

q -ADDITIVE FUNCTIONS ON POLYNOMIAL SEQUENCES

MANFRED G. MADRITSCH

ABSTRACT. The present paper deals with functions acting only on the digits of an q -ary expansion. In particular let n be a positive integer, then we call

$$n = \sum_{r=0}^{\ell} d_r(n)q^r \quad \text{with} \quad d_r(n) \in \{0, \dots, q-1\}$$

its q -ary expansion. We call a function f strictly q -additive if it acts only on the digits of a representation, *i.e.*,

$$f(n) = \sum_{r=0}^{\ell} f(d_r(n)).$$

The goal is to prove that if p is a polynomial having at least one coefficient with bounded continued fraction expansion, then

$$\sum_{n \leq N} f(\lfloor p(n) \rfloor) = \mu_f N \log_q(p(N)) + NF_{q,k}(\log_q(p(N))) + \mathcal{O}\left(\frac{N}{\log N}\right).$$

This result is motivated by the asymptotic distribution result of Bassily and Kátai and a similar result of M. Peter.

1. INTRODUCTION

Let $q \geq 2$ be an integer. Then we can represent every positive integer n in a unique way as follows

$$(1.1) \quad n = \sum_{r=0}^{\ell} d_r(n)q^r \quad \text{with} \quad d_r(n) \in \{0, \dots, q-1\}.$$

We call this the q -ary representation of n with q the base and $\{0, \dots, q-1\}$ the set of digits.

If a function f acts only on the digits of a representation, *i.e.*,

$$f(n) = \sum_{r=0}^{\ell} f(d_r(n)q^r),$$

where n is as in (1.1), then we call it q -additive. If this action of f is independent of the position of the digit, *i.e.*,

$$f(n) = \sum_{r=0}^{\ell} f(d_r(n)),$$

then we call f strictly q -additive. In the following we will concentrate on these type of functions.

A simple example of a strictly q -additive function is the sum-of-digits function s_q which is defined by

$$s_q(n) = \sum_{r=0}^{\ell} d_r(n),$$

Date: July 25, 2011.

2000 Mathematics Subject Classification. 11N37 (11A63).

Key words and phrases. q -additive functions, polynomial sequences, summatory function.

Supported by the Austrian Research Foundation (FWF), Project S9603, that is part of the Austrian Research Network “Analytic Combinatorics and Probabilistic Number Theory”.

where n is as in (1.1). Strictly q -additive functions have been investigated from several points of view. In the present paper we want to concentrate on distributional properties and, in particular, on the summatory function of f on polynomial values.

Before we present the result we want to give an overview on what is known in this area. For the case of the sum-of-digits function together with the identity function Delange [5] was able to show

$$\sum_{n \leq N} s_q(n) = \frac{q-1}{2} N \log_q N + NF(\log_q N),$$

where \log_q is the logarithm to base q and F a 1-periodic, continuous and nowhere differentiable function.

The moments of the sum-of-digits function were considered by Kirschenhofer [11] and by Grabner *et al.* [10]. For different methods originating from analytic number theory such as Mellin's formula elegant proofs have been shown by Flajolet *et al.* [6] and Grabner and Hwang [9]. Generalizations to number systems in number fields were done by Thuswaldner [22] and Gittenberger and Thuswaldner [8].

Apart from the sequence of the positive integers there are the primes or the integer values of polynomials of interest. For the case of primes the summatory function of the sum-of-digits function was considered by Shiokawa [21]. Mauduit and Rivat considered the distribution in residue classes and the uniform distribution modulo 1 of $(\alpha s_q(\lfloor n^c \rfloor))_{n \geq 1}$ (with α an irrational and $\lfloor \cdot \rfloor$ the floor function) with $1 \leq c \leq \frac{4}{3}$ in [13] and with $c = 2$ in [14]. The distribution for the case $c = 1$ goes back to Gelfond [7].

Now we draw our attention towards general q -additive functions and observe the considerations of Bassily and Kátai [1] for the asymptotic distribution of the values of $(f(p(n)))_{n \geq 1}$ with p a polynomial of degree k having integer coefficients. In particular they were able to show that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \# \left\{ n < N : \frac{f(p(n)) - k\mu_f \log_q N}{\sqrt{k\sigma_f^2 \log_q N}} < x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt,$$

where

$$(1.2) \quad \mu_f = \frac{1}{q} \sum_{a=0}^{q-1} f(a) \quad \text{and} \quad \sigma_f^2 = \frac{1}{q} \sum_{a=0}^{q-1} f(a)^2 - \mu_f^2.$$

This provides us with the motivation to consider the summatory function of arbitrary q -additive functions for polynomial sequences. To this end we see that not much is known. However, there is a partial result by Peter [20] for k th powers and where α , the leading coefficient, is an irrational of finite type. In particular he could show that there are $c \in \mathbb{R}$ and $\varepsilon > 0$ such that

$$\sum_{n \leq N} s_q(\lfloor \alpha n^k \rfloor) = \frac{q-1}{2} N \log_q(\alpha N^k) + cN + NF(\log_q(\alpha N^k)) + \mathcal{O}(N^{1-\varepsilon})$$

where F is a 1-periodic function and $\alpha = 1$ or $\alpha > 0$ an irrational of finite type.

Similar ideas are needed for the construction of pseudo-random sequences. These sequences were introduced by Mauduit and Sárközy in a series of papers [2–4, 15–18]. To this end we define $\chi(x)$ by

$$\chi(x) = \begin{cases} +1 & \text{for } 0 \leq \{x\} < 1/2, \\ -1 & \text{for } 1/2 \leq \{x\} < 1, \end{cases}$$

where $\{ \cdot \}$ denotes the fractional part. In particular, in Part V [17] and VI [18] they considered sequences $E_N = \{e_1, \dots, e_N\}$ with

$$e_n = \chi(\alpha n^k),$$

where α is an irrational number, whose partial quotients in the continued fraction expansion are bounded, *i.e.*, there exists a $K \in \mathbb{N}$ such that

$$(1.3) \quad \alpha = [a_0; a_1, a_2, \dots], \quad a_i \leq K, \quad \text{for } i \geq 1.$$

Our aim is now to perform several tasks. First we want to get rid of the requirement on α and to consider the sequence $(\lfloor p(n) \rfloor)_{n \geq 0}$ where p is a polynomial of degree k with real coefficients. Furthermore we are interested in arbitrary q -additive functions instead of the sum-of-digits function. Finally we want to use the bound in the continued fraction expansion of the coefficients.

In order to state the result in a more convenient way we need some definitions. For $x \in \mathbb{R}$ we denote by $\lfloor x \rfloor$ the floor function and by $\{x\} := x - \lfloor x \rfloor$ the fractional part of x , respectively. Let ψ be centralized fraction function defined by

$$(1.4) \quad \psi(x) = x - \lfloor x \rfloor - \frac{1}{2} = \{x\} - \frac{1}{2}.$$

Moreover, we define

$$(1.5) \quad \begin{aligned} J_{q,k}(x) &:= \int_0^x \sum_{a=0}^{q-1} f(a) \left(\psi \left(t - \frac{a+1}{q} \right) - \psi \left(t - \frac{a}{q} \right) \right) t^{\frac{1}{k}-1} dt, \quad x \geq 0, \quad q, k \in \mathbb{N}, \text{ and} \\ F_{q,k}(t) &:= \mu_f(1 - \{t\}) + \frac{1}{k} q^{(1-\{t\})/k} \sum_{n \geq 0} q^{-\frac{n}{k}} J_{q,k}(q^{n-1+\{t\}}), \quad t \in \mathbb{R}, \end{aligned}$$

where μ_f is the mean of the values of f defined in (1.2). Finally let p be the polynomial under consideration with real coefficients of degree k , *i.e.*,

$$(1.6) \quad p(x) = \alpha_k x^k + \dots + \alpha_1 x + \alpha_0.$$

Now we are able to state our result.

Theorem 1.1. *Let $q \in \mathbb{N} \setminus \{1\}$ and f be a strictly q -additive function and p be a polynomial of degree $k \geq 1$ as in (1.6). Suppose that there exists a $1 \leq i \leq k$ such that the coefficient α_i is an irrational number whose partial quotients satisfy (1.3) with some $K \in \mathbb{N}$. Then*

$$(1.7) \quad \sum_{n \leq N} f(\lfloor p(n) \rfloor) = \mu_f N \log_q(p(N)) + N F_{q,k}(\log_q(p(N))) + \mathcal{O}\left(\frac{N}{\log N}\right).$$

Remark 1.2. We can show a similar result if f is q -additive but not necessarily strictly q -additive. In this more general setting one has to keep track of the position of a digit in the q -ary expansion throughout the whole proof. This leads to a more delicate periodic function $F_{q,k}$ which then depends on the integer value of $\log_q(p(N))$ (*i.e.*, the position) also.

We will show this theorem in two steps. In Section 2 we will provide all the tools which we will apply in Section 3 in order to prove the theorem. The idea of proof is to extract each digit of an expansion and divide them up into three parts according to their relative position in the expansion. In particular we will distinguish between most significant digits, “middle” digits, and least significant digits. Since the “middle” ones are the easiest to treat we will start with them. This treatment provides us with the bounds for least, “middle” and most. The problem which led to the restriction in Peter’s result lies in the low significant digits. Therefore we will use a different approach in order to face them. Furthermore we will extract the periodic function $F_{q,k}$ from the most significant digits by improving and shortening the methods used by Peter.

2. PREREQUISITES

In this section we want to present the lemmas which will help us on the way to prove the theorem. As described above we will rewrite the sum and divide it according to the position of the digit in the q -ary expansion into three parts.

To extract a single digit we will need the indicator function which we will write as a combination of ψ functions. Let $\mathcal{I}_{r,a}(x)$ be the indicator function which indicates if the digit at position r is a , *i.e.*,

$$\mathcal{I}_{r,a}(x) = \begin{cases} 1 & \text{if } d_r(x) = a, \\ 0 & \text{else.} \end{cases}$$

We will extract the main term and use two different representations for $\mathcal{I}_{r,a}(x)$. Thus

$$(2.1) \quad \begin{aligned} \mathcal{I}_{r,a}(x) &= \frac{1}{q} + \frac{1}{q} (q\mathcal{I}_{r,a}(x) - 1) \\ &= \frac{1}{q} + \psi\left(\frac{x}{q^{r+1}} - \frac{a+1}{q}\right) - \psi\left(\frac{x}{q^{r+1}} - \frac{a}{q}\right) \end{aligned}$$

Then we use the following lemma in order to calculate its Fourier transformation and thus get some exponential sums which are easier to treat.

Lemma 2.1 ([12, Theorem 1.8]). *Let $p(t)$ be a real function in $[a, b]$. Then for $\delta > 0$ the estimation*

$$\sum_{a < n \leq b} \psi(p(n)) \ll \frac{b-a}{\delta} + \sum_{\nu=1}^{\infty} \min\left(\frac{\delta}{\nu^2}, \frac{1}{\nu}\right) \left| \sum_{a < n \leq b} e(\nu p(n)) \right|$$

holds.

In order to properly estimate the exponential sums occurring in Lemma 2.1 we will subdivide the range of digits into three parts according to their position in the expansion. The range of the “middle” digits will be chosen such that the following lemma always applies.

Lemma 2.2 ([19, Lemma 2]). *Let $p(t)$ be a polynomial of the form*

$$p(t) = \alpha_k t^k + \alpha_{k-1} t^{k-1} + \dots + \alpha_1 t.$$

Let $G > 0$ be any constant and $N \geq 2$. Let s be an integer with $1 \leq s \leq k$, let H_i, K_i ($i = s+1, \dots, k$) be any positive constants, and let H_s^*, K_s^* be constants such that

$$\begin{aligned} H_s^* &\geq 2^{s+1} (G + \max_{s < i \leq k} H_i + 1) + k_s \sum_{s+1}^k K_i, \\ K_s^* &\geq 2^{s+1} (G + \max_{s < i \leq k} H_i + 1) + 2k_s \sum_{s+1}^k K_i. \end{aligned}$$

Suppose that there are rational numbers A_i/B_i ($s < i \leq k$) such that

$$1 \leq B_i \leq (\log N)^{K_i} \quad \text{and} \quad \left| \alpha_i - \frac{A_i}{B_i} \right| \leq \frac{(\log N)^{H_i}}{B_i N^i} \quad (s < i \leq k)$$

and that there is no rational number A_s/B_s with $(A_s, B_s) = 1$ such that

$$1 \leq B_s \leq (\log N)^{K_s^*} \quad \text{and} \quad \left| \alpha_s - \frac{A_s}{B_s} \right| \leq \frac{(\log N)^{H_s^*}}{B_s N^s}.$$

Then, for any real P and Q with $|P| \ll Q \leq N$,

$$\left| \sum_{P < n \leq P+Q} e(p(n)) \right| \ll N (\log N)^{-G}.$$

Now we are done with the “middle” digits and turn our attention to the least significant digits. Their estimation will be established through the assumption that there is a coefficient which has bounded continued fraction expansion.

Lemma 2.3 ([17, Lemma 3]). *Suppose that α is an irrational number whose partial quotients satisfy (1.3) with some $K \in \mathbb{N}$. Then for all $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ we have*

$$\left| \alpha - \frac{p}{q} \right| > \frac{1}{(K+2)q^2}.$$

The two lemmas above consider the case of the “middle” and the least significant digits, respectively. For the most significant digits, however, we need a different approach. We will treat those digits by first transforming the sum into an integral with the following lemma.

Lemma 2.4 ([12, Theorem 1.5]). *Let g in $[a, b]$ ($0 \leq a < b$) be a non-negative, strictly decreasing function with a continuous derivative in (a, b) . If $g^{-1}(t)$ denotes the inverse function of $g(t)$, then*

$$\begin{aligned} \sum_{a < n \leq b} \psi(g(n)) - \int_a^b \psi(t)g'(t)dt - \psi(a)\psi(g(a)) \\ = \sum_{g(a) < m \leq g(b)} \psi(g^{-1}(m)) - \int_{g(a)}^{g(b)} \frac{\psi(t)}{g'(g^{-1}(t))}dt - \psi(b)\psi(g(b)). \end{aligned}$$

The interesting integral is the first one. For the second one we need the following to deal with.

Lemma 2.5 ([12, Theorem 2.3]). *Let $p(t)$ be a real function in $[a, b]$, twice continuously differentiable. Let $p''(t)$ be monotonic and be either positive or negative throughout. Then*

$$\sum_{a < n \leq b} \psi(p(n)) \ll \int_a^b |p''(t)|^{\frac{1}{3}} dt + \frac{1}{\sqrt{|p''(a)|}} + \frac{1}{\sqrt{|p''(b)|}}.$$

Finally we will extract the periodic function $F_{q,k}$ (defined in (1.5) by standard methods.

3. PROOF OF THEOREM 1.1

In the section above we have collected all the tools and roughly described the road we will follow in order to proof Theorem 1.1. We will proceed in six steps in Sections 3.1-3.6. The first will be a reformulation of the problem in Section 3.1. Then we divide the range of summation up into three parts according to the position of the digit under consideration. In the Sections 3.3, 3.2 and 3.4 we consider “middle” digits, least and most significant digits, respectively. The next step (Section 3.5) will help us giving the error term a shape as in the result of Delange [5]. Finally, in Section 3.6, we put everything together and gain the desired result.

3.1. Rewriting the problem. First of all we split the sum in (1.7) up as follows

$$(3.1) \quad S(N) := \sum_{N < n \leq 2N} f(\lfloor p(n) \rfloor)$$

Now the idea is to use the q -additivity of f and exchange the sum over n with the sum over the digits.

In order to perform the exchange of the sum over n and the sum over the digits we set

$$\begin{aligned} (3.2) \quad T_{r,a}(N) &:= \sum_{N < n \leq 2N} \left(\mathcal{I}_{r,a}(p(n)) - \frac{1}{q} \right) \\ &= \sum_{N < n \leq 2N} \psi \left(\frac{p(n)}{q^{r+1}} - \frac{a+1}{q} \right) - \psi \left(\frac{p(n)}{q^{r+1}} - \frac{a}{q} \right) \\ U_{r,a}(N) &:= \sum_{N < n \leq 2N} \psi \left(\frac{p(n)}{q^{r+1}} - \frac{a}{q} \right). \end{aligned}$$

Now using the two representations in (2.1) and plugging (3.2) into (3.1) we get that

$$\begin{aligned} (3.3) \quad S(N) &= \sum_{N < n \leq 2N} \sum_{0 \leq r \leq \log_q(p(2N))} \sum_{a=0}^{q-1} f(a) \mathcal{I}_{r,a}(p(n)) \\ &= \mu_f N \log_q(p(2N)) + \sum_{a=0}^{q-1} f(a) \sum_{0 \leq r \leq \log_q(p(2N))} T_{r,a}(N) + \mathcal{O}(1) \\ &= \mu_f N \log_q(p(2N)) + \sum_{a=0}^{q-1} f(a) \sum_{0 \leq r \leq \log_q(p(2N))} U_{r,a+1}(N) - U_{r,a}(N) + \mathcal{O}(1). \end{aligned}$$

The goal of the following sections will be the estimation of $U_{r,a}(N)$ for fixed values of r and a . Our choice of the two different representations will depend on the size of r .

3.2. “Middle” digits. We will use Lemma 2.2 with $s = k$ and show that the requirements are always fulfilled.

Lemma 3.1. *Let $H = 2^{k+1}5 + 8$ and $N \geq 1$ be a real. Then*

$$(3.4) \quad \sum_{(\log N)^H \leq q^r \leq N^{k-1}} U_{r,a}(N) \ll \frac{N}{(\log N)^2}.$$

This lemma provides us with the bounds for the least and most significant digits. In order to make Lemma 3.2 always applicable for the “middle” digits we define the three ranges to be

$$(3.5) \quad 0 \leq q^r \leq (\log N)^H, \quad (\log N)^H \leq q^r \leq N^{k-1}, \quad \text{and} \quad N^{k-1} \leq q^r \leq N^k,$$

where $H = 2^{k+1}5 + 8$.

Proof. We start by applying Lemma 2.1 with $\delta = (\log N)^4$, which yields

$$(3.6) \quad U_{r,a}(N) \ll N(\log N)^{-4} + \sum_{\nu=1}^{(\log N)^8} \frac{1}{\nu} |S(N, r, \nu)|,$$

where

$$(3.7) \quad S(N, r, \nu) = \sum_{N < n \leq 2N} e\left(\frac{\nu p(n)}{q^{r+1}}\right).$$

We rewrite the middle bound on q^r in (3.5) into one for r and get

$$(3.8) \quad H \frac{\log \log N}{\log q} \leq r \leq (k-1) \frac{\log N}{\log q}.$$

Now we fix r and ν and show that we always can apply Lemma 2.2 with $s = k$. Let $h = H - 8 = 2^{k+1}5$. Then by Diophantine approximation there always exists a rational a/b with $b > 0$, $(a, b) = 1$,

$$1 \leq b \leq N^k (\log N)^{-h} \quad \text{and} \quad \left| \frac{\nu \alpha_k}{q^{r+1}} - \frac{a}{b} \right| \leq \frac{(\log N)^h}{b N^k}.$$

We distinguish three cases according to the size of b .

Case 1, $(\log N)^h \leq b$: In this case we may apply Lemma 2.2 with $s = k$ and get

$$(3.9) \quad S(N, r, \nu) \ll N(\log N)^{-3}.$$

Case 2, $2 \leq b < (\log N)^h$: In this case we get that

$$\left| \frac{\nu \alpha_k}{q^{r+1}} \right| \geq \left| \frac{a}{b} \right| - \frac{1}{b^2} \geq \frac{1}{2b} \geq \frac{1}{2} (\log N)^{-h}.$$

Thus in view of r we have

$$r + 1 \leq (h + 2) \frac{\log \log N}{\log q} + o(1),$$

which contradicts the lower bound in (3.8) for N sufficiently large.

Case 3, $b = 1$: Now we have to distinguish two cases according to whether $a = 0$ or not.

Case 3.1, $\left| \frac{\nu \alpha_k}{q^{r+1}} \right| \geq \frac{1}{2}$: This implies that

$$r + 1 \leq 2 \frac{\log \log N}{\log q} + o(1)$$

contradicting the lower bound in (3.8) again.

Case 3.2, $\left| \frac{\nu \alpha_k}{q^{r+1}} \right| < \frac{1}{2}$: We clearly have $a = 0$, which implies

$$\left| \frac{\nu \alpha_k}{q^{r+1}} \right| \leq \frac{(\log N)^h}{N^k}.$$

Noting that $\nu \leq (\log N)^8$ yields for r that

$$k \frac{\log N}{\log q} - (h-8) \frac{\log \log N}{\log q} + o(1) \leq r+1,$$

this time contradicting the upper bound in (3.8).

The only possibility satisfying (3.8) is **Case 1**. Thus plugging (3.9) into (3.6)

$$U_{r,a}(N) \ll \frac{N}{(\log N)^4} + \frac{N}{(\log N)^4} \sum_{\nu=1}^{(\log N)^8} \frac{1}{\nu} \ll \frac{N}{(\log N)^3}.$$

This together with the middle range of r in (3.5) proves the lemma. \square

3.3. Least significant digits. In this section we concentrate on the least significant digits. As here Lemma 2.2 might not be applicable we will use the fact that there is a coefficient having bounded continued fraction expansion. The aim of this section is the following

Lemma 3.2. *Let $H = 2^{k+1}5 + 8$ and N be a positive integer. Then*

$$(3.10) \quad \sum_{0 \leq q^r \leq (\log N)^H} U_{r,a}(N) \ll \frac{N}{(\log N)^2}.$$

Proof. This time we start by applying Lemma 2.1 with $\delta = (\log N)^3$, which yields

$$(3.11) \quad U_{r,a}(N) \ll N(\log N)^{-3} + \sum_{\nu=1}^{(\log N)^6} \frac{1}{\nu} |S(N, r, \nu)|,$$

where $S(N, r, \nu)$ is defined in (3.7).

We fix r and ν . By our assumption there is a $1 \leq j \leq k$ such that α_j satisfies (1.3) with a $K \in \mathbb{N}$. The idea is to consider the coefficients one after the other whether we can apply Lemma 2.2 or not. The bounded continued fraction expansion together with Lemma 2.3 will guarantee that we can apply Lemma 2.2 at least for α_j . Therefore we recursively set

$$\begin{aligned} H_k^* &= 2^{k+3}, \\ H_k &= H_k^* + 2^{k+3}, \\ H_i^* &= 2^{i+1}(4 + H_{i+1}) + 2i(H_{i+1} + \dots + H_k) \quad (i = 1, \dots, k-1), \\ H_i &= H_i^* + 2(H \log q + 3) \quad (i = 1, \dots, k-1), \\ h_i &= H_i^* + H \log q + 3 \quad (i = 1, \dots, k). \end{aligned}$$

Now we distinguish whether Lemma 2.2 is applicable or not. In particular we consider whether there exists a $j < s \leq k$ such that there are rational numbers $a_{s+1}/b_{s+1}, \dots, a_k/b_k$ with

$$(3.12) \quad 1 \leq b_i \leq (\log N)^{2h_i} \quad \text{and} \quad \left| \alpha_i - \frac{a_i}{b_i} \right| \leq \frac{(\log N)^{h_i}}{b_i N^i} \quad (s < i \leq k),$$

but there is no rational number a_s/b_s with $(a_s, b_s) = 1$ such that

$$(3.13) \quad 1 \leq b_s \leq (\log N)^{2h_s} \quad \text{and} \quad \left| \alpha_s - \frac{a_s}{b_s} \right| \leq \frac{(\log N)^{h_s}}{b_s N^s}$$

or not.

Case 1: There exists a $j < s \leq k$ such that (3.12) and (3.13) hold. Noting that $1 \leq \nu \leq (\log N)^6$ we get that

$$1 \leq q^{r+1} b_i \leq (\log N)^{2H_i} \quad \text{and} \quad \left| \frac{\nu \alpha_i}{q^{r+1}} - \frac{\nu a_i}{q^{r+1} b_i} \right| \leq \frac{(\log N)^{h_i}}{q^{r+1} b_i N^i}$$

for $s < i \leq k$, but there is no rational number A_s/B_s with $(A_s, B_s) = 1$ such that

$$1 \leq B_s \leq (\log N)^{2H_s^*} \quad \text{and} \quad \left| \frac{\nu \alpha_s}{q^{r+1}} - \frac{A_s}{B_s} \right| \leq \frac{(\log N)^{H_s^*}}{B_s N^s},$$

since otherwise we would get that

$$1 \leq \nu B_s \leq (\log N)^{2H_s^*+6} \leq (\log N)^{2h_0}$$

and

$$\left| \alpha_s - \frac{q^{r+1} A_s}{\nu B_s} \right| \leq \frac{(\log N)^{H_0^*} + H \log q + 1}{\nu B_s X^s} \leq \frac{(\log N)^{h_s}}{\nu B_s X^s},$$

which contradicts our assumption (3.13) on s .

Case 2: There is no $j < s \leq k$ such that (3.12) and (3.13) hold. In order to apply Lemma 2.2 it suffices to show that (3.13) holds with $s = j$. We will prove this by contradiction: assume that there is a rational number a_j/b_j with $(a_j, b_j) = 1$ such that

$$1 \leq b_j \leq (\log N)^{2h_j} \quad \text{and} \quad \left| \alpha_j - \frac{a_j}{b_j} \right| \leq \frac{(\log N)^{h_j}}{b_j N^j}$$

holds. Now an application of Lemma 2.3 yields

$$\frac{1}{(K+1)b_j^2} < \left| \alpha_j - \frac{a_j}{b_j} \right| \leq \frac{(\log N)^{h_j}}{b_j N^j},$$

which implies that

$$\frac{N^j}{(K+1)(\log N)^{h_j}} < b_j$$

contradicting our assumption.

Thus in both cases we may apply Lemma 2.2 with $K_i = 2H_i$ and $K_s^* = 2H_s^*$ and get

$$S(N, r, \nu) \ll \frac{N}{(\log N)^3}.$$

Plugging this in (3.11) and (3.10) yields

$$(3.14) \quad \sum_{0 \leq q^r \leq (\log N)^H} U_{r,a}(N) \ll \sum_{0 \leq r \leq H \frac{\log \log N}{\log q}} \left(\frac{N}{(\log N)^3} + \frac{N}{(\log N)^3} \sum_{\nu=1}^{(\log N)^6} \frac{1}{\nu} \right) \ll \frac{N}{(\log N)^2},$$

which proves the lemma. \square

3.4. Most significant digits. For the digits at high positions we will use a different method which goes back to Peter [20].

Lemma 3.3. *Let $H > 0$ be real and N be a positive integer. If*

$$(k-1) \log_q N \leq r \leq k \log_q N$$

then

$$(3.15) \quad U_{r,a}(N) = \frac{1}{k} \alpha_k^{-\frac{1}{k}} q^{\frac{r+1}{k}} \int_{q^{-r-1}p(N)}^{q^{-r-1}p(2N)} \psi \left(t - \frac{a}{q} \right) t^{1/k-1} dt + \mathcal{O} \left(N^{\frac{2}{3}} \right),$$

Proof. In order to apply Lemma 2.4 we need that the function under consideration is monotonically decreasing. We guarantee this by rewriting $U_{r,a}$ as follows.

$$U_{r,a}(N) = \sum_{n \leq N} \psi \left(\frac{p(\lfloor 2N \rfloor - n)}{q^{r+1}} - \frac{a}{q} \right).$$

Then an application of Lemma 2.4 yields

$$(3.16) \quad \begin{aligned} U_{r,a}(N) &= -q^{-r-1} \int_0^N \psi(t) p'(\lfloor 2N \rfloor - t) dt \\ &+ \sum_{q^{-r-1}p(\lfloor 2N \rfloor - N) - \frac{a}{q} < m \leq q^{-r-1}p(\lfloor 2N \rfloor) - \frac{a}{q}} \psi \left(\lfloor 2N \rfloor - p^{-1} \left(q^{r+1} \left(m + \frac{a}{q} \right) \right) \right) \\ &+ q^{r+1} \int_{q^{-r-1}p(\lfloor 2N \rfloor) - N}^{q^{-r-1}p(\lfloor 2N \rfloor)} \frac{\psi \left(t - \frac{a}{q} \right)}{p'(p^{-1}(q^{r+1}t))} dt + \mathcal{O}(1). \end{aligned}$$

We will write \asymp if both \ll and \gg hold. Since p is a polynomial of degree k we can write $p(t) = \alpha_k t^k + \mathcal{O}(t^{k-1})$, which yields for the inverse of p

$$p^{-1}(t) = \left(\frac{t}{\alpha_k} \right)^{\frac{1}{k}} + \mathcal{O}(1).$$

Implicitly calculating the derivatives gives

$$(3.17) \quad \begin{aligned} (p^{-1})'(t) &= \frac{1}{p'(p^{-1}(t))} = \frac{1}{k} \alpha_k^{-\frac{1}{k}} t^{\frac{1}{k}-1} + \mathcal{O}(t^{-1}), \\ (p^{-1})''(x) &= -\frac{p''(p^{-1}(t))}{(p'(p^{-1}(t)))^3} \asymp -\frac{k-1}{k^2} \alpha_k^{-\frac{1}{k}} t^{\frac{1}{k}-2}. \end{aligned}$$

Now we consider the three parts of (3.16) and define $\psi_1(x) := \int_0^x \psi(t) dt$ for $x \in \mathbb{R}$. It is clear, that $\psi_1(x)$ is continuous, bounded and piecewise continuously differentiable. Integration by parts for the first integral in (3.16) yields the estimate

$$(3.18) \quad \int_0^N \psi(t) p'(\lfloor 2N \rfloor - t) dt = \psi_1(t) p'(\lfloor 2N \rfloor - t) \Big|_0^N + \int_0^N \psi_1(t) p''(\lfloor 2N \rfloor - t) dt \ll N^{k-1}.$$

In order to estimate the sum we define $g(t) := \lfloor 2N \rfloor - p^{-1}(q^{r+1}(t + \frac{a}{q}))$ for $t \geq 0$. For the range $q^{-r-1}p(\lfloor 2N \rfloor - N) - \frac{a}{q} < t \leq q^{-r-1}p(\lfloor 2N \rfloor) - \frac{a}{q}$, we have $t \asymp q^{-r-1}N^k$ and thus by (3.17)

$$|g''(t)| \asymp q^{(r+1)/k} t^{1/k-2} \asymp q^{2(r+1)} N^{1-2k}.$$

Then an application of Lemma 2.5 yields for the sum in (3.16) that

$$(3.19) \quad \sum_{q^{-r-1}g(\lfloor 2N \rfloor - N) - \frac{a}{q} < m \leq q^{-r-1}g(\lfloor 2N \rfloor) - \frac{a}{q}} \psi(g(m)) \ll q^{-(r+1)/3} N^{(k+1)/3} + q^{-r-1} N^{k-\frac{1}{2}} \ll N^{\frac{2}{3}}.$$

Now we aim for the main term which we extract from the second integral of (3.16). Therefore we define $I_{r,a}(N)$ to be the following integral

$$(3.20) \quad I_{r,a}(N) := \int_{q^{-r-1}p(N)}^{q^{-r-1}p(2N)} \psi\left(t - \frac{a}{q}\right) t^{\frac{1}{k}-1} dt.$$

Noting that $\psi(t) \ll 1$ we get with (3.17)

$$(3.21) \quad \begin{aligned} & q^{-(r+1)} \int_{q^{-r-1}p(\lfloor 2N \rfloor - N)}^{q^{-r-1}p(\lfloor 2N \rfloor)} \frac{\psi\left(t - \frac{a}{q}\right)}{p'(p^{-1}(q^{r+1}t))} dt \\ &= q^{-(r+1)} \int_{q^{-r-1}p(\lfloor 2N \rfloor - N)}^{q^{-r-1}p(\lfloor 2N \rfloor)} \psi\left(t - \frac{a}{q}\right) \left(\frac{1}{k} \alpha_k^{-\frac{1}{k}} (q^{r+1}t)^{\frac{1}{k}-1} + \mathcal{O}((q^{r+1}t)^{-1}) \right) dt \\ &= \frac{1}{k} \alpha_k^{-\frac{1}{k}} q^{\frac{r+1}{k}} \int_{q^{r+1}p(\lfloor 2N \rfloor - N)}^{q^{-r-1}p(\lfloor 2N \rfloor)} \psi\left(t - \frac{a}{q}\right) t^{\frac{1}{k}-1} dt + \mathcal{O}\left(\int_{q^{-r-1}p(\lfloor 2N \rfloor - N)}^{q^{-r-1}p(\lfloor 2N \rfloor)} \psi\left(t - \frac{a}{q}\right) t^{-1} dt \right) \\ &= \frac{1}{k} \alpha_k^{-\frac{1}{k}} q^{\frac{r+1}{k}} \left(I_{r,a}(N) + \left\{ \int_{q^{-r-1}p(\lfloor 2N \rfloor - N)}^{q^{-r-1}p(N)} + \int_{q^{r+1}p(2N)}^{q^{-r-1}p(\lfloor 2N \rfloor)} \right\} \psi\left(t - \frac{a}{q}\right) t^{\frac{1}{k}-1} dt \right) \\ &\quad + \mathcal{O}(\log N) \\ &= \frac{1}{k} \alpha_k^{-\frac{1}{k}} q^{\frac{r+1}{k}} I_{r,a}(N) + \mathcal{O}(\log N) \end{aligned}$$

Plugging (3.18), (3.19) and (3.21) into (3.16) proves the lemma. \square

3.5. Extracting the periodic function. We will now use the integral representation as in (3.15) in order to extract the periodic function. Since Lemma 3.3 looks similar to Lemma 2.3 of [20] we will follow the lines there.

For $x \geq 0$ and a an integer we define

$$(3.22) \quad I_a(x) := \int_0^x \psi\left(t - \frac{a}{q}\right) t^{\frac{1}{k}-1} dt = \int_0^x \psi\left(t + \frac{q-a}{q}\right) t^{\frac{1}{k}-1} dt,$$

where we have used that ψ is 1-periodic. Then J is defined to be the following function

$$(3.23) \quad J(x) := \sum_{a=0}^{q-1} f(a) (I_{a+1}(x) - I_a(x)) = \sum_{a=0}^{q-1} f(a) \int_0^x \left(\psi\left(t - \frac{a+1}{q}\right) - \psi\left(t - \frac{a}{q}\right) \right) t^{\frac{1}{k}-1} dt.$$

Integration by parts of the second representation in (3.22) yields $I_a(x) \ll 1$ and thus

$$(3.24) \quad J(x) \ll 1, \quad x \geq 0.$$

Now we concentrate on the sum over the r of the $U_{r,a}$. For $N \geq 1$ and r and a positive integers we clearly have $I_{r,a} = I_a(q^{-r-1}p(2N)) - I_a(q^{-r-1}p(N))$ with $I_{r,a}$ defined in (3.20). Thus

$$(3.25) \quad \begin{aligned} & \sum_{r \leq \log_q(g(2N))} q^{(r+1)/k} \sum_{a=0}^{q-1} f(a) (I_{r,a+1} - I_{r,a}) \\ &= \sum_{r \leq \log_q(g(2N))} q^{(r+1)/k} J(q^{-r-1}p(2N)) - \sum_{r \leq \log_q(g(2N))} q^{(r+1)/k} J(q^{-r-1}p(N)) \\ &=: S_1(N) - S_2(N), \quad N \geq 1. \end{aligned}$$

Define

$$R(x) := \lfloor \log_q(p(x)) \rfloor \quad \text{and} \quad F(x) := \sum_{n \geq 0} q^{-\frac{n}{k}} J(q^{x+n-\frac{1}{2}}), \quad x \in \mathbb{R}.$$

By noting that $p(t) = \alpha_k t^k + \mathcal{O}(t^{k-1})$ we get

$$(3.26) \quad \begin{aligned} S_1(N) &= \sum_{r \leq R(2N)} q^{(R(2N)-r+1)/k} J(q^{r-R(2N)-1}p(2N)) \\ &= p(2N)^{\frac{1}{k}} q^{(\frac{1}{2}-\psi(R(2N)))/k} \sum_{r \leq R(2N)} q^{-\frac{r}{k}} J(q^{\psi(R(2N))+r-\frac{1}{2}}) \\ &= \alpha_k^{\frac{1}{k}} 2N q^{(\frac{1}{2}-\psi(R(2N)))/k} F(\psi(R(2N))) + \mathcal{O}(1), \end{aligned}$$

where we have used (3.24) and

$$\sum_{r \leq R(2N)} q^{-\frac{r}{k}} \ll N^{-1}, \quad \sum_{r > R(2N)} q^{-\frac{r}{k}} \ll N^{-1}.$$

For $S_2(N)$ we get

$$(3.27) \quad S_2(N) = S_1\left(\frac{N}{2}\right) + \sum_{R(N/2) < r \leq R(N)} q^{\frac{r+1}{k}} J(q^{-r-1}p(N)).$$

For the second sum we have that $R(N/2) < r \leq R(N)$, which implies $q^{-r-1}p(N) < \frac{1}{q}$. Thus connecting the representations in (2.1) and (3.23) we gain for $0 \leq x < \frac{1}{q}$

$$J(x) = \int_0^x \sum_{a=0}^{q-1} f(a) \left(\mathcal{I}_a(t) - \frac{1}{q} \right) t^{\frac{1}{k}-1} dt = -\mu_f k x^{\frac{1}{k}}.$$

Now putting (3.26) and (3.27) together yields

$$(3.28) \quad S_2(N) = \alpha_k^{\frac{1}{k}} N q^{(\frac{1}{2}-\psi(R(N)))/k} F(\psi(R(N))) - \mu_f k \alpha_k^{\frac{1}{k}} N (R(N) - R(N/2)) + \mathcal{O}(1).$$

3.6. Putting all together. Now we plug (3.10), (3.4), and (3.15) into (3.3) to get

$$\begin{aligned}
 S(N) &= \mu_f NR(2N) + \sum_{a=0}^{q-1} f(a) \sum_{0 \leq r \leq \log_q(p(2N))} U_{r,a+1}(N) - U_{r,a}(N) + \mathcal{O}(1) \\
 &= \mu_f 2NR(2N) - \mu_f NR(N) + \mathcal{O}\left(\frac{N}{(\log N)^2}\right) \\
 (3.29) \quad &+ \frac{1}{k} \alpha_k^{-\frac{1}{k}} \sum_{a=0}^{q-1} f(a) \sum_{0 \leq r \leq \log_q(p(2N))} q^{(r+1)/k} (I_{r,a+1}(N) - I_{r,a}(N)) \\
 &- \frac{1}{k} \alpha_k^{-\frac{1}{k}} \sum_{a=0}^{q-1} f(a) \sum_{q^r \leq N^{k-1}} q^{(r+1)/k} (I_{r,a+1}(N) - I_{r,a}(N)) \\
 &+ \mathcal{O}(1)
 \end{aligned}$$

Now integration by parts gives for positive integers a and r , and $N \geq 1$

$$(3.30) \quad I_{r,a} \ll q^{\left(1-\frac{1}{k}\right)(r+1)} N^{1-k}.$$

Thus using (3.25), (3.26), (3.28) and (3.30) we get

$$\begin{aligned}
 S(N) &= \mu_f 2NR(2N) - \mu_f NR(N) \\
 &+ 2Nq^{\left(\frac{1}{2}-\psi(R(2N))\right)/k} F(\psi(R(2N))) \\
 &- Nq^{\left(\frac{1}{2}-\psi(R(N))\right)/k} F(\psi(R(N))) \\
 &+ \mathcal{O}\left(\frac{N}{(\log N)^2}\right)
 \end{aligned}$$

Finally summing up over $N = 2^{-i}x$ for $1 \leq i \leq \log_2 x$ yields

$$\begin{aligned}
 \sum_{n \leq x} f(\lfloor p(n) \rfloor) &= \sum_{1 \leq i \leq \log_2 x} S(2^{-i}x) + \mathcal{O}(1) \\
 &= \mu_f x \lfloor \log_q p(x) \rfloor + \frac{1}{k} x q^{(1/2-\psi(\log_q(p(x))))/k} F(\psi(\lfloor \log_q(p(x)) \rfloor)) \\
 &+ \mathcal{O}\left(\frac{x}{\log x}\right),
 \end{aligned}$$

which proves the theorem.

REFERENCES

- [1] N. L. Bassily and I. Kátai, *Distribution of the values of q -additive functions on polynomial sequences*, Acta Math. Hungar. **68** (1995), no. 4, 353–361.
- [2] J. Cassaigne, S. Ferenczi, C. Mauduit, J. Rivat, and A. Sárközy, *On finite pseudorandom binary sequences. III. The Liouville function. I*, Acta Arith. **87** (1999), no. 4, 367–390.
- [3] ———, *On finite pseudorandom binary sequences. IV. The Liouville function. II*, Acta Arith. **95** (2000), no. 4, 343–359.
- [4] J. Cassaigne, C. Mauduit, and A. Sárközy, *On finite pseudorandom binary sequences. VII. The measures of pseudorandomness*, Acta Arith. **103** (2002), no. 2, 97–118.
- [5] H. Delange, *Sur la fonction sommatoire de la fonction “somme des chiffres”*, Enseignement Math. (2) **21** (1975), no. 1, 31–47.
- [6] P. Flajolet, P. Grabner, P. Kirschenhofer, H. Prodinger, and R. F. Tichy, *Mellin transforms and asymptotics: digital sums*, Theoret. Comput. Sci. **123** (1994), no. 2, 291–314.
- [7] A. O. Gel’fond, *Sur les nombres qui ont des propriétés additives et multiplicatives données*, Acta Arith. **13** (1967/1968), 259–265.
- [8] B. Gittenberger and J. M. Thuswaldner, *The moments of the sum-of-digits function in number fields*, Canad. Math. Bull. **42** (1999), no. 1, 68–77.
- [9] P. J. Grabner and H.-K. Hwang, *Digital sums and divide-and-conquer recurrences: Fourier expansions and absolute convergence*, Constr. Approx. **21** (2005), no. 2, 149–179.

- [10] P. J. Grabner, P. Kirschenhofer, H. Prodinger, and R. F. Tichy, *On the moments of the sum-of-digits function*, Applications of Fibonacci numbers, Vol. 5 (St. Andrews, 1992), Kluwer Acad. Publ., Dordrecht, 1993, pp. 263–271.
- [11] P. Kirschenhofer, *On the variance of the sum of digits function*, Number-theoretic analysis (Vienna, 1988–89), Lecture Notes in Math., vol. 1452, Springer, Berlin, 1990, pp. 112–116.
- [12] E. Krätzel, *Lattice points*, Mathematics and its Applications (East European Series), vol. 33, Kluwer Academic Publishers Group, Dordrecht, 1988.
- [13] C. Mauduit and J. Rivat, *Répartition des fonctions q -multiplicatives dans la suite $([n^c])_{n \in \mathbb{N}}$, $c > 1$* , Acta Arith. **71** (1995), no. 2, 171–179.
- [14] ———, *La somme des chiffres des carrés*, Acta Math. **203** (2009), no. 1, 107–148.
- [15] C. Mauduit and A. Sárközy, *On finite pseudorandom binary sequences. I. Measure of pseudorandomness, the Legendre symbol*, Acta Arith. **82** (1997), no. 4, 365–377.
- [16] ———, *On finite pseudorandom binary sequences. II. The Champernowne, Rudin-Shapiro, and Thue-Morse sequences, a further construction*, J. Number Theory **73** (1998), no. 2, 256–276.
- [17] ———, *On finite pseudorandom binary sequences. V. On $(n\alpha)$ and $(n^2\alpha)$ sequences*, Monatsh. Math. **129** (2000), no. 3, 197–216.
- [18] ———, *On finite pseudorandom binary sequences. VI. On $(n^k\alpha)$ sequences*, Monatsh. Math. **130** (2000), no. 4, 281–298.
- [19] Y. Nakai and I. Shiokawa, *Discrepancy estimates for a class of normal numbers*, Acta Arith. **62** (1992), no. 3, 271–284.
- [20] M. Peter, *The summatory function of the sum-of-digits function on polynomial sequences*, Acta Arith. **104** (2002), no. 1, 85–96.
- [21] I. Shiokawa, *On the sum of digits of prime numbers*, Proc. Japan Acad. **50** (1974), 551–554.
- [22] J. M. Thuswaldner, *The sum of digits function in number fields*, Bull. London Math. Soc. **30** (1998), no. 1, 37–45.

(M. G. Madritsch) DEPARTMENT OF ANALYSIS AND COMPUTATIONAL NUMBER THEORY
GRAZ UNIVERSITY OF TECHNOLOGY
8010 GRAZ, AUSTRIA
E-mail address: madritsch@math.tugraz.at