

## Syntactic Complexity of Ultimately Periodic Sets of Integers and Application to a Decision Procedure

**Anne Lacroix**

**Narad Rampersad**

**Michel Rigo**

**Élise Vandomme**

*Institute of Mathematics, University of Liège, Grande Traverse 12 (B 37),  
B-4000 Liège, Belgium.*

---

**Abstract.** We compute the cardinality of the syntactic monoid of the language  $0^* \text{rep}_b(m\mathbb{N})$  made of base  $b$  expansions of the multiples of the integer  $m$ . We also give lower bounds for the syntactic complexity of any (ultimately) periodic set of integers written in base  $b$ . We apply our results to a well studied problem: decide whether or not a  $b$ -recognizable set of integers is ultimately periodic.

### 1. Introduction

This paper is dedicated to Tero Harju’s 60th birthday. As we shall explain, it can be related to a decision problem studied by Tero a long time ago [13, 12]. Happy birthday Tero!

Syntactic complexity has received some recent and renewed interest. See for instance [7] for some background, and we quote: “*In spite of suggestions that syntactic semigroups deserve to be studied further, relatively little has been done on the syntactic complexity of a regular language.*”

In this paper, which is an extended version of the conference paper [18], we first compute the syntactic complexity of the set  $m\mathbb{N}$  written in base  $b$ , i.e., the cardinality  $M_{b,m}$  of the syntactic monoid of the language  $0^* \text{rep}_b(m\mathbb{N})$  made of base  $b$  expansions of the multiples of the integer  $m$ . A similar problem was solved for the state complexity of the language  $0^* \text{rep}_b(m\mathbb{N})$ , i.e., the number of states of its minimal automaton. As usual  $(m, n)$  denotes the greatest common divisor of  $m$  and  $n$ .

#### **Theorem 1.1. (B. Alexeev [1])**

Let  $b, m \geq 2$  be integers. Let  $N, M$  be such that  $b^N < m \leq b^{N+1}$  and  $(m, 1) < (m, b) < \dots <$

$(m, b^M) = (m, b^{M+1}) = (m, b^{M+2}) = \dots$ . The minimal automaton of  $0^* \text{rep}_b(m\mathbb{N})$  has exactly

$$\frac{m}{(m, b^{N+1})} + \sum_{t=0}^{\inf\{N, M-1\}} \frac{b^t}{(m, b^t)} \text{ states.}$$

For the binary system, the first few values of  $M_{2,m}$  are given in Table 1. Let  $b \geq 2$ . An explicit

$m$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$M_{2,m}$	3	6	5	20	13	21	7	54	41	110	20	156	43	60	9	136	109	342	62
$m$	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			
$M_{2,m}$	126	221	253	27	500	313	486	65	812	121	155	11	330	273	420	164			

Table 1. The first values of the syntactic complexity in the base 2 case.

formula for  $M_{b,m}$  is obtained as a consequence of Theorems 3.1, 3.2 and 3.3 for the following three cases: the constant  $m$  and the base  $b$  are coprime or,  $m$  is a power of  $b$  or,  $m = b^n q$  with  $(q, b) = 1$ ,  $q \geq 2$  and  $n \geq 1$ . Extending the results of [18], we provide lower bounds for the syntactic complexity of any ultimately periodic set of integers written in base  $b$ , i.e., any finite union of arithmetic progressions. In particular, we include the general case where the period  $m$  is of the kind  $db^n q$  with  $(q, b) = 1$  and the set of prime factors of  $d$  is a subset of the one of  $b$ . In the framework of numeration systems, syntactic complexity has an advantage in comparison to left or right quotients, we have the opportunity to work simultaneously on prefixes and suffixes of base  $b$  expansions, that is on most and least significant digits.

Motivation for this work comes from the following decision problem. Let  $S$  be an abstract numeration system built on a regular language. See [4, Chap. 3] for background. It is well-known that any ultimately periodic set is  $S$ -recognizable, i.e., it has a regular language of representations within the system  $S$ . An instance of the decision problem is given by an abstract numeration system  $S$  and a DFA accepting some  $S$ -recognizable set  $X \subseteq \mathbb{N}$ . The question is therefore to decide whether  $X$  is ultimately periodic or not. This problem was settled positively for integer base systems by Honkala in [14]. See also [2] and in particular [5] for a first order logic approach. Recently this decision problem was settled positively in [3] for a large class of numeration systems based on linear recurrence sequences. Considering this decision problem for any abstract numeration system turns out to be equivalent to the so-called  $\omega$ -HD0L ultimate periodicity decision problem, see again [4, 15]. For instance, the pure morphic case, i.e., the  $\omega$ -D0L ultimate periodicity decision problem was solved in [13]. Relying on combinatorics on words techniques, the general case has very recently received a positive answer [9, 10, 16].

Since syntactic complexity provides an alternative measure for the complexity of a regular language, one could try to develop new decision procedures based on the syntactic complexity instead of the state complexity of the corresponding languages. A step in that direction is to first consider integer base numeration systems. As a consequence of our results, we present such a procedure in Section 5, using a similar approach as in [14].

In the next section, we recall basic definitions, fix notation and discuss the case of  $n$ -definite languages. Let  $b \geq 2$ . Any integer  $m$  can be uniquely written as  $m = db^n q$  with  $(q, b) = 1$ , such that  $n \geq 0$

and  $q \geq 1$  are chosen maximal. In Section 3, when  $m = db^nq$  with  $d = 1$ , we present exact formulae for the syntactic complexity of  $m\mathbb{N}$  and a lower bound on the number of infinite classes of the syntactic monoid of a periodic set of period  $m$ . In Section 4, we provide such a lower bound in the general case, i.e.,  $d \geq 1$ . We end the paper with a procedure for the decision problem described above.

## 2. Basics

For  $i \leq j$ , we denote by  $\llbracket i, j \rrbracket$  the interval of integers  $\{i, i+1, \dots, j-1, j\}$ . A *deterministic finite automaton* (or DFA) over the alphabet  $A$  is a 5-tuple  $\mathcal{A} = (Q, q_0, F, A, \delta)$  where  $Q$  is the set of states,  $q_0$  is the initial state,  $F$  is the set of final states and  $\delta : Q \times A^* \rightarrow Q$  is the (extended) transition function. We denote by  $|u|$  the length of the word  $u \in A^*$  and by  $\#P$  the cardinality of  $P$ .

### 2.1. Integer base numeration systems

Let  $b \geq 2$  be an integer. We denote by  $A_b$  the canonical alphabet of digits  $\llbracket 0, b-1 \rrbracket$ . For any word  $u = u_\ell \cdots u_0 \in A_b^*$ , we define the *numerical value* of  $u$  as

$$\text{val}_b(u) = \sum_{i=0}^{\ell} u_i b^i.$$

Note that  $\text{val}_b(uv) = \text{val}_b(u) b^{|v|} + \text{val}_b(v)$  for all  $u, v \in A_b^*$ . For any integer  $n > 0$ , we denote the usual base  $b$  expansion of  $n$  by  $\text{rep}_b(n)$ . We assume that such a greedy expansion does not start with 0. By convention,  $\text{rep}_b(0)$  is the empty word  $\varepsilon$ . A set  $X$  of integers is said to be *b-recognizable* if the language  $\text{rep}_b(X) \subseteq A_b^*$  is a regular language accepted by some DFA.

A set  $X \subseteq \mathbb{N}$  is *periodic* of *period*  $p$  if for all  $n \in \mathbb{N}$ ,  $n \in X \Leftrightarrow n+p \in X$ . The period is always understood to be the minimal period of  $X$ . In particular, if  $X \subseteq \mathbb{N}$  is periodic of period  $p$ , then for all  $i, j \in \mathbb{N}$ ,

$$i \not\equiv j \pmod{p} \Rightarrow \exists r \in \llbracket 0, p-1 \rrbracket : (i+r \in X, j+r \notin X) \text{ or } (i+r \notin X, j+r \in X). \quad (1)$$

A set  $X \subseteq \mathbb{N}$  is *ultimately periodic* of *period*  $p$  and *index*  $I > 0$  if, for all  $n \geq I$ ,  $n \in X \Leftrightarrow n+p \in X$  and exactly one of the two elements  $I-1, I+p-1$  is in  $X$ . Again, index and period are always understood to be minimal. It is easy to see that any ultimately periodic set is *b-recognizable* for all bases  $b \geq 2$ .

### 2.2. Syntactic complexity and $n$ -definite languages

Let  $L$  be a language over the finite alphabet  $A$ . The *context* of a word  $u \in A^*$  with respect to  $L$  is given by the set of pairs

$$\mathbf{C}_L(u) = \{(x, y) \in A^* \times A^* \mid xuy \in L\}.$$

If  $L$  is clearly understood, we will simply write  $\mathbf{C}(u)$ . Define the *Myhill congruence* [17] of  $L$  by  $u \leftrightarrow_L v$  if and only if, for all  $x, y \in A^*$ ,  $xuy \in L \Leftrightarrow xvy \in L$ . In other words,  $u \leftrightarrow_L v$  if and only if  $\mathbf{C}_L(u) = \mathbf{C}_L(v)$ . This congruence is also known as the *syntactic congruence* of  $L$ . The monoid  $A^*/\leftrightarrow_L$  made of the equivalence classes of the relation  $\leftrightarrow_L$ , is the *syntactic monoid* of  $L$ . It is well-known that

$L$  is a regular language if and only if  $A^*/\leftrightarrow_L$  is finite. The *syntactic complexity* of  $L$  is the cardinality of its syntactic monoid. If  $X \subseteq \mathbb{N}$  is a  $b$ -recognizable set of integers, by extension we define the *syntactic complexity* of  $X$  (w.r.t.  $b$ ) as the syntactic complexity of the language  $0^* \text{rep}_b(X)$ .

**Proposition 2.1.** Let  $L$  be a language over  $A$ . Two words  $u, v \in A^*$  are such that  $u \leftrightarrow_L v$  if and only if they perform the same transformation on the set of states of the minimal automaton  $\mathcal{M}_L = (Q_L, q_{0,L}, F_L, A, \delta_L)$  of  $L$ , i.e., for all  $r \in Q_L$ ,  $\delta_L(r, u) = \delta_L(r, v)$ . In particular, if  $u, v$  are such that  $\delta_L(q_{0,L}, u) \neq \delta_L(q_{0,L}, v)$ , then  $u \not\leftrightarrow_L v$ .

**Definition 2.1.** Let  $Q_L = \{q_1, \dots, q_n\}$  be the set of states of the minimal automaton of  $L$ . The *transformation automaton* of  $L$  has  $(Q_L)^n$  as set of states,  $(q_1, \dots, q_n)$  as initial state and for all  $w \in A^*$ ,  $(r_1, \dots, r_n) \in (Q_L)^n$ , its transition function  $\tau : (Q_L)^n \times A^* \rightarrow (Q_L)^n$  is given by  $\tau((r_1, \dots, r_n), w) = (\delta_L(r_1, w), \dots, \delta_L(r_n, w))$ . We will only consider the accessible part of this automaton. Thanks to the above proposition,  $u \leftrightarrow_L v$  if and only if  $\tau((q_1, \dots, q_n), u) = \tau((q_1, \dots, q_n), v)$ .

**Definition 2.2.** A language  $L \subseteq A^*$  is *weakly  $n$ -definite*, if for any  $x, y \in A^*$  satisfying  $|x| \geq n$ ,  $|y| \geq n$  and having the same suffix of length  $n$ ,  $x \in L$  if and only if  $y \in L$  [19, 6]. In other words,  $L$  can be written as  $G \cup A^*F$  where  $F$  (resp.  $G$ ) is finite and contains only words of length  $n$  (resp. less than  $n$ ). Let  $n \geq 1$ . A language is  *$n$ -definite* if it is weakly  $n$ -definite and not weakly  $(n-1)$ -definite. One also finds the terminology *suffix testable* in the literature, see [20].

It is well-known that if a language  $L$  is  $n$ -definite, then its minimal automaton has at least  $n+1$  states [19]. In particular, its syntactic monoid also has at least  $n+1$  elements. We adapt this result to get extra information about such a monoid in the case of sets of integers.

**Lemma 2.1.** Let  $X$  be a set of integers. If  $L = 0^* \text{rep}_b(X)$  is  $n$ -definite, then there exist arbitrarily large integers  $t_1, \dots, t_{n+1}$  such that the  $n+1$  words  $\text{rep}_b(t_1), \dots, \text{rep}_b(t_{n+1})$  belong to different equivalence classes of  $\leftrightarrow_L$ .

**Proof:**

Note that there exist  $k$  words  $s_1, \dots, s_k$  of length  $n$  such that a word of length at least  $n$  belongs to  $L$  if and only if it belongs to  $A_b^* \{s_1, \dots, s_k\}$ . Define an accessible DFA  $\mathcal{A} = (Q, q_\varepsilon, F, A_b, \delta)$  where

$$Q_n = \{q_w \mid |w| = n\} \text{ and } Q = Q_n \cup \{q_w \mid |w| < n\}$$

and for all  $u \in A_b^*$  such that  $|u| < n$  and  $a \in A_b$ , we have  $\delta(q_u, a) = q_{ua}$ . Now if  $|u| = n$ , then  $u = cx$  for some  $c \in A_b$ ,  $|x| = n-1$  and we have  $\delta(q_u, a) = q_{xa}$ . Notice that  $\mathcal{A}$  restricted to the states in  $Q_n$  is a strongly connected component isomorphic to the de Bruijn graph of order  $n$  over  $A_b$ . The set of final states of  $\mathcal{A}$  is easily defined in such a way that the language accepted by  $\mathcal{A}$  is  $L$ . In particular, a state in  $Q_n$  is final if and only if it is of the form  $q_{s_i}$  for  $i \in \llbracket 1, k \rrbracket$ .

Consider the minimal automaton of  $L$  denoted by  $\mathcal{M}_L = (Q_L, q_{0,L}, F_L, A_b, \delta_L)$  and the canonical morphism [11] of automata  $\Phi : Q \rightarrow Q_L$  from  $\mathcal{A}$  to  $\mathcal{M}_L$  such that  $\Phi(\delta(r, w)) = \delta_L(\Phi(r), w)$  for all  $r \in Q$  and  $w \in A_b^*$ . Let  $R := \Phi(Q_n)$ . In other words,  $R$  is the set of states of  $\mathcal{M}_L$  reached by words of length at least  $n$ . Using the same arguments as in [19], let us show that  $\#R \geq n+1$ . For all  $r, r' \in R$  and  $i \geq 0$ , define

$$E_i(r, r') \Leftrightarrow (\forall x \in A_b^*) [ |x| \geq i \Rightarrow (\delta_L(r, x) \in F_L \Leftrightarrow \delta_L(r', x) \in F_L) ].$$

This equivalence relation  $E_i$  over  $R$  induces a partition  $P_i$  of  $R$  into  $\#P_i$  equivalence classes. It is clear that  $E_i(r, r')$  implies  $E_{i+1}(r, r')$  and thus  $\#P_i \geq \#P_{i+1}$ . Since  $L$  is  $n$ -definite, there exist a word  $V = v_1 \cdots v_{n-1}$  of length  $n-1$  and two symbols  $\sigma, \tau \in A_b$  such that  $\sigma V \in L$  and  $\tau V \notin L$ . Let  $T \geq n$  and  $i \in \llbracket 0, n-1 \rrbracket$ . Take the two states  $r = \Phi(\delta(q_\varepsilon, 10^T \sigma v_1 \cdots v_{n-i-1}))$  and  $r' = \Phi(\delta(q_\varepsilon, 10^T \tau v_1 \cdots v_{n-i-1}))$  in  $R$ . By considering the word  $v_{n-i} \cdots v_{n-1}$  of length  $i$ , the states  $r$  and  $r'$  do not satisfy  $E_i(r, r')$  but for all words  $u$  of length at least  $i+1$ , we have

$$\delta(q_\varepsilon, 10^T \sigma v_1 \cdots v_{n-i-1} u) = q_S = \delta(q_\varepsilon, 10^T \tau v_1 \cdots v_{n-i-1} u)$$

where  $S$  is the suffix of length  $n$  of  $v_1 \cdots v_{n-i-1} u$  and thus  $E_{i+1}(r, r')$ . We have just shown that  $E_i$  is a refinement of  $E_{i+1}$  and  $\#P_0 > \#P_1 > \cdots > \#P_{n-1} > \#P_n \geq 1$ . Consequently,  $\#R \geq \#P_0 \geq n+1$ .

The minimal automaton  $\mathcal{M}_L$  of  $L$  contains at least  $n+1$  distinct states of the kind  $\Phi(q_{u_1}), \dots, \Phi(q_{u_{n+1}}) \in R$  for some words  $u_1, \dots, u_{n+1} \in A_b^*$  of length  $n$ . Let  $I > 0$ . Take a large enough  $T$  such that, for all  $i \in \llbracket 1, n+1 \rrbracket$ ,  $\text{val}_b(10^T u_i) > I$  and observe that

$$\Phi(\delta(q_\varepsilon, 10^T u_i)) = \Phi(q_{u_i}) = \delta_L(q_{0,L}, 10^T u_i) \in R.$$

The words  $10^T u_i$ ,  $i = 1, \dots, n+1$ , perform pairwise distinct transformations on the set of states of the minimal automaton  $\mathcal{M}_L$  and the syntactic monoid of  $L$  contains at least  $n+1$  classes (see Proposition 2.1).  $\square$

### 3. First results on the syntactic complexity

Let  $m, x \geq 2$  be integers such that  $(m, x) = 1$ . We denote by  $\text{ord}_m(x)$  the order of  $x$  in the multiplicative group  $U(\mathbb{Z}/m\mathbb{Z})$  made of the invertible elements in  $\mathbb{Z}/m\mathbb{Z}$ . That is  $\text{ord}_m(x)$  is the smallest positive integer  $j$  such that  $x^j \equiv 1 \pmod{m}$ . In particular,  $\text{ord}_m(x)$  is the period of the sequence  $(x^n \pmod{m})_{n \geq 0}$ .

We first consider the case where the base and the period are coprime. Interestingly, the syntactic complexity depends only on the period and not on the structure of the periodic set.

**Theorem 3.1.** Let  $m, b \geq 2$  be integers such that  $(m, b) = 1$ . If  $X \subseteq \mathbb{N}$  is periodic of (minimal) period  $m$ , then the syntactic complexity of  $0^* \text{rep}_b(X)$  is given by  $m \cdot \text{ord}_m(b)$ . In particular, this result holds for  $X = m\mathbb{N}$ . Moreover, for each class of  $\leftrightarrow_{0^* \text{rep}_b(X)}$ , there exists an arbitrarily large integer  $k$  such that  $\text{rep}_b(k)$  belongs to this class.

**Proof:**

Let  $X \subseteq \mathbb{N}$  be a periodic set of period  $m$ . Let  $u, v \in A_b^*$ . Let us first show that we have

$$u \leftrightarrow_{0^* \text{rep}_b(X)} v \Leftrightarrow \mathcal{C}(u) = \mathcal{C}(v) \Leftrightarrow \begin{cases} \text{val}_b(u) \equiv \text{val}_b(v) \pmod{m}, \\ |u| \equiv |v| \pmod{\text{ord}_m(b)}. \end{cases} \quad (2)$$

Let  $\alpha$  be a multiple of  $\text{ord}_m(b)$  such that  $b^\alpha > m$ . Since  $(b^i \pmod{m})_{i \geq 0}$  is a purely periodic sequence of period  $\text{ord}_m(b)$ , it follows that  $\text{val}_b(u0^\alpha) \equiv \text{val}_b(u) \pmod{m}$ . Assume that  $\text{val}_b(u) \not\equiv \text{val}_b(v) \pmod{m}$ . Using (1) there exists  $r \in \llbracket 0, m-1 \rrbracket$  such that  $\text{val}_b(u) + r \in X$  and  $\text{val}_b(v) + r \notin X$  (the other case is treated similarly). So  $(\varepsilon, 0^{\alpha-|\text{rep}_b(r)|} \text{rep}_b(r))$  belongs to  $\mathcal{C}(u)$  and not to  $\mathcal{C}(v)$ . Now assume that  $\text{val}_b(u) \equiv \text{val}_b(v) \pmod{m}$  and  $|u| \not\equiv |v| \pmod{\text{ord}_m(b)}$ . In that case, we obtain that  $\text{val}_b(1u) \not\equiv$

$\text{val}_b(1v) \bmod m$  and we can proceed as in the first situation, there exists some  $r \in \llbracket 0, m-1 \rrbracket$  such that  $(1, 0^{\alpha-|\text{rep}_b(r)|} \text{rep}_b(r))$  belongs to  $\mathcal{C}(u)$  and not to  $\mathcal{C}(v)$ .

Now proceed to the converse and assume that  $u, v$  are such that  $\text{val}_b(u) \equiv \text{val}_b(v) \bmod m$  and  $|u| \equiv |v| \bmod \text{ord}_m(b)$ . For all  $x, y \in A_b^*$ , we have

$$\begin{aligned} \text{val}_b(xuy) &= \text{val}_b(x) b^{|u|+|y|} + \text{val}_b(u) b^{|y|} + \text{val}_b(y) \\ &\equiv \text{val}_b(x) b^{|v|+|y|} + \text{val}_b(v) b^{|y|} + \text{val}_b(y) \equiv \text{val}_b(xvy) \bmod m \end{aligned}$$

and we have again used the fact that the sequence  $(b^i \bmod m)_{i \geq 0}$  is purely periodic of period  $\text{ord}_m(b)$ .

To conclude the proof, by considering words of the kind  $0^{\alpha-|\text{rep}_b(r)|+j} \text{rep}_b(r)$ , for  $r \in \llbracket 0, m-1 \rrbracket$  and  $j \in \llbracket 1, \text{ord}_m(b) \rrbracket$ , it follows from (2) that we have  $m \cdot \text{ord}_m(b)$  non-empty classes of  $\leftrightarrow_{0^* \text{rep}_b(X)}$ . Moreover each class contains representations of arbitrarily large integers. For all  $T \geq 0$  and for all  $u \in A_b^*$ , using (2) we have

$$u \leftrightarrow_{0^* \text{rep}_b(X)} (1 0^{\text{ord}_m(b)-1})^{Tm} u.$$

□

**Remark 3.1.** Note that the characterization of equation (2) implies that two periodic sets of minimal period  $m$  coprime with  $b$  have isomorphic syntactic monoids.

**Example 3.1.** Consider the set  $X = 3\mathbb{N}$  in base 2. The minimal automaton of  $0^* \text{rep}_2(X)$  is depicted in Figure 1. Table 2 gives the corresponding multiplication table of the syntactic monoid of  $0^* \text{rep}_2(X)$  where each class is given by one of its representatives.

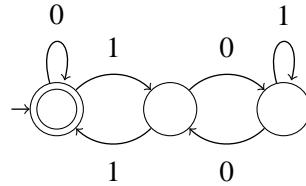


Figure 1. The minimal automaton of  $0^* \text{rep}_2(3\mathbb{N})$ .

Now consider the case where the period is a power of the base.

**Theorem 3.2.** Let  $b \geq 2$  and  $m = b^n$  with  $n \geq 1$ . Then the syntactic complexity of  $0^* \text{rep}_b(m\mathbb{N})$  is given by  $2n+1$ .

**Proof:**

The words  $\varepsilon, 0, \dots, 0^n, 1, 10, \dots, 10^{n-1}$  have pairwise different contexts w.r.t. the language  $0^* \text{rep}_b(m\mathbb{N})$ . For  $i = 0, \dots, n$ ,  $(10^{n-i}, \varepsilon)$  belongs to  $\mathcal{C}(0^{i+\ell})$ , for all  $\ell \geq 0$ , but does not belong to  $\mathcal{C}(0^j)$  for  $j < i$  nor  $\mathcal{C}(10^k)$ , for  $0 \leq k \leq n-1$ . In the same way, for  $i = 1, \dots, n$ ,  $(\varepsilon, 0^i)$  belongs to  $\mathcal{C}(10^{n-j})$ , for  $0 \leq j \leq i$ , but not to  $\mathcal{C}(10^j)$  for  $j < n-i$ . So the syntactic monoid of  $0^* \text{rep}_b(m\mathbb{N})$  has at least  $2n+1$  elements. Now consider some word  $u \in A_b^+$ . Write  $u = v0^i$  where  $v$  is either empty or ends with a non-zero digit. If  $i \geq n$ , then  $u \leftrightarrow_{0^* \text{rep}_b(m\mathbb{N})} 0^n$ . If  $v \neq \varepsilon$  and  $i < n$ , then  $u \leftrightarrow_{0^* \text{rep}_b(m\mathbb{N})} 10^i$ . If  $v = \varepsilon$  and  $i < n$ , the case  $u = 0^i$  was already considered. □

	$\varepsilon$	0	1	01	10	101
$\varepsilon$	$\varepsilon$	0	1	01	10	101
0	0	$\varepsilon$	01	1	101	10
1	1	10	$\varepsilon$	101	0	01
01	01	101	0	10	$\varepsilon$	1
10	10	1	101	$\varepsilon$	01	0
101	101	01	10	0	1	$\varepsilon$

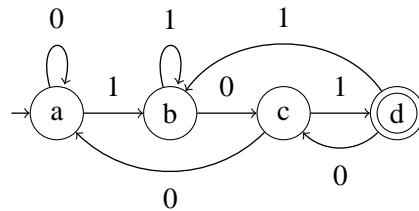
Table 2. The multiplication table of the syntactic monoid of  $0^* \text{rep}_2(3\mathbb{N})$ .

**Proposition 3.1.** Let  $b \geq 2$ . If  $X \subseteq \mathbb{N}$  is a periodic set of (minimal) period  $m = b^n$  with  $n \geq 1$ , then the syntactic complexity of  $L = 0^* \text{rep}_b(X)$  is greater than or equal to  $n + 1$ . Moreover there exist arbitrarily large integers  $t_1, \dots, t_{n+1}$  such that the  $n + 1$  words  $\text{rep}_b(t_1), \dots, \text{rep}_b(t_{n+1})$  belong to different equivalence classes of  $\leftrightarrow_L$ .

**Proof:**

Let  $X \subseteq \mathbb{N}$  be a periodic set of period  $b^n$ ,  $n \geq 1$ . By minimality of the period there exist  $V \in A_b^*$ ,  $\sigma, \tau \in A_b$  such that  $\sigma \neq \tau$ ,  $|V| = n - 1$ , and for all  $u \in A_b^*$ , we have  $\text{val}_b(u\sigma V) \in X$  and, for all  $u \in A_b^*$ ,  $\text{val}_b(u\tau V) \notin X$ . (If that was not the case, the fact that a word  $w$  belongs to  $L$  would only depend on its suffix of length  $n - 1$ , so in particular, we would have  $\text{val}_b(w) \in X$  if and only if  $\text{val}_b(w) + b^{n-1} \in X$  for all words  $w$ . This contradicts the fact that  $b^n$  is the period of  $X$ .) In other words,  $L$  is a  $n$ -definite language. One can conclude using Lemma 2.1.  $\square$

**Remark 3.2.** The bound in Proposition 3.1 is tight. One can for instance consider the set  $5 + 8\mathbb{N}$  written in base 2. The minimal automaton of  $0^* \text{rep}_2(5 + 8\mathbb{N})$  is depicted in Figure 2. The corresponding

Figure 2. The minimal automaton of  $0^* \text{rep}_2(5 + 8\mathbb{N})$ .

transformation automaton (see Definition 2.1) is given in Figure 3 displaying one strongly connected component with four elements. The syntactic monoid has exactly four infinite equivalence classes.

For the sake of completeness, we state the following result. See [18, Theorem 4].

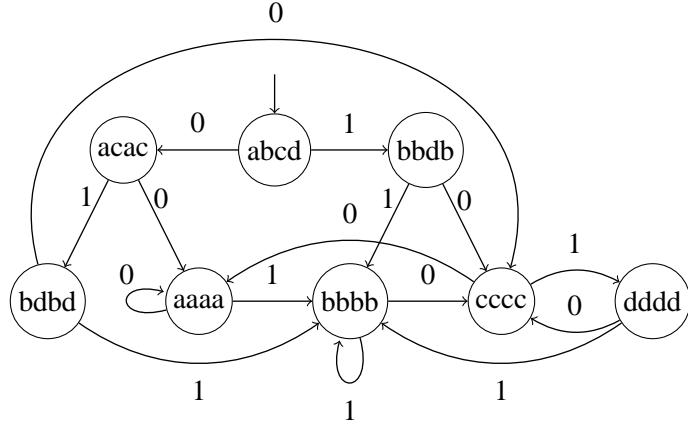


Figure 3. The transformation automaton of  $0^* \text{rep}_2(5 + 8\mathbb{N})$ .

**Theorem 3.3.** Let  $b \geq 2$  and  $m = b^n q$  where  $n \geq 1$  and  $(q, b) = 1$  and  $q \geq 2$ . Then the syntactic complexity of  $0^* \text{rep}_b(m\mathbb{N})$  is given by  $(n + 1) \cdot q \cdot \text{ord}_q(b) + n$ .

#### 4. Lower bound in the general case

Let  $b \geq 2$ . Notice that any integer  $m$  can be uniquely written as

$$m = db^n q \quad \text{with } (q, b) = 1 \quad (3)$$

such that  $n \geq 0$  and  $q \geq 1$  are chosen maximal. Hence, if  $p_1^{\alpha_1} \cdots p_k^{\alpha_k}$  denotes the prime decomposition of  $b$  with  $\alpha_i > 0$ , for all  $i \in \llbracket 1, k \rrbracket$ , then  $d = p_1^{\delta_1} \cdots p_k^{\delta_k}$  with  $\delta_i \geq 0$ , for all  $i \in \llbracket 1, k \rrbracket$  and there exists  $r$  such that  $\delta_r < \alpha_r$ . For convenience, we set  $s := \text{ord}_q(b)$  in all what follows. Let

$$\eta := \min\{j \in \mathbb{N} \mid \forall i \in \llbracket 1, k \rrbracket, n\alpha_i + \delta_i \leq j\alpha_i\}.$$

In other words,  $\eta$  is the smallest  $j$  such that  $db^n | b^j$ . Note that  $\eta \geq n$ .

**Remark 4.1.** Since  $db^n | b^k$ , for all  $k \geq \eta$ , a word has its numerical value modulo  $db^n$  completely determined by its suffix of length  $\eta$ . In particular, if two words  $u, v \in A_b^*$  of length at least  $\eta$  have same suffix of length  $\eta$  and are such that  $\text{val}_b(u) \equiv \text{val}_b(v) \pmod{q}$ , then  $\text{val}_b(u) \equiv \text{val}_b(v) \pmod{m}$ .

**Lemma 4.1.** Let  $b \geq 2$  and  $m = db^n q$  be given as in (3). Let  $X \subseteq \mathbb{N}$  be a periodic set of (minimal) period  $m$ . For any words  $u, v \in A_b^*$  of length at least  $\eta$ , we have

$$(|u| \equiv |v| \pmod{s}) \wedge (\text{val}_b(u) \equiv \text{val}_b(v) \pmod{m}) \Rightarrow u \leftrightarrow_{0^* \text{rep}_b(X)} v. \quad (4)$$

**Proof:**

Let  $u, v \in A_b^*$  be two words of length at least  $\eta$ . Using the fact that  $b^k \equiv b^{k+s} \pmod{q}$ , for all  $k \geq 0$  and  $b^k \equiv b^{k+s} \pmod{db^n}$ , for all  $k \geq \eta$ , we have that  $b^k \equiv b^{k+s} \pmod{m}$ , for all  $k \geq \eta$ . Notice that if  $u, v$  are such that  $|u| \equiv |v| \pmod{s}$  and  $\text{val}_b(u) \equiv \text{val}_b(v) \pmod{m}$ , then, for all  $x, y \in A_b^*$ ,  $\text{val}_b(xuy) \equiv \text{val}_b(xvy) \pmod{m}$  which means that  $u \leftrightarrow_{0^* \text{rep}_b(X)} v$ .  $\square$

**Lemma 4.2.** Let  $b \geq 2$  and  $m = db^nq$  given as in (3). Let  $X \subseteq \mathbb{N}$  be a periodic set of (minimal) period  $m$ . For any words  $u, v \in A_b^*$ , we have

$$(|u| \not\equiv |v| \pmod{s}) \vee (\text{val}_b(u) \not\equiv \text{val}_b(v) \pmod{q}) \Rightarrow u \not\leftrightarrow_{0^* \text{rep}_b(X)} v. \quad (5)$$

The reader may notice that the main difference between (4) and (5) is that congruences of numerical values are considered modulo  $m$  and  $q$  respectively.

**Proof:**

As a first case, suppose that  $\text{val}_b(u) \not\equiv \text{val}_b(v) \pmod{q}$ . Take  $\alpha \in \mathbb{N}$  such that  $b^\alpha > m$ . Since  $(b, q) = 1$ , we get that  $\text{val}_b(u)b^\alpha \not\equiv \text{val}_b(v)b^\alpha \pmod{q}$ . Hence  $\text{val}_b(u)b^\alpha \not\equiv \text{val}_b(v)b^\alpha \pmod{m}$  and, using (1), there exists  $r \in \llbracket 0, m-1 \rrbracket$  such that  $\text{val}_b(u)b^\alpha + r \in X$  and  $\text{val}_b(v)b^\alpha + r \notin X$  (the other case is treated similarly). We can conclude that  $(\varepsilon, 0^{\alpha-|\text{rep}_b(r)|} \text{rep}_b(r))$  belongs to  $\mathcal{C}(u)$  and not to  $\mathcal{C}(v)$ . As a second case, suppose that  $\text{val}_b(u) \equiv \text{val}_b(v) \pmod{q}$  but that  $|u| \not\equiv |v| \pmod{s}$ . This implies that  $b^{|u|} \not\equiv b^{|v|} \pmod{q}$ . Therefore  $\text{val}_b(1u) \not\equiv \text{val}_b(1v) \pmod{q}$  and we proceed as in the first case. There exists  $r \in \llbracket 0, m-1 \rrbracket$  such that  $(1, 0^{\alpha-|\text{rep}_b(r)|} \text{rep}_b(r))$  belongs to  $\mathcal{C}(u)$  and not to  $\mathcal{C}(v)$ .  $\square$

**Lemma 4.3.** For all words  $w \in A_b^*$ ,  $i \in \llbracket 0, q-1 \rrbracket$ ,  $\ell \in \llbracket 0, s-1 \rrbracket$  and  $I > 0$ , there exists a word  $u$  having  $w$  as suffix and such that  $\text{val}_b(u) \equiv i \pmod{q}$ ,  $|u| \equiv \ell \pmod{s}$  and  $\text{val}_b(u) > I$ .

**Proof:**

Let  $L \geq |w|$  be a multiple of  $s$ . We set  $z = \text{val}_b(w) \pmod{q}$ . Using the periodicity of  $(b^k \pmod{q})_{k \geq 0}$ , we have  $b^{L+ks} \equiv 1 \pmod{q}$ , for all  $k \geq 0$ , and

$$\text{val}_b[(0^{s-1}1)^{i+q-z} 0^{L-|w|} w] \equiv \sum_{k=0}^{i+q-z-1} b^{L+ks} + z \equiv i \pmod{q}.$$

To conclude the proof, one has to add a prefix of the kind  $0^\ell(0^{s-1}1)^{tq}$ .  $\square$

**Definition 4.1.** Take a periodic set  $X \subseteq \mathbb{N}$  of period  $m = db^nq$  given as in (3). Consider the characteristic word  $(x_t)_{t \geq 0} \in \{0, 1\}^\omega$  of  $X$  where  $x_t = 1$  if and only if  $t \in X$ . This infinite word is periodic of period  $m$ . The  $q$  infinite words  $(x_{qt})_{t \geq 0}$ ,  $(x_{qt+1})_{t \geq 0}$ ,  $\dots$ ,  $(x_{qt+q-1})_{t \geq 0}$  are periodic and each of their periods divides  $db^n$ . As  $m$  is the minimal period of  $X$  (for details, see [14]), there exist a non-empty set  $J \subseteq \llbracket 0, q-1 \rrbracket$  such that, for all  $j \in J$ ,  $(x_{qt+j})_{t \geq 0}$  has exactly period  $P_j$  satisfying

$$P_j := p_1^{\beta_{j,1}} \cdots p_k^{\beta_{j,k}} \text{ where } \max_{r \in \llbracket 1, k \rrbracket} \beta_{j,r} = \max_{r \in \llbracket 1, k \rrbracket} (n\alpha_r + \delta_r). \quad (6)$$

We set

$$\gamma_j := \min\{g \in \mathbb{N} \mid \forall r \in \llbracket 1, k \rrbracket, g\alpha_r \geq \beta_{j,r}\}. \quad (7)$$

In other words,  $\gamma_j$  is the smallest  $g$  such that  $P_j \mid b^g$ . Now choose  $i \in J$  such that the corresponding  $\gamma_i$  is maximal (if several  $i$ 's fulfill this condition, to avoid ambiguity, we take the smallest such  $i$ ) and we will denote  $P$  and  $\gamma$  the corresponding  $P_i$  and  $\gamma_i$ . Clearly,  $\gamma$  is a function of  $d, n$  and  $q$ . In particular,  $b^{\gamma-1} \not\equiv 0 \pmod{P}$ .

**Remark 4.2.** Note that if  $n \rightarrow +\infty$  or  $d \rightarrow +\infty$ , then  $\gamma \rightarrow +\infty$ .

**Example 4.1.** Let  $b = 75$ , i.e.,  $p_1 = 3, p_2 = 5, \alpha_1 = 1$  and  $\alpha_2 = 2$ . Consider two infinite words over  $\{0, 1\}$ ,  $v$  having a 1 at multiples of 9,  $w$  at multiples of 25. We define  $x$  by interleaving these two words. This word  $x$  has period 2.9.25, i.e.,  $d = 3, n = 1$  and  $q = 2$ . Hence  $\delta_1 = 1, \delta_2 = 0$  and we get  $n\alpha_1 + \delta_1 = n\alpha_2 + \delta_2 = 2$ . Observe that  $(x_{2t})_{t \geq 0} = v$  (resp.  $(x_{2t+1})_{t \geq 0} = w$ ) has period  $P_0 = 3^2$  (resp.  $P_1 = 5^2$ ) which satisfies condition (6). We get  $J = \{0, 1\}, \gamma_0 = 2$  and  $\gamma_1 = 1$ . So we choose  $i = 0$  and set  $P = 3^2$  and  $\gamma = 2$ .

**Proposition 4.1.** Let  $b \geq 2$  and  $m = db^nq$  be given as in (3). If  $X \subseteq \mathbb{N}$  is a periodic set of (minimal) period  $m$ , then the syntactic complexity of  $0^* \text{rep}_b(X)$  is greater than or equal to  $M = \max\{q.s, \frac{\gamma+1}{q.s}\}$ . Moreover there exist arbitrarily large integers  $t_1, \dots, t_M$  such that  $\text{rep}_b(t_1), \dots, \text{rep}_b(t_M)$  belong to different equivalence classes of  $\leftrightarrow_{0^* \text{rep}_b(X)}$ .

**Proof:**

Thanks to Lemma 4.3, for all  $i \in \llbracket 0, q-1 \rrbracket, \ell \in \llbracket 0, s-1 \rrbracket$  and  $I > 0$ , we can construct a word  $u \in A_b^*$  such that  $\text{val}_b(u) \equiv i \pmod{q}, |u| \equiv \ell \pmod{s}$  and  $\text{val}_b(u) > I$ . Thus, by (5), any two such words are in different classes for  $\leftrightarrow_{0^* \text{rep}_b(X)}$ . Hence we have at least  $q.s$  different classes and each of them contains representations of arbitrary large elements.

Consider  $i \in \llbracket 0, q-1 \rrbracket, P$  and  $\gamma$  given in Definition 4.1 such that  $(x_{qt+i})_{t \geq 0}$  has exactly period  $P$ . Take  $T \geq 0$  such that  $|\text{rep}_b(qT+i)| \geq \gamma$ . Note that if  $j \not\equiv k \pmod{P}$ , then  $\text{rep}_b(q(T+j)+i)$  and  $\text{rep}_b(q(T+k)+i)$  have different suffixes of length  $\gamma$ . Indeed, assume that  $\text{rep}_b(q(T+j)+i)$  and  $\text{rep}_b(q(T+k)+i)$  have the same suffix of length  $\gamma$ , then  $q(k-j) \equiv 0 \pmod{b^\gamma}$ . Since  $(q, b) = 1$ , we get  $k \equiv j \pmod{b^\gamma}$ . As  $P$  divides  $b^\gamma$ , we have  $k \equiv j \pmod{P}$ .

Since  $(x_{qt+i})_{t \geq 0}$  has (minimal) period  $P$  and  $b^{\gamma-1} \not\equiv 0 \pmod{P}$ , then  $T \not\equiv T + b^{\gamma-1} \pmod{P}$  and there exists  $j \in \llbracket 0, P-1 \rrbracket$  such that  $g_1 := q(T+j)+i \in X$  and  $g_2 := q(T+j+b^{\gamma-1})+i \notin X$  (or equivalently  $g_1 \notin X$  and  $g_2 \in X$ ). Since  $g_2 = g_1 + qb^{\gamma-1}$ , then  $\text{rep}_b(g_1)$  and  $\text{rep}_b(g_2)$  have the same suffix  $V = v_1 \cdots v_{\gamma-1}$  of length  $\gamma-1$ . But since  $g_1 \not\equiv g_2 \pmod{P}$ ,  $\text{rep}_b(g_1)$  and  $\text{rep}_b(g_2)$  do not have the same suffix of length  $\gamma$ . We can assume that there exist some prefixes  $p, p' \in A_b^*$  and two distinct symbols  $\sigma, \tau \in A_b$  such that

$$\text{rep}_b(g_1) = p\sigma V \in 0^* \text{rep}_b(X) \text{ and } \text{rep}_b(g_2) = p'\tau V \notin 0^* \text{rep}_b(X).$$

We are now ready to show that the minimal automaton of  $0^* \text{rep}_b(X \cap (q\mathbb{N} + i))$  has at least  $\gamma + 1$  states reached by words of length at least  $\eta$  (this last requirement is always fulfilled by adding leading zeroes if necessary). Let  $L = 0^* \text{rep}_b(X)$  and  $L' = 0^* \text{rep}_b(q\mathbb{N} + i)$  having respectively  $\mathcal{M}_L$  and  $\mathcal{M}_{L'}$  as minimal automaton. We proceed as in the proof of Lemma 2.1. On the set of states of the minimal automaton  $\mathcal{M}_{L \cap L'}$  of  $L \cap L'$ , for all  $k \geq 0$ , we define the relation

$$E_k(r, r') \Leftrightarrow (\forall x \in A_b^*)[|x| \geq k \Rightarrow (\delta_{L \cap L'}(r, x) \in F \Leftrightarrow \delta_{L \cap L'}(r', x) \in F)]$$

where  $F$  is the set of final states of  $\mathcal{M}_{L \cap L'}$ . Let  $k \in \llbracket 0, \gamma - 1 \rrbracket$ . Consider the states  $r$  and  $r'$  reached when reading respectively  $p\sigma v_1 \cdots v_{\gamma-k-1}$  and  $p'\tau v_1 \cdots v_{\gamma-k-1}$ . They do not satisfy  $E_k(r, r')$  but we can observe that they satisfy  $E_{k+1}(r, r')$ . Indeed, take any  $x \in A_b^*$  of length  $k + 1$ . Since  $\text{val}_b(p\sigma V) \equiv \text{val}_b(p'\tau V) \pmod{q}$  then<sup>1</sup>, for all words  $W \in A_b^*$ ,  $\text{val}_b(p\sigma W) \equiv \text{val}_b(p'\tau W) \pmod{q}$ . So, in particular,

$$X_k := p\sigma v_1 \cdots v_{\gamma-k-1}x \text{ and } Y_k := p'\tau v_1 \cdots v_{\gamma-k-1}x$$

have the same value modulo  $q$  and the same suffix of length  $\gamma$ . This means<sup>2</sup> that either they both belong to  $L \cap L'$ , or they both do not belong to  $L \cap L'$ . Following the same lines as in the proof of Lemma 2.1 we conclude that the minimal automaton of  $L \cap L'$  has at least  $\gamma + 1$  states.

For any DFA  $\mathcal{M}$  having  $Q$  as set of states and  $\delta$  as transition function, we set

$$\mathcal{T}_\eta(\mathcal{M}) = \{f_w : Q \rightarrow Q, q \mapsto \delta(q, w) \mid w \in A_b^*, |w| \geq \eta\}$$

as the set of actions on  $Q$  derived by words of length at least  $\eta$ . Consider the product automaton  $\mathcal{P} = \mathcal{M}_L \times \mathcal{M}_{L'}$  recognizing  $L \cap L'$ . We clearly have  $\#\mathcal{T}_\eta(\mathcal{P}) \leq \#\mathcal{T}_\eta(\mathcal{M}_L) \cdot \#\mathcal{T}_\eta(\mathcal{M}_{L'})$ . Since there is a canonical morphism of automata from  $\mathcal{P}$  onto  $\mathcal{M}_{L \cap L'}$ , we get  $\#\mathcal{T}_\eta(\mathcal{M}_{L \cap L'}) \leq \#\mathcal{T}_\eta(\mathcal{P})$ .

Using Proposition 2.1 and Theorem 3.1, we get  $\#\mathcal{T}_\eta(\mathcal{M}_{L'}) = q.s$ . With the above discussion about the number of states of  $\mathcal{M}_{L \cap L'}$ , we have  $\#\mathcal{T}_\eta(\mathcal{M}_{L \cap L'}) \geq \gamma + 1$ . Putting all of this together,

$$\gamma + 1 \leq \#\mathcal{T}_\eta(\mathcal{M}_{L \cap L'}) \leq \#\mathcal{T}_\eta(\mathcal{P}) \leq \#\mathcal{T}_\eta(\mathcal{M}_L).q.s.$$

Therefore,  $\#\mathcal{T}_\eta(\mathcal{M}_L) \geq \frac{\gamma+1}{q.s}$ . Since  $\mathcal{T}_\eta(\mathcal{M}_L)$  is isomorphic to a subset of the syntactic monoid of  $L$ , this monoid has at least  $\frac{\gamma+1}{q.s}$  classes, each containing a word of length at least  $\eta$ . Since for all words  $w$  of length at least  $\eta$ , we can use Lemma 4.1 to get  $w \leftrightarrow_{0^* \text{rep}_b(X)} (0^{s-1}1)^m w$ , integers corresponding to elements of the different equivalence classes can be chosen arbitrarily large.  $\square$

## 5. Application to a decision procedure

Let  $X \subseteq \mathbb{N}$  be a  $b$ -recognizable set of integers such that  $0^* \text{rep}_b(X)$  is accepted by some DFA  $\mathcal{A}$ . A usual technique for deciding whether or not  $X$  is ultimately periodic is to prove that whenever  $X$  is ultimately periodic, then its period and its preperiod must be bounded by some quantities depending only on the size of the DFA  $\mathcal{A}$ . Therefore, one has a finite number of admissible periods and preperiods to test, leading to a decision procedure. For details, see [3]. In particular, the following result [3, Prop. 44] stated in full generality for any abstract numeration system (i.e., the language of numeration is a regular language) shows that we have only to obtain an upper bound on the admissible periods.

**Proposition 5.1.** Let  $S = (L, \Sigma, <)$  be an abstract numeration system. If  $X \subseteq \mathbb{N}$  is an ultimately periodic set of period  $p_X$  such that  $\text{rep}_S(X)$  is accepted by a DFA with  $d$  states, then the preperiod of  $X$  is bounded by an effectively computable constant  $C$  depending only on  $d$  and  $p_X$ .

<sup>1</sup>Subtracting  $\text{val}_b(V)$  on both sides, we get  $\text{val}_b(p\sigma 0^{|V|}) \equiv \text{val}_b(p'\tau 0^{|V|}) \pmod{q}$ . Since  $(q, b) = 1$ , we can multiply or divide a convenient number of times by  $b$  and get  $\text{val}_b(p\sigma 0^{|W|}) \equiv \text{val}_b(p'\tau 0^{|W|}) \pmod{q}$ . Hence, one adds  $\text{val}_b(W)$  to both sides.

<sup>2</sup>If  $\text{val}_b(X_k)$  and  $\text{val}_b(Y_k)$  are both congruent to a value not equal to  $i$  modulo  $q$ , then  $X_k$  and  $Y_k$  do not belong to  $L'$ . As a second case, we may assume that  $X_k, Y_k \in L'$ , i.e.,  $\text{val}_b(X_k)$  and  $\text{val}_b(Y_k)$  are respectively of the form  $q(T + j) + i$  and  $q(T + j') + i$  for some  $j, j'$ . As discussed in the second paragraph of this proof, since  $X_k$  and  $Y_k$  have the same suffix of length  $\gamma$ , then  $j \equiv j' \pmod{P}$  which means that  $q(T + j) + i \in X \Leftrightarrow q(T + j') + i \in X$ , or  $X_k \in L \Leftrightarrow Y_k \in L$ .

The following result is a consequence of Proposition 4.1.

**Theorem 5.1.** Let  $b \geq 2$ . If  $X \subseteq \mathbb{N}$  is an ultimately periodic set of period  $p_X = db^nq$  given as in (3), then the syntactic complexity of  $0^* \text{rep}_b(X)$  is greater than or equal to  $\max\{q.s, \frac{\gamma+1}{q.s}\}$  where  $\gamma$  is defined as in (7).

**Proof:**

Let  $I$  be the preperiod of  $X$ . Even if Proposition 4.1 is about (purely) periodic sets of integers, if we consider instead an ultimately periodic set, since we can choose words belonging to different equivalence classes in such a way that their numerical value is greater than  $I$ , then the lower bound on the number of classes is still valid for the ultimately periodic case.  $\square$

Assume that  $b \geq 2$ . Therefore, giving a DFA  $\mathcal{A}$  accepting  $0^* \text{rep}_b(X)$  and so the corresponding syntactic monoid, if  $X$  is ultimately periodic, then we get an upper bound on its period. Indeed, if  $m$  written  $db^nq$  as in (3) tends to infinity, then at least one of the three quantities  $d$ ,  $n$  or  $q$  tends to infinity and thanks to Remark 4.2, the corresponding function  $\max\{q.s, \frac{\gamma+1}{q.s}\}$  tends to infinity. As soon as a bound on the possible period is obtained, one can use Proposition 5.1 to get a bound on the possible preperiod. It suffices then to test a finite number of admissible pairs preperiod/period and compare those with the given set  $X$ .

## Acknowledgments

We thank one of the referees for improving our Definition 4.1 and giving the corresponding example.

## References

- [1] B. Alexeev, Minimal DFA for testing divisibility, *J. Comput. Syst. Sci.* **69** (2004), 235–243.
- [2] J.-P. Allouche, N. Rampersad, J. Shallit, Periodicity, repetitions, and orbits of an automatic sequence, *Theoret. Comput. Sci.* **410** (2009), 2795–2803.
- [3] J. P. Bell, E. Charlier, A. S. Fraenkel, M. Rigo, A decision problem for ultimately periodic sets in non-standard numeration systems, *Int. J. Algebra and Computation* **19** (2009), 809–839.
- [4] V. Berthé, M. Rigo (Eds), *Combinatorics, Automata and Number Theory*, Encyclopedia of Mathematics and its Applications **135**, Cambridge University Press (2010).
- [5] V. Bruyère, G. Hansel, C. Michaux, R. Villemaire, Logic and  $p$ -recognizable sets of integers, *Bull. Belg. Math. Soc.* **1** (1994), 191–238.
- [6] J. A. Brzozowski, Canonical Regular Expressions and Minimal State Graphs for Definite Events, Proceedings of the Symposium on Mathematical Theory of Automata New York, NY, April 24-26, 1962 J. Fox, ed., MRI Symposia Series, Vol. 12 Polytechnic Press of the Polytechnic Institute of Brooklyn, Brooklyn, NY pp. 529–561, (1963).
- [7] J. Brzozowski, Y. Ye, Syntactic Complexity of Ideal and Closed Languages, preprint (2010) [arXiv:1010.3263v1](https://arxiv.org/abs/1010.3263v1)
- [8] E. Charlier, N. Rampersad, M. Rigo, L. Waxweiler, The minimal automaton recognizing  $m\mathbb{N}$  in a linear numeration system, to appear in *INTEGERS*.

- [9] F. Durand, HD0L  $\omega$ -equivalence and periodicity problems in the primitive case (to the memory of G. Rauzy), to appear in *J. of Uniform Distribution Theory*.
- [10] F. Durand, Decidability of the HD0L ultimate periodicity problem, preprint (2011) [arXiv:1111.3268](https://arxiv.org/abs/1111.3268).
- [11] S. Eilenberg, *Automata, Languages and Machines*, Vol. A, Academic Press, New York, 1974.
- [12] V. Halava, T. Harju, T. Kärki, A new proof for the decidability of D0L ultimate periodicity, in Proceedings of WORDS 2011, *EPTCS* **63**, (2011), 147–151.
- [13] T. Harju, M. Linna, On the periodicity of morphisms on free monoids, *RAIRO Inform. Théor. Appl.* **20** (1986), 47–54.
- [14] J. Honkala, A decision method for the recognizability of sets defined by number systems, *Theor. Inform. Appl.* **20** (1986), 395–403.
- [15] J. Honkala, M. Rigo, A note on decidability questions related to abstract numeration systems, *Discrete Math.* **285** (2004), 329–333.
- [16] I. Mitrofanov, A proof for the decidability of HD0L ultimate periodicity, preprint (2011) [arXiv:1110.4780](https://arxiv.org/abs/1110.4780).
- [17] J. Myhill, Finite automata and representation of events, Wright Air Development Center Technical Report **57-624** (1957).
- [18] M. Rigo, E. Vandomme, Syntactic complexity of ultimately periodic sets of integers, Proceedings of the fifth conference LATA (Languages and Automata, Theory and Applications), *Lect. Notes in Comput. Sci.* **6638**, Springer Verlag (2011), 477–488, A.-H. Dediu, S. Inenaga, C. Martin-Vide (Eds.).
- [19] M. Perles, M. O. Rabin, E. Shamir, The theory of definite automata, *IEEE Trans. Electron. Comp.* (1963) 233–243.
- [20] J.-E. Pin, *Syntactic semigroups*, Chap. 10 in Handbook of language theory, Vol. I, G. Rozenberg and A. Salomaa (ed.), Springer Verlag, (1997), 679–746.