JOURNAL OF NUMBER THEORY **18**, 121–134 (1984)

On the Distribution of Digits in Cantor Representations of Integers

P. KIRSCHENHOFER AND R. F. TICHY

Department of Mathematics,
Technical University of Vienna, A-1040 Vienna, Austria

Communicated by E. Hlawka

Received May 3, 1982

Asymptotic formulas on the average values of the "sum of digits" function and the average numbers of occurrences of fixed subblocks in Cantor representations of integers are established. The theorems generalize a result by H. Delange [*Enseign. Math.* **21** (1975), 31–47].

1. INTRODUCTION AND NOTATION

The problem of the average value of the "sum of digits" function of (positive) integers has been studied by a number of authors:

Let q denote an integer >1 and $S_q(n)$ the sum of digits of n in q-ary representation. For the average value

$$M(m) := \frac{1}{m} \sum_{n=0}^{m-1} S_q(n) \tag{1.1}$$

the following results have been established: Bush [2] has shown

$$M(m) \sim \frac{q-1}{2 \log q} \log m \qquad (m \to \infty),$$

Bellman and Shapiro [1] derived

$$M(m) = \frac{q-1}{2 \log q} \log m + O(\log \log m)$$

and Mirsky [5] improved the O-term to $O(1)$.

An explicit formula for the error term is due to Delange [3]:

$$M(m) = \frac{q-1}{2 \log q} \log m + F\left(\frac{\log m}{\log q}\right), \tag{1.2}$$

0022-314X/84 $3.00

where F is a continuous and periodic function with period 1, the Fourier coefficients of which are determined in [3], too. (It is even shown that F is nowhere differentiable.)

The above cited problem is also related to the question of the average number of occurrences of a fixed digit or a fixed block of digits in the q-ary representation of a natural number. If we denote by $B_q(w, n)$ the number of blocks w in the q-ary representation of n, where overlapping is allowed, the following result holds, if the first and last digits of n differ from zero:

$$\frac{1}{m} \sum_{n=0}^{m-1} B_q(w, n) = \left(\frac{\log m}{\log q} - |w| + 1\right) q^{-|w|} + H_w\left(\frac{\log m}{\log q}\right) + O\left(\frac{1}{m}\right). \quad (1.3)$$

Here $|w|$ is the length of w and H_w is continuous and periodic with period 1.

In the special case $q = 2$, $w = 1^s$ this is proved by Prodinger [6], the general case is due to Kirschenhofer [4].

The present paper deals with a generalization of the above results to the case of Cantor representation of integers. Let $(q(i))$ be a sequence of natural numbers with $q(0) = 1$, $q(i) > 1$ for all $i \geqslant 1$, and $Q(j) := q(0) \cdot q(1) \cdot \cdots \cdot q(j)$. Any natural number n has a unique representation

$$n = \sum_{j \geqslant 0} a_j(Q; n)\, Q(j), \qquad 0 \leqslant a_j(Q; n) \leqslant q(j+1) - 1, \quad (1.4)$$

which we will call the Q-Cantor representation of n induced by $(q(i))$. If we define the average value

$$M(Q; m) = \frac{1}{m} \sum_{n=0}^{m-1} S(Q; n), \quad (1.5)$$

where $S(Q; n) = \sum_{j \geqslant 0} a_j(Q; n)$ is the sum of digits in the Q-Cantor representation, we obtain in the case $1 = q(0) < q(1) \leqslant q(2) \leqslant q(3) \leqslant \cdots$

$$\begin{aligned} M(Q; m) = {} & \frac{1}{2} \sum_{j=1}^{Q^*(m)} (q(j) - 1) + \frac{P(m)}{2} + \frac{q(Q^*(m))\{P(m)\}^2}{2P(m)} \\ & - \frac{1}{2} - \frac{H(P(m))}{P(m)} - \frac{\{P(m)\}}{2P(m)} \\ & + \frac{q(Q^*(m-1) - 1)\{P(m)\, q(Q^*(m))\}^2}{2q(Q^*(m))\, P(m)} + O\left(\frac{1}{q(Q^*(m))}\right). \end{aligned} \quad (1.6)$$

In this formula $Q^*(m) = i$ denotes the uniquely determined integer $i \geqslant 0$ with $Q(i) \leqslant m < Q(i+1)$, and $P(m) = m/Q(Q^*(m))$; H is the periodic function

$$H(x) = \int_0^x (\{v\} - \tfrac{1}{2})\, dv$$

($\{x\} = x - [x]$ is the fractional part of x).

In the case $\lim_{i\to\infty} q(i)=\infty$ the O-term tends to 0, while all other terms need not do so.

If $1=q(0)<q(1)\leqslant q(2)\leqslant\cdots\leqslant q(s)\leqslant q(s+1)=q(s+2)=\cdots=q$ an easy modification of [3] yields

$$\begin{aligned} M(Q;m)=\frac{q-1}{2}\frac{\log(m/c)}{\log q}+\frac{1}{2}\sum_{j=1}^{s}(q(j)-1) \\ +F\left(\frac{\log(m/c)}{\log q}\right)+O\left(\frac{1}{m}\right) \end{aligned} \tag{1.7}$$

with $c=q(1)\cdots q(s)$ and F the periodic function from (1.6).

For a bounded, but not necessarily monotone sequence $(q(i))$ we obtain in general

$$M(Q;m)=\tfrac{1}{2}\sum_{j=1}^{Q^*(m)}(q(j)-1)+O(1). \tag{1.8}$$

If the sequence $q(i)$ is periodic of length k, i.e., $q(0)=1$, $q(i)=q_j$ where $i\equiv j(k)$ $(1\leqslant j\leqslant k)$, $q=q_1\cdot\cdots\cdot q_k$ a result similar to Delange (1.2) is derived:

$$M(Q;m)=\frac{\log m}{2\log q}\sum_{j=1}^{k}(q_j-1)+G\left(\frac{\log m}{\log q}\right), \tag{1.9}$$

where G is a continuous function, periodic of period 1, the Fourier expansion of which will be given.

In the problem of the average occurrences of single digits or blocks we will show that for a sequence $(q(i))$ with $\lim_{i\to\infty} q(i)=\infty$

$$\frac{1}{m}\sum_{n=0}^{m-1}B(Q;w,n)=\sum_{j=|w|}^{Q^*(m)+|w|}\frac{Q(j-|w|)}{Q(j)}+O(1), \tag{1.10}$$

holds for the numbers $B(Q;w,n)$ of occurrences of w (with first and last digit different from 0) as a subblock of n in Q-Cantor representation. (From our lemmata the $O(1)$-term could be analyzed more in detail, too, which we will not do for the sake of brevity).

2. The "Sum of Digits" Function in Cantor Representation

In order to express the "sum of digits" function $S(Q;n)$ we prove the following identity for the digits $a_j(Q;n)$ of n in Q-Cantor representation (compare 1.4).

LEMMA 1. $a_j(Q; n) = \lfloor n/Q(j) \rfloor - q(j+1)\lfloor n/Q(j+1) \rfloor$.

Proof. We have $n = \sum_{j \geqslant 0} a_j(Q; n)\, Q(j)$ and so

$$\left[\frac{n}{Q(j)}\right] = \sum_{k \geqslant j} a_k(Q; n) \frac{Q(k)}{Q(j)}$$

and

$$\left[\frac{n}{Q(j+1)}\right] = \sum_{k \geqslant j+1} a_k(Q; n) \frac{Q(k)}{Q(j)\, q(j+1)}$$

from which the result is immediate.

For the following calculations observe that

$$\left[\frac{n}{Q(j)}\right] = \left[\frac{t}{Q(j)}\right] \qquad \text{for} \quad n \leqslant t < n+1. \tag{2.1}$$

Therefore

$$\begin{aligned} S(Q; n) &= \sum_{j \geqslant 0} \left(\left[\frac{n}{Q(j)}\right] - q(j+1) \left[\frac{n}{Q(j+1)}\right] \right) \\ &= \sum_{j \geqslant 0} \int_n^{n+1} \left(\left[\frac{t}{Q(j)}\right] - q(j+1) \left[\frac{t}{Q(j+1)}\right] \right) dt \end{aligned}$$

and

$$\sum_{n=0}^{m-1} S(Q; n) = \sum_{j \geqslant 0} \int_0^m \left(\left[\frac{t}{Q(j)}\right] - q(j+1) \left[\frac{t}{Q(j+1)}\right] \right) dt. \tag{2.2}$$

For $j \geqslant Q^*(m) + 1$ and $n \leqslant m - 1$

$$\frac{n}{Q(j)} \leqslant \frac{m-1}{Q(Q^*(m)+1)} < 1$$

because of the definition of Q^* (below 1.6); hence it suffices in (2.2) to sum over $0 \leqslant j \leqslant Q^*(m)$. Let

$$g_j(x) = \int_0^x \left(\lfloor q(j)\, t \rfloor - q(j) \lfloor t \rfloor - \frac{q(j)-1}{2} \right) dt. \tag{2.3}$$

Then g_j is continuous periodic with period 1 and $g_j(n) = 0$ for $n \in \mathbb{Z}$.

With this abbreviation (2.2) reads

$$\sum_{n=0}^{m-1} S(Q; n) = \sum_{j=0}^{Q^*(m)} \left(m \frac{q(j+1)-1}{2} + g_{j+1}\left(\frac{m}{Q(j+1)}\right) Q(j+1) \right)$$

and so we have proved the following

LEMMA 2. *$M(Q; m) = (1/2) \sum_{j=1}^{u(m)} (q(j) - 1) + (1/m) \sum_{j=1}^{u(m)} g_j(m/Q(j))$ with $u(m) = Q^*(m) + 1$ and $M(Q; m)$ as in* (1.5).

For the following let $(q(i))$ be a monotone sequence and $\lim_{i \to \infty} q(i) = \infty$. In order to derive (1.6) we first estimate

$$\frac{1}{m} \sum_{j=1}^{u(m)-3} g_j\left(\frac{m}{Q(j)}\right) Q(j) = O\left(\frac{1}{q(Q^*(m))}\right), \tag{2.4}$$

which is a consequence of

$$\left| g_j\left(\frac{m}{Q(j)}\right) \right| \leqslant \left\{\frac{m}{Q(j)}\right\} \frac{q(j)-1}{2} \leqslant \frac{q(j)-1}{2}$$

and

$$\frac{1}{m} \sum_{j=1}^{u(m)-3} (q(j)-1)\, Q(j) \leqslant \frac{1}{m} Q(Q^*(m)-1)$$

$$= \frac{1}{q(Q^*(m))} \cdot \frac{Q(Q^*(m))}{m} \leqslant \frac{1}{q(Q^*(m))}.$$

After having established (2.4) it remains to consider the contribution of the cases $j = u(m) - 2$, $u(m) - 1$, $u(m)$ to the second sum in Lemma 2. We start with $j = u(m)$:

$$g_{u(m)}\left(\frac{m}{Q(u(m))}\right) = \int_0^{m/Q(u(m))} [tq(u(m))]\, dt - q(u(m)) \cdot \int_0^{m/Q(u(m))} [t]\, dt$$

$$- \frac{q(u(m))-1}{2} \frac{m}{Q(u(m))}.$$

By the substitution $v = t \cdot q(u(m))$ the first integral becomes

$$\frac{1}{q(u(m))} \int_0^{P(m)} [v]\, dv = \frac{P(m)^2}{2q(u(m))} - \frac{P(m)}{2q(u(m))} - \frac{H(P(m))}{q(u(m))}$$

with

$$P(m) = \frac{m}{Q(Q^*(m))} \qquad \text{and} \qquad H(x) = \int_0^x \left(\{v\} - \frac{1}{2}\right) dv.$$

The second integral is zero since $m/Q(u(m)) \leqslant 1$. Putting all together the contribution of $j = u(m)$ is

$$\frac{P(m)}{2} - \frac{1}{2} - \frac{H(P(m))}{P(m)} - \frac{q(u(m))-1}{2}. \tag{2.5}$$

A similar calculation yields for $j = u(m) - 1 = Q^*(m)$

$$g_{u(m)-1}\left(\frac{m}{Q(u(m)-1)}\right) = \int_0^{\{P(m)\}} \left([t \cdot q(u(m)-1)] - \frac{q(u(m)-1}{2}\right) dt$$

and with

$v = t \cdot q(u(m)-1)$

$$= \frac{q(u(m)-1)\{P(m)\}^2}{2} - \frac{\{P(m)\}}{2} - \frac{1}{q(u(m)-1)} H(\{P(m)\}\, q(u(m)-1))$$

and finally for the contribution of $j = u(m) - 1$

$$\frac{q(u(m)-1)\{P(m)\}^2}{2P(m)} - \frac{\{P(m)\}}{2P(m)} + O\left(\frac{1}{q(u(m)-1)}\right). \tag{2.6}$$

In the same way

$$g_{u(m)-2}\left(\frac{m}{Q(u(m)-2)}\right)$$
$$= \frac{g(u(m)-2)\{P(m)\, q(u(m)-1)\}^2}{2} - \frac{\{P(m)\, q(u(m)-1)\}}{2}$$
$$- \frac{1}{q(u(m)-2)} H(\{P(m)\, q(u(m)-1)\}\, q(u(m)-2);$$

or for the contribution of $j = u(m) - 2$

$$\frac{\{P(m)\, q(u(m)-1)\}^2}{2P(m)} \cdot \frac{q(u(m)-2)}{q(u(m)-1)} + O\left(\frac{1}{q(u(m)-1)}\right). \tag{2.7}$$

Combining Lemma 2 with (2.4), (2.5), (2.6) and (2.7) the following theorem results:

THEOREM 1. *The average value $M(Q; m) = (1/m) \sum_{n=0}^{m-1} S(Q; n)$ of the "sum of digits" function in Q-Cantor representation* (1.4) *induced by a monotone sequence $1 = q(0) < q(1) \leqslant q(2) \leqslant \cdots$ (with $\lim_{i\to\infty} q(i) = \infty$) satisfies the asymptotic relation*

$$M(Q; m) = \frac{1}{2} \sum_{j=1}^{Q^*(m)} (q(j) - 1) + \frac{P(m)}{2} + \frac{q(Q^*(m))\{P(m)\}^2}{2P(m)}$$
$$- \frac{1}{2} - \frac{H(P(m))}{P(m)} - \frac{\{P(m)\}}{2P(m)}$$
$$+ \frac{q(Q^*(m)-1)\{P(m)\, q(Q^*(m))\}^2}{2q(Q^*(m))\, P(m)} + O\left(\frac{1}{q(Q^*(m))}\right)$$

with $H(x) = \int_0^x (\{v\} - \frac{1}{2})\, dv$, $P(m) = m/Q(Q^*(m))$ *and* $Q^*(m) = i$ *the integer with* $Q(i) \leqslant m < Q(i+1)$.

If $q(i)$ is monotone and bounded (1.7) results immediately from Lemma 2 by Delange's method of [3]. In the following we want to give some examples:

EXAMPLE 1. We take $q(i) = i + 1$ $(i \geqslant 0)$; then $Q(i) = (i+1)!$. By Stirling's approximation of the Γ-function

$$Q^*(m) = \frac{\log m}{\log \log m} + O\left(\frac{\log m \cdot \log \log \log m}{(\log \log m)^2}\right)$$

and furthermore

$$P(m) = \frac{mq(Q^*(m)+1)}{Q(Q^*(m)+1)} < Q^*(m) + 2 = O\left(\frac{\log m}{\log \log m}\right).$$

These estimates combined with Theorem 1 lead to

$$M(Q; m) = \frac{1}{4}\left(\frac{\log m}{\log \log m}\right)^2 + O\left(\frac{(\log m)^2 \log \log \log m}{(\log \log m)^3}\right).$$

EXAMPLE 2. We take $q(i) = 2^i$ $(i \geqslant 0)$ and obtain $Q(i) = 2^{i(i+1)/2}$. It is not difficult to see that

$$Q^*(m) = \left[\frac{-1 + \sqrt{1 + 8ld(m)}}{2}\right] =: [a_m],$$

where ld denotes the logarithm with respect to the base 2. Then

$$\frac{1}{2} \sum_{j=1}^{Q^*(m)} (q(j) - 1) = 2^{a_m - \{a_m\}} - \frac{1}{2} a_m + O(1),$$

$$\frac{P(m)}{2} = 2^{a_m\{a_m\} + \frac{1}{2}(\{a_m\} - \{a_m\}^2) - 1},$$

$$\frac{q(Q^*(m))}{2P(m)} = 2^{a_m(1 - \{a_m\}) + \frac{1}{2}(\{a_m\}^2 - 3\{a_m\}) - 1}.$$

All other terms appearing in Theorem 1 are $O(1)$, too, and could be determined explicitly if desired. ∎

For a bounded but not necessarily monotone sequence $(q(i))$ Lemma 2 holds, too. The second sum in Lemma 2 can be estimated by

$$\frac{1}{m}\sum_{j=1}^{u(m)} g_j\left(\frac{m}{Q(j)}\right)Q(j) \leqslant \frac{\text{const}}{m}\, Q(Q^*(m)+2) = O(1)$$

and so (1.8) results.

3. The "Sum of Digits" Function for Periodic Bases

In this section we consider periodic sequences $(q(i))$ of the following type:

$$\begin{aligned} &q(0) = 1 \\ &q(i) = q_j \qquad \text{if} \quad i \equiv j(k) \ \text{ with } \ 1 \leqslant j \leqslant k \ \text{ and fixed } \ q_1, \ldots, q_k. \end{aligned} \tag{3.1}$$

Then

$$Q(i) = \prod_{j=1}^{k} q_j^{[(i+k-j)/k]}. \tag{3.2}$$

In order to determine the average values $M(Q; m)$ we shall use Lemma 2 again. Therefore we calculate $Q^*(m)$ at first:

Lemma 3. *If the sequence $(q(i))$ is periodic as in (3.1),*

$$Q^*(m) = \sum_{j=1}^{k} [lqm + lq(q_{j+1} \cdot \cdots \cdot q_k)],$$

where lq is the logarithm to the base $q := q_1 \cdot \cdots \cdot q_k$.

Proof. We choose l with $2 \leqslant l \leqslant k$ such that

$$[lqm + lq(q_{l+1} \cdot \cdots \cdot q_k)] = [lqm],$$

but

$$[lqm + lq(q_l \cdots q_k)] = [lqm] + 1. \tag{3.3}$$

Let $A(m)$ denote the right-hand expression in the equation of Lemma 3; then by (3.2), (3.3)

$$Q(A(m)) = q^{[lqm]+1-lq(q_l \cdots q_k)} \leqslant m$$

and

$$m < q^{[lqm]+1-lq(q_{l+1}\cdots q_k)} = Q(A(m)+1)$$

from which the Lemma follows immediately.

By Lemma 2 the first term of $M(Q; m)$ is

$$\tfrac{1}{2}\sum_{i=1}^{k}(q_i-1)\operatorname{card}\{r : 1 \leqslant r \leqslant Q^*(m)+1 \text{ and } r \equiv i(k)\}$$
$$= \tfrac{1}{2}\sum_{i=1}^{k}(q_i-1)A(m,i).$$

LEMMA 4. $A(m,i) = [lqm + lq(q_i\cdots q_k)]$.

Proof. Let l be defined as in (3.3). Then by Lemma 3

$$Q^*(m)+1 = k[lqm]+1$$

and therefore

$$A(m,i) = [lqm] + \begin{cases} 0 & \text{for } \ i > l \\ 1 & \text{for } \ i \leqslant l \end{cases}$$

which equals $[lqm + lq(q_i\cdots q_k)]$ by (3.3).

So the first term of $M(Q; m)$ is

$$\tfrac{1}{2}\sum_{i=1}^{k}(q_i-1)[lq(m)+lq(q_i\cdots q_k)]. \tag{3.4}$$

It remains for us to investigate

$$\frac{1}{m}\sum_{j=1}^{u(m)} g_j\left(\frac{m}{Q(j)}\right)Q(j) = \sum_{i=1}^{k} R_i(m)$$

with

$$R_i(m) = \frac{1}{m}\sum_{j=1,\, j\equiv i(k)}^{Q^*(m)+1} g_i(mq^{-[j/k]}(q_1\cdots q_i)^{-1})\, q^{[j/k]}(q_1\cdots q_i).$$

By the substitution $r = [j/k]$

$$R_i(m) = (q_{i+1}\cdots q_k)^{-1}\frac{1}{m}\sum_{r=0}^{[lq(m/q_1\cdots q_{i-1})]} g_i\left(\frac{m(q_{i+1}\cdots q_k)}{q^{r+1}}\right)q^{r+1}$$

and by $s = [lq(m/q_1\cdots q_{i-1})] - r$ (observing $g_i(n) = 0$ for integers n):

$$R_i(m) = q_i\sum_{s\geqslant 0} g_i\left(\frac{1}{q_i}\cdot q^{s+\{lq(m/q_1\cdots q_{i-1})\}}\right)\cdot q^{-s-\{lq(m/q_1\cdots q_{i-1})\}}.$$

Similarily to Delange [3] we introduce

$$h_i(x) = \sum_{j \geq 0} q^{-j} g_i(q^j x) \qquad (1 \leq i \leq k) \tag{3.5}$$

and get

$$R_i(m) = q_i h_i \left(\frac{1}{q_i} q^{\{lq(m/q_1 \cdots q_{i-1})\}} \right) \cdot q^{-\{lq(m/q_1 \cdots q_{i-1})\}}.$$

With the abbreviation

$$\begin{aligned} G_i(x) = \frac{1}{2} (q_i - 1)(1 - \{x - lq(q_1 \cdots q_{i-1})\} - lq(q_1 \cdots q_{i-1})) \\ + q_i q^{-\{x - lq(q_1 \cdots q_{i-1})\}} h_i \left(\frac{1}{q_i} \cdot q^{\{x - lq(q_1 \cdots q_{i-1})\}} \right) \end{aligned} \tag{3.6}$$

we obtain the following theorem:

THEOREM 2. *Let $(q(i))$ be a periodic sequence as in* (3.1). *Then the average value of the "sum of digits" function in Q-Cantor representation* (1.4) *induced by the sequence $(q(i))$ is given by*

$$M(Q; m) = \frac{1}{2} \frac{\log m}{\log q} \sum_{j=1}^{k} (q_j - 1) + \sum_{j=1}^{k} G_j \left(\frac{\log m}{\log q} \right)$$

where $q = q_1 \cdots q_k$, and G_j are continuous and periodic functions with period 1 *as in* (3.6).

In a similar way as in [3] it is possible to expand the function $G(x) = \sum_{j=1}^{k} G_j(x)$ in a Fourier series

$$G(x) = \sum_{r \in \mathbb{Z}} c_r e^{2\pi i r x}$$

with

$$c_0 = \sum_{j=1}^{k} \frac{q_j - 1}{2 \log q} (\log 2\pi - 1 - \log(q_1 \cdot \cdots \cdot q_j)) - \frac{1}{2} + \sum_{j=1}^{k} \frac{q_j - 1}{4}$$

and for $r \neq 0$:

$$c_r = i \frac{\zeta \left(\frac{2r\pi i}{\log q} \right)}{2r\pi \left(1 + \frac{2r\pi i}{\log q} \right)} \sum_{j=1}^{k} (q_j - 1) \, e^{-2\pi i r (\log(q_1 \cdots q_j)/\log q)},$$

where ζ denotes Riemann's ζ-function.

4. Occurrences of Blocks in Cantor Representation

In this section we want to study the average occurrences of blocks in Cantor representations of integers. Let $B(Q; w, n)$ denote the number of occurrences of the string w (with first and last digit different from 0) as a block in Q-Cantor representation of n.

LEMMA 5.

$$B(Q; w, n) = \sum_k \left(\left[\frac{n}{Q(k)} + b(k, w) + \frac{Q(k-l-1)}{Q(k)} \right] - \left[\frac{n}{Q(k)} + b(k, w) \right] \right),$$

where $l+1$ denotes the length of the string $w = w_0 w_1 \cdots w_l$, the sum is taken over all $k \geqslant l+1$ such that $w_i < q(k-i)$ $(0 \leqslant i \leqslant l)$ and

$$b(k, w) = \sum_{j=0}^{l} \frac{Q(k-j-1)}{Q(k)} (q(k-j) - 1 - w_j).$$

The part of the sum, where $Q(k-l-1) > n$ is zero.

Proof. Let $n = \sum_{j \geqslant 0} a_j(Q; n)\, Q(j)$ be the Q-Cantor representation of n, then

$$\begin{aligned} \frac{n}{Q(k)} + b(k, w) = {} & \sum_{j \geqslant k} a_j(Q; n) \frac{Q(j)}{Q(k)} \\ & + \sum_{j=k-(l+1)}^{k-1} (a_j(Q; n) + q(j+1) - 1 - w_{k-j-1}) \frac{Q(j)}{Q(k)} \\ & + \sum_{j=0}^{k-(l+1)-1} a_j(Q; n) \frac{Q(j)}{Q(k)}. \end{aligned}$$

The third sum is less then

$$\frac{Q(k-l-1)}{Q(k)} = \frac{1}{q(k-l) \cdots q(k)}$$

and therefore cannot contribute to the value of the sum in Lemma 5; trivially the same holds for the first sum because it is a natural number. The second sum $\sum_2$ satisfies the inequality

$$0 \leqslant \sum_2 < 2 - \frac{Q(k-l-1)}{Q(k)};$$

so $\lfloor \sum_2 + Q(k-l-1)/Q(k) \rfloor - \lfloor \sum_2 \rfloor$ can only take the values 0 or 1, and is 1 if and only if

$$\sum_2 + \frac{Q(k-l-1)}{Q(k)} = 1.$$

So the k-term of the sum in Lemma 5 takes the value 1, if

$$w = a_{k-1}(Q; n) \cdots a_{k-(l+1)}(Q; n)$$

and takes the value 0 if not.

The lower bound for k is trivial; the upper bound follows from the fact that for $Q(k-l-1) > n$ we have $a_j(n) = 0$ for $j \geqslant k-l-1$. (This will be used in Lemma 7.) Now the proof of Lemma 5 is complete.

LEMMA 6. *For any t, $b \in \mathbb{R}$ and k a positive integer:*

$$\left\lfloor \frac{t}{k} + b \right\rfloor = \left\lfloor \frac{n}{k} + b \right\rfloor \qquad \textit{with} \quad n = \lfloor t + \{kb\} \rfloor.$$

Proof.

$$\left\lfloor \frac{t}{k} + b \right\rfloor = \left\lfloor \frac{\lfloor t + kb \rfloor}{k} \right\rfloor = \left\lfloor \frac{\lfloor t + kb \rfloor + \{kb\}}{k} \right\rfloor = \left\lfloor \frac{\lfloor t + \{kb\} \rfloor}{k} + b \right\rfloor.$$

Now we can prove the following identity for the desired average values of $B(Q; w, n)$:

LEMMA 7.

$$\frac{1}{m} \sum_{n=0}^{m-1} B(Q; w, n) = \sum_k \frac{Q(k-l-1)}{Q(k)} + \frac{1}{m} \sum_k f_{k,w}\left(\frac{m}{Q(k)}\right) Q(k),$$

where the sums are taken over all $l+1 \leqslant k \leqslant l+1+Q^*(m)$ such that $w_i < q(k-i)$ $(0 \leqslant i \leqslant l)$ and $f_{k,w}$ are continuous periodic functions with period 1 *defined by*

$$f_{k,w}(x) = \int_0^x \left(\left\lfloor u + b(k,w) + \frac{Q(k-l-1)}{Q(k)} \right\rfloor \right.$$
$$\left. - \lfloor u + b(k,w) \rfloor - \frac{Q(k-l+1)}{Q(k)} \right) du.$$

Proof. By Lemmata 5 and 6:

$$\sum_{n=0}^{m-1} B(Q; w, n)$$

$$= \sum_k \left(\int_{-\{Q(k)b(k,w)+Q(k-l-1)\}}^{m-\{Q(k)b(k,w)+Q(k-l-1)\}} \left[\frac{t}{Q(k)} + b(k,w) + \frac{Q(k-l-1)}{Q(k)} \right] dt \right.$$

$$\left. - \int_{-\{Q(k)b(k,w)\}}^{m-\{Q(k)b(k,w)\}} \left[\frac{t}{Q(k)} + b(k,w) \right] dt \right)$$

$$= \sum_k \int_0^m \left(\left[\frac{t}{Q(k)} + b(k,w) + \frac{Q(k-l-1)}{Q(k)} \right] - \left[\frac{t}{Q(k)} + b(k,w) \right] \right) dt.$$

The upper bound for k follows from the fact that for all $k > l + 1 + Q^*(m)$ we have $Q(k-l-1) > n$ $(0 \leqslant n \leqslant m-1)$ and therefore no contribution to the sum by the last part of Lemma 5.

In the following we consider a sequence $(q(i))$ such that $\lim_{i \to \infty} q(i) = \infty$ and

$$\lim_{m \to \infty} \sum_{k=l+1}^{l+1+Q^*(m)} \frac{Q(k-l-1)}{Q(k)} = \infty. \tag{4.1}$$

In this case the first sum in the expression for

$$\frac{1}{m} \sum_{n=0}^{m-1} B(Q; w, n)$$

in Lemma 8 is the main contribution:

THEOREM 3. *The average value of the number of occurrences of the string w (with first and last digit different from zero) as a block in Q-Cantor representation of n induced by a sequence q(i) such that* (4.1) *holds, satisfies the asymptotic relation:*

$$\frac{1}{m} \sum_{n=0}^{m-1} B(Q; w, n) = \sum_{k=|w|}^{|w|+Q^*(m)} \frac{Q(k-|w|)}{Q(k)} + O(1),$$

where $|w|$ is the length of w.

Proof. By Lemma 7 we have to show that

$$\frac{1}{m} \sum_k f_{k,w} \left(\frac{m}{Q(k)} \right) Q(k) = O(1).$$

Now, $|f_{k,w}(x)| \leqslant \{x\}$ and so

$$\left| \frac{1}{m} \cdot \sum_{k=l+1}^{Q^*(m)-1} f_{k,w}\left(\frac{m}{Q(k)}\right) Q(k) \right| \leqslant \frac{1}{m} Q(Q^*(m)) \leqslant 1,$$

$$\left| \frac{1}{m} \cdot f_{Q^*(m),w}\left(\frac{m}{Q(Q^*(m))}\right) Q(Q^*(m) \right| \leqslant 1,$$

$$\left| \frac{1}{m} \sum_{k=Q^*(m)+1}^{l+1+Q^*(m)} f_{k,w}\left(\frac{m}{Q(k)}\right) Q(k) \right| \leqslant \frac{1}{m} \sum_{k=Q^*(m)+1}^{l+1+Q^*(m)} \frac{m}{Q(k)} Q(k) \leqslant l+1.$$

It should be mentioned that by a more detailed study of the $O(1)$-term it is possible to derive a result similar to Theorem 1 for the average number of block occurrences, too. Since the method is very similar to Section 2, we will not present it here.

REFERENCES

1. R. BELLMAN AND H. N. SHAPIRO, On a problem in additive number theory, *Ann. Math. Princeton II* **49** (1948), 333–340.
2. L. E. BUSH, An asymptotic formula for the average sum of the digits of integers, *Amer. Math. Monthly* **47** (1940), 154–156.
3. H. DELANGE, Sur la fonction sommatoire de la fonction "somme de chiffres," *Enseign. Math.* **21** (1975), 31–47.
4. P. KIRSCHENHOFER, Subblock occurrences in the q-ary representation of n, *SIAM J. Alg. Disc. Meth.* **4** (1983), 231–236.
5. L. MIRSKY, A theorem on representations of integers in the scale of r, *Scripta Math. New York* **15** (1949), 11–12.
6. H. PRODINGER, Generalizing the "sum of digits" function, *SIAM J. Alg. Disc. Meth.* **3** (1982), 35–42.

Printed in Belgium