SORTING ALGORITHMS FOR BROADCAST COMMUNICATIONS: MATHEMATICAL ANALYSIS

PETER J. GRABNER† AND HELMUT PRODINGER

ABSTRACT. We study three algorithms for sorting under the broadcast communication model. Two algorithms are based on the maximum finding and the loser selection strategy, the third one is the naïve strategy based on loser selection only. While a precise description of the algorithms can be found in the paper [15], we concentrate here on the mathematical aspects of the analysis. From these elementary considerations it follows that the average time complexities of the non-naïve algorithms are $\Theta(n)$. We give precise expressions for the constants involved by writing them as contour integrals involving zeta functions. The numerical evaluation via residues leads to slowly converging series, and the acceleration of them is a nontrivial task that is done in a slightly more general fashion in order to fit all the applications.

1. INTRODUCTION

The purpose of this paper is two-fold: firstly we discuss a probabilistic model for conflict resolution in broadcast communication and secondly we present a method to compute certain integrals which occur in the asymptotic analysis of this model (and also elsewhere) numerically to a high precision.

Let us first describe the model: Assume that n persons communicate via a channel. Several algorithms related to the transmitted data ("numbers") are of interest. A particularly simple one is to find the maximum of the n items (communicated by the n persons). This algorithm was formulated and analyzed in [13, 5]. A more complex task is sorting: in the end every party should have a complete sorted list of the n items. This implies of course that every item has to be communicated at some stage.

We will analyze three algorithms to achieve this; the interest is in the average number of discrete time slots needed for communicating data (number of "rounds").

When more than one person sends at a time, a *conflict* arises. The conflict resolution schemes that are considered here are all based on coin flippings (with a fair coin): The persons involved in a conflict flip a coin, and those who flipped tail step out and wait until those who flipped head have resolved their conflict.

This principle is often described as "selection of a *loser*", see [6, 11], or as *leader election*, see [12, 3, 8]. (The loser is determined by consecutive rounds of coin flippings; those who flip heads create smaller and smaller populations until a single person remains. There is a special rule to be applied however when nobody flips head; then this round has to be

Date: January 16, 2001.

1991 *Mathematics Subject Classification.* Primary: 68Q25; Secondary: 05A15, 05A16, 60C05.

† This author is supported by the START-project Y96-MAT of the Austrian Science Foundation.

repeated.) The loser then successfully transmits his data. By different recursive strategies we obtain the variants of the algorithm studied in this paper.

Section 2: The first algorithm uses the maximum finding strategy as studied in [13, 5]. Finding the maximum needs several rounds of recursive coin flippings; after each successful broadcast of a number only the persons with larger numbers than the announced one go on recursively. Thus, in the process of finding the maximum, several data items have already been transmitted successfully (and are known to all participants). The remaining set of persons, who have not yet transmitted their data successfully, follows the same strategy recursively, until no persons are left. See [14] for a more detailed description of the algorithm.

Section 3: The second algorithm does not use the maximum finding strategy, but loser selection instead. The loser transmits his number, and the participants are split into two subfiles (larger resp. smaller numbers). These two subfiles are processed recursively, until all persons have broadcasted. See [15] for a detailed description.

Section 4: Finally, we discuss a very naïve strategy, where we only record the communication provided by the loser, and find the next loser from the reduced set of $n-1$ persons, etc. This is naturally slower.

Loser selection works by coin flippings; the first two algorithms use additionally a splitting à la binary search, whereas the naïve strategy only works sequentially. It turns out that the second algorithm is the best, the intuitive reason being perhaps that it is good to split the file as soon as possible, i. e. when the first loser has been found.

We remark that in our probabilistic model we use independent unbiased coins for selecting the losers. Furthermore, all permutations of data are equally likely (as it si customary when analyzing sorting algorithms).

The average number of rounds to select a loser is given by the recursion

$$S_n(1-2^{-n}) = 1 + 2^{-n}\sum_{k=0}^{n}\binom{n}{k}S_k, \quad n\ge 2, \qquad S_0, S_1 = 0.$$

(Note that this algorithm itself is recursive).

From [11] we know the solution

$$S_n = 1 - \sum_{k=1}^{n-1}\binom{n}{k}\frac{B_k}{2^k-1}, \qquad n\ge 2$$

with Bernoulli numbers B_k.

There is also the very useful representation of S_n as a contour integral, viz.

$$S_n = 1 - \frac{1}{2\pi i}\int_{\frac12-i\infty}^{\frac12+i\infty}\frac{\Gamma(1-s)\Gamma(n+1)\zeta(1-s)}{\Gamma(n+1-s)(2^s-1)}\,ds.$$

See [4] for some background on such integrals.

The average number of rounds M_n to find the maximum [13, 5] is given by

$$M_n = 1 + \sum_{k=0}^{n-1} \frac{2B_k}{1-2^{-k-1}} \left[\binom{n}{k+1} - 1 \right], \qquad n \geq 0;$$

it was also proved that

$$M_n - M_{n-1} = \frac{1}{2\pi i} \int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} \frac{\Gamma(n)\Gamma(1-s)}{\Gamma(n-s)} \frac{2\zeta(1-s)}{1-2^{-s-1}} \, ds. \tag{1}$$

Several parameters of the probabilistic model described above can be expressed in terms of series involving the sequences M_n and S_n. These series turn out to converge very slowly so that they are not even suitable to compute three digits of the values in question. In the specific case here the second possibility to compute these constants is also not applicable: deform the line of integration in the contour integral representation and collect residues (cf. section 5). Here the resulting series is again converging too slowly. In section 6 we will describe how to use the integral representation to obtain a rapidly converging algorithm for numerical computation of these integrals. This algorithm is formulated rather general and could be applied to similar problems immediately.

Remark. If the only task were to *transmit* all the elements, then the most obvious strategy is the conflict resolution scheme as described in [9]; all participants try to broadcast in the beginning, and then the conflict is resolved by consecutive rounds of coin flippings. The average number of rounds for this procedure is (apart from small fluctuations) asymptotically equal to $\frac{2}{\log 2} n$; the constant $\frac{2}{\log 2} = 2.885390082$ is better than the constants for the algorithms in this paper. However, in this way the data arrive in random order, whereas the strategies in this paper produce the sorted file (one can think about the data as arranged in the Quicksort tree [12]).

2. The maximum finding strategy

Denote by T_n the average number of rounds with this method. Shiau and Yang [14] argue like this: There are extra costs of M_n for finding the maximum on the first level of recursion. The first successful broadcast, which can be any element with the same probability $\frac{1}{n}$, splits the file into a subfile of $n-k-1$ smaller elements to which the recursive strategy is applied at a later stage, with costs T_{n-k-1}. In the file of k larger elements, the search for the maximum is still continuing, but we have taken care already for the extra cost for maximum searching, thus a contribution $T_k - M_k$. If $k = 0$, it costs however 0, and if $k = 1$, it costs 1, so that we have the recursion for $n \geq 2$

$$T_n = M_n + \frac{1}{n} \sum_{k=0}^{n-1} T_k + \frac{1}{n} \sum_{k=2}^{n-1} (T_k - M_k) + \frac{1}{n},$$

with $T_0 = 1$ and $T_1 = 1$.

We multiply that by n:

$$nT_n = nM_n + \sum_{k=0}^{n-1} T_k + \sum_{k=2}^{n-1}(T_k - M_k) + 1,$$

write this with n replaced by $n-1$:

$$(n-1)T_{n-1} = (n-1)M_{n-1} + \sum_{k=0}^{n-2} T_k + \sum_{k=2}^{n-2}(T_k - M_k) + 1,$$

and take differences:

$$nT_n - (n-1)T_{n-1} = nM_n - (n-1)M_{n-1} + T_{n-1} + T_{n-1} - M_{n-1},$$

or (replacing n by k)

$$\frac{T_k}{k+1} - \frac{T_{k-1}}{k} = \frac{M_k - M_{k-1}}{k+1}.$$

Summing this from $k = 3, \dots, n$ we get

$$\frac{T_n}{n+1} - \frac{T_2}{3} = \sum_{k=3}^{n} \frac{M_k - M_{k-1}}{k+1}$$

or

$$T_n = \frac{13}{6}(n+1) + (n+1)\sum_{k=3}^{n} \frac{M_k - M_{k-1}}{k+1} \quad \text{for } n \geq 2.$$

Since asymptotics of M_n are well–known [5] we know *a priori* that $T_n \sim A_1 \cdot n$, where the constant A_1 is given by

$$A_1 = \frac{13}{6} + \sum_{k=3}^{\infty} \frac{M_k - M_{k-1}}{k+1}.$$

(With more effort, one could derive a full asymptotic expansion; see section 6 for such computations.)

Using (1) we obtain

$$\sum_{k=3}^{\infty} \frac{1}{k+1}(M_k - M_{k-1}) = \frac{1}{2\pi i} \int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} \sum_{k=3}^{\infty} \frac{\Gamma(k)\Gamma(1-s)}{(k+1)\Gamma(k-s)} \frac{2\zeta(1-s)}{1-2^{-s-1}}\, ds. \tag{2}$$

Define

$$f(s) := \Gamma(1-s) \sum_{n=3}^{\infty} \frac{\Gamma(n)}{(n+1)\Gamma(n-s)} = -\frac{6s^2 - 3s - 5}{6(s-1)} - s(s+1)\psi'(-s), \tag{3}$$

where as usual $\psi(s) = \frac{\Gamma'(s)}{\Gamma(s)}$. Equation (3) is a immediate consequence of the following chain of equations (cf. [2]):

$$\begin{aligned}
\psi'(z) &= -\int_0^1 \frac{t^{z-1}}{1-t}\log t\,dt = -\int_0^1 \frac{(1-t)^{z-1}}{t}\log(1-t)\,dt \\
&= \sum_{n=1}^{\infty}\frac{1}{n}\int_0^1 (1-t)^{z-1}t^{n-1}\,dt = \sum_{n=1}^{\infty}\frac{1}{n}\frac{\Gamma(z)\Gamma(n)}{\Gamma(n+z)} \qquad (4)\\
&= \frac{1}{z} + \sum_{n=2}^{\infty}\left(\frac{\Gamma(z)\Gamma(n-1)}{\Gamma(n+z)} - \frac{\Gamma(z)\Gamma(n-1)}{n\Gamma(n+z)}\right) = \frac{1}{z} + \frac{1}{z^2} - \sum_{n=1}^{\infty}\frac{\Gamma(z)\Gamma(n)}{(n+1)\Gamma(n+z)}.
\end{aligned}$$

As $\psi(s)$ is given by the uniformly convergent expansion

$$\psi'(s) = \sum_{n=0}^{\infty}\frac{1}{(n+s)^2}$$

it is easy to see that ψ has double poles at the negative integers. For further details on ψ we refer to [2, 10, 16]. Furthermore, we note that (cf. [1]) for $|\arg(-s)| < \pi$

$$\psi'(-s) = -\frac{1}{s} + \frac{1}{2s^2} - \frac{1}{6s^3} + \frac{1}{30s^5} + \mathcal{O}(|s|^{-7}).$$

Using the identity

$$\psi'(s) + \psi'(-s) = \frac{\pi^2}{\sin^2 \pi s} + \frac{1}{s^2}$$

we can find an estimate valid for all s with $\forall n \in \mathbb{N}_0 : |s-n| > \varepsilon$:

$$f(s) = \mathcal{O}\left(\frac{1}{|s|^2} + \frac{1}{\varepsilon^2}e^{-2\pi|\Im s|}\right). \qquad (5)$$

Our goal is to compute the integral

$$\frac{1}{2\pi i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} f(s)\frac{2\zeta(1-s)}{1-2^{-s-1}}\,ds.$$

Expanding the ζ–function and the geometric series it is

$$\sum_{k\geq 1}\frac{1}{k}\sum_{j\geq 0}2^{1-j}\frac{1}{2\pi i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} f(s)\left(\frac{k}{2^j}\right)^s\,ds.$$

For the evaluation of the integral we need now

$$\Phi(x) := \frac{1}{2\pi i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} f(s)x^s\,ds = \begin{cases} 0 & \text{for } x \geq 1 \\ \dfrac{(3-x)(1+x+x^2)}{3(1-x)^2} + \dfrac{2x\log x}{(1-x)^3} & \text{for } 0 < x < 1. \end{cases} \qquad (6)$$

This equation is achieved as follows (this is actually a standard technique in analytic number theory): to obtain the formula for $x \geq 1$ consider the integral over the segment $\{-\frac{1}{2}+it \mid -R \leq t \leq R\}$ and the arc $\{\sigma+it \mid \sigma^2+t^2=R^2, \sigma<-\frac{1}{2}\}$. Since the integrand has no poles in this region, the integral vanishes; now let R tend to ∞ and observe that the estimate (5) implies that the integral over the arc tends to 0. For $x<1$ we shift the line of integration to the right (technically, we again truncate the integral at $t=\pm R$ and show that the contribution tends to 0 for $R \to \infty$) and collect residues:

$$\frac{1}{2\pi i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} f(s)x^s\,ds = -\sum_{k=1}^{K}\operatorname*{Res}_{s=k}\left\{f(s)x^s\right\} + \frac{1}{2\pi i}\int_{K+\frac{1}{2}-i\infty}^{K+\frac{1}{2}+i\infty} f(s)x^s\,ds.$$

By (5) the integral is bounded by $\mathcal{O}(x^K)$, which allows the limit $K \to \infty$. Computing the residues and summing up yields the desired result.

Using (6) we obtain

$$\sum_{k\geq 1}\frac{1}{k}\sum_{j\geq 0}2^{1-j}\frac{1}{2\pi i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} f(s)\left(\frac{k}{2^j}\right)^s ds = 2\sum_{j\geq 0}\sum_{1\leq k<2^j}\frac{2^{-j}}{k}\Phi\left(\frac{k}{2^j}\right).$$

Observe that $\Phi(1)=0$, so that we can add the term $k=2^j$ to the sum.

In section 6 we will explain how to compute such sums numerically to a high precision. Therefore we get

$$A_1 = \frac{13}{6} + 2\Sigma_1 = 3.67982\,61095\,85742\,29512\,70603\,33235\,96302\,23675\ldots$$

Note that the paper [14] provides the elementary bounds $\frac{7}{2}=3.5 \leq A_1 \leq \frac{23}{6}=3.833\cdots$.

3. The splitting strategy

It is not hard to find a recursion for T_n, the average number of rounds for this method. (We use the notation T_n again, but there is no chance for confusion.) The values $T_0=1$ and $T_1=1$ are self-explanatory, and for $n \geq 2$ we have

$$T_n = 1 + S_n + \frac{1}{n}\sum_{k=0}^{n-1}(T_k + T_{n-1-k});$$

the 1 counts the initial broadcast, causing the conflict, which needs S_n to be resolved. Then, the probability is uniformly $\frac{1}{n}$ that the file splits. (See [15] for a more elaborate description.)

The traditional method to solve that is as in the previous section:

$$nT_n = n + nS_n + 2\sum_{k=0}^{n-1}T_k;$$

replace n by $n-1$

$$(n-1)T_{n-1} = n-1+(n-1)S_{n-1} + 2\sum_{k=0}^{n-2} T_k$$

and subtract to get

$$nT_n - (n+1)T_{n-1} = 1 + nS_n - (n-1)S_{n-1}.$$

Divide this by $n(n+1)$:

$$\frac{T_n}{n+1} - \frac{T_{n-1}}{n} = \frac{1}{n} - \frac{1}{n+1} + \frac{S_n}{n+1} - \frac{n-1}{n(n+1)}S_{n-1}$$

and sum this from $n = 3, \dots, N$, but write again n for N:

$$\frac{T_n}{n+1} - \frac{5}{3} = \frac{1}{3} - \frac{1}{n+1} + \sum_{k=3}^{n} \frac{S_k}{k+1} + \sum_{k=3}^{n} \Big(\frac{1}{k} - \frac{2}{k+1}\Big) S_{k-1}$$

or

$$\frac{T_n}{n+1} = 2 - \frac{1}{n+1} - \frac{S_n}{n+1} + \frac{2}{3} + 2\sum_{k=3}^{n} \frac{S_k}{k+1} - 2\sum_{k=3}^{n} \frac{S_{k-1}}{k+1}$$

from which we conclude that

$$T_n = \frac{8}{3}n + \frac{5}{3} - S_n + 2(n+1)\sum_{k=3}^{n} \frac{S_k - S_{k-1}}{k+1} \quad \text{for } n \ge 2.$$

(In [15] this is slightly wrong.) We know again *a priori* that $T_n \sim A_2 \cdot n$. Our aim is compute this constant A_2:

$$A_2 = \frac{8}{3} + 2\sum_{k=3}^{\infty} \frac{S_k - S_{k-1}}{k+1}. \tag{7}$$

From the integral formula for S_k we conclude that

$$\begin{aligned} S_k - S_{k-1} &= -\frac{1}{2\pi i}\int_{\frac12 - i\infty}^{\frac12 + i\infty} \frac{\Gamma(1-s)\Gamma(k+1)\zeta(1-s)}{\Gamma(k+1-s)(2^s-1)}\,ds + \frac{1}{2\pi i}\int_{\frac12 - i\infty}^{\frac12 + i\infty} \frac{\Gamma(1-s)\Gamma(k)\zeta(1-s)}{\Gamma(k-s)(2^s-1)}\,ds \\ &= -\frac{1}{2\pi i}\int_{\frac12 - i\infty}^{\frac12 + i\infty} \frac{s\Gamma(1-s)\Gamma(k)\zeta(1-s)}{\Gamma(k+1-s)(2^s-1)}\,ds. \end{aligned}$$

Define

$$f(s) = \Gamma(1-s)\sum_{k\ge 3} \frac{1}{k+1}\frac{\Gamma(k)}{\Gamma(k+1-s)},$$

which makes sense for $\Re s < 1$. The identity

$$f(s) = \frac{6s^2 - 15s + 4}{6(s-1)(s-2)} - \frac{1}{s} + s\psi'(-s)$$

is again a consequence of (4) and provides the analytic continuation of $f(s)$. Then

$$A_2 = \frac{8}{3} + 2 \cdot \frac{1}{2\pi i} \int\limits_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \frac{sf(s)\zeta(1-s)}{2^s - 1}\, ds.$$

Note that

$$\frac{1}{2^s - 1} = \sum_{k \geq 1} 2^{-ks}.$$

Therefore

$$\frac{1}{2\pi i} \int\limits_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \frac{sf(s)\zeta(1-s)}{2^s - 1}\, ds = \sum_{k \geq 1} \frac{1}{2\pi i} \int\limits_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} sf(s)\zeta(1-s)2^{-ks}\, ds.$$

The integral

$$\frac{1}{2\pi i} \int\limits_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} sf(s)\zeta(1-s)2^{-ks}\, ds$$

can be computed by shifting the line of integration to $\Re s = -\frac{1}{2}$ (again this is justified by the growth rate of the integrand) and expanding the ζ–function:

$$\sum_{n=1}^{\infty} \frac{1}{2\pi i} \int\limits_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} sf(s)n^{s-1}2^{-ks}\, ds.$$

Thus it remains to compute the function

$$\Psi(x) = \frac{1}{2\pi i} \int\limits_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} sf(s)x^s\, ds.$$

This is done in a similar way as above and gives

$$\Psi(x) = \begin{cases} \dfrac{x(17 - 14x + 13x^2 - 4x^3)}{6(x-1)^2} - \dfrac{x + x^2}{(x-1)^3} \log x & \text{for } 0 < x \leq 1 \\ 0 & \text{for } x > 1. \end{cases}$$

With this notation, the constant A_2 evaluates to

$$A_2 = \frac{8}{3} + 2 \sum_{j\ge 0} \sum_{1\le k<2^j} \frac{1}{k} \Psi\Big(\frac{k}{2^j}\Big) = 3.54551\,78132\,67322\,88513\,75825\,40856\,52901\,34887\ldots$$

(see the last section for such evaluations).

Note also that the paper [15] gives the bounds (after correcting the simple error) $3.33\cdots \le A_2 \le 4$.

4. The naïve strategy

The average cost of selecting *one* loser is given by $c_n = 1 + S_n$ with the quantities S_n from the introduction. Thus we have

$$c_n = 2 - \sum_{k=1}^{n-1} \binom{n}{k} \frac{B_k}{2^k - 1}.$$

In total, we need to evaluate

$$C_n = c_1 + \cdots + c_n.$$

We know

$$\sum_{k=1}^{n-1} \binom{n}{k} \frac{B_k}{2^k - 1} = \frac{1}{2\pi i} \int_{\frac12 - i\infty}^{\frac12 + i\infty} \frac{\Gamma(n+1)\Gamma(1-s)}{\Gamma(n+1-s)} \frac{\zeta(1-s)}{2^s - 1}\, ds;$$

since $c_1 = 1$ and $c_2 = 3$ this leads to

$$C_n = 2n - \sum_{k=3}^{n} \frac{1}{2\pi i} \int_{\frac12 - i\infty}^{\frac12 + i\infty} \frac{\Gamma(k+1)\Gamma(1-s)}{\Gamma(k+1-s)} \frac{\zeta(1-s)}{2^s - 1}\, ds.$$

As before we sum up the terms in the integral

$$\sum_{k=3}^{n} \frac{\Gamma(k+1)\Gamma(1-s)}{\Gamma(k+1-s)} = \frac{\Gamma(n+2)\Gamma(1-s)}{(s+1)\Gamma(n+1-s)} - \frac{6}{(s-1)(s-2)(s+1)}$$

and obtain

$$\begin{aligned} C_n = 2n &+ \frac{1}{2\pi i} \int_{\frac12 - i\infty}^{\frac12 + i\infty} \frac{6}{(s-1)(s-2)(s+1)} \frac{\zeta(1-s)}{2^s - 1}\, ds \\ &- \frac{1}{2\pi i} \int_{\frac12 - i\infty}^{\frac12 + i\infty} \frac{\Gamma(n+2)\Gamma(1-s)}{(s+1)\Gamma(n+1-s)} \frac{\zeta(1-s)}{2^s - 1}\, ds. \end{aligned} \tag{8}$$

For further reference we denote

$$I=\frac{1}{2\pi i}\int\limits_{\frac12-i\infty}^{\frac12+i\infty}\frac{6}{(s-1)(s-2)(s+1)}\frac{\zeta(1-s)}{2^s-1}\,ds.$$

Now we use

$$\frac{\Gamma(n+2)}{\Gamma(n+1-s)}=n^{s+1}\left(1-\frac{(s-3)(s+1)}{2n}+\mathcal{O}\Big(\frac{|s|^4}{n^2}\Big)\right) \tag{9}$$

(cf. [1]) to obtain

$$\begin{aligned}C_n=2n+I-&\frac{1}{2\pi i}\int\limits_{\frac12-i\infty}^{\frac12+i\infty}\Gamma(1-s)\frac{\zeta(1-s)}{2^s-1}n^{s+1}\,ds\\&+\frac{1}{2\pi i}\int\limits_{\frac12-i\infty}^{\frac12+i\infty}\Gamma(1-s)\frac{(s-3)(s+1)}{2}\frac{\zeta(1-s)}{2^s-1}n^{s}\,ds+\mathcal{O}(n^{-\frac12}).\end{aligned} \tag{10}$$

We shift the line of integration to $\Re s=-\frac12$ and obtain by calculating the residues at the poles on the line $\Re s=0$

$$\begin{aligned}C_n&=2n+I-\frac{1}{2\pi i}\int\limits_{-\frac12-i\infty}^{-\frac12+i\infty}\Gamma(1-s)\frac{\zeta(1-s)}{2^s-1}n^{s+1}\,ds\\&+\frac{1}{2\pi i}\int\limits_{-\frac12-i\infty}^{-\frac12+i\infty}\Gamma(1-s)\frac{(s-3)(s+1)}{2}\frac{\zeta(1-s)}{2^s-1}n^{s}\,ds+n\log_2 n-\frac12 n+nF(\log_2 n)\\&+\frac32\log_2 n+\frac{1}{\log 2}-\frac34+G(\log_2 n)+\mathcal{O}(n^{-\frac12}),\end{aligned}$$

where F and G denote two periodic continuous functions of period 1 and mean 0 given by their Fourier-expansions ($\chi_k=\frac{2k\pi i}{\log 2}$)

$$\begin{aligned}F(x)&=-\frac{1}{\log 2}\sum_{k\in\mathbb{Z}\setminus\{0\}}\Gamma(1-\chi_k)\zeta(1-\chi_k)e^{2k\pi ix}\\G(x)&=-\frac{1}{2\log 2}\sum_{k\in\mathbb{Z}\setminus\{0\}}(1+\chi_k)(3-\chi_k)\Gamma(1-\chi_k)\zeta(1-\chi_k)e^{2k\pi ix}.\end{aligned} \tag{11}$$

It remains to compute the integral I. This is done as follows:

$$\begin{aligned}
I &= \sum_{k=1}^{\infty} \frac{1}{2\pi i} \int_{\frac12 - i\infty}^{\frac12 + i\infty} \frac{6}{(s-1)(s-2)(s+1)} \zeta(1-s) 2^{-ks}\, ds \\
&= \sum_{k=1}^{\infty} \left(\frac{1}{2\pi i} \int_{-\frac12 - i\infty}^{-\frac12 + i\infty} \frac{6}{(s-1)(s-2)(s+1)} \zeta(1-s) 2^{-ks}\, ds - 3 \right) \\
&= \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \left(\frac{1}{2\pi i} \int_{-\frac12 - i\infty}^{-\frac12 + i\infty} \frac{6}{(s-1)(s-2)(s+1)} n^{s-1} 2^{-ks}\, ds - 3 \right).
\end{aligned}$$

As before we compute the integral

$$\frac{1}{2\pi i} \int_{-\frac12 - i\infty}^{-\frac12 + i\infty} \frac{6}{(s-1)(s-2)(s+1)} x^s\, ds = \begin{cases} 3x - 2x^2 & \text{for } 0 < x \le 1 \\ \frac{1}{x} & \text{for } x > 1. \end{cases}$$

We obtain

$$I = \sum_{k=1}^{\infty} \left(\sum_{n=1}^{2^k} \frac{1}{n} \left(3 \left(\frac{n}{2^k} \right) - 2 \left(\frac{n}{2^k} \right)^2 \right) + 2^k \sum_{n=2^k+1}^{\infty} \frac{1}{n^2} - 3 \right).$$

After summing up the first sum and some obvious cancellation we arrive at

$$I = \sum_{k=1}^{\infty} \left(2^k \sum_{n=2^k}^{\infty} \frac{1}{n^2} - 1 \right) - 2.$$

This can be computed by the methods outlined in section 6; the numerical value is given by

$$I = -1.44637\,64113\,48039\,93349\,42661\,97916\,01589\,56591\,06916\ldots.$$

This gives

$$C_n = n \log_2 n + \frac{3}{2} n + nF(\log_2 n) + \frac{3}{2} \log_2 n + B + G(\log_2 n) + \mathcal{O}(n^{-\frac12})$$

with

$$B = \frac{1}{\log 2} - \frac{3}{4} + I = -0.75368\,13704\,59076\,52613\,43415\,16914\,12375\,82324\,60962\ldots.$$

5. Mellin Integrals

One possible way to give a reasonably convergent series expansion for the value of an integral of the form

$$\frac{1}{2\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} f(s)\frac{\zeta(1-s)}{2^s-1}\,ds \tag{12}$$

would be to move the line of integration to the left. This yields

$$\frac{1}{2\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} f(s)\frac{\zeta(1-s)}{2^s-1}\,ds = \sum_{k\in\mathbb{Z}}\operatorname*{Res}_{s=\chi_k}\left\{f(s)\frac{\zeta(1-s)}{2^s-1}\right\} + \frac{1}{2\pi i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} f(s)\frac{\zeta(1-s)}{2^s-1}\,ds, \tag{13}$$

where again $\chi_k = \frac{2k\pi i}{\log 2}$. Under the growth condition $|f(s)| = \mathcal{O}(|s|^{-2})$ the last integral is zero by the same arguments as used above. Thus we have a series expansion for the integral (12)

$$\frac{1}{\log 2}\sum_{k\in\mathbb{Z}\setminus\{0\}} f(\chi_k)\zeta(1-\chi_k) + \operatorname*{Res}_{s=0}\left\{f(s)\frac{\zeta(1-s)}{2^s-1}\right\}. \tag{14}$$

Unfortunately, the rate of convergence of this sum is not much better than the rate of convergence of the sums (2) and (7). Thus this method does not yield a numerically feasible method to compute integrals of the form (12) to high precision, although it is possible to get a few digits by using (14), as pointed by one of the referees. The next section will describe how to find high precision approximations to the numerical value of such integrals.

6. How to compute the integrals numerically

In this section we discuss how to compute a sums like

$$\Sigma_1 = \sum_{j\geq 0}\sum_{1\leq k<2^j}\frac{2^{-j}}{k}\Phi\Big(\frac{k}{2^j}\Big) \quad\text{and}\quad \Sigma_2 = \sum_{j\geq 0}\sum_{1\leq k<2^j}\frac{1}{k}\Phi\Big(\frac{k}{2^j}\Big) \tag{15}$$

numerically. We assume that $\Phi\in C^{2m}(0,1]$, $\Phi(0)=D$, and

$$\frac{\Phi(t)-D}{t} = \sum_{\ell=0}^{2m} C_\ell t^\ell \log t + Q_{2m}(t) + O(t^{2m+1}\log t), \tag{16}$$

where $Q_{2m}\in C^{2m+1}[0,1]$ and C_ℓ are constants (these two properties are satisfied for the functions to which these studies will be applied later).

In the sequel we will use the Bernoulli-polynomials $P_k(t)$ defined by

$$\sum_{\ell=0}^{\infty} P_\ell(t)\frac{z^\ell}{\ell!} = \frac{ze^{zt}}{e^z-1}$$

and the Bernoulli numbers $B_\ell = P_\ell(0)$. Throughout this section we use the notation $\{x\}$ to indicate the fractional part of x.

Then we have

$$\begin{aligned}
\sum_{1\le k\le 2^j} \frac{1}{k} f\Big(\frac{k}{2^j}\Big) &= \sum_{k=1}^{2^j} \frac{1}{k}\left(f\Big(\frac{k}{2^j}\Big) - D\right) + D\sum_{k=1}^{2^j}\frac{1}{k} \\
&= \sum_{k=1}^{2^j} \frac{1}{k}\left(f\Big(\frac{k}{2^j}\Big) - D\right) \\
&\quad + D\bigg(j\log 2 + \gamma + \frac{1}{2^{j+1}} - \sum_{\ell=1}^{m}\frac{B_{2\ell}}{2\ell}\frac{1}{2^{2\ell j}} + 2^{-(2m+1)j}\int_1^\infty P_{2m+1}(\{2^j x\})x^{-2m-2}\,dx\bigg) \\
&= 2^{-j}\sum_{k=1}^{2^j}\left[\frac{2^j}{k}\left(f\Big(\frac{k}{2^j}\Big) - D\right) - \sum_{\ell=0}^{2m} C_\ell\Big(\frac{k}{2^j}\Big)^\ell \log\frac{k}{2^j}\right] \\
&\quad + \sum_{\ell=0}^{2m} C_\ell 2^{-j}\sum_{k=1}^{2^j}\Big(\frac{k}{2^j}\Big)^\ell\log\frac{k}{2^j} + D\bigg(j\log 2 + \gamma + \frac{1}{2^{j+1}} - \sum_{\ell=1}^{m}\frac{B_{2\ell}}{2\ell}\frac{1}{2^{2\ell j}} + R_j\bigg)
\end{aligned}$$

where $|R_j| \le \frac{1}{2m+1}\|P_{2m+1}\|_\infty 2^{-(2m+1)j}$.

We set $g(t) = g_{2m}(t) = \frac{f(t)-D}{t} - \sum_{\ell=0}^{2m} C_\ell t^\ell \log t$; under our assumptions on f, the function g is $2m$ times differentiable on $[0,1]$ and the $(2m+1)$–st derivative is in $L^1[0,1]$. Then we have

$$\begin{aligned}
2^{-j}\sum_{k=1}^{2^j} g\Big(\frac{k}{2^j}\Big) &= \int_0^1 g(t)\,dt + \frac{1}{2}\big(g(1)-g(0)\big)2^{-j} + \sum_{\ell=1}^{m}\frac{B_{2\ell}}{(2\ell)!}\left(g^{(2\ell-1)}(1) - g^{(2\ell-1)}(0)\right)2^{-2\ell j} \\
&\quad + \frac{1}{(2m+1)!}2^{-(2m+1)j}\int_0^1 P_{2m+1}(\{2^j x\})g^{(2m+1)}(x)\,dx \qquad (17)\\
&= \int_0^1 g(t)\,dt + \frac{1}{2}\big(g(1)-g(0)\big)2^{-j} + \sum_{\ell=1}^{m}\frac{B_{2\ell}}{(2\ell)!}\left(g^{(2\ell-1)}(1) - g^{(2\ell-1)}(0)\right)2^{-2\ell j} + R_j'
\end{aligned}$$

with $|R_j'| \le \frac{1}{(2m+1)!}\|P_{2m+1}\|_\infty\|g^{(2m+1)}\|_1$.

Finally, we compute the sums

$$2^{-j}\sum_{k=1}^{2^j}\Big(\frac{k}{2^j}\Big)^r\log\frac{k}{2^j} \qquad (18)$$

for $r \in \mathbb{N}_0$ using the Euler-MacLaurin summation formula. For this purpose we first compute the sum

$$\begin{aligned}2^{-(s+1)j}\sum_{k=1}^{2^j}k^s &= \frac{1}{s+1}+\frac{1}{2^{1+j}}+\sum_{\ell=1}^{m}\frac{B_{2\ell}}{2\ell}\binom{s}{2\ell-1}2^{-2\ell j}+2^{-js}\zeta(-s)\\ &\quad-2^{-(2m+1)j}\binom{s}{2m+1}\int_1^\infty P_{2m+1}(\{2^jx\})x^{s-2m-1}\,dx;\end{aligned}\tag{19}$$

here we have used the identity (cf. [7])

$$\zeta(-s)=-\frac{1}{s+1}+\frac{1}{2}-\sum_{\ell=1}^{m}\frac{B_{2\ell}}{2\ell}\binom{s}{2\ell-1}+\binom{s}{2m+1}\int_1^\infty P_{2m+1}(\{x\})x^{s-2m-1}\,dx,$$

which is valid for $\Re s<2m$.

Differentiating (19) and setting $s=r\in\mathbb{N}_0$ yields

$$\begin{aligned}2^{-j}\sum_{k=1}^{2^j}\left(\frac{k}{2^j}\right)^r\log\frac{k}{2^j} &= -\frac{1}{(r+1)^2}+\sum_{\ell=1}^{m}\frac{B_{2\ell}}{2\ell}\frac{d}{ds}\binom{s}{2\ell-1}\bigg|_{s=r}2^{-2\ell j}\\ &\quad+j2^{-jr}\zeta(-r)\log 2-2^{-jr}\zeta'(-r)\\ &\quad-2^{-(2m+1)j}\frac{d}{ds}\binom{s}{2m+1}\bigg|_{s=r}\int_1^\infty P_{2m+1}(\{2^jx\})x^{r-2m-1}\,dx;\end{aligned}\tag{20}$$

here we have used that $\binom{r}{2m+1}=0$ for $r<2m$. We note also that $\zeta(-r)=-\frac{B_{r+1}}{r+1}$ for $r\geq 1$ and $\zeta(0)=-\frac{1}{2}$.

Finally, we observe that

$$\frac{d}{ds}\binom{s}{2\ell-1}\bigg|_{s=r}=\begin{cases}(-1)^r\dfrac{r!(2\ell-r-2)!}{(2\ell-1)!} & \text{for } r\leq 2\ell-2\\ \dbinom{r}{2\ell-1}\displaystyle\sum_{t=0}^{2\ell-2}\frac{1}{r-t} & \text{for } r>2\ell-2.\end{cases}$$

Inserting this into (20) and estimating the integral trivially yields

$$\begin{aligned}2^{-j}\sum_{k=1}^{2^j}\left(\frac{k}{2^j}\right)^r\log\frac{k}{2^j} &= -\frac{1}{(r+1)^2}+\sum_{\ell=1}^{\lfloor\frac{r+1}{2}\rfloor}\frac{B_{2\ell}}{2\ell}\binom{r}{2\ell-1}\sum_{t=0}^{2\ell-2}\frac{1}{r-t}2^{-2\ell j}\\ &\quad+j2^{-jr}\zeta(-r)\log 2-2^{-jr}\zeta'(-r)\\ &\quad-(-1)^r\sum_{\ell=\lfloor\frac{r+3}{2}\rfloor}^{m}\frac{r!(2\ell-r-2)!}{(2\ell-1)!}\frac{B_{2\ell}}{2\ell}2^{-2\ell j}+R_{j,r}\end{aligned}\tag{21}$$

with $|R_{j,r}|\leq\frac{r!(2m-r-1)!}{(2m+1)!}2^{-(2m+1)j}\|P_{2m+1}\|_\infty$.

Extending the last sum (21) to $m+2$ for $r=2m$ and putting everything together we obtain (the $C_0 j 2^{-j}\log 2$–term in the fourth line is to compensate the difference between $\zeta(0)$ and B_1)

$$\begin{aligned}
&\sum_{k=1}^{2^j}\frac{1}{k}f\Big(\frac{k}{2^j}\Big)=D\bigg(j\log 2+\gamma+\frac{1}{2^{j+1}}-\sum_{\ell=1}^{m}\frac{B_{2\ell}}{2\ell}2^{-2\ell j}+R_j\bigg)+\\
&\int_0^1\frac{f(t)-D}{t}\,dt+\frac12\big(g(1)-g(0)\big)2^{-j}+\sum_{\ell=1}^{m}\frac{B_{2\ell}}{(2\ell)!}\left(g^{(2\ell-1)}(1)-g^{(2\ell-1)}(0)\right)2^{-2\ell j}+\\
&C_0 j2^{-j}\log 2+\sum_{r=0}^{2m}C_r\Bigg(\sum_{\ell=1}^{\lfloor\frac r2\rfloor}\frac{B_{2\ell}}{2\ell}\binom{r}{2\ell-1}\sum_{t=0}^{2\ell-2}\frac{1}{r-t}2^{-2\ell j}+j2^{-j(r+1)}\frac{B_{r+1}}{r+1}\log 2-\\
&2^{-j(r+1)}\zeta'(-r)-(-1)^r\sum_{\ell=\lfloor\frac{r+2}{2}\rfloor}^{m}\frac{r!(2\ell-r-2)!}{(2\ell-1)!}\frac{B_{2\ell}}{2\ell}2^{-2\ell j}\Bigg)+\widetilde{R}_j
\end{aligned}\tag{22}$$

with

$$\begin{aligned}
|\widetilde{R}_j|\le 2^{-(2m+1)j}\|P_{2m+1}\|_\infty&\left(\frac{D}{2m+1}+\frac{1}{(2m+1)!}\|g^{(2m+1)}\|_1+\sum_{r=0}^{2m-1}C_r\frac{r!(2m-r-1)!}{(2m+1)!}\right)\\
&+C_{2m}\left(|\zeta'(-2m)|2^{-(2m+1)j}+|B_{2m+2}|2^{-2j(m+1)}+\frac{\|P_{2m+3}\|}{2m+1}2^{-j(2m+3)}\right);
\end{aligned}$$

here we have used that

$$\int_0^1\frac{f(t)-D}{t}\,dt=\int_0^1 g(t)\,dt-\sum_{r=0}^{2m}\frac{C_r}{(r+1)^2}.$$

For the function

$$f(t)=\Phi(t)=\frac{(3-t)(1+t+t^2)}{3(1-t)^2}+\frac{2t\log t}{(1-t)^3}$$

we have $D=1$ and $C_r=(r+1)(r+2)$, and $\int_0^1\frac{f(t)-1}{t}\,dt=-\frac{11}{6}$.

For a numerical approximation of (15) we split summation at J and replace the infinite part of the sum by the asymptotic estimate (22). For $J=10$ and $m=10$ this gives an error estimate of $2.152\cdot 10^{-45}$. Thus we have

$$S_1=0.75657\,97214\,59537\,81423\,01968\,71433\,28464\,81778\,24250\,446\ldots$$

In a similar way we can treat the sum

$$S_2=\sum_{j\ge 0}\sum_{1\le k<2^j}\frac1k\Psi\Big(\frac{k}{2^j}\Big)$$

In this case $D = 0$ and $\int_0^1 \frac{\Psi(x)}{x}\,dx = 0$. For $J = 8$ and $m = 10$ we obtain an error estimate of $2.0765 \cdot 10^{-48}$. Thus we have

$$S_2 = 0.43942\,55733\,00328\,10923\,54579\,37094\,93117\,34110\,28053\ldots$$

7. Concluding remarks

The method for the numerical computation of Mellin integrals as described in section 6 can be easily generalized to integrals of the form

$$\int_{c-i\infty}^{c+i\infty} f(s)\frac{\zeta(1-s)}{A^s-1}\,ds \quad \text{for } A>1,$$

where $f(s)$ is a meromorphic function in the whole complex plane satisfying the following properties:

- $f(s) = \mathcal{O}(|s|^{-1-\varepsilon})$ for some positive ε and $|\arg(-s)| < \pi$
- $f(s)$ has no poles left of a line $\Re s = \sigma_0$
- the sum over the residues $\operatorname{Res} f(s)x^s$ converges for $0 < x < 1$.

If A is not an integer one has to use a slightly changed version of the Euler-MacLaurin summation formula:

$$\sum_{1\le k<A^n} g\left(\frac{k}{A^n}\right) = A^n\int_0^1 g(x)\,dx - \frac{1}{2}g(0) + \left(\{A^n\}-\frac{1}{2}\right)g(1)$$
$$+\sum_{\ell=1}^{m}\frac{B_{2\ell}}{(2\ell)!}\left(g^{(2\ell-1)}(1)-g^{(2\ell-1)}(0)\right)A^{-(2\ell-1)n}$$
$$+\frac{1}{(2m+1)!}\int_0^1 P_{2m+1}(\{A^n x\})g^{(2m+1)}(x)\,dx.$$

Again singularities of g have to be subtracted as shown in section 6.

Acknowledgment. We thank two anonymous referees and Hsien–Kuei Hwang for their valuable remarks.

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions.* Dover, New York, 1965.
[2] G. Andrews, R. Askey, and R. Roy. *Special Functions.* Number 71 in Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1999.
[3] J. A. Fill, H. M. Mahmoud, and W. Szpankowski. On the distribution for the duration of a randomized leader election algorithm. *Ann. Appl. Probab.*, 6:1260–1283, 1996.
[4] P. Flajolet and R. Sedgewick. Mellin transforms and asymptotics: finite differences and Rice's integrals. *Theor. Comput. Sci.*, 144(1–2):101–124, June 1995.
[5] P. Grabner and H. Prodinger. An asymptotic study of a recursion occurring in the analysis of an algorithm on broadcast communication. *Inf. Process. Lett.*, 65:89–93, 1998.

[6] P. J. Grabner. Searching for losers. *Random Structures and Algorithms*, 4:99–110, 1993.
[7] A. Ivić. *The Riemann Zeta-Function.* J. Wiley, New York, 1985.
[8] S. Janson and W. Szpankowski. Analysis of an asymmetric leader election algorithm. *Electron. J. Comb.*, 4:#R17, 17 pages, 1997.
[9] P. Mathys and P. Flajolet. Q–ary collision resolution algorithms in random access systems with free or blocked channel access. *IEEE Transactions on Information Theory*, IT-31(2):217–243, March 1985.
[10] N. Nielsen. *Handbuch der Theorie der Gammafunktion.* Teubner, Leipzig, 1906.
[11] H. Prodinger. How to select a loser. *Discrete Math.*, 120:149–159, 1993.
[12] R. Sedgewick and P. Flajolet. *Introduction to the Analysis of Algorithms.* Addison-Wesley, Reading, Massachusetts, 1996.
[13] S.-H. Shiau and C.-B. Yang. A fast maximum finding algorithm on broadcast communications. *Inf. Process. Lett.*, 60:81–89, 1996.
[14] S.-H. Shiau and C.-B. Yang. A fast sorting algorithm and its generalization on broadcast communications. preprint, 18 pages, 2000.
[15] S.-H. Shiau and C.-B. Yang. Sorting algorithms on broadcast communications: Algorithm and fundamental analysis. in preparation, 2000.
[16] E. C. Titchmarsh. *The Theory of Functions.* Oxford University Press, 1939.

(P. Grabner) Institut für Mathematik A, Technische Universität Graz, Steyrergasse 30, 8010 Graz, Austria
E-mail address: grabner@weyl.math.tu-graz.ac.at

(H. Prodinger) The John Knopfmacher Centre for Applicable Analysis and Number Theory, Department of Mathematics, University of the Witwatersrand, P. O. Wits, 2050 Johannesburg, South Africa
E-mail address: helmut@gauss.cam.wits.ac.za