ON SOME APPLICATIONS OF FORMULAE OF RAMANUJAN IN THE ANALYSIS OF ALGORITHMS

P. KIRSCHENHOFER AND H. PRODINGER

Abstract. Using several transformation formulae from Ramanujan's second Notebook we achieve distribution results on random variables related to dynamic data structures (so-called "tries"). This continues research of Knuth, Flajolet and others *via* an approach that is completely new in this subject.

§1. *Introduction.* Our aim in this paper is to demonstrate the extensive applicability of several series relations occurring in Ramanujan's Notebook [**13**] in the analysis of special data structures and algorithms.

The identities we refer to are the following. Let α and β be positive numbers with $\alpha\beta = \pi^2$.

IDENTITY 1. [Ramanujan's Formula for $\zeta(2N+1)$]. *Let N be a positive integer. Then*

$$\alpha^{-N}\left\{\tfrac{1}{2}\zeta(2N+1)+\sum_{k\geq 1}\frac{k^{-2N-1}}{e^{2\alpha k}-1}\right\}=(-\beta)^{-N}\left\{\tfrac{1}{2}\zeta(2N+1)+\sum_{k\geq 1}\frac{k^{-2N-1}}{e^{2\beta k}-1}\right\}$$
$$-2^{2N}\sum_{k=0}^{N+1}(-1)^k\frac{B_{2k}}{(2k)!}\times\frac{B_{2N+2-2k}}{(2N+2-2k)!}\alpha^{N+1-k}\beta^k,$$

where B_n indicates the n-th Bernoulli number defined by

$$\frac{z}{e^z-1}=\sum_{n\geq 0}B_n\frac{z^n}{n!}.$$

IDENTITY 2.

$$\sum_{k\geq 1}\frac{1}{k(e^{2\alpha k}-1)}-\tfrac{1}{4}\log\alpha+\frac{\alpha}{12}=\sum_{k\geq 1}\frac{1}{k(e^{2\beta k}-1)}-\tfrac{1}{4}\log\beta+\frac{\beta}{12}.$$

IDENTITY 3.

$$\alpha\sum_{k\geq 1}\frac{k}{e^{2\alpha k}-1}+\beta\sum_{k\geq 1}\frac{k}{e^{2\beta k}-1}=\frac{\alpha+\beta}{24}-\frac{1}{4}.$$

IDENTITY 4. *Let $N \geq 2$ be an integer. Then*

$$\alpha^N \sum_{k \geq 1} \frac{k^{2N-1}}{e^{2\alpha k}-1} - (-\beta)^N \sum_{k \geq 1} \frac{k^{2N-1}}{e^{2\beta k}-1} = (\alpha^N - (-\beta)^N) \frac{B_{2N}}{4N}.$$

It should be noticed that these 4 identities cover transformation formulae for the series

$$\sum_{k \geq 1} \frac{k^m}{e^{2\alpha k}-1}$$

for all odd integers m.

These (and several additional formulae) can be found in the excellent survey by Berndt [**2**] which also offers historical remarks on the proofs.

In the following sections we are concerned with random variables X_N describing certain parameters of a data structure called *tries* (defined in Section 2) which are of importance in Theoretical Computer Science.

For the reader's convenience it seems to be appropriate to *outline the mathematical methods* before going into technical details. In the first step recurrence relations for the probability generating functions of the random variables in question are established. These recursions lead to explicit expressions for the *expectations*

$$EX_N = \sum_{k \geq 2} \binom{N}{k} (-1)^k f(k),$$

where f is a complicated function which can be continued analytically. In order to get asymptotic information it is convenient to apply the following lemma from the calculus of finite differences (compare [**12**]).

LEMMA 1. *Let C be a curve surrounding the points $2, \ldots, N$ and let $f(z)$ be analytic within C. Then*

$$\sum_{k \geq 2} \binom{N}{k} (-1)^k f(k) = \frac{-1}{2\pi i} \int_C [N; z] f(z) dz$$

with

$$[N; z] = \frac{(-1)^{N-1} N!}{z(z-1)\ldots(z-N)}.$$

In our applications f is a special meromorphic function and the asymptotic expansion of EX_N is obtained *via*

$$EX_N \sim \sum \text{Res}\,([N; z] f(z)),$$

where the sum is taken over all poles different from $2, \ldots, N$. The results are of the type

$$EX_N \sim AN + N\delta(\log_2 N),$$

where $\delta(x)$ is a continuous periodic function with period 1, very small amplitude, mean 0 and known Fourier expansion.

In order to investigate the *variance* (which turns out to be much more delicate) one starts with the formula

$$\operatorname{Var} X_N = EX_N(X_N - 1) + EX_N - (EX_N)^2.$$

The first term can be treated in a similar fashion as the expectation (although with more tedious computations) and yields for the leading terms

$$EX_N(X_N - 1) \sim BN^2 + N^2 \delta_1(\log_2 N).$$

Thus

$$\operatorname{Var} X_N \sim (B - A^2)N^2 + N^2(\delta_1(\log_2 N) - 2A\delta(\log_2 N) - \delta^2(\log_2 N)).$$

Using Ramanujan's formulae it can be shown that in the two following problems (Sections 3 and 4), surprisingly enough, the non-fluctuating part of the N^2-term, namely

$$B - A^2 - \text{zeroeth Fourier coefficient of } \delta^2(x),$$

vanishes. Consequently (since the involved periodic functions are continuous and the variance is non-negative) even the fluctuating part of the N^2-term vanishes; as a byproduct this yields identities for the Fourier coefficients of $\delta^2(x)$, *e.g.*, for $k \neq 0$,

$$\sum_{\substack{l+m=k\\ l,m\neq 0}} \Gamma(1-\chi_l)\Gamma(1-\chi_m) = -2\Gamma(1-\chi_k)$$
$$+ (\log 2)\Gamma(2-\chi_k)\left\{2\sum_{j\geq 0}\binom{\chi_k - 2}{j}\frac{1}{2^{j+1}-1} + \frac{1}{4}\right\},$$

where $\chi_n = 2n\pi i/\log 2$.

As a final result the variances are established to be of order N.

It should be emphasized that periodic fluctuations occur frequently in distribution results in Number Theory, Combinatorics and Computer Science. Henceforth we believe that the techniques of this paper might be useful in the computation of higher centralized moments in other problems (compare Section 5 for more specific remarks).

Notational remarks. (1) $[z^n]f(z)$ denotes the n-th coefficient in the Laurent series.

(2) We use the following abbreviations

$$L = \log 2, \qquad \chi_k = \frac{2k\pi i}{L}.$$

§2. *Tries.* In this section we give a short description of the data structure in consideration as well as the parameters in which we are interested.

An important class of algorithms is concerned with storing and searching for data in well designed data structures.

One of the most prominent examples are the so-called (*binary*) *tries* (from information re*trie*val). N data are stored in external nodes of a binary tree.

It is assumed that each item has a "key" being an infinite sequence of 0 and 1 (where all such sequences are regarded equally likely). In the tree each left (resp. right) branch is labelled with 0 (resp. 1). This yields an encoding of each external node by means of the (finite) sequence of labels describing the path from the *root* to this node. Each item is stored in *that* external node corresponding to the shortest *unique* prefix of its key.

Example. We consider 5 data $A, \ldots, E$ with keys starting as follows (the shortest prefixes are indicated):

A: *0 1* 0 1 1 . . . .

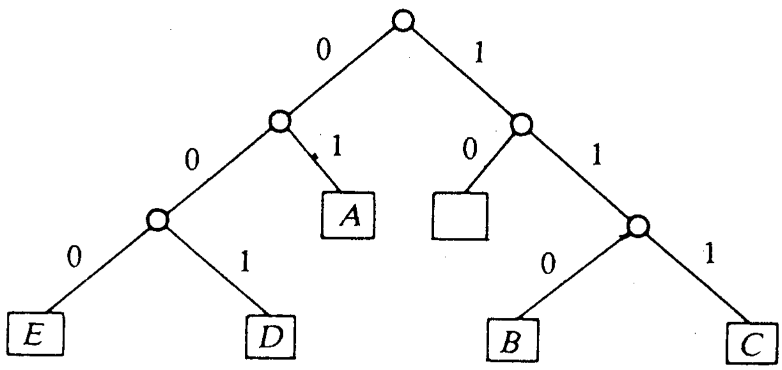
B: *1 1 0* 1 0 . . . .

C: *1 1 1* 0 0 . . . .

D: *0 0 1* 0 1 . . . .

E: *0 0 0* 0 0 . . . .

The corresponding trie is now:



Observe that some external nodes in a trie may be empty so that there is no general dependency between N (number of items) and the number of internal nodes. For this reason it is of great importance for the analysis of this data structure to gain results about the distribution of the random variable

$$X_N = \text{number of internal nodes of a trie built from } N \text{ data},$$

where all key sequences are regarded equally likely.

The expectation of X_N has been analysed by Knuth in his famous book [11]. However, the variance was not considered up to now; it is investigated in Section 3.

Another random variable which is of interest in Computer Science is the number of internal nodes with two non-empty external nodes as immediate successors ("external internal nodes"). Here we study a slightly more general parameter. Let $s \geqslant 2$ be an integer. Then we consider the random variable

$$Y_N^{(s)} = \text{number of minimal subtrees containing exactly } s \text{ non-empty external nodes.}$$

The parameter from above is the instance $s = 2$. In our example these variables take the values $Y_5^{(2)}=2$, $Y_5^{(3)}=1$, $Y_5^{(4)}=0$, $Y_5^{(5)}=1$. Distribution results on these variables are obtained in Section 4.

The results of Sections 3 and 4 may be seen as a continuation of the work of Flajolet and Sedgewick on the analysis of trie algorithms [7].

§3. *Internal nodes in tries.* Let $F_N(z)$ be the probability generating function for the random variable X_N described in Section 2. Then the following recursion holds:

$$F_N(z)=z\sum_{k=0}^{N}2^{-N}\binom{N}{k}F_k(z)F_{N-k}(z),\qquad N\geq 2,\qquad F_0(z)=F_1(z)=1. \tag{3.1}$$

This follows from the observation that the probability for a trie of N data to have k data in the left subtree is $2^{-N}\binom{N}{k}$ and that, for $N\geq 2$, $X_N=1$ (for the root) plus the number of internal nodes in the two subtrees.

According to the introduction we need precise asymptotics for the expectation $l_N=F'_N(1)$. From (3.1) we have

$$l_N=1+2^{1-N}\sum_{k=0}^{N}\binom{N}{k}l_k,\qquad N\geq 2,\qquad l_0=l_1=0, \tag{3.2}$$

so that the exponential generating function

$$L(z)=\sum_{N\geq 0}l_N\frac{z^N}{N!}$$

satisfies

$$L(z)=e^z-1-z+2e^{z/2}L(\tfrac{1}{2}z). \tag{3.3}$$

In order to simplify we consider

$$\tilde{L}(z)=e^{-z}L(z)=\sum_{N\geq 0}\tilde{l}_N\frac{z^N}{N!} \tag{3.4}$$

and get

$$\tilde{L}(z)=2\tilde{L}(\tfrac{1}{2}z)+1-e^{-z}-ze^{-z}, \tag{3.5}$$

so that

$$\tilde{l}_N=\frac{(-1)^N(N-1)}{1-2^{1-N}},\qquad N\geq 2,\qquad \tilde{l}_0=\tilde{l}_1=0, \tag{3.6}$$

and finally

$$l_N=\sum_{k\geq 2}\binom{N}{k}(-1)^k\frac{k-1}{1-2^{1-k}},\qquad N\geq 2. \tag{3.7}$$

Thus we may apply Lemma 1, with

$$f(z)=\frac{z-1}{1-2^{1-z}},$$

and get

$$l_N = \frac{-1}{2\pi i} \int_C [N; z] f(z)\,dz,$$

with

$$[N; z] = \frac{(-1)^{N-1} N!}{z(z-1)\ldots(z-N)},$$

where C surrounds the points $2, \ldots, N$. If we change C to a rectangle with the corners $-M \pm iY$, $X \pm iY$, $Y \neq 2k\pi/L$ for all $k \in \mathbb{Z}$, and let X, Y tend to infinity we get the following expression for l_N (for the estimate of the remainder term compare [7] and [**14**]).

$$l_N = \sum \operatorname{Res}([N; z] f(z)) + O(N^{-M}), \qquad \text{any} \qquad M > 0,$$

where the sum is taken over all poles different from $2, \ldots, N$, *i.e.*, $z = 1 + \chi_k$, $k \in \mathbb{Z}$ and $z = 0$. The residues are computed by standard techniques; we find

$$\operatorname{Res}([N; z] f(z); z = 1) = \frac{N}{L}$$

$$\operatorname{Res}([N; z] f(z); z = 1 + \chi_k) = \frac{N}{L} N^{\chi_k} \chi_k \Gamma(-1 - \chi_k) \times \left\{1 - \frac{(1+\chi_k)\chi_k}{2N} + O\left(\frac{1}{N^2}\right)\right\}, \qquad k \neq 0,$$

$$\operatorname{Res}([N; z] f(z); z = 0) = -1.$$

Thus we have rederived

THEOREM 2. *The expected number of internal nodes in a trie built from N data is for* $N \to \infty$

$$l_N = \frac{N}{L}(1 + \tau_1(\log_2 N)) - 1 - \frac{1}{2L} \tau_2(\log_2 N) + O\left(\frac{1}{N}\right),$$

where

$$\tau_1(x) = \sum_{k \neq 0} \chi_k \Gamma(-1 - \chi_k) e^{2k\pi i x},$$

and

$$\tau_2(x) = \sum_{k \neq 0} \chi_k \Gamma(1 - \chi_k) e^{2k\pi i x}.$$

Here τ_1 and τ_2 are continuous periodic functions with period 1 and mean 0 and very small amplitude (observe that

$$|\Gamma(iy)|^2 = \frac{\pi}{y \sinh \pi y}, \tag{3.8}$$

[**1**; (6.1.29)]).

We remark that Knuth [11] has proved this result (with a less accurate expansion) by means of the *Mellin transform technique.* Although this might be feasible in the following investigations too, we prefer the approach from above (sometimes called *Rice's method*) since it makes the computations a lot easier.

For later use we notice

$$l_N^2 = \frac{N^2}{L^2}(1+2\tau_1+\tau_1^2) - \frac{2N}{L}\left(1+\tau_1+\frac{1}{2L}\tau_2+\frac{1}{2L}\tau_1\tau_2\right)+O(1), \qquad (3.9)$$

where the argument $\log_2 N$ is omitted in the periodic functions.

The variance will be computed *via*

$$\operatorname{Var} X_N = w_N + l_N - l_N^2, \qquad (3.10)$$

with

$$w_N = EX_N(X_N-1) = F_N''(1). \qquad (3.11)$$

From the recursion (3.1) we get, for all $N \geqslant 0$,

$$w_N = 2^{1-N}\sum_{k\geqslant 0}\binom{N}{k} l_k l_{N-k} + 2^{2-N}\sum_{k\geqslant 0}\binom{N}{k} l_k + 2^{1-N}\sum_{k\geqslant 0}\binom{N}{k} w_k.$$

Let

$$W(z) = \sum_{N\geqslant 0} w_N \frac{z^N}{N!};$$

then

$$W(z) = 2(L(\tfrac{1}{2}z))^2 + 4e^{z/2}L(\tfrac{1}{2}z) + 2e^{z/2}W(\tfrac{1}{2}z).$$

Now we set

$$\tilde{W}(z) = e^{-z}W(z) = \sum_{N\geqslant 0}\tilde{w}_N\frac{z^N}{N!} \qquad \text{and} \qquad \hat{L}(z) = (\tilde{L}(z))^2$$

and find

$$\tilde{W}(z) = 2\hat{L}(\tfrac{1}{2}z) + 4\tilde{L}(\tfrac{1}{2}z) + 2\tilde{W}(\tfrac{1}{2}z). \qquad (3.12)$$

In order to obtain an explicit expression for the coefficients $\hat{l}_N$ of $\hat{L}(z)$ we square equation (3.5) and get after some straightforward algebraic manipulations

$$\hat{l}_N = \frac{2(-1)^N}{1-2^{2-N}}\left[\frac{N-1}{1-2^{1-N}} - \sum_{j\geqslant 2}\binom{N}{j}\frac{j-1}{2^{j-1}-1} + N\sum_{j\geqslant 2}\binom{N-1}{j}\frac{j-1}{2^{j-1}-1} + 2^{N-3}(N-1)(N-4)\right], \qquad N\geqslant 3, \qquad (3.13)$$

$$\hat{l}_0 = \hat{l}_1 = \hat{l}_2 = 0.$$

From (3.12) we have

$$\tilde{w}_N(1-2^{1-N}) = 2^{1-N}\hat{l}_N + 2^{2-N}\tilde{l}_N.$$

Since $\tilde{w}_2 = 4$ we get for $N \geq 2$

$$\frac{w_{N+1}}{N+1} = -\sum_{k \geq 2} \binom{N}{k} (-1)^k f(k) + 2N, \tag{3.14}$$

with

$$f(z) = \frac{1}{(z+1)(1-2^{-z})} \left[\frac{2z}{1-2^{-z}} - 2z + \frac{2^{1-z}}{1-2^{1-z}} \right.$$
$$\times \left\{ \frac{z}{1-2^{-z}} - \sum_{j \geq 2} \binom{z+1}{j} \frac{j-1}{2^{j-1}-1} \right.$$
$$\left.\left. + (z+1) \sum_{j \geq 2} \binom{z}{j} \frac{j-1}{2^{j-1}-1} + 2^{z-2} z(z-3) \right\} \right]. \tag{3.15}$$

Following our general approach we have to determine the residues of $f(z)[N; z]$ at the poles $z = 1 + \chi_k$, $z = \chi_k$ and $z = -1$. (Observe that all poles are actually simple.) After some tedious but straightforward computations we find the residues.

$$\operatorname{Res}\,(f(z)[N; z]; z = 1) = -N \left[1 + \frac{2}{L} \sum_{j \geq 2} \frac{(-1)^j}{(j+1)(j-1)(2^j-1)} \right.$$
$$\left. + \frac{2}{L} \sum_{j \geq 1} \frac{(-1)^j}{(j+1)(2^j-1)} \right]. \tag{3.16}$$

$$\operatorname{Res}\,(f(z)[N; z]; z = 1 + \chi_k)$$
$$= -\frac{2}{L} N^{1+\chi_k} \Gamma(-\chi_k) \left\{ \frac{1}{2} + \sum_{j \geq 0} \frac{1}{(j+2)(2^{j+1}-1)} \right.$$
$$\left. \times \left[\chi_k \binom{\chi_k - 1}{j} - \binom{\chi_k}{j} \right] \right\}, \qquad k \neq 0. \tag{3.17}$$

$$\operatorname{Res}\,(f(z)[N; z]; z = 0) = \frac{5}{2L} + \frac{2}{L} \sum_{j \geq 1} \frac{(-1)^j}{2^j - 1}. \tag{3.18}$$

$$\operatorname{Res}\,(f(z)[N; z]; z = \chi_k) = \frac{2}{L} N^{\chi_k} \Gamma(1-\chi_k) \left\{ \frac{5+\chi_k}{4(1+\chi_k)} + \sum_{j \geq 0} \frac{1}{(j+2)(2^{j+1}-1)} \right.$$
$$\left. \times \left[(\chi_k - 1) \binom{\chi_k - 2}{j} - \binom{\chi_k - 1}{j} \right] \right\}, \qquad k \neq 0, \tag{3.19}$$

$$\operatorname{Res}\,(f(z)[N; z]; z = -1) = -2/(N+1). \tag{3.20}$$

Now we have derived all necessary asymptotic expansions to treat Var X_N according to equation (3.10). Besides of periodic fluctuations with mean zero the coefficient of N^2 in Var X_N is

$$3 + \frac{2}{L} \sum_{j \geq 2} \frac{(-1)^j}{(j+1)(j-1)(2^j-1)} + \frac{2}{L} \sum_{j \geq 1} \frac{(-1)^j}{(j+1)(2^j-1)}$$
$$- \frac{1}{L^2} - \frac{1}{L^2} \sum_{l \neq 0} |\chi_l|^2 |\Gamma(-1+\chi_l)|^2, \tag{3.21}$$

where the last sum is the zeroeth Fourier coefficient of the periodic function $(\tau_1(x))^2$, with $\tau_1(x)$ from Theorem 2. We start with the treatment of this sum which is the most difficult part of this section and will make use of Ramanujan's identities from Section 1.

PROPOSITION 3.

$$\sum_{l\neq 0} |\chi_l|^2 |\Gamma(-1+\chi_l)|^2 = -1+3L^2+2L\sum_{j\geq 2}\frac{(-1)^j}{(j+1)(j-1)(2^j-1)} + 2L\sum_{j\geq 1}\frac{(-1)^j}{(j+1)(2^j-1)}.$$

Proof. We have

$$\sum_{l\neq 0} |\chi_l|^2 |\Gamma(-1+\chi_l)|^2 = 2\sum_{j\geq 0}(-1)^j \sum_{l\geq 1} |\Gamma(\chi_l)|^2 \frac{1}{|\chi_l|^{2j}}$$

and by (3.8) this becomes

$$L\sum_{j\geq 0}(-1)^j\left(\frac{L^2}{4\pi^2}\right)^j 2(g_j(\alpha)-g_j(2\alpha)), \tag{3.22}$$

with

$$\alpha=\frac{\pi^2}{L} \qquad \text{and} \qquad g_j(\alpha)=\sum_{l\geq 1}\frac{1}{l^{2j+1}(e^{2\alpha l}-1)}.$$

Using Ramanujan's Identity 1 from the introduction and observing $\beta=\pi^2/\alpha=L$ we obtain, for $j\geq 0$,

$$\begin{aligned}
&2(g_j(\alpha)-g_j(2\alpha))\\
&=(-1)^j\left(\frac{\pi^2}{L^2}\right)^j\Bigg[\zeta(2j+1)(1-4^j)+2\left[g_j(L)-4^j g_j\left(\frac{L}{2}\right)\right]\\
&\quad -2^{2j+1}\sum_{k=0}^{j+1}(-1)^{k+j}\frac{B_{2k}}{(2k)!}\frac{B_{2j+2-2k}}{(2j+2-2k)!}\pi^{2j+2-2k}L^{2k-1}[1-2^{2j+1-2k}]\Bigg].
\end{aligned}$$

For $j=0$ we use Identity 2.

$$2(g_0(\alpha)-g_0(2\alpha))=2g_0(L)-2g_0\left(\frac{L}{2}\right)-\frac{11}{12}L+\frac{\pi^2}{6L}.$$

Therefore (3.22) becomes

$$\sum_{j\geq 1}\Bigg\{-L\zeta(2j+1)-2L\sum_{l\geq 1}\frac{1}{l^{2j+1}(2^l-1)} + (2\pi)^{2j+2}\sum_{k=0}^{j+1}(-1)^{k+j}B'_{2k}\left(\frac{L}{2\pi}\right)^{2k}B'_{2j+2-2k}\Bigg\}$$

$$+\sum_{j\geq 1}\left\{L4^{-j}\zeta(2j+1)+2L4^{-j}\sum_{l\geq 1}\frac{1}{l^{2j+1}(2^{2l}-1)}\right.$$

$$\left.-2\pi^{2j+2}\sum_{k=0}^{j+1}(-1)^{k+j}B'_{2k}\left(\frac{L}{\pi}\right)^{2k}B'_{2j+2-2k}\right\}$$

$$+2L\sum_{l\geq 1}\frac{1}{l(2^{2l}-1)}-2L\sum_{l\geq 1}\frac{1}{l(2^{l}-1)}-\frac{11}{12}L^2+\frac{\pi^2}{6}, \tag{3.23}$$

with the abbreviation $B'_n = B_n/n!$.

We treat the first sum in (3.23). It is

$$L\sum_{j\geq 1}\left(1-\zeta(2j+1)-2\sum_{l\geq 2}\frac{1}{l^{2j+1}(2^l-1)}\right)+\sum_{j\geq 1}A_j, \tag{3.24}$$

with

$$A_j=-3L+(-1)^j(2\pi)^{2j+2}\sum_{k=0}^{j+1}(-1)^k B'_{2k}\left(\frac{L}{2\pi}\right)^{2k}B'_{2j+2-2k}.$$

Observe that the sums are now convergent.

From [8; (54.1.4)]

$$\sum_{j\geq 0}(1-\zeta(2j+1))=-\tfrac{1}{4}.$$

Also

$$\sum_{j\geq 1}\sum_{l\geq 2}\frac{1}{l^{2j+1}(2^l-1)}=\sum_{l\geq 2}\frac{1}{(l+1)_3(2^l-1)},$$

where $(x)_n = x(x-1)\dots(x-n+1)$. It remains to treat $\sum A_j$.

By [8; (50.5.10) and (50.5.16)] we have

$$A_j=-3L-(2\pi i)^{2j+2}[t^{2j+2}]\frac{Lt}{4\pi}\cot\left(\frac{Lt}{4\pi}\right)\frac{t}{2}\coth\left(\frac{t}{2}\right)$$

$$=-3L-[t^{2j}]\frac{L\pi}{2}\cot(\pi t)\coth\left(\frac{Lt}{2}\right).$$

Thus

$$\sum_{j\geq 1}A_j=\lim_{t\to 1-}\left\{\frac{-3Lt^2}{1-t^2}-\frac{L\pi}{2}\cot(\pi t)\coth\left(\frac{Lt}{2}\right)+\frac{1}{t^2}+\frac{L^2}{12}-\frac{\pi^2}{3}\right\}$$

$$=\frac{9L}{4}+2L^2+1+\frac{L^2}{12}-\frac{\pi^2}{3}.$$

Therefore (3.24) equals

$$-\tfrac{1}{4}L-2L\sum_{l\geq 2}\frac{1}{(l+1)_3(2^l-1)}+1+\tfrac{9}{4}L+2L^2+\tfrac{1}{12}L^2-\tfrac{1}{3}\pi^2. \tag{3.25}$$

Now we turn to the second sum in (3.23). It is

$$L\sum_{j\geq 1}4^{-j}\zeta(2j+1)+2L\sum_{j\geq 1}4^{-j}\sum_{l\geq 1}\frac{1}{l^{2j+1}(2^{2l}-1)}-\sum_{j\geq 1}C_j, \tag{3.26}$$

with

$$C_j = 2(-1)^j \pi^{2j+2} \sum_{k=0}^{j+1} (-1)^k B'_{2k} \left(\frac{L}{\pi}\right)^{2k} B'_{2j+2-2k}.$$

From [8; (54.3.4)] we know

$$\sum_{j\geq 1} 2^{-(2j+1)} \zeta(2j+1) = \tfrac{1}{2} - \tfrac{1}{2}\psi(\tfrac{3}{2}) - \tfrac{1}{4}\pi \cot(\tfrac{1}{2}\pi) - \tfrac{1}{2}\gamma = -\tfrac{1}{2} + L.$$

Furthermore

$$\sum_{j\geq 1} 2^{-(2j+1)} \sum_{l\geq 1} \frac{1}{l^{2j+1}(2^{2l}-1)} = \sum_{l\geq 1} \frac{1}{(2l+1)_3(2^{2l}-1)}.$$

Now we treat $\sum C_j$. By [8; (50.5.10) and (50.5.16)]

$$C_j = -2(\pi i)^{2j+2} [t^{2j+2}] \frac{Lt}{2\pi} \cot\left(\frac{Lt}{2\pi}\right) \frac{t}{2} \coth\left(\frac{t}{2}\right),$$

and therefore

$$\begin{aligned}\sum_{j\geq 1} C_j &= -\frac{\pi L}{2} \sum_{j\geq 1} [t^{2j}] \coth\left(\frac{Lt}{2}\right) \cot\left(\frac{t\pi}{2}\right)\\ &= -\frac{L\pi}{2}\left[\coth\left(\frac{L}{2}\right)\cot\left(\frac{\pi}{2}\right) - \frac{4}{L\pi} + \frac{\pi}{3L} - \frac{L}{3\pi}\right] = 2 - \tfrac{1}{6}\pi^2 + \tfrac{1}{6}L^2.\end{aligned}$$

So we have obtained for (3.26)

$$-L + 2L^2 + 4L \sum_{l\geq 1} \frac{1}{(2l+1)_3(2^{2l}-1)} - 2 + \tfrac{1}{6}\pi^2 - \tfrac{1}{6}L^2. \tag{3.27}$$

Altogether (3.23) (and thus (3.22)) becomes

$$-1 + L + 3L^2 - 2L \sum_{l\geq 2} \frac{(-1)^{l-1}}{(l+1)_3(2^l-1)} - 2L \sum_{l\geq 1} \frac{(-1)^{l-1}}{l(2^l-1)},$$

and by a short rearrangement of the sums we finally get the proposed formula.

LEMMA 4. $\operatorname{Var} X_N = O(N)$.

Proof. Inserting the result of Proposition 3 for the zeroeth Fourier coefficient of τ_1^2 into (3.21) we find that the non-fluctuating part of the N^2-term in $\operatorname{Var} X_N$ vanishes. Thus

$$\operatorname{Var} X_N = N^2 \tau_3(\log_2 N) + O(N),$$

where $\tau_3(x)$ is continuous since its Fourier series is absolutely convergent, which can be seen by obvious estimates using (3.8), and has mean 0. If $\tau_3(x)$ would not vanish identically we could find an $\varepsilon > 0$ and an interval, say $[a, b] \subseteq [0, 1]$, such that $\tau_3(x) < -\varepsilon$ for $x \in [a, b]$. Since $\log_2 N$ is dense modulo 1, the variance of X_N would be negative for an infinity of values N, an obvious contradiction. This finishes the proof of Lemma 4.

As a byproduct of the last proof we have shown that $\tau_3(x) \equiv 0$, so that all Fourier coefficients of τ_3 must vanish. Collecting the contributions to τ_3 from w_N (*via* (3.14) and (3.17)) and l_N^2 (compare (3.9)) we get the following

COROLLARY 5. *For all integers $k \neq 0$*

$$\sum_{\substack{l+m=k \\ l,m\neq 0}} \chi_l \Gamma(-1-\chi_l)\chi_m\Gamma(-1-\chi_m)$$

$$= 2L\Gamma(-\chi_k)\left\{\frac{1}{2} + \sum_{j\geq 0} \frac{1}{(j+2)(2^{j+1}-1)}\left[\chi_k\binom{\chi_k-1}{j} - \binom{\chi_k}{j}\right]\right\}$$

$$-2\chi_k\Gamma(-1-\chi_k).$$

For the non-fluctuating part of the N-term of Var X_N we gain by a careful collection of all contributions in (3.10)

$$-3-\frac{2}{L}\sum_{j\geq 2}\frac{(-1)^j}{(j+1)(j-1)(2^j-1)}-\frac{2}{L}\sum_{j\geq 1}\frac{(-1)^j}{(j+1)(2^j-1)}$$

$$+\frac{1}{2L}+\frac{2}{L}\sum_{j\geq 0}\frac{(-1)^j}{2^{j+1}-1}+\frac{1}{L^2}[\tau_1\tau_2]_0,$$

where $[\tau_1\tau_2]_0$ indicates the zeroeth Fourier coefficient of $\tau_1(x)\tau_2(x)$. With the same abbreviation we find from Proposition 3 that the first 3 terms may be rewritten as

$$-\frac{1}{L^2}-\frac{1}{L^2}[\tau_1^2]_0.$$

$[\tau_1\tau_2]_0$ could be analysed analogously to our treatment of $[\tau_1^2]_0$ from before. However even weak estimates show that this time the coefficient in question does *not* vanish. For numerical purposes it is even more convenient in this instance *not* to rewrite terms like $[\tau_1^2]_0$ and $[\tau_1\tau_2]_0$, because they are extremely small.

Thus we have proved

THEOREM 6. *The variance* Var X_N *of the number of internal nodes in a trie built from N data is, for* $N\to\infty$,

$$\text{Var } X_N = \frac{N}{L}\left(\frac{1}{2}-\frac{1}{L}+2\sum_{j\geq 0}\frac{(-1)^j}{2^{j+1}-1}-\frac{1}{L}[\tau_1^2]_0 + \frac{1}{L}[\tau_1\tau_2]_0+\tau_4(\log_2 N)\right)+O(1)\approx 0.8461\ldots N,$$

where $L=\log 2$, τ_1, τ_2 and τ_4 are continuous periodic functions with period 1, *mean zero and very small amplitude*; $[\ ]_0$ *means the zeroeth Fourier coefficient.*

Remark. τ_1 and τ_2 are defined in Theorem 2; τ_4 follows from our residue calculation from above to be

$$\tau_4(x) = \sum_{k \neq 0} d_k e^{2k\pi i x},$$

with

$$d_k = -2(1+\chi_k)\Gamma(-\chi_k)\left\{\frac{1}{2} + \sum_{j\geq 0} \frac{1}{(j+2)(2^{j+1}-1)}\left[\chi_k\binom{\chi_k - 1}{j} - \binom{\chi_k}{j}\right]\right\}$$
$$-2\Gamma(1-\chi_k)\left\{\frac{5+\chi_k}{4(1+\chi_k)} + \sum_{j\geq 0}\frac{1}{(j+2)(2^{j+1}-1)}\right.$$
$$\left.\times\left[(\chi_k - 1)\binom{\chi_k - 2}{j} - \binom{\chi_k - 1}{j}\right]\right\}$$
$$+3\chi_k\Gamma(-1-\chi_k) + \frac{1}{L}\chi_k\Gamma(1-\chi_k) + \frac{1}{L}\sum_{\substack{l+m=k\\ l,m\neq 0}} \chi_l\Gamma(-1-\chi_l)\chi_m\Gamma(1-\chi_m).$$

§4. *External internal nodes and generalizations.* Let $F_N(z)$ be the probability generating function for the random variable $Y_N^{(s)}$ described in Section 2. Then

$$\begin{aligned} F_N(z) &= \sum_{k\geq 0} 2^{-N}\binom{N}{k} F_k(z)F_{N-k}(z) \qquad \text{for} \qquad N \geq s+1; \\ F_0(z) &= F_1(z) = \ldots = F_{s-1}(z) = 1, \qquad F_s(z) = z. \end{aligned} \tag{4.1}$$

Since the further treatment is similar to Section 3 we may confine ourselves to a shorter presentation. (We keep all notations for generating functions and coefficients as in Section 3.)

For the expectations $l_N = F'_N(1)$ we have

$$\begin{aligned} l_N &= 2^{1-N}\sum_{k\geq 0}\binom{N}{k} l_k \qquad \text{for} \qquad N\geq s+1, \\ l_0 &= l_1 = \ldots = l_{s-1} = 0, \qquad l_s = 1. \end{aligned} \tag{4.2}$$

From this we get

$$\frac{l_{N+s-2}}{(N+s-2)_{s-2}} = \frac{1-2^{1-s}}{s!}\sum_{k\geq 2}\binom{N}{k}(-1)^k f(k) \tag{4.3}$$

with

$$f(z) = \frac{z(z-1)}{1-2^{3-s-z}}. \tag{4.4}$$

An application of Lemma 1 yields

THEOREM 7. *The expectation $l_N = EY_N^{(s)}$ of the number of minimal subtrees containing exactly $s(\geq 2)$ non-empty external nodes in a trie built from N data is*

$$l_N = \frac{N}{L}\frac{1-2^{1-s}}{s!}\left((s-2)! + \tau_5(\log_2 N) - \frac{1}{2N}\tau_6(\log_2 N)\right) + O\left(\frac{1}{N}\right)$$

where

$$\tau_5(x)=\sum_{k\neq 0}\Gamma(s-1-\chi_k)e^{2k\pi ix}$$

and

$$\tau_6(x)=\sum_{k\neq 0}\chi_k(1+\chi_k)\Gamma(s-1-\chi_k)e^{2k\pi ix}.$$

Remark that the instance $s=2$ covers the number of *external internal nodes.*

For the computation of Var $Y_N^{(s)}$ we need

$$l_N^2=\frac{N^2}{L^2}\frac{(1-2^{1-s})^2}{s!^2}((s-2)!^2+2(s-2)!\tau_5+\tau_5^2)$$
$$-\frac{N}{L^2}\frac{(1-2^{1-s})^2}{s!^2}((s-2)!\tau_6+\tau_5\tau_6)+O(1). \tag{4.5}$$

Again we have

$$\text{Var } Y_N^{(s)}=w_N+l_N-l_N^2,$$

and $w_N=F_N''(1)$ fulfills $(N\geq 0)$

$$w_N=2^{1-N}\sum_{k\geq 0}\binom{N}{k}w_k+2^{1-N}\sum_{k\geq 0}\binom{N}{k}l_kl_{N-k}. \tag{4.6}$$

We derive the solution

$$\frac{w_{N+s-1}}{(N+s-1)_{s-1}}=(-1)^{s+1}\frac{(1-2^{1-s})^2}{s!^2}\sum_{k\geq 2}\binom{N}{k}(-1)^kf(k) \tag{4.7}$$

with

$$f(z)=\frac{(z)_{s+1}2^{2-s-z}}{(1-2^{2-s-z})(1-2^{3-s-z})}\left\{2\sum_{j\geq 0}\binom{z-s-1}{j}\frac{1}{2^{s-1+j}-1}+2^{z-s-1}\right\}. \tag{4.8}$$

Calculating the residues and applying Lemma 1 we obtain besides of periodic fluctuations with mean zero for the coefficient of N^2 in Var $Y_N^{(s)}$

$$\frac{1}{L}\frac{(1-2^{1-s})^2}{s!^2}\left\{(2s-3)!2\sum_{j\geq 0}\binom{2-2s}{j}\frac{1}{2^{s-1+j}-1}+(2s-3)!2^{2-2s}\right.$$
$$\left.-\frac{1}{L}((s-2)!)^2-\frac{1}{L}\sum_{l\neq 0}|\Gamma(s-1+\chi_l)|^2\right\}. \tag{4.9}$$

A further application of Ramanujan's identities from Section 1 allows to derive the following

PROPOSITION 8. *For $s\geq 2$ we have*

$$\sum_{l\neq 0}|\Gamma(s-1+\chi_l)|^2=L\left\{2(2s-3)!\sum_{j\geq 0}\binom{2-2s}{j}\frac{1}{2^{s-1+j}-1}+(2s-3)!2^{2-2s}\right\}$$
$$-\{(s-2)!\}^2.$$

Proof. We have

$$\sum_{l\neq 0} |\Gamma(s-1+\chi_l)|^2 = 2\sum_{l\geq 1} |(s-2+\chi_l)_{s-1}|^2 |\Gamma(\chi_l)|^2$$

$$= 2\sum_{l\geq 1}\left(\prod_{j=0}^{s-2}(|\chi_l|^2+j^2)\right)\frac{L}{2l\sinh\dfrac{2l\pi^2}{L}}$$

$$= L\sum_{t=1}^{s-1} c_t(s)\left(\frac{4\pi^2}{L^2}\right)^t \sum_{l\geq 1}\frac{l^{2t-1}}{\sinh\dfrac{2l\pi^2}{L}}, \tag{4.10}$$

with

$$\sum_{t=1}^{s-1} c_t(s)x^t = \prod_{j=0}^{s-2}(x+j^2). \tag{4.11}$$

Now we consider the last sum in (4.10).

$$\sum_{l\geq 1}\frac{l^{2t-1}}{\sinh(2l\alpha)} = 2\sum_{l\geq 1}\frac{l^{2t-1}}{e^{2l\alpha}-1} - 2\sum_{l\geq 1}\frac{l^{2t-1}}{e^{4l\alpha}-1}$$

Using Ramanujan's Identities 3 and 4 from the introduction we have, with $\alpha = \pi^2/L$ and $\beta = L$,

$$\sum_{l\geq 1}\frac{l^{2t-1}}{\sinh\dfrac{2l\pi^2}{L}} = 2(-1)^{t+1}\left(\frac{L^2}{4\pi^2}\right)^t\left\{\sum_{l\geq 1}\frac{(-1)^{l-1}l^{2t-1}}{2^l-1} + \frac{B_{2t}}{4t}(2^{2t}-1) - \delta_{t,1}\frac{1}{2L}\right\},$$

so that (4.10) can be written as

$$2L\left\{\sum_{t=1}^{s-1} c_t(s)(-1)^{t+1}\sum_{l\geq 1}\frac{(-1)^{l-1}l^{2t-1}}{2^l-1}\right.$$
$$\left. + \sum_{t=1}^{s-1} c_t(s)(-1)^{t+1}\frac{B_{2t}}{4t}(2^{2t}-1)\right\} - c_1(s), \tag{4.12}$$

and

$$c_1(s) = \{(s-2)!\}^2.$$

Now we treat the sums in (4.12). The more complicated sum is the second one. From (4.11)

$$\sum_{t=1}^{s-1} c_t(s)(-1)^t u^{2t} = \sum_{j=0}^{s-2}(j-u)(j+u) = (-1)^{s-1}u(u+s-2)_{2s-3},$$

so that

$$\frac{c_t(s)(-1)^t}{(2s-3)!(-1)^{s-1}} = [u^{2t-1}]\binom{u+s-2}{2s-3} = \sum_j \binom{s-2}{2s-3-j}[u^{2t-1}]\binom{u}{j}$$
$$= \sum_j \binom{s-2}{2s-3-j}\frac{S_1(j,2t-1)}{j!},$$

where $S_1(n, k)$ denote the Stirling numbers of the first kind,

$$[x^{2s-3}](1+x)^{s-2}\frac{(\log(1+x))^{2t-1}}{(2t-1)!}.$$

Therefore we have

$$\sum_{t=1}^{s-1} c_t(s)(-1)^{t+1}\frac{B_{2t}}{4t}(2^{2t}-1)$$

$$=\frac{(2s-3)!}{2}(-1)^{s-1}[x^{2s-3}](1+x)^{s-2}\sum_{r\geq 2}\frac{B_r}{r!}(\log(1+x))^{r-1}(1-2^r). \quad (4.13)$$

Since

$$\sum_{r\geq 0}\frac{B_r}{r!}u^r=\frac{u}{e^u-1},$$

the last sum equals

$$\frac{1}{x}-\frac{2}{2x+x^2}+B_1,$$

and (4.13) turns into

$$\frac{(2s-3)!}{2}(-1)^{s-2}[x^{2s-3}](1+x^{s-2})\frac{1}{x(1+\frac{1}{2}x)}$$

$$=\frac{(2s-3)!}{2}(-1)^{s-2}\sum_i\binom{s-2}{i}(-1)^i 2^{-2s+2+i}=\frac{(2s-3)!}{2}2^{-2s+2}. \quad (4.14)$$

Thus we have found a short expression for the second sum in (4.12). The treatment of the first sum is similar but easier.

$$\sum_{t=1}^{s-1} c_t(s)(-1)^{t+1}\sum_{l\geq 1}\frac{(-1)^{l-1}l^{2t-1}}{2^l-1}=\sum_{l\geq 1}\frac{(-1)^l}{2^l-1}(-1)^{s-1}(l+s-2)_{2s-3},$$

and with $j=l-s+1$, this is

$$(2s-3)!\sum_{j\geq 0}(-1)^j\binom{j+2s-2}{j}\frac{1}{2^{s-1+j}-1}$$

$$=(2s-3)!\sum_{j\geq 0}\binom{2-2s}{j}\frac{1}{2^{s-1+j}-1}. \quad (4.15)$$

Inserting (4.14) and (4.15) into (4.12) completes the proof of the proposition.

With the same argument as in Section 3 we have as an immediate consequence

LEMMA 9. $\operatorname{Var} Y_N^{(s)} = O(N)$.

Again we derive as a byproduct that all Fourier coefficients of the fluctuating part of the N^2-term in $\operatorname{Var} Y_N^{(s)}$ must vanish. Collecting the corresponding contributions from the residues in $z = 1 + \chi_k$, $(k \neq 0)$ of $[N; z]f(z)$ with $f(z)$ from (4.8) we obtain the following identities.

COROLLARY 10. *For integer k, s with $k \neq 0$ and $s \geq 2$:*

$$\sum_{\substack{l+m=k \\ l,m \neq 0}} \Gamma(s-1-\chi_l)\Gamma(s-1-\chi_m)$$
$$= -2(s-2)!\Gamma(s-1-\chi_k) + L\Gamma(2s-2-\chi_k)$$
$$\times \left\{ 2 \sum_{j \geq 0} \binom{\chi_k + 2 - 2s}{j} \frac{1}{2^{s-1+j}-1} + 2^{2-2s} \right\}.$$

By some tedious but straightforward computations the N-term of $\operatorname{Var} Y_N^{(s)}$ can be determined.

THEOREM 11. *The variance of the number of minimal subtrees containing exactly $s(\geq 2)$ non-empty external nodes in a trie built from N data is*

$$\operatorname{Var} Y_N^{(s)} = N \frac{(1-2^{1-s})^2}{L(s!)^2} \left\{ \frac{s!(s-2)!}{1-2^{1-s}} - 2^{2-2s}(2s-3)!s \right.$$
$$- 2(2s-3)! \sum_{j \geq 0} \left[\binom{2-2s}{j} + (2s-2) \binom{1-2s}{j} \right] \frac{1}{2^{s-1+j}-1}$$
$$\left. - \frac{1}{L} [\tau_5 \tau_6]_0 \right\} + N\tau_7(\log_2 N) + O(1)$$

where τ_5, τ_6, τ_7 are continuous periodic functions of period 1, mean 0 and very small amplitude.

Remark. τ_5 and τ_6 appear in Theorem 7; the Fourier coefficients of τ_7 are rather complicated and therefore omitted for brevity. $[\tau_5\tau_6]_0$ might be expressed in terms of sums similar to $[\tau_5^2]_0$, but since the N-term of the variance does *not* vanish it is more convenient to stay with the original expression because it can be safely neglected for numerical purposes.

§5. *Concluding remarks.* In Proposition 8 and Corollary 10 we derived convolution formulae for the Γ-function, namely for the expressions

$$\sum_{\substack{l+m=k \\ l,m \neq 0}} \Gamma(n-\chi_l)\Gamma(n-\chi_m), \qquad n \in \mathbb{N}. \tag{5.1}$$

The instance $n = 0$ occurs also in a problem in trie statistics (compare [**10**]), but it is significantly easier than the instances $n \geq 1$ treated in this paper.

For $k=0$ it follows immediately from the functional equation for Dedekind's η-function that

$$\sum_{l\neq 0}|\Gamma(\chi_l)|^2=\tfrac{1}{6}\pi^2-2L\sum_{j\geq 1}\frac{(-1)^{j-1}}{j(2^j-1)}-\tfrac{11}{12}L^2 \tag{5.2}$$

(compare [9]).

In a similar way as in Sections 3 and 4 we get as a corollary for $k\neq 0$:

$$\sum_{\substack{l+m=k\\ l,m\neq 0}}\Gamma(-\chi_l)\Gamma(-\chi_m)=2\Gamma(-\chi_k)\left\{\gamma+L\sum_{j\geq 1}\binom{\chi_k}{j}\frac{1}{2^j-1}+\psi(-\chi_k)\right\}. \tag{5.3}$$

The instance $n=-1$ of (5.1) is an immediate consequence of (5.2) and Proposition 3, resp. (5.3) and Corollary 5. Since

$$\begin{aligned}\sum_{\substack{l+m=k\\ l,m\neq 0}}&\chi_l\Gamma(-1-\chi_l)\chi_m\Gamma(-1-\chi_m)\\ &=\sum(1+\chi_l)\Gamma(-1-\chi_l)(1+\chi_m)\Gamma(-1-\chi_m)\\ &\quad-\sum(1+\chi_l+\chi_m)\Gamma(-1-\chi_l)\Gamma(-1-\chi_m)\\ &=\sum\Gamma(-\chi_l)\Gamma(-\chi_m)-(1+\chi_k)\sum\Gamma(-1-\chi_l)\Gamma(-1-\chi_m),\end{aligned}$$

we have, for $k=0$,

$$\sum_{l\neq 0}|\Gamma(-1+\chi_l)|^2=1+\tfrac{1}{6}\pi^2-L-2L\sum_{j\geq 2}\frac{(-1)^j}{(j+1)_3(2^j-1)}-\tfrac{47}{12}L^2 \tag{5.4}$$

and, for $k\neq 0$,

$$\begin{aligned}\sum_{\substack{l+m=k\\ l,m\neq 0}}&\Gamma(-1-\chi_l)\Gamma(-1-\chi_m)\\ &=2\Gamma(-1-\chi_k)\left\{1-\gamma+\tfrac{1}{2}L-L\sum_{j\geq 1}\binom{1+\chi_k}{j}\frac{1}{(j+1)(2^j-1)}-\psi(-1-\chi_k)\right\}.\end{aligned} \tag{5.5}$$

Finally we want to mention that periodic fluctuations occur frequently in distribution results. We give some examples.

(i) *Sum-of-digits.* A by now classical result of Delange [3] gives the following expression for the average of the sum of digits $S_q(n)$ in the q-ary representation of n

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

where

$$F(x)=\sum_{k\in \mathbf{Z}}c_k e^{2k\pi i x}$$

with

$$c_0=\frac{q-1}{2\log q}(\log 2\pi-1)-\frac{q+1}{4}$$

and $(k \neq 0)$

$$c_k = i\frac{q-1}{2k\pi}\left(1+\frac{2k\pi i}{\log q}\right)^{-1}\zeta\left(\frac{2k\pi i}{\log q}\right).$$

(ii) *Register allocation.* Flajolet, Raoult and Vuillemin [6] have investigated the average R_N of the optimal number of registers needed to evaluate an arithmetic expression with N binary operators.

$$R_N = \log_4 N + R(\log_4 N) + O(N^{-\frac{1}{2}+\epsilon})$$

where

$$R(x) = \sum_{k \in \mathbf{Z}} r_k e^{2k\pi i x},$$

with

$$r_0 = -\frac{1}{2} - \frac{1}{L} - \frac{\gamma}{2L} + \log_2 (2\pi),$$

and $(k \neq 0)$

$$r_k = \frac{1}{L}(\chi_k - 1)\Gamma\left(\frac{\chi_k}{2}\right)\zeta(\chi_k).$$

Flajolet [4] has also computed the variance; an evaluation of $(R(x))^2$ similar to our treatment was not performed since it was not necessary in order to show the cancellation of the $\log^2 N$ term.

Similar results concerning the Gray code representation of natural numbers and an algorithm called "odd-even-merging" were obtained by Flajolet and Ramshaw [5].

We hope that the methods used in this paper might be a useful attempt in order to get precise results on the centralized moments in these and several other problems.

References

1. M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions* (Dover, New York, 1970).
2. B. C. Berndt. Modular transformations and generalizations of several formulae of Ramanujan. *Rocky Mountain J. Math*, 7 (1977), 147-189.
3. H. Delange. Sur la fonction sommatoire de la fonction somme des chiffres. *l'Enseignement Mathématique*, 21 (1975), 31-47.
4. P. Flajolet. Analyse d'algorithmes de manipulation d'arbres et de fichiers. *Cahier de BURO*, 34-35 (1981), 1-209.
5. P. Flajolet and L. Ramshaw. A note on Gray-Code and Odd-Even Merge. *SIAM J. Computing*, 9 (1980), 142-158.
6. P. Flajolet, J.-C. Raoult and J. Vuillemin. The number of registers required for evaluating arithmetic expressions. *Theoret. Comp. Sci.*, 9 (1979), 99-125.
7. P. Flajolet and R. Sedgewick. Digital search trees revisited. *SIAM J. Comput.*, 15 (1986), 748-767.
8. E. R. Hansen. *A table of series and products* (Prentice-Hall, Englewood Cliffs, 1975).
9. P. Kirschenhofer, H. Prodinger and J. Schoissengeier. Zur Auswertung gewisser Reihen mit

:0. P. Kirschenhofer, H. Prodinger and W. Szpankowski. On the variance of the external path length in a symmetric digital trie. Discrete Applied Math., 25 (1989), 129–143.
1. D. E. Knuth. *The art of computer programming*, Vol. 3: "*Sorting and searching*" (Addison Wesley, Reading Mass, 1973).
'2. N. E. Nörlund. *Vorlesungen über Differenzenrechnung* (Chelsea, New York, 1954).
·3. S. Ramanujan. *Notebooks of Srinivasa Ramanujan* (2 volumes) (Tata Institute of Fundamental Research, Bombay 1957).
'4. U. Schmid. *Analyse von Collision-Resolution Algorithmen in Random-Access Systemen mit dominanten Übertragungskanälen* (Dissertation, T. U. Wien, 1986).

)r. P. Kirschenhofer,
nstitut für Algebra und Diskrete Mathematik,
'echnische Universität Wien,
Viedner Hauptstrasse 8-10,
\-1040 Wien,
\ustria.

)r. H. Prodinger,
nstitut für Algebra und Diskrete Mathematik,
'echnische Universität Wien,
Viedner Hauptstrasse 8-10,
\-1040 Wien,
\ustria.

68R10: *COMPUTER SCIENCE; Discrete mathematics; Graph theory.*

Received on the 26th of May, 1987.