ACTA ARITHMETICA
104.1 (2002)

The summatory function of the sum-of-digits function on polynomial sequences

by

Manfred Peter (Freiburg)

1. Introduction. Let q be an integer > 1. For $n \in \mathbb{N}_0$, let $n = \sum_{r\ge 0} e_r(n)q^r$, with $e_r(n) \in \{0,\ldots,q-1\}$, be the q-ary representation of n and

$$s_q(n) := \sum_{r\ge 0} e_r(n)$$

the sum of digits of n in base q. The distribution properties of the function s_q have been investigated from many points of view. Delange [2] showed that the summatory function of s_q can be written in the form

$$\sum_{0\le n<N} s_q(n) = \frac{q-1}{2\log q} N\log N + NF\left(\frac{\log N}{\log q}\right), \qquad N \in \mathbb{N}, \tag{1.1}$$

where $F : \mathbb{R} \to \mathbb{R}$ is 1-periodic, continuous and nowhere differentiable. Shiokova [20] proved that

$$\sum_{p\le N} s_q(p) = \frac{q-1}{2\log q} N + O\left(N\left(\frac{\log\log N}{\log N}\right)^{1/2}\right), \qquad N \ge 1, \tag{1.2}$$

where p runs through prime numbers. Heppner [8] generalized this result to arbitrary subsets of $\mathbb{N}$ whose counting function has a certain asymptotics.

Since s_q can be seen as the sum of the "independent random variables" e_r, $r \ge 0$, the central limit theorem gives

$$\lim_{N\to\infty} \frac{1}{2^N} \#\left\{0 \le n < 2^N \;\middle|\; \frac{s_q(n) - N/2}{\sqrt{N}/2} \le x\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\,dt, \qquad x \in \mathbb{R}.$$

2000 *Mathematics Subject Classification*: Primary 11N37; Secondary 11A63.

Key words and phrases: sum of digits, mean value, lattice points.

Under the density hypothesis for the Riemann zeta function, Kátai and Mogyoródi [10] proved an analogous result where n is restricted to primes. Under the same hypothesis, Kátai [9] proved that the kth absolute moments of the left hand distribution functions converge to the kth absolute moment of the standard normal distribution. Kirschenhofer [12] proved an analogue of (1.1) for the summatory function of s_q^2. In [7] higher powers of s_q were investigated with several different methods.

There are results analogous to (1.1) for number systems other than the q-ary (see, e.g., [3], [14]) and for the summatory function of $A(w,\cdot)$ where $A(w,n)$ gives the number of occurrences of the word w among the sequence of digits of n (see, e.g., [11], [13], [14]). It is also possible to prove (1.1) over certain number fields (see, e.g., [6], [21]). Mean value formulae like (1.1) have applications to the average running time analysis of certain algorithms (see [4] for references to the literature).

Restricting s_q to subsequences of $\mathbb{N}_0$ generally comes with a loss of precision (see (1.2)). In the present paper, the function s_q is averaged over polynomial sequences and an asymptotic formula of type (1.1) is proved. Mauduit and Rivat [17] already showed that on sequences of the form $([n^c])_{n\geq 1}$ the function s_q is uniformly distributed in residue classes and $\alpha\cdot s_q$ is uniformly distributed modulo 1 (where α is an irrational) if $1\leq c<4/3$ (the case $c=1$ goes back to Gelfond [5]). Related results were obtained by Mauduit and Sárközy [18], [19]. They investigated the pseudorandom behaviour of the function χ defined by $\chi(x)=1$ for $0\leq\{x\}<1/2$, $\chi(x)=-1$ for $1/2\leq\{x\}<1$, on sequences $(\alpha n^k)_{n\geq 1}$.

For $x\in\mathbb{R}$, let $[x]$ be the floor of x, $\lceil x\rceil$ the ceiling of x, $\{x\}:=x-[x]$ the fractional part of x and $\psi(x):=\{x\}-1/2$. Define

$$\begin{aligned} J_{q,k}(x) &:= \int_0^x (q\psi(t)-\psi(qt))t^{1/k-1}\,dt, \qquad x\geq 0,\ q,k\in\mathbb{N},\\ F_{q,k}(t) &:= \frac{1}{k}\,q^{(1-\{t\})/k}\sum_{n\geq 0} q^{-n/k}J_{q,k}(q^{n-1+\{t\}})\\ &\qquad + \frac{q-1}{2}(1-\{t\}), \qquad t\in\mathbb{R}. \end{aligned}$$

Theorem. *Let $q,k\in\mathbb{N}\setminus\{1\}$, and $\alpha=1$ or $\alpha>0$ an irrational of finite type. There are $c\in\mathbb{R}$ and $\varepsilon>0$ such that*

$$\begin{aligned} \sum_{0\leq n\leq N} s_q([\alpha n^k]) &= \frac{q-1}{2}\,N\,\frac{\log(\alpha N^k)}{\log q}+cN\\ &\qquad + NF_{q,k}\left(\frac{\log(\alpha N)^k}{\log q}\right)+O(N^{1-\varepsilon}), \qquad N\geq 1. \end{aligned}$$

For $k=1$, the functions $F_{q,1}$ and F in (1.1) are the same. So apart from the term cN and the error term $O(N^{1-\varepsilon})$, the above theorem is a straightforward generalization of (1.1) (but in the proof below we must assume $k \geq 2$).

The essential parts of the proof are contained in Lemmas 2.1–2.3. There the summatory function of $e_r([\alpha n^k])$ is evaluated asymptotically depending on the size of r. For r small and $\alpha = 1$ number-theoretic fluctuations in the distribution of n^k in residue classes give the term cN. The proof is straightforward and elementary. For $\alpha > 0$ of finite type a result of van der Corput and Vinogradov about the discrepancy of αn^k modulo 1 is used. In this case $c = 0$. For r in a middle range there are no fluctuations and an exponential sum estimate of van der Corput is sufficient. For large r a transformation formula from lattice point theory is applied. Thus a second oscillating main term is isolated. This technique is well known in the case where lattice points are counted in large planar sets with zeros of curvature on the boundary.

Acknowledgements. I would like to thank Prof. Dr. D. Wolke for bringing this problem to my attention.

2. Reduction to lattice-sums and lattice-integrals. Set $f(x) := \alpha x^k$. For $x \geq 0$ and $r \in \mathbb{N}_0$, we have

$$e_r([x]) = \left[q\left\{\frac{x}{q^{r+1}}\right\}\right] = \frac{q-1}{2} + q\psi\left(\frac{x}{q^{r+1}}\right) - \psi\left(\frac{x}{q^r}\right).$$

This gives

$$S(N) := \sum_{N<n\leq 2N} s_q([f(n)]) = \sum_{0\leq r\leq R_N} T_r(N), \qquad N \geq 1, \tag{2.1}$$

where $R_N := [\log(\alpha(2N)^k)/\log q]$ and

$$\begin{aligned} T_r(N) &:= \sum_{N<n\leq 2N} e_r([f(n)]) \\ &= \frac{q-1}{2}N + qU_r(N) - U_{r-1}(N) + O(1), \qquad (2.2) \\ U_r(N) &:= \sum_{N<n\leq 2N} \psi\left(\frac{f(n)}{q^{r+1}}\right). \end{aligned}$$

In the following three lemmas $T_r(N)$ is evaluated asymptotically depending on how large q^r is in relation to N.

LEMMA 2.1. *For $r \in \mathbb{N}_0$ and $N \geq 1$ with $q^r \ll N^{1/2}$, we have*

$$T_r(N) = \frac{c_r}{q^{r+1}} N + O(N^{1-\delta})$$

with constants $c_r = c_r(\alpha)$ *and* $\delta = \delta(\alpha) > 0$. *Furthermore,*

$$c_r = \frac{q-1}{2}\, q^{r+1} + O(q^r e^{-\delta r}), \qquad r \in \mathbb{N}_0.$$

Proof. First we handle $\alpha = 1$. Since $n \mapsto e_r(n^k)$ is q^{r+1}-periodic, we have

$$\begin{aligned} T_r(N) &= \sum_{0\le l<q^{r+1}} e_r(l^k) \sum_{N<n\le 2N:\, n\equiv l \,(\mathrm{mod}\, q^{r+1})} 1 \\ &= \sum_{0\le l<q^{r+1}} e_r(l^k)\Big(\frac{N}{q^{r+1}} + O(1)\Big) = \frac{c_r}{q^{r+1}}\, N + O(N^{1/2}), \end{aligned}$$

where

$$\begin{aligned} c_r &:= \sum_{0\le l<q^{r+1}} e_r(l^k) = \sum_{1\le e<q} e\cdot \#\{0\le l<q^{r+1} \mid e_r(l^k)=e\} \\ &= \sum_{1\le e<q} e \sum_{0\le a<q^r} \#\{l \bmod q^{r+1} \mid l^k \equiv eq^r + a \pmod{q^{r+1}}\}. \end{aligned} \tag{2.3}$$

Denote the last cardinality by $c_r(e,a)$. Let $1\le e<q$, $0\le a<q^r$. Let $q = \prod_{i=1}^t p_i^{e_i}$ with $e_1,\dots,e_t \ge 1$ be the prime decomposition of q. Then

$$c_r(e,a) = \prod_{i=1}^t c_{r,i}(e,a) \tag{2.4}$$

by the Chinese Remainder Theorem, where

$$c_{r,i}(e,a) := \#\{x \bmod p_i^{e_i(r+1)} \mid x^k \equiv eq^r + a \pmod{p_i^{e_i(r+1)}}\}. \tag{2.5}$$

If $\mathrm{ord}_{p_i}(a) \ge r/2$ for some $1\le i\le t$ (where $\mathrm{ord}_{p_i}(0) := \infty$) and x is counted in (2.5) then $x^k \equiv 0 \pmod{p_i^{\lceil r/2\rceil}}$ and thus $\mathrm{ord}_{p_i}(x) \ge r/(2k)$. Consequently, we have

$$\begin{aligned} &\sum_{0\le a<q^r:\,\mathrm{ord}_{p_i}(a)\ge r/2 \text{ for some } 1\le i\le t} c_r(e,a) \\ &\le \sum_{1\le i\le t}\sum_{0\le a<q^r} \#\{l \bmod q^{r+1} \mid l^k \equiv eq^r + a \pmod{q^{r+1}},\ l\equiv 0 \pmod{p_i^{\lceil r/(2k)\rceil}}\} \\ &\ll q^r 2^{-r/(2k)}. \end{aligned} \tag{2.6}$$

Now assume $b_i := \mathrm{ord}_{p_i}(a) < r/2$ for $1\le i\le t$. Set $a_i := a p_i^{-b_i}$, $1\le i\le t$. If x is counted in (2.5) then $x^k \equiv 0 \pmod{p_i^{b_i}}$ and thus $x = p_i^{\lceil b_i/k\rceil} y$, $y\in\mathbb{Z}$. Therefore $p_i^{\lceil b_i/k\rceil k - b_i} y^k \equiv eq^r p_i^{-b_i} + a_i \pmod{p_i^{e_i(r+1)-b_i}}$. Since $p_i \nmid a_i$ and $p_i \mid q^r p_i^{-b_i}$, we have $k \mid b_i$. Therefore from now on we assume that $k\mid b_i$ for $1\le i\le t$. Then

$$c_{r,i}(e,a) = \#\{y \bmod p_i^{e_i(r+1)-b_i/k} \mid y^k \equiv eq^r p_i^{-b_i} + a_i \pmod{p_i^{e_i(r+1)-b_i}}\}.$$

For $1 \le i \le t$, let $k = p_i^{u_i} k_i$, $p_i \nmid k_i$, $u_i \in \mathbb{N}_0$. Assume that $r \ge 4u_i + 2e_i$ for $1 \le i \le t$. If $y_1 \equiv y_2 \pmod{p_i^{e_i(r+1)-u_i-b_i}}$ then $y_1 = y_2 + z p_i^{e_i(r+1)-u_i-b_i}$ and

$$\begin{aligned} y_1^k &\equiv y_2^k + k y_2^{k-1} z p_i^{e_i(r+1)-u_i-b_i} + p_i^{2e_i(r+1)-2u_i-2b_i}(\dots) \\ &\equiv y_2^k \pmod{p_i^{e_i(r+1)-b_i}}. \end{aligned}$$

Thus

$$c_{r,i}(e,a) = p_i^{u_i+b_i-b_i/k} c'_{r,i}(e,a), \tag{2.7}$$

where

$$\begin{aligned} &c'_{r,i}(e,a) \\ &\quad := \#\{y \bmod p_i^{e_i(r+1)-u_i-b_i} \mid y^k \equiv e q^r p_i^{-b_i} + a_i \pmod{p_i^{e_i(r+1)-b_i}}\}. \end{aligned}$$

Now

$$c'_{r,i}(e,a) = \sum_{z \bmod p_i^{e_i r-u_i-b_i} : z^k \equiv a_i \pmod{p_i^{e_i r-b_i}}} c''_{r,i}(e,a,y),$$

where

$$\begin{aligned} c''_{r,i}(e,a,y) := \#\{y \bmod p_i^{e_i(r+1)-u_i-b_i} \mid y \equiv z \pmod{p_i^{e_i r-u_i-b_i}}, \\ y^k \equiv e q^r p_i^{-b_i} + a_i \pmod{p_i^{e_i(r+1)-b_i}}\}. \end{aligned}$$

Fix z. Set $y := z + t p_i^{e_i r-u_i-b_i}$. Then

$$\begin{aligned} &y^k \equiv e q^r p_i^{-b_i} + a_i \pmod{p_i^{e_i(r+1)-b_i}} \\ &\qquad \Leftrightarrow \; k_i z^{k-1} t \equiv (a_i - z^k) p_i^{-(e_i r-b_i)} + e q^r p_i^{-e_i r} \pmod{p_i^{e_i}}. \end{aligned}$$

Since $p_i \nmid k_i z^{k-1}$ the last congruence has exactly one solution in t modulo $p_i^{e_i}$. Thus $c''_{r,i}(e,a,y) = 1$ and

$$c'_{r,i}(e,a) = \#\{z \bmod p_i^{e_i r-u_i-b_i} \mid z^k \equiv a_i \pmod{p_i^{e_i r-b_i}}\}. \tag{2.8}$$

Equations (2.4) and (2.6)–(2.8) now give

$$\sum_{0 \le a < q^r} c_r(e,a) = \sum_{0 \le b_i < r/2,\, k \mid b_i \ (1 \le i \le t)} \prod_{i=1}^{t} p_i^{u_i+b_i-b_i/k} \sum_{0 \le a' < q^r \prod_{i=1}^t p_i^{-b_i} : (a',q)=1}$$

$$\prod_{i=1}^{t} \#\Big\{ z \bmod p_i^{e_i r-u_i-b_i} \;\Big|\; z^k \equiv a' \Big(\prod_{j \ne i} p_j^{b_j/k} \Big)^k \pmod{p_i^{e_i r-b_i}} \Big\} + O(q^r 2^{-r/(2k)})$$

for $r \ge \max_{1 \le i \le t}(4u_i + 2e_i) =: r_0$. For $1 \le i \le t$ the ith factor in the second product is $\#\{z \bmod p_i^{e_i r-u_i-b_i} \mid z^k \equiv a' \pmod{p_i^{e_i r-b_i}}\}$. The Chinese Remainder Theorem gives

$$\sum_{0\le a<q^r} c_r(e,a) = \sum_{0\le b_i<r/2,\, k|b_i\ (1\le i\le t)} \prod_{i=1}^t p_i^{u_i+b_i-b_i/k} \sum_{0\le a'<\prod_{i=1}^t p_i^{e_ir-b_i}:(a',q)=1}$$

$$\prod_{i=1}^t p_i^{-u_i} \#\Big\{ z \bmod \prod_{i=1}^t p_i^{e_ir-b_i} \;\Big|\; z^k \equiv a' \pmod{p_i^{e_ir-b_i}} \Big\} + O(q^r 2^{-r/(2k)})$$

$$= \sum_{0\le b_i<r/2,\, k|b_i\ (1\le i\le t)} \prod_{i=1}^t p_i^{b_i(1-1/k)} \prod_{i=1}^t p_i^{e_ir-b_i} \prod_{i=1}^t \left(1-\frac{1}{p_i}\right) + O(q^r 2^{-r/(2k)})$$

$$= q^r + O(q^r 2^{-r/(2k)}).$$

Plugging this result into (2.3) gives $c_r = (q-1)q^{r+1}/2 + O(q^r 2^{-r/(2k)})$.

Next we handle the case of an irrational α of finite type. Define

$$T_r^*(N) := \sum_{1\le n\le N} \left[q \left\{ \frac{\alpha n^k}{q^{r+1}} \right\} \right], \qquad r\in\mathbb{N}_0,\ N\ge 1.$$

Then $T_r(N) = T_r^*(2N) - T_r^*(N)$. Koksma's inequality ([16, Theorem 5.1]) gives

$$\left| \frac{1}{N} T_r^*(N) - \int_0^1 [q\{x\}]\,dx \right| \ll D_{N,r}^*, \tag{2.9}$$

where $D_{N,r}^*$ is the star-discrepancy of the sequence $(\alpha q^{-r-1} n^k)_{1\le n\le N}$ modulo 1. Since α is of finite type there are $\eta\ge 2$ and $C>0$ such that $|\alpha - a/b| \ge Cq^{-\eta}$ for all $a\in\mathbb{Z}$, $b\in\mathbb{N}$. Now let $r\in\mathbb{N}_0$ and $N\ge 1$ with $q^r \ll N^{1/2}$. Dirichlet's Approximation Theorem shows that there are $a\in\mathbb{Z}$ and $b\in\mathbb{N}$ with $(a,b)=1$, $b\le N^{k-1}$, such that

$$\left| \frac{\alpha k!}{q^{r+1}} - \frac{a}{b} \right| \le \frac{1}{bN^{k-1}} \le \frac{1}{b^2}.$$

Now we need a lower bound for b. Since

$$(k!\,b)^{-\eta} \ll \left| \alpha - \frac{aq^{r+1}}{k!\,b} \right| \le \frac{q^{r+1}}{k!\,bN^{k-1}},$$

we get $b \gg (N^{k-1}q^{-r})^{1/(\eta-1)}$. A theorem of Vinogradov and van der Corput [1] now shows that $D_{N,r}^* \ll N^{-\delta}$ for some constant $\delta>0$. From (2.9) it follows that $T_r^*(N)/N = (q-1)/2 + O(N^{-\delta})$ and thus $T_r(N) = (q-1)N/2 + O(N^{1-\delta})$. ∎

LEMMA 2.2. *For $r\in\mathbb{N}_0$ and $N\ge 1$ with $N^{1/2} \ll q^r \ll N^{k-1}$, we have*

$$T_r(N) = \frac{q-1}{2} N + O(N^{1-\delta})$$

with some constant $\delta = \delta(\alpha) > 0$.

Proof. Set $g(x) := q^{-r-1}f(x)$, $a := N$, $b := 2N$, $K := 2^k$, $\lambda := q^{-r}N^k$. For $x \in [a,b]$ and $2 \le \kappa \le k$, we have

$$|g^{(\kappa)}(x)| = |q^{-r-1}\alpha k(k-1)\ldots(k-\kappa+1)x^{k-\kappa}| \asymp q^{-r}x^{k-\kappa} \asymp \lambda a^{-\kappa}.$$

Furthermore, $a \ll \lambda$ by assumption. A classical exponential sum estimate ([15, Theorem 2.8]) now gives

$$U_r(N) = \sum_{a<n\le b} \psi(g(n)) \ll (a^{K-k-1}\lambda)^{1/(K-1)} \ll N^{1-1/(2(K-1))}.$$

Together with (2.2) this proves the lemma. ∎

The above two lemmas show that as long as $q^r \ll N^{k-1}$, the function $T_r(N)$ has a smooth main term with no oscillations. For the range $N^{k-1} \ll q^r \ll N^k$ this is no longer true. Since ψ is oscillating and has mean value 0 one expects the sum $U_r(N)$ to be small in comparison to the length of the summation range. Sums of this type are well known in lattice point theory where they occur in the asymptotic evaluation of the cardinality

$$\#\left\{(x,y) \in \mathbb{N}^2 \;\middle|\; N < x \le 2N,\ 0 < y \le \frac{f(x)}{q^{r+1}}\right\}.$$

Here x is far away from 0 but since $q^r \gg N^{k-1}$ the high order of the zero of $f(x)$ at 0 creates a problem. The usual way is to interchange x and y and count lattice points below the graph of $x = f^{-1}(q^{r+1}y)$. The singularity of $d(f^{-1})/dy$ at 0 now creates an additional main term, and the remainder can be estimated much better since the order of the singularity of $d^2(f^{-1})/dy^2$ is about 2 and thus almost fixed. For $r \in \mathbb{N}_0$ and $N \ge 1$, define

$$I_r(N) := \int_{q^{-r-1}\alpha N^k}^{q^{-r-1}\alpha(2N)^k} x^{1/k-1}\psi(x)\,dx.$$

Lemma 2.3. *For $r \in \mathbb{N}_0$ and $N \ge 1$ with $N^{k-1} \ll q^r \ll N^k$, we have*

$$T_r(N) = \frac{q-1}{2}N + \frac{1}{k}\alpha^{-1/k}(q^{(r+1)/k+1}I_r(N) - q^{r/k}I_{r-1}(N)) + O(N^{2/3}).$$

Proof. The essence of the above idea is contained in Theorem 1.5 of [15]. Applying this transformation formula for ψ-sums to $U_r(N)$ gives

$$\begin{aligned}(2.10)\quad U_r(N) &= \sum_{0<n\le N} \psi(q^{-r-1}f([2N]-n)) + O(1)\\ &= -q^{-r-1}\int_0^N \psi(t)\,f'([2N]-t)\,dt\\ &\quad+ \sum_{q^{-r-1}f([2N]-N)<m\le q^{-r-1}f([2N])} \psi([2N]-f^{-1}(q^{r+1}m))\end{aligned}$$

$$+ q^{r+1} \int\limits_{q^{-r-1}f([2N]-N)}^{q^{-r-1}f([2N])} \frac{\psi(t)}{f'(f^{-1}(q^{r+1}t))}\, dt + O(1).$$

For $x \in \mathbb{R}$, define $\psi_1(x) := \int_0^x \psi(t)\, dt$. The function ψ_1 is continuous, bounded and piecewise continuously differentiable. Integration by parts gives for the first integral in (2.10) the estimate

$$\alpha k \int\limits_0^N \psi(t)([2N]-t)^{k-1}\, dt \tag{2.11}$$

$$= \alpha k \psi_1(t)([2N]-t)^{k-1}\Big|_0^N + \alpha k(k-1) \int\limits_0^N \psi_1(t)([2N]-t)^{k-2}\, dt \ll N^{k-1}.$$

Define $g(t) := [2N] - \alpha^{-1/k} q^{(r+1)/k} t^{1/k}$, $t \geq 0$. For $q^{-r-1}f([2N]-N) \leq t \leq q^{-r-1}f([2N])$, we have $t \asymp q^{-r-1}N^k$ and thus

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

A van der Corput estimate (e.g. [15, Theorem 2.3]) now gives for the ψ-sum in (2.10) the bound

$$\sum_{q^{-r-1}f([2N]-N) < m \leq q^{-r-1}f([2N])} \psi(g(m)) \ll q^{-(r+1)/3} N^{(k+1)/3} + q^{-r-1} N^{k-1/2}$$

$$\ll N^{2/3}.$$

Together with (2.10) and (2.11) we get

$$U_r(N) = O(q^{-r-1}N^{k-1}) + O(N^{2/3}) \tag{2.12}$$

$$+ \frac{1}{k}\alpha^{-1/k} q^{(r+1)/k} \int\limits_{q^{-r-1}f([2N]-N)}^{q^{-r-1}f([2N])} \psi(t) t^{1/k-1}\, dt + O(1).$$

The integral in (2.12) equals

$$I_r(N) + \int\limits_{q^{-r-1}\alpha([2N]-N)^k}^{q^{-r-1}\alpha N^k} \psi(t) t^{1/k-1}\, dt - \int\limits_{q^{-r-1}\alpha[2N]^k}^{q^{-r-1}\alpha(2N)^k} \psi(t) t^{1/k-1}\, dt.$$

The trivial estimate $\psi(t) \ll 1$ shows the last two integrals are $O(q^{-(r+1)/k})$. Thus (2.12) becomes

$$U_r(N) = \frac{1}{k}\alpha^{-1/k} q^{(r+1)/k} I_r(N) + O(N^{2/3}).$$

Plugging this into (2.2) proves the lemma. ■

3. Evaluation of the second main term. Our goal now is to give the second main term a shape as in [2]. For $x \ge 0$, define

$$I(x) := \int_0^x \psi(t) t^{1/k-1}\,dt.$$

Then for $x \ge 0$, we have

$$J(x) := qI(x) - q^{-1/k} I(qx) = \int_0^x (q\psi(t) - \psi(qt)) t^{1/k-1}\,dt.$$

Integration by parts gives for $x \ge 1$ the estimate

$$I(x) = \psi_1(t)t^{1/k-1}\Big|_1^x - \int_1^x \psi_1(t)\Big(\frac{1}{k} - 1\Big) t^{1/k-2}\,dt + \int_0^1 \psi(t)t^{1/k-1}\,dt \ll 1;$$

for $0 \le x < 1$ this estimate can be seen immediately. Thus

$$J(x) \ll 1, \qquad x \ge 0. \tag{3.1}$$

For $N \ge 1$ and $r \in \mathbb{N}_0$, we have $I_r(N) = I(q^{-r-1}\alpha(2N)^k) - I(q^{-r-1}\alpha N^k)$. Therefore

$$\begin{aligned} &\sum_{0\le r\le R_N} (q^{(r+1)/k+1} I_r(N) - q^{r/k} I_{r-1}(N)) \\ &= \sum_{0\le r\le R_N} q^{(r+1)/k} J(q^{-r-1}\alpha(2N)^k) - \sum_{0\le r\le R_N} q^{(r+1)/k} J(q^{-r-1}\alpha N^k) \\ &=: S_1(N) - S_2(N), \qquad N \ge 1. \end{aligned} \tag{3.2}$$

Define

$$y^* := \log(\alpha y^k)/\log q, \quad y > 0, \qquad H(y) := \sum_{n\ge 0} q^{-n/k} J(q^{y+n-1/2}), \quad y \in \mathbb{R}.$$

Then

$$\begin{aligned} S_1(N) &= \sum_{0\le n\le R_N} q^{(R_N-n+1)/k} J(q^{n-R_N-1}\alpha(2N)^k) \\ &= \alpha^{1/k} 2N q^{(1/2-\psi((2N)^*))/k} \sum_{0\le n\le R_N} q^{-n/k} J(q^{\psi((2N)^*)+n-1/2}). \end{aligned}$$

From (3.1) it follows that the sum differs from $H(\psi((2N)^*))$ by

$$O\Big(\sum_{n>R_N} q^{-n/k}\Big) = O(N^{-1})$$

and thus

$$S_1(N) = \alpha^{1/k} 2N q^{(1/2-\psi((2N)^*))/k} H(\psi((2N)^*)) + O(1). \tag{3.3}$$

Furthermore,

$$S_2(N) = S_1\left(\frac{N}{2}\right) + \sum_{R_{N/2}<r\le R_N} q^{(r+1)/k} J(q^{-r-1}\alpha N^k). \tag{3.4}$$

For $R_{N/2} < r \le R_N$, we have $q^{-r-1}\alpha N^k < 1/q$. For $0 \le x < 1/q$, we have

$$J(x) = \int_0^x \left(q\left(t-\frac{1}{2}\right) - \left(qt - \frac{1}{2}\right)\right) t^{1/k-1}\,dt = -\frac{q-1}{2}\,kx^{1/k}.$$

Thus (3.3) and (3.4) give

$$\begin{aligned} S_2(N) = \alpha^{1/k} N q^{(1/2-\psi(N^*))/k} H(\psi(N^*)) \\ - \frac{q-1}{2}\,k\alpha^{1/k} N(R_N - R_{N/2}) + O(1). \end{aligned} \tag{3.5}$$

Integration by parts shows that for $r \in \mathbb{N}_0$ and $N \ge 1$, we have

$$I_r(N) \ll q^{(1-1/k)(r+1)} N^{1-k}. \tag{3.6}$$

Now everything is put together. From (2.1) and Lemmas 2.1–2.3 it follows that

$$\begin{aligned} S(N) = {} & N\frac{q-1}{2} \sum_{0\le r\le R_N} 1 + N \sum_{q^r \le N^{1/2}} \left(\frac{c_r}{q^{r+1}} - \frac{q-1}{2}\right) \\ & + \frac{1}{k}\,\alpha^{-1/k} \sum_{0\le r\le R_N} (q^{(r+1)/k+1} I_r(N) - q^{r/k} I_{r-1}(N)) \\ & - \frac{1}{k}\,\alpha^{-1/k} \sum_{q^r\le N^{k-1}} (q^{(r+1)/k+1} I_r(N) - q^{r/k} I_{r-1}(N)) \\ & + O(N^{1-\delta}\log N). \end{aligned} \tag{3.7}$$

The first sum equals $R_N + 1 = [(2N)^*] + 1$. From Lemma 2.1 it follows that

$$c := \sum_{r\ge 0} \left(\frac{c_r}{q^{r+1}} - \frac{q-1}{2}\right)$$

is absolutely convergent and that the second sum in (3.7) equals

$$c + O\Big(\sum_{q^r > N^{1/2}} e^{-\delta r}\Big) = c + O(N^{-\delta/(2\log q)}).$$

From (3.2), (3.3) and (3.5) it follows that the third sum in (3.7) is

$$\begin{aligned} \alpha^{1/k} 2N q^{(1/2-\psi((2N)^*))/k} H(\psi((2N)^*)) - \alpha^{1/k} N q^{(1/2-\psi(N^*))/k} H(\psi(N^*)) \\ + \frac{q-1}{2}\,k\alpha^{1/k} N(R_N - R_{N/2}) + O(1). \end{aligned}$$

Finally, (3.6) shows that the last sum in (3.7) is $O(1)$. Consequently, there is some constant $\varepsilon > 0$ such that

$$\begin{aligned} S(N) = {} & \frac{q-1}{2}\, 2N[(2N)^*] - \frac{q-1}{2}\, N[N^*] \\ & + \frac{1}{k}\, 2Nq^{(1/2-\psi((2N)^*))/k} H(\psi((2N)^*)) \\ & - \frac{1}{k}\, Nq^{(1/2-\psi(N^*))/k} H(\psi(N^*)) \\ & + 2N\left(c + \frac{q-1}{2}\right) - N\left(c + \frac{q-1}{2}\right) + O(N^{1-\varepsilon}), \qquad N \geq 1. \end{aligned}$$

Summing up over $N = 2^{-i}x$, $1 \leq i \leq \log x/\log 2$, gives, for $x \geq 1$,

$$\begin{aligned} \sum_{0 \leq n \leq x} s_q([\alpha n^k]) &= \sum_{1 \leq i \leq \log x/\log 2} S(2^{-i}x) + O(1) \\ &= \frac{q-1}{2}\, x[x^*] + \frac{1}{k}\, xq^{(1/2-\psi(x^*))/k} H(\psi(x^*)) \\ &\quad + \left(c + \frac{q-1}{2}\right) x + O(x^{1-\varepsilon}). \end{aligned}$$

This proves the Theorem. ■

References

[1] J. G. van der Corput et C. Pisot, *Sur la discrépance modulo m*, Indag. Math. 1 (1939), 184–195.

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

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

[4] P. Flajolet, P. Grabner, P. Kirschenhofer, H. Prodinger and R. F. Tichy, *Mellin transforms and asymptotics: digital sums*, Theoret. Comput. Sci. 123 (1994), 291–314.

[5] A. O. Gelfond, *Sur les nombres qui ont des propriétés additives et multiplicatives données*, Acta Arith. 13 (1968), 259–265.

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

[7] P. J. Grabner, P. Kirschenhofer, H. Prodinger and R. F. Tichy, *On the moments of the sum-of-digits function*, in: Applications of Fibonacci Numbers 5, G. E. Bergum *et al.* (eds.), Kluwer, Dordrecht, 1993, 263–271.

[8] E. Heppner, *Über die Summe der Ziffern natürlicher Zahlen*, Tagungsbericht 44/1975 Zahlentheorie, Math. Forschungsinstitut Oberwohlfach (2. 11. bis 8. 11. 1975).

[9] I. Kátai, *On the sum of digits of primes*, Acta Math. Acad. Sci. Hungar. 30 (1977), 169–173.

[10] I. Kátai and J. Mogyoródi, *On the distribution of digits*, Publ. Math. Debrecen 15 (1968), 57–68.

[11] P. Kirschenhofer, *Subblock occurrences in the q-ary representation of n*, SIAM J. Algebraic Discrete Methods 4 (1983), 231–236.

[12] —, *On the variance of the sum of digits function*, in: Number Theoretic Analysis, Seminar Wien 1988–89, Lecture Notes in Math. 1452, Springer, Berlin, 1990, 112–116.

[13] P. Kirschenhofer and H. Prodinger, *Subblock occurrences in positional number systems and Gray code representations*, J. Inform. Optim. Sci. 5 (1984), 29–42.

[14] P. Kirschenhofer and R. F. Tichy, *On the distribution of digits in Cantor representations of integers*, J. Number Theory 18 (1984), 121–134.

[15] E. Krätzel, *Lattice Points*, Kluwer, Dordrecht, 1988.

[16] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974.

[17] C. Mauduit et J. Rivat, *Répartition des fonctions q-multiplicatives dans la suite* $([n^c])_{n\in\mathbb{N}}$, $c > 1$, Acta Arith. 71 (1995), 171–179.

[18] C. Mauduit and A. Sárközy, *On finite pseudorandom binary sequences, V. On $(n\alpha)$ and $(n^2\alpha)$ sequences*, Monatsh. Math. 129 (2000), 197–216.

[19] —, —, *On finite pseudorandom binary sequences, VI. On $(n^k\alpha)$ sequences*, ibid. 130 (2000), 281–298.

[20] I. Shiokova, *On the sum of digits of prime numbers*, Proc. Japan Acad. 50 (1974), 551–554.

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

Mathematisches Institut
Albert-Ludwigs-Universität
Eckerstr. 1
D-79104 Freiburg, Germany
E-mail: manfred.peter@math.uni-freiburg.de

Received on 15.6.2001
and in revised form on 9.11.2001 (4054)