

ON THE MOMENTS OF THE SUM-OF-DIGITS FUNCTION

P.J. Grabner†, P. Kirschenhofer, H. Prodinger and R.F. Tichy†

1. INTRODUCTION

In a recent paper in *The Fibonacci Quarterly* R.E. Kennedy and C.N. Cooper [12] dealt with the second moment of the sum-of-digits function in the decimal number system, explicitly stating the question for higher moments as an open problem. We consider this question in the sequel.

Let $v(n)$ represent the total number of 1-digits in the binary representation of the integer n . It is not hard to see that

$$M_1(N) = \sum_{n < N} v(n) = \frac{1}{2}N \log_2 N + o(N \log N), \quad (1.1)$$

since asymptotically the binary representations contain roughly as many 0's as 1's. The Trollope-Delange formula is more surprising: It expresses $S(n)$ in an exact formula (cf. [21], [5])

$$M_1(n) = \frac{1}{2}N \log_2 N + N \delta_1(\log_2 N), \quad (1.2)$$

where $\delta_1(u)$, a 'fractal function', is a continuous, periodic, nowhere differentiable function with an explicit Fourier expansion involving the Riemann zeta function.

The argument given by Delange is based on a combinatorial decomposition of binary representations of integers, followed by a computation of the Fourier coefficients of the fractal

†These authors were supported by the Austrian Science Foundation Project Nr. P8274-PHY.

function. This exact formula was used to analyze register allocation algorithms, or equivalently the order of random 'channel networks' (cf.[9]). It was later extended to some non-standard digital representations of integers. Cantor representations were considered in [17], and Gray code [8] was studied for the purpose of analyzing sorting networks. In [13] occurrences of blocks of digits in standard q -ary representations, in [15] subblock occurrences in Gray code representation and in [3] the subsequence $u(3n)$ of the classical Thue-Morse sequence $u(n)$ were investigated. In a recent paper [11] this result was extended to digit expansions with respect to linear recurrences. As general surveys we refer to Stolarsky's article [20] and to [16]. As an especially important paper in this area we mention Brillhart, Erdős and Morton's paper [2].

J. Coquet [4] studied in detail higher moments of the binary sum-of-digits function. Later in [14] the second moment $M_2(N)$ (and the corresponding centered moment) was investigated in detail and the following exact formula was proved by Delange's method:

$$M_2(N) = \sum_{n < N} v(n)^2 = \frac{N}{4} \log_2^2 N + N \log_2 N \left(\frac{1}{4} + \delta_1(\log_2 N) \right) + N \delta_2(\log_2 N), \quad (1.3)$$

where δ_1, δ_2 are continuous nowhere differentiable functions of period 1 and δ_1 is the function that occurred in (1.2). This formula can already be found in [4] and Coquet could also establish similar formulae for higher moments but without proving the continuity of the remainder terms. By the same approach such exact formulae can be proved for arbitrary q -ary number systems.

$$\sum_{n < N} v_q(n)^2 = \left(\frac{q-1}{2} \right)^2 N \log_q^2 N + N \log_q N \eta_1(\log_q N) + N \eta_2(\log_q N), \quad (1.4)$$

where η_1 and η_2 are continuous nowhere differentiable functions of period 1 and \log_q denotes the logarithm to base q , as usual. In [12] R.E. Kennedy and C.N. Cooper rediscovered a weaker result for the decimal system:

$$\sum_{n < N} v_{10}(n)^2 = 20.25 N \log_{10} N + o(N \log N).$$

In Section 2 we use Delange's approach to compute the third moment in the binary number system, the q -ary case can be settled by similar but more elaborate computations. In Section 3 we sketch this approach in the case of higher moments¹. In the final Section 4 we present an exposition of a purely analytic approach using Mellin transform. In [7] this approach was applied to establish exact and asymptotical formulae in various cases of digital sums. By this

¹While finishing the present paper, the authors became aware of a manuscript due to J.-M. Dumont and A. Thomas where the corresponding result is proved for digital systems generated by substitution automata. This more general result is proved by difficult and less elementary methods.

method we get automatically the mean and the Fourier-coefficients of the periodic functions. In detail we apply this method to the expectation (first moment) of the q -ary sum-of-digits function. In the case of higher moments some analytic problems remain to be solved.

2. DELANGE'S APPROACH: THE THIRD MOMENT

In order to compute the third moment M_3 of the binary sum-of-digits function we introduce the following notation:

$$f_j(t) = \left\lfloor \frac{t}{s^j} \right\rfloor - 2 \left\lfloor \frac{t}{2^{j+1}} \right\rfloor, \quad g_j(t) = f_j(t) - \frac{1}{2}.$$

Thus we have

$$\begin{aligned} M_3(N) &= \int_0^N \sum_{(j_1, j_2, j_3)} f_{j_1}(t) f_{j_2}(t) f_{j_3}(t) dt \\ &= \sum_{k=0}^3 \binom{3}{k} \frac{1}{2^{3-k}} \int_0^N \sum_{(j_1, \dots, j_k)} g_{j_1}(t) \dots g_{j_k}(t) dt, \end{aligned}$$

where the summation is extended over all k -tuples (j_1, \dots, j_k) of non negative integers $\leq L = \lfloor \log_2 N \rfloor$. Denoting the integral for $k = 3$ by I_3 we obtain

$$\begin{aligned} M_3(N) &= I_3 + \frac{3}{2}(L+1)M_2(N) - \frac{3}{4}(L+1)^2M_1(N) + \frac{N}{8}(L+1)^3 \\ &= I_3 + N \left(\frac{1}{8}K^3 + K^2 \left(\frac{3}{8} + \frac{3}{4}\delta_1 \right) + K \left(\frac{3}{2}\delta_2 + \frac{3}{8}y \right) + \left(\frac{3}{2}y\delta_2 - \frac{3}{4}y^2\delta_1 + \frac{1}{8}y^3 \right) \right), \end{aligned} \tag{2.1}$$

where $\delta_1 = \delta_1(\log_2 N)$ and $\delta_2 = \delta_2(\log_2 N)$ are the periodic fluctuations of the first and second moments occurring in (1.2) and (1.3); $K = \log_2 N$ and $y = 1 - \{\log_2 N\}$. We note that $\delta_1(0) = \delta_1(1) = \delta_2(0) = \delta_2(1) = 0$.

Now we compute the remaining integral

$$\begin{aligned} I_3 &= \left(\sum_{j_1=j_2=j_3} + 3 \sum_{j_1=j_2 \neq j_3} + 6 \sum_{j_1 < j_2 < j_3} \right) \int_0^N g_{j_1}(t) g_{j_2}(t) g_{j_3}(t) dt \\ &= \frac{1}{4} \sum_j \int_0^N g_j(t) dt + \frac{3}{4}L \sum_j \int_0^N g_j(t) dt + 6 \sum_{j_1 < j_2 < j_3} \int_0^N g_{j_1}(t) g_{j_2}(t) g_{j_3}(t) dt. \end{aligned}$$

Notice that the first two parts of this equality follow from the fact that $g_j(t)$ is piecewise constant and $g_j(t)^2 = \frac{1}{4}$. For the last summation we obtain

$$\sum_{j_1 < j_2 < j_3} = \sum_{0 \leq j_1 < j_2 < j_3 \leq L} 2^{j_3+1} \int_0^N g(2^{j_3-j_1}u) g(2^{j_3-j_2}u) g(u) du,$$

where $g(u) = \lfloor 2u \rfloor - 2\lfloor u \rfloor - \frac{1}{2}$. Now transform the index of summation as follows:

$j: = j_3, d_1: = j_3 - j_1, d_2: = j_3 - j_2$. Thus we obtain

$$0 \leq j_1 < j_2 < j_3 \leq L = \sum_{j=0}^L 2^{j+1} \sum_{1 \leq d_1 < d_2 \leq j} \int_0^{\frac{N}{2^{j+1}}} g(2^{d_1}u)g(2^{d_2}u)g(u)du.$$

Notice that the inner summation can be extended to infinity without changing the value and set $k = L - j$. Hence

$$\begin{aligned} 0 \leq j_1 < j_2 < j_3 \leq L &= 2^{L+1} \sum_{k \geq 0} \sum_{1 \leq d_1 < d_2} \int_0^{2^{k+(\log_2 N)-1}} g(u)g(2^{d_1}u)g(2^{d_2}u)du \\ &= 2^{1-(\log_2 N)} \sum_{k \geq 0} 2^{-k} h_3(2^{\{\log_2 N\}-1} 2^k), \end{aligned}$$

where h_3 is the periodic function defined by

$$h_3(x) = \sum_{1 \leq d_1 < d_2} \int_0^x g(u)g(2^{d_1}u)g(2^{d_2}u)du. \tag{2.2}$$

We observe that the integrals occurring in (2.2) are integrals over distinct Rademacher functions; hence $h_3(1) = h_3(\frac{1}{2}) = 0$. Thus setting

$$\psi(x) = 2^{2-\{x\}} \sum_{k \geq 0} 2^{-k} h_3(2^k 2^{\{x\}-1}) \tag{2.3}$$

we obtain a continuous and 1-periodic function with $\psi(0) = \psi(1) = 0$. Therefore we have

$$I_3 = N \left(K \frac{3}{4} \delta_1 - \frac{3}{8} y \right) + \left(-\frac{1}{2} \delta_1 + \frac{x}{4} + \frac{3}{4} x \delta_1 - \frac{3}{8} x^2 + 6\psi(1-y) \right). \tag{2.4}$$

By combining (2.1) and (2.4) we derive

$$M_3(N) = N \left(\frac{1}{8} K^3 + K^2 \left(\frac{3}{8} + \frac{3}{4} \delta_1 \right) + K \left(\frac{3}{4} \delta_1 + \frac{3}{2} \delta_2 \right) + \delta_3 \right), \tag{2.5}$$

where

$$\delta_3 = \delta_3(\log_2 N) = \frac{3}{2} y \delta_2 - \frac{3}{4} y^2 \delta_1 + \frac{1}{8} y^3 - \frac{1}{2} \delta_1 + \frac{y}{4} + \frac{3}{4} y \delta_1 - \frac{3}{8} y^2 + 6\psi(1-y). \tag{2.6}$$

Thus we have proved

Theorem 1: The third moment of the binary sum-of-digits function satisfies

$$\begin{aligned} \frac{1}{N} \sum_{n < N} v_2(n)^3 &= \frac{1}{8} (\log_2 N)^3 + \left(\frac{3}{8} + \frac{3}{4} \delta_1(\log_2 N) \right) (\log_2 N)^2 \\ &\quad + \left(\frac{3}{4} \delta_1(\log_2 N) + \frac{3}{2} \delta_2(\log_2 N) \right) \log_2 N + \delta_3(\log_2 N), \end{aligned}$$

where δ_3 is a continuous and periodic function with period 1 defined by (2.6).

Remark: Obviously this result can be extended to q -ary representations. However, the computations are much more involved.

3. DELANGE'S APPROACH: THE GENERAL CASE

In the following we sketch the computation of the s -th moments M_s of the binary sum-of-digits function:

Adopting analogous notations as in the previous section we start from

$$M_s(N) = \int_0^N \sum_{(j_1, \dots, j_s)} f_{j_1}(t) \dots f_{j_s}(t) dt$$

$$= \sum_{k=0}^s \binom{s}{k} \frac{1}{2^{s-k}} \int_0^N \sum_{(j_1, \dots, j_k)} g_{j_1}(t) \dots g_{j_k}(t) dt.$$

In the following we evaluate the integral I_k appearing in the last expression. To this end we observe that $g_j^2(t) \equiv \frac{1}{4}$, so that it is meaningful to group together multiple occurrences of g 's with the same index j and to distinguish odd and even and even frequencies since $g_j^{2n}(t) \equiv \frac{1}{2^{2n}}$, while $g_j^{2n+1}(t) \equiv \frac{1}{2^{2n}} g_j(t)$.

The above described grouping of g 's with the same index induces a set partition of the subindex set $\{1, \dots, k\}$, where we have to count separately the numbers p resp. r of blocks of odd resp. even cardinality. Let us denote by

$$S(k; p, r)$$

the number of the above described set partitions. Then we may compute I_k as follows: (We use the abbreviation $(x)_k = x(x-1) \dots (x-k+1)$.)

$$I_k = \sum_{p,r} \frac{p!}{2^{k-p}} S(k; p, r) (L+1-p)_r \int_0^N \sum_{0 \leq i_1 < \dots < i_p \leq L} g_{i_1}(t) \dots g_{i_p}(t) dt.$$

(Observe that for given index values i_1, \dots, i_p belonging to the p blocks of odd cardinality, there are $(l+1-p)_r$ possibilities to choose the r index values belonging to the blocks of even cardinality differently from i_1, \dots, i_p .)

The sums

$$\Sigma_p(N) = \sum_{0 \leq i_1 < \dots < i_p \leq L} \int_0^N g_{i_1}(t) \dots g_{i_p}(t) dt$$

can be treated in exactly the same manner as in Section 2 for the instance $p = 3$ and yield

$$\frac{1}{N} \Sigma_p(N) = \tau_p(\log_2 N),$$

where

$$\tau_p(x) = 2^{1-\{x\}} \sum_{i \geq 0} 2^{-ih_p} (2^{i2^{-1} + \{x\}})$$

with

$$h_p(x) = \sum_{1 \leq i_1 < i_2 < \dots < i_{p-1}} \int_0^x g(u)g(u2^{i_1}) \dots g(u2^{i_{p-1}}) du.$$

Note that $h_p(x)$ is a continuous function with $h_p(m) = 0$ for all $m \in \mathbb{Z}$, so that $\tau_p(x)$ is a periodic function of period 1.

Collecting all contributions and rewriting $L = \lfloor \log_2 N \rfloor$ as $L = \log_2 N - \{\log_2 N\}$ we obtain

Theorem 2: The s -th moment of the binary sum-of-digits function satisfies

$$\frac{1}{N} \sum_{n < N} v_2(n)^s = \frac{1}{2^s} (\log_2 N)^s + \sum_{i=0}^{s-1} (\log_2 N)^i \eta_i(\log_2 N)$$

where $\eta_i(x)$ are continuous periodic functions of period 1 that can be expressed explicitly as described before.

Remarks: 1) the numbers $S(k; p, r)$ do not seem to have been studied in the literature. However similar numbers satisfying interesting recurring relations occur already in [4]. Of course the numbers $S(k; p, r)$ are related to the classical Stirling numbers of the second kind [19] via

$$\sum_{p+r=j} S(k; p, r) = S(k, j).$$

The trivariate generating function of the $S(k; p, r)$ is easily obtained via the operator method for combinatorial constructions as described systematically in [10]:

$$\sum_{p, r, k} S(k; p, r) x^p y^r \frac{z^k}{k!} = \exp(x \sinh z + y(\cosh z - 1)).$$

2) From Theorem 2 it is not difficult to compute expressions for the higher centralized moments.

4. MELLIN PERRON APPROACH: AN EXPOSITION

As announced in the introduction we present an analytic method to compute digital sums. The method makes use of the so called Mellin Perron summation formula:

$$\sum_{n < N} a_n \left(1 - \frac{n}{N}\right) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \sum_{n=1}^{\infty} \frac{a_n}{n^s} N^s \frac{ds}{s(s+1)}.$$

This formula is well known in analytic number theory and can be proved by an application of Mellin transform (cf. [1], [6], [7]). We demonstrate the use of this method in the case of the expectation of the q -ary sum-of-digits function.

Theorem 3: [Delange] The sum-of-digits function $v_q(n)$ satisfies

$$S_q(N) = \sum_{n < N} v_q(n) = \frac{q-1}{2} N \log_q N + N \delta(\log_q N),$$

where $\delta(u)$ is representable by the Fourier series $\delta(u) = \sum_{k \in \mathbb{Z}} \delta_k e^{2\pi i k x}$ and

$$\delta_0 = \frac{q-1}{2} \left(\log_q 2\pi - \frac{1}{\log q} \right) - \frac{q+1}{4},$$

$$\delta_k = -\frac{q-1}{\log q} \zeta(\chi_k) \text{ for } k \neq 0,$$

where $\chi_k = \frac{2k\pi i}{\log q}$.

Proof: Let $v_q(k)$ be the largest exponent such that $q^{v_q(k)}$ divides k and $v_q(k)$ be the sum-of-digits function in the q -ary representation. We have $v_q(k) - v_q(k-1) = 1 - (q-1)v_q(k)$, so that $S_q(N)$ resembles a double summation of $v_q(k)$. Furthermore it is well-known and easy to see that

$$\sum_{k \geq 1} \frac{v_q(k)}{k^s} = \frac{\zeta(s)}{q^s - 1}. \tag{4.1}$$

Thus, from (4.1), with $a_k = 1 - (q-1)v_q(k)$, we get the basic integral representation

$$S_q(n) = \frac{n}{2\pi i} \int_{2-i\infty}^{2+i\infty} \frac{q^s - q}{q^s - 1} \zeta(s) n^s \frac{ds}{s(s+1)}. \tag{4.2}$$

The integrand in (4.2) has a double pole at $s = 0$ and simple poles at $s = \chi_k$ (note that the pole $s = 1$ of $\zeta(s)$ is canceled out). Shifting the line of integration to $\Re(s) = -\frac{1}{4}$ and taking residues into account we get

$$S(n) = \frac{1}{2} n \log_q n + n F_0(\log_q n) - n R(n), \tag{4.3}$$

where the remainder term is

$$R(n) = \frac{1}{2\pi i} \int_{-\frac{1}{4}-i\infty}^{-\frac{1}{4}+i\infty} \frac{q^s - q}{q^s - 1} n^s \frac{ds}{s(s-1)}, \tag{4.4}$$

so that there only remains to prove that $R(n) \equiv 0$ when n is an integer. The integral converges since $|\zeta(-\frac{1}{4} + it)| \ll |t|^{\frac{3}{4}}$ (cf. [22]).

Using the expansion

$$\frac{q^s - q}{q^s - 1} = q + (q-1)(q^s + q^{2s} + q^{3s} \dots)$$

in (4.4) which is legitimate since now $\Re(s) < 0$, we find that $R(n)$ is a sum of terms of the form

$$\frac{1}{2\pi i} \int_{-\frac{1}{4}-i\infty}^{-\frac{1}{4}+i\infty} \zeta(s) (q^k n)^s \frac{ds}{s(s+1)},$$

and each of these terms is 0 by virtue of

$$\frac{1}{2\pi i} \int_{-\frac{1}{4}-i\infty}^{-\frac{1}{4}+i\infty} \zeta(s)n^s \frac{ds}{s(s+1)} = 0,$$

which is an immediate consequence of the summation formula for arithmetic progressions and (4.1). \square

Remark: It is clear from the discussion above that an exact formula for a sum-of-digits function is obtained each time the analytical behavior of the corresponding Dirichlet generating function is sufficiently known. In the case of higher moments of the sum-of-digits function this procedure can be worked out only formally: serious convergence problems occur. In order to overcome these difficulties one would have to know precise bounds for the rate of growth of the analytic continuation of the Dirichlet series

$$\sum_{n=1}^{\infty} \frac{v(n)^k}{n^s}$$

along vertical lines.

REFERENCES

- [1] Apostol, T.M. Introduction to Analytic Number Theory, Springer Verlag, Berlin-Heidelberg-New York, 1984.
- [2] Brillhart, J., Erdős, P. and Morton, P. "On Sums of Rudin-Shapiro coefficients II." *Pacific J. Math.*, Vol. 107 (1983): pp. 39-69.
- [3] Coquet, J. "A Summation Formula Related to the Binary Digits." *Invent. Math.*, Vol. 73 (1983): pp. 107-115.
- [4] Coquet, J. "Power Sums of Digital Sums." *J. Number Th.*, Vol. 22 (1986): pp. 161-176.
- [5] Delange, H. *Sur la fonction sommatoire de la fonction* "Somme des Chiffres." *Enseignement Math. (2)*, Vol. 21 (1975): pp. 31-47.
- [6] Doetsch, G. Handbuch der Laplace Transformation. Birkhäuser Verlag, Basel, 1950.
- [7] Flajolet, P., Grabner, P.J., Kirschenhofer, P., Prodinger, H. and Tichy, R.F. "Mellin Transform and Asymptotics: Digital Sums." *Theor. Comput. Sci.*, to appear.
- [8] Flajolet, P. and Ramshaw, L. "A Note on Gray Code and Odd-Even Merge." *SIAM J. Comput.*, Vol. 9 (1980): pp. 142-158.
- [9] Flajolet, P., Raoult, J.C. and Vuillemin, J. "The Number of Registers Required for Evaluating Arithmetic Expressions." *Theor. Comput. Sci.*, Vol. 9 (1979): pp. 99-125.

- [10] Flajolet, P. and Vitter, J. "Average-Case Analysis of Algorithms and Data Structures," Handbook of Theoretical Computer Science Vol. A "Algorithms and Complexity." North-Holland, (1990): pp. 431-524.
- [11] Grabner, P.J. and Tichy, R.F. " α -Expansions, Linear Recurrences and the Sum-of-Digits Function." *Manuscripta Math.*, Vol. 70 (1991): pp. 311-324.
- [12] Kennedy, R.E. and Cooper, C.N. "An Extension of a Theorem by Cheo and Yien Concerning Digital Sums." *The Fibonacci Quarterly*, Vol. 29 (1991): pp. 145-149.
- [13] Kirschenhofer, P. "Subblock Occurrences in the q-ary Representation of n." *SIAM J. Alg. Discr. Math.*, Vol. 4 (1983): pp. 231-236.
- [14] Kirschenhofer, P. "On the Variance of the Sum of Digits Function." Number-Theoretic Analysis, Lecture Notes in Mathematics 1452 (E. Hlawka and R.F. Tichy, editors.) (1990): pp. 112-116.
- [15] Kirschenhofer, P. and Prodinger, H. "Subblock Occurrences in Positional Number Systems and Gray Code Representation." *J. Inf. Opt. Sci.*, Vol. 5 (1984): pp. 29-42.
- [16] Kirschenhofer, P., Prodinger, H. and Tichy, R.F. "Über die Ziffernsumme natürlicher Zahlen und verwandte Probleme." Zahlentheoretische Analysis, Lecture Notes in Mathematics 1114 (E. Hlawka, editor) (1985): pp. 55-65.
- [17] Kirschenhofer, P. and Tichy, R.F. "On the distribution of digits in Cantor representations of integers." *J. Number Th.*, Vol. 18 (1984): pp. 121-134.
- [17*] Mauclaire, J.-L. and Murata, L. "On q-Additive Functions." *Proc. Japan Acad.*, Vol. 59, Series A (1987): pp. 441-444.
- [18] Prodinger, H. "Generalizing the sum of digits function." *SIAM J. Alg. Discr. Math.*, Vol. 3 (1982): pp. 35-42.
- [19] Riordan, J. An Introduction to Combinatorial Theory. J. Wiley, New York, 1958.
- [20] Stolarsky, K.B. "Power and Exponential Sums of Digital Sums Related to Binomial Coefficient Parity." *SIAM J. Appl. Math.*, Vol. 32 (1977): pp. 717-730.
- [21] Trollope, H. "An explicit expression for binary digital Sums." *Math. Mag.*, Vol. 41 (1968): pp. 21-25.
- [22] Whittaker, E.T. and Watson, G.N. A Course in Modern Analysis, Cambridge University Press, 1927.

*After this paper was presented, refereed, and accepted for publication, the authors became aware of the paper by Mauclaire and Murata where explicit formulas as derived in Theorem 3 were obtained in a more general context.