ON THE SUM OF DIGITS FUNCTION FOR NUMBER SYSTEMS WITH NEGATIVE BASES

PETER J. GRABNER AND JÖRG M. THUSWALDNER

Abstract. Let $q \geq 2$ be an integer. Then $-q$ gives rise to a number system in $\mathbb{Z}$, i.e., each number $n \in \mathbb{Z}$ has a unique representation of the form $n = c_0 + c_1(-q) + \ldots + c_h(-q)^h$, with $c_i \in \{0, \ldots, q-1\}$ $(0 \leq i \leq h)$. The aim of this paper is to investigate the sum of digits function $\nu_{-q}(n)$ of these number systems. In particular, we derive an asymptotic expansion for

$$\sum_{n<N} |\nu_{-q}(n) - \nu_{-q}(-n)|$$

and obtain a Gaussian asymptotic distribution result for $\nu_{-q}(n) - \nu_{-q}(-n)$. Furthermore, we prove non-differentiability of certain continuous occurring in this context. We use automata and analytic methods to derive our results.

1. Introduction and Statement of Results

In this paper we study the sum of digits function of number systems with negative integer base. These number systems were first studied intensively by Knuth [19]. It is easy to prove that for an integer q, $q \geq 2$, every $n \in \mathbb{Z}$ has a unique representation of the form

$$n = c_0 + c_1(-q) + c_2(-q)^2 + \cdots + c_h(-q)^h \tag{1.1}$$

with $c_i \in \{0, 1, \ldots q-1\}$ $(0 \leq i \leq h)$ and $c_h \neq 0$ for $h \neq 0$. Expression (1.1) is called the $(-q)$-adic representation of n. We denote the string of digits $c_0, c_1, \ldots, c_h$ of this representation by $(n)_{-q}$. The sum of digits function of n is defined by

$$\nu_{-q}(n) = c_0 + c_1 + \ldots + c_h.$$

The sum of digits function of ordinary q-adic number systems $(q \geq 2)$ gained interest in the last thirty years. Delange [6] found an exact formula for its summatory function. He proved that

$$\sum_{n<N} \nu_q(n) = \frac{q-1}{2} N \log_q N + N F(\log_q N),$$

where F is a continuous, nowhere differentiable function of period one. The Fourier coefficients of F were computed by elementary methods in [6] and by means of standard methods from analytic number theory in [11].

Date: August 29, 2011.
1991 *Mathematics Subject Classification.* Primary: 11B85; Secondary: 11A63, 05C50.
Key words and phrases. digital expansions, sum of digits, finite automata, non-differentiability.
The first author is supported by the START project Y96-MAT of the Austrian Science Fund.

Also for the sum of digits function with respect to negative bases an exact formula for the sum of digits function is known. One can show that (cf. [12, 25])

$$\sum_{|n|<N} \nu_{-q}(n) = (q-1)N\log_q N + N\Phi(\log_{q^2} N), \tag{1.2}$$

where Φ is again a continuous function of period 1.

The moments of the sum of digits function were calculated in [5, 18, 23] in the case of positive q and in [15] for negative q. Furthermore, we mention that Gelfond [14] studied the distribution of $\nu_q(n)$ in residue classes. This result has recently been extended to more general number systems in [26].

Dumont and Thomas [8] studied the moments of the sum of digits function in generalized number systems arising from substitutions and finite automata. They study the behaviour of sums of $f(\nu(n))$, where $\nu(n)$ denotes the sum of digits function of a number system associated to a substitution and f is a real function subject to some growth conditions. Their results imply that $\nu(n)$ is asymptotically normally distributed. Similar results for number systems whose base sequences satisfy linear recurrences were obtained in [7]. Limit distributions for the q-adic sum of digits function on polynomial values are derived in [2].

In this paper we prove a result on the asymptotic distribution of $\nu_{-q}(n) - \nu_{-q}(-n)$. We use an estimate for certain exponential sums needed for the distribution result to give an asymptotic expansion for the sum of $|\nu_{-q}(n) - \nu_{-q}(-n)|$.

As we will show in Section 2, $\nu_{-q}(n) - \nu_{-q}(-n)$ can be generated by a finite state automaton from the q-adic expansion of n. But since its values can be arbitrarily large, it is no q-automatic function. On the other hand, it is a $(-q)$-additive function which is not q-additive, despite of the fact that it satisfies a weaker additivity property, which will be exploited in the proof of Theorem 2. Thus despite $\nu_{-q}(n) - \nu_{-q}(-n)$ is closely related to q-automatic as well as q-additive functions it does not belong to one of these classes.

For a general description of q-automatic and q-additive functions and the asymptotic of their summatory functions we refer to [4]; for an analytic approach to exact and asymptotic formulæ for summatory functions of q-additive functions we refer to [11, 21, 22]. A discussion of periodicity phenomena and non-differentiability of the remainder functions occurring in this context is given in [24].

Theorem 1. *Let $\nu_{-q}(n)$ denote the sum of digits function with respect to the base $-q$ for an integer $q \geq 2$. Then the asymptotic formulæ*

$$\sum_{n<N} (\nu_{-q}(n) - \nu_{-q}(-n)) = NG(\log_{q^2} N) + \mathcal{O}\left((\log N)^2\right) \tag{1.3}$$

and

$$\sum_{n<N} (\nu_{-q}(n) - \nu_{-q}(-n))^2 = 2r_1(q)N\log_q N + NH(\log_q N) + \mathcal{O}\left((\log N)^3\right) \tag{1.4}$$

hold with continuous periodic functions $G(x)$ and $H(x)$ of period 1. *$G(x)$ satisfies $G(x+\frac{1}{2}) = -G(x)$ and therefore has mean* 0. *Furthermore,*

$$\sum_{n<N} |\nu_{-q}(n) - \nu_{-q}(-n)| = 2\sqrt{\frac{r_1(q)}{\pi}} N \sqrt{\log_q N} + \frac{N}{\sqrt{\log_q N}} \Psi\left(\log_q N\right) + \mathcal{O}\left(\frac{N}{(\log N)^{\frac{3}{2}}}\right), \tag{1.5}$$

hold, where

$$\Psi(x) = \frac{1}{2\sqrt{\pi r_1(q)}} H(x) + \frac{r_2(q)}{2\sqrt{\pi r_1(q)^3}} - \frac{1}{3Q^2\sqrt{\pi r_1(q)}} \tag{1.6}$$

and

$$\begin{aligned} r_1(q) &= \frac{1}{6}\frac{(q-1)(q^2-4q+7)}{(q+1)} \\ r_2(q) &= \frac{1}{180}\frac{(q-1)(q^6-11q^5+82q^4-322q^3+577q^2-371q+76)}{(q+1)^3}; \end{aligned} \tag{1.7}$$

$Q = 2$ if q is even and $Q = 1$ if q is odd.

Theorem 2. *The functions $G(x)$ and $H(x)$ in* (1.3) *and* (1.4) *are nowhere differentiable; therefore, by* (1.6), *$\Psi(x)$ in* (1.5) *is nowhere differentiable.*

Corollary 1.

$$\begin{aligned} \sum_{n<N} \min(\nu_{-q}(n), \nu_{-q}(-n)) &= \frac{q-1}{2} N \log_q N - \sqrt{\frac{r_1(q)}{\pi}} N (\log_q N)^{1/2} + \frac{1}{2}\Phi(\log_{q^2} N) N \\ &\quad -\frac{1}{2}\Psi(\log_q N)\frac{N}{(\log_q N)^{1/2}} + \mathcal{O}\left(\frac{N}{(\log N)^{3/2}}\right), \end{aligned}$$

and

$$\begin{aligned} \sum_{n<N} \max(\nu_{-q}(n), \nu_{-q}(-n)) &= \frac{q-1}{2} N \log_q N + \sqrt{\frac{r_1(q)}{\pi}} N (\log_q N)^{1/2} + \frac{1}{2}\Phi(\log_{q^2} N) N \\ &\quad +\frac{1}{2}\Psi(\log_q N)\frac{N}{(\log_q N)^{1/2}} + \mathcal{O}\left(\frac{N}{(\log N)^{3/2}}\right), \end{aligned}$$

where Φ and Ψ are the periodic functions occurring in (1.2) *and* (1.5), *respectively.*

Theorem 3. *Let $\nu_{-q}(n)$ denote the sum of digits function with respect to the base $-q$ for an integer $q \geq 2$. Then the quantity $\nu_{-q}(n) - \nu_{-q}(-n)$ is asymptotically normally distributed with mean* 0 *and variance $2r_1(q)\log_q N$, i.e.*

$$\frac{1}{N}\#\left\{n < N \;\middle|\; \frac{\nu_{-q}(n) - \nu_{-q}(-n)}{\sqrt{2r_1(q)\log_q N}} < x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{\xi^2}{2}}\, d\xi + \mathcal{O}\left(\frac{1}{\sqrt{\log N}}\right), \tag{1.8}$$

where $r_1(q)$ is given by (1.7).

In order to prove these theorems we construct an automaton that maps the string of digits of the q-adic representation of n, denoted by $(n)_q$, to a string of digits, whose sum is $\nu_{-q}(n)-\nu_{-q}(-n)$. Using the accompanying matrix of this automaton we get an asymptotic expansion for the sum

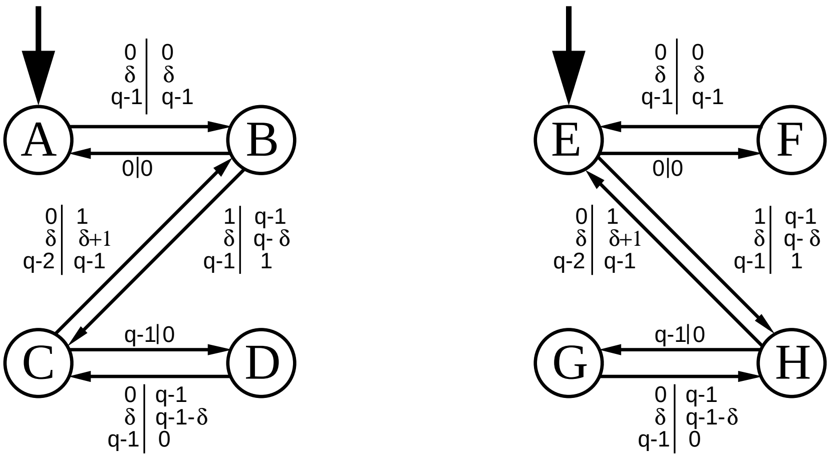
$$\sum_{n<N} e^{it(\nu_{-q}(n)-\nu_{-q}(-n))}. \tag{1.9}$$

Starting from this expansion we will derive the result via Fourier transform.

2. Construction of the Automaton

Let $\delta_j(n,q)$ be the j-th digit in the q-adic representation of n. In this section we construct a transducer $\mathcal{A}$ (cf. [9, 20]), which produces the sequence of differences of digits $\delta_j(n,-q)-\delta_j(-n,-q)$ from the sequence of digits of n. For the use of automata in the study of digital expansions we refer to [1, 13]

We start with a description of the maps $(n)_q \mapsto (n)_{-q}$ and $(n)_q \mapsto (-n)_{-q}$ by means of transducers. We use the following notation: let P and Q be two states of a transducer. Then $(n\delta_1)^P = (n)^Q\delta_2$ indicates that there is an edge from P to Q, marked with $\delta_1|\delta_2$. This means, if P is the actual state and the automaton reads δ_1 as the next digit, it moves to state Q writing the digit δ_2.



The automaton $\mathcal{A}^+$ The automaton $\mathcal{A}^-$

FIGURE 1

$\mathcal{A}^+$ is the transducer, corresponding to the map $(n)_q \mapsto (n)_{-q}$. It has four states A ,B, C and D (A being the initial state). It is characterized by the following rules:

$$\begin{array}{llll}
(n\delta)^A = (n)^B\delta & (0\le\delta\le q-1), & (n0)^B = (n)^A 0, & \\
(n\delta)^B = (n)^C\overline{q-\delta} & (1\le\delta\le q-1), & (n\delta)^C = (n)^B\overline{\delta+1} & (0\le\delta\le q-2), \\
(n\overline{q-1})^C = (n)^D 0, & & (n\delta)^D = (n)^C\overline{q-1-\delta}, & (0\le\delta\le q-1)
\end{array}$$

($\overline{a+b}$ indicates that $a+b$ is only one digit). Note that the states correspond to the possible carries that can occur. Since the sign of $(-q)^k$ changes, the automaton has to distinguish between even and odd indices of digits: if the j-th digit has to be processed, then the automaton rests at A or C for odd j and at B or D for even j.

The automaton $\mathcal{A}^-$, corresponding to the map $(n)_q \mapsto (-n)_{-q}$ also consists of four states E, F, G, H (E being the initial state). It is defined by

$$\begin{array}{lll} (n0)^E = (n)^F 0, & (n\delta)^F = (n)^E \delta & (0 \le \delta \le q-1), \\ (n\delta)^E = (n)^H \overline{q-\delta} \quad (1 \le \delta \le q-1), & (n\delta)^H = (n)^E \overline{\delta+1} & (0 \le \delta \le q-2), \\ (n\overline{q-1})^H = (n)^G 0, & (n\delta)^G = (n)^H \overline{q-1-\delta}, & (0 \le \delta \le q-1). \end{array}$$

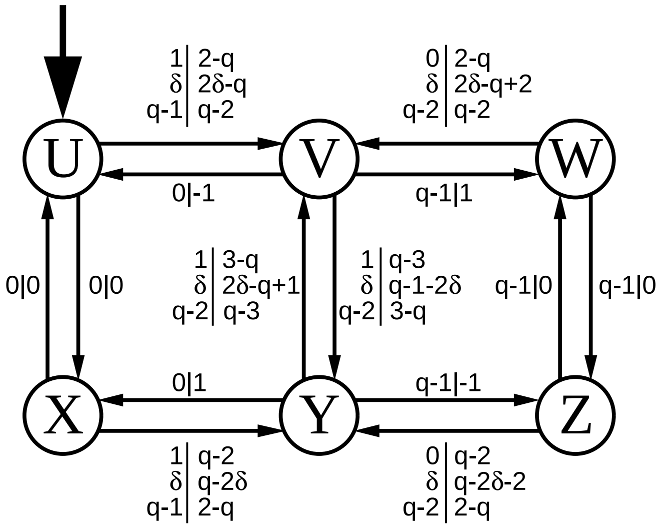


FIGURE 2. The automaton $\mathcal{A}$ producing $\nu_{-q}(n) - \nu_{-q}(-n)$

The transducer $\mathcal{A}$ is defined as the product automaton of $\mathcal{A}^+$ and $\mathcal{A}^-$ defined by Figure 1, whose outputs are the differences of the outputs of $\mathcal{A}^+$ and $\mathcal{A}^-$. The states of $\mathcal{A}$ are defined by

$$U = (A, E), \quad V = (B, H), \quad W = (C, G), \quad X = (B, F), \quad Y = (C, E), \quad Z = (D, H)$$

and $\mathcal{A}$ is depicted in Figure 2 (the remaining two states are never reached).

Remark 1. *It can be immediately read off from this automaton that for odd q the difference $\nu_{-q}(n) - \nu_{-q}(-n)$ is always even. This is a consequence of the fact that the sum of the outputs along any closed path from U to U is even.*

In order to get the asymptotic of the sum (1.9) we need certain accompanying matrices of the automaton $\mathcal{A}$. For $0 \le k \le q-1$ define the 6×6 matrices $M_k = (m^{(k)}_{j_1,j_2})$ in the following way (the states are numbered in their alphabetical order). If there is an edge from j_1 to j_2 in $\mathcal{A}$, marked with $k|\ell$, then $m^{(k)}_{j_1,j_2} = e^{it\ell}$. Otherwise $m^{(k)}_{j_1,j_2} = 0$. Furthermore we define, using the abbreviation $z = e^{it}$,

$$M(t) := \sum_{k=0}^{q-1} M_k = \begin{pmatrix} 0 & s_1 & 0 & 1 & 0 & 0 \\ z^{-1} & 0 & z & 0 & s_2 & 0 \\ 0 & s_1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & s_1 & 0 \\ 0 & s_2 & 0 & z & 0 & z^{-1} \\ 0 & 0 & 1 & 0 & s_1 & 0 \end{pmatrix}$$

with

$$s_1 = z^{2-q} + z^{4-q} + \cdots + z^{q-4} + z^{q-2} = \frac{\sin(q-1)t}{\sin t}$$

and

$$s_2 = z^{3-q} + z^{5-q} + \cdots + z^{q-5} + z^{q-3} = \frac{\sin(q-2)t}{\sin t}.$$

The eigenvalues of M are given by

$$\begin{aligned}
\lambda_1 &= \frac{(1+s_2) + \sqrt{(1-s_2)^2 + 8s_1\cos t}}{2}, \\
\lambda_2 &= \frac{-(1+s_2) - \sqrt{(1-s_2)^2 + 8s_1\cos t}}{2}, \\
\lambda_3 &= \frac{-(1+s_2) + \sqrt{(1-s_2)^2 + 8s_1\cos t}}{2}, \\
\lambda_4 &= 1, \\
\lambda_5 &= \frac{(1+s_2) - \sqrt{(1-s_2)^2 + 8s_1\cos t}}{2}, \\
\lambda_6 &= -1.
\end{aligned}$$

For $t = 0$, $\lambda_3 = \lambda_4 = 1$ and $\lambda_5 = \lambda_6 = -1$, and the limit of the transformation matrices $S(t)$, formed by the eigenvectors of $M(t)$, is a singular matrix for $t \to 0$. Since we are especially interested in values of t close to 0, we replace $S(t)$ by the slightly different matrix

$$T(t) = \begin{pmatrix}
\frac{1-s_2+\sqrt{a_1}}{4\cos t} & \frac{1-s_2+\sqrt{a_1}}{4\cos t} & \frac{1-s_2-\sqrt{a_1}}{4\cos t} & \frac{1}{2}a_2 & \frac{1-s_2-\sqrt{a_1}}{4\cos t} & \frac{1}{2}a_2 \\
1 & -1 & -1 & -\frac{1}{z+1} & 1 & \frac{1}{z+1} \\
\frac{1-s_2+\sqrt{a_1}}{4\cos t} & \frac{1-s_2+\sqrt{a_1}}{4\cos t} & \frac{1-s_2-\sqrt{a_1}}{4\cos t} & \frac{1}{2}a_3 & \frac{1-s_2-\sqrt{a_1}}{4\cos t} & \frac{1}{2}a_3 \\
\frac{1-s_2+\sqrt{a_1}}{4\cos t} & -\frac{1-s_2+\sqrt{a_1}}{4\cos t} & -\frac{1-s_2-\sqrt{a_1}}{4\cos t} & \frac{1}{2}a_2 & \frac{1-s_2-\sqrt{a_1}}{4\cos t} & -\frac{1}{2}a_2 \\
1 & 1 & 1 & -\frac{1}{z+1} & 1 & -\frac{1}{z+1} \\
\frac{1-s_2+\sqrt{a_1}}{4\cos t} & -\frac{1-s_2+\sqrt{a_1}}{4\cos t} & -\frac{1-s_2+\sqrt{a_1}}{4\cos t} & \frac{1}{2}a_3 & \frac{1-s_2-\sqrt{a_1}}{4\cos t} & -\frac{1}{2}a_3
\end{pmatrix}$$

with

$$\begin{aligned} a_1 &= (1-s_2)^2 + 8s_1 \cos t, \\ a_2 &= \tfrac{4s_1 \cos t - 4z^{-1}(1+s_2)\cos t + 2i(1-s_2-\sqrt{a_1})\sin t}{8 \sin^2 t \cos t}, \\ a_3 &= \tfrac{4s_1 \cos t - 4z(1+s_2)\cos t - 2i(1-s_2-\sqrt{a_1})\sin t}{8 \sin^2 t \cos t}. \end{aligned}$$

These matrices satisfy $\lim_{t\to 0} T(t) = T(0)$ with $T(0)$ regular, and furthermore, $D = (d_{j_1,j_2}) := T(t)^{-1}M(t)T(t)$ is a diagonal matrix apart from the two nonzero entries $d_{3,4}$ and $d_{5,6}$. Note that for the entries on the diagonal we have $d_{j,j} = \lambda_j$ $(1 \le j \le 6)$. It will be of importance later that none of the nonzero entries outside the diagonal is located in the first or second column or row. In fact, the first and the second eigenvalue are the dominant ones.

3. Asymptotic of an Exponential Sum

In this section we derive an asymptotic expansion of the sum (1.9). To this matter we use the matrices $M_k(t)$, $M(t)$, and $T(t)$ of the previous section. The method of using matrices in order to get asymptotic expansions of digital sums is classical. It was for example applied in [16] in order to compute asymptotic for certain rarefied sums of the Thue–Morse Sequence. Similar ideas are also used in [23].

Proposition 1. *There exists a $\delta(q) > 0$ such that*

$$\sum_{n<N} e^{i(t+Qa\pi)(\nu_{-q}(n)-\nu_{-q}(-n))} = N^{\alpha_q(t)} \left(F_t^{(1)}\left(\log_q N\right) + itF_t^{(2)}\left(\log_{q^2} N\right)\right) + \mathcal{O}\left(\log N\right), \tag{3.1}$$

for $a \in \mathbb{Z}$ and $|t| \le \delta(q)$, where

$$\alpha_q(t) = \log_q \lambda_1(t) = 1 - \frac{r_1(q)}{\log q} t^2 - \frac{r_2(q)}{\log q} t^4 + \mathcal{O}\left(t^6\right)$$

with

$$r_1(q) := \frac{1}{6} \frac{(q-1)(q^2-4q+7)}{(q+1)}$$

and

$$r_2(q) := \frac{1}{180} \frac{(q-1)(q^6-11q^5+82q^4-322q^3+577q^2-371q+76)}{(q+1)^3};$$

furthermore, $Q = 2$, if q is even, and $Q = 1$ if q is odd. The functions $F_t^{(1,2)}(x)$ are continuous and periodic with period 1 (with respect to x) and are even and C^∞ with respect to t; furthermore, the following expansion around $t = 0$ holds:

$$F_t^{(1)}(\log_q N) + itF_t^{(2)}(\log_{q^2} N) = 1 + itG(\log_{q^2} N) - \frac{t^2}{2} H(\log_q N) + \mathcal{O}\left(t^3\right). \tag{3.2}$$

The functions G and H are the same as those occurring in the statement of Theorem 3 in (1.3) *and* (1.4). *Finally,*

$$\sum_{n<N} e^{i(t+Qa\pi)(\nu_{-q}(n)-\nu_{-q}(-n))} = \mathcal{O}\left(N^{\alpha}\right)$$

for $\alpha = \alpha_q(\delta(q)) < 1$ *for* $\delta(q) < |t| < \frac{Q\pi}{2}$.

Proof. Since by Remark 1 everything is periodic with period $Q\pi$, we restrict ourselves to $a = 0$. A simple discussion of the function $|\lambda_1(t)|$ shows that $|\lambda_1(t)|$ attains its maximum q precisely for $t = 0$ in the interval $[-\frac{Q\pi}{2}, \frac{Q\pi}{2}]$. Thus there exists a $\delta(q)$ such that $|\lambda_1(t)| \leq |\lambda_1(\delta(q))|$ for $\delta(q) \leq |t| \leq \frac{Q\pi}{2}$.

First we want to represent the sum in terms of the matrices $T := T(t)$, $M_j := M_j(t)$ and $M := M(t)$. Let $n = \sum_{j=1}^{J} \eta_j q^j$ be the q-adic representation of n. Then

$$e^{it(\nu_{-q}(n)-\nu_{-q}(-n))} = \mathbf{v}_1^T \left(\prod_{j=0}^{J} M_{\eta_j}\right) M_0^3 \mathbf{v}_2 \tag{3.3}$$

for $\mathbf{v}_1^T := (1,0,0,0,0,0)$ and $\mathbf{v}_2^T := (1,0,0,1,0,0)$. Multiplication by M_0^3 adds three leading zeros to the input, which takes care of the fact that the number of digits of $(n)_{-q}$ and $(-n)_{-q}$ can differ by at most three from the number of digits of $(n)_q$. This also forces the automaton to either stop at state U or state X. Since we start from state U, the lefthand side of (3.3) is equal to the sum of the first and fourth entry of the first line in the matrix product. These entries are extracted by multiplication by the vectors $\mathbf{v}_1^T$ and $\mathbf{v}_2$.

In order to get a representation for the sum we define

$$\tilde{f}(n) := \left(\prod_{j=0}^{J} M_{\eta_j}\right) \quad \text{and} \quad f(n) := \left(\prod_{j=0}^{J} M_{\eta_j}\right) M_0^3.$$

Furthermore, let

$$N = \sum_{k=0}^{K} \delta_k q^k \tag{3.4}$$

be the q-adic representation of the integer N and define

$$N_\ell := \sum_{k=\ell}^{K} \delta_k q^k \quad (0 \leq \ell \leq K).$$

Since $f(N_{\ell+1}+n)=\tilde f(n)f(N_{\ell+1})$ for $n<q^{\ell+1}$ we have

$$\begin{aligned}\sum_{n<N} f(n) &= \sum_{\ell=0}^{K}\sum_{N_{\ell+1}\le n<N_\ell} f(n)\\ &= \sum_{\ell=0}^{K}\sum_{n<\delta_\ell q^\ell}\tilde f(n)f(N_{\ell+1})\\ &= \sum_{\ell=0}^{K} M^\ell \sum_{j=0}^{\delta_\ell-1} M_j\left(\prod_{k=\ell+1}^{K} M_{\delta_k}\right)M_0^3.\end{aligned}$$

Here we set $\sum_{j=0}^{-1}M_j=0$. Write $M^\ell=TD^\ell T^{-1}$ with D as before. Since D is “almost” a diagonal matrix, having only two nonzero entries $d_{3,4}$ and $d_{5,6}$ apart from the diagonal entries, we can split it in the following way. Write $D:=D_1+D_2+D_3$, where $D_1:=\operatorname{diag}(\lambda_1,0,0,0,0,0)$, $D_2:=\operatorname{diag}(0,\lambda_2,0,0,0,0)$ and $D_3:=D-D_1-D_2$ is a matrix having only zero entries in the first and second column and row. With that we have

$$\sum_{n<N} f(n)=\sum_{p=1}^{3} T\sum_{\ell=0}^{2K}\tilde D_p^\ell T^{-1}\sum_{j=0}^{\delta_\ell-1}M_j\left(\prod_{k=\ell+1}^{2K}M_{\delta_k}\right)M_0^3=P_1+P_2+P_3. \tag{3.5}$$

First we deal with the sum P_1, corresponding to the matrix D_1. In order to extract the main term, we define for $x=\sum_{\ell=0}^{\infty}\frac{\delta_\ell}{q^\ell}$

$$F_{1,t}(\log_q x)=E_1^{\log_q x}\sum_{\ell=0}^{\infty}E_1^\ell T^{-1}\sum_{j=0}^{\delta_\ell-1}M_j\prod_{k=0}^{\ell-1}M_{\delta_{\ell-k-1}}M_0^3 \tag{3.6}$$

with $E_1:=\operatorname{diag}(\lambda_1^{-1},0,0,0,0,0)$, where $F_{1,t}$ is a continuous periodic function of period 1; continuity is proved along the same lines as in [17]. With this definition we get

$$P_1:=TD_1^{\log_q N}F_{1,t}\left(\log_q N-K\right).$$

Further properties of $F_{1,t}$ will be collected in Lemma 1 below.

In order to get an expansion for the sum P_2 we have to take into account the fact that the powers of λ_2 have alternating signs for $t=0$. This causes a doubling of period lengths in the occurring periodic functions and yields mean 0 for the fluctuation (cf. [17]). To derive these fluctuations we write the expansion of N in the form

$$N=\sum_{k=0}^{2L}\delta_k q^k$$

with $L:=\left[\frac{K-1}{2}\right]+1$ (i.e., the representation has a leading zero, if K is odd, and coincides with the former representation of N in (3.4) if K is even). Now we have

$$P_2=\sum_{n<N}f(n)=T\sum_{\ell=0}^{2L}D_2^\ell T^{-1}\sum_{j=0}^{\delta_\ell-1}M_j\left(\prod_{k=\ell+1}^{2L}M_{\delta_k}\right)M_0^3$$

Again we define a function $F_{2,t}$ in order to extract the main term. Let for $x = \sum_{\ell=0}^{\infty} \frac{\delta_\ell}{q^\ell}$

$$F_{2,t}\left(\log_q x\right) = \left(E_2^2\right)^{\log_q x} \sum_{\ell=0}^{\infty} E_2^\ell T^{-1} \sum_{j=0}^{\delta_\ell - 1} M_j \prod_{k=0}^{\ell-1} M_{\delta_{\ell-k-1}} M_0^3 \tag{3.7}$$

with $E_2 := \operatorname{diag}(0, \lambda_2^{-1}, 0, 0, 0, 0)$. This definition gives

$$P_2 = T D_2^{2L} F_{2,t}\left(\log_{q^2} N - L\right). \tag{3.8}$$

Further properties of $F_{2,t}$ will be collected in Lemma 1.

It remains to deal with the sum

$$T \sum_{\ell=0}^{2K} D_3^\ell T^{-1} \sum_{j=0}^{\delta_\ell - 1} M_j \left(\prod_{k=\ell+1}^{2K} M_{\delta_k} \right) M_0^3.$$

This sum does not contribute to the leading terms of the expansion. In fact, we can estimate it in the following way. Since the matrix norms of the matrices M_j $(j = 1, .., 6)$ are less than or equal to 1 the norm of the product is $\mathcal{O}(1)$. Moreover, all the entries of the matrices T are independent of N and periodic in t. Thus their norms can also be bounded by an absolute constant. What remains is the matrix D_3^ℓ. Since all its eigenvalues are ≤ 1 for $|t| < \delta(q)$ and ℓ is less than or equal to $\log_q N$, we conclude that $\|D_3^\ell\| = \mathcal{O}(1)$ for $|t| < \delta(q)$. Taking into account the sum over ℓ we arrive at

$$\|P_3\| = \mathcal{O}(\log N)$$

for $|t| < \delta(q)$.

Lemma 1. *$F_{1,t}(\log_q x)\mathbf{v_2}$ is an even, $F_{2,t}(\log_{q^2} x)\mathbf{v_2}$ an odd function of t.*

Proof. We start with the function $F_{1,t}(\log_q x)$. Let

$$A := \begin{pmatrix} e^{ita} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & e^{-itc} & 0 & 0 \\ e^{itb} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & e^{-ita} & 0 & 0 \\ e^{itc} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & e^{-itb} & 0 & 0 \end{pmatrix}$$

and denote by $I_{1,4}$ the permutation matrix that exchanges the first and fourth columns of a matrix, if it is multiplied on the righthand side. With this notation the product occurring in $F_{1,t}$ is of the shape

$$M_j \prod_{k=0}^{\ell-1} M_{\delta_{\ell-k-1}} M_0^3 = A \quad \text{or} \quad M_j \prod_{k=0}^{\ell-1} M_{\delta_{\ell-k-1}} M_0^3 = A I_{1,4}$$

for a suitable choice of a, b, c. This can easily be proved by induction on the number of factors and is reflected by the shape of the automaton $\mathcal{A}$. We only deal with the first case, the second one being totally similar. Next we examine the matrix $B := E_1^{\log_q x} E_1^\ell T^{-1}$. Since each factor of B contains only even functions as entries, the same holds for B. Denote the entries of B by $(b_{\mu,\rho})$. Direct calculation yields that $b_{\mu,\rho} = b_{\mu,\rho+3}$ for $1 \leq \mu \leq 6$ and $1 \leq \rho \leq 3$. If we are able to show that the product $BA\mathbf{v_2}$ is an even function, we are done, since $F_{1,t}(\log_q x)\mathbf{v_2}$ is a sum of vectors of this structure and the sum of even functions is again an even function.

$$\begin{aligned} BA\mathbf{v_2} &= (b_{\mu 1}(t)e^{iat} + b_{\mu 4}(t)e^{-iat} + b(t)_{\mu 3}e^{ibt} + b_{\mu 6}(t)e^{-ibt} + b_{\mu 5}(t)e^{ibt} + b_{\mu 2}(t)e^{-ibt})_{\mu=1,\dots,6} \\ &= (2b_{\mu 1}(t)\cos(at) + 2b_{\mu 3}(t)\cos(bt) + 2b_{\mu 5}(t)\cos(ct))_{\mu=1,\dots,6}. \end{aligned}$$

Since $b_{\mu\rho}$ $(1 \leq \mu, \rho \leq 6)$ are even functions, the first assertion is proved. The second assertion can be proved in the same way. The only difference is that, instead of B, one uses the matrix $C := E_2^{\log_q x} E_2^\ell T^{-1}$. C again consists only of even functions. Furthermore, for the entries $(c_{\mu\rho})$ of C, one can show that $c_{\mu\rho} = -c_{\mu,\rho+3}$, where $1 \leq \mu \leq 6$ and $1 \leq \rho \leq 3$. Arguing in the same way as before yields

$$CA\mathbf{v_2} = (2ib_{\mu 1}(t)\sin(at) + 2ib_{\mu 3}(t)\sin(bt) + 2ib_{\mu 5}(t)\sin(ct))_{\mu=1,\dots,6}.$$

Since $F_{2,t}(\log_q x)\mathbf{v_2}$ is a sum of vectors of this shape, we have shown that it is an odd function. $\square$

Now we are in a position to describe the asymptotic behaviour of the sum in the statement of the proposition. Summing up the amounts contributed by the three sums considered above and multiplying by $\mathbf{v}_1^T$ on the left and by $\mathbf{v}_2$ on the right yields

$$\sum_{n<N} e^{it(\nu_{-q}(n)-\nu_{-q}(-n))} = N^{\alpha_q(t)}\left(F_t^{(1)}\left(\log_q N\right) + itF_t^{(2)}\left(\log_{q^2} N\right)\right) + \mathcal{O}\left(\log N\right),$$

for $|t| < \delta(q)$. Since the sums in (3.6) and (3.7) are exponentially convergent for $|t| \leq \delta(q)$, the functions $F_{1,t}$ and $F_{2,t}$ are C^∞ with respect to t. The same holds for $F_t^{(1,2)}$. $\square$

4. Proof of the Theorems

Proof of Theorem 1. We derive the expansions (1.3) and (1.4) with help of Proposition 1. First we differentiate the expansion in Proposition 1 once and twice, respectively. Setting $t = 0$ and ignoring the error term $\mathcal{O}(\log N)$ for the moment, yields (1.3) and (1.4) apart from the error terms. It remains to determine the contribution of the error term $\mathcal{O}(\log N)$ after the differentiation. Equation (3.5) shows that this error term comes from the matrix P_3. From the shape of the matrices M_j it is clear that the product of the matrices M_j contained in P_3 is a matrix containing sums of at most $\log_q N$ terms of the form $\exp(it\ell)$ with $|\ell| \leq (q-1)\log_q N$ (cf. the proof of Lemma 1). Each differentiation of such an expression yields one extra $\log N$. Thus we arrive at the error terms indicated in (1.3) and (1.4).

So we are left with the sum over the absolute values. In order to derive its asymptotic expansion we use Fourier transform techniques. First define

$$g(x) := \begin{cases} 1 - |x| & \text{for } |x| < 1 \\ 0 & \text{for } |x| \geq 1. \end{cases}$$

It is easily seen that the Fourier transform of $g(x)$ is

$$\int_{-\infty}^{\infty} g(x)e^{itx}\,dx = \frac{2-2\cos t}{t^2}. \tag{4.1}$$

Now we apply the inverse transform. Since $|\nu_{-q}(n)-\nu_{-q}(-n)| \le q\log_q N$ for $n<N$, substituting $x=\frac{|\nu_{-q}(n)-\nu_{-q}(-n)|}{q\log_q N}$ and summing up yields

$$\begin{aligned} L(N) &:= \sum_{n<N}\left(1-\left|\frac{|\nu_{-q}(n)-\nu_{-q}(-n)|}{q\log_q N}\right|\right) && (4.2)\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{2-2\cos t}{t^2}\sum_{n<N}\exp\left(\frac{it(\nu_{-q}(n)-\nu_{-q}(-n))}{q\log_q N}\right)dt. && (4.3)\end{aligned}$$

Now we want to apply the asymptotic expansions of Proposition 1 to the sum in the integrand. Since these expansions can be used only in the neighbourhood of certain points, we have to split the line of integration. To this matter we define the sets

$$I_a = [Q\pi aq\log_q N-(\log_q N)^{11/16}, Q\pi aq\log_q N+(\log_q N)^{11/16}] \quad (a\in\mathbb{Z}), \quad R=\mathbb{R}\setminus\bigcup_{a\in\mathbb{Z}} I_a.$$

Let

$$J_a = \frac{1}{2\pi}\int_{I_a}\frac{2-2\cos t}{t^2}\sum_{n<N}\exp\left(\frac{it(\nu_{-q}(n)-\nu_{-q}(-n))}{q\log_q N}\right)dt$$

and let E be the corresponding integral over the range R, then (4.2) reads

$$L(N) = \sum_{a=-\infty}^{\infty} J_a + E.$$

Now we derive the asymptotic of the integrals J_a and E. It will turn out that the main term comes from J_0, and for $a\neq 0$ the corresponding integrals only affect terms of smaller order. E turns out to be of smaller order. We start with the integral J_0. Applying Proposition 1 and expanding $\exp\left(-r_2(q)\frac{t^4}{q^4\log_q^3 N}\right)$ we arrive at

$$\begin{aligned} J_0 &= \frac{N}{2\pi}\int_{I_0}\frac{2-2\cos t}{t^2}\exp\left(-r_1(q)\frac{t^2}{q^2\log_q N}\right)\\ &\quad\times\left(1-r_2(q)\frac{t^4}{q^4(\log_q N)^3}+\mathcal{O}\left(\frac{t^6}{(\log N)^5}\right)\right)\\ &\quad\times\left(1+\frac{it}{q\log_q N}G(\log_{q^2} N)-\frac{t^2}{2q^2(\log_q N)^2}H(\log_q N)+\mathcal{O}\left(\frac{t^3}{(\log N)^3}\right)\right)dt.\end{aligned}$$

In order to evaluate the integral, we split it in the following way.

$$\begin{aligned}
J_0 &= \frac{N}{2\pi}\int_{I_0}\frac{2-2\cos t}{t^2}\exp\left(-r_1(q)\frac{t^2}{q^2\log_q N}\right)dt\\
&\quad +i\frac{N}{2\pi}\frac{G(\log_{q^2}N)}{q\log_q N}\int_{I_0}\frac{2-2\cos t}{t}\exp\left(-r_1(q)\frac{t^2}{q^2\log_q N}\right)dt\\
&\quad -\frac{N}{4\pi}\frac{H(\log_q N)}{q^2(\log_q N)^2}\int_{I_0}(2-2\cos t)\exp\left(-r_1(q)\frac{t^2}{q^2\log_q N}\right)dt\\
&\quad -\frac{N}{2\pi}\frac{r_2(q)}{q^5(\log_q N)^4}\int_{I_0}(2-2\cos t)t^2\exp\left(-r_1(q)\frac{t^2}{q^2\log_q N}\right)dt\\
&\quad -i\frac{N}{2\pi}\frac{r_2(q)G(\log_{q^2}N)}{q^4(\log_q N)^3}\int_{I_0}(2-2\cos t)t^3\exp\left(-r_1(q)\frac{t^2}{q^2\log_q N}\right)dt\\
&\quad +\mathcal{O}\left(\frac{1}{(\log_q N)^5}\frac{N}{2\pi}\int_{I_0}(2-2\cos t)t^4\exp\left(-r_1(q)\frac{t^2}{q^2\log_q N}\right)dt\right)\\
&= J_{0,1}+J_{0,2}+J_{0,3}+J_{0,4}+J_{0,5}+J_{0,6}.
\end{aligned}$$

We can extend the range of integration of $J_{0,j}$ $(1\le j\le 6)$ to infinity making an error of order $\mathcal{O}\left(N(\log N)^{-d}\right)$ for d arbitrary. Evaluating the resulting integrals yields

$$\begin{aligned}
J_{0,1} &= N-2\sqrt{\frac{r_1(q)}{\pi}}\frac{N}{q(\log_q N)^{1/2}}+\mathcal{O}\left(N^{1-\frac{\log q}{4r_1(q)}}\right),\\
J_{0,2} &= 0,\\
J_{0,3} &= -H(\log_q N)\frac{1}{2\sqrt{\pi r_1(q)}}\frac{N}{q(\log_q N)^{3/2}}+\mathcal{O}\left(N^{1-\frac{\log q}{4r_1(q)}}\right),\\
J_{0,4} &= -\frac{r_2(q)}{2\sqrt{\pi r_1(q)^3}}\frac{N}{q(\log_q N)^{3/2}}+\mathcal{O}\left(N^{1-\frac{\log q}{4r_2(q)}}\right),\\
J_{0,5} &= 0,\\
J_{0,6} &= \mathcal{O}\left(\frac{N}{(\log N)^{5/2}}\right).
\end{aligned}$$

Next we deal with the integrals J_a for $a\neq 0$. Applying Proposition 1 and substituting $u=(t-Q\pi aq\log_q N)/\sqrt{\log_q N}$ yields

$$\begin{aligned}
J_a &= \frac{N}{2\pi}\int_{-(\log_q N)^{3/16}}^{(\log_q N)^{3/16}}\frac{2-2\cos(Q\pi aq\log_q N+\sqrt{\log_q N}u)}{(Q\pi aq\log_q N+\sqrt{\log_q N}u)^2}\\
&\quad\times\exp\left(-\frac{r_1u^2}{q^2}-\frac{r_2u^4}{q^4\log_q N}+\mathcal{O}\left(\frac{u^6}{(\log N)^2}\right)\right)\\
&\quad\times\left(1+i\frac{G(\log_{q^2}N)u}{q\sqrt{\log_q N}}-\frac{H(\log_q N)u^2}{2q^2\log_q N}+\mathcal{O}\left(\frac{u^3}{(\log N)^{3/2}}\right)\right)\sqrt{\log_q N}du.
\end{aligned}$$

Now we expand $(Q\pi aq\log_q N+\sqrt{\log_q N}u)^{-2}$ and $\exp(-r_2u^4/q^4\log_q N)$. Furthermore, we integrate over the error terms keeping in mind that $\int_{-\alpha}^{\alpha}u^pe^{-u^2}du$ is bounded uniformly in α. This yields

$$\begin{aligned}J_a &= \frac{N}{2\pi}\frac{2\sqrt{\log_q N}}{(Q\pi aq\log_q N)^2}\int\limits_{-(\log_q N)^{3/16}}^{(\log_q N)^{3/16}}\left(1-\cos\left(Q\pi aq\log_q N+\sqrt{\log_q N}u\right)\right)\exp\left(-\frac{r_1u^2}{q^2}\right)\\ &\times\left(1-\frac{2u}{Q\pi aq\sqrt{\log_q N}}+i\frac{G(\log_{q^2}N)u}{q\sqrt{\log_q N}}\right)du+\mathcal{O}\left((\log N)^{-5/2}\right).\end{aligned}$$

If we extend the range of integration to the whole real line, the error is smaller than $\mathcal{O}\left((\log N)^{-3}\right)$ because of the factor $\exp(-r_1u^2q^{-2})$ in the integrand. It is easy to see that $(\beta>0)$

$$\int_{-\infty}^{\infty}e^{-\beta u^2}du=\sqrt{\frac{\pi}{\beta}},\quad \int_{-\infty}^{\infty}ue^{-\beta u^2}du=0$$

and

$$\int_{-\infty}^{\infty}\cos\left(\sqrt{\log_q N}u+Q\pi aq\log_q N\right)e^{-\beta u^2}du=\mathcal{O}\left(N^{-1/4}\right).$$

Observing that J_a is a sum of integrals of these types enables us to find expressions for the integral J_a. Summing up over all $a\neq 0$ finally yields

$$\sum_{a\neq 0}J_a=\frac{1}{3Q^2q\sqrt{\pi r_1(q)}}(\log_q N)^{-3/2}+\mathcal{O}\left((\log N)^{-5/2}\right).$$

It remains to estimate the integral E. Here the range of integration is extended over intervals $E_a=[Qaq\pi\log_q N+(\log_q N)^{11/16},Q(a+1)q\pi\log_q N-(\log_q N)^{11/16}]$, where the exponential sum in (4.3) is bounded by $\mathcal{O}\left(N^{\alpha_q((\log_q N)^{-5/16})}\right)$. Thus we have

$$\begin{aligned}&\left|\frac{1}{2\pi}\int_{I_a}\frac{2-2\cos t}{t^2}\sum_{n<N}\exp\left(\frac{it(\nu_{-q}(n)-\nu_{-q}(-n))}{q\log_q N}\right)dt\right|\leq\\ &\frac{4\log_q N}{\min(|a|,|a+1|)^2Q^2\pi^2(\log_q N)^2}N^{\alpha_q((\log_q N)^{-5/16})},\end{aligned}$$

except for $a=0,-1$, where the denominator has to be replaced by $(\log_q N)^{11/8}$. Since $\alpha_q(t)\leq 1-\frac{r_1(q)}{\log q}t^2$ this yields an error term $\mathcal{O}\left(N(\log N)^{-d}\right)$ for any $d>0$.

Summing up what we proved until now, we get

$$L(N)=N-2\sqrt{\frac{r_1}{\pi}}\frac{N}{q(\log_q N)^{1/2}}-\Psi(\log_q N)\frac{N}{q(\log_q N)^{3/2}}+\mathcal{O}\left(\frac{N}{(\log N)^{5/2}}\right)$$

with

$$\Psi(x)=\frac{1}{2\sqrt{\pi r_1(q)}}H(x)+\frac{r_2(q)}{2\sqrt{\pi r_1(q)^3}}-\frac{1}{3Q^2\sqrt{\pi r_1(q)}}.$$

and the theorem is proved. ☐

Remark 2. *The proof of Theorem 1 shows that one could obtain a full asymptotic expansion for the sums involving terms of the form* $\frac{N}{(\log N)^{k+\frac{1}{2}}}\Psi_k(\log_{q^2} N)$ *for* $k \geq 1$ *with periodic functions* Ψ_k*.*

Proof of Corollary 1. We use the representations

$$\sum_{n<N} \min(\nu_{-q}(n), \nu_{-q}(-n)) = \frac{1}{2}\left(\sum_{|n|<N} \nu_{-q}(n) - \sum_{n<N} |\nu_{-q}(n) - \nu_{-q}(-n)|\right), \tag{4.4}$$

$$\sum_{n<N} \max(\nu_{-q}(n), \nu_{-q}(-n)) = \frac{1}{2}\left(\sum_{|n|<N} \nu_{-q}(n) + \sum_{n<N} |\nu_{-q}(n) - \nu_{-q}(-n)|\right). \tag{4.5}$$

and the known fact (cf. [12, 25])

$$\sum_{|n|<N} \nu_{-q}(n) = (q-1)N\log_q N + N\Phi(\log_{q^2} N). \tag{4.6}$$

Together with Theorem 1, this gives the desired result. ☐

Proof of Theorem 2. The proof follows the same lines as the proof of Théorème 3 in [24], but has to use some additional ideas, because in the present situation the error term is weaker.

For a fixed $x \in [0,1)$ we write the q-adic expansion of q^x as

$$q^x = \sum_{j=0}^{\infty} \frac{\delta_j}{q^j}.$$

Let $L(k) = \mathcal{O}(\log k)$ be an integer-valued function to be fixed later. With the help of this function we define x_k and N_k by

$$N_k = q^{x_k+k+L(k)} = q^{L(k)} \sum_{j=0}^{k} \delta_j q^{k-j}.$$

For a function $g(N)$ with $g(N) = o(N)$ and $\log^3 N = o(g(N))$ we define y_k by

$$N_k + g(N_k) = q^{y_k+k+L(k)}.$$

From these definitions and the requirements for $g(N)$ and $L(k)$ we easily see that $x_k, y_k \to x$ and

$$y_k - x_k = \frac{g(N_k)}{N_k \log q} + \mathcal{O}\left(\frac{g(N_k)^2}{N_k^2}\right).$$

We first prove the non-differentiability of $G(x)$. For this purpose we use the fact that the arithmetic function $h(n) = \nu_{-q}(n) - \nu_{-q}(-n)$ has the property $h(q^{2k+2}a + b) = h(a) + h(b)$

for $b < q^{2k}$, which can be seen from the automaton in Figure 2. In the following we assume k and $L(k)$ to be even. This yields

$$\sum_{N_k \le n < N_k + g(N_k)} h(n) = g(N_k)h(N_k) + \sum_{n < g(N_k)} h(n). \tag{4.7}$$

On the other hand, we derive from (1.3) that

$$\begin{aligned} \sum_{N_k \le n < N_k + g(N_k)} h(n) &= N_k(G(\log_{q^2}(N_k + g(N_k))) - G(\log_{q^2} N_k)) + \\ & g(N_k)G(\log_{q^2}(N_k + g(N_k))) + \mathcal{O}\left((\log N_k)^2\right). \end{aligned} \tag{4.8}$$

Assume now that G is differentiable at x. Then we can write the difference on the righthand side of (4.8) in terms of the derivative; combining this with (4.7) and dividing by $g(N_k)$ we derive

$$\frac{1}{g(N_k)} \sum_{n < g(N_k)} h(n) = h(N_k) + G(\log_{q^2} g(N_k)) = \frac{1}{\log q} G'(x) + G(\log_{q^2}(N_k + g(N_k))) + o(1). \tag{4.9}$$

Letting k tend to infinity in (4.9), the last expression tends to $\frac{1}{\log q} G'(x) + G(x)$. Taking $g(N_k) = k^4$ and $L(k) = 4[\log_q k] + 8$ we observe that the middle expression in (4.9) is then the sum $h(N_k) + G(4 \log_{q^2} k)$. The first summand takes only integer values, the second takes values in a dense subset of the image of G. This argument would break down ,if G would be a constant. Since G is an odd function this would imply that $G(x) \equiv 0$, which is easily seen to be impossible (for instance, it would be a contradiction to our distribution result).

In order to prove the non-differentiability of $H(x)$ we compute using (1.4)

$$\begin{aligned} \sum_{N \le n < N + g(N)} f(n) &= 2r_1(q)g(N)\log_q N + N\left(H(\log_q(N + g(N))) - H(\log_q N)\right) + \\ & 2r_1(q)g(N) + g(N)H(\log_q(N + g(N))) + \mathcal{O}\left((\log N)^3\right) \end{aligned} \tag{4.10}$$

for $f(n) = (\nu_{-q}(n) - \nu_{-q}(-n))^2$.

We insert $N = N_k$ and the above estimates into (4.10) and use the periodicity of $H(t)$ to obtain

$$\begin{aligned} \sum_{N_k \le n < N_k + g(N_k)} f(n) &= 2r_1(q)g(N_k)\log_q N_k + N_k\left(H(y_k) - H(x_k)\right) + \\ & 2r_1(q)g(N_k) + g(N_k)H(y_k) + \mathcal{O}\left((\log N_k)^3\right). \end{aligned}$$

Assume now that H is differentiable at x. Then we could replace $H(y_k) - H(x_k)$ in this last expression by $\frac{g(N_k)}{N_k \log q} H'(x) + o(\frac{g(N_k)}{N_k})$ to obtain

$$\frac{1}{g(N_k)} \sum_{0 \le n < g(N_k)} f(N_k + n) - 2r_1(q)\log_q N_k = \frac{1}{\log q} H'(x) + 2r_1(q) + H(x) + o(1). \tag{4.11}$$

Now we take $g(N_k)=k^4$, suppose that $L(k)$ is even for any k and $4\log_q k+3<L(k)=\mathcal{O}(\log k)$. This assures, that adding $(N_k)_{-q}$ to the representation $(n)_{-q}$ of an integer n in the range $0<n<g(N_k)$ does not produce any carries. Thus the value of $\sum_{0<n<g(N_k)} f(N_k+n)$ is independent of the choice of $L(k)$ (in the given range). On the other hand, the term $2r_1(q)\log_q N_k$ on the left hand side of (4.11) depends linearly on $L(k)$.

If (4.11) were true, the limit of its lefthand side should be independent of the choice of $L(k)$ for $k\to\infty$, since the same is obviously true for its righthand side. This will lead us to a contradiction: consider the choices

$$L(k)=2\left[2\log_q k+3\right] \quad \text{and} \quad L(k)=2\left[3\log_q k\right].$$

By the above arguments the sum in (4.11) remains the same for both choices of $L(k)$. On the other hand, $2r_1(q)\log_q N_k$ depends linearly on $L(k)$. Thus if the limit of the lefthand side of (4.11) exists for the first choice of $L(k)$ it can not exist for the second choice, and vice versa. This yields the desired contradiction. □

Proof of Theorem 3. The proof uses Vaaler's version (cf. [27]) of the Berry-Esseen inequality (cf. [3], [10]). Since this procedure is totally standard, we will only give a sketched proof. We use (3.1) for $a=0$ and replace t by $\frac{t}{\sqrt{2r_1(q)\log_q N}}$. Together with (3.2) this gives

$$\begin{aligned}\hat{g}_N(t) = \quad & \frac{1}{N}\sum_{n<N}\exp\left(\frac{it}{\sqrt{\log_q N}}(\nu_{-q}(n)-\nu_{-q}(-n))\right)= \\ & e^{-\frac{t^2}{2}}\left(1+\mathcal{O}\left(\frac{t}{\sqrt{\log N}}\right)\right)\left(1+\mathcal{O}\left(\frac{t^4}{\log N}\right)\right)\end{aligned} \tag{4.12}$$

for $t=\mathcal{O}\left(\sqrt{\log N}\right)$. The function on the lefthand side is the characteristic function of the distribution function of

$$\frac{\nu_{-q}(n)-\nu_{-q}(-n)}{\sqrt{r_1(q)\log_q N}};$$

we will denote this distribution function by $g_N(x)$. Theorem 13 in [27] implies that

$$\begin{aligned}|g_N(x)-g(x)| \ \leq\ & \int_{-T}^{T}\hat{J}(T^{-1}t)\frac{1}{2\pi t}\left|\hat{g}_N(t)-e^{-\frac{t^2}{2}}\right|dt \\ +\ & \frac{1}{2T}\left(1+\int_{-T}^{T}\hat{K}(T^{-1}t)\left(\hat{g}_N(t)-e^{-\frac{t^2}{2}}\right)dt\right),\end{aligned} \tag{4.13}$$

where $g(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{\xi^2}{2}}d\xi$; the functions $\hat{J}$ and $\hat{K}$ are given by

$$\begin{aligned}\hat{J}(t) &= \begin{cases}\pi t(1-|t|)\cot\pi t+|t| & \text{for } |t|<1\\ 0 & \text{otherwise,}\end{cases}\\ \hat{K}(t) &= \begin{cases}1-|t| & \text{for } |t|\leq 1\\ 0 & \text{otherwise.}\end{cases}\end{aligned}$$

In order to use the full information in the asymptotic of $\hat{g}_N(t)$ in (4.12), we have to split the range of integration in (4.13) at $|t| = (\log N)^{\frac{1}{6}}$ (this is the point where the two $\mathcal{O}$-terms in (4.12) change their rôles). We choose $T = \sqrt{\log N}$ and estimate the integrals trivially to obtain the desired result. $\square$

Acknowledgment. The authors are indebted to the anonymous referee for many valuable remarks.

References

[1] J.-P. Allouche, E. Cateland, W. J. Gilbert, H.-O. Peitgen, J. O. Shallit, and G. Skordev. Automatic maps in exotic numeration systems. *Theory Comput. Syst.*, 30:285–331, 1997.

[2] N. L. Bassily and I. Katai. Distribution of the values of q-additive functions on polynomial sequences. *Acta Math. Hungar.*, 68(3):355–363, 1995.

[3] A. C. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc*, 49:122–136, 1941.

[4] E. Cateland. *Suites digitales et suites k-régulières*. PhD thesis, Université Bordeaux I, 1992.

[5] J. Coquet. Power sums of digital sums. *J. Number Theory*, 22:161–176, 1986.

[6] H. Delange. Sur la fonction sommatoire de la fonction "somme des chiffres". *l'Enseignement Math.(2)*, 21:31–47, 1975.

[7] M. Drmota and J. Gajdosik. The distribution of the sum-of-digits function. *J. Théor. Nombres Bordx.*, 10:17–32, 1998.

[8] J.-M. Dumont and A. Thomas. Digital sum moments and substitutions. *Acta Arith.*, 64:205–225, 1993.

[9] S. Eilenberg. *Automata, Languages and Machines*, volume A. Academic Press, 1974.

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

[11] P. Flajolet, P. J. Grabner, P. Kirschenhofer, H. Prodinger, and R. F. Tichy. Mellin transform and asymptotics: Digital sums. *Theoret. Comput. Sci.*, 123:291–314, 1994.

[12] P. Flajolet and L. Ramshaw. A note on Gray code and odd-even merge. *SIAM J. Comput.*, 9:142–158, 1980.

[13] C. Frougny. Representation of numbers and finite automata. *Math. Systems Theory*, 25:37–60, 1992.

[14] A. O. Gelfond. Sur les nombres qui ont des propriétés additives et multiplicatives données. *Acta Arith.*, 13:259–266, 1968.

[15] B. Gittenberger and J. M. Thuswaldner. The moments of the sum of digits function in number fields. *Canadian Math. Bull.*, 42(1):68–77, 1999.

[16] S. Goldstein, K. A. Kelley, and E. R. Speer. The fractal structure of rarefied sums of the Thue–Morse sequence. *J. Number Theory*, 42:1–19, 1992.

[17] P. J. Grabner. Completely q-multiplicative functions: the Mellin-transform approach. *Acta Arith.*, 65:85–96, 1993.

[18] P. J. Grabner, P. Kirschenhofer, H. Prodinger, and R. F. Tichy. On the moments of the sum-of-digits function. In G. E. Bergum, A. N. Philippou, and A. F. Horadam, editors, *Applications of Fibbonacci Numbers*, volume 5, pages 263–273. Kluwer Academic Publishers, 1993.

[19] D. E. Knuth. *The Art of Computer Programming, Vol 2: Seminumerical Algorithms*. Addison Wesley, London, 1981.

[20] W. Kuich and A. Salomaa. *Semirings, Automata, Languages*. Springer, Berlin, 1986.

[21] J.-L. Mauclaire and L. Murata. On q-additive functions, I. *Proc. Japan Acad.*, 59:274–276, 1983.

[22] J.-L. Mauclaire and L. Murata. On q-additive functions, II. *Proc. Japan Acad.*, 59:441–444, 1983.

[23] T. Okada, T. Sekiguchi, and Y. Shiota. Applications of binomial measures to power sums of digital sums. *J. Number Theory*, 52:256–266, 1995.

[24] G. Tenenbaum. Sur la non-dérivabilité de fonctions périodiques associées à certaines formules sommatoires. In R. L. Graham et al., editor, *The Mathematics of Paul Erdős*, volume 13 of *Algorithms Comb.* Springer, Berlin, 1997.

[25] J. M. Thuswaldner. The sum of digits function in number fields. *Bull. London Math. Soc.*, 30:37–45, 1998.

[26] J. M. Thuswaldner. The sum of digits function in number fields: Distribution in residue classes. *J. Number Theory*, 74:111–125, 1999.

[27] J. D. Vaaler. Some extremal functions in Fourier analysis. *Bull. Amer. Math. Soc.*, 12:183–216, 1985.

Institut für Mathematik A, Technische Universität Graz, Steyrergasse 30, 8010 Graz, Austria

E-mail address: `grabner@weyl.math.tu-graz.ac.at`

Institut für Mathematik und Angewandte Geometrie, Abteilung für Mathematik und Statistik, Montanuniversität Leoben, Franz-Josef-Strasse 18, A–8700 Leoben, Austria

E-mail address: `thuswald@g3809u.unileoben.ac.at`