ON THE NUMBER OF OPTIMAL BASE 2 REPRESENTATIONS OF INTEGERS

PETER J. GRABNER AND CLEMENS HEUBERGER‡

Abstract. We study representations of integers n in binary expansions using the digits $0, \pm 1$. We analyze the average number of such representations of minimal "weight" (= number of non-zero digits). The asymptotic main term of this average involves a periodically oscillating function, which is analyzed in some detail. The main tool is the construction of a measure on $[-1, 1]$, which encodes the number of representations.

1. Introduction

In many public key cryptosystems, raising one or more elements of a given group to large powers plays an important role (cf. for instance [6, 13]). In practice, the underlying groups are often chosen to be the multiplicative group of a finite field $\mathbb{F}_q$ or the group law of an elliptic curve (elliptic curve cryptosystems).

Let P be an element of a given group, whose group law will be written additively. What we need is to form nP for large $n \in \mathbb{N}$ in a short amount of time. One way to do this is the *binary method* (cf. [18]). This method uses the operations of "doubling" and "adding P". If we write n in its binary representation, the number of doublings is fixed by $\lfloor \log_2 n \rfloor$ and each *one* in this representation corresponds to an addition. Thus the cost of the multiplication depends on the length of the binary representation of n and the number of ones in this representation.

If addition and subtraction are equally costly in the underlying group, it makes sense to work with *signed binary representations*, i.e., binary representations with digits $\{0, \pm 1\}$. The advantage of these representations is their redundancy: in general, n has many different signed binary representations. Let n be written in a signed binary representation. Then the number of non-zero digits is called the *Hamming weight* of this representation. Since each non-zero digit causes a group addition (1 causes addition of P, -1 causes subtraction of P), one is interested in finding a representation of n having minimal Hamming weight. Such a minimal representation was exhibited by Reitwiesner [16]. Since it has no adjacent non-zero digits, this type of representation is often called *non-adjacent form* or NAF, for short. On average, only one third of the digits of a NAF is different from zero. Morain and Olivos [14] first observed that NAFs are useful for calculating nP for large n quickly.

In this paper we show that in an average sense every integer n has "many" signed binary representations of minimal weight. We give sharp upper bounds for the number $f(n)$ of such representations and study the summatory function of $f(n)$. In order to prove that this summatory function exhibits a periodically fluctuating main term, we develop a new approach to summatory functions of digital functions. This new approach allows to study sums of digital functions without having "nice" explicit formulæ for these functions. We construct a purely singular continuous measure, which encodes the distribution of the number of minimal weight expansions after rescaling. This measure is then used to describe the periodic fluctuation in the asymptotic expansion. Furthermore, we describe a method to compute the Fourier coefficients of the periodic fluctuation numerically to high precision.

Date: April 20, 2004.

2000 *Mathematics Subject Classification.* Primary: 11A63 Secondary: 11K16, 11K55, 68W40, 94A60.

Key words and phrases. Signed digit expansions, minimal Hamming weight, elliptic curve cryptography.

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

‡ This author is supported by the grant S8307-MAT of the Austrian Science Fund.

2. Counting frequencies

We consider binary $\{0, \pm 1\}$-expansions $\boldsymbol{\varepsilon} = (\varepsilon_k, \dots, \varepsilon_0)$ of integers. The Hamming weight of $\boldsymbol{\varepsilon}$ is defined as the number of non-zero digits ε_j. An expansion is said to be *optimal* or *minimal*, if it has minimal Hamming weight amongst all expansions of the same integer. One example of a minimal expansion is the *non-adjacent form* introduced by Reitwiesner [16]: this is the unique binary expansion of an integer which satisfies $\varepsilon_j \varepsilon_{j+1} = 0$. The number of minimal expansions of an integer n will be denoted by $f(n)$.

In [11, Remark 20, Figure 16] it has been proved that an expansion $\boldsymbol{\varepsilon}$ is optimal if and only if it is accepted by the automaton in Figure 1 (reading the digits from right to left), cf. also [10, Theorem 12].

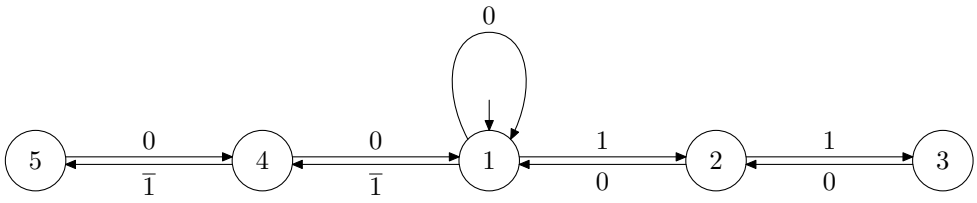

Figure 1. Automaton recognizing optimal signed binary expansions from right to left

As a first result we give a sharp upper bound for the counting function $f(n)$. This estimate will also be useful for the average case analysis in Section 3.

Theorem 1. *For all integers ℓ, the number of optimal expansions can be bounded by*

$$f(\ell) \le F_{\lfloor \log_4 |\ell| \rfloor + 3}, \tag{2.1}$$

where F_n denotes the Fibonacci sequence $F_0 = 0$, $F_1 = 1$, $F_{n+2} = F_{n+1} + F_n$. This bound is sharp for infinitely many values of ℓ. Less precisely, we have

$$f(n) = \mathcal{O}(n^{\log_4 \varphi}) \text{ with } \varphi = \frac{1+\sqrt{5}}{2}. \tag{2.2}$$

Proof. The automaton in Figure 1 gives rise to four auxiliary functions $a_j(n)$, $j = 2, \dots, 5$ which count the number of admissible runs in the automaton representing n and starting in state j. We set $a_1(n) = f(n)$. These functions satisfy the following relations, which can be read off from the automaton

$$\begin{aligned}
a_1(2n) &= a_1(n) & a_1(2n+1) &= a_2(n) + a_4(n+1)\\
a_2(2n) &= a_1(n) & a_2(2n+1) &= a_3(n)\\
a_3(2n) &= a_2(n) & a_3(2n+1) &= 0\\
a_4(2n) &= a_1(n) & a_4(2n+1) &= a_5(n+1)\\
a_5(2n) &= a_4(n) & a_5(2n+1) &= 0.
\end{aligned} \tag{2.3}$$

We claim that for $s \ge 0$ and $0 \le \ell < 4^s$, there are constants c_1, c_2, i_1, i_2 (depending on s and ℓ) such that for all n,

$$\begin{aligned}
a_1(4^s n + \ell) &= c_1 a_{i_1}(n) + c_2 a_{i_2}(n+1),\\
&\quad \{i_1, i_2\} \in \{(1,0), (0,1), (3,1), (2,4), (1,5)\},\\
\min\{c_1, c_2\} &\le F_s,\\
\max\{c_1, c_2\} &\le F_{s+1},
\end{aligned}$$

where $i_j = 0$ means that $c_j = 0$. Furthermore, for each s, there is an ℓ such that $(c_1, i_1, c_2, i_2) = (F_s, 2, F_{s+1}, 4)$ and an ℓ such that $(c_1, i_1, c_2, i_2) = (F_{s+1}, 2, F_s, 4)$.

To prove the claim by induction on s, we consider the relations given in Table 1. We remark here that the table shows that the function $f(n)$ is 2-regular in the sense of Allouche and Shallit (cf. [1]).

	$m=0$	$m=1$
$c_1a_1(4n+m)$	$c_1a_1(n)$	$c_1a_1(n)+c_1a_5(1+n)$
$c_2a_1(4n+m+1)$	$c_2a_1(n)+c_2a_5(1+n)$	$c_2a_2(n)+c_2a_4(1+n)$
$c_1a_3(4n+m)+c_2a_1(4n+m+1)$	$(c_1+c_2)a_1(n)+c_2a_5(1+n)$	$c_2a_2(n)+c_2a_4(1+n)$
$c_1a_2(4n+m)+c_2a_4(4n+m+1)$	$c_1a_1(n)$	$(c_1+c_2)a_2(n)+c_2a_4(1+n)$
$c_1a_1(4n+m)+c_2a_5(4n+m+1)$	$c_1a_1(n)$	$c_1a_1(n)+(c_1+c_2)a_5(1+n)$
	$m=2$	$m=3$
$c_1a_1(4n+m)$	$c_1a_2(n)+c_1a_4(1+n)$	$c_1a_3(n)+c_1a_1(1+n)$
$c_2a_1(4n+m+1)$	$c_2a_3(n)+c_2a_1(1+n)$	$c_2a_1(1+n)$
$c_1a_3(4n+m)+c_2a_1(4n+m+1)$	$(c_1+c_2)a_3(n)+c_2a_1(1+n)$	$c_2a_1(1+n)$
$c_1a_2(4n+m)+c_2a_4(4n+m+1)$	$c_1a_2(n)+(c_1+c_2)a_4(1+n)$	$c_2a_1(1+n)$
$c_1a_1(4n+m)+c_2a_5(4n+m+1)$	$c_1a_2(n)+c_1a_4(1+n)$	$c_1a_3(n)+(c_1+c_2)a_1(1+n)$

TABLE 1. Recurrence relations for a_j

We see that if $a_1(4^sn+\ell)=c_1a_{i_1}(n)+c_2a_{i_2}(n+1)$, then we have $a_1(4^{s+1}n+m4^s+\ell)=c_1'a_{i_1'}(n)+c_2'a_{i_2'}(n+1)$ with

$$\min\{c_1',c_2'\}\le\max\{c_1,c_2\}\le F_{s+1},$$
$$\max\{c_1',c_2'\}\le c_1+c_2\le F_s+F_{s+1}=F_{s+2}$$

and such that (i_1',i_2') is again one of the pairs considered. Moreover, there are the two pairs where the inequalities are sharp.

Since $a_j(0)=1$ for $j\in\{1,2,3,4,5\}$ and $a_1(1)=a_2(1)=1$ and $a_j(1)=0$ for $j\in\{3,4,5\}$, the assertion of the theorem follows by setting $n=0$. □

3. Construction of a measure and average case analysis

This section is devoted to the precise study of the summatory function $\sum_{n<N}f(n)$, which describes the average behavior of $f(n)$. In order to exhibit the fluctuating main term of this sum we introduce a measure μ on $[-1,1]$, which will turn out to be purely singular continuous in Section 5. The construction of this measure is similar to the distribution measures of infinite Bernoulli convolutions as studied in [4]. There it encodes the number of representations of integers as sums of Fibonacci numbers.

Let $f_n(k)$ denote the number of representations of an integer k of minimal weight and length at most n. Since any representation of minimal weight is at most 1 digit longer than the usual binary expansion, $f_n(k)=f_{\lfloor\log_2|k|\rfloor+2}(k)=f(k)$ for $n\ge\lfloor\log_2|k|\rfloor+2$. We define a sequence of measures by

$$\mu_n=\frac{1}{M_n}\sum_{k\in\mathbb{Z}}f_n(k)\delta_{k2^{-n}},\tag{3.1}$$

where δ_x denotes the unit point mass concentrated in x and

$$M_n=\sum_{k\in\mathbb{Z}}f_n(k).$$

We notice that all points $k2^{-n}$ with $f_n(k)>0$ lie in the interval $[-1,1]$.

In order to compute the characteristic function of μ_n we consider the weighted adjacency matrix of the automaton in Figure 1:

$$A(z)=\begin{pmatrix}1&z&0&\frac{1}{z}&0\\1&0&z&0&0\\0&1&0&0&0\\1&0&0&0&\frac{1}{z}\\0&0&0&1&0\end{pmatrix}.$$

In the matrix $A(z)$ a transition with label d is represented by an entry z^d. Then we have (using the notation $e(t) = e^{2\pi i t}$)

$$\widehat{\mu}_n(t) = \frac{1}{M_n} \sum_{k\in\mathbb{Z}} f_n(k) e\left(k2^{-n}t\right) = \frac{1}{M_n} v_1 A\left(e(t2^{-n})\right) A\left(e(t2^{-n+1})\right) \cdots A\left(e(t/2)\right) v_2$$

with $v_1 = (1,0,0,0,0)$ and $v_2 = (1,1,1,1,1)^T$.

We notice that

$$M_n = (1,0,0,0,0)A(1)^n(1,1,1,1,1)^T = C\alpha^n + \mathcal{O}(|\alpha_2|^n), \tag{3.2}$$

where α and α_2 are the largest and second largest roots of the characteristic polynomial of $A(1)$ given by

$$(x-1)(x+1)(x^3 - x^2 - 3x + 1),$$

and $C = \frac{1}{37}(14\alpha^2 + 5\alpha - 22)$, numerically

$$\alpha = 2.17009\ldots, \qquad \alpha_2 = -1.48119\ldots, \qquad C = 1.48055\ldots.$$

We will prove that (μ_n) weakly tends to a limit measure by showing that $\widehat{\mu}_n(t)$ tends to a limit $\widehat{\mu}(t)$.

Lemma 1. *The sequence of measures μ_n defined by* (3.1) *converges weakly to a probability measure μ. The characteristic functions satisfy the inequality*

$$|\widehat{\mu}_n(t) - \widehat{\mu}(t)| = \begin{cases} \mathcal{O}\left(|t|2^{-\eta n}\right) & \text{for } |t| \le 1 \\ \mathcal{O}\left(|t|^\eta 2^{-\eta n}\right) & \text{for } |t| \ge 1 \end{cases} \tag{3.3}$$

with

$$\eta = \frac{\log\alpha - \log|\alpha_2|}{\log 2 + \log\alpha - \log|\alpha_2|} = 0.355251\ldots.$$

The constants implied by the $\mathcal{O}$-symbol are absolute.

Proof. We study the product

$$P_n(t) = \alpha^{-n} \prod_{j=1}^{n} A\left(e(t2^{-j})\right),$$

where $\coprod_{j=1}^n p_j = p_n p_{n-1} \cdots p_1$. We have $\|A(e(t)) - A(1)\| \le \pi\sqrt{8}|t|$, where $\|\cdot\|$ denotes the spectral norm. Since the characteristic polynomial of $A(e(-t))^T A(e(t))$ turns out to be independent of t, if t is real, we have $\|A(e(t))\| = \alpha$. For $|t| \le 1$ we estimate

$$\begin{aligned}\|P_n(t) - P_n(0)\| &\le \frac{1}{\alpha^n} \sum_{\ell=1}^{n} \alpha^{n-\ell} \sum_{1\le j_1<\cdots<j_\ell\le n} \prod_{m=1}^{\ell} \left(\frac{\pi\sqrt{8}|t|}{2^{j_m}}\right) \\ &\le \sum_{\ell=1}^{n} \frac{(\pi\sqrt{8})^\ell}{\alpha^\ell \ell!} |t|^\ell \left(\sum_{j=1}^{n} 2^{-j}\right)^\ell \le \exp\left(\frac{\pi\sqrt{8}}{\alpha}|t|\right) - 1 \le 60|t|.\end{aligned}$$

Furthermore, we have for $m > n > \ell$ and $1 \le |t| \le 2^\ell$

$$\begin{aligned}\|P_n(t) - P_m(t)\| &= \left\|P_{n-\ell}(t2^{-\ell})P_\ell(t) - P_{m-\ell}(t2^{-\ell})P_\ell(t)\right\| \\ &\le \left\|P_{n-\ell}(t2^{-\ell}) - P_{n-\ell}(0)\right\| + \left\|P_{m-\ell}(t2^{-\ell}) - P_{m-\ell}(0)\right\| + \|P_{n-\ell}(0) - P_{m-\ell}(0)\| \\ &\le 120|t|2^{-\ell} + 6\left(\frac{|\alpha_2|}{\alpha}\right)^{n-\ell} = \mathcal{O}\left(|t|^\eta 2^{-\eta n}\right).\end{aligned} \tag{3.4}$$

In the last step we have set $\ell = \lceil (1-\eta)\log_2|t| + \eta n \rceil$. The inequality is valid for $m > n > \log_2|t|$.

We now assume that $|t| \le 1$ and $m > n > \ell$. Then we have

$$\begin{aligned}
|\widehat{\mu}_n(t) - \widehat{\mu}_m(t)| &= \left| \frac{\alpha^n}{M_n} v_1 P_n(t) v_2 - \frac{\alpha^m}{M_m} v_1 P_m(t) v_2 \right| \\
&= \left| \frac{\alpha^n}{M_n} v_1 P_{n-\ell}\left(t2^{-\ell}\right) P_\ell(t) v_2 - \frac{\alpha^m}{M_m} v_1 P_{m-\ell}\left(t2^{-\ell}\right) P_\ell(t) v_2 \right| \\
&\le \left| \frac{\alpha^n}{M_n} v_1 P_{n-\ell}(0) P_\ell(t) v_2 - \frac{\alpha^m}{M_m} v_1 P_{m-\ell}(0) P_\ell(t) v_2 \right| + \mathcal{O}\left(|t| 2^{-\ell}\right) \\
&= \left| \frac{\alpha^n}{M_n} v_1 P_{n-\ell}(0) \left(P_\ell(t) - P_\ell(0)\right) v_2 - \frac{\alpha^m}{M_m} v_1 P_{m-\ell}(0) \left(P_\ell(t) - P_\ell(0)\right) v_2 \right| + \mathcal{O}\left(|t| 2^{-\ell}\right) \\
&= |t| \mathcal{O}\left(2^{-\ell} + \left(\frac{|\alpha_2|}{\alpha} \right)^{n-\ell} \right),
\end{aligned}$$

where we have used $\frac{\alpha^n}{M_n} v_1 P_n(0) v_2 = 1$ in the fourth line. Setting $\ell = \lfloor \eta n \rfloor$ gives

$$|\widehat{\mu}_n(t) - \widehat{\mu}_m(t)| = \mathcal{O}\left(|t| 2^{-\eta n}\right).$$

Thus $\widehat{\mu}_n(t)$ converges uniformly on compact subsets of $\mathbb{R}$ to a continuous limit $\widehat{\mu}(t)$, and the measures μ_n tend to a measure μ weakly. The two inequalities (3.3) are immediate. □

In the next lemma we prove continuity of the measure μ. Our study of the Fourier expansion of the periodic main term as well as the remainder term estimate in (3.8) will depend on the modulus of continuity given here.

Lemma 2. *The measure μ satisfies*

$$\mu([x, y]) = \mathcal{O}\left((y - x)^\beta\right) \tag{3.5}$$

for $\beta = \log_2 \alpha - \log_4 \varphi = 0.770632\ldots > \frac{1}{2}$.

Proof. We first notice that every interval $[x, y] \subseteq [-1, 1]$ can be covered by an interval $[x', y']$, which is the union of at most two elementary binary intervals, *i.e.* intervals of the form $[a2^{-n}, (a+1)2^{-n}]$ with $a \in \mathbb{Z}$, $n \in \mathbb{N}$, such that

$$\frac{1}{5}(y' - x') \le (y - x) \le (y' - x'). \tag{3.6}$$

For a proof just consider the interval

$$\left[2^{-n} \lfloor x 2^n \rfloor, 2^{-n} \lceil y 2^n \rceil\right]$$

with $n = -\lfloor \log_2(y - x) \rfloor - 1$.

Thus it suffices to prove (3.5) for elementary binary intervals. We have

$$\mu\left([a2^{-n}, (a+1)2^{-n}]\right) = \lim_{m \to \infty} \frac{1}{M_m} \sum_{a2^{m-n} \le k \le (a+1)2^{m-n}} f_m(k) = \lim_{m \to \infty} \frac{1}{M_m} \sum_{0 \le k \le 2^{m-n}} f_m(a2^{m-n} + k).$$

Let $a2^{m-n} + k = \sum_{j \ge 0} \varepsilon_j 2^j$ be an optimal expansion. Then for some $\varepsilon \in \{0, \pm 1\}$, $\sum_{j=0}^{m-n-1} \varepsilon_j 2^j$ and $\sum_{j \ge m-n} \varepsilon_j 2^{j-m+n}$ are optimal expansions of $k - \varepsilon 2^{m-n}$ and $a + \varepsilon$, respectively. Therefore we have

$$\sum_{0 \le k < 2^{m-n}} f_m(a2^{m-n} + k) \le \sum_{\substack{0 \le k \le 2^{m-n} \\ \varepsilon \in \{0, \pm 1\}}} f(a + \varepsilon) f(k - \varepsilon 2^{m-n}) \ll (a+1)^{\log_4 \varphi} \sum_{-2^{m-n} \le k \le 2^{m-n+1}} f(k)$$

by Theorem 1. Since the last sum is bounded by M_{m-n+2} we have

$$\mu\left([a2^{-n}, (a+1)2^{-n}]\right) \ll (a+1)^{\log_4 \varphi} \lim_{m \to \infty} \frac{M_{m-n+2}}{M_m} \ll \left(\frac{\sqrt{\varphi}}{\alpha} \right)^n.$$

Combining this with (3.6) gives (3.5). □

In order to give an error bound for the rate of convergence of the measures μ_n to the measure μ, we will use the following version of the Berry-Esseen inequality, which was proved in [7].

Proposition 1. *Let μ_1 and μ_2 be two probability measures with their Fourier transforms defined by*

$$\widehat{\mu}_k(t) = \int_{-\infty}^{\infty} e^{2\pi i t x}\, d\mu_k(x), \quad k = 1, 2.$$

Suppose that $(\widehat{\mu}_1(t) - \widehat{\mu}_2(t))t^{-1}$ is integrable on a neighborhood of zero and μ_2 satisfies

$$\mu((x, y)) \le c|x - y|^{\beta}$$

for some $0 < \beta < 1$. Then the following inequality holds for all real x and all $T > 0$

$$|\mu_1((-\infty, x)) - \mu_2((-\infty, x))| \le \left| \int_{-T}^{T} \hat{J}(T^{-1}t)(2\pi i t)^{-1} \left(\widehat{\mu}_1(t) - \widehat{\mu}_2(t)\right) e^{-2\pi i x t}\, dt \right| + \left(c + \frac{1}{\pi^2}\right) T^{-\frac{2\beta}{2+\beta}} + \left| \frac{1}{2T} \int_{-T}^{T} \left(1 - \frac{|t|}{T}\right) \left(\widehat{\mu}_1(t) - \widehat{\mu}_2(t)\right) e^{-2\pi i x t}\, dt \right|,$$

where

$$\hat{J}(t) = \pi t(1 - |t|)\cot \pi t + |t|.$$

Lemma 3. *The measures μ_n satisfy*

$$|\mu_n((x, y)) - \mu((x, y))| = \mathcal{O}\left(2^{-\theta n}\right) \tag{3.7}$$

uniformly for all $x, y \in \mathbb{R}$ with $\theta = \frac{2\beta\eta}{\eta(\beta+2)+2\beta} = 0.2168\ldots$.

Proof. We apply Proposition 1 to the measures μ_n and μ. For this purpose we use the inequalities (3.3) to obtain

$$|\mu_n((-\infty, x)) - \mu((-\infty, x))| \ll 2^{-\eta n} \int_{-1}^{1} dt + 2^{-\eta n} \int_{1 \le |t| \le T} |t|^{\eta - 1}\, dt + T^{-\frac{2\beta}{2+\beta}} + 2^{-\eta n} \frac{1}{T} \int_{-1}^{1} |t|\, dt + 2^{-\eta n} \frac{1}{T} \int_{1 \le |t| \le T} |t|^{\eta}\, dt \ll 2^{-\theta n}$$

by choosing $T = 2^{\theta \frac{2+\beta}{2\beta} n}$. □

In Section 5 we will prove that μ is purely singular with respect to Lebesgue measure. As a first step we prove the following lemma.

Lemma 4. *The measure μ is not absolutely continuous with respect to Lebesgue measure.*

Proof. We observe that

$$\widehat{\mu}(2^k) = \lim_{n \to \infty} v_1 \frac{\alpha^n}{M_n} P_n(2^k) v_2 = \lim_{n \to \infty} v_1 \frac{\alpha^n}{M_n} P_{n-k}(1) A(1)^k v_2.$$

The matrices $P_m(1)$ converge to a limiting matrix $P(1)$ with rate $120 \cdot 2^{-m/2} + 6\left(\frac{|\alpha_2|}{\alpha}\right)^{m/2}$ by (3.4). A numerical computation shows that

$$P(1) = \begin{pmatrix} 0.2350 & -0.2552 + 0.04946i & -0.01256 - 0.1395i & -0.2552 - 0.04946i & -0.01256 + 0.1395i \\ 0.1375 & -0.1493 + 0.02893i & -0.007352 - 0.08165i & -0.1493 - 0.02893i & -0.007352 + 0.08165i \\ 0.06337 & -0.06880 + 0.01333i & -0.003388 - 0.03762i & -0.06880 - 0.01333i & -0.003388 + 0.03762i \\ 0.1375 & -0.1493 + 0.02893i & -0.007352 - 0.08165i & -0.1493 - 0.02893i & -0.007352 + 0.08165i \\ 0.06337 & -0.06880 + 0.01333i & -0.003388 - 0.03762i & -0.06880 - 0.01333i & -0.003388 + 0.03762i \end{pmatrix}$$

The matrices $\alpha^{-k} A(1)^k$ tend to

$$B = \begin{pmatrix} 0.546474 & 0.319711 & 0.147326 & 0.319711 & 0.147326 \\ 0.319711 & 0.187045 & 0.0861923 & 0.187045 & 0.0861923 \\ 0.147326 & 0.0861923 & 0.0397184 & 0.0861923 & 0.0397184 \\ 0.319711 & 0.187045 & 0.0861923 & 0.187045 & 0.0861923 \\ 0.147326 & 0.0861923 & 0.0397184 & 0.0861923 & 0.0397184 \end{pmatrix}.$$

Thus

$$\lim_{k\to\infty} \widehat{\mu}(2^k) = \frac{1}{C} v_1 P(1) B v_2 = -0.0703223\ldots$$

Thus μ is not absolutely continuous by the Riemann-Lebesgue lemma. □

Now the statement of the asymptotic behavior of the summatory function $\sum_{n<N} f(n)$ is a consequence of the preceding discussion of the properties of μ.

Theorem 2. *The counting function $f(n)$ of the representations of n with minimal weight satisfies*

$$\sum_{n<N} f(n) = N^{\log_2 \alpha} \Phi(\log_2 N) + \mathcal{O}(N^{\log_2 \alpha - \theta}), \tag{3.8}$$

where Φ denotes a continuous periodic function of period 1 and θ is given in Lemma 3. Furthermore, Φ is Hölder continuous with exponent $\beta = \log_2 \alpha - \log_4 \varphi = 0.770632\ldots$. The function Φ is differentiable almost everywhere and singular in the sense that it is not the integral of its derivative.

Proof. By the definition of μ_n in (3.1) we have for $2^{k-2} \le N < 2^{k-1}$

$$\sum_{n<N} f(n) = \sum_{n<N} f_k(n) = M_k \mu_k([0, N2^{-k})) = C\alpha^k \mu([0, N2^{-k})) + \mathcal{O}(|\alpha_2|^k) + \mathcal{O}(\alpha^k 2^{-\theta k}).$$

Setting $\Phi(x) = C\alpha^{2-\{x\}} \mu([0, 2^{\{x\}-2}])$ we obtain the desired result. We notice that $\Phi(0) = \lim_{x\to 1-} \Phi(x)$ by the fact that the measure μ satisfies the relation $\mu([0, 2x)) = \alpha\mu([0, x))$ for $0 < x < \frac{1}{4}$ by definition.

The Hölder exponent of Φ follows from Lemma 2. The function Φ is differentiable almost everywhere as a quotient of an increasing function and a differentiable function. The singularity of Φ follows from Lemma 4. □

4. Dirichlet Series and Fourier coefficients

In order to compute the Fourier coefficients of the periodic function Φ occurring in Theorem 2, we introduce the Dirichlet generating functions

$$\Psi_j(s) = \sum_{n=1}^{\infty} \frac{a_j(n)}{n^s},$$

which converge absolutely for $\Re(s) > \log_2 \alpha$ by Theorem 2 (cf. [9]).

Using (2.3) we derive

$$\begin{aligned}
\Psi_1(s) &= 2^{-s}\left(\Psi_1(s) + \Psi_2(s) + \Psi_4(s)\right) + 1 + 2^{-s}\left(H_2^+(s) + H_4^-(s)\right)\\
\Psi_2(s) &= 2^{-s}\left(\Psi_1(s) + \Psi_3(s)\right) + 1 + 2^{-s} H_3^+(s)\\
\Psi_3(s) &= 2^{-s}\Psi_2(s)\\
\Psi_4(s) &= 2^{-s}\left(\Psi_1(s) + \Psi_5(s)\right) + 2^{-s} H_5^-(s)\\
\Psi_5(s) &= 2^{-s}\Psi_4(s),
\end{aligned} \tag{4.1}$$

where

$$H_j^{\pm}(s) = \sum_{n=1}^{\infty} a_j(n) \left(\frac{1}{(n \pm \frac{1}{2})^s} - \frac{1}{n^s} \right).$$

The Dirichlet series $H_j^{\pm}(s)$ are absolutely convergent for $\Re(s) > \log_2 \alpha$ and convergent for $\Re(s) > \log_2 \alpha - 1$. By general properties of Dirichlet series (cf. [9]) we have the following growth estimates along vertical lines

$$H_j^{\pm}(\sigma + it) = \mathcal{O}(|t|^{\log_2 \alpha - \sigma}) \text{ for } \log_2 \alpha - 1 < \sigma \le \log_2 \alpha. \tag{4.2}$$

From (4.1) we get

$$\Psi_1(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \frac{8^s + 4^s - 2^s + (4^s - 1)(H_2^+(s) + H_4^-(s)) + 2^s H_3^+(s) + 2^s H_5^-(s)}{8^s - 4^s - 3\cdot 2^s + 1}, \tag{4.3}$$

which provides the analytic continuation of $\Psi_1(s)$ to the region $\Re(s) > \log_2 \alpha - 1$ and shows that the poles of $\Psi_1(s)$ in this region lie at $s = \log_2 \alpha + \frac{2k\pi i}{\log 2}$ and $s = \log_2 |\alpha_2| + \frac{(2k+1)\pi i}{\log 2}$ $(k \in \mathbb{Z})$.

We now apply the Mellin-Perron summation formula to obtain

$$\sum_{n<N} f(n)\left(1-\frac{n}{N}\right) = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} \Psi_1(s)\frac{N^s\,ds}{s(s+1)}.$$

Shifting the line of integration to $\Re(s) = \log_2 \alpha - \frac{1}{2}$, using (4.3), and collecting residues yields

$$\sum_{n<N} f(n)\left(1-\frac{n}{N}\right) = N^{\log_2\alpha}\sum_{k\in\mathbb{Z}} \frac{c_k}{\chi_k+1} e\left(k\log_2 N\right) + \frac{1}{2\pi i}\int_{\log_2\alpha-\frac{1}{2}-i\infty}^{\log_2\alpha-\frac{1}{2}+i\infty} \Psi_1(s)\frac{N^s\,ds}{s(s+1)},$$

where

$$c_k = \frac{2\alpha^2+2\alpha-1+(\alpha^2-1)(H_2^+(\chi_k)+H_4^-(\chi_k))+\alpha H_3^+(\chi_k)+\alpha H_5^-(\chi_k)}{(\alpha^2+6\alpha-3)\chi_k\log 2} \tag{4.4}$$

and $\chi_k = \log_2\alpha + \frac{2k\pi i}{\log 2}$. The integral and the sum converge by the growth estimate (4.2).

We use an argument given in [8] to compute the Fourier coefficients of the periodic function Φ. First we cite a pseudo-Tauberian argument stated in [5, Proposition 2]

Proposition 2. *Let p be a continuous function and periodic with period 1 and let τ be a complex number with $\Re(\tau) > 0$. Then there exists a continuously differentiable function q of period 1 such that*

$$\frac{1}{N^{\tau+1}}\sum_{n<N} n^\tau p(\log_2 n) = q(\log_2 N) + o(1).$$

Furthermore,

$$\int_0^1 q(u)du = \frac{1}{\tau+1}\int_0^1 p(u)du.$$

An application of this proposition to $p(u) = \Phi(u)e(-\ell u)$ and $\tau = \chi_\ell$ shows that the c_k in (4.4) are indeed the Fourier coefficients of Φ.

In order to obtain the convergence of the Fourier series of Φ, we cite Bernstein's theorem for Fourier series (see [20, p. 240])

Proposition 3. *If f is a real-valued function defined on $[0,1]$ and satisfies a Hölder condition of order $\beta > 1/2$, namely,*

$$|f(x)-f(y)| \le K|x-y|^\beta \qquad (x,y\in[0,1]),$$

for some positive constant K, then the Fourier series of f converges absolutely and uniformly.

We summarize.

Theorem 3. *The function Φ defined in (3.8) admits an absolutely and uniformly convergent Fourier series*

$$\Phi(t) = \sum_{k\in\mathbb{Z}} c_k e(kt),$$

where c_k is given by (4.4).

In [8] a method for the numerical computation of the Fourier coefficients c_k, or more generally for the computation of the values of Dirichlet generating functions of digital functions was described. The basic idea is to expand the functions $H_j^\pm(s)$ in terms of function evaluations $\Psi_j(s+k)$ $(k \ge 1)$; this is done by writing $(n \pm \frac{1}{2})^{-s}$ as a binomial series. Since $\Psi_j(s+k) \approx a_j(1)$ for k large enough, the infinite sum can be computed numerically.

We have computed the first seven Fourier coefficients by this method:

$$\begin{aligned}
c_0 =& 1.00640\,74723\,03529\,37352\,02842\,85855\,81336\,33055\,57035\,48188\ldots\\
c_1 =& 0.00734\,84453\,42244\,68089\,95364\,74294\,73583\,52315\,12670\,18294\ldots\\
&+ 0.02689\,11696\,16758\,68783\,02281\,99800\,38391\,33382\,04336\,33025\ldots i\\
c_2 =& -0.00430\,40242\,79775\,54322\,96219\,62111\,36973\,03149\,13327\,29671\ldots\\
&- 0.00267\,43837\,30021\,02109\,40115\,62991\,88249\,06171\,44235\,24279\ldots i\\
c_3 =& -0.00127\,39427\,57534\,41610\,01651\,96139\,66214\,08173\,27366\,30382\ldots\\
&+ 0.00109\,88363\,17314\,47930\,07972\,98256\,69412\,99603\,80485\,12526\ldots i\\
c_4 =& 0.00394\,88393\,42163\,50681\,33279\,73298\,99918\,67013\,01876\,22678\ldots\\
&- 0.00161\,07854\,95299\,29954\,46287\,08562\,73623\,84651\,87192\,47177\ldots i\\
c_5 =& -0.00277\,30499\,06965\,12243\,95529\,90477\,88154\,74114\,71154\,01015\ldots\\
&+ 0.00258\,99972\,84840\,72727\,87528\,23704\,81316\,63001\,64706\,16579\ldots i\\
c_6 =& -0.00003\,84537\,28840\,80211\,49042\,02548\,65870\,42924\,48539\,84912\ldots\\
&- 0.00025\,86306\,19932\,25562\,45234\,74086\,81085\,61564\,92721\,30980\ldots i
\end{aligned}$$

In Figure 2, we compare plots of the function Φ and the trigonometric polynomial formed with the first 7 Fourier coefficients.

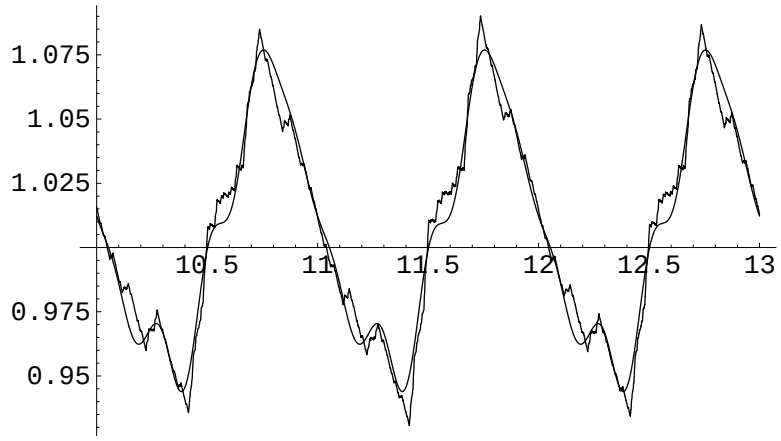

FIGURE 2. Plot of Φ compared with the trigonometric polynomial formed with the first seven Fourier coefficients

5. Purity of the measure μ

In this section we study the measure μ introduced in Section 3 in further detail. In particular, we show that it is purely singular continuous. As it is the case for Bernoulli convolutions (cf. [4]) the measure turns out to be pure as a consequence of the Jessen-Wintner theorem.

Lemma 5 ([12, Theorem 35], [3, Lemma 1.22 (ii)])**.** *Let $Q = \prod_{n=0}^{\infty} Q_n$ be an infinite product of discrete spaces equipped with a measure ν, which satisfies Kolmogorov's 0-1-law (i.e. every tail event has either measure 0 or 1). Furthermore, let X_n be a sequence of random variables defined on the spaces Q_n, such that the series $X = \sum_{n=0}^{\infty} X_n$ converges ν-almost everywhere. Then the distribution of X is either purely discrete, or purely singular continuous, or absolutely continuous with respect to Lebesgue measure.*

Remark 1. We notice that in [3] and [12] the additional assumption of mutual independence of the random variables X_n is made in the statement of the result instead of the 0-1-law. The proofs however only depend on the 0-1-law.

In the following we will study a measure ν on the space

$$\mathcal{K} = \left\{\mathbf{x} \in \{0, \pm 1\}^{\mathbb{N}} \mid \forall n \in \mathbb{N} : (x_1, x_2, \ldots, x_n) \text{ is an optimal expansion}\right\}.$$

We define ν on cylinders

$$[\varepsilon_1, \ldots, \varepsilon_n] = \{\mathbf{x} \in \mathcal{K} \mid x_1 = \varepsilon_1, \ldots, x_n = \varepsilon_n\}$$

by

$$\nu([\varepsilon_1, \ldots, \varepsilon_n]) = \lim_{k\to\infty} \frac{1}{M_k} \#\left(\{(x_1, \ldots, x_k) \text{ is optimal}\} \cap [\varepsilon_1, \ldots, \varepsilon_n]\right).$$

We notice that the measure μ studied in Section 3 is the image of ν under the map $\mathbf{x} \mapsto \sum_{n=1}^{\infty} x_n 2^{-n}$.

In order to give an explicit expression for $\nu([\varepsilon_1, \ldots, \varepsilon_n])$ we introduce the adjacency matrices A_ε associated with transitions with label ε in the automaton in Figure 1:

$$A_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad A_{-1} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We note that $A(1) = A_0 + A_1 + A_{-1}$. Then we have

$$\begin{aligned} \nu([\varepsilon_1, \ldots, \varepsilon_n]) &= \lim_{k\to\infty} \frac{1}{M_k} v_1 A(1)^{k-n} A_{\varepsilon_n} \cdots A_{\varepsilon_1} v_2 \\ &= v_1 \left(\lim_{k\to\infty} \frac{\alpha^k}{M_k} \left(\alpha^{-1} A(1)\right)^{k-n} \right) \alpha^{-n} A_{\varepsilon_n} \cdots A_{\varepsilon_1} v_2 \\ &= \frac{1}{C} v_1 B \alpha^{-n} A_{\varepsilon_n} \cdots A_{\varepsilon_1} v_2, \end{aligned}$$

where B is the matrix introduced in the proof of Lemma 4.

Furthermore, we introduce a measure F on $\mathcal{K}$ by defining it on cylinder sets as

$$F([\varepsilon_1, \ldots, \varepsilon_n]) := w^T \alpha^{-n} A_{\varepsilon_n} \cdots A_{\varepsilon_1} w,$$

where w is the normalized positive eigenvector of $A(1)$ for the eigenvalue α. Notice that $B = w \cdot w^T$. By definition we have

$$\frac{\|v_1 B\|_2}{C} F([\varepsilon_1, \ldots, \varepsilon_n]) \le \nu([\varepsilon_1, \ldots, \varepsilon_n]) \le \frac{\|v_1 B\|_2}{C \min_j w_j} F([\varepsilon_1, \ldots, \varepsilon_n])$$

for all cylinders and therefore $C_1 F(S) \le \nu(S) \le C_2 F(S)$ for all measurable sets $S \subseteq \mathcal{K}$ and positive constants C_1 and C_2. Thus the measures ν and F are equivalent.

Lemma 6. *The random variables $Y_n(x_1, x_2, \ldots) = x_n$ on $(\mathcal{K}, F)$ form a mixing sequence, i. e.*

$$|F(A \cap B) - F(A)F(B)| \le 6 \left(\frac{|\alpha_2|}{\alpha}\right)^k F(A)F(B)$$

for $A \in \sigma(Y_1, \ldots, Y_n)$ and $B \in \sigma(Y_{n+k+1}, \ldots)$ for all $n, k \in \mathbb{N}$.

Proof. It suffices to prove the lemma for cylinder sets A and B. Then we have

$$\begin{aligned} &|F\left(\{x \in \mathcal{K} \mid x_1 = \varepsilon_1, \ldots, x_n = \varepsilon_n, x_{n+k+1} = \delta_1, \ldots, x_{n+k+\ell} = \delta_\ell\}\right) - F([\varepsilon_1, \ldots, \varepsilon_n])F([\delta_1, \ldots, \delta_\ell])| \\ &\quad = \left| w^T \alpha^{-\ell-1} A_{\delta_\ell} \cdots A_{\delta_1} \left(\left(\alpha^{-1} A(1)\right)^k - w \cdot w^T \right) \alpha^{-n} A_{\varepsilon_n} \cdots A_{\varepsilon_1} w \right| \\ &\quad \le 6 \left(\frac{|\alpha_2|}{\alpha}\right)^k F([\varepsilon_1, \ldots, \varepsilon_n])F([\delta_1, \ldots, \delta_\ell]). \end{aligned}$$

□

Since mixing sequences of random variables satisfy a 0-1-law (cf. [17, Section V. b, p. 110] or [19, § 1.7]), and since the measures ν and F are equivalent, ν satisfies the hypotheses of Lemma 5. Therefore the measure μ is of pure type; since we already know that μ is continuous (Lemma 2) but not absolutely continuous (Lemma 4) we have proved

Theorem 4. *The measure μ is purely singular continuous.*

6. Concluding remarks

In [15] and [2] an algorithm for number representation is suggested, which uses a randomized perturbation of the classical Reitwiesner system (cf. [16]) to prevent differential power attacks on cryptographic devices. All the algorithms presented there increase the weight of the representation.

Our results show that on average there exists a large number of representations of minimal weight. Therefore a countermeasure against power attacks which does not increase the costs of the operation is possible.

Acknowledgement. This research was initiated while the authors were invited to the John Knopfmacher Centre, University of the Witwatersrand, Johannesburg.

References

1. J.-P. Allouche and J. Shallit, *Automatic sequences, theory, applications, generalizations*, Cambridge University Press, Cambridge, 2003.
2. N. Ebeid and A. Hasan, *Analysis of dpa countermeasures based on randomizing the binary algorithm*, Tech. Report CORR 2003-14, Centre for Applied Cryptographic Research (CACR), University of Waterloo, 2003, available at `http://www.cacr.math.uwaterloo.ca/techreports/2003/corr2003-14.ps`.
3. P. D. T. A. Elliott, *Probabilistic number theory. I, mean-value theorems*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science], vol. 239, Springer-Verlag, New York, 1979.
4. P. Erdős, *On a family of symmetric Bernoulli convolutions*, Amer. J. Math. **61** (1939), 974–976.
5. P. Flajolet, P. J. Grabner, P. Kirschenhofer, H. Prodinger, and R. F. Tichy, *Mellin transforms and asymptotics: digital sums*, Theor. Comput. Sci. **123** (1994), 291–314.
6. D. M. Gordon, *A survey of fast exponentiation methods*, J. Algorithms **27** (1998), 129–146.
7. P. J. Grabner, *Functional iterations and Brownian motion on the Sierpiński gasket*, Mathematika **44** (1997), 374–400.
8. P. J. Grabner and H.-K. Hwang, *Digital sums and divide-and-conquer recurrences: Fourier expansions and absolute convergence*, Constructive Approximation **?** (2004), ?–?
9. G. H. Hardy and M. Riesz, *The general theory of Dirichlet's series*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 18, Stechert-Hafner, Inc., New York, 1964.
10. C. Heuberger, *Minimal expansions in redundant number systems: Fibonacci bases and greedy algorithms*, to appear in Period. Math. Hungar.; Preprint available at `http://www.opt.math.tu-graz.ac.at/~cheub/publications/minredfibo.pdf`, 2004.
11. C. Heuberger and H. Prodinger, *Analysis of alternative digit sets for nonadjacent representations*, Preprint, available at `http://www.opt.math.tu-graz.ac.at/~cheub/publications/dnaf-1.pdf`.
12. B. Jessen and A. Wintner, *Distribution functions and the Riemann zeta function*, Trans. Amer. Math. Soc. **38** (1935), 48–88.
13. N. Koblitz, A. Menezes, and S. Vanstone, *The state of elliptic curve cryptography*, Des. Codes Cryptogr. **19** (2000), 173–193.
14. F. Morain and J. Olivos, *Speeding up the computations on an elliptic curve using addition-subtraction chains*, Inform Theory Appl. **24** (1990), 531–543.
15. E. Oswald and M. Aigner, *Randomized addition-subtraction chains as a countermeasure against power attacks*, Cryptographic hardware and embedded systems—CHES 2001 (Paris), Lecture Notes in Comput. Sci., vol. 2162, Springer, Berlin, 2001, pp. 39–50.
16. G. W. Reitwiesner, *Binary arithmetic*, Advances in computers, Vol. 1, Academic Press, New York, 1960, pp. 231–308.
17. M. Rosenblatt, *Random processes*, Springer-Verlag, New York, 1974, Graduate Texts in Mathematics, No. 17.
18. J. von zur Gathen and J. Gerhard, *Modern computer algebra*, Cambridge University Press, New York, 1999.
19. P. Walters, *Ergodic theory*, Springer, Berlin, 1982.
20. A. Zygmund, *Trigonometric Series. Vol. I, II*, Cambridge University Press, Cambridge, 1988.

(P. Grabner) Institut für Mathematik A, Technische Universität Graz, Steyrergasse 30, 8010 Graz, Austria
E-mail address: peter.grabner@tugraz.at

(C. Heuberger) Institut für Mathematik B, Technische Universität Graz, Steyrergasse 30, 8010 Graz, Austria
E-mail address: clemens.heuberger@tugraz.at