Scheicher, K. and Thuswaldner, J.M.
Osaka J. Math.
41 (2004), 327–351

ON THE CHARACTERIZATION OF CANONICAL NUMBER SYSTEMS

KLAUS SCHEICHER and JÖRG M. THUSWALDNER

(Received September 24, 2002)

1. Introduction

It is well known that each positive integer n can be expressed uniquely as a sum $n = d_0 + d_1 b + \cdots + d_h b^h$ with an integral base number $b \geq 2$, $d_h \neq 0$ and $d_i \in \{0, \ldots, b-1\}$. This concept can be generalized in several directions.

On the one hand the base sequence 1, b, b^2, ... can be replaced by a sequence $1 = u_0 < u_1 < u_2 < \cdots$ to obtain representations of positive integers. Of special interest is the case where the sequence $\{u_i\}_{i=0}^{\infty}$ is defined by a linear recurrence. A famous example belonging to this class is the so-called Zeckendorf representation.

On the other hand, one can generalize the set of numbers which can be represented. We mention two kinds of number systems belonging to this class:

The so called β-expansions introduced by Rényi [27] which are representations of real numbers in the unit interval as sums of powers of a real base number β. These digit representations of real numbers are strongly related to digit representations of positive integers if β is a zero of the characteristic polynomial of a linear recurring base sequence $\{u_i\}_{i=0}^{\infty}$. Of special interest is the case where β is a Pisot number. These expansions have been extensively studied. We mention here the papers Berend-Frougny [6] , Frougny [12, 13], Frougny-Solomyak [14, 15] and Loraud [25] and refer to the references given there.

Another kind of number systems which admit the representation of a set which is different from $\mathbb{N}$ are the so-called *canonical number systems* (for short CNS). Since CNS form the main object studied in the present paper we recall their definition (cf. Akiyama-Pethő [2]).

DEFINITION 1.1. Let

$$P(x) := b_n x^n + b_{n-1} x^{n-1} + \cdots + b_0 \in \mathbb{Z}[x]$$

be such that $n \geq 1$ and $b_n = 1$ (set $b_j = 0$ for $j > n$). Let $\mathcal{N} = \{0, 1, \ldots, |b_0| - 1\}$ and

The first author was supported by the Austrian Science Foundation Project S8305.
The second author was supported by the Austrian Science Foundation Project P-14200-MAT.

$\mathcal{R}$ be the quotient ring

$$\mathcal{R} = \mathbb{Z}[x]/P(x)\mathbb{Z}[x].$$

• We say that $\gamma \in \mathcal{R}$ has a finite representation if it admits a representation of the shape

$$\gamma = d_0 + d_1 x + \cdots + d_h x^h$$

with $d_j \in \mathcal{N}$ for $0 \leq j \leq h$ and $d_h \neq 0$ for $h \neq 0$ (set $d_j = 0$ for $j > h$).

• The numbers $d_j = d_j(\gamma)$, $j \geq 0$, are called the digits of γ with respect to $(P(x), \mathcal{N})$.

• The pair $(P(x), \mathcal{N})$ is called canonical number system or CNS in $\mathcal{R}$, if each $\gamma \in \mathcal{R}$ has a finite representation. $\mathcal{N}$ is called digit set of this CNS.

• If $P(x)$ is irreducible, then let α be one of its zeros. In this case $\mathcal{R}$ is isomorphic to $\mathbb{Z}[\alpha]$, the ring generated by $\mathbb{Z}$ and α. Therefore we may replace x by α in the above expansions. In this case, we simplify the notation $(P(x), \mathcal{N})$ to $(\alpha, \mathcal{N})$ and α is called base of this CNS.

Unlike for "ordinary" number systems, where it is clear that each integer $b \geq 2$ can serve as base, it is a difficult problem to determine which polynomials provide CNS. Despite there are many papers dealing with the characterization of possible polynomials — we will give a detailed overview in the next section — there does not exist a complete characterization up to now. Some of the known results provide the characterization of the polynomials for some classes of CNS, others give algorithms that allow to decide whether a given polynomial $P(x)$ forms a CNS or not. In the present paper we have two aims. First we want to present a fast algorithm for deciding whether a given $P(x)$ is a CNS polynomial or not, in a second step we use this new algorithm to characterize a large class of CNS polynomials. Our results prove Conjecture 1 of Akiyama-Pethő [2] and provide a conditional proof of Conjecture 2 of the same paper.

The paper is organized as follows. In the next section we discuss earlier results on the characterization of CNS and state some easy facts about them. Section 3 is devoted to the definition of certain graphs and automata which reflect many important properties of CNS and thus are very important for the proofs of our results. In Section 4 we establish a fast algorithm which allows to decide whether a given $P(x)$ is a CNS polynomial or not (Theorem 4.4). In Section 5 we use our algorithm to characterize a large class of CNS polynomials (Theorem 5.8). Section 6 contains characterizations of cubic (Theorem 6.1) and quartic (Theorem 6.2) CNS under a certain condition. In Section 7 some numerically found examples are presented dealing with conjectures on CNS polynomials. We finish the paper with some remarks in Section 8.

2. Some facts about canonical number systems

In this section we want to review some earlier results on CNS. Some of them will be used in the proofs of our results.

As mentioned in the introduction it is an open problem to give a complete characterization of all polynomials that provide a CNS. Nevertheless, there are many partial results. Knuth [21] considered certain examples of bases, one of them the Gaussian integer $-1+i$ which is intimately related to the famous "twin dragon" fractal (as for the connection between fractals and CNS cf. also Akiyama-Thuswaldner [4, 5], Kátai [17] and Scheicher-Thuswaldner [29]). The first systematic treatment was given in Kátai-Szabó [20], where all Gaussian integers which are CNS bases are characterized. This result was generalized to quadratic integers in Kátai-Kovács [18, 19] and independently in Gilbert [16]. Kőrnyei [22] dealt with a special class of cubic integers and very recently Brunotte [9, 10] characterized all CNS whose bases are roots of trinomials. As for the general case Kovács [23] proved that an algebraic integer α gives rise to a CNS if its minimal polynomial $P(x) = x^n + b_{n-1}x^{n-1} + \cdots + b_1x + b_0$ satisfies

$$2 \leq b_0 \geq b_1 \geq \cdots \geq b_{n-1} \geq 1.$$

However, this characterization is not complete. Examples not contained in this class are provided in Kovács-Pethő [24], where also an algorithm is established that decides whether a given α is a CNS base or not. Recently, Akiyama-Pethő [2] developed a much faster algorithm than the one in [24]. In particular, they proved the following result. Let M be a positive integer and α be an algebraic integer of degree n with minimal polynomial $P(x)$ as in Definition 1.1. If $b_0 \geq (1 + M^{-1}) \sum_{j=1}^{n} |b_j|$ then it suffices to check for $(2M)^n$ elements of $\mathcal{R}$ whether they have finite representation in order to decide if $P(x)$ is a CNS polynomial or not. These $(2M)^n$ numbers are given explicitly. Since this criterion works for all polynomials satisfying

$$\sum_{j=1}^{n} |b_j| < b_0 \tag{2.1}$$

we will call (2.1) the *Akiyama-Pethő-condition* or the AP condition, for short. This condition will also play an important role in our paper.

It is the aim of the present paper to give further improvements of the Akiyama-Pethő-Algorithm. These improvements allow us to characterize a large class of CNS polynomials. In principle we even get a complete characterization of all CNS under the AP condition. But unfortunately this characterization contains many awkward sets of inequalities which surely can be simplified considerably. For the case of cubic and quartic CNS we achieve such simplifications and are thus able to give easy criteria in order to decide whether $P(x)$ is a CNS polynomial or not.

In our proofs we use finite automata and graphs. These objects turned out to be useful for the treatment of characterization problems (cf. for instance Scheicher [28]

and Thuswaldner [31, 32]).

It is easy to see that a necessary condition for $(P(x), \mathcal{N})$ to be a CNS is that $\mathcal{N}$ forms a complete residue system $\mathbb{Z}[x]$ modulo $(x, P(x))$, the ideal generated by x and $P(x)$. Indeed, if the residue class $r \bmod (x, P(x))$ does not occur in $\mathcal{N}$ then no $\gamma \in \mathcal{R}$ with $\gamma \equiv r \bmod (x, P(x))$ has a representation. Due to this fact we may define the mapping $\Phi(z)$ from $\mathcal{R}$ to itself by

$$\Phi(z) = x^{-1}(z - r),$$

where r is the unique element of $\mathcal{N}$ satisfying $z \equiv r \bmod (x, P(x))$. Kovács-Pethő [24] remarked that $P(x)$ can serve as a polynomial for a CNS only if all its zeros are greater than 1 in modulus. This ensures that the iterates z, $\Phi(z)$, $\Phi^2(z)$, ... end up in the finite set

$$S := \{z = z_0 + z_1 x + \cdots + z_{n-1}x^{n-1} \in \mathcal{R} \colon |z_i| \le C\},$$

where $C > 0$ is a certain computable constant. The finiteness of S implies that the sequence $\{\Phi^j(z)\}_{j\ge 0}$ becomes ultimately periodic for each $z \in \mathcal{R}$. Let $\mathcal{P}$ be the set of all points which can occur in such a period, i.e.

$$\mathcal{P} := \{p \in S \mid \exists \omega \in \mathbb{N} : p = \Phi^{\omega}(p)\}.$$

It is clear from the definition of Φ and $\mathcal{P}$ that each $z \in \mathcal{R}$ admits a unique representation of the shape

$$z = \sum_{l=0}^{L} d_l x^l + x^{L+1} p \tag{2.2}$$

with $p \in \mathcal{P}$, $d_l \in \mathcal{N}$ and $L \in \mathbb{N}$ as small as possible. Since p is a periodic point, it has a representation of the shape

$$p = \sum_{j=0}^{N\omega-1} c_{j \bmod \omega} x^j + x^{N\omega} p \tag{2.3}$$

for each $N \in \mathbb{N}$. For this reason we will use the abbreviation

$$z = ([c_{\omega-1} \cdots c_0]^{\infty} d_L \cdots d_0)$$

for the representation (2.2). Here $[c_{\omega-1} \cdots c_0]^{\infty}$ is the infinite repetition of the string $c_{\omega-1} \cdots c_0$. Instead of $([0]^{\infty} d_L \cdots d_0)$ we will simply write $(d_L \cdots d_0)$. As mentioned in Definition 1.1 in this case the representation is called finite. Generalizing Definition 1.1 we will use the notation $d_j(z)$ for the j-th digit in the representation of z, i.e. the coefficient of x^j in this representation.

The following lemma contains an easy criterion for $P(x)$ to be the polynomial of a CNS.

Lemma 2.1. *Suppose that each zero of $P(x)$ has modulus greater than* 1 *and $\mathcal{N}$ is a complete residue system* mod$(x, P(x))$. *Then $(P(x), \mathcal{N})$ is a CNS if and only if $\mathcal{P} = \{0\}$.*

Proof. If $\mathcal{P} = \{0\}$ then (2.2) ensures that $(P(x), \mathcal{N})$ is a CNS. If $0 \neq p \in \mathcal{P}$ and $\Phi^{\omega}(p) = p$ then p has a representation of the shape (2.3). If all c_k $(0 \leq k \leq \omega - 1)$ were equal to zero this would imply $p = 0$, a contradiction. Since the representation of p is $([c_{\omega-1} \cdots c_0]^{\infty})$ this representation can not be finite. □

We want to mention here that $\gamma \in \mathcal{R}$ has finite representation if and only if the sequence $\{\Phi^j(\gamma)\}_{j \geq 0}$ is ultimately zero. This is an easy consequence of the definition of Φ.

In the remaining part of the paper we will always assume that $\mathcal{N} = \{0, 1, \ldots, |b_0| - 1\}$ is a complete residue system mod$(x, P(x))$ and that all zeros of $P(x)$ are greater than 1 in modulus. By the above considerations this ensures that each $\gamma \in \mathcal{R}$ admits a representation of the shape (2.2).

3. Definition of graphs and automata

In this section we want to define certain classes of directed graphs. These graphs will be used to perform the addition of fixed numbers on the space of representations. Furthermore, we will state some properties of these graphs and discuss their relation to so-called transducer automata. First of all we want to give a definition of this kind of automata.

DEFINITION 3.1 (cf. Berstel [8] or Eilenberg [11]). The 6-tuple $\mathcal{A} = (Q, \Sigma, \Delta, T, v_0, R)$ is called a finite state transducer automaton if

- Q, Σ and Δ are nonempty, finite sets, and
- $T\colon Q \times \Sigma \to Q$ and $R\colon Q \times \Sigma \to \Delta$ are unique mappings.

The sets Σ and Δ are called input and output alphabet, respectively. Q is called the set of states and v_0 is the starting state. The mappings T and R are called transition and result function, respectively.

A finite automaton works as follows. The automaton starts at time 0 at the state v_0. At each discrete time t, the automaton reads an input digit l_t and determines the corresponding output digit $l'_t = R(v_t, l_t)$ as well as the next state $v_{t+1} = T(v_t, l_t)$.

A transducer automaton can be interpreted as a labeled directed graph G in the following way. The vertices of G are the elements of the set of states Q. Furthermore, there exists an edge from a vertex v_1 to a vertex v_2 labeled by $l_1 \mid l'_1$, $v_1 \xrightarrow{l_1|l'_1} v_2$ for

short, if $T(v_1, l_1) = v_2$ and $R(v_1, l_1) = l_1'$. We will need this interpretation frequently in the present paper.

DEFINITION 3.2. Let $P(x)$ be as in Definition 1.1. We say that the addition of the number $z \in \mathcal{R}$ is computable by the finite state transducer $A(z)$, if

- for any $\gamma \in \mathcal{R}$ the transducer $A(z)$ is able to read the digits $d_j(\gamma)$ of the representation of γ as input string and returns the digits $d_j(\gamma + z)$ of the representation of $\gamma + z$ as output string.
- a finite representation of γ results in a finite representation of $\gamma + z$.

The transducer $A(1)$ is called the counting automaton or adding machine of $(P(x), \mathcal{N})$.

We will need the transducers $A(z)$ for certain numbers $z \in \mathcal{R}$ in order to derive our characterization results. These transducers will emerge from the following infinite labeled directed graph. Let $\mathcal{A}(\mathcal{R})$ be the labeled directed graph with set of vertices $\mathcal{R}$. The edges connecting two vertices are defined as follows. Let v_0, v_1 be two vertices of $\mathcal{R}$. Then there exists an edge from v_0 to v_1 labeled by $l_0|l_0'$ with l_0, $l_0' \in \mathcal{N}$ if and only if

$$v_0 + l_0 = l_0' + x v_1. \tag{3.1}$$

This edge will be denoted by

$$v_0 \xrightarrow{l_0|l_0'} v_1.$$

Of course, relation (3.1) can be iterated. If one starts at the vertex v_0 and uses as input the digits of the representation of a number $l_0 + \cdots + l_k x^k$, one ends at the state v_{k+1} where

$$v_0 + l_0 + \cdots + l_k x^k = l_0' + \cdots + l_k' x^k + x^{k+1} v_{k+1}. \tag{3.2}$$

The l_i and l_i' can be interpreted as input and output digits, respectively. Suppose that we use the digits of the finite representation

$$y = (l_L \cdots l_0)$$

as input digits (from right to left) starting at a vertex $z \in \mathcal{R}$. Then we can see from (3.2) that the sequence l_0', l_1', ... of output digits is the representation

$$y + z = \left([l_{L'+\omega}' \cdots l_{L+1}']^{\infty} l_{L'}' \cdots l_0'\right).$$

The representation formed from the output string is obviously always finite, i.e. $l_{L'+1}' = 0$ and $\omega = 1$, if $(P(x), \mathcal{N})$ is a CNS.

The above procedure becomes more clear if we set up a new type of representation for the elements of $\mathcal{R}$. This representation has been found independently in Brunotte [9, 10].

Lemma 3.3. *Let $P(x)$ be as in* Definition 1.1. *Each $q \in \mathcal{R}$ has a unique representation*

$$q = \sum_{j=0}^{n-1} x^j q_j \tag{3.3}$$

with

$$q_j = \sum_{i=1}^{n} \varepsilon_i b_{i+j}, \quad (\varepsilon_i \in \mathbb{Z}, \ \ j = 0, \ldots, n-1). \tag{3.4}$$

For such sums we will use the notation

$$q = (\varepsilon_1, \ldots, \varepsilon_n)_\varepsilon. \tag{3.5}$$

This representation will be called the ε-representation of q. The change $(q_0, \ldots, q_{n-1}) \to (\varepsilon_1, \ldots, \varepsilon_n)_\varepsilon$ corresponds to a linear base transformation of the lattice $\mathcal{R}$.

Proof. The equations (3.4) provide a linear system of n equations

$$\begin{bmatrix} b_n & 0 & \cdots & 0 \\ b_{n-1} & b_n & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ b_1 & b_2 & \cdots & b_n \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \vdots \\ \varepsilon_n \end{bmatrix} = \begin{bmatrix} q_{n-1} \\ \vdots \\ \vdots \\ q_0 \end{bmatrix}.$$

This system has a unique solution such that $\varepsilon_i \in \mathbb{Z}$ for all i, since $b_n = 1$ and b_i, $q_i \in \mathbb{Z}$. □

Let $q = (\varepsilon_1, \ldots, \varepsilon_n)_\varepsilon$. We will now examine how this ε-representation changes if we move along the edge $q \xrightarrow{l|l'} q'$ in $\mathcal{A}(\mathcal{R})$, i.e. we will determine the ε-representation of q' from the ε-representation of q. By the above lemma we have

$$q = \sum_{j=0}^{n-1} x^j \sum_{i=1}^{n} \varepsilon_i b_{i+j}. \tag{3.6}$$

By the definition of the edges of $\mathcal{A}(\mathcal{R})$ there exists a unique $k \in \mathbb{Z}$ such that

$$\varepsilon_1 b_1 + \cdots + \varepsilon_n b_n + l = k b_0 + l'. \tag{3.7}$$

Since

$$b_n x^n + b_{n-1}x^{n-1} + \cdots + b_0 = 0$$

we can subtract k times this minimal polynomial from (3.6) to obtain

$$q = \sum_{j=0}^{n} x^j \left(-kb_j + \sum_{i=1}^{n} \varepsilon_i b_{i+j} \right).$$

Since $q + l = xq' + l'$ this implies that

$$q' = \sum_{j=0}^{n-1} x^j \left(-kb_{j+1} + \sum_{i=1}^{n} \varepsilon_i b_{i+j+1} \right) = (-k, \varepsilon_1, \ldots, \varepsilon_{n-1})_\varepsilon.$$

Thus the ε-representation of q' emerges from the ε-representation of q by canceling ε_n, shifting $\varepsilon_1, \ldots, \varepsilon_{n-1}$ to the right and inserting $-k$, which is defined according to (3.7), as the first element.

DEFINITION 3.4. An edge of the shape $(\varepsilon_1, \ldots, \varepsilon_n)_\varepsilon \to (t, \varepsilon_1, \ldots, \varepsilon_{n-1})_\varepsilon$ in $\mathcal{A}(\mathcal{R})$ is called an *edge of type t*. If we emphasize on the type of an edge we will use the notation

$$(\varepsilon_1, \ldots, \varepsilon_n)_\varepsilon \xrightarrow{\text{type } t} (t, \varepsilon_1, \ldots, \varepsilon_{n-1})_\varepsilon.$$

We will be interested in subgraphs of $\mathcal{A}(\mathcal{R})$ which are closed in a certain sense. To this matter we need the following definition.

DEFINITION 3.5. A number v_{k+1} is called reachable from v_0 if there exist $l_0, \ldots, l_k \in \mathcal{N}$ such that (3.2) holds. We will denote this by $v_0 \rightsquigarrow v_{k+1}$. The series of states

$$v_0 \xrightarrow{l_0|l_0'} v_1 \xrightarrow{l_1|l_1'} \cdots \xrightarrow{l_k|l_k'} v_{k+1}$$

is called the path connecting v_0 with v_{k+1}. Let $M \subset \mathcal{R}$. Then $\overline{M}$ denotes the set

$$\overline{M} = \{q' \mid q \rightsquigarrow q',\ q \in M\}.$$

In the following lemma we show the existence of a rather "small" set E, which is closed in the sense of the above definition.

Lemma 3.6. *Let $P(x)$ be as in* Definition 1.1 *and suppose that $P(x)$ fulfills the AP condition* (2.1). *Then $E = \{q \mid q = (\varepsilon_1, \ldots, \varepsilon_n)_\varepsilon,\ \varepsilon_i \in \{-1, 0, 1\}\}$ satisfies $\overline{E} = E$. Thus $\overline{E}$ has 3^n elements.*

Proof. We have to show that for each $q \in E$ the existence of an edge $q \xrightarrow{l|l'} q'$ in $\mathcal{A}(\mathcal{R})$ implies $q' \in E$. Suppose that $q = (\varepsilon_1, \ldots, \varepsilon_n)_\varepsilon$ with $\varepsilon_j \in \{-1, 0, 1\}$. From the AP condition it follows that

$$\left| \sum_{i=1}^{n} \varepsilon_i b_i \right| \leq \sum_{i=1}^{n} |b_i| < b_0,$$

and thus

$$-b_0 < \varepsilon_1 b_1 + \cdots + \varepsilon_n b_n < b_0.$$

Since $l \in \mathcal{N}$, we obtain

$$-b_0 < \varepsilon_1 b_1 + \cdots + \varepsilon_n b_n + l < 2b_0 - 1.$$

Hence there exists a $k \in \{-1, 0, 1\}$ such that

$$\varepsilon_1 b_1 + \cdots + \varepsilon_n b_n + l = kb_0 + l'. \tag{3.8}$$

Thus the edge $q \xrightarrow{l|l'} q'$ is an edge of type $-k$ and we obtain

$$q' = (-k, \varepsilon_1, \ldots, \varepsilon_{n-1})_\varepsilon.$$

Since $|k| \leq 1$ this implies $q' \in E$ and we are done. □

Let $\mathcal{A}(\overline{E}) = \mathcal{A}(E)$ be the restriction of $\mathcal{A}(\mathcal{R})$ to the set of vertices E. Since E is closed in the sense of Definition 3.5 we conclude that for any pair $q \in E$, $l \in \mathcal{N}$ there exist $q' \in E$, $l' \in \mathcal{N}$ such that $q \xrightarrow{l|l'} q'$ is an edge in $\mathcal{A}(E)$.

We end this section with a definition that relates certain paths in the graph $\mathcal{A}(\mathcal{R})$ to the iterates of the function Φ defined in the previous section.

DEFINITION 3.7. Let G be a subgraph of $\mathcal{A}(\mathcal{R})$. Suppose we start at a vertex v_0. If we use an input string consisting only of zeros, the corresponding walk $v_0 \to v_1 \to v_2 \to \cdots$ is called the zero walk starting in v_0. Note that by the definition of Φ we have $v_j = \Phi^j(v_0)$ for each $j \in \mathbb{N}$.

Definition 3.7 implies that zero walk in a subgraph of $\mathcal{A}(\mathcal{R})$ ends up in a cycle whose vertices are contained in $\mathcal{P}$ after finitely many steps.

4. The algorithmic characterization

In the present section we give a fast algorithm which decides whether a given polynomial $P(x)$ provides a CNS, provided that $P(x)$ satisfies the AP condition. With

help of this algorithm we will be able to determine a large class of polynomials which provide a CNS. First we need some preparatory results.

Proposition 4.1. *The addition of a number $z \in E$ is computable by a finite transducer automaton if each element of $\overline{\{z\}}$ has finite representation.*

Proof. Since $\overline{\{z\}} \subset E$, Lemma 3.6 shows that $\overline{\{z\}}$ is finite. If we start in a vertex $z \in E$ of the graph $\mathcal{A}(\mathcal{R})$ and move through its edges using the digits $d_j(y)$ of $y \in \mathcal{R}$ as input string we get the digits $d_j(z+y)$ of $z+y$ as output string. Since E is closed in the sense of Definition 3.5 during this procedure we never leave the subgraph $\mathcal{A}(E)$. Thus the addition of z to an arbitrary number $y \in \mathcal{R}$ can be performed with help of the finite graph $\mathcal{A}(E)$. Using the notation of Definition 3.1 we now define the transducer automaton $A(z)$ in the following way. Set

$$\begin{aligned} Q &:= E, \\ \Sigma = \Delta &:= \mathcal{N}, \\ T(v,l) &:= v', \text{ where } v' \text{ is the solution of } v + l = l' + xv' \quad (l' \in \mathcal{N}), \\ R(v,l) &:= l', \text{ where } l' \in \mathcal{N} \text{ is the solution of } v + l = l' + xv', \\ v_0 &:= z. \end{aligned}$$

Thus $A(z)$ regarded as a graph in the sense explained after Definition 3.1 is equal to $\mathcal{A}(E)$. So $A(z)$ is able to read the input digits $d_j(y)$ and returns the output digits $d_j(z+y)$.

In order to fulfill the requirements of Definition 3.2 it remains to check that $A(z)$ transforms finite representations to finite representations. Since the input string has only finitely many nonzero digits after finitely many steps we enter a zero walk. On entering this zero walk the automaton rests at a certain state $v \in \overline{\{z\}}$. By Definition 3.7 this zero walk runs through the vertices $\Phi^j(v)$. Since by assumption v has finite representation this walk reaches zero after finitely many steps. Because the result function R of the automaton $A(z)$ fulfills $R(0,0) = 0$, from this point on the output digits are all equal to zero. Thus the output string is the digit string of a finite representation, and we are done. □

This proposition allows us to prove a first algorithmic criterion to check whether a given polynomial $P(x)$ provides a CNS or not.

Proposition 4.2. *Let $P(x)$ be as in* Definition 1.1 *and suppose that $P(x)$ fulfills the AP condition. Set $s_1 := (1,0,\ldots,0)_\varepsilon, \ldots, s_n := (0,\ldots,0,1)_\varepsilon$ and $F := \{\pm s_1, \ldots, \pm s_n\}$.*

Then $(P(x), \mathcal{N})$ is a CNS if and only if each $z \in \overline{F}$ has a finite representation.

Proof. Since $\overline{F} \subset E$, by Proposition 4.1 the addition of each $\pm s_j \in F$ is computable by a finite transducer automaton $A(\pm s_j)$. We have to show that each $y \in \mathcal{R}$ has a finite representation. By Lemma 3.3, y has an ε-representation of the form

$$y = (\varepsilon_1, \ldots, \varepsilon_n)_\varepsilon \quad (\varepsilon_i \in \mathbb{Z}).$$

We will now build up y in finitely many steps starting from $z_0 = 0 = (0, \ldots, 0)_\varepsilon$. Obviously, 0 has the finite representation (0). Putting the digits of this trivial representation in the automaton $A(\text{sign}(\varepsilon_1)s_1)$ we obtain a finite representation of $z_1 = (\text{sign}(\varepsilon_1) \cdot 1, 0, \ldots, 0)_\varepsilon$. Putting the finite representation of z_1 again in the automaton $A(\text{sign}(\varepsilon_1)s_1)$ and repeating this procedure $|\varepsilon_1|$ times produces a finite representation of

$$z_{|\varepsilon_1|} = (\text{sign}(\varepsilon_1)|\varepsilon_1|, 0, \ldots, 0)_\varepsilon = (\varepsilon_1, 0, \ldots, 0)_\varepsilon.$$

The finiteness of this representation is assured by the fact that the automaton $A(\text{sign}(\varepsilon_1)s_1)$ sends finite representations to finite representations. Now we put $z_{|\varepsilon_1|}$ for $|\varepsilon_2|$ times in the automaton $A(\text{sign}(\varepsilon_2)s_2)$ which yields that $z_{|\varepsilon_1|+|\varepsilon_2|} = (\varepsilon_1, \varepsilon_2, 0, \ldots, 0)_\varepsilon$ has finite representation. Treating the other coordinates in the same way we finally arrive at

$$y = z_{|\varepsilon_1|+\cdots+|\varepsilon_n|} = (\varepsilon_1, \ldots, \varepsilon_n)_\varepsilon.$$

This implies that y has finite representation. Since y was arbitrary we conclude that $(P(x), \mathcal{N})$ is a CNS. □

REMARK 4.3. This proves Conjecture 2. of [2] under the AP condition.

We are now in a position to prove our first main result.

Theorem 4.4. *Let $P(x)$ be as in* Definition 1.1 *and suppose that $P(x)$ satisfies the AP condition. Then $(P(x), \mathcal{N})$ is a CNS if and only if each element of the set*

$$D := \{q \mid q = (\varepsilon_1, \ldots, \varepsilon_n)_\varepsilon,\ \varepsilon_i \in \{0, 1\}\} \tag{4.1}$$

has a finite representation.

Proof. By Lemma 2.1, $(P(x), \mathcal{N})$ is a CNS if $\mathcal{P} = \{0\}$, i.e. if all zero walks in $\mathcal{A}(\mathcal{R})$ end up in the cycle at the vertex zero. By Proposition 4.2 it is sufficient to check this only for the zero paths starting in $z \in E$, since $\overline{F} \subset E$. What we have to show is that we can even confine ourselves to checking all zero paths starting in $z \in D$. We will prove this in the following way. Suppose that there exists an element $z := (\varepsilon_1, \ldots, \varepsilon_n)_\varepsilon \in E$ with at least one $\varepsilon_j = -1$ having infinite representation. If we can

show that this implies the existence of an element $(\varepsilon_1', \dots, \varepsilon_n')_\varepsilon \in D$ having infinite representation we are done. Indeed, this would imply that if there exist elements with infinite representation then some of them must lie in D.

Suppose now that $z \in E$ has infinite representation. We will look more closely on what happens if we follow the zero walk starting at z. Since the input digit $l = 0$, the AP condition implies that

$$-b_0 < \varepsilon_1 b_1 + \cdots + \varepsilon_n b_n + l < b_0.$$

Hence, there exist $k \in \{-1, 0\}$ and $l' \in \mathcal{N}$ such that

$$\varepsilon_1 b_1 + \cdots + \varepsilon_n b_n + l = k b_0 + l'.$$

Arguing in the same way as in the paragraphs preceding the statement of Lemma 3.6 yields that the first edge on the zero walk is

$$(\varepsilon_1, \dots, \varepsilon_n)_\varepsilon \to (-k, \varepsilon_1, \dots, \varepsilon_{n-1})_\varepsilon.$$

Thus after one step on the zero walk the first coordinate in the ε-representation is zero or one. Iterating this procedure n times yields

$$(\varepsilon_1, \dots, \varepsilon_n)_\varepsilon \to \underbrace{\cdots\cdots}_{n \text{ steps}} \to (\varepsilon_1', \dots, \varepsilon_n')_\varepsilon =: z'$$

with $\varepsilon_j' \in \{0, 1\}$ for $1 \le j \le n$. Thus $z' \in D$. Note that by assumption z has an infinite representation. Thus $\{\Phi^j(z)\}_{j \ge 0}$ is not ultimately zero. Since $z' = \Phi^n(z)$ this implies that $\{\Phi^j(z')\}_{j \ge 0} = \{\Phi^{j+n}(z)\}_{j \ge 0}$ is not ultimately zero. Hence, z' has an infinite representation. Thus we found an element of D having an infinite representation. This ends the proof. □

5. The characterization of a class of CNS polynomials

In what follows we want to exhibit algebraic conditions which will allow us to decide whether a given polynomial $P(x)$ provides a CNS or not. These conditions will enable us to characterize a large class of CNS polynomials. First we want to give the following definition.

DEFINITION 5.1 (cf. Berlekamp [7, p. 84], or Lothaire [26]).

- We say a string w' emerges from the string w by digit rotation if there exist two words u and v such that $w = uv$ and $w' = vu$.
- A string $w_1 \cdots w_k$ has period $p \in \{1, \dots, k\}$ if $w_j = w_{j+p}$ for $j = 1, \dots, k - p$.
- A string $w_1 \cdots w_k$ has primitive period $p_0 \in \{1, \dots, k\}$ if it has period p_0 and no period $p < p_0$.

• A *necklace* of length k is an equivalence class of strings of length k under rotation.

• A *primitive necklace* of length k is a necklace with primitive period k (i.e. an aperiodic necklace).

• A *Lyndon word* is the lexicographically smallest representative of a primitive necklace.

From the proof of Theorem 4.4 follows that it is sufficient to check if all elements of the set D have finite representation in order to decide whether a given $P(x)$ forms a CNS or not. By the definition of Φ this is equivalent to $\mathcal{P}_D = \{0\}$, where

$$\mathcal{P}_D := \{p \in D \mid \exists \omega \in \mathbb{N} : p = \Phi^{\omega}(p)\}.$$

Since each $p \in \mathcal{P}_D$ generates a zero cycle in the graph $\mathcal{A}(D)$, which is the restriction of $\mathcal{A}(\mathcal{R})$ to the set of vertices D, we get the following simple result.

Lemma 5.2. *Let $P(x)$ be as in* Definition 1.1 *fulfilling the AP condition. Then $(P(x), \mathcal{N})$ is a CNS if and only if $\mathcal{A}(D)$ contains no zero cycle apart from $(0, \ldots, 0)_\varepsilon \to (0, \ldots, 0)_\varepsilon$.*

Note that the longest zero cycle contained in $\mathcal{A}(D)$ can not be longer than $2^n - 1$ because $\mathcal{A}(D)$ has 2^n states and $(0, \ldots, 0)_\varepsilon$ must not occur in a nontrivial cycle. Suppose that the state

$$(\varepsilon_1, \ldots, \varepsilon_k, \varepsilon_1, \ldots, \varepsilon_k, \ldots, \varepsilon_{n \bmod k})_\varepsilon, \quad \varepsilon_j \in \{0, 1\}$$

of $\mathcal{A}(D)$ has — regarded as a binary string — primitive period k for a $k \in \{1, \ldots, 2^n - 1\}$. (Here $n \bmod k$ is chosen from the residue system $\{1, \ldots, k\}$ modulo k.) If this state belongs to a zero cycle of length k, this is the zero cycle generated by the period $\varepsilon_1 \varepsilon_2 \cdots \varepsilon_k$, i.e.

$$\begin{array}{llll}
(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k, & \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k, & \ldots\ldots & \varepsilon_{n \bmod k})_\varepsilon \xrightarrow{\text{type } \varepsilon_k} \\
(\varepsilon_k, \varepsilon_1, \ldots, \varepsilon_{k-1}, & \varepsilon_k, \varepsilon_1, \ldots, \varepsilon_{k-1}, & \ldots\ldots & \varepsilon_{n-1 \bmod k})_\varepsilon \xrightarrow{\text{type } \varepsilon_{k-1}} \\
\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots & \ldots\ldots\ldots \quad \ldots\ldots\ldots \\
(\varepsilon_2, \varepsilon_3, \ldots, \varepsilon_1, & \varepsilon_2, \varepsilon_3, \ldots, \varepsilon_1, & \ldots\ldots & \varepsilon_{n-k+1 \bmod k})_\varepsilon \xrightarrow{\text{type } \varepsilon_1} \\
(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k, & \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k, & \ldots\ldots & \varepsilon_{n \bmod k})_\varepsilon
\end{array}$$

(note that $\varepsilon_{n \bmod k} = \varepsilon_{n-k \bmod k}$). Since $(0, \ldots, 0)_\varepsilon$ must not occur in a nontrivial cycle, the state

$$(\varepsilon_1, \ldots, \varepsilon_k, \varepsilon_1, \ldots, \varepsilon_k, \ldots, \varepsilon_{n \bmod k})_\varepsilon, \quad \varepsilon_j \in \{0, 1\},$$

regarded as a binary string must not contain more than $n-1$ consecutive zeros. Let

$$J_0 := \{j \mid 1 \le j \le k \text{ and } \varepsilon_{j-1 \bmod k} = 0\},$$
$$J_1 := \{j \mid 1 \le j \le k \text{ and } \varepsilon_{j-1 \bmod k} = 1\}.$$

Then by the definition of type t edges this cycle exists if and only if the system of inequalities

$$\begin{aligned} 0 &\le \varepsilon_j b_1 + \cdots + \varepsilon_k b_{k-j+1} + \varepsilon_1 b_{k-j+2} + \cdots + \varepsilon_{j-1} b_k < b_0 \text{ for } j \in J_0, \\ -b_0 &\le \varepsilon_j b_1 + \cdots + \varepsilon_k b_{k-j+1} + \varepsilon_1 b_{k-j+2} + \cdots + \varepsilon_{j-1} b_k < 0 \quad \text{for } \in J_1 \end{aligned} \tag{5.1}$$

holds (note that $b_l = 0$ for $l > n$). Summing up what we have proved we get the following result.

Lemma 5.3. *Let $P(x)$ be as in* Definition 1.1 *satisfying the AP condition. Then there exists a zero cycle of length $k \in \{1, \ldots, 2^n - 1\}$ generated by the period $\varepsilon_1 \cdots \varepsilon_k$ if and only if the inequalities* (5.1) *hold simultaneously. Thus to each zero cycle of length k there corresponds a set of k inequalities.*

REMARK 5.4. Note that (5.1) provides a full characterization of all CNS which fulfill the AP condition. Let $P(x)$ be as in Definition 1.1 fulfilling the AP condition. Then $P(x)$ provides a CNS if and only if the set of inequalities (5.1) does not hold simultaneously for any cycle of length $\le 2^n - 1$. Of course, this criterion is very hard to survey. In Section 6, however, we will show that it can be used to derive simple algebraic criteria in certain cases.

In what follows we will have to deal with concrete cycles of small length. Thus we are interested in how many cycles of a given length k exist in the graph $\mathcal{A}(D)$. To this matter we need the following result.

Lemma 5.5. *Let L_k be the number of binary Lyndon words of length k. Then*

$$L_k = \frac{1}{k} \sum_{d \mid k} \mu\left(\frac{k}{d}\right) 2^d \tag{5.2}$$

where μ denotes the Möbius function.

Proof. The number of binary words of length k is 2^k. Each word of length k has primitive period d with $d \mid k$. Therefore

$$2^k = \sum_{d \mid k} d L_d.$$

Möbius inversion yields (5.2). □

REMARK 5.6. Concerning the sequence L_k we refer to sequence *A001037* in Sloane's database of integer sequences [30].

The above notion is useful for counting the possible zero cycles in $\mathcal{A}(D)$.

Lemma 5.7. *For each $k \in \{1, \dots, 2^n - 1\}$ there exist at most L_k possible zero cycles of length k in $\mathcal{A}(D)$.*

Proof. Let $\mathcal{L}_k$ be the set of Lyndon words of length k which do not have more than n leading zeros and do not contain two equal subwords of length n. If we regard the ε-representations of the states of $\mathcal{A}(D)$ as strings, each vertex is a binary string of length n. A cycle of length $k \in \{1, \dots, 2^n - 1\}$ is then generated by an equivalence class of binary strings of primitive period k under rotation. Thus by Definition 5.1 there is a one-to-one correspondence between the nontrivial cycles of length k which do not contain $(0, \dots, 0)_\varepsilon$ and $\mathcal{L}_k$. The lemma now follows from Lemma 5.5. □

We are now in a position to state the criterion.

Theorem 5.8. *Let $P(x)$ be as in* Definition 1.1 *satisfying the AP condition. If*

$$\sum_{j=1}^{n} b_j \geq 0 \tag{5.3}$$

and

$$b_2 \geq 0, \dots, b_{n-1} \geq 0 \tag{5.4}$$

then $(P(x), \mathcal{N})$ is a CNS.

Proof. By Lemma 5.2 we have to show that under the assumptions of the theorem there exists no zero cycle in $\mathcal{A}(D)$ apart from $0 \to 0$. We start with the examination of the cycle

$$(1, \dots, 1)_\varepsilon \xrightarrow{\text{type } 1} (1, \dots, 1)_\varepsilon. \tag{5.5}$$

By Lemma 5.3 this is the only possible nontrivial zero cycle of length one. This zero cycle exists if and only if

$$b_1 + \cdots + b_n < 0$$

holds. Since this inequality contradicts (5.3) we conclude that the zero cycle (5.5) does not exist in $\mathcal{A}(D)$.

Now suppose that there exists a zero cycle of length $k \geq 2$ in $\mathcal{A}(D)$. Let $\varepsilon_1 \cdots \varepsilon_k$ be the corresponding Lyndon word. Hence $\varepsilon_1 \cdots \varepsilon_k$ is lexicographic minimal and $\varepsilon_k = 1$. Otherwise one could rotate $\varepsilon_1 \cdots \varepsilon_k$ by one digit to get one more leading zero. Furthermore, there exists a minimal $m \in \{2, \ldots, k\}$ with $\varepsilon_m = 1$ such that

$$\varepsilon_1 \cdots \varepsilon_k = 0 \cdots 0 \varepsilon_m \cdots \varepsilon_k.$$

Since $\varepsilon_k = 1$, there must be a type 1 edge

$$\begin{array}{l}(0, 0, \ldots, 0, \ \varepsilon_m, \ \varepsilon_{m+1}, \ \ldots \ \varepsilon_k, \quad \ldots\ldots)_\varepsilon \xrightarrow{\text{type } 1} \\ (1, 0, \ldots, 0, \ 0, \quad \varepsilon_m, \quad \ldots \ \varepsilon_{k-1}, \ \ldots\ldots)_\varepsilon.\end{array}$$

This is only possible if

$$\varepsilon_m b_m + \cdots + \varepsilon_n b_n < 0.$$

Since $\varepsilon_i \geq 0$, this inequality contradicts (5.4). □

REMARK 5.9. Theorem 5.8 proves Conjecture 1 of Akiyama-Pethő [2].

6. Characterization of cubic and quartic CNS

In order to characterize cubic and quartic CNS we have to look more closely to the possible zero cycles up to length $2^3 - 1 = 7$ and $2^4 - 1 = 15$, respectively. First of all we give a complete list of these cycles up to length 4 together with their associated inequalities. To this matter let $P(x)$ be as in Definition 1.1 satisfying the AP condition. Furthermore, set $b_n = 1$ and $b_j = 0$ for $j > n$. Now we determine the possible cycles of the graph $\mathcal{A}(D)$ associated to $P(x)$ up to length 4 as well as one important cycle of length 5.

- There exist two possible cycles of length one:
  (1.a) $(0, \ldots, 0)_\varepsilon \xrightarrow{\text{type } 0} (0, \ldots, 0)_\varepsilon \xrightarrow{\text{type } 0} \ldots$.
  This cycle is the trivial cycle.
  (1.b) $(1, \ldots, 1)_\varepsilon \xrightarrow{\text{type } 1} (1, \ldots, 1)_\varepsilon \xrightarrow{\text{type } 1} \ldots$.
  To this cycle there corresponds the inequality

$$-b_0 < b_1 + \cdots + b_n < 0.$$

- There exists one possible cycle of length two:
  (2) $(0, 1, 0, \ldots)_\varepsilon \xrightarrow{\text{type } 1} (1, 0, 1, \ldots)_\varepsilon \xrightarrow{\text{type } 0} (0, 1, 0, \ldots)_\varepsilon \xrightarrow{\text{type } 1} \ldots$.
  This cycle occurs if

$$\begin{aligned} -b_0 &< b_2 + b_4 + b_6 + \cdots < 0, \\ 0 &\leq b_1 + b_3 + b_5 + \cdots < b_0. \end{aligned}$$

• There exist two possible cycles of length three:
(3.a) $(\overline{1,0,0},\ldots)_\varepsilon \xrightarrow{\text{type } 0} (0,1,0,\ldots)_\varepsilon \xrightarrow{\text{type } 0} (0,0,1,\ldots)_\varepsilon \xrightarrow{\text{type } 1}$
$(1,0,0,\ldots)_\varepsilon \xrightarrow{\text{type } 0} \ldots$.
The related inequalities are

$$\begin{aligned} 0 &\le b_1 + b_4 + b_7 + \cdots < b_0,\\ 0 &\le b_2 + b_5 + b_8 + \cdots < b_0,\\ -b_0 &< b_3 + b_6 + b_9 + \cdots < 0. \end{aligned}$$

(3.b) $(0,1,1,\ldots)_\varepsilon \xrightarrow{\text{type } 1} (1,0,1,\ldots)_\varepsilon \xrightarrow{\text{type } 1} (1,1,0,\ldots)_\varepsilon \xrightarrow{\text{type } 0}$
$(0,1,1,\ldots)_\varepsilon \xrightarrow{\text{type } 1} \ldots$.
The related inequalities are

$$\begin{aligned} -b_0 &< b_2 + b_3 + b_5 + b_6 \ldots < 0,\\ -b_0 &< b_1 + b_3 + b_4 + b_6 \ldots < 0,\\ 0 &\le b_1 + b_2 + b_4 + b_5 \ldots < b_0. \end{aligned}$$

• There exist three possible cycles of length four:
(4.a) $(1,0,0,0,\ldots)_\varepsilon \xrightarrow{\text{type } 0} (0,1,0,0,\ldots)_\varepsilon \xrightarrow{\text{type } 0} (0,0,1,0,\ldots)_\varepsilon \xrightarrow{\text{type } 0}$
$(0,0,0,1,\ldots)_\varepsilon \xrightarrow{\text{type } 1} \ldots$.
The related inequalities are

$$\begin{aligned} 0 &\le b_1 + b_5 + b_9 + \cdots < b_0,\\ 0 &\le b_2 + b_6 + b_{10} + \cdots < b_0,\\ 0 &\le b_3 + b_7 + b_{11} + \cdots < b_0,\\ -b_0 &< b_4 + b_8 + b_{12} + \cdots < 0. \end{aligned}$$

(4.b) $(1,1,0,0,\ldots)_\varepsilon \xrightarrow{\text{type } 0} (0,1,1,0,\ldots)_\varepsilon \xrightarrow{\text{type } 0} (0,0,1,1,\ldots)_\varepsilon \xrightarrow{\text{type } 1}$
$(1,0,0,1,\ldots)_\varepsilon \xrightarrow{\text{type } 1} \ldots$.
The related inequalities are

$$\begin{aligned} 0 &\le b_1 + b_2 + b_5 + \cdots < b_0,\\ 0 &\le b_2 + b_3 + b_6 + \cdots < b_0,\\ -b_0 &\le b_3 + b_4 + b_7 + \cdots < 0,\\ -b_0 &< b_1 + b_4 + b_5 + \cdots < 0. \end{aligned}$$

(4.c) $(1,1,1,0,\ldots)_\varepsilon \xrightarrow{\text{type } 0} (0,1,1,1,\ldots)_\varepsilon \xrightarrow{\text{type } 1} (1,0,1,1,\ldots)_\varepsilon \xrightarrow{\text{type } 1}$
$(1,1,0,1,\ldots)_\varepsilon \xrightarrow{\text{type } 1} \ldots$.

The related inequalities are

$$\begin{aligned} 0 &\le b_1 + b_2 + b_3 + \cdots < b_0,\\ -b_0 &\le b_2 + b_3 + b_4 + \cdots < 0,\\ -b_0 &\le b_1 + b_3 + b_4 + \cdots < 0,\\ -b_0 &< b_1 + b_2 + b_4 + \cdots < 0. \end{aligned}$$

• By Lemma 5.7 there exist six possible cycles of length five. Since only one of them will play a prominent role in the forthcoming calculations we will confine ourselves to writing down only this one:

$$\begin{aligned} (5)\quad & (1,0,0,1,0,1,\ldots)_\varepsilon \xrightarrow{\text{type } 0} (0,1,0,0,1,0,\ldots)_\varepsilon \xrightarrow{\text{type } 1}\\ & (1,0,1,0,0,1,\ldots)_\varepsilon \xrightarrow{\text{type } 0} (0,1,0,1,0,0,\ldots)_\varepsilon \xrightarrow{\text{type } 0}\\ & (0,0,1,0,1,0,\ldots)_\varepsilon \xrightarrow{\text{type } 1} (1,0,0,1,0,1,\ldots)_\varepsilon \xrightarrow{\text{type } 0} \ldots. \end{aligned}$$

This cycle occurs if

$$\begin{aligned} 0 &\le b_1 + b_4 + b_6 + \cdots < b_0,\\ -b_0 &< b_2 + b_5 + b_7 + \cdots < 0,\\ 0 &\le b_1 + b_3 + b_6 + \cdots < b_0,\\ 0 &\le b_2 + b_4 + b_7 + \cdots < b_0,\\ -b_0 &< b_3 + b_5 + b_8 + \cdots < 0. \end{aligned}$$

Of course it is an easy task to extend this list up to cycles of arbitrary length with increasing effort. To this matter one needs to know explicitly all Lyndon words up to a certain length.

6.1. The cubic case. Let $P(x)$ be as in Definition 1.1 satisfying the AP condition. In this subsection we want to find simple algebraic conditions under which there do not exist cycles in $\mathcal{A}(D)$. This will lead to a complete characterization of cubic CNS under the AP condition. Since the condition related to the cycle (1.b) must not be fulfilled we get the necessary condition

$$b_1 + b_2 + 1 \ge 0 \tag{6.1}$$

for $P(x)$ to provide a cubic CNS. Next we deal with the zero cycle (2). Suppose that $b_2 < 0$. Then, in order to avoid this zero cycle, we must have $b_1 + 1 < 0$ because otherwise both inequalities for the cycle (2) would be fulfilled. But adding these two inequalities gives an inequality which contradicts (6.1). Thus we get

$$b_2 \ge 0. \tag{6.2}$$

(6.1) and (6.2) exclude the occurrence of the cycles (1.b) and (2). Since there can occur zero cycles up to length seven we have to check next whether (3.a) and (3.b) can exist. The third inequality of (3.a) reads $-b_0 < 1 < 0$ in the cubic case. It can ob-

viously never be fulfilled. Thus (3.a) can not occur. The first inequality of (3.b) reads $-b_0 < b_2+1 < 0$. By (6.2) it can never be fulfilled. Thus also this cycle can not occur.

Similarly, one can exclude the occurrence of all the other cycles up to length seven. By Lemma 5.7 one has to check all cycles which do not contain more than three consecutive zeros. These cycles correspond to all binary Lyndon words with length $\leq$ seven with less than three leading zeros. This are $\sum_{k=1}^{7} L_k = 41$ words. Thus we have to check 41 sets of inequalities. This can be done easily with help of a short computer program. One can considerably diminish the number of sets of inequalities by arguing in a similar way as in the quartic case below.

So the necessary conditions (6.1) and (6.2) assure that there does not exist a zero cycle in $\mathcal{A}(D)$ apart from $0 \to 0$. Thus they are also sufficient. Summing up we proved the following result.

Theorem 6.1. *Let* $P(x) = x^3 + b_2x^2 + b_1x + b_0$ *satisfying the AP condition. Then* $(P(x), \mathcal{N})$ *is a CNS if and only if*

$$b_1 + b_2 + 1 \geq 0 \quad \textit{and} \quad b_2 \geq 0$$

hold.

6.2. The quartic case. For quartic polynomials we get the following characterization result.

Theorem 6.2. *Let* $P(x) = x^4 + b_3x^3 + b_2x^2 + b_1x + b_0$ *satisfying the AP condition. Then* $(P(x), \mathcal{N})$ *is a CNS if and only if*

$$b_1 + b_2 + b_3 + 1 \geq 0,\ b_2 + b_3 \geq -1,\ b_2 \geq -1,\ b_1 \geq -1,\ b_3 \geq 0 \quad \textit{or}$$
$$b_1 + b_2 + b_3 + 1 \geq 0,\ b_2 + b_3 \geq -1,\ b_2 \geq -1,\ b_1 < -1,\ b_3 \geq -1$$

holds.

In order to prove this theorem we need the following preparatory lemma.

Lemma 6.3. *Let* $P(x)$ *be a quartic polynomial* $P(x) = x^4 + b_3x^3 + b_2x^2 + b_1x + b_0$ *which fulfills*

$$b_3 + 1 \geq 0,$$
$$b_2 + 1 \geq 0.$$

Suppose that no cycle of length less than or equal to 6 *exists in* $\mathcal{A}(D)$. *Then the related graph* $\mathcal{A}(D)$ *contains no cycle of length greater than* 6.

Proof. Note that $\mathcal{A}(D)$ has 16 states. If we can show that there exist 10 states which can not be contained in a cycle of length greater than 6 the result follows. For abbreviation we call a cycle of length greater than 6 a *long* cycle.

Suppose that $(0000)_\varepsilon$ is contained in a long cycle. Then this cycle contains either the edge $(0000)_\varepsilon \xrightarrow{\text{type } 0} (0000)_\varepsilon$ or the edge $(0000)_\varepsilon \xrightarrow{\text{type } 1} (1000)_\varepsilon$. In the first case $(0000)_\varepsilon$ forms a cycle of length 1, the second case would yield the inequalities $-b_0 < 0 < 0$ which are never fulfilled. Thus $(0000)_\varepsilon$ can not be contained in a long cycle.

Suppose that $(0001)_\varepsilon$ were contained in a long cycle. Then, since $(0000)_\varepsilon$ is excluded, we must have an edge $(0001)_\varepsilon \xrightarrow{\text{type } 1} (1000)_\varepsilon$ in this long cycle. But this implies the inequalities $-b_0 < 1 < 0$ which are never fulfilled. Thus $(0001)_\varepsilon$ can not be contained in a long cycle.

$(1000)_\varepsilon$ can also not be an element of a long cycle because it would imply that one of the edges $(0000)_\varepsilon \to (1000)_\varepsilon$ or $(0001)_\varepsilon \to (1000)_\varepsilon$ were contained in it, which is impossible by the above paragraphs.

Now we show that neither $(0101)_\varepsilon$ nor $(0010)_\varepsilon$ can belong to a long cycle. To this matter we distinguish three cases.

- *The case* $b_3 \geq 0$. $(0101)_\varepsilon \xrightarrow{\text{type } 1} (1010)_\varepsilon$ is impossible because it would lead to $-b_0 < b_2 + 1 < 0$, a contradiction. Thus we must have $(0101)_\varepsilon \xrightarrow{\text{type } 0} (0010)_\varepsilon$. But $(0010)_\varepsilon$ has no successor since $(0001)_\varepsilon$ can not belong to a long cycle and the existence of the edge $(0010)_\varepsilon \xrightarrow{\text{type } 1} (1001)_\varepsilon$ would lead to the contradiction $b_3 < 0$.
- *The case* $b_3 = -1$ *and* $b_2 \geq 0$. In this case we must have $b_1 + 1 < 0$, because otherwise the cycle (3.a) would exist. Furthermore, $b_1 + b_2 \geq 0$ and $b_2 + b_3 \geq 0$ must hold, because otherwise the cycle (1.b) would exist. But under these conditions the only walk starting from $(0101)_\varepsilon$ is

$$\begin{aligned}(0101)_\varepsilon &\xrightarrow{\text{type } 0} (0010)_\varepsilon \xrightarrow{\text{type } 1} (1001)_\varepsilon \xrightarrow{\text{type } 1} (1100)_\varepsilon \\ &\xrightarrow{\text{type } 0} (0110)_\varepsilon \xrightarrow{\text{type } 0} (0011)_\varepsilon \xrightarrow{\text{type } 0} (0001)_\varepsilon.\end{aligned}$$

But since $(0001)_\varepsilon$ can not exist in a long cycle, neither $(0101)_\varepsilon$ nor $(0010)_\varepsilon$ can belong to a long cycle.

- *The case* $b_3 = -1$ *and* $b_2 = -1$. In this case we must have $b_1 + 1 > 0$ and $b_1 + b_3 \geq 0$ in order to avoid cycle (1.b). Thus the only walk leading away from $(0101)_\varepsilon$ is

$$\begin{aligned}(0101)_\varepsilon &\xrightarrow{\text{type } 0} (0010)_\varepsilon \xrightarrow{\text{type } 1} (1001)_\varepsilon \xrightarrow{\text{type } 0} (0100)_\varepsilon \\ &\xrightarrow{\text{type } 1} (1010)_\varepsilon \xrightarrow{\text{type } 0} (0101)_\varepsilon.\end{aligned}$$

This is a cycle of length 5. Thus also in this case neither $(0101)_\varepsilon$ nor $(0010)_\varepsilon$ can belong to a long cycle.

Since the above mentioned 5 elements can not be contained in a long cycle, the

only path of length 3 which can lead to $(1010)_\varepsilon$ in a long cycle is given by

$$(0011)_\varepsilon \xrightarrow{\text{type } 1} (1001)_\varepsilon \xrightarrow{\text{type } 0} (0100)_\varepsilon \xrightarrow{\text{type } 1} (1010)_\varepsilon$$

But the first edge yields $-b_0 < b_3 + 1 < 0$, a contradiction. Thus $(1001)_\varepsilon$, $(0100)_\varepsilon$ and $(1010)_\varepsilon$ can not exist in a long cycle.

It is now easy to see that $(0011)_\varepsilon$ can have no successors and $(1100)_\varepsilon$ can have no predecessors in a long cycle. Thus, summing up we get that the elements

$$(0000)_\varepsilon,\ (0001)_\varepsilon,\ (0010)_\varepsilon,\ (0100)_\varepsilon,\ (1000)_\varepsilon$$
$$(1001)_\varepsilon,\ (1010)_\varepsilon,\ (0101)_\varepsilon,\ (1100)_\varepsilon,\ (0011)_\varepsilon$$

can not occur in a long cycle. Thus there can not exist cycles of length greater than 6 if the conditions of the lemma are fulfilled and we are done. □

After this preparation we are in a position to prove Theorem 6.2.

Proof. By similar arguments as in the cubic case we find that the cycle (1.b) does not exist if

$$b_1 + b_2 + b_3 + 1 \geq 0. \tag{6.3}$$

Again in the same way as above we see that the cycle (2) does not exist if

$$b_2 + 1 \geq 0. \tag{6.4}$$

Furthermore, we see that cycle (3.b) does not exist if one of the following conditions is true.

(i) $b_2 + b_3 \geq 0$,
(ii) $b_1 + b_3 + 1 \geq 0$,
(iii) $b_1 + b_2 + 1 < 0$.

If (iii) holds then $b_3 \geq 1$ by (6.3). Together with (6.4) this implies $b_2 + b_3 \geq 0$ and we reduced this case to (i).

In order to treat (i) we have to distinguish four cases.

- *The case* $b_2 \geq 0$ *and* $b_1 + 1 \geq 0$. In this case we must have

$$b_3 \geq 0$$

because otherwise the cycle (3.a) would exist.

- *The case* $b_2 \geq 0$ *and* $b_1 + 1 < 0$. In this case we must have

$$b_3 \geq -1$$

because otherwise the cycle (4.b) would exist.

• *The case* $b_2 = -1$ *and* $b_1 + 1 \geq 0$. In this case we must have either $b_3 \geq 0$ or $b_1 + b_3 < 0$ because otherwise the cycle (5) would exist. Since $b_1 + b_3 < 0$ together with $b_2 = -1$ would imply the existence of the cycle (1.b) we conclude that

$$b_3 \geq 0$$

has to hold in this case.

• *The case* $b_2 = -1$ *and* $b_1 + 1 < 0$. In this case we must have

$$b_3 \geq 0$$

because otherwise the cycle (1.b) would exist.

It remains to deal with (ii). Since (i) is already treated we can assume that

$$b_2 + b_3 < 0. \tag{6.5}$$

Suppose first that $b_2 \geq 0$. Then (6.5) implies $b_3 < 0$ and thus by (ii) we also have $b_1 \geq 0$. But the last three inequalities yield the existence of the cycle (3.a). Thus $b_2 = -1$ must hold. In this case exactly the same arguments as used in (i) yield the additional condition $b_3 \geq 0$.

Summing up we get that the condition

$$\begin{array}{l} b_1 + b_2 + b_3 + 1 \geq 0,\ b_2 + b_3 \geq -1,\ b_2 \geq -1,\ b_1 \geq -1,\ b_3 \geq 0 \quad \text{or} \\ b_1 + b_2 + b_3 + 1 \geq 0,\ b_2 + b_3 \geq -1,\ b_2 \geq -1,\ b_1 < -1,\ b_3 \geq -1 \end{array} \tag{6.6}$$

is necessary for the quartic polynomial $P(x)$ to provide a CNS.

Furthermore, it is easy to check that this condition ensures that none of the cycles up to length 6 can exist. To this matter by Lemma 5.7 one has to check $\sum_{k=1}^{6} L_k = 23$ sets of inequalities. Since $b_1 - 1 \geq 0$ and $b_3 - 1 \geq 0$, Lemma 6.3 yields that there does not exist any cycle of length greater than 6. Thus condition (6.6) is also sufficient and the proof is finished. □

7. Some interesting examples

In this section we present some numerically constructed examples of polynomials which have interesting properties. The polynomials in this examples do not fulfill the AP condition. However, they are expanding in the sense that each of their roots lie outside the unit circle.

• It was conjectured for a long time that if $P(x)$ is a CNS polynomial, then $P(x)+1$ is also a CNS polynomial. A counterexample to this conjecture is given by

$$P(x) = x^3 + 173x^2 + 257x + 198.$$

$P(x)$ is a CNS polynomial (which can be proved by Brunotte's method; cf. [9, 10])

but $P(x)+1$ possesses the cycle

$$\begin{aligned}&(-1,\ 3,\ -3)_\varepsilon \to (\ -1,\ -1,\ 3)_\varepsilon \to (\ \ 3,\ -1,\ -1)_\varepsilon\ \to\\ &(-3,\ 3,\ -1)_\varepsilon \to (\ \ 2,\ -3,\ 3)_\varepsilon \to (\ \ 1,\ \ 2,\ -3)_\varepsilon\ \to\\ &(-3,\ 1,\ \ 2)_\varepsilon \to (\ \ 3,\ -3,\ 1)_\varepsilon \to (\ -1,\ \ 3,\ -3)_\varepsilon.\end{aligned}$$

• Next we give an expanding cubic polynomial having a long cycle. Let

$$P(x) = x^3 + 196x^2 + 341x + 199.$$

Then $P(x)$ has a cycle of length 84. One element of this cycle is $(-11, 10, -6)_\varepsilon$. We conjecture that already cubic polynomials can have arbitrary long cycles.

• Let

$$P(x) = x^3 + 192x^2 + 272x + 199.$$

Then $P(x)$ has a cycle of length eight which consists of elements $(x_1, x_2, x_3)_\varepsilon$ with $|x_i| \geq 2$. One element of this cycle is $(-6, 3, 2)_\varepsilon$. We conjecture that to each $k \in \mathbb{N}$ one can find a cubic polynomial having a cycle all of whose elements $(y_1, y_2, y_3)_\varepsilon$ fulfill $|y_i| \geq k$.

8. Concluding remarks

In the present paper all results apart from the examples in the previous section are subject to the AP condition. Of course it would be desirable to get unconditional results. We fear that this will be hard in general. There are two forthcoming papers concerning related topics:

In Akiyama-Brunotte-Pethő [1], a conjecture of W.J. Gilbert on cubic CNS polynomials is partially proved, and it is shown, that this conjecture is not complete.

In Akiyama-Rao [3], an efficient algorithm is given to determine whether or not $P(x)$ is a CNS polynomial by Brunotte's method [9]. Furthermore, large classes of CNS polynomials are characterized.

Regarding the results under the AP condition we are sure that the sets of inequalities in (5.1) can be considerably simplified also for higher degrees. However the calculations necessary to obtain such a simplification become very hard to survey. Up to now we were not able to find a general principle that allows us to derive simplifications for the characterization of CNS of arbitrary degrees.

References

[1] S. Akiyama, H. Brunotte and A. Pethő: *Cubic CNS polynomials, notes on a conjecture of W.J. Gilbert*, J. Math. Anal. Appl. **281** (2003), 402–415.

[2] S. Akiyama and A. Pethő: *On canonical number systems*, Theoret. Comput. Sci., **270** (2002), 921–933.
[3] S. Akiyama and H. Rao: *New criteria on canonical number systems*, Acta Arith. **111** (2004), 5–25.
[4] S. Akiyama and J.M. Thuswaldner: *Topological properties of two-dimensional number systems*, J. Théor. Nombres Bordeaux, **12** (2000), 69–79.
[5] S. Akiyama and J.M. Thuswaldner: *Topological structure of fractal tilings generated by quadratic number systems*, preprint.
[6] D. Berend and C. Frougny: *Computability by finite automata and Pisot bases*, Math. Systems Theory, **27** (1994), 274–282.
[7] E.R. Berlekamp: Algebraic coding theory, McGraw-Hill, New York, 1968.
[8] J. Berstel: Transductions and context-free languages, Teubner, 1979.
[9] H. Brunotte: *On trinomial bases of radix representations of algebraic integers*, Acta Sci. Math. (Szeged), **67** (2001), 553–559.
[10] H. Brunotte: *Characterisation of CNS trinomials*, Acta Sci. Math. (Szeged), **68** (2002), 673–679.
[11] S. Eilenberg: Automata, languages and machines Academic Press, New York, 1974.
[12] C. Frougny: *Representation of numbers and finite automata*, Math. Systems Theory, **25** (1992), 37–60.
[13] C. Frougny: Number Systems; Chapter 7 in: M. Lothaire, editor, Algebraic combinatorics on words. Cambridge University Press, 2002.
[14] C. Frougny and B. Solomyak: *Finite beta-expansions*, Ergodic Theory Dynam. Systems, **12** (1992), 45–82.
[15] C. Frougny and B. Solomyak: *On representation of integers in linear numeration systems*, In M. Pollicott and K. Schmidt, editors, Ergodic theory of $\mathbf{Z}^d$-Actions. London Mathematical Society Lecture Note Series, **228** (1996), 345–368. Cambridge University Press, Cambridge UK.
[16] W.J. Gilbert: *Radix representations of quadratic fields*, J. Math. Anal. Appl., **83** (1981), 264–274.
[17] I. Kátai: Number systems and fractal geometry, Jannus Pannonius University Pecs, 1995.
[18] I. Kátai and B. Kovács: *Kanonische Zahlensysteme in der Theorie der Quadratischen Zahlen*, Acta Sci. Math. (Szeged), **42** (1980), 99–107.
[19] I. Kátai and B. Kovács: *Canonical number systems in imaginary quadratic fields*, Acta Math. Hungar., **37** (1981), 159–164.
[20] I. Kátai and J. Szabó: *Canonical number systems for complex integers*, Acta Sci. Math. (Szeged), **37** (1975), 255–260.
[21] D.E. Knuth: The art of computer programming, **2** Seminumerical Algorithms, Addison-Wesley, Reading, Mass., 1973.
[22] S. Kőrmendi: *Canonical number systems in* $\mathbb{Q}(3\sqrt{2})$, Acta Sci. Math. (Szeged), **50** (1986), 351–357.
[23] B. Kovács: *Canonical number systems in algebraic number fields*, Acta Math. Hungar., **37** (1981), 405–407.
[24] B. Kovács and A. Pethő: *Number systems in integral domains, especially in orders of algebraic number fields*, Acta Sci. Math. (Szeged), **55** (1991), 286–299.
[25] N. Loraud: *β-shift, systèmes de numération et automates*, J. Théor. Nombres Bordeaux, **7** (1995), 473–498.
[26] M. Lothaire: Combinatorics on words, volume 17 of Encyclopaedia of Mathematics and its Applications, Addison-Wesley, Reading, Mass., 1983. Reprinted in the Cambridge Mathematical Library. Cambridge University Press, Cambridge UK, 1997.
[27] A. Rényi: *Representations for real numbers and their ergodic properties*, Acta Math. Acad. Sci. Hung., **8** (1957), 477–493.
[28] K. Scheicher: *Kanonische Ziffernsysteme und Automaten*, Grazer Math. Ber., **333** (1997), 1–17.
[29] K. Scheicher and J.M. Thuswaldner: *Canonical number systems, counting automata and fractals*, Math. Proc. Cambridge Philos. Soc., **133** (2002), 163–182.
[30] N.J.A. Sloane: Database of integer sequences,

http://www.research.att.com/~njas/sequences.

[31] J.M. Thuswaldner: Elementary properties of canonical number systems in quadratic fields, In G.E. Bergum et. al., editor, Applications of Fibonacci Numbers, **7**, 405–414. Kluwer Academic Publisher, 1998.

[32] J.M. Thuswaldner: *Attractors of invertable expanding mappings and number systems in* $\mathbb{Z}^2$, Publ. Math. Debrecen, **58** (2001), 423–440.

Klaus Scheicher
Institut für Analysis, Abteilung für Finanzmathematik
Johannes Kepler Universität Linz
Altenbergerstraße 69, A-4040 Linz
AUSTRIA
e-mail: klaus.scheicher@jku.at

Current address:
Johann Radon Institute for Computational and Applied Mathematics
Austrian Academy of Sciences
Altenbergerstraße 69, A-4040 Linz
AUSTRIA
e-mail: klaus.scheicher@oeaw.ac.at

Jörg M. Thuswaldner
Institut für Mathematik und Angewandte Geometrie
Abteilung für Mathematik und Statistik
Montanuniversität Leoben
Franz-Josef-Strasse 18, A-8700 Leoben
AUSTRIA
e-mail: joerg.thuswaldner@unileoben.ac.at