

Consecutive records in geometrically distributed words

Kamilla Oliver · Helmut Prodinger

Received: 1 September 2010 / Accepted: 19 April 2011 / Published online: 13 May 2011
© African Mathematical Union and Springer-Verlag 2011

Abstract Words $a_1a_2 \dots a_n$ with independent letters a_k taken from the set of natural numbers, and a weight (probability) attached via the geometric distribution pq^{i-1} ($p + q = 1$) are considered. The parameter $\mathcal{K}(a_1a_2 \dots a_n)$, (the number of weak consecutive records), has proved to be essential in the analysis of a skip list structure. Related to it is the (new) parameter \mathcal{M} , i.e., the largest consecutive record in a random word of length n . Exact and asymptotic formulæ are derived for the expectation and the variance.

Keywords Alternating sum · Heine's transformation · q -series · Asymptotics · Rice's method

Mathematics Subject Classification (2010) 60C05 · 05A16

1 Introduction

We consider words $a_1a_2 \dots a_n$ with letters a_k taken from the set of natural numbers, and a weight (probability) attached to it by saying that the letter $i \in \mathbb{N}$ occurs with probability pq^{i-1} ($p + q = 1$) and that the letters are independent. The parameter $\mathcal{K}(a_1a_2 \dots a_n)$, which we call the number of *weak consecutive records*, has proved to be essential in the analysis of

K. Oliver
91052 Erlangen, Germany
e-mail: olikamilla@gmail.com

H. Prodinger (✉)
Department of Mathematics, University of Stellenbosch,
Stellenbosch 7602, South Africa
e-mail: hproding@sun.ac.za

a skip list structure [10]. The word is scanned from left to right, and assuming that the current record (maximum) is value k , any letter different from $k, k + 1$ is ignored. If, however, the symbol scanned is one of these, we call it a weak consecutive record, and set the value of the current maximum to it. So, the current record either stays at k or advances to the next value $k + 1$. The skip list version assumes that the first letter of the word defines the first record.

For the sake of clarity, we consider the word 1 3 1 1 2 4 3 5 3 5 1 4 1 2 3 4 6 5 1 and underline each consecutive maximum: 1 3 1 1 2 4 3 5 3 5 1 4 2 1 3 4 6 5 1. The number of underlined symbols (9 in this case) is the parameter \mathcal{K} of interest.

In [10], the average of the parameter $\mathcal{K}(n)$ was shown to be (with $Q = 1/q$)

$$\mathbb{E}\mathcal{K}(n) = 1 + (Q + 1) \sum_{j=1}^n \binom{n}{j} \frac{(-1)^{j-1} (q)_{j-1} p^{j+1} q^j}{1 - q^{j+1}},$$

which was also evaluated asymptotically.

Theorem 1 (Old theorem) *The expectation of the $\mathcal{K}(n)$ -parameter is asymptotic to*

$$\mathbb{E}\mathcal{K}(n) \sim (Q + 1) \log_Q n + \frac{(Q + 1)\gamma}{L} + \frac{Q + 1}{L} \log(p) - (Q + 1)\alpha - \frac{(1 + q)^2}{2pq} + 1 + \delta(\log_Q n).$$

The constant α is given by

$$\alpha = \sum_{i \geq 1} \frac{q^i}{1 - q^i};$$

$\delta(x)$ is a small periodic function. Its Fourier coefficients could be given in principle.

In this paper we investigate the parameter \mathcal{M} , which is the *maximum of the underlined values*. Now, clearly, for that, we do not need to underline repetitions of the current maximum, as in the instance of the \mathcal{K} -parameter. So, when our current maximum is $k - 1$, we ignore all letters different from k , and when it occurs (with probability pq^{k-1}) we set the current maximum to k . We will find explicit and asymptotic expressions for average and variance, assuming random words of length n .

This parameter is related to other ones that appear in the literature:

- If the current maximum is updated (to new value j) whenever a larger element j (than the current k) occurs (with probability pq^{j-1}), then the resulting parameter is simply the maximum of the word, and this parameter is well understood [8, 16].
- If the local counter is updated from k to $k + 1$ with probability q^k , and the process starts with a counter value 1 before any letter is read, then this is called *approximate counting*, and it is also very well understood [3, 5, 7, 11–13].

This contribution belongs to the area called *combinatorics of geometrically distributed words* which was started with [14]; this area attracted also other people’s attention, notably the team lead by A. Knopfmacher in Johannesburg contributed to it, see e.g. [2].

We use (standard) notation from q -analysis: $(x)_n = (x; q)_n = \prod_{i=0}^{n-1} (1 - xq^i)$ and $(x)_\infty = \prod_{i \geq 0} (1 - xq^i)$. Note that $(x)_n = (x)_\infty / (xq^n)_\infty$, and the latter form makes sense also for n a complex number.

Furthermore, we use $Q = 1/q$ and $L = \log Q$.

2 Analysis of the \mathcal{M} -parameter

Let $b_k(z)$ be the generating function, such that the coefficient of z^n is the probability that a random word of length n has parameter \mathcal{M} equal to k . We derive a recursion:

$$b_k(z) = b_{k-1}(z) \frac{zpq^{k-1}}{1 - z[1 - pq^k]} + \frac{zpq^{k-1}}{1 - z[1 - pq^k]}.$$

It holds for $k \geq 1$, when we assume that $b_0 = 0$. It is easy to understand: A new consecutive record k is attached to an existing word, followed by letters different from $k + 1$; thus, all these letters are no consecutive records. The second term of the right-hand side corresponds to the situation that there was no previous record, i.e., the word starts with letter k .

We simplify:

$$b_k(z)[1 - z(1 - pq^k)] = b_{k-1}(z)zpq^{k-1} + zpq^{k-1}.$$

With the useful substitution $z = w/(w - 1)$ (“Euler transform” [6]), this reads as

$$b_k(1 - wpq^k) = -b_{k-1}wpq^{k-1} - wpq^{k-1},$$

where we write $b_k := b_k(z)$ for convenience, or

$$\frac{b_k(wpq; q)_k (-1)^k}{w^k p^k q^{\binom{k}{2}}} = \frac{b_{k-1}(wpq; q)_{k-1} (-1)^{k-1}}{w^{k-1} p^{k-1} q^{\binom{k-1}{2}}} - \frac{(wpq; q)_{k-1} (-1)^k}{w^{k-1} p^{k-1} q^{\binom{k-1}{2}}}.$$

This can now be summed, and we get an explicit formula for b_k that we write as a lemma, for further reference.

Lemma 2 For all $k \geq 1$,

$$b_k = \sum_{j=0}^{k-1} \frac{q^{\binom{k}{2} - \binom{j}{2}} (-pw)^{k-j}}{(wpq^{j+1}; q)_{k-j}}.$$

□

We notice that

$$\sum_{k \geq 1} \frac{q^{\binom{k}{2}} (-pq^j w)^k}{(wpq^{j+1}; q)_k} = -pq^j w;$$

the corresponding computation appears for instance in [13], and it uses Heine’s transformation formula. Therefore

$$\sum_{k \geq 1} b_k = \sum_{j \geq 0} \sum_{k \geq 1} \frac{q^{\binom{k}{2}} (-pq^j w)^k}{(wpq^{j+1}; q)_k} = - \sum_{j \geq 0} pq^j w = -w = \frac{z}{1 - z},$$

which is combinatorially clear, since it simply describes all words.

Now, we compute the average of the \mathcal{M} -parameter:

$$\begin{aligned} \sum_{k \geq 1} kb_k &= \sum_{k \geq 1} k \sum_{j=0}^{k-1} \frac{q^{\binom{k}{2} - \binom{j}{2}} (-pw)^{k-j}}{(wpq^{j+1}; q)_{k-j}} \\ &= \sum_{j \geq 0} \sum_{k \geq 1} (k+j) \frac{q^{\binom{k}{2}} (-pq^j w)^k}{(wpq^{j+1}; q)_k} \\ &= \sum_{j \geq 0} \sum_{k \geq 1} k \frac{q^{\binom{k}{2}} (-pq^j w)^k}{(wpq^{j+1}; q)_k} - \sum_{j \geq 0} j p q^j w \\ &= \sum_{j \geq 0} \frac{d}{dt} \sum_{k \geq 0} \frac{q^{\binom{k}{2}} (-pq^j wt)^k}{(wpq^{j+1}; q)_k} \Big|_{t=1} - \frac{q}{p} w. \end{aligned}$$

The inner sum we will attack with Heine’s transform: (see [1])

$$\sum_{n \geq 0} \frac{(a)_n (b)_n t^n}{(q)_n (c)_n} = \frac{(b)_\infty (at)_\infty}{(c)_\infty (t)_\infty} \sum_{n \geq 0} \frac{(c/b)_n (t)_n b^n}{(q)_n (at)_n}.$$

We replace $q^{\binom{k}{2}} (-1)^k$ by $\lim_{\varepsilon \rightarrow 0} \varepsilon^k (1/\varepsilon)_k$:

$$\begin{aligned} \sum_{k \geq 0} \frac{q^{\binom{k}{2}} (-pq^j wt)^k}{(wpq^{j+1}; q)_k} &= \lim_{\varepsilon \rightarrow 0} \sum_{k \geq 0} \frac{(1/\varepsilon)_k (q)_k (pq^j wt \varepsilon)^k}{(wpq^{j+1}; q)_k (q)_k} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{(q)_\infty (pq^j wt)_\infty}{(wpq^{j+1})_\infty (pq^j wt \varepsilon)_\infty} \sum_{k \geq 0} \frac{(wpq^j)_k (pq^j wt \varepsilon)_k q^k}{(q)_k (pq^j wt)_k} \\ &= \frac{(q)_\infty (pq^j wt)_\infty}{(wpq^{j+1})_\infty} \sum_{k \geq 0} \frac{(0)_k (wpq^j)_k q^k}{(q)_k (pq^j wt)_k} \\ &= \frac{(q)_\infty (pq^j wt)_\infty}{(wpq^{j+1})_\infty} \frac{(wpq^j)_\infty}{(pq^j wt)_\infty (q)_\infty} \sum_{k \geq 0} \frac{(t)_k (q)_k (wpq^j)^k}{(q)_k} \\ &= (1 - wpq^j) \sum_{k \geq 0} (t)_k (wpq^j)^k. \end{aligned}$$

Now the differentiation w.r.t. t , followed by $t = 1$ simplifies this very much:

$$-(1 - wpq^j) \sum_{k \geq 1} (q)_{k-1} (wpq^j)^k.$$

We are at

$$\begin{aligned} \sum_{k \geq 1} kb_k &= - \sum_{j \geq 0} (1 - wpq^j) \sum_{k \geq 0} (q)_k (wpq^j)^{k+1} - \frac{q}{p} w \\ &= - \sum_{j \geq 0} \sum_{k \geq 0} (q)_k (wpq^j)^{k+1} + \sum_{j \geq 0} \sum_{k \geq 0} (q)_k (wpq^j)^{k+2} - \frac{q}{p} w \\ &= - \sum_{k \geq 0} \frac{(q)_k (wp)^{k+1}}{1 - q^{k+1}} + \sum_{k \geq 1} \frac{(q)_{k-1} (wp)^{k+1}}{1 - q^{k+1}} - \frac{q}{p} w \end{aligned}$$

$$\begin{aligned}
 &= -\sum_{k \geq 1} \frac{(q)_k (wp)^{k+1}}{1 - q^{k+1}} + \sum_{k \geq 1} \frac{(q)_{k-1} (wp)^{k+1}}{1 - q^{k+1}} - \frac{w}{p} \\
 &= \sum_{k \geq 1} \frac{(q)_{k-1} q^k (wp)^{k+1}}{1 - q^{k+1}} - \frac{w}{p}.
 \end{aligned}$$

We note that for $j \geq 1$

$$[w^{j+1}] \sum_{k \geq 1} kb_k = \frac{(q)_{j-1} q^j p^{j+1}}{1 - q^{j+1}}$$

and

$$[w^1] \sum_{k \geq 1} kb_k = -\frac{1}{p}.$$

Therefore (see [6] about how $[z^n]f(z)$ and $[w^n]f(w/(w - 1))$ are related)

$$\begin{aligned}
 [z^{n+1}] \sum_{k \geq 1} kb_k &= (-1)^{n+1} [w^{n+1}] (1 - w)^n \sum_{k \geq 1} kb_k \\
 &= (-1)^{n+1} \sum_{j=0}^n \binom{n}{n-j} (-1)^{n-j} [w^{j+1}] \sum_{k \geq 1} kb_k \\
 &= \sum_{j=0}^n \binom{n}{j} (-1)^{j+1} [w^{j+1}] \sum_{k \geq 1} kb_k \\
 &= \frac{1}{p} + \sum_{j=1}^n \binom{n}{j} (-1)^{j+1} \frac{(q)_{j-1} q^j p^{j+1}}{1 - q^{j+1}}.
 \end{aligned}$$

This expression possesses an integral representation (Rice’s method [4]):

$$[z^{n+1}] \sum_{k \geq 1} kb_k = \frac{1}{p} + \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{(-1)^{n+1} n!}{z(z-1) \dots (z-n)} \frac{(q)_{z-1} q^z p^{z+1}}{1 - q^{z+1}} dz.$$

The curve \mathcal{C} encircles the poles $1, 2, \dots, n$ and no others. To get asymptotics, we must collect the (negative) residues at 0 and at $z = \chi_k = \frac{2\pi i k}{L}$. We need the local expansion of the integrand around $z = 0$. We notice that $(q)_{z-1} = (q)_\infty / (q^z)_\infty$ and $(q^z)_\infty = (1 - q^z)(q^{z+1})_\infty$ and

$$(q^{z+1})_\infty = \prod_{k \geq 1} (1 - q^{z+k}) \sim (q)_\infty + (q)_\infty \sum_{k \geq 1} \frac{Lq^k}{1 - q^k} \cdot z.$$

Providing this, the residue at $z = 0$ can be computed by a computer. For the residue at $z = \chi_k$, we notice that $(q^{\chi_k+1})_\infty = (q)_\infty$, and

$$\left. \frac{(-1)^{n+1} n!}{z(z-1) \dots (z-n)} \right|_{z=\chi_k} = \frac{n! \Gamma(-\chi_k)}{\Gamma(n+1-\chi_k)} \sim n^{\chi_k} \Gamma(-\chi_k).$$

Therefore the negative residue at $z = \chi_k$ is

$$-n^{\chi_k} \Gamma(-\chi_k) \frac{1}{L} p^{\chi_k}.$$

Summarizing, we have asymptotically evaluated the average of the \mathcal{M} -parameter:

Theorem 3 *The average of the \mathcal{M} -parameter is asymptotically given by*

$$\mathbb{E}(\mathcal{M}) \sim \log_Q n - \alpha + \frac{\log(p)}{L} + \frac{\gamma}{L} + \frac{1}{2} + \delta(\log_Q(pn)),$$

where the periodic function $\delta(x)$ is given as

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

To obtain the second (factorial) moment, we compute

$$\begin{aligned} \sum_{k \geq 2} k(k-1)b_k &= \sum_{k \geq 1} k(k-1) \sum_{j=0}^{k-1} \frac{q^{\binom{k}{2} - \binom{j}{2}} (-pw)^{k-j}}{(wpq^{j+1}; q)_{k-j}} \\ &= \sum_{j \geq 0} \sum_{k \geq 1} (k+j)(k+j-1) \frac{q^{\binom{k}{2}} (-pq^j w)^k}{(wpq^{j+1}; q)_k} \\ &= \sum_{j \geq 0} \sum_{k \geq 1} k(k-1) \frac{q^{\binom{k}{2}} (-pq^j w)^k}{(wpq^{j+1}; q)_k} + 2 \sum_{j \geq 0} j \sum_{k \geq 1} k \frac{q^{\binom{k}{2}} (-pq^j w)^k}{(wpq^{j+1}; q)_k} - \frac{2q^2}{p^2} w. \end{aligned}$$

We need

$$\begin{aligned} \sum_{j \geq 0} j \sum_{k \geq 1} k \frac{q^{\binom{k}{2}} (-pq^j w)^k}{(wpq^{j+1}; q)_k} &= - \sum_{j \geq 0} j(1 - wpq^j) \sum_{k \geq 1} (q)_{k-1} (wpq^j)^k \\ &= - \sum_{j \geq 0} j \sum_{k \geq 1} (q)_{k-1} (wpq^j)^k + \sum_{j \geq 0} j \sum_{k \geq 1} (q)_{k-1} (wpq^j)^{k+1} \\ &= - \sum_{k \geq 1} \frac{(q)_{k-1} (wp)^k q^k}{(1 - q^k)^2} + \sum_{k \geq 1} \frac{(q)_{k-1} (wp)^{k+1} q^{k+1}}{(1 - q^{k+1})^2} \\ &= - \frac{wq}{p} + \sum_{k \geq 1} \frac{(q)_{k-1} (wp)^{k+1} q^{2k+1}}{(1 - q^{k+1})^2}. \end{aligned}$$

Thus far we are at

$$\sum_{k \geq 2} k(k-1)b_k = \sum_{j \geq 0} \sum_{k \geq 1} k(k-1) \frac{q^{\binom{k}{2}} (-pq^j w)^k}{(wpq^{j+1}; q)_k} + 2 \sum_{k \geq 1} \frac{(q)_{k-1} (wp)^{k+1} q^{2k+1}}{(1 - q^{k+1})^2} - \frac{2q}{p^2} w.$$

We compute

$$\begin{aligned} \sum_{j \geq 0} \sum_{k \geq 1} k(k-1) \frac{q^{\binom{k}{2}} (-pq^j w)^k}{(wpq^{j+1}; q)_k} &= \sum_{j \geq 0} \frac{d}{dt^2} (1 - wpq^j) \sum_{k \geq 0} (t)_k (wpq^j)^k \Big|_{t=1} \\ &= 2 \sum_{j \geq 0} (1 - wpq^j) \sum_{k \geq 2} (q)_{k-1} T(k-1) (wpq^j)^k \end{aligned}$$

with

$$T(k) = \sum_{i=1}^k \frac{q^i}{1 - q^i}.$$

We simplify

$$\begin{aligned} \sum_{j \geq 0} \sum_{k \geq 1} k(k-1) \frac{q^{\binom{k}{2}} (-pq^j w)^k}{(wpq^{j+1}; q)_k} \\ = 2 \sum_{k \geq 1} \frac{(q)_k T(k) (wp)^{k+1}}{1 - q^{k+1}} - 2 \sum_{k \geq 1} \frac{(q)_{k-1} T(k-1) (wp)^{k+1}}{1 - q^{k+1}} \\ = 2 \sum_{k \geq 1} \frac{(q)_{k-1} [1 - T(k-1)] (wp)^{k+1} q^k}{1 - q^{k+1}}. \end{aligned}$$

So

$$\sum_{k \geq 2} k(k-1) b_k = 2 \sum_{k \geq 1} \frac{(q)_{k-1} [1 - T(k-1)] (wp)^{k+1} q^k}{1 - q^{k+1}} + 2 \sum_{k \geq 1} \frac{(q)_{k-1} (wp)^{k+1} q^{2k+1}}{(1 - q^{k+1})^2} - \frac{2q}{p^2} w.$$

and

$$[w^{j+1}] \sum_{k \geq 2} k(k-1) b_k = 2 \frac{(q)_{j-1} [1 - T(j-1)] p^{j+1} q^j}{1 - q^{j+1}} + 2 \frac{(q)_{j-1} p^{j+1} q^{2j+1}}{(1 - q^{j+1})^2}$$

for $j \geq 1$, and $\frac{-2q}{p^2}$ for $j = 0$. Therefore

$$\begin{aligned} [z^{n+1}] \sum_{k \geq 1} k(k-1) b_k &= (-1)^{n+1} [w^{n+1}] (1-w)^n \sum_{k \geq 1} k(k-1) b_k \\ &= (-1)^{n+1} \sum_{j=0}^n \binom{n}{n-j} (-1)^{n-j} [w^{j+1}] \sum_{k \geq 1} k(k-1) b_k \\ &= \sum_{j=0}^n \binom{n}{j} (-1)^{j+1} [w^{j+1}] \sum_{k \geq 1} k(k-1) b_k \\ &= \frac{2q}{p^2} + 2 \sum_{j=1}^n \binom{n}{j} (-1)^{j+1} (q)_{j-1} p^{j+1} q^j \left[\frac{-T(j-1)}{1 - q^{j+1}} + \frac{1}{(1 - q^{j+1})^2} \right]. \end{aligned}$$

Again, there is an integral representation for it:

$$\begin{aligned} [z^{n+1}] \sum_{k \geq 1} k(k-1) b_k \\ = \frac{2q}{p^2} + \frac{2}{2\pi i} \int_{\mathcal{C}} \frac{(-1)^{n+1}}{z(z-1) \dots (z-n)} (q)_{z-1} p^{z+1} q^z \left[\frac{T(z-1)}{1 - q^{z+1}} - \frac{1}{(1 - q^{z+1})^2} \right] dz. \end{aligned}$$

Note that

$$T(k) = \sum_{j=1}^k \frac{q^j}{1 - q^j} = \alpha - \sum_{j>k} \frac{q^j}{1 - q^j} = \alpha - \sum_{j \geq 1} \frac{q^{j+k}}{1 - q^{j+k}},$$

whence we can take

$$T(z) = \alpha - \sum_{j \geq 1} \frac{q^{j+z}}{1 - q^{j+z}}.$$

We expect a triple pole at $z = 0$, and so we need to expand $T(z - 1)$:

$$\begin{aligned} T(z - 1) &= T(z) - \frac{q^z}{1 - q^z} = \alpha - \sum_{j \geq 1} \frac{q^{j+z}}{1 - q^{j+z}} - \frac{q^z}{1 - q^z} \\ &\sim \beta Lz - \frac{1}{Lz} + \frac{1}{2} - \frac{Lz}{12}, \end{aligned}$$

with

$$\beta = \sum_{j \geq 1} \frac{q^j}{(1 - q^j)^2}.$$

Also note the expansion

$$(q)_z = \frac{(q)_\infty}{(q^{z+1})_\infty} \sim 1 - \alpha Lz + \frac{\alpha^2 + \beta}{2} L^2 z^2.$$

Reading off the negative residue at $z = 0$ gives

$$(\log_Q n)^2 + 2 \left(\frac{\gamma + \log(p)}{L} - \alpha \right) \log_Q n + \frac{\log^2(p) + \gamma^2 + \pi^2/6 + \log(p)}{L^2} - \frac{2\alpha(\log(p) + \gamma)}{L} + \alpha^2 - \beta - \frac{1}{6}.$$

To compute the variance, we add the expectation and subtract the square of the expectation and find (many simplifications occur!)

$$\frac{\pi^2}{6L^2} + \frac{1}{12} - \beta.$$

We have always computed the coefficient of z^{n+1} for convenience, but it does not matter for the asymptotic formula. We summarize:

Theorem 4 *The average and variance of the \mathcal{M} -parameter are asymptotically given by*

$$\begin{aligned} \mathbb{E}\mathcal{M}(n) &\sim \log_Q n - \alpha + \frac{\log(p)}{L} + \frac{\gamma}{L} + \frac{1}{2} + \delta_E(\log_Q n), \\ \mathbb{V}\mathcal{M}(n) &\sim \frac{\pi^2}{6L^2} + \frac{1}{12} - \beta + \delta_V(\log_Q n). \end{aligned}$$

Here, $\delta(x)$ is an unspecified periodic function of period 1 and small amplitude. Its Fourier coefficients could be computed in principle. The residues come from the poles at $z = \chi_k = \frac{2\pi ik}{L}$.

3 Conclusion

While the \mathcal{M} -parameter was challenging, there are other interesting ones. We mention the *sum of the positions of the underlined elements*. In our running example

$$\underline{1} \ 3 \ 1 \ 1 \ 2 \ 4 \ 3 \ 5 \ 3 \ 5 \ 1 \ 4 \ 2 \ 1 \ 3 \ 4 \ 6 \ 5 \ 1,$$

the underlined elements are in positions 1, 3, 4, 5, 7, 9, 12, 16, 18, therefore this parameter of the word is

$$1 + 3 + 4 + 5 + 7 + 9 + 12 + 16 + 18 = 75.$$

This parameter was introduced for ordinary records in permutations in [9] and adapted to words in [15]. For the reader's convenience, we repeat the results about the expectation.

Theorem 5 *The expected value of the sum of the positions of records, in random words of length n , is given by*

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

The expected value of the sum of the positions of records, in random words of length n , has the asymptotic expansion

$$E_n = \frac{pn}{\log Q} (1 + \delta(\log_Q n)) + O(1),$$

where the periodic function (of small amplitude) is given by

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

However, in the present context, the relevant recursions (that are not too hard to set up) seem to be quite involved. It might not be hopeless, but at that stage we decided not to go further.

Acknowledgments Thanks are due to Guy Louchard who spotted an error in an earlier version of this paper.

References

1. Andrews, G.: The Theory of Partitions. In: Encyclopedia of Mathematics and its Applications, vol. 2. Addison-Wesley, Reading (1976)
2. Archibald, M., Knopfmacher, A.: The average position of the d th maximum in a sample of geometric random variables. Stat. Probab. Lett. **79**(7), 864–872 (2009)
3. Charalambides, Ch.A.: Moments of a class of discrete q -distributions. J. Stat. Plan. Inference **135**, 64–76 (2005)
4. Flajolet, P., Sedgewick, R.: Mellin transforms and asymptotics: Finite differences and Rice's integrals. Theor. Comput. Sci. **144**, 101–124 (1995)
5. Flajolet, Ph.: Approximate counting: a detailed analysis. BIT **25**(1), 113–134 (1985)
6. Graham, R.L., Knuth, D.E., Patashnik, O.: Concrete Mathematics, 2nd edn. Addison-Wesley, Reading (1994)
7. Kirschenhofer, P., Prodinger, H.: Approximate counting: an alternative approach. RAIRO Theor. Inf. Appl. **25**, 43–48 (1991)
8. Kirschenhofer, P., Prodinger, H.: A result in order statistics related to probabilistic counting. Computing **51**, 15–27 (1993)
9. Kortchemski, I.: Asymptotic behavior of permutation records. J. Combin. Theory Ser. A (6):1154–1166 (2009)
10. Louchard, G., Prodinger, H.: Analysis of a new skip list variant. Discrete Math. Theor. Comput. Sci. Proc. AG:365–374 (2006)

11. Louchard, G., Prodinger, H.: Generalized approximate counting revisited. *Theor. Comput. Sci.* **391**, 109–125 (2008)
12. Prodinger, H.: Hypothetic analyses: approximate counting in the style of Knuth, path length in the style of Flajolet. *Theor. Comput. Sci.* **100**, 243–251 (1992)
13. Prodinger, H.: Approximate counting via Euler transform. *Math. Slovaca* **44**, 569–574 (1994)
14. Prodinger, H.: Combinatorics of geometrically distributed random variables: left-to-right maxima. *Discrete Math.* **153**, 253–270 (1996)
15. Prodinger, H.: Records in geometrically distributed words: sum of positions. *Appl. Anal. Discrete Math.* **2**, 234–240 (2008)
16. Szpankowski, W., Rego, V.: Yet another application of a binomial recurrence. *Order statistics. Computing* **43**(4), 401–410 (1990)