$$(q,\delta)-\text{Numeration Systems with Missing Digits}$$

Frédérique Bassino
Helmut Prodinger

July 25, 2002
(q, δ)-NUMERATION SYSTEMS WITH MISSING DIGITS

FRÉDÉRIQUE BASSINO† AND HELMUT PRODINGER†

ABSTRACT. We consider the (q, δ) numeration system, with basis q ≥ 2 and the set of digits \( \{ \delta, \delta + 1, \ldots, q + \delta - 1 \} \) where \(- (q - 1) \leq \delta \leq 0 \). We study properties of numbers where some digits do not occur. This is analogous to the Cantor set \( \{ 0.a_1 a_2 \cdots | a_i \in \{ 0, 2 \} \} \).

We compute an asymptotic equivalent of the nth moment of the “Cantor (q, D)-distribution” which can be described as the numbers \( 0.w_1 w_2 \ldots \) with \( w_i \in D \subseteq \{ \delta, \ldots, q + \delta - 1 \} \), and each such letter can occur with the same probability \( 1/\text{Card}D \).

Furthermore, we consider \( n \) random strings according to this distribution and the expected minimum of them. We find a recursion which we solve asymptotically.

1. Introduction

We consider the (q, δ) numeration system, with basis q ≥ 2 and the set of digits \( \{ \delta, \delta + 1, \ldots, q + \delta - 1 \} \) where \(- (q - 1) \leq \delta \leq 0 \). Every real number \( x \) has an essentially unique\(^2\) representation

\[
x = \sum_{k \leq n} a_k q^k
\]

with \( a_k \in \{ \delta, \delta + 1, \ldots, q + \delta - 1 \} \). In particular, we are interested in properties of numbers where some digits do not occur. This is analogous to the Cantor set, which can be described as

\[
\{ 0.a_1 a_2 \cdots | a_i \in \{ 0, 2 \} \}.
\]

The Cantor distribution with parameter \( \vartheta \), \( 0 < \vartheta \leq \frac{1}{2} \), was introduced in [11] by the random series

\[
\frac{\vartheta}{\vartheta} \sum_{i \geq 1} X_i \vartheta^i,
\]

where the \( X_i \) are independent with the distribution

\[
P \{ X_i = 0 \} = P \{ X_i = 1 \} = \frac{1}{2},
\]

Date: August 9, 2002.

1991 Mathematics Subject Classification. Primary: 11A63, 11K16; Secondary: 60C05.

Key words and phrases. Digits, Depoissonization, Mellin Transform, Asymptotic Expansion, Moments.

†This author was supported by the CNRS/NRF-project no.10959.

Part of this work was done during the first author’s visit to the John Knopfmacher Centre for Applicable Analysis and Number Theory at the University of the Witwatersrand, Johannesburg, South Africa.

\(^1\)Often, the letter \( d \) is used instead of \( \delta \). In this paper, however, we need the letter \( d \) for other purposes.

\(^2\)Some numbers have two representations, which is the analogue of \( 1 = 0.999\ldots \).
and $\bar{\vartheta} = 1 - \vartheta$. The name stems from the instance $\vartheta = \frac{1}{3}$, since then exactly those numbers from the interval $[0,1]$ appear that have a ternary expansion solely consisting of the digits 0 and 2.

The moments of this distribution where considered in [11], and, more recently in [5], where an asymptotic formula for the $n$th moment was derived using a combination of analytic techniques, notably depoissonization ("de-Poissonization") and Mellin transforms.

In the first part of the present paper we derive analogous results: Let $D$ be a (given) subset of the set of digits $\{\delta, \delta + 1, \ldots, q + \delta - 1\}$; we set $d = \text{Card}D$ and $D = \{d_1 < d_2 < \ldots < d_d\}$.

We consider infinite (random) words $w_1 w_2 \ldots$ over the alphabet $D = \{d_1, d_2, \ldots, d_d\}$ and a mapping value, defined by

$$\text{value}(w_1 w_2 \ldots) = \sum_{i \geq 1} w_i q^{-i}.$$ 

Each letter can appear with probability $\frac{1}{q}$.

In this way we obtain a probability distribution on the interval $[\delta/(q-1), \delta/(q-1)+1]$, which we will call the Cantor--$(q,D)$ distribution. In Section 2 we study its moments.

Another interesting topic related to the Cantor distribution was introduced in [6]; one assumes that $n$ (independent) elements are drawn according to the Cantor distribution. One is interested in the expected value of the minimum of them. Hosking gave a recursion for these expectations, which was eventually solved in [9], both exactly and asymptotically. For the exact solution (involving Bernoulli numbers) a neat trick of Knuth’s was essential; for the asymptotics one could then rely on Rice’s method [3].

In Section 4 we are going to solve the analogous question in our model of the $(q,\delta)$-system with allowed set of digits $D$. Unfortunately, the nice trick does no longer work in this more general case, and we thus have to use the technique of depoissonization; for more details about this technique, one can refer to [7] and [13]; the present approach is modelled after the analysis in [8], which is also covered in [7] and [13].

2. The Moments

Observe the recursion formula, valid for all $i \in \{1, \ldots, d\}$

$$\text{value}(d_i w) = d_i \cdot q^{-1} + q^{-1} \cdot \text{value}(w). \quad (2.1)$$

Here, $dw$ is the concatenation of the digit $d$ and the infinite string $w$. We will start with a finite model: The set of (finite) non-empty words $D^+$ over the alphabet $D$ is given by

$$D^+ = D + DD^+.$$ 

It is also useful to have $D(z)$, the generating function of the words of $D^+$, according to their lengths:

$$D(z) = \frac{dz}{1 - dz} = \sum_{m \geq 1} d^m z^m.$$
Now we want to analyze the $n$th moments of the Cantor–$(q, D)$ distribution. We will work with finite words (of length $m$), and then letting the length $m$ tend to infinity. We do this by considering the generating functions

$$G_n(z) := \sum_{w \in D^+} \left( \text{value}(w) \right)^n z^{|w|};$$

where $|w|$ denotes the length of the word $w$. The quantity

$$\frac{[z^m]G_n(z)}{[z^m]D(z)}$$

is then the $n$th moment, when considering words of length $m$; finally we consider the limit of this for $m \to \infty$ and call it $M_n$.

Using the recursion $D^+ = D + DD^+$, we find for $n \geq 1$

$$G_n(z) = \sum_{i=1}^{d} d_i^n q^{-n} z + \sum_{i=1}^{d} \sum_{w \in D^+} \left( d_i q^{-1} + q^{-1} \text{value}(w) \right)^n z^{|w|+1}$$

$$= \sum_{i=1}^{d} d_i^n q^{-n} z + z q^{-n} \sum_{i=1}^{d} \sum_{j=0}^{n} \binom{n}{j} d_i^{n-j} G_j(z).$$

This recursion can be made explicit by rewriting it as

$$G_n(z) = \frac{z q^{-n}}{1 - d q^{-n} z} \sum_{i=1}^{d} \left[ d_i^n + \sum_{j=0}^{n-1} \binom{n}{j} d_i^{n-j} G_j(z) \right].$$

Note the special instance $G_0(z) = D(z)$ and

$$G_1(z) = \frac{q^{-1} z}{1 - d q^{-1} z} \left( 1 + D(z) \right) \sum_{i=1}^{d} d_i = \frac{q^{-1} z}{(1 - d q^{-1} z)(1 - dz)} \sum_{i=1}^{d} d_i.$$

From this we could get (by partial fraction decomposition) the coefficient of $z^m$ explicitly.

However, since we only consider the limit for $m \to \infty$, life is easier. Both, $G_1(z)$ and $D(z)$ have the dominant singularity at $z = 1/d$, and it is a simple pole. Consequently,

$$[z^m]G_1(z) \sim A d^m, \quad \text{with} \quad A = \lim_{z \to 1/d} (1 - zd) G_1(z)$$

and

$$[z^m]D(z) \sim B d^m, \quad \text{with} \quad B = \lim_{z \to 1/d} (1 - zd) D(z).$$

Therefore

$$\lim_{m \to \infty} \frac{[z^m]G_1(z)}{[z^m]D(z)} = \frac{A}{B},$$

which we can compute as

$$\lim_{z \to 1/d} \frac{q^{-1} z \sum_{i=1}^{d} d_i}{dz(1 - d q^{-1} z)} = \frac{1}{d(q - 1)} \sum_{i=1}^{d} d_i.$$
The fact that \( z = 1/d \) is a simple pole remains true for all the functions \( G_n(z) \). We can thus divide \( G_n(z) \) by \( D(z) \), cancel the pole, and insert \( z = 1/d \) in the rest, to obtain the \( n \)th moment \( M_n \) as

\[
M_n = \frac{d^{-1}q^{-n}}{1-q^{-n}} \sum_{i=1}^{d} \sum_{j=0}^{n-1} d_i^{n-j} \binom{n}{j} M_j.
\]

**Theorem 1.** The moments of the Cantor-\((q,D)\) distribution satisfy the following recursion: \( M_0 = 1 \) and for \( n \geq 1 \)

\[
M_n = \frac{1}{d(q^n - 1)} \sum_{j=0}^{n-1} \sum_{i=1}^{d} d_i^{n-j} \binom{n}{j} M_j.
\]

For instance

\[
M_1 = \frac{1}{d(q-1)} \sum_{i=1}^{d} d_i,
\]

\[
M_2 = \frac{1}{d(q^2-1)} \sum_{i=1}^{d} d_i^2 + \frac{1}{d^2(q-1)^2(q+1)} \left( \sum_{i=1}^{d} d_i \right)^2,
\]

\( (2.3) \)

\[
\text{Variance} = M_2 - M_1^2 = \frac{1}{d(q^2-1)} \sum_{i=1}^{d} d_i^2 - \frac{q}{d^2(q-1)^2(q+1)} \left( \sum_{i=1}^{d} d_i \right)^2.
\]

3. **The Asymptotic Behaviour of the Moments**

The next problem is to investigate the asymptotic behaviour of the moments \( M_n \), as \( n \to \infty \). Remember that this investigation for the classical case was done in [5].

A rough estimation shows us that the moments decrease exponentially. Indeed, if we set \( M_n \approx \lambda^{-n} \), we might infer that \( \lambda = (q-1)/d_M \), where \( d_M \) is the digit of maximal modulus in \( D \).

First, we assume that there is only one digit of maximal modulus; without loss of generality we may further assume that it is positive, since otherwise we would have simply to multiply the moments by \( (-1)^n \) and work with the set of digits \(-D\) instead.

We set

\[
m_n := M_n \cdot \lambda^n
\]

and show that this sequence has nicer properties. It satisfies the modified recurrence

\[
m_n = \frac{1}{d(q^n - 1)} \sum_{i=1}^{d} \sum_{j=0}^{n-1} \binom{n}{j} (\lambda d_i)^{n-j} m_j.
\]

To study this sequence further, we rewrite it as

\[
m_n \cdot d(q^n - 1) = \sum_{i=1}^{d} \left( \sum_{j=0}^{n} \binom{n}{j} (\lambda d_i)^{n-j} m_j - m_n \right)
\]

or

\[
m_n = \frac{1}{d q^n} \sum_{i=1}^{d} \sum_{j=0}^{n} \binom{n}{j} (\lambda d_i)^{n-j} m_j.
\]
and note that this holds for all $n \geq 0$. Then we introduce the *exponential generating function* 

$$m(z) = \sum_{n \geq 0} m_n \frac{z^n}{n!}$$

and get 

$$m(z) = \frac{1}{d} \sum_{i=1}^{d} e^{\frac{z}{d} \lambda_i} \sum_{k \geq 1} \frac{1}{d} \sum_{i=1}^{d} e^{\frac{z}{d} \lambda_i / d^k}.$$

As in [5], we have to consider the *Poisson transformed function* $\hat{m}(z) = e^{-z} m(z)$, which satisfies the functional equation 

$$\hat{m}(z) = \frac{1}{d} \sum_{i=1}^{d} \frac{e^{\frac{z}{d} \lambda_i + \frac{1}{q-1}}}{d^k} \hat{m}\left(\frac{z}{q}\right). \quad (3.1)$$

This functional equation (3.1) can also be solved by iteration: 

$$\hat{m}(z) = \prod_{k \geq 1} \frac{1}{d} \sum_{i=1}^{d} e^{\frac{z}{d} \lambda_i + \frac{1}{q-1}} d^k.$$

However, this product is not too useful, and we have to go back to the functional equation.

The next step is to consider the behaviour of $\hat{m}(z)$ for $z \to \infty$. The reason is that $m_n \sim \hat{m}(n)$. The justification for this is a technique called *depoissonization*.

The general references for that are [7] and [13]. We follow [5], where the technique is explained in more detail.

It is suggestive to use a new name $R(z)$ for $\frac{1}{d} \sum_{i \neq M} e^{\frac{z}{d} \lambda_i + \frac{1}{q-1}} \hat{m}\left(\frac{z}{q}\right)$ and consider it to be an auxiliary (and known) function; 

$$\hat{m}(z) = \frac{1}{d} \hat{m}\left(\frac{z}{q}\right) + R(z). \quad (3.2)$$

We compute the *Mellin transform* of (3.2) (see [2] for definitions and properties): 

$$\hat{m}^*(s) = q^s \hat{m}^*(s) + R^*(s) = \frac{R^*(s)}{1 - \frac{s}{d}}.$$

The function $m(z)$ can be recovered from this by *Mellin's inversion formula*, 

$$\hat{m}(z) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} R^*(s) \frac{z^{-s}}{1 - \frac{s}{d}} ds,$$

where $0 < c < \log_q d$. By shifting the integral to the right and taking the *negative residues* into account, we get the desired asymptotic behaviour of $\hat{m}(z)$. There are simple poles at $s = \log_q d + \frac{2k\pi i}{\log_q d}$, $k \in \mathbb{Z}$. The negative residue there is 

$$\frac{1}{\log q} R^* \left( \log_q d + \frac{2k\pi i}{\log_q d} \right) z^{-k\log_q d - \frac{2k\pi i}{\log_q d}}.$$

The value for $k = 0$ is of special interest; it is, to make it more explicit, 

$$\frac{1}{\log q} z^{-\log_q d} \int_0^\infty \frac{1}{d} \sum_{i \neq M} e^{\frac{z}{d} \lambda_i + \frac{1}{q-1}} \hat{m}\left(\frac{z}{q}\right) z^{-1} dz.$$
Traditionally, one collects all the terms into a periodic function.

**Theorem 2.** The $n$th moment $M_n$ of the Cantor--$(q,D)$ distribution has for $n \to \infty$ the following asymptotic behaviour

$$M_n = \left( \frac{d_M}{q - 1} \right)^n \Phi(- \log_q n) n^{-\log_q d} \left( 1 + O \left( \frac{1}{n} \right) \right),$$

where $\Phi(x)$ is a periodic function with period 1 and known Fourier coefficients. The mean (zeroth Fourier coefficient) is given by

$$\frac{1}{\log q} \int_0^{\infty} \frac{1}{d} \sum_{i=\mathbb{N}} \mathbb{E}^{(d,\lambda+1-1)} \left( \frac{z}{q} \right)^{\log_q d - 1} dz.$$

This integral can be computed numerically by replacing $\hat{m} \left( \frac{z}{q} \right)$ by the first few values of its Taylor expansion, which can be obtained by iterating the recurrence for the numbers $m_n$.

**Example.** We consider $q = 5$ and $D = \{-1,1,3\}$, so $d = 3$, $d_M = 3$, and $\lambda = \frac{4}{3}$. Then

$$m_n = \frac{1}{3(5^n - 1)} \sum_{j=0}^{n-1} \binom{n}{j} \left( \left( \frac{4}{3} \right)^{n-j} + \left( \frac{4}{3} \right)^{n-j} + 4^{n-j} \right) m_j.$$

So we replace $R(z)$ by

$$R_{\text{approx}} = \frac{1}{3} \left( e^{-16z/15} + e^{-8z/15} \right) e^{-z/5} \left( 1 + \frac{1}{20} z + \frac{1}{288} z^2 + \frac{19}{14400} z^3 + \frac{587}{13458000} z^4 + \ldots \right)$$

and compute

$$\frac{1}{\log 5} \int_0^{\infty} R_{\text{approx}} z^{\log_3/\log_5 - 1} dz = 0.59896 \ldots .$$

We find $M_{100} \left( \frac{4}{3} \right)^{100} \left( \frac{4}{3} \right)^{100} \approx 0.60351$.

The case when $-d_1 = d_d$ requires special care. If one has e.g. a symmetric set of digits $D$, then all odd moments vanish. A similar phenomenon occurs if the largest positive and negative digits coincide. Depoisonization, as it is described in [7] and in [13], does not cover this instance. But one could just extract coefficients $n! [z^n] m(z)$ using Cauchy’s integral formula using as the path of integration a circle of radius $n$. We have

$$n! [z^n] m(z) = \frac{1}{2\pi i} \oint_{|z| = n} \frac{m(z)}{z^{n+1}} dz.$$

After a separation of the integral between positive and negative half plane and a change of variable in the second term of the sum, we get

$$n! [z^n] m(z) = \frac{n!}{2\pi i} \int_{|z| = n, \Re z \geq 0} \frac{m(z)}{z^{n+1}} dz + (-1)^n \frac{1}{2\pi i} \int_{|z| = n, \Re z \leq 0} \frac{m(-z)}{z^{n+1}} dz.$$

We now use the fact that the saddle point in the integral

$$\frac{n!}{2\pi i} \int_{C} \frac{e^z}{z^{n+1}} dz = 1$$

...
lies at $z = n + O(1)$. For more details about the saddle point method, one can refer to [1] or to [10]. As $m(z) = e^z (e^{-z} m(z))$ where
\[ e^{-z} m(z) = \prod_{k \geq 1} \frac{\sum_{i=1}^d e^{z \delta_{d,k} + 1 - \delta_{d,k}}/\delta^k}{d} \]
is bounded, the previous saddle point is asymptotically not affected by multiplying $e^z$ by this infinite product. Thus we get
\[ \frac{n!}{2\pi i} \int_{|z|=n, \Re z \geq 0} \frac{m(z)}{z^{n+1}} \, dz \sim e^{-n} m(n) \left( 1 + O\left( \frac{1}{n} \right) \right), \]
and similarly
\[ \frac{n!}{2\pi i} \int_{|z|=n, \Re z \geq 0} \frac{m(-z)}{z^{n+1}} \, dz \sim e^{-n} m(-n) \left( 1 + O\left( \frac{1}{n} \right) \right). \]

Compare [4] for such an approach.

We next study both terms of the sum by making use of the Mellin transform as in the previous instance.

**Theorem 3.** In the instance of two dominant digits $d_1, d_2$, with $-d_1 = d_2 = d_M$, the $n$th moment $M_n$ of the Cantor–$(q, D)$ distribution has for $n \to \infty$ the following asymptotic behaviour
\[ M_n = \left( \frac{d_M}{q-1} \right)^n \frac{\Phi_1(-\log_q n) + (-1)^n \Phi_2(-\log_q n)}{n^{\log_q d} \left( 1 + O\left( \frac{1}{n} \right) \right)}, \]
where $\Phi_1(x)$ and $\Phi_2(x)$ are periodic functions with period 1 and known Fourier coefficients. The means (zeroth Fourier coefficients) are given by
\[ \frac{1}{\log q} \int_0^\infty \frac{1}{d} \sum_{i \neq d} \frac{e^{\frac{z}{d_i} \lambda + 1 - \frac{1}{d_i}}}{\frac{z}{q}} e^{-z/q} \, m \left( \frac{z}{q} \right) \, z^{\log_q d-1} \, dz, \]
\[ \frac{1}{\log q} \int_0^\infty \frac{1}{d} \sum_{i \neq 1} \frac{e^{\frac{z}{d_i} \lambda + 1 - \frac{1}{d_i}}}{\frac{z}{q}} e^{-z/q} \, m \left( \frac{z}{q} \right) \, z^{\log_q d-1} \, dz, \]
respectively.

While the even moments are always positive, the sign of the odd ones depends on the largest (in modulus) digit $d_i \in D$ such that not both $d_i$ and $-d_i$ are in $D$.

**Example.** We consider $q = 7$ and $D = \{-3, 2, 3\}$, so $d = 3$, $d_M = 3$, and $\lambda = 2$. Then
\[ m_n = \frac{1}{3(7^n - 1)} \sum_{j=0}^{n-1} \binom{n}{j} \left( (-6)^n - 4^n - 6^n \right) m_j. \]
So we replace $R(z)$ by
\[ R_{\text{approx}} = \frac{1}{3} \left( e^{-1z/7} + e^{-2z/7} \right) e^{-z/7} \left( 1 + \frac{1}{63} z + \frac{101}{6370} z^2 + \frac{251}{325770} z^3 + \ldots \right) \]
and compute
\[ \frac{1}{\log 7} \int_0^\infty R_{\text{approx}} z^{\log 3/\log 7-1} \, dz = 0.63967 \ldots . \]
This is the first contribution; call it $C_1$. Now we do the same for the set of digits $-D = \{-3, -2, 3\}$. Then we use

$$R_{\text{approx}} = \frac{1}{3} \left( e^{-12z/7} + e^{-10z/7} \right) e^{-z/7} \left( 1 - \frac{1}{63} z + \frac{101}{6301} z^2 - \frac{251}{5261762} z^3 + \ldots \right)$$

and find

$$\frac{1}{\log t} \int_0^\infty R_{\text{approx}} \frac{z}{\log^3 \log^7 t-1} \, dz = 0.39769 \ldots .$$

This is the second contribution; call it $C_2$.

We find $M_{100} \approx 100 \log 3 / \log 7 \approx 1.04057$; this is close to $C_1 + C_2 = 1.03737$. On the other hand, $M_{200} \approx 200 \log 3 / \log 7 \approx 0.26770$; this is to be compared with $C_1 - C_2 = 0.24197$.

4. Expected value of the minimum order statistic of the $(q, D)$-distribution

We consider random strings $c_1 c_2 c_3 \ldots$ where the $c_i$'s $ \in D = \{d_1, d_2, \ldots, d_d\}$ are equally likely. We then consider the random variable value that maps $c_1 c_2 c_3 \ldots$ to the real number

$$\text{value}(c_1 c_2 c_3 \ldots) = \sum_{i \geq 1} c_i q^{-i} \in \left[ \frac{d_1}{q-1}, \frac{d_d}{q-1} \right].$$

The strings now have a natural order from the usual ordering of the real numbers. This is easily seen to be equivalent to the lexicographic ordering of strings, i.e. $c_1 c_2 c_3 \ldots < c_1' c_2' c_3' \ldots$ iff there is a $k$ such that $c_i = c_i'$ for $i = 1, \ldots, k-1$ and $c_k < c_k'$. It thus makes sense to speak of order statistics for strings. Suppose that $n$ independent random strings $w_1, \ldots, w_n$ are produced. We denote by $a_n$ the average value of the minimum of the $n$ real numbers $\text{value}(w_1), \ldots, \text{value}(w_n)$. We derive the following recursion for the expected minimum

$$a_n = \frac{1}{qd^n} \left[ \sum_{i=1}^{d-1} \left( \sum_{k=1}^{n} \binom{n}{k} (d-i)^{n-k} (d_i + a_k) \right) + d_d + a_n \right], \quad n \geq 1 ,$$

or

$$(d^n q - d) a_n = \sum_{i=1}^{d} \left[ (d-i+1)^n - (d-i)^n \right] d_i + \sum_{k=1}^{n-1} \sum_{i=1}^{d-1} \binom{n}{k} (d-i)^{n-k} d_k , \quad n \geq 1 .$$

This recursion is obtained by considering the smallest digit $d_i$ that at least one of the $n$ random strings has in its first position. The minimum value will be one of these, and be determined recursively; the first position adds the quantity $\frac{d_i}{q}$ to the recursively determined minimum.

Now, if $n$ is large, it is almost certain that there is a string starting with $d_1 d_1 d_1 \ldots$ among the $n$ random strings, producing the minimal value (in the limit) $\frac{d_1}{q-1}$. Remember that in the classical Cantor case $d_1 = 0$, and the question was to analyze how fast $a_n$ approaches zero. In order to obtain meaningful results, we define $\alpha_n := a_n - \frac{d_1}{q-1}$ and rewrite the recursion:
\[(d^n q - d) \left( \alpha + \frac{d_1}{q - 1} \right) = \sum_{i=1}^{d} \left[ (d - i + 1)^n - (d - i)^n \right] d_i + \sum_{k=1}^{n-1} \sum_{i=1}^{d-1} \binom{n}{k} (d - i)^{n-k} \left( \alpha + \frac{d_1}{q - 1} \right) \]

or

\[(d^n q - d) \alpha = -(d - 1)^n d_1 + \sum_{i=2}^{d} \left[ (d - i + 1)^n - (d - i)^n \right] d_i + \sum_{k=1}^{n-1} \sum_{i=1}^{d-1} \binom{n}{k} (d - i)^{n-k} \alpha_k.\]

Then we introduce the exponential generating function (with \(\alpha_0 := 0\))

\[a(z) = \sum_{n \geq 0} \alpha_n \frac{z^n}{n!}\]

and get

\[qa'(z) - da(z) = (1 - e^{(d-1)z})d_1 + \sum_{i=2}^{d} \left( e^{(d-i+1)z} - e^{(d-i)z} \right) d_i + \sum_{i=1}^{d-1} \left( e^{(d-i)z} - 1 \right) a(z),\]

or

\[a(z) = \frac{1}{q} (1 - e^{dz}) a(z) + \frac{d_1}{q} (1 - e^{(d-1)z}) + \frac{1}{q} \sum_{i=1}^{d} \left( e^{(d-i+1)z} - e^{(d-i)z} \right) d_i.\]

We have to consider the Poisson transformed function \(\widehat{a}(z) = e^{-z} a(z)\), which satisfies the functional equation

\[\widehat{a}(dz) = \frac{1}{q} \left( 1 - e^{-dz} \right) \widehat{a}(z) + \frac{d_1}{q} (e^{-dz} - e^{-z}) + \frac{1}{q} (e^z - 1) \sum_{i=2}^{d} e^{-iz} d_i.\]

The next step is to consider the behaviour of \(\widehat{a}(z)\) for \(z \to \infty\). The reason is that \(\alpha_n \sim \widehat{a}(n)\). The justification for this is again the technique of depoissonization. We set

\[b(z) = \frac{1 - e^{-dz}}{1 - e^{-z}}, \quad \text{and} \quad \phi(z) = \prod_{j=0}^{\infty} b(zd^j) = \frac{1}{1 - e^{-z}}\]

and

\[c(z) = \frac{d_1}{q} (e^{-z} - e^{-z/d}) + \frac{1}{q} (e^{-z/d} - 1) \sum_{i=2}^{d} e^{-iz/d} d_i.\]

We then get

\[\widehat{a}(z) = \sum_{n=0}^{\infty} q^{-n} c(q^{-n} z) \prod_{k=1}^{n} b(q^{-k} z).\]

As \(\widehat{a}(0) = 0\), we finally obtain

\[\widehat{a}(z) \phi(z) = \sum_{n=0}^{\infty} q^{-n} c(q^{-n} z) \phi(d^{-n} z). \quad (4.1)\]

We compute the Mellin transform of (4.1); since it is a harmonic sum (see [2] for more background), we obtain

\[\left( \widehat{a}(z) \phi(z) \right)^*(s) = \sum_{n \geq 0} q^{-n} d^{\alpha s} \left( c(z) \phi(z) \right)^*(s) = \frac{1}{1 - \frac{d^s}{q} \left( c(z) \phi(z) \right)^*(s)},\]
The function $\hat{a}(z)$ can be recovered from this by Mellin’s inversion formula,

$$
\hat{a}(z)\phi(z) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{(c(z)\phi(z))^*}{1 - \frac{d^2}{q}} z^{-s} \, ds,
$$

where $0 < c < \log_d q$. By shifting the integral to the right and taking the negative residues into account, we get the desired asymptotic behaviour of $\hat{a}(z)$. There are simple poles at $s = \log_d q + \frac{2k\pi i}{\log d}$, $k \in \mathbb{Z}$. The negative residue there is

$$
\frac{1}{\log d} \left(c(z)\phi(z)\right)^* \left(\log_d q + \frac{2k\pi i}{\log d}\right) z^{-\log_d q - \frac{2k\pi i}{\log d}}.
$$

The value for $k = 0$ is of special interest; it is, to make it more explicit,

$$
\frac{1}{\log d} z^{-\log_d q} \int_0^\infty c(z)\phi(z) z^{\log_d q - 1} \, dz.
$$

Moreover $\phi(z) \sim 1$ as $z \to \infty$. One collects all the terms into a periodic function.

**Theorem 4.** The expected value of the minimum order statistics of the Cantor–$(q, D)$ distribution has for $n \to \infty$ the following asymptotic behaviour

$$
a_n = \frac{d_1}{q - 1} + \Phi(-\log_d q) n^{-\log_d q} \left(1 + O\left(\frac{1}{n}\right)\right),
$$

where $\Phi(x)$ is a periodic function with period 1 and known Fourier coefficients. The mean (zeroth Fourier coefficient) is given by

$$
\frac{1}{\log d} \int_0^\infty c(z)\phi(z) z^{\log_d q - 1} \, dz.
$$

As before, the integral can be computed numerically. However, one can say even more, since

$$
\int_0^\infty \frac{e^{-kz} z^{s-1} \, dz}{e^z - 1} = \Gamma(s) \left(\zeta(s) - \sum_{j=1}^{k} \frac{1}{j^s}\right),
$$

so the integral could be expressed in terms of the Gamma and zeta function.

**Example.** We consider again the example $q = 5$ with $D = \{-1, 1, 3\}$ and $d = 3$. Then $a_{100} + \frac{1}{4} \approx 0.00265441$. Further,

$$
c(z) = \frac{1}{5} \left(-4e^{-z} + 2e^{-z/3} + 2e^{-2z/3}\right),
$$

and

$$
\frac{1}{\log 3} \int_0^\infty c(z)\phi(z) z^{\log_3 5 / \log 3 - 1} \, dz \approx 1.77099.
$$

Eventually, $1.77099 \cdot 10^{\log_3 5 / \log 3} \approx 0.00208$. 
References


Frédérique Bassino, Institut Gaspard Monge, Université de Marne-la-Vallée, 77454 Marne-la-Vallée Cedex 2, France

E-mail address: bassino@univ-mlv.fr

Helmut Prodinger, The John Knopfmacher Centre for Applicable Analysis and Number Theory, School of Mathematics, University of the Witwatersrand, P. O. Wits, 2050 Johannesburg, South Africa

E-mail address: helmut@staff.ms.wits.ac.za