{0} = +\infty$) (cf. Theorem 1.21 of [9]).

We set $(x_{n,1}, x_{n,2}) = \left(\frac{P(n)}{\alpha^j(\alpha+1)}, \frac{P(n)}{\alpha^{j+1}(\alpha+1)}\right)$ and $M = (\log N)^{2\lambda}$. Then we have, since $U_b(\Delta)$ is the union of 4 convex subsets and the conditions of Lemmata 9 and 10 hold,

$$\begin{aligned} E_{k,b}(\Delta) &\le 4J_N N + \lambda_2(U_b(\Delta))N \\ &\ll \frac{N}{(\log N)^\lambda} + \left(\log(\log N)^{2\lambda}\right)^2 N(\log N)^{-\tau_0/2} + \Delta N. \end{aligned}$$

Similarly we get, with Lemma 11,

$$F_{k,b}(\Delta) \ll \frac{N}{(\log N)^\lambda} + \left(\log(\log N)^{2\lambda}\right)^2 N(\log N)^{-\tau_0/2} + \Delta\pi(N).$$

We can choose $\tau_0 > 2\lambda$ and the inequalities are proved. □

7. Proof of Main Lemma

For $b \in \{0, \dots, a\}$ let $\varphi_b(x, y)$ be a function periodic mod 1, defined explicitly in $[0,1] \times [0,1]$ by

$$\varphi_b(x_1, x_2) := \begin{cases} 1 & \text{if } (x_1, x_2) \in A_b \setminus \partial A_b \\ \frac{1}{2} & \text{if } (x_1, x_2) \in \partial A_b \\ 0 & \text{otherwise} \end{cases}.$$

Its Fourier expansion $\sum\sum c_{m_1,m_2}(b) e(m_1x_1 + m_2x_2)$ is given by

$$c_{0,0}(b) = \lambda_2(A_b),$$

$$c_{m_1,m_2}(b) = \sum_{(x_1,x_2)\in V(A_b)} \frac{\left|\det((x_1 - y_1, x_2 - y_2))_{(y_1,y_2)\in\Gamma(x_1,x_2)}\right|}{\prod_{(y_1,y_2)\in\Gamma(x_1,x_2)} -2\pi i(m_1(x_1 - y_1) + m_2(x_2 - y_2))} \times e(-m_1x_1 - m_2x_2),$$

where $V(A_b)$ denotes the set of vertices of the rectangle A_b and $\Gamma(x_1, x_2)$ the set of vertices adjacent to $(x_1, x_2) \in V(A_b)$ (cf. Drmota [7], Lemma 1). This can be bounded by (cf. Lemma 2 of Drmota [7])

$$\begin{aligned} (7.1)\quad & |c_{m_1,m_2}(b)|^2 \\ &\ll \sum_{(x_1,x_2)\in V(A_b)} \prod_{(y_1,y_2)\in\Gamma(x_1,x_2)} \frac{1}{(1 + |m_1(x_1 - y_1) + m_2(x_2 - y_2)|)^2} \\ &\ll \frac{1}{\left(1 + \left|m_1 + \frac{1}{\alpha}m_2\right|\right)^2 \left(1 + \left|m_2 - \frac{1}{\alpha}m_1\right|\right)^2} \\ &\ll \min\left(1, \frac{1}{\tilde{m}_1^2}\right) \min\left(1, \frac{1}{\tilde{m}_2^2}\right) \end{aligned}$$

uniformly for all (m_1, m_2), where the constants implied by $\ll$ only depend on A_b and $\tilde{m}_1 := m_1 + \frac{1}{\alpha} m_2, \tilde{m}_2 := m_2 - \frac{1}{\alpha} m_1$.

For (small) $\Delta > 0$ we consider the function

$$\psi_b(x_1, x_2) := \frac{1}{\Delta^2} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \varphi_b(x_1 + z_1 - \frac{1}{\alpha} z_2, x_2 + \frac{1}{\alpha} z_1 + z_2) dz_1 dz_2.$$

The Fourier expansion $\sum\sum d_{m_1,m_2}(b) e(m_1 x_1 + m_2 x_2)$ of this function is given by

$$d_{m_1,m_2}(b) = c_{m_1,m_2}(b) \frac{\left(e(\frac{\tilde{m}_1 \Delta}{2}) - e(-\frac{\tilde{m}_1 \Delta}{2})\right)\left(e(\frac{\tilde{m}_2 \Delta}{2}) - e(-\frac{\tilde{m}_2 \Delta}{2})\right)}{-4\pi^2 \tilde{m}_1 \tilde{m}_2 \Delta^2}$$

if $(m_1, m_2) \neq (0,0)$ and

$$d_{0,0}(b) = c_{0,0}(b) = \lambda_2(A_b).$$

Hence

$$|d_{m_1,m_2}(b)| \ll \min\left(1, \frac{1}{|\tilde{m}_1|}, \frac{1}{\Delta \tilde{m}_1^2}\right) \min\left(1, \frac{1}{|\tilde{m}_2|}, \frac{1}{\Delta \tilde{m}_2^2}\right) \tag{7.2}$$

and

$$d_{m_1,m_2}(b) = c_{m_1,m_2}(b) \left(1 + \mathcal{O}\left(\tilde{m}_1^2 \Delta^2\right)\right)\left(1 + \mathcal{O}\left(\tilde{m}_2^2 \Delta^2\right)\right) \tag{7.3}$$

as $\tilde{m}_i \Delta \to 0$.

It is clear that $0 \le \psi_b(x_1, x_2) \le 1$ for every pair (x_1, x_2) and that

$$\psi_b(x_1, x_2) = \begin{cases} 1 & \text{if } (x_1, x_2) \in A_b \setminus U_b(\Delta) \\ 0 & \text{if } (x_1, x_2) \notin A_b \cup U_b(\Delta) \end{cases}$$

We define

$$F\Big((x_{1,1}, x_{1,2}), \ldots, (x_{h,1}, x_{h,2})\Big) := \psi_{b_1}(x_{1,1}, x_{1,2}) \cdots \psi_{b_h}(x_{h,1}, x_{h,2})$$

and

$$t(n) := F\left(\left(\frac{n}{\alpha^{k_1}(\alpha+1)}, \frac{n}{\alpha^{k_1+1}(\alpha+1)}\right), \ldots, \left(\frac{n}{\alpha^{k_h}(\alpha+1)}, \frac{n}{\alpha^{k_h+1}(\alpha+1)}\right)\right).$$

We set

$$\Sigma_1 := \#\{n < N \mid \epsilon_{k_1}(P(n)) = b_1, \ldots, \epsilon_{k_h}(P(n)) = b_h\}$$
$$\Sigma_2 := \#\{p < N \mid \epsilon_{k_1}(P(p)) = b_1, \ldots, \epsilon_{k_h}(P(p)) = b_h\}$$

and get, with (4.2) and Lemma 13,

$$\left|\Sigma_1 - \sum_{n<N} t(P(n))\right| \le E_{k_1,b_1}(\Delta) + \cdots + E_{k_h,b_h}(\Delta),$$

$$\left|\Sigma_2 - \sum_{n<N} t(P(n))\right| \le F_{k_1,b_1}(\Delta) + \cdots + F_{k_h,b_h}(\Delta),$$

for Δ greater than the error terms $\mathcal{O}\left(\alpha^{-k_i}\right) = \mathcal{O}\left(\alpha^{-(\log N)^{\eta}}\right)$ of (4.2).

Furthermore, set $\mathbf{V} := \left(\frac{1}{\alpha^{k_1}(\alpha+1)}, \frac{1}{\alpha^{k_1+1}(\alpha+1)}, \dots, \frac{1}{\alpha^{k_h}(\alpha+1)}, \frac{1}{\alpha^{k_h+1}(\alpha+1)}\right)^t$ and let $\mathcal{M}$ be the set of vectors $\mathbf{M} = (m_{1,1}, m_{1,2}, \dots, m_{h,1}, m_{h,2})$ with integer entries $m_{i,j}$.

Then we have

$$t(n) = \sum_{\mathbf{M}\in\mathcal{M}} T_{\mathbf{M}} e(\mathbf{MV}n),$$

where

$$T_{\mathbf{M}} = d_{m_{1,1},m_{1,2}}(b_1)\dots d_{m_{h,1},m_{h,2}}(b_h).$$

and

$$\sum_{n<N} t(P(n)) = \sum_{\mathbf{M}\in\mathcal{M}} T_{\mathbf{M}} \sum_{n<N} e(\mathbf{MV}P(n)), \tag{7.4}$$

$$\sum_{p<N} t(P(p)) = \sum_{\mathbf{M}\in\mathcal{M}} T_{\mathbf{M}} \sum_{p<N} e(\mathbf{MV}P(p)).$$

If $|m_{i,j}| \le (\log N)^{2\delta}$ for all i,j, Lemmata 9 and 10 provide

$$\sum_{n<N} e(\mathbf{MV}P(n)) \ll N(\log N)^{-\tau_0},$$

if $\mathbf{MV} \ne 0$. Lemma 11 provides a similar result for primes.
Since $(m_{i,1}, m_{i,2}) \mapsto (\tilde{m}_{i,1}, \tilde{m}_{i,2})$ is, up to a constant, an orthogonal transformation, we have

$$\min\left(1, \frac{1}{|\tilde{m}_{i,1}|}\right)\min\left(1, \frac{1}{|\tilde{m}_{i,2}|}\right) \ll \min\left(1, \frac{1}{|m_{i,1}|}\right)\min\left(1, \frac{1}{|m_{i,2}|}\right)$$

and, with (7.2),

$$\begin{aligned}
&\sum_{\mathbf{M}\in\mathcal{M}:|m_{i,j}|\le(\log N)^{2\delta}} |T_{\mathbf{M}}| \\
&\qquad \ll \prod_{i=1}^{h} \sum_{m_{i,1},m_{i,2}=-[(\log N)^{2\delta}]}^{[(\log N)^{2\delta}]} \min\left(1, \frac{1}{|\tilde{m}_{i,1}|}\right)\min\left(1, \frac{1}{|\tilde{m}_{i,2}|}\right) \\
&\qquad \ll \left(\sum_{m_1,m_2=-[(\log N)^{2\delta}]}^{[(\log N)^{2\delta}]} \min\left(1, \frac{1}{|m_1|}\right)\min\left(1, \frac{1}{|m_2|}\right)\right)^h \\
&\qquad \ll \left(\log(\log N)^{2\delta}\right)^{2h}.
\end{aligned}$$

For the $\mathbf{M}$ with $|m_{i,j}| > (\log N)^{2\delta}$ for some i, j, we get similarly

$$\sum_{\mathbf{M}\in\mathcal{M}:\exists i,j \text{ with } |m_{i,j}|>(\log N)^{2\delta}} |T_{\mathbf{M}}|$$

$$\ll \left(\sum_{m=[(\log N)^{2\delta}]}^{\infty} \frac{1}{m^2\Delta}\right)\left(\sum_{m=1}^{\infty} \min\left(\frac{1}{m}, \frac{1}{m^2\Delta}\right)\right)^{2h-1}$$

$$\ll \frac{1}{(\log N)^{2\delta}\Delta}\left(\log\frac{1}{\Delta}+\Delta\right)^{2h-1} \ll \frac{(\log(\log N)^{\delta})^{2h-1}}{(\log N)^{\delta}},$$

if we set $\Delta = (\log N)^{-\delta}$. Therefore we have

$$\Sigma_1 = N \sum_{\mathbf{M}\in\mathcal{M}:\mathbf{MV}=0} T_{\mathbf{M}} + \mathcal{O}\left(N(\log N)^{-\tau_0/2} + N(\log N)^{-\delta/2}\right) \tag{7.5}$$

(and a similar expression for Σ_2). Since the main term depends on Δ, we want to replace $T_{\mathbf{M}}$ by

$$T'_{\mathbf{M}} = c_{m_{1,1},m_{1,2}}(b_1)\dots c_{m_{h,1},m_{h,2}}(b_h).$$

Hence we have to estimate the difference $\sum_{\mathbf{M}\in\mathcal{M}:\mathbf{MV}=0}(T_{\mathbf{M}} - T'_{\mathbf{M}})$.

By (7.3), we have

$$T_{\mathbf{M}} = T'_{\mathbf{M}}\left(1 + \mathcal{O}\left(\max_{i,j} \tilde{m}_{i,j}^2\Delta^2\right)\right). \tag{7.6}$$

First assume $|m_{i,j}| < (\log N)^{\delta/2}$ for all i, j. Then we obtain from (7.6) and (7.1)

$$\sum_{\mathbf{M}\in\mathcal{M}:|m_{i,j}|<(\log N)^{\delta/2}} |T_{\mathbf{M}} - T'_{\mathbf{M}}| \ll \sum_{\mathbf{M}\in\mathcal{M}:|m_{i,j}|<(\log N)^{\delta/2}} |T'_{\mathbf{M}}|(\log N)^{-\delta}$$

$$\ll \left(\sum_{m=1}^{[(\log N)^{\delta/2}]} \frac{1}{m}\right)^{2h} (\log N)^{-\delta} \le \frac{\left(\log(\log N)^{\delta/2}\right)^{2h}}{(\log N)^{\delta}} \ll (\log N)^{-\delta/2}.$$

and it remains to estimate the sum of the $T_{\mathbf{M}}$ and $T'_{\mathbf{M}}$ with $|m_{i,j}| > (\log N)^{\delta/2}$ for some i, j which satisfy $\mathbf{MV} = 0$, i.e.

$$m_{1,1}G'_{k_h-k_1+1} + m_{1,2}G'_{k_h-k_1} + \dots + m_{h,1} = 0, \tag{7.7}$$

$$m_{1,1}G'_{k_h-k_1} + m_{1,2}G'_{k_h-k_1-1} + \dots + m_{h,2} = 0. \tag{7.8}$$

This is done by the following lemma, where only one of the equations is needed.

Lemma 14. *We have*

$$\sum{}' \prod_{i=1}^{H} \min\left(1, \frac{1}{|m_i|}\right) \ll (\log N)^{-\frac{\delta}{2(H-1)^2}}, \tag{7.9}$$

where $\sum'$ *denotes the sum over all integer solutions* $(m_1, \dots, m_H)$ *of the linear equation*

$$\gamma_1 m_1 + \dots + \gamma_{H-1} m_{H-1} + m_H = 0, \tag{7.10}$$

(with integers $\gamma_i \neq 0$*) such that* $|m_i| > (\log N)^{\delta/2}$ *for some* i*. The constant implied by* $\ll$ *does not depend on the* γ_i*.*

Proof. First we remark that $m_i = 0$ for some i reduces the problem to a smaller one. For $H = 1$ (as well as for $H = 2$), the lemma is trivial. Hence we assume $H > 1$ and $m_i \neq 0$ for all i.

For every choice of $(m_1, \dots, m_{H-1})$, let m_H be the corresponding solution of (7.10). First we sum up over all choices with $|m_H| \geq |m_1 \dots m_{H-1}|^{1/(H-1)^2}$ and obtain

$$\sum \frac{1}{|m_1 \dots m_H|} \leq 2^{H-1} \sum_{m_1=1}^{\infty} \cdots \sum_{m_{H-1}=1}^{\infty} \frac{1}{m_1 \dots m_{H-1}} \frac{1}{(m_1 \dots m_{H-1})^{\frac{1}{(H-1)^2}}}$$
$$= 2^{H-1} \left(\sum_{m=1}^{\infty} \frac{1}{m^{1+\frac{1}{(H-1)^2}}} \right)^{H-1}.$$

If we consider only $|m_i| \geq (\log N)^{\delta/2}$ for some $i \leq H-1$, we have thus

$$\sum \frac{1}{|m_1 \dots m_H|} \ll (\log N)^{-\frac{\delta}{2(H-1)^2}}.$$

For $|m_H| \geq (\log N)^{\delta/2}$ and $|m_i| < (\log N)^{\delta/2}$ for $i \leq H-1$, we get

$$\sum \frac{1}{|m_1 \dots m_H|} \leq 2^{H-1} \left(\sum_{m=1}^{[(\log N)^{\delta/2}]} \frac{1}{m} \right)^{H-1} \frac{1}{(\log N)^{\delta/2}}$$
$$\ll \frac{\left(\log(\log N)^{\delta/2}\right)^{H-1}}{(\log N)^{\delta/2}}.$$

It remains to estimate the sum over the choices $(m_1, \dots, m_{H-1})$ with $|m_H| < |m_1 \dots m_{H-1}|^{1/(H-1)^2}$. W.l.o.g., assume $|\gamma_1 m_1| = \max_{1 \leq i \leq H-1} |\gamma_i m_i|$. Then we have

$$|m_H| < |\gamma_1 m_1 \dots \gamma_{H-1} m_{H-1}|^{\frac{1}{(H-1)^2}} \leq |\gamma_1 m_1|^{\frac{1}{H-1}} \tag{7.11}$$

and

$$|\gamma_2 m_2 + \dots + \gamma_{H-1} m_{H-1}| \in \left[|\gamma_1 m_1| - |\gamma_1 m_1|^{\frac{1}{H-1}}, |\gamma_1 m_1| + |\gamma_1 m_1|^{\frac{1}{H-1}} \right].$$

We split the possible range of $|\gamma_2 m_2|$ into

$$I_2 = \left(0, |\gamma_1 m_1| - |\gamma_1 m_1|^{(H-2)/(H-1)}\right] \text{ and}$$
$$J_2 = \left(|\gamma_1 m_1| - |\gamma_1 m_1|^{(H-2)/(H-1)}, |\gamma_1 m_1|\right].$$

For J_2, we obtain

$$(7.12) \qquad \sum_{m_2 : |\gamma_2 m_2| \in J_2} \frac{1}{|m_2|} \leq \frac{2|\gamma_1 m_1|^{\frac{H-2}{H-1}}}{|\gamma_2|} \frac{|\gamma_2|}{|\gamma_1 m_1| - |\gamma_1 m_1|^{\frac{H-2}{H-1}}} \leq \frac{4}{|\gamma_1 m_1|^{\frac{1}{H-1}}}.$$

Summing up over all such $(m_1, \ldots, m_H)$ with $|m_i| \geq (\log N)^{\delta/2}$ for some i, we get

$$(7.13) \quad \sum \frac{1}{|m_1 \ldots m_H|} \leq \sum_{\substack{m_1: \\ |\gamma_1 m_1| \geq (\log N)^{\delta/2}}} \frac{2^{H-1}}{|m_1||\gamma_1 m_1|^{\frac{1}{H-1}}} (\log |\gamma_1 m_1|)^{H-3}$$
$$\leq 2^H (\log N)^{-\frac{\delta}{2H}}.$$

Thus it suffices to consider m_2 with $|\gamma_2 m_2| \in I_2$ from now on. This implies

$$|\gamma_3 m_3 + \cdots + \gamma_{H-1} m_{H-1}| \in \Big[|\gamma_1 m_1 + \gamma_2 m_2| - |\gamma_1 m_1|^{\frac{1}{H-1}},$$
$$|\gamma_1 m_1 + \gamma_2 m_2| + |\gamma_1 m_1|^{\frac{1}{H-1}}\Big]$$

with

$$|\gamma_1 m_1 + \gamma_2 m_2| \geq |\gamma_1 m_1|^{\frac{H-2}{H-1}}.$$

We split the possible range of $|\gamma_3 m_3|$ into

$$J_3 = \Big(|\gamma_1 m_1 + \gamma_2 m_2| - |\gamma_1 m_1|^{(H-3)/(H-1)},$$
$$|\gamma_1 m_1 + \gamma_2 m_2| + |\gamma_1 m_1|^{(H-3)/(H-1)}\Big]$$

and $I_3 = (0, 2|\gamma_1 m_1|] \setminus J_3$. Similarly to (7.12), we obtain

$$\sum_{m_3 : |\gamma_3 m_3| \in J_3} \frac{1}{|m_3|} \leq \frac{8}{|\gamma_1 m_1|^{\frac{1}{H-1}}}$$

and the sum over these $(m_1, \ldots, m_H)$ can be estimated as in (7.13). For all other m_3, we have

$$|\gamma_1 m_1 + \gamma_2 m_2 + \gamma_3 m_3| \geq |\gamma_1 m_1|^{\frac{H-3}{H-1}}.$$

We can proceed inductively and in the only remaining case we would have

$$|\gamma_1 m_1 + \cdots + \gamma_{H-1} m_{H-1}| \geq |\gamma_1 m_1|^{\frac{1}{H-1}}$$

which contradicts (7.11). Thus the lemma is proved. □

We apply Lemma 14 for (7.7) with $H = 2h - 1$. Multiplying each term of the sum in (7.9) by $\min(1, 1/|m_{h,2}|)$ (where $m_{h,2}$ is determined by (7.8)), gives

$$\sum_{\mathbf{M}\in\mathcal{M}:\mathbf{MV}=0,\exists i,j:|m_{i,j}|>(\log N)^{\delta/2}} T'_{\mathbf{M}} \ll (\log N)^{-\frac{\delta}{8(h-1)^2}}$$

and the same estimate for $T_{\mathbf{M}}$.

Hence

$$\sum_{\mathbf{M}\in\mathcal{M}:\mathbf{MV}=0} T_{\mathbf{M}} = q'_{k_1,\ldots,k_h,b_1,\ldots,b_h} + \mathcal{O}\left((\log N)^{-\frac{\delta}{8(h-1)^2}}\right),$$

where

$$q'_{k_1,\ldots,k_h,b_1,\ldots,b_h} = \sum_{\mathbf{M}\in\mathcal{M}:\mathbf{MV}=0} T'_{\mathbf{M}}.$$

Together with (7.5), we obtain

$$\Sigma_1 = Nq'_{k_1,\ldots,k_h,b_1,\ldots,b_h} + \mathcal{O}\left(N(\log N)^{-\lambda}\right),$$

if we choose $\tau_0 = 2\lambda$ and $\delta = 8(h-1)^2\lambda$.

The result does not depend on the choice of the polynomial $P(n)$. If we set $P(n) = n$, Lemma 4 implies

$$q'_{k_1,\ldots,k_h,b_1,\ldots,b_h} = q_{k_1,\ldots,k_h,b_1,\ldots,b_h}.$$

Similarly we get

$$\Sigma_2 = \pi(N)q_{k_1,\ldots,k_h,b_1,\ldots,b_h} + \mathcal{O}\left(N(\log N)^{-\lambda}\right).$$

Remark. In the case $h = 1$ we have $\mathbf{MV} = 0$ only for $(m_1, m_2) = (0, 0)$ and $c_{0,0}(b) = \lambda_2(A_b) = p_b = q_{k,b}$.

8. Proof of Theorem 4

In order to prove independence of different digital expansions we can proceed essentially along the same lines as for the proof of Theorem 3. We just have to replace the Main Lemma (Lemma 8) by the following three (main) lemmas (corresponding to the three parts of Theorem 4) which imply

$$\frac{1}{N}\sum_{n<N}\prod_{\ell=1}^{2}\left(\frac{\overline{f}_\ell(P_\ell(n)) - \overline{M}_\ell(N^{r_\ell})}{\overline{D}_\ell(N^{r_\ell})}\right)^{h_\ell} - \prod_{\ell=1}^{2}\left(\frac{1}{N}\sum_{n<N}\left(\frac{\overline{f}_\ell(P_\ell(n)) - \overline{M}_\ell(N^{r_\ell})}{\overline{D}_\ell(N^{r_\ell})}\right)^{h_\ell}\right) \to 0$$

and the corresponding statement for primes. Therefore the twodimensional moments converge to those of the twodimensional normal law and Theorem 4 is proved.

Lemma 15. *Let q_1, q_2 be two positive coprime integers and $P_1(x)$, $P_2(x)$ two integer polynomials of degrees r_1 resp. r_2 with positive leading terms. Then for every $h_1, h_2 \ge 1$ and for every $\lambda > 0$ we have*

$$\begin{aligned}\frac{1}{N}\#\{n < N \mid \epsilon_{q_1,k_1}(P_1(n)) = b_1, \ldots, \epsilon_{q_1,k_{h_1}}(P_1(n)) = b_{h_1},\\ \epsilon_{q_2,l_1}(P_2(n)) = c_1, \ldots, \epsilon_{q_2,l_{h_2}}(P_2(n)) = c_{h_2}\}\\ = q_1^{-h_1} q_2^{-h_2} + \mathcal{O}\left(\frac{1}{(\log N)^\lambda}\right)\end{aligned}$$

and

$$\begin{aligned}\frac{1}{\pi(N)}\#\{p < N \mid \epsilon_{q_1,k_1}(P_1(p)) = b_1, \ldots, \epsilon_{q_1,k_{h_1}}(P_1(p)) = b_{h_1},\\ \epsilon_{q_2,l_1}(P_2(p)) = c_1, \ldots, \epsilon_{q_2,l_{h_2}}(P_2(p)) = c_{h_2}\}\\ = q_1^{-h_1} q_2^{-h_2} + \mathcal{O}\left(\frac{1}{(\log N)^\lambda}\right)\end{aligned}$$

uniformly for all integers

$$(\log N)^\eta \le k_1 < k_2 < \cdots < k_{h_1} \le r_1 \log_{q_1} N - (\log N)^\eta,$$

$$(\log N)^\eta \le l_1 < l_2 < \cdots < l_{h_2} \le r_2 \log_{q_2} N - (\log N)^\eta,$$

and $b_1, b_2, \ldots, b_{h_1} \in \{0, 1, \ldots, q_1 - 1\}$ resp. $c_1, c_2, \ldots, c_{h_2} \in \{0, 1, \ldots, q_2 - 1\}$.

Lemma 16. *Let $q \ge 2$ and $a \ge 1$ be two integers and $P_1(x)$, $P_2(x)$ two integer polynomials of degrees r_1 resp. r_2 with positive leading terms. Then for every $h_1, h_2 \ge 1$ and for every $\lambda > 0$ we have*

$$\begin{aligned}\frac{1}{N}\#\{n < N \mid \epsilon_{q,k_1}(P_1(n)) = b_1, \ldots, \epsilon_{q,k_{h_1}}(P_1(n)) = b_{h_1},\\ \epsilon_{G,l_1}(P_2(n)) = c_1, \ldots, \epsilon_{G,l_{h_2}}(P_2(n)) = c_{h_2}\}\\ = q^{-h_1} q_{l_1,\ldots,l_{h_2},c_1,\ldots,c_{h_2}} + \mathcal{O}\left(\frac{1}{(\log N)^\lambda}\right)\end{aligned}$$

and

$$\begin{aligned}\frac{1}{\pi(N)}\#\{p < N \mid \epsilon_{q,k_1}(P_1(p)) = b_1, \ldots, \epsilon_{q,k_{h_1}}(P_1(p)) = b_{h_1},\\ \epsilon_{G,l_1}(P_2(p)) = c_1, \ldots, \epsilon_{G,l_{h_2}}(P_2(p)) = c_{h_2}\}\\ = q^{-h_1} q_{l_1,\ldots,l_{h_2},c_1,\ldots,c_{h_2}} + \mathcal{O}\left(\frac{1}{(\log N)^\lambda}\right)\end{aligned}$$

uniformly for all integers

$$(\log N)^\eta \le k_1 < k_2 < \cdots < k_{h_1} \le r_1 \log_q N - (\log N)^\eta,$$
$$(\log N)^\eta \le l_1 < l_2 < \cdots < l_{h_2} \le r_2 \log_\alpha N - (\log N)^\eta,$$

and $b_1, b_2, \dots, b_{h_1} \in \{0, 1, \dots, q-1\}$ *resp.* $c_1, c_2, \dots, c_{h_2} \in \{0, 1, \dots, a\}$.

Lemma 17. *Let* $a_1, a_2 \ge 1$ *be two integers such that* $\sqrt{\frac{a_1^2+4}{a_2^2+4}}$ *is irrational and let* $G = (G_j)$ *and* $H = (H_j)$ *denote the corresponding second order recurrent sequences. Furthermore, let* $P_1(x)$, $P_2(x)$ *be two integer polynomials of degrees* r_1 *resp.* r_2 *with positive leading terms. Then for every* $h_1, h_2 \ge 1$ *and for every* $\lambda > 0$ *we have*

$$\begin{aligned}\frac{1}{N}\#\{n < N \mid \epsilon_{G,k_1}(P_1(n)) = b_1, \dots, \epsilon_{G,k_{h_1}}(P_1(n)) = b_{h_1},\\ \epsilon_{H,l_1}(P_2(n)) = c_1, \dots, \epsilon_{H,l_{h_2}}(P_2(n)) = c_{h_2}\}\\ = q^{(G)}_{k_1,\dots,k_{h_1},b_1,\dots,b_{h_1}} q^{(H)}_{l_1,\dots,l_{h_2},c_1,\cdots,c_{h_2}} + \mathcal{O}\left(\frac{1}{(\log N)^\lambda}\right)\end{aligned}$$

and

$$\begin{aligned}\frac{1}{\pi(N)}\#\{p < N \mid \epsilon_{G,k_1}(P_1(p)) = b_1, \dots, \epsilon_{G,k_{h_1}}(P_1(p)) = b_{h_1},\\ \epsilon_{H,l_1}(P_2(p)) = c_1, \dots, \epsilon_{H,l_{h_2}}(P_2(p)) = c_{h_2}\}\\ = q^{(G)}_{k_1,\dots,k_{h_1},b_1,\dots,b_{h_1}} q^{(H)}_{l_1,\dots,l_{h_2},c_1,\dots,c_{h_2}} + \mathcal{O}\left(\frac{1}{(\log N)^\lambda}\right)\end{aligned}$$

uniformly for all integers

$$(\log N)^\eta \le k_1 < k_2 < \cdots < k_{h_1} \le r_1 \log_{\alpha_1} N - (\log N)^\eta,$$
$$(\log N)^\eta \le l_1 < l_2 < \cdots < l_{h_2} \le r_2 \log_{\alpha_2} N - (\log N)^\eta,$$

and $b_1, b_2, \dots, b_{h_1} \in \{0, 1, \dots, a_1\}$ *resp.* $c_1, c_2, \dots, c_{h_2} \in \{0, 1, \dots, a_2\}$.

The proofs of these lemmas run along the same lines as the previous Main Lemma (compare also with [1] and [6]). We have to consider sums of the type

$$\sum_{\mathbf{M}_1,\mathbf{M}_2} T_{\mathbf{M}_1} T_{\mathbf{M}_2} \sum_{n<N} e(\mathbf{M}_1\mathbf{V}_1 P_1(n) + \mathbf{M}_2\mathbf{V}_2 P_2(n))$$

(cf. (7.4), where, in the q-ary case, $\mathbf{M}_\ell$, $\mathbf{V}_\ell$ and $T_{\mathbf{M}_\ell}$ are defined by

$$\begin{aligned}\mathbf{M}_\ell &= (m_1^{(\ell)}, \dots, m_{h_\ell}^{(\ell)}),\\ \mathbf{V}_\ell &= (q^{-k_1+1}, \dots, q^{-k_{h_\ell}+1}),\\ T_{\mathbf{M}_\ell} &= d_{m_1^{(\ell)},q_\ell}(b_1) \dots d_{m_{h_\ell}^{(\ell)},q_\ell}(b_{h_\ell})\end{aligned}$$

with

$$d_{m,q}(b) = \frac{e(-\frac{mb}{q}) - e(-\frac{m(b+1)}{q})}{2\pi i m} \frac{e(\frac{m\Delta}{2}) - e(-\frac{m\Delta}{2})}{2\pi i m\Delta}.$$

Especially, if $r_1 \neq r_2$, then the proof is straightforward and very similar to that of Proposition 1 in [6]. The reason is that there are no cancellations in the leading coefficient of the polynomial $\mathbf{M}_1\mathbf{V}_1P_1(n) + \mathbf{M}_2\mathbf{V}_2P_2(n)$ and consequently one can directly apply Lemmata 10 and 11 in order to estimate the corresponding exponential sums.

Therefore we concentrate on the case $r_1 = r_2$. Here we have to adapt certain properties.

Lemma 18. *Suppose that $q_1, q_2 \geq 2$ are coprime integers and c_1, c_2, r positive integers. For arbitrary (but fixed) integers h_1, h_2 let $m_j^{(\ell)}$ ($1 \leq j \leq h_\ell$, $\ell \in \{1,2\}$) be satisfying $m_j^{(\ell)} \not\equiv 0 \bmod q$ and $|m_j^{(\ell)}| \leq (\log N)^\delta$, where $\delta > 0$ is any given constant. Set*

$$S_\ell := \frac{m_1^{(\ell)}}{q_\ell^{k_1^{(\ell)}+1}} + \cdots + \frac{m_h^{(\ell)}}{q_\ell^{k_{h_\ell}^{(\ell)}+1}}.$$

Then, for

$$(\log N)^\eta \leq k_1^{(\ell)} < k_2^{(\ell)} < \cdots < k_{h_\ell}^{(\ell)} \leq \log_{q_\ell} N^r - (\log N)^\eta$$

we uniformly have

$$\frac{q^{(\log N)^{\eta'}}}{N^r} \ll |c_1 S_1 + c_2 S_2| \ll q^{-(\log N)^{\eta'}}$$

for all given $0 < \eta' < \eta$, where $q = \max\{q_1, q_2\}$.

This lemma is implicitly contained in the proof of Proposition 2 of [6], the statement of which is that of Lemma 15 for $r = 1$. However, by using Lemmata 10, 11 (which have not been used in this generality in [6]) and 18, Lemma 15 follows as Proposition 2 of [6].

Lemma 19. *Let $q \geq 2$ and $a \geq 1$ be two integers and c_1, c_2, r positive integers. For arbitrary (but fixed) integers h_1, h_2, let $m_j^{(1)}$ ($1 \leq j \leq h_1$) be integers satisfying $m_j^{(1)} \not\equiv 0 \bmod q$ and $|m_j^{(1)}| \leq (\log N)^\delta$ and let $m_{i,j}^{(2)}$ ($1 \leq i \leq h_2$, $j \in \{1,2\}$) be integers satisfying $|m_{i,j}^{(2)}| \leq (\log N)^\delta$, where $\delta > 0$ is any given constant. Let*

$$S_1 := \frac{m_1^{(1)}}{q^{k_1^{(1)}+1}} + \cdots + \frac{m_{h_1}^{(1)}}{q^{k_{h_1}^{(1)}+1}}$$

and

$$S_2 := \frac{m_{1,1}^{(2)}}{\alpha^{k_1^{(2)}}} + \frac{m_{1,2}^{(2)}}{\alpha^{k_1^{(2)}+1}} + \cdots + \frac{m_{h_2,1}^{(2)}}{\alpha^{k_{h_2}^{(2)}}} + \frac{m_{h_2,2}^{(2)}}{\alpha^{k_{h_2}^{(2)}+1}}.$$

Then, for

$$(\log N)^\eta \le k_1^{(1)} < k_2^{(1)} < \cdots < k_{h_1}^{(1)} \le \log_q N^r - (\log N)^\eta$$

and for

$$(\log N)^\eta \le k_1^{(2)} < k_2^{(2)} < \cdots < k_{h_2}^{(2)} \le \log_\alpha N^r - (\log N)^\eta$$

we uniformly have

$$\frac{q^{(\log N)^{\eta'}}}{N^r} \ll \left| c_1 S_1 + c_2 \frac{S_2}{\alpha+1} \right| \ll q^{-(\log N)^{\eta'}}$$

for all given $0 < \eta' < \eta$.

Proof. The upper bound is trivial. Thus, we concentrate on the lower bound. We have, with (5.1) and $\alpha^k(\alpha+1) = G_k\alpha + G_{k-1}$,

$$S := c_1 S_1 + c_2 \frac{S_2}{\alpha+1} = \frac{c_1 \hat{m}^{(1)}}{q^{k_{h_1}^{(1)}+1}} + \frac{c_2\left(\hat{m}_1^{(2)}\alpha + \hat{m}_2^{(2)}\right)}{G_{k_{h_2}^{(2)}+1}\alpha + G_{k_{h_2}^{(2)}}}$$

with integers $\hat{m}^{(1)}, \hat{m}_1^{(2)}, \hat{m}_2^{(2)}$ and therefore $S = 0$ if and only if the equations

$$\begin{aligned} c_1 \hat{m}^{(1)} G_{k_{h_2}^{(2)}+1} + c_2 \hat{m}_1^{(2)} q^{k_{h_1}^{(1)}+1} &= 0 \\ c_1 \hat{m}^{(1)} G_{k_{h_2}^{(2)}} + c_2 \hat{m}_2^{(2)} q^{k_{h_1}^{(1)}+1} &= 0 \end{aligned}$$

hold. Since $(G_k, G_{k+1}) = 1$ for all k, we obtain $q^{k_{h_1}^{(1)}+1} | c_1 \hat{m}^{(1)}$ and hence $q | \hat{m}^{(1)}$ (for sufficiently large $k_{h_1}^{(1)}$) which is not possible for $m_{h_1}^{(1)} \not\equiv 0 \bmod q$.

Hence we may assume $S \ne 0$. In order to get a lower bound for S, we use Baker's theorem (see [21]) saying that for non-zero algebraic numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ and integers $b_1, b_2, \ldots, b_n$ we have either

$$\alpha_1^{b_1} \cdots \alpha_n^{b_n} = 1$$

or

$$\left| \alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1 \right| \ge \exp(-U),$$

where

$$U = 2^{6n+32} n^{3n+6} d^{n+2} (1 + \log d)(\log B + \log d) \log A_1 \cdots \log A_n$$

with $d = [\mathbb{Q}(\alpha_1 \ldots, \alpha_n) : \mathbb{Q}]$,

$$B = \max\{2, |b_1|, |b_2|, \ldots, |b_n|\}.$$

and real numbers $A_1, A_2, \dots, A_n \geq e$ with $\log A_j \geq h(\alpha_j)$, where $h(\cdot)$ denotes the absolute logarithmic height.

Set $\varepsilon = \eta/(h_1 + h_2 - 1)$. Then there exists an integer K with $0 \leq K \leq h_1 + h_2 - 2$ such that for all j, ℓ

$$k_{j+1}^{(\ell)} - k_j^{(\ell)} \notin \left[(\log N)^{K\varepsilon}, (\log N)^{(K+1)\varepsilon}\right).$$

So fix K with this property. First suppose $k_{j+1}^{(\ell)} - k_j^{(\ell)} < (\log N)^{K\varepsilon}$ for all j, ℓ. Then we have $\log |\hat{m}^{(1)}| \ll (\log N)^{K\varepsilon}$, $\log |\hat{m}_i^{(2)}| \ll (\log N)^{K\varepsilon}$ and we can apply Baker's theorem for $r = 6$ with $\alpha_1 = q$, $\alpha_2 = \alpha$, $\alpha_3 = \hat{m}^{(1)}$, $\alpha_4 = \hat{m}_1^{(2)}\alpha + \hat{m}_2^{(2)}$, $\alpha_5 = -c_1$, $\alpha_6 = c_2/(\alpha+1)$ and $b_1 = -k_{h_1}^{(1)} - 1$, $b_2 = k_{h_2}^{(2)} + 1$, $b_3 = b_5 = 1$, $b_4 = b_6 = -1$ and obtain

$$\left| \frac{-c_1 \hat{m}^{(1)} (\alpha+1) \alpha^{k_{h_2}^{(2)}+1}}{c_2 (\hat{m}_1^{(2)}\alpha + \hat{m}_2^{(2)}) q^{k_{h_1}^{(1)}+1}} - 1 \right| \geq e^{-C \log(\max(k_{h_1}^{(1)}, k_{h_2}^{(2)})) \log|\hat{m}^{(1)}| \log(|\hat{m}_1^{(2)}| + |\hat{m}_2^{(2)}|)}$$

for a certain constant $C > 0$. Of course, this implies

$$|S| \geq \max\left(q^{-k_{h_1}^{(1)}}, \alpha^{-k_{h_2}^{(2)}}\right) e^{-c \log\log N \, (\log N)^{K\varepsilon}} \geq \frac{(\log N)^{\tau}}{N^{\tau}}$$

for some constant $c > 0$ and all $\tau > 0$.

Otherwise we have some s_1, s_2 such that $k_{j+1}^{(\ell)} - k_j^{(\ell)} < (\log N)^{K\varepsilon}$ for all $j < s_\ell$ and $k_{s_\ell+1}^{(\ell)} - k_{s_\ell}^{(\ell)} \geq (\log N)^{(K+1)\varepsilon}$. Here we get by Baker's theorem, as above,

$$\begin{aligned} \overline{S} &:= c_1 \left(\frac{m_1^{(1)}}{q^{k_1^{(1)}+1}} + \cdots + \frac{m_{s_1}^{(1)}}{q^{k_{s_1}^{(1)}+1}} \right) + \frac{c_2}{\alpha+1} \left(\frac{m_{1,1}^{(2)}}{\alpha^{k_1^{(2)}+1}} + \cdots + \frac{m_{s_2,2}^{(2)}}{\alpha^{k_{s_2}^{(2)}}} \right) \\ &\geq \max\left(q^{-k_{s_1}^{(1)}}, \alpha^{-k_{s_2}^{(2)}}\right) e^{-c \log\log N \, (\log N)^{K\varepsilon}} \end{aligned}$$

and can estimate $S - \overline{S}$ by

$$|S - \overline{S}| \ll (\log N)^{\delta} \left(q^{-k_{s_1}^{(1)} - (\log N)^{(K+1)\varepsilon}} + \alpha^{-k_{s_2}^{(2)} - (\log N)^{(K+1)\varepsilon}} \right).$$

Hence we have

$$\begin{aligned} |S| \geq \max\left(q^{-k_{s_1}^{(1)}}, \alpha^{-k_{s_2}^{(2)}}\right) \Big(& e^{-c \log\log N \, (\log N)^{K\varepsilon}} \\ & - \mathcal{O}\left((\log N)^{\delta} e^{-\log(\min(q,\alpha)) (\log N)^{(K+1)\varepsilon}} \right) \Big) \geq \frac{(\log N)^{\tau}}{N^{\tau}}. \end{aligned}$$

□

Lemma 20. *Let $a_1, a_2 \geq 1$ be two integers such that $\sqrt{\frac{a_1^2+4}{a_2^2+4}}$ is irrational, let $G = (G_j)$ and $H = (H_j)$ denote the corresponding second order recurrent sequences and c_1, c_2, r be positive integers.*

For arbitrary (but fixed) integers h_1, h_2 let $m_{i,j}^{(\ell)}$ $(1 \leq j \leq h_\ell,\ j, \ell \in \{1,2\})$ be integers satisfying $|m_{i,j}^{(\ell)}| \leq (\log N)^\delta$ (where $\delta > 0$ is any given constant) such that

$$S_1 := \frac{m_{1,1}^{(1)}}{\alpha^{k_1^{(1)}}} + \frac{m_{1,2}^{(1)}}{\alpha^{k_1^{(1)}+1}} + \cdots + \frac{m_{h_2,1}^{(1)}}{\alpha^{k_{h_1}^{(1)}}} + \frac{m_{h_1,2}^{(1)}}{\alpha^{k_{h_1}^{(1)}+1}} \neq 0$$

and

$$S_2 := \frac{m_{1,1}^{(2)}}{\alpha^{k_1^{(2)}}} + \frac{m_{1,2}^{(2)}}{\alpha^{k_1^{(2)}+1}} + \cdots + \frac{m_{h_2,1}^{(2)}}{\alpha^{k_{h_2}^{(2)}}} + \frac{m_{h_2,2}^{(2)}}{\alpha^{k_{h_2}^{(2)}+1}} \neq 0$$

Then, for

$$(\log N)^\eta \leq k_1^{(\ell)} < k_2^{(\ell)} < \cdots < k_{h_1}^{(\ell)} \leq \log_{\alpha_\ell} N^r - (\log N)^\eta$$

we uniformly have

$$\frac{\alpha^{(\log N)^{\eta'}}}{N^r} \ll \left| c_1 \frac{S_1}{\alpha_1 + 1} + c_2 \frac{S_2}{\alpha_2 + 1} \right| \ll \alpha^{-(\log N)^{\eta'}}$$

for all given $0 < \eta' < \eta$, where $\alpha = \max\{\alpha_1, \alpha_2\}$.

Proof. Again we can concentrate on the lower bound and have

$$S := c_1 \frac{S_1}{\alpha_1 + 1} + c_2 \frac{S_2}{\alpha_2 + 1} = \frac{c_1 \left(\hat{m}_1^{(1)} \alpha_1 + \hat{m}_2^{(1)} \right)}{G_{k_{h_1}^{(1)}+1} \alpha_1 + G_{k_{h_1}^{(1)}}} + \frac{c_2 \left(\hat{m}_1^{(2)} \alpha_2 + \hat{m}_2^{(2)} \right)}{H_{k_{h_2}^{(2)}+1} \alpha_2 + H_{k_{h_2}^{(2)}}}$$

The assumption that $\sqrt{\frac{a_1^2+4}{a_2^2+4}}$ is irrational ensures $\alpha_2 \notin \mathbb{Q}(\alpha_1)$. Hence S is zero if and only if the equations

$$\begin{aligned}
c_1 \hat{m}_1^{(1)} H_{k_{h_2}^{(2)}+1} + c_2 \hat{m}_1^{(2)} G_{k_{h_1}^{(1)}+1} &= 0 \\
c_1 \hat{m}_1^{(1)} H_{k_{h_2}^{(2)}} + c_2 \hat{m}_2^{(2)} G_{k_{h_1}^{(1)}+1} &= 0 \\
c_1 \hat{m}_2^{(1)} H_{k_{h_2}^{(2)}+1} + c_2 \hat{m}_1^{(2)} G_{k_{h_1}^{(1)}} &= 0 \\
c_1 \hat{m}_2^{(1)} H_{k_{h_2}^{(2)}} + c_2 \hat{m}_2^{(2)} G_{k_{h_1}^{(1)}} &= 0
\end{aligned}$$

hold. Then we must have e.g.

$$\hat{m}_1^{(1)} = -\hat{m}_1^{(2)} \frac{c_2}{c_1} \frac{G_{k_{h_1}^{(1)}+1}}{H_{k_{h_2}^{(2)}+1}} = \hat{m}_2^{(1)} \frac{G_{k_{h_1}^{(1)}+1}}{G_{k_{h_1}^{(1)}}}$$

and $G_{k_{h_1}^{(1)}+1}|\hat{m}_1^{(1)}$ because of $(G_{k_{h_1}^{(1)}+1}, G_{k_{h_1}^{(1)}}) = 1$. With $|m_{i,j}^{(\ell)}| \le (\log N)^\delta$ we get $\hat{m}_1^{(1)} = 0$ and thus $\hat{m}_2^{(1)} = \hat{m}_1^{(2)} = \hat{m}_2^{(2)} = S_1 = S_2 = 0$.

Hence $S \neq 0$ and the lower bound is obtained similarly to Lemma 19. □

Acknowledgement. The authors are grateful to an anonymous referee for his careful reading of a previous version of this paper and for many valuable suggestions to improve the presentation and the proofs.

References

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

[2] J. COQUET, *Corrélation de suites arithmétiques.* Sémin. Delange-Pisot-Poitou, 20e Année 1978/79, Exp. 15, 12 p. (1980).

[3] H. DELANGE, *Sur les fonctions q-additives ou q-multiplicatives.* Acta Arith. **21** (1972), 285–298.

[4] R. L. DOBRUŠIN, *Central limit theorem for nonstationary Markov chains* II. Theory Prob. Applications **1** (1956), 329–383. (Translated from: Teor. Vareojatnost. i Primenen. **1** (1956), 365–425.)

[5] M. DRMOTA, *The distribution of patterns in digital expansions.* In: Algebraic Number Theory and Diophantine Analysis (F. Halter-Koch and R. F. Tichy eds.), de Gruyter, Berlin, 2000, 103–121.

[6] M. DRMOTA, *The joint distribution of q-additive functions.* Acta Arith. **100** (2001), 17–39.

[7] M. DRMOTA, *Irregularities of Distributions with Respect to Polytopes.* Mathematika **43** (1996), 108–119.

[8] M. DRMOTA, M. FUCHS, E. MANSTAVIČIUS, *Functional Limit Theorems for Digital Expansions.* Acta Math. Hung., to appear,

[9] M. DRMOTA, R. F. TICHY, *Sequences, Discrepancies and Applications.* Lecture Notes in Mathematics **1651**, Springer Verlag, Berlin, 1998.

[10] J. M. DUMONT, A. THOMAS, *Systèmes de numération et fonctions fractales relatifs aux substitutions.* J. Theoret. Comput. Sci. **65** (1989), 153–169.

[11] J. M. DUMONT, A. THOMAS, *Gaussian asymptotic properties of the sum-of-digits functions.* J. Number Th. **62** (1997), 19–38.

[12] G. FARINOLE, *Représentation des nombres réels sur la base du nombre d'or, Application aux nombres de Fibonacci.* Prix Fermat Junior 1999, Quadrature **39** (2000).

[13] P. GRABNER, R. F. TICHY, *α-expansions, linear recurrences and the sum-of-digits function.* Manuscripta Math. **70** (1991), 311–324.

[14] L. K. HUA, *Additive Theory of Prime Numbers.* Translations of Mathematical Monographs Vol. 13, Am. Math. Soc., Providence, 1965.

[15] B. A. LIFŠIC, *On the convergence of moments in the central limit theorem for non-homogeneous Markov chains.* Theory Prob. Applications **20** (1975), 741–758. (Translated from: Teor. Vareojatnost. i Primenen. **20** (1975), 755–772.)

[16] E. MANSTAVIČIUS, *Probabilistic theory of additive functions related to systems of numerations.* Analytic and Probabilistic Methods in Number Theory, VSP, Utrecht 1997, 413–430.

[17] E. MANSTAVIČIUS, *Sums of digits obey the Strassen law.* In: Proceedings of the 38-th Conference of the Lithuanian Mathematical Society, R. Ciegis et al (Eds), Technika, Vilnius, 1997, 33–38.

[18] M. MENDÈS FRANCE, *Nombres normaux. Applications aux fonctions pseudo-aléatoires.* J. Analyse Math. **20** (1967) 1–56.

[19] A. RÉNYI, *Representations for real numbers and their ergodic properties.* Acta Math. Acad. Sci. Hung. 8 (1957), 477–493.

[20] I. M. VINOGRADOV, *The method of trigonometrical sums in the theory of numbers.* Interscience Publishers, London.

[21] M. WALDSCHMIDT, *Minorations de combinaisons linéaires de logarithmes de nombres algébriques*. Can. J. Math. **45** (1993), 176–224.

Michael DRMOTA, Wolfgang STEINER
Department of Geometry
Technische Universität Wien
Wiedner Hauptstraße 8-10/113
A-1040 Wien
Austria
E-mail : michael.drmota@tuwien.ac.at, steiner@geometrie.tuwien.ac.at