DISSERTATION

The Distribution of Digital Expansions on Polynomial Sequences

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der technischen Wissenschaften unter der Leitung von

Ao.Univ.Prof. Dipl.-Ing. Dr. Michael Drmota
Institut für Geometrie, E113

eingereicht an der Technischen Universität Wien
Fakultät für Technische Naturwissenschaften und Informatik

von

Wolfgang Steiner
Matr.Nr. 9425329
Premreinergasse 19/2/5, 1130 Wien

Wien, am 12. Februar 2002 ....

Kurzfassung

Ziel der Dissertation ist es, die Verteilung der Ziffernsummenfunktion (und ähnlicher Funktionen) auf polynomiellen Folgen von natürlichen Zahlen und Primzahlen zu studieren.

Von Bassily und Kátai [2] wurde ein zentraler Grenzwertsatz für die Verteilung der Folgen $f(P(n))$, $n \in \mathbb{N}$, und $f(P(p))$, $p \in \mathbb{P}$, gezeigt, wobei $f(n)$ eine q-additive Funktion und $P(n)$ ein Polynom mit ganzzahligen Koeffizienten (und positivem führenden Koeffizienten) ist. Da die Ideen in ihrem Beweis grundlegend für die Beweise aller anderen Sätze in dieser Dissertation sind, wird der Beweis in Kapitel 1 präsentiert.

Kapitel 2 behandelt gemeinsame Verteilungen mehrerer q-additiver Funktionen. Drmota [11] hat Bassily und Kátais Ergebnis auf die gemeinsame Verteilung von Folgen $f_\ell(P_\ell(n))$ (beziehungsweise $f_\ell(P_\ell(p))$) verallgemeinert, wobei die $f_\ell(n)$ q_ℓ-additive Funktionen sind und die Grade der Polynome $P_\ell(n)$ alle verschieden sein müssen. Der Beweis ist relativ kurz und wird in Sektion 2.1 geführt. Im Fall, dass der Grad der Polynome gleich ist, konnte Drmota nur für zwei Folgen $f_1(P_1(n))$ und $f_2(P_2(n))$ einen zentralen Grenzwertsatz beweisen, wobei q_1 und q_2 teilerfremd und die Polynome $P_1(n), P_2(n)$ linear sein müssen. Diese Ergebnisse können leicht auf beliebige Polynome $P_1(n), P_2(n)$ und Primzahlfolgen erweitert werden, indem Resultate Vinogradovs und Huas über Exponentialsummen polynomieller Folgen adaptiert werden (siehe Lemmata 1.2, 1.3, 2.5 und 2.6). Theorem 2.3 erweitert diese Ergebnisse außerdem auf multiplikativ unabhängige q_1, q_2.

Für stark q-additive Funktionen zu einem gemeinsamen q liefert Theorem 2.2 eine Charakterisierung der Verteilung beliebiger polynomieller Folgen. Es gilt immer ein zentraler Grenzwertsatz, wobei die Kovarianzmatrix nicht Diagonalgestalt hat, wenn zwei Polynomgrade identisch sind. Dieses Resultat ist auch für multiplikativ abhängige q_1, q_2 anwendbar, da es dann natürliche Zahlen s_1, s_2 gibt, sodass $q_1^{s_2} = q_2^{s_1} = q$ gilt, und $f_1(n), f_2(n)$ daher stark q-addditive Funktionen sind.

Als wichtigen Spezialfall dieser Sätze erhalten wir für die Ziffernsummenfunktion $s_q(n)$

$$\frac{1}{N}\#\left\{n<N \,\middle|\, \frac{s_{q_1}(n)-\frac{q_1-1}{2}\log_{q_1}N}{\sqrt{\frac{q_1^2-1}{12}\log_{q_1}N}}<x_1, \frac{s_{q_2}(n)-\frac{q_2-1}{2}\log_{q_2}N}{\sqrt{\frac{q_2^2-1}{12}\log_{q_2}N}}<x_2\right\}$$

$$\to \begin{cases} \frac{1}{2\pi}\int\limits_{-\infty}^{x_1} e^{-t^2/2}dt \int\limits_{-\infty}^{x_2} e^{-t^2/2}dt & \text{wenn } q_1,q_2 \text{ multiplikativ unabhängig sind} \\ \frac{1}{2\pi\sqrt{1-C^2}}\int\limits_{-\infty}^{x_2}\int\limits_{-\infty}^{x_1} e^{-\frac{1}{2(1-C^2)}(t_1^2+t_2^2-2Ct_1t_2)}dt_1dt_2 & \text{sonst,} \end{cases}$$

wobei die Kovarianz durch

$$C=\frac{\tilde{q}+1}{\tilde{q}-1}\sqrt{\frac{(q_1-1)(q_2-1)}{s_1s_2(q_1+1)(q_2+1)}} \text{ für } q_1=\tilde{q}^{s_1}, q_2=\tilde{q}^{s_2}, (s_1,s_2)=1$$

gegeben ist. Dieses Ergebnis ist für nicht teilerfremde q_1,q_2 neu. Für Polynomfolgen und Primzahlen gelten ähnliche Aussagen.

Im Kapitel 3 werden G-additive Funktionen betrachtet, die von der G-adischen Entwicklung natürlicher Zahlen abhängen, wobei G eine durch eine lineare Rekursion erzeugte Folge natürlicher Zahlen ist. Das Hauptergebnis (Theorem 3.2) ist ein Analogon zu Bassily und Kátais Resultat. Ein großer Unterschied zu q-additiven Funktionen ist dabei, dass die Ziffern durch eine Markoffkette dargestellt werden statt durch eine Folge unabhängiger Zufallsvariablen. Außerdem ist für q-adische Entwicklungen die k-te Ziffer von n durch den Wert von $\{n/q^{k+1}\}$ bestimmt. Für G-adische Entwicklungen benötigen wir dazu Fliesen des Torus $\mathbb{T}^d$, wobei d der Grad der linearen Rekursion ist. Für $d=2$ sind diese Fliesen Rechtecke, für $d\geq 3$ hingegen haben sie fraktalen Rand, und es handelt sich dabei um Rauzyfraktale.

Die letzten Ergebnisse (Theoreme 3.3 und 3.4) betreffen die Unabhängigkeit verschiedener G-additiver Funktionen (und q-additiver Funktionen), die unter ähnlichen Bedingungen wie für q-additive Funktionen gezeigt werden kann.

Abstract

The aim of this thesis is to study the distribution of the sum-of-digits function (and similar functions) on polynomial sequences of integers and primes.

Bassily and Kátai [2] proved a central limit theorem for the distribution of sequences $f(P(n))$, $n \in \mathbb{N}$, and $f(P(p))$, $p \in \mathbb{P}$, where $f(n)$ is a q-additive function and $P(n)$ an arbitrary polynomial with integer coefficients (and positive leading term). Since the ideas in their proof are fundamental for the proofs of all other theorems in this thesis, the proof is presented in Chapter 1.

Chapter 2 deals with joint distributions of several q-additive functions. Drmota [11] generalised Bassily and Kátai's result on the joint distribution of sequences $f_\ell(P_\ell(n))$ (and $f_\ell(P_\ell(p))$ respectively), where the f_ℓ are q_ℓ-additive functions and the $P_\ell(n)$ polynomials with different degrees. The proof is rather short and can be found in Section 2.1. For polynomials with equal degrees, Drmota could prove a central limit theorem only for two sequences $f_1(P_1(n)), f_2(P_2(n))$ with coprime q_1, q_2 and linear polynomials $P_1(n), P_2(n)$. By adapting results on exponential sums of polynomial sequences of Vinogradov and Hua (see Lemmata 1.2, 1.3, 2.5 and 2.6), this result can be easily extended to arbitrary polynomials $P_1(n), P_2(n)$ and sequences of primes. Theorem 2.3 extends this result to multiplicatively independent q_1, q_2.

For strongly q-additive functions with respect to the same q, a characterisation for the distribution of arbitrary polynomial sequences is given by Theorem 2.2. We always have a central limit theorem, but the covariance matrix is not diagonal, if any two degrees of the polynomials are equal. This result can also be used for multiplicatively dependent q_1, q_2. Then we have positive integers s_1, s_2 such that $q_1^{s_2} = q_2^{s_1} = q$ and $f_1(n), f_2(n)$ are therefore strongly q-additive functions.

In particular, we obtain for the sum-of-digits function $s_q(n)$

$$\frac{1}{N}\#\left\{n<N \,\middle|\, \frac{s_{q_1}(n)-\frac{q_1-1}{2}\log_{q_1}N}{\sqrt{\frac{q_1^2-1}{12}\log_{q_1}N}}<x_1, \frac{s_{q_2}(n)-\frac{q_2-1}{2}\log_{q_2}N}{\sqrt{\frac{q_2^2-1}{12}\log_{q_2}N}}<x_2\right\}$$
$$\to \begin{cases} \frac{1}{2\pi}\int\limits_{-\infty}^{x_1} e^{-t^2/2}dt \int\limits_{-\infty}^{x_2} e^{-t^2/2}dt & \text{if } q_1, q_2 \text{ are multiplicatively independent} \\ \frac{1}{2\pi\sqrt{1-C^2}} \int\limits_{-\infty}^{x_2}\int\limits_{-\infty}^{x_1} e^{-\frac{1}{2(1-C^2)}(t_1^2+t_2^2-2Ct_1t_2)}dt_1dt_2 & \text{else,} \end{cases}$$

where the covariance is given by

$$C=\frac{\tilde{q}+1}{\tilde{q}-1}\sqrt{\frac{(q_1-1)(q_2-1)}{s_1s_2(q_1+1)(q_2+1)}} \text{ for } q_1=\tilde{q}^{s_1}, q_2=\tilde{q}^{s_2}, (s_1,s_2)=1.$$

For q_1, q_2 which are not coprime, this result is new. Similar statements hold for polynomial sequences of integers and primes.

In Chapter 3, G-additive functions are considered, which depend on G-ary expansions of integers, where G is a sequence of integers generated by a linear recurrence. The main result (Theorem 3.2) is an analogue to Bassily and Kátai's result. An important difference to q-ary expansions is that the digits are represented by a Markov chain instead of a sequence of independent random variables. Furthermore, for q-ary expansions the value of the k-th digit of n is determined by the value of $\{n/q^{k+1}\}$. For G-ary expansions, we need tilings of the torus $\mathbb{T}^d$, where d is the degree of the linear recurrence, to obtain a similar characterisation. For $d=2$, these tilings are rectangles, whereas for $d\geq 3$ they have fractal boundary. More precisely, they are Rauzy fractals.

The last results (Theorems 3.3 and 3.4) deal with the independence of joint distributions of G-additive functions (and q-additive functions) which can be proved under similar conditions as for q-additive functions.

Contents

Chapter 1

Introduction

The purpose of this chapter is to present a work of Bassily and Kátai ([2]) on the distribution of the values of q-additive functions on polynomial sequences and some other results on this topic. Since the ideas in Bassily and Kátai's work will be a main ingredient of the proofs of all other theorems, we recall them in Section 1.3. Given that their proof is very succinct (and at one point wrong), we will not stick to their words and notation.

First we have to define q-additive functions.

1.1 q-ary expansions and q-additive functions

Let $q > 1$ be a given integer. Then every non-negative integer n has a unique *q-ary expansion*

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

with $\epsilon_{k,q}(n) \in \{0, 1, \ldots, q-1\}$, where we will omit the index q when there is no risk of confusion. Then the *sum-of-digits function* is given by

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

The sum-of-digits function is a special case of a *q-additive function*, i.e. a real-valued function f defined on the non-negative integers which satisfies $f(0) = 0$ and

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

Such a function is said to be *strongly q-additive*, if

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

1.2 Some results on the distribution of q-additive functions

We start with the mean value of the sum-of-digits function. The first asymptotic formula is due to Bush [5]. After some other works on this topic, Delange [8] proved

$$\frac{1}{N}\sum_{n<N} s_q(n) = \frac{q-1}{2}\log_q N + \gamma(\log_q N), \tag{1.1}$$

where γ is a continuous, nowhere differentiable and periodic function with period 1. Higher moments of $s_q(n)$ were considered for example by Grabner, Kirschenhofer, Prodinger and Tichy [22].

The most general result concerning the mean value of q-additive functions is due to Manstavičius [27] (extending earlier work of Coquet [6]). Let

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

and

$$M(N) = \sum_{k=0}^{[\log_q N]} \mu_k, \qquad B(N)^2 = \sum_{k=0}^{[\log_q N]} \mu_{2;k}^2.$$

Then

$$\frac{1}{N}\sum_{n<N} (f(n) - M(N))^2 \le cB(N)^2,$$

which implies

$$\frac{1}{N}\sum_{n<N} f(n) = M(N) + \mathcal{O}\left(B(N)\right).$$

Now we turn to distributional results for q-additive functions. Delange [7] proved an analogue to the Erdős-Wintner theorem. There exists a distribution function $F(x)$ such that, as $N \to \infty$

$$\frac{1}{N}\#\left\{n < N \,|\, f(n) < x\right\} \to F(x) \tag{1.2}$$

if and only if the two series $\sum_{k\ge 0}\mu_k$, $\sum_{k\ge 0}\mu_{2;k}^2$ converge. This theorem was generalised by Kátai [25] who proved that there exists a distribution function $F(x)$ such that, as $N \to \infty$,

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

if and only if the series $\sum_{k\ge 0}\mu_{2;k}^2$ converges.

The most general known theorem concerning a central limit theorem is again due to Manstavičius [27]. Suppose that, as $N \to \infty$,

$$\max_{bq^j < N} |f(bq^j)| = o(B(N))$$

and that $D(N) \to \infty$, where

$$D(N)^2 = \sum_{k=0}^{[\log_q N]} \sigma_k^2 \quad \text{and} \quad \sigma_k^2 = \frac{1}{q}\sum_{b=0}^{q-1} f(bq^k)^2 - m_k^2.$$

Then, as $N \to \infty$,

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

where $\Phi(x)$ is the normal distribution function.

Finally we turn to polynomial sequences and Bassily and Kátai [2].

Theorem 1.1. *Let f be a q-additive function such that $f(bq^k) = \mathcal{O}(1)$ as $k \to \infty$ for all $b \in \{0, 1, \ldots, q-1\}$. Assume $\frac{D(N)}{(\log N)^\eta} \to \infty$ 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 $N \to \infty$,*

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

and

$$\frac{1}{\pi(N)}\#\left\{p \in \mathbb{P}, p < N \,\middle|\, \frac{f(P(p)) - M(N^r)}{D(N^r)} < x\right\} \to \Phi(x), \tag{1.4}$$

where $\Phi(x)$ denotes the distribution function of the normal law.

Note that this theorem was only stated for $\eta = \frac{1}{3}$. However, a short inspection of the proof shows that $\eta > 0$ is sufficient.

Corollary 1.1. *Let $P(n)$ be a polynomial with integer coefficients, degree r and positive leading term. Then, as $N \to \infty$,*

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \frac{s_q(P(n)) - \frac{q-1}{2} r \log_q N}{\sqrt{\frac{q^2-1}{12} r \log_q N}} < x\right\} \to \Phi(x)$$

and

$$\frac{1}{\pi(N)}\#\left\{p < N \,\middle|\, \frac{s_q(P(p)) - \frac{q-1}{2} r \log_q N}{\sqrt{\frac{q^2-1}{12} r \log_q N}} < x\right\} \to \Phi(x)$$

1.3 Bassily and Kátai's proof

The main idea is to compare the moments of $\frac{X_N - \mathbf{E}\, X_N}{\sqrt{\mathbf{Var} X_N}}$ and $\frac{Y_{N^r} - \mathbf{E}\, Y_{N^r}}{\sqrt{\mathbf{Var} Y_{N^r}}}$, where X_N and Y_N are defined by

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

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

and to apply the Fréchet-Shohat theorem.

If we define random variables $\xi_{k,N}$ by

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

then

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

i.e. Y_N is a weighted sum of $\xi_{k,N}$. For $N = q^j$, we have $\mathbf{Pr}[\xi_{k,q^j} = b] = \frac{1}{q}$ (if $k < j$) and $(\xi_{k,q^j})_{0 \le k < q}$ is a chain of (identically distributed) independent random variables. Hence Y_{q^j} is a sum of independent random variables.

For arbitrary N, we have $\mathbf{Pr}[\xi_{k,N} = b] = \frac{1}{q} + \mathcal{O}\left(\frac{q^k}{N}\right)$. Lemma 1.1 will allow us to restrict to the truncated function

$$\overline{f}^{(N)} = \sum_{k=A(N)}^{B(N)} f_k(\epsilon_k(n)) \text{ with } A(N) = [(\log N)^\eta],\ B(N) = [\log_q N] - [(\log N)^\eta]$$

for some $\eta > 0$. In the range $A(N) \le k \le B(N)$, we have

$$\mathbf{Pr}[\xi_{k,N} = b] = \frac{1}{q} + \mathcal{O}\left(q^{-(\log N)^\eta}\right)$$

and, for an arbitrary number h of k_i,

$$\mathbf{Pr}[\xi_{k_1,N} = b_1, \ldots, \xi_{k_h,N} = b_h] = \frac{1}{q^h} + \mathcal{O}\left(q^{-(\log N)^\eta}\right).$$

This means that $\overline{Y}_N$ is a sum of asymptotically independent random variables.

Analogously to $\overline{f}^{(N)}$, we define

$$\overline{M}(N) = \sum_{k=A(N)}^{B(N)} \mu_k, \quad \overline{D}(N)^2 = \sum_{k=A(N)}^{B(N)} \sigma_k^2 \quad \text{and} \quad \overline{Y}_N = \sum_{k=A(N)}^{B(N)} f_k(\xi_{k,N}).$$

Because of $f_k(b) = \mathcal{O}(1)$, we have

$$\overline{M}(N) - M(N) = \mathcal{O}\left((\log N)^\eta\right) \quad \text{and} \quad \overline{D}(N)^2 - D(N)^2 = \mathcal{O}\left((\log N)^\eta\right).$$

With these definitions, we can state the following lemma.

Lemma 1.1. *Assume* $\frac{D(N)}{(\log N)^{\eta}} \to \infty$ *for some* $\eta > 0$ *and let* $P(n)$ *be a polynomial with integer coefficients, degree* r *and positive leading term. Then we have*

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

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

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

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

Furthermore, if for all $h \geq 0$

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

then we also have

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

and conversely.

Proof. From the condition stated for $D(N)$, we get $\frac{\overline{D}(N^r)}{D(N^r)} \to 1$ and

$$\max_{n<N}\left|\frac{(f(P(n)) - M(N^r)) - (\overline{f}^{(N^r)}(P(n)) - \overline{M}(N^r))}{\overline{D}(N^r)}\right| \to 0 \tag{1.5}$$

as $N \to \infty$. Therefore we have, for fixed $x \geq 0$,

$$\begin{aligned}
&\frac{1}{N}\#\left\{n < N \,\middle|\, \frac{f(P(n)) - M(N^r)}{D(N^r)} < x\right\}\\
&\qquad \leq \frac{1}{N}\#\left\{n < N \,\middle|\, \frac{f(P(n)) - M(N^r)}{\overline{D}(N^r)} < x(1+\delta)\right\}\\
&\qquad\quad \leq \frac{1}{N}\#\left\{n < N \,\middle|\, \frac{\overline{f}^{(N^r)}(P(n)) - \overline{M}(N^r)}{\overline{D}(N^r)} < x + \delta x + \delta\right\}
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{N}\#\left\{n < N \,\middle|\, \frac{f(P(n)) - M(N^r)}{D(N^r)} < x\right\}\\
&\qquad\quad \geq \frac{1}{N}\#\left\{n < N \,\middle|\, \frac{\overline{f}^{(N^r)}(P(n)) - \overline{M}(N^r)}{\overline{D}(N^r)} < x - \delta x - \delta\right\}
\end{aligned}$$

for all $\delta > 0$ and $N \ge N_1$ (for some N_1). Since the limit function is continuous, one direction of the equivalence is proved for $x \ge 0$. The case $x < 0$ and the converse implication can be proved similarly.

Now suppose that X, Y are two random variables with $|X - Y| \le \kappa \le 1$ then

$$\mathbf{E}\,|X|^h \le \sum_{l=0}^{h} \binom{h}{l} \mathbf{E}\,|Y|^l \kappa^{h-l}$$

and conversely. Hence, if $\mathbf{E}Y^h$ exists for some even h then $\mathbf{E}\,|Y|^l$ exist, too, for all $l \le h$ and consequently $\mathbf{E}\,X^h \le \mathbf{E}\,Y^h + \mathcal{O}\,(\kappa)$. In the same way, we get the converse inequality and we can obviously extend this property for odd h.

In order to complete the proof of Lemma 1.1, we just have to apply this observation to

$$X = \frac{f(P(n)) - M(N^r)}{\overline{D}(N^r)} \quad \text{and} \quad Y = \frac{\overline{f}^{(N^r)}(P(n)) - \overline{M}(N^r)}{\overline{D}(N^r)}.$$

We also use the fact $|X - Y| \le \kappa \to 0$ (cf. (1.5)) and the property $\frac{\overline{D}(N^r)}{D(N^r)} \to 1$. $\square$

Remark 1.1. Lemma 1.1 is stated for polynomial sequences of all integers $n < N$. Clearly the corresponding statements for primes hold too.

Remark 1.2. Bassily and Kátai [2] used the approximation

$$\max_{n \le N} \left| \frac{f(n) - M(N)}{D(N)} - \frac{\overline{f}^{(N)}(n) - \overline{M}(N)}{\overline{D}(N)} \right| \to 0$$

as $N \to \infty$ (cf. (1.5)), but this is wrong in general and the sum-of-digits function provides a counterexample. Lemma 1.1 corrects their proof.

Since the $\overline{Y}_N$ are sums of asymptotically independent random variables, they satisfy a central limit theorem with convergence of moments and the variance is asymptotically $\mathbf{Var}\,\overline{Y}_N \sim \overline{D}(N)^2$. Hence

$$\frac{\overline{Y}_N - \overline{M}(N)}{\overline{D}(N)} \to \mathcal{N}(0,1)$$

and, for all $h \ge 0$,

$$\left(\frac{\overline{Y}_N - \overline{M}(N)}{\overline{D}(N)}\right)^h \to \int_{-\infty}^{\infty} x^h \, d\Phi(x).$$

It remains to compare the moments of $\overline{X}_N$ to those of $\overline{Y}_{N^r}$, where $\overline{X}_N$ is defined by

$$\mathbf{Pr}[\overline{X}_N \le x] = \frac{1}{N} \# \left\{ n < N : \overline{f}^{(N^r)}(P(n)) \le x \right\}.$$

We have

$$\frac{1}{N}\sum_{n<N}\left(\frac{\overline{f}^{(N^r)}(P(n))-\overline{M}(N^r)}{\overline{D}(N^r)}\right)^h=\frac{1}{N}\sum_{n<N}\left(\frac{\sum_{k=A(N^r)}^{B(N^r)}\left(f_k(\epsilon_k(P(n)))-\mu_k\right)}{\overline{D}(N^r)}\right)^h$$

$$=\frac{1}{N}\sum_{n<N}\sum_{A(N^r)\le k_1,\ldots,k_h\le B(N^r)}\prod_{j=1}^{h}\frac{f_{k_j}(\epsilon_{k_j}(P(n)))-\mu_{k_j}}{\overline{D}(N^r)}$$

$$=\prod_{j=1}^{h}\sum_{k_j=A(N^r)}^{B(N^r)}\sum_{b_j=0}^{q-1}\frac{f_{k_j}(b_j)-\mu_{k_j}}{\overline{D}(N^r)}\frac{1}{N}\#\{n<N\mid \epsilon_{k_1}(P(n))=b_1,\ldots,\epsilon_{k_h}(P(n))=b_h\}.$$

For primes, we get similarly

$$\frac{1}{\pi(N)}\sum_{p<N}\left(\frac{\overline{f}^{(N^r)}(P(p))-\overline{M}(N^r)}{\overline{D}(N^r)}\right)^h$$
$$=\prod_{j=1}^{h}\sum_{k_j=A(N^r)}^{B(N^r)}\sum_{b_j=0}^{q-1}\frac{f_{k_j}(b_j)-\mu_{k_j}}{\overline{D}(N^r)}\frac{1}{\pi(N)}\#\{p<N:\epsilon_{k_1}(P(p))=b_1,\ldots,\epsilon_{k_h}(P(p))=b_h\}$$

and the moments of $\overline{Y}(N^r)$ are

$$\frac{1}{N^r}\sum_{n<N^r}\left(\frac{\overline{f}^{(N)}(n)-\overline{M}(N^r)}{\overline{D}(N^r)}\right)^h$$
$$=\prod_{j=1}^{h}\sum_{k_j=A(N^r)}^{B(N^r)}\sum_{b_j=0}^{q-1}\frac{f_{k_j}(b_j)-\mu_{k_j}}{\overline{D}(N^r)}\frac{1}{N^r}\#\{n<N^r:\epsilon_{k_1}(n)=b_1,\ldots,\epsilon_{k_h}(n)=b_h\}$$

The next proposition assures that these moments converge to the same limit. This is the essential and most difficult part of the proof. Note that it suffices to consider different k_j, because for $k_i=k_j$ and $b_i\neq b_j$ obviously the numbers are zero and for $k_i=k_j, b_i=b_j$ just make h smaller.

Proposition 1.1. *Let $P(n)$ be a polynomial with integer coefficients, degree r and positive leading term. Then, for every $h\geq 1$ and for every $\lambda>0$, we have*

$$\frac{1}{N}\#\{n<N|\epsilon_{k_1}(P(n))=b_1,\ldots,\epsilon_{k_h}(P(n))=b_h\}=\frac{1}{q^h}+\mathcal{O}\left((\log N)^{-\lambda}\right) \tag{1.6}$$

and

$$\frac{1}{\pi(N)}\#\{p<N|\epsilon_{k_1}(P(p))=b_1,\ldots,\epsilon_{k_h}(P(p))=b_h\}=\frac{1}{q^h}+\mathcal{O}\left((\log N)^{-\lambda}\right) \tag{1.7}$$

uniformly for all integers

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

and $b_1, b_2, \ldots, b_h \in \{0, 1, \ldots, q-1\}$.

The proof of Proposition 1.1 uses the next three lemmata. The first two can be found in Hua [24].

Lemma 1.2. *Let* $P(n)$ *be a polynomial of degree* r *of the form*

$$P(n) = \frac{a}{b} n^r + \gamma_1 n^{r-1} + \cdots + \gamma_r$$

with $\gcd(a,b) = 1$. *Let* τ *be a positive number satisfying*

$$\tau \ge \left(2^k + 1\right) \tau_0 + 2^{3(k-2)}$$

and

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

Then, as $N \to \infty$,

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

where $e(x) = \exp(2\pi i x)$.

Lemma 1.3. *Let* $P(n)$ *be as in Lemma 1.2 and* τ_0, τ *arbitrary positive numbers satisfying*

$$\tau \ge 2^{6k} \tau_0$$

and

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

Then, as $N \to \infty$,

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

Lemma 1.4. *Let* $0 < \Delta < 1$ *and*

$$U_{b,q,\Delta} = [0, \Delta] \cup \bigcup_{b=1}^{q-1} \left[\frac{b}{q} - \Delta, \frac{b}{q} + \Delta\right] \cup [1-\Delta, 1].$$

Then, for every $\eta > 0$ *and arbitrary* $\lambda > 0$, *we have uniformly for* $(\log N)^\eta \le k \le [\log_q N^r] - (\log N)^\eta$ *and* $0 < \Delta < 1/(2q)$, *as* $N \to \infty$,

$$\frac{1}{N} \# \left\{ n < N \,\middle|\, \left\{\frac{P(n)}{q^{k+1}}\right\} \in U_{b,q,\Delta} \right\} \ll \Delta + (\log N)^{-\lambda}$$

and

$$\frac{1}{\pi(N)} \# \left\{ p < N \,\middle|\, \left\{\frac{P(p)}{q^{k+1}}\right\} \in U_{b,q,\Delta} \right\} \ll \Delta + (\log N)^{-\lambda},$$

where $\{x\}$ *denotes the fractional part of* x.

Proof. We use the inequality of Erdős-Turán: The discrepancy of the real numbers $x_1, \dots, x_N \bmod 1$

$$D_N = \sup \left| \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{[\alpha,\beta]}(\{x_n\}) - (\beta - \alpha) \right|,$$

where the supremum is taken over intervals $[\alpha, \beta] \subseteq [0, 1]$ and $\mathbf{1}_{[\alpha,\beta]}$ is the characteristic function of $[\alpha, \beta]$, can be estimated by

$$D_N \ll \frac{1}{M} + \sum_{m=1}^{M} \frac{1}{m} \left| \frac{1}{N} \sum_{n=1}^{N} e(mx_n) \right| \tag{1.8}$$

uniformly for all positive integers M (cf. [14], Theorem 1.21).

$U_{b,q,\Delta}$ is the union of $q+1$ subintervals, its measure is $2q\Delta$. Let $M = (\log N)^{\lambda+1}$ and apply (1.8) to the sequences $x_n = \frac{P(n-1)}{q^{k+1}}$ for each subinterval of $U_{b,q,\Delta}$. The conditions of Lemmata 1.2 and 1.3 clearly hold for the polynomials $\frac{hP(n)}{q^{k+1}}$. This gives the stated inequalities. □

Proof of Proposition 1.1. Let $\sum_{m\in\mathbb{Z}} c_{m,b,q} e(mx)$ be the Fourier series of $\mathbf{1}_{[\frac{b}{q},\frac{b+1}{q}]}$, i.e.

$$c_{0,b,q} = \frac{1}{q}, \quad c_{m,b,q} = \frac{e\left(-\frac{mb}{q}\right) - e\left(-\frac{m(b+1)}{q}\right)}{2\pi i m} \text{ for } m \neq 0. \tag{1.9}$$

Let $\psi_{b,q,\Delta}(x)$ be defined by

$$\psi_{b,q,\Delta}(x) = \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} \mathbf{1}_{[\frac{b}{q},\frac{b+1}{q}]}(\{x+z\})\, dz.$$

The Fourier coefficients of $\psi_{b,q,\Delta}(x)$ are $d_{0,b,q,\Delta} = \frac{1}{q}$ and

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

for $m \neq 0$. Note that $d_{m,b,q,\Delta} = 0$ if $m \neq 0$ and $m \equiv 0 \bmod q$ and that

$$|d_{m,b,q,\Delta}| \le \min\left(\frac{1}{\pi|m|}, \frac{1}{\Delta \pi m^2}\right). \tag{1.10}$$

By definition, we have $0 \le \psi_{b,q,\Delta}(x) \le 1$ and

$$\psi_{b,q,\Delta}(x) = \begin{cases} 1 & \text{if } x \in \left[\frac{b}{q} + \Delta, \frac{b+1}{q} - \Delta\right], \\ 0 & \text{if } x \in [0,1] \setminus \left[\frac{b}{q} - \Delta, \frac{b+1}{q} + \Delta\right]. \end{cases}$$

If we set

$$t(x) = \psi_{b_1,q,\Delta}\left(\frac{x}{q^{k_1+1}}\right)\cdots\psi_{b_h,q,\Delta}\left(\frac{x}{q^{k_h+1}}\right),$$

then we get for $\Delta < 1/(2q)$

$$\left|\#\{n < N | \epsilon_{k_1}(P(n)) = b_1, \ldots, \epsilon_{k_h}(P(n)) = b_h\} - \sum_{n<N} t(P(n))\right| \leq$$

$$\leq \sum_{j=1}^{h} \#\left\{n < N \left| \left\{\frac{P(n)}{q^{k_j+1}}\right\} \in U_{b_j,q,\Delta}\right.\right\} \ll \Delta N + N(\log N)^{-\lambda}$$

For convenience, denote by $\mathcal{M}$ the set of integer vectors $\mathbf{m} = (m_1, \ldots, m_h)$ and set $\mathbf{v} = \left(q^{-k_1-1}, \ldots, q^{-k_h-1}\right)$, $T_{\mathbf{m}} = d_{m_1,b_1,q,\Delta} \cdots d_{m_h,b_h,q,\Delta}$. Then $t(x)$ has Fourier series expansion

$$t(x) = \sum_{\mathbf{m}\in\mathcal{M}} T_{\mathbf{m}} e(\mathbf{m}\cdot\mathbf{v}x)$$

and

$$\sum_{n<N} t(P(n)) = \sum_{\mathbf{m}\in\mathcal{M}} T_{\mathbf{m}} \sum_{n\leq N} e\left(\frac{A_{\mathbf{m}}}{H_{\mathbf{m}}}P(n)\right).$$

We check that Lemma 1.2 can be applied to the polynomials $\frac{A_{\mathbf{m}}}{H_{\mathbf{m}}}P(n)$: We can omit those $\mathbf{m}$ for which there is a j such that $q|m_j$, $m_j \neq 0$, since $d_{m_j,b_j,q,\Delta} = 0$ implies $T_{\mathbf{m}} = 0$. Let $q = p_1^{e_1}\cdots p_s^{e_s}$ and assume $p_t^{e_t} \not| m_h$. Then we have $p_t^{k_h e_t} | H_{\mathbf{m}}$ because of

$$H_{\mathbf{m}}(m_h + q^{k_h-k_{h-1}} m_{h-1} + \cdots + m_1^{k_h-k_1}) = A_{\mathbf{m}} q^{k_h+1}.$$

Thus there exists an $\kappa > 0$ depending only on q, such that $H_{\mathbf{m}} \geq q^{\kappa k_h}$. We can prove similarly $H_{\mathbf{m}} \geq q^{\kappa k_s}$ if $q \not| m_s$ and $m_{s+1} = \cdots = m_h = 0$.

Hence Lemma 1.2 can be applied if $\mathbf{m} \neq \mathbf{0}$ and we obtain

$$\frac{1}{N}\#\{n < N|\epsilon_{k_1}(P(n)) = b_1, \ldots, \epsilon_{k_h}(P(n)) = b_h\}$$
$$= \frac{1}{q^h} + \mathcal{O}\left((\log N)^{-\tau_0}\sum_{\mathbf{m}\neq\mathbf{0}} |T_{\mathbf{m}}|\right) + \mathcal{O}\left(\Delta + (\log N)^{-\lambda}\right)$$

The main term $1/q^h$ comes from the choice $\mathbf{m} = \mathbf{0}$. From (1.10), we obtain

$$\sum_{\mathbf{m}\neq\mathbf{0}} |T_{\mathbf{m}}| \leq \left(\frac{1}{q} + 2\sum_{m=1}^{\infty}\min\left(\frac{1}{\pi m}, \frac{1}{\pi\Delta m^2}\right)\right)^h \ll \left(\log\frac{1}{\Delta}\right)^h.$$

Let $\Delta = (\log N)^{-\lambda}$ and $\tau_0 > \lambda$. Then (1.6) follows immediately from the above relation and Lemma 1.4. (1.7) can be shown by the same arguments with Lemma 1.3 instead of Lemma 1.2. $\square$

Thus we have proved

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

and

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

for all $h \geq 0$. The Fréchet-Shohat theorem (see e.g. Billingsley [3], p. 390) implies (1.3) and (1.4).

Chapter 2

Joint Distributions

In this chapter, we generalise Theorem 1.1 to the joint distribution of different polynomials and different q-additive functions with possibly different q.

We are able to prove a central limit theorem for sequences $f_\ell(P_\ell(n))$ (and $f_\ell(P_\ell(p))$ respectively), where f_ℓ are q_ℓ-additive functions, if all $P_\ell(n)$ have different degrees (Theorem 2.1, Section 2.1), if all q_ℓ are equal and the f_ℓ are strongly q-additive functions (Theorem 2.2, Section 2.2) and for two sequences, if q_1 and q_2 are multiplicatively independent (Theorem 2.3, Section 2.3).

2.1 Polynomials of different degrees

First we prove the following theorem due to Drmota [11].

Theorem 2.1. *Let f_ℓ, $1 \le \ell \le d$, be q_ℓ-additive functions such that $f_\ell(bq_\ell^k) = \mathcal{O}(1)$ as $k \to \infty$ for all $b \in \{0, \ldots, q_\ell - 1\}$. Assume that $\frac{D_\ell(N)}{(\log N)^\eta} \to \infty$ as $N \to \infty$ for some $\eta > 0$. Let $P_\ell(n)$ be polynomials of different degrees r_ℓ with integer coefficients and positive leading terms. Then, as $N \to \infty$,*

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \frac{f_\ell(P_\ell(n)) - M_\ell(N^{r_\ell})}{D_\ell(N^{r_\ell})} < x_\ell, \ell = 1, 2, \ldots, d\right\} \to \Phi(x_1)\ldots\Phi(x_d)$$

and

$$\frac{1}{\pi(N)}\#\left\{p < N \,\middle|\, \frac{f_\ell(P_\ell(p)) - M_\ell(N^{r_\ell})}{D_\ell(N^{r_\ell})} < x_\ell, \ell = 1, \ldots, d\right\} \to \Phi(x_1)\ldots\Phi(x_d)$$

Remark 2.1. Drmota stated this theorem for pairwise coprime q_ℓ but did not use this assumption in the proof. Thus it is not necessary.

As already mentioned, Theorem 2.1 is a direct generalisation of Bassily and Kátai [2]. It turns out that it suffices to prove the following lemma, which states the asymptotic independence of all digits. Then the independence of the distributions is an easy corollary.

Proposition 2.1 (cf. Proposition 1.1). *Let $P_\ell(n)$, $1 \le \ell \le d$, be polynomials of different degrees r_ℓ with integer coefficients and positive leading terms. Let $\lambda > 0$ be an arbitrary constant and h_ℓ, $1 \le \ell \le d$, non-negative integers. Then, as $N \to \infty$,*

$$\frac{1}{N} \# \left\{ n < N \,\middle|\, \epsilon_{q_\ell, k_j^{(\ell)}}(P_\ell(n)) = b_j^{(\ell)}, 1 \le j \le h_\ell, 1 \le \ell \le d \right\} = \frac{1}{q_1^{h_1} q_2^{h_2} \cdots q_d^{h_d}} + \mathcal{O}\left((\log N)^{-\lambda} \right) \quad (2.1)$$

and

$$\frac{1}{\pi(N)} \# \left\{ p < N \,\middle|\, \epsilon_{q_\ell, k_j^{(\ell)}}(P_\ell(n)) = b_j^{(\ell)}, 1 \le j \le h_\ell, 1 \le \ell \le d \right\} = \frac{1}{q_1^{h_1} q_2^{h_2} \cdots q_d^{h_d}} + \mathcal{O}\left((\log N)^{-\lambda} \right) \quad (2.2)$$

uniformly for integers

$$(\log N^{r_\ell})^\eta \le k_1^{(\ell)} < k_2^{(\ell)} < \cdots < k_{h_\ell}^{(\ell)} \le \log_{q_\ell} N^{r_\ell} - (\log N^{r_\ell})^\eta \quad (1 \le \ell \le d)$$

(with some $\eta > 0$) and $b_j^{(\ell)} \in \{0, 1, \ldots, q_\ell - 1\}$.

Proof. We follow the proof of Lemma 1.1 and point out only the differences. Set

$$t(n_1, \ldots, n_d) = \prod_{\ell=1}^{d} \prod_{j=1}^{h_\ell} \psi_{b_j^{(\ell)}, q_\ell, \Delta} \left(\frac{n_\ell}{q_\ell^{k_j^{(\ell)}+1}} \right).$$

Then we get, for $\Delta < 1/(2q)$,

$$\left| \# \left\{ n < N \,\middle|\, \epsilon_{q_\ell, k_j^{(\ell)}}(P_\ell(n)) = b_j^{(\ell)} \text{ for all } j, \ell \right\} - \sum_{n<N} t(P_1(n), \ldots, P_d(n)) \right|$$
$$\le \sum_{\ell=1}^{d} \sum_{j=1}^{h_\ell} \# \left\{ n < N \,\middle|\, \left\{ \frac{P_\ell(n)}{q_\ell^{k_j^{(\ell)}+1}} \right\} \in U_{b_j^{(\ell)}, q_\ell, \Delta} \right\} \ll \Delta N + N (\log N)^{-\lambda}$$

Set $h = (h_1 + \cdots + h_d)$ and denote by $\mathcal{M}$ the set of h-dimensional integer vectors $\mathbf{M} = (\mathbf{m}_1, \ldots, \mathbf{m}_d)$ with $\mathbf{m}_\ell = (m_1^{(\ell)}, \ldots, m_{h_\ell}^{(\ell)})$, $1 \le \ell \le d$. Furthermore, set

$$T_{\mathbf{M}} = \prod_{\ell=1}^{d} \prod_{j=1}^{h_\ell} d_{m_j^{(\ell)}, b_j^{(\ell)}, q_\ell, \Delta}.$$

With $\mathbf{v}_\ell = (q_\ell^{-k_1^{(\ell)}-1}, \dots, q_\ell^{-k_{h_\ell}^{(\ell)}-1})$, $t(x_1, \dots, x_d)$ has Fourier series expansion

$$t(x_1, \dots, x_d) = \sum_{\mathbf{M} \in \mathcal{M}} T_{\mathbf{M}} e\left(\mathbf{m}_1 \cdot \mathbf{v}_1 x_1 + \dots + \mathbf{m}_d \cdot \mathbf{v}_d x_d\right).$$

Thus we are led to consider the exponential sums

$$\sum_{\mathbf{M} \in \mathcal{M}} T_{\mathbf{M}} \sum_{n<N} e\left(\mathbf{m}_1 \cdot \mathbf{v}_1 P_1(n) + \dots + \mathbf{m}_d \cdot \mathbf{v}_d P_d(n)\right).$$

$\mathbf{m}_1 = \dots = \mathbf{m}_d = \mathbf{0}$ provides the leading term $1/(q_1^{h_1} \cdots q_d^{h_d})$. If there exists ℓ and j with $m_j^{(\ell)} \neq 0$ and $m_j^{(\ell)} \equiv 0 \bmod q_\ell$, then $T_{\mathbf{M}} = 0$. So it remains to consider the case where there exists ℓ and j with $m_j^{(\ell)} \not\equiv 0 \bmod q_\ell$. Here the exponent is of the form

$$\mathbf{m}_1 \cdot \mathbf{v}_1 P_1(n) + \dots + \mathbf{m}_d \cdot \mathbf{v}_d P_d(n) = \frac{A_{\mathbf{m}_1}}{H_{\mathbf{m}_1}} P_1(n) + \dots + \frac{A_{\mathbf{m}_d}}{H_{\mathbf{m}_d}} P_d(n),$$

with at least one $A_{\mathbf{m}_\ell} \neq 0$. Let $P_\ell(n)$ be the (unique) polynomial with maximal degree r_ℓ such that $A_{\mathbf{m}_\ell} \neq 0$. Then

$$\mathbf{m}_1 \cdot \mathbf{v}_1 P_1(n) + \dots + \mathbf{m}_d \cdot \mathbf{v}_d P_d(n) = \frac{A_{\mathbf{m}_\ell}}{H_{\mathbf{m}_\ell}} P_\ell(n) + \gamma_1 n^{r_\ell - 1} + \dots + \gamma_{r_\ell}.$$

By the same arguments as in Lemma 1.1, we can therefore apply Lemma 1.2 and obtain

$$\frac{1}{N} \# \left\{ n < N \,\middle|\, \epsilon_{q_\ell, k_j^{(\ell)}}(P(n)) = b_j^{(\ell)}, 0 \le j \le h_\ell, 1 \le \ell \le d \right\}$$
$$= \frac{1}{q_1^{h_1} q_2^{h_2} \cdots q_d^{h_d}} + \mathcal{O}\left((\log N)^{-\tau_0} \sum_{\mathbf{M} \in \mathcal{M} \setminus \mathbf{0}} |T_{\mathbf{M}}| \right) + \mathcal{O}\left(\Delta + (\log N)^{-\lambda}\right).$$

With

$$\sum_{\mathbf{M} \in \mathcal{M} \setminus \mathbf{0}} |T_{\mathbf{M}}| \le (2 + 2\log(1/\Delta))^{h_1 + \dots + h_d},$$

(2.1) is proved.

The proof of (2.2) runs along the same lines. □

Corollary 2.1. *With the definitions of Proposition 2.1, as $N \to \infty$,*

$$\frac{1}{N} \# \left\{ n < N \,\middle|\, \epsilon_{q_\ell, k_j^{(\ell)}}(P_\ell(n)) = b_j^{(\ell)}, 1 \le j \le h_\ell, 1 \le \ell \le d \right\}$$
$$= \prod_{\ell=1}^{d} \left(\frac{1}{N} \# \left\{ n < N \,\middle|\, \epsilon_{q_\ell, k_j^{(\ell)}}(P_\ell(n)) = b_j^{(\ell)}, 1 \le j \le h_\ell \right\} \right) + \mathcal{O}\left((\log N)^{-\lambda}\right)$$

and

$$\frac{1}{\pi(N)}\#\left\{p<N\,\middle|\,\epsilon_{q_\ell,k_j^{(\ell)}}(P_\ell(p))=b_j^{(\ell)}, 1\le j\le h_\ell, 1\le \ell\le d\right\}$$
$$=\prod_{\ell=1}^{d}\left(\frac{1}{\pi(N)}\#\left\{p<N\,\middle|\,\epsilon_{q_\ell,k_j^{(\ell)}}(P_\ell(p))=b_j^{(\ell)}, 1\le j\le h_\ell\right\}\right)+\mathcal{O}\left((\log N)^{-\lambda}\right)$$

uniformly for integers

$$(\log N^{r_\ell})^\eta \le k_1^{(\ell)}, k_2^{(\ell)}, \ldots, k_{h_\ell}^{(\ell)} \le \log_{q_\ell} N^{r_\ell} - (\log N^{r_\ell})^\eta \quad (1\le \ell\le d)$$

(with some $\eta>0$) and $b_j^{(\ell)}\in\{0,1,\ldots,q_\ell-1\}$.

Proof. If there exists ℓ and j_1, j_2 with $k_{j_1}^{(\ell)}=k_{j_2}^{(\ell)}$ but $b_{j_1}^{(\ell)}\ne b_{j_2}^{(\ell)}$ then both sides are zero.

So it remains to consider the case, where for every ℓ the integers $k_j^{(\ell)}$, $1\le j\le h_\ell$, are different and without loss of generality we can assume that they are increasing. Hence we can directly apply Proposition 2.1. □

Corollary 2.2. *For any choice of non-negative integers h_ℓ, $1\le\ell\le d$, we have, as $N\to\infty$,*

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

and

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

Proof. We have

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

$$=\sum_{A_1(N^{r_1})\le k_1^{(1)},\dots,k_{h_1}^{(1)}\le B_1(N^{r_1})}\cdots\sum_{A_d(N^{r_d})\le k_1^{(d)},\dots,k_{h_d}^{(d)}\le B_d(N^{r_d})}$$

$$\frac{1}{N}\sum_{n<N}\prod_{\ell=1}^{d}\prod_{j=1}^{h_\ell}\frac{f_{\ell,k_j^{(\ell)}}\left(\epsilon_{q_\ell,k_j^{(\ell)}}(P(n))\right)-\mu_{\ell,k_j^{(\ell)}}}{\overline{D}_\ell(N^{r_\ell})}$$

$$=\sum_{A_1(N^{r_1})\le k_1^{(1)},\dots,k_{h_1}^{(1)}\le B_1(N^{r_1})}\cdots\sum_{A_d(N^{r_d})\le k_1^{(d)},\dots,k_{h_d}^{(d)}\le B_d(N^{r_d})}\sum_{b_1^{(1)}=0}^{q_1-1}\cdots\sum_{b_{h_d}^{(d)}=0}^{q_d-1}$$

$$\frac{1}{N}\#\left\{n<N\,\middle|\,\epsilon_{q_\ell,k_j^{(\ell)}}(P_\ell(n))=b_j^{(\ell)}\text{ for all }j,\ell\right\}\prod_{\ell=1}^{d}\prod_{j=1}^{h_\ell}\frac{f_{\ell,k_j^{(\ell)}}(b_j^{(\ell)})-\mu_{\ell,k_j^{(\ell)}}}{\overline{D}_\ell(N^{r_\ell})}$$

and

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

$$=\prod_{\ell=1}^{d}\left(\sum_{A_\ell(N^{r_\ell})\le k_1^{(\ell)},\dots,k_{h_\ell}^{(\ell)}\le B_\ell(N^{r_\ell})}\frac{1}{N}\sum_{n<N}\prod_{j=1}^{h_\ell}\frac{f_{\ell,k_j^{(\ell)}}\left(\epsilon_{q_\ell,k_j^{(\ell)}}(P(n))\right)-\mu_{\ell,k_j^{(\ell)}}}{\overline{D}_\ell(N^{r_\ell})}\right)$$

$$=\sum_{A_1(N^{r_1})\le k_1^{(1)},\dots,k_{h_1}^{(1)}\le B_1(N^{r_1})}\cdots\sum_{A_d(N^{r_d})\le k_1^{(d)},\dots,k_{h_d}^{(d)}\le B_d(N^{r_d})}\sum_{b_1^{(1)}=0}^{q_1-1}\cdots\sum_{b_{h_d}^{(d)}=0}^{q_d-1}$$

$$\prod_{\ell=1}^{d}\left(\frac{1}{N}\#\left\{n<N\,\middle|\,\epsilon_{q_\ell,k_j^{(\ell)}}(P_\ell(n))=b_j^{(\ell)}\text{ for all }j\right\}\prod_{j=1}^{h_\ell}\frac{f_{\ell,k_j^{(\ell)}}(b_j^{(\ell)})-\mu_{\ell,k_j^{(\ell)}}}{\overline{D}_\ell(N^{r_\ell})}\right)$$

By Corollary 2.1, the two terms are equal up to an error term of the form $\mathcal{O}\left((\log N)^{-\lambda+h-h\eta}\right)$. The result for primes is obtained analogously. □

By combining Theorem 1.1, Lemma 1.1 and Corollary 2.2, we obtain

$$\frac{1}{N}\sum_{n<N}\prod_{\ell=1}^{d}\left(\frac{f_\ell(P_\ell(n))-M_\ell(N^{r_\ell})}{D_\ell(N^{r_\ell})}\right)^{h_\ell}\to\int x_1^{h_1}\dots x_d^{h_d}d\Phi(x_1)\dots d\Phi(x_d),$$

$$\frac{1}{\pi(N)}\sum_{p<N}\prod_{\ell=1}^{d}\left(\frac{f_\ell(P_\ell(p))-M_\ell(N^{r_\ell})}{D_\ell(N^{r_\ell})}\right)^{h_\ell}\to\int x_1^{h_1}\dots x_d^{h_d}d\Phi(x_1)\dots d\Phi(x_d)$$

and the Fréchet-Shohat theorem implies the statements of Theorem 2.1.

2.2 Strongly q-additive functions with the same q

The next theorem is a generalisation of Theorem 1.1 for the case $q_1 = \ldots = q_d = q$ for polynomials of not necessarily different degrees. If the degrees of the polynomials are not different, we do not have asymptotic independence of all digits as in Corollary 2.1, but we can show that random vectors which represent the digits form a Markov chain (Subsection 2.2.2). Hence we obtain a central limit theorem for these random vectors and a comparison of the moments (Subsection 2.2.3) gives the central limit theorem for $f_\ell(P_\ell(n))$. For simplicity, we restrict to strongly q-additive functions.

2.2.1 Results

Theorem 2.2. *Let f_ℓ, $1 \le \ell \le d$, be strongly q-additive functions with $\sigma_\ell^2 = \sum\limits_{b=1}^{q-1} \frac{f_\ell(b)^2}{q} - \left(\sum\limits_{b=1}^{q-1} \frac{f_\ell(b)}{q} \right)^2 > 0$ and $P_\ell(n) = g_{r_\ell}^{(\ell)} n^{r_\ell} + \cdots + g_1^{(\ell)} n + g_0^{(\ell)}$ polynomials with integer coefficients and positive leading terms. Then, as $N \to \infty$,*

$$\frac{1}{N} \# \left\{ n < N \,\middle|\, \frac{f_\ell(P_\ell(n)) - M_\ell(N^{r_\ell})}{D_\ell(N^{r_\ell})} < x_\ell, \ell = 1, \ldots, d \right\} \to \Phi_V(x_1, \ldots, x_d) \tag{2.3}$$

and

$$\frac{1}{\pi(N)} \# \left\{ p < N \,\middle|\, \frac{f_\ell(P_\ell(p)) - M_\ell(N^{r_\ell})}{D_\ell(N^{r_\ell})} < x_\ell, \ell = 1, 2, \ldots, d \right\} \to \Phi_V(x_1, \ldots, x_d) \tag{2.4}$$

where $\Phi_V(x_1, \ldots, x_d)$ denotes the distribution function of the d-dimensional normal law with covariance matrix $V = (v_{i,j})_{1 \le i,j \le d}$ given by

$$v_{i,j} = \begin{cases} 1 & \text{if } i = j \\ C_{i,j}\left(\frac{g_{r_i}^{(i)}}{(g_{r_i}^{(i)}, g_{r_j}^{(j)})}, \frac{g_{r_j}^{(j)}}{(g_{r_i}^{(i)}, g_{r_j}^{(j)})} \right) & \text{if } g_{r_j}^{(j)} P_i(n) \equiv g_{r_i}^{(i)} P_j(n) \\ \frac{r_i - \max\left\{ s \,\middle|\, g_{r_i}^{(i)} g_s^{(j)} \neq g_{r_j}^{(j)} g_s^{(i)} \right\}}{r_i} C_{i,j}\left(\frac{g_{r_i}^{(i)}}{(g_{r_i}^{(i)}, g_{r_j}^{(j)})}, \frac{g_{r_j}^{(j)}}{(g_{r_i}^{(i)}, g_{r_j}^{(j)})} \right) & \text{if } r_i = r_j \\ 0 & \text{else,} \end{cases}$$

where

$$\begin{aligned} C_{i,j}(g_i, g_j) = &\frac{1}{\sigma_i \sigma_j} \sum_{l=0}^{R_j - 1} \sum_{b_i=1}^{q-1} \sum_{b_j=1}^{q-1} \left(\pi_{b_i, b_j, g_i q^l, g_j} - \frac{1}{q^2} \right) f_i(b_i) f_j(b_j) \\ &+ \frac{1}{\sigma_i \sigma_j} \sum_{l=1}^{R_i - 1} \sum_{b_i=1}^{q-1} \sum_{b_j=1}^{q-1} \left(\pi_{b_i, b_j, g_i, g_j q^l} - \frac{1}{q^2} \right) f_i(b_i) f_j(b_j) \end{aligned}$$

with R_ℓ such that $q|\frac{q^{R_\ell}}{(q^{R_\ell},g^{(\ell)}_{r_\ell})}$ and

$$\pi_{b_i,b_j,g_iq^l,g_j} = \pi_{b_i,b_j,g,g'} = \frac{1}{q^2} - \frac{\left(\overline{(b_i+1)g'} - \overline{b_ig'}\right)\left(\overline{(b_j+1)g} - \overline{b_jg}\right)}{gg'q^2}$$
$$+\frac{\min\left(\overline{b_ig'},\overline{b_jg}\right) + \min\left(\overline{(b_i+1)g'},\overline{(b_j+1)g}\right) - \min\left(\overline{(b_i+1)g'},\overline{b_jg}\right) - \min\left(\overline{b_ig'},\overline{b_jg}\right)}{gg'q}$$

where $g = \frac{g_iq^l}{(q^l,g_j)}$, $g' = \frac{g_j}{(q^l,g_j)}$ and $\overline{y}$ denotes the representative y' of $y' \equiv y\,(q)$ with $0 \le y' < q$. (π_{b_i,b_j,g_i,g_jq^l} is given symmetrically.)

Remark 2.2. If V is positive definite, we have, with $\mathbf{t} = (t_1,\dots,t_d)$,

$$\Phi_V(x_1,\dots,x_d) = \frac{1}{(2\pi)^{d/2}\sqrt{\det V}}\int_{-\infty}^{x_d}\dots\int_{-\infty}^{x_1} e^{-\frac{1}{2}\mathbf{t}V^{-1}\mathbf{t}^t}dt_1\dots dt_d.$$

Remark 2.3. If $g^{(\ell)}_{r_\ell}$ is coprime to q, then we have $R_\ell = 1$.

$l \ge R_j$ implies $\pi_{b_i,b_j,g_iq^l,g_j} = \frac{1}{q^2}$ for all b_i, b_j.

The π_{b_i,b_j,g_iq^l,g_j} are the joint probabilities of digits $k+l$ and k of g_in and g_jn (which do not depend on k):

$$\pi_{b_i,b_j,g_iq^l,g_j} = \mathbf{Pr}[\epsilon_k(g_iq^ln) = b_i, \epsilon_k(g_j) = b_j] = \mathbf{Pr}[\epsilon_{k+l}(g_in) = b_i, \epsilon_k(g_j) = b_j].$$

Note that we need $C_{i,j}(g_i,g_j)$ only for coprime g_i, g_j.

Remark 2.4. As in all other theorems, the constant term of the polynomials plays no role.

Corollary 2.3. *Let $P_\ell(n) = g^{(\ell)}_{r_\ell}n^{r_\ell} + \dots + g^{(\ell)}_1 n + g^{(\ell)}_0$ be polynomials with integer coefficients and positive leading terms. Then, as $N \to \infty$,*

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \frac{s_q(P_\ell(n)) - \frac{q-1}{2}\log_q N^{r_\ell}}{\sqrt{\frac{q^2-1}{12}\log_q N^{r_\ell}}} < x_\ell, \ell = 1,\dots,d\right\}$$
$$\to \frac{1}{(2\pi)^{d/2}\sqrt{\det V}}\int_{-\infty}^{x_d}\dots\int_{-\infty}^{x_1} e^{-\frac{1}{2}\mathbf{t}V^{-1}\mathbf{t}^t}dt_1\dots dt_d$$

with the positive definite matrix $V = (v_{i,j})_{1\le i,j\le d}$ given by

$$v_{i,j} = \begin{cases} 1 & \text{if } i = j \\ C_{i,j}\left(\frac{g^{(i)}_{r_i}}{(g^{(i)}_{r_i},g^{(j)}_{r_j})}, \frac{g^{(j)}_{r_j}}{(g^{(i)}_{r_i},g^{(j)}_{r_j})}\right) & \text{if } g^{(j)}_{r_j}P_i(n) \equiv g^{(i)}_{r_i}P_j(n) \\ \frac{r_i - \max\left\{s \,\middle|\, g^{(i)}_{r_i}g^{(j)}_s \ne g^{(j)}_{r_j}g^{(i)}_s\right\}}{r_i} C_{i,j}\left(\frac{g^{(i)}_{r_i}}{(g^{(i)}_{r_i},g^{(j)}_{r_j})}, \frac{g^{(j)}_{r_j}}{(g^{(i)}_{r_i},g^{(j)}_{r_j})}\right) & \text{if } r_i = r_j \\ 0 & \text{else,} \end{cases}$$

and

$$C_{i,j}(g_i,g_j) = \frac{q^2-(q,g_i)^2-(q,g_j)^2+1}{g_ig_j(q^2-1)} + \frac{1}{g_ig_j(q^2-1)}\left(\sum_{l=1}^{R_j-1}\frac{q^2-\left(q,\frac{g_iq^l}{(q^l,g_j)}\right)}{q^l}+\sum_{l=1}^{R_i-1}\frac{q^2-\left(q,\frac{g_jq^l}{(g_i,q^l)}\right)}{q^l}\right)$$

Remark 2.5. For monomials $P_\ell(n)=g_\ell n^r$ with $(g_\ell,q)=1$ we just have

$$v_{i,j}=\frac{(g_i,g_j)^2}{g_ig_j}.$$

For $q=2$ and $r=1$, this was proved by W.M. Schmidt [35]. Schmid [33] obtained a local limit law in this case.

Furthermore, we can calculate the joint distribution of the sum-of-digits functions for multiplicatively dependent q_1,q_2.

Corollary 2.4. *For $q_1=\tilde{q}^{s_1}, q_2=\tilde{q}^{s_2}$ with positive integers $\tilde{q},s_1,s_2$ and $(s_1,s_2)=1$, we have, as $N\to\infty$,*

$$\frac{1}{N}\#\left\{n<N \,\middle|\, \frac{s_{q_1}(n)-\frac{q_1-1}{2}\log_{q_1}N}{\sqrt{\frac{q_1^2-1}{12}\log_{q_1}N}}<x_1, \frac{s_{q_2}(n)-\frac{q_2-1}{2}\log_{q_2}N}{\sqrt{\frac{q_2^2-1}{12}\log_{q_2}N}}<x_2\right\} \to \frac{1}{2\pi\sqrt{1-C^2}}\int_{-\infty}^{x_2}\int_{-\infty}^{x_1}e^{-\frac{1}{2(1-C^2)}(t_1^2+t_2^2-2Ct_1t_2)}dt_1dt_2$$

with

$$C=\frac{\tilde{q}+1}{\tilde{q}-1}\sqrt{\frac{(q_1-1)(q_2-1)}{s_1s_2(q_1+1)(q_2+1)}}.$$

For general strongly q_ℓ-additive functions similar statements can easily be derived. The case of multiplicatively independent q_1,q_2 is treated in Section 2.3.

2.2.2 A Markov chain and calculation of the covariance

We define the polynomials

$$P_\ell^{(s)}(n)=g_{r_\ell}^{(\ell)}n^{r_\ell}+\cdots+g_s^{(\ell)}n^s \quad \text{for} \quad 1\le s\le r=\max_{1\le\ell\le d}r_\ell,$$

which will be needed in the next subsection. In this subsection we fix s.

Furthermore, we define the vectors

$$\mathbf{w}_k^{(s)}(n)=(w_{k,s},\ldots,w_{k,r})=\left(\left\{\frac{n^s}{q^{k+1}}\right\},\left\{\frac{n^{s+1}}{q^{k+1}}\right\},\ldots,\left\{\frac{n^r}{q^{k+1}}\right\}\right)$$

for $0 \le n < N$ and see, by Proposition 2.1, that they asymptotically form a net to the base q if $k \in \left[(\log N)^\eta, \log_q N^s - (\log N)^\eta\right]$ (but not for $k > \log_q N^s$). Proposition 2.1 gives rather bad error terms if we want to calculate the number of $\mathbf{w}_k^{(s)}(n)$ in an arbitrary set of $\mathbb{T}^{r-s+1}$. Nevertheless this will give reason to use the Lebesgue measure as probability measure on $\mathbb{T}^{r-s+1}$.

We have $\epsilon_k(P_\ell^{(s)}(n)) = b$ if and only if

$$\left\{g_{r_\ell}^{(l)} w_{k,r_\ell} + \cdots + g_s^{(l)} w_{k,s}\right\} \in \left[\frac{b}{q}, \frac{b+1}{q}\right).$$

This means that, for each digit b, $\{\mathbf{w}_k^{(s)}(n) \mid \epsilon_k(P_\ell^{(s)}(n)) = b\}$ (as a set of $\mathbb{T}^{r-s+1}$) is contained in the stripe $S_{b,\ell}^{(s)}$ between the hyperplanes $g_{r_\ell}^{(\ell)} x_{r_\ell} + \cdots + g_s^{(\ell)} x_s = \frac{b}{q}$ (included) and $g_{r_\ell}^{(\ell)} x_{r_\ell} + \cdots + g_s^{(\ell)} x_s = \frac{b+1}{q}$ (excluded). If $P_\ell^{(s)}(n) = 0$, set $S_{0,\ell}^{(s)} = \mathbb{T}^{r-s+1}$ and $S_{b,\ell}^{(s)} = \emptyset$ for $b \neq 0$.

Thus each set $\{\mathbf{w}_k^{(s)}(n) \mid \epsilon_k(P_1^{(s)}(n)) = b_1, \ldots, \epsilon_k^{(s)}(P_1(n)) = b_d\}$ is contained in $S_{b_1,1}^{(s)} \cap \cdots \cap S_{b_d,d}^{(s)}$ and each of this intersections consists of a finite number of convex sets, the boundaries of which are the above hyperplanes. Let $(W_j^{(s)})_{1 \le j \le \kappa_s}$ be the partition of $\mathbb{T}^r$ induced by these sets (or equivalently by the hyperplanes). Then $f_\ell|_{W_j^{(s)}}$ is constant for all ℓ, j.

Furthermore, we have $\epsilon_{k-j}(P_\ell^{(s)}(n)) = b$ if and only if $T^j(\mathbf{w}_k^{(s)}(n)) \in S_{b,\ell}^{(s)}$ with the map $T : \mathbb{T}^r \to \mathbb{T}^r$, $T(w_{k,s}, \ldots, w_{k,r}) = (qw_{k,s}, \ldots, qw_{k,r})$. Hence

$$\left\{n \,\middle|\, \epsilon_0(P_\ell^{(s)}(n)) = b_0^{(\ell)}, \ldots, \epsilon_k(P_\ell^{(s)}(n)) = b_k^{(\ell)}\right\}$$
$$= \left\{n \,\middle|\, \mathbf{w}_k^{(s)}(n) \in T^{-k} S_{b_0^{(\ell)},\ell}^{(s)}, \ldots, \mathbf{w}_k^{(s)}(n) \in S_{b_k^{(\ell)},\ell}^{(s)}\right\}$$

and we define a sequence of random variables $(Y_k^{(s)})_{k \ge 0}$ on $\{W_1^{(s)}, W_2^{(s)}, \ldots, W_{\kappa_s}^{(s)}\}$ by

$$\mathbf{Pr}[Y_0^{(s)} = W_{j_0}^{(s)}, \ldots, Y_k^{(s)} = W_{j_k}^{(s)}] = \lambda_{r-s+1}(T^{-k} W_{j_0}^{(s)} \cap \ldots T^{-1} W_{j_{k-1}}^{(s)} \cap W_{j_k}^{(s)})$$

for $1 \le j_i \le \kappa_s$, $0 \le i \le k$. (λ_n denotes the n-dimensional Lebesgue measure.)

Lemma 2.1. $\left(Y_k^{(s)}\right)_{k \ge 0}$ *is a Markov chain.*

Proof. It is easy to see that $T|_{W_j^{(s)}}$ is injective for $1 \le j \le \kappa_s$ and that $TW_j^{(s)}$ is the (disjoint) union of sets $W_i^{(s)}$, since the image of the hyperplane

$g_{r_\ell}^{(\ell)} x_{r_\ell} + \cdots + g_s^{(\ell)} x_s = \frac{b}{q}$ is the hyperplane $g_{r_\ell}^{(\ell)} x_{r_\ell} + \cdots + g_s^{(\ell)} x_s = 0$. Hence we have

$$\begin{aligned}
\mathbf{Pr}[Y_0^{(s)} = W_{j_0}^{(s)}, \ldots, Y_{k+1}^{(s)} = W_{j_{k+1}}^{(s)}] &= \lambda_{r-s+1}(T^{-(k+1)} W_{j_0}^{(s)} \cap \cdots \cap W_{j_{k+1}}^{(s)}) \\
&= \frac{1}{q^{r-s+1}} \lambda_{r-s+1}(T^{-k} W_{j_0}^{(s)} \cap \cdots \cap W_{j_k}^{(s)} \cap T W_{j_{k+1}}^{(s)}) \\
&= \begin{cases} \frac{1}{q^{r-s+1}} \lambda_{r-s+1}(T^{-k} W_{j_0}^{(s)} \cap \cdots \cap W_{j_k}^{(s)}) & \text{if } W_{j_k}^{(s)} \subseteq T W_{j_{k+1}}^{(s)} \\ 0 & \text{else.} \end{cases}
\end{aligned}$$

Thus

$$\begin{aligned}
&\mathbf{Pr}[Y_{k+1}^{(s)} = W_{j_{k+1}}^{(s)} | Y_0^{(s)} = W_{j_0}^{(s)}, \ldots, Y_k^{(s)} = W_{j_k}^{(s)}] \\
&\qquad = \begin{cases} \frac{1}{q^r} & \text{if } W_{j_k}^{(s)} \subseteq T W_{j_{k+1}}^{(s)} \\ 0 & \text{else} \end{cases} = \mathbf{Pr}[Y_{k+1}^{(s)} = W_{j_{k+1}}^{(s)} | Y_k^{(s)} = W_{j_k}^{(s)}],
\end{aligned}$$

i.e. the Markov chain property is fulfilled. □

As already noted, each f_ℓ is constant on each $W_j^{(s)}$ because of $W_j^{(s)} \subseteq S_{b_1,1}^{(s)} \cap \cdots \cap S_{b_d,d}$ for some b_i. Therefore we define the d-dimensional function f on $(W_j^{(s)})_{1 \le j \le \kappa_s}$ by

$$f(W_j^{(s)}) = \Big(f_1(W_j^{(s)}), \ldots, f_d(W_j^{(s)}) \Big) = (f_1(b_1), \ldots, f_d(b_d)).$$

Before stating a central limit theorem for $f(Y_k^{(s)})$, we study the covariance $\mathbf{Cov}(f_i(Y_{k_i}^{(s)}), f_j(Y_{k_j}^{(s)}))$. To this effect, the following lemma, which will be proved together with Proposition 2.3, will be very useful. Note that $Y_k^{(s)} \subseteq S_{b,\ell}^{(s)}$ is equivalent to $f_\ell(Y_k^{(s)}) = b$.

Lemma 2.2.

$$\mathbf{Pr}[Y_{k_i}^{(s)} \subseteq S_{b_i,i}^{(s)}, Y_{k_j}^{(s)} \subseteq S_{b_j,j}^{(s)}] = \sum_{m_i, m_j : \frac{m_i P_i^{(s)}(n)}{q^{k_i}} + \frac{m_j P_j^{(s)}(n)}{q^{k_j}} \equiv 0} c_{m_i,b_i,q} c_{m_j,b_j,q}, \tag{2.5}$$

where $c_{m_i,b_i,q}$ are the Fourier coefficients in (1.9).

By Lemma 2.2, we have

$$\mathbf{Pr}[Y_{k_i}^{(s)} \subseteq S_{b_i,i}^{(s)}, Y_{k_j}^{(s)} \subseteq S_{b_j,j}^{(s)}] = c_{0,b_i,q} c_{0,b_j,q} = \mathbf{Pr}[Y_{k_i}^{(s)} \subseteq S_{b_i,i}^{(s)}] \mathbf{Pr}[Y_{k_j}^{(s)} \subseteq S_{b_j,j}^{(s)}],$$

if the polynomials do not have the same degree or are not proportional and $\mathbf{Cov}(f_i(Y_{k_i}^{(s)}), f_j(Y_{k_j}^{(s)})) = 0$.

Now assume $r_i = r_j$ and that the polynomials are proportional. Furthermore, let w.l.o.g. $k_i \ge k_j$. Then the m_i in (2.5) must satisfy

$m_i g_r^{(i)} \equiv 0\,(q^{k_i-k_j})$, i.e. $m_i \equiv 0\left(\frac{q^{k_i-k_j}}{(q^{k_i-k_j},g_r^{(i)})}\right)$. If $k_i-k_j \ge R_i$, this implies $m_i \equiv 0\,(q)$. Hence $c_{m_i,b_i,q}c_{m_j,b_j,q}=0$ for $(m_i,m_j)\neq(0,0)$ and

$$\mathbf{Cov}\left(f_i(Y_{k_i}^{(s)}), f_j(Y_{k_j}^{(s)})\right) = 0 \quad \text{ if } k_i-k_j \ge R_i \text{ or } k_j-k_i \ge R_j.$$

(For $k_j \ge k_i$, we get the result by the symmetry of the covariance.)

Since the chain $(Y_k^{(s)})_{k\ge 0}$ is homogeneous, we obtain

$$\begin{aligned}
&\mathbf{Cov}\left(\sum_{k=A(N)}^{B(N)} f_i(Y_k^{(s)}), \sum_{k=A(N)}^{B(N)} f_j(Y_k^{(s)})\right)\\
&\qquad = \sum_{k=A(N)}^{B(N)} \sum_{l=\max(-R_i+1,A(N)-k)}^{\min(R_j-1,B(N)-k)} \mathbf{Cov}\left(f_i(Y_k^{(s)}), f_j(Y_{k+l}^{(s)})\right)\\
&\qquad = (B(N)-A(N)) \sum_{l=-R_i+1}^{R_j-1} \mathbf{Cov}\left(f_i(Y_k^{(s)}), f_j(Y_{k+l}^{(s)})\right) + \mathcal{O}\left(1\right).
\end{aligned}$$

Now we can state the central limit theorem.

Proposition 2.2. *The sums of the random variables $f(Y_k^{(s)})$ satisfy a multidimensional central limit theorem with convergence of moments. More precisely, we have, for all $\mathbf{a}=(a_1,\ldots,a_d)\in\mathbb{R}^d$, as $N\to\infty$,*

$$\frac{\sum_{k=A(N)}^{B(N)}\sum_{\ell=1}^{d} a_\ell f_\ell(Y_k^{(s)}) - \sum_{\ell=1}^{d} a_\ell \overline{M}_\ell(N)}{\sigma_\ell\sqrt{B(N)-A(N)}} \to \mathcal{N}\left(0, \mathbf{a}V^{(s)}\mathbf{a}^t\right), \tag{2.6}$$

where the covariance matrix $V^{(s)} = \left(v_{i,j}^{(s)}\right)_{1\le i,j\le d}$ is given by

$$v_{i,j}^{(s)} = \frac{1}{\sigma_i\sigma_j}\sum_{l=-R_i+1}^{R_j-1} \mathbf{Cov}\left(f_i(Y_k^{(s)}), f_j(Y_{k+l}^{(s)})\right)$$

and for all integers $h_\ell \ge 0$ we have

$$\mathbf{E}\prod_{\ell=1}^{d}\left(\frac{\sum_{k=A(N)}^{B(N)} f_\ell(Y_k^{(s)}) - \overline{M}_\ell(N)}{\overline{D}_\ell(N)}\right)^{h_\ell} \to \int x_1^{h_1}\cdots x_d^{h_d}\, d\Phi_{V^{(s)}}(x_1,\ldots,x_d). \tag{2.7}$$

Proof. We have

$$\begin{aligned}
\mathbf{Var}\sum_{\ell=1}^{d}\sum_{k=A(N)}^{B(N)} a_\ell f_\ell(Y_k^{(s)}) &= \sum_{i=1}^{d}\sum_{j=1}^{d}\mathbf{Cov}\left(\sum_{k=A(N)}^{B(N)} a_i f_i(Y_k^{(s)}), \sum_{k=A(N)}^{B(N)} a_j f_j(Y_k^{(s)})\right) \\
&= (B(N)-A(N))\sum_{i=1}^{d}\sum_{j=1}^{d} a_i a_j \sum_{l=-R_i+1}^{R_j-1}\mathbf{Cov}\left(f_i(Y_k^{(s)}), f_j(Y_{k+l}^{(s)})\right) + \mathcal{O}\left(1\right) \\
&= (B(N)-A(N))\sigma_i\sigma_j \mathbf{a}V^{(s)}\mathbf{a}^t + \mathcal{O}\left(1\right).
\end{aligned}$$

If $\mathbf{a}V^{(s)}\mathbf{a}^t = 0$, then we have $\sum_{\ell=1}^{d}\sum_{k=A(N)}^{B(N)} a_\ell f_\ell(Y_k^{(s)}) = \mathcal{O}\left(1\right)$ and both sides in (2.6) are zero.

Otherwise we use the central limit theorem for stationary and homogeneous Markov chains or φ-mixing sequences (see e.g. Billingsley [3], p. 364). We need that all states are recurrent and aperiodic. For $Y_k^{(s)}$, this condition is satisfied, since we clearly have an integer E such that $T^E W_j^{(s)} = \mathbb{T}^{r-s+1}$ for all $W_j^{(s)}$ and hence $\mathbf{Pr}[Y_{k+l}^{(s)} = W_{j_{k+l}}^{(s)} | Y_k^{(s)} = W_{j_k}^{(s)}] > 0$ for all $l \geq E$. This implies that $X_k = \sum_{\ell=1}^{d} a_\ell f_\ell(Y_k^{(s)})$ is φ-mixing too and the central limit theorem holds for X_k. (Note that X_k need not be a Markov chain, if $\sum_{\ell=1}^{d} a_\ell f_\ell$ is not injective.)

For the convergence of moments it suffices to show that they exist. The onedimensional moments $\mathbf{E}\left(\frac{\sum_{k=A(N)}^{B(N)} f_\ell(Y_k^{(s)}) - \overline{M}_\ell(N)}{\overline{D}_\ell(N)}\right)^{h_\ell}$ are just the moments of $\overline{f}_\ell^{(N)}(n)$ (cf. Section 1.3) and converge therefore. With the relation $\mathbf{E}\left|X_N^r \tilde{X}_N^s\right| \leq \left(\mathbf{E}\, X_N^{2r}\right)^{\frac{1}{2}}\left(\mathbf{E}\, \tilde{X}_N^{2s}\right)^{\frac{1}{2}}$ for all random variables $X_N, \tilde{X}_N$, we obtain the convergence of the multidimensional moments. □

For the calculation of $\mathbf{Cov}(f_i(Y_k^{(s)}), f_j(Y_j^{(s)}))$, it suffices to consider $Y_k = Y_k^{(1)}$ and linear polynomials because of Lemma 2.2 and the succeeding remarks. For the sum-of-digits function, we get explicit expressions.

Lemma 2.3. *Let $P_1(n) = g_1 n$, $P_2(n) = g_2 n$ and $f_1(n) = f_2(n)$ the sum-of-digits function. Then the covariance of $f_1(Y_k)$ and $f_2(Y_k)$ is given by*

$$\mathbf{Cov}(f_1(Y_k), f_2(Y_k)) = \frac{(q^2 - d_1^2 - d_2^2 + 1)(g_1, g_2)^2}{12 g_1 g_2} \tag{2.8}$$

where $d_1 = \left(q, \frac{g_1}{(g_1,g_2)}\right)$ and $d_2 = \left(q, \frac{g_2}{(g_1,g_2)}\right)$.

Proof. Because of Lemma 2.2, the digit probability does not change if we replace g_1, g_2 by $\frac{g_1}{(g_1,g_2)}, \frac{g_2}{(g_1,g_2)}$. Therefore we assume $(g_1, g_2) = 1$.

The covariance is given by

$$
\begin{aligned}
&\mathbf{Cov}(f_1(Y_k), f_2(Y_k)) \\
&= \sum_{b_1=0}^{q-1}\sum_{b_2=0}^{q-1} \mathbf{Pr}[\epsilon_k(g_1 n) = b_1, \epsilon_k(g_2 n) = b_2] b_1 b_2 - \mathbf{E}\, f_1(Y_k) \mathbf{E}\, f_2(Y_k). \quad (2.9)
\end{aligned}
$$

In order to get integer numbers, we define

$$
\begin{aligned}
a_{b_1,b_2} &= q g_1 g_2 \mathbf{Pr}[\epsilon_k(g_1 n) = b_1, \epsilon_k(g_2 n) = b_2] \\
&= \#\left\{ x \in \{0, 1, \ldots, q g_1 g_2 - 1\} \,\middle|\, \left[\frac{x}{g_2}\right] \equiv b_1\,(q), \left[\frac{x}{g_1}\right] \equiv b_2\,(q) \right\}.
\end{aligned}
$$

Because of

$$
\sum_{b_1=0}^{q-1}\sum_{b_2=0}^{q-1} a_{b_1,b_2} b_1 b_2 = \sum_{i=1}^{q-1}\sum_{b_1=q-i}^{q-1}\sum_{j=1}^{q-1}\sum_{b_2=q-j}^{q-1} a_{b_1,b_2}
$$

we study $A_{i,j} = \sum_{b_1=q-i}^{q-1}\sum_{b_2=q-j}^{q-1} a_{b_1,b_2}$.

For every x in the set corresponding to a_{b_1,b_2}, $(q g_1 g_2 - 1 - x)$ is in the set corresponding to $a_{q-1-b_1,q-1-b_2}$. Therefore we have $a_{b_1,b_2} = a_{q-1-b_1,q-1-b_2}$ and

$$
\begin{aligned}
A_{i,j} &= \sum_{b_1=0}^{i-1}\sum_{b_2=0}^{j-1} a_{b_1,b_2} \\
&= \#\{x \in \{0, \ldots, q g_1 g_2 - 1\} \mid x \equiv 0, \ldots, i g_2 - 1 (q g_2), x \equiv 0, \ldots, j g_1 - 1 (q g_1)\}
\end{aligned}
$$

Since $(q g_1, q g_2) = q$, the system of congruences $x \equiv x_1\,(q g_2)$ and $x \equiv x_2\,(q g_1)$ has no solution x if $x_1 \not\equiv x_2\,(q)$ and a unique solution modulo $q g_1 g_2$ for $x_1 \equiv x_2\,(q)$. Denote by $\overline{y}^{(q)}$ the representative y' of $y' \equiv y\,(q)$ with $0 \le y' < q$. Then

$$
\begin{aligned}
A_{i,j} &= i g_2 \frac{j g_1 - \overline{j g_1}^{(q)}}{q} + \overline{j g_1}^{(q)} \frac{i g_2 - \overline{i g_2}^{(q)}}{q} + \min(\overline{i g_2}^{(q)}, \overline{j g_1}^{(q)}) \\
&= \frac{i g_2 j g_1}{q} - \frac{\overline{i g_2}^{(q)}\overline{j g_1}^{(q)}}{q} + \min(\overline{i g_2}^{(q)}, \overline{j g_1}^{(q)}).
\end{aligned}
$$

Hence

$$
\sum_{i=1}^{q-1}\sum_{j=1}^{q-1} A_{i,j} = \frac{q(q-1)^2}{4} g_1 g_2 - \frac{q(q-d_1)(q-d_2)}{4} + d_1 d_2 \sum_{i=1}^{q''-1}\sum_{j=1}^{q'-1} \min(i d_2, j d_1),
$$

where $q' = q/d_1$ and $q'' = q/d_2$. We have

$$\begin{aligned}
&\sum_{i=1}^{q''-1}\sum_{j=1}^{q'-1}\min(id_2, jd_1)\\
&= \sum_{i=1}^{q''-1} id_2\left(q'-1-\left[\frac{id_2}{d_1}\right]\right) + \sum_{j=1}^{q'-1} jd_1\left(q''-1-\left[\frac{jd_1}{d_2}\right]\right) + \sum_{i=1}^{\frac{q''}{d_1}-1} id_1d_2
\end{aligned}$$

and

$$\begin{aligned}
\sum_{i=1}^{q''-1} i\left(q'-1-\left[\frac{id_2}{d_1}\right]\right) &= (q'-1)\sum_{i=1}^{q''-1} i - \frac{d_2}{d_1}\sum_{i=1}^{q''-1} i^2 - \frac{1}{d_1}\sum_{i=1}^{q''-1}\overline{id_2}^{(d_1)} i\\
&= \frac{(q'-1)(q''-1)q''}{2} - \frac{q'(q''-1)(2q''-1)}{6} + \frac{1}{d_1}\sum_{j=0}^{\frac{q''}{d_1}-1}\sum_{i=1}^{d_1-1}(jd_1+i)\overline{id_2}^{(d_1)}\\
&= \frac{q'(q''^2-1)}{6} + \frac{q''}{4}\left(-q'' - \frac{q''}{d_1} - d_1 + 3\right) + \frac{q''}{d_1^2}\sum_{i=1}^{d_1-1}\overline{id_2}^{(d_1)} i.
\end{aligned}$$

With

$$\begin{aligned}
\frac{d_2}{d_1}\sum_{i=1}^{d_1-1}\overline{id_2}^{(d_1)} i =& \frac{d_2}{d_1}\left(\sum_{i=1}^{d_1-1} d_2 i^2 - \sum_{i=\left[\frac{d_1}{d_2}\right]+1}^{\left[\frac{2d_1}{d_2}\right]} d_1 i - \cdots - \sum_{i=\left[\frac{(d_2-1)d_1}{d_2}\right]+1}^{d_1-1} (d_2-1)d_1 i\right)\\
=& d_2\left(\sum_{i=1}^{d_1-1}\frac{d_2}{d_1} i^2 - (d_2-1)\sum_{i=1}^{d_1-1} i + \sum_{i=1}^{\left[\frac{(d_2-1)d_1}{d_2}\right]} i + \cdots + \sum_{i=1}^{\left[\frac{d_1}{d_2}\right]} i\right)\\
=& \frac{d_2^2(d_1-1)(2d_1-1)}{6} - \frac{d_2(d_2-1)(d_1-1)d_1}{2}\\
&+ \sum_{j=1}^{d_2-1}\frac{(jd_1 - \overline{jd_1}^{(d_2)} + d_2)(jd_1 - \overline{jd_1}^{(d_2)})}{2d_2}\\
=& \frac{d_2(d_1-1)(3d_1 - d_1d_2 - d_2)}{6} + \frac{(d_1^2+1)(d_2-1)(2d_2-1)}{12}\\
&+ \frac{(d_1-1)(d_2-1)d_2}{4} - \frac{d_1}{d_2}\sum_{j=1}^{d_2-1}\overline{jd_1}^{(d_2)} j\\
=& \frac{d_1^2+d_2^2+1}{12} + \frac{d_1^2d_2 + d_1d_2^2 - 3d_1d_2}{4} - \frac{d_1}{d_2}\sum_{j=1}^{d_2-1}\overline{jd_1}^{(d_2)} j
\end{aligned}$$

we obtain

$$
\begin{aligned}
g_1g_2\mathbf{Cov}(f_1(Y_k), f_2(Y_k)) =& \frac{1}{q}\sum_{i=1}^{q-1}\sum_{j=1}^{q-1} A_{i,j} - g_1g_2\frac{(q-1)^2}{4} \\
=& -\frac{(q-d_1)(q-d_2)}{4} + \frac{q^2-d_2^2}{6} + \frac{-d_1q - q - d_1^2d_2 + 3d_1d_2}{4} \\
& + \frac{d_1^2+d_2^2+1}{12} + \frac{d_1^2d_2 + d_1d_2^2 - 3d_1d_2}{4} \\
& + \frac{q^2-d_1^2}{6} + \frac{-d_2q - q - d_1d_2^2 + 3d_1d_2}{4} + \frac{q-d_1d_2}{2} \\
=& \frac{q^2-d_1^2-d_2^2+1}{12}
\end{aligned}
$$

and the lemma is proved. □

Clearly we have

$$\mathbf{Pr}[\epsilon_k(g_1n) = b_1, \epsilon_k(g_2n) = b_2] = \frac{A_{b_i+1,b_j+1} - A_{b_i,b_j+1} - A_{b_i+1,b_j} + A_{b_i,b_j}}{qg_1g_2}$$

for $(g_1, g_2) = 1$. Thus

$$\mathbf{Pr}[\epsilon_k(g_1n) = b_1, \epsilon_k(g_2n) = b_2] = \pi_{b_1,b_2,g_1,g_2}$$

first for $(g_1, g_2) = 1$, and, with Lemma 2.2, for general g_1, g_2. With Remark 2.3, we get

$$v_{i,j}^{(s)} = \begin{cases} C_{i,j}\left(\frac{g_{r_i}^{(i)}}{(g_{r_i}^{(i)},g_{r_j}^{(j)})}, \frac{g_{r_j}^{(j)}}{(g_{r_i}^{(i)},g_{r_j}^{(j)})}\right) & \text{if } g_{r_j}^{(j)}P_i^{(s)}(n) = g_{r_i}^{(i)}P_j^{(s)}(n) \\ 0 & \text{else.} \end{cases}$$

For $q_1 = \tilde{q}^{s_1}$ and $q_2 = \tilde{q}^{s_2}$, $f_1(n) = s_{q_1}(n)$ and $f_2(n) = s_{q_2}(n)$ are strongly q-additive functions with $q = q_1^{s_2} = q_2^{s_1}$. Then, for $P_1(n) = P_2(n) = n$, $(Y_k)_{k\geq 0}$ is clearly a sequence of independent random variables and

$$f_1(Y_k) = X_0 + \tilde{q}X_1 + \cdots + \tilde{q}^{s_1-1}X_{s_1-1} + X_{s_1} + \cdots + \tilde{q}^{s_1-1}X_{2s_1-1} + \cdots + \tilde{q}^{s_1-1}X_{s_1s_2-1},$$

$$f_2(Y_k) = X_0 + \tilde{q}X_1 + \cdots + \tilde{q}^{s_2-1}X_{s_2-1} + X_{s_2} + \cdots + \tilde{q}^{s_2-1}X_{2s_2-1} + \cdots + \tilde{q}^{s_2-1}X_{s_1s_2-1},$$

where $(X_j)_{0\leq j\leq s_1s_2-1}$ is a sequence of identically distributed independent random variables on $\{0, 1, \ldots, \tilde{q}-1\}$.

Hence we have

$$\mathbf{Cov}(f_1(Y_k), f_2(Y_k)) = \sum_{j=0}^{s_1s_2-1} c_j\mathbf{Var}X_j,$$

where c_j runs through $\{\tilde{q}^{ab} : 0 \le a \le s_1 - 1, 0 \le b \le s_2 - 1\}$ because of $(s_1, s_2) = 1$. This implies

$$\begin{aligned}\mathbf{Cov}(f_1(Y_k), f_2(Y_k)) &= \frac{\tilde{q}^2 - 1}{12} \left(1 + \tilde{q} + \cdots + \tilde{q}^{s_1 - 1}\right) \left(1 + \tilde{q} + \cdots + \tilde{q}^{s_2 - 1}\right) \\ &= \frac{(\tilde{q} + 1)(\tilde{q}^{s_1} - 1)(\tilde{q}^{s_2} - 1)}{12(\tilde{q} - 1)}.\end{aligned}$$

With $\sigma_1^2 = \mathbf{Var}\, f_1(Y_k) = s_2(q_1^2 - 1)/12$ and $\sigma_2^2 = \mathbf{Var}\, f_2(Y_k) = s_1(q_2^2 - 1)/12$, we get for the normalized covariance

$$\frac{\mathbf{Cov}(f_1(Y_k), f_2(Y_k))}{\sigma_1 \sigma_2} = \frac{\tilde{q} + 1}{\tilde{q} - 1} \frac{(q_1 - 1)(q_2 - 1)}{\sqrt{s_1 s_2 (q_1^2 - 1)(q_2^2 - 1)}}.$$

2.2.3 Comparison of moments

It remains to compare the moments of $f_\ell(P_\ell(n))$ to those in (2.7). We need the following proposition (cf. Proposition 2.1).

Proposition 2.3. *Let $P_\ell(x)$, $1 \le \ell \le d$, be integer polynomials with positive leading terms, $\lambda > 0$ an arbitrary constant and h_ℓ, $1 \le \ell \le d$, non-negative integers. Then for integers*

$$(\log N)^\eta \le k_1^{(\ell)} < k_2^{(\ell)} < \cdots < k_{h_\ell}^{(\ell)} \le \log_q N^{r_\ell} - (\log N)^\eta \quad (1 \le \ell \le d)$$

(with some $\eta > 0$) which satisfy

$$k_j^{(\ell)} \notin \Big(\log_q N^s - (\log N)^\eta, \log_q N^s + (\log N)^\eta \Big)$$

for all $1 \le s \le r_\ell - 1$, we have uniformly, as $N \to \infty$,

$$\begin{aligned}\frac{1}{N} \# \left\{ n < N \,\middle|\, \epsilon_{k_j^{(\ell)}}(P_\ell(n)) = b_j^{(\ell)}, 1 \le j \le h_\ell, 1 \le \ell \le d \right\} \\ = \prod_{s=1}^{r} p^{(s)}_{k_1^{(1)}, \cdots, k_{h_d}^{(d)}, b_1^{(1)}, \ldots, b_{h_d}^{(d)}} + \mathcal{O}\left((\log N)^{-\lambda} \right)\end{aligned}$$

and

$$\begin{aligned}\frac{1}{\pi(N)} \# \left\{ p < N \,\middle|\, \epsilon_{k_j^{(\ell)}}(P_\ell(n)) = b_j^{(\ell)}, 1 \le j \le h_\ell, 1 \le \ell \le d \right\} \\ = \prod_{s=1}^{r} p^{(s)}_{k_1^{(1)}, \cdots, k_{h_d}^{(d)}, b_1^{(1)}, \ldots, b_{h_d}^{(d)}} + \mathcal{O}\left((\log N)^{-\lambda} \right)\end{aligned}$$

with

$$p^{(s)}_{k_1^{(1)}, \cdots, k_{h_d}^{(d)}, b_1^{(1)}, \ldots, b_{h_d}^{(d)}} = \begin{cases} \mathbf{Pr}\left[Y^{(s)}_{k_j^{(\ell)}} \subseteq S^{(s)}_{b_j^{(\ell)}, \ell} \text{ for all } (j, \ell) \in K_s \right] & \text{if } K_s \ne \emptyset \\ 1 & \text{else,} \end{cases}$$

where

$$K_s = \left\{ (j,\ell) \,\middle|\, k_j^{(\ell)} \in \left[\log_q N^{s-1} + (\log N)^\eta, \log_q N^s - (\log N)^\eta\right] \right\}.$$

Proof. We follow the proof of Propositions 1.1 and 2.1 and point out the differences.

We have to consider the sums

$$\Sigma = \sum_{\mathbf{M}\in\mathcal{M}} T_{\mathbf{M}} \sum_{n<N} e\left(\mathbf{m}_1 \cdot \mathbf{v}_1 P_1(n) + \cdots + \mathbf{m}_d \cdot \mathbf{v}_d P_d(n)\right).$$

First of all set $\Delta = (\log N)^{-\delta}$ with an arbitrary (but fixed) constant $\delta > 0$. Then we can restrict to those $\mathbf{M}$ for which $|m_j^{(\ell)}| < (\log N)^{2\delta}$ for all j, ℓ because of

$$\begin{aligned}\sum_{\exists \ell, j : |m_j^{(\ell)}| \ge (\log N)^{2\delta}} |T_{\mathbf{M}}| &\ll \left(\sum_{m=[(\log N)^{2\delta}]}^{\infty} \frac{1}{\Delta m^2} \right) \left(\sum_{m=0}^{\infty} \min\left(1, \frac{1}{m}, \frac{1}{\Delta m^2}\right) \right)^{h-1} \\ &\ll \frac{1}{\Delta} (\log N)^{-\delta} \left(\log \frac{1}{\Delta}\right)^{h-1} \ll (\log N)^{-\delta/2},\end{aligned}$$

where $h = (h_1 + \cdots + h_d)$. Furthermore, it is sufficient to consider just the case where $m_j^{(\ell)} \neq 0$ for all j, ℓ. (Otherwise, we just reduce h_ℓ to a smaller value.)

Set

$$Q_{\mathbf{M}}(n) = \mathbf{m}_1 \cdot \mathbf{v}_1 P_1(n) + \cdots + \mathbf{m}_d \cdot \mathbf{v}_d P_d(n).$$

We have to check whether $Q_{\mathbf{M}}(n)$ has degree r and satisfies the conditions of Lemmata 1.2 and 1.3.

The coefficient of n^r is, if we set $k_{\max} = \max_\ell k_{h_\ell}^{(\ell)}$,

$$\frac{A_{\mathbf{M}}}{H_{\mathbf{M}}} = \sum_{(j,\ell)\in K_r} \frac{g_r^{(\ell)} m_j^{(\ell)} q^{k_{\max} - k_j^{(\ell)}}}{q^{k_{\max}}} + \sum_{(j,\ell)\notin K_r} \frac{g_r^{(\ell)} m_j^{(\ell)} q^{k_{\max} - k_j^{(\ell)}}}{q^{k_{\max}}} \tag{2.10}$$

with $(A_{\mathbf{M}}, H_{\mathbf{M}}) = 1$. If $A_{\mathbf{M}} \neq 0$, then the conditions of Lemmata 1.2 and 1.3 are obviously satisfied. If $A_{\mathbf{M}} = 0$, assume $k_{\max} \in K_r$. Then we obtain

$$\sum_{(j,\ell)\in K_r} g_r^{(\ell)} m_j^{(\ell)} q^{k_{\max} - k_j^{(\ell)}} \equiv 0 \left(q^{k_{\max} - (\log_q N^{r-1} - (\log N)^\eta)}\right).$$

Because of $|m_j^{(\ell)}| < (\log N)^{2\delta}$ this implies $\sum_{(j,\ell)\in K_r} g_r^{(\ell)} m_j^{(\ell)} q^{k_{\max} - k_j^{(\ell)}} = 0$.

Hence $A_{\mathbf{M}} = 0$ if and only if both sums in (2.10) are zero and we have

$$\begin{aligned}
&\frac{1}{N}\#\left\{n < N \,\middle|\, \epsilon_{k_j^{(\ell)}}(P_\ell(n)) = b_j^{(\ell)}, 1 \le j \le h_\ell, 1 \le \ell \le d\right\} \\
&=\frac{1}{N}\#\left\{n < N \,\middle|\, \epsilon_{k_j^{(\ell)}}(P_\ell(n)) = b_j^{(\ell)}, (j,\ell) \in K_r\right\} \\
&\quad\times \frac{1}{N}\#\left\{n < N \,\middle|\, \epsilon_{k_j^{(\ell)}}(P_\ell(n)) = b_j^{(\ell)}, (j,\ell) \notin K_r\right\} + \mathcal{O}\left((\log N)^{-\lambda}\right).
\end{aligned}$$

Now we can repeat the arguments for $(j,\ell) \in K_{r-1}$ and get inductively

$$\begin{aligned}
&\frac{1}{N}\#\left\{n < N \,\middle|\, \epsilon_{k_j^{(\ell)}}(P_\ell(n)) = b_j^{(\ell)}, 1 \le j \le h_\ell, 1 \le \ell \le d\right\} \\
&= \prod_{s=1}^{r} \frac{1}{N}\#\left\{n < N \,\middle|\, \epsilon_{k_j^{(\ell)}}(P_\ell(n)) = b_j^{(\ell)}, (j,\ell) \in K_s\right\} + \mathcal{O}\left((\log N)^{-\lambda}\right).
\end{aligned}$$

Hence we may assume from now on that all $k_j^{(\ell)}$ are contained in one set K_s for some $s \le r$.

If the degree of $Q_{\mathbf{M}}(n)$ is smaller than s, we have

$$|Q_{\mathbf{M}}(n)| \ll \frac{(\log N)^{2\delta} N^{s-1}}{q^{\log_q N^{s-1} + (\log N)^\eta}} = \frac{(\log N)^{2\delta}}{q^{(\log N)^\eta}}$$

for all $n < N$ and, with $e(y) = 1 + \mathcal{O}(y)$,

$$\sum_{\left|m_j^{(\ell)}\right| < (\log N)^{2\delta}, \deg(Q_{\mathbf{M}}(n)) < s} T_{\mathbf{M}} \left(\sum_{n<N} e(Q_{\mathbf{M}}(n)) - N\right) \ll \frac{N(\log N)^{2\delta(h+1)}}{q^{(\log N)^\eta}}.$$

Thus we can treat these $Q_{\mathbf{M}}(n)$ as if they were the zero polynomial and it suffices to regard the polynomials $P_\ell^{(s)}(n)$ and

$$Q_{\mathbf{M}}^{(s)}(n) = \mathbf{m}_1 \cdot \mathbf{v}_1 P_1^{(s)}(n) + \cdots + \mathbf{m}_d \cdot \mathbf{v}_d P_d^{(s)}(n).$$

The conditions of Lemmata 1.2 and 1.3 are satisfied if and only if $Q_{\mathbf{M}}^{(s)}(n) \not\equiv 0$ and we obtain

$$\begin{aligned}
\Sigma = N \sum_{\mathbf{M} \in \mathcal{M}: Q_{\mathbf{M}}^{(s)}(n) \equiv 0} T_{\mathbf{M}} + \mathcal{O}\left(N(\log N)^{-\tau_0} \sum_{\mathbf{M} \in \mathcal{M}: |m_j^{(\ell)}| < (\log N)^{2\delta}, Q_{\mathbf{M}}^{(s)}(n) \not\equiv 0} |T_{\mathbf{M}}| \right) \\
+ \mathcal{O}\left(N(\log N)^{-\delta/2}\right) + \mathcal{O}\left(\max_{j,\ell} U_{b_j^{(\ell)}, q_\ell, \Delta}\right).
\end{aligned}$$

Since the main term $\sum_{\mathbf{M}\in\mathcal{M}:Q_{\mathbf{M}}^{(s)}(n)\equiv 0} T_{\mathbf{M}}$ depends on Δ, we want to replace $T_{\mathbf{M}}$ by

$$T'_{\mathbf{M}} = \prod_{\ell=1}^{d}\prod_{j=1}^{h_\ell} c_{m_j^{(\ell)},b_j^{(\ell)},q}.$$

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

We clearly have

$$d_{m_j^{(\ell)},b_j^{(\ell)},q} = c_{m_j^{(\ell)},b_j^{(\ell)},q}\left(1+\mathcal{O}\left(m_j^{(\ell)}\Delta\right)\right)$$

as $\Delta \to 0$ and therefore

$$T_{\mathbf{M}} = T'_{\mathbf{M}}\left(1+\mathcal{O}\left(\max_{j,\ell} m_j^{(\ell)}\Delta\right)\right). \tag{2.11}$$

First assume $|m_j^{(\ell)}| < (\log N)^{\delta/2}$ for all j,ℓ. From (2.11) and $c_{m_j^{(\ell)},b_j^{(\ell)},q} \le \min\left(1, \frac{1}{m_j^{(\ell)}}\right)$, we obtain

$$\begin{aligned}\sum_{\mathbf{M}\in\mathcal{M}:|m_j^{(\ell)}|<(\log N)^{\delta/2}} |T_{\mathbf{M}} - T'_{\mathbf{M}}| &\ll \sum_{\mathbf{M}\in\mathcal{M}:|m_j^{(\ell)}|<(\log N)^{\delta/2}} |T'_{\mathbf{M}}|(\log N)^{-\delta/2}\\ &\ll \left(\sum_{m=1}^{[(\log N)^{\delta/2}]}\frac{1}{m}\right)^h (\log N)^{-\delta/2} \le \frac{\left(\log(\log N)^{\delta/2}\right)^h}{(\log N)^{\delta/2}} \ll (\log N)^{-\delta/3}\end{aligned}$$

It remains to estimate the $T_{\mathbf{M}}$ and $T'_{\mathbf{M}}$ with $|m_j^{(\ell)}| > (\log N)^{\delta/2}$ for some j,ℓ which satisfy the equation $Q_{\mathbf{M}}^{(s)}(n)\equiv 0$, i.e.

$$\sum_{j,\ell} g_r^{(\ell)} q^{k_{\max}-k_j^{(\ell)}} m_j^{(\ell)} = 0. \tag{2.12}$$

Assume first $g_r^{(\ell)} \ne 0$ for all j,ℓ. For simplicity, let us rewrite (2.12) as

$$\gamma_1 m_1 + \gamma_2 m_2 + \cdots + \gamma_h m_h = 0,$$

where m_h is an $m_j^{(\ell)}$ with $k_j^{(\ell)} = k_{\max}$ and the other m_i are arbitrary permutations of the $m_j^{(\ell)}$. Hence $\gamma_h = g_r^{(\ell)}$ is bounded by $\max_\ell g_r^{(\ell)}$ which is a constant. We may assume $\gamma_h = 1$.

Then for every choice of m_i, $1 \le i \le h-1$, we get a unique $m_h(= -\gamma_1 m_1 - \cdots - \gamma_{h-1}m_{h-1})$ which satisfies (2.12). If we sum up the $T'_{\mathbf{M}}$ with $|m_h| \ge |m_1 \dots m_{h-1}|^{1/(h-1)^2}$, we obtain

$$\sum T'_{\mathbf{M}} \ll \sum_{m_1=1}^{\infty}\cdots\sum_{m_{h-1}=1}^{\infty}\frac{1}{|m_1|\dots|m_{h-1}|}\frac{1}{|m_1\dots m_{h-1}|^{\frac{1}{(h-1)^2}}} = \prod_{i=1}^{h-1}\sum_{m_i=1}^{\infty}\frac{1}{|m_i|^{1+\frac{1}{(h-1)^2}}}$$

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

$$\sum T'_{\mathbf{M}} \ll (\log N)^{\frac{\delta}{2(h-1)^2}}.$$

For $|m_h| \geq (\log N)^{\delta/2}$, we get

$$\sum_{|m_i| \leq (\log N)^{\delta/2} \text{ for all } i \leq h-1} T'_{\mathbf{M}} \ll \prod_{i=1}^{h-1} \left(\sum_{m_i=1}^{[(\log N)^{\delta/2}]} \frac{1}{m_i} \right) \frac{1}{(\log N)^{\delta/2}} \ll \frac{\log(\log N)^{\delta/2}}{(\log N)^{\delta/2}}$$

and $|m_i| \geq (\log N)^{\delta/2}$ for some $i \leq h-1$ else.

For the remaining m_i, we have $|m_h| < |m_1 \dots m_{h-1}|^{1/(h-1)^2}$. We fix m_1 and consider all m_i, $2 \leq i \leq h-1$, with $|\gamma_i m_i| \leq |\gamma_1 m_1|$. Then we have

$$|m_h| < |\gamma_1 m_1 \dots \gamma_{h-1} m_{h-1}|^{\frac{1}{(h-1)^2}} \leq |\gamma_1 m_1|^{\frac{1}{h-1}}$$

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 two intervals and get

$$\sum_{m_2 : |\gamma_2 m_2| \in \left(|\gamma_1 m_1| - |\gamma_1 m_1|^{(h-2)/(h-1)}, |\gamma_1 m_1| \right]} \frac{1}{|m_2|}$$
$$\leq \frac{2|\gamma_1 m_1|^{\frac{h-2}{h-1}}}{|\gamma_2|} \frac{|\gamma_2|}{|\gamma_1 m_1| - |\gamma_1 m_1|^{\frac{h-2}{h-1}}} \leq \frac{4}{|\gamma_1 m_1|^{\frac{1}{h-1}}}$$

for all (not too small) m_1. From now on we consider only $|\gamma_2 m_2| \in \left(0, |\gamma_1 m_1| - |\gamma_1 m_1|^{(h-2)/(h-1)}\right]$. This implies

$$|\gamma_3 m_3 + \dots + \gamma_{h-1} m_{h-1}| \in \left[|\gamma_1 m_1 + \gamma_2 m_2| - |\gamma_1 m_1|^{1/h}, |\gamma_1 m_1 + \gamma_2 m_2| + |\gamma_1 m_1|^{1/h} \right]$$

with

$$|\gamma_1 m_1 + \gamma_2 m_2| \in \left[|\gamma_1 m_1|^{\frac{h-2}{h-1}}, 2|\gamma_1 m_1| - |\gamma_1 m_1|^{\frac{h-2}{h-1}} \right].$$

For m_3, we get

$$\sum_{m_3 : |\gamma_3 m_3| \in \left(|\gamma_1 m_1 + \gamma_2 m_2| - |\gamma_1 m_1|^{(h-3)/(h-1)}, |\gamma_1 m_1 + \gamma_2 m_2| + |\gamma_1 m_1|^{(h-3)/(h-1)} \right]} \frac{1}{|m_2|} \leq \frac{4}{|\gamma_1 m_1|^{\frac{1}{h-1}}}$$

and the remaining m_3 imply

$$|\gamma_1 m_1 + \gamma_2 m_2 + \gamma_3 m_3| \in \left[|\gamma_1 m_1|^{\frac{h-3}{h-1}}, 3|\gamma_1 m_1| \right].$$

We proceed inductively and the remaining m_{h-1} imply

$$|\gamma_1 m_1+\gamma_2 m_2+\gamma_3 m_3+\gamma_{h-1}m_{h-1}| \in \left[|\gamma_1 m_1|^{\frac{1}{h-1}},(h-2)|\gamma_1 m_1|\right],$$

but this contradicts $|m_h|<|\gamma_1 m_1|^{1/(h-1)}$ and no m_{h-1} are left. Thus the sum over all m_i, $2\le i\le h-1$, with $|\gamma_i m_i|\le|\gamma_1 m_1|$, can be split into $h-2$ sums, where the sum over one m_i is always bounded by $|\gamma_1 m_1|^{-1/(h-1)}$ and the sum over the other m_i can be bounded by $\log|\gamma_1 m_1|$. Hence we obtain

$$\sum_{Q_{\mathbf{M}}^{(s)}(n)\equiv 0,m_1\,\mathrm{fixed},|\gamma_1 m_1|\ge|\gamma_i m_i|,|m_h|<|m_1\ldots m_{h-1}|^{1/(h-1)^2}} T'_{\mathbf{M}} \le \frac{1}{|m_1|}\frac{4(h-3)(\log|\gamma_1 m_1|)^{h-3}}{|\gamma_1 m_1|^{\frac{1}{h-1}}}$$

for all m_1. If we consider only $|m_i|\ge(\log N)^{\delta/2}$ for some $i\le h$, then we also have $|\gamma_1 m_1|\ge(\log N)^{\delta/2}$ and get

$$\sum T'_{\mathbf{M}} \ll \sum_{m_1:|\gamma_1 m_1|\ge(\log N)^{\delta/2}} \frac{(\log|\gamma_1 m_1|)^{h-3}}{|m_1||\gamma_1 m_1|^{1/(h-1)}} \ll \frac{1}{|\gamma_1 m_1|^{1/h}} \ll (\log N)^{-\delta/2h}.$$

Summing up, we have

$$\sum_{\mathbf{M}\in\mathcal{M}:Q_{\mathbf{M}}^{(s)}(n)\equiv 0,|m_j^{(\ell)}|\ge(\log N)^{\delta/2}\ \text{for some}\ j,\ell} T'_{\mathbf{M}} \ll (\log N)^{-\frac{\delta}{2(h-1)^2}}. \tag{2.13}$$

If we consider only $m_j^{(\ell)}\ge(\log N)^{\delta/2}$ for some ℓ with $g_r^{(\ell)}=0$ and $g_i^{(\ell)}\ne 0$ for some $i\ge s$, then replace all $g_r^{(\ell)}$ in (2.12) by $g_i^{(\ell)}$ and we get the same estimate. If all $g_i^{(\ell)}$ are zero, then we have the zero polynomial and all digits $b_j^{(\ell)}$ must be zero. Clearly (2.13) also holds if we replace $T'_{\mathbf{M}}$ by $T_{\mathbf{M}}$.

Hence

$$\sum_{\mathbf{M}\in\mathcal{M}:Q_{\mathbf{M}}^{(s)}(n)\equiv 0} T_{\mathbf{M}} = \tilde{p}^{(s)}_{k_1^{(1)},\ldots,k_{h_d}^{(d)},b_1^{(1)},\ldots,b_{h_d}^{(d)}} + \mathcal{O}\left((\log N)^{-\frac{\delta}{2(h-1)^2}}\right),$$

where

$$\tilde{p}^{(s)}_{k_1^{(1)},\ldots,k_{h_d}^{(d)},b_1^{(1)},\ldots,b_{h_d}^{(d)}} = \sum_{\mathbf{M}\in\mathcal{M}:Q_{\mathbf{M}}^{(s)}(n)\equiv 0} T'_{\mathbf{M}}$$

and we get

$$\Sigma = Np^{(s)}_{k_1^{(1)},\ldots,k_{h_d}^{(d)},b_1^{(1)},\ldots,b_{h_d}^{(d)}} + \mathcal{O}\left((\log N)^{-\lambda}\right),$$

for $\delta=2(h-1)^2\lambda$ and $\tau_0>\lambda$.

It remains to prove that the $\tilde{p}^{(s)}_{k_1^{(1)},\cdots,k_{h_d}^{(d)},b_1^{(1)},\ldots,b_{h_d}^{(d)}}$ are the probabilities defined by the Markov chain.

We have

$$\left\{n < N \,\middle|\, \epsilon_{k_j^{(\ell)}}(P_\ell^{(s)}(n)) = b_j^{(\ell)} \text{ for all } (j,l) \in K_s\right\}$$
$$= \left\{n < N \,\middle|\, \mathbf{w}_{k_{\max}}^{(s)}(n) \in \bigcap_{(j,\ell)\in K_s} T^{k_j^{(\ell)}-k_{\max}} S^{(s)}_{b_j^{(\ell)},\ell}\right\}$$

and this intersection consists of a finite number of convex sets, which can be arbitrarily well approximated by elementary rectangles

$$\prod_{i=s}^{r}\left[\sum_{j=1}^{J_i}\tilde{b}_j^{(i)}q^{-j}, \sum_{j=1}^{J_i}\tilde{b}_j^{(i)}q^{-j}+q^{-J_i}\right).$$

We have

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \mathbf{w}_{k_{\max}}^{(s)}(n) \in \prod_{i=s}^{r}\left[\sum_{j=1}^{J_i}\tilde{b}_j^{(i)}q^{-j}, \sum_{j=1}^{J_i}\tilde{b}_j^{(i)}q^{-j}+q^{-J_i}\right)\right\}$$
$$= \frac{1}{N}\#\left\{n < N \,\middle|\, \epsilon_{k_{\max}-j+1}(n^i) = \tilde{b}_j^{(i)}, 1 \le j \le J_i, s \le i \le r\right\} \to \frac{1}{q^{J_s}\dots q^{J_r}},$$

if $k_{\max} \le \log N - (\log N)^\eta$ and $J_i \le k_{\max} - (\log N)^\eta$ because of Proposition 2.1. This means that the density in each of this rectangles converges to its Lebesgue measure. Since we do not change the sets if we shift all $k_j^{(\ell)}$ and increase N, the J_i can be arbitrarily large. Therefore $\tilde{p}^{(s)}_{k_1^{(1)},\cdots,k_{h_d}^{(d)},b_1^{(1)},\dots,b_{h_d}^{(d)}}$ must be the Lebesgue measure of $\bigcap_{j,\ell} T^{k_j^{(\ell)}-k_{\max}} S^{(s)}_{b_j^{(\ell)},\ell}$, which is just $p^{(s)}_{k_1^{(1)},\cdots,k_{h_d}^{(d)},b_1^{(1)},\dots,b_{h_d}^{(d)}}$.

This also implies Lemma 2.2 ($d = 2, h_1 = h_2 = 1$). □

Proposition 2.3 shows that we have to change the definition of $\overline{f}^{(N^{r_\ell})}$ slightly, namely

$$\overline{f}_\ell^{(N^{r_\ell})}(P_\ell(n)) = \sum_{s=1}^{r_\ell}\ \sum_{k=(s-1)\log_q N+A(N)}^{(s-1)\log_q N+B(N)} f_\ell(\epsilon_k(P_\ell(n))).$$

We still have $\overline{f}_\ell^{(N^{r_\ell})}(P_\ell(n)) = f_\ell(P_\ell(n)) + \mathcal{O}\left((\log N)^\eta\right)$. The definitions of $\overline{M}(N^{r_\ell})$ and $\overline{D}(N^{r_\ell})$ are adapted similarly.

Corollary 2.5. *We have*

$$\frac{1}{N}\sum_{n<N}\prod_{\ell=1}^{d}\left(\frac{\overline{f}_\ell^{(N^{r_\ell})}(P_\ell(n))-\overline{M}_\ell(N^{r_\ell})}{\overline{D}_\ell(N^{r_\ell})}\right)^{h_\ell}$$
$$-\mathbf{E}\prod_{\ell=1}^{d}\left(\frac{\sum_{s=1}^{r_\ell}\sum_{k=(s-1)\log_q N+A(N)}^{(s-1)\log_q N+B(N)} f_\ell\left(Y_k^{(s)}\right)-\overline{M}_\ell(N^{r_\ell})}{\overline{D}_\ell(N^{r_\ell})}\right)^{h_\ell}\to 0$$

and

$$\frac{1}{\pi(N)}\sum_{p<N}\prod_{\ell=1}^{d}\left(\frac{\overline{f}_\ell^{(N^{r_\ell})}(P_\ell(p))-\overline{M}_\ell(N^{r_\ell})}{\overline{D}_\ell(N^{r_\ell})}\right)^{h_\ell}$$
$$-\mathbf{E}\prod_{\ell=1}^{d}\left(\frac{\sum_{s=1}^{r_\ell}\sum_{k=(s-1)\log_q N+A(N)}^{(s-1)\log_q N+B(N)} f_\ell\left(Y_k^{(s)}\right)-\overline{M}_\ell(N^{r_\ell})}{\overline{D}_\ell(N^{r_\ell})}\right)^{h_\ell}\to 0,$$

where the $Y_k^{(s)}$ and $Y_{k'}^{(s')}$ are independent if $s\neq s'$.

Proof. The second terms are the sum over all integers

$$k_1^{(\ell)},\ldots,k_{h_\ell}^{(\ell)}\in[A(N),\log_q N^{r_\ell}-A(N)]\setminus\bigcup_{s=1}^{r_\ell-1}[\log_q N^s-A(N),\log_q N^s+A(N)],$$

$1\le\ell\le d$, of

$$\mathbf{E}\prod_{\ell=1}^{d}\prod_{j=1}^{h_\ell}\frac{f_\ell\left(Y_{k_j^{(\ell)}}^{(s)}\right)-\mu_{\ell,k_j^{(\ell)}}}{D_\ell(N^{r_\ell})}$$
$$=\sum_{b_1^{(1)}=0}^{q-1}\cdots\sum_{b_{h_d}^{(d)}=0}^{q-1}\prod_{\ell=1}^{d}\prod_{j=1}^{h_\ell}\frac{f_\ell(b_j^\ell)-\mu_{\ell,k_j^{(\ell)}}}{\overline{D}_\ell(N^{r_\ell})}\mathbf{Pr}\left[Y_{k_j^{(\ell)}}^{(s)}\subseteq S_{b_j^{(\ell)}}^{(s)}\text{ for all }j,\ell\right],$$

where the s are such that $k_j^{(\ell)}\in K_s$. Since the $Y_{k_j^{(\ell)}}^{(s)}$ are independent for different s, we have

$$\mathbf{Pr}\left[Y_{k_j^{(\ell)}}^{(s)}\subseteq S_{b_j^{(\ell)}}^{(s)}\text{ for all }(j,\ell)\right]=\prod_{s=1}^{r}\mathbf{Pr}\left[Y_{k_j^{(\ell)}}^{(s)}\subseteq S_{b_j^{(\ell)}}^{(s)}\text{ for all }(j,\ell)\in K_s\right]$$

and, by Proposition 2.3, the corresponding first terms are the same up to an error term of $\mathcal{O}\left((\log N)^{-\lambda}\right)$. Hence the convergences are valid with error terms $\mathcal{O}\left((\log N)^{-\lambda+h-h\eta}\right)$. $\square$

Similarly to Lemma 1.1, we obtain

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

and therefore

$$\frac{1}{N}\#\left\{n<N \,\middle|\, \frac{f_\ell(P_\ell(n))-M_\ell(N^{r_\ell})}{D_\ell(N^{r_\ell})} < x_\ell, \ell=1,2,\ldots,d\right\}$$
$$\to \mathbf{Pr}\left[\frac{\sum_{s=1}^{r_\ell}\sum_{k=(s-1)\log_q N+A(N)}^{(s-1)\log_q N+B(N)} f_\ell\left(Y_k^{(s)}\right)-\overline{M}_\ell(N^{r_\ell})}{\overline{D}_\ell(N^{r_\ell})} < x_\ell, \ell=1,\ldots,d\right].$$

Clearly we have $\overline{M}_\ell(N^{r_\ell}) = r_\ell\overline{M}_\ell(N)$, $\overline{D}_\ell(N^{r_\ell})^2 = r_\ell\overline{D}_\ell(N)^2$ and

$$\frac{\sum_{s=1}^{r_\ell}\sum_{k=(s-1)\log_q N+A(N)}^{(s-1)\log_q N+B(N)} f_\ell\left(Y_k^{(s)}\right)-\overline{M}_\ell(N^{r_\ell})}{\overline{D}_\ell(N^{r_\ell})}$$
$$= \frac{1}{\sqrt{r_\ell}}\sum_{s=1}^{r_\ell}\frac{\sum_{k=A(N)}^{B(N)} f_\ell\left(Y_k^{(s)}\right)-\overline{M}_\ell(N)}{\sigma_\ell\sqrt{B(N)-A(N)+1}} \to \frac{1}{\sqrt{r_\ell}}\left(Z_\ell^{(1)}+\cdots+Z_\ell^{(r)}\right)$$

by Proposition 2.2, where the $Z^{(s)} = (Z_1^{(s)},\ldots,Z_d^{(s)})$ are independent normally distributed random vectors with covariance matrices $V^{(s)}$. (For $s>r_\ell$, we have $f_\ell(Y_k^{(s)}) = 0 = Z_\ell^{(s)}$ because of $P_\ell^{(s)}(n)\equiv 0$ and $S_{0,\ell}^{(s)} = \mathbb{T}^{r-s+1}$.)

Hence the sum is normally distributed and the elements of the covariance matrix V are given by

$$v_{i,j} = \frac{1}{\sqrt{r_i r_j}}\left(v_{i,j}^{(1)}+\cdots+v_{i,j}^{(r)}\right).$$

For $r_i\neq r_j$, all $v_{i,j}^{(s)}$ are zero, as well as for all $s>r_i$. If $g_{r_j}^{(j)}P_i(n)\equiv g_{r_i}^{(i)}P_j(n)$, then $v_{i,j}^{(1)} = \cdots = v_{i,j}^{(r_i)} = v_{i,j}$. If we just have $r_i = r_j$ and $g_{r_j}^{(j)}g_s^{(i)} = g_{r_i}^{(i)}g_s^{(j)}$ for all $s>s'$, then $v_{i,j}^{(s'+1)} = \cdots = v_{i,j}^{(r_i)}$ and $v_{i,j}^{(s)} = 0$ for $s\leq s'$. Therefore $v_{i,j} = \frac{r_i-s'}{r_i}v_{i,j}^{(r_i)}$ and the covariance matrix has the stated form. This concludes the proof of (2.3).

The proof of (2.4) runs along the same lines.

2.3 Two polynomials of the same degree

For different q_ℓ and equal degrees r_ℓ, up to now only the case $d=2$ can be handled exhaustively. Theorem 2.3 was stated by Drmota [11] for coprime integers q_1, q_2, linear polynomials and only for sequences of all integers (not primes). In a joint work with Drmota ([13]) the theorem was stated for all polynomials and sequences of primes, but still only for coprime integers.

Theorem 2.3. *Let $q_1, q_2 > 1$ be multiplicatively independent integers and let f_ℓ, $\ell = 1, 2$, be q_ℓ-additive functions such that $f_\ell(bq_\ell^j) = \mathcal{O}(1)$ as $j \to \infty$ for all $b \in \{0, 1, \dots, q_\ell - 1\}$. Assume that $\frac{D_\ell(N)}{(\log N)^\eta} \to \infty$ as $N \to \infty$, for some $\eta > 0$. Let $P_\ell(n)$ be polynomials of degree r_ℓ with integer coefficients and positive leading terms. Then, as $N \to \infty$,*

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \frac{f_\ell(P_\ell(n)) - M_\ell(N^{r_\ell})}{D_\ell(N^{r_\ell})} < x_\ell,\ \ell = 1, 2\right\} \to \Phi(x_1)\Phi(x_2)$$

and

$$\frac{1}{\pi(N)}\#\left\{p < N \,\middle|\, \frac{f_\ell(P_\ell(p)) - M_\ell(N^{r_\ell})}{D_\ell(N^{r_\ell})} < x_\ell,\ \ell = 1, 2\right\} \to \Phi(x_1)\Phi(x_2)$$

Note that for multiplicatively dependent q_1, q_2, the distributions of $f_1(P_1(n)), f_2(P_2(n))$ are dependent (cf. Corollary 2.4).

We have to prove the following proposition.

Proposition 2.4 (cf. Proposition 2.1). *Let q_1, q_2 be multiplicatively independent integers and $P_1(n), P_2(n)$ integer polynomials with positive leading terms. Let $\lambda > 0$ be an arbitrary constant and h_1, h_2 non-negative integers. Then for integers*

$$(\log N^{r_\ell})^\eta \le k_1^{(\ell)} < k_2^{(\ell)} < \dots < k_{h_\ell}^{(\ell)} \le \log_{q_\ell} N^{r_\ell} - (\log N^{r_\ell})^\eta \quad (\ell = 1, 2)$$

(with some $\eta > 0$), we have, as $N \to \infty$,

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \epsilon_{q_1,k_j^{(1)}}(P_1(n)) = b_j^{(1)}, \epsilon_{q_2,k_j^{(2)}}(P_2(n)) = b_j^{(2)}, 1 \le j \le h_\ell\right\}$$
$$= \frac{1}{q_1^{h_1}q_2^{h_2}} + \mathcal{O}\left((\log N)^{-\lambda}\right)$$

and

$$\frac{1}{\pi(N)}\#\left\{p < N \,\middle|\, \epsilon_{q_1,k_j^{(1)}}(P_1(n)) = b_j^{(1)}, \epsilon_{q_2,k_j^{(2)}}(P_2(n)) = b_j^{(2)}, 1 \le j \le h_\ell\right\}$$
$$= \frac{1}{q_1^{h_1}q_2^{h_2}} + \mathcal{O}\left((\log N)^{-\lambda}\right)$$

uniformly for $b_j^{(\ell)} \in \{0, \dots, q_\ell - 1\}$ and $k_j^{(\ell)}$ in the given range, where the implicit constant of the error term may depend on q_ℓ, on the polynomials P_ℓ, on h_ℓ and on λ.

For the proof we need the corollary to the following lemma, which is a proper version of Baker's theorem on linear forms, due to Waldschmidt [37].

Lemma 2.4. *Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be non-zero algebraic numbers and $b_1, b_2, \ldots, b_n$ integers such that*

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

and let $A_1, A_2, \ldots, A_n \geq e$ real numbers with $\log A_j \geq h(\alpha_j)$, where $h(\cdot)$ denotes the absolute logarithmic height. Set $d = [\mathbb{Q}(\alpha_1 \ldots, \alpha_n) : \mathbb{Q}]$. Then

$$\left|\alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1\right| \geq e^{-U},$$

where

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

and

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

Corollary 2.6. *Let k_1, k_2 be positive integers, q_1, q_2 positive real numbers and m_1, m_2 real numbers such that $\frac{m_1}{q_1^{k_1}} + \frac{m_2}{q_2^{k_2}} \neq 0$. Then there exists a constant $C > 0$ such that*

$$\left|\frac{m_1}{q_1^{k_1}} + \frac{m_2}{q_2^{k_2}}\right| \geq \max\left(\frac{|m_1|}{q_1^{k_1}}, \frac{|m_2|}{q_2^{k_2}}\right) e^{-C \log q_1 \log q_2 \log(\max(k_1,k_2)) \log(\max(|m_1|,|m_2|))}.$$

Proof. Because of $m_1 q_1^{-k_1} + m_2 q_2^{-k_2} \neq 0$, we can apply Lemma 2.4 for $n = 3$, $\alpha_1 = q_1$, $\alpha_2 = q_2$, $\alpha_3 = -m_2/m_1$, $b_1 = k_1$, $b_2 = -k_2$, $b_3 = 1$ and directly obtain

$$\begin{aligned}\left|\frac{m_1}{q_1^{k_1}} + \frac{m_2}{q_2^{k_2}}\right| &= |m_1| q_1^{k_1} \left|-q_1^{k_1} q_2^{-k_2} \frac{m_2}{m_1} - 1\right| \\ &\geq |m_1| q_1^{k_1} e^{-C \log q_1 \log q_2 \log(\max(k_1,k_2)) \log \max(|m_1|,|m_2|)}.\end{aligned}$$

Since the problem is symmetric it is no loss of generality to assume that $|m_1| q_1^{-k_1} \geq |m_2| q_2^{-k_2}$. □

Furthermore, we need the following adapted versions of Lemmata 1.2 and 1.3.

Lemma 2.5 (cf. Lemma 1.2). *Let $P(n)$ be a polynomial of degree r with leading coefficient β. Let τ_0, τ be arbitrary positive numbers satisfying*

$$\tau \geq 2^k \tau_0 + 2^{3(k-2)}$$

and

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

Then, as $N \to \infty$,

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

Lemma 2.6 (cf. Lemma 1.3). *Let* $P(n)$ *be as in Lemma 1.2 and* τ_0, τ *arbitrary positive numbers satisfying*

$$\tau \geq 2^{6k}(\tau_0 + 1)$$

and

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

Then, as $N \to \infty$,

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

To prove Lemmata 2.5 and 2.6, we just have to replace q by $\frac{1}{\beta}$ in the proofs of Hua and use the following lemma:

Lemma 2.7.

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

where $\|x\| = \min(\{x\}, 1 - \{x\})$.

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

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

$\square$

Now we can prove Proposition 2.4.

Proof of Proposition 2.4. As for Propositions 2.1 and 2.3, we have to estimate the sums

$$\Sigma = \sum_{(\mathbf{m}_1,\mathbf{m}_2)\in\mathcal{M}} T_{\mathbf{m}_1,\mathbf{m}_2} \sum_{n<N} e\left(\mathbf{m}_1 \cdot \mathbf{v}_1 P_1(n) + \mathbf{m}_2 \cdot \mathbf{v}_2 P_2(n)\right).$$

The case of different degrees of the polynomials is treated by Proposition 2.1. So we can assume that they have the same degree $r_1 = r_2 = r$.

As in the proof of Proposition 2.3 we fix $\Delta = (\log N)^{-\delta}$ and restrict to those $(\mathbf{m}_1, \mathbf{m}_2)$ for which $|m_j^{(\ell)}| < (\log N)^{2\delta}$ and $m_j^{(\ell)} \not\equiv 0\,(q_\ell)$ for all j, ℓ.

Suppose now $g_r^{(1)}\mathbf{m}_1 \cdot \mathbf{v}_1 + g_r^{(2)}\mathbf{m}_2 \cdot \mathbf{v}_2 \neq 0$ and 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 and $\ell = 1, 2$

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

So fix K with this property. Before discussing the general case, let us consider two extremal ones.

First suppose $k_{j+1}^{(\ell)} - k_j^{(\ell)} < (\log N)^{K\varepsilon}$ for all j, ℓ. Then we set

$$\overline{m}_\ell = g_r^{(\ell)} \sum_{j=1}^{h_\ell} m_j^{(\ell)} q_\ell^{k_{h_\ell}^{(\ell)} - k_j^{(\ell)}} \qquad (\ell = 1, 2)$$

and have $\log |\overline{m}_\ell| \ll (\log N)^{K\varepsilon}$. We can apply Corollary 2.6 to

$$g_r^{(1)}\mathbf{m}_1 \cdot \mathbf{v}_1 + g_r^{(2)}\mathbf{m}_2 \cdot \mathbf{v}_2 = \frac{\overline{m}_1}{q_1^{k_{h_1}^{(1)}+1}} + \frac{\overline{m}_2}{q_2^{k_{h_2}^{(2)}+1}}$$

and obtain

$$\begin{aligned}
\left|g_r^{(1)}\mathbf{m}_1 \cdot \mathbf{v}_1 + g_r^{(2)}\mathbf{m}_2 \cdot \mathbf{v}_2\right| &\geq \max\left(q_1^{-k_{h_1}^{(1)}-1}, q_2^{-k_{h_2}^{(1)}-1}\right) e^{-c \log\log N\,(\log N)^{K\varepsilon}} \\
&\geq \frac{\max(q_1, q_2)^{(\log N)^\eta} e^{-c\log\log N\,(\log N)^{K\varepsilon}}}{N^r} \\
&\geq \frac{e^{\log(\max(q_1,q_2))(\log N)^\eta - c\log\log N\,(\log N)^{\eta\frac{h_1+h_2-2}{h_1+h_2-1}}}}{N^r} \geq \frac{(\log N)^\tau}{N^r}
\end{aligned}$$

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

$$\left|g_r^{(1)}\mathbf{m}_1 \cdot \mathbf{v}_1 + g_r^{(2)}\mathbf{m}_2 \cdot \mathbf{v}_2\right| \leq \frac{(h_1 + h_2)(\log N)^{2\delta}}{\min(q_1, q_2)^{-(\log N)^\eta}},$$

Lemmata 2.5 and 2.6 can be applied.

Next suppose $k_{j+1}^{(\ell)} - k_j^{(\ell)} \geq (\log N)^{(K+1)\varepsilon}$ for all j, ℓ. Here we set

$$\overline{m}_\ell = g_r^{(\ell)} m_1^{(\ell)} \quad (\ell = 1, 2)$$

and obtain

$$
\begin{aligned}
|g_r^{(1)}\mathbf{m}_1\cdot\mathbf{v}_1+g_r^{(2)}\mathbf{m}_2\cdot\mathbf{v}_2| &\geq \left|\frac{\overline{m}_1}{q_1^{k_1^{(1)}+1}}+\frac{\overline{m}_2}{q_2^{k_1^{(2)}+1}}\right| - \left|\sum_{j_1=2}^{h_1}\frac{m_{j_1}^{(1)}}{q_1^{k_{j_1}^{(1)}+1}}\right| - \left|\sum_{j_2=2}^{h_2}\frac{m_{j_2}^{(2)}}{q_2^{k_{j_2}^{(2)}+1}}\right| \\
&\geq \max\left(q_1^{-k_{h_1}^{(1)}-1}, q_2^{-k_{h_2}^{(1)}-1}\right) e^{-c(\log\log N)^2} \\
&\quad - \mathcal{O}\left((\log N)^{2\delta}\max\left(q_1^{-k_{h_1}^{(1)}-1}, q_2^{-k_{h_2}^{(1)}-1}\right) e^{-(\log N)^{(K+1)\varepsilon}}\right) \\
&\gg \max\left(q_1^{-k_{h_1}^{(1)}-1}, q_2^{-k_{h_2}^{(1)}-1}\right) e^{-c(\log\log N)^2}.
\end{aligned}
$$

Thus again Lemmata 2.5 and 2.6 can be applied.

In general, we assume that for some s_ℓ ($\ell=1,2$) $k_{j+1}^{(\ell)}-k_j^{(\ell)}<(\log N)^{K\varepsilon}$ for all $j<s_\ell$ and $k_{s_\ell+1}^{(\ell)}-k_{s_\ell}^{(\ell)}\geq(\log N)^{(K+1)\varepsilon}$. Here we set

$$
\overline{m}_\ell = g_r^{(\ell)}\sum_{j=1}^{s_\ell} m_j^{(\ell)} q_\ell^{k_{s_\ell}^{(\ell)}-k_j^{(\ell)}} \qquad (\ell=1,2).
$$

and have again $\log|\overline{m}_\ell|\ll(\log N)^{K\varepsilon}$. Furthermore, we can estimate the sums

$$
\sum_{j=s_\ell+1}^{h_\ell}\frac{m_j^{(\ell)}}{q_\ell^{k_j^{(\ell)}+1}} = \mathcal{O}\left((\log N)^{2\delta} q_\ell^{-k_{s_\ell}-(\log N)^{(K+1)\varepsilon}}\right).
$$

Thus we get

$$
\begin{aligned}
\left|g_r^{(1)}\mathbf{m}_1\cdot\mathbf{v}_1+g_r^{(2)}\mathbf{m}_2\cdot\mathbf{v}_2\right| &\geq \left|\frac{\overline{m}_1}{q_1^{k_{s_1}^{(1)}+1}}+\frac{\overline{m}_2}{q_2^{k_{s_2}^{(2)}+1}}\right| - \left|\sum_{j_1=s_1+1}^{h_1}\frac{m_{j_1}^{(1)}}{q_1^{k_{j_1}^{(1)}+1}}\right| - \left|\sum_{j_2=s_2+1}^{h_2}\frac{m_{j_2}^{(2)}}{q_2^{k_{j_2}^{(2)}+1}}\right| \\
&\geq \max\left(q_1^{-k_{s_1}^{(1)}-1}, q_2^{-k_{s_2}^{(2)}-1}\right) e^{-c\log\log N\,(\log N)^{K\varepsilon}} \\
&\quad - \mathcal{O}\left((\log N)^{2\delta}\max\left(q_1^{-k_{s_1}^{(1)}-1}, q_2^{-k_{s_2}^{(2)}-1}\right) e^{-(\log N)^{(K+1)\varepsilon}}\right) \\
&\gg \max\left(q_1^{-k_{s_1}^{(1)}-1}, q_2^{-k_{s_2}^{(2)}-1}\right) e^{-c\log\log N\,(\log N)^{K\varepsilon}}
\end{aligned}
$$

and the conditions of Lemmata 2.5 and 2.6 are satisfied.

If q_1 and q_2 are coprime, then we have $g_r^{(1)}\mathbf{m}_1\cdot\mathbf{v}_1+g_r^{(2)}\mathbf{m}_2\cdot\mathbf{v}_2=0$ only for $\mathbf{m}_1=\mathbf{m}_2=\mathbf{0}$. Otherwise we may have other choices of $(\mathbf{m}_1,\mathbf{m}_2)$.

Set $q=(q_1,q_2)$ and $\tilde{q}_1=q_1/q$, $\tilde{q}_2=q_2/q$. Assume, w.l.o.g., $k_{h_1}^{(1)}\geq k_{h_2}^{(2)}$.

Then we have

$$\sum_{j_1=1}^{h_1} \frac{g_r^{(1)} m_{j_1}^{(1)}}{q_1^{k_{j_1}^{(1)}}} + \sum_{j_2=1}^{h_2} \frac{g_r^{(2)} m_{j_2}^{(2)}}{q_2^{k_{j_2}^{(2)}}}$$

$$= g_r^{(1)} \frac{m_1^{(1)} \tilde{q}_1^{k_{h_1}^{(1)}-k_1^{(1)}} \tilde{q}_2^{k_{h_2}^{(2)}} q^{k_{h_1}^{(1)}-k_1^{(1)}} + \cdots + m_{h_1-1}^{(1)} \tilde{q}_1^{k_{h_1}^{(1)}-k_{h_1-1}^{(1)}} \tilde{q}_2^{k_{h_2}^{(2)}} q^{k_{h_1}^{(1)}-k_{h_1-1}^{(1)}} + m_{h_1}^{(1)} \tilde{q}_2^{k_{h_2}^{(2)}}}{\tilde{q}_1^{k_{h_1}^{(1)}} \tilde{q}_2^{k_{h_2}^{(2)}} q^{k_{h_1}^{(1)}}}$$

$$+ g_r^{(2)} \frac{m_1^{(2)} \tilde{q}_1^{k_{h_1}^{(1)}} \tilde{q}_2^{k_{h_2}^{(2)}-k_1^{(2)}} q^{k_{h_1}^{(1)}-k_1^{(2)}} + \cdots + m_{h_2}^{(2)} \tilde{q}_1^{k_{h_1}^{(1)}} q^{k_{h_1}^{(1)}-k_{h_2}^{(2)}}}{\tilde{q}_1^{k_{h_1}^{(1)}} \tilde{q}_2^{k_{h_2}^{(2)}} q^{k_{h_1}^{(1)}}}, \quad (2.14)$$

where we have omit the "+1" in the denominator for simplicity. (Just consider $k_j^{(\ell)} - 1$ instead of $k_j^{(\ell)}$.) Hence we must have

$$g_r^{(1)} \left(m_1^{(1)} \tilde{q}_1^{k_{h_1}^{(1)}-k_1^{(1)}} q^{k_{h_1}^{(1)}-k_1^{(1)}} + \cdots + m_{h_1-1}^{(1)} \tilde{q}_1^{k_{h_1}^{(1)}-k_{h_1-1}^{(1)}} q^{k_{h_1}^{(1)}-k_{h_1-1}^{(1)}} + m_{h_1}^{(1)} \right) \equiv 0 \left(\tilde{q}_1^{k_{h_1}^{(1)}} \right). \quad (2.15)$$

Of course this is useful only if $\tilde{q}_1 > 1$, which we assume first. As above, we have to distinguish several cases. (2.15) implies

$$m_{j+1}^{(1)} q_1^{k_{h_1}^{(1)}-k_{j+1}^{(1)}} + \cdots + m_{h_1-1}^{(1)} q_1^{k_{h_1}^{(1)}-k_{h_1-1}^{(1)}} + \cdots + m_{h_1}^{(1)} \equiv 0 \left(\tilde{q}_1^{k_{h_1}^{(1)}-k_j^{(1)}} \right) \quad (2.16)$$

for all j, $1 \le j \le h_1 - 1$. If $k_{j+1}^{(1)} - k_j^{(1)} \ge (\log N)^{\varepsilon}$ for some j, then $|m_j^{(\ell)}| < (\log N)^{2\delta}$ implies that the left hand side of (2.16) must be zero. Hence $m_{h_1}^{(1)} \equiv 0\,(q_1)$ which implies $T_{\mathbf{m}_1,\mathbf{m}_2} = 0$ since we have excluded $m_{h_1}^{(1)} = 0$. If $k_{j+1}^{(1)} - k_j^{(1)} \le (\log N)^{\varepsilon}$ for all j, then the left hand side of (2.15) must be zero and $m_{h_1}^{(1)} \equiv 0\,(q_1)$.

Now consider the case $\tilde{q}_1 = 1$, i.e. $q_1 | q_2$. Then we have to check

$$g_r^{(1)} \left(m_1^{(1)} \tilde{q}_2^{k_{h_2}^{(2)}} q^{k_{h_1}^{(1)}-k_1^{(1)}} + \cdots + m_{h_1-1}^{(1)} \tilde{q}_2^{k_{h_2}^{(2)}} q^{k_{h_1}^{(1)}-k_{h_1-1}^{(1)}} + m_{h_1}^{(1)} \tilde{q}_2^{k_{h_2}^{(2)}} \right) +$$

$$g_r^{(2)} \left(m_1^{(2)} \tilde{q}_2^{k_{h_2}^{(2)}-k_1^{(2)}} q^{k_{h_1}^{(1)}-k_1^{(2)}} + \cdots + m_{h_2-1}^{(2)} \tilde{q}_2^{k_{h_2}^{(2)}-k_{h_2-1}^{(2)}} q^{k_{h_1}^{(1)}-k_{h_2-1}^{(2)}} + m_{h_2}^{(2)} q^{k_{h_1}^{(1)}-k_{h_2}^{(2)}} \right) = 0. \quad (2.17)$$

This implies

$$g_r^{(2)} q^{k_{h_1}^{(1)}-k_{h_2}^{(2)}} \left(m_{j+1}^{(2)} q_2^{k_{h_2}^{(2)}-k_{j+1}^{(2)}} + \cdots + m_{h_2-1}^{(2)} q_2^{k_{h_2}^{(2)}-k_{h_2-1}^{(2)}} + m_{h_2}^{(2)} \right) \equiv 0 \left(\tilde{q}_2^{k_{h_2}^{(2)}-k_j^{(2)}} \right) \quad (2.18)$$

for $1 \le j \le h_2 - 1$ and for $j = 0$, if we set $k_0^{(2)} = 0$.

Assume first $k_{h_1}^{(1)} - k_{h_2}^{(2)} \le (\log N)^{\varepsilon/2}$. Then we can do the same reasonings as above and obtain $m_{h_2}^{(2)} \equiv 0\,(q_2)$.

The last (and most difficult) case is $k_{h_1}^{(1)} - k_{h_2}^{(2)} \ge (\log N)^{\varepsilon/2}$. First suppose that $\tilde{q}_2$ has some prime divisor $\tilde{p}_2 \not| q$. Then we get from (2.18)

$$g_r^{(2)} \left(m_{j+1}^{(2)} q_2^{k_{h_2}^{(2)} - k_{j+1}^{(2)}} + \cdots + m_{j+1}^{(2)} q_2^{k_{h_2}^{(2)} - k_{h_2-1}^{(2)}} + m_{h_2}^{(2)} \right) \equiv 0 \left(\tilde{p}_2^{k_{h_2}^{(2)} - k_j^{(2)}} \right)$$

for $0 \le j \le h_2 - 1$ and again $m_{h_2}^{(2)} \equiv 0\,(q_2)$. Suppose next that q has some prime divisor $p \not| \tilde{q}_2$. Then we have

$$g_r^{(1)} \left(m_1^{(1)} q^{k_{h_1}^{(1)} - k_1^{(1)}} + \cdots + m_{h_1-1}^{(1)} q^{k_{h_1}^{(1)} - k_{h_1-1}^{(1)}} + m_{h_1}^{(1)} \right) \equiv 0 \left(p^{k_{h_1}^{(1)} - k_{h_2}^{(2)}} \right)$$

and we can do the same reasonings with ε/h_1 instead of ε.

It remains to consider q and $\tilde{q}_2$ with prime factorisations $q = p_1^{e_1} \ldots p_s^{e_s}$, $\tilde{q}_2 = p_1^{\tilde{e}_1} \ldots p_s^{\tilde{e}_s}$, where all e_i and $\tilde{e}_i$ are positive. Let us rewrite (2.17):

$$\begin{aligned} & g_r^{(1)} \left(m_1^{(1)} \prod_{i=1}^{s} p_i^{k_{h_2}^{(2)} \tilde{e}_i + (k_{h_1}^{(1)} - k_1^{(1)}) e_i} + \cdots + m_{h_1}^{(1)} \prod_{i=1}^{s} p_i^{k_{h_2}^{(2)} \tilde{e}_i} \right) \\ + & g_r^{(2)} \left(m_1^{(2)} \prod_{i=1}^{s} p_i^{(k_{h_2}^{(2)} - k_1^{(2)}) \tilde{e}_i + (k_{h_1}^{(1)} - k_1^{(2)}) e_i} + \cdots + m_{h_2}^{(2)} \prod_{i=1}^{s} p_i^{(k_{h_1}^{(1)} - k_{h_2}^{(2)}) e_i} \right) = 0. \end{aligned}$$

By assumption, q_1 and q_2 are multiplicatively independent. Thus we have $s \ge 2$ and $e_i/\tilde{e}_i \ne e_j/\tilde{e}_j$ for some i, j. Therefore $k_{h_2}^{(2)} \tilde{e}_i - (k_{h_1}^{(1)} - k_{h_2}^{(2)}) e_i$ cannot be zero for all i and the difference must be at least $\frac{1}{2}(\log N)^{\varepsilon/2}$ for some i. Let

$$(k_{h_1}^{(1)} - k_{h_2}^{(2)}) e_{i_0} - k_{h_2}^{(2)} \tilde{e}_{i_0} \ge \frac{1}{2} (\log N)^{\varepsilon/2}.$$

Then we have

$$g_r^{(1)} \left(m_1^{(1)} \prod_{i=1}^{s} p_i^{(k_{h_1}^{(1)} - k_1^{(1)}) e_i} + \cdots + m_{h_1}^{(1)} \right) \equiv 0 \left(p_{i_0}^{(k_{h_1}^{(1)} - k_{h_2}^{(2)}) e_{i_0} - k_{h_2}^{(2)} \tilde{e}_{i_0}} \right)$$

and we can again do the same reasonings. Similarly

$$k_{h_2}^{(2)} \tilde{e}_{i_0} - (k_{h_1}^{(1)} - k_{h_2}^{(2)}) e_{i_0} \ge \frac{1}{2} (\log N)^{\varepsilon/2}$$

leads to

$$g_r^{(2)} \left(m_1^{(2)} \prod_{i=1}^{s} p_i^{(k_{h_2}^{(2)} - k_1^{(2)})(\tilde{e}_i + e_i)} + \cdots + m_{h_2}^{(2)} \right) \equiv 0 \left(p_{i_0}^{\frac{1}{2} (\log N)^{\varepsilon/2}} \right)$$

and the same result.

Hence, we finally get

$$\sum_{(\mathbf{m}_1,\mathbf{m}_2)\neq(\mathbf{0},\mathbf{0})} |T_{\mathbf{m}_1,\mathbf{m}_2}| \cdot \left| \frac{1}{N} \sum_{n<N} e\left((g_r^{(1)}\mathbf{m}_1 \cdot \mathbf{v}_1 + g_r^{(2)}\mathbf{m}_2 \cdot \mathbf{v}_2)n \right) \right|$$
$$= \mathcal{O}\left((\log N)^{-\delta/2} \right) + \mathcal{O}\left((\log N)^{2(h_1+h_2)\delta-\lambda} \right),$$

which completes the proof of Proposition 2.4. □

Chapter 3

Parry Expansions

Now we turn to digital expansions which are slightly different from q-ary expansions.

3.1 G-ary expansions and G-ary functions

Let the sequence $G = (G_k)_{k\geq 0}$ be defined by the linear recurrence

$$G_k = a_1 G_{k-1} + a_2 G_{k-2} + \cdots + a_d G_{k-d} \text{ for } k \geq d$$

and

$$G_k = a_1 G_{k-1} + a_2 G_{k-2} + \cdots + a_k G_0 \text{ for } 1 \leq k < d,\ G_0 = 1,$$

with non-negative integers a_i which satisfy the relations

$$(a_j, a_{j+1}, \ldots, a_d) \leq (a_1, a_2, \ldots, a_{d-j+1}) \text{ for } 2 \leq j \leq d$$

(where "$<$" denotes the lexicographical order) and $a_d > 0$.

Then every non-negative integer n has a unique *G-ary digital expansion*

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

with integer digits $\epsilon_k(n) \geq 0$ satisfying

$$(\epsilon_k(n), \epsilon_{k-1}(n), \ldots, \epsilon_{k-d+1}(n)) < (a_1, a_2, \ldots, a_d) \text{ for all } k \geq 0. \tag{3.1}$$

For $d = 1$, we just get q-ary expansions with $q = a_1$. Therefore assume $d > 1$. The best known example of these expansions is the Zeckendorf expansion with $d = 2$ and $a_1 = a_2 = 1$. Then the G_k are the Fibonacci numbers.

Let

$$\chi(x) = x^d - a_1 x^{d-1} - \cdots - a_{d-1} x - a_d$$

be the characteristic polynomial of the linear recurrence. It is easy to show that it has a unique dominant root $\alpha \in \mathbb{R}^+$ (e.g. consider its (primitive) companion matrix and apply the Perron-Frobenius theorem). If $\chi(x)$ is irreducible over $\mathbb{Z}$, denote by $\alpha_2, \ldots, \alpha_d$ the (distinct) algebraic conjugates of α. Then we have, for some constants $c_1, \ldots, c_d$,

$$G_k = c_1\alpha^k + c_2\alpha_2^k + \cdots + c_d\alpha_d^k. \tag{3.2}$$

(We will show $c_1 = \frac{\alpha^d-1}{\alpha-1}\frac{1}{\prod_{j>1}(\alpha-\alpha_j)}$ in Section 3.3 and get, for reasons of symmetry, $c_i = \frac{\alpha_i^d-1}{\alpha_i-1}\frac{1}{\prod_{j\neq i}(\alpha_i-\alpha_j)}$ for all $i \geq 1$, where $\alpha_1 = \alpha$.)

(3.1) and (3.2) show that these G-ary expansions of integers are strongly related to Parry's α-expansions of real numbers (with simple α-numbers) (cf. Parry [30], Grabner and Tichy [23]). Therefore they are called Parry expansions.

The analogue to q-additive functions are *G-additive functions*, i.e.

$$f(n) = \sum_{k\geq 0} f(\epsilon_k(n)G_k) = \sum_{k\geq 0} f_k(\epsilon_k(n)) \text{ for all } n \in \mathbb{N},\ f_k(0) = 0,$$

a special case of which is the sum-of-digits function

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

Several authors have studied these functions, e.g. Grabner and Tichy [23] proved the following analogue to (1.1):

$$\frac{1}{N}\sum_{n<N} s_G(n) = c_G \log_\alpha N + F(\log_\alpha N) + \mathcal{O}\left(N^{\beta-1}\log N\right),$$

where c_G is a positive constant (expressions for which will be given in Theorem 3.1), $0 \leq \beta < 1$ and F is a continuous, nowhere differentiable function, the graph of which has Haussdorf dimension 1.

Dumont and Thomas [16] obtained similar results for the moments. They used the more general framework of numeration systems associated with a substitution which we will present in Section 3.2. In [17] they prove a central limit theorem for the sum-of-digits function.

Our aim is to prove a theorem similar to Theorem 1.1, i.e. to generalise the central limit theorem on G-additive functions and on polynomial sequences.

3.2 Central limit theorem for $P(n) = n$

First we have to prove the following theorem on the distribution of the sequence $f(n)$, $0 \leq n < N$. For $d = 2$, all theorems in this chapter can be found in a joint paper with Drmota ([13]).

Theorem 3.1. *Let f be a G-additive function such that $f_k(e) = \mathcal{O}(1)$ as $k \to \infty$ for all $e \in \{0, 1, \ldots, a_1\}$. Then, for all $\eta > 0$, the expected value of $f(n)$, $0 \le n < N$, is given by*

$$E_N = \frac{1}{N} \sum_{n<N} f(n) = M(N) + \mathcal{O}\left((\log N)^\eta\right), \tag{3.3}$$

where

$$M(N) = \sum_{k=0}^{[\log_\alpha N]} \mu_k \quad \text{with} \quad \mu_k = \sum_{e=1}^{a_1} p_e f_k(e).$$

and the constants p_e are the asymptotic probabilities of the digits e, the values of which are determined by equation (3.11). Furthermore, set

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

with

$$\sigma_{k,k'}^{(2)} = \begin{cases} \sum\limits_{e=1}^{a_1} p_e f_k(e)^2 - \mu_k^2 & \text{if } k = k' \\ \sum\limits_{i=2}^{d} \left(\frac{\alpha_i}{\alpha}\right)^{|k-k'|} \sum\limits_{e=1}^{a_1} \sum\limits_{e'=1}^{a_1} p_{e,e'}^{(i)} f_{\min(k,k')}(e) f_{\max(k,k')}(e') & \text{if } k \ne k' \end{cases}$$

and constants $p_{e,e'}^{(i)}$ described on page 58.

Assume that there exists a constant $c > 0$ such that $\sigma_{k,k}^{(2)} \ge c$ for all $k \ge 0$. Then we have

$$\frac{1}{N} \sum_{n<N} (f(n) - E_N)^2 \sim D(N)^2, \tag{3.4}$$

$$\frac{1}{N} \# \left\{ n < N \,\middle|\, \frac{f(n) - M(N)}{D(N)} < x \right\} \to \Phi(x) \tag{3.5}$$

and

$$\frac{1}{N} \sum_{n<N} \left(\frac{f(n) - M(N)}{D(N)} \right)^h \to \int_{-\infty}^{\infty} x^h \, d\Phi(x) \tag{3.6}$$

as $N \to \infty$.

Remark 3.1. In case $d = 2$ we give more explicit expressions for μ_k and $\sigma_{k,k'}^{(2)}$:

$$\mu_k = \frac{\alpha+1}{\alpha D} \sum_{b=1}^{a_2-1} f_k(b) + \frac{1}{D} \sum_{b=a_2}^{a_1-1} f_k(b) + \frac{a_2}{\alpha D} f_k(a_1),$$

$$\sigma_{k,k'}^{(2)} = \begin{cases} \frac{\alpha+1}{\alpha D} \sum_{b=1}^{a_2-1} f_k(b)^2 + \frac{1}{D} \sum_{b=a_2}^{a_1-1} f_k(b)^2 + \frac{a_2}{\alpha D} f_k(a_1)^2 - \mu_k^2 & \text{if } k = k' \\ \left(-\frac{a_2}{\alpha^2}\right)^{|k-k'|} \mu_{\min(k,k')} \overline{\mu}_{\max(k,k')} & \text{if } k \ne k', \end{cases}$$

where $D = \sqrt{a_1^2 + 4a_2}$ and

$$\overline{\mu}_k = \frac{\alpha - a_2}{a_2 D} \sum_{b=1}^{a_2-1} f_k(b) - \frac{1}{D} \sum_{b=a_2}^{a_1-1} f_k(b) + \frac{\alpha}{D} f_k(a_1).$$

The proof relies on the fact that the digits of the possible G-ary expansions can be represented by random variables which form a Markov chain (of order $d-1$). For convenience, we reduce this Markov chain to a Markov chain of order 1 by using a representation of the digital expansions in terms of substitutions, like Dumont and Thomas [15, 17], who studied strongly G-additive functions, i.e. $f(n) = \sum_{k \geq 0} f(\epsilon_k(n))$.

So let σ be the substitution on $\mathcal{A} = \{1, \ldots, d\}$ defined by

$$\begin{aligned} \sigma : i &\to 1^{a_i}(i+1) \quad \text{for } 1 \leq i \leq d-1 \\ d &\to 1^{a_d} \end{aligned}$$

and let σ also stand for its extension on the set of words $\mathcal{A}^* = \bigcup_{i=1}^{\infty} \mathcal{A}^i \cup \{\Lambda\}$ with Λ the empty word. We denote by $|m|$ the length of the word m, and $m' < m$ means that m' is a strict prefix of m.

A sequence of words $m_{j-1} m_{j-2} \ldots m_0$ is said to be b-admissible, if there exist (unique) letters $b_j = b, b_{j-1}, \ldots, b_0$ such that $m_k b_k \leq \sigma(b_{k+1})$ for all $k < j$. The admissible representation of an integer $n \geq 1$ is the (unique) 1-admissible sequence $m_{j-1}(n) m_{j-2}(n) \ldots m_0(n)$, with $m_{j-1}(n) \neq \Lambda$, such that

$$n = \left|\sigma^{j-1}(m_{j-1}(n))\right| + \cdots + \left|\sigma^0(m_0(n))\right|.$$

Denote by $b_k(n)$ the letter b_k corresponding to this 1-admissible sequence. It is easy to show (by induction) that the numbers $\left|\sigma^k(1)\right|$ are just the G_k defined by the linear recurrence in the Introduction, and we have $m_k(n) = 1^{\epsilon_k(n)}$.

The matrix of the substitution

$$M = \Big(\#\{\text{occurrences of } b \text{ in } \sigma(b')\}\Big)_{b,b' \in \mathcal{A}} = \begin{pmatrix} a_1 & a_2 & \cdots & \cdots & a_d \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

is the companion matrix of the characteristic polynomial of the linear recurrence.

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

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

If we define $Y_{k,N}$ by

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

and $\xi_{k,N}$ by

$$\mathbf{Pr}[\xi_{k,N} = (m,b)] = \frac{1}{N}\#\{n < N : (m_k(n), b_k(n)) = (m,b)\},$$

we have, with $f_k(m,b) = f_k(|m|)$,

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

i.e. Y_N is a weighted sum of the $\xi_{k,N}$. Therefore we will first have a detailed look at the $\xi_{k,N}$.

Dumont and Thomas [17] showed that, for fixed j, the sequence $(\xi_{j-1,G_j}, \xi_{j-2,G_j}, \dots, \xi_{0,G_j})$ constitutes a Markov chain with transition probabilities

$$\begin{aligned} &\mathbf{Pr}[\xi_{k,G_j} = (m,b) | \xi_{k+1,G_j} = (m',b')] = \mathbf{Pr}[\xi_{k,G_j} = (m,b) | \xi_{k+1,G_j} = (.,b')] \\ &\quad = \begin{cases} \frac{|\sigma^k(b)|}{|\sigma^{k+1}(b')|} = p_{(.,b'),(m,b)} + o(\rho^k) & \text{if } mb \le \sigma(b') \\ 0 & \text{otherwise,} \end{cases} \end{aligned}$$

where $(.,b)$ denotes the set of states $\{(m,b) : m \in \mathcal{A}^*\}$, $p_{(.,b'),(m,b)} = \frac{\nu_b}{\nu_{b'}\alpha}$,

$$(\nu_1, \dots, \nu_d) = (1, \alpha - a_1, \alpha^2 - a_1\alpha - a_2, \dots, \alpha^{d-1} - a_1\alpha^{d-2} - \cdots - a_{d-1})$$

is a left eigenvector of M to the eigenvalue α, and $\rho < 1$ a constant such that all roots of $\chi(x)$ except α have modulus less than $\alpha\rho$. (For Pisot numbers α, we can set $\rho = \alpha^{-1}$.)

Furthermore, denote by $P_{k,j}$ the matrix of transition probabilities $\mathbf{Pr}[\xi_{k,G_j} = (.,b) | \xi_{k+1,G_j} = (.,b')]$. Then we have $P_{k,j} = P + \mathcal{O}\left(\rho^k\right)$ with

$$P = \left(p_{(.,b'),(.,b)}\right)_{b',b\in\mathcal{A}} = \begin{pmatrix} \frac{a_1}{\alpha} & \frac{a_2}{\alpha^2 - a_1\alpha} & \cdots & \frac{a_{d-1}}{\alpha^{d-1} - a_1\alpha^{d-2} - \cdots - a_{d-2}\alpha} & 1 \\ \frac{\alpha - a_1}{\alpha} & 0 & \cdots & \cdots & 0 \\ 0 & \frac{\alpha^2 - a_1\alpha - a_2}{\alpha^2 - a_1\alpha} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{\alpha^{d-1} - a_1\alpha^{d-2} - \cdots - a_{d-1}}{\alpha^{d-1} - a_1\alpha^{d-2} - \cdots - a_{d-2}\alpha} & 0 \end{pmatrix}.$$

P is similar to

$$\begin{pmatrix} \frac{a_1}{\alpha} & \frac{a_2}{\alpha^2} & \cdots & \cdots & \frac{a_d}{\alpha^d} \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

and its eigenvalues are therefore $1, \frac{\alpha_2}{\alpha}, \ldots, \frac{\alpha_d}{\alpha}$. Hence we have

$$\mathbf{Pr}[\xi_{k,G_j} = (.,b)] = p_{(.,b)} + \mathcal{O}\left(\rho^{\min(k,j-k)}\right), \tag{3.7}$$

where the probability vector $(p_{(.,1)}, \ldots, p_{(.,d)})^t$ is the right eigenvector of P to the eigenvalue 1 with $\sum_{b=1}^{d} p_{(.,b)} = 1$:

$$\begin{pmatrix} p_{(.,1)} \\ \vdots \\ p_{(.,d)} \end{pmatrix} = \frac{1}{\chi'(\alpha)} \left(\alpha^{d-1}, \alpha^{d-1} - a_1\alpha^{d-2}, \alpha^{d-1} - a_1\alpha^{d-2} - a_2\alpha^{d-3}, \ldots, \frac{a_d}{\alpha}\right)^t,$$

where $\chi'(x)$ denotes the derivative of $\chi(x)$. We deduce

$$\begin{aligned}\mathbf{Pr}[\xi_{k,G_j} = (m,b)] &= \sum_{b':mb\le\sigma(b')} \mathbf{Pr}[\xi_{k,G_j} = (m,b) | \xi_{k+1,G_j} = (.,b')]\mathbf{Pr}[\xi_{k+1,G_j} = (.,b')] \\ &= p_{(m,b)} + \mathcal{O}\left(\rho^{\min(k,j-k)}\right)\end{aligned}$$

with

$$p_{(m,b)} = \sum_{b':mb\le\sigma(b')} p_{(.,b'),(m,b)} p_{(.,b')}.$$

In case $d = 2$ we have $\sigma(1) = 1^{a_1}2$ and $\sigma(2) = 1^{a_2}$, thus

$$p_{(.,1),(1^e,1)} = \frac{1}{\alpha} \text{ for } 0 \le e < a_1,\ p_{(.,1),(1^{a_1},2)} = \frac{a_2}{\alpha^2},\ p_{(.,2),(1^e,1)} = \frac{1}{a_2} \text{ for } 0 \le e < a_2$$

and

$$p_{(.,1)} = \frac{\alpha}{D},\ p_{(.,2)} = \frac{a_2}{\alpha D}.$$

These asymptotics suggests to approximate the digital distribution by a stationary Markov chain $(X_k, k \ge 0)$ with the probability distribution $\mathbf{Pr}[X_k = (m,b)] = p_{(m,b)}$ and the transition probabilities $\mathbf{Pr}[X_k = (m,b) | X_{k+1} = (.,b')] = p_{(.,b'),(m,b)}$. The next lemma shows how we can quantify this approximation for the finite-dimensional distributions.

Lemma 3.1. *For every $h \ge 1$ and integers $0 \le k_1 < k_2 < \cdots < k_h < j$ we have*

$$\mathbf{Pr}[\xi_{k_1,G_j} = (.,b_1), \ldots, \xi_{k_h,G_j} = (.,b_h)] = \hat{p}_{k_1,\ldots,k_l,(.,b_1),\ldots,(.,b_l)} + \mathcal{O}\left(\rho^{\min(k_1,j-k_h)}\right),$$

where

$$\hat{p}_{k_1,\ldots,k_l,(.,b_1),\ldots,(.,b_h)} = \mathbf{Pr}[X_{k_1} = (.,b_1), \ldots, X_{k_h} = (.,b_h)].$$

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

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

and consequently

$$\mathbf{Pr}[\xi_{k,G_j} = (.,b) | \xi_{k',G_j} = (.,b')] = \mathbf{Pr}[X_k = (.,b) | X'_k = (.,b')] + \mathcal{O}\left(\rho^k\right). \tag{3.8}$$

Since

$$\begin{aligned}
&\mathbf{Pr}[\xi_{k_1,G_j} = (.,b_1), \ldots, \xi_{k_h,G_j} = (.,b_h)] \\
&= \mathbf{Pr}[\xi_{k_1,G_j} = (.,b_1) | \xi_{k_2,G_j} = (.,b_2)] \mathbf{Pr}[\xi_{k_2,G_j} = (.,b_2) | \xi_{k_3,G_j} = (.,b_3)] \cdots \\
&\qquad \cdots \mathbf{Pr}[\xi_{k_{h-1},G_j} = (.,b_{h-1}) | \xi_{k_h,G_j} = (.,b_h)] \mathbf{Pr}[\xi_{k_h,G_j} = (.,b_h)]
\end{aligned}$$

we just have to apply (3.8) and (3.7) and the lemma follows. □

Hence we have

$$\mathbf{Pr}[\xi_{k,G_j} = (m,b), \xi_{l,G_j} = (m',b')] = \hat{p}_{k,l,(m,b),(m',b')} + \mathcal{O}\left(\rho^{\min(k,j-k')}\right)$$

$(0 \le k < k' < j)$ with

$$\hat{p}_{k,k',(m,b),(m',b')} = \sum_{c: mb \le \sigma(c)} \frac{p_{(m',b')}}{p_{(.,b')}} \hat{p}_{k+1,k',(.,c),(.,b')} p_{(.,c),(m,b)} \tag{3.9}$$

because of

$$\begin{aligned}
&\mathbf{Pr}[\xi_{k,G_j} = (m,b) | \xi_{k',G_j} = (m',b')] \\
&= \sum_{c: mb \le \sigma(c)} \mathbf{Pr}[\xi_{k,G_j} = (m,b) | \xi_{k+1,G_j} = (.,c)] \mathbf{Pr}[\xi_{k+1,G_j} = (.,c) | \xi_{k',G_j} = (.,b')].
\end{aligned}$$

For finite dimensional distributions we have

$$\begin{aligned}
&\mathbf{Pr}[\xi_{k_1,G_j} = (m_1,b_1), \ldots, \xi_{k_h,G_j} = (m_h,b_h)] \\
&\qquad = \hat{p}_{k_1,\ldots,k_h,(m_1,b_1),\ldots,(m_h,b_h)} + \mathcal{O}\left(\rho^{\min(k_1,j-k_h)}\right),
\end{aligned} \tag{3.10}$$

where the $\hat{p}_{k_1,\ldots,k_h,(m_1,b_1),\ldots,(m_h,b_h)}$ are defined similarly to (3.9).

The next lemma shows that, for general N, $\xi_{k,N}$ is similar to ξ_{k,G_j} where G_j is the largest element of G not exceeding N ($j \approx [\log_\alpha N]$). Here we set $\rho = \alpha^{-1}$, if α is a Pisot number.

Lemma 3.2. *The probability distribution of $\xi_{k,N}$ for $G_j \le N < G_{j+1}$ with $k < j$ is given by*

$$\mathbf{Pr}[\xi_{k,N} = (m,b)] = \mathbf{Pr}[\xi_{k,G_j} = (m,b)] + \mathcal{O}\left(\rho^{(j-k)/2}\right).$$

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

$$\begin{aligned}
&\mathbf{Pr}[\xi_{k_1,N} = (m_1,b_1),\dots,\xi_{k_h,N} = (m_h,b_h)] \\
&\qquad = \mathbf{Pr}[\xi_{k_1,G_j} = (m_1,b_1),\dots,\xi_{k_h,G_j} = (m_h,b_h)] + \mathcal{O}\left(\rho^{(j-k_h)/2}\right)
\end{aligned}$$

Proof. For $N = \sum_{k=0}^{j} \epsilon_k G_k$, we have

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

Therefore

$$\begin{aligned}
\mathbf{Pr}[\xi_{k,N} = b] =& \frac{1}{N}\Bigg(\#\{n < \epsilon_j G_j \mid \epsilon_k(n) = b\} + \#\{n < \epsilon_{j-1} G_{j-1} \mid \epsilon_k(n) = b\} + \cdots \\
&+ \#\{n < \epsilon_{k+1} G_{k+1} \mid \epsilon_k(n) = b\} + \begin{cases} \sum\limits_{i=0}^{k-1} \epsilon_i G_i & \text{if } \epsilon_k = b \\ 0 & \text{otherwise} \end{cases} \Bigg) \\
=& \frac{1}{N}\left(\epsilon_j G_j \mathbf{Pr}[\xi_{k,G_j} = b] + \cdots + \epsilon_{[\frac{k+j}{2}]} G_{[\frac{k+j}{2}]} \mathbf{Pr}[\xi_{k,G_{[\frac{k+j}{2}]}} = b] \right) \\
&+ \mathcal{O}\left(\frac{1}{N} G_{[\frac{k+j}{2}]} \right) \\
=& \mathbf{Pr}[\xi_{k,G_j} = b] + \mathcal{O}\left(\frac{1}{\alpha^{\frac{j-k}{2}}} \right),
\end{aligned}$$

where we have used

$$\begin{aligned}
\#\{n < G_j \mid \epsilon_k(n) = b\} &= \#\{G_j \le n < 2G_j \mid \epsilon_k(n) = b\} = \cdots \\
&= \#\{(a_1 - 1)G_j \le n < a_1 G_j \mid \epsilon_k = b\}.
\end{aligned}$$

and

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

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

$$\mathbf{Pr}[\xi_{k,N} = b, \xi_{k',N} = b'] = \frac{1}{N} \sum_{i=k'+1}^{j} \epsilon_i G_i \mathbf{Pr}[\xi_{k,G_i} = b, \xi_{k',G_i} = b']$$

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

Thus, we can proceed in the same way. $\square$

As in Section 1.3, we can concentrate on the digits $\epsilon_k(n)$ with $A(N) \le k \le B(N)$, where $A(N) = [(\log N)^\eta]$, $B(N) = [\log_\alpha N] - [(\log N)^\eta]$ and $\eta > 0$ is a sufficiently small number (to be chosen in the sequel), in order to obtain uniform estimates.

The following lemma is a direct consequence of Lemma 3.2 and (3.10). Note that it is not necessary that $k_1, \ldots, k_h$ are ordered and that they are distinct.

Lemma 3.3. *For every $h \ge 1$ and for every $\lambda > 0$ we have*

$$\frac{1}{N}\#\{n < N \mid \epsilon_{k_1}(n) = e_1, \ldots, \epsilon_{k_l}(n) = e_h\} = \hat{p}_{k_1,\ldots,k_h,e_1,\ldots,e_h} + \mathcal{O}\left((\log N)^{-\lambda}\right)$$

uniformly for

$$A(N) \le k_1, k_2, \cdots, k_h \le B(N),$$

where

$$\hat{p}_{k_1,\ldots,k_h,e_1,\ldots,e_h} = \sum_{(m_i,b_i):|m_i|=e_i} \hat{p}_{k_1,\ldots,k_h,(m_1,b_1),\ldots,(m_h,b_h)}.$$

As in Section 1.3 we define

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

Now, we turn to the derivation of $E_N = \mathbf{E}\, Y_N$, i.e. to the proof of (3.3). For $Y_{k,N}$, we get

$$\mathbf{E}\, Y_{k,N} = \sum_{m,b} \mathbf{Pr}[\xi_{k,N} = (m,b)] f_k(|m|) = \sum_{e=0}^{a_1} p_e f_k(e) + \mathcal{O}\left(\rho^{\min(k,(j-k)/2)}\right),$$

where

$$p_e = \sum_{m,b:|m|=e} p_{(m,b)}. \tag{3.11}$$

In case $d = 2$, we have

$$p_e = \begin{cases} p_{(1^e,1)} + p_{(1^e,2)} = \frac{\alpha+1}{\alpha D} & \text{if } e < a_2 \\ p_{(1^e,1)} = \frac{1}{D} & \text{if } a_2 \le e < a_1 \\ p_{(1^{a_1},2)} = \frac{a_2}{\alpha D} & \text{if } e = a_1 \end{cases}.$$

Since $f_k(e)$ is bounded, we have

$$E_N = \sum_{k=0}^{[\log_\alpha N]} \mathbf{E}\, Y_{k,N} = \sum_{k=A(N)}^{B(N)} \mathbf{E}\, Y_{k,N} + \mathcal{O}\left((\log N)^\eta\right)$$

and get (with $f_k(0)=0$)

$$E_N = \sum_{k=A(N)}^{B(N)} \mu_k + \mathcal{O}\left(\rho^{(\log N)^\eta/2}\right) + \mathcal{O}\left((\log N)^\eta\right) = M(N) + \mathcal{O}\left((\log N)^\eta\right).$$

The variance is clearly given by

$$\mathbf{Var}\left(\sum_{k=0}^{[\log_\alpha N]} f_k(X_k)\right) = \sum_{k,k'=0}^{[\log_\alpha N]} \Big(\mathbf{E}\left(f_k(X_k)f_{k'}(X_{k'})\right) - \mathbf{E}\, f_k(X_k)\mathbf{E}\, f_{k'}(X_{k'})\Big)$$

and

$$\mathbf{E}\left(f_k(X_k)f_{k'}(X_{k'})\right) - \mathbf{E}\, f_k(X_k)\mathbf{E}\, f_{k'}(X_{k'}) = \sum_{e,e'=0}^{a_1} (\hat{p}_{k,k',e,e'} - p_e p_{e'}) f_k(e) f_{k'}(e').$$

Since the eigenvalues of M are $\frac{\alpha_1}{\alpha}, \ldots, \frac{\alpha_d}{\alpha}$ (with $\alpha_1 = \alpha$), we have, for $k < k'$,

$$\hat{p}_{j,k,(.,b),(.,b')} = \sum_{i=1}^{d} p^{(i)}_{(.,b),(.,b')} \left(\frac{\alpha_i}{\alpha}\right)^{k-j}$$

with (easily determined) constants $p^{(i)}_{(.,b),(.,b')}$ and $p^{(1)}_{(.,b),(.,b')} = p_{(.,b)}p_{(.,b')}$. Since the $\hat{p}_{k,k',e,e'}$ are (weighted) sums of $\hat{p}_{k,k',(.,b),(.,b')}$, we have

$$\hat{p}_{k,k',e,e'} = \sum_{i=1}^{d} p^{(i)}_{e,e'} \left(\frac{\alpha_i}{\alpha}\right)^{k-k'},$$

where the constants $p^{(i)}_{e,e'}$ are the respective sums of $p^{(i)}_{(.,b),(.,b')}$. Note that $p^{(1)}_{e,e'} = p_e p_{e'}$. With these $p^{(i)}_{e,e'}$ we get $D(N)^2 = \mathbf{Var}\left(\sum_{k=0}^{[\log_\alpha N]} f_k(X_k)\right)$.

In Lemma 1.1, which is also valid for G-additive functions, we need $\frac{D(N)}{(\log N)^\eta} \to \infty$ for some $\eta > 0$. We prove $D(N) \gg \log N$ if the variances of $f_k(X_k)$ have a uniform lower bound.

Lemma 3.4. *Suppose that there exists a constant $c > 0$ such that $\sigma^{(2)}_{k,k} \geq c$ for all $k \geq 0$. Then we have a constant w such that*

$$\mathbf{Var}\left(\sum_{k=s}^{s'-1} f_k(X_k)\right) \geq w(s'-s) \tag{3.12}$$

for all $s, s' \geq 0$ with $s' - s \geq 3d$.

Proof. Set $X'_k = f_k(X_k) - \mathbf{E}\, f_k(X_k)$ and $S = \sum_{k=s}^{s'-1} X'_k$. Then $\mathbf{Var}\, X'_k = \sigma^{(2)}_{k,k} \geq c$ and $\mathbf{Var}\left(\sum_{k=s}^{s'-1} f_k(X_k)\right) = \mathbf{E}\, S^2$.

In [13], Dobrušin's work [9] is used to prove $\mathbf{E}\, S^2 \geq c(s'-s)\beta/100$, where β is the ergodicity coefficient

$$\beta = 1 - \sup_{m,b,m',b',\mathcal{A}} \left| \mathbf{Pr}[X_k \in \mathcal{A} | X_{k+1} = (m,b)] - \mathbf{Pr}[X_k \in \mathcal{A} | X_{k+1} = (m',b')] \right|$$

(which does not depend on k). Hence the lemma is proved, if we have $\beta > 0$.

If all a_i are non-zero, we have $p_{(m,b),(\Lambda,1)} > 0$ for all possible (m,b). Therefore, if $(\Lambda,1) \in \mathcal{A}$, we have

$$\mathbf{Pr}[X_k \in \mathcal{A} | X_{k+1} = (m,b)] > 0 \text{ for all } (m,b)$$

and the difference cannot be 1. If $(\Lambda,1) \notin \mathcal{A}$, the difference cannot be 1, because we have

$$\mathbf{Pr}[X_k \in \mathcal{A} | X_{k+1} = (m,b)] < 1 \text{ for all } (m,b).$$

By construction, the transition probabilities attain just finitely many values. Therefore we have $\beta > 0$.

If $a_b = 0$ for some b $(1 < b < d)$, then

$$\mathbf{Pr}[X_k = (\Lambda, b+1) | X_{k+1} = (1^{a_{b-1}}, b)] = 1$$

and

$$\mathbf{Pr}[X_k = (\Lambda, b+1) | X_{k+1} = (\Lambda, 1)] = 0,$$

hence $\beta = 0$. Then we need a result of Giesbrecht. In [19] he proved $\mathbf{E}\, S^2 \geq \frac{b_0}{2} \sum\limits_{k=s}^{s'-1} \tilde{p}_{k-u}^{(v)} \tilde{\lambda}_k^{(u)}$ with a constant b_0 and the following definitions:

$$\tilde{p}_k^{(v)} = \begin{cases} \sum_{m,b} \varphi_k^{(v)}(m,b) & \text{if } k \in [s+u+v, s'] \\ 0 & \text{else} \end{cases},$$

where $\varphi_k^{(v)}$ satisfies $P[X_{k-v} = (m,b) | X_k = (m',b')] \geq \varphi_k^{(v)}(m,b)$ for all (m,b) and (m',b'),

$$\tilde{\lambda}_k^{(u)} = \sup_{\gamma \geq 0} g_{k-u}^{(v)}(\gamma) h_k^{(u)}(\gamma),$$

$$g_k^{(v)}(\gamma) = \sup_t \inf_{m,b,\mathcal{A}} \mathbf{Pr}\left[\left| \sum_{j=k-u}^{k-1} X_k' - t \right| \leq \gamma | X_k = (m,b), X_{k-v} \in \mathcal{A} \right],$$

$$h_k^{(u)}(\gamma) = \inf_{m,b} \max\{0, (\chi_k^{(u)}(m,b) - 4\gamma\psi_k^{(u)}(m,b))\},$$

$$\chi_k^{(u)}(m,b) = \int_{-\infty}^{\infty} t^2 \mu_k^{(u)}(m,b,dt), \qquad \psi_k^{(u)}(m,b) = \int_{-\infty}^{\infty} |t| \mu_k^{(u)}(m,b,dt),$$

$$\mu_k^{(u)}(m,b,t) = \mathbf{Pr}\left[\left(\sum_{j=k-u}^{k-1} X'_k\right) - \left(\sum_{j=k-u}^{k-1} \tilde{X}'_k\right) < t | X_k = (m,b)\right],$$

where $(\tilde{X}'_j)_{j\geq 0}$ is a copy of $(X'_j)_{j\geq 0}$ and the two chains are independent from each other.

We have $\mathbf{Pr}[X_{k-d+1} = (m,b)|X_k = (m',b')] > 0$ for all k,m,b,m',b' and this probability takes only finitely many values. Therefore we obtain $\varphi_k^{(d-1)} \geq c'$ for some constant c'. For $g_k^{(d-1)}(\gamma)$ we get

$$\begin{aligned} g_k^{(d-1)}(\gamma) \geq \inf_{m,b,\mathcal{A}} \max_{m_{k-d+1},b_{k-d+1},\ldots,m_{k-1},b_{k-1}} & \mathbf{Pr}[X_{k-d+1} = (m_{k-d+1}, b_{k-d+1}), \\ & \ldots, X_{k-1} = (m_{k-1}, b_{k-1}) | X_k = (m,b), X_{k-1} \in \mathcal{A}] \end{aligned}$$

for all $\gamma \geq 0$. There always exist values $m_{k-d+1}, b_{k-d+1}, \ldots, m_{k-1}, b_{k-1}$ for which this probability is non-zero and this probability takes again only finitely many values. Therefore we have $g_k^{(d-1)}(\gamma) \geq c''$ for some constant c''.

The f_k are bounded, say $f_k(e) \leq C$ for all $k \geq 0$ and $e \in \{0,\ldots,a_1\}$. Therefore $\psi_k^{(d)}(x) \leq \left|\left(\sum_{j=k-d}^{k-1} X'_k\right) - \left(\sum_{j=k-d}^{k-1} \tilde{X}'_k\right)\right| \leq 2dC$. On the other hand, we have $|f_{k-d}(e) - f_{k-d}(0)| > \sqrt{c}$ for some digit e and

$$\begin{aligned} \mathbf{Pr}[X_{k-1} = (\Lambda, b_{k-1}), \ldots, X_{k-d+2} = (\Lambda, b_{k-d+2}), X_{k-d+1} = (\Lambda, 1), \\ X_{k-d} = (\Lambda, 1) | X_k = (m,b)] > 0, \end{aligned}$$

$$\begin{aligned} \mathbf{Pr}[\tilde{X}_{k-1} = (\Lambda, b_{k-1}), \ldots, \tilde{X}_{k-d+2} = (\Lambda, b_{k-d+2}), \tilde{X}_{k-d+1} = (\Lambda, 1), \\ \tilde{X}_{k-d} = (1^e, b_{k-d}) | X_k = (m,b)] > 0 \end{aligned}$$

for some (unique) b_i and all (m,b). Hence $\mu_k^{(d)}(m,b,t)$ jumps at some point $t > \sqrt{c}$ at least by some constant c''' and we have $\chi_k^{(d)}(m,b) \geq cc'''$.

These results do not depend on γ. With $\gamma = \frac{cc'''}{16dC}$ we obtain $h_k^{(d)} \geq cc''c'''/2$ and (with $u = d$, $v = d-1$) $\mathbf{E}\, S^2 \geq (s'-s-3d+1)cc'c''c'''b_0/4$. Hence the lemma is proved and $w \geq cc'c''c'''b_0/(12d)$. $\square$

Immediately we get the following corollary.

Corollary 3.1. *Suppose that there exists a constant $c > 0$ such that $\sigma_{k,k}^{(2)} \geq c$ for all $k \geq 0$. Then we have*

$$D(N)^2 \gg \log N \text{ and } \overline{D}(N)^2 \gg \log N.$$

In order to prove (3.6) it suffices, because of Lemma 1.1, to show that the moments

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

with

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

converge to the corresponding moments of the normal law. This implies

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

and, by Lemma 1.1, (3.5). First we prove a central limit theorem (with convergence of moments) for the exact Markov chain.

Lemma 3.5. *Suppose that there exists a constant $c > 0$ such that $\sigma_{k,k}^{(2)} \geq c$ for all $k \geq 0$. Then the sums of the random variables $f_k(X_k)$ satisfy a central limit theorem. More precisely*

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

and for all $h \geq 0$ we have

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

as $N \to \infty$.

Proof. If all a_i are non-zero, then the ergodicity coefficient β is positive and the lemma can be proved with the help of Theorem 4 of Lifšic [26]. If $\beta = 0$, we have to adapt this theorem.

An inspection of Lifšic' proof and Dobrušin's work [9] shows that we get the same result if we replace the ergodicity coefficient β by a constant $\theta > 0$ that satisfies

$$\gamma_j = \frac{1}{2} \sup_{m,b} \sum_{m',b'} \left| \mathbf{Pr}[X_k = (m',b') | X_{k+j} = (m,b)] - \mathbf{Pr}[X_k = (m',b')] \right| \leq (1-\theta)^j \quad (3.13)$$

for all $j \geq 1$ and

$$\mathbf{Var} \left(\sum_{k=s}^{s'-1} f_k(X_k) \right) \geq c(s'-s)\theta \qquad (3.14)$$

for all $s, s' \geq 0$ with $s' - s \geq s_0$ for some constant s_0.

We have $\gamma_j > 0$ for all $j \geq 1$ since the sum in (3.13) is always less than 1 and we only have a finite number of states (m, b). Dobrušin [9] proved $\gamma_j \leq 1 - \beta_j$, where

$$\beta_j = 1 - \sup_{m,b,m',b',\mathcal{A}} \left| \mathbf{Pr}[X_k \in \mathcal{A} | X_{k+j} = (m,b)] - \mathbf{Pr}[X_k \in \mathcal{A} | X_{k+j} = (m',b')] \right|.$$

For some j_0 with $1 < j_0 < d$ we have $\mathbf{Pr}[X_k = (\Lambda, 1) | X_{k+j} = (m,b)] > 0$ for all possible (m, b) and all $j \geq j_0$. This implies $\beta_j > 0$ for all $j \geq j_0$ and we define

$$\theta = \min \left(1 - \max_{1 \leq k < j_0} \gamma_k^{1/k}, 1 - \max_{j_0 \leq k < 2j_0} (1 - \beta_k)^{1/k}, \frac{w}{c} \right).$$

Then (3.14) holds because of (3.12). Because of $\gamma_j \leq 1 - \beta_j$, (3.13) holds for $j < 2j_0$. For $j \geq 2j_0$, we apply the inequality $1 - \beta_{i+j} \leq (1 - \beta_i)(1 - \beta_j)$ (see Dobrušin [9]) and get, by induction on q,

$$1 - \beta_{qj_0+t} \leq (1 - \beta_{j_0})(1 - \beta_{(q-1)j_0+t}) \leq (1-\theta)^{j_0}(1-\theta)^{(q-1)j_0+t} = (1-\theta)^{qj_0+t}$$

for $q \geq 2$, $t < j_0$. Hence θ satisfies the required properties, we can apply the (adapted) theorem of Lifšic and the lemma is proved. □

The next lemma concludes the proof of Theorem 3.1. In particular, for $h = 2$, it implies together with Lemma 1.1 and (3.3) the asymptotics for the variance (3.4).

Lemma 3.6. *For every $h \geq 1$ and every $\lambda > 0$ we have*

$$\frac{1}{N} \sum_{n<N} \left(\frac{\overline{f}^{(N)}(n) - \overline{M}(N)}{\overline{D}(N)} \right)^h - \mathbf{E} \left(\frac{\sum_{k=A(N)}^{B(N)} f_k(X_k) - \overline{M}(N)}{\overline{D}(N)} \right)^h \to 0.$$

Proof. The first term is the sum over all integers $A(N) \leq k_1, \ldots, k_h \leq B(N)$ of

$$\frac{1}{N} \sum_{n<N} \prod_{j=1}^{h} \frac{f_{k_j}(\epsilon_{k_j}(n)) - \mu_{k_j}}{\overline{D}(N)}$$
$$= \sum_{e_1=0}^{a_1} \cdots \sum_{e_h=0}^{a_1} \#\{n < N \mid \epsilon_{k_1}(n) = e_1, \ldots, \epsilon_{k_h}(n) = e_h\} \prod_{j=1}^{h} \frac{f_{k_j}(e_j) - \mu_{k_j}}{\overline{D}(N)}.$$

The second term is the sum over all integers $A(N) \leq k_1, \ldots, k_h \leq B(N)$ of

$$\mathbf{E} \left(\prod_{j=1}^{h} \frac{f_{k_j}(X_{k_j}) - \mu_{k_j}}{\overline{D}(N)} \right)$$
$$= \sum_{e_1=0}^{a_1} \cdots \sum_{e_h=0}^{a_1} \mathbf{Pr}[X_1 = e_1, \ldots, X_h = e_h] \prod_{j=1}^{h} \frac{f_{k_j}(e_j) - \mu_{k_j}}{\overline{D}(N)}.$$

Hence, with Lemma 3.3, the convergence is valid with an error term of the form $\mathcal{O}\left((\log N)^{\lambda+h-h\eta}\right)$. $\square$

3.3 Tilings

The proof of Theorem 1.1 relies essentially on the fact that the value of $\epsilon_k(n)$ can be determined without using the greedy algorithm, namely by

$$\epsilon_k(n) = e \Longleftrightarrow \left\{\frac{n}{q^{k+1}}\right\} \in \left[\frac{e}{q}, \frac{e+1}{q}\right).$$

In order to get an analogue to Theorem 1.1 for G-ary expansions, we need a similar characterisation of the digits. It turns out that we need a tiling of the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$, i.e. a family of sets $(\Omega_e)_{e\in\{0,\ldots,a_1\}}$ such that

- $\bigcup_{e=0}^{a_1} \Omega_e = \mathbb{T}^d$,
- each of the Ω_e is the closure of its interior,
- the intersection of two different Ω_e has Lebesgue measure zero,

and vectors $\mathbf{v}(n,k) \in \mathbb{T}^d$ such that

$$\epsilon_k(n) = e \Longleftrightarrow \mathbf{v}(n,k) \in \Omega_e. \tag{3.15}$$

For q-ary expansions we have $\frac{1}{q^j}\#\{n < q^j \mid \epsilon_k(n) = e\} = \frac{1}{q}$ for all $j > k$. In our case we have $\frac{1}{G_j}\#\{n < G_j \mid \epsilon_k(n) = e\} = p_e + \mathcal{O}\left(\alpha^{-\min(k,j-k)}\right)$ for all $j > k$. Therefore we obtain tilings with $\lambda_d(\Omega_e) = p_e$, where λ_d denotes the d-dimensional Lebesgue measure, which satisfy (3.15) only up to an error term of $\mathcal{O}\left(\alpha^{-k}\right)$.

Unfortunately, we have to make some restrictions on the sequence G: we need $a_d = 1$, α has to be a Pisot number with minimal polynomial $\chi(x)$, i.e. $|\alpha_i| < 1$ for $2 \le i \le d$, and

$$\mathrm{Fin}(\alpha) = \mathbb{Z}[\alpha^{-1}] \cap \mathbb{R}^+, \tag{3.16}$$

where $\mathrm{Fin}(\alpha)$ denotes the set of non-negative real numbers with finite α-expansion, i.e.

$$\left\{x \in \mathbb{R}^+ \,\middle|\, x = \sum_{k=-L}^{M} \epsilon_k \alpha^k \text{ with } (\epsilon_k, \ldots, \epsilon_{k-d+1}) < (a_1, \ldots, a_d) \text{ for all } k \le M\right\}.$$

Proposition 3.1. *Let G be as in Section 3.1 with $a_d = 1$, irreducible characteristic polynomial $\chi(x)$ and its dominant root α a Pisot number which satisfies (3.16).*

$$\mathbf{v}(n,k) = \frac{n}{\alpha^k}\frac{\alpha-1}{\alpha^d-1}\left(\alpha^{d-1}, \ldots, \alpha, 1\right)^t \in \mathbb{T}^d.$$

Then we have a tiling $(\Omega_e)_{e\in\{0,\dots,a_1\}}$ of $\mathbb{T}^d$ with

$$d(\mathbf{v}(n,k),\Omega_{\epsilon_k(n)}) = \mathcal{O}\left(\alpha^{-k}\right) \text{ for all } k,n\in\mathbb{N}, \tag{3.17}$$

where $d(\mathbf{x},S)=\inf_{\mathbf{y}\in S}\|\mathbf{x}-\mathbf{y}\|_\infty$.

Remark 3.2. We have $d(\mathbf{v}(n,k),\Omega_{\epsilon_k(n)})>0$, i.e. $\mathbf{v}(n,k)\notin\Omega_{\epsilon_k(n)}$ only for a small number of n and k (see Lemma 3.10).

Proof. We regard the linear map

$$\phi=\begin{pmatrix} a_1 & a_2 & \cdots & \cdots & a_d\\ 1 & 0 & \cdots & \cdots & 0\\ 0 & 1 & \ddots & & \vdots\\ \vdots & \ddots & \ddots & \ddots & \vdots\\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}\in \mathrm{GL}(d,\mathbb{Z})$$

with eigenvalues $\alpha,\alpha_2,\dots,\alpha_d$. Since α is a Pisot number and $a_d=1$, ϕ is a hyperbolic toral automorphism and we have a ϕ-invariant decomposition of $\mathbb{R}^d$ into the unstable eigenspace $E_u=\mathbb{R}(\alpha^{d-1},\dots,\alpha,1)^t$ and the stable eigenspace E_s (of dimension $d-1$). Let $\mathbf{e}_u=\pi_u\left((1,0,\dots,0)^t\right)$ and $\mathbf{e}_s=\pi_s\left((1,0,\dots,0)^t\right)$ with $\pi_u:\mathbb{R}^d\to E_u$ the projection along E_s to E_u and $\pi_s:\mathbb{R}^d\to E_s$ the projection along E_u to E_s. Set $\mathbf{e}_u=c_1'(\alpha^{d-1},\dots,\alpha,1)^t$.

Then the sequence $(G_j')_{j\ge 0}$ defined by the linear recurrence

$$G_j'=a_1G_{j-1}'+\cdots+a_dG_{j-d}' \text{ for } j\ge d$$

with initial values $G_0'=0,\ \dots,\ G_{d-2}'=0,\ G_{d-1}'=1$ satisfies

$$G_j'=c_1'\alpha^j+c_2'\alpha_2^j+\cdots+c_d'\alpha_d^j$$

for some constants $c_2',\dots,c_d'$. By induction on j, we can prove the equation

$$G_j=G_j'+G_{j+1}'+\cdots+G_{j+d-1}'.$$

Because of $G_j\to c_1\alpha^j$, $G_j'\to c_1'\alpha^j$ for $j\to\infty$, we have

$$c_1=c_1'(1+\alpha+\cdots+\alpha^{d-1}).$$

With

$$n=c_1\sum_{j=0}^{\infty}\epsilon_j(n)\alpha^j+\mathcal{O}\left(1\right),$$

we obtain

$$\mathbf{v}(n,k)=\frac{n}{c_1\alpha^k}\mathbf{e}_u=\sum_{j=0}^{\infty}\epsilon_j(n)\alpha^{j-k}\mathbf{e}_u+\mathcal{O}\left(\alpha^{-k}\right)=\sum_{j=0}^{\infty}\epsilon_j(n)\phi^{j-k}(\mathbf{e}_u)+\mathcal{O}\left(\alpha^{-k}\right).$$

Clearly we have

$$\phi^j(\mathbf{e}_u) + \phi^j(\mathbf{e}_s) = \phi^j\left((1,0,\dots,0)^t\right) \in \mathbb{Z}^d \text{ for all } j \geq 0$$

and thus

$$\mathbf{v}(n,k) \equiv \underbrace{\sum_{j=0}^{k-1} \epsilon_j(n)\phi^{j-k}(\mathbf{e}_u) - \sum_{j=k}^{\infty} \epsilon_j(n)\phi^{j-k}(\mathbf{e}_s)}_{\mathbf{v}'(n,k)} + \mathcal{O}\left(\alpha^{-k}\right) \bmod \mathbb{Z}^d. \tag{3.18}$$

Set

$$\Omega_e = \operatorname{Clos}\left\{\mathbf{v}'(n,k) : k,n \in \mathbb{N} \text{ with } \epsilon_k(n) = e\right\}.$$

Then we know by Praggastis [31] that $(\Omega_e)_{e\in\{0,\dots,a_1\}}$ is a tiling of $\mathbb{T}^d$ if $\operatorname{Fin}(\alpha) = \mathbb{Z}[\alpha] \cap \mathbb{R}_+$. We have $\mathbb{Z}[\alpha] = \mathbb{Z}[\alpha^{-1}]$, because the characteristic polynomial is monic and $a_d = 1$. Hence (3.16) implies that $(\Omega_e)_{e\in\{0,\dots,a_1\}}$ is a tiling and (3.17) holds because of (3.18). □

Remark 3.3. For $d = 2$, these tilings consist of rectangles which are given in the following example. For $d \geq 3$, the involved sets always have fractal boundary.

Example. Figure 3.1 gives an example of the rectangles in case $d = 2$. Here we have $a_1 = 3$ (and clearly $a_2 = 1$). Figure 3.2 makes clear that $(\Omega_e)_{0\leq e\leq 3}$ is a tiling of $\mathbb{T}^2$. Note that for these pictures $\mathbf{v}(n,k)$ is slightly modified, namely

$$\mathbf{v}(n,k) = \frac{n}{\alpha^{k+1}(\alpha+1)}(\alpha,1)^t.$$

We can give the rectangles as the convex hull of their corners (see [13]):

$$\Omega_0 = \operatorname{convhull}\left(\left(-\frac{1}{D},\frac{\alpha}{D}\right),\ (0,1),\ \left(\frac{\alpha^{-1}+1}{D},\frac{\alpha^{-1}-1}{D}\right),\ \left(\frac{\alpha^{-1}}{D},-\frac{1}{D}\right)\right)$$

$$\Omega_e = \operatorname{convhull}\left(\left(\frac{\alpha^{-1}+e-1}{D}, 1-\frac{1+\alpha^{-1}(1-e)}{D}\right),\ \left(\frac{\alpha^{-1}+e}{D}, 1-\frac{1-\alpha^{-1}e}{D}\right),\right.$$
$$\left.\left(\frac{\alpha^{-1}+e+1}{D},\frac{\alpha^{-1}(e+1)-1}{D}\right),\ \left(\frac{\alpha^{-1}+e}{D},\frac{\alpha^{-1}e-1}{D}\right)\right) \text{ for } e \in \{1,\dots,a_1-1\}$$

$$\Omega_{a_1} = \operatorname{convhull}\left(\left(\frac{\alpha-1}{D},\frac{a_1}{D}\right),\ \left(1-\frac{1}{D},\frac{\alpha}{D}\right),\ (1,0),\ \left(\frac{\alpha}{D},-\frac{1}{\alpha^2 D}\right)\right).$$

Example. Figure 3.3 shows the sets Ω_e for the Tribonacci expansion ($d = 3$, $a_1 = a_2 = a_3 = 1$). Ω_0 is the largest of the three prisms and Ω_1 is

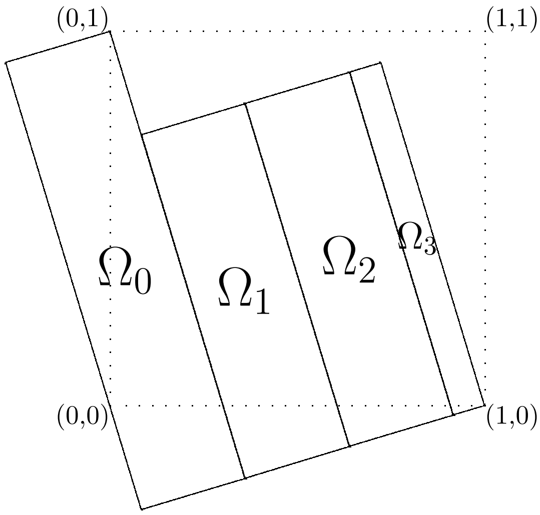

Figure 3.1: $\Omega_0, \Omega_1, \Omega_2, \Omega_3$ for $d = 2$, $a_1 = 3$

the union of the two smaller ones. $\pi_s(\Omega_0)$ is the Rauzy fractal (for details on the Rauzy fractal see Messaoudi [28, 29] and Rauzy [32] for the original work). Figure 3.4 illustrates how (Ω_0, Ω_1) tiles $\mathbb{R}^3$. These figures were drawn by Siegel, who studied in [36] substitutions of Pisot type.

For the proof of Proposition 3.3, we will need a covering of Ω_e and its boundary by convex sets. Since the boundary of Ω_e has fractal structure for $d > 2$, we approximate it by parallelepipeds.

Each Ω_e is the union of sets

$$\Omega_{e_0,\dots,e_{d-2}} = \mathrm{Clos}\left\{\mathbf{v}'(n,k) : k, n \in \mathbb{N} \text{ with } (\epsilon_k(n), \dots, \epsilon_{k+d-2}(n)) = (e_0, \dots, e_{d-2})\right\}$$

(with $e_0 = e$), which are prisms:

$$\Omega_{e_0,\dots,e_{d-2}} = \pi_s(\Omega_{e_0,\dots,e_{d-2}}) \oplus [0, \sup_{k,n \text{ as above}} \sum_{j=0}^{k-1} \epsilon_j(n)\alpha^{j-k}]\mathbf{e}_u.$$

Therefore we study the boundary of $\pi_s(\Omega_{e_0,\dots,e_{d-2}})$.

The problem of determining all points on the boundary is equivalent to determining all points with more than one ϕ-representation, which can be done with the help of a finite automaton. This method is adapted from Messaoudi [29] who examined the Rauzy fractal. Siegel [36] studied similar problems with similar automata.

Let $\mathcal{N}$ be the set of sequences $(b_j)_{j\in\mathbb{Z}}$ with

$$(b_j, b_{j-1}, \dots, b_{j-d+1}) < (a_1, a_2, \dots, a_d) \text{ for all } j \in \mathbb{Z}$$

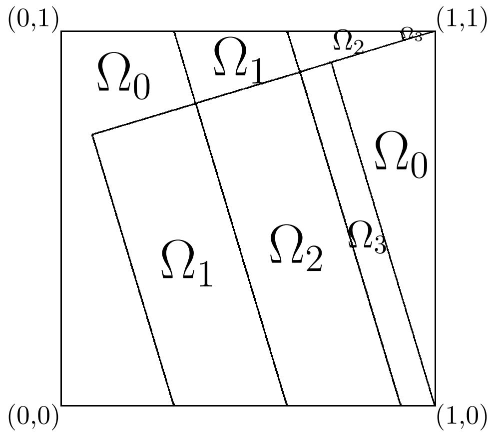

Figure 3.2: $\Omega_0, \Omega_1, \Omega_2, \Omega_3$ for $d = 2$, $a_1 = 3$ in $\mathbb{T}^2$

and an integer K such that $b_j = 0$ for $j \geq K$. Let $\mathcal{N}_f$ be the set of sequences $(b_j)_{j\in\mathbb{Z}} \in \mathcal{N}$ with an integer J such that $b_j = 0$ for $j \leq J$. With $\mathcal{E} = \left\{ \sum_{j=1}^{\infty} \epsilon_j \phi^j(\mathbf{e}_s) \,|(\epsilon_j)_{j\geq 1} \in \mathcal{N}_f \right\}$, we get the following proposition.

Proposition 3.2 (cf. [29], Théorème 1). *Let* $\mathbf{x} = \sum_{j=-L}^{\infty} b_j \phi^j(\mathbf{e}_s)$ *and* $\mathbf{y} = \sum_{j=-L}^{\infty} b'_j \phi^j(\mathbf{e}_s)$*, where* $(b_j)_{j\geq -L} \in \mathcal{N}$ *and* $(b'_j)_{j\geq -L} \in \mathcal{N}$*, then* $\mathbf{x} = \mathbf{y}$ *if and only if we have, for all* $i \geq -L$*,*

$$\mathbf{x}_i - \mathbf{y}_i \in S$$

where $\mathbf{x}_i = \phi^{-i}\left(\sum_{j=-L}^{i} b_j \phi^j(\mathbf{e}_s)\right)$, $\mathbf{y}_i = \phi^{-i}\left(\sum_{j=-L}^{i} b'_j \phi^j(\mathbf{e}_s)\right)$ *and*

$$S = \left\{ \pm \sum_{j=-s}^{0} \epsilon_j \phi^j(\mathbf{e}_s) : (\epsilon_j)_{-s\leq j\leq 0} \in \mathcal{N}_f, \mathcal{E} \cap \left(\mathcal{E} \pm \sum_{j=-s}^{0} \epsilon_j \phi^j(\mathbf{e}_s) \right) \neq \emptyset \right\}.$$

for some (fixed) integer s.

We need two small lemmata for the proof of Proposition 3.2.

Lemma 3.7. *For all integers* $j \geq d - 1$*, we have*

$$\alpha^j = \alpha^{d-1} G'_j + \alpha^{d-2}(a_2 G'_{j-1} + a_3 G'_{j-2} + \cdots + a_d G'_{j-d+1}) + \cdots + \alpha(a_{d-1} G'_{j-1} + a_d G'_{j-2}) + a_d G'_{j-1}, \quad (3.19)$$

where the sequence $(G'_j)_{j\geq 0}$ *is defined in the proof of Proposition 3.1.*

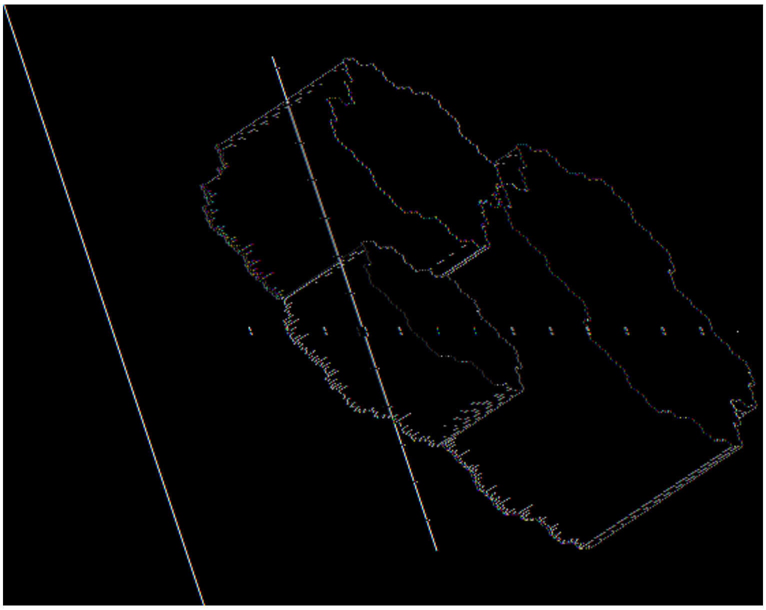

Figure 3.3: Ω_0, Ω_1 for the Tribonacci expansion

Proof. Induction on j. ☐

Lemma 3.8. *Define the linear map*

$$\kappa : \left\{ \pm \sum_{j=-\infty}^{\infty} \epsilon_j \phi^j(\mathbf{e}_s) : (\epsilon_j)_{j\in\mathbb{Z}} \in \mathcal{N}_f \right\} \to \pm Fin(\alpha)$$

by $\kappa(\phi^j(\mathbf{e}_s)) = \alpha^j$ *for all* $j \in \mathbb{Z}$. *Then* κ *is well defined and a bijection.*

Proof. Clearly κ is a bijection, if it is well defined.

We show that all elements on the left side are distinct. Suppose that two representations $\varepsilon \sum_{j=-\infty}^{\infty} \epsilon_j \phi^j(\mathbf{e}_s)$ and $\varepsilon' \sum_{j=-\infty}^{\infty} \epsilon'_j \phi^j(\mathbf{e}_s)$ with $(\epsilon_j)_{j\in\mathbb{Z}}, (\epsilon'_j)_{j\in\mathbb{Z}} \in \mathcal{N}_f$ and $\varepsilon, \varepsilon' \in \{\pm 1\}$ represent the same vector. Hence we have $Q(\phi)(\mathbf{e}_s) = \mathbf{0}$ for some polynomial $Q = q_m x^m + \cdots + q_1 x + q_0 \not\equiv 0$ (after applying some power of ϕ).

The proof of Proposition 3.1 shows $\phi^j(\mathbf{e}_s) = \sum_{i=2}^{d} c'_i \alpha_i^j (\alpha_i^{d-1}, \ldots, \alpha_i, 1)^t$. Hence $\sum_{j=0}^{m} q_j \sum_{i=2}^{d} c'_i \alpha_i^j (\alpha_i^{d-1}, \ldots, \alpha_i, 1)^t = \mathbf{0}$. By easy calculations (solution of a linear equation system), we get $c'_i = \left(\prod_{k\neq i}(\alpha_i - \alpha_k)\right)^{-1} \neq 0$. If $\alpha_i \in \mathbb{R}$ for all $i \in \{2, \ldots, d\}$, then the $(\alpha_i^{d-1}, \ldots, \alpha_i, 1)^t$ are linearly independent vectors of $\mathbb{R}^d$ and we must have $Q(\alpha_i) = 0$ for all $i \in \{2, \ldots, d\}$. For $\alpha_i \notin \mathbb{R}$ we get $Q(\alpha_i) = 0$ similarly.

This implies $Q(\alpha) = 0$ and $\varepsilon \sum_{j=-\infty}^{\infty} \epsilon_j \alpha^j = \varepsilon' \sum_{j=-\infty}^{\infty} \epsilon'_j \alpha^j$. Therefore we have $\varepsilon = \varepsilon'$ and, since finite α-representations are unique, $(\epsilon_j)_{j\in\mathbb{Z}} = (\epsilon'_j)_{j\in\mathbb{Z}}$.

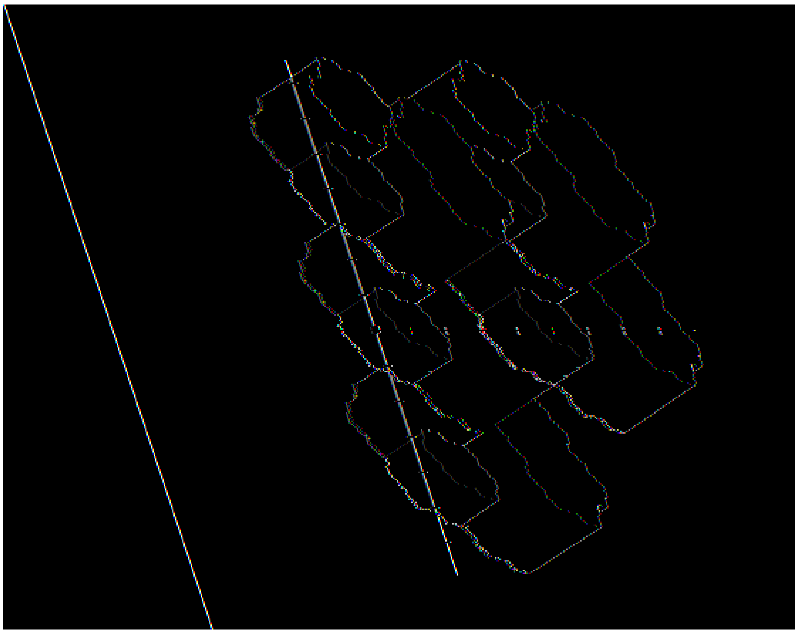

Figure 3.4: Tiling of $\mathbb{R}^3$ for the Tribonacci expansion

Thus κ is well defined and the lemma proved. $\square$

Proof of Proposition 3.2. Since $\phi|_{E_s}$ is contracting, we have

$$\mathbf{x} - \mathbf{y} = \lim_{i\to\infty} \phi^{i-d+1}(\mathbf{x}_i - \mathbf{y}_i) = \mathbf{0},$$

if $\mathbf{x}_i - \mathbf{y}_i \in S$.

To show the other direction of the implication, we suppose $\mathbf{x} = \mathbf{y}$. Hence $\phi^{-i}(\mathbf{x}) = \phi^{-i}(\mathbf{y})$ and

$$\mathbf{x}_i - \mathbf{y}_i = \sum_{j=i+1}^{\infty} (b'_j - b_j)\phi^{j-i}(\mathbf{e}_s) = \sum_{j=1}^{\infty} (b'_{j+i} - b_{j+i})\phi^j(\mathbf{e}_s).$$

On the other hand we have

$$\mathbf{x}_i - \mathbf{y}_i = \phi^{-i}\left(\sum_{j=-L}^{i} (b_j - b'_j)\phi^j(\mathbf{e}_s)\right) = \phi^{-L-i-d+1}\left(\sum_{j=d-1}^{L+i+d-1} g_j\phi^j(\mathbf{e}_s)\right),$$

where $g_j = b_{j-L-d+1} - b'_{j-L-d+1}$. We apply κ and get by (3.19)

$$\kappa(\mathbf{x}_i - \mathbf{y}_i) = \alpha^{-L-i-d+1}\left(g'_{d-1}\alpha^{d-1} + \cdots + g'_1\alpha + g'_0\right)$$

with integers g'_j which are easily seen to be all positive if

$$(b_i, b_{i-1}, \ldots, b_{-L}) > (b'_i, b'_{i-1}, \ldots, b'_{-L})$$

and all negative if "<" holds. Hence we have $\kappa(\mathbf{x}_i - \mathbf{y}_i) \in \mathbb{Z}_+[\alpha^{-1}]$ and $\kappa(\mathbf{x}_i - \mathbf{y}_i) \in \mathbb{Z}_-[\alpha^{-1}]$ respectively. Because of (3.16), we have

$$\kappa(\mathbf{x}_i - \mathbf{y}_i) = \pm \sum_{j=-s}^{m} \epsilon_j \alpha^j \text{ with } (\epsilon_j)_{-s \le j \le m} \in \mathcal{N}_f. \tag{3.20}$$

Assume, w.l.o.g., $\kappa(\mathbf{x}_i) = \kappa(\mathbf{y}_i) + \sum_{j=-s}^{m} \epsilon_j \alpha^j$. Then (3.16) implies

$$\kappa(\mathbf{x}_i) = \sum_{j=-s'}^{m'} \epsilon'_j \alpha^j \text{ with } (\epsilon'_j)_{-s' \le j \le m'} \in \mathcal{N}_f \text{ and } m' \ge m.$$

Since $\kappa(\mathbf{x}_i) = \sum_{j=-L}^{i} b_j \alpha^{j-i}$ and finite α-expansions are unique, we have $m' = 0$ which implies $m \le 0$.

By applying κ^{-1} to (3.20), we get

$$\sum_{j=1}^{\infty} (b'_{j+i} - b_{j+i}) \phi^j(\mathbf{e}_s) = \pm \sum_{j=-s}^{0} \epsilon_j \phi^j(\mathbf{e}_s)$$

and

$$\sum_{j=1}^{\infty} b_{j+i} \phi^j(\mathbf{e}_s) \in \mathcal{E} \cap \left(\mathcal{E} \pm \sum_{j=-s}^{0} \epsilon_j \phi^j(\mathbf{e}_s) \right).$$

Lemma 2.10 of Praggastis [31] shows that we have an integer s such that $\left(\mathcal{E} \pm \sum_{j=-\infty}^{0} \epsilon_j \phi^j(\mathbf{e}_s)\right) = \emptyset$, if $\epsilon_j \neq 0$ for some $j < -s$. This concludes the proof of the proposition. $\square$

If we set $\mathbf{z}_i = \mathbf{x}_i - \mathbf{y}_i$, then

$$\mathbf{z}_{i+1} = \phi^{-1}(\mathbf{z}_i) + (b_{i+1} - b'_{i+1})\mathbf{e}_s.$$

Therefore the points with two representations are determined by a finite automaton, the states of which are the elements of S and two states $\mathbf{z}, \mathbf{z}'$ are connected by an edge labeled by (b, b'), if $\mathbf{z}' = \phi^{-1}(\mathbf{z}) + (b - b')\mathbf{e}_s$ or, equivalently, $\kappa(\mathbf{z}') = \kappa(\mathbf{z})/\alpha + (b - b')$. (The starting point is $\mathbf{0}$.)

As Gilbert [20] for the twin dragon, we obtain a ν-th approximation to the boundary by determining all paths of length ν in the automaton and drawing for each such path p a parallelepiped that contains the image of all paths which start with p. This is the idea of the following lemma.

Lemma 3.9. *For all $\nu \in \mathbb{N}$ and $e \in \{0, \ldots, a_1\}$, the boundary of Ω_e is contained in sets $U_{e,\nu}$ which are the union of $\mathcal{O}(\gamma^\nu)$ parallelepipeds of size $C\alpha^{-\nu}$ for some constants $\gamma < \alpha$ and C, with edges parallel to $\mathbf{a}_1, \ldots, \mathbf{a}_d$, where $\mathbf{a}_i = (\alpha_i^{d-1}, \ldots, \alpha_i, 1)^t$ for the real eigenvalues α_i ($\alpha_1 = \alpha$) and $\mathbf{a}_i = (\Re\alpha_i^{d-1}, \ldots, \Re\alpha_i, 1)^t$, $\mathbf{a}_{i+1} = (\Im\alpha_i^{d-1}, \ldots, \Im\alpha_i, 0)^t$ for the pairs of complex eigenvalues $(\alpha_i, \alpha_{i+1} = \overline{\alpha_i})$.*

Proof. A point can be on the boundary of Ω_e if its π_s-image has at least two ϕ-representations $\sum_{j=0}^{\infty} b_j \phi^j(\mathbf{e}_s) = \sum_{j=-L}^{\infty} b'_j \phi^j(\mathbf{e}_s)$ with $(b_0, \ldots, b_{d-2}) \neq (b'_0, \ldots, b'_{d-2})$, $b_0 = e$ and j_0 the smallest integer with $b_{j_0} \neq b'_{j_0}$. Denote by B_ν the number of different initial sequences $(b_0, \ldots, b_\nu)$ of points on the boundary. We show that these sequences cannot have $2s+2$ subsequent zeros.

Suppose on the contrary that $(b_{j_1+1}, \ldots, b_{j_1+2s+2}) = (0, \ldots, 0)$ for some $j_1 \geq j_0$. Set $\mathbf{z}_i = \sum_{j=j_0}^{i} (b_j - b'_j) \phi^{j-i}(\mathbf{e}_s)$. We have $\mathbf{z}_{j_0} \neq \mathbf{0}$ by definition and $\mathbf{z}_i \neq \mathbf{0}$ for all $i > j_0$, because $\kappa(\mathbf{z}_i) = 0$ would imply that two different finite α-representations are equal.

Assume $\kappa(\mathbf{z}_{j_1}) < 0$. Then we have $\kappa(\mathbf{z}_i) < 0$ for all $j_1 < i \leq j_1 + 2s + 2$ and the uniqueness of finite α-representations implies $\mathbf{z}_i \notin S$ for some i, $j_1 < i \leq j_1 + s + 1$, which contradicts Proposition 3.2. If $\kappa(\mathbf{z}_{j_1}) > 0$, then the uniqueness of finite α-representations implies $\kappa(\mathbf{z}_i) < 0$ or $\mathbf{z}_i \notin S$ for some i, $j_1 < i \leq j_1 + s + 1$. As above $\kappa(\mathbf{z}_i) < 0$ implies $\mathbf{z}_{i'} \notin S$ for some $i' \leq i + s + 1$.

Therefore $2s + 2$ subsequent zeros are not possible and $B_\nu = \mathcal{O}(\gamma^\nu)$ for some $\gamma < \alpha$.

The $\mathbf{a}_i$ are the real eigenvectors of ϕ and the real and imaginary parts of the complex eigenvectors respectively. Let c be the size of the parallelepiped that covers $\mathcal{E}$ and all its images of rotations in the planes spanned by the complex eigenvectors ($\mathcal{E}$ is a bounded set). Then all points on the boundary with same initial sequence $(b_0, \ldots, b_\nu)$ are covered by a parallelepiped of size $C|\alpha_2|^\nu \ldots |\alpha_d|^\nu = C\alpha^{-\nu}$ and we have B_ν of these parallelepipeds.

This concludes the proof of the lemma. □

3.4 Central limit theorem for polynomial sequences

Now, we can state the analogue to Theorem 1.1. Clearly we have to make the same restrictions on f as in Theorem 3.1 and the same restrictions on G as in Proposition 3.1.

Theorem 3.2. *Let G be as in Section 3.1 with $a_d = 1$, irreducible characteristic polynomial $\chi(x)$ and its dominant root α a Pisot number which satisfies (3.16). Let f be a G-additive function such that $f_k(e) = \mathcal{O}(1)$ as $k \to \infty$ for all $e \in \{0, 1, \ldots, a_1\}$ and assume that there exists a constant $c > 0$ such that $\sigma_{k,k}^{(2)} \geq c$ for all $k \geq 0$. Let $P(n)$ be a polynomial of degree r with integer coefficients and positive leading term. Then, as $N \to \infty$,*

$$\frac{1}{N} \# \left\{ n < N \,\middle|\, \frac{f(P(n)) - M(N^r)}{D(N^r)} < x \right\} \to \Phi(x) \tag{3.21}$$

and

$$\frac{1}{\pi(N)}\#\left\{p<N\,\middle|\,\frac{f(P(p))-M(N^r)}{D(N^r)}<x\right\}\to\Phi(x) \tag{3.22}$$

Remark 3.4. For $a_1 \geq a_2 \geq \cdots \geq a_d > 0$, we know from Brauer [4] that α is a Pisot number with minimal polynomial $\chi(x)$. (3.16) has been shown in this case by Frougny and Solomyak [18]. Hence Theorem 3.2 holds for these sequences.

For $d = 3, a_2 = 0, a_3 = 1$, α_2 and α_3 are complex numbers and have therefore absolute value $1/\sqrt{\alpha}$. For these a_i, the equation (3.16) was shown by Akiyama [1]. Hence Theorem 3.2 holds for these sequences too and the only restriction in case $d = 3$ is $a_3 = 1$.

Remark 3.5. α may not be a Pisot number (e.g. the dominant root of $x^6 - x^5 - 1$). We also have α which are Pisot units, but do not satisfy (3.16): let α be the dominant root of $x^4 - x^3 - 1$. Then the α-expansion of 2 is $10.010(00001)^\infty$.

We have to prove the following analogue to Proposition 1.1.

Proposition 3.3. *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*

$$\frac{1}{N}\#\{n<N:\epsilon_{k_1}(P(n))=e_1,\ldots,\epsilon_{k_h}(P(n))=e_h\}=\hat{p}_{k_1,\ldots,k_h,e_1,\ldots,e_h} + \mathcal{O}\left((\log N)^{-\lambda}\right)$$

and

$$\frac{1}{\pi(N)}\#\{p<N:\epsilon_{k_1}(P(p))=e_1,\ldots,\epsilon_{k_h}(P(p))=e_h\}=\hat{p}_{k_1,\ldots,k_h,e_1,\ldots,e_h} + \mathcal{O}\left((\log N)^{-\lambda}\right)$$

uniformly for all integers

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

and $e_1, e_2, \ldots, e_h \in \{0, 1, \ldots, a_1\}$. (The $\hat{p}_{k_1,\ldots,k_h,e_1,\ldots,e_h}$ are as in Lemma 3.3.)

We adapt the proof of Proposition 1.1 and include some elements of the proof of Gittenberger and Thuswaldner [21], who proved a similar theorem for digital expansions of the Gaussian integers. There the digits are also determined by tilings with fractal boundary.

Denote by $U_{e,\nu}$ the union of parallelepipeds of Lemma 3.9 containing the boundary of Ω_e. Let $\mathbf{1}_{\Omega_e \cup U_{e,\nu}}$ the characteristic function of $\Omega_e \cup U_{e,\nu}$ on the torus $\mathbb{T}^d$ and $\sum_{\mathbf{m}\in\mathbb{Z}^d} c_{\mathbf{m},e,\nu} e(\mathbf{m}\cdot\mathbf{x})$ its Fourier expansion. In order to calculate $c_{\mathbf{m},e,\nu}$ we split up $\Omega_e \cup U_{e,\nu}$ into parallelepipeds with edges parallel to $\mathbf{a}_1, \dots, \mathbf{a}_d$. Then we clearly have $c_{\mathbf{0},e,\nu} = \lambda_d(\Omega_e \cup U_{e,\nu})$ and, by Lemma 1 of Drmota [10], the $\mathbf{m}$-th Fourier coefficient of such a parallelepiped is

$$\sum_{\mathbf{x}\in V} \frac{\left|\det(\mathbf{x}-\mathbf{y})_{\mathbf{y}\in\Gamma(\mathbf{x})}\right|}{\prod_{\mathbf{y}\in\Gamma(\mathbf{x})}(-2\pi i)\mathbf{m}\cdot(\mathbf{x}-\mathbf{y})} e(-\mathbf{m}\cdot\mathbf{x}) = \sum_{\mathbf{x}\in V} \frac{\left|\det(\pm\mathbf{a}_j)_{1\le j\le d}\right|}{\prod_{j=1}^{d}(-2\pi i)\mathbf{m}\cdot(\pm\mathbf{a}_j)} e(-\mathbf{m}\cdot\mathbf{x}),$$

where V denotes the set of vertices of the parallelepiped and $\Gamma(\mathbf{x})$ the set of vertices adjacent to $\mathbf{x}$. As in Gittenberger and Thuswaldner [21], the contributions of the inner parallelepipeds cancel out and only the $\mathcal{O}(\gamma^\nu)$ corners of the boundary of $\Omega_e \cup U_{e,\nu}$ play a role. The contribution of a corner can be bounded by (cf. Drmota [10], Lemma 2)

$$\left| \frac{\left|\det(\pm\mathbf{a}_j)_{1\le j\le d}\right|}{\prod_{j=1}^{d}(-2\pi i)\mathbf{m}\cdot(\pm\mathbf{a}_j)} \right| \ll \prod_{i=1}^{d} \frac{1}{(1+|\mathbf{m}\cdot\mathbf{a}_i|)^2}$$

uniformly for all $\mathbf{m}$. Hence we define $\tilde{m}_i = \mathbf{m}\cdot\mathbf{a}_i$ and have

$$|c_{\mathbf{m},e,\nu}| \ll \gamma^\nu \prod_{i=1}^{d} \min\left(1, \frac{1}{|\tilde{m}_i|}\right)$$

As in Section 1.3, we consider the function

$$\psi_{e,\nu,\Delta}(\mathbf{x}) = \frac{1}{\Delta^d} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \cdots \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \mathbf{1}_{\Omega_e\cup U_{e,\nu}}(\mathbf{x}+z_1\mathbf{a}_1+\cdots+z_d\mathbf{a}_d)dz_1\dots dz_d.$$

By enlarging the parallelepipeds of $U_{e,\nu}$, we obtain sets $Q_{e,\nu}$ which are again unions of $\mathcal{O}(\gamma^\nu)$ parallelepipeds with $\lambda_d(Q_{e,\nu}) = \mathcal{O}\left(\left(\frac{\gamma}{\alpha}\right)^\nu\right)$ such that

$$\psi_{e,\nu,\Delta}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x}\in\Omega_e\setminus Q_{e,\nu} \\ 0 & \text{if } \mathbf{x}\notin\Omega_e\cup Q_{e,\nu}, \end{cases}$$

if we assume $\Delta < \alpha^{-\nu}$.

For the Fourier expansion $\psi_{e,\nu,\Delta}(\mathbf{x}) = \sum_{\mathbf{m}\in\mathbb{Z}^d} d_{\mathbf{m},e,\nu,\Delta} e(\mathbf{m}\cdot\mathbf{x})$, we get

$$|d_{\mathbf{m},e,\nu,\Delta}| \ll \gamma^\nu \prod_{i=1}^{d} \min\left(1, \frac{1}{|\tilde{m}_i|}, \frac{1}{\Delta\tilde{m}_i^2}\right).$$

We set

$$t(n) = \psi_{e_1,\nu,\Delta}(\mathbf{v}(n,k_1))\dots\psi_{e_h,\nu,\Delta}(\mathbf{v}(n,k_h)).$$

Then we have $t(n) = 1$ if $\mathbf{v}(n,k_i)\in\Omega_{e_i}\setminus Q_{e_i,\nu}$ for all i, $1\le i\le h$ and $t(n) = 0$ if $\mathbf{v}(n,k_i)\notin\Omega_{e_i}\cup Q_{e_i,\nu}$ for some i. Therefore we estimate the number of integers with $\mathbf{v}(n,k_i)\in Q_{e_i,\nu}$ by the following lemma.

Lemma 3.10. *Let*

$$E_{k,e,\nu} = \#\left\{n \leq N \,|\mathbf{v}(P(n),k) \in Q_{e,\nu}\right\}, \; F_{k,e,\nu} = \#\left\{p \leq N \,|\mathbf{v}(P(p),k) \in Q_{e,\nu}\right\}$$

and λ *an arbitrary positive constant. Then, uniformly in* k, $(\log N^r)^\eta \leq k \leq r \log_\alpha N^r - (\log N^r)^\eta$, *we have*

$$E_{k,e,\nu} \ll \left(\frac{\gamma}{\alpha}\right)^\nu N + N(\log N)^{-\lambda}, \;\; F_{k,e,\nu} \ll \left(\frac{\gamma}{\alpha}\right)^\nu \pi(N) + N(\log N)^{-\lambda}.$$

Proof. The proof of this lemma uses the isotropic discrepancy

$$J_N = \sup_{C \subseteq \mathbb{T}^d} \left| \frac{1}{N} \sum_{n=1}^{N} \chi_C(\{\mathbf{x}_n\}) - \lambda_d(C) \right|,$$

where the supremum is taken over all convex subsets C of $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ and $\mathbf{x}_1, \dots, \mathbf{x}_N \in \mathbb{R}^d$. It can be estimated by the normal discrepancy

$$D_N = \sup_{I \subseteq \mathbb{T}^d} \left| \frac{1}{N} \sum_{n=1}^{N} \chi_I(\{\mathbf{x}_n\}) - \lambda_d(I) \right|$$

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

$$D_N \leq J_N \leq (4d\sqrt{d} + 1) D_N^{\frac{1}{d}}$$

(see Drmota and Tichy [14], Theorem 1.12).

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

$$D_N \ll \frac{1}{M} + \sum_{\mathbf{h} \in \mathbb{Z}^d : 0 < \|\mathbf{h}\|_\infty < M} \prod_{i=1}^{d} \frac{1}{\max\{1, |h_i|\}} \left| \frac{1}{N} \sum_{n=1}^{N} e(\mathbf{h} \cdot \mathbf{x}_n) \right|,$$

where M is an arbitrary positive integer (cf. [14], Theorem 1.21).

We set $\mathbf{x}_n = \mathbf{v}(P(n), k)$ and $M = (\log N)^{d\lambda}$. Then we have, since $Q_{e,\nu}$ is the union of $\mathcal{O}(\gamma^\nu)$ convex subsets and the conditions of Lemmata 2.5 and 2.6 hold,

$$\begin{aligned} E_{k,e,\nu} &\ll \gamma^\nu J_N N + \lambda_d(Q_{e,\nu}) N \\ &\ll N\gamma^\nu \left((\log N)^{-\lambda} + \log(\log N)^{d\lambda} (\log N)^{-\tau_0/d} \right) + \left(\frac{\gamma}{\alpha}\right)^\nu N. \end{aligned}$$

Similarly we get with Lemma 2.6,

$$F_{k,e,\nu} \ll \pi(N)\gamma^\nu \left((\log N)^{-\lambda} + \log(\log N)^{d\lambda} (\log N)^{-\tau_0/d} \right) + \left(\frac{\gamma}{\alpha}\right)^\nu \pi(N).$$

We can choose $\tau_0 > d\lambda$ and the inequalities are proven. □

We define

$$\Sigma_1 = \#\{n < N : \epsilon_{k_1}(P(n)) = e_1, \ldots, \epsilon_{k_h}(P(n)) = e_h\}$$

and

$$\Sigma_2 = \#\{n < N : \epsilon_{k_1}(P(n)) = e_1, \ldots, \epsilon_{k_h}(P(n)) = e_h\}.$$

For $\nu \ll \log\log N$ and $(\log N^r)^\eta \leq k \leq r \log_\alpha N^r - (\log N^r)^\eta$, the error term $\mathcal{O}\left(\alpha^{-k}\right)$ of Proposition 3.1 is negligible compared to the size of each parallelepiped in $Q_{e,\nu}$ and we have

$$\left|\Sigma_1 - \sum_{n<N} t(P(n))\right| \leq E_{k_1,e_1,\nu} + \cdots + E_{k_h,e_h,\nu},$$

$$\left|\Sigma_2 - \sum_{n<N} t(P(p))\right| \leq F_{k_1,e_1,\nu} + \cdots + F_{k_h,e_h,\nu}.$$

As usual, we will consider only Σ_1 since Σ_2 can be treated similarly.

Let $\mathcal{M}$ be the set of vectors $\mathbf{M} = (\mathbf{m}_1, \ldots, \mathbf{m}_h)$ with integer vectors $\mathbf{m}_i = (m_1^{(i)}, \ldots, m_d^{(i)})$. Then we have

$$\sum_{n<N} t(P(n)) = \sum_{\mathbf{M}\in\mathcal{M}} T_{\mathbf{M},\nu} \sum_{n<N} e\Big(\big(\mathbf{m}_1 \cdot \mathbf{v}(1,k_1) + \cdots + \mathbf{m}_h \cdot \mathbf{v}(1,k_h)\big)P(n)\Big),$$

with

$$T_{\mathbf{M},\nu} = d_{\mathbf{m}_1,e_1,\nu,\Delta} \cdots d_{\mathbf{m}_h,e_h,\nu,\Delta}.$$

Because of

$$\sum_{i=1}^{h} \mathbf{m}_i \cdot \mathbf{v}(1,k_i) = \frac{\alpha-1}{\alpha^d-1}\left(\frac{m_1^{(1)}\alpha^{d-1} + \cdots + m_d^{(1)}}{\alpha^{k_1}} + \cdots + \frac{m_1^{(h)}\alpha^{d-1} + \cdots + m_d^{(h)}}{\alpha^{k_h}}\right)$$

we have to estimate

$$S = \frac{\alpha^{k_h-k_1+d-1}m_1^{(1)} + \cdots + m_d^{(1)}\alpha^{k_h-k_1} + \cdots + m_1^{(h)}\alpha^{d-1} + \cdots + m_d^{(h)}}{\alpha^{k_h}} \tag{3.23}$$

in order to estimate these exponential sums, where we may assume $k_1 < k_2 < \cdots < k_h$.

Since the $\mathbf{a}_i$ form a basis of $\mathbb{R}^d$, we have

$$\frac{1}{|\mathbf{a}_1 \cdot \mathbf{m}|} \cdots \frac{1}{|\mathbf{a}_d \cdot \mathbf{m}|} \ll \frac{1}{|m_1|} \cdots \frac{1}{|m_d|}$$

and therefore

$$\sum_{\mathbf{M}\in\mathcal{M}} |T_{\mathbf{M},\nu}| \ll \left(\sum_{m_1=-\infty}^{\infty}\cdots\sum_{m_d=-\infty}^{\infty}\gamma^{\nu}\prod_{i=0}^{d-1}\min\left(1,\frac{1}{|\tilde{m}_i|},\frac{1}{\tilde{m}_i^2\Delta}\right)\right)^h$$
$$\ll \left(\sum_{\tilde{m}_1=-\infty}^{\infty}\cdots\sum_{\tilde{m}_d=-\infty}^{\infty}\gamma^{\nu}\prod_{i=0}^{d-1}\min\left(1,\frac{1}{|\tilde{m}_i|},\frac{1}{\tilde{m}_i^2\Delta}\right)\right)^h$$
$$\ll (\log N(\log\frac{1}{\Delta}))^{dh}$$

If $|m_i^{(j)}| > (\log N)^{2\delta}$ for some i,j, then we have $|\tilde{m}_i^{(j)}| > \tilde{c}(\log N)^{2\delta}$ for some i,j and

$$\sum_{\exists i,j \text{ with } |m_i^{(j)}|>(\log N)^{2\delta}} |T_{\mathbf{M},\nu}|$$
$$\ll \gamma^{h\nu}\left(\sum_{m=[\tilde{c}(\log N)^{2\delta}]}^{\infty}\frac{1}{m^2\Delta}\right)\left(\sum_{m=1}^{[\tilde{c}(\log N)^{2\delta}]}\frac{1}{|m|}\right)^{dh-1} \ll \gamma^{h\nu}\frac{\left(\log(\log N)^{2\delta}\right)^{dh}}{(\log N)^{\delta}}$$

if we set $\Delta = (\log N)^{\delta}$. Hence we need estimates of S for all $\mathbf{M}$ with $|m_i^{(j)}| \le (\log N)^{2\delta}$ for all i,j.

We use the following lemma due to W.M. Schmidt:

Lemma 3.11 (W.M. Schmidt [34], p. 153). *Suppose $1,\beta_1,\beta_2,\ldots,\beta_v$ are linearly independent over $\mathbb{Q}$, and they generate an algebraic number field of degree d. Then*

$$|\beta_1 q_1 + \cdots + \beta_v q_v - p| > cq^{-d+1}$$

for arbitrary integers $q_1,\ldots,q_v,p$ having $q = \max(|q_1|,\ldots,|q_v|) > 0$ and some constant c.

Lemma 3.12. *Let $|m_i^{(j)}| \le (\log N)^{2\delta}$ for all i,j,*

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

and arbitrary constants $\delta > 0, \eta > 0$. Then the S defined by (3.23) satisfy

$$S = 0 \quad \text{or} \quad \frac{\alpha^{(\log N)^{\eta'}}}{N^r} \ll |S| \ll \frac{1}{\alpha^{(\log N)^{\eta'}}} \qquad (3.24)$$

for all $\eta' < \eta$.

Proof. Assume $S \neq 0$. Because of Lemma 3.7 we have

$$S = \frac{\hat{m}_1 \alpha^{d-1} + \cdots + \hat{m}_{d-1}\alpha + \hat{m}_d}{\alpha^{k_h}}$$

with integers $\hat{m}_i$ which satisfy

$$|\hat{m}_i| \ll (\log N)^{2\delta} \alpha^{k_h - k_1} \quad (1 \leq i \leq d).$$

Therefore

$$|S| \ll \frac{(\log N)^{2\delta} \alpha^{k_h - k_1}}{\alpha^{k_h}} \leq \frac{(\log N)^{2\delta}}{\alpha^{(\log N)^{\eta}}} \ll \frac{1}{\alpha^{(\log N)^{\eta'}}}.$$

To obtain the lower bound we start by setting $\varepsilon = \eta/h$. Then there exists an integer K, $0 \leq K \leq h-1$, such that for all j

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

So fix K with this property. We have to distinguish two cases.

If $k_{j+1} - k_j \leq (\log N)^{K\varepsilon}$ for all j, we apply Lemma 3.11 and get

$$|S| \gg \frac{1}{\max\limits_{i \in \{0,\ldots,d-1\}} |\hat{m}_i|^{d-1} \alpha^{k_h}} \gg \frac{1}{(\log N)^{2(d-1)\delta} \alpha^{k_h + (d-1)(h-1)(\log N)^{K\varepsilon}}}$$

$$\gg \frac{\alpha^{(\log N)^{\eta} - (d-1)(h-1)(\log N)^{\frac{h\eta}{h+1}}}}{N^r (\log N)^{2(d-1)\delta}} \gg \frac{\alpha^{(\log N)^{\eta'}}}{N^r}.$$

Otherwise we have a $j < h$ such that $k_{j+1} - k_j \geq (\log N)^{(K+1)\varepsilon}$ and $k_j - k_1 \leq (j-1)(\log N)^{K\varepsilon}$. Then we split up the sum into two terms

$$\begin{aligned} S = {} & \frac{\alpha^{k_j - k_1 + d - 1} m_1^{(1)} + \cdots + m_d^{(1)} \alpha^{k_j - k_1} + \cdots + m_1^{(j)} \alpha^{d-1} + \cdots + m_d^{(j)}}{\alpha^{k_j}} \\ & + \frac{\alpha^{k_h - k_{j+1} + d - 1} m_1^{(j+1)} + \cdots + m_d^{(j+1)} \alpha^{k_h - k_{j+1}} + \cdots + m_1^{(h)} \alpha^{d-1} + \cdots + m_d^{(h)}}{\alpha^{k_h}} \\ = {} & S_1 + S_2. \end{aligned}$$

If $S_1 = 0$, then $S = S_2$ and we are concerned with a problem containing less terms. By using induction on h (which is not made explicit here), we may assume that this case has already been treated. Otherwise we have

$$|S_1| \gg \frac{1}{(\log N)^{2(d-1)\delta} \alpha^{k_j + (d-1)(j-1)(\log N)^{K\varepsilon}}},$$

whereas

$$|S_2| \ll \frac{(\log N)^{2\delta} \alpha^{k_h - k_{j+1}}}{\alpha^{k_h}} \leq \frac{(\log N)^{2\delta}}{\alpha^{k_j + (\log N)^{(K+1)\varepsilon}}}.$$

Hence

$$|S| \gg \frac{\alpha^{(\log N)^{\eta} - (j-1)(d-1)(\log N)^{K\varepsilon}}}{N^r (\log N)^{2\delta}} \gg \frac{\alpha^{(\log N)^{\eta'}}}{N^r}$$

and the lemma is proved. ∎

Hence we have

$$\Sigma_1 = \sum_{\mathbf{M}\in\mathcal{M}:\sum \mathbf{m}_i\cdot\mathbf{v}(1,k_i)=0} T_{\mathbf{M},\nu} + \mathcal{O}\left(\gamma^{h\nu} N(\log N)^{-\tau_0} + \gamma^{h\nu} N(\log N)^{-\delta/2} + N\left(\frac{\gamma}{\alpha}\right)^{\nu}\right)$$

We set

$$T'_{\mathbf{M},\nu} = c_{\mathbf{m}_1,e_1,\nu}\cdots c_{\mathbf{m}_h,e_h,\nu}$$

and have to compare $T_{\mathbf{M},\nu}$ to $T'_{\mathbf{M},\nu}$. Here we have

$$T_{\mathbf{M},\nu} = T'_{\mathbf{M},\nu} + \mathcal{O}\left(\gamma^{\nu}\max_{i,j}\left|\tilde{m}_i^{(j)}\right|\Delta\right)$$

and

$$\sum_{\mathbf{M}\in\mathcal{M}:|\tilde{m}_i^{(j)}|<(\log N)^{\frac{\delta}{2dh}}\text{ for all } i,j} \left|T_{\mathbf{M},\nu} - T'_{\mathbf{M},\nu}\right| \ll \gamma^{\nu}(\log N)^{-\delta/3}.$$

For the other $\mathbf{M}$ we obtain by the same methods as in Section 2.1

$$\sum_{\mathbf{M}\in\mathcal{M}:\sum \mathbf{m}_i\cdot\mathbf{v}(1,k_i)=0,|\tilde{m}_i^{(j)}|\geq(\log N)^{\frac{\delta}{2dh}}\text{ for some } i,j} T'_{\mathbf{M}} \ll (\log N)^{-\frac{\delta}{2dh(dh-1)^2}}.$$

If we set

$$\tilde{p}_{k_1,\ldots,k_h,e_1,\ldots,e_h,\nu} = \sum_{\mathbf{M}\in\mathcal{M}:\sum \mathbf{m}_i\cdot\mathbf{v}(1,k_i)=0} T'_{\mathbf{M},\nu},$$

we get

$$\Sigma_1 = N\tilde{p}_{k_1,\ldots,k_h,e_1,\ldots,e_h,\nu} + \mathcal{O}\left(\gamma^{\nu}N(\log N)^{-\frac{\delta}{2dh(dh-1)^2}}\right) + \mathcal{O}\left(N\left(\frac{\gamma}{\alpha}\right)^{\nu}\right)$$

Remark 3.6. In case of one variable k, we have $\mathbf{m}\cdot\mathbf{v}(1,k_i) = 0$ only for $\mathbf{m} = \mathbf{0}$. Hence $\tilde{p}_{k,e,\nu} = c_{\mathbf{0},e,\nu} \to \lambda_d(\Omega_e) = p_e = \hat{p}_{k,e}$ as $\nu \to \infty$.

We set $\nu = [C\log\log N]$ for some constant C which satisfies $\left(\frac{\gamma}{\alpha}\right)^{\nu} \ll (\log N)^{-\lambda}$, choose δ such that $(\log N)^{-\frac{\delta}{2dh(dh-1)^2}} \ll \alpha^{-\nu}$ and get

$$\Sigma_1 = N\tilde{p}_{k_1,\ldots,k_h,e_1,\ldots,e_h,[C\log\log N]} + \mathcal{O}\left(N(\log N)^{-\lambda}\right).$$

For $P(n) = n$ and $(\log N)^{\eta} \leq k_1,\ldots,k_h \leq \log_\alpha N - (\log N)^{\eta}$, Lemma 3.3 implies

$$\Sigma_1 = N\hat{p}_{k_1,\ldots,k_h,e_1,\ldots,e_h} + \mathcal{O}\left(N(\log N)^{-\lambda}\right)$$

and therefore

$$\tilde{p}_{k_1,\ldots,k_h,e_1,\ldots,e_h,[C\log\log N]} = \hat{p}_{k_1,\ldots,k_h,e_1,\ldots,e_h} + \mathcal{O}\left((\log N)^{-\lambda}\right).$$

For $(\log N^r)^{\eta} \leq k_1,\ldots,k_h \leq \log_\alpha N^r - (\log N^r)^{\eta}$, we obtain this result by considering Σ_1 for $P(n) = n$ and N^r.

As already noted, we get the corresponding result for primes by the same arguments. Thus Proposition 3.3 and Theorem 3.2 are proved.

3.5 Joint distributions of G-additive and q-additive functions

Finally, we generalise Theorems 2.1 and 2.3 on G-ary expansions.

Theorem 3.3. *Let f_ℓ, $1 \le \ell \le L$, be either q_ℓ-additive functions as in Theorem 1.1 or G_ℓ-additive functions as in Theorem 3.2. Let $P_\ell(n)$ be polynomials of different degrees r_ℓ with integer coefficients and positive leading terms. Then, as $N \to \infty$,*

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \frac{f_\ell(P_\ell(n)) - M_\ell(N^{r_\ell})}{D_\ell(N^{r_\ell})} < x_\ell, \ell = 1, 2, \ldots, L\right\} \to \Phi(x_1)\ldots\Phi(x_L)$$

and

$$\frac{1}{\pi(N)}\#\left\{p < N \,\middle|\, \frac{f_\ell(P_\ell(p)) - M_\ell(N^{r_\ell})}{D_\ell(N^{r_\ell})} < x_\ell, \ell = 1, \ldots, L\right\} \to \Phi(x_1)\ldots\Phi(x_L)$$

The strategy of the proof of Theorem 3.3 is exactly the same as that of Theorem 2.1 and the changes which have to be made are obvious. Therefore they will not be presented.

Theorem 3.4. *Let f_1 be a G_1-additive function as in Theorem 3.2 with dominant root α_1 of degree d_1 and f_2 either a q-additive function as in Theorem 1.1 or a G_2-additive function as in Theorem 3.2 with dominant root α_2 of degree d_2 such that $[\mathbb{Q}(\alpha_1, \alpha_2) : \mathbb{Q}] = d_1 d_2$. Let $P_1(n), P_2(n)$ be polynomials with integer coefficients, degrees r_1, r_2 and positive leading term. Then, as $N \to \infty$,*

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \frac{f_\ell(P_\ell(n)) - M_\ell(N^{r_\ell})}{D_\ell(N^{r_\ell})} < x_\ell \ (\ell = 1, 2)\right\} \to \Phi(x_1)\Phi(x_2) \tag{3.25}$$

and

$$\frac{1}{\pi(N)}\#\left\{p < N \,\middle|\, \frac{f_\ell(P_\ell(p)) - M_\ell(N^{r_\ell})}{D_\ell(N^{r_\ell})} < x_\ell \ (\ell = 1, 2)\right\} \to \Phi(x_1)\Phi(x_2). \tag{3.26}$$

Remark 3.7. If $(d_1, d_2) = 1$, then $[\mathbb{Q}(\alpha_1, \alpha_2) : \mathbb{Q}] = d_1 d_2$ is always satisfied. If $d_1 = d_2 = 2$, this condition is equivalent to $\frac{D_1}{D_2} = \frac{\sqrt{{a_1^{(1)}}^2 + 4}}{\sqrt{{a_1^{(2)}}^2 + 4}}$ being irrational.

As usual it suffices to prove Propositions 3.4 and 3.5.

Proposition 3.4 (cf. Proposition 2.4). *Let G be a sequence as in Theorem 3.2 with dominant root α, q an integer ($q \ge 2$) and $P_1(n), P_2(n)$ integer polynomials with positive leading terms and degrees r_1, r_2. Let $\lambda > 0$ be an arbitrary constant and h_1, h_2 non-negative integers. Then for integers*

$$(\log N^{r_1})^\eta \le k_1^{(1)} < k_2^{(1)} < \cdots < k_{h_1}^{(1)} \le \log_\alpha N^{r_1} - (\log N^{r_1})^\eta$$

(with some $\eta > 0$) and

$$(\log N^{r_2})^\eta \le k_1^{(2)} < k_2^{(2)} < \cdots < k_{h_2}^{(2)} \le \log_q N^{r_2} - (\log N^{r_2})^\eta$$

we have, as $N \to \infty$,

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \epsilon_{G_1,k_j^{(1)}}(P_1(n)) = b_j^{(1)}, \epsilon_{q,k_j^{(2)}}(P_2(n)) = b_j^{(2)}, 1 \le j \le h_\ell\right\}$$
$$= \hat{p}_{k_1^{(1)},\ldots,k_{h_1}^{(1)},b_1^{(1)},\ldots,b_{h_1}^{(1)}}\frac{1}{q^{h_2}} + \mathcal{O}\left((\log N)^{-\lambda}\right)$$

and

$$\frac{1}{\pi(N)}\#\left\{p < N \,\middle|\, \epsilon_{G_1,k_j^{(1)}}(P_1(p)) = b_j^{(1)}, \epsilon_{q,k_j^{(2)}}(P_2(p)) = b_j^{(2)}, 1 \le j \le h_\ell\right\}$$
$$= \hat{p}_{k_1^{(1)},\ldots,k_{h_1}^{(1)},b_1^{(1)},\ldots,b_{h_1}^{(1)}}\frac{1}{q^{h_2}} + \mathcal{O}\left((\log N)^{-\lambda}\right)$$

uniformly for $b_j^{(1)} \in \{0,\ldots,a_1\}$, $b_j^{(2)} \in \{0,\ldots,q-1\}$ and $k_j^{(\ell)}$ in the given range, where the implicit constant of the error term may depend on q_ℓ, on the polynomials P_ℓ, on h_ℓ and on λ.

Proof. The proof is similar to that of Proposition 2.4. We have to estimate the exponential sums

$$\sum_{n<N} e\left(\sum_{i=1}^{h_1} \mathbf{m}_i^{(1)} \cdot \mathbf{v}^{(1)}(1,k_i^{(1)})P_1(n) + \mathbf{m}^{(2)} \cdot \mathbf{v}^{(2)}\right).$$

If the degrees of r_1, r_2 are different then we are in the same situation as in Proposition 2.1. So assume $r_1 = r_2 = r$.

Denote by $g_r^{(1)}, g_r^{(2)}$ the leading terms of the polynomials and set

$$S = S_1 + S_2 = \sum_{i=1}^{h_1} \mathbf{m}_i^{(1)} \cdot \mathbf{v}^{(1)}(1,k_i^{(1)})g_r^{(1)} + \mathbf{m}^{(2)} \cdot \mathbf{v}^{(2)} g_r^{(2)}$$
$$= \frac{m_1^{(1,1)}\alpha^{k_{h_1}^{(1)}-k_1^{(1)}+d-1} + \cdots + m_d^{(1,1)}\alpha^{k_{h_1}^{(1)}-k_1^{(1)}} + \cdots + m_1^{(1,h_1)}\alpha^{d-1} + \cdots + m_d^{(1,h_1)}}{\alpha^{k_{h_1}^{(1)}}} g_r^{(1)}$$
$$+ \frac{m_1^{(2)}q^{k_{h_2}^{(2)}-k_1^{(2)}} + \cdots + m_{h_2}^{(2)}}{q^{k_{h_2}^{(2)}}} g_r^{(2)} = \frac{\hat{m}_1^{(1)}\alpha^{d-1} + \cdots + \hat{m}_{d-1}^{(1)}\alpha + \hat{m}_d^{(1)}}{\tilde{G}_{k_{h_1}^{(1)},1}\alpha^{d-1} + \cdots + \tilde{G}_{k_{h_1}^{(1)},d-1}\alpha + \tilde{G}_{k_{h_1}^{(1)},d}} g_r^{(1)} + \frac{\hat{m}^{(2)}}{q^{k_{h_2}^{(2)}}} g_r^{(2)}$$

(cf. Lemma 3.7) with

$$\tilde{G}_{k,i} = a_i G'_{k-1} + \cdots + a_d G'_{k-1-d+i}. \tag{3.27}$$

Hence we have $S=0$ if and only if

$$\hat{m}_i^{(1)} g_r^{(1)} q^{k_{h_2}^{(2)}} = \hat{m}^{(2)} g_r^{(2)} \tilde{G}_{k_{h_1}^{(1)},i} \text{ for all } i \in \{1,2,\dots,d\}.$$

Now we show that the $\tilde{G}_{k,i}$, $1 \le i \le d$, have no common divisor. First assume $\gcd(G'_{k-1}, G'_{k-2}, \dots, G'_{k-d}) = g > 1$. Because of $a_d = 1$ we get $g|G'_{k-d-1}$ and inductively $g|G'_j$ for all $j < k$, but this is not possible because of the choice of the initial values ($G'_{d-1} = 1$). Hence we have $\gcd(G'_{k-1}, G'_{k-2}, \dots, G'_{k-d}) = 1$. Now assume $\gcd(\tilde{G}_{k,1}, \tilde{G}_{k,2}, \dots, \tilde{G}_{k,d}) = g > 1$ for some k. Then (3.27) with $i=d$ gives $g|G'_{k-1}$, with $i = d-1$ we get $g|G'_{k-2}$ and inductively $g|G'_{k-i}$ for all $i \in \{1,\dots,d\}$ which contradicts $\gcd(G'_{k-1}, G'_{k-2}, \dots, G'_{k-d}) = 1$. Thus we have

$$\gcd(\tilde{G}_{k,1}, \tilde{G}_{k,2}, \dots, \tilde{G}_{k,d}) = 1 \text{ for all } k \ge 1. \tag{3.28}$$

Therefore we have, for every prime divisor p of q, some i such that $p \not| G_{k_{h_1}^{(1)},i}$. Hence $p^{k_{h_2}^{(2)}} | \hat{m}^{(2)}$ and $q|\hat{m}^{(2)}$. This implies $q|m_{h_2}^{(2)}$ and either $m_{h_2}^{(2)} = 0$, i.e. we have a smaller problem, or $d_{m_{h_2}^{(2)}, b_{h_2}^{(2)}, q, \Delta} = 0$. Thus we may assume $S \neq 0$ if $S_1 \neq 0$.

Now we can proceed as in the proof of Proposition 2.4. It suffices to consider those $\mathbf{m}_i^{(1)}$ and $\mathbf{m}^{(2)}$ with $0 < |m_j^{(1,i)}| < (\log N)^{2\delta}$, $0 < |m_j^{(2)}| < (\log N)^{2\delta}$ for all i,j and $S_1 \neq 0$. Clearly we have

$$S \ll \frac{(\log N)^{2\delta}}{\min(\alpha,q)^{(\log N)^\eta}}.$$

For the lower bound set $\varepsilon = \eta/(h_1+h_2-1)$. Then there exists an integer K with $0 \le K \le h_1+h_2-2$ 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. Set

$$\overline{m}_1 = \hat{m}_1^{(1)} \alpha^{d-1} + \dots + \hat{m}_{d-1}^{(1)} \alpha + \hat{m}_d^{(1)}, \quad \overline{m}_2 = g_r^{(2)} \hat{m}^{(2)}.$$

Then we have $\log |\overline{m}_\ell| \ll (\log N)^{K\varepsilon}$ because of $\hat{m}_i^{(1)} \ll (\log N)^{2\delta} \alpha^{k_{h_1}^{(1)} - k_1^{(1)}}$, we can apply Corollary 2.6 to

$$S = \frac{\overline{m}_1}{\alpha^{k_{h_1}^{(1)}}} + \frac{\overline{m}_2}{q^{k_{h_2}^{(2)}+1}}$$

and obtain

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

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

Otherwise we have some s_ℓ, $\ell = 1, 2$, 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 set

$$\overline{m}_1 = g_r^{(1)} \sum_{j=1}^{s_1} \left(m_1^{(1,j)} \alpha^{k_{s_1}^{(1)} - k_j^{(1)} + d - 1} + \cdots + m_d^{(1,j)} \alpha^{k_{s_1}^{(1)} - k_j^{(1)}} \right),$$

$$\overline{m}_2 = g_r^{(2)} \sum_{j=1}^{s_2} m_j^{(2)} q^{k_{s_2}^{(2)} - k_j^{(2)}}$$

Then we have again $\log|\overline{m}_\ell| \ll (\log N)^{K\varepsilon}$. Furthermore, we can estimate the sums

$$\sum_{j=s_1+1}^{h_1} \frac{m_1^{(1,j)}\alpha^{d-1} + \cdots + m_d^{(1,j)}}{\alpha^{k_j^{(2)}}} \ll (\log N)^{2\delta} q^{-k_{s_1}^{(1)} - (\log N)^{(K+1)\varepsilon}},$$

$$\sum_{j=s_2+1}^{h_2} \frac{m_j^{(2)}}{q^{k_j^{(2)}+1}} \ll (\log N)^{2\delta} q^{-k_{s_2}^{(2)} - (\log N)^{(K+1)\varepsilon}}.$$

Thus we get

$$\begin{aligned} S &\geq \left| \frac{\overline{m}_1}{\alpha^{k_{s_1}^{(1)}}} + \frac{\overline{m}_2}{q^{k_{s_2}^{(2)}+1}} \right| - \mathcal{O}\left((\log N)^{2\delta} \left(\alpha^{-k_{s_1}^{(1)} - (\log N)^{(K+1)\varepsilon}} + q^{-k_{s_1}^{(1)} - (\log N)^{(K+1)\varepsilon}} \right) \right) \\ &\geq \max\left(\alpha^{-k_{s_1}^{(1)}}, q^{-k_{s_2}^{(2)}-1}\right) \left(e^{-c\log\log N\,(\log N)^{K\varepsilon}} - \mathcal{O}\left((\log N)^{2\delta} e^{-\log(\min(\alpha,q))(\log N)^{(K+1)\varepsilon}} \right) \right) \\ &\geq \frac{(\log N)^\tau}{N^r} \end{aligned}$$

and the conditions of Lemmata 2.5 and 2.6 are satisfied.

Therefore the limits of the joint probabilities are just the products of the simple probabilities. □

Proposition 3.5. *Let G_1, G_2 be sequences as in Theorem 3.2 with dominant roots α_1, α_2 and $P_1(n), P_2(n)$ integer polynomials with positive leading terms and degrees r_1, r_2. Let $\lambda > 0$ be an arbitrary constant and h_1, h_2 non-negative integers. Then for integers*

$$(\log N^{r_\ell})^\eta \leq k_1^{(\ell)} < k_2^{(\ell)} < \cdots < k_{h_\ell}^{(\ell)} \leq \log_{\alpha_\ell} N^{r_\ell} - (\log N^{r_\ell})^\eta \quad (\ell = 1, 2)$$

(with some $\eta > 0$) we have, as $N \to \infty$,

$$\frac{1}{N}\#\left\{n < N \,\middle|\, \epsilon_{G_1,k_j^{(1)}}(P_1(n)) = b_j^{(1)}, \epsilon_{G_2,k_j^{(2)}}(P_2(n)) = b_j^{(2)}, 1 \le j \le h_\ell\right\}$$
$$= \hat{p}_{k_1^{(1)},\dots,k_{h_1}^{(1)},b_1^{(1)},\dots,b_{h_1}^{(1)}}\hat{p}_{k_1^{(2)},\dots,k_{h_2}^{(2)},b_1^{(2)},\dots,b_{h_2}^{(2)}} + \mathcal{O}\left((\log N)^{-\lambda}\right)$$

and

$$\frac{1}{\pi(N)}\#\left\{p < N \,\middle|\, \epsilon_{G_1,k_j^{(1)}}(P_1(p)) = b_j^{(1)}, \epsilon_{G_2,k_j^{(2)}}(P_2(p)) = b_j^{(2)}, 1 \le j \le h_\ell\right\}$$
$$= \hat{p}_{k_1^{(1)},\dots,k_{h_1}^{(1)},b_1^{(1)},\dots,b_{h_1}^{(1)}}\hat{p}_{k_1^{(2)},\dots,k_{h_2}^{(2)},b_1^{(2)},\dots,b_{h_2}^{(2)}} + \mathcal{O}\left((\log N)^{-\lambda}\right)$$

uniformly for $b_j^{(\ell)} \in \{0,\dots,q_\ell - 1\}$ and $k_j^{(\ell)}$ in the given range, where the implicit constant of the error term may depend on q_ℓ, on the polynomials P_ℓ, on h_ℓ and on λ.

Proof. The proof is almost the same as that of Proposition 3.5. It remains to prove that $S = 0$ only if $S_1 = S_2 = 0$, where

$$S = S_1 + S_2 = \sum_{i=1}^{h_1} \mathbf{m}_i^{(1)} \cdot \mathbf{v}^{(1)}(1, k_i^{(1)}) g_r^{(1)} + \sum_{i=1}^{h_2} \mathbf{m}_i^{(2)} \cdot \mathbf{v}^{(2)}(1, k_i^{(2)}) g_r^{(2)}$$
$$= \frac{m_1^{(1,1)}\alpha_1^{k_{h_1}^{(1)}-k_1^{(1)}+d_1-1} + \cdots + m_{d_1}^{(1,h_1)}}{\alpha_1^{k_{h_1}^{(1)}}} g_r^{(1)} + \frac{m_1^{(2,1)}\alpha_2^{k_{h_2}^{(2)}-k_1^{(2)}+d_2-1} + \cdots + m_{d_2}^{(2,h_1)}}{\alpha_2^{k_{h_2}^{(2)}}} g_r^{(2)}$$
$$= \frac{\hat{m}_1^{(1)}\alpha_1^{d_1-1} + \cdots + \hat{m}_{d_1}^{(1)}}{\tilde{G}_{k_{h_1}^{(1)},1}^{(1)}\alpha_1^{d_1-1} + \cdots + \tilde{G}_{k_{h_1}^{(1)},d_1}^{(1)}} g_r^{(1)} + \frac{\hat{m}_1^{(2)}\alpha_2^{d_2-1} + \cdots + \hat{m}_{d_2}^{(2)}}{\tilde{G}_{k_{h_2}^{(2)},1}^{(2)}\alpha_2^{d_2-1} + \cdots + \tilde{G}_{k_{h_2}^{(2)},d_2}^{(2)}} g_r^{(2)}$$

Because of $[\mathbb{Q}(\alpha_1, \alpha_2) : \mathbb{Q}] = d_1 d_2$, the $\alpha_1^i \alpha_2^j$, $0 \le i < d_1$, $0 \le j < d_2$, are linearly independent over $\mathbb{Q}$ and we get the equation system

$$\hat{m}_i^{(1)} g_r^{(1)} \tilde{G}_{k_{h_2}^{(2)},j}^{(2)} = \hat{m}_j^{(2)} g_r^{(2)} \tilde{G}_{k_{h_1}^{(1)},i}^{(1)} \quad (1 \le i \le d_1, 1 \le j \le d_2).$$

Hence we have

$$\hat{m}_i^{(1)} = -\hat{m}_j^{(2)} \frac{g_r^{(2)}}{g_r^{(1)}} \frac{\tilde{G}_{k_{h_1}^{(1)},i}^{(1)}}{\tilde{G}_{k_{h_2}^{(2)},j}^{(2)}} = \hat{m}_{i'}^{(1)} \frac{\tilde{G}_{k_{h_1}^{(1)},i}^{(1)}}{\tilde{G}_{k_{h_1}^{(1)},i'}^{(1)}} \quad (1 \le i, i' \le d_1). \tag{3.29}$$

Therefore the system of $d_1 d_2$ equations can be reduced to $d_1 + d_2 - 1$ equations and we have non-trivial solutions, but they must satisfy

$$\hat{m}_i^{(\ell)} \equiv 0 \left(\tilde{G}_{k_{h_\ell}^{(\ell)},i}^{(\ell)}\right) \tag{3.30}$$

for all i, ℓ. This implies $m_i^{(\ell)} = 0$ for all i, ℓ because of

$$\tilde{G}^{(\ell)}_{k^{(\ell)}_{h_\ell},i} \geq G'_{k^{(\ell)}_{h_\ell}-d} \sim {c'_1}^{(\ell)} \alpha_\ell^{k^{(\ell)}_{h_\ell}-d} \text{ and } \left|\hat{m}_i^{(\ell)}\right| \ll (\log N)^{2\delta} \alpha_\ell^{k^{(\ell)}_{h_\ell}-k^{(\ell)}_1}$$

and thus $S_1 = S_2 = 0$. To show (3.30), let $p_1^{e_1} \dots p_t^{e_t}$ be the prime factorisation of $\tilde{G}^{(\ell)}_{k^{(\ell)}_{h_\ell},i}$. For each p_j we have $p_j \nmid \tilde{G}^{(\ell)}_{k^{(\ell)}_{h_\ell},i'}$ for some i' because of (3.28). Hence (3.29) implies $p_j^{e_j} | \hat{m}_1^{(1)}$ for all j and (3.30) is proved.

This concludes the proof of Theorem 3.4. $\square$

Bibliography

[1] S. Akiyama, Cubic pisot units with finite beta expansions, in: *Algebraic Number Theory and Diophantine Analysis (F. Halter-Koch and R. F. Tichy eds)*, de Gruyter, Berlin, 2000, 11–26.

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

[3] P. Billingsley, *Probability and Measure* (3rd ed.), Wiley, New York, 1995.

[4] A. Brauer, On algebraic equations with all but one root in the interior of the unit circle, *Math. Nach.* **4** (1951) 250–257.

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

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

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

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

[9] 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.)

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

[11] M. Drmota, The joint distribution of q-additive functions, *Acta Arith.* **100** (1) (2001), 17-39.

[12] M. Drmota and J. Gajdosik, The Distribution of the Sum-of-Digits Function, *J. Theor. Nombres Bordx.* **10** (1998), 17–32.

[13] M. Drmota and W. Steiner, The Zeckendorf expansion of polynomial sequences, submitted,
http://www.geometrie.tuwien.ac.at/steiner/ze.ps.gz

[14] M. Drmota and R. F. Tichy, *Sequences, Discrepancies and Applications*, volume 1651 of *Lecture Notes in Mathematics*, Springer, Berlin, 1998.

[15] J. M. Dumont and A. Thomas, Systèmes de numération et fonctions fractales relatifs aux substitutions, *J. Theoret. Comput. Sci.* **65** (1989), 153–169.

[16] J. M. Dumont and A. Thomas, Digital sum moments and substitutions, *Acta Arith.* **64** (3) (1993), 205–225.

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

[18] C. Frougny and B. Solomyak, Finite β-expansions, *Ergodic Theory and Dynamical Systems* **12** (4) (1992), 713–723.

[19] N. Giesbrecht, Bounds for sums of random variables over a Markov chain, *Stochastic Process. Appl.* **53** (2) (1994), 269–283.

[20] W. J. Gilbert, Fractal dimension of sets derived from complex bases, *Ann. Sci. Math. Quebec* **11** (1) (1987), 65–77.

[21] B. Gittenberger and J. M. Thuswaldner, Asymptotic normality of b-additive functions on polynomial sequences in the gaussian number field, J. Number Theory **84** (2000), 317–341.

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

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

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

[25] I. Kátai, Distribution of q-additive function, in: *Probability theory and applications, Essays to the Mem. of J. Mogyorodi, Math. Appl.* **80**, Kluwer, Dortrecht, 309–318 (1992).

[26] 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.)

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

[28] A. Messaoudi, Autour du fractal de Rauzy, *Thèse de doctorat*, Université de la Méditerranée, 1996.

[29] A. Messaoudi, Propriétés arithmétiques et dynmamiques du fractal de Rauzy, *J. Théor. Nombres Bordx.* **10** (1998), 135–162.

[30] W. Parry, On the β-expansion of real numbers, *Acta Math. Acad. Sci. Hung.* **11** (1960), 401-416.

[31] B. Praggastis, Numeration Systems and Markov Partitions from Self Similar Tilings, *Trans. Am. Math. Soc.* **351** (8) (1999), 3315–3349.

[32] G. Rauzy, Nombres algébriques et substitutions, *Bull. Soc. Math. Fr.* **110** (2) (1982), 147–178.

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

[34] W. M. Schmidt, Diophantine Approximation, *Lecture Notes in Math.* 785, Springer, Berlin, 1980.

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

[36] A. Siegel, Représentations géometrique, combinatoire et arithmétique des systèmes susbstitutifs de type Pisot, *Thèse de doctorat*, Université de la Méditerranée, 2000.

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

Acknowledgements

First of all, I want to express my sincere thanks to Prof. Michael Drmota for giving me the possibility to work within one of his research projects, for attracting my attention to the problems treated within this thesis and for providing continously support. I would also like to thank my parents without whom my studies would have been much harder or impossible. Finally, I am grateful to the Austrian Science Foundation FWF (grant S8302-MAT) for the financial support.

Lebenslauf

Ich wurde am 30. März 1976 als Sohn von Dr. Ingrid Steiner, geb. Hofreiter, und Dipl.-Ing. Peter Steiner in Wien geboren. Von 1982 bis 1986 besuchte ich die Volksschule der Piaristen in Wien VIII, anschließend das Bundesgymnasium Wien XIII, Fichtnergasse, wo ich am 21. Juni 1994 mit Auszeichnung maturierte.

Im September 1994 immatrikulierte ich an der Technischen Universität Wien und inskribierte für das Studium der Technischen Mathematik, Studienzweig Mathematik in den Naturwissenschaften. Von 1996 bis 1998 absolvierte ich den "Tronc Commun" der Ecole Centrale Paris im Rahmen des Doppeldiplomprogramms TIME. Am 9. Mai 2000 schloss ich das Diplomstudium der Technischen Mathematik mit Auszeichnung ab, wodurch ich auch das Ingenieursdiplom der Ecole Centrale Paris erhielt.

Seit Mai 2000 arbeite ich als Forschungsassistent in dem FWF-Projekt S8302-MAT, *Statistical Properties of Number Systems*, unter der Leitung von Prof. Michael Drmota.

Wien, im Februar 2002 — Wolfgang Steiner