SIAM J. ALG. DISC. METH.
Vol. 3, No. 1, March 1982

0196-5212/82/0301-0004 $01.00/0

GENERALIZING THE SUM OF DIGITS FUNCTION*

HELMUT PRODINGER†

Abstract. The number theoretic function $G_{q,\alpha}(n)=\sum_{k\geq 1}\sum_{j=0}^{q-1}\lfloor n/q^k+j\alpha\rfloor$ has appeared in the literature for some special values of α. Some properties of this function are investigated. Since $G_{q,0}(n)$ is closely related to the sum of digits of the q-ary representation of n, a generalized "sum of digits" function can be defined via $G_{q,\alpha}$. For $q=2$ and $\alpha=2^{-s}$ the summing function of this "sum of digits" function is analyzed using a technique of Delange.

1. Introduction and elementary results. Let $q\in\mathbb{N}$, $q\neq 1$ and define the functions $G_{q,\alpha}:\mathbb{N}_0\to\mathbb{N}_0$ by

$$G_{q,\alpha}(n):=\sum_{k\geq 1}\sum_{1\leq j<q}\left\lfloor\frac{n}{q^k}+j\alpha\right\rfloor. \tag{1}$$

($\lfloor x\rfloor$ denotes the greatest integer less or equal to x.)

To make this definition meaningful, α must be in the range $\alpha\in[0,(q-1)^{-1})$. But for all considerations (except for Theorem 5) it is better to restrict α to the range $[0,q^{-1}]$, especially because the generalized "sum of digits" function (see § 2) takes then only nonnegative values, which is very desirable.

In [6] an alternative expression for $G_{2,1/4}$ is given by a complicated method; the same method applies to $G_{2,1/2}$ showing that this function is the identity.

The last result can be found in [4, p. 43] in the general form

$$G_{q,1/q}(n)=\sum_{k\geq 1}\sum_{1\leq j<q}\left\lfloor\frac{n}{q^k}+\frac{j}{q}\right\rfloor=n. \tag{2}$$

In the sequel it will be shown that, starting from (2), some formulas for $G_{q,\alpha}$ can be derived in an easy way. To be able to formulate this result adequately, it is useful to use the following denotation.

If ξ is a string of integers in the range $[0,q-1]$, let $B_q(\xi,n)$ denote the number of subblocks ξ in the q-ary representation of n (subblocks are allowed to overlap).

THEOREM 1.

$$G_{q,q^{-s}}(n)=n-\sum_{1\leq j<q}jB_q(j,n)+\sum_{1\leq j<q}jB_q((q-1)^{s-1}j,n).$$

(For instance, $G_{2,1/4}(n)=n-B_2(1,n)+B_2(11,n)$.)

Proof. It is sufficient to show that the number of indices k, l such that

$$\left\lfloor\frac{n}{q^k}+\frac{l}{q}\right\rfloor=1+\left\lfloor\frac{n}{q^k}+\frac{l}{q^s}\right\rfloor$$

equals

$$\sum_{1\leq j<q}jB_q(j,n)-\sum_{1\leq t<q}tB_q((q-1)^{s-1}t,n).$$

* Received by the editors May 29, 1980, and in revised form June 1, 1981.
† Institut für Algebra und Diskrete Mathematik, Technische Universität Wien, Guszhausstrasze 27-29, A-1040 Vienna, Austria.

This can only happen if

$$\left\lfloor \frac{n}{q^k} + \frac{l}{q} \right\rfloor = 1 + \left\lfloor \frac{n}{q^k} \right\rfloor \quad \text{and} \quad s \geqq 2.$$

(For $s = 1$ the theorem is trivially fulfilled.) Now assume that the fractional part of n/q^k starts with the digit j, $1 \leqq j < q - 1$ (with respect to the q-ary representation of n). Then in (1) each l with $j + l \geqq q$ is possible (there are j of such l's), and furthermore

$$\left\lfloor \frac{n}{q^k} + \frac{l}{q^s} \right\rfloor = \left\lfloor \frac{n}{q^k} \right\rfloor.$$

Now assume $j = q - 1$. Then there are again $q - 1$ indices in (1) such that $j + l \geqq q$, but

$$\left\lfloor \frac{n}{q^k} + \frac{l}{q^s} \right\rfloor = 1 + \left\lfloor \frac{n}{q^k} \right\rfloor$$

is also possible, and this happens if and only if the fractional part of n/q^k starts with $q-1, q-1, \cdots, q-1, t$; in (1) each l with $t + l \geqq q$ is possible (there are t of such l's).

Since the sum over k in (1) means that every digit is exactly one time the leading digit of the fractional part of n/q^k, the proof is finished.

Remark that the formula holds also for $\alpha = 0$ where the second sum vanishes, which can be seen as a "limiting case".

In the sequel it will be shown that $G_{q,\alpha}$ for $0 \leqq \alpha < q^{-1}$ has a rather erratic behavior, which contrasts to the case $\alpha = q^{-1}$.

LEMMA 2. *Assume $a \neq q^{-1}$. Then there exists an n such that*

$$G_{q,\alpha}(n) = G_{q,\alpha}(n + 1).$$

Since the proof of this lemma is rather long and not too interesting, we just indicate that an appropriate choice for n is (with respect to the q-ary representation) of the form $(1000 \cdots 0)_q$.

THEOREM 3. *For $0 \leqq \alpha < q^{-1}$ the function $G_{q,\alpha}$ is not surjective.*

Proof. By Theorem 1,

$$q^t = G_{q,1/q}(q^t) \geqq G_{q,\alpha}(q^t) \geqq G_{q,0}(q^t) = q^t - 1.$$

Thus there are numbers $t_1 < t_2$ such that

$$G_{q,\alpha}(q^{t_2}) - G_{q,\alpha}(q^{t_1}) = q^{t_2} - q^{t_1}.$$

Because of the monotony of $G_{q,\alpha}$, surjectivity in the interval $[t_1, t_2]$ means also injectivity, but this property is not fulfilled.

Remark. If α is allowed to be in the range $\alpha \in [0, (q-1)^{-1})$, $\alpha \neq q - 1$ (compare the comments after the definition of $G_{q,\alpha}$), Lemma 2 and Theorem 3 are still true.

It is clear that from $\alpha \leqq \beta$ it follows that $G_{q,\alpha}(n) \leqq G_{q,\beta}(n)$. The following stronger result is easily obtained.

THEOREM 4. *If $\alpha < \beta$ then there is an n such that $G_{q,\alpha}(n) < G_{q,\beta}(n)$.*

Proof. Choose numbers n, k such that

$$1 - \beta \leqq \frac{n}{q^k} < 1 - \alpha;$$

then

$$\left\lfloor \frac{n}{q^k} + \beta \right\rfloor = 1 \quad \text{and} \quad \left\lfloor \frac{n}{q^k} + \alpha \right\rfloor = 0.$$

As D. E. Knuth has pointed out [5], it would be interesting to investigate $G_{q,\alpha}(n)$ for fixed n, where α is the variable. A first result in this direction is the following theorem.

THEOREM 5.

$$\int_0^1 G_{2,\alpha}(n)\, d\alpha = n.$$

Proof. Since

$$\int_0^1 \lfloor x+\alpha \rfloor\, d\alpha = x,$$

it follows that

$$\begin{aligned}\int_0^1 G_{2,\alpha}(n)\, d\alpha &= \int_0^1 \sum_{k\geqq 1} \left\lfloor \frac{n}{2^k}+\alpha \right\rfloor d\alpha \\ &= \sum_{k\geqq 1} \int_0^1 \left\lfloor \frac{n}{2^k}+\alpha \right\rfloor d\alpha = \sum_{k\geqq 1} \frac{n}{2^k} = n.\end{aligned}$$

(It is not very hard to see that the integration and the summation can be interchanged.)

2. The summing function of the function "generalized sum of digits". In Delange [1] the summing function of the function "sum of digits to the base q" is considered: The sum of digits is

$$\begin{aligned}S_q(n) &= \sum_{r=0}^{\infty} \left(\left\lfloor \frac{n}{q^r} \right\rfloor - q \left\lfloor \frac{n}{q^{r+1}} \right\rfloor \right) \\ &= n-(q-1) \sum_{r=1}^{\infty} \left\lfloor \frac{n}{q^r} \right\rfloor \\ &= n - \sum_{r=0}^{\infty} \sum_{j=1}^{q-1} \left\lfloor \frac{n}{q^{r+1}} \right\rfloor \\ &= n - \sum_{1\leqq j<q} jB_q(j,n) = n - G_{q,0}(n).\end{aligned} \tag{3}$$

In view of § 1 it is natural to define the generalized function "sum of digits" by

$$S_{q,\alpha}(n) := n - G_{q,\alpha}(n).$$

In [1] it is proved that

$$\frac{1}{m} \sum_{n=0}^{m-1} S_q(n) = \frac{q-1}{2} \log_q m + F(\log_q m),$$

where $F(x)$ is continuous, periodic with period 1 and thus bounded. ($\log_q m$ means the logarithm to the base q.) Further information in this area can be found in the beautiful thesis of Flajolet [2].

In the rest of this paper the summing function of $S_{2,2^{-s}}(n)$ is treated, but I hope to do further work in this direction in the future. The ordinary sum of digits appears as the limit for $s \to \infty$.

From Theorem 1, we know that

$$G_{2,2^{-s}}(n) = n - B_2(1,n) + B_2(1^s, n).$$

Plugging this into the definition of $S_{q,\alpha}$, we find that

$$S_{2,2^{-s}}(n) = B_2(1, n) - B_2(1^s, n) \tag{4}$$
$$= S_2(n) - B_2(1^s, n).$$

So $S_{2,2^{-s}}(n)$ is just the number of ones in the binary expansion of n minus the number of blocks of s consecutive ones in that expansion.

The rest of this paper is an analysis of the summing function of the function $B_2(1^s, n)$; then by (4), an analogue to Delange's result of $S_{2,2^{-s}}(n)$ can be formulated as a corollary.

THEOREM 6. *Let $B_2(1^s, n)$ denote the number of subblocks of s consecutive ones appearing in the binary representation of n, where overlapping is allowed. Then the summation of $B_2(1^s, n)$ is given by the formula*

$$\frac{1}{m}\sum_{n=0}^{m-1} B_2(1^s, n) = \frac{\log_2 m - (s-1)}{2^s} + H_s(\log_2 m) + \frac{E}{m},$$

where H_s is continuous, periodic with period 1, *and satisfies $H_s(0) = 0$, and where E is bounded by $0 \leqq E < 1$.*

Proof. A crucial point in Delange's derivation is the property

$$\left\lfloor \frac{t}{q^r} \right\rfloor = \left\lfloor \frac{n}{q^r} \right\rfloor \quad \text{for } n \leqq t < n+1. \tag{5}$$

For $\alpha \neq 0$ it is not trivial to find an appropriate analogue.

An analogue to property (5) can be written as follows:

$$\left\lfloor \frac{t}{2^r} + \frac{1}{2^s} \right\rfloor = \left\lfloor \frac{n}{2^r} + \frac{1}{2^s} \right\rfloor \tag{6}$$

holds for $n \leqq t < n+1$ and $r \geqq s$ and also for $n - (1/2^{s-r}) \leqq t < n + 1 - (1/2^{s-r})$ and $r < s$.

Let $l = \log_2 m$. We have

$$\begin{aligned}
\sum_{n=0}^{m-1} B_2(1^s, n) &= \sum_{n=0}^{m-1} G_{2,2^{-s}}(n) - \sum_{n=0}^{m-1} G_{2,0}(n) \\
&= \sum_{r=1}^{s-1} \int_{2^{r-s}}^{m-2^{r-s}} \left\lfloor \frac{t}{2^r} + \frac{1}{2^s} \right\rfloor dt + \sum_{r=s}^{\lfloor l \rfloor + 1} \int_0^m \left\lfloor \frac{t}{2^r} + \frac{1}{2^s} \right\rfloor dt - \sum_{r=0}^{\lfloor l \rfloor} \int_0^m \left\lfloor \frac{t}{2^{r+1}} \right\rfloor dt \\
&= -\sum_{r=1}^{s-1} 2^{r-s} \left\lfloor \frac{m}{2^r} + \frac{1}{2^s} \right\rfloor + \sum_{r=0}^{\lfloor l \rfloor} \int_0^m \left(\left\lfloor \frac{t}{2^{r+1}} + \frac{1}{2^s} \right\rfloor - \left\lfloor \frac{t}{2^{r+1}} \right\rfloor \right) dt.
\end{aligned}$$

Now define

$$C = \sum_{r=1}^{s-1} 2^{r-s} \left\lfloor \frac{m}{2^r} + \frac{1}{2^s} \right\rfloor,$$

$$g_s(x) = \int_0^x \left(\left\lfloor t + \frac{1}{2^s} \right\rfloor - \lfloor t \rfloor - \frac{1}{2^s} \right) dt.$$

$g_s(x)$ is periodic with period 1, continuous and $g_s(0) = 0$. With this notation we can write

$$\begin{aligned}\sum_{n=0}^{m-1} B_2(1^s, n) &= \sum_{r=0}^{\lfloor l \rfloor} \int_0^m \left(\left\lfloor \frac{t}{2^{r+1}} + \frac{1}{2^s} \right\rfloor - \left\lfloor \frac{t}{2^{r+1}} \right\rfloor - \frac{1}{2^s} \right) dt + (\lfloor l \rfloor + 1) \frac{m}{2^s} - C \\ &= \sum_{r=0}^{\lfloor l \rfloor} 2^{r+1} g_s\left(\frac{m}{2^{r+1}}\right) + (\lfloor l \rfloor + 1) \frac{m}{2^s} - C \\ &= \sum_{r=-\infty}^{\lfloor l \rfloor} 2^{r+1} g_s\left(\frac{m}{2^{r+1}}\right) + (\lfloor l \rfloor + 1) \frac{m}{2^s} - C \\ &= \sum_{k=0}^{\infty} 2^{1+\lfloor l \rfloor - k} g_s(m 2^{k-\lfloor l \rfloor - 1}) + (\lfloor l \rfloor + 1) \frac{m}{2^s} - C.\end{aligned}$$

Now remember that $m = 2^l$ and define $\{l\} = l - \lfloor l \rfloor$ and

$$h_s(x) = \sum_{k \geqq 0} 2^{-k} g_s(x 2^k).$$

Then

$$\sum_{n=0}^{m-1} B_2(1^s, n) = m 2^{1-\{l\}} \cdot h_s(2^{\{l\}-1}) + (\lfloor l \rfloor + 1) \frac{m}{2^s} - C.$$

Now defining

$$H_s(l) = 2^{1-\{l\}} h_s(2^{\{l\}-1}) - \frac{1}{2^s} (\{l\} - 1),$$

it remains only to analyze the quantity C to complete the proof.

$$C = \sum_{r=1}^{s-1} 2^{r-s} \left\lfloor \frac{m}{2^r} + \frac{1}{2^s} \right\rfloor = \sum_{r=1}^{s-1} 2^{r-s} \left\lfloor \frac{m}{2^r} \right\rfloor,$$

since r lies in the range $1 \leqq r \leqq s - 1$. Thus

$$C = \sum_{r=1}^{s-1} 2^{r-s} \left(\frac{m}{2^r}\right) - \sum_{r=1}^{s-1} 2^{r-s} \left\{\frac{m}{2^r}\right\} = m \frac{s-1}{2^s} - E.$$

Since $\{x\}$ lies always in the interval $[0, 1)$, we can deduce that the remaining error term E must also lie in that interval.

Using Delange's result on the summing function of $S_2(n)$ from [1] we get immediately:

COROLLARY 7. *If $S_{q,\alpha}(n)$ denotes the generalized "sum of digits" function defined above, then the summing function of the quantity $S_{2,2^{-s}}(n)$ is given by*

$$\frac{1}{m} \sum_{n=0}^{m-1} S_{2,2^{-s}}(n) = \left(\frac{1}{2} - \frac{1}{2^s}\right) \log_2 m + \frac{s-1}{2^s} + F(\log_2 m) - H_s(\log_2 m) - \frac{E}{m},$$

where both F and H_s are continuous, periodic with period 1, and take the value 0 on the integers, and where E is bounded by $0 \leqq E < 1$.

3. The Fourier series for $H_s(x)$. Delange [1] has already determined the Fourier series for $F(x)$. Similar methods apply to $H_s(x)$.

THEOREM 8. *The Fourier expansion* $H_s(x)=\sum_k h_k e^{2k\pi ix}$ *of the function* $H_s(x)$ *converges absolutely, and its coefficients are given by*

$$h_0=\log_2 \Gamma\left(1-\frac{1}{2^s}\right)-\frac{1}{2^s \log 2}-\frac{1}{2^{s+1}},$$

$$h_k=\frac{\zeta\left(\frac{2k\pi i}{\log 2}, 1-\frac{1}{2^s}\right)-\zeta\left(\frac{2k\pi i}{\log 2}\right)}{2k\pi i\left(1+\frac{2k\pi i}{\log 2}\right)} \quad \text{for } k \neq 0.$$

Proof. Let $0 \leqq x<1$. Since

$$H_s(x)=2^{1-x} h_s(2^{x-1})+\frac{1-x}{2^s},$$

the determination of the Fourier coefficients decomposes as:

$$h_k=\int_0^1 2^{1-x} h_s(2^{x-1})\, e^{-2k\pi ix}\, dx+\frac{1}{2^s} \int_0^1(1-x)\, e^{-2k\pi ix}\, dx=a_k+b_k.$$

It is easily seen that

$$b_k=\frac{1}{2^s} \cdot \frac{1}{2k\pi i} \quad \text{for } k \neq 0, \quad b_0=\frac{1}{2^s} \cdot \frac{1}{2},$$

$$a_k=\int_0^1 2^{1-x} \sum_{r=0}^{\infty} 2^{-r} g_s(2^{r+x-1})\, e^{-2k\pi ix}\, dx,$$

and as in [1], the integration and the summation can be interchanged:

$$a_k=\sum_{r=0}^{\infty} \int_0^1 2^{1-r-x} g_s(2^{r+x-1})\, e^{-2k\pi ix}\, dx.$$

The change of variable $x=1-r+\log_2 u$ gives

$$\int_0^1 2^{1-r-x} g_s(2^{r+x-1})\, e^{-2k\pi ix}\, dx=\frac{1}{\log 2} \int_{2^{r-1}}^{2^r} \frac{g_s(u)}{u^2} \exp(-2k\pi i \cdot \log_2 u)\, du.$$

Thus

$$a_k=\frac{1}{\log 2} \int_{1/2}^{\infty} \frac{g_s(u)}{u^{2+2k\pi i/\log 2}}\, du.$$

As in [1], the integral

$$\Phi_s(z)=\int_{1/2}^{\infty} \frac{g_s(u)}{u^{z+1}}\, du$$

should be studied; then

$$a_k=\frac{1}{\log 2} \Phi_s\left(1+\frac{2k\pi i}{\log 2}\right).$$

Since

$$g_s(u)=\int_0^u\left(\left\lfloor t+\frac{1}{2^s}\right\rfloor-\lfloor t\rfloor-\frac{1}{2^s}\right) dt,$$

by partial integration for Re $z > 0$,

$$\Phi_s(z) = -\frac{1}{2^s} \cdot \frac{2^{z-1}}{z} + \frac{1}{z} \int_{1/2}^{\infty} \left(\left\lfloor u + \frac{1}{2^s} \right\rfloor - \lfloor u \rfloor - \frac{1}{2^s} \right) \frac{du}{u^z}.$$

For Re $z > 2$, the integral can be split into three parts. The third one is

$$-\frac{1}{2^s} \cdot \frac{1}{z} \int_{1/2}^{\infty} \frac{du}{u^z} = -\frac{1}{2^s} \cdot \frac{1}{z} \cdot \frac{2^{z-1}}{z-1}.$$

The second one is

$$-\frac{1}{z} \int_{1/2}^{\infty} \lfloor u \rfloor \frac{du}{u^z} = -\frac{1}{z} \cdot \frac{1}{z-1} \cdot \zeta(z-1).$$

The first one is

$$\frac{1}{z} \int_{1/2}^{\infty} \left\lfloor u + \frac{1}{2^s} \right\rfloor \frac{du}{u^z} = \frac{1}{z} \int_{1-2^{-s}}^{\infty} \left\lfloor u + \frac{1}{2^s} \right\rfloor \frac{du}{u^z} = \frac{1}{z(z-1)} \zeta\left(z-1, 1-\frac{1}{2^s}\right),$$

where $\zeta(z-1, a)$ is the z-function of Hurwitz (see [3]). This gives

$$\Phi_s(z) = -\frac{1}{2^s} \cdot \frac{2^{z-1}}{z-1} + \frac{\zeta\left(z-1, 1-\frac{1}{2^s}\right) - \zeta(z-1)}{z(z-1)}.$$

This holds for Re $z > 0$ by analytical continuation. This gives

$$a_k = -\frac{1}{2^s} \cdot \frac{1}{2k\pi i} + \frac{1}{2k\pi i} \left(1 + \frac{2k\pi i}{\log 2}\right)^{-1} \left(\zeta\left(\frac{2k\pi i}{\log 2}, 1 - \frac{1}{2^s}\right) - \zeta\left(\frac{2k\pi i}{\log 2}\right) \right)$$

for $k \neq 0$. Now a_0 must be computed. From [7],

$$\zeta(z-1, a) = \tfrac{1}{2} - a + (z-1)(\log \Gamma(a) - \tfrac{1}{2} \log(2\pi)) + O((z-1)^2) \quad \text{for } z \to 1.$$

Thus

$$\zeta\left(z-1, 1-\frac{1}{2^s}\right) = -\frac{1}{2} + \frac{1}{2^s} + (z-1)\left(\log \Gamma\left(1 - \frac{1}{2^s}\right) - \frac{1}{2} \log(2\pi)\right) + O((z-1)^2),$$

$$\zeta(z-1) = \zeta(z-1, 1) = -\tfrac{1}{2} - (z-1)\tfrac{1}{2} \log(2\pi) + O((z-1)^2),$$

$$2^{z-1} = 1 + (\log 2)(z-1) + O((z-1)^2),$$

$$\frac{1}{z} = 1 - (z-1) + O((z-1)^2).$$

This yields after some manipulations

$$\Phi_s(z) = -\frac{1}{2^s}(1 + \log 2) + \log \Gamma\left(1 - \frac{1}{2^s}\right) + O((z-1)^2) \quad \text{for } z \to 1.$$

Hence

$$a_0 = -\frac{1}{2^s \log 2} - \frac{1}{2^s} + \log_2 \Gamma\left(1 - \frac{1}{2^s}\right).$$

Finally, since $\zeta(it, a) = O(|t|^{1/2} \log |t|)$ [7], the Fourier series of H_s will converge absolutely.

Acknowledgment. The author would like to express his deep gratitude to an anonymous referee who went very carefully through the manuscript; he pointed out several errors and made numerous suggestions which increase both the readability and the contents of this paper.

REFERENCES

[1] H. DELANGE, *Sur la fonction sommatoire de la fonction somme des chiffres*, l'Enseignement Mathématique, 21 (1975), pp. 31–47.

[2] P. FLAJOLET, *Analyse d'algorithmes de manipulation d'arbres et de fichiers*, Thèse, Université Paris-Sud, 1979.

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

[4] D. E. KNUTH, *The Art of Computer Programming*, vol. 1, Addison-Wesley, Reading, MA, 1972.

[5] ———, personal communication.

[6] H. PRODINGER AND F. J. URBANEK, *Infinite* 0-1-*sequences without long adjacent identical blocks*, Disc. Math., 28 (1979), pp. 277–289.

[7] E. T. WHITTAKER AND G. N. WATSON, *A Course of Modern Analysis*, Cambridge University Press, Cambridge, 1927.