COMBINATORIAL AND ARITHMETICAL PROPERTIES OF LINEAR NUMERATION SYSTEMS

PETER J. GRABNER†, PETER KIRSCHENHOFER‡, AND ROBERT F. TICHY*

Dedicated to the memory of Pál Erdős

Abstract. We extend a result of J. Alexander and D. Zagier on the Garsia entropy of the Erdős measure. Our investigation heavily relies on methods from combinatorics on words. Furthermore, we introduce a new singular measure related to the Farey tree.

1. Introduction

This paper is devoted to the investigation of multiplicities of representations and related combinatorial and probabilistic questions for a special class of linear numeration systems. Non-uniqueness of radix expansions was extensively studied by Pál Erdős and his coauthors, see for instance [14].

We will study numeration systems given by a linear recurring base sequence

$$G_{n+m} = G_{n+m-1} + \cdots + G_n \quad \text{for } n \geq 0 \tag{1.1}$$

$$G_k = G_{k-1} + \cdots + G_0 + 1 \quad \text{for } 0 \leq k < m. \tag{1.2}$$

Any positive integer n can be represented in a digital expansion

$$n = \sum_{\ell=0}^{L} \delta_\ell(n) G_\ell \tag{1.3}$$

with digits $\delta_\ell\{0,1\}$ for $0 \leq \ell \leq L$, where the digits are computed by the greedy algorithm: there is a unique integer L such that $G_L \leq n < G_{L+1}$. Then n can be written as $n = \delta_L G_L + n_L$ with $0 \leq n_L < G_L$ and by iterating this procedure with n_L the expansion (1.3) is obtained. An extensive description of digital expansions with respect to linear recurring base sequences is given in [15, 19, 20, 21]. In [19], especially dynamical properties of such expansions are investigated. The corresponding shift transformation is the classical β-shift investigated by A. Rényi and W. Parry [33, 35].

Recurrence (1.1) has the property that its dominating characteristic root $\beta > 1$ satisfying

$$\beta^m = \beta^{m-1} + \cdots + \beta + 1 \quad m \geq 2$$

Date: July 5, 2002.

1991 *Mathematics Subject Classification.* Primary: 11K55; Secondary: 05A17, 11A63, 28D20.

Key words and phrases. digital representations, entropy, generating functions.

† This author is supported by the START-project Y96-MAT of the Austrian Science Fund.

‡ This author is supported by the Austrian Science Fund (FWF) grant P14200-MAT.

* This author is supported by the Austrian Science Fund (FWF) grant S8307-MAT.

is a PV-number, i. e. β is an algebraic integer all of whose conjugates lie inside the unit circle. This is a particular case of a classical result of A. Brauer [3] on polynomials with decreasing coefficients.

In the present paper we will deal with questions related to the investigation of the distribution $\mu_{\beta,b}$ of sums of the form

$$\sum_{n=1}^{\infty} \delta_n \beta^{-n},$$

where δ_n are independent identically b-distributed random variables taking values in the set $\{0, 1, \dots, \lceil\beta\rceil - 1\}$. Certainly, this distribution is obtained as an infinite convolution of the measures $b(\beta^n dx)$, where b is a probability measure on the set of digits. For instance $\beta = 2$ and $b(\{0\}) = b(\{1\}) = \frac{1}{2}$ yields the Lebesgue measure, and $\beta = 3$ and $b(\{0\}) = b(\{2\}) = \frac{1}{2}$ describes the classical Cantor measure. By a theorem of Jessen and Wintner [24] $\mu_{\beta,b}$ is either atomic, or purely singular, or absolutely continuous. P. Erdős [12] proved that for $\beta = \frac{1+\sqrt{5}}{2}$, the golden ratio, and $b(\{0\}) = b(\{1\}) = \frac{1}{2}$, the uniform distribution, $\mu_{\beta,b}$ is purely singular. A year later [13] he proved that there is a $\gamma > 1$ such that for almost all $\beta < \gamma$ the measure $\mu_{\beta,b}$, for b the uniform distribution, is absolutely continuous; B. Solomyak [40] has proved that one can take $\gamma = 2$, which is clearly best possible. We refer to the survey [34] for further details and more recent developments. In [6] functional equations satisfied by the density function are investigated. In a recent paper Lalley [28] considered digit distributions b with $b(\{0\}) = p$ and $b(\{1\}) = 1 - p$ $(0 < p < 1)$. In this case he proved for $1 < \beta < 2$ a PV-number that $\mu_{\beta,b}$ is again purely singular.

In our instance b is the equidistribution on the digit set $\{0, 1\}$; we denote the corresponding measure by μ_β. From the classical proof of Erdős [12] it follows that μ_β is purely singular whenever β is a PV-number [37]. Garsia [17] introduced a new concept in the study of μ_β, which is now called Garsia entropy. He considered the finite convolutions $b(\beta dx) * b(\beta^2 dx) * \cdots * b(\beta^N dx)$. Let G_N^β be the support of this atomic measure which assigns weight $p(x)$ to a point x, and define the Garsia entropy by

$$H(G_N^\beta) = -\sum_{x\in G_N^\beta} p(x) \log_2 p(x). \tag{1.4}$$

Then set

$$H_\beta = \lim_{N\to\infty} \frac{H(G_N^\beta)}{N \log_2 \beta}. \tag{1.5}$$

Garsia proved that $H_\beta < 1$ implies that μ_β is purely singular (for $\beta < 2$). Furthermore, he observed that $H_\beta < 1$ for any PV-number β; however, his arguments do not yield the numerical value of H_β.

Alexander and Zagier [1] found a constructive approach in the case of the golden ratio $\beta = \frac{1+\sqrt{5}}{2}$, which leads to sharp bounds for H_β in that case. In their proof they make use of a graph-theoretic encoding of the different representations of a given number. The corresponding graph is called the Fibonacci graph. We remark here that Garsia's entropy is the entropy of the random walk on this graph in the sense of Avez [2], Kaĭmanovich and Vershik [25]. By the procedure studied in [1] this graph can be reduced in a well defined manner to the so called Euclidean tree. This is the complete binary tree with each node associated a pair (a, b) of labels in the following way: the root is labelled $(2, 1)$ and, inductively, given a node at level n labelled with (a, b) its two successors at level $n+1$ are labelled $(a+b, a)$ resp. $(a+b, b)$. Thus this tree corresponds to the subtractive Euclidean algorithm.

In Section 2 of the present paper we extend this constructive approach to the dominating characteristic root $\beta = \beta_m \in (1, 2)$ of the so-called m-bonacci recurrence (1.1) satisfying

$$\beta^m = \beta^{m-1} + \cdots + \beta + 1. \tag{1.6}$$

The associated m-bonacci graph is a planar graph whose combinatorial structure turns out to be considerably more complicated than in the special instance $m = 2$ discussed above. Therefore our discussion will avoid this graph theoretic description and will mainly use the language of 0,1-sequences. We make use of well-suited generating functions and the method of Guibas and Odlyzko for counting 0,1-strings with forbidden subwords (cf. [22]).

Our main result is the following theorem, which will be proved in Section 2.

Theorem 1. *Let $e(k, i)$ denote the number of steps in the subtractive Euclidean algorithm applied to the pair (k, i) (i. e. $e(i, i) = 0$ and $e(k+i, i) = e(k, i) + 1$). Let furthermore*

$$\kappa_n = \sum_{\substack{0<i<k \\ \gcd(k,i)=1 \\ e(k,i)=n}} k \log_2 k \tag{1.7}$$

and

$$\mathcal{T}(x) = 1 - \frac{1}{2}\left(\frac{1-3x}{1-x}\right)^2 \sum_{n=1}^{\infty} \kappa_n x^n. \tag{1.8}$$

If $\beta = \beta_m$ is the PV-number fulfilling $\beta^m = \beta^{m-1} + \cdots + \beta + 1$, then the constant H_{β_m} in the asymptotic leading term of the Garsia entropy (cf. (1.5)) is given by

$$H_{\beta_m} = \frac{1}{\log_2 \beta_m} \mathcal{T}(2^{-m}). \tag{1.9}$$

In Section 3 a measure related to Minkowski's singular function is defined. This measure is used later to derive precise error bounds for the numerical values of the entropy.

In Section 4 estimates for the numerical values of H_{β_m} are deduced from this theorem.

Section 5 is devoted to the discussion of connections between the Euclidean tree and the Farey tree and related singular measures, see also [27]. This gives interesting connections to Minkowski's $?(x)$-function. We furthermore refer to other arithmetically defined singular functions, such as Takagi's function, which can be used for describing digital sum problems such as (5.1).

2. Proof of the Theorem

We start with the remark that because of (1.6) and the fundamental observations of A. Rényi and W. Parry [33, 35] in general a given real number will have several radix expansions in base $\beta = \beta_m$, which are encoded by 0,1-sequences. In the following we call two finite 0,1-sequences of equal length $N \geq 0$ equivalent, if they represent the same number in base β. The size of the corresponding equivalence classes will be called *frequency* in this paper. We will use the notation $\varphi(w)$ for the frequency of an equivalence class represented by a word $w \in \{0,1\}^*$. Let $F_N(k)$ denote the number of classes of 0,1-sequences of length N with frequency k. Since the weights in the definition of $H(G_N^\beta)$ (cf. (1.4)) are 2^{-N} times the frequencies, we have

$$H(G_N^\beta) = -\sum_{k=1}^{\infty} kF_N(k)2^{-N}\log_2\frac{k}{2^N} = N - \sum_{k=1}^{\infty} kF_N(k)2^{-N}\log_2 k. \tag{2.1}$$

For convenience we introduce the generating functions of the multiplicities $F_N(k)$

$$\begin{aligned} f_k(x) &= \sum_{N=0}^{\infty} F_N(k)x^N, \\ \Phi(x,s) &= \sum_{k=1}^{\infty} k^s f_k(x). \end{aligned} \tag{2.2}$$

We observe that

$$\begin{aligned} \Phi(x,1) &= \sum_{N=0}^{\infty} 2^N x^N = \frac{1}{1-2x}, \\ \left.\frac{\partial\Phi(x,s)}{\partial s}\right|_{s=1} &= \sum_{\substack{k\ge1\\ N\ge1}} kF_N(k)(\log k)x^N. \end{aligned} \tag{2.3}$$

Therefore the generating function of the quantities (2.1) is given by

$$\mathcal{H}(x) = \sum_{N=0}^{\infty} H(G_N^\beta)x^N = \frac{x}{(1-x)^2} - \frac{1}{\log 2}\left.\frac{\partial\Phi(x/2,s)}{\partial s}\right|_{s=1}. \tag{2.4}$$

In order to analyze the functions $f_k(x)$ we investigate the structure of the equivalence classes of 0,1-words in detail. Observe that two strings are equivalent, if they can be transformed into each other by a finite number of replacements of the subblock $01^{(m)}$ by $10^{(m)}$ and vice versa. We call these blocks the *characteristic blocks*. The class of a 0,1-string, which ends with either of these two blocks, is called *relational*.

Remark 1. *Note that* $\varphi(\delta_1\dots\delta_{n-1}\delta_n) = \varphi(\delta_1\dots\delta_{n-1})$, $\delta_i \in \{0,1\}$, *if and only if* $\delta_1\dots\delta_{n-1}\delta_n$ *is not in a relational class. Furthermore, for a relational class* w *with representatives* $w_1 0$ *and* $w_2 1$ *the frequency satisfies* $\varphi(w) = \varphi(w_1) + \varphi(w_2)$.

We denote all (possibly empty) sequences, which do not contain the subblocks $0^{(m)}$ resp. $1^{(m)}$, *m-free*. Note that all classes of 0,1-strings of frequency ≥ 2 can be generated from the relational classes by appending an m-free 0,1-string. Furthermore, the classes of frequency 1 are generated by appending an m-free string to any of the sequences from $(0^{(m)})^* = \{\varepsilon, 0^{(m)}, 0^{(2m)}, \dots\}$ resp. $(1^{(m)})^*$ (where ε denotes the empty string).

The generating function of m-free strings can be computed immediately by the method of Guibas and Odlyzko (cf. [22, 30]) to be

$$S(x) = \frac{1-x^m}{1-2x+x^m}. \tag{2.5}$$

With this notation we have (by the preceding paragraph)

$$f_1(x) = \frac{1+x^m}{1-x^m} S(x) = \frac{1+x^m}{1-2x+x^m}, \tag{2.6}$$

and

$$f_k(x) = S(x) r_k(x), \quad k \geq 2, \tag{2.7}$$

where $r_k(x)$ denotes the generating function of all strings in relational classes of frequency k.

We proceed by analyzing the relational classes in further detail. For each relational class there is a unique "shortest relational prefix", i. e. a relational class whose representatives are prefixes of elements of the given class and whose length is minimal with respect to this property. These are exactly the relational classes of frequency 2.

In the following we will make use of a simple observation:

Remark 2. *Every relational class has a representative of the form* $0\dots 0^{(m)}$ *or* $1\dots 1^{(m)}$ *except for the class of strings equivalent to* $01^{(m)}$ *and* $10^{(m)}$. *Similarly, it has a representative of the form* $1\dots 0^{(m)}$ *or* $0\dots 1^{(m)}$.

The generating function of all shortest relational prefixes can be determined as follows: each of these classes has a representative of the form 0^*w, with $w = 01\dots 0^{(m)}$ or 1^*w with $w = 10\dots 1^{(m)}$, where w does not contain the subblocks $0^{(m)}$ and $1^{(m)}$ (except for the end); furthermore, $01^{(m)}$ is a shortest relational prefix. But we have to take care of the fact that among those strings there are also strings that do not correspond to a shortest relational prefix, namely the strings of the form 0^*w, with $w = 01\dots 0^{(m-1)}10^{(m)}$, and the strings 1^*w with $w = 10\dots 1^{(m-1)}01^{(m)}$. The latter strings are in bijective correspondence with the strings $01\dots 0^{(m)}1^{(m)}$ and $10\dots 1^{(m)}0^{(m)}$ by using the replacement rule $10^{(m)} \mapsto 01^{(m)}$. In order to count all words as described above, we introduce the generating function of all m-free strings, which begin and end with the same digit, $S_e(x)$; similarly, we define $S_d(x)$ as the generating function of m-free strings with different digits at the beginning and the end. Then we have

$$1 + S_e(x) + S_d(x) = S(x) \quad \text{and } S_d(x) = S_e(s)(x + x^2 + \cdots + x^{m-1})$$

and by inserting the second equation into the first one

$$\frac{1-x^m}{1-x} S_e(x) + 1 = S(x). \tag{2.8}$$

All strings of the form $01\dots 0^{(m)}$ and $10\dots 1^{(m)}$ can be obtained by adding one digit at the beginning and m digits at the end of a string counted by $S_e(x)$. Furthermore, by the above mentioned bijection, removing all strings of the form $01\dots 0^{(m-1)}10^{(m)}$ and $10\dots 1^{(m-1)}01^{(m)}$

amounts a factor of $(1-x^m)$ in the generating function, and adding 0^* or 1^* in the beginning a further $\frac{1}{1-x}$-factor. Thus by (2.8)

$$x^{m+1}\frac{1-x^m}{1-x}S_e(x)+x^{m+1}=x^{m+1}S(x)$$

is the generating function of all shortest relational prefixes, where the $+x^{m+1}$-term corresponds to the string $01^{(m)}$.

Observe that the deletion of the shortest relational prefix is an invariant (shift) operation on the set of strings. Thus we have

$$f_k(x)=S(x)r_k(x)=S(x)^2 l_k(x), \tag{2.9}$$

where $l_k(x)$ is the generating function of all relational classes of frequency k having a representative of the form $w=10^{(m)}w'$. In the following we will give a more detailed description of these words.

Remark 3. *For every such relational class we can find a representative of the form*

$$10^{(m)}e_1z_1e_2z_2\dots z_{s-1}e_s$$

where $e_\ell\in\{0^{(m)},1^{(m)}\}^*$ *and* z_ℓ *is a word starting with* 0 *and ending with* $10^{(m)}$ *but not with* $0^{(m-1)}10^{(m)}$ *resp. a word starting with* 1 *and ending with* $01^{(m)}$ *but not with* $1^{(m-1)}01^{(m)}$ *resp. the word* $10^{(m)}$ *(and* m*-free except for the explicitly mentioned occurrences).*

Now we have to take care of the frequency of a word $w=10^{(m)}e_1z_1e_2z_2\dots e_s$. For this purpose we write $z_\ell=z'_\ell v_\ell$ with $v_\ell\in\{10^{(m)},01^{(m)}\}$, so that

$$w=(10^{(m)}e_1)z'_1(v_1e_2)z'_2\dots z'_{s-1}(v_{s-1}e_s). \tag{2.10}$$

Observe that the intermediate subblocks z'_i cannot be altered by the replacement rule $01^{(m)}\leftrightarrow 10^{(m)}$ and are therefore unique. Thus the representatives for the classes of the subblocks $10^{(m)}e_i$ can be chosen independently. Altogether we have ($v_0=10^{(m)}$)

$$\varphi((10^{(m)}e_1)z'_1(v_1e_2)z'_2\dots z'_{s-1}(v_{s-1}e_s))=\prod_{\ell=1}^{s}\varphi(v_{\ell-1}e_\ell). \tag{2.11}$$

In order to compute the frequencies $\varphi(v_{\ell-1}e_\ell)$ for $e_\ell\in\{0^{(m)},1^{(m)}\}^*$ we write $e_\ell=\eta_1\eta_2\dots\eta_n$, where $\eta_j\in\{0^{(m)},1^{(m)}\}$. We introduce two auxiliary functions φ_1 and φ_2 defined by

$$\begin{aligned}
\varphi_1(v\eta_1\eta_2\dots\eta_n)&=\varphi_1(v\eta_1\dots\eta_{n-1})+\varphi_2(v\eta_1\dots\eta_{n-1}),\\
\varphi_2(v\eta_1\eta_2\dots\eta_n)&=\begin{cases}\varphi_1(v\eta_1\dots\eta_{n-1}) & \text{for } \eta_{n-1}\neq\eta_n\\ \varphi_2(v\eta_1\dots\eta_{n-1}) & \text{for } \eta_{n-1}=\eta_n,\end{cases}\\
\varphi_1(v)=1,&\quad \varphi_2(v)=1
\end{aligned} \tag{2.12}$$

for $v\in\{01^{(m)},10^{(m)}\}$. Then

$$\varphi(v\eta_1\eta_2\dots\eta_n)=\varphi_1(v\eta_1\eta_2\dots\eta_n)+\varphi_2(v\eta_1\eta_2\dots\eta_n). \tag{2.13}$$

This follows from the fact that for a given representative of a relational class w there exist two words $w_1,w_2\in\{0,1\}^*$ such that w_10 and w_21 are equivalent to w. The frequency $\varphi(w)$

is clearly given by the sum of the frequencies of the words w_1 and w_2 by Remark 1. In order to compute these frequencies, we have to find w_1 and w_2: without loss of generality let us consider a word $w = 10^{(m)}\eta_1 \dots \eta_n$ with $\eta_i \in \{0^{(m)}, 1^{(m)}\}$ and $\eta_n = 0^{(m)}$. Assume furthermore that $\eta_{n-k-1} \neq \eta_{n-k} = \eta_{n-k+1} = \cdots = \eta_n$. Then w_1 and w_2 are given by $w_1 = 10^{(m)}\eta_1 \dots \eta_{n-1}0^{(m-1)}$ and $w_2 = 10^{(m)}\eta_1 \dots \eta_{n-k-2}(1^{(m-1)}0)^{(k+1)}1^{(m-1)}$, where w_2 can be found by repeated application of the replacement rule $10^{(m)} \mapsto 01^{(m)}$ (because of symmetry for $\eta_n = 1^{(m)}$ the same transformations can be made with interchanging 0 and 1). The frequency of w_1 is equal to the frequency of $10^{(m)}\eta_1 \dots \eta_{n-1}$ and the frequency of w_2 is equal to the frequency of $10^{(m)}\eta_1 \dots \eta_{n-k-2}$, since by Remark 1 we have

$$\varphi(10^{(m)}\eta_1 \dots \eta_{n-k-2}(1^{(m-1)}0)^{(k-2)}1^{(m-1)}) = \varphi(10^{(m)}\eta_1 \dots \eta_{n-k-2}(1^{(m-1)}0)^{(k-3)}1^{(m-1)}) = \cdots = \varphi(10^{(m)}\eta_1 \dots \eta_{n-k-2}1^{(m-1)}) = \varphi(10^{(m)}\eta_1 \dots \eta_{n-k-2}).$$

By the definition of φ_1 (2.12) $\varphi_1(w) = \varphi(w_1)$. Furthermore, for $k = 0$ in the above argument we have $\varphi(w_2) = \varphi(10^{(m)}\eta_1 \dots \eta_{n-2}) = \varphi_1(10^{(m)}\eta_1 \dots \eta_{n-1})$. If $k > 0$, which is $\eta_{n-1} = \eta_n$, we have $\varphi(w_2) = \varphi(10^{(m)}\eta_1 \dots \eta_{n-k-2})$; since by Remark 1 this implies that none of the words $\eta_1 \dots \eta_{n-\ell-2}$ for $\ell = 0, \dots, k$ is relational, we have

$$\varphi(w_2) = \varphi(10^{(m)}\eta_1 \dots \eta_{n-k-2}) = \varphi_1(10^{(m)}\eta_1 \dots \eta_{n-k-1}) = \varphi_2(10^{(m)}\eta_1 \dots \eta_{n-k}) = \cdots = \varphi_2(10^{(m)}\eta_1 \dots \eta_{n-1}).$$

By the above construction a labelled complete binary tree can be defined as follows: every node at level $\ell \geq 0$ can be represented by a string of length ℓ: $\eta_1 \dots \eta_\ell$, and the two successors are given by $\eta_1 \dots \eta_\ell 0^{(m)}$ and $\eta_1 \dots \eta_\ell 1^{(m)}$. Every node e is labelled by the pair $(\varphi_1(ve), \varphi_2(ve))$. By the rules given in (2.12) the two successors of a node labelled with a pair (a, b) has successors labelled with $(a+b, a)$ and $(a+b, b)$. From this rule it is clear that the labels satisfy $\gcd(a, b) = 1$. Thus, starting with root labelled $(1, 1)$ at level 0, the labels of the nodes in the ℓ-th level of this tree are all pairs (a, b) with $\gcd(a, b) = 1$ which need ℓ steps of the subtractive Euclidean algorithm to compute the gcd. In this way we define for any coprime a and b a function by setting $e(a, b) = \ell$. This tree is called Euclidean tree.

Now we encode the above combinatorial descriptions in generating functions. Let $\alpha_k(x)$ denote the generating functions of all strings $v\delta_1 \dots \delta_n$ corresponding to nodes with frequency $k \geq 2$ in the Euclidean tree (obviously, n has to be a multiple of m). This function is given by

$$\alpha_k(x) = x^{m+1} \sum_{\substack{0<i<k/2 \\ \gcd(k,i)=1}} x^{me(k-i,i)} = \sum_{\substack{0<i<k \\ \gcd(k,i)=1}} x^{1+me(k,i)}, \quad k \geq 2, \tag{2.14}$$

where the factor x^{m+1} corresponds to the prefix v above.

The generating function $g(x)$ for the words z'_i can be determined as follows: first observe that $x^{m+1}g(x)$ is the generating function of the words z_i. The z_i are formed from all m-free strings which begin and end with the same digit ε by adding between 1 and $m-1$ digits $1-\varepsilon$ in the beginning and m digits $1-\varepsilon$ in the end. This gives the generating function $(x+\cdots+x^{m-1})S_e(x)x^m$. Then we have to remove all strings ending with $0^{(m-1)}10^{(m)}$ or $1^{(m-1)}01^{(m)}$; furthermore, we have to add the string $10^{(m)}$ and subtract the strings $0^{(m-1)}10^{(m)}$ and

$1^{(m-1)}01^{(m)}$, which are produced additionally by the above procedure. This gives for the generating function

$$x^{m+1}g(x) = (x + \cdots + x^{m-1})S_e(x)x^m + x^{m+1} - 2x^{2m}$$

and using (2.8)

$$g(x) = \frac{1 - 2x^{m-1} + x^m}{1 - 2x + x^m} = (1 - x^{m-1})S(x) - x^{m-1}. \tag{2.15}$$

By (2.11) the generating function $l_k(x)$ satisfies the recurrence relation

$$\begin{aligned} l_k(x) &= \sum_{\substack{d|k \\ d\neq 1,k}} \alpha_d(x) l_{\frac{k}{d}}(x) g(x) + \alpha_k(x), \quad k \geq 2 \\ l_1(x) &= 1. \end{aligned} \tag{2.16}$$

We introduce the following Dirichlet generating functions

$$\begin{aligned} \mathcal{A}(x,s) &= \sum_{k=2}^{\infty} k^s \alpha_k(x) g(x) \\ \mathcal{L}(x,s) &= 1 + \sum_{k=2}^{\infty} k^s l_k(x) g(x). \end{aligned} \tag{2.17}$$

Because of (2.16) we have

$$\mathcal{L}(x,s) = \frac{1}{1 - \mathcal{A}(x,s)}. \tag{2.18}$$

Alternatively, the last relation can be obtained by observing that $\mathcal{L}$ is the generating function of the "sequence construction" (cf. [38]) applied to the set $\{10^{(m)}ez'\}$ where e and z are as in (2.10).

In order to evaluate $\mathcal{H}(x)$ in (2.4) we need $\frac{\partial \Phi}{\partial s}$. By (2.9), (2.17), and (2.18) we have

$$\begin{aligned} \frac{\partial \Phi}{\partial s}(x,1) &= \frac{S(x)^2}{g(x)} \frac{\partial \mathcal{L}}{\partial s}(x,1) \\ &= \frac{S(x)^2}{g(x)} \frac{1}{(1 - \mathcal{A}(x,1))^2} \frac{\partial \mathcal{A}}{\partial s}(x,1) \\ &= \frac{S(x)^2}{g(x)} \mathcal{L}(x,1)^2 \frac{\partial \mathcal{A}}{\partial s}(x,1). \end{aligned} \tag{2.19}$$

Now we use (2.3):

$$\frac{1}{1-2x} = \Phi(x,1) = f_1(x) + \frac{S(x)^2}{g(x)} \left(\mathcal{L}(x,1) - 1\right).$$

Inserting for $S(x)$, $g(x)$ and $f_1(x)$ we obtain

$$S(x)\mathcal{L}(x,1) = \frac{g(x)}{S(x)}\frac{1}{1-2x} - \frac{g(x)}{S(x)}f_1(x) + S(x) = \frac{1-3x^m}{(1-x^m)(1-2x)}. \tag{2.20}$$

Using (2.20) we obtain from (2.19)

$$\frac{\partial\Phi}{\partial s}(x,1) = \frac{(1-3x^m)^2}{(1-x^m)^2(1-2x)^2}\frac{1}{g(x)}\frac{\partial\mathcal{A}}{\partial s}(x,1) = \frac{(1-3x^m)^2}{(1-x^m)^2(1-2x)^2}a(x),$$

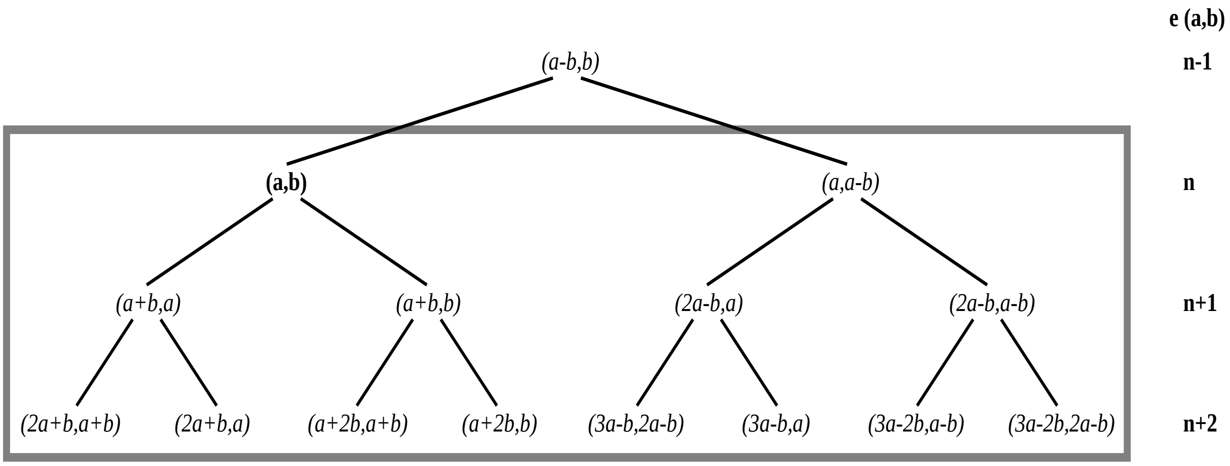
where (by (2.14))

$$a(x) = \sum_{\substack{0<i<k \\ \gcd(k,i)=1}} x^{1+me(k,i)} k\log_2 k = \sum_{N=1}^{\infty} \kappa_N x^{mN+1}.$$

Inserting the last result in (2.4) yields the theorem.

3. A singular measure related to Minkowski's function

In order to give precise estimates for the κ_n's we study the sequence $\nu_1 = \kappa_1 = 2$, $\nu_2 = \kappa_2 - 6\kappa_1 = 6\log_2 3 - 12$, $\nu_{n+2} = 9\kappa_n - 6\kappa_{n+1} + \kappa_{n+2}$ for $n \geq 1$, which are the coefficients of the function $(1-3x)^2\sum_{n=1}^{\infty}\kappa_n x^n$. For the computation of the ν_n we look at the local structure of the Euclidean tree:



We collect the "boxed" terms in the above picture to obtain

$$\nu_{n+2} = \sum_{\substack{0<b<a\\ e(a,b)=n\\ \gcd(a,b)=1}} af\left(\frac{b}{a}\right)$$

with

$$f(x) = x\log_2\left(\frac{(2-x)^6(2+x)(1+2x)^2}{(3-2x)^2(3-x)(1+x)^6}\right) + \log_2\left(\frac{(3-2x)^3(3-x)^3(2+x)^2(1+2x)}{(2-x)^{12}(1+x)^6}\right).$$

Thus ν_n can be written as an integral with respect to a linear combination of point measures:

$$\nu_{n+2} = 3^n\int_0^1 f(x)\,dF_n(x),$$

where

$$F_n(t) = 3^{-n}\sum_{\substack{0<b<a\\ e(a,b)=n\\ \gcd(a,b)=1\\ \frac{b}{a}<t}} a.$$

The normalizing factor 3^{-n} is chosen in order to obtain weak convergence of F_n. For the proof of this fact we need an auxiliary sequence of distribution functions

$$G_n(t) = 3^{-n}\sum_{\substack{0<b<a\\ e(a,b)=n\\ \gcd(a,b)=1\\ \frac{b}{a}<t}} b.$$

Now it is easy to see that

$$\begin{aligned} F_{n+1}(t) &= \frac{1}{3}\left(F_n\left(\frac{t}{1-t}\right) + G_n\left(\frac{t}{1-t}\right)\right) \\ G_{n+1}(t) &= \frac{1}{3}G_n\left(\frac{t}{1-t}\right) \end{aligned} \qquad \text{for } t \le \frac{1}{2} \tag{3.1}$$

and

$$\begin{aligned} F_{n+1}(t) &= 2 - \frac{1}{3}\left(F_n\left(\frac{1-t}{t}+0\right) + G_n\left(\frac{1-t}{t}+0\right)\right) \\ G_{n+1}(t) &= 1 - \frac{1}{3}F_n\left(\frac{1-t}{t}+0\right) \end{aligned} \qquad \text{for } t > \frac{1}{2}; \tag{3.2}$$

the initial distributions are given by

$$F_0(t) = \begin{cases} 0 & \text{for } t \le \frac{1}{2} \\ 2 & \text{for } t > \frac{1}{2} \end{cases}$$

$$G_0(t) = \begin{cases} 0 & \text{for } t \le \frac{1}{2} \\ 1 & \text{for } t > \frac{1}{2} \end{cases}$$

Using the matrices

$$M_A = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad M_B = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$

writing $\vec{F}_n(t) = (F_n(t), G_n(t))^{\mathrm{T}}$ and $\vec{v} = (2,1)^{\mathrm{T}}$ this can be rewritten as

$$\vec{F}_{n+1}(t) = \begin{cases} M_A \vec{F}_n(\frac{t}{1-t}) & \text{for } t \le \frac{1}{2} \\ \vec{v} - M_B \vec{F}_n(\frac{1-t}{t} + 0) & \text{for } t > \frac{1}{2}. \end{cases} \tag{3.3}$$

Notice that this recurrence is rather similar to the functional equation of Minkowski's $?(x)$-function:

$$?(x) = \begin{cases} \frac{1}{2}?\left(\frac{x}{1-x}\right) & \text{for } x \le \frac{1}{2} \\ 1 - \frac{1}{2}?\left(\frac{1-x}{x}\right) & \text{for } x > \frac{1}{2}. \end{cases}$$

Now similar arguments as those used in [36] to obtain an exact formula for $?(x)$ in terms of the continued fraction expansion of x can be used to obtain an expression for $\vec{F}_n(t)$. Let $0 < t < 1$ be given by its continued fraction expansion $t = [a_1, a_2, \ldots]$ and let k and ℓ be given by $a_1 + \cdots + a_k + \ell = n$ with $0 \le \ell < a_{k+1}$. Then $\vec{F}_n(t)$ can be computed from (3.3)

$$\begin{aligned} \vec{F}_n([a_1, a_2, \ldots]) = \sum_{m=1}^{k} (-1)^{m-1} M_A^{a_1-1} M_B M_A^{a_2-1} M_B \cdots M_A^{a_m-1} \vec{v} \\ + (-1)^k M_A^{a_1-1} M_B \cdots M_A^{a_k-1} M_B M_A^{\ell} \vec{F}_0([a_{k+1} - \ell, a_{k+2}, \ldots]). \end{aligned} \tag{3.4}$$

We note now that the matrix product can be given in terms of the convergents of t:

$$M_A^{a_1-1} M_B M_A^{a_2-1} M_B \cdots M_A^{a_m-1} = 3^{-(a_1+\cdots+a_m-1)} \begin{pmatrix} q_{m-1} & q_m - q_{m-1} \\ p_{m-1} & p_m - p_{m-1} \end{pmatrix}.$$

Furthermore, the last summand is non-zero only if $a_{k+1} - \ell = 1$. Thus we obtain

$$\begin{aligned} F_n([a_1, a_2, \ldots]) &= \sum_{m=1}^{k} (-1)^{m-1} 3^{-(a_1+\cdots+a_m-1)} (q_m + q_{m-1}) \\ &\quad + (-1)^k 3^{-(a_1+\cdots+a_{k+1}-1)} (q_{k+1} + q_k) \delta_{\ell, a_{k+1}-1} \\ G_n([a_1, a_2, \ldots]) &= \sum_{m=1}^{k} (-1)^{m-1} 3^{-(a_1+\cdots+a_m-1)} (p_m + p_{m-1}) \\ &\quad + (-1)^k 3^{-(a_1+\cdots+a_{k+1}-1)} (p_{k+1} + p_k) \delta_{\ell, a_{k+1}-1}. \end{aligned}$$

Now we define

$$\begin{aligned} F([a_1, a_2, \ldots]) &= \sum_{k=1}^{\infty} (-1)^{k-1} 3^{-(a_1+\cdots+a_k-1)} (q_k + q_{k-1}) \quad q_0 = 1 \\ G([a_1, a_2, \ldots]) &= \sum_{k=1}^{\infty} (-1)^{k-1} 3^{-(a_1+\cdots+a_k-1)} (p_k + p_{k-1}) \quad p_0 = 0; \end{aligned} \tag{3.5}$$

these series are easily seen to converge, since $p_k, q_k < \vartheta^{a_1+\cdots+a_k}$ ($\vartheta = \frac{1+\sqrt{5}}{2}$, cf. [36]). Furthermore, these functions are continuous, which can be proved along the same lines as the continuity of $?(x)$. Finally, $dG(x) = x dF(x)$ by definition.

In [1], formula (3.4), the following rigorous bounds for the $\mu_n = \frac{1}{2}(\kappa_{n+1} - 3\kappa_n)$ are established

$$\log_2 \frac{3}{2} < \frac{\mu_n}{3^{n-1}} < \frac{2}{3}. \tag{3.6}$$

The integral

$$\int_0^1 f(x)\, dF(x) \tag{3.7}$$

has to be 0, since otherwise $3^{-n}\nu_n$ would have a nonzero limit, which would imply that κ_n would be of order of magnitude $n^2 3^n$; this contradicts (3.6).

Using the fact that the integral (3.7) vanishes we derive

$$3^{-n}\nu_{n+2} = \int_0^1 f(x)\, dF_n(x) = \int_0^1 (F(x) - F_n(x))\, df(x). \tag{3.8}$$

From the rate of convergence of the series (3.5) it is clear that

$$|F(x) - F_n(x)| = \mathcal{O}\left(\left(\frac{1+\sqrt{5}}{6}\right)^n\right);$$

in the following we will work out a much better bound which also answers a question posed in [1] p. 133: "There is no reason to expect them (the coefficients) to be positive and decreasing to zero, although the table indicates that they are. We can prove only that λ_n (the Taylor coefficients of $\mathcal{T}(x)$ in (1.8)) is $\mathcal{O}((\frac{4}{3})^{\frac{n}{2}})$."

We study the centralized moments of the distributions F_n:

$$m_n^{(k)} = \int_0^1 \left(x - \frac{1}{2}\right)^k dF_n(x) = 3^{-n} \sum_{\substack{0<b<a \\ (a,b)=1 \\ e(a,b)=n}} a \left(\frac{b}{a} - \frac{1}{2}\right)^k; \tag{3.9}$$

it is clear from our previous knowledge on the functions F_n that $m_n^{(0)} = 2$, $m_n^{(2k+1)} = 0$ for all $n \geq 0$, and $m_0^{(2k)} = 0$ for $k \geq 1$. Thus it suffices to study the even moments.

In order to get precise information on the rate of convergence of the $m_n^{(2k)}$'s for $n \to \infty$ we derive a recurrence formula for these moments. Since a pair (a, b) at level n produces two pairs $(a+b,a)$ and $(a+b,b)$ at level $n+1$, we have the following recursion

$$
\begin{aligned}
m_{n+1}^{(2k)} &= 3^{-n-1}\sum_{\substack{0<b<a\\(a,b)=1\\e(a,b)=n}}(a+b)\left(\left(\frac{a}{a+b}-\frac{1}{2}\right)^{2k}+\left(\frac{b}{a+b}-\frac{1}{2}\right)^{2k}\right)\\
&= 3^{-n-1}\sum_{\substack{0<b<a\\(a,b)=1\\e(a,b)=n}}a\left(\frac{3}{2}+\left(\frac{b}{a}-\frac{1}{2}\right)\right)\left(\left(\frac{a}{a+b}-\frac{1}{2}\right)^{2k}+\left(\frac{b}{a+b}-\frac{1}{2}\right)^{2k}\right).
\end{aligned}
\tag{3.10}
$$

Now we expand the summand into its Taylor series around $\frac{1}{2}$ to obtain (for $k\geq 1$)

$$
m_{n+1}^{(2k)} = 2\cdot 6^{-2k}+\sum_{\ell=1}^{\infty}b_{2\ell}^{(2k)}m_n^{(2\ell)}
\tag{3.11}
$$

with

$$
b_{2\ell}^{(2k)} = 2^{-2k}\left(\frac{2}{3}\right)^{2\ell}\sum_{r=0}^{2k}\binom{2k}{r}\binom{2\ell+r-2}{2k-2}3^{-r}.
\tag{3.12}
$$

Now the computation of the moments $m_n^{(2k)}$ can be viewed as an iteration of a linear map with positive coefficients. In order to make this sensible we have to introduce the space on which the iteration has to be performed:

$$
\begin{aligned}
\tilde{\ell}^{\infty} &= \left\{(x^{(2)},x^{(4)},\ldots)\mid \sup_k |x^{(2k)}|2^{2k}<\infty\right\}\\
\|(x^{(2)},x^{(4)},\ldots)\| &= \sup_k |x^{(2k)}|2^{2k}.
\end{aligned}
\tag{3.13}
$$

A simple computation shows that

$$
2^{2k}\sum_{t=1}^{\infty}b_{2t}^{(2k)}2^{-2t}=\frac{1}{3}-3^{-2k},
$$

which implies that the norm of the linear operator defined by $b_{2\ell}^{(2k)}$ is $\frac{1}{3}$. Furthermore, by positivity of this operator, the sequences $m_n^{(2k)}$ are monotonically increasing to the moments of dF which we denote by $m^{(2k)}$. Using the error estimate in Banach's fixed point theorem we obtain

$$
0\leq m^{(2k)}-m_n^{(2k)}\leq 2^{-2k}3^{-n-1}.
$$

We use the Taylor expansion of $f(x)$ around $x=\frac{1}{2}$:

$$
f(x)=\log_2\frac{409600000}{387420489}-\frac{8}{5\log 2}\left(x-\frac{1}{2}\right)^2+\sum_{k=2}^{\infty}f_{2k}\left(x-\frac{1}{2}\right)^{2k}\quad\text{with } f_{2k}>0\qquad\text{for } k\geq 2.
$$

Putting everything together $3^{-n}\nu_n$ can be bounded:

$$
\begin{aligned}
&-0.00151\ldots\cdot 3^{-n} = \frac{f(0)-f(\frac{1}{2})-\frac{1}{4}f_2}{3}3^{-n}\\
&=\log_2\frac{409600000}{387420489}m^{(0)}+\frac{8}{5\log 2}m^{(2)}+\sum_{k=2}^{\infty}f_{2k}\left(m^{(2k)}-2^{-2k}3^{-n-1}\right)\\
&\le\int_0^1 f(x)\,dF_n(x)\\
&\le\log_2\frac{409600000}{387420489}m^{(0)}-\frac{8}{5\log 2}\left(m^{(2)}-\frac{1}{12}3^{-n}\right)+\sum_{k=2}^{\infty}f_{2k}m^{(2k)}=\frac{2}{15\log 2}3^{-n}.
\end{aligned}
$$

Therefore we have

$$-0.00151\ldots\cdot\le\nu_n\le\frac{2}{15\log 2}=0.192359\ldots\quad\text{for } n\ge 3. \tag{3.14}$$

4. Bounds for the Entropy

The function $\mathcal{T}(x)$ introduced in (1.8) was extensively studied in section 4 of [1]. In particular it follows from these studies that the constant H_{β_m} may be expressed as

$$H_{\beta_m}\log_2\beta_m = 1-\frac{1}{2}\left(\frac{1-3\cdot 2^{-m}}{1-2^{-m}}\right)^2\sum_{n=1}^{\infty}\kappa_n 2^{-mn} \tag{4.1}$$

$$= 1-\frac{1-3\cdot 2^{-m}}{(1-2^{-m})^2}\sum_{n=1}^{\infty}\mu_n 2^{-mn} \tag{4.2}$$

$$= 1-\frac{1}{(1-2^{-m})^2}\sum_{n=1}^{\infty}\nu_n 2^{-mn}, \tag{4.3}$$

where $\mu_n=\frac{1}{2}(\kappa_n-3\kappa_{n-1})$.

If (4.3) is truncated after N terms, the error $E_{N,m}$ can bounded using (3.14)

$$-0.00151\cdot\frac{2^{-m(N+1)}}{(1-2^{-m})^3}\le E_{N,m}\le\frac{1}{15\log 2}\frac{2^{-m(N+1)}}{(1-2^{-m})^3}.$$

By computing 26 coefficients κ_n the following table of numerical values for H_{β_m} $(2\le m\le 13)$ can be established.

m	H_{β_m}
2	0.995713126685555
3	0.980409319534731
4	0.986926474333800
5	0.992585300274171
6	0.996032591584967
7	0.997937445507094
8	0.998944915449832
9	0.999465368055570
10	0.999730606878347
11	0.999864704467762
12	0.999932181983893
13	0.999966043207405
14	0.999983008336978
15	0.999991500519328

It is clear from (4.1) and the fact that

$$\beta_m = 2 - m2^{-m} + \mathcal{O}(2^{-m})$$

that $H_{\beta_m} \to 1$ (exponentially) for $m \to \infty$. This reflects the observation that the dependence between digits becomes weaker for larger values of m.

5. Concluding Remarks

In this section we want to indicate some further problems of arithmetical and dynamical nature concerning linear numeration systems. The first kind of such problems is related to the asymptotic behaviour of digital sums. The summatory function of the sum-of-digits function and even more generally completely G-additive functions are studied in a series of papers (cf. [20, 21, 9, 10, 11, 7, 8]). Note that a G-additive function is an arithmetic function satisfying the equation

$$f\left(\sum_{k=0}^{K} \delta_k G_k\right) = \sum_{k=0}^{K} f(\delta_k G_k),$$

whereas an arithmetic function is called completely G-additive, if it satisfies

$$f\left(\sum_{k=0}^{K} \delta_k G_k\right) = \sum_{k=0}^{K} f(\delta_k).$$

For instance, numeration systems generated by substitutions over finite alphabets are studied in [10]. It is shown there that for a given completely G-additive function f the following asymptotic formula holds

$$\sum_{n<N} f(n)^k = CN(\log N)^k + \sum_{\ell=1}^{k} N(\log N)^{k-\ell} F_{\ell,k}(\log N) + o(N), \tag{5.1}$$

where $F_{\ell,k}$ denote continuous periodic functions. For the special case of the summatory function of the sum-of-digits with respect to the Fibonacci expansion (i.e. $G_{n+2} = G_{n+1}+G_n$, $G_0 = 1$, $G_1 = 2$) such a result is due to Coquet and van den Bosch [5]. More general expansions in connection with formal languages and substitutions were investigated in [16, 29, 23].

In the case of binary and more generally q-adic radix expansions the moments of the sum-of-digits function were studied in [4]. However, the continuity of the corresponding fluctuation functions $F_{\ell,k}$ in (5.1) remained an open problem for $k \geq 3$ until it could be proved in the special case $k = 3$ in [18] and in general in [31, 32]. The proof of continuity in the general case is based on a new approach using Takagi's singular function related to the so-called binomial measure (cf. [41]).

A singular function corresponding to the subtractive Euclidean algorithm is the Minkowski $?(x)$-function. This function is defined by an invariant measure related to the map

$$x \mapsto \begin{cases} \frac{x}{1+x} & \text{for } 0 \leq x \leq \frac{1}{2} \\ \frac{1}{1+x} & \text{for } \frac{1}{2} < x \leq 1. \end{cases}$$

We note here that based on (3.5) a general class of related functions can be introduced. An important tool for the study of these functions is the Farey-tree which is a graph of the same structure as the Euclidean tree; the only difference being that the nodes of a fixed level are permuted. For detailed information we refer to [27, 42, 43]. In particular, various dynamical properties of the above shift map are investigated. For classical papers in this direction we refer to [26, 36].

In a remarkable recent paper N. Sidorov and A. Vershik [39] studied the Erdős measure with the help of a dynamic approach including the investigation of various properties of the so called golden shift related to the Fibonacci number system ($\beta = \frac{1+\sqrt{5}}{2}$). They consider $\Sigma = \{0,1\}^{\mathbb{N}}$ (equipped with the infinite product λ' of the uniform distribution on the factors) and independent random variables $E_k : \Sigma \to \{0,1\}$, the k-th projections. They define a map $L' : \Sigma \to [0,1]$ given by

$$L'(\varepsilon_1, \varepsilon_2, \ldots) = \sum_{k=1}^{\infty} \varepsilon_k \beta^{-k-1}.$$

Let $X \subset \Sigma$ denote the shift-invariant subspace

$$X = \{(\varepsilon_1, \varepsilon_2, \ldots) \mid \forall k \geq 1 (\varepsilon_k, \varepsilon_{k+1}, \ldots, \varepsilon_{k+m-1}) <_{\texttt{lex}} (a_0, \ldots, a_{m-1})\}$$

with respect to the one-sided shift $\tau(\varepsilon_1, \varepsilon_2, \ldots) = (\varepsilon_2, \varepsilon_3, \ldots)$. Furthermore, they consider the map $L : X \to [0,1]$ defined by $L(\varepsilon_1, \varepsilon_2, \ldots) = \sum_{k=1}^{\infty} \varepsilon_k \beta^{-k}$ and set $T = L\tau L^{-1}$. It is easy to see that $T(x) = \{\beta x\}$, and this transformation on $[0,1]$ is known as the β-shift, cf. [33]. Parry has found the unique T-invariant measure π which is equivalent to Lebesgue measure. Observe that λ' is τ-invariant on Σ and $L^{-1}(\pi)$ is τ-invariant on X. In these terms the Erdős measure μ_β can be written as $L'(\lambda')$. Now using the normalization mapping $\Sigma \to X$, which identifies distinct digital representations of one number, an application of the ergodic theorem proves that the Erdős measure is singular with respect to Lebesgue measure. For more details in the we refer to [39]. It would be very interesting to extend this approach to more general linear numeration systems. However, it seems that specific properties of the golden ratio influence the method significantly.

Acknowledgement. We are indebted to two anonymous referees for many valuable remarks concerning the presentation of the paper.

References

[1] J. C. Alexander and D. B. Zagier. The entropy of certain infinitely convolved Bernoulli measures. *J. London Math. Soc.*, 44:121–134, 1991.
[2] A. Avez. Théorème de Choquet-Deny pour les groupes à croissance non exponentielle. *C. R. Acad. Sci. Paris Sér. A*, 279:25–28, 1974.
[3] A. Brauer. On algebraic equations with all but one root in the interior of the unit circle. *Math. Nachr.*, 4:250–257, 1951.
[4] J. Coquet. Power sums of digital sums. *J. Number Th.*, 22:161–176, 1986.
[5] J. Coquet and P. van den Bosch. A summation formula involving Fibonacci digits. *J. Number Th.*, 22:139–146, 1986.
[6] G. Derfel and R. Schilling. Spatially chaotic configurations and functional equations with rescaling. *J. Phys. A*, 29(15):4537–4547, 1996.
[7] J.-M. Dumont. Formules sommatoires et systèmes de numération lies aux substitutions. *Séminaires de Théorie des Nombres, Bordeaux*, 1987/88.
[8] J.-M. Dumont. Summation formulae for substitutions on a finite alphabet. In J.-M. Luck, P. Moussa, and M. Waldschmidt, editors, *Number Theory and Physics*, volume 47 of *Proc. Physics*, pages 185–194. Springer, Berlin, 1990.
[9] J.-M. Dumont and A. Thomas. Digital sum problems and substitutions on a finite alphabet. *J. Number Th.*, 39:351–366, 1991.
[10] J.-M. Dumont and A. Thomas. Digital sum moments and substitutions. *Acta Arith.*, 64:205–225, 1993.
[11] J.-M. Dumont and A. Thomas. Gaussian asymptotic properties of the sum-of-digits function. *J. Number Th.*, 62:19–38, 1997.
[12] P. Erdős. On a family of symmetric Bernoulli convolutions. *Amer. J. Math.*, 61:974–976, 1939.
[13] P. Erdős. On the smoothness properties of a family of Bernoulli convolutions. *Amer. J. Math.*, 62:180–186, 1940.
[14] P. Erdős, M. Joo, and I. Joo. On a problem of Tamas Varga. *Bull. Soc. Math. Fr.*, 120:507–521, 1992.
[15] A. S. Fraenkel. Systems of numeration. *Amer. Math. Monthly*, 92:105–114, 1985.
[16] C. Frougny. Representation of numbers and finite automata. *Math. Systems Theory*, 25:37–60, 1992.
[17] A. M. Garsia. Entropy and singularity of infinite convolutions. *Pacific J. Math.*, 13:1159–1169, 1963.
[18] P. J. Grabner, P. Kirschenhofer, H. Prodinger, and R. F. Tichy. On the moments of the sum-of-digits function. In G.E. Bergum, editor, *Applications of Fibonacci Numbers*, volume 5, pages 263–271. Kluwer, 1993.
[19] P. J. Grabner, P. Liardet, and R. F. Tichy. Odometers and systems of numeration. *Acta Arith.*, 70:103–123, 1995.
[20] P. J. Grabner and R. F. Tichy. Contributions to digit expansions with respect to linear recurrences. *J. Number Th.*, 36:160–169, 1990.
[21] P. J. Grabner and R. F. Tichy. α-expansions, linear recurrences and the sum-of-digits function. *Manuscr. Math.*, 70:311–324, 1991.
[22] L. J. Guibas and A. M. Odlyzko. String overlaps, pattern matching, and nontransitive games. *J. Comb. Theory, Ser. A*, 30:183–208, 1981.
[23] M. Hollander. Greedy numeration systems and recognizability. *Theory Comput. Syst.*, 1998. to appear.
[24] B. Jessen and A. Wintner. Distribution functions and the Riemann zeta function. *TRans. Amer. Math. Soc.*, 38:48–88, 1938.

[25] V. A. Kaĭmanovich and A. M. Vershik. Random walks on discrete groups: boundary and entropy. *Ann. Probab.*, 11:457–490, 1983.

[26] J.R. Kinney. Note on a singular function of Minkowski. *Proc. Am. Math. Soc.*, 11:788–794, 1960.

[27] J.C. Lagarias. Number theory and dynamical systems. In *The unreasonable effectiveness of number theory*, volume 46 of *Proc. Symp. Appl. Math.*, pages 35–72. Amer. Math. Soc., 1992.

[28] S. P. Lalley. Random series in powers of algebraic integers: Hausdorff dimension of the limit distribution. *J. London Math. Soc.*, 57:629–654, 1998.

[29] N. Loraud. β-shift, systèmes de numération et automates. *J. de Theorie des Nombres Bordeaux*, 7:473–498, 1995.

[30] A.M. Odlyzko. Enumeration of strings. In A. Apostolico and Z. Galil, editors, *Combinatorial algorithms on words*, number F 12 in NATO ASI Ser., pages 205–228. Proc. NATO Adv. Res. Workshop, Maratea, Italy 1984, Springer, 1985.

[31] T. Okada, T. Sekiguchi, and Y. Shiota. Applications of binomial measures to power sums of digital sums. *J. Number Theory*, 52:256–266, 1995.

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

[33] W. Parry. On the β-expansions of real numbers. *Acta Math. Acad. Sci. Hung.*, 11:401–416, 1960.

[34] Y. Peres, W. Schlag, and B. Solomyak. Sixty years of Bernoulli convolutions. In C. Bandt, S. Graf, and M. Zähle, editors, *Fractal Geometry and Stochastics II*, number 46 in Progress in Probability, pages 39–65. Birkhäuser, 1999.

[35] A. Rényi. Representation for real numbers and their ergodic properties. *Acta Math. Acad. Sci. Hung.*, 8:477–493, 1957.

[36] R. Salem. On some singular monotonic functions which are strictly increasing. *Trans. Am. Math. Soc.*, 53:427–439, 1943.

[37] R. Salem. *Albebraic numbers and Fourier analysis.* Heath, Lexington, 1963.

[38] R. Sedgewick and P. Flajolet. *An Introduction to the Analysis of Algorithms.* Addison-Wesley, Amsterdam, 1996.

[39] N. Sidorov and A. Vershik. Ergodic properties of the Erdős measure, the entropy of the goldenshift, and related problems. *Mh. Math.*, 126:215–261, 1998.

[40] B. Solomyak. On the random series $\sum \pm\lambda^n$ (an Erdős problem). *Ann. of Math. (2)*, 142(3):611–625, 1995.

[41] T. Takagi. A simple example of the continuous function without derivative. *Phys.-Math. Soc. Japan*, 1:176–177, 1903. The Collected Papers of T. Takagi, S. Kuroda ed., Iwanami (1973), 5–6.

[42] R. F. Tichy and J. Uitz. An extension of Minkowski's singular function. *Appl. Math. Lett.*, 8:39–46, 1995.

[43] J. Uitz. Chaotische dynamische Systeme und Minkowski Funktionen. Master's thesis, Technical University Graz, 1995.

(P. Grabner & R. Tichy) Institut für Mathematik A, Technische Universität Graz, Steyrergasse 30, 8010 Graz, Austria

E-mail address: {grabner, tichy}@weyl.math.tu-graz.ac.at

(P. Kirschenhofer) Institut für Mathematik und Angewandte Geometrie, Montanuniversität Leoben, 8700 Leoben, Austria

E-mail address: kirsch@unileoben.ac.at