MOMENTS OF DISTRIBUTIONS RELATED TO DIGITAL EXPANSIONS

LIGIA L. CRISTEA* AND HELMUT PRODINGER†

Abstract. The paper studies certain (probability) measures of binomial type defined in a recursive way on the unit interval. These measures are related to the sum-of-digit function and similar quantities. In particular, we undertake an asymptotic analysis of the moments of the corresponding distributions. This is done by a combination of a method based on the Mellin transform and the depossonisation technique.

1. Introduction

Let $n = \sum_{i\geq 0} a_i q^i$ with $a_i \in \{0,1,\dots,q-1\}$ be the q-adic expansion of the natural number n and $S_q(n) = \sum_{i\geq 0} a_i$ the sum-of-digits function. The formula

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

with a periodic function $F(x)$ of period 1 appeared together with an elementary proof in [3]; compare also [7]. In order to study higher moments like $\sum_{n<N} S_q^d(n)$ and related quantities, the authors Kobayashi, Muramoto, Okada, Sekiguchi and Shiota designed a method which they used in a series of papers [16, 18, 13, 17, 19, 14, 15]. This method uses a *binomial measure* (and, of course, generalisations). In its simplest form, it is defined in the following manner:
Let, $0 < r < 1$ and $I = I_{0,0} = [0,1]$,

$$\begin{aligned} I_{n,j} &= \Big[\frac{j}{2^n}, \frac{j+1}{2^n}\Big), \text{ for } j = 0,1,\dots,2^n-2,\\ I_{n,2^n-1} &= \Big[\frac{2^n-1}{2^n}, 1\Big], \end{aligned}$$

for $n = 1,2,3,\dots$. The binomial measure μ_r is a probability measure on I uniquely determined by the conditions

$$\begin{aligned} \mu_r(I_{n+1,2j}) &= r\mu_r(I_{n,j}),\\ \mu_r(I_{n+1,2j+1}) &= (1-r)\mu_r(I_{n,j}), \end{aligned}$$

Date: May 11, 2004.
2000 *Mathematics Subject Classification.* 30E05, 60C05.
Key words and phrases. moments, binomial measure, depoissonisation, Mellin transform, Gray code, Cantor distribution.
*This author was supported by the Austrian Research Foundation (FWF), Project S 8308-N05.
†This author was supported by the NRF grant n^o2053748.
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.

for $n = 0, 1, 2, \dots$ and $j = 0, 1, \dots, 2^n - 1$.
The aim of the present paper is to study the asymptotic behaviour of the moments of these singular moments, which is a problem different from the moments of the sum-of-digits function.
In Section 2 we treat the (simplest) case $q = 2$, namely the binomial distribution (depending on the parameter r) on the unit interval, in more detail. Later on we use the same method for more general cases. We follow the approach presented in [1], [10]. The binomial distribution has already been studied, by means of a different method, in [9].
Section 3 studies the more general case of the multinomial distribution on the unit interval. Here $q \geq 2$ and the distribution depends on $q - 1$ parameters. We skip some of the details, as we follow the same ideas as in Section 2.
Section 4 is a direct application of the previous section: instead of the unit interval one takes the unit square, divided into N^2 congruent squares of side-length $1/N$. We define and study the asymptotic behaviour of the moments vector of the multinomial distribution on the unit square. The reasons why we restrict our computations to the case $N = 3$ are twofold: on the one hand, it is a straightforward exercise to generalise the result for any $N \geq 2$, on the other hand, by an appropriate choice of the parameters in the case $N = 3$ one gets all informations about the planar analogon of the Cantor distribution (studied in [10]).
In Section 5 we study another distribution on the unit interval: the Gray code distribution. Although our considerations are based on ideas that are analogous to those of the sections 2 and 3, in this case we have a system of recursions, not just one recursion between the moments of different orders.
For the binomial, multinomial and Gray distribution respectively we check numerically the formulae describing the asymptotic behaviour of the moments. Our asymptotic formulae are of the form $n^\theta \Psi(\log n)$, where the periodic function Ψ usually appears in the form $\Phi(-\log n)$, with a "nice" function $\Phi(x)$.

2. The binomial distribution and its moments

2.1. **The binomial measure.** In [16] the authors define the *binomial measure* μ_r $(0 < r < 1)$ on the unit interval as follows.
Let $I = I_{0,0} = [0, 1]$ and

$$\begin{aligned} I_{n,j} &= \Big[\frac{j}{2^n}, \frac{j+1}{2^n}\Big), \text{ for } j = 0, 1, \dots, 2^n - 2, \\ I_{n,2^n-1} &= \Big[\frac{2^n - 1}{2^n}, 1\Big], \end{aligned}$$

for $n = 1, 2, 3, \dots$. The binomial measure μ_r is a probability measure on I uniquely determined by the conditions

$$\begin{aligned} \mu_r(I_{n+1,2j}) &= r\mu_r(I_{n,j}), \\ \mu_r(I_{n+1,2j+1}) &= (1 - r)\mu_r(I_{n,j}), \end{aligned}$$

for $n = 0, 1, 2, \dots$ and $j = 0, 1, \dots, 2^n - 1$.

For every $n \in \{1, 2, 3, \dots\}$ we call the intervals $I_{n,j}$, with $j \in \{0, 1, \dots, 2^n - 1\}$ *elementary intervals of level* n, in short n*-intervals*. They will also occur in the considerations of Section 5.

Let us now consider $\mathcal{W}$ to be the set of all infinite words over the alphabet $D = \{0, 1\}$ and $\mathcal{W}_m$ the set of all words of length m ($m \geq 1$) over the alphabet D. For every word $\omega \in \mathcal{W}$, $\omega = \omega_1\omega_2 \dots \omega_n \dots$ we define its *value*

$$val(\omega) = \sum_{i \geq 1} \omega_i \cdot 2^{-i}.$$

In this way we assign to every infinite word $\omega = \omega_1\omega_2 \dots \omega_n \dots$ the binary fraction $0.\omega_1\omega_2 \dots \omega_n \dots$. Analogously we define the value of any word of $\mathcal{W}_m$ in the obvious way.

Remark It is easy to show that in the case of choosing in a random way (with respect to μ_r) a word $\omega \in \mathcal{W}$, we have $\mathbb{P}(\omega_k = 0) = \mathbb{P}(\omega_k = 1) = \frac{1}{2}$, for $k = 1, 2, \dots$, thus these probabilities depend neither on the value of the parameter r nor on k.

2.2. **The moments of the binomial distribution.** In the following we study the moments of the function val with respect to the distribution defined by μ_r.

We denote by M_n the moment of order n. We have

$$M_n = \sum_{\omega \in \mathcal{W}} \mu_r(\omega) \cdot (val(\omega))^n.$$

It is easy to verify that the relation

$$val(d\omega) = d \cdot 2^{-1} + 2^{-1} \cdot val(\omega)$$

holds for all infinite strings $\omega \in \mathcal{W}$ and $d \in \{0, 1\}$. Here $d\omega$ denotes the concatenation of the digit d and the string ω.

Let now

$$M_n^m = \sum_{\omega \in \mathcal{W}_m} \mu_r(\omega) \cdot (val(\omega))^n.$$

With the above notations we have

$$M_n = \lim_{m \to \infty} M_n^m.$$

We use this relation in order to deduce a recurrence relation between the moments of different orders. In order to do this we introduce the notation $\mathcal{W}_m^k$ for the set of words of $\mathcal{W}_m$ containing exactly k times the character 0. We obtain then

$$M_n^m = \sum_{k=0}^{m} r^k (1 - r)^{m-k} \sum_{\omega \in \mathcal{W}_m^k} (val(\omega))^n.$$

By analysing the first character of the words occurring in the last sum we get

$$\begin{aligned}M_n^m &= r\cdot\frac{1}{2^n}\sum_{k=0}^{m-1} r^k(1-r)^{m-1-k}\sum_{\omega\in\mathcal{W}_{m-1}^k}(val(\omega))^n\\
&+(1-r)\cdot\frac{1}{2^n}\sum_{k=0}^{m-1} r^k(1-r)^{m-1-k}\sum_{\omega\in\mathcal{W}_{m-1}^k}(1+val(\omega))^n\\
&= r\cdot\frac{1}{2^n}\sum_{k=0}^{m-1} r^k(1-r)^{m-1-k}\sum_{\omega\in\mathcal{W}_{m-1}^k}(val(\omega))^n\\
&+(1-r)\cdot\frac{1}{2^n}\sum_{k=0}^{m-1} r^k(1-r)^{m-1-k}\sum_{j=0}^{n-1}\binom{n}{j}\sum_{\omega\in\mathcal{W}_{m-1}^k}(val(\omega))^j\\
&= \frac{1}{2^n}M_n^{m-1}+(1-r)\cdot\frac{1}{2^n}\sum_{j=0}^{n-1}\binom{n}{j}M_j^{m-1}.\end{aligned}$$

By taking the limit for $m\to\infty$, after a rearrangement of terms we get the recursion

$$M_n=\frac{r}{2^n}M_n+\frac{1-r}{2^n}\sum_{j=0}^{n}\binom{n}{j}M_j,\ \text{for all integers } n\ge 1. \tag{2.1}$$

It is easy to verify that $M_0=1$. Thus we have proven the following

Theorem 1. *The moments of the binomial distribution μ_r satisfy the relations:*

$$\begin{aligned}M_0&=1,\\ M_n&=\frac{r}{2^n}M_n+\frac{1-r}{2^n}\sum_{j=0}^{n}\binom{n}{j}M_j,\ \textit{for all integers } n\ge 1.\end{aligned}$$

Remark The above recursion also holds for $n=0$. One can use this recursion in order to compute a list of the first moments M_1, M_2, M_3,....

From (2.1) one can express M_n with the help of the previous moments:

$$M_n=\frac{1-r}{2^n-1}\sum_{j=0}^{n-1}\binom{n}{j}M_j.$$

2.3. **The asymptotics of the moments M_n.** The recursion (2.1) enables us to study the asymptotic behaviour of the moments M_n, for $n\to\infty$.
Earlier references about the asymptotic behaviour of the moments of singular measures of the Cantor type are, e.g., [10, 1, 5, 8]. Our approach follows [10, 1].
First, we define the *exponential generating function*

$$M(z)=\sum_{n\ge 0}M_n\frac{z^n}{n!}.$$

From (2.1) we get, by multiplying by $\frac{z^n}{n!}$ and summing up over all integers $n \geq 0$

$$M(z) = r \cdot M(\frac{z}{2}) + (1-r) \cdot M(\frac{z}{2}) \cdot e^{-\frac{z}{2}}.$$

This implies that the *Poisson transformed function* $\widehat{M}(z) = M(z) \cdot e^{-z}$ satisfies

$$\widehat{M}(z) = r \cdot \widehat{M}(\frac{z}{2}) \cdot e^{-\frac{z}{2}} + (1-r) \cdot \widehat{M}(\frac{z}{2}). \tag{2.2}$$

Herefrom we obtain by iteration:

$$\widehat{M}(z) = \prod_{k \geq 1} \big(r \cdot e^{\frac{z}{2^k}} + (1-r) \big).$$

As we are looking for the asymptotics of the moments M_n we are going to study the behaviour of $\widehat{M}(z)$ for $z \to \infty$. This is based on the fact that $M_n \sim \widehat{M}(n)$, which can be justified by using depoissonisation (for details about depoissonisation we refer to [11], [20]).

The basic idea is to extract the coefficients M_n from $M(z)$ using Cauchy's integral formula and the saddle point method. This leads in our applications to an approximation $M_n = \widehat{M}(n)(1 + \mathcal{O}\frac{1}{n})$, with more terms being available in principle. Let us rewrite (2.2) as

$$\widehat{M}(z) = (1-r) \cdot \widehat{M}(\frac{z}{2}) + R(z), \tag{2.3}$$

where $R(z) = r \cdot \widehat{M}(\frac{z}{2}) \cdot e^{-\frac{z}{2}}$ is considered to be an auxiliary function which we treat as a known function. We compute the Mellin transform $\widehat{M}^*(s)$ (for details regarding the Mellin transform we refer to [6]) of the function $\widehat{M}(z)$ in (2.3). We get

$$\widehat{M}^*(s) = (1-r) \cdot 2^s \cdot \widehat{M}^*(s) + R^*(s) = \frac{R^*(s)}{1-(1-r) \cdot 2^s}.$$

Now the function $\widehat{M}(z)$ can be obtained by applying the Mellin inversion formula, namely

$$\widehat{M}(z) = \frac{1}{2\pi \mathrm{i}} \int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} \widehat{M}^*(s) \cdot z^{-s} ds = \frac{1}{2\pi \mathrm{i}} \int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} \frac{R^*(s)}{1-(1-r) \cdot 2^s} \cdot z^{-s} ds, \tag{2.4}$$

where $0 < c < \log_2 \frac{1}{1-r}$.

We shift the integral to the right and take the residues with negative sign into account in order to estimate $\widehat{M}(z)$ in (2.4). The function under the integral has simple poles at $s_k = \log_2 \frac{1}{1-r} + \frac{2k\pi \mathrm{i}}{\log 2}$, $k \in \mathbb{Z}$. For these the residues with negative sign are

$$\frac{1}{\log 2} R^* \Big(\log_2 \frac{1}{1-r} + \frac{2k\pi\mathrm{i}}{\log 2} \Big) z^{-\log_2 \frac{1}{1-r} - \frac{2k\pi\mathrm{i}}{\log 2}},$$

with $R^*(s) = \int_0^\infty r\widehat{M}(\frac{z}{2}) \cdot e^{-\frac{z}{2}} \cdot z^{s-1} dz$.

For $k = 0$ the residue with negative sign is, using the definition of $R(z)$,

$$\frac{1}{\log 2} \cdot z^{\log_2 (1-r)} \int_0^\infty r\widehat{M}(\frac{z}{2}) \cdot e^{-\frac{z}{2}} \cdot z^{\log_2 \frac{1}{1-r} - 1} dz.$$

This term plays an important role in the asymptotic behaviour of the nth moment M_n of the binomial distribution. In order to get this one collects all mentioned residues into a periodic function. With the above considerations we get

Theorem 2. *The nth moment M_n of the binomial distribution μ_r admits the asymptotic estimate*

$$M_n = \Phi(-\log_2 n) \cdot n^{\log_2(1-r)}\Big(1 + \mathcal{O}\big(\frac{1}{n}\big)\Big), \tag{2.5}$$

for $n \to \infty$, where $\Phi(x)$ is a periodic function having period 1 and known Fourier coefficients. The mean (zeroth Fourier coefficient) of Φ is given by the expression

$$\frac{1}{\log 2}\int_0^\infty r\widehat{M}\big(\frac{z}{2}\big) \cdot e^{-\frac{z}{2}} \cdot z^{\log_2 \frac{1}{1-r}-1} dz. \tag{2.6}$$

Remark One can compute this integral numerically by taking for $\widehat{M}(\frac{z}{2})$ the first few terms of its Taylor expansion. These can be found from the recurrence (2.1) for the numbers M_n. The integral in (2.6) can be written as an infinite series

$$\begin{aligned}\int_0^\infty r\widehat{M}\big(\frac{z}{2}\big) \cdot e^{-\frac{z}{2}} \cdot z^{\log_2 \frac{1}{1-r}-1} dz &= r\int_0^\infty e^{-z}\sum_{k\ge 0} M_k \frac{z^k}{2^k k!} z^{\log_2 \frac{1}{1-r}-1} dz \\ &= r\sum_{k\ge 0}\frac{M_k}{2^k k!}\cdot \Gamma\Big(k + \log_2 \frac{1}{1-r}\Big).\end{aligned}$$

This series is well suited for numerical computations. For example, let $r = 0.6$, then $M_{100} = 0.002453\ldots$ and the value predicted by (2.5) (without the oscillation) is $0.002491\ldots$.

3. The multinomial distribution and its moments

3.1. **The multinomial measure.** Let $q \geq 2$ be a positive integer. Analogously as in Section 2 we introduce the notation: $I = I_{0,0} = [0, 1]$ and

$$\begin{aligned} I_{n,j} &= \Big[\frac{j}{q^n}, \frac{j+1}{q^n}\Big), \text{ for } j = 0, 1, \ldots, q^n - 2, \\ I_{n,q^n-1} &= \Big[\frac{q^n-1}{q^n}, 1\Big], \end{aligned}$$

for $n = 1, 2, 3, \ldots$. Let $\mathbf{r} = (r_0, r_1, \ldots, r_{q-1})$ with $0 \leq r_i \leq 1$ and $\sum_{k=0}^{q-1} r_k = 1$.
The *multinomial measure* $\mu_{q,\mathbf{r}}$ is the probability measure on I defined by

$$\mu_{q,\mathbf{r}}(I_{n+1,qj+k}) = r_k \cdot \mu_{q,\mathbf{r}}(I_{n,j})$$

for $n = 0, 1, 2, \ldots,\ j = 0, 1, \ldots, q^n - 1,\ k = 0, 1, \ldots, q - 1$. For further details about properties of the multinomial measure we refer to [18].

Remark Obviously, this measure generalises the binomial measure defined earlier: by taking $q = 2$, $r_0 = r$ and $r_1 = 1-r$ (with $0 < r < 1$) in the definition of the multinomial measure we get the binomial measure.
Moreover, for $q = 3$, $r_2 = 0$ and $r_1 = r_3 = 1/2$ we get the well-known Cantor distribution (for details we refer to [10]).

We let $\mathcal{W}(q)$ be the set of all infinite words over the alphabet $D = \{0, 1, \ldots, q-1\}$ and $\mathcal{W}_m(q)$ the set of all words of length m ($m \geq 1$) over the alphabet D. For every word $\omega \in \mathcal{W}(q)$, $\omega = \omega_1\omega_2 \ldots \omega_n \ldots$ we define its *value*

$$val(\omega) = \sum_{i \geq 1} \omega_i \cdot q^{-i}.$$

3.2. **The moments of the multinomial distribution.** As in the previous section, we want to estimate the moments M_n of the function val with respect to the multinomial distribution. We have

$$M_n = \sum_{\omega \in \mathcal{W}(q)} \mu_r(\omega) \cdot (val(\omega))^n.$$

Let

$$M_n^m = \sum_{\omega \in \mathcal{W}_m(q)} \mu_r(\omega) \cdot (val(\omega))^n.$$

Now we want to deduce a recurrence formula relating moments of different orders to each other. For this we proceed analogously to Section 2. Let us denote, for $m \geq 1$,

$$\mathcal{K}(m, q) = \{\mathbf{k} = (k_0, k_1, \ldots, k_{q-1}) \in \{0, 1, 2, \ldots, m\}^q : k_0 + k_1 + \cdots + k_{q-1} = m\}$$

and, for a given integer $m \geq 1$, let $\mathcal{W}_m^{\mathbf{k}}(q)$ denote the set of all words ω of $W_m(q)$ with the property that for every $d_i = i \in \{0, 1, \ldots, q-1\}$ the word ω contains exactly k_i times the digit d_i. With the above notations, we have

$$M_n^m = \sum_{\substack{\mathbf{k} \in \mathcal{K}(m,q) \\ w \in \mathcal{W}_m^{\mathbf{k}}(q)}} \prod_{i=0}^{q-1} r_i^{k_i} \cdot (val(\omega))^n.$$

By analysing the first character (digit) of the words ω occurring in the above sum we obtain

$$\begin{aligned} M_n^m &= \sum_{i=0}^{q-1} r_i \frac{1}{q^n} \sum_{\substack{\mathbf{k} \in \mathcal{K}(m-1,q) \\ \omega \in \mathcal{W}_{m-1}^{\mathbf{k}}(q)}} \left(\prod_{l=0}^{q-1} r_l^{k_l} \right) (d_i + val(\omega))^n \\ &= \frac{1}{q^n} \sum_{i=0}^{q-1} r_i \sum_{j=0}^{n} \binom{n}{j} d_i^{n-j} \cdot M_j^{m-1}. \end{aligned}$$

By taking the limit for $m \to \infty$ we get

Theorem 3. *The moments of the multinomial distribution $\mu_{q,\mathbf{r}}$ satisfy the relations:*

$$M_0 = 1,$$

$$M_n = \frac{1}{q^n} \sum_{i=0}^{q-1} r_i \sum_{j=0}^{n} \binom{n}{j} d_i^{n-j} \cdot M_j, \text{ for all integers } n \geq 1. \tag{3.1}$$

Remark The above recursion holds also for $n = 0$. As in the previous case, one can use the above recursion in order to compute the first consecutive moments $M_1, M_2, M_3, \dots$.

From (3.1) one can express M_n with the help of the previous moments:

$$M_n = \frac{1}{q^n - 1} \sum_{i=0}^{q-1} r_i \sum_{j=0}^{n-1} \binom{n}{j} d_i^{n-j} \cdot M_j.$$

3.3. **The asymptotics of the moments M_n.** The recursion (3.1) provides useful information for the study of the asymptotic behaviour of the moments M_n, for $n \to \infty$. Let us consider the *exponential generating function* $M(z) = \sum_{n \geq 0} M_n \frac{z^n}{n!}$. From (3.1) we get then

$$M(z) = \sum_{i=0}^{q-1} r_i \cdot e^{d_i \cdot \frac{z}{q}} \cdot M(\frac{z}{q}).$$

This yields for the *Poisson transformed function* $\widehat{M}(z) = M(z) \cdot e^{-z}$,

$$\widehat{M}(z) = \sum_{i=0}^{q-1} r_i \cdot e^{\frac{z}{q}(d_i + 1 - q)} \cdot \widehat{M}(\frac{z}{q}) \tag{3.2}$$

and, by iteration in (3.2),

$$\widehat{M}(z) = \prod_{k \geq 1} \sum_{i=0}^{q-1} r_i \cdot e^{z(d_i + 1 - q)q^{-k}}.$$

By applying depoissonisation, $M_n \sim \widehat{M}(n)$ for $n \to \infty$ and thus we are going to analyse the behaviour of $M(z)$ for $z \to \infty$.
In order to do this we rewrite (3.2) as

$$\widehat{M}(z) = r_{q-1} \cdot \widehat{M}(\frac{z}{q}) + R(z), \tag{3.3}$$

where the auxiliary function R (here presumed to be a known function) has the expression

$$R(z) = \sum_{i=0}^{q-2} r_i \cdot e^{\frac{z}{q}(d_i + 1 - q)} \cdot \widehat{M}(\frac{z}{q}).$$

Thus the Mellin transform $\widehat{M^*}(s)$ of the function $\widehat{M}(z)$ in (3.3) satisfies

$$M^*(s) = r_{q-1} \cdot q^s \cdot \widehat{M}^*(s) + R^*(s) = \frac{R^*(s)}{1 - r_{q-1} \cdot q^s}.$$

Now we apply the Mellin inversion formula in order to recover the function $\widehat{M}(z)$ and obtain

$$\widehat{M}(z) = \frac{1}{2\pi \mathrm{i}} \int_{c - \mathrm{i}\infty}^{c + \mathrm{i}\infty} \widehat{M}^*(s) \cdot z^{-s} ds = \frac{1}{2\pi \mathrm{i}} \int_{c - \mathrm{i}\infty}^{c + \mathrm{i}\infty} \frac{R^*(s)}{1 - r_{q-1} \cdot q^s} \cdot z^{-s} ds, \tag{3.4}$$

where $0 < c < \log_q \frac{1}{r_{q-1}}$.
We evaluate $\widehat{M}(z)$ aymptotically in (3.4) by shifting the integral to the right and taking

the residues with negative sign. The function under the above integral has simple poles at $s_k = \log_q \frac{1}{r_{q-1}} + \frac{2k\pi \mathrm{i}}{\log q}$, $k \in \mathbb{Z}$. The residues with negative sign at the points s_k are

$$\frac{1}{\log q} R^*\Big(\log_q \frac{1}{r_{q-1}} + \frac{2k\pi \mathrm{i}}{\log q}\Big) \cdot z^{-(\log_q \frac{1}{r_{q-1}} + \frac{2k\pi \mathrm{i}}{\log q})},$$

where $R^*\Big(\log_q \frac{1}{r_{q-1}} + \frac{2k\pi \mathrm{i}}{\log q}\Big) = \int_0^\infty \sum_{i=0}^{q-2} r_i \cdot e^{\frac{z}{q}(d_i+1-q)} \cdot \widehat{M}(\frac{z}{q}) \cdot z^{\log_q \frac{1}{r_{q-1}} + \frac{2k\pi \mathrm{i}}{\log q} - 1} dz.$

Theorem 4. *The nth moment M_n of the multinomial distribution $\mu_{q,\mathbf{r}}$ admits the asymptotic estimate*

$$M_n = \Phi(-\log_q n) \cdot n^{\log_q r_{q-1}} \Big(1 + \mathcal{O}\big(\frac{1}{n}\big)\Big), \tag{3.5}$$

for $n \to \infty$, where $\Phi(x)$ is a periodic function having period 1 and known Fourier coefficients. The mean (zeroth Fourier coefficient) of Φ is given by the expression

$$\frac{1}{\log q} \int_0^\infty \sum_{i=0}^{q-2} r_i \cdot e^{\frac{z}{q}(d_i+1-q)} \cdot \widehat{M}(\frac{z}{q}) \cdot z^{\log_q \frac{1}{r_{q-1}} - 1} dz. \tag{3.6}$$

Remark The above integral can be computed numerically by replacing $\widehat{M}(\frac{z}{q})$ by the first few terms of its Taylor expansion which can found iterating the recurrence relation (3.1). As before we can rewrite (3.6) as

$$\frac{1}{\log q} \sum_{i=0}^{q-2} r_i \sum_{k \geq 0} \frac{M_k}{q^k k!} \cdot \Big(\frac{q}{q-i}\Big)^{k + \log_q \frac{1}{r_{q-1}}} \cdot \Gamma\Big(k + \log_q \frac{1}{r_{q-1}}\Big).$$

The numerical computations for $r_0 = 0.2$, $r_1 = 0.33$, $r_2 = 0.25$, $r_3 = 0.22$ yield $M_{100} = 0.006887\ldots$, whereas the value predicted by (3.5) (without the oscillation) is $0.006970\ldots$.

4. The multinomial distribution on the unit square and its moments-vector

4.1. **The multinomial measure on the unit square.** First we define a unit-square analogon of the multinomial measure. In order to do this we consider the following IFS (Iterated Functions System = system of contractions, for more details on iterated functions systems we refer to [4]):

$$\Big\{ [0,1]^2, \phi_0(x_1,x_2) = \left(\frac{1}{3}x_1, \frac{1}{3}x_2\right), \phi_1(x_1,x_2) = \left(\frac{1}{3}x_1, \frac{1}{3}x_2 + \frac{1}{3}\right),$$
$$\phi_2(x_1,x_2) = \left(\frac{1}{3}x_1, \frac{1}{3}x_2 + \frac{2}{3}\right), \phi_3(x_1,x_2) = \left(\frac{1}{3}x_1 + \frac{1}{3}, \frac{1}{3}x_2\right),$$
$$\phi_4(x_1,x_2) = \left(\frac{1}{3}x_1 + \frac{1}{3}, \frac{1}{3}x_2 + \frac{1}{3}\right), \phi_5(x_1,x_2) = \left(\frac{1}{3}x_1 + \frac{1}{3}, \frac{1}{3}x_2 + \frac{2}{3}\right),$$
$$\phi_6(x_1,x_2) = \left(\frac{1}{3}x_1 + \frac{2}{3}, \frac{1}{3}x_2\right), \phi_7(x_1,x_2) = \left(\frac{1}{3}x_1 + \frac{2}{3}, \frac{1}{3}x_2 + \frac{1}{3}\right),$$
$$\phi_8(x_1,x_2) = \left(\frac{1}{3}x_1 + \frac{2}{3}, \frac{1}{3}x_2 + \frac{2}{3}\right) \Big\}$$

$$= \left\{ [0,1]^2, w_h(x_1,x_2) = \left(\frac{1}{3}x_1 + \frac{\alpha_1}{3}, \frac{1}{3}x_2 + \frac{\alpha_2}{3} \right),\ h \in \{0,\ldots,8\},\right.$$
$$\left. h = \sum_{i=1}^{2} \alpha_i \cdot 3^{2-i},\ \alpha_i \in \{0,1,2\} \right\}.$$

Let us now consider the Cartesian coordinates of the points in the unit square $[0,1]^2$ with their ternary expansion

$$x_i = \sum_{l=1}^{\infty} \epsilon_l^{(i)} \cdot 3^{-l},\ i = 1,2.$$

For any $h \in \{0,1,\ldots,8\}$, $h = \sum_{i=1}^{2} \alpha_i \cdot 3^{2-i}$, all points situated in the interior (with respect to the topology induced by the Euclidean metric) of the square $\phi_h([0,1]^2)$ have as the first ternary digit of their Cartesian coordinate x_i the number α_{2-i}, i.e., $\epsilon_1^{(i)} = \alpha_{2-i}$, for $i = 1,2$.

$\underbrace{0\ 2}_{2}$	$\underbrace{1\ 2}_{5}$	$\underbrace{2\ 2}_{8}$
$\underbrace{0\ 1}_{1}$	$\underbrace{1\ 1}_{4}$	$\underbrace{2\ 1}_{7}$
$\underbrace{0\ 0}_{0}$	$\underbrace{1\ 0}_{3}$	$\underbrace{2\ 0}_{6}$

FIGURE 1

Figure 1 shows the connection between $h \in \{0,1,\ldots,8\}$ and its ternary digits α_1, α_2 (with $h = \alpha_1 \cdot 3 + \alpha_2$).

Remark It is easy to see (for details we refer to [4]) that the IFS $\{[0,1]^2, \phi_i, i = 0,1,2,3,5,6,7,8\}$, with ϕ_i mentioned above, defines the Sierpiński carpet, a fractal which is a planar analogon of the well-known two-thirds Cantor set. For more details about the Sierpiński carpet we refer, e.g., to [2].

Let $\mathbf{r} = (r_0, r_1, \ldots, r_8)$. The *multinomial measure on the unit square* is the probability measure on $[0,1]^2$ denoted by $\mu_{3,\mathbf{r}}$, defined by

$$\mu_{3,\mathbf{r}}\left(\phi_h([0,1)^*)\right) = r_h \cdot \mu_{3,\mathbf{r}}([0,1)^*),\ \text{for } h = 0,1,\ldots,8,$$
$$\mu_{3,\mathbf{r}}\left(\phi_{h_1}(\phi_{h_2}(\ldots(\phi_{h_n}([0,1]^*))))\right) = r_{h_1} \cdot r_{h_2} \ldots r_{h_{n-1}} \cdot \mu_{3,\mathbf{r}}(\phi_{h_n}([0,1)^*)),\ \text{for } n \geq 2,$$

where by $[0,1)^*$ we denote one of the sets $[0,1) \times [0,1)$, $[0,1) \times [0,1]$, $[0,1] \times [0,1)$ and $[0,1] \times [0,1])$ correspondingly to the situation that the right or/and the upper side of the square $\phi_h([0,1]^2)$ lies on the right/upper side of the unit square (analogously to the intervals occurring in the definition of the multinomial measure on the unit interval defined in Section 3).

Remark In the above construction we have considered the iterated functions system $\{[0,1]^2, \phi_h,\ h = 0, 1, \dots, 8\}$ that "divides" the unit square into nine congruent squares of side length 1/3. In fact one can define, for any integer $N \geq 2$ the measure $\mu_{N,\mathbf{r}}$ by taking an analogous IFS $\{[0,1]^2, \phi_h, h = 0, 1, \dots, N^2-1\}$ that "divides" the unit square into N^2 squares of side-length $1/N$ and an N^2-dimensional vector $\mathbf{r} = (r_0, r_1, \dots, r_{N^2-1})$ with $r_h \geq 0$, for $h = 0, 1, \dots, N^2-1$ and $\sum_{h=0}^{N^2-1} r_h = 1$.

All considerations that we make for the multinomial measure $\mu_{3,\mathbf{r}}$ on the unit square can be made in the analogous way for $\mu_{N,\mathbf{r}}$, for any $N \geq 2$. In the following we will restrict our study to the case $N = 3$, as it is straightforward to extend these results for the multinomial measure $\mu_{N,\mathbf{r}}$.

Remark By taking $r_4 = 0$ in the definition of $\mu_{3,\mathbf{r}}$ we get a planar analogon of the Cantor distribution on the unit interval (for details regarding this distribution we refer, e.g., to [10]).

Let us denote by $\mathcal{W}$ the set $\mathcal{W}(3) \times \mathcal{W}(3)$, where $\mathcal{W}(3)$ has already been defined in Section 3 as the set of all infinite words over the alphabet $D = \{0, 1, 2\}$. Let $\mathcal{W}_{m_1,m_2}$ denote the set $\mathcal{W}_{m_1}(3) \times \mathcal{W}_{m_2}(m_3)$, where $\mathcal{W}_{m_i}(3)$, $i = 1, 2$ have the same meaning ($\mathcal{W}_m(q)$ with $m = m_i$ respectively, and $q = 3$) as in the previous section. For every $\omega \in \mathcal{W}$, $\omega = (\omega^1, \omega^2) = (\omega_1^{(1)}\omega_2^{(1)} \dots \omega_k^{(1)} \dots, \omega_1^{(2)}\omega_2^{(2)} \dots \omega_k^{(2)} \dots)$ we define its *value* as the vector

$$\begin{aligned}\mathbf{val}(\omega) &= (val(\omega^{(1)}), val(\omega^{(2)})) \\ &= \left(\sum_{j\geq 1} \omega_j^{(1)} \cdot 3^{-j}, \sum_{j \geq 1} \omega_j^{(2)} \cdot 3^{-j} \right).\end{aligned}$$

Analogously we define the value of any $\omega \in \mathcal{W}_{m_1,m_2}$ as the vector

$$\mathbf{val}(\omega) = \left(\sum_{j=1}^{m_1} \omega_j^{(1)} \cdot 3^{-j}, \sum_{j=1}^{m_2} \omega_j^{(2)} \cdot 3^{-j} \right).$$

4.2. **The moments vector of the multinomial distribution on the unit square.** In the following we approach the moments vector of the function **val** with respect to the distribution defined on the unit square by $\mu_{3,\mathbf{r}}$. Let us first introduce, for any two positive integers n_1, n_2 the notation

$$\mathbf{val}^{(n_1,n_2)} = \left((val(\omega^{(1)}))^{n_1}, (val(\omega^{(2)}))^{n_2} \right),$$

for $\omega = (\omega^{(1)}, \omega^{(2)}) \in \mathcal{W}$ or $\omega = (\omega^{(1)}, \omega^{(2)}) \in \mathcal{W}_{m_1,m_2}$, for some integers $m_1, m_2 \geq 1$.

For the integers $n_1, n_2 \geq 0$ we define the *moment vector of order* (n_1, n_2) with respect to $\mu_{3,\mathbf{r}}$, denoted by $M_{(n_1,n_2)}$,

$$\begin{aligned} M_{(n_1,n_2)} &= \sum_{\omega\in\mathcal{W}} \mu_{3,\mathbf{r}}(\omega)\cdot \mathbf{val}^{(n_1,n_2)}(\omega) \\ &= \sum_{\omega\in\mathcal{W}} \mu_{3,\mathbf{r}}(\omega)\cdot \big((val(\omega^{(1)}))^{n_1}, (val(\omega^{(2)}))^{n_2}\big)\,. \end{aligned}$$

This is equivalent to

$$M_{(n_1,n_2)} = \Big(\sum_{\omega\in\mathcal{W}} \mu_{3,\mathbf{r}}(\omega)\cdot (val(\omega^{(1)}))^{n_1}, \sum_{\omega\in\mathcal{W}} \mu_{3,\mathbf{r}}(\omega)\cdot (val(\omega^{(2)}))^{n_2}\Big) = (M_{(n_1,\cdot)}, M_{(\cdot,n_2)}),$$

where $M_{(n_1,\cdot)}$, $M_{(\cdot,n_2)}$ denote the first and respectively second component of the moments vector of order (n_1, n_2) with respect to $\mu_{3,\mathbf{r}}$.

Let now

$$M^{(m_1,m_2)}_{(n_1,n_2)} = \sum_{\omega\in\mathcal{W}_{m_1,m_2}} \mu_{3,\mathbf{r}}(\omega)\cdot \big((val(\omega^{(1)}))^{n_1}, (val(\omega^{(2)}))^{n_2}\big)\,.$$

We have

$$M_{(n_1,n_2)} = \lim_{\substack{m_1\to\infty\\ m_2\to\infty}} M^{(m_1,m_2)}_{(n_1,n_2)},$$

and

$$M_{(n_1,n_2)} = \left(\lim_{m_1\to\infty} M^{(m_1,\cdot)}_{(n_1,\cdot)}, \lim_{m_2\to\infty} M^{(\cdot,m_2)}_{(\cdot,n_2)}\right),$$

where $M^{(m_1,\cdot)}_{(n_1,\cdot)}$ and $M^{(\cdot,m_2)}_{(\cdot,n_2)}$ are the components of the moments vector $M^{(m_1,m_2)}_{(n_1,n_2)}$.

Let us introduce the following useful notations

$$\begin{aligned} r'_0 &= r_0 + r_1 + r_2, & r'_1 &= r_3 + r_4 + r_5, & r'_2 &= r_6 + r_7 + r_8, \\ r''_0 &= r_0 + r_3 + r_6, & r''_1 &= r_1 + r_4 + r_7, & r''_2 &= r_2 + r_5 + r_8, \end{aligned}$$

i.e.,

$$r'_k = \sum_{l=3k}^{3(k+1)-1} r_l \quad\text{and}\quad r''_k = \sum_{l=0}^{2} r_{k+3l}, \quad \text{for } k = 0, 1, 2.$$

Remark With these notations, it is easy to see that one can reduce the study of $M_{(n_1,\cdot)}$ and $M_{(\cdot,n_2)}$ to the considerations presented in Section 3, as we studied the moments of the multinomial distribution on the unit interval.
We get all information about $M_{(n_1,\cdot)}$ by taking $q = 3$, $\mathbf{r} = \mathbf{r}' = (r'_0, r'_1, r'_2)$ in the definition of $\mu_{q,\mathbf{r}}$ given in Section 3.
The same works for $M_{(\cdot,n_2)}$: here we take $q = 3$, $\mathbf{r} = \mathbf{r}'' = (r''_0, r''_1, r''_2)$ in the above mentioned definition.

With this remark we get, by applying Theorem 3.5, the following

Corollary 1. *The moments vector of order (n_1, n_2) of the multinomial distribution $\mu_{3,\mathbf{r}}$ on the unit square admits, for $n_1 \to \infty$, and $n_2 \to \infty$, the estimate*

$$M_{(n_1,n_2)} =$$
$$\Big(\Phi_1(-\log_3 n_1) \cdot n_1^{\log_3 r_2'}\big(1 + \mathcal{O}(\frac{1}{n_1})\big), \Phi_2(-\log_3 n_2) \cdot n^{\log_3 r_2''}\big(1 + \mathcal{O}(\frac{1}{n_2})\big)\Big), \qquad (4.1)$$

where Φ_1 and Φ_2 are periodic functions having period 1 and known Fourier coefficients.

Special cases.

1. The case $r_1 = 1/8$, for $i = 0, 1, 2, 3, 5, 6, 7, 8$ and $r_4 = 0$.
This case corresponds to the analogon of the (symmetric) Cantor distribution on the (planar) Sierpiński carpet. From Corollary 1 we get:

$$M_{(n_1,n_2)} =$$
$$\Big(\Phi_1(-\log_3 n_1) \cdot n_1^{\alpha-1}\big(1 + \mathcal{O}(\frac{1}{n_1})\big), \Phi_2(-\log_3 n_2) \cdot n_2^{\alpha-1}\big(1 + \mathcal{O}(\frac{1}{n_2})\big)\Big), \qquad (4.2)$$

where $\alpha = \log_3 8$ is the Hausdorff dimension of the Sierpiński carpet and Φ_1 and Φ_2 are as in the above Corollary.

2. The case $r_l = \rho_{\epsilon_0(l)} \cdot \rho_{\epsilon_1(l)}$, for $l \in \{0, 1, \ldots, 8\}$ with the ternary digit expansion $l = \epsilon_0(l) + \epsilon_1(l) \cdot 3$, where $\rho_1, \rho_2, \rho_3 \in (0, 1)$ and $\rho_1 + \rho_2 + \rho_3 = 1$.
The numbers ρ_1, ρ_2, ρ_3 can be seen as probabilities associated to the digits $0, 1, 2$ respectively. It is easy to see that in this case we have, with the above notations, $r_k' = \rho_k$ and $r_k'' = \rho_k$ for $k = 0, 1, 2$ and thus the study of this case easily reduces itself to the 1-dimensional case studied in the previous section (with $q = 3$ and $r_k = \rho_k$ in the definition of the multinomial measure on the unit interval).

5. The Gray code distribution and its moments

The Gray code measure is defined on the unit interval. We call it *Gray code measure* because it is strongly related to the Gray code (see [12], [21]). This measure occurs (not under this name) in connection with the estimation of digital sum functions (in [13]), namely when one is looking for explicit formulae of the power sum and the exponential sum of digital sums of the Gray code representation of natural numbers.

5.1. **A few words about the Gray code.** The Gray code was patented by Frank Gray in 1953, and was a solution to a communications problem involving digitalisation of analogue data. Later on, Gray codes have been used in many other applications (see e.g., [12]). Let us now give a definition of the Gray code:
Let $n \in \mathbb{N}$ have the binary expansion $n = \sum_{k \geq 0} \epsilon_k(n) \cdot 2^k$ with $\epsilon_k(n) \in \{0, 1\}$. The n-th *Gray code* is defined as the infinite sequence

$$(\ldots, \epsilon_{k+1}(n) \oplus \epsilon_k(n), \ldots, \epsilon_2(n) \oplus \epsilon_1(n), \epsilon_1(n) \oplus \epsilon_0(n)),$$

where $\oplus$ denotes the addition modulo 2. Thus, the Gray code is an encoding of the integers as sequences of 0's and 1's. It has the property that representations of adjacent integers differ in exactly one position. The integer n is encoded as the binary representation of $g(n)$ where $g : \mathbb{N} \to \mathbb{N}$ (called the *Gray map*) is defined by

$$g(0) = 0, \;\; g(2^p + j) = 2^p + g(2^p - 1 - j) \quad \text{for } 0 \le j < 2^p.$$

5.2. **The Gray code measure.** In [13] the measure $\tilde{\mu}_r$ (which we call here the *Gray code measure*) is defined in the following way.
Let again $I = I_{0,0} = [0, 1]$ and

$$\begin{aligned} I_{n,j} &= \Big[\frac{j}{2^n}, \frac{j+1}{2^n}\Big), \text{ for } j = 0, 1, \dots, 2^n - 2, \\ I_{n,2^n-1} &= \Big[\frac{2^n - 1}{2^n}, 1\Big], \end{aligned}$$

for $n = 1, 2, 3, \dots$. For each $0 < r < 1$ there exists a unique probability measure $\tilde{\mu}_r$ on I such that

$$\begin{aligned} \tilde{\mu}_r(I_{n+1,2j}) &= \begin{cases} r\tilde{\mu}_r(I_{n,j}) & j: \text{ even}, \\ (1-r)\tilde{\mu}_r(I_{n,j}) & j: \text{ odd}, \end{cases} \\ \tilde{\mu}_r(I_{n+1,2j+1}) &= \begin{cases} (1-r)\tilde{\mu}_r(I_{n,j}) & j: \text{ even}, \\ r\tilde{\mu}_r(I_{n,j}) & j: \text{ odd}, \end{cases} \end{aligned}$$

for $j = 0, 1, \dots, 2^n - 1$ and $n = 0, 1, 2, \dots$.
It is easy to see that for $r = \frac{1}{2}$ the measure $\tilde{\mu}$ coincides with the ordinary Lebesgue measure and otherwise it is a continuous singular measure (for more details we refer to [13]). Let us again denote by $\mathcal{W}$ the set of all infinite words over the alphabet $D = \{0, 1\}$ and by $\mathcal{W}_m$ the set of all words of length m ($m \ge 1$) over the alphabet D. For every word, we have defined (in Section 2) its *value*, $val(\omega) = \sum_{i \ge 1} \omega_i \cdot 2^{-i}$, assigning to any infinite word (the value of) a binary fraction and analogously for the finite words.

5.3. **The moments of the Gray code distribution.** Our next aim is to study the moments of the function val with respect to the distribution defined by $\tilde{\mu}_r$. We have

$$M_n = \sum_{\omega \in \mathcal{W}} \tilde{\mu}_r(\omega) \cdot (val(\omega))^n$$

and

$$M_n = \lim_{m \to \infty} M_n^m,$$

where $M_n^m = \sum_{\omega \in \mathcal{W}_m} \tilde{\mu}_r(\omega) \cdot (val(\omega))^n$.
As in the previous sections we are now looking for a recurrence relation between the moments of different orders. The idea is again to study first the recursive behaviour of the moments of finite words, i.e., M_n^m. In order to do this, we will first do some observations regarding finite words of given length m, where $m \ge 1$ is an arbitrarily fixed integer.
Let us first imagine reading a word $\omega \in \mathcal{W}_m$, $\omega = \omega_1\omega_2 \dots \omega_m$, starting with its first digit, ω_1, ending with the last, ω_n: If $\omega_1 = 0$ then $val(\omega)$ lies in the left elementary interval of level 1 and if $\omega_1 = 1$ it lies in the right elementary interval of level 1. Now we go on reading the second digit, ω_2. As we already know in which interval of level 1 we are, ω_2 indicates whether $val(\omega)$ lies in the left or the right interval of level 2 inside

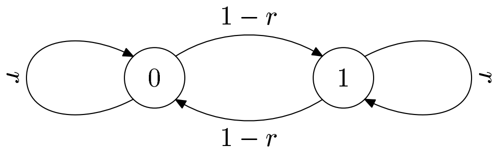

FIGURE 2. A Markov chain model of the problem

interval of level 1 indicated by ω_1. Continuing in this way, the digit ω_k will indicate the position of the k-interval ($k \le m$) that contains $\omega \in \mathcal{W}_m$. Moreover, these positions of the k-intervals (for $1 \le k \le m$) are the same for any $\omega' \in \mathcal{W}$ whose block of the first m digits coincides with ω.

Remark *A Markov chain model of the problem*
Let us now consider the Markov chain with state space $X = \{0, 1\} = D$ and transition probabilities

$$p(x, y) = \begin{cases} r, & x = y, \\ 1 - r, & x \neq y, \end{cases}$$

where $x, y \in \{0, 1\}$ and r is the parameter in the definition of the Gray measure $\tilde{\mu}_r$. We assume that the Markov chain is in the state 0 at the moment of time $t = 0$, with probability 1 and we analyse $X(t)$, i.e., the state of the Markov chain at the moments of time $t = 1, 2, \dots, m, \dots$. For more details regarding Markov chains we refer,e.g., to [22]. Generating finite random words ω (with respect to the distribution $\tilde{\mu}_r$) over the alphabet $D = \{0, 1\}$, i.e., words of $\mathcal{W}_m$, for some $m \ge 1$, is equivalent to the random walk described by the above Markov chain: $X(t)$ for some $t = k \in \mathbb{N}$ indicates the k-th digit of ω, i.e., with the above notations, $X(k) = \omega_k$, $k \ge 1$.

By iterating the matrix of transition probabilities of the Markov chain and applying a known basic property of Markov chains one can easily obtain the following results:

$$\mathbb{P}(X(k) = 0) = \frac{1}{2}\big((2r - 1)^k + 1\big), \ \mathbb{P}(X(k) = 1) = \frac{1}{2}\big(1 - (2r - 1)^k\big),$$

and thus,

$$\mathbb{P}(\omega_k = 0) = \frac{1}{2}\big((2r - 1)^k + 1\big), \ \mathbb{P}(\omega_k = 1) = \frac{1}{2}\big(1 - (2r - 1)^k\big),$$

for $k = 1, 2, \dots$.
Thus the distribution of the k-th digit ($k \ge 1$) of a word $\omega \in \mathcal{W}$ depends in the case of the Gray code distribution explicitly on r and k (excepting the case when $r = \frac{1}{2}$), which was not the case for the binomial measure.

Let us now introduce some useful definitions and notations. For any $\omega \in \mathcal{W}_m$, $\omega = \omega_1\omega_2 \dots \omega_m$ we denote by $\overline{\omega} = \overline{\omega}_1\overline{\omega}_2 \dots \overline{\omega}_m$ the word of $\mathcal{W}_m$ with the property $\overline{\omega}_k = 1 \oplus \omega_k$, for $k = 1, 2, \dots, m$, where $\oplus$ is the addition modulo 2. Analogously, for $m \in \mathcal{W}$: $\overline{\omega} \in \mathcal{W}$, $\overline{\omega} = \overline{\omega}_1 \dots \overline{\omega}_m \dots$ with the property $\overline{\omega}_k = 1 \oplus \omega_k$, for $k = 1, 2, \dots$.
For any integer $k \ge 0$, with the binary representation (of length m) $k = \sum_{j=1}^{m} \epsilon_j(k) \cdot 2^j$,

$\epsilon_j \in \{0,1\}$, $j = 1,2,\ldots,m$ and $0 < r < 1$ we denote by $\pi_{r,m}(k)$ the product

$$r^{m-\tilde{s}(k)} \cdot (1-r)^{\tilde{s}(k)},$$

where $\tilde{s}(k)$ denotes the number of digits 1 occurring in the Gray code $g(k)$ of k, i.e., $\tilde{s}(k)$ is the sum of digits of $g(k)$ ($\tilde{s}$ is called the *Gray digital sum*, see e.g., [13]). Moreover, for any integer $m \geq 1$ and any word $\omega \in \mathcal{W}_m$ we define

$$\pi_r(\omega) := \pi_{r,m}(2^m \cdot val(\omega)).$$

Remark It is easy to show by induction that for any positive integer m, any integer $0 \leq k \leq 2^m - 1$ and any $0 < r < 1$ we have

$$\tilde{\mu}_r(I_{m,k}) = \pi_{r,m}(k).$$

In other words, $\pi_{r,m}(k)$ is the probability that a random word $\omega \in \mathcal{W}$ has for the starting block $\omega_1\omega_2\ldots\omega_m$ of length m, the property $\omega_j = \epsilon_j(k) \in \{0,1\}$ for $j = 1,2,\ldots,m$, where $k = \sum_{j=1}^{m} \epsilon_j(k) \cdot 2^j$. Moreover, with the above notations,

$$\tilde{\mu}_r(\omega_1\omega_2\ldots\omega_m) = \pi_r(\omega), \text{ for any } \omega = \omega_1\omega_2\ldots\omega_m \in \mathcal{W}_m.$$

Let ϕ be the (bijective) function $\phi : \mathcal{W} \to \mathcal{W}$, $\phi(\omega) = \overline{\omega}$. We associate to ϕ for any $m \geq 1$ the obvious (bijective) restriction $\phi_m : \mathcal{W}_m \to \mathcal{W}_m$. It is straightforward to show that $\phi(\phi(\omega)) = \omega$, for all $\omega \in \mathcal{W}$, as well as $\phi_m(\phi_m(\omega)) = \omega$, for $\omega \in \mathcal{W}_m$, and any $m \geq 1$. This function will play a role in the deduction of the recurrence for the moments M_n defined earlier.
We have

$$\begin{aligned}
M_n^m &= \sum_{\omega \in \mathcal{W}_m} \pi_r(\omega) \cdot (val(\omega))^n \\
&= \frac{r}{2^n} \sum_{\omega' \in \mathcal{W}_{m-1}} \pi_r(\omega') \cdot (val(\omega'))^n + \frac{1-r}{2^n} \sum_{\omega' \in \mathcal{W}_{m-1}} \pi_r(\omega') \cdot (1 + val(\overline{\omega'}))^n.
\end{aligned}$$

Let us now also consider the moments $\overline{M}_n$ and $\overline{M}_n^m$ of the composed function $val \circ \phi$, with respect to the Gray code distribution:

$$\begin{aligned}
\overline{M}_n^m &= \sum_{\omega \in \mathcal{W}_m} \tilde{\mu}_r(\omega) \cdot (val(\phi(\omega)))^n = \sum_{\omega \in \mathcal{W}_m} \pi_r(\omega) \cdot (val(\overline{\omega}))^n, \\
\overline{M}_n &= \sum_{\omega \in \mathcal{W}} \tilde{\mu}_r(\omega) \cdot (val(\phi(\omega)))^n = \sum_{\omega \in \mathcal{W}} \tilde{\mu}_r(\omega) \cdot (val(\overline{\omega}))^n.
\end{aligned}$$

We have

$$\overline{M}_n = \lim_{m \to \infty} \overline{M}_n^m.$$

With $\overline{r} = 1 - r$ the above relations yield

$$\begin{aligned}
M_n^m &= \frac{r}{2^n} \sum_{\omega' \in \mathcal{W}_{m-1}} \pi_r(\omega') \cdot (val(\omega'))^n + \frac{\overline{r}}{2^n} \sum_{j=0}^{n} \binom{n}{j} \sum_{\omega' \in \mathcal{W}_{m-1}} \pi_r(\omega') \cdot (val(\overline{\omega'}))^j \\
&= \frac{r}{2^n} \cdot M_n^{m-1} + \frac{\overline{r}}{2^n} \sum_{j=0}^{n} \binom{n}{j} \overline{M}_j^{m-1}, \qquad (5.1)
\end{aligned}$$

and

$$\begin{aligned}\overline{M}_n^m =&\frac{r}{2^n}\sum_{\omega'\in\mathcal{W}_{m-1}}\pi_r(\omega')\cdot(1+val(\phi(\phi(\omega'))))^n+\frac{\overline{r}}{2^n}\sum_{\omega'\in\mathcal{W}_{m-1}}\pi_r(\omega')\cdot(val(\phi(\omega')))^n\\ =&\frac{r}{2^n}\sum_{\omega'\in\mathcal{W}_{m-1}}\pi_r(\omega')\cdot(1+val(\omega'))^n+\frac{\overline{r}}{2^n}\sum_{\omega'\in\mathcal{W}_{m-1}}\pi_r(\omega')\cdot(val(\overline{\omega'}))^n,\end{aligned}$$

which produces

$$\overline{M}_n^m=\frac{\overline{r}}{2^n}\cdot\overline{M}_n^{m-1}+\frac{r}{2^n}\sum_{j=0}^{n}\binom{n}{j}\cdot M_j^{m-1}. \tag{5.2}$$

By taking the limit for $m\to\infty$ in (5.1) and (5.2) we obtain the recurrences:

$$M_n=\frac{r}{2^n}M_n+\frac{\overline{r}}{2^n}\sum_{j=0}^{n}\binom{n}{j}\overline{M}_j, \tag{5.3}$$

$$\overline{M}_n=\frac{\overline{r}}{2^n}\overline{M}_n+\frac{r}{2^n}\sum_{j=0}^{n}\binom{n}{j}M_j. \tag{5.4}$$

It is easy to verify that $M_0=1$ and $\overline{M}_0=M_0=1$. Thus we have proven the following

Theorem 5. *The moments of the Gray distribution $\tilde{\mu}_r$ satisfy the relations:*

$$M_0=\overline{M}_0=1,$$

$$M_n=\frac{r}{2^n}M_n+\frac{\overline{r}}{2^n}\sum_{j=0}^{n}\binom{n}{j}\overline{M}_j,$$

$$\overline{M}_n=\frac{\overline{r}}{2^n}\overline{M}_n+\frac{r}{2^n}\sum_{j=0}^{n}\binom{n}{j}M_j,$$

for all integers $n\geq 1$.

As in the previous sections, one can use these recursion relations in order to compute a list of the first few moments $M_1, M_2,\ldots$, and $\overline{M}_1,\overline{M}_2,\ldots$.

5.4. **The asymptotics of the moments M_n.** The recursions (5.3) and (5.4) enable us to study the asymptotic behaviour of the moments M_n, for $n\to\infty$. We will proceed analogously to the previous cases. Let us first define the exponential generating functions

$$A(z)=\sum_{n\geq 0}M_n\frac{z^n}{n!},\ \ B(z)=\sum_{n\geq 0}\overline{M}_n\frac{z^n}{n!}.$$

From the above recursions we get

$$\begin{aligned}A(z)&=r\cdot A(\frac{z}{2})+\overline{r}\cdot e^{\frac{z}{2}}\cdot B(\frac{z}{2}),\\ B(z)&=\overline{r}\cdot B(\frac{z}{2})+r\cdot e^{\frac{z}{2}}\cdot A(\frac{z}{2}),\end{aligned}$$

and thus for the *Poisson transformed functions* $\widehat{A}(z) = A(z)\cdot e^{-z}$ and $\widehat{B}(z) = B(z)\cdot e^{-z}$,

$$\widehat{A}(z) = \overline{r}\cdot \widehat{B}(\frac{z}{2}) + r\cdot e^{-\frac{z}{2}}\cdot \widehat{A}(\frac{z}{2}), \tag{5.5}$$

$$\widehat{B}(z) = r\cdot \widehat{A}(\frac{z}{2}) + \overline{r}\cdot e^{-\frac{z}{2}}\cdot \widehat{B}(\frac{z}{2}). \tag{5.6}$$

Herefrom we have

$$\widehat{A}(z) = \overline{r}r\cdot \widehat{A}(\frac{z}{4}) + \overline{r}^2\cdot e^{-\frac{z}{4}}\cdot \widehat{B}(\frac{z}{4}) + r\cdot e^{-\frac{z}{2}}\cdot \widehat{A}(\frac{z}{2}). \tag{5.7}$$

In order to study the asymptotics of the moments M_n, we are going to study the behaviour of $\widehat{A}(z)$ for $z\to\infty$, as $M_n \sim \widehat{A}(n)$, which can be again justified by using depoissonisation. First let us rewrite (5.7) as

$$\widehat{A}(z) = \overline{r}r\widehat{A}(\frac{z}{4}) + H(z), \tag{5.8}$$

where $H(z) = \overline{r}^2\cdot e^{-\frac{z}{4}}\cdot \widehat{B}(\frac{z}{4}) + r\cdot e^{-\frac{z}{2}}\cdot \widehat{A}(\frac{z}{2})$ is an auxiliary function which we presume to be known. We compute the Mellin transform $\widehat{A}^*(s)$ of the function $\widehat{A}(z)$ in (5.8). We obtain

$$\widehat{A}^*(s) = \overline{r}r\cdot 4^s\cdot \widehat{A}^*(s) + H^*(s) = \frac{H^*(s)}{1-\overline{r}r\cdot 4^s}.$$

Now we can obtain the function $\widehat{A}(z)$ by applying the Mellin inversion formula, i.e.,

$$\widehat{A}(z) = \frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} \widehat{A}^*(s)\cdot z^{-s}ds = \frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} \frac{H^*(s)}{1-\overline{r}r\cdot 4^s}\cdot z^{-s}ds, \tag{5.9}$$

where $0 < c < \log_4\frac{1}{\overline{r}r}$ and $H^*(s) = \int_0^\infty H(z)\cdot z^{s-1}dz$.
We estimate $\widehat{A}(z)$ in (5.9) by shifting the integral to the right and taking the residues with negative sign. The function under the integral has simple poles at $s_k = \log_4\frac{1}{\overline{r}r} + \frac{2k\pi\mathrm{i}}{\log 4}$, $k\in\mathbb{Z}$. Its residues with negative sign at s_k are

$$\frac{1}{\log 4}H^*\Big(\log_4\frac{1}{\overline{r}r} + \frac{2k\pi\mathrm{i}}{\log 4}\Big)z^{-\log_4\frac{1}{\overline{r}r}-\frac{2k\pi\mathrm{i}}{\log 4}}.$$

For $k=0$ the residue with negative sign is

$$\frac{1}{\log 4}\cdot z^{\log_4\overline{r}}\int_0^\infty \Big(\overline{r}^2\cdot e^{-\frac{z}{4}}\cdot B(\frac{z}{4}) + r\cdot e^{-\frac{z}{2}}\cdot \widehat{A}(\frac{z}{2})\Big)\cdot z^{\log_4\frac{1}{\overline{r}r}-1}dz$$

and plays an important role in the asymptotics of the nth moment M_n of the Gray code distribution.
We get an estimate for M_n by summing up all mentioned residues into a periodic function. With the above notations we have

Theorem 6. *The nth moment M_n of the Gray code distribution $\tilde{\mu}_r$ admits the asymptotic estimate*

$$M_n = \Phi(-\log_4 n)\cdot n^{\log_4(\overline{r}r)}\Big(1+\mathcal{O}(\frac{1}{n})\Big), \tag{5.10}$$

for $n\to\infty$, where $\Phi(x)$ is a periodic function having period 1 and known Fourier coefficients. The mean (zeroth Fourier coefficient) is given by the expression

$$\frac{1}{\log 4}\int_0^\infty \Big(\overline{r}^2\cdot e^{-\frac{z}{4}}\cdot \widehat{B}(\frac{z}{4}) + r\cdot e^{-\frac{z}{2}}\cdot \widehat{A}(\frac{z}{2})\Big)\cdot z^{\log_4\frac{1}{\overline{r}r}-1}dz. \tag{5.11}$$

Remark One can numerically compute this integral by taking for $\widehat{A}(\frac{z}{2})$ and $\widehat{B}(\frac{z}{4})$ the first terms of their Taylor expansions. These terms can be found from the recurrences (5.3) and (5.4). By rewriting (5.11) as

$$\frac{1}{\log 4}\bigg(\frac{\overline{r}^2}{\sqrt{\overline{r}r}}\sum_{k\geq 0}\frac{\overline{M}_k}{2^k k!}\cdot\Gamma\Big(k+\log_4\frac{1}{\overline{r}r}\Big)+r\cdot\sum_{k\geq 0}\frac{M_k}{2^k k!}\Gamma\Big(k+\log_4\frac{1}{\overline{r}r}\Big)\bigg)$$

we obtain, for $r=0.4$, $M_{100}=0.009601\ldots$, whereas the value predicted by (5.10) (without the oscillation) is $0.009695\ldots$.

References

[1] F. Bassino and H. Prodinger. (q,δ)-numeration systems with missing digits. *Monatshefte für Mathematik*, 141:89–99, 2004.

[2] G. David and S. Semmes. *Fractured fractals and broken dreams*, volume 7 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press Oxford University Press, New York, 1997. Self-similar geometry through metric and measure.

[3] H. Delange. Sur la fonction sommatoire de la fonction"somme des chiffres". *Enseignement Math. (2)*, 21(1):31–47, 1975.

[4] K.J. Falconer. *Fractal Geometry. Mathematical Foundations and Applications*. John Wiley & Sons, Chichester, 1990.

[5] H.-J. Fischer. On the paper: "Asymptotics for the moments of singular distributions" [J. Approx. Theory **74** (1993), no. 3, 301–334; MR 94h:44002] by W. Goh and J. Wimp. *J. Approx. Theory*, 82(3):362–374, 1995.

[6] Ph. Flajolet, X. Gourdon, and Ph. Dumas. Mellin transforms and asymptotics: Harmonic sums. *Theoret. Comput. Sci.*, 144:3–58, 1995.

[7] Ph. Flajolet, P. Grabner, P. Kirschenhofer, H. Prodinger, and R. F. Tichy. Mellin transforms and asymptotics: digital sums. *Theoret. Comput. Sci.*, 123(2):291–314, 1994.

[8] W. Goh and J. Wimp. Asymptotics for the moments of singular distributions. *J. Approx. Theory*, 74(3):301–334, 1993.

[9] W. Goh and J. Wimp. A generalized Cantor-Riesz-Nagy function and the growth of its moments. *Asymptotic Analysis*, 8:379–392, 1994.

[10] P. J. Grabner and H. Prodinger. Asymptotic analysis of the moments of the Cantor distribution. *Statist. Probab. Lett.*, 26(3):243–248, 1996.

[11] Ph. Jacquet and W. Szpankowski. Analytical de-Poissonization and its applications. *Theoret. Comput. Sci.*, 201(1-2):1–62, 1998.

[12] D. Knuth. *The Art of Computer Programming, Volume 4, Pre-Fascicle 2a: Generating all n-tuples (version of 29 Aug 2003).* http://www-cs-faculty.stanford.edu/ knuth/news.html, 2004.

[13] Z. Kobayashi. Digital Sum Problems for the Gray Code Representation of Natural Numbers. *Interdisciplinary Information Sciences*, 8:167–175, 2002.

[14] K. Muramoto, T. Okada, T. Sekiguchi, and Yasunobu Shiota. Digital sum problems for the p-adic expansion of natural numbers. *Interdiscip. Inform. Sci.*, 6(2):105–109, 2000.

[15] K. Muramoto, T. Okada, T. Sekiguchi, and Yasunobu Shiota. An explicit formula of subblock occurrences for the p-adic expansion. *Interdiscip. Inform. Sci.*, 8(1):115–121, 2002.

[16] T. Okada, T. Sekiguchi, and Y. Shiota. Applications of Binomial Measures to Power Sums of Digital Sums. *Journal of Number Theory*, 52:256–266, 1995.

[17] T. Okada, T. Sekiguchi, and Y. Shiota. An explicit formula of the exponential sums of digital sums. *Japan J. Indust. Appl. Math.*, 12(3):425–438, 1995.

[18] T. Okada, T. Sekiguchi, and Y. Shiota. A Generalization of Hata-Yamaguti's Results on the Takagi Function II: Multinomial Case. *Japan J. Indust. Appl. Math.*, 13:435–463, 1996.

[19] Y. Shiota, T. Sekiguchi, and T. Okada. Hata-Yamaguti's result on [the] Takagi function and its applications to digital sum problems. *Sūrikaisekikenkyūsho Kōkyūroku*, (961):73–80, 1996. Analytic number theory (Japanese) (Kyoto, 1995).

[20] W. Szpankowski. *Average case analysis of algorithms on sequences.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2001. With a foreword by Philippe Flajolet.
[21] E. Weisstein. *Mathworld.* http://mathworld.wolfram.com/GrayCode.html.
[22] W. Woess. *Random walks on infinite graphs and groups*, volume 138 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge, 2000.

Ligia-Loretta Cristea, Graz University of Technology, Institute for Mathematics, Steyrergasse 30/ III, 8010-Graz, Austria, and Department of Mathematics, Politehnica University of Timişoara, Romania

E-mail address: cristea@finanz.math.tu-graz.ac.at

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@maths.wits.ac.za