ON THE m-ENUMERATION OF MAXIMUM AND LEFT-TO-RIGHT MAXIMA IN GEOMETRICALLY DISTRIBUTED WORDS

HELMUT PRODINGER

Abstract. Words with letters being natural numbers equipped with geometric probabilities are counted using a novel idea of Cichon and Macyna's, namely by sending the letters at random to one of m subwords, where the parameters *maximum* and number of *left-to-right maxima* are counted as usual. The final result is the sum of the m individual results. A technique is described how to translate *explicit* orginal formulæ as alternating sums into similar sums including the parameter m. The technique of choice for the asymptotic enumeration of moments is *Rice's method*.

1. Introduction

Cichon, together with his coauthor Macyna, had the seminal idea [1] to generalize *approximate counting* to *approximate counting with m counters*.

We briefly explain approximate counting, as in the original version [2]. A counter is initially set to 1; when new items (to be counted) arrive, the counter either keeps its value or is increased by 1, according to a random experiment with geometric probability: if the counter value is k, then an increase happens with probability 2^{-k}. After n random increments, the counter value is approximately $\log_2 n$ (much more precise results are known), so that it can be used to estimate the number of random items.

The new version uses m counters, and chooses for each incoming item one of these counters at random (with probability $\frac{1}{m}$) where it is dealt with as usual. The *result* of the procedure is the sum of the individual results of the m counters. This total count has been analyzed in [10, 4, 7] and behaves on average like $m \log_2 n$.

This fundamental idea should, however, not be restricted to approximate counting, as there are other important parameters which have been analyzed under the assumption that random words of length n are given. The letters are typically *integers*, and integer k appears with the *geometric probability* pq^{k-1}, where $p+q=1$.

We will deal with 3 such parameters in this paper, namely the *maximum*, and the number of *left-to-right maxima* in the strict/weak sense. Recall that a letter in a word is a left-to-right maximum in the strict/weak sense if it is strictly larger or just larger or equal than all the letters that have been seen so far. Here, we assume again m devices to which

Date: September 28, 2012.

2010 *Mathematics Subject Classification.* 05A16, 60C05.

Key words and phrases. Geometrically distributed words, maximum, records, Rice's method, asymptotic expansion.

The author was supported by an incentive grant from the NRF in South Africa.

each incoming symbol is randomly associated. Each device computes its maximum resp. left-to-right maximum, and at the end the results are added.

Intuitively, we assume that each device sees roughly n/m letters, and since each parameter has typically an average of $c\cdot\log n$, we should get a result of the form $c\cdot m\log\frac{n}{m}$. This will be made precise in the present paper.

Here is the general set-up. Let $P_{n,l}$ be the probability that the parameter of interest is l after seeing n random letters, and $\mathcal{P}_{n,l}$ the probability that, in the m-model, the total (accumulated) value is l. This follows a multinomial distribution:

$$\mathcal{P}_{N,l}=m^{-N}\sum_{n_1+\cdots+n_m=N}\binom{N}{n_1,\ldots,n_m}\sum_{l_1+\cdots+l_m=l}P_{n_1,l_1}\cdots P_{n_m,l_m}.$$

The first two moments are in all cases computed via *probability generating functions*.

Let

$$G_n(u)=\sum_{l\ge0}P_{n,l}u^l\quad\text{and}\quad\mathcal{G}_n(u)=\sum_{l\ge0}\mathcal{P}_{n,l}u^l.$$

It is beneficial to work with exponential generating functions

$$F(z,u)=\sum_{n\ge0}G_n(u)\frac{z^n}{n!}\quad\text{and}\quad\mathcal{F}(z,u)=\sum_{n\ge0}\mathcal{G}_n(u)\frac{z^n}{n!}.$$

A simple rearrangement gives us

$$\mathcal{F}(mz,u)=\big(F(z,u)\big)^m.$$

Let us denote differentiations with respect to u by a prime. Then

$$\mathcal{F}'(mz,1)=m\big(F(z,1)\big)^{m-1}F'(z,1)=me^{z(m-1)}F'(z,1)$$

and

$$\mathcal{F}''(mz,1)=me^{z(m-1)}F''(z,1)+m(m-1)e^{z(m-2)}\big(F'(z,1)\big)^2.$$

We use the following notations for moments:

$$E_n=n![z^n]F'(z,1),\qquad\mathcal{E}_n=n![z^n]\mathcal{F}'(z,1),$$
$$E_n^{(2)}=n![z^n]F''(z,1),\qquad\mathcal{E}_n^{(2)}=n![z^n]\mathcal{F}''(z,1).$$

Comparing coefficients, we get

$$\mathcal{E}_N=m^{1-N}\sum_{n=1}^{N}\binom{N}{n}(m-1)^{N-n}E_n$$

and

$$\mathcal{E}_N^{(2)}=m^{1-N}\sum_{n=1}^{N}\binom{N}{n}(m-1)^{N-n}E_n^{(2)}+(m-1)m^{1-N}\sum_{n=1}^{N}\binom{N}{n}(m-2)^{N-n}\sum_{j=0}^{n}\binom{n}{j}E_jE_{n-j}.$$

In all our examples, there is a formula

$$E_n = \sum_{k=1}^{n} \binom{n}{k} (-1)^k \omega(k).$$

Therefore

$$\begin{aligned}
\mathcal{E}_N &= m^{1-N} \sum_{n=1}^{N} \binom{N}{n} (m-1)^{N-n} \sum_{k=1}^{n} \binom{n}{k} (-1)^k \omega(k) \\
&= m^{1-N} \sum_{k=1}^{N} \binom{N}{k} (-1)^k \omega(k) \sum_{n=k}^{N} (m-1)^{N-n} \binom{N-k}{n-k} \\
&= \sum_{k=1}^{N} \binom{N}{k} (-1)^k \omega(k) m^{1-k}.
\end{aligned}$$

There is also

$$E_n^{(2)} = \sum_{k=1}^{n} \binom{n}{k} (-1)^k \theta(k).$$

Here, the translation is more complicated. First,

$$\begin{aligned}
\sum_{j=0}^{n} \binom{n}{j} E_j E_{n-j} &= \sum_{j=0}^{n} \binom{n}{j} \sum_{a=1}^{j} \binom{j}{a} (-1)^a \omega(a) \sum_{b=1}^{n-j} \binom{n-j}{b} (-1)^b \omega(b) \\
&= \sum_{a,b=1}^{n} \binom{n}{a+b} \binom{a+b}{b} (-1)^{a+b} \omega(a) \omega(b) 2^{n-a-b}.
\end{aligned}$$

Second,

$$\begin{aligned}
&\sum_{n=1}^{N} \binom{N}{n} (m-2)^{N-n} \sum_{j=0}^{n} \binom{n}{j} E_j E_{n-j} \\
&\quad = \sum_{n=1}^{N} \binom{N}{n} (m-2)^{N-n} \sum_{a,b=1}^{n} \binom{n}{a+b} \binom{a+b}{b} (-1)^{a+b} \omega(a) \omega(b) 2^{n-a-b} \\
&\quad = \sum_{n=1}^{N} (m-2)^{N-n} \sum_{a,b=1}^{n} \binom{N-a-b}{N-n} \binom{N}{a+b} \binom{a+b}{b} (-1)^{a+b} \omega(a) \omega(b) 2^{n-a-b} \\
&\quad = \sum_{a,b=1}^{N} m^{N-a-b} \binom{N}{a+b} \binom{a+b}{b} (-1)^{a+b} \omega(a) \omega(b) \\
&\quad = m^N \sum_{k=1}^{N} \binom{N}{k} (-1)^k m^{-k} \sum_{j=1}^{k-1} \binom{k}{j} \omega(j) \omega(k-j).
\end{aligned}$$

Consequently

$$\mathcal{E}_N^{(2)} = m\sum_{k=1}^{N}\binom{N}{k}(-1)^k\theta(k)m^{-k} + (m-1)m\sum_{k=1}^{N}\binom{N}{k}(-1)^k m^{-k}\sum_{j=1}^{k-1}\binom{k}{j}\omega(j)\omega(k-j).$$

The paper [10] has already essentially these developments, but in a less systematic way.

The further analysis depends on the specific form of $\omega(k)$ and $\theta(k)$.

We need the following abbreviations, which are traditional in this area: $Q := 1/q$, $L := \log Q$, $\chi_k = \frac{2\pi i k}{L}$,

$$\alpha = \sum_{k\ge1}\frac{1}{Q^k-1}, \quad \beta = \sum_{k\ge1}\frac{1}{(Q^k-1)^2}, \quad \text{and} \quad \tau = \sum_{k\ge1}\frac{(-1)^{k-1}}{k(Q^k-1)}.$$

These constants appear in the local expansion around $z \sim 0$ of

$$\sum_{k\ge1}\frac{1}{Q^{z+k}-1} \quad \text{and} \quad \sum_{k\ge1}\binom{z}{k}\frac{1}{Q^k-1}.$$

Our asymptotic method of choice is often called *Rice's method* and described in great detail in [3]. It allows to write a sum

$$\sum_{k=1}^{n}\binom{n}{k}(-1)^k\omega(k)$$

as a contour integral

$$-\frac{1}{2\pi i}\int_{\mathcal{C}}\frac{\Gamma(n+1)\Gamma(-z)}{\Gamma(n+1-z)}\omega(z)dz$$

where $\omega(z)$ is an analytic extension of the sequence $\omega(k)$, and the curve of integration includes the poles at $k = 1,\dots,n$ and no others. Changing the contour brings the residues outside of the curve into the game (with a negative sign), and they constitute the asymptotic expansion. In our application, this residue is located at $z = 0$. There are also poles at $z = \chi_k$ (also with real part $= 0$) which contribute; this contribution is a Fourier series and thus establishes a fluctuating function. It is of small amplitude since the Gamma function decays very quickly along the imaginary axis.

There is another interesting aspect, namely that such a periodic function of mean zero, which appears in the expected value, must be squared when computing the variance. This function is again a periodic function, but no longer of mean zero, although this mean is extremally small. In various earlier analyses of algorithms, it was crucial to find an alternative representation for this tiny quantity. There is a survey article [9] about that. One of these identities will be used in the present paper as well.

Several residues must be computed. After due preparations, this can be done nowadays with a computer algebra system. This is in contrast with earlier efforts, e. g. [8], where such expansions had to be performed by hand.

2. Maximum

In the instance of the maximum (the largest letter in the word) we know from [11] and [6] that

$$\omega(k) = \frac{-1}{1-q^k} \quad \text{and} \quad \theta(k) = \frac{-2q^k}{(1-q^k)^2}.$$

We must compute the residue at $z=0$ of

$$\frac{\Gamma(N+1)\Gamma(-z)}{\Gamma(N+1-z)}\frac{-1}{1-q^z}m^{1-z},$$

which is (a computer can do that)

$$m\log_Q N - m\log_Q m + \frac{m}{2} + \frac{m\gamma}{L} + O(N^{-1}).$$

The contribution from the residues at $z=\chi_k$ is given by

$$-\frac{m}{L}\sum_{k\neq 0}\Gamma(-\chi_k)e^{2\pi i\cdot\log_Q\frac{N}{m}} =: -m\delta(\log_Q \tfrac{N}{m}).$$

Now we look at the residue at $z=0$ of

$$\frac{\Gamma(N+1)\Gamma(-z)}{\Gamma(N+1-z)}\frac{-2q^z}{(1-q^z)^2}m^{1-z},$$

which is

$$\begin{aligned}&m\log_Q^2 N - 2m\log_Q m\log_Q N + \frac{2m\gamma}{L}\log_Q N\\ &-\frac{m}{6} + \frac{m\pi^2}{6L^2} + \frac{m\gamma^2}{L^2} + m\log_Q^2 m - \frac{2m\gamma}{L}\log_Q m + O(N^{-1}).\end{aligned}$$

The other part of the second factorial moment requires some preparation. We have

$$\begin{aligned}&\sum_{j=1}^{k-1}\binom{k}{j}\frac{1}{1-q^j}\frac{1}{1-q^{k-j}}\\ &=\frac{1}{1-q^k}\sum_{j=1}^{k-1}\binom{k}{j}\left[\frac{1}{1-q^j}+\frac{1}{1-q^{k-j}}-1\right]\\ &=\frac{2}{1-q^k}\sum_{j=1}^{k-1}\binom{k}{j}\frac{1}{1-q^j}-\frac{1}{1-q^k}\sum_{j=1}^{k-1}\binom{k}{j}\\ &=\frac{2}{1-q^k}\sum_{j=1}^{k-1}\binom{k}{j}\frac{q^j}{1-q^j}+\frac{2}{1-q^k}\sum_{j=1}^{k-1}\binom{k}{j}-\frac{1}{1-q^k}\sum_{j=1}^{k-1}\binom{k}{j}\\ &=\frac{2}{1-q^k}\sum_{j=1}^{k-1}\binom{k}{j}\frac{q^j}{1-q^j}+\frac{1}{1-q^k}\sum_{j=1}^{k-1}\binom{k}{j}\end{aligned}$$

$$= \frac{2}{1-q^k}\sum_{j=1}^{k}\binom{k}{j}\frac{q^j}{1-q^j} - \frac{2q^k}{(1-q^k)^2} + \frac{2^k-2}{1-q^k}.$$

Hence we have to compute the residue at $z=0$ of

$$\frac{\Gamma(N+1)\Gamma(-z)}{\Gamma(N+1-z)}(m-1)m^{1-z}\Big[\frac{2}{1-q^z}\sum_{j\ge1}\binom{z}{j}\frac{q^j}{1-q^j} - \frac{2q^z}{(1-q^z)^2} + \frac{2^z-2}{1-q^z}\Big].$$

It is

$$m(m-1)\log_Q^2 N + \frac{m(m-1)2\gamma}{L}\log_Q N - 2m(m-1)\log_Q m\log_Q N + m(m-1)\log_Q N$$
$$+ m(m-1)\log_Q^2 m - m(m-1)\log_Q m - \frac{2m(m-1)\gamma}{L}\log_Q m$$
$$+ m(m-1)\Big[\frac13 + \frac{\pi^2}{6L^2} + \frac{\gamma}{L} - \frac{\log 2}{L} + \frac{\gamma^2}{L^2}\Big] - \frac{2m(m-1)}{L}\tau.$$

Collecting all the ingredients for the variance, we get

$$\frac{m^2}{12} + \frac{m^2\pi^2}{6L^2} - m(m-1)\log_Q 2 - \frac{2m(m-1)}{L}\tau - m^2[\delta^2]_0 + m\delta_V(\log_Q \tfrac{N}{m}) + O(N^{-1}),$$

where the periodic function $\delta_V(x)$ could also be expressed as a Fourier series. The quantity $[\delta^2]_0$ is the constant term in $\delta^2(x)$; it is a tiny quantity and possessed an alternative form [8]:

$$[\delta^2]_0 = \frac{\pi^2}{6L^2} + \frac{1}{12} - \log_Q 2 - \frac{2}{L}\tau.$$

This allows to simplify the variance. We collect the results of this section.

Theorem 1. *Mean and variance of the m-version of the maximum parameter of geometrically distributed words satisfy*

$$\mathcal{E}_N = m\log_Q N - m\log_Q m + \frac{m}{2} + \frac{m\gamma}{L} - m\delta(\log_Q \tfrac{N}{m}) + O(N^{-1})$$

and

$$\mathcal{V}_N = m\log_Q 2 + \frac{2m}{L}\tau + m\delta_V(\log_Q \tfrac{N}{m}) + O(N^{-1}).$$

Notice that the constant in the variance is very close to

$$\frac{m\pi^2}{6L^2} + \frac{m}{12}.$$

3. Left-to-right maxima in the strict sense

A left-to-right maximum in the strict sense is a letter that is strictly larger than all the letters to the left. A probability generating function is available [8]:

$$\prod_{k\ge1}\Big(1 + \frac{zupq^{k-1}}{1-z(1-q^k)}\Big);$$

the coefficient of $z^n u^k$ in it is the probability that a random word of length n has k left-to-right maxima.

From this an explicit expression for the expectation E_n as an alternating sum is derived, with

$$\omega(k) = \frac{-p}{1-q^k}.$$

Twice differentiating w.r.t. u, followed by $u = 1$ leads to the second factorial moments. Since there are small inaccuracies in the original paper [8], we take the chance to correct them here:

$$\begin{aligned} E_n^{(2)} &= 2p^2 \sum_{0\le i<j} \Big[1 + \frac{1}{q^{i-j}-1}(1-q^i)^n + \frac{1}{q^{j-i}-1}(1-q^j)^n\Big] \\ &= 2p^2 \sum_{k=1}^{n} \binom{n}{k} (-1)^k \sum_{0\le i<j} \Big[\frac{1}{q^{i-j}-1} q^{ik} + \frac{1}{q^{j-i}-1} q^{jk}\Big]. \end{aligned}$$

Therefore

$$\begin{aligned} \frac{\theta(k)}{2p^2} &= \sum_{0\le i<j} \frac{1}{q^{i-j}-1} q^{ik} + \sum_{0\le i<j} \frac{1}{q^{j-i}-1} q^{jk} \\ &= \alpha \sum_{0\le i} q^{ik} - \sum_{0\le i<j} \frac{1}{1-q^{j-i}} q^{jk} \\ &= \alpha \frac{1}{1-q^k} - \sum_{0\le i<j} q^{jk} - \sum_{0\le i<j} \frac{1}{Q^{j-i}-1} q^{jk} \\ &= \alpha \frac{1}{1-q^k} - \sum_{0<j} j q^{jk} - \sum_{0<j} \sum_{h=1}^{j} \frac{1}{Q^h-1} q^{jk} \\ &= \alpha + \alpha \sum_{j\ge 1} q^{jk} - \frac{q^k}{(1-q^k)^2} - \sum_{j\ge 1} \sum_{h=1}^{j} \frac{1}{Q^h-1} q^{jk} \\ &= \alpha - \frac{q^k}{(1-q^k)^2} + \sum_{j\ge 1} \sum_{h>j} \frac{1}{Q^h-1} q^{jk} \\ &= \alpha - \frac{q^k}{(1-q^k)^2} + \sum_{h\ge 1} \frac{1}{Q^h-1} \sum_{j=1}^{h-1} q^{jk} \\ &= \alpha - \frac{Q^k}{(Q^k-1)^2} + \sum_{h\ge 1} \frac{1}{Q^h-1} \frac{q^k - q^{hk}}{1-q^k} \\ &= \alpha \frac{1}{1-q^k} - \frac{Q^k}{(Q^k-1)^2} - \sum_{h\ge 1} \frac{1}{Q^h-1} \frac{q^{hk}}{1-q^k} \end{aligned}$$

$$= \alpha \frac{1}{1-q^k} - \frac{Q^k}{(Q^k-1)^2} - \frac{1}{1-q^k} \sum_{h,j\ge1} q^{hk+hj}$$
$$= \alpha \frac{1}{1-q^k} - \frac{Q^k}{(Q^k-1)^2} - \frac{1}{1-q^k} \sum_{j\ge1} \frac{1}{Q^{k+j}-1}.$$

Furthermore, we get as in the previous section

$$\sum_{j=1}^{k-1} \binom{k}{j} \frac{1}{1-q^j} \frac{1}{1-q^{k-j}} = \frac{2}{1-q^k} \sum_{j=1}^{k} \binom{k}{j} \frac{q^j}{1-q^j} - \frac{2q^k}{(1-q^k)^2} + \frac{2^k-2}{1-q^k}.$$

So we must look at the residue at $z=0$ of

$$\frac{\Gamma(N+1)\Gamma(-z)}{\Gamma(N+1-z)} 2p^2 m^{1-z} \Big[\alpha \frac{1}{1-q^z} - \frac{Q^z}{(Q^z-1)^2} - \frac{1}{1-q^z} \sum_{j\ge1} \frac{1}{Q^{z+j}-1} \Big]$$
$$+ \frac{\Gamma(N+1)\Gamma(-z)}{\Gamma(N+1-z)} (m-1) p^2 m^{1-z} \Big[\frac{2}{1-q^z} \sum_{j\ge1} \binom{z}{j} \frac{q^j}{1-q^j} - \frac{2q^z}{(1-q^z)^2} + \frac{2^z-2}{1-q^z} \Big].$$

There are again many terms, which we decided not to display, but when computing the variance, thanks to many cancellations, what remains is

$$mpq \log_Q N - mpq \log_Q m + \frac{mpq\gamma}{L} - \frac{2m(m-1)p^2\tau}{L} - 2mp^2(\alpha+\beta) + \frac{m^2p^2\pi^2}{6L^2}$$
$$- m(m-1)p^2 \log_Q 2 + \frac{p^2m^2}{12} + \frac{mpq}{2} - p^2m^2[\delta^2]_0 + m\varpi(\log \tfrac{N}{m}) + O(N^{-1}).$$

Note that the leading term does not contain a fluctuation, since such terms also cancel out, but that was not stated properly in [8]. Likewise, the term including $\alpha+\beta$ was missed earlier.

The terms involving m^2 are

$$-\frac{2p^2\tau}{L} + \frac{p^2\pi^2}{6L^2} - p^2 \log_Q 2 + \frac{p^2}{12} - p^2[\delta^2]_0,$$

but this is zero, as discussed earlier.

Theorem 2. *Mean and variance of the m-version of the number of left-to-right maxima (in the strict sense) of geometrically distributed words satisfy*

$$\mathcal{E}_N = mp \log_Q N - mp \log_Q m + \frac{mp}{2} + \frac{mp\gamma}{L} - mp\delta(\log_Q \tfrac{N}{m}) + O(N^{-1})$$

and

$$\mathcal{V}_N = mpq \log_Q N - mpq \log_Q m + \frac{mpq\gamma}{L} + \frac{2mp^2\tau}{L} - 2mp^2(\alpha+\beta)$$
$$+ mp^2 \log_Q 2 + \frac{mpq}{2} + m\varpi(\log \tfrac{N}{m}) + O(N^{-1}).$$

The Fourier coefficients of $\varpi(x)$ are not computed, but they can be expressed with $\Gamma(-\chi_k)$ and $\Gamma'(-\chi_k)$.

4. Left-to-right maxima in the weak sense

This section is a variation of the previous one; this time an element is already a left-to-right maximum if it is larger *or equal* than all elements to the left of it. In [8] we find

$$E_n = \frac{p}{q}\sum_{k=1}^{n}\binom{n}{k}(-1)^{k-1}\frac{1}{Q^k-1},$$

so

$$\omega(k) = \frac{p}{q}\frac{-1}{Q^k-1}.$$

Further,

$$E_n^{(2)} = g_n + h_n$$

with

$$g_n = [z^n]\frac{2p^2}{q^2}\frac{z^2}{1-z}\sum_{1\le i<j}\frac{q^{i+j}}{(1-z(1-q^i))(1-z(1-q^j))}$$

and

$$h_n = [z^n]\frac{2p^2}{q^2}\frac{z^2}{1-z}\sum_{j\ge 1}\frac{q^{2j}}{(1-z(1-q^j))^2}.$$

A similar computation as in the previous section leads to

$$g_n = \frac{2p^2}{q^2}\sum_{k=1}^{n}\binom{n}{k}(-1)^k\left[\alpha\frac{1}{Q^k-1} - \frac{1}{(Q^k-1)^2} - \frac{1}{Q^k-1}\sum_{h\ge 1}\frac{1}{Q^{k+h}-1}\right].$$

For h_n we are able to compute a more concise representation than in [8]:

$$\begin{aligned}
h_n &= [z^n]\frac{2p^2}{q^2}\sum_{j\ge 1}\left[\frac{1}{1-z} - \frac{1}{Q^j-1}\frac{1}{(1-z(1-q^j))^2} + \frac{2-Q^j}{Q^j-1}\frac{1}{1-z(1-q^j)}\right]\\
&= \frac{2p^2}{q^2}\sum_{j\ge 1}\left[1 - \frac{1}{Q^j-1}(n+1)(1-q^j)^n + \frac{2-Q^j}{Q^j-1}(1-q^j)^n\right]\\
&= \frac{2p^2}{q^2}\sum_{j\ge 1}\left[1 - \frac{1}{Q^j-1}(n+1)\sum_{k=0}^{n}\binom{n}{k}(-1)^k q^{jk} + \frac{2-Q^j}{Q^j-1}\sum_{k=0}^{n}\binom{n}{k}(-1)^k q^{jk}\right]\\
&= \frac{2p^2}{q^2}\sum_{j\ge 1}\left[1 - \frac{1}{Q^j-1}\sum_{k=0}^{n}\binom{n+1}{k+1}(k+1)(-1)^k q^{jk} + \frac{2-Q^j}{Q^j-1}\sum_{k=0}^{n}\binom{n}{k}(-1)^k q^{jk}\right]\\
&= \frac{2p^2}{q^2}\sum_{j\ge 1}\left[1 - \frac{1}{Q^j-1}\sum_{k=0}^{n}\binom{n}{k+1}(k+1)(-1)^k q^{jk}\right.\\
&\qquad\left. - \frac{1}{Q^j-1}\sum_{k=0}^{n}\binom{n}{k}(k+1)(-1)^k q^{jk} + \frac{2-Q^j}{Q^j-1}\sum_{k=0}^{n}\binom{n}{k}(-1)^k q^{jk}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{2p^2}{q^2}\sum_{j\ge1}\bigg[1+\frac{Q^j}{Q^j-1}\sum_{k=1}^{n}\binom{n}{k}k(-1)^kq^{jk}\\
&\qquad\qquad -\frac{1}{Q^j-1}\sum_{k=0}^{n}\binom{n}{k}(k+1)(-1)^kq^{jk}+\frac{2-Q^j}{Q^j-1}\sum_{k=0}^{n}\binom{n}{k}(-1)^kq^{jk}\bigg]\\
&= \frac{2p^2}{q^2}\sum_{k=1}^{n}\binom{n}{k}(-1)^k\sum_{j\ge1}\bigg[\frac{Q^j}{Q^j-1}kq^{jk}-\frac{1}{Q^j-1}(k+1)q^{jk}+\frac{2-Q^j}{Q^j-1}q^{jk}\bigg]\\
&= \frac{2p^2}{q^2}\sum_{k=1}^{n}\binom{n}{k}(-1)^k\sum_{j\ge1}(k-1)q^{jk}\\
&= \frac{2p^2}{q^2}\sum_{k=1}^{n}\binom{n}{k}(-1)^k\frac{k-1}{Q^k-1}.
\end{aligned}
$$

Therefore

$$
\theta(k)=\frac{2p^2}{q^2}\bigg[\alpha\frac{1}{Q^k-1}-\frac{1}{(Q^k-1)^2}-\frac{1}{Q^k-1}\sum_{h\ge1}\frac{1}{Q^{k+h}-1}+\frac{k-1}{Q^k-1}\bigg].
$$

We also need this formula

$$
\sum_{j=1}^{k-1}\binom{k}{j}\frac{1}{Q^j-1}\frac{1}{Q^{k-j}-1}=\frac{2}{Q^k-1}\sum_{j\ge1}\binom{k}{j}\frac{1}{Q^j-1}-\frac{2}{(Q^k-1)^2}+\frac{2^k-2}{Q^k-1},
$$

which is again very similar to a formula from the previous section.

So, for the expected value $\mathcal{E}_N$, we must compute the residue at $z=0$ (and at $z=\chi_k$) of

$$
\frac{\Gamma(N+1)\Gamma(-z)}{\Gamma(N+1-z)}m^{1-z}\frac{p}{q}\frac{-1}{Q^z-1},
$$

which leads to

$$
\mathcal{E}_N\sim\frac{mp}{q}\log_Q N-\frac{mp}{q}\log_Q m+\frac{mp\gamma}{qL}-\frac{mp}{2q}-\frac{mp}{q}\delta(\log_Q \tfrac{N}{m}).
$$

For the asymptotic equivalent of $\mathcal{E}_N^{(2)}$, we must compute the residue at $z=0$ of

$$
\begin{aligned}
&\frac{\Gamma(N+1)\Gamma(-z)}{\Gamma(N+1-z)}m^{1-z}\frac{2p^2}{q^2}\bigg[\alpha\frac{1}{Q^z-1}-\frac{1}{(Q^z-1)^2}-\frac{1}{Q^z-1}\sum_{h\ge1}\frac{1}{Q^{z+h}-1}+\frac{z-1}{Q^z-1}\bigg]\\
&+\frac{\Gamma(N+1)\Gamma(-z)}{\Gamma(N+1-z)}m^{1-z}(m-1)\frac{p^2}{q^2}\bigg[\frac{2}{Q^z-1}\sum_{j\ge1}\binom{z}{j}\frac{1}{Q^j-1}-\frac{2}{(Q^z-1)^2}+\frac{2^z-2}{Q^z-1}\bigg].
\end{aligned}
$$

Again we only display the variance which is much shorter, thanks to cancellations:

$$
\frac{mp}{q^2}\log_Q N-\frac{mp}{q^2}\log_Q m+\frac{mp\gamma}{Lq^2}-\frac{2m(m-1)p^2\tau}{q^2L}-\frac{2mp^2}{q^2}(\alpha+\beta)+\frac{m^2p^2\pi^2}{6q^2L^2}
$$

$$-\frac{m(m-1)p^2}{q^2}\log_Q 2+\frac{p^2m^2}{12q^2}-\frac{mp}{2q^2}-\frac{p^2m^2}{q^2}[\delta^2]_0+m\upsilon(\log\tfrac{N}{m}).$$

And again, there is simplification, since the term of order m^2 cancels:

$$\frac{mp}{q^2}\log_Q N-\frac{mp}{q^2}\log_Q m+\frac{mp\gamma}{Lq^2}+\frac{2mp^2\tau}{q^2L}-\frac{2mp^2}{q^2}(\alpha+\beta)+\frac{mp^2}{q^2}\log_Q 2-\frac{mp}{2q^2}+m\upsilon(\log\tfrac{N}{m}).$$

We collect the results.

Theorem 3. *Mean and variance of the m-version of the number of left-to-right maxima (in the weak sense) of geometrically distributed words satisfy*

$$\mathcal{E}_N=\frac{mp}{q}\log_Q N-\frac{mp}{q}\log_Q m+\frac{mp\gamma}{qL}-\frac{mp}{2q}-\frac{mp}{q}\delta(\log_Q\tfrac{N}{m})+O(N^{-1})$$

and

$$\mathcal{V}_N=\frac{mp}{q^2}\log_Q N-\frac{mp}{q^2}\log_Q m+\frac{mp\gamma}{Lq^2}+\frac{2mp^2\tau}{q^2L}-\frac{2mp^2}{q^2}(\alpha+\beta)\\+\frac{mp^2}{q^2}\log_Q 2-\frac{mp}{2q^2}+m\upsilon(\log\tfrac{N}{m})+O(N^{-1}).$$

The Fourier coefficients of $\upsilon(x)$ are not computed, but they can be expressed with $\Gamma(-\chi_k)$ and $\Gamma'(-\chi_k)$.

5. Conclusion

It is likely that alternative approaches, as in [4] (based on [5]) or [7] will also work here.

References

[1] J. Cichoń and W. Macyna. Approximate counters for flash memory. *17th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA), Toyama, Japan*, 2011.

[2] P. Flajolet. Approximate counting: a detailed analysis. *BIT*, 25:113–134, 1985.

[3] P. Flajolet and R. Sedgewick. Mellin transforms and asymptotics: Finite differences and Rice's integrals. *Theoretical Computer Science*, 144:101–124, 1995.

[4] M. Fuchs, C.-K. Lee, and H. Prodinger. Approximate counting via the Poisson-Laplace-Mellin method. *DMTCS, proceedings AofA12*, AQ:13–28, 2012.

[5] H.-K. Hwang, M. Fuchs, and V. Zacharovas. Asymptotic variance of random symmetric digital search trees. *Discrete Mathematics and Theoretical Computer Science*, 12:103–166, 2010.

[6] P. Kirschenhofer and H. Prodinger. A result in order statistics related to probabilistic counting. *Computing*, 51:15–27, 1993.

[7] G. Louchard and H. Prodinger. Approximate counting with m counters: a probabilistic analysis. *submitted*, xx:xxx–xxx, 2012.

[8] H. Prodinger. Combinatorics of geometrically distributed random variables: Left-to-right maxima. *Discrete Mathematics*, 153:253–270, 1996.

[9] H. Prodinger. Periodic oscillations in the analysis of algorithms—and their cancellations. *Journal of the Iranian Statistical Society*, 3:251–270, 2004.

[10] H. Prodinger. Approximate counting with m counters: a detailed analysis. *Theoretical Computer Science*, 439:58–68, 2012.

[11] W. Szpankowski and V. Rego. Yet another application of a binomial recurrence. Order statistics. *Computing*, 43(4):401–410, 1990.

Helmut Prodinger, Department of Mathematics, University of Stellenbosch, 7602 Stellenbosch, South Africa
E-mail address: hproding@sun.ac.za