A Note on Alternating Sums*

Peter Kirschenhofer
Department of Algebra and Discrete Mathematics
Technical University of Vienna †

Submitted: February 19, 1995; Accepted: June 9, 1995

Dedicated to Professor Dominique Foata on the occasion of his 60th birthday

Abstract

We present some results on a certain type of alternating sums which frequently arise in connection with the average-case analysis of algorithms and data structures. Whereas the so-called Rice's method for treating such sums uses complex contour integration we perform manipulations of generating functions in order to get explicit results from which asymptotic estimates follow immediately.

1 Introduction

In the present paper we deal with alternating sums of the type

$$\sum_{k=a}^{N} \binom{N}{k} (-1)^k f(k), \tag{1}$$

where a is a natural number fulfilling $0 \le a \le N$. If $a = 0$, this expression is an N-th order difference of the sequence $(f(n))$. Sums of the referred type occur frequently in the average case analysis of divide-and-conquer algorithms resp. data structures like tries or digital search trees. We refer to [2], [7], [9], [10] and [13] for a number of examples.

A standard method for the treatment of sums of that type, attributed to S.O.Rice by D.E.Knuth [7], but in fact already contained e.g. in Nörlund's classical book on *Differenzenrechnung* [12], is to rewrite the finite sum as a complex contour integral

$$\sum_{k=a}^{N} \binom{N}{k} (-1)^k f(k) = -\frac{1}{2\pi i} \int_{\mathcal{C}} B(N+1, -z) f(z)\, dz, \tag{2}$$

where $B(x, y) = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ is the Beta function, $\mathcal{C}$ is a positively oriented closed curve surrounding the points $a, a+1, \ldots, N$ and $f(z)$ is an analytic continuation of the discrete sequence $f(k)$ to the complex plane, with no poles within the region surrounded by $\mathcal{C}$.

*This research was supported by the Austrian National Bank Project Nr.4995.
†Wiedner Hauptstraße 8–10/118.4, A-1040 Vienna, Austria. e-mail: kirsch@rsmb.tuwien.ac.at

If the integrand decreases sufficiently fast towards $\pm i\infty$, the asymptotic evaluation of this expression can be achieved by extending the contour of the integral to the left and collecting the residues at the newly encountered poles. The residue computations can often be performed automatically using a Computer algebra system like MAPLE. In a forthcoming paper [3] Flajolet and Sedgewick give an excellent overview of this technique. In particular it is proved that for meromorphic functions $f(z)$ which are of polynomial growth on concentric circles around the origin not passing through a pole, the summation of residues leads to explicit summation formulæ.

In this note we will confine ourselves to functions $f(z)$ of that type and concentrate on generating function manipulations that allow to gain explicit results. We start with the observation that by Mittag-Leffler's theorem a meromorphic function $f(z)$ allows a partial fraction decomposition of the type

$$f(z) = \sum_i (f_i(z) - p_i(z)) \; + \; g(z), \tag{3}$$

where the $f_i(z)$ are the principal parts of $f(z)$, the $p_i(z)$ are polynomials and $g(z)$ is an entire function. In [6] an algorithmic approach for establishing this decomposition is discussed.

Considering functions of polynomial growth in the above defined sense we will assume that $g(z)$ is a polynomial.

The summation of the polynomial parts of f is trivial, since, with Δ denoting the difference operator $\Delta f(x) = f(x+1) - f(x)$,

$$\sum_{0\le k\le N} \binom{N}{k} (-1)^k k^m = \Big((-1)^N \Delta^N x^m\Big)_{x=0} = (-1)^N N! S(m,N), \tag{4}$$

where $S(i,j)$ denote the Stirling numbers of the second kind. In particular the sum equals 0 for $N > m$.

The main problem is therefore to be able to sum up the polar contributions

$$\sum_{0\le k\le N, k\ne\xi} \binom{N}{k} (-1)^k \frac{1}{(k-\xi)^m}, \tag{5}$$

for positive integers m.

In the next section we give an explicit expression for sums of that type, discuss special instances and analyze the contribution of regularly spaced poles in the complex plane.

In the final section we apply these results to some concrete examples originating from the average case analysis of an algorithm called "Approximate Counting" as well as from the analysis of the path length in digital search trees.

2 Some Explicit Summation Formulæ

Let us start with the instance $\xi = K \in \{0,\ldots,N\}$ in expression (5).

Proposition 2.1. *Let K be an integer satisfying $0 \le K \le N$ and m be a nonnegative integer. Then*

$$\sum_{0\le k\le N, k\ne K} \binom{N}{k} (-1)^k \frac{1}{(k-K)^m} = \binom{N}{K} (-1)^{K+1} \frac{1}{m!} Y_m(\ldots,(i-1)!(H_{N-K}^{(i)} + (-1)^i H_K^{(i)}),\ldots),$$

where $Y_m(\ldots,x_i,\ldots)$ are the Bell polynomials and $H_r^{(i)}=\sum_{j=1}^r \frac{1}{j^i}$ the Harmonic numbers of the i-th order.

Proof. We start from the well-known formula (compare e.g. [5] eqn.(5.41))

$$\sum_{k=0}^N \binom{N}{k}(-1)^k \frac{1}{x+k} = \frac{N!}{x(x+1)\ldots(x+N)}, \tag{6}$$

which simply follows from the fact that the left-hand side equals $(-1)^N \Delta^N \frac{1}{x}$. Therefore

$$\sum_{0\le k\le N, k\ne K} \binom{N}{k}(-1)^k \frac{1}{x+k} = \frac{1}{x+K}\left(\frac{N!}{\prod_{0\le j\le N, j\ne N}(x+j)} - \binom{N}{K}(-1)^K\right). \tag{7}$$

Applying the operator $\mathrm{E}^{-t}=e^{-tD}$ (where D denotes the differential operator with respect to x) to the last equation we find

$$\sum_{m\ge 0} t^m \sum_{0\le k\le N, k\ne K} \binom{N}{k}(-1)^k \frac{1}{(x+k)^{m+1}} =$$
$$\frac{1}{x+K-t}\left(\frac{N!}{\prod_{0\le j\le N, j\ne N}(x+j)} \prod_{0\le j\le N, j\ne K} \frac{1}{1-\frac{t}{x+j}} - \binom{N}{K}(-1)^K\right).$$

Inserting $x=-K$ and denoting by $[t^n]h(t)$ the coefficient of t^n in the power series $h(t)$ we get

$$\sum_{0\le k\le N, k\ne K} \binom{N}{k}(-1)^k \frac{1}{(x+k)^m} = [t^{m-1}]\frac{1}{t}\binom{N}{K}(-1)^{K+1}\left(\prod_{0\le j\le N, j\ne K}\frac{1}{1-\frac{t}{j-K}} - 1\right). \tag{8}$$

The last factor on the right-hand side can be rewritten as follows

$$\begin{aligned}
\prod_{0\le j\le N, j\ne K}\frac{1}{1-\frac{t}{j-K}} - 1 &= \exp\left(-\sum_{0\le j\le N, j\ne K}\log(1-\frac{t}{j-K})\right) - 1 \\
&= \exp\left(\sum_{i\ge 1}\frac{t^i}{i}\sum_{0\le j\le N, j\ne K}\frac{1}{(j-K)^i}\right) - 1 \\
&= \sum_{m\ge 1}\frac{t^m}{m!} Y_m(\ldots,(i-1)!(H_{N-K}^{(i)}+(-1)^i H_K^{(i)}),\ldots).
\end{aligned}$$

Extracting the coefficient of t^m we find the announced formula. ∎

It is worthwhile to mention the following most important special instances of Proposition 2.1:

Corollary 2.1. *Let K be an integer fulfilling $0\le K\le N$. Then we have for*
(1) m=1:

$$\sum_{0\le k\le N, k\ne K}\binom{N}{k}(-1)^k\frac{1}{k-K} = \binom{N}{K}(-1)^{K+1}(H_{N-K}-H_K),$$

(2) m=2:

$$\sum_{0\le k\le N, k\ne K} \binom{N}{k}(-1)^k \frac{1}{(k-K)^2} = \binom{N}{K}(-1)^{K+1}\frac{1}{2}\left((H_{N-K}-H_K)^2 + H^{(2)}_{N-K} + H^{(2)}_K\right),$$

For m a nonegative integer we have for
(3) K=0:

$$\sum_{k=1}^{N} \binom{N}{k}(-1)^k \frac{1}{k^m} = -\frac{1}{m!}Y_m(\ldots,(i-1)!H^{(i)}_N),\ldots),$$

(4) K=1:

$$\sum_{k=2}^{N} \binom{N}{k}(-1)^k \frac{1}{(k-1)^m} = (-1)^{m-1} + N\frac{1}{m!}Y_m(\ldots,(i-1)!(H^{(i)}_{N-1}+(-1)^i),\ldots),$$

(5) K=2:

$$\sum_{k=3}^{N} \binom{N}{k}(-1)^k \frac{1}{(k-2)^m} = (-1)^m(N-\frac{1}{2^m}) - \binom{N}{2}\frac{1}{m!}Y_m(\ldots,(i-1)!(H^{(i)}_{N-2}+(-1)^i(1+\frac{1}{2^i})),\ldots)$$

The instance $\xi \notin \{0,1,\ldots\}$ in expression (5) can be treated by an even more direct calculation. Since it will appear (in different but equivalent formulation) in the forthcoming paper [3] (where the special instance (3) of Corollary 2.1 may be found, too), we only cite the final formula.

Proposition 2.2. *Let $\xi \notin \{0,1,\ldots\}$ be a complex number and m be a positive integer. Then*

$$\sum_{0\le k\le N} \binom{N}{k}(-1)^k \frac{1}{(k-\xi)^m} = \Gamma(-\xi)\frac{\Gamma(N+1)}{\Gamma(N+1-\xi)}\frac{1}{(m-1)!}Y_{m-1}(\ldots,(i-1)!\zeta_N(i,-\xi),\ldots),$$

where $\zeta_N(i,-\xi) = \sum_{j=0}^{N}\frac{1}{(j-\xi)^i}$ are partial sums of the Hurwitz zeta function with parameter $-\xi$.

The special instance $\xi = -K$, where K is a positive integer, reads as follows:

Corollary 2.2. *Let K and m be positive integers. Then*

$$\sum_{0\le k\le N} \binom{N}{k}(-1)^k \frac{1}{(k+K)^m} = \frac{1}{K\binom{N+K}{K}}\frac{1}{(m-1)!}Y_{m-1}(\ldots,(i-1)!(H^{(i)}_{N+K}-H^{(i)}_{K-1}),\ldots).$$

Propositions 2.1 and 2.2 express the alternating sums in a manner that is also very well suited for the purpose of deriving asymptotic expansions for N getting large. The asymptotics of the i-th order Harmonic numbers is well-known, and the same holds for the partial sums of the Hurwitz zeta function, compare [3]. Although we do not want to concentrate on asymptotic aspects in this note, it seems worthwhile to mention that the quotient of values of the Gamma function appearing in Proposition 2.2 can be rewritten using the classical formula

$$\frac{1}{\Gamma(z)} = z\mathrm{e}^{\gamma z}\prod_{k\ge 1}\left[(1+\frac{z}{k})\mathrm{e}^{-z/k}\right] \tag{9}$$

in the following way:

$$\frac{\Gamma(N+1)}{\Gamma(N+1-\xi)} = \mathrm{e}^{\xi(H_N-\gamma)} \prod_{k\geq N+1} \left[(1-\frac{\xi}{k})\mathrm{e}^{\xi/k}\right], \tag{10}$$

where γ is Euler's constant.

It is not difficult to see that this expression is asymptotic to $N^{\xi}(1+O(\frac{1}{N}))$ for N getting large.

An interesting situation arises if the meromorphic function $f(z)$ that continues the discrete sequence $f(k)$, has *regularly spaced poles* on a line parallel to the imaginary axis. Let us consider for instance the situation of first order poles $\xi_n = \sigma + in\tau$, with σ, τ fixed real numbers and n running through the integers. We denote the residues of $f(z)$ in ξ_n by c_n. The sum

$$\sum_{0\leq k\leq N, k\neq \xi_0} \binom{N}{k} (-1)^k \frac{c_0}{k-\xi_0}$$

is evaluated according to Corollary 2.1(1) resp. equation (6) depending on ξ_0 being an element of $\{0,\ldots,N\}$ or not. Therefore we focus our attention on the contribution of the remaining poles: According to our considerations from above we have the following explicit formula:

Corollary 2.3. *For $\xi_n = \sigma + \mathrm{i}n\tau$, a complex number with σ, τ being reals, $\tau \neq 0$, we have*

$$\sum_{k=0}^{N} \binom{N}{k} (-1)^k \sum_{n\in Z\setminus\{0\}} \frac{c_n}{k-\xi_n} = e^{\sigma(H_N-\gamma)} \sum_{n\in Z\setminus\{0\}} c_n \Gamma(-\xi_n) e^{\mathrm{i}n\tau(H_N-\gamma)} \prod_{k\geq N+1} \left[(1-\frac{\xi_n}{k})e^{\frac{\xi_n}{k}}\right].$$

For N getting large this expression is asymptotic to

$$N^{\sigma} \sum_{n\in Z\setminus\{0\}} c_n \Gamma(-\xi_n) \mathrm{e}^{\mathrm{i}n\tau \log N} = N^{\sigma} \delta(\log_{\varphi(\tau)} N), \qquad \varphi(\tau) = \mathrm{e}^{2\pi/\tau}, \tag{11}$$

where $\delta(x)$ is a periodic function with period 1, integral mean zero, and Fourier expansion

$$\delta(x) = \sum_{n\in Z\setminus\{0\}} c_n \Gamma(-\xi_n) \mathrm{e}^{2\pi \mathrm{i}nx}. \tag{12}$$

3 Examples

Approximate Counting is a probabilistic algorithm proposed by R. Morris [11] that allows an approximate count in the interval 1 to N using only about $\log_2 \log_2 N$ bits. The algorithm uses a counter C, which is set to 1 at the beginning. Then, in each incremental step, the actual value C is incremented by 1 with probability 2^{-C}, resp. remains as it is with probability $1-2^{-C}$. The idea is that after N increments C should contain a good approximation to $\log_2 N$.

In [1] Ph. Flajolet has presented a detailed analysis of this algorithm using real analysis and the Mellin integral transform. In [8] H. Prodinger and the author presented an alternative approach using Rice's integral method in order to gain asymptotic results on the expectation E_N and the variance of the content of the counter after N incremental steps. In the following we present explicit results for these expectations in the style of Section 2.

We start from the following alternating sum expression of the expectation, which is proved in [8], Prop. 3. (A different method to establish this formula using q-hypergeometric series may be found in [14].)

$$\mathrm{E}_N = 1 - \sum_{k=1}^{N} \binom{N}{k} (-1)^k 2^{-k} Q_{k-1}, \qquad \text{where} \qquad Q_j = \prod_{n=1}^{j} (1 - \frac{1}{2^n}). \tag{13}$$

The continuation of the sequence (Q_k) to the complex plane is standard: With

$$Q(t) = \prod_{n\geq 1} (1 - \frac{t}{2^n}), \qquad Q_\infty = Q(1), \tag{14}$$

we have

$$Q_k = \frac{Q_\infty}{Q(2^{-k})}, \tag{15}$$

and now k can be replaced by z being a complex number different from a negative integer.

According to Euler's identity

$$\frac{1}{Q(t)} = \sum_{n\geq 0} \frac{t^n}{2^n Q_n} \tag{16}$$

from the theory of q-hypergeometric series we have

$$\frac{Q_\infty}{Q(2^{-z})} = Q_\infty \sum_{j\geq 0} \frac{2^{-j(z+1)}}{Q_j} = \sum_{j\geq 0} 2^{-j(z+1)} Q(2^{-j}).$$

Now we apply Euler's second identity

$$Q(t) = \prod_{n\geq 1} (1 - \frac{t}{2^n}) = \sum_{j\geq 0} a_{j+1} t^j, \quad \text{where} \quad a_{j+1} = (-1)^j 2^{-\binom{j+1}{2}} / Q_j, \tag{17}$$

and derive

$$\frac{Q_\infty}{Q(2^{-z})} = \sum_{j\geq 0} a_{j+1} \sum_{n\geq 0} 2^{-(z+1+j)n} = \sum_{j\geq 1} \frac{a_j}{1 - 2^{-(z+j)}}. \tag{18}$$

Therefore

$$2^{-k} Q_{k-1} = \sum_{j\geq 0} \frac{a_{j+1} 2^j}{2^{k+j} - 1} = \frac{1}{2^k - 1} + \sum_{j\geq 1} \frac{a_{j+1} 2^j}{2^{k+j} - 1} \tag{19}$$

and

$$\begin{aligned} \sum_{k=1}^{N} \binom{N}{k} (-1)^k 2^{-k} Q_{k-1} \quad &= \quad \sum_{k=1}^{N} \binom{N}{k} (-1)^k \frac{1}{2^k - 1} + \\ &\quad + \sum_{j\geq 1} a_{j+1} 2^j \left(\sum_{k=0}^{N} \binom{N}{k} (-1)^k \frac{1}{2^{k+j} - 1} - \frac{1}{2^j - 1} \right) \end{aligned} \tag{20}$$

The partial fraction decomposition of the occuring meromorphic functions is surprisingly simple and follows from the well-known result (compare e.g. [6] eqn.(7.10-10))

$$\frac{2\pi}{\mathrm{e}^{2\pi z} - 1} = -\pi + \frac{1}{z} + \sum_{n=1}^{\infty} \left(\frac{1}{z - in} + \frac{1}{z + in} \right). \tag{21}$$

(The usual argument after the determination of the principal parts is that the remaining part is a bounded, entire function and therefore constant.) We find

$$\frac{1}{2^z-1}=-\frac{1}{2}+\frac{1}{Lz}+\frac{1}{L}\sum_{n\in Z\setminus\{0\}}\frac{1}{z-\chi_n},\quad \text{where}\quad L=\log 2,\ \chi_n=\frac{2n\pi i}{L}. \tag{22}$$

Let us treat the alternating sum

$$\Sigma_1=\sum_{k=1}^{N}\binom{N}{k}(-1)^k\frac{1}{2^k-1} \tag{23}$$

first. Inserting the partial fraction decomposition (20) we have

$$\Sigma_1=-\frac{1}{2}\sum_{k=1}^{N}\binom{N}{k}(-1)^k+\frac{1}{L}\sum_{k=1}^{N}\binom{N}{k}(-1)^k\frac{1}{k}+\frac{1}{L}\sum_{k=1}^{N}\binom{N}{k}(-1)^k\frac{1}{k-\chi_n}$$

Now we apply Corollary 2.1(3) and Corollary 2.3 and get

$$\Sigma_1=\frac{1}{2}(1-\delta_{N,0})-\frac{H_N}{L}+\frac{1}{L}\sum_{n\in Z\setminus\{0\}}\Gamma(-\chi_n)\mathrm{e}^{\chi_n(H_N-\gamma)}\prod_{k\geq N+1}\left[(1-\frac{\chi_n}{k})\mathrm{e}^{\frac{\chi_n}{k}}\right]. \tag{24}$$

In the next step we treat the sums

$$\Sigma_2(j)=\sum_{k=0}^{N}\binom{N}{k}(-1)^k\frac{1}{2^{k+j}-1}. \tag{25}$$

Applying Corollaries 2.2 and 2.3 we find

$$\Sigma_2(j)=\frac{1}{Lj\binom{N+j}{j}}+\frac{1}{L}\mathrm{e}^{-j(H_N-\gamma)}\sum_{n\in Z\setminus\{0\}}\Gamma(j-\chi_n)\mathrm{e}^{\chi_n(H_N-\gamma)}\prod_{k\geq N+1}\left[(1+\frac{j-\chi_n}{k})\mathrm{e}^{\frac{\chi_n-j}{k}}\right]. \tag{26}$$

Inserting the expressions for Σ_1 and $\Sigma_2(j)$ in (20) we gain the following final formula for the expectations E_N:

Proposition 3.1. *The average content of the counter in the algorithm "Approximate Counting" after $N\geq 1$ successive increments is given by the explicit formula*

$$\mathrm{E}_N=\frac{H_N}{L}+\frac{1}{2}-\alpha-\frac{1}{L}\varphi_0(N)-\frac{1}{L}\sum_{j\geq 1}a_{j+1}\frac{2^j}{j\binom{N+j}{j}}-\frac{1}{L}\sum_{j\geq 1}a_{j+1}2^j\varphi_j(N),$$

where $L=\log 2,\ \alpha=-\sum_{j\geq 1}a_{j+1}\frac{2^j}{2^j-1},\ a_{j+1}=(-1)^j2^{-\binom{j+1}{2}}/Q_j$, and the fluctuating functions $\varphi_j(N)$ are defined by

$$\varphi_j(N)=\mathrm{e}^{-j(H_N-\gamma)}\sum_{n\in Z\setminus\{0\}}\Gamma(j-\chi_n)\mathrm{e}^{\chi_n(H_N-\gamma)}\prod_{k\geq N+1}\left[(1+\frac{j-\chi_n}{k})\mathrm{e}^{\frac{\chi_n-j}{k}}\right].$$

It should be remarked that the explicit formula from above gives the terms in decreasing asymptotic order in N for increasing j. It should further be remarked that the constant α, which often occurs in this type of analysis, allows a more familiar, though less rapidly converging, different representation:

Using the decomposition (18) we have

$$\begin{aligned}
\alpha &= -\sum_{j\geq 1} a_{j+1}\frac{2^j}{2^j-1} = -\sum_{j\geq 1} a_{j+1}\frac{1}{1-2^{-j}} \\
&= -\lim_{t\to 1-}\left(\frac{Q_\infty}{Q(2t)} - \frac{1}{1-t}\right) = -\lim_{t\to 1-}\left(\frac{Q_\infty}{(1-t)Q(t)} - \frac{1}{1-t}\right) \\
&= \sum_{n\geq 1}\frac{1}{2^n-1}. \qquad (27)
\end{aligned}$$

A similar treatment is possible for the variances.

Digital Search Trees are a prominent data structure for storing and retrieving information with keys from random bit streams. N items, represented by 0,1-sequences with bits chosen independently with equal probabilities, are stored in the internal nodes of a binary tree according to the following principle. The first item is stored in the root, and then, recursively, the next item is stored in the first nonoccupied internal node of the tree, where the path to this node is determined by the prefix of the key, a 0 reading for going left and a 1 reading for going right. One of the most important parameters for the performance of algorithms on this data structure is the internal path length, i.e. the sum of distances of all occupied nodes from the root. The asymptotics of the expectation E_N of this parameter (adopting the symmetric Bernoulli model for the input keys) has been analyzed using different techniques by a number of authors. Compare [10] for references, as well as for information on the variance of the referred parameter.

In the context of this note we want to achieve an explicit formula for the expectations E_N of the path length, from which the asymptotics is immediate, too. Again we may start from an alternating sum expression in the style of eqn.(1) (compare [10]):

$$\mathrm{E}_N = \sum_{k=2}^{N}\binom{N}{k}(-1)^k Q_{k-2}, \qquad (28)$$

where the Q_k are defined as in (13). Proceeding as in the analysis of Approximate Counting we find

$$Q_{k-2} = \sum_{j\geq 0} a_{j+1}\frac{2^{k+j-1}}{2^{k+j-1}-1} = Q_\infty + \sum_{j\geq 0} a_{j+1}\frac{1}{2^{k+j-1}-1}. \qquad (29)$$

Therefore we have for $N \geq 1$

$$\begin{aligned}
\mathrm{E}_N &= Q_\infty(N-1) + \sum_{k=2}^{N}\binom{N}{k}(-1)^k\frac{1}{2^{k-1}-1} - \sum_{k=1}^{N}\binom{N}{k}(-1)^k\frac{1}{2^k-1} - N + \\
&\quad + \sum_{j\geq 2} a_{j+1}\left(\sum_{k=0}^{N}\binom{N}{k}(-1)^k\frac{1}{2^{k+j-1}-1} - \frac{1}{2^{j-1}-1} + \frac{N}{2^j-1}\right). \qquad (30)
\end{aligned}$$

Observe that we know already how to rewrite the sums

$$\Sigma_1 = \sum_{k=1}^{N} \binom{N}{k} (-1)^k \frac{1}{2^k - 1}, \qquad \Sigma_2(j) = \sum_{k=0}^{N} \binom{N}{k} (-1)^k \frac{1}{2^{k+j} - 1}$$

from eqs. (24) and (26). Therefore it remains to treat the sum

$$\Sigma_0 = \sum_{k=2}^{N} \binom{N}{k} (-1)^k \frac{1}{2^{k-1} - 1}. \tag{31}$$

This can be performed in the same style as we treated Σ_1, but it should be noted that in fact Σ_0 is the classical instance of such alternating sums studied by Rice himself. The solution may be found in Knuth's book [7], Ex. 5.2.2-54 and reads in our representation of the fluctuating parts as

$$\Sigma_0 = \frac{N(H_{N-1} - 1)}{L} - \frac{N}{2} + 2 + \frac{\mathrm{e}^{H_N - \gamma}}{L} \sum_{n \in Z \setminus \{0\}} \Gamma(-1 - \chi_n) \mathrm{e}^{\chi_n (H_N - \gamma)} \prod_{k \geq N+1} \left[(1 - \frac{1 + \chi_n}{k}) \mathrm{e}^{\frac{1+\chi_n}{k}} \right] \tag{32}$$

The constant

$$C_1 = \sum_{j \geq 2} a_{j+1} \frac{1}{2^j - 1} \tag{33}$$

is easily expressed in terms of α from Proposition 3.1:

$$C_1 = 1 - \sum_{j \geq 1} a_{j+1} + \sum_{j \geq 1} a_{j+1} \frac{1}{1 - 2^{-j}} = 2 - Q_\infty - \alpha \tag{34}$$

according to Euler's second identity (17). The second constant that must be computed is

$$C_2 = \sum_{j \geq 2} a_{j+1} \frac{1}{2^{j-1} - 1}. \tag{35}$$

Proceeding similarly as with α in (27) we get

$$\begin{aligned} C_2 &= -\sum_{j \geq 2} a_{j+1} + \sum_{j \geq 2} a_{j+1} \frac{1}{1 - 2^{1-j}} \\ &= -Q_\infty + \lim_{t \to 1-} \left(\frac{Q_\infty}{Q(4t)} - \frac{1}{1 - 2t} + \frac{1}{1 - t} \right) \\ &= -Q_\infty + \lim_{t \to 1-} \left(\frac{Q_\infty}{(1 - 2t)(1 - t) Q(t)} - \frac{1}{1 - 2t} + \frac{1}{1 - t} \right) \\ &= -Q_\infty + \alpha + 1 \end{aligned} \tag{36}$$

Collecting all contributions we find

Proposition 3.2. *The expected internal path length* E_N *under the symmetric Bernoulli model of a Digital Search Tree built from* N *records is given by the explicit formula*

$$\begin{aligned} \mathrm{E}_N &= \frac{N H_N}{L} + N(-\frac{1}{L} + \frac{1}{2} - \alpha) + \frac{1}{L} \varphi_{-1}(N) \\ &\quad + \frac{H_N}{L} + \frac{5}{2} - \alpha - \frac{1}{L} - \frac{1}{L} \varphi_0(N) \\ &\quad + \frac{1}{L} \sum_{j \geq 2} a_{j+1} \frac{1}{(j-1)\binom{N+j-1}{j-1}} + \frac{1}{L} \sum_{j \geq 2} a_{j+1} \varphi_{j-1}(N), \end{aligned}$$

with L, α and $\varphi_j(N)$ defined as in Proposition 3.1.

References

[1] Ph. Flajolet. Approximate counting: a detailed analysis. *BIT* 25(1985), 113-134.

[2] Ph. Flajolet and R. Sedgewick. Digital search trees revisited. *SIAM J. Comput.* 15(1986), 748-767.

[3] Ph. Flajolet and R. Sedgewick. Mellin transforms and asymptotics: Finite differences and Rice's integrals. *Theoretical Computer Science*, 1995. To appear.

[4] I. Goulden and D. Jackson. *Combinatorial Enumeration.* J. Wiley, 1983.

[5] R.L. Graham, D.E. Knuth and O. Patashnik. *Concrete Mathematics*, 2nd Edition. Addison Wesley, 1994.

[6] P. Henrici. *Applied and Computational Complex Analysis*, Vol.1. J. Wiley, 1988.

[7] D. E. Knuth. *The Art of Computer Programming*, Vol.3. Addison Wesley, 1973.

[8] P. Kirschenhofer and H. Prodinger. Approximate counting: an alternative approach. *Informatique théorique et Appl.* 25(1991), 43-48.

[9] P. Kirschenhofer and H. Prodinger. The path length of random skip lists. *Acta Informatica* 31(1994), 775-792.

[10] P. Kirschenhofer, H. Prodinger and W. Szpankowski. Digital search trees again revisited: the internal path length perspective. *SIAM J. Comput.* 23(1994), 598-616.

[11] R. Morris. Counting large numbers of events in small registers. *Comm. A.C.M.* 21(1978), 840-842.

[12] N. E. Nörlund. Vorlesungen über Differenzenrechnung. Chelsea, 1954.

[13] H. Prodinger. Combinatorics of geometrically distributed random variables: Left-to-right maxima. *Discrete Mathematics*, 1995. To appear.

[14] H. Prodinger. Approximate Counting via Euler Transform. *Mathematica Slovaca.* To appear.