#A54 INTEGERS 11 (2011)

DIGITAL SUMS AND FUNCTIONAL EQUATIONS

Roland Girgensohn[1]
Zentrum Mathematik, Technische Universität München, Germany

Received: 8/30/10, Revised: 6/14/11, Accepted: 8/22/11, Published: 9/19/11

Abstract

Let $S(n)$ denote the total number of digits '1' in the binary expansions of the integers between 1 and $n-1$. The Trollope-Delange formula is a classical result which provides an explicit representation for $S(n)$ in terms of the continuous, nowhere differentiable Takagi function. Recently, connections have been established between digital sums such as $S(n)$ and certain functional equations associated with the Takagi function and its relatives. In the present paper we explore such a connection to derive a new, simple proof for the Trollope-Delange formula as well as for some of its generalizations involving power and exponential sums.

1. Introduction

Let $S(n)$ denote the total number of digits '1' in the binary expansions of the integers between 1 and $n-1$. Since that is roughly half the number of all of those digits, it is not far-fetched that $S(n)$ must be of the order $S(n) = \frac{1}{2}n\log_2 n + O(n)$. Interestingly, it turns out that the capital-O term in this expansion can be given explicitly as n times a continuous, 1-periodic function of $\log_2 n$. This was first proved by J.R. Trollope in [19]; subsequently, in [4], H. Delange gave a very short and direct proof of this representation. The continuous function appearing in their representation is a slight modification of the well-known continuous, nowhere differentiable Takagi function, investigated by T. Takagi already in 1903 [18] and often presented as one of the simplest examples of a nowhere differentiable function.

The Trollope-Delange formula has been investigated and generalized intensively in the intervening years. Bases other than 2 have been examined, the occurrence of subblocks other than the digit '1' has been counted, and other modifications have been applied to these quantities. Always it turned out that a representation of Trollope-Delange type of the quantity in question could be given, often with explicit continuous functions. Some references to such work are given below.

[1]Postal address: Kepserstr. 5, 85356 Freising, Germany

The purpose of the present note is to give a new and simple proof of the Trollope-Delange formula and then to turn this proof into a method by applying it to some other variations and generalizations of these sums. The method proceeds by extracting functional equations from the digital sum sequences such as $S(n)$, identifying their solutions and then using this process to prove a formula of Trollope-Delange type in just a few lines.

To fix notation, let

$$j = \sum_{i \geq 0} a_i(j)\, 2^i \quad \text{with } a_i(j) \in \{0, 1\} \tag{1}$$

be the binary expansion of $j \in \mathbb{N}_0$, let

$$s(j) = \sum_{i \geq 0} a_i(j) \tag{2}$$

be the number of digits '1' in the binary expansion of j, and let

$$S(n) = \sum_{j=0}^{n-1} s(j), \tag{3}$$

$$S(n; t) = \sum_{j=0}^{n-1} \exp(t \cdot s(j)) \quad \text{for } t \in \mathbb{R} \text{ and} \tag{4}$$

$$S_k(n) = \sum_{j=0}^{n-1} s(j)^k \quad \text{for } k \in \mathbb{N} \tag{5}$$

denote the digital sum in question as well as its so-called exponential and power sums. Then the Trollope-Delange formula for $S(n)$ is

$$\frac{1}{n} S(n) = \frac{1}{2} \log_2 n + \frac{1}{2} \widetilde{F}(\log_2 n) \tag{6}$$

where the 1-periodic function $\widetilde{F}$ is given by

$$\widetilde{F}(u) = 1 - u - 2^{1-u}\, T\left(\frac{1}{2^{1-u}}\right) \quad \text{for } 0 \leq u \leq 1. \tag{7}$$

Here, T is the Takagi function

$$T(x) = \sum_{n=0}^{\infty} \frac{1}{2^n}\, d(2^n x) \quad \text{with } d(x) = \operatorname{dist}(x, \mathbb{Z}) \text{ and } x \in \mathbb{R}. \tag{8}$$

While this is the "direct" definition of the Takagi function, T can just as well be defined "indirectly" on $[0, 1]$ as the only continuous solution of the system of two functional equations

$$f\left(\frac{x}{2}\right) = \frac{1}{2} f(x) + \frac{x}{2}, \quad f\left(\frac{x+1}{2}\right) = \frac{1}{2} f(x) + \frac{1-x}{2} \quad \text{for } x \in [0, 1]. \tag{9}$$

The system (9) is a specific example of the type of functional equations which, as mentioned above, appear in the digital sum sequences and can therefore be used to prove the Trollope-Delange formula. Thus, we will in Section 2 of this note review some basic properties of these functional equations, before we use them in Section 3 to give a simple proof of the Trollope-Delange formula. This proof establishes and then exploits a series of simple identities for the sequence $S(n)$ which will turn out to be equivalent to the functional equations (9). In Section 4 we use similar methods to derive a representation for $S(n;t)$ in an analogous manner. The representation itself is known; its derivation is new. In Section 5 we note that $S(n;t)$ is the exponential generating function for the quantities $S_k(n)$, so that generating function techniques can be used to extract information about the $S_k(n)$ from the representation for $S(n;t)$. Finally, in Section 6, we will use the same techniques to derive analogous formulas for the number of zeros in the binary expansions of the integers. The continuous functions appearing in all of these representations will always be given as solutions of functional equations of the same type as (9); this representation is explicit in the sense that these solutions can be computed explicitly from the functional equations.

All of these digital sums have been considered before, and representations for them have been given. Beyond the classical papers by Trollope and Delange, some newer examples are [3], [5], [13], [14], [11], [9] and [10]. In [9], the idea to use functional equations of type (9) to treat the digital sums is put forward, and in [10], this idea is explored to the fullest. The present paper has the same focus and scope as [10]; the methods however are different. In [10] as in all of the previous work, the functions appearing in the representations are taken as a given, and the representations are then proved from properties of these functions. Here, we go the other way round: we start with the sequences themselves, discover the functional equations within these sequences and then identify the functions from the functional equations. This is therefore a direct way of deriving Trollope-Delange type formulas.

2. Functional Equations

Before proving representations for the digital sums, some basic facts about the functional equations used in the representations must be recounted.

Let $g_0, g_1 : [0,1] \to \mathbb{R}$ be given *perturbation functions*, and let $|a_0|, |a_1| < 1$. We consider the system of two functional equations on $[0,1]$,

$$f\left(\frac{x}{2}\right) = a_0 f(x) + g_0(x), \tag{10}$$

$$f\left(\frac{x+1}{2}\right) = a_1 f(x) + g_1(x), \tag{11}$$

where f is an unknown function satisfying these equations for all $x \in [0,1]$.

It is clear that if such a function f exists, then it must satisfy $f(0) = g_0(0)/(1-a_0)$ (put $x = 0$ into (10)) and $f(1) = g_1(1)/(1-a_1)$ (put $x = 1$ into (11)). Moreover, by putting $x = 1$ into (10) and $x = 0$ into (11), it follows that $f\left(\frac{1}{2}\right)$ must equal the two expressions

$$a_0 \frac{g_1(1)}{1-a_1} + g_0(1) = a_1 \frac{g_0(0)}{1-a_0} + g_1(0). \tag{12}$$

This is only possible if the two expressions are equal themselves. Thus, (12) is a nessecary condition for the existence of a solution f.

It is also sufficient. Assume that g_0, g_1 are continuous and that (12) holds. Then it can be proved as an easy application of Banach's fixed point theorem that there exists a unique continuous solution f of (10),(11) on $[0,1]$. (The earliest proof of this was given in much more generality by M.F. Barnsley in connection with his fractal interpolation functions, [2].)

If (12) holds, then any solution f is uniquely determined on the dyadic rationals $i/2^n$, $n \in \mathbb{N}_0$, $i = 0, \ldots, 2^n$. In fact, we have already seen above that (10),(11) fix the values $f(0)$, $f(1)$ and $f\left(\frac{1}{2}\right)$. This process can be continued. If, for $n \in \mathbb{N}$, the values $f\left(\frac{2i+1}{2^n}\right)$ $(i = 0, \ldots, 2^{n-1}-1)$ are already known, then the values $f\left(\frac{2i+1}{2^{n+1}}\right)$ can be computed from (10), and the values $f\left(\frac{2i+1+2^n}{2^{n+1}}\right)$ can be computed from (11). This covers all the dyadic rationals, and since they are dense in $[0,1]$, the functional equations in this sense allow the explicit computation of their continuous solution.

These functional equations have been discussed in a different context in [6],[7]. They were used there to characterize classes of nowhere differentiable functions as solutions of these functional equations, and then to prove the non-differentiability of the solutions directly from the functional equations. This applies to a large variety of nowhere differentiable functions which previously had been investigated separately; among them are the Takagi functions as well as the famous Weierstraß functions $C(x) = \sum_{n=0}^{\infty} a^n \cos(2^n \cdot 2\pi x)$, which satisfy $(10), (11)$ with $a_0, a_1 = a$, $g_0(x) = \cos(\pi x)$, $g_1(x) = -\cos(\pi x)$.

3. The Trollope-Delange Formula

First, we need a simple lemma which connects functions defined on the integers with functions defined on the dyadic rationals. As a basic notation for the rest of this note, set

$$p(n) = 2^{[\log_2 n]} \quad \text{for } n \in \mathbb{N} \tag{13}$$

where $[u]$ is the largest integer less than or equal to u. In other words, $p(n)$ is the largest power of 2 less than or equal to n. We always have

$$p(n) \leq n < 2p(n). \tag{14}$$

The identities

$$\begin{aligned} p(2n) &= 2p(n), && (15)\\ p(n+p(n)) &= 2p(n), && (16)\\ p(n+2p(n)) &= 2p(n) && (17) \end{aligned}$$

are simple but will be important throughout the rest of this note.

Lemma 1. *Let $G : \mathbb{N} \to \mathbb{R}$ be a function on the integers. For $n \in \mathbb{N}$, set*

$$x := \frac{n-p(n)}{p(n)} \in [0,1) \quad \textit{and} \quad F(x) = F\left(\frac{n-p(n)}{p(n)}\right) := G(n). \qquad (18)$$

Then F is a well-defined function on the dyadic rationals in $[0,1)$ if and only if $G(2n) = G(n)$ for all $n \in \mathbb{N}$.

Proof. Note that when n ranges through all integers, then $\frac{n-p(n)}{p(n)}$ ranges through all dyadic rationals in $[0,1)$. Every dyadic rational is hit infinitely often. Thus, F is well-defined if and only if $G(n_1) = G(n_2)$ for all $n_1, n_2 \in \mathbb{N}$ with $\frac{n_1-p(n_1)}{p(n_1)} = \frac{n_2-p(n_2)}{p(n_2)}$.

This last equality implies $p(n_2)\, n_1 = p(n_1)\, n_2$. Thus n_1 and n_2 can only differ by a factor which is a power of 2. The reverse is also true, so that $\frac{n_1-p(n_1)}{p(n_1)} = \frac{n_2-p(n_2)}{p(n_2)}$ if and only if $n_1 = 2^{\ell} n_2$ with an $\ell \in \mathbb{Z}$.

Thus, F is well defined if and only if $G(n_1) = G(n_2)$ for all $n_1, n_2 \in \mathbb{N}$ whose quotient is an integer power of 2. This condition in turn holds if and only if $G(n) = G(2n)$ for all $n \in \mathbb{N}$. □

The Trollope-Delange formula can now be proved as a consequence of a series of simple identities. First, let p be a power of 2 and note that

$$s(2j) = s(j) \quad \text{and} \quad s(2j+1) = s(j)+1 \quad \text{for all } j = 0, 1, \ldots, \qquad (19)$$

$$\begin{aligned} s(j+p) &= s(j)+1 \quad \text{for } j = 0, 1, \ldots, p-1, \quad \text{and} && (20)\\ s(j+p) &= s(j) \quad \text{for } j = p, p+1, \ldots, 2p-1. && (21) \end{aligned}$$

Next, note that these identities imply, for all $n \in \mathbb{N}$,

$$\begin{aligned} S(2n) &= 2S(n) + n, && (22)\\ S(n+p(n)) &= S(n) + S(p(n)) + p(n) \quad \text{and} && (23)\\ S(n+2p(n)) &= S(n) + S(2p(n)) + n. && (24) \end{aligned}$$

To derive (22), split the sum for $S(2n)$ into two parts, summing over the even and the odd integers, and use (19). To derive (23), split the sum for $S(n+p(n))$ into three parts starting at 0, $p(n)$ and $2p(n)$, identify the second part as $S(2p(n)) - S(p(n))$

and use (20) on the third part. To derive (24), split the sum for $S(n+2p(n))$ into two parts starting at 0 and $2p(n)$, and use (20) (with $2p(n)$ instead of p) on the second part.

Now set

$$G(n) := \frac{1}{p(n)}\left(S(n) - \frac{n}{p(n)}S(p(n))\right). \tag{25}$$

Then, using formulas (15)–(17) and (22)–(24), it is easy to compute

$$\begin{aligned} G(2n) &= G(n), & (26)\\ G(n+p(n)) &= \frac{1}{2}G(n) - \frac{n-p(n)}{4p(n)} \quad \text{and} & (27)\\ G(n+2p(n)) &= \frac{1}{2}G(n) + \frac{n}{4p(n)}. & (28)\end{aligned}$$

If we set $x = x(n) := \frac{n-p(n)}{p(n)}$, then, using (16) and (17),

$$\begin{aligned} \frac{x(n)}{2} &= \frac{n-p(n)}{2p(n)} = \frac{n+p(n)-p(n+p(n))}{p(n+p(n))} = x(n+p(n)) \quad \text{and} & (29)\\ \frac{x(n)+1}{2} &= \frac{n}{2p(n)} = \frac{n+2p(n)-p(n+2p(n))}{p(n+2p(n))} = x(n+2p(n)). & (30)\end{aligned}$$

Moreover, by Lemma 1 the function F given by $F(x) = F\left(\frac{n-p(n)}{p(n)}\right) := G(n)$ is well-defined on the dyadic rationals in $[0,1)$ and by (27),(28) satisfies

$$\begin{aligned} F\left(\frac{x}{2}\right) &= G(n+p(n)) = \frac{1}{2}G(n) - \frac{n-p(n)}{4p(n)} = \frac{1}{2}F(x) - \frac{x}{4} \quad \text{and} & (31)\\ F\left(\frac{x+1}{2}\right) &= G(n+2p(n)) = \frac{1}{2}G(n) + \frac{n}{4p(n)} = \frac{1}{2}F(x) + \frac{x+1}{4}. & (32)\end{aligned}$$

This means that F satisfies the system (10),(11) with $a_0 = a_1 = \frac{1}{2}$, $g_0(x) = -\frac{x}{4}$ and $g_1(x) = \frac{x+1}{4}$. Since condition (12) is satisfied (with $F(0) = 0$, $F(1) = 1$ and $F(\frac{1}{2}) = \frac{1}{4}$), it follows that F is the restriction to the dyadic rationals of a continuous function on $[0,1]$ (also denoted by F), which is a solution of (10),(11) on $[0,1]$.

By solving (25) for $S(n)$, we get a representation for $S(n)$ involving $G(n)$ and $S(p(n))$. The latter quantity can be given explicitly: Since $p(n)$ is a power of 2, the binary expansion of the $p(n)$ numbers between 0 and $p(n)-1$ can be thought of as being '0'-'1'-strings, each of length $\log_2 p(n)$ and precisely half of whose digits are '1's. Therefore

$$S(p(n)) = \frac{1}{2}\,p(n)\log_2 p(n). \tag{33}$$

From this we get the representation

$$S(n) = \frac{n}{p(n)}S(p(n)) + p(n)G(n) = \frac{n\log_2 p(n)}{2} + p(n)\,F\left(\frac{n-p(n)}{p(n)}\right), \tag{34}$$

and this (via $p(n) = n/(x+1)$) is equivalent to

$$\frac{1}{n}S(n) = \frac{\log_2 n}{2} - \frac{\log_2(x+1)}{2} + \frac{1}{x+1}F(x) \quad \text{where } x = \frac{n-p(n)}{p(n)} \tag{35}$$

and where F is the countinuous function on $[0,1]$ which is given by the system (31),(32). This is the Trollope-Delange formula.

It is, however, given in a form slightly different from (6),(7). The equivalence of the two forms can be seen quickly by noting that

$$F(x) = x - \frac{1}{2}T(x) = \frac{x+1}{2} - T\left(\frac{x+1}{2}\right). \tag{36}$$

The first equality in (36) follows because $F(x)$ and $x - \frac{1}{2}T(x)$ satisfy the same system of type (10),(11), and the second equality follows by applying the second equation of (9). Now set u equal to the fractional part of $\log_2 n$, i.e., $u := \{\log_2 n\}$, and note that $x = 2^u - 1$ (resp. $u = \log_2(x+1)$) to arrive at (6),(7).

The core of the proof of the Trollope-Delange formula presented here consists of deriving identities for $S(n+p(n))$ and $S(n+2p(n))$, (23) and (24). With these two identities (and the value $S(1) = 1$), the whole sequence $S(n)$ is determined without it ever being necessary to compute a single value of $s(n)$. The Trollope-Delange formula follows because (23) and (24) determine the sequence $S(n)$ in the same way as the functional equations (10) and (11) allow the computation of their continuous solution; both perspectives are in fact equivalent by (29) and (30). This also explains the approximation process which is visible in Figure 1 of [5].

Of course, identities for $S(2n)$ and $S(2n+1)$ would similarly determine the whole sequence $S(n)$. An identity for $S(2n)$ is given above in (22), and for $S(2n+1)$ it can be proved in a similar fashion that

$$S(2n+1) = 2S(n) + n + s(n). \tag{37}$$

This means that although (22) and (37) also determine the whole sequence $S(n)$, they only do so after computing the values $s(n)$. In this sense, (23) and (24) are "simpler" or "more natural" than (22) and (37).

Also, note that the proof presented here does not require any advance knowledge of the main term $\frac{\log_2 n}{2}$. This term (in the form $\frac{\log_2 p(n)}{2}$) arises in the computations in a natural way, where it comes from the term $S(p(n))$. One might say that the values $S(p)$, where p is a power of 2, act as "stepping stones" for the complete sequence $S(n)$. We will see such a behavior again for the other types of digital sums considered below.

4. Exponential Sums

An explicit formula for the exponential sums $S(n;t)$ has been derived in [14] and (by different methods) in [10]; in our notation it is given in formula (50), below. This representation involves the so-called de Rham (sometimes Lebesgue) singular function $L_a(x)$. It can be defined as the continuous solution of the system

$$f\left(\frac{x}{2}\right) = a\, f(x) \quad \text{and} \quad f\left(\frac{x+1}{2}\right) = (1-a)\, f(x) + a \tag{38}$$

on $[0,1]$ for fixed $a \in (0,1)$. This function was first constructed by R. Salem in [16] as a simple example of a singular function (it is strictly monotone and for $a \neq \frac{1}{2}$ has derivative 0 almost everywhere). In [15], G. de Rham used the functional equations (38) to characterize $L_a(x)$ and to prove its singularity.

The goal now is to give another proof of the representation from [14] and [10], using the method developed in the previous section, i.e., by deriving analogs for the formulas (23)–(35). In particular, we do not have to know in advance that the function $L_a(x)$ will appear in the representation.

We get, using (19)–(21) in the same way as before,

$$\begin{aligned} S(2n;t) &= \left(e^t+1\right) S(n;t), && (39)\\ S(n+p(n);t) &= S(n;t) + e^t\, S(p(n);t) \quad \text{and} && (40)\\ S(n+2p(n);t) &= e^t\, S(n;t) + S(2p(n);t). && (41) \end{aligned}$$

Now set

$$G(n;t) := \frac{S(n;t)}{S(p(n);t)}. \tag{42}$$

Then it follows that

$$\begin{aligned} G(2n;t) &= G(n;t), && (43)\\ G(n+p(n);t) &= \frac{1}{e^t+1}\, G(n;t) + \frac{e^t}{e^t+1} \quad \text{and} && (44)\\ G(n+2p(n);t) &= \frac{e^t}{e^t+1}\, G(n;t) + 1. && (45) \end{aligned}$$

Thus the function $x \mapsto F(x;t)$ given by $F(x;t) = F\left(\frac{n-p(n)}{p(n)};t\right) := G(n;t)$ is well-defined on the dyadic rationals in $[0,1)$ and by (44),(45) satisfies

$$\begin{aligned} F\left(\frac{x}{2};t\right) &= \frac{1}{e^t+1} F(x;t) + \frac{e^t}{e^t+1} \quad \text{and} && (46)\\ F\left(\frac{x+1}{2};t\right) &= \frac{e^t}{e^t+1} F(x;t) + 1. && (47) \end{aligned}$$

Since $e^t > 0$, we have $\frac{1}{e^t+1} < 1$ and $\frac{e^t}{e^t+1} < 1$. Moreover, condition (12) is satisfied with

$$F(0;t) = 1,\ F(1;t) = e^t + 1 \text{ and } F\left(\frac{1}{2};t\right) = \frac{2e^t+1}{e^t+1}. \tag{48}$$

Therefore $F(x,t)$ is a continuous function in x.

Again, $S(p(n);t)$ can be computed explicitly; the easiest way would be to iterate formula (39) to get

$$S(p(n);t) = (e^t+1)^{\log_2 p(n)}. \tag{49}$$

Altogether, we get

$$S(n;t) = (e^t+1)^{\log_2 p(n)} \cdot F\left(\frac{n-p(n)}{p(n)};t\right) = (e^t+1)^{\log_2 n} \cdot (e^t+1)^{-\log_2(x+1)} F(x;t) \tag{50}$$

where $x = \frac{n-p(n)}{p(n)}$. If we let

$$\widetilde{F}(x;t) := (e^t+1)^{-\log_2(x+1)} F(x;t), \tag{51}$$

then $\widetilde{F}$ satisfies $\widetilde{F}(0;t) = \widetilde{F}(1;t) = 1$ for every $t \in \mathbb{R}$, so that for every t, $\widetilde{F}(\cdot;t)$ can be continued to a continuous, 1-periodic function on $\mathbb{R}$. Thus (50) is the exponential analog of the Trollope-Delange formula.

Now, the function $F(x;t)$ appearing in (50) is just a rescaled and re-parametrized version of de Rham's singular function $L_a(x)$. Explicitly, we have

$$F(x;t) = 1 + e^t\, L_{\frac{1}{1+e^t}}(x), \tag{52}$$

because both sides of this identity satisfy the same functional equations of type (10),(11).

5. Power Sums

Now set $S_0(n) := n$ and consider the power sums $S_k(n)$. The first explicit representation for $S_k(n)$ was given by J. Coquet in [3]. He proved that there exist 1-periodic functions $G_{k,\ell}$ such that

$$\frac{1}{n}S_k(n) = \left(\frac{\log_2 n}{2}\right)^k + \sum_{\ell=0}^{k-1} (\log_2 n)^\ell G_{k,\ell}(\log_2 n) \tag{53}$$

holds; he also found certain recurrence relations between the functions $G_{k,\ell}$. In [13], a more explicit representation for the functions $G_{k,\ell}$ was given and their continuity was proved. In [10], a still more explicit representation for the $G_{k,\ell}$ was given.

The basic building block in the latter two papers is de Rham's function $L_a(x)$; the functions $G_{k,\ell}$ are then given as certain combinations of the partial derivatives (with respect to a) of $L_a(x)$. These partial derivatives appear in the representations because

$$S_k(n) = \frac{\partial^k}{\partial t^k} S(n;t)\bigg|_{t=0}, \tag{54}$$

and $S(n;t)$ is related to $L_a(x)$ via $F(x;t)$ and (52). This fact is noted and made good use of in [14] and [10].

In the present note, we will also use (54) as a starting point, but we will then follow a line of argument different from that given in [14] and [10]. It will lead us to a representation which is different from (although of course equivalent to) Coquet's (53) in that it uses different, maybe slightly simpler, building blocks. These building blocks appear in the course of the argument in a natural way; prior knowledge of them is not necessary.

This argument uses the method of (exponential) generating functions. In fact, (54) directly translates into

$$S(n;t) = \sum_{k=0}^{\infty} \frac{S_k(n)}{k!}\, t^k, \tag{55}$$

so that all information about $S_k(n)$ is already contained in $S(n;t)$ and can be extracted by expanding the result (50) of the previous section into a power series and comparing coefficients. Note that for every $n \in \mathbb{N}$ the sum (55) has an infinite radius of convergence, because $S_k(n) \leq n^{k+1}$.

Before beginning with the actual argument, we need a few preparations. In general, if $c = (c_k)$ is an arbitrary sequence, then its exponential generating function is

$$C(t) = \sum_{k=0}^{\infty} \frac{c_k}{k!}\, t^k. \tag{56}$$

For later reference, note the formula for the Cauchy product of exponential generating functions: If C_1 and C_2 are the exponential generating functions of $(c_{1,k})$ and $(c_{2,k})$, then the exponential generating function of $C_1 \cdot C_2$ is

$$C_1(t) \cdot C_2(t) = \sum_{k=0}^{\infty} \frac{c_{3,k}}{k!}\, t^k \quad \text{where } c_{3,k} = \sum_{\nu=0}^{k} \binom{k}{\nu} c_{1,\nu}\, c_{2,k-\nu}. \tag{57}$$

To finish the preparations, set the auxiliary function $D(t)$ (which appears in (50)) equal to

$$D(t) := \frac{1}{1+e^t} \tag{58}$$

and expand $D(t)^{-u}$ into an exponential series,

$$D(t)^{-u} = (1+e^t)^u = \sum_{k=0}^{\infty} \frac{d(k;u)}{k!}\, t^k \tag{59}$$

where

$$d(k;u) = \left. \frac{\partial^k}{\partial t^k}(1+e^t)^u \right|_{t=0} \tag{60}$$

and with radius of convergence at least π. From (60) we get the recursion

$$d(0;u) = 2^u \quad \text{and} \quad d(k+1;u) = u\, d(k;u) - u\, d(k;u-1), \tag{61}$$

so that (inductively) $d(k;u) = 2^{u-k}\, q_k(u)$ where q_k is a polynomial of degree k and leading coefficient 1, satisfying the recursion

$$q_0(u) = 1 \quad \text{and} \quad q_{k+1}(u) = u\,(2q_k(u) - q_k(u-1)) \text{ for } k \geq 0. \tag{62}$$

We will also need the particular values $d_k := d(k;-1)$; since

$$D(t)^1 = \frac{1}{1+e^t} = \frac{1}{2}\left(1 - \tanh\left(\frac{t}{2}\right)\right), \tag{63}$$

they have the explicit representation

$$d_k = -\frac{2^{k+1}-1}{k+1}\, B_{k+1}, \tag{64}$$

where the numbers B_k are the Bernoulli numbers.

Now the result (50) of the previous section reads,

$$S(n;t) = D(t)^{-\log_2 p(n)} \cdot F(x;t) = D(t)^{-\log_2 n} \cdot D(t)^{\log_2(x+1)} F(x;t) \tag{65}$$

where, by (46),(47), the function $F(x;t)$ satisfies, for every t, the system of functional equations

$$F\left(\frac{x}{2};t\right) = D(t)F(x;t) + 1 - D(t) \quad \text{and} \quad F\left(\frac{x+1}{2};t\right) = (1-D(t))F(x;t) + 1. \tag{66}$$

Note here that if $x \in [0,1)$ is a dyadic rational, then by induction $F(x;t)$ is a polynomial in $D(t)$, so that the series for $F(x;t)$ has radius of convergence at least π. Moreover, it follows from results in [8] that for every $x \in [0,1]$, $F(x;t)$ is analytic around $t = 0$.

Thus writing

$$F(x;t) = \sum_{k=0}^{\infty} \frac{F_k(x)}{k!} t^k, \tag{67}$$

(66) implies by (57) that, for $k \geq 0$, $F_k(x)$ is a solution of

$$F_k\left(\frac{x}{2}\right) = \frac{1}{2}F_k(x) + \delta_{k,0} - d_k + \sum_{\nu=0}^{k-1}\binom{k}{\nu} d_{k-\nu}\, F_\nu(x), \tag{68}$$

$$F_k\left(\frac{x+1}{2}\right) = \frac{1}{2}F_k(x) + \delta_{k,0} - \sum_{\nu=0}^{k-1}\binom{k}{\nu} d_{k-\nu}\, F_\nu(x). \tag{69}$$

Here, $\delta_{k,\ell}$ is the Kronecker symbol which equals 1 if $k = \ell$ and 0 otherwise.

This is a recursive definition of the functions F_k: If the solutions F_ν of the functional equations are known for $\nu = 0, \ldots, k-1$, then these solutions enter the functional equations for F_k as perturbation functions. Also, the solution for $k = 0$ is $F_0(x) = x + 1$. It can be seen recursively that condition (12) is satisfied (with, for $k \geq 1$, $F_k(0) = 0$, $F_k(1) = 1$ and $F_k(\frac{1}{2}) = -d_k$), so that each F_k is a continuous function on $[0, 1]$. At the end of this section, we will say more about these functions.

Now, $S_k(n)$ can be given explicitly in terms of the continuous functions F_k and the polynomials q_k. One just has to convert (65) via (57) into its exponential series and compare coefficients on both sides.

Writing

$$D(t)^{\log_2(x+1)} F(x;t) = \sum_{k=0}^{\infty} \frac{f_k(x)}{k!}\, t^k, \tag{70}$$

we have

$$f_k(x) = \frac{1}{x+1} \sum_{\mu=0}^{k} \binom{k}{\mu} 2^{-\mu} q_\mu(-\log_2(x+1)) \cdot F_{k-\mu}(x). \tag{71}$$

Note here that the function $t \mapsto D(t)^{\log_2(x+1)} F(x;t)$ is, for $x = 0$ and $x = 1$, identical to 1 (use (48) and (58)). Therefore, $f_k(0) = f_k(1) = \delta_{k,0}$ for all k, and the continuous function f_k can be continuously extended to a 1-periodic function on $\mathbb{R}$.

Next from (65) we get, expanding $D(t)^{-\log_2 n} \cdot D(t)^{\log_2(x+1)} F(x;t)$ via (57),

$$S_k(n) = \sum_{\nu=0}^{k} \binom{k}{\nu} n\, 2^{-\nu}\, q_\nu(\log_2 n) \cdot f_{k-\nu}(x). \tag{72}$$

Note here that, if the right-hand side of (72) is sorted by powers of $\log_2 n$, then each of these powers gets a factor which is a linear combination of the functions f_k, i.e., a continuous function in x which can be continued 1-periodically to a continuous function on $\mathbb{R}$.

Now putting everything together, we get the following analog of the Trollope-Delange formula for power sums.

Theorem 2. *Fix $k \geq 0$. Then for every $n \geq 0$ the identity*

$$\frac{1}{n} S_k(n) = \frac{1}{x+1} \sum_{\nu+\mu+\lambda=k} \frac{k!}{\nu!\,\mu!\,\lambda!}\, 2^{-\nu-\mu}\, q_\nu(\log_2 n)\, q_\mu(-\log_2(x+1))\, F_\lambda(x) \tag{73}$$

holds, where $x = \frac{n-p(n)}{p(n)}$*, where the polynomials* q_k *are given by the recursion (62), and where the functions* F_k *are the uniquely determined continuous solutions on* $[0, 1]$ *of the (recursive) system (68),(69).*

In this representation, every power of $\log_2 n$ *is multiplied by some continuous function in* x *which can be extended to a continuous, 1-periodic function on* $\mathbb{R}$*. The main term is* $\left(\frac{\log_2 n}{2}\right)^k$*.*

For $k = 0$, we get $S_0(n) = n$; for $k = 1$, we recover the original Trollope-Delange formula in the form (35). For $k = 2$, we get

$$\begin{aligned}\frac{1}{n} S_2(n) \quad &= \quad \frac{1}{4}(\log_2 n)^2 + \left(\frac{1}{4} - \frac{1}{2}\log_2(x+1) + \frac{1}{x+1}F_1(x)\right)\log_2 n \\ &\quad + \frac{1}{4}\log_2(x+1)^2 - \frac{1}{4}\log_2(x+1) - \frac{\log_2(x+1)}{x+1}F_1(x) + \frac{1}{x+1}F_2(x).\end{aligned} \tag{74}$$

The functions F_k are interesting; Figures 1–2 show F_1–F_4. Of course, F_1 is just the function F appearing in the original Trollope-Delange formula; it is therefore a relation of the Takagi function T. In fact, all of the functions F_k are linear combinations of certain functions T_k investigated in [17], [1] and [10]. These are the partial derivatives of the de Rham function,

$$T_k(x) = \left.\frac{\partial^k}{\partial a^k} L_a(x)\right|_{a=\frac{1}{2}}, \tag{75}$$

that is,

$$L_a(x) = \sum_{k=0}^{\infty} \frac{T_k(x)}{k!}\left(a - \frac{1}{2}\right)^k. \tag{76}$$

It was proved in [1] that the continuous functions T_k are nowhere differentiable; pictures of T_1–T_4 (resp. rescaled versions thereof) can also be found in [17] and [1]. Note that $T_0(x) = x$ and $T_1(x) = 2T(x)$ where T is the Takagi function. An alternative recursive definition of the functions T_k by way of functional equations of type (10),(11) is also given in [1] (in a different but equivalent form) and [10], namely, for $k \geq 1$ and $x \in [0, 1]$,

$$T_k\left(\frac{x}{2}\right) \quad = \quad \frac{1}{2}T_k(x) + kT_{k-1}(x) \quad \text{and} \tag{77}$$

$$T_k\left(\frac{x+1}{2}\right) \quad = \quad \frac{1}{2}T_k(x) + \delta_{k,1} - kT_{k-1}(x). \tag{78}$$

Since $F(x; t)$ and $L_a(x)$ are related by (52), we can set $a = \frac{1}{1+e^t}$ and then expand both sides of (52) into a power series around $t = 0$. Comparing coefficients, we find that F_k is a linear combination of $T_0, \dots, T_k$, namely

$$F_k(x) = \sum_{m=0}^{k} \frac{r_{k,m}}{m!}\, T_m(x) \quad \text{for } k \geq 1, \tag{79}$$

where the coefficients $r_{k,m}$ come from the power series

$$e^t \left(\frac{1}{1+e^t} - \frac{1}{2} \right)^m = \sum_{k=0}^{\infty} \frac{r_{k,m}}{k!} t^k. \tag{80}$$

From this, an explicit representation for the $r_{k,m}$ can be worked out, namely

$$r_{k,m} = \left(-\frac{1}{2} \right)^m \sum_{\nu=0}^{k} \binom{k}{\nu} \frac{1}{2^\nu} \sum_{\mu=0}^{m} \binom{m}{\mu} q_\nu(-\mu)\,(-1)^\mu\,. \tag{81}$$

Alternatively, the $r_{k,m}$ can also be computed recursively, as again follows from (80):

$$r_{k,0} = 1 \text{ for } k \geq 0, \quad r_{0,1} = 0 \text{ and } r_{k,1} = -d_k - \frac{1}{2} \text{ for } k \geq 1 \text{ and} \tag{82}$$

$$r_{k,m} = \sum_{\nu=m-1}^{k-1} \binom{k}{\nu} r_{\nu,m-1} d_{k-\nu} \text{ for } m \geq 2. \tag{83}$$

Explicitly, we have

$$F_1(x) = T_0(x) - \frac{1}{4}T_1(x), \quad F_2(x) = T_0(x) - \frac{1}{2}T_1(x) + \frac{1}{16}T_2(x), \tag{84}$$

$$F_3(x) = T_0(x) - \frac{5}{8}T_1(x) + \frac{3}{16}T_2(x) - \frac{1}{64}T_3(x), \tag{85}$$

and so on.

6. The Number of Zeros

For comparison with the results in the previous section, we now give analogous representations for the exponential and power sums of the number of digits '0' in the binary expansions of the integers. Since the computations run along the same lines as those in the previous sections, they will be omitted here.

Denote

$$s^{(0)}(j) \;:=\; \sum_{i=0}^{\log_2 p(j)} \delta_{a_i(j),0} \quad \text{for } j \geq 1 \quad \text{and } s^{(0)}(0) := -1, \tag{86}$$

$$S^{(0)}(n;t) \;=\; \sum_{j=0}^{n-1} \exp(t \cdot s^{(0)}(j)) \quad \text{for } t \in \mathbb{R} \text{ and} \tag{87}$$

$$S_k^{(0)}(n) \;=\; \sum_{j=0}^{n-1} s^{(0)}(j)^k \quad \text{for } k \in \mathbb{N}. \tag{88}$$

Setting $s^{(0)}(0) = -1$ in (86) has the effect of simplifying the formulas greatly. (For example, in our notation, the analog of formula (39) is again, $S^{(0)}(2n;t) = (e^t+1)\, S^{(0)}(n;t)$, while, without that normalization, we would get an extra summand of $1-e^t$.) Note, however, that this leads to values which may take some getting used to, such as

$$S_1^{(0)}(1) = -1,\ S_1^{(0)}(2) = -1,\ S_1^{(0)}(3) = 0,\ S_1^{(0)}(4) = 0, \tag{89}$$

and so on.

Define the function $F^{(0)}(x;t)$ as the continuous solution of the system

$$F^{(0)}\left(\frac{x}{2};t\right) = (1-D(t))\, F^{(0)}(x;t) + D(t) \quad \text{and} \tag{90}$$

$$F^{(0)}\left(\frac{x+1}{2};t\right) = D(t)\, F^{(0)}(x;t) + e^t. \tag{91}$$

This function is continuous in x since condition (12) is satisfied with

$$F^{(0)}(0;t) = 1,\ F^{(0)}(1;t) = e^t+1 \text{ and } F^{(0)}\left(\frac{1}{2};t\right) = \frac{e^{2t}+e^t+1}{e^t+1}. \tag{92}$$

It is in fact just a rescaled and re-parametrized version of the function $F(x;t) =: F^{(1)}(x;t)$ from Section 4:

$$F^{(0)}(x;t) = e^{2t}\, F^{(1)}(x;-t) + 1 - e^{2t}. \tag{93}$$

Now the Trollope-Delange formula for $S^{(0)}(n;t)$ is

$$S^{(0)}(n;t) = e^{-t} D(t)^{-\log_2 p(n)} \cdot F^{(0)}(x;t) = e^{-t} D(t)^{-\log_2 n} \cdot D(t)^{\log_2(x+1)} \cdot F^{(0)}(x;t) \tag{94}$$

with $x = \frac{n-p(n)}{p(n)}$.

For the power sums, define for $k \geq 0$ the function $F_k^{(0)}(x)$ as the continuous solution of

$$F_k^{(0)}\left(\frac{x}{2}\right) = \frac{1}{2} F_k^{(0)}(x) + d_k - \sum_{\nu=0}^{k-1} \binom{k}{\nu} d_{k-\nu}\, F_\nu^{(0)}(x), \tag{95}$$

$$F_k^{(0)}\left(\frac{x+1}{2}\right) = \frac{1}{2} F_k^{(0)}(x) + 1 + \sum_{\nu=0}^{k-1} \binom{k}{\nu} d_{k-\nu}\, F_\nu^{(0)}(x). \tag{96}$$

Again, for $k = 0$ the solution is $F_0^{(0)}(x) = x+1$, and for $k \geq 1$ we get $F_k^{(0)}(0) = 0$, $F_k^{(0)}(1) = 1$ and $F_k^{(0)}\left(\frac{1}{2}\right) = 1 + d_k$. Condition (12) is satisfied, so that all of these functions are continuous.

Now define a sequence $(q_k^{(0)})$ of polynomials of degree k and with leading coefficient 1 by

$$q_0^{(0)}(u) = 1 \quad \text{and} \quad q_{k+1}^{(0)}(u) = 2\,(u-1)\, q_k^{(0)}(u) - u\, q_k^{(0)}(u-1) \text{ for } k \geq 0. \tag{97}$$

If we write $q_k^{(1)}(u) := q_k(u)$ for the polynomials from the previous section, then we get the following Trollope-Delange analog for the power sums of the zero-counting function:

$$\frac{1}{n}S_k^{(0)}(n) = \frac{1}{x+1}\sum_{\nu+\mu+\lambda=k}\frac{k!}{\nu!\,\mu!\,\lambda!}\,2^{-\nu-\mu}\,q_\nu^{(0)}(\log_2 n)\,q_\mu^{(1)}(-\log_2(x+1))\,F_\lambda^{(0)}(x) \tag{98}$$

where $x = \frac{n-p(n)}{p(n)}$.

In particular, for $k=1$, we get

$$\frac{1}{n}S_1^{(0)}(n) = \frac{1}{2}\log_2 n - 1 - \frac{1}{2}\log_2(x+1) + \frac{F_1^{(0)}(x)}{x+1}, \tag{99}$$

which is also, in a slightly different form, one of the results in [9]. Note that

$$F_1^{(0)}(x) = x + \frac{1}{2}T(x), \tag{100}$$

because both sides of the equality are continuous functions which satisfy the same system of functional equations.

For $k=2$, we get

$$\begin{aligned}\frac{1}{n}S_2^{(0)}(n) &= \frac{1}{4}(\log_2 n)^2 + \left(-\frac{3}{4} - \frac{1}{2}\log_2(x+1) + \frac{F_1^{(0)}(x)}{x+1}\right)\log_2 n + \frac{1}{4}\log_2(x+1)^2\\ &\quad + 1 + \frac{3}{4}\log_2(x+1) - \frac{(\log_2(x+1)+2)}{x+1}F_1^{(0)}(x) + \frac{1}{x+1}F_2^{(0)}(x).\end{aligned} \tag{101}$$

Some papers dealing with the number of occurrences of “words” (‘0’-‘1’-strings longer than just one letter) in the binary expansions of the integers distinguish between the number of occurrences “with overhang” and those “without overhang”. Examples are [5], [12] or [11]. The quantity $S^{(0)}(n;t)$ computed in this section would correspond to the formulas “without overhang”. The corresponding quantity “with overhang” could be defined as

$$s_n^{(0)}(j) \quad := \quad \sum_{i=0}^{\log_2 p(n)} \delta_{a_i(j),0} \quad \text{for } j \geq 0,\ n \geq 1, \tag{102}$$

$$\widetilde{S}^{(0)}(n;t) \quad := \quad \sum_{j=0}^{n-1}\exp(t\cdot s_n^{(0)}(j)) \quad \text{for } t\in\mathbb{R},\ n\geq 1. \tag{103}$$

It is, however, easy to see that

$$\widetilde{S}^{(0)}(n;t) = S^{(0)}(n;t) + \left(e^t + e^{-t}\right)\left(1+e^t\right)^{\log_2 p(n)}, \tag{104}$$

so that there is no real difference between the two quantities in this case.

References

[1] Allaart, P.C., Kawamura, K.: Extreme values of some continuous nowhere differentiable functions, *Math. Proc. Camb. Philos. Soc.* **140** (2006), 269–295.

[2] Barnsley, M.F.: Fractal functions and interpolation, *Constr. Approx.* **2** (1986), 303–329.

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

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

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

[6] Girgensohn, R.: Functional equations and nowhere differentiable functions, *Aeq. Math.* **46** (1993), 243–256.

[7] Girgensohn, R.: Nowhere differentiable solutions of a system of functional equations, *Aeq. Math.* **47** (1994), 89–99.

[8] Hata, M., Yamaguti, M.: The Takagi function and its generalization, *Japan J. Appl. Math.* **1** (1984), 183–199.

[9] Krüppel, M.: Takagi's continuous nowhere differentiable function and binary digital sums, *Rostock. Math. Kolloq.* **63** (2008), 37–54.

[10] Krüppel, M.: De Rham's singular function, its partial derivatives with respect to the parameter and binary digital sums, *Rostock. Math. Kolloq.* **64** (2009), 57–74.

[11] Kobayashi, Z., Kuzumaki, T., Okada, T., Sekiguchi, T., Shiota, Y.: A probability measure which has Markov property, *Acta Math. Hungar.* **121** (2008), 45–71.

[12] Muramoto, K., Okada, T., Sekiguchi, T., Shiota, Y.: An explicit formula of subblock occurrences for the p-adic expansion, *Interdisciplinary Information Science* **8** (2002), 115–121.

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

[14] Okada, T., Sekiguchi, T., Shiota, Y.: An explicit formula of the exponential sums of digital sums, *Japan J. Indust. Appl. Math.* **12** (1995), 425–438.

[15] de Rham, G.: Sur certaines équations fonctionelles, *École Polytechnique de l'Université de Lausanne, Ouvrage publié à l'occasion de son centenaire* (1953), 95–97.

[16] Salem, R.: On some singular monotonic functions which are strictly increasing, *Trans. Amer. Math. Soc.* **53** (1943), 427–439.

[17] Sekiguchi, T., Shiota, Y.: A generalization of Hata-Yamaguti's results on the Takagi function, *Japan J. Indust. Appl. Math.* **8** (1991), 203–219.

[18] Takagi, T.: A simple example of the continuous function without derivative, *Proc. Phys. Math. Soc. Japan* **1** (1903), 176–177.

[19] Trollope, E.: An explicit expression for binary digital sums, *Math. Mag.* **41** (1968), 21–25.

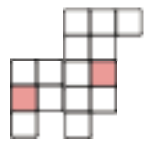

Figure 1: The functions F_1 (left) and F_2 (right).

Figure 2: The functions F_3 (left) and F_4 (right).