ANALYSIS OF ALTERNATIVE DIGIT SETS FOR NONADJACENT REPRESENTATIONS

CLEMENS HEUBERGER[‡] AND HELMUT PRODINGER[*]

Abstract. It is known that every positive integer n can be represented as a finite sum of the form $\sum_i a_i 2^i$, where $a_i \in \{0, 1, -1\}$ and no two consecutive a_i's are non-zero ("nonadjacent form", NAF). Recently, Muir and Stinson [12, 13] investigated other digit sets of the form $\{0, 1, x\}$, such that each integer has a nonadjacent representation (such a number x is called admissible). The present paper continues this line of research.

The topics covered include transducers that translate the standard binary representation into such a NAF and a careful topological study of the (exceptional) set (which is of fractal nature) of those numbers where no finite look-ahead is sufficient to construct the NAF from *left-to-right*, counting the number of digits 1 (resp. x) in a (random) representation, and the non-optimality of the representations if x is different from 3 or -1.

1. Introduction

Redundant number systems have been studied by many people, and for various reasons. Probably the most famous one uses the base 2, but instead of the traditional digits $0, 1$, it allows a third digit -1. One can make this system *unique* by superimposing a condition: no two adjacent digits different from zero are permitted. This *non-adjacent-form* ("NAF") goes back (at least) to Reitwiesner [15]. For pointers to the earlier literature we refer the reader to [10] and [7].

Recently, Muir and Stinson [12, 13] asked for other sets of digits, in particular of the form $\{0, 1, x\}$, such that every (positive) integer has a unique representation using these digits, base 2, and obeying again the condition that two adjacent digits cannot both be different from zero.

The present paper analyzes their number systems with respect to frequencies of digits, description of characteristic sets and similar questions. Our main focus is to provide results for *general* x as much as possible; we give explicit results as well as algorithms to compute some characteristic quantities for specific x.

We summarize here what follows in the later sections.

In order to describe the digits x that work, we define a suitable graph. The answer depends on whether each node can be reached from the starting node. We propose

[‡] This paper was written while the first author was a visitor at the John Knopfmacher Centre for Applicable Analysis and Number Theory, School of Mathematics, University of the Witwatersrand, Johannesburg. He thanks the centre for its hospitality. He was also supported by the grant S8307-MAT of the Austrian Science Fund.

[*] This author is supported by the grant NRF 2053748 of the South African National Research Foundation.

an $O(|x|)$-time algorithm for this decision. The list of admissible numbers x starts like $3, -1, -5, -13, \dots$ (they must all be $\equiv 3 \bmod 4$).

We then define transducers (finite automata with output) that translate the standard binary representation into a NAF using digits $\{0, 1, x\}$ by reading the input word from right to left (thus starting with the least significant digit). Such a transducer has at most $|x| + 4$ states. It can be constructed even for such x's that don't lead to NAFs; in such cases, not every number has a representation, and one can "see" from the transducer when this happens. (For the standard NAF, such a transducer is of course well known.)

One of the advantages of the (standard) NAF, especially for applications in cryptography, is, that many digits are zero "on average." To be more precise, if one considers all admissible words of length N, with digits (letters) from the set $\{0, 1, -1\}$, then a random word has on average $\frac{N}{6}$ digits '1' resp. '-1' and $\frac{2N}{3}$ zeros. Interestingly, these frequencies persist even in the general case, with error terms depending on the extra digit x; in principle (for any particular x), they can be computed. The variance follows a similar pattern, with the constant $\frac{1}{6}$ being replaced by $\frac{11}{108}$. From this one can conclude that the underlying distribution is Gaussian.

For various reasons (on-line computations!) it would be nice to generate the NAF when reading the digits of the binary expansion from left-to-right. It is known for the standard case $x = -1$ that this is not possible, but there exist equivalent representations (with respect to the large number of zeros, i.e., small "Hamming weight"). There are always situations where a finite look-ahead will not be sufficient to decide which digit should be taken. We scale this exceptional set B down to the unit interval and study it as a function of the length of the look-ahead. While this set is just $\{\frac{1}{6}, \frac{2}{6}, \frac{4}{6}, \frac{5}{6}\}$ in the standard instance, we enter the realm of fractals for $x = -5, -13, \dots$. The (Lebesgue-) measure of B is zero, but its Hausdorff dimension is positive and strictly less than 1. A similar situation occurred in the study of joint expansions in Grabner, Heuberger, and Prodinger [5]. For particular x, the Hausdorff dimension can be explicitly computed from the dominant eigenvalue of the adjacency matrix of a certain auxiliary automaton, but it is quite challenging to say something in general. We manage, however, to derive nontrivial lower and upper bounds.

As already mentioned, for $x = -1$, the NAF is optimal, as there is no representation of any integer with fewer nonzero digits. We show in an algorithmic fashion that this is also true for $x = 3$. This is no longer true for the other possible values of x. In most cases, the number 3 serves as a counter example; in a few exceptional cases one must take another one. As a byproduct of our analysis, we can describe (for $x = 1$ resp. $x = 3$) *all* optimal representations; recall that in these cases a representation of an integer m is optimal if it has the same number on nonzero digits as the NAF of m. These representations can be described by finite automata.

It is always beneficial in order to understand how the arithmetic of a number system works to understand first how addition of 1 (or any other fixed integer) works. Knuth's book [10] has several such examples. Here, we confine ourselves to the transducer describing the addition of 1 in the instance $x = -5$.

For the standard NAF case ($x = -1$) it is possible to give an explicit formula for the digits of the expansion, cf. Prodinger [14]. This is somehow a reflection of the fact that the above mentioned exceptional set is so simple. In the instance $x = -5$, there is still an explicit formula for the digits. However, it sometimes predicts a nonzero digit when it actually is zero, but is correct otherwise. This corresponds to Figure 9, where one can see that the intervals that contain green resp. blue points, are separated. For other values of x nothing like that exists anymore.

Before entering into the details, we will fix a minimum amount of notation. Let n be an integer and $D \subset \mathbb{Z}$. A *(binary) D-expansion of n* is a sequence $(\ldots, d_2, d_1, d_0) \in D^{\mathbb{N}_0}$ such that only finitely many digits are nonzero and such that $n = \sum_{j\geq 0} d_j 2^j$. The *standard binary expansion* of a nonnegative integer n is its unique $\{0,1\}$-expansion. We will usually denote sequences by boldface letters. Where appropriate, we will also think about D-expansions as finite or infinite words over the alphabet D. By $\mathsf{value}(\ldots, d_2, d_1, d_0)$, we denote $\sum_{j\geq 0} d_j 2^j$. The *position of the most significant digit* $\mathrm{MSB}(\mathbf{d})$ is $\max\{j : d_j \neq 0\}$.

2. Nonadjacent Digit Sets

We recall some definitions of Muir and Stinson [12, 13].

Definition 1. (1) Let $0 \in D \subset \mathbb{Z}$ and $n \in \mathbb{Z}$. Then a D-expansion $\boldsymbol{\eta}$ of n is called a *D-nonadjacent form (D-NAF)* of n, if

$$\eta_j \eta_{j+1} = 0 \text{ for all } j \geq 0.$$

(2) Let $0 \in D \subset \mathbb{Z}$. If there is a D-NAF for every positive integer n, then D is called a *nonadjacent digit set* (NADS).

Muir and Stinson [12, 13] study digit sets $D = \{0, 1, x\}$ for integers x. The following results have been proved in their paper:

Proposition 2. *Let $D = \{0, 1, x\}$.*

(1) *If D is a NADS, then $x \equiv 3 \pmod 4$.*
(2) *If $x \equiv 3 \pmod 4$, then each positive integer has at most one D-NAF.*
(3) *The only NADS $D = \{0, 1, x\}$ with positive x is $\{0, 1, 3\}$.*

Definition 3. Let $D = \{0, 1, x\}$ with $x \equiv 3 \pmod 4$. Then we define $\eta_0 : \mathbb{Z} \to D$ and $r : \mathbb{Z} \to \mathbb{Z}$ by setting

$$\eta_0(n) := \begin{cases} 0, & \text{if } n \equiv 0 \pmod 2, \\ 1, & \text{if } n \equiv 1 \pmod 4, \\ x, & \text{if } n \equiv 3 \pmod 4, \end{cases} \qquad r(n) := \frac{n - \eta_0(n)}{2}.$$

It is easily seen (cf. Muir and Stinson [12, 13]) that $n \in \mathbb{N}$ admits a D-NAF if and only if $r(n)$ admits a D-NAF and that if $(\ldots, \eta_1', \eta_0')$ is the D-NAF of $r(n)$ then $(\ldots, \eta_1', \eta_0', \eta_0(n))$ is the D-NAF of n. (Note that $r(n)$ is even if n is odd).

Since $r(n) < n$ for all positive $n \not\equiv 3 \pmod 4$ and $r^2(n) := r(r(n)) < n$ for all positive $n \equiv 3 \pmod 4$ with $n > |x|/3$, the set D is a NADS if and only if all $0 < n \leq |x|/3$ with $n \equiv 3 \pmod 4$ have a D-NAF.

Proposition 4. *Let $D = \{0, 1, x\}$ with $x < 0$ and $x \equiv 3 \pmod 4$. Define the directed graph $G := (V, A)$ by $V = \{0, \dots, \lfloor |x| /3 \rfloor\}$ and*

$$A := \{(m, n) \in V^2 : n \in \{2m, 4m + 1, 4m + x\}\}.$$

Then D is a NADS if and only if every $n \in V$ is reachable from 0.

Proof. We observe that if $0 \neq n \in V$, then there is exactly one edge with head n, namely $(r^i(n), n)$, where $i = (n \bmod 2) + 1$. Therefore, there is a path[1] from 0 to n if and only if n has a D-NAF. □

Therefore, we may check whether a set D is a NADS by simple breadth-first search (cf. Algorithm 1).

Algorithm 1 Check for NADS

Input: $D = \{0, 1, x\}$ with $x < 0$ and $x \equiv 3 \pmod 4$
Output: Decision, whether D is a NADS

```
S ← (0)
c ← 1
while S ≠ () do
  m ← First element of S
  Remove m from S
  T ← {2m, 4m + 1, 4m + x} ∩ {0, ..., ⌊|x| /3⌋}
  c ← c + #T
  Append T to S
end while
if c = ⌊|x| /3)⌋ + 1 then
  Return(True)
else
  Return(False)
end if
```

It is clear that the run-time of this algorithm is $O(|x|)$. Muir and Stinson [12, 13] give a list of all NADS $\{0, 1, x\}$ with $|x| \leq 10\,000$. The list starts with

$$3, -1, -5, -13, -17, -25, -29, -37, -53, -61, -65, \dots$$

They also gave some necessary and some sufficient conditions on x such that $\{0, 1, x\}$ is a NADS.

We remark that for negative integers n, we always have $r(n) > n$ if $x < 0$. This implies that for some finite positive k, we have $r^k(n) \geq 0$.

Proposition 5. *Let $D := \{0, 1, x\}$ with $x < 0$ be a NADS. Then every integer $n \in \mathbb{Z}$ has a D-NAF.*

[1] In this paper, a *path* in a transducer, in an automaton or in a directed graph is not required to have distinct vertices.

3. Calculating a D-NAF from right to left

Let $D = \{0, 1, x\}$ with $x \equiv 3 \pmod 4$ be fixed throughout this section. Our aim is to give a transducer which transforms the binary expansion of n into its D-NAF from right to left.

Since the least significant digit $\eta_0(n)$ depends on $n \bmod 4$, a transducer will need a look-ahead of 1 in order to be able to make a decision.

The transducer $\mathcal{T}_0$ over the input alphabet $\{0, 1\}$ and the output alphabet D is defined as follows: It has the set of states $Q_0 = \{0, \dots, 2+|x|\} \cup \{\mathcal{I}\}$, where the states $0 \le m \le 2+|x|$ represent carries and $\mathcal{I}$ denotes the initial state. The sets of initial and terminal states are $\{\mathcal{I}\}$ and $\{0\}$, respectively. The set of transitions is

$$E_0 = \{\mathcal{I} \xrightarrow{0|\varepsilon} 0, \mathcal{I} \xrightarrow{1|\varepsilon} 1\} \cup \left\{m \xrightarrow{d|\eta_0(2d+m)} r(2d+m) : 0 \le m \le 2 + |x|\,, d \in \{0,1\}\right\}, \tag{3.1}$$

where ε denotes the empty word. Since

$$r(2d+m) = \frac{2d + m - \eta_0(2d+m)}{2} \le \frac{2 + (2 + |x|) + |x|}{2} = 2 + |x|\,,$$

this finite transducer is well defined. Of course, we only have to consider the accessible states in $\mathcal{T}_0$; we therefore define $\mathcal{T}$ to be the subgraph of $\mathcal{T}_0$ spanned by the accessible states. The sets of states and transitions of $\mathcal{T}$ will be denoted by Q and E, respectively.

Note that this is a right-to-left transducer, i.e., a path $m \xrightarrow{d_{J-1}\cdots d_0|\eta_{J-1}\cdots\eta_0} m'$ is $m \xrightarrow{d_0|\eta_0} m_1 \xrightarrow{d_1|\eta_1} m_2 \xrightarrow{d_2|\eta_2} \cdots \xrightarrow{d_{J-1}|\eta_{J-1}} m'$.

To prove that the transducer indeed calculates a D-NAF of an integer when reading its standard binary expansion, the following lemma is useful.

Lemma 6. (1) *Let $\mathcal{I} \xrightarrow{i_J\cdots i_0|o_{J-1}\cdots o_0} m$ be a path in $\mathcal{T}$. Then*

$$(o_{J-1}, \dots, o_0) \text{ is the } D\text{-NAF of } \mathsf{value}(i_J, \dots, i_0) - m2^J.$$

(2) *If there is a path $m \xrightarrow{d_{J-1}\cdots d_0|\eta_{J-1}\cdots\eta_0} m'$ in $\mathcal{T}$ for some $m \ne \mathcal{I}$, then we have*

$$2^J m' + \mathsf{value}(\eta_{J-1}, \dots, \eta_0) = 2\,\mathsf{value}(d_{J-1}, \dots, d_0) + m. \tag{3.2}$$

Proof. The lemma is easily proved by using the definition of the transducer and induction. □

The following theorem states that the transducer $\mathcal{T}$ does what we promised.

Theorem 7. *Let $D = \{0, 1, x\}$ with $x \equiv 3 \pmod 4$ and $\mathcal{T} = (Q, E, \{\mathcal{I}\}, \{0\})$ the transducer constructed above. Then the following holds:*

1. $\#Q \le |x| + 4$.
2. *An integer n with binary expansion $\mathbf{d}$ has a D-NAF if and only if there is a successful path with input label $(d_{\mathrm{MSB}(\mathbf{d})+\#Q-2}, \dots, d_0)$ in $\mathcal{T}$. In this case, the output label of this successful path is the D-NAF of n.*
3. *The set D is a NADS if and only if the only cycle in $\mathcal{T}$ with input label $0\cdots0$ is $0 \xrightarrow{0|0} 0$.*

Proof. (1) Follows from the definition.

(2) Assume that $n > 0$ has a D-NAF $\boldsymbol{\eta}$ and a standard binary expansion $\mathbf{d}$. Let $J := \max\{\mathrm{MSB}(\boldsymbol{\eta}), \mathrm{MSB}(\mathbf{d})\}$. Consider the path $\mathcal{I} \xrightarrow{d_{J+2}d_{J+1}d_J\cdots d_0|o_{J+1}o_J\cdots o_0} m$ in $\mathcal{T}$. From Lemma 6 we conclude that $n - m2^{J+2}$ has a D-NAF $(o_{J+1}, o_J, \dots, o_0)$. This implies that $\mathsf{value}(\eta_J, \dots, \eta_0) \equiv \mathsf{value}(o_{J+1}, \dots, o_0) \pmod{2^{J+2}}$. Since both expansions are D-NAFs, we infer that $\eta_j = o_j$ for $0 \le j \le J$. This yields $o_{J+1}2^{J+1} = -m2^{J+2}$ from which we conclude that $o_{J+1} = m = 0$. This implies that $\mathcal{I} \xrightarrow{d_{J+2}d_{J+1}d_J\cdots d_0|0\eta_J\cdots\eta_0} 0$ is a successful path in $\mathcal{T}$. We will prove that the length of a successful path is at most $\mathrm{MSB}(\mathbf{d}) + \#Q - 1$ after proving the third part of the theorem.

On the other hand, if there is some successful path, Lemma 6 shows that it corresponds to a D-NAF of the value of its input.

(3) When processing the binary expansion of n on the transducer, we can distinguish between two phases: in the first phase, we read "significant" input, in the second phase, we just read leading zeros of the binary expansion. If we reach the terminal state 0 in this second phase, we are successful and got a D-NAF of our input. However, if we enter a cycle in the second phase apart from the trivial cycle $0 \xrightarrow{0|0} 0$, it is clear that we will never reach the terminal state, i.e., there is no D-NAF.

This implies that after reading $d_{\mathrm{MSB}(\mathbf{d})}$, we will reach each of the states $Q \setminus \{\mathcal{I}, 0\}$ at most once, hence we need at most $\#Q - 2$ leading zeros to reach the terminal state 0.

□

For $x = 3$, $x = -1$, $x = -5$, $x = -9$, and $x = -13$, the transducers $\mathcal{T}$ are shown in Figures 1, 2, 3, 4, and 5, respectively. Note that for $x = -9$, there is a cycle $3 \xrightarrow{0|\bar{9}} 6 \xrightarrow{0|0} 3$, therefore, $\{0, 1, -9\}$ is not a NADS.

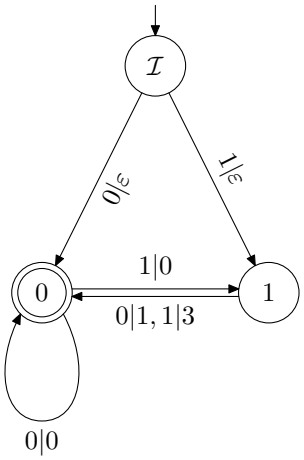

FIGURE 1. Transducer $\mathcal{T}$ for calculating the $\{0, 1, 3\}$-NAF of n from its standard binary expansion from right to left.

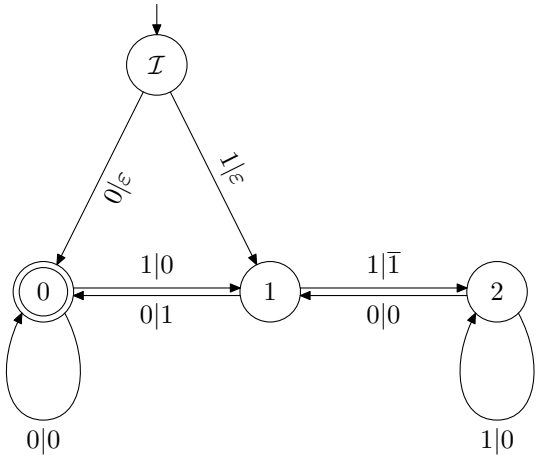

FIGURE 2. Transducer $\mathcal{T}$ for calculating the $\{0, 1, -1\}$-NAF of n from its standard binary expansion from right to left.

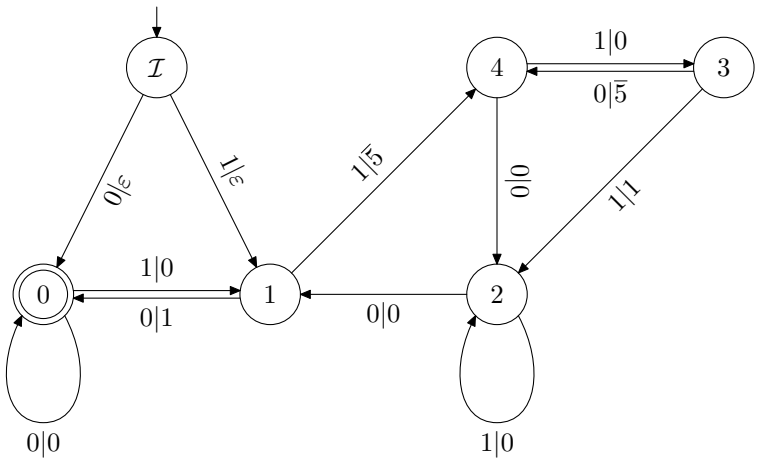

FIGURE 3. Transducer $\mathcal{T}$ for calculating the $\{0, 1, -5\}$-NAF of n from its standard binary expansion from right to left.

In some parts of this paper, we will study the input automaton $\mathcal{A}$ of $\mathcal{T}$, i.e., we only consider the input labels of the transitions in $\mathcal{T}$. By construction, the automaton $\mathcal{A}$ is deterministic. We will use the notation $m' = (d_{J-1} \cdots d_0) \cdot m$ for the transition function in this automaton, which means that there is a path in $\mathcal{A}$ from m to m' with (input) label $(d_{J-1} \cdots d_0)$. Furthermore, we will apply these transitions to sets of states also, i.e. $(d_{J-1} \cdots d_0) \cdot M := \{(d_{J-1} \cdots d_0) \cdot m : m \in M\}$ for $M \subseteq Q$.

The following lemma will be used several times:

Lemma 8. *Let $x < 0$,*

$$k_0 := 2 + \max\{\mathrm{MSB}(\boldsymbol{\eta}(n)) : -2 \le n \le 2 + |x|\} \tag{3.3}$$

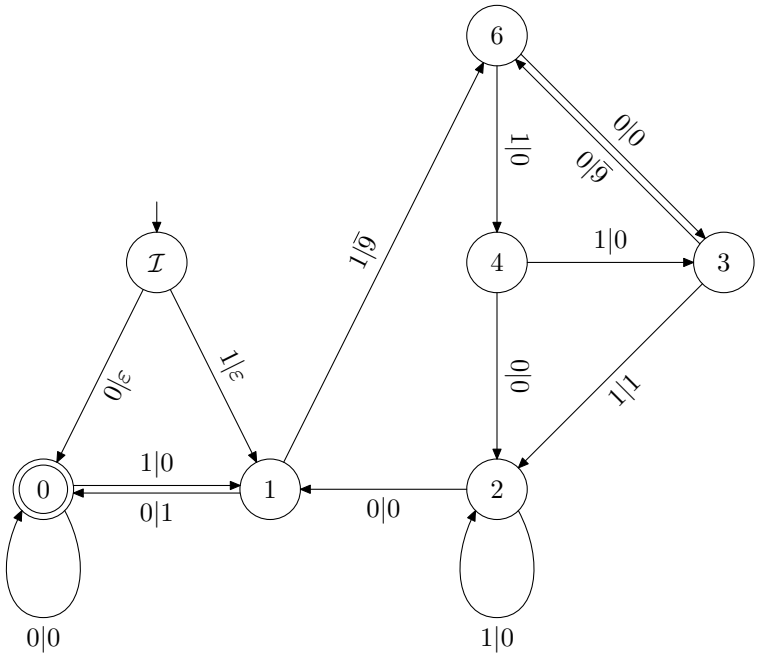

FIGURE 4. Transducer $\mathcal{T}$ for calculating the $\{0,1,-9\}$-NAF of n from its standard binary expansion from right to left.

and $m \in Q \setminus \{\mathcal{I}\}$. Then for any $k \geq k_0$, we have

$$(0^k) \cdot m = 0,$$
$$(1^k) \cdot m = 2,$$

where d^k means the word consisting of k repetitions of the letter d.

Proof. Let $d \in \{0,1\}$. We consider the path

$$m \xrightarrow{d^k|\eta_{k-1}\cdots\eta_0} m'$$

in $\mathcal{T}$. By (3.2), we have

$$2^k m' + \mathsf{value}(\eta_{k-1}\cdots\eta_0) = 2d(2^k-1) + m = d2^{k+1} + m - 2d. \tag{3.4}$$

By definition of k_0, the D-NAF $\boldsymbol{\eta}'$ of $m-2d$ satisfies $\mathrm{MSB}(\boldsymbol{\eta}') \leq k-2$. Then (3.4) implies that $\mathsf{value}(\eta_{k-1}\cdots\eta_0) \equiv \mathsf{value}(\eta'_{k-1}\cdots\eta'_0) \pmod{2^k}$, which yields $\eta_j = \eta'_j$ for $0 \leq j \leq k-2$. Inserting this in (3.4) we immediately see that $\eta_{k-1} = 0$ and $m' = 2d$. □

Furthermore, we can also construct a transducer $\widetilde{\mathcal{T}}_0$ which takes an arbitrary binary $\{0,1,x\}$-expansion and transforms it to the D-NAF: The set of states $\widetilde{Q}_0$ is $\{\mathcal{I}\} \cup \{-2\,|x|,$

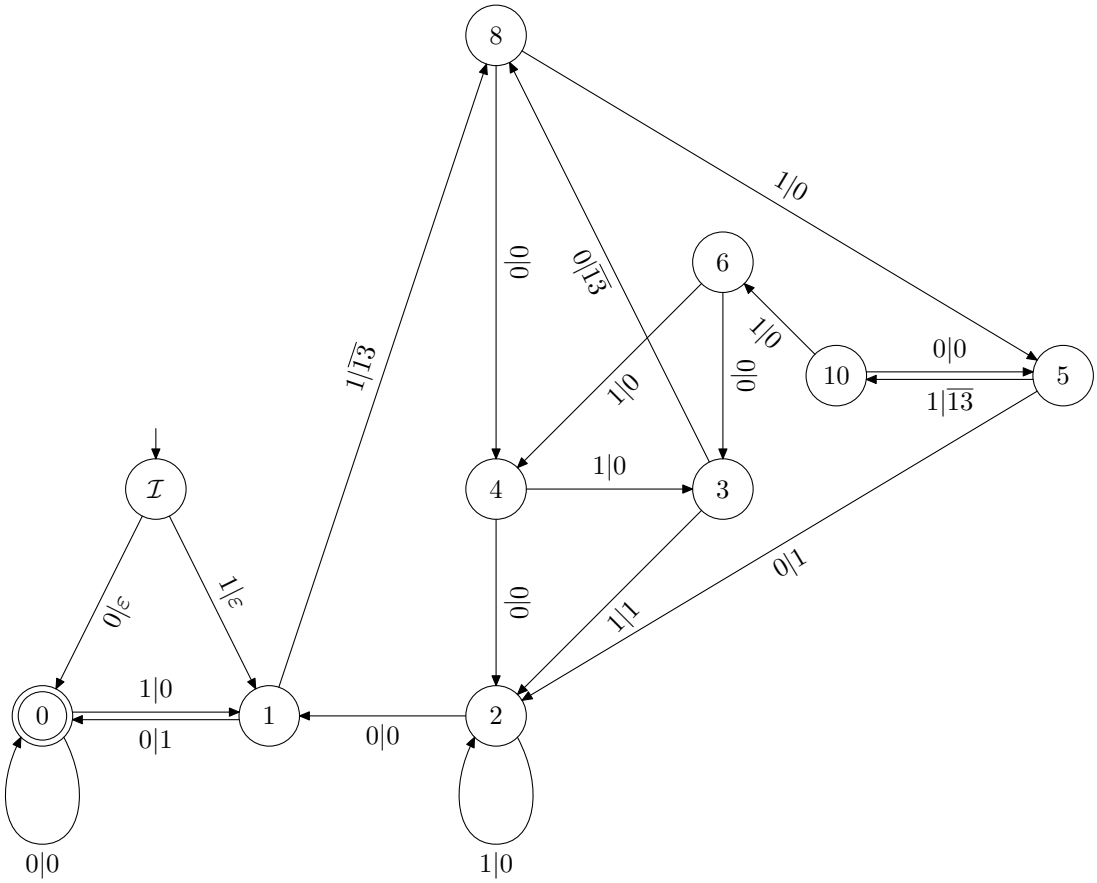

FIGURE 5. Transducer $\mathcal{T}$ for calculating the $\{0, 1, -13\}$-NAF of n from its standard binary expansion from right to left.

$\ldots, |x| + 2\}$, the set of transitions $\widetilde{E}_0$ is

$$\widetilde{E}_0 = \{\mathcal{I} \xrightarrow{0|\varepsilon} 0, \mathcal{I} \xrightarrow{1|\varepsilon} 1, \mathcal{I} \xrightarrow{x|\varepsilon} x\} \cup \left\{m \xrightarrow{d|\eta_0(2d+m)} r(2d+m) : -2|x| \le m \le 2 + |x|, d \in \{0, 1, x\}\right\}. \tag{3.5}$$

The transducer $\widetilde{\mathcal{T}}$ is obtained by removing inaccessible states. The transducers $\widetilde{\mathcal{T}}$ for $x = 3$ and $x = -1$ are shown in Figures 13 and 15, respectively.

4. Frequency of digits

Let $D = \{0, 1, x\}$ be a fixed NADS. We denote the D-NAF of a nonnegative integer n by $\boldsymbol{\eta}(n)$. For $d \in \{1, x\}$, we denote the number of occurrences of the digit d in the D-NAF of $n \in \mathbb{N}_0$ by

$$f_d(n) := \sum_{j \ge 0} [\eta_j(n) = d], \tag{4.1}$$

where we use Iverson's notation $[expr] = 1$ if $expr$ is true and $[expr] = 0$ otherwise, cf. [6].

Let X_N be a uniformly distributed random variable on $\{0, \dots, 2^N - 1\}$. We are interested in the distribution of the random variable $F_{d,N} := f_d(X_N)$. We will first calculate the main terms of mean and variance using a recurrence approach. In a second step, we study the probability generating function and derive a central limit law using Hwang's [8] "quasi power theorem." For concrete values of x, the generating function approach gives also second order terms for the mean and the variance.

We derive recursive formulæ for the first two moments

$$E_d(N) := \mathbb{E}(F_{d,N}) = \frac{1}{2^N} \sum_{n=0}^{2^N-1} f_d(n),$$

$$S_d(N) := \mathbb{E}((F_{d,N})^2) = \frac{1}{2^N} \sum_{n=0}^{2^N-1} f_d^2(n)$$

for $N \geq N_0$, where

$$N_0 := \max\{\mathrm{MSB}(\boldsymbol{\eta}(n)) : 0 \leq n < \lceil -x/4 \rceil\} + 4.$$

We split the sum into three parts, depending on the residue class of n modulo 4, writing $n = 2m$, $n = 4m + 1$, and $n = 4m + x$, respectively:

$$\begin{aligned} E_d(N) &= \frac{1}{2^N} \sum_{m=0}^{2^{N-1}-1} f_d(m) + \frac{1}{2^N} \sum_{m=0}^{2^{N-2}-1} f_d(m) + \frac{1}{2^N} \sum_{m=y}^{2^{N-2}+y-1} f_d(m) + \frac{1}{4} \\ &= \frac{1}{2} E_d(N-1) + \frac{1}{2} E_d(N-2) + \frac{1}{4} \\ &\quad + \frac{1}{2^N} \sum_{m=0}^{y-1} \sum_{j \geq 0} \left(\left[\eta_j(2^{N-2} + m) = d\right] - \left[\eta_j(m) = d\right] \right), \end{aligned}$$

where $y = \lceil -x/4 \rceil$. By our choice of N_0, we conclude that for $0 \leq m < y$, the D-NAF of $2^{N-2} + m$ is simply the D-NAF of m with the digit at position $N - 2$ set to 1, i.e.,

$$\eta_j(2^{N-2} + m) = \eta_j(m) + \left[j = N - 2\right].$$

This yields

$$E_d(N) = \frac{1}{2} E_d(N-1) + \frac{1}{2} E_d(N-2) + \frac{1}{4} + \frac{y\,[d=1]}{2^N}, \qquad N \geq N_0.$$

Solving this linear recurrence with constant coefficients, we get

$$E_d(N) = \mathbb{E}(F_{d,N}) = \frac{1}{6} N + e_d + O\left(\frac{1}{2^N}\right), \tag{4.2}$$

where e_d is some constant.

A recurrence relation for the second moment can be derived similarly, we get

$$S_d(N) = \frac{1}{2} S_d(N-1) + \frac{1}{2} S_d(N-2) + \frac{1}{2} E_d(N-2) + \frac{1}{4} + \frac{C}{2^N}, \qquad N \geq N_0,$$

where $C = [d = 1] \sum_{m=0}^{y-1} (2f_d(m) + 1)$ is a constant. Inserting (4.2) (with undetermined constants), we again get a linear recurrence with constant coefficients. Solving it, we get

$$S_d(N) = \mathbb{E}((F_{d,N})^2) = \frac{1}{36}N^2 + \left(\frac{11}{108} + \frac{e_d}{3}\right) N + (v_d + e_d^2) + O\left(\frac{N}{2^N}\right)$$

for some constant v_d. Subtracting $(E_d(N))^2$, we get

$$\mathbb{V}(F_{d,N}) = \mathbb{E}((F_{d,N})^2) - (\mathbb{E}(F_{d,N}))^2 = \frac{11}{108}N + v_d + O\left(\frac{N}{2^N}\right). \tag{4.3}$$

We now consider the probability generating function

$$G_d(Y, Z) = \sum_{N \geq 0} \sum_{m \geq 0} \mathbb{P}(F_{d,N} = m) Y^m Z^N.$$

We define the $(\#Q \times \#Q)$-matrices $A_d := A_d(Y) = (a_{ij})_{i,j \in Q}$ and $B_d = B_d(Y) = (b_{ij})_{i,j \in Q}$ by

$$a_{ij} = \sum_{\substack{v \in \{0,1\} \\ i \xrightarrow{v|\eta} j \in E}} \frac{1}{2} Y^{[\eta = d]}, \qquad b_{ij} = \begin{cases} Y^{[\eta=d]}, & \text{if there is a transition } i \xrightarrow{0|\eta} j \in E, \\ 0, & \text{otherwise,} \end{cases}$$

where the rows and columns of A_d and B_d are ordered as $\mathcal{I}, 0, 1, \ldots$. Then $G_d(Y, Z)$ can be expressed as

$$G_d(Y, Z) = (1, 0, \ldots, 0)(I - A_d Z)^{-1} B_d^{\#Q-2} (0, 1, 0, \ldots, 0)^t,$$

where the factor $B_d^{\#Q-2}$ ensures that we come back to the terminal state 0. For concrete x, G_d can be calculated explicitly, for instance, we have

$$G_1(Y, Z) = \frac{16 - 8Z + 8YZ - 4Z^2 - YZ^4 + Y^2Z^4}{2(-2 + Z)(-4 + 2Z + Z^2 + YZ^2)},$$
$$G_{-5}(Y, Z) = \frac{-(8 + 4Z - YZ^3 + Y^2Z^3)}{2(-4 + 2Z + Z^2 + YZ^2)}$$

for $x = -5$. Then we have

$$\mathbb{E}(F_{d,N}) = [Z^N] \left(\frac{\partial G_d}{\partial Y}\bigg|_{Y=1} \right),$$
$$\mathbb{V}(F_{d,N}) = [Z^N] \left(\frac{\partial^2 G_d}{\partial Y^2}\bigg|_{Y=1} \right) + \mathbb{E}(F_{d,N}) - (\mathbb{E}(F_{d,N}))^2.$$

For the first ten values of x, we calculated means and variances by this approach. Comparing with (4.2) and (4.3) for the first few values of x gives the constants e_d and v_d in Table 1.

From the definition (3.1) of the transducer $\mathcal{T}$ it is clear that

$$A_d(Y) = \begin{pmatrix} 0 & 1 & 1 & 0 & \ldots & 0 \\ 0 & & & \widetilde{A}_d(Y) & & \end{pmatrix}$$

x	e_1	v_1	e_x	v_x
3	$\frac{2}{9} \approx 0.2222$	$\frac{8}{81} \approx 0.0987$	$-\frac{1}{9} \approx -0.1111$	$-\frac{4}{81} \approx -0.0493$
-1	$\frac{5}{9} \approx 0.5555$	$-\frac{16}{81} \approx -0.1975$	$-\frac{1}{9} \approx -0.1111$	$-\frac{4}{81} \approx -0.0493$
-5	$\frac{23}{36} \approx 0.6388$	$-\frac{253}{1296} \approx -0.1952$	$-\frac{1}{36} \approx -0.0277$	$\frac{155}{1296} \approx 0.1195$
-13	$\frac{467}{576} \approx 0.8107$	$-\frac{61609}{331776} \approx -0.1856$	$-\frac{25}{576} \approx -0.0434$	$\frac{35279}{331776} \approx 0.1063$
-17	$\frac{319}{288} \approx 1.1076$	$-\frac{13825}{82944} \approx -0.1666$	$\frac{71}{576} \approx 0.1232$	$\frac{176783}{331776} \approx 0.5328$
-25	$\frac{7553}{9216} \approx 0.8195$	$-\frac{20629249}{84934656} \approx -0.2428$	$\frac{637}{4608} \approx 0.1382$	$\frac{8343287}{21233664} \approx 0.3929$
-29	$\frac{2183}{2304} \approx 0.9474$	$-\frac{1241137}{5308416} \approx -0.2338$	$-\frac{37}{2304} \approx -0.0160$	$\frac{893351}{5308416} \approx 0.1682$
-37	$\frac{47939}{36864} \approx 1.3004$	$-\frac{197600137}{1358954496} \approx -0.1454$	$\frac{26143}{73728} \approx 0.3545$	$\frac{6190893119}{5435817984} \approx 1.1389$
-53	$\frac{22009}{18432} \approx 1.1940$	$-\frac{54592561}{339738624} \approx -0.1606$	$\frac{14561}{36864} \approx 0.3949$	$\frac{1292596799}{1358954496} \approx 0.9511$
-61	$\frac{10289}{9216} \approx 1.1164$	$-\frac{19291489}{84934656} \approx -0.2271$	$-\frac{187}{9216} \approx -0.0202$	$\frac{13417319}{84934656} \approx 0.1579$

TABLE 1. Constant terms in means and variances of $F_{d,N}$

for some $(\#Q-1)\times(\#Q-1)$-matrix $\widetilde{A}_d$. By definition of $\mathcal{T}$ and Lemma 8, the matrix $\tilde{A}_d$ is the adjacency matrix of a strongly connected directed graph. Moreover, since there is a loop $0 \xrightarrow{0|0} 0$ in $\mathcal{T}$, the matrix $\widetilde{A}_d^\ell(1)$ is positive for some positive integer ℓ. Therefore, $\widetilde{A}_d(1)$ is primitive (cf. for instance [11, §15.6]) and has a unique simple eigenvalue of maximal modulus by the Perron-Frobenius theorem. Since the spectrum $\sigma(A_d(1)) = \sigma(\widetilde{A}_d(1)) \cup \{0\}$, the same holds for $A_d(1)$ and, by continuity, for $A_d(Y)$ for Y in some neighbourhood of 1.

Thus, the generating function G_d is a rational function whose numerator has a unique simple pole of minimal modulus for Y sufficiently close to 1. We obtain

$$\mathbb{E}(e^{F_{d,N}s}) = e^{u(s)N+v(s)}\left(1+O\left(\rho^N\right)\right)$$

for some analytic functions $u(s)$ and $v(s)$ and some $0<\rho<1$ for $|s|$ sufficiently small. By Hwang's [8] "quasi power theorem" we conclude the central limit theorem (4.4).

Theorem 9. *Let $D=\{0,1,x\}$ be a NADS, $d\in\{1,x\}$ and $F_{d,N}$ the number of occurrences of the digit d in the D-NAF of a randomly chosen integer in the interval $\{0,\dots,2^N-1\}$. Then we have*

$$\mathbb{E}(F_{d,N}) = \frac{1}{6}N + e_d + O\left(\frac{1}{2^N}\right),$$
$$\mathbb{V}(F_{d,N}) = \frac{11}{108}N + v_d + O\left(\frac{N}{2^N}\right)$$

for some constants e_d and v_d depending on x, which can be computed explicitly. For $|x|\le 61$, they are given in Table 1.

Furthermore, we have

$$\mathbb{P}\left(F_{d,N} \le \frac{1}{6}N + z\sqrt{\frac{11}{108}N}\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-t^2/2}\,dt + O\left(\frac{1}{\sqrt{N}}\right) \tag{4.4}$$

uniformly with respect to z, $z \in \mathbb{R}$.

5. "Calculating" Digits from Left To Right

The aim of this section is to give a description of arbitrary digits in a NADS $D = \{0, 1, x\}$. If n has standard binary expansion $(\dots d_{\ell+1}d_\ell d_{\ell-1}\cdots d_0)$ and D-NAF $(\dots \eta_{\ell+1}\eta_\ell\eta_{\ell-1}\cdots\eta_0)$, then η_ℓ depends on $n \bmod 2^{\ell+2}$ only. Alternatively, η_ℓ depends on $\left\{n/2^{\ell+2}\right\}$, where $\{z\}$ denotes the fractional part of z.

In Figures 7, 8, 9, and 10, we draw the second digit from the left for $x = 3$, -1, -5, and -13 respectively: For $k = 3, \dots, 12$ and $n = 0, \dots, 2^k - 1$, we calculated the ℓth digit of the D-NAF of all integers m such that

$$\frac{n}{2^k} \le \left\{\frac{m}{2^{\ell+2}}\right\} < \frac{n+1}{2^k};$$

these are those integers whose first (from the left!) digits in the standard binary expansion agree with the first digits of n. In some cases, the ℓth digit was the same for all integers m which lie in the given interval, in some cases, several digits could occur. We mapped the sets of possible digits to colors according to Figure 6 and filled a rectangle of width $1/2^k$ and height 1 with this color at position $(n/2^k, 12 - k)$. We remark that for $x = -1$, the picture is rather regular, whereas for $x = -13$, many decisions are still open after 12 digits.

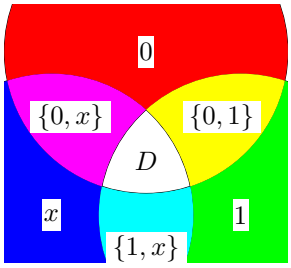

Figure 6. Colors for Figures 7–10.

In what follows, we will describe this phenomenon in terms of the states of the transducer $\mathcal{T}$.

Since the digit η_ℓ depends on the state $m_{\ell+1} = (d_\ell \cdots d_0) \cdot \mathcal{I}$ and on $d_{\ell+1}$, we will try to predict the state $m_{\ell+1}$ from the knowledge of $d_\ell \cdots d_{\ell-r+1}$ for small r, wherever possible. Since we do not know in which state we are after reading the unknown digits $d_{\ell-r} \cdots d_0$, we have to assume that we are in any state (apart from the initial state). So we denote the set $Q \setminus \{\mathcal{I}\}$ by Q^*.

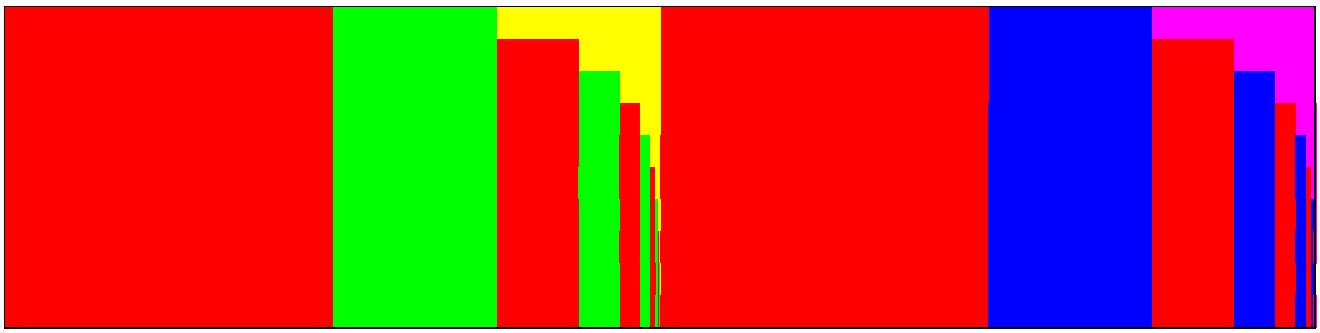

FIGURE 7. Second digit from the left for $x = 3$.

FIGURE 8. Second digit from the left for $x = -1$.

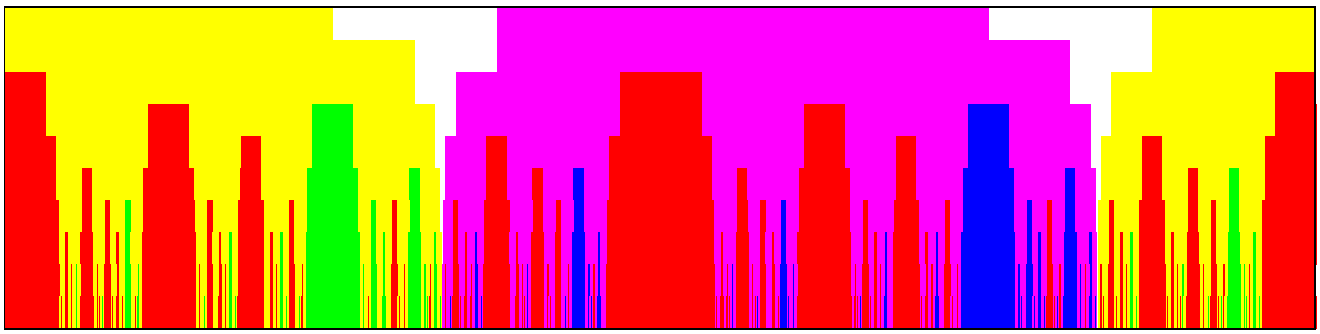

FIGURE 9. Second digit from the left for $x = -5$.

We define the map $\Phi_r : \{0,1\}^r \to Q^*$ by

$$\Phi_r(x_1x_2\cdots x_r) := (x_1x_2\cdots x_r)\cdot Q^*.$$

It is clear that $\Phi_{r+1}(x_1x_2\cdots x_rx_{r+1}) \subseteq \Phi_r(x_1x_2\cdots x_r)$. Therefore, for any $\mathbf{x} = (x_1x_2\cdots) \in \{0,1\}^{\mathbb{N}}$, the limit

$$\Phi(x_1x_2\cdots) := \lim_{r\to\infty} \Phi_r(x_1\dots x_r)$$

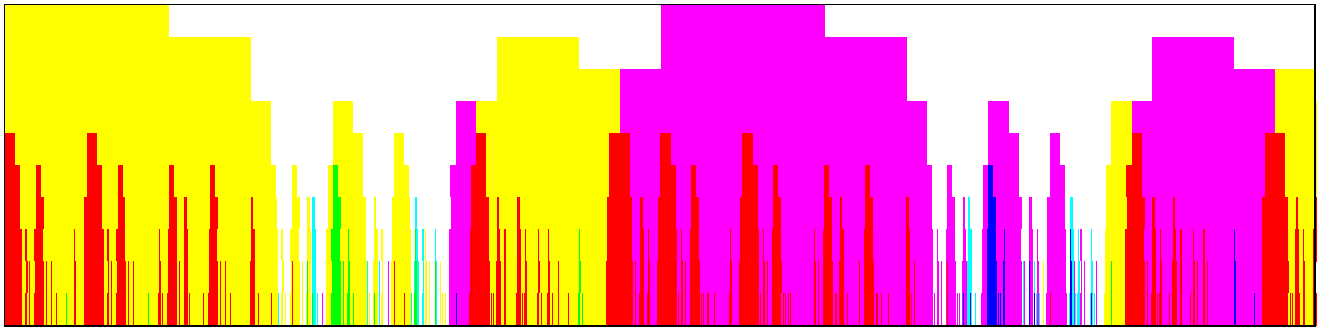

FIGURE 10. Second digit from the left for $x = -13$.

exists. We note that if $\Phi(\mathbf{x})$ is a singleton $\{m\}$ for some $\mathbf{x} \in \{0,1\}^{\mathbb{N}}$, we have $\Phi_r(x_1x_2\cdots x_r) = \{m\}$ for some $r \in \mathbb{N}$. In this case, we have $\Phi(\mathbf{y}) = \Phi(\mathbf{x})$ for any $\mathbf{y} \in \{0,1\}^{\mathbb{N}}$ with the property that the prefixes of length r of $\mathbf{x}$ and $\mathbf{y}$ coincide. This means that Φ is continuous in all points where it has a singleton image.

We now assume $x < 0$. From Lemma 8, we see that for any $(x_1, \ldots, x_j) \in \{0,1\}^j$ and $k \geq k_0$, we have

$$\begin{aligned}\Phi_{k+j+1}(x_1\cdots x_j10^k) &= (x_1\cdots x_j1)\cdot\{0\} = (x_1\cdots x_j)\cdot\{1\}\\ &= (x_1\cdots x_j0)\cdot\{2\} = \Phi_{k+j+1}(x_1\cdots x_j01^k).\end{aligned}$$

This implies that $\Phi(x_1x_2\cdots x_j0111\cdots) = \Phi(x_1x_2\cdots x_j1000\cdots)$.

Therefore, Φ descends to a function on the unit interval $[0,1]$ (which we call Φ again) by

$$\Phi\left(\sum_{j\geq 1} x_j 2^{-j}\right) := \Phi(x_1x_2\cdots).$$

This function is continuous at all points $z \in [0,1]$ where $\Phi(z)$ is a singleton. This implies that the sets $A_j := \Phi^{-1}(\{j\})$ are open for $j \in Q^*$. The set $A = \bigcup_{j\in Q^*} A_j$ is therefore open. Since the set of real numbers in $[0,1]$ whose binary expansion contains a block (0^{k_0}) has Lebesgue measure 1 and is a subset of A by Lemma 8, we conclude that $\lambda(A) = 1$, too.

We translate our results back to the ℓth digit: We have $\eta_\ell(n) = \eta_0(2d_{\ell+1} + m_{\ell+1}) = \eta$ for $m_{\ell+1} = (d_\ell\cdots d_0)\cdot\mathcal{I}$ and the standard binary expansion $\mathbf{d}$. The relation $m_{\ell+1} = (d_\ell\cdots d_0)\cdot\mathcal{I}$ is equivalent to $\{n/2^{\ell+1}\} \in A_{m_{\ell+1}}$, since $\Phi_{k_0+\ell+1}(\{n/2^{\ell+1}\})$ is a singleton by Lemma 8. To include $d_{\ell+1}$ in our description, we define $\varphi_d(z) := (d+z)/2$ for $d \in \{0,1\}$ and see that

$$\left\{\frac{n}{2^{\ell+2}}\right\} \in \varphi_{d_{\ell+1}}(A_{m_{\ell+1}}).$$

We now collect all sets $\varphi_d(A_m)$ which lead to the same digit $\eta \in D$ in the set

$$W_\eta := \operatorname{int}\left(\overline{\bigcup_{\substack{d\in\{0,1\}\\ m\in Q^*\\ \eta_0(2d+m)=\eta}} \varphi_d(A_m)}\right),$$

where the bar denotes closure and int denotes the interior. We take the interior of the closure for "aesthetical reasons": We want to avoid "holes" in the sets which would only be meaningful at the level of the transducer, but not on the level of the digits. Since $\{n/2^{\ell+1}\}$ is in the interior of some A_m anyway, the dyadic points are not affected by this operation.

This yields the following theorem.

Theorem 10. *Let $D = \{0, 1, x\}$ be a NADS with $x < 0$. There are disjoint open subsets W_η, $\eta \in D$, of the unit interval $[0, 1]$ such that*

$$\eta_\ell(n) = \eta \iff \{n/2^{\ell+2}\} \in W_\eta$$

for $\ell \geq 0$. The sum of the Lebesgue measures of the W_η, $\eta \in D$, equals 1.

6. Dimension of the boundary

Theorem 10 shows that

$$[0, 1] \setminus (W_0 \cup W_1 \cup W_x) = \partial(W_0 \cup W_1 \cup W_x) = \partial W_0 \cup \partial W_1 \cup \partial W_x$$

has Lebesgue measure zero. However, Figures 9 and 10 demonstrate that this "exceptional set" is quite "irregular." The aim of this section is to quantify this "irregularity" in terms of Hausdorff dimension.

We will calculate the Hausdorff dimension of $\partial(W_0 \cup W_1 \cup W_x)$ as the spectral radius of the adjacency matrix of an auxiliary automaton as follows.

We construct auxiliary automata $\mathcal{A}_m$ for $m \geq 2$. The set of states is the set $Q_m := \{I \subseteq Q^* : \#I = m\}$ of m element subsets of Q^*. The alphabet is still $\{0, 1\}$. The set of transitions

$$E_m := \{I \xrightarrow{d} J : J = d \cdot I = \{d \cdot i : i \in I\}, d \in \{0, 1\}, I, J \in Q_m\}$$

is defined in terms of the transitions in the input automaton underlying $\mathcal{T}$. All states are initial and terminal states. The adjacency matrix of $\mathcal{A}_m$ will be denoted by M_m. For $x = -5$ and $m = 2$, this automaton is shown in Figure 11.

We will mainly work with $\mathcal{A}_2$, which is the only automaton to occur in the statement of our result, however in some cases, we may be forced to use $\mathcal{A}_m$ for some $m > 2$ instead.

The remaining part of this section is devoted to the proof of the following theorem.

Theorem 11. *Let $D = \{0, 1, x\}$ with $x < 0$ be a NADS and W_η, $\eta \in D$, be the sets described in Theorem 10. Let $\rho(M_2)$ denote the spectral radius of the adjacency matrix M_2 of $\mathcal{A}_2$. Then we have*

$$\dim_H(\partial(W_0 \cup W_1 \cup W_x)) = \log_2 \rho(M_2)$$

and

$$0 < \mathcal{H}^{\log_2 \rho(M_2)}(\partial(W_0 \cup W_1 \cup W_x)) < \infty.$$

For $x = -1$, we have $\log_2 \rho(M_2) = 0$, for $|x| \geq 5$, we have

$$0 < 0.347121 - \frac{1.04136}{\ell + 2} \leq \log_2 \rho(M_2) \leq 1 - \frac{1}{2^{k_0} \log 2} + \frac{1}{2^{2k_0+1} \log 2} < 1,$$

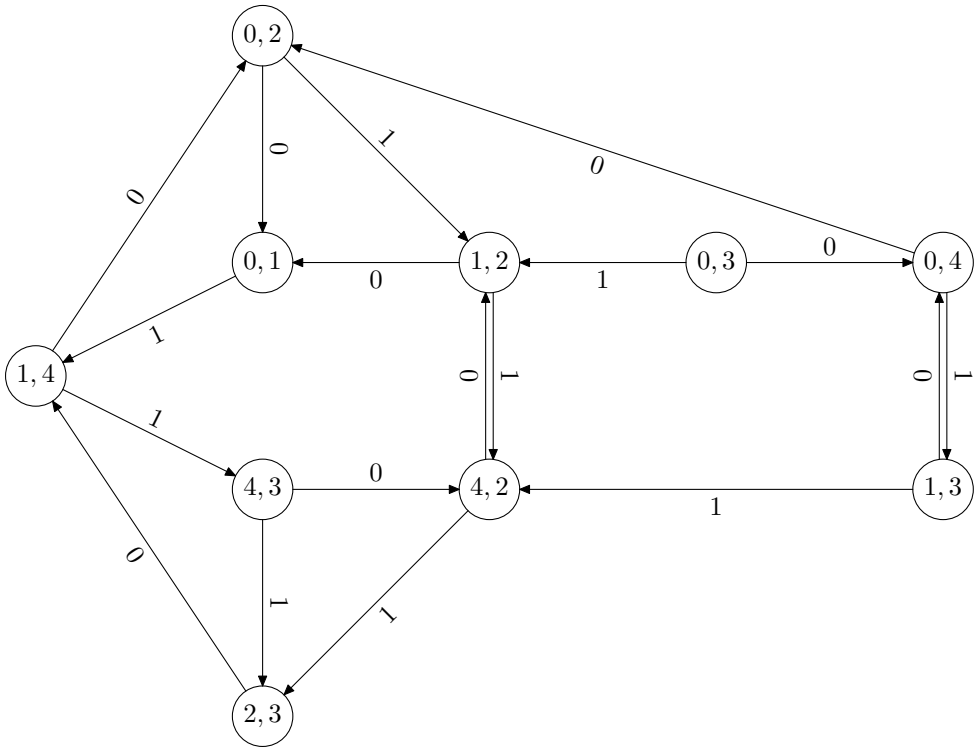

FIGURE 11. Automaton $\mathcal{A}_2$ for $x = -5$.

where k_0 has been defined in (3.3) *and* $\ell = \lfloor \log_2(|x| + 3) \rfloor - 1$.

For $|x| \leq 29$, the values $\log_2 \rho(M_2)$ are shown in Table 2.

The transitions in $\mathcal{A}_m$ will again be denoted by $J = d \cdot I$, since this coincides with the earlier definition for subsets of Q^*. However, this transition is not defined in $\mathcal{A}_m$ for every d and I, since it might be the case that I is mapped to some J with $\#J < m$. We will also reuse the notation $\mathbf{d} \cdot I$ for strings $\mathbf{d} \in \{0, 1\}^n$.

As in the previous section, we first study the problem at the level of the states of the transducer and will translate them back to the level of output digits afterwards. We consider the set $B := [0, 1] \setminus A$ of those real numbers which do not allow us to decide about the final state from the knowledge of a finite number of digits. To calculate the Hausdorff dimension of B for general x, we proceed as follows: We first derive an upper bound for the box dimension using a suitable covering of B. In a second step, we construct a lower bound for the box dimension. Finally we will show that a subset of B can be interpreted as a finite union of graph directed sets (cf. Falconer [4]) satisfying an open set condition, which implies (cf. Edgar [3]) that the Hausdorff, box and similarity dimensions are equal. Finally, we use this fact for calculating the Hausdorff Dimension of B explicitly for small values of $|x|$.

x	$\#Q$	$\#Q_2$	$r(\mathcal{A}_2)$	$\dim_H(\partial(W_0 \cup W_1 \cup W_x))$
-1	3	3	2	0
-5	5	10	3	0.7618
-13	9	36	7	0.9003
-17	10	45	6	0.9037
-25	14	91	17	0.9364
-29	14	91	8	0.9357
-37	18	153	34	0.9492
-53	22	231	25	0.9606
-61	25	300	55	0.9547
-65	25	300	33	0.9477
-113	37	666	79	0.9733
-121	41	820	147	0.9729
-125	39	741	83	0.9649
-137	42	861	98	0.9757
-145	48	1128	228	0.9763
-149	46	1035	130	0.9790

TABLE 2. Order of $\mathcal{T}$ and of $\mathcal{A}_2$, number $r(\mathcal{A}_2)$ of strongly connected components of $\mathcal{A}_2$, and Hausdorff dimension of $\partial(W_0 \cup W_1 \cup W_x)$.

So we start by giving an upper bound for the upper box dimension

$$\overline{\dim}_B B = \limsup_{n\to\infty} \frac{\log_2 N_{2^{-n}}(B)}{n},$$

where $N_{2^{-n}}(B)$ denotes the number of 2^{-n}-mesh intervals that intersect B. We have

$$N_{2^{-n}}(B) = \#\{\mathbf{d} \in \{0,1\}^n : \Phi_n(\mathbf{d}) \text{ is not a singleton}\}, \tag{6.1}$$

since if $\Phi_n(\mathbf{d})$ is not a singleton, there are two points z_1 and z_2 in the interval $[\sum_j d_j 2^{-j}, \sum_j d_j 2^{-j} + 2^{-n}]$ lying in disjoint open sets A_{m_1} and A_{m_2}. Therefore, there must be a point of B between z_1 and z_2.

Lemma 12. *For k_0 defined in* (3.3)*, we have*

$$\overline{\dim}_B(B) \le 1 - \frac{1}{2^{k_0}\log 2} + \frac{1}{2^{2k_0+1}\log 2}. \tag{6.2}$$

Proof. By Lemma 8, we have $N_{2^{-n}}(B) \le \#U_n$, where

$$U_n = \big\{(x_1,\dots,x_n) \in \{0,1\}^n : (x_j,\dots,x_{j+k_0-1}) \notin \{(0^{k_0}),(1^{k_0})\} \text{ for all } 1 \le j \le n-k_0+1\big\}.$$

The strings in U_n can be described by a regular expression

$$(\varepsilon + 0 + \cdots + 0^{k_0-1})\big((1 + 1^2 + \cdots + 1^{k_0-1})(0 + 0^2 + \cdots + 0^{k_0-1})\big)^*(\varepsilon + 1 + \cdots + 1^{k_0-1}),$$

which can be translated to the generating function

$$\begin{aligned} G(z) &:= \sum_{n\ge 0} \#U_n z^n \\ &= (1+\cdots+z^{k_0-1})\cdot\frac{1}{1-(z+\cdots+z^{k_0-1})^2}\cdot(1+\cdots+z^{k_0-1}) = \frac{1-z^{k_0}}{1-2z+z^{k_0}}. \end{aligned}$$

Let $q(z) := 1-2z+z^{k_0}$. We note that we have $k_0 \ge 4$ by (3.3). Then $|q(z)-(1-2z)| = |z^{k_0}| < 2|z|-1 \le |1-2z|$ for $1-\delta < |z| < 1$ for a suitable δ. Therefore, $q(z)$ has exactly one zero with modulus less than 1, say ρ, by Rouché's Theorem. Since $q(1/2) > 0$ and $q(1/2+1/(2k_0)) < 0$, we know that ρ is real and $\rho = 1/2 + O({k_0}^{-1})$. By bootstrapping, we obtain

$$\rho = \frac{1+\rho^{k_0}}{2} = \frac{1}{2} + O(2^{-k_0}), \qquad \rho = \frac{1}{2} + \frac{1}{2^{k_0+1}} + O\left(\frac{k_0}{4^{k_0}}\right).$$

In fact, we have

$$\rho \ge \frac{1}{2} + \frac{1}{2^{k_0+1}} \tag{6.3}$$

for $k_0 \ge 4$, since $q(2^{-1}+2^{-(k_0+1)}) > 0$.

Defining the constant

$$c_{k_0} := \lim_{z\to\rho} G(z)\left(1-\frac{z}{\rho}\right),$$

we have

$$\#U_n = [z^n]G(z) = \frac{c_{k_0}}{\rho^n} + O(1).$$

Therefore, we conclude that the upper box dimension of B can be bounded from above by $\overline{\dim}_B(B) \le -\log_2\rho$, hence (6.3) yields (6.2). □

We now derive lower bounds for the box dimension.

Lemma 13. *Let $|x| \ge 5$. Then we have*

(6.4)

$$\underline{\dim}_B(B) = \liminf_{n\to\infty}\frac{\log_2 N_{2^{-n}}(B)}{n} \ge \left(\frac{1}{2}-\frac{3}{2\ell+4}\right)\log_2\frac{1+\sqrt{5}}{2} \approx 0.347121 - \frac{1.04136}{\ell+2} > 0,$$

where $\ell = \lfloor \log_2(|x|+3)\rfloor - 1$.

Proof. From the definition of $\mathcal{T}$, we deduce that 0 and 1 always have the neighbourhood shown in Figure 12, where $s := (3+|x|)/2$. We further notice that any path from s to 1 has at least length $\ell = \lfloor \log_2 s\rfloor$, since $m \ge 2^i$ implies $r(2d+m) \ge 2^{i-1}$ for $d \in \{0,1\}$. We consider the set of sequences

$$\begin{aligned} L_n := \{\mathbf{d} = (d_n,\dots,d_1) \in \{0,1\}^n : d_i d_{i+1} = 0 \text{ for } 1 \le i < n \\ \text{and if } (d_i\cdots d_1)\cdot s = 1 \text{ then } d_{i+1} = 1 \text{ for } 1 \le i < n\}. \end{aligned} \tag{6.5}$$

We claim that $\mathbf{d}\cdot s \ne \mathbf{d}\cdot 0$ for all $\mathbf{d} \in L_n$. To prove this, we observe that $\mathbf{d}\cdot 0 = d_n$. For any $1 \le i \le n$, we have $(d_i\cdots d_1)\cdot s \ge 1$ by definition (and the known neighbourhood of 0

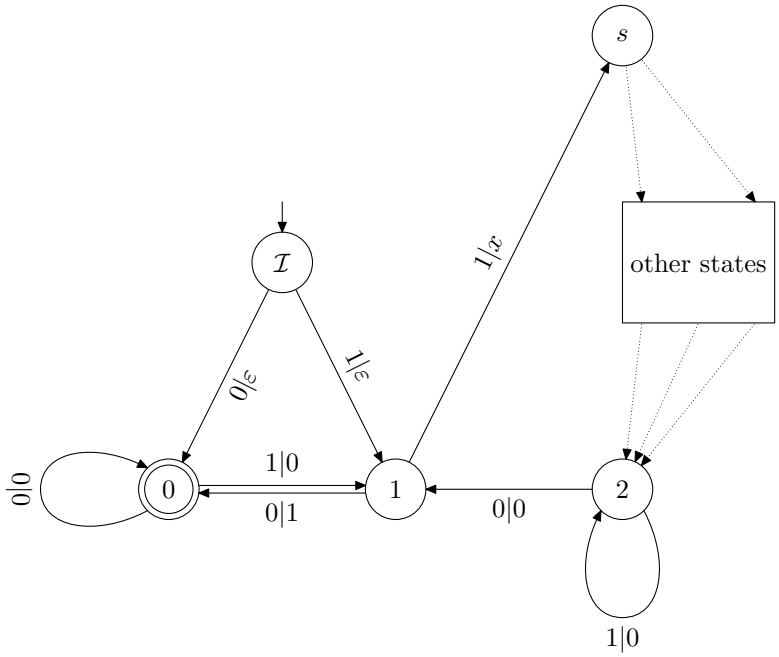

FIGURE 12. Structure of $\mathcal{T}$ for general x (Neighbourhoud of $\mathcal{I}$, 0, 1 is correct.)

and 1). If $(d_i \cdots d_1) \cdot s = 1$, we have $d_i = 0$, which implies $(d_i \cdots d_1) \cdot 0 = 0$. This proves the claim. By (6.1), this implies that $N_{2^{-n}}(B) \geq \#L_n$.

We now derive a lower bound for $\#L_n$. For $1 \leq t \leq \ell$ we have

$$L_t := \{(d_t, \dots, d_1) \in \{0,1\}^t : d_i d_{i+1} = 0 \text{ for } 1 \leq i < t\},$$

since the second condition is no restriction for sequences of that length. The regular expression for sequences which avoid adjacent 1s is $0^*(100^*)^*(\varepsilon + 1)$, the corresponding generating function equals

$$\frac{1}{1-z}\frac{1}{1-\frac{z^2}{1-z}}(1+z) = \frac{1+z}{1-z-z^2} = \sum_{t\geq 0} F_{t+2} z^t,$$

where F_t denotes the Fibonacci sequence $F_{t+2} = F_{t+1} + F_t$ with $F_0 = 0$ and $F_1 = 1$. Therefore, we get

$$\#L_t = F_{t+2} = \frac{\alpha^{t+2}}{\sqrt{5}} - O(\alpha^{-t}) \geq \alpha^t.$$

for $\alpha = (1+\sqrt{5})/2$ and $1 \leq t \leq \ell$. For $\mathbf{d} = (d_{\ell-2}, \dots, d_0) \in L_{\ell-1}$ we define

$$\psi(\mathbf{d}) := 1\mathbf{h}((\mathbf{d}0) \cdot s)\mathbf{d}0,$$

where $\mathbf{h}(m)$ is the input label of a shortest path from state m to state 1 whose input label does not start (at the right) with 1 and does not contain adjacent ones. We note that if $(\mathbf{d}0) \cdot s = 1$, we have $d_{\ell-2} = 0$, $\mathbf{h}(1) = \varepsilon$, and $\psi(\mathbf{d})$ does not contain adjacent ones.

If $m \neq 0$, we write $\mathbf{h}(m) = \mathbf{h}'(0 \cdot m)$, where $\mathbf{h}'(m')$ is the input label of a shortest path from state m' to state 1 whose input label does not contain adjacent ones. We claim that

$$\lfloor \log_2 m' \rfloor \leq |\mathbf{h}'(m')| \leq \lfloor \log_2 m' \rfloor + 1, \tag{6.6}$$

where $|\mathbf{h}'(m')|$ denotes the length of $\mathbf{h}'(m')$. The left inequality follows from the argument given before (6.5). For the right inequality, we choose the input digit d to be 0 unless $m' \equiv 3 \pmod 4$, in which case we take $d = 1$. Thus $\eta_0(2d + m') \in \{0, 1\}$ and $r(2d + m') = d + \lfloor m'/2 \rfloor$. If we choose $d = 1$, it is clear that $r(2d + m')$ is even so that we will not choose input 1 in the next step. Furthermore, we have $\lfloor \log_2(r(2d + m')) \rfloor = \lfloor \log_2(m') \rfloor - 1$ unless $m' = 2^g - 1$ for some g. But in the latter case, we have $r(2d + m') = 2^{g-1}$ with $|\mathbf{h}'(m')| = g$. Therefore, (6.6) follows by induction.

We conclude that

$$|\mathbf{h}'(m')| \leq \lfloor \log_2 m' \rfloor + 1 \leq \lfloor \log_2(|x| + 2) \rfloor + 1 \leq \ell + 2$$

and therefore $|\psi(\mathbf{d})| \leq 2\ell + 4$. We now define an injective map $L_{\ell-1}^{\lfloor n/(2\ell+4) \rfloor} \to L_n$ by

$$(\mathbf{d}_{\lfloor n/(2\ell+4) \rfloor}, \ldots, \mathbf{d}_1) \mapsto \mathbf{d}_0\Bigg(n - \sum_{j=1}^{\lfloor n/(2\ell+4) \rfloor} |\psi(\mathbf{d}_j)|\Bigg)\psi(\mathbf{d}_{\lfloor n/(2\ell+4) \rfloor}) \cdots \psi(\mathbf{d}_1),$$

where $\mathbf{d}_0(m) \in L_m$ is some fixed admissible string of length m which is used to obtain the required length n. This yields

$$\#L_n \geq \alpha^{(\ell-1)\lfloor n/(2\ell+4) \rfloor} \geq \frac{1}{\alpha^{\ell-1}} \cdot \left(\alpha^{(\ell-1)/(2\ell+4)}\right)^n.$$

For $|x| \geq 5$, the bound (6.4) follows. □

We want to prove that the lower box, upper box, Hausdorff and similarity dimensions of B agree. To this aim, we first collect a few simple consequences of Perron-Frobenius theory.

We denote the spectral radius of a $(n \times n)$-matrix M by $\rho(M)$. If $I, J \subseteq \{1, \ldots, n\}$, we denote the submatrix of M with rows I and columns J by $M(I, J)$.

If $M = (m_{ij})_{1 \leq i,j \leq n}$ is a nonnegative $(n \times n)$-matrix, we will consider the digraph G induced by M, i.e., the directed graph with set of vertices $\{1, \ldots, n\}$ and set of arcs $\{(i, j) : m_{ij} > 0\}$. We will call any strongly connected component $\mathcal{C} \subseteq \{1, \ldots, n\}$ of G which fulfills

$$\rho(M(\mathcal{C}, \mathcal{C})) = \rho(M) \tag{6.7}$$

a *dominating component* of M. We will identify subsets of $\{1, \ldots, n\}$ and the subgraph of G induced by these vertices when speaking about strongly connected components.

Lemma 14. *Let M be a nonnegative $(n \times n)$-matrix. Then there is a nonnegative vector $0 \neq x \in \mathbb{R}^n$ such that*

$$Mx = \rho(M)x.$$

Let G be the directed graph induced by M. Then M has a dominating component.

Proof. By continuity, the assertion on the nonnegative eigenvector follows from the Perron-Frobenius theorem, cf. [11, Theorem 15.5.1], for instance.

Let $\mathcal{C}_1, \dots, \mathcal{C}_r$ be the strongly connected components of G. We consider the auxiliary digraph G' with set of vertices $\{\mathcal{C}_1, \dots, \mathcal{C}_r\}$ and an arc from $\mathcal{C}_i$ to $\mathcal{C}_j$ if there is an arc from some vertex in $\mathcal{C}_i$ to some vertex in $\mathcal{C}_j$ in the original digraph G. By construction, the auxiliary digraph G' has no directed cycle. This implies that we may sort the components in such a way that the adjacency matrix of G' is upper triagonal. If we permute the rows and columns of M according to the strongly connected components, we get a matrix which is block upper triagonal. Therefore, the spectrum $\sigma(M)$ of M equals

$$\sigma(M) = \bigcup_{i=1}^{r} \sigma(M(\mathcal{C}_i, \mathcal{C}_i)).$$

This implies that $\rho(M) = \max_{i=1,\dots,r} \rho(M(\mathcal{C}_i, \mathcal{C}_i))$. This proves the existence of a dominating component. □

Lemma 15. *Let M be a nonnegative $(n \times n)$-matrix and $e = (1, \dots, 1)^t \in \mathbb{R}^n$. Then*

$$\lim_{m\to\infty} \frac{\log_2(e^t M^m e)}{m} = \log_2 \rho(M),$$

which may equal $-\infty$.

Proof. Each entry of M^m is bounded by a constant times $m^{n-1}\rho(M)^m$, so the same holds for the nonnegative number $e^t M^m e$.

On the other hand, let x be a nonnegative eigenvector of M for $\rho(M)$, which exists by Lemma 14. Without loss of generality, $0 \le x_j \le 1$ for $j = 1, \dots, n$. Therefore,

$$e^t M^m e \ge x^t M^m x = \rho(M)^m x^t x.$$

Taking logarithms and limits yields the result. □

In general, we cannot expect the automaton $\mathcal{A}_m$ to be strongly connected. Therefore, we cannot apply the results on graph directed sets directly, but we have to apply it on the strongly connected components. To cover the remaining parts of B, we give an upper bound for the upper box dimension (and therefore for the Hausdorff dimension) using the spectral radius $\rho(M_2)$ of M_2 at this point.

Lemma 16. *We have*

$$\dim_H(B) \le \overline{\dim}_B(B) \le \log_2 \rho(M_2).$$

For $|x| \ge 5$, we have $\rho(M_2) > 1$.

Proof. Consider $\mathbf{d} \in \{0,1\}^n$ such that $\Phi_n(\mathbf{d})$ is not a singleton. Then there are $u \ne v \in Q^*$ such that $\mathbf{d} \cdot u \ne \mathbf{d} \cdot v$, i.e., there is at least one path in $\mathcal{A}_2$ from some $\{u, v\} \in Q_2$ to some $\{u', v'\} \in Q_2$ with label $\mathbf{d}$. Thus the number $N_{2^{-n}}(B)$ can be bounded from above by the number of paths of length n in $\mathcal{A}_2$, which equals $e^t M_2^n e$. From Lemma 15 we conclude that

$$\overline{\dim}_B(B) = \limsup_{n\to\infty} \frac{\log_2 N_{2^{-n}}(B)}{n} \le \limsup_{n\to\infty} \frac{\log_2 e^t M_2^n e}{n} = \log_2 \rho(M_2).$$

It is well known that the Hausdorff dimension is always less or equal the lower (and therefore the upper) box dimension.

Since $\overline{\dim}_B(B) > 0$ for $|x| \geq 5$ by Lemma 13, we conclude that $\rho(M_2) > 1$ in that case. □

Proof of Theorem 11. We assume that $|x| \geq 5$.

Since $0 \notin 1 \cdot Q^*$ (cf. Figure 12), the automaton $\mathcal{A}_{\#Q^*}$ has one state and at most one transition, hence $\rho(M_{\#Q^*}) \leq 1 < \rho(M_2)$ by Lemma 16. We now choose $m \geq 2$ maximal such that $\rho(M_m) \geq \rho(M_2)$. The preceeding observation implies that $m < \#Q^*$.

We define sets $V_I := \{x \in [0,1] : I \subseteq \Phi(x)\}$ for $I \in Q_m$. By definition of Φ, these sets are compact. We consider the contracting similarities $\varphi_d(z) := (d+z)/2$ for $d \in \{0,1\}$. We clearly have $\Phi(\varphi_d(z)) = d \cdot \Phi(z)$ for $d \in \{0,1\}$ and $z \in [0,1]$. Then it is easily proved that

$$V_I = \bigcup_{\substack{d\in\{0,1\}\\ J\in Q_m\\ I=d\cdot J}} \varphi_d(V_J), \qquad I \in Q_m. \tag{6.8}$$

Let now $\mathcal{C} \subseteq Q_m$ be a dominant component of M_m, i.e., $\rho(M_m(\mathcal{C},\mathcal{C})) \geq \rho(M_2) > 1$.

First, we claim that

$$V_I \neq \emptyset \text{ for every } I \in \mathcal{C}. \tag{6.9}$$

To this aim, we note that there is a directed cycle $I \xrightarrow{\mathbf{d}} I$ in $\mathcal{C}$, where $\mathbf{d} \in \{0,1\}^r$ for some $1 \leq r \leq \#\mathcal{C}$. It follows that

$$\Phi_{sr}(\mathbf{d}^s) = \mathbf{d}^s \cdot Q^* \supseteq \mathbf{d}^s \cdot I = I,$$

which implies that $\Phi(\mathbf{d}^\omega) \supseteq I$, where $\mathbf{d}^\omega$ denotes the infinite word $\mathbf{ddd}\cdots$. The real number z with binary expansion $0.\mathbf{ddd}\cdots$ is therefore an element of V_I. This proves (6.9).

Next, we claim that

$$U_I := \{z \in [0,1] : I = \Phi(z)\} \neq \emptyset \text{ for every } I \in \mathcal{C}. \tag{6.10}$$

To prove this, we assume that $U_I = \emptyset$ for some $I \in \mathcal{C}$. Let now $\mathbf{d} \in \{0,1\}^n$ for some $n \in \mathbb{N}$ be the label of a path of length n in $\mathcal{C}$, i.e., there are K_1, $K_2 \in \mathcal{C}$ such that $K_1 \xrightarrow{\mathbf{d}} K_2$ is a path in $\mathcal{C}$. We take some $z \in V_{K_1}$ and a path $K_2 \xrightarrow{\mathbf{d}'} I$ in $\mathcal{C}$. We have

$$\Phi(\varphi_{\mathbf{d}'\mathbf{d}}(z)) = \mathbf{d}'\mathbf{d} \cdot \Phi(z) \supseteq \mathbf{d}'\mathbf{d} \cdot K_1 = I,$$

where $\varphi_{\boldsymbol{\delta}}$ is defined by $\varphi_{\delta_1\cdots\delta_n} := \varphi_{\delta_1} \circ \cdots \circ \varphi_{\delta_n}$.

Since $U_I = \emptyset$ by assumption, this implies that

$$\#\left(\mathbf{d} \cdot Q^*\right) \geq \#\left(\mathbf{d}'\mathbf{d} \cdot Q^*\right) \geq \#\Phi(\varphi_{\mathbf{d}'\mathbf{d}}(z)) > \#I = m,$$

which yields $\#\left(\mathbf{d} \cdot Q^*\right) \geq m+1$. This means that there is a path $K_3 \xrightarrow{\mathbf{d}} K_4$ of length n in $\mathcal{A}_{m+1}$. The above construction shows that there is an injective map from the set of labels of paths of length n in $\mathcal{C}$ to the set of paths of length n in $\mathcal{A}_{m+1}$. We obtain

$$e^t M_{m+1}^n e \geq \#\{\mathbf{d} \in \{0,1\}^n : \mathbf{d} \text{ is label of a path of length } n \text{ in } \mathcal{C}\} \geq \frac{1}{\#\mathcal{C}^2} e^t M_m(\mathcal{C},\mathcal{C})^n e.$$

By Lemma 15, we conclude that

$$\log_2 \rho(M_{m+1}) = \lim_{n\to\infty} \frac{\log_2(e^t M_{m+1}^n e)}{n} \geq \lim_{n\to\infty} \frac{\log_2\left(\frac{1}{\#\mathcal{C}^2} e^t M_m(\mathcal{C},\mathcal{C})^n e\right)}{n} = \log_2 \rho(M_m(\mathcal{C},\mathcal{C})) = \log_2(\rho(M_2)).$$

This yields $\rho(M_{m+1}) \geq \rho(M_2)$, a contradiction to the choice of m. Hence our claim (6.10) is proved.

We now restrict (6.8) to the component $\mathcal{C}$: There is a unique collection of nonempty compact sets $(V_I^{\mathcal{C}})_{I\in\mathcal{C}}$, such that

$$V_I^{\mathcal{C}} = \bigcup_{\substack{d\in\{0,1\}\\ J\in\mathcal{C}\\ I=d\cdot J}} \varphi_d(V_J^{\mathcal{C}}), \qquad I \in \mathcal{C}, \tag{6.11}$$

cf. for instance Edgar [3, Theorem 4.3.5]. These sets $V_I^{\mathcal{C}}$, $I \in Q_m$, are the *graph directed sets* defined by $\mathcal{C}$ and the contractions φ_d, cf. Falconer [4].

It is clear that $V_I^{\mathcal{C}} \subseteq V_I$ for each $I \in \mathcal{C}$, since the fixed point $(V_I^{\mathcal{C}})_{I\in\mathcal{C}}$ of (6.11) can be obtained by iterating the right hand side of (6.11) starting with the collection $(V_I)_{I\in\mathcal{C}}$, which yields a sequence of tuples of compact sets which is nonincreasing in each component by (6.8).

For $I \in \mathcal{C}$, we define

$$O_I := \{z \in [0,1] : d(z,V_I) < d(z,V_J) \text{ for all } J \in Q_m \cup \{\partial\} \text{ with } J \neq I\},$$

where $V_\partial := \{0,1\}$ is the boundary of the unit interval, $d(z,K) := \inf\{d(z,y) : y \in K\}$ for compact sets K and the Euclidean distance $d(z,y)$. The sets O_I, $I \in \mathcal{C}$, are open. Since $\Phi(0) = \{0\}$ and $\Phi(1) = \{2\}$ by Lemma 8, we have $U_I \subseteq O_I$, hence $O_I \neq \emptyset$ for $I \in \mathcal{C}$ by (6.10). It is easily checked that for fixed $I \in \mathcal{C}$, we have

$$\begin{gathered}\varphi_d(O_J) \subseteq O_I, \text{ for all } d \in \{0,1\}, J \in \mathcal{C} \text{ with } I = d\cdot J,\\ \text{the sets } \varphi_d(O_J) \text{ for } d \in \{0,1\}, J \in \mathcal{C} \text{ with } I = d\cdot J \text{ are disjoint.}\end{gathered} \tag{6.12}$$

This is the *open set condition*, cf. Edgar [3, Definition 6.4.7].

Therefore, by Edgar [3, Theorem 6.4.8], we have $\dim_H(V_I^{\mathcal{C}}) = s$ and $\mathcal{H}^s(V_I^{\mathcal{C}}) > 0$ for all $I \in \mathcal{C}$, where s is the unique *similarity dimension* such that $\rho(2^{-s}M_m(\mathcal{C},\mathcal{C})) = 1$. Thus we have $s = \log_2 \rho(M_m)$. Since $V_I^{\mathcal{C}} \subseteq V_I \subseteq B$ and

$$0 < \mathcal{H}^{\log_2\rho(M_m)}(V_I^{\mathcal{C}}) \leq \mathcal{H}^{\log_2\rho(M_m)}(B) \leq \mathcal{H}^{\log_2\rho(M_2)}(B) < \infty$$

by definition of m and Lemma 16, we have $\dim_H(B) = \dim_B(B) = \log_2 \rho(M_2)$.

Finally, we translate this result to the sets W_η defined in Theorem 10. From the definition of W_η we conclude that for $\eta \in \{0,1,x\}$, we have

$$\mathcal{H}^{\log_2\rho(M_2)}(\partial W_\eta) \leq \sum_{\substack{d\in\{0,1\}\\ m\in Q^*\\ \eta_0(2d+m)=\eta}} \frac{1}{\rho(M_2)} \mathcal{H}^{\log_2\rho(M_2)}(\partial A_m) < \infty. \tag{6.13}$$

We now choose $I \in \mathcal{C}$ and $\{u, v\} \subseteq I$ in such a way that $u \neq v$ and $v_2(u-v)$ is minimal, where $v_2(n)$ denotes the maximal t such that 2^t divides n. Since $\rho(M_2) > 1$ by Lemma 16, there is a transition $J = d{\cdot}I$ in $\mathcal{C}$. We set $u' = d{\cdot}u$ and $v' = d{\cdot}v$. As $\#I = \#J = m$, we have $u' \neq v'$. If $u \equiv v \pmod 4$, then $u'-v' = 1/2(u-\eta_0(2d+u))-1/2(v-\eta_0(2d+v)) = 1/2(u-v)$, so $v_2(u'-v') < v_2(u-v)$, a contradiction. If u and v are both even, we similarly get $u'-v' = 1/2(u-v)$, which is also a contradiction. Therefore, we have $\eta_0(2d+u) \neq \eta_0(2d+v)$. Let now $z \in V_I$. By definition of Φ, every neighborhood of z contains a point $z_1 \in A_u$ and a point $z_2 \in A_v$. We clearly have $\varphi_d(z_1) \in W_{\eta_0(2d+u)}$ and $\varphi_d(z_2) \in W_{\eta_0(2d+v)}$, which implies that $\varphi_d(z) \in \partial(W_0 \cup W_1 \cup W_x)$. We conclude that

$$0 < \frac{1}{\rho(M_2)}\mathcal{H}^{\log_2 \rho(M_2)}(V_I^{\mathcal{C}}) = \mathcal{H}^{\log_2 \rho(M_2)}(\varphi_d(V_I^{\mathcal{C}})) \leq \mathcal{H}^{\log_2 \rho(M_2)}(\partial(W_0 \cup W_1 \cup W_x)). \tag{6.14}$$

This concludes the proof of the theorem for $|x| \geq 5$.

For $x = -1$, we note that $\rho(M_2) = 1$, which implies that $\dim_H(\partial(W_0 \cup W_1 \cup W_x)) = 0$ by Lemma 16. It can easily be checked that in this case, we have

$$\begin{aligned} W_0 &= [0, 1/6) \cup (2/6, 4/6) \cup (5/6, 1], \\ W_1 &= (1/6, 2/6), \\ W_{-1} &= (4/6, 5/6), \end{aligned}$$

and therefore $\mathcal{H}^0(\partial(W_0 \cup W_1 \cup W_x)) = 4$.

The remaining dimensions in Table 2 have been computed using Mathematica®. □

7. Geometric Approach for Calculating the Frequency of Digits

We give a geometric approach (going back to Delange [2]) using the results in the preceeding sections to compute the summatory function of the frequency of digits. In contrast to the results in Section 4, we are now able to compute the summatory function up to some integer N instead of considering the full block length $\{0, \ldots, 2^L - 1\}$ as in Section 4. However, we will need the results of that section to compute the Lebesgue measures of the sets W_d, $d \in \{0, 1, x\}$.

For $d \in \{1, x\}$ and positive $N \in \mathbb{Z}$, let

$$H_d(N) := \sum_{n=0}^{N-1} f_d(n) = \sum_{n=0}^{N-1} \sum_{k \geq 0} [\eta_k(n) = d],$$

where $f_d(n)$ has already been defined in (4.1). Since $\eta_k(n) = 0$ for $k > \lfloor \log_2(n) \rfloor + \#Q - 2$ by Theorem 7, Theorem 10 implies that

$$H_d(N) = \sum_{n=0}^{N-1} \sum_{k=0}^{K+1} \left[\left\{ \frac{n}{2^{k+2}} \right\} \in W_d \right],$$

where $K = \lfloor \log_2 N \rfloor + \#Q - 2$. We proceed as in Section 4 of [5]: We replace W_d by an appropriate approximation $W_{d,k}$, replace the sum by an integral and pull out the main

term given by the Lebesgue measure of W_d. Since the technical modifications are straight forward, we skip the details. We get

$$H_d(N) = \lambda(W_d) N \log_2 N + N\psi_d(\{\log_2 N\}) + O(N^{\log_2 \rho(M_2)}), \tag{7.1}$$

where $\rho(M_2)$ is the dominant eigenvalue of the auxiliary automaton as in Theorem 11,

$$\begin{aligned}
\psi_d(z) &:= \lambda(W_d)(\#Q - \{z\}) + 2^{\{z\}+\#Q-2}\Psi_d(2^{\{z\}+2-\#Q}) + \sum_{k=-1}^{\infty} \beta_k,\\
\Psi_d(t) &:= \sum_{k=0}^{\infty} \int_0^t \left(\left[\left\{x2^{k-2}\right\} \in W_d\right] - \lambda(W_d)\right)\, dx,\\
\beta_k &:= \int_0^1 \Big([\{x\} \in W_{d,k}] - [\{x\} \in W_d] \Big)\, dx,\\
W_{d,k} &:= \bigcup_{m \in 2^{k+2} W_d \cap \mathbb{Z}} \left[\frac{m}{2^{k+2}}, \frac{m+1}{2^{k+2}} \right).
\end{aligned}$$

As in [5] we see that $\psi_d(z)$ is continous in $[0,1)$ and that $\lim_{z\to 1^-} \psi_d(z)$ exists. Of course, ψ_d is 1-periodic. We cannot conclude that ψ_d is continuous at $z=1$ by direct computation as in [5]. Instead, we get continuity by considering

$$O(L) = H_d(2^L) - H_d(2^L - 1) = 2^L\Big(\psi_d(0) - \psi_d\Big(\Big\{ \log_2\Big(1 - \frac{1}{2^L}\Big)\Big\}\Big)\Big) + O(L).$$

Comparing with Theorem 9, we see that (7.1) implies that $\lambda(W_d) = 1/6$ for $d \in \{1, x\}$. Since $\lambda(W_0 \cup W_1 \cup W_x) = 1$ we must have $\lambda(W_0) = 2/3$.

We summarize this result in the following theorem.

Theorem 17. *Let $D = \{0, 1, x\}$ be a NADS, $d \in \{1, x\}$ and N be a positive integer. Then the number of occurrences of the digit d in the D-NAFS of the integers $0, \dots, N-1$ equals*

$$H_d(N) = \frac{1}{6} N \log_2 N + N\psi_d(\{\log_2 N\}) + O(N^{\log_2 \rho(M_2)}),$$

where ψ_d is a 1-periodic continuous function and $\rho(M_2) < 2$ is the spectral radius of the adjacency matrix of the auxiliary automaton described in Section 6.

Moreover, the Lebesgue measures of the sets W_d described in Theorem 10 equal

$$\lambda(W_1) = \lambda(W_x) = \frac{1}{6}, \qquad \lambda(W_0) = \frac{2}{3}.$$

8. Non-Optimality

For $D = \{0, 1, -1\}$, the D-NAF of n has minimal Hamming weight amongst all $\{0, 1, -1\}$-expansions of n, cf. Reitwiesner [15], where the Hamming weight $c(\boldsymbol{\eta})$ of an expansion $\boldsymbol{\eta} \in D^{\mathbb{N}_0}$ equals the number of nonzero digits $\#\{j : \eta_j \neq 0\}$.

For $x \leq -5$, this is no longer the case:

Theorem 18. *Let $D = \{0, 1, x\}$ be a NADS. Then the following two conditions are equivalent:*

- *For every positive integer n, the Hamming weight of the D-NAF of n is minimum amongst all D-expansions of n.*
- $x \in \{-1, 3\}$.

We first consider the case $x = 3$ and give an algorithmic proof, which can also be used in more general situations. For instance, the case $x = -1$ follows along the same lines.

Lemma 19. *The $\{0, 1, 3\}$-NAF is the $\{0, 1, 3\}$-expansion of minimal Hamming weight.*

Proof of the lemma. We consider the transducer $\widetilde{\mathcal{T}}$ (cf. Figure 13). We introduce weights

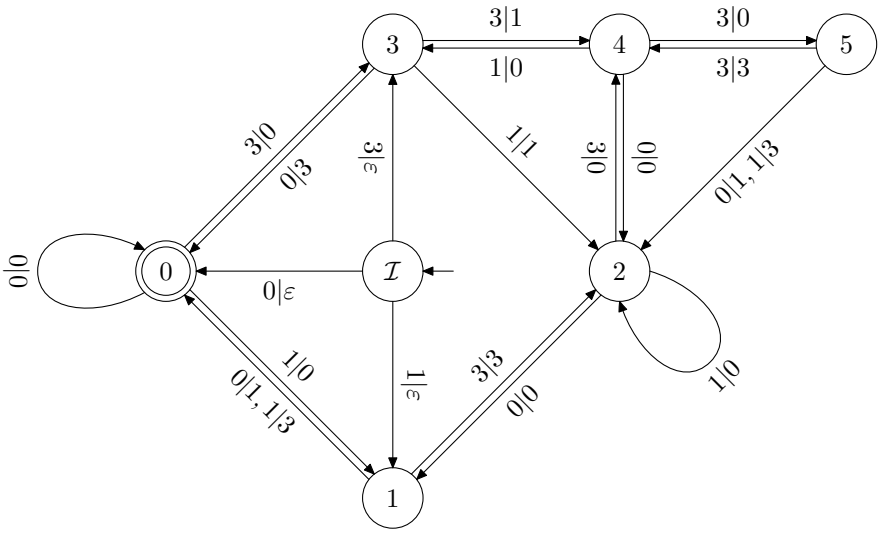

FIGURE 13. Transducer $\widetilde{\mathcal{T}}$ calculating a $\{0, 1, 3\}$-NAF of n from any $\{0, 1, 3\}$-expansion of n from right to left.

for the transitions:

$$w(i \xrightarrow{d|o} j) := c(d) - c(o),$$

where c denotes the Hamming weight and $c(\varepsilon) = 0$. If we have a successful path $\mathcal{I} \xrightarrow{\mathbf{d}|\boldsymbol{\eta}} 0$, the weight of this path equals $c(\mathbf{d}) - c(\boldsymbol{\eta})$. We calculate the shortest path from $\mathcal{I}$ to 0 in this weighted digraph using the Ford-Bellman algorithm (cf. Cook et al. [1]). It turns out that the shortest path has weight 0, i.e. $c(\mathbf{d}) - c(\boldsymbol{\eta}) \geq 0$, as requested. □

Remark 20. We note that the above proof of Lemma 19 also yields a complete description of all $\{0, 1, 3\}$-expansions of minimal weight: The Ford-Bellman algorithm calculates a feasible potential $\pi : V \to \mathbb{Z}$, where $\pi(i)$ is the weight of a shortest path from $\mathcal{I}$ to i. In our case, we have

i	$\mathcal{I}$	0	1	2	3	4	5,
$\pi(i)$	0	0	1	1	1	1	2.

Thus if we have $\pi(i) + w(i \xrightarrow{d|o} j) > \pi(j)$ for some transition, the transition corresponds to an actual gain when modifying the given representation $\mathbf{d}$ to the D-NAF. Therefore,

we remove all those transitions and all output labels and get an automaton which accepts minimal $\{0,1,3\}$-expansions, see Figure 14. In this figure, we also identified states 0 and $\mathcal{I}$, since they only differed in the output labels of the transitions leaving them.

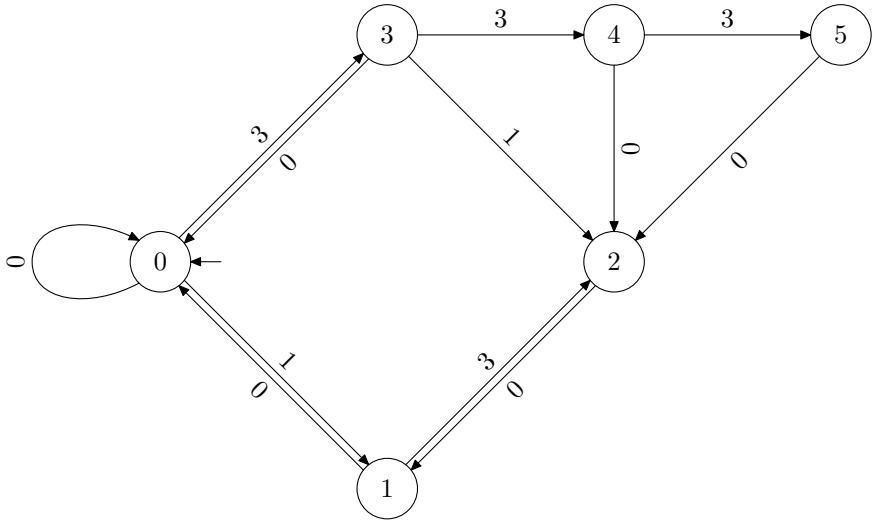

FIGURE 14. Automaton accepting all binary $\{0,1,3\}$-expansions of minimal Hamming weight.

If we use the same approach for $x=-1$, we get the transducer in Figure 15 and the automaton in Figure 16, respectively. The transducer corresponds to the algorithm due to Jedwab and Mitchell [9], the automaton to the syntactical rules described in Heuberger [**?**].

Proof of Theorem 18. By Lemma 19 and Reitwiesner [15], we only have to consider the case $x \le -5$. If $|x|+3 = 2^g$ for some $g \ge 4$, we consider $n = 2^{g+1}+7$. We have

$$n = 1\cdot 2^{2g} + x\cdot 2^g + 1\cdot 2^2 + x\cdot 2^0, \tag{8.1}$$

$$n = 1\cdot 2^{g+2} + x\cdot 2 + 1. \tag{8.2}$$

Both are $\{0,1,x\}$-expansions, (8.1) is the D-NAF with Hamming weight 4, whereas (8.2) is an expansion of Hamming weight 3.

If $x=-5$, we have

$$23 = \mathsf{value}(100\bar{5}0\bar{5}0\bar{5}), \tag{8.3}$$

$$23 = \mathsf{value}(\quad 1000\bar{5}1). \tag{8.4}$$

The first expansion (8.3) is the D-NAF, again with Hamming weight 4, the second expansion has Hamming weight 3.

Next, if $|x|+3$ is not a power of 2, we consider $n=3$. We have $\eta_0(3)=x$, $\eta_1(3)=0$ and $r^2(3) = (3+|x|)/4$. By assumption, $(3+|x|)/4$ is not a power of 2. Since $(3+|x|)/4 < |x|$, we conclude that any D-expansion of $(3+|x|)/4$ has Hamming weight at least 2, therefore,

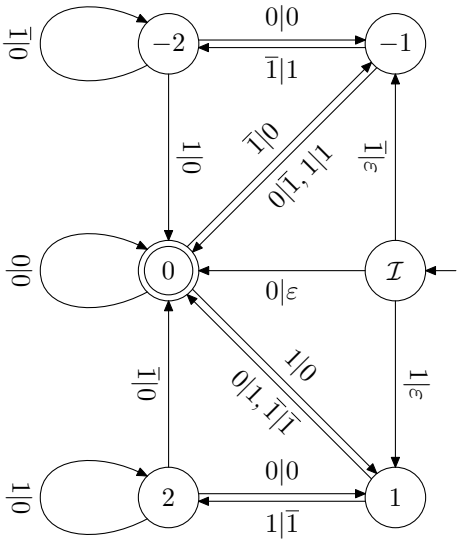

FIGURE 15. Transducer $\widetilde{\mathcal{T}}$ calculating a $\{0, 1, -1\}$-NAF of n from any $\{0, 1, -1\}$-expansion of n from right to left.

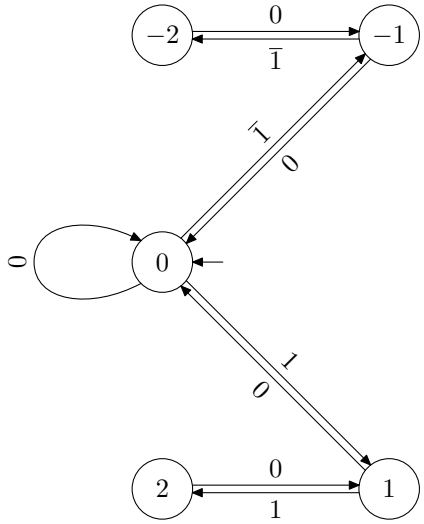

FIGURE 16. Automaton accepting all binary $\{0, 1, -1\}$-expansions of minimal Hamming weight.

the D-NAF of 3 has Hamming weight at least 3. However, the standard binary expansion $3 = \mathsf{value}(11)$ has lower Hamming weight. □

9. Addition of 1

The transducer for calculating the addition of 1 for $x = -5$ is shown in Figure 17.

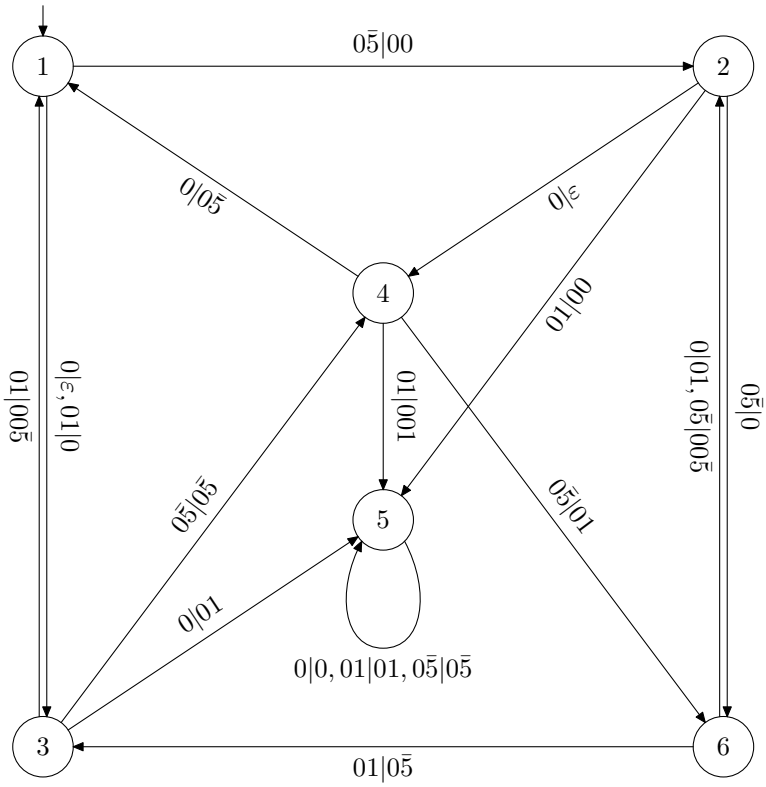

FIGURE 17. Add $+1$ or -1 in $\{0, 1, -5\}$-NAF from right to left. Start node is 1 for addition of $+1$ and 2 for addition of -1

10. Digit formulæ

It was proved in [14, 7] that for the digit set $\{0, 1, -1\}$, the formula

$$n = \sum_{k\geq 0} \left(\left\lfloor \frac{n}{2^{k+2}} + \frac{5}{6} \right\rfloor - \left\lfloor \frac{n}{2^{k+2}} + \frac{4}{6} \right\rfloor - \left\lfloor \frac{n}{2^{k+2}} + \frac{2}{6} \right\rfloor + \left\lfloor \frac{n}{2^{k+2}} + \frac{1}{6} \right\rfloor \right) 2^k$$

produces the NAF, i. e., that if we write the NAF of n as $\dots a_2(n)a_1(n)a_0(n)$, the digits are given by

$$a_k(n) = \left\lfloor \frac{n}{2^{k+2}} + \frac{5}{6} \right\rfloor - \left\lfloor \frac{n}{2^{k+2}} + \frac{4}{6} \right\rfloor - \left\lfloor \frac{n}{2^{k+2}} + \frac{2}{6} \right\rfloor + \left\lfloor \frac{n}{2^{k+2}} + \frac{1}{6} \right\rfloor.$$

In the instance $\{0, 1, -5\}$, there is a similar formula, viz.

$$a_k(n) = \left\lfloor \frac{n}{2^{k+2}} + \frac{23}{24} \right\rfloor - \left\lfloor \frac{n}{2^{k+2}} + \frac{1}{24} \right\rfloor - 6 \left\lfloor \frac{n}{2^{k+2}} + \frac{16}{24} \right\rfloor + 6 \left\lfloor \frac{n}{2^{k+2}} + \frac{4}{24} \right\rfloor.$$

If $a_k(n) \in \{1, -5\}$, this formula produces the right result, but it sometimes happens that if $a_k(n) = 0$, the formula gives either 1 or -5.

For other systems like $\{1, 0, -13\}$ etc. not even such a formula exists.

References

1. W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, and A. Schrijver, *Combinatorial optimization*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons Inc., New York, 1998. MR **99b:**90098
2. H. Delange, *Sur la fonction sommatoire de la fonction "somme des chiffres"*, Enseignement Math. (2) **21** (1975), no. 1, 31–47. MR 52 #319
3. G. A. Edgar, *Measure, topology, and fractal geometry*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1990. MR **92a:**54001
4. K. Falconer, *Techniques in fractal geometry*, John Wiley & Sons Ltd., Chichester, 1997. MR **99f:**28013
5. P. J. Grabner, C. Heuberger, and H. Prodinger, *Distribution results for low-weight binary representations for pairs of integers*, Theoret. Comput. Sci. **319** (2004), 307–331.
6. R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete mathematics. A foundation for computer science*, second ed., Addison-Wesley, 1994. MR **97d:**68003
7. C. Heuberger and H. Prodinger, *On minimal expansions in redundant number systems: Algorithms and quantitative analysis*, Computing **66** (2001), 377–393.
8. H.-K. Hwang, *On convergence rates in the central limit theorems for combinatorial structures*, European J. Combin. **19** (1998), 329–343. MR **99c:**60014
9. J. Jedwab and C. J. Mitchell, *Minimum weight modified signed-digit representations and fast exponentiation*, Electron. Lett. **25** (1989), 1171–1172.
10. D. E. Knuth, *Seminumerical algorithms*, third ed., The Art of Computer Programming, vol. 2, Addison-Wesley, 1998.
11. P. Lancaster and M. Tismenetsky, *The theory of matrices*, second ed., Computer Science and Applied Mathematics, Academic Press Inc., Orlando, FL, 1985. MR **87a:**15001
12. J. A. Muir and D. R. Stinson, *Alternative digit sets for nonadjacent representations*, to appear in Lecture Notes in Computer Science (Selected Areas in Cryptography, 2003), Preprint available at `http://www.cacr.math.uwaterloo.ca/~dstinson/papers/na_digitsets-pv.ps`.
13. ———, *Alternative digit sets for nonadjacent representations*, Journal Version of [12]; Preprint available at `http://www.cacr.math.uwaterloo.ca/~dstinson/papers/na_digitsets-j.ps`.
14. H. Prodinger, *On binary representations of integers with digits* $-1, 0, 1$, Integers **0** (2000), A08, available at `http://www.integers-ejcnt.org/vol0.html`.
15. G. W. Reitwiesner, *Binary arithmetic*, Advances in computers, vol. 1, Academic Press, New York, 1960, pp. 231–308.

(C. Heuberger) Institut für Mathematik B, Technische Universität Graz, Steyrergasse 30, 8010 Graz, Austria

E-mail address: `clemens.heuberger@tugraz.at`

(H. 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`