AofA'10 DMTCS proc. **AM**, 2010, 93–106

Analyzing a Weighted Digital Sum Variant †

Y. K. Cheung[1‡] and Mordecai Golin[2]

[1] *Computer Science Department, Courant Institute of Mathematical Sciences, New York University.* Email: ykcheung@cims.nyu.edu
[2] *Dept. of Computer Science & Engineering, Hong Kong Univ. of Science and Technology.* Email: golin@cs.ust.hk

Consider the following weighted digital sum (WDS) variant: write integer n as $n = 2^{i_1} + 2^{i_2} + \cdots + 2^{i_k}$ with $i_1 > i_2 > \cdots > i_k \geq 0$ and set $W_M(n) := \sum_{t=1}^{k} t^M 2^{i_t}$. This type of weighted digital sum arises (when $M = 1$) in the analysis of bottom-up mergesort but is not "smooth" enough to permit a clean analysis. We therefore analyze its average $TW_M(n) := \frac{1}{n} \sum_{j<n} W_M(j)$.

We show that $TW_M(n)$ has a solution of the form

$$nG_M(\lg n) + d_M \lg^M n + \sum_{d=0}^{M-1} \left(\lg^d n\right) G_{M,d}(\lg n),$$

where d_M is a constant and $G_M(u)$, $G_{M,d}(u)$'s are periodic functions with period one (given by absolutely convergent Fourier series).

Keywords: Mellin Transform, Digital Sum

1 Introduction

Exact formulas of different kinds of Digital Sums and Weighted Digital Sums (WDSs) have been studied before, e.g. in [Del75], [FGK$^+$94], [GH05] and [CFGL09]. In this paper we use Mellin transform techniques to analyze a new type of WDS. The simplest form of this sum arises in the analysis of the worst-case running time of bottom-up mergesort.

Assume[(i)] that the worst-case running time to merge two sorted lists of sizes n_1 and n_2 into one sorted list is $n_1 + n_2$.

Now, define $C_w(n)$ to be the worst-case running time of bottom-up mergesort with n elements. Bottom-up mergesort essentially splits a list of n items into two sublists, sorts each recursively, and then merges them back together. If n is a power of 2, then it splits the list into two even parts and sorts those recursively.

† Work of both authors partially supported by Hong Kong CERG grant 613507.

‡ Work done while at Department of Mathematics, Hong Kong University of Science and Technology.

(i) The actual worst-case time is $n_1 + n_2 - 1$. But, since any mergesort uses exactly $n - 1$ merges, the running time derived with cost $n_1 + n_2$ is exactly $n - 1$ more than the real worst-case running time.

This yields $C_w(2^k) = k2^k$. If n is not a power of 2 though, i.e., $n = 2^k + j$ with $1 \le j \le 2^k - 1$, then the algorithm splits the items into one list of size 2^k, one list of size j. Therefore, $C_w(n)$ satisfies:

$$\begin{aligned} C_w(2^k) &= k2^k. \\ C_w(2^k + j) &= C_w(2^k) + C_w(j) + (2^k + j), \quad \text{for } 1 \le j \le 2^k - 1. \end{aligned}$$

Panny and Prodinger [PP95] derived an exact solution for $C_w(n)$ containing a term $G(\log n)$, where $G(x)$, defined by a Fourier series, is periodic with period one. However, the Fourier series is only Cesàro summable. Furthermore, $G(x)$ is discontinuous at all dyadic points (points of the form $x = n/2^m$, where n is integer, m is non-negative integer), which are exactly the points of interest.

For a better understanding of this function, start with the binary representation of n, ignore the 0 bits and write n as the sum of descending powers of 2, i.e. $n = 2^{i_1} + 2^{i_2} + \cdots + 2^{i_k}$ with $i_1 > i_2 > \cdots > i_k \ge 0$. Iterating the above recurrence for $C_w(n)$ gives

$$C_w(n) = \sum_{t=1}^{k} i_t 2^{i_t} + \sum_{t=1}^{k} t 2^{i_t} - 2^{i_k}.$$

$\sum_{t=1}^{k} i_t 2^{i_t}$ is a type of WDS analyzed in [CFGL09]. (It also arises elsewhere, e.g., in the analysis of binomial queues [Bro78].) This motivates the analysis of the other sum, which we define as

$$W_1(n) := \sum_{t=1}^{k} t 2^{i_t}. \tag{1}$$

$W_1(n)$ is not smooth enough to be analyzed directly (see Figure 1(a)), so we instead study its average

$$TW_1(n) := \frac{1}{n} \sum_{j<n} W_1(j). \tag{2}$$

(See Figure 1(c).) As in [CFGL09] where the analysis of $\sum_{t=1}^{k} i_t 2^{i_t}$ was generalized to the analysis of $\sum_{t=1}^{k} (i_t)^M 2^{i_t}$ we may also generalize this problem by weighting the powers of 2 with polynomial weights, i.e. by defining $W_0(n) := n$ and, $\forall M \ge 1$ (see Figure 1(b)),

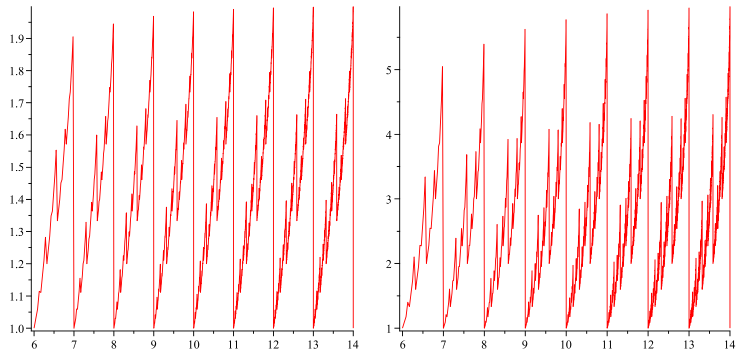
$$W_M(n) := \sum_{t=1}^{k} t^M 2^{i_t} \tag{3}$$

and then introducing the average functions (see Figure 1(d)),

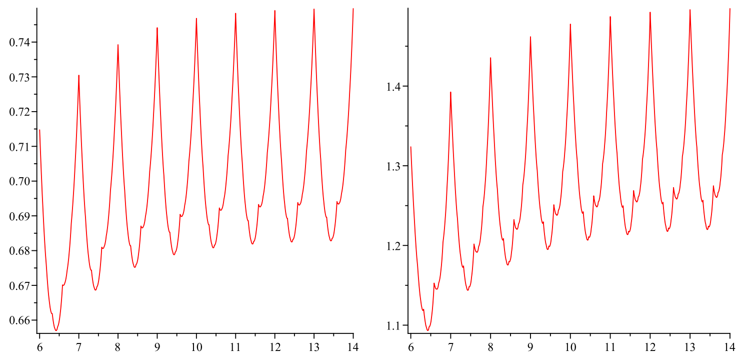
$$TW_M(n) := \frac{1}{n} \sum_{j<n} W_M(j). \tag{4}$$

We will show that $TW_M(n)$ has an exact closed-form formula, which is in the form of

$$TW_M(n) = nG_M(\lg n) + d_M \lg^M n + \sum_{d=0}^{M-1} \left(\lg^d n\right) G_{M,d}(\lg n), \tag{5}$$



(a) $W_1(n)/n$ vs. $\lg n$ (b) $W_2(n)/n$ vs. $\lg n$



(c) $TW_1(n)/n$ vs. $\lg n$ (d) $TW_2(n)/n$ vs. $\lg n$

Fig. 1: All figures are plotted on a $\lg n$ scale. (a) and (b) illustrate $W_1(n)/n$ and $W_2(n)/n$. These functions appear periodic but they are not smooth and contain "large" fluctuations, making direct analysis difficult. (c) and (d) are the averaged versions, $TW_1(n)/n$ and $TW_2(n)/n$, which are analyzed in this paper.

where d_M is a constant, $G_M(u)$ and $G_{M,d}(u)$'s are periodic functions with period one given by absolutely convergent Fourier series. The derivation of a solution in this form is done via standard techniques, e.g., the use of the Mellin-Perron Formula. What is novel here is the proof that the resulting Fourier series are absolutely convergent.

Note: Due to space considerations in this extended abstract, some straightforward proofs of Theorems and Lemmas have been omitted. These can be found in the 2nd part of [CFGL10], along with much more background information.

2 Background

2.1 The Mellin-Perron Formula

The main tools used in this paper are Dirichlet generating functions and the Mellin-Perron formula. For more background see, [Rie96, pp.13-23], [FGK$^+$94] and [FS08, pp.762-767].

Theorem 1 (The Mellin-Perron formula) *Let $\{\lambda_j\}$, $j = 1, 2, \ldots$ be a sequence and $c > 0$ lies in the half-plane of absolute convergence of $\sum_{j=1}^{\infty} \lambda_j j^{-s}$. Then for any $m \geq 1$,*

$$\frac{1}{m!}\sum_{j<n}\lambda_j\left(1-\frac{j}{n}\right)^m = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\left(\sum_{j=1}^{\infty}\frac{\lambda_j}{j^s}\right)\frac{n^s ds}{s(s+1)(s+2)\cdots(s+m)}. \tag{6}$$

In particular, when $m = 1$ and $m = 2$,

$$\frac{1}{n}\sum_{j<n}\lambda_j(n-j) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\left(\sum_{j=1}^{\infty}\frac{\lambda_j}{j^s}\right)\frac{n^s ds}{s(s+1)}, \tag{7}$$

$$\frac{1}{2n^2}\sum_{j<n}\lambda_j(n-j)^2 = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\left(\sum_{j=1}^{\infty}\frac{\lambda_j}{j^s}\right)\frac{n^s ds}{s(s+1)(s+2)}. \tag{8}$$

Rewriting $TW_M(n)$ in terms of summations in the form of the left hand side of (7) and (8) will let us use the Mellin-Perron formula to evaluate $TW_M(n)$ by evaluating the associated line integrals instead, as shown in the following lemma:

Lemma 1 *Let A be a function with $A(0) = 0$ and $TA(n) := \frac{1}{n}\sum_{j<n} A(j)$. Then*

$$TA(n) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\left(\sum_{j=1}^{\infty}\frac{\nabla A(j)}{j^s}\right)\frac{n^s ds}{s(s+1)}, \tag{9}$$

where the backward difference function ∇A *is defined by* $\nabla A(j) = A(j) - A(j-1)$ *for any function A, and $c > 0$ lies in the half-plane of absolute convergence of* $\sum_{j=1}^{\infty} \nabla A(j) j^{-s}$.

Note that in the right hand side of (9), $\sum_{j=1}^{\infty} \nabla A(j) j^{-s}$, the *Dirichlet generating function* (DGF) of $\{\nabla A(j)\}$, is the only factor depending upon $\{\lambda_j\}$.

2.2 Useful Facts Involving the Riemann-Zeta Function

The Riemann-Zeta function $\zeta(s)$ will appear in the kernels of the integrals that we will be evaluating. We therefore list some basic facts concerning $\zeta(s)$ [Tit86, WW63] that we will need.

First, $\zeta(s)$ is analytic in the whole complex plane with the exception of a simple pole at $s = 1$ with residue 1.

Next, in [FGK+94], Flajolet et. al. proved the identity

$$\frac{1}{2\pi i}\int_{-1/4-i\infty}^{-1/4+i\infty}\zeta(s)\frac{n^s ds}{s(s+1)} = 0. \tag{10}$$

By mimicking their proof, we can easily prove the similar identity

$$\frac{1}{2\pi i}\int_{-5/4-i\infty}^{-5/4+i\infty}\zeta(s)\frac{n^s ds}{s(s+1)(s+2)} = 0. \tag{11}$$

Finally, when integrating $\zeta(s)$, the following asymptotic bounds will be useful:

Lemma 2 *[WW63] If $s = \sigma + it$, where $\sigma, t \in \mathbb{R}$, the Riemann-Zeta function satisfies the bound*

$$\zeta(s) = O(|t|^{\tau(\sigma)} \log |t|) \tag{12}$$

where

$$\tau(\sigma) = \begin{cases} \frac{1}{2} - \sigma & \textit{for } \sigma \leq 0 \\ \frac{1}{2} & \textit{for } 0 \leq \sigma \leq \frac{1}{2} \\ 1 - \sigma & \textit{for } \frac{1}{2} \leq \sigma \leq 1 \\ 0 & \textit{for } 1 \leq \sigma. \end{cases} \tag{13}$$

2.3 Useful Formulae Involving Some DGFs

We will evaluate the line-integral representation of $TW_M(n)$ using the Cauchy Residue Theorem. This will require knowing the zeros and (residues of) the poles of the integral's kernel's components, especially those of the DGF.

We now list some basic results that will help us understand these.

Definition 1 *Express $n = (b_i b_{i-1} \cdots b_1 b_0)_2$ in its binary representation.*

Set $v(n) := \sum_{t=0}^{i} b_t$ to be the number of "1"s in the binary representation of n and $v_2(n)$ to be the number of trailing "0"s in the binary representation of n.

Definition 2 *$\forall M \geq 0$, denote the DGFs of $v(n)^M$ and $(v(n) + v_2(n))^M$ by*

$$V_M(s) := \sum_{j=1}^{\infty} \frac{v(j)^M}{j^s}, \qquad Z_M(s) := \sum_{j=1}^{\infty} \frac{(v(j) + v_2(j))^M}{j^s}.$$

Lemma 3 *Let n be a positive integer. Then (a) $v(2n) = v(n)$ and $v(2n+1) = v(n) + 1$; (b) $v_2(2n) = v_2(n) + 1$; (c) if n is odd, $v_2(n) = 0$; (d) $v(n) - v(n-1) = 1 - v_2(n)$.*

The following lemma gives three basic formulae expressing some special DGFs in terms of $\zeta(s)$ and $V_M(s)$.

Lemma 4 *For $M \geq 1$,*

$$\sum_{\textit{odd } j} \frac{v(j)^M}{j^s} = \left(1 - \frac{1}{2^s}\right) V_M(s). \tag{14}$$

The following two DGFs have closed-form formulae in terms of $\zeta(s)$:

$$\sum_{j=1}^{\infty} \frac{v_2(j)}{j^s} = \frac{1}{2^s - 1}\zeta(s), \quad \sum_{j=1}^{\infty} \frac{\nabla v(j)}{j^s} = \frac{2^s - 2}{2^s - 1}\zeta(s). \tag{15}$$

2.4 Absolute Convergence of Fourier Series

Evaluating the line integrals of the kernels will reduce to summations of residues at poles regularly spaced along a vertical line. These summations will best be expressed as Fourier series. For these representations to be useful, we will need to show that these Fourier series converge absolutely. Our major tools will be the following two lemmas which are quite technical but straightforward to prove.

Lemma 5 *Let $\epsilon > 0$, $\sigma_0, t_0 \in \mathbb{R}$, $t_0 \geq 1 + \epsilon$ and f be a complex function. If*

1. *f is analytic in $X = \{s = \sigma + it : \sigma \geq \sigma_0 - \epsilon, |t| \geq t_0 - \epsilon\}$ and*
2. *$\exists A, B > 0$ such that $\forall \sigma + it \in X$, $|f(\sigma + it)| = O(|t|^A \log^B |t|)$,*

then, for every fixed integer $q > 0$,

$$\forall \sigma \geq \sigma_0, \forall |t| \geq t_0, \quad |f^{(q)}(\sigma + it)| = O(|t|^A \log^B |t|).$$

Lemma 6 *Let $g(s) = L(s) f(s) \frac{n^s}{s(s+1)}$. $\forall j \in \mathbb{Z}$, set $\theta_j = \sigma + \frac{2\pi j}{\ln 2} i$. If*

1. *$\forall j \in \mathbb{Z} \setminus \{0\}$, f is analytic at $s = \theta_j$,*
2. *$\exists A < 1$, $B \geq 0$, such that for all integers positive integers q,*
3. *$|f^{(q)}(\theta_j)| = O(|j|^A \log^B |j|)$ (where the constant in the big O may depend upon q),*
4. *$L(s)$ is a meromorphic function,*
5. *$\forall j \in \mathbb{Z}$, $L(s)$ has a pole of order at most[(ii)] n_1 at $s = \theta_j$;*
   furthermore, the coefficients of the Laurent series of $L(s)$ are identical at each $s = \theta_j$,
6. *$\frac{f(s)}{s(s+1)}$ has a pole of order at most n_2 at $s = \theta_0$,*

then the sum of residues at $s = \theta_j$ can be written in the form[(iii)]

$$\sum_{j \in \mathbb{Z}} Res(g(s), s = \theta_j) = \sum_{i=n_1}^{n_1+n_2-1} \lambda_i n^\sigma \lg^i n + \sum_{i=0}^{n_1-1} F_i(\lg n) n^\sigma \lg^i n, \tag{16}$$

where the λ_i's are constants and $F_i(u)$'s are periodic functions with period one given by their Fourier series $F_i(u) = \sum_{j \in \mathbb{Z}} a_{i,j} e^{2\pi i j u}$. Furthermore, all the Fourier series $F_i(u)$ are absolutely convergent.

3 Analyzing $TW_1(n)$

We can use Lemma 1 to rewrite

$$TW_M(n) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} B_M(s) \frac{n^s ds}{s(s+1)}, \tag{17}$$

where $B_M(s)$ is the DGF of $\nabla W_M(j)$;

$$B_M(s) := \sum_{j=1}^{\infty} \frac{\nabla W_M(j)}{j^s}. \tag{18}$$

Similar to [FGK$^+$94], [Hwa98], [GH05] and [CFGL09], the main step for evaluating the integral in (17) is finding the poles of $B_M(s)$, the DGF in its kernel. The complication here is that $B_M(s)$ is not "nice" enough to permit integrating the kernel directly. We will have to split $B_M(s)$ into two parts, using the $m = 1$ case of (6) to evaluate the first part and the $m = 2$ case of (6) to evaluate the second part.

(ii) "$h(s)$ has a pole of order *at most* N at $s = s_0$", permits $h(s)$ to be analytic at $s = s_0$.

(iii) Since we are only *upper-bounding* the order of poles but do not know their exact order, λ_i may be zero and the $F_i(u)$'s may be constant functions, or even zero functions.

3.1 Deriving the DGF

We start by noticing that $B_M(s)$ can be rewritten in terms of DGFs $V_M(s)$ and $Z_M(s)$ introduced in Definition 2.

Lemma 7

$$B_M(s) = \frac{2^s - 1}{2^s - 2} V_M(s) - \frac{1}{2^s - 2} Z_M(s).$$

Proof: Observe that if n is expressed as $n = 2^{i_1} + 2^{i_2} + \cdots + 2^{i_k}$ with $i_1 > i_2 > \cdots > i_k \geq 0$, then

$$\begin{aligned} W_M(2n) &= \sum_{t=1}^{k} t^M 2^{i_t+1} = 2\sum_{t=1}^{k} t^M 2^{i_t} = 2W_M(n), \\ W_M(2n+1) &= \sum_{t=1}^{k} t^M 2^{i_t+1} + (k+1)^M = 2W_M(n) + (v(n)+1)^M. \end{aligned}$$

Recalling from Lemma 3 that $v(n) - v(n-1) = 1 - v_2(n)$ and $v(2n+1) = v(n) + 1$ gives

$$\nabla W_M(2n) = 2\nabla W_M(n) - (v(n) + v_2(n))^M \qquad \text{and} \qquad \nabla W_M(2n+1) = v(2n+1)^M.$$

Then, (14) in Lemma 4 permits writing

$$\begin{aligned} B_M(s) &= \sum_{\text{odd } j} \frac{v(j)^M}{j^s} + \sum_{l=1} \frac{2\nabla W_M(l) - (v(l) + v_2(l))^M}{(2l)^s} \\ &= \left(1 - \frac{1}{2^s}\right) V_M(s) + \frac{1}{2^{s-1}} B_M(s) - \frac{1}{2^s} Z_M(s). \end{aligned}$$

Solving for $B_M(s)$ proves the lemma. □

For the remainder of this section we will restrict ourselves to the case $M = 1$; $M > 1$ will be analyzed in the next section. For $M = 1$, we can use the identities in Lemma 4 to refine this further to $B_1(s) = V_1(s) - \frac{\zeta(s)}{(2^s-1)(2^s-2)}$. Substituting into (17) yields

$$TW_1(n) = \frac{1}{2\pi i} \int_{3-i\infty}^{3+i\infty} V_1(s) \frac{n^s ds}{s(s+1)} - \frac{1}{2\pi i} \int_{3-i\infty}^{3+i\infty} \frac{\zeta(s)}{(2^s-1)(2^s-2)} \frac{n^s ds}{s(s+1)}. \tag{19}$$

To evaluate the first integral in terms of values of the Riemann-Zeta function, we note that $V_1(s)$ is the DGF of $v(j)$, so the first integral in (19), when transformed from integral back to summation by (7), is a double summation of $v(j)$. A double summation of $v(j)$ is also a triple summation of $\nabla v(j)$, and we can write a closed-form formula for the DGF of $\nabla v(j)$ in terms of $\zeta(s)$. Equation (8) then provides an exact formula of the triple summation of $\nabla v(j)$, and we can evaluate the first integral in (19).

We now present the details. Define $TV(n) := \frac{1}{n} \sum_{j=1}^{n} \sum_{i=1}^{j-1} v(i)$. Algebraic manipulations permit writing $TV(n)$ in two different ways:

$$TV(n) = \frac{1}{n} \sum_{k<n} v(k)(n-k) \quad \text{and} \quad TV(n) = \frac{1}{n} \sum_{k<n} \nabla v(k) \left[\frac{(n-k)^2 + (n-k)}{2}\right]. \tag{20}$$

Applying (7) to the left equation in (20), yields

$$TV(n) = \frac{1}{2\pi i}\int_{3-i\infty}^{3+i\infty} V_1(s)\frac{n^s ds}{s(s+1)}, \tag{21}$$

where the right side is exactly the first integral in (19).

Applying (8) and (7) to the right equation in (20) gives the alternate expression

$$TV(n) = \frac{n}{2\pi i}\int_{3-i\infty}^{3+i\infty}\left(\sum_{j=1}^{\infty}\frac{\nabla v(j)}{j^s}\right)\frac{n^s ds}{s(s+1)(s+2)} + \frac{1}{4\pi i}\int_{3-i\infty}^{3+i\infty}\left(\sum_{j=1}^{\infty}\frac{\nabla v(j)}{j^s}\right)\frac{n^s ds}{s(s+1)}. \tag{22}$$

Setting (21) equal to (22), substituting into (19) and finally using the right equation of (15) yields a "nicer" integral representation for $TW_1(n)$.

$$\begin{aligned} TW_1(n) &= \frac{n}{2\pi i}\int_{3-i\infty}^{3+i\infty}\frac{2^s-2}{2^s-1}\zeta(s)\frac{n^s ds}{s(s+1)(s+2)} \\ &\quad + \frac{1}{4\pi i}\int_{3-i\infty}^{3+i\infty}\frac{2^s-2}{2^s-1}\zeta(s)\frac{n^s ds}{s(s+1)} \\ &\quad - \frac{1}{2\pi i}\int_{3-i\infty}^{3+i\infty}\frac{1}{(2^s-1)(2^s-2)}\zeta(s)\frac{n^s ds}{s(s+1)}. \end{aligned} \tag{23}$$

All the three integrals can be evaluated *exactly* using standard methods. e.g. [FGK+94], [CFGL09]. Briefly speaking, each of the integrals are evaluated by integrating the kernel along a rectangular contour in the complex plane, which can be shown to have, asymptotically, the same value as the desired integral. The integration along the closed rectangular contour can be computed exactly using the Cauchy Residue Theorem, after finding the poles and their residues of the kernel. This last step is straightforward because the poles/residues of the Riemann-Zeta function are well understood (see the references in Section 2.2).

Theorem 2

$$TW_1(n) = nF_{W,1}(\lg n) - \frac{1}{4}\lg n + F_{W,0}(\lg n). \tag{24}$$

where $F_{W,1}(u)$ and $F_{W,0}(u)$ are two absolutely convergent Fourier series, whose coefficients are given by

$$\begin{aligned} F_{W,1}(u) &= \frac{\ln\pi - \gamma_0 + 2\ln 2}{4\ln 2} - \frac{1}{2\ln 2}\sum_{j\in\mathbb{Z}\setminus\{0\}}\left(\frac{2\zeta(\beta_j)}{\beta_j(\beta_j+1)(\beta_j+2)} + \frac{\zeta(\alpha_j)}{\alpha_j(\alpha_j+1)}\right)e^{2\pi iju} \\ F_{W,0}(u) &= \frac{2-2\ln\pi - 5\ln 2}{8\ln 2} + \frac{1}{2\ln 2}\sum_{j\in\mathbb{Z}\setminus\{0\}}\frac{\zeta(\beta_j)}{\beta_j(\beta_j+1)}e^{2\pi iju}. \end{aligned}$$

4 Analyzing $TW_M(n)$ for $M > 1$

We will now evaluate (17) for $M > 1$. The DGF $B_M(s)$ for $M > 1$ has a more complicated structure than $B_1(s)$. We will therefore have to introduce new techniques to evaluate it.

4.1 Properties of Poles of the DGF

To start, we will need the following semi-recursive formula of $B_M(s)$. Again, it can be derived through straightforward manipulations using the definitions and Lemmas of Section 2.3.

Lemma 8 *For $M > 1$, $B_M(s)$ satisfies*

$$B_M(s) = \frac{2^s}{2^s - 1}\zeta(s) + \frac{1}{2^s - 1}\sum_{r=1}^{M-1}\binom{M}{r}B_r(s) - \frac{2^s}{2^s - 2}\sum_{r=1}^{M}\binom{M}{r}R_r(s), \tag{25}$$

where $R_r(s)$ is defined as

$$R_r(s) := \sum_{i=1}^{\infty} v(i)^r \left[\frac{1}{(2i)^s} - \frac{1}{(2i+1)^s}\right]. \tag{26}$$

Denote

$$I_r(s) := -\frac{1}{2}\sum_{i=1}^{\infty} v(i)^r \left(\frac{1}{(2i)^s} - \frac{2}{(2i+1)^s} + \frac{1}{(2i+2)^s}\right). \tag{27}$$

Standard algebraic manipulations, e.g. those in [Hwa98, pp.536], let us rewrite a summation of this form as an integral, making its analytic properties more apparent.

$$I_r(s) = \frac{s}{2^s}\int_1^{\infty} \frac{v(\lfloor x \rfloor)^r}{x^{s+1}}\xi(x)dx, \tag{28}$$

where $\xi(x) = -\frac{1}{2}$ when $0 \le x - \lfloor x \rfloor < \frac{1}{2}$ and $\xi(x) = \frac{1}{2}$ otherwise.

After noting $R_r(s)$ can be expressed in terms of $I_r(s)$ and another function well studied in [GH05], Lemma 8 can be rewritten in terms of $I_k(s)$ and $\zeta(s)$.

Lemma 9

$$B_M(s) = \frac{P_{M,1}(2^s)}{(2^s - 1)^M}\zeta(s) + \sum_{k=1}^{M} \frac{P_{M,2,k}(2^s)}{(2^s - 1)^{M-k}(2^s - 2)} I_k(s), \tag{29}$$

where $P_{M,1}(x)$ and $P_{M,2,k}(x)$ are two polynomials, with $P_{1,1}(1) = \frac{1}{2}$, $P_{M,1}(1) = MP_{M-1,1}(1) - M!/2^M$ for $M \ge 2$ and $P_{M,2,k}(0) = 0$ for $k = 1, 2, \cdots, M$.

This now permits identifying the poles of $B_M(s)$ inside Γ. (See Figure 2 for illustration.)

Corollary 1 *For $M \ge 1$, The singularities of $B_M(s)$ inside Γ are*
(i) poles of order at most 1 at $s = 1$ and $s = \alpha_j$; *(ii) poles of order M at $s = 0$ and $s = \beta_j$.*

Hence, the singularities of $\frac{B_M(s)n^s}{s(s+1)}$ inside Γ are

(i) poles of order at most 1 at $s = 1$ and $s = \alpha_j$;

(ii) a pole of order $M + 1$ at $s = 0$; and

(iii) poles of order M at $s = \beta_j$.

Proof: (29) permits us to identify the singularities by working through the various terms and recalling that $I_k(s)$ is analytic when $\Re(s) > -1$.

The recurrence relations $P_{M,1}(1) = MP_{M-1,1}(1) - M!/2^M$ with initial condition $P_{1,1}(1) = 1/2$ give $P_{M,1}(1) > 0$ for $M \geq 1$. Hence at $s = 0, \beta_j$, $P_{M,1}(2^s)/(2^s-1)^M$ has poles of order *exactly* M, while $\zeta(s)$ is analytic (but is not zero).

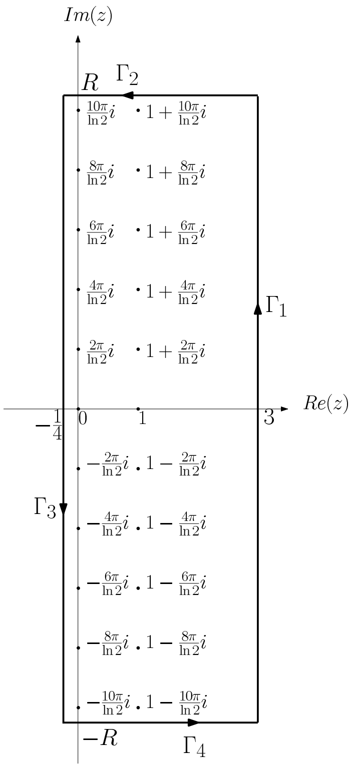
At $s = \alpha_j$, $P_{M,1}(2^s)/(2^s-1)^M$ and $\zeta(s)$ are all analytic.
At $s = 1$, $P_{M,1}(2^s)/(2^s-1)^M$ is analytic, but $\zeta(s)$ has a simple pole.

At $s = 0, \beta_j$, the order of poles of $P_{M,2,k}(2^s)(2^s-1)^{-(M-k)}(2^s-2)^{-1}I_k(s)$ is *at most* $M-1$.

At $s = 1, \alpha_j$, $P_{M,2,k}(2^s)(2^s-1)^{-(M-k)}(2^s-2)^{-1}I_k(s)$ has poles of order at most 1 (due to the term $(2^s-2)^{-1}$). □

4.2 A Formula for $TW_M(n)$

To evaluate the integral in (17), we consider integrating the kernel on the rectangular contour $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \cup \Gamma_4$, defined in Figure 2.



$$\begin{aligned} \Gamma_1 &= \{3 + iy : -R \leq y \leq R\}, \\ \Gamma_2 &= \{x + iR : -1/4 \leq x \leq 3\}, \\ \Gamma_3 &= \{-1/4 + iy : -R \leq y \leq R\}, \\ \Gamma_4 &= \{x - iR : -1/4 \leq x \leq 3\}. \end{aligned} \tag{30}$$

Fig. 2: Contour Γ. The dots represent the poles of $\frac{B_M(s)n^s}{s(s+1)}$ inside Γ.

We then show in the two following lemmas that the integrals along the top, bottom and left contours vanish as $R \to \infty$. To prove these two lemmas, we need two basic observations. Suppose $H(s) = P(2^s)(2^s-1)^{-N_1}(2^s-2)^{-N_2}$, where P is a polynomial and N_1, N_2 are non-negative integers.

Fact 1: When $\Re(s) < 0$, $H(s)$ can be expressed as a power series of 2^s, and this series is absolutely and uniformly convergent on the line $\Re(s) + (-\infty, +\infty)i$. Furthermore, if $P(0) = 0$, i.e. the constant term of P is zero, then the constant term of the power series is also zero.

Fact 2: $H(s)$ is bounded along the line segment $(-1/4, 3) + iR_j$ independently of j.

Lemma 10

$$\lim_{j\to\infty} \int_{-1/4+iR_j}^{3+iR_j} \frac{B_M(s)n^s}{s(s+1)} ds = 0.$$

Proof: For $s \in (-1/4, 3) + iR_j$, Grabner and Hwang [GH05] proved that

$$|I_M(s)| = O(|j|^{3/4} \log^{2M} |j|) = o(|j|).$$

Furthermore, Lemma 2 gives

$$|\zeta(s)| = O(|j|^{3/4} \log |j|) = o(|j|).$$

By (29) and Fact 2, $|B_M(s)|$ is bounded by $o(|j|)$ along $(-1/4, 3) + iR_j$. Hence, as $j \to \infty$,

$$\begin{aligned} \left| \int_{-1/4+iR_j}^{3+iR_j} \frac{B_M(s)n^s}{s(s+1)} ds \right| &\leq \int_{-1/4+iR_j}^{3+iR_j} \left| \frac{B_M(s)n^s}{s(s+1)} \right| ds \\ &\leq \int_{-1/4+iR_j}^{3+iR_j} \left(o(|j|) \times O(|j|^{-2}) \times n^3 \right) ds \\ &= \frac{13}{4} o(|j|^{-1}) n^3 \to 0 \end{aligned}$$

□

Lemma 11 *For any positive integer M,*

$$\int_{-1/4-i\infty}^{-1/4+i\infty} B_M(s) \frac{n^s ds}{s(s+1)} = 0.$$

Proof: Grabner and Hwang [GH05] proved the bound

$$\left| I_r\left(-\frac{1}{4} + it\right) \right| = O(|t|^{3/4} \log^{2r} |t|) = O(|t|^{3/4+\delta})$$

for any $\delta > 0$. we can then plug this into a result of Hwang [Hwa98][Proposition 2] to get

$$\int_{-1/4-i\infty}^{-1/4+i\infty} \frac{(2^k n)^s}{s(s+1)} I_r(s) ds = 0 \tag{31}$$

for positive integers k, n, r.

(29) shows that $B_M(s)$ can be expressed in the form of

$$B_M(s) = \frac{P_{M,1}(2^s)}{(2^s-1)^M}\zeta(s) + \sum_{k=1}^{M} \frac{P_{M,2,k}(2^s)}{(2^s-1)^{M-k}(2^s-2)} I_k(s),$$

while $P_{M,2,k}(0) = 0$. By Fact 1, when $\Re(s) = -1/4$, $P_{M,1}(2^s)(2^s-1)^{-M}$ and $P_{M,2,k}(2^s)(2^s-1)^{-(M-k)}(2^s-2)^{-1}$ can be expressed as power series of 2^s, and the power series for $P_{M,2,k}(2^s)(2^s-1)^{-(M-k)}(2^s-2)^{-1}$ have zero constant terms. Hence, when $\Re(s) = -1/4$, we may rewrite $B_M(s)$ to be

$$B_M(s) = \sum_{j=0}^{\infty} p_j 2^{js}\zeta(s) + \sum_{k=1}^{M}\sum_{j=1}^{\infty} q_{k,j} 2^{js} I_k(s)$$

for some $\{p_j\}$ and $\{q_{k,j}\}$. Hence

$$\begin{aligned} &\int_{-1/4-i\infty}^{-1/4+i\infty} B_M(s)\frac{n^s ds}{s(s+1)} \\ = \; &\int_{-1/4-i\infty}^{-1/4+i\infty} \left(\sum_{j=0}^{\infty} p_j (2^j n)^s \zeta(s) + \sum_{k=1}^{M}\sum_{j=1}^{\infty} q_{k,j}(2^j n)^s I_k(s)\right) \frac{ds}{s(s+1)}. \end{aligned}$$

However, the power series $\sum_{j=0}^{\infty} p_j(2^j n)^s$ and $\sum_{j=1}^{\infty} q_{k,j}(2^j n)^s$ are uniformly convergent on $-1/4 + (-\infty,\infty)i$, by Fact 1. This allows interchange of the integral sign and the summation signs.

Hence, $\int_{-1/4-i\infty}^{-1/4+i\infty} B_M(s)\frac{n^s ds}{s(s+1)}$ can be expressed as a series, in which each term is either a constant multiplied by an integral in the form of (31), or a constant multiplied by an integral in the form of (10). □

Theorem 3

$$TW_M(n) = nG_M(\lg n) + d_M \lg^M n + \sum_{d=0}^{M-1}\left(\lg^d n\right) G_{M,d}(\lg n), \tag{32}$$

where d_M is a constant, $G_M(u)$ and $G_{M,d}(u)$'s are periodic functions with period one given by absolutely convergent Fourier series.

Proof: Consider the contour Γ in Figure 2, taking $R \to \infty$. Lemma 10 and Lemma 11 show that $\frac{1}{2\pi i}\int_{\Gamma_q} \frac{B_M(s)n^s}{s(s+1)} ds = 0$ for $q = 2,3,4$. Hence

$$TW_M(n) = \frac{1}{2\pi i}\int_{\Gamma_1} B_M(s)\frac{n^s}{s(s+1)} ds,$$

is the sum of residues at the poles of $\frac{B_M(s)n^s}{s(s+1)}$ inside Γ, by the Cauchy Residue Theorem.

By Lemma 2, we have the bound $|\zeta(\sigma+it)| = O(|t|^{1/2+\epsilon}\log|t|)$ when $\sigma \ge -\epsilon$ for sufficiently small ϵ. Grabner and Hwang [GH05] also proved that $|I_r(\sigma+it)| = O(|t|^{1/2+\epsilon}\log^{2r}|t|)$ when $\sigma \ge -\epsilon$ for sufficiently small ϵ. Hence by Lemma 5,

$$|\zeta^{(q)}(\alpha_j)|,\ |\zeta^{(q)}(\beta_j)| = O(|j|^{1/2+\epsilon}\log|j|)$$

and

$$|I_r^{(q)}(\alpha_j)|,\ |I_r^{(q)}(\beta_j)| = O(|j|^{1/2+\epsilon}\log^{2r}|j|)$$

for any fixed positive integer q.

$B_M(s)$ can be expressed in the form of (29). Knowing that each function in the form of $P(2^s)(2^s - 1)^{-N_1}(2^s - 2)^{-N_2}$ will have a Laurent series with identical coefficients at $\theta_j = \sigma + \frac{2\pi j}{\ln 2}i$ for any fixed σ, together with the results from the last paragraph and Corollary 1, we use Lemma 6 when $\sigma = 0, 1$ to derive

$$\sum_{j\in\mathbb{Z}} \operatorname{Res}\left(\frac{B_M(s)n^s}{s(s+1)}, s = \alpha_j\right) = nG_M(\lg n)$$

and

$$\sum_{j\in\mathbb{Z}} \operatorname{Res}\left(\frac{B_M(s)n^s}{s(s+1)}, s = \beta_j\right) = d_M \lg^M n + \sum_{d=0}^{M-1}\left(\lg^d n\right) G_{M,d}(\lg n),$$

where $G_M(u)$ and $G_{M,d}(u)$'s are periodic functions with period one given by absolutely convergent Fourier series. □

References

[Bro78] Mark R. Brown. Implementation and analysis of binomial queue algorithms. *SIAM Journal on Computing*, 7(3):298–319, 1978.

[CFGL09] Y. K. Cheung, Phillippe Flajolet, Mordecai Golin, and C. Y. James Lee. Multidimensional divide-and-conquer and weighted digital sums (extended abstract). In *ANALCO*, pages 58–65, 2009.

[CFGL10] Y. K. Cheung, Phillippe Flajolet, Mordecai Golin, and C. Y. James Lee. Multidimensional divide-and-conquer and weighted digital sums. *arXiv: 1003.0150 [cs.DS]*, 2010.

[Del75] Hubert Delange. Sur la fonction sommatoire de la fonction somme des chiffres. *Enseignement Math.*, 21(2):31–47, 1975.

[FGK+94] Phillippe Flajolet, Peter J. Grabner, Peter Kirschenhofer, Helmut Prodinger, and Robert F. Tichy. Mellin transforms and asymptotics: Digital sums. *Theoretical Computer Science*, 123(2):291–314, 1994.

[FS08] Phillippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2008.

[GH05] Peter J. Grabner and Hsien-Kuei Hwang. Digital sums and divide-and-conquer recurrences: Fourier expansions and absolute convergence. *Constructive Approximation*, 21(2):149–179, 2005.

[Hwa98] Hsien-Kuei Hwang. Asymptotics of divide-and-conquer recurrences: Batcher's sorting algorithm and a minimum euclidean matching heuristic. *Algorithmica*, 22(4):529–546, 1998.

[PP95] Wolfgang Panny and Helmut Prodinger. Bottom-up mergesort - a detailed analysis. *Algorithmica*, 14(4):340–354, 1995.

[Rie96] Marko R. Riedel. Applications of the mellin-perron formula in number theory. Master's thesis, University of Toronto, 1996.

[Tit86] Edward C. Titchmarsh. *The Theory of the Riemann Zeta-Function*. Oxford Science Publications, 1986.

[WW63] George N. Watson and Edmund T. Whittaker. *A Course of Modern Analysis*. Cambridge University Press, 1963.