Journal of Integer Sequences, Vol. 13 (2010),
Article 10.2.2

On the Recognizability of Self-Generating Sets

Tomi Kärki[1]
University of Turku
Department of Mathematics
FI-20014 Turku
Finland
topeka@utu.fi

Anne Lacroix and Michel Rigo
University of Liège
Department of Mathematics
Grande Traverse 12 (B37)
B-4000 Liège
Belgium
A.Lacroix@ulg.ac.be
M.Rigo@ulg.ac.be

Abstract

Let I be a finite set of integers and F be a finite set of maps of the form $n \mapsto k_i\, n + \ell_i$ with integer coefficients. For an integer base $k \geq 2$, we study the k-recognizability of the minimal set X of integers containing I and satisfying $\varphi(X) \subseteq X$ for all $\varphi \in F$. We answer an open problem of Garth and Gouge by showing that X is k-recognizable when the multiplicative constants k_i are all powers of k and additive constants ℓ_i are chosen freely. Moreover, solving a conjecture of Allouche, Shallit and Skordev, we prove under some technical conditions that if two of the constants k_i are multiplicatively independent, then X is not k-recognizable for any $k \geq 2$.

[1]Part of this research was done when the author was working in the Department of Mathematics at the University of Liège. The financial support from the Osk. Huttunen Foundation is gratefully acknowledged.

1 Introduction

In the general framework of numeration systems, the so-called recognizable sets of integers have been extensively studied. Let $k \geq 2$ be an integer. The function $\mathrm{rep}_k \colon \mathbb{N} \to \{0, \ldots, k-1\}^*$ maps a non-negative integer onto its k-ary representation (without leading zeros). A set $X \subseteq \mathbb{N}$ is *k-recognizable* if the language $\mathrm{rep}_k(X) = \{\mathrm{rep}_k(n) \mid n \in X\}$ is regular; see, for instance, [3]. A similar definition can be given for the k-recognizable subsets of $\mathbb{Z}$ using convenient conventions to represent negative numbers, like adding a symbol "$-$" to the alphabet or considering the positive and the negative elements separately. Since the seminal work of Cobham [5], it is well-known that the recognizability of a set depends on the choice of the base k — except for the ultimately periodic sets, i.e., the union of a finite set and a finite number of infinite arithmetic progressions, which are easily seen to be k-recognizable for all $k \geq 2$. The celebrated theorem of Cobham can be stated as follows. Let $k, \ell \geq 2$ be two multiplicatively independent bases, i.e., $\log k / \log \ell$ is irrational. If a set $X \subseteq \mathbb{N}$ is both k-recognizable and ℓ-recognizable, then it is ultimately periodic.

Kimberling [13] introduced the so-called *self-generating* sets of integers. They can be defined as follows. Let $r \geq 1$ and $G = \{\varphi_1, \varphi_2, \ldots, \varphi_r\}$ be a set of affine maps where $\varphi_i \colon n \mapsto k_i\, n + \ell_i$ with $k_i, \ell_i \in \mathbb{Z}$ and $2 \leq k_1 \leq k_2 \leq \cdots \leq k_r$. The set generated by G and a finite set of integers I is the minimal subset X of $\mathbb{Z}$ containing I and such that $\varphi_i(X) \subseteq X$ for all $i = 1, \ldots, r$. For any subset $S \subseteq \mathbb{Z}$, we set $G(S) := \{\varphi(s) \mid s \in S, \varphi \in G\}$, $G^0(S) := S$ and $G^{m+1}(S) := G(G^m(S))$ for all $m \geq 0$. Otherwise stated, $X = \bigcup_{m \geq 0} G^m(I)$ is the set of all integers n such that there exist $m \geq 0$, $a \in I$ and a finite sequence $(\varphi_{i_1}, \varphi_{i_2}, \ldots, \varphi_{i_m})$ of maps in G such that

$$n = \varphi_{i_m} \circ \varphi_{i_{m-1}} \circ \cdots \circ \varphi_{i_1}(a) = \varphi_{i_m}(\varphi_{i_{m-1}}(\cdots \varphi_{i_1}(a) \cdots)). \tag{1}$$

Example 1. Kimberling [13] showed for $G = \{n \mapsto 2n, n \mapsto 4n - 1\}$ and $I = \{1\}$ that the corresponding self-generating set

$$\mathcal{K}_1 = \{1, 2, 3, 4, 6, 7, 8, 11, 12, 14, 15, 16, \ldots\}$$

is closely related to the Fibonacci word. This relationship will be developed in Section 4, where with our techniques we obtain again Kimberling's original result. Notice that for $I = \{0\}$, we get a subset containing negative integers: $\mathcal{K}_0 = \{0, -1, -2, -4, -5, -8, -9, \ldots\}$. In particular, for $I = \{0, 1\}$, the corresponding self-generating set is $\mathcal{K}_0 \cup \mathcal{K}_1$.

Self-generating sets are also called *affinely recursive* in [14] where the correspondence between words $i_1 i_2 \cdots i_m$ over the alphabet $\{1, 2, \ldots, r\}$ and integers $\varphi_{i_m}(\varphi_{i_{m-1}}(\cdots \varphi_{i_1}(1) \cdots))$ is studied. For example, conditions under which this correspondence is one-to-one are given, which in turn implies that the natural ordering of the integers induces an ordering on the set of non-empty words over $\{1, 2, \ldots, r\}$ providing a kind of abstract numeration system [15]. Note that in the definition of affinely recursive sets [14] the set of generating functions G can be an infinite set of maps of the form $\varphi_i \colon n \mapsto k_i\, n + \ell_i$, where $k_i, \ell_i \in \mathbb{N}$.

Allouche, Shallit and Skordev [2] consider a general framework for self-generating sets. The k-ary representations of the elements of some self-generating sets are related to words over $\Sigma_k = \{0, 1, \ldots, k-1\}$ where some fixed block of digits is missing. As an illustration, one can notice that the set $\mathcal{K}_1 - 1 = \{0, 1, 2, 3, 5, 6, 7, 10, \ldots\}$ introduced in Example 1 consists of all integers whose binary expansion does not contain "00" as factor. Recall that the *characteristic sequence* $(\mathbf{c}_X(n))_{n\geq 0}$ of a set $X \subseteq \mathbb{N}$ is defined by $\mathbf{c}_X(n) = 1$, if $n \in X$ and $\mathbf{c}_X(n) = 0$, otherwise. In particular, X is k-recognizable (resp., ultimately periodic) if and only if $(\mathbf{c}_X(n))_{n\geq 0}$ is k-automatic (resp., an ultimately periodic infinite word). Self-generating sets are consequently studied from the point of view of automatic and morphic sequences as well as in relation to non-standard numeration systems; for the definitions and further information, see [1, 16]. Moreover, Allouche, Shallit and Skordev ask the following question: *Under what conditions is the characteristic sequence of a self-generating set k-automatic?* They also present the following conjecture.

Conjecture 2. With "mixed base" rules, such as $G = \{n \mapsto 2n + 1, n \mapsto 3n\}$, the set generated from $I = \{1\}$ is not k-recognizable for any integer base $k \geq 2$.

Let us fix the notation once and for all.

Definition 3. In this paper, instead of considering a set G of maps as described above, we will moreover consider the extended set of $r + 1 \geq 2$ maps

$$F = G \cup \{\varphi_0\} = \{\varphi_0, \varphi_1, \ldots, \varphi_r\}$$

where $\varphi_0 \colon n \mapsto n$ and $\varphi_i \colon n \mapsto k_i n + \ell_i$ with $k_i, \ell_i \in \mathbb{Z}$ and

$$2 \leq k_1 \leq k_2 \leq \cdots \leq k_r.$$

Having the identity function φ_0 at our disposal, for any set $S \subseteq \mathbb{Z}$, we have $F^m(S) \subseteq F^{m+1}(S)$. Therefore, for any finite set I of integers, the set

$$F^\omega(I) := \lim_{m\to\infty} F^m(I)$$

is exactly the *self-generating set* with respect to G and I.

This article is an extended version of our presentation given in the MFCS conference 2009 [12]. The content of the paper is the following. In Section 2 we give some simple observations on self-generating sets. For example, if we add to F an extra map $\psi \colon n \mapsto n + \ell$ with $\ell \neq 0$, then the corresponding self-generating set $F^\omega(I)$ is ultimately periodic and therefore k-recognizable for all $k \geq 2$. We also show that we can restrict our considerations to subsets of $\mathbb{N}$ and assume that all additive constants ℓ_i for the maps $\varphi_i \in F$ are non-negative.

In sections 3 and 4 we consider the multiplicatively dependent case. The results are based on Frougny's normalization transducer; see, e.g., Chapter 7 in [16]. If all multiplicative constants k_i are pairwise multiplicatively dependent, then we give a general method to build a finite automaton recognizing $\mathrm{rep}_k(F^\omega(I))$ for any k that is multiplicatively dependent on every k_i. This allows us to generalize a recognizability result of Garth and Gouge [9]. Moreover, a new proof of the relation between the Kimberling set $\mathcal{K}_1$ and the infinite Fibonacci word is given in Section 4; for other proofs, see [2, 13].

In the multiplicatively independent case of Section 5 we study differences and ratios of consecutive elements in the considered self-generating set. The results rely on a classical gap theorem; see Theorem 14. We prove that if there exist i, j such that k_i and k_j are multiplicatively independent and if $\sum_{i=1}^{r} k_i^{-1} < 1$, then $F^\omega(I)$ is not k-recognizable for any $k \geq 2$. In particular, this condition always holds for sets F where $r = 2$ and $k_1 < k_2$ are multiplicatively independent, answering Conjecture 2 in the affirmative.

2 Some reductions

First we show that assuming $k_i \geq 2$ for every $i = 1, 2, \ldots, r$ is not a real restriction from the point of view of recognizability.

Lemma 4. *If we add to F in Definition 3 an extra map $\psi\colon n \mapsto n + \ell$ with $\ell \neq 0$, then the corresponding self-generating set $F^\omega(I)$ is ultimately periodic of period ℓ.*

Proof. Denote by $F^j(I) \bmod \ell$ the set $\{n \bmod \ell \mid n \in F^j(I)\}$. Recall that the identity function φ_0 belongs to F. Since there are finitely many congruence classes modulo ℓ and $F^j(I) \bmod \ell \subseteq F^{j+1}(I) \bmod \ell$, there must exist an integer J such that $F^{J+1}(I) \bmod \ell = F^J(I) \bmod \ell$. Moreover, this means that $F^j(I) \bmod \ell = F^J(I) \bmod \ell$ for every $j \geq J$, and, consequently,

$$F^\omega(I) \bmod \ell = F^J(I) \bmod \ell. \tag{2}$$

On the other hand, if $n \in F^\omega(I)$, then $\psi^t(n) = n + t\,\ell \in F^\omega(I)$. Since $n + t\,\ell \equiv n \bmod \ell$, we conclude by (2), for any $n \geq \max F^J(I)$, that

$$\mathbf{c}_{F^\omega(I)}(n) = \begin{cases} 1, & \text{if } n \bmod \ell \in F^J(I) \bmod \ell; \\ 0, & \text{otherwise.} \end{cases}$$

Hence, the characteristic sequence of $F^\omega(I)$ is ultimately periodic with preperiod $\max F^J(I)$ and period ℓ. $\square$

Remark 5. In Definition 3 and in what follows, we always assume that all multiplicative constants k_i of the affine maps $\varphi_1, \ldots, \varphi_r$ in F are at least 2. This condition does not guarantee that the corresponding self-generating set is not ultimately periodic. For example, if $\varphi_i(x) = r\,x + i$ for $i = 1, \ldots, r$, then we easily see that $F^\omega(\{0\}) = \mathbb{N}$.

The next lemma justifies that we may restrict our consideration to non-negative integers.

Lemma 6. *Let $F^\omega(I)$ be a self-generating set as given in Definition 3. One can effectively construct two finite sets of non-negative integers I^+ and I^- such that*

$$F^\omega(I) \cap \mathbb{N} = F^\omega(I^+) \cap \mathbb{N} \quad \text{and} \quad F^\omega(I) \cap -\mathbb{N} = -(\overline{F}^\omega(I^-) \cap \mathbb{N}),$$

where $-\mathbb{N}$ is the set of all non-positive integers and $\overline{F} = \{\varphi_0, \overline{\varphi}_1, \overline{\varphi}_2, \ldots, \overline{\varphi}_r\}$ with $\overline{\varphi}_i\colon n \mapsto k_i\,n - \ell_i$ for $i = 1, 2, \ldots, r$.

Proof. Let $m = \max\{|\ell_i| \mid i = 1, 2, \ldots, r\}$ and denote by M the interval of integers $[\![-m, m]\!]$. Define $I_j := F^j(I) \cap M$ for $j \geq 0$. Since $k_i \geq 2$ for all $i \in \{1, 2, \ldots, r\}$, it follows that if n does not belong to M, then $\varphi_i(n) \notin M$ for all $i \in \{0, 1, \ldots, r\}$. By this property and since $F^j(I) \subseteq F^{j+1}(I)$, there must exist an integer J such that $I_j = I_J$ for all $j \geq J$. Hence, the integers of $F^\omega(I)$ falling into the interval M are exactly the ones in I_J and we can find the set $I^+ := ((F^\omega(I) \cap M) \cup I) \cap \mathbb{N}$ effectively.

Next we show that $F^\omega(I) \cap \mathbb{N} = F^\omega(I^+) \cap \mathbb{N}$. Since $I^+ \subseteq F^\omega(I)$, it is clear by definition that $F^\omega(I^+) \cap \mathbb{N} \subseteq F^\omega(I) \cap \mathbb{N}$. Assume now that there exists an integer x belonging to $(F^\omega(I) \cap \mathbb{N}) \setminus (F^\omega(I^+) \cap \mathbb{N})$. Since I^+ contains all non-negative elements of I, the element

x must be generated from some negative element $a \in I$. In other words, there exists a finite sequence $(\varphi_{i_1}, \varphi_{i_2}, \ldots, \varphi_{i_t})$ of maps in F such that $x = \varphi_{i_t} \circ \varphi_{i_{t-1}} \circ \cdots \circ \varphi_{i_1}(a)$. Since a is negative and x is positive, there exists j such that $\varphi_{i_{j-1}} \circ \varphi_{i_{j-2}} \circ \cdots \circ \varphi_{i_1}(a) = y < 0$ and $\varphi_{i_j}(y) = z \geq 0$. Hence, we have $k_{i_j} y < 0$ and $z = k_{i_j} y + \ell_{i_j} < m$. This means that $z \in (F^\omega(I) \cap M) \cap \mathbb{N}$ and therefore $x = \varphi_{i_t} \circ \varphi_{i_{t-1}} \circ \cdots \circ \varphi_{i_{j+1}}(z) \in F^\omega(I^+) \cap \mathbb{N}$. This is a contradiction.

Similarly, by defining $I^- := -(((F^\omega(I) \cap M) \cup I) \cap -\mathbb{N})$, we obtain $F^\omega(I) \cap -\mathbb{N} = F^\omega(-I^-) \cap -\mathbb{N}$. If $\overline{F} = \{\varphi_0, \overline{\varphi}_1, \overline{\varphi}_2, \ldots, \overline{\varphi}_r\}$, where $\overline{\varphi}_i \colon n \mapsto k_i n - \ell_i$ for $i = 1, 2, \ldots, r$, then we clearly have $F^\omega(I) \cap -\mathbb{N} = -(\overline{F}^\omega(I^-) \cap \mathbb{N})$, which concludes the proof. □

Let $y \geq 0$. Recall (for instance, see [3]) that a set $Y \subseteq \mathbb{N}$ is k-recognizable if and only if $Y + y$ is k-recognizable. As explained by the following lemma, from the point of view of recognizability of subsets of $\mathbb{N}$, one can also assume that all additive constants ℓ_i are non-negative.

Lemma 7. *Let $F^\omega(I)$ be a self-generating set as given in Definition 3. There exist a non-negative integer y and a self-generating set $\widehat{F}^\omega(I - y)$ such that $F^\omega(I) = \widehat{F}^\omega(I - y) + y$ and $\widehat{F} = \{\varphi_0, \widehat{\varphi}_1, \ldots, \widehat{\varphi}_r\}$, where $\widehat{\varphi}_i \colon n \mapsto k_i n + \widehat{\ell}_i$ for every $i = 1, 2, \ldots, r$ with some non-negative constants $\widehat{\ell}_i$ completely determined by F.*

Proof. Assume that at least for some function $\varphi_i \in F$ the constant ℓ_i is negative. Otherwise, the claim is trivial. Let $y = \max\{|\ell_i| \mid \ell_i < 0\}$ and set

$$\widehat{\ell}_i := \ell_i + (k_i - 1)y$$

for $i = 1, 2, \ldots, r$. Since $k_i \geq 2$, the constants $\widehat{\ell}_i$ are non-negative for every i. Let $\widehat{F} = \{\varphi_0, \widehat{\varphi}_1, \ldots, \widehat{\varphi}_r\}$ where $\widehat{\varphi}_i \colon n \mapsto k_i n + \widehat{\ell}_i$ for $i = 1, \ldots, r$. We show by induction on the number of applied maps m that x belongs to $F^m(I)$ if and only if $x - y$ belongs to $\widehat{F}^m(I - y)$.

First, for any $x \in I$, it is obvious that $x - y$ belongs to $I - y$ and vice versa. Assume now that $x \in F^m(I)$ for some $m \geq 1$. In other words, there exist $z \in F^{m-1}(I)$ and $i \in \{0, \ldots, r\}$ such that $x = \varphi_i(z)$. By induction hypothesis, $z - y$ belongs to $\widehat{F}^{m-1}(I - y)$. If $\varphi_i = \varphi_0$, then $x = z$ and $x - y \in \widehat{F}^{m-1}(I - y) \subseteq \widehat{F}^m(I - y)$. Hence, assume that $\varphi_i \neq \varphi_0$. We have $\varphi_i(z) = k_i z + \ell_i$ and $\widehat{\varphi}_i(z - y) = k_i(z - y) + \ell_i + (k_i - 1)y = \varphi_i(z) - y$. This proves that $x - y$ belongs to $\widehat{F}^m(I - y)$.

Next assume that $x - y \in \widehat{F}^m(I - y)$ for some $m \geq 1$, i.e., $x - y = \widehat{\varphi}_i(z)$ for some $z \in \widehat{F}^{m-1}(I - y)$ and $i \in \{0, \ldots, r\}$. As above, we may assume that $\varphi_i \neq \varphi_0$. Then we have $x = \widehat{\varphi}_i(z) + y = k_i(z + y) + \ell_i = \varphi_i(z + y)$, where $z + y$ belongs to $F^{m-1}(I)$ by induction hypothesis. Hence, x belongs to $F^m(I)$. □

Example 8. Consider the set $\mathcal{K}_1$ of Example 1 generated from $\{1\}$ by the maps $n \mapsto 2n$ and $n \mapsto 4n - 1$. Applying the construction given in the previous proof, set $y = 1$ and consider the maps $2n + 1$ and $4n + 2$. These two maps generate from $\{1\} - 1 = \{0\}$, the set $\{0, 1, 2, 3, 5, 6, 7, 10, \ldots\}$ which is equal to $\mathcal{K}_1 - 1$.

3 Multiplicatively Dependent Case

In this section we assume that the multiplicative coefficients k_i appearing in Definition 3 are all pairwise multiplicatively dependent, i.e., for every pair (i,j), there exist positive integers e_i and e_j such that $k_i^{e_i} = k_j^{e_j}$. Note that k_i and k_j are multiplicatively dependent if and only if there exist an integer $n \geq 2$ and two integers $d_i, d_j \geq 1$ such that $k_i = n^{d_i}$ and $k_j = n^{d_j}$. By this characterization, it is easy to see that if the coefficients k_i are pairwise multiplicatively dependent, then there exists an integer k such that every k_i is a power of k. Our aim is to build a finite automaton showing that the set $F^\omega(I)$ is k-recognizable.

Recall that $\Sigma_k = \{0, 1, \ldots, k-1\}$ and that $\mathrm{rep}_k \colon \mathbb{N} \to \Sigma_k^*$ maps an integer n to its k-ary representation without leading zeros. For any finite alphabet $A \subseteq \mathbb{Z}$, the function $\mathrm{val}_{A,k} \colon A^* \to \mathbb{Z}$ maps a word $w = w_n w_{n-1} \cdots w_0$ over A to the corresponding numerical value

$$\mathrm{val}_{A,k}(w) = \sum_{i=0}^{n} w_i\, k^i.$$

The function defined over the set of words $w \in A^*$ such that $\mathrm{val}_{A,k}(w) \geq 0$ and which maps w to $\mathrm{rep}_k(\mathrm{val}_{A,k}(w))$ is called *normalization* over A. In the special case $A = \Sigma_k$, we simply write val_k instead of $\mathrm{val}_{\Sigma_k,k}$.

Theorem 9. *Let F given in Definition 3 be such that the multiplicative coefficients $k_1, \ldots, k_r$ are all pairwise multiplicatively dependent. For any finite $I \subset \mathbb{Z}$, the self-generating set $F^\omega(I)$ is k-recognizable if k_i is a power of k for every $i = 1, 2, \ldots, r$.*

We give a proof relying on Frougny's normalization theorem. Another proof is given in [12].

Proof. Assume that the maps in F are of the kind $\varphi_i \colon n \mapsto k^{e_i} n + \ell_i$ with $e_i \geq 1$ for all $i \in \{1, \ldots, r\}$. Since in the constructions of $\overline{F}$ and $\widehat{F}$ of Lemma 6 and Lemma 7 the multiplicative constants k_i are not modified, it suffices to consider only non-negative elements of $F^\omega(I)$ and, moreover, we may assume that all initial values in I and all additive constants ℓ_i are non-negative. Thus, we assume $F^\omega(I) \subseteq \mathbb{N}$ and show that this self-generating set is k-recognizable.

Let n be an element of $F^\omega(I)$. In other words, there exists a finite sequence $(\varphi_{i_1}, \varphi_{i_2}, \ldots, \varphi_{i_m})$ of maps in F such that $n = \varphi_{i_m}(\varphi_{i_{m-1}}(\cdots \varphi_{i_1}(a) \cdots))$ for some $a \in I$. With that integer, we associate the word

$$w = a\, 0^{e_{i_1}-1} \ell_{i_1} \cdots 0^{e_{i_m}-1} \ell_{i_m}$$

over the finite alphabet $A = I \cup \{0, \ell_1, \ldots, \ell_r\} \subset \mathbb{N}$. One can notice that $\mathrm{val}_{A,k}(w) = n$ and $\mathrm{val}_{A,k}(I\{0^{e_1-1}\ell_1, \ldots, 0^{e_r-1}\ell_r\}^*) = F^\omega(I)$. Frougny's normalization theorem ([16, Proposition 7.1.4], see also [8]) says that normalization over A is computable by a finite transducer T. It is also well-known (see, e.g., [1, Theorem 4.3.6]) that if a regular language L is an input of a transducer then the output language is also regular. Hence, feeding the transducer T with the language $I\{0^{e_1-1}\ell_1, \ldots, 0^{e_r-1}\ell_r\}^*$ gives us the regular language $\mathrm{rep}_k(F^\omega(I))$, which proves the claim. $\square$

Remark 10. The set $F^\omega(I)$ considered in the above theorem is k-recognizable and therefore k^n-recognizable for all $n \geq 1$; again, see [3] for details. But usually this set is not ultimately periodic and therefore, by Cobham's Theorem, not ℓ-recognizable for any $\ell \geq 2$ such that k and ℓ are multiplicatively independent. Indeed, if Theorem 15 described below can be applied, then $F^\omega(I)$ contains arbitrarily large gaps.

Remark 11. Garth and Gouge [9] consider the sequence S_F which is the increasing sequence of the elements in $F^\omega(I)$ in the case where $I = \{1\}$, $k_i = k^{e_i}$, $1 = e_1 \leq e_2 \leq \cdots \leq e_r$, $\ell_1 = 0$ and $-k^{e_i} < \ell_i \leq 0$ for each $i = 1, 2, \ldots, r$. They prove that this sequence reduced modulo positive integer m is *morphic*. In other words, there exists a morphism f satisfying $f(a) = ax$ for some letter a and some word $x \neq \varepsilon$ such that S_F mod m is the image under a coding of the infinite word

$$f^\omega(a) = \lim_{n\to\infty} f^n(a) = axf(x)f^2(x)\cdots,$$

which is a fixed point of f. Moreover, they show that the characteristic sequence of $F^\omega(I)$ is k-automatic.

The authors of [9] ask whether their results hold for more general families of functions, for example, allowing $\ell_i \leq -k^{e_i}$. The answer for the case where the multiplicative constants k_i are powers of a fixed k but additive constants ℓ_i are chosen freely follows easily from Theorem 9. Namely, as was mentioned in the introduction, the set of non-negative integers $F^\omega(I)$ is k-recognizable if and only if its characteristic sequence $(\mathbf{c}_{F^\omega(I)}(n))_{n\geq 0}$ is k-automatic. Note that in the general case $F^\omega(I) \subseteq \mathbb{Z}$ we should consider two-sided k-automatic sequences and two-sided infinite fixed points (see Section 5.3 and Section 7.4 in [1] for more information) or consider non-negative and non-positive integers separately. In any case, by Lemma 6, the general case can be reduced to subsets of $\mathbb{N}$.

Hence, let us consider a self-generating set $F^\omega(I) \subseteq \mathbb{N}$ where the multiplicative constants k_i are powers of some k. By Theorem 9, the characteristic sequence $(\mathbf{c}_{F^\omega(I)}(n))_{n\geq 0}$ is k-automatic. Since $(n \bmod m)_{n\geq 0}$ is clearly k-automatic for any $k \geq 2$, then also the sequence

$$\mathbf{u} = ([\mathbf{c}_{F^\omega(I)}(n), n \bmod m])_{n\geq 0}$$

over the alphabet $\Sigma_2 \times \Sigma_m$ is k-automatic. Thus, by the result of Cobham [6], it is the image under a coding of a fixed point of a k-uniform morphism. Define a morphism $f\colon (\Sigma_2 \times \Sigma_m)^* \to \Sigma_m^*$ by

$$f([a,b]) = \begin{cases} \varepsilon, & \text{if } a = 0; \\ b, & \text{otherwise.} \end{cases}$$

Since the image of a morphic sequence by any morphism is either finite or morphic [4] (see also [1, Corollary 7.7.5]) and $(\mathbf{c}_{F^\omega(I)}(n))_{n\geq 0}$ contains infinitely many ones, we conclude that $f(\mathbf{u})$ is morphic. Since $f(\mathbf{u})$ is clearly the sequence S_F reduced modulo m, we have answered the open question of Garth and Gouge [9] by generalizing their results for any additive constants ℓ_i.

Remark 12. Sequences with missing blocks are considered in [2, 9, 14]. For example, if $\varphi_1\colon n \mapsto 2n+1, \varphi_2\colon n \mapsto 4n+2$ and $I = \{0\}$, then the set $F^\omega(I)$ is the set of integers that do not contain the block “00” in their normalized binary expansion. Recall that this set is $\mathcal{K}_1 - 1$; see Example 8. In [2] the authors ask whether or not the sequences with missing blocks are always particular cases of affinely recursive sets. We want to make a remark that, if $F^\omega(I)$ is a sequence with missing blocks, then all constants k_i must be multiplicatively dependent. Otherwise, assume that k_1 and k_2 are multiplicatively independent. Consider now the subset $X_i \subseteq F^\omega(I)$ generated from I by only applying the map φ_i. By Theorem 9, this subset is k_i-recognizable. Consider now the language $0^*\mathrm{rep}_k(X_i)$, where k is multiplicatively independent to k_i. It is known that this language is *right dense* meaning that every word over the alphabet Σ_k appears as a prefix of some word in $0^*\mathrm{rep}_k(X_i)$; for a proof, see [1, Lemma 11.1.1]. Hence, it follows that any block of digits over Σ_k is a factor of $\mathrm{rep}_k(n)$ for some integer $n \in X_i$. For any integer $k \geq 2$, either k_1 or k_2 is multiplicatively independent with k, and therefore the set X_1 or X_2, and consequently also $F^\omega(I)$, cannot be a set of integers that do not have some block of digits in their normalized base-k representation.

In order to obtain a self-contained proof for Theorem 9, we may tailor Frougny's normalization transducer for the language $0^*I\{0^{e_1-1}\ell_1, \ldots, 0^{e_r-1}\ell_r\}^*$ and directly conclude that the output language over Σ_k is regular. Next we describe this in more detail. The following construction is needed to prove the result relating $\mathcal{K}_1$ and the infinite Fibonacci word in the next section. By Lemma 6, it suffices to consider the set $F^\omega(I) \cap \mathbb{N}$.

Let $C \subset \mathbb{Z}$ be a finite input alphabet and let Σ_k be the output alphabet. Denote $m = \max\{|c-a| \mid c \in C, a \in \Sigma_k\}$ and let $\gamma = m/(k-1)$. Note that by the Euclidean division, for every $s \in \mathbb{Z}$ and $c \in C$, there exist a unique $a \in \Sigma_k$ and $s' \in \mathbb{Z}$ such that $s + c = s'k + a$. Moreover, if $|s| < \gamma$, then $|s'| \leq (|s| + |c-a|)/k < (\gamma + m)/k = \gamma$. This justifies that we may define a finite right subsequential transducer, where the set of states $Q = \{s \in \mathbb{Z} \mid |s| < \gamma\}$ corresponds to possible carries, the initial state is 0 and the set of edges is

$$E = \{s \xrightarrow{c/a} s' \mid s + c = s'k + a\}. \tag{3}$$

A *right subsequential* transducer is a transducer that reads the input from right to left and the underlying automaton where only inputs are considered is deterministic. Moreover, we have a partial terminal function $\omega\colon Q \to \Sigma_k^*$ mapping a state $s \geq 0$ onto its normalized representation $\mathrm{rep}_k(s)$. Let $w = c_n c_{n-1} \cdots c_0 \in C^* \setminus 0C^*$ be a representation of an integer $N = \mathrm{val}_{C,k}(w) \geq 0$. If we enter w into the transducer, there is a unique path

$$0 = s_0 \xrightarrow{c_0/a_0} s_1 \xrightarrow{c_1/a_1} s_2 \xrightarrow{c_2/a_2} \cdots \xrightarrow{c_n/a_n} s_{n+1}$$

such that $N = \sum_{i=0}^n c_i k^i = \sum_{i=0}^n a_i k^i + s_{n+1}k^{n+1}$. Hence, $\omega(s_{n+1})a_n a_{n_1} \cdots a_0$ is the normalized representation in base k of the integer N. This transducer is Frougny's normalization transducer for an input not containing leading zeros; see the proof of Lemma 7.1.1 in [16].

Next we adapt the above construction to our specific case of self-generating sets. Let the input alphabet be $C = I \cup \{0, \ell_1, \ldots, \ell_r\}$. We want to restrict the accepted input to the words $w \in 0^*I\{0^{e_1-1}\ell_1, \ldots, 0^{e_r-1}\ell_r\}^*$ such that $\mathrm{val}_{C,k}(w) \geq 0$. As was shown in the proof of Theorem 9, these words represent exactly the numbers in $F^\omega(I) \cap \mathbb{N}$. Hence, we build a transducer $\mathcal{T}$ such that from each carry state $q \in Q = \{s \in \mathbb{Z} \mid |s| < \gamma\}$ we may read only words of the form $0^{e_i-1}\ell_i$ from right to left, output the corresponding output of Frougny's transducer and end up in some carry state $q' \in Q$. This can be achieved by introducing chains of intermediate states where each state has only one incoming and outgoing edge simulating the behavior of Frougny's transducer. For example, assume that $k = 2$, $q = 1$ and we want to read 003 from right to left. This corresponds to the map $\varphi\colon n \mapsto 8n + 3$. By the construction, in our modified transducer there is a path

$$1 \xrightarrow{3/0} \hat{2} \xrightarrow{0/0} \hat{1} \xrightarrow{0/1} 0,$$

where $\hat{2}$ and $\hat{1}$ are additional intermediate states and the starting state 1 and the ending state 0 belong to the original set Q. From each state $q \in Q$ there are exactly r paths of this kind corresponding to the r maps $\varphi_i \in F$.

In addition, we need transitions corresponding to the initial values I. Let $t \notin Q$ be a unique final state. For each $q \in Q$ and $a \in I$ such that $q + a \geq 0$, we add extra states and transitions which form a separate path from q to t such that it simulates Frougny's transducer with input $0^i a$, where i is the maximum integer satisfying $k^i \leq q + a$. Padding with sufficiently many zeros insures that the carry is 0 after entering the final state t. Note that since we consider only non-negative elements of $F^\omega(I)$, we do not build a path from q to the final state t for an initial value $a \in I$ such that $q + a < 0$. For example, in the case $k = 2$, $q = 1$ and $a = 5$ we have $i = 2$, since $k^2 < q + a = 6 < k^3$, and the path from q to t is

$$1 \xrightarrow{5/0} \hat{3} \xrightarrow{0/1} \hat{1} \xrightarrow{0/1} t,$$

where $\hat{3}$ and $\hat{1}$ are new intermediate states. There is also a loop from the final state t onto itself with input 0 and output 0. This corresponds to allowing leading zeros after the most significant non-zero digit.

By our construction, each path from the initial state 0 to the final state t corresponds to reading some word of the language $0^*I\{0^{e_1-1}\ell_1, \ldots, 0^{e_r-1}\ell_r\}^*$. Therefore, the output of such an accepted path in our transducer $\mathcal{T}$ corresponds to some normalized representation (with possibly leading zeros) of a number in the self-generating set $F^\omega(I)$. Conversely, the normalized representation of a number in $F^\omega(I)$ padded with sufficiently many zeros corresponds to the input of some accepted path in our transducer $\mathcal{T}$. Therefore, we may forget the input and consider a finite automaton $\mathcal{A}$ where the edges are labeled only with the output. Moreover, let us define that if in $\mathcal{A}$ there is a path from a state q to the state t with a label belonging to 0^*, then the set q is an accepting state. This allows us to accept all normalized representations with an arbitrary number of leading zeros. We may also change the reading direction by turning the arrows and changing the roles of the initial and final states. Of course, the automaton obtained this way need not be complete and deterministic, but it can be made complete by adding missing edges which end up in a sink state and it can be made deterministic by the subset construction. Hence, we have constructed this way a deterministic finite automaton $\mathcal{B}$ which recognizes $0^*\mathrm{rep}_k(F^\omega(I) \cap \mathbb{N})$ and, by Lemma 6, we conclude that $F^\omega(I)$ is k-recognizable.

4 Kimberling set and the Fibonacci word

In this section we show a result connecting the Kimberling set $\mathcal{K}_1$ considered in Example 1 and the infinite Fibonacci word defined as the fixed point $\varphi^\omega(0) = 01\varphi(1)\varphi^2(1)\cdots =$

$01001010\cdots$ of the morphism $\varphi\colon 0 \mapsto 01, 1 \mapsto 0$. Recall that $\mathcal{K}_1 = F^\omega(I)$, where $F = \{\varphi_0, \varphi_1, \varphi_2\}$, $\varphi_1\colon n \mapsto 2n$, $\varphi_2\colon n \mapsto 4n-1$ and $I = \{1\}$.

Theorem 13. *Let S be the increasing sequence of elements of $\mathcal{K}_1$. Omitting the first term, the sequence S reduced modulo 2, is the Fibonacci word $\varphi^\omega(0)$.*

This was the main result in [13] and it was reproved in [2]. Here we give a third proof based on the transducer construction of the previous section and on some technical manipulation of morphisms.

Proof. Let us first build the transducer $\mathcal{T}$ for the set $\mathcal{K}_1 = F^\omega(I)$ as explained in the end of Section 3. This transducer and the corresponding reduced automaton $\mathcal{A}$ are illustrated in Figure 1. Using the same notation as above, we have $k = 2$, $C = \{1, 0, -1\}$, $m = 2$ and $\gamma = 2$. Since we never reach a carry state 1 from the initial state 0, our set $Q = \{-1, 0, 1\}$ can be reduced to $\{-1, 0\}$. The input 0 corresponds to the map φ_1 and the input $0(-1)$ corresponds to the map φ_2. When we read $0(-1)$ from right to left starting from either state 0 or -1, we introduce an intermediate state $\widehat{-1}$. Namely, for $s = 0$ and $c = -1$, we have $s + c = (-1) \cdot k + 1$ and, for $s = -1$ and $c = -1$, we have $s + c = (-1) \cdot k + 0$. Then from the state $\widehat{-1}$ we must read 0 and, since $-1 + 0 = -1 \cdot k + 1$, we output 1 and end up in $-1 \in Q$. Moreover, we can read the initial value $1 \in I$ starting from any state in Q. For example, there is an edge with label $1/0$ from -1 to F, since $-1 + 1 = 0 \cdot k + 0$.

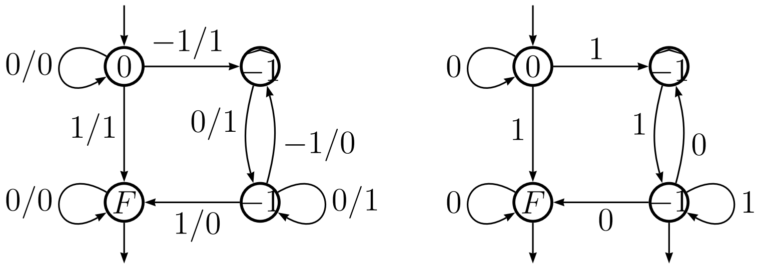

Figure 1: Transducer $\mathcal{T}$ and automaton $\mathcal{A}$ corresponding to the Kimberling set.

Using standard techniques we may easily build from $\mathcal{A}$ a deterministic automaton $\mathcal{B}$ accepting $0^*\mathrm{rep}_2(\mathcal{K}_1)$ when reading digits from left to right. This automaton is described in Figure 2. A number in $\mathcal{K}_1$ such that its binary representation is accepted by b (the corresponding path ends in the final state b) must be odd, since all incoming edges of b are labeled by 1. Similarly, we conclude that a number having a binary representation accepted by c or d must be even. Hence, with an output function $\tau\colon A^* \mapsto \Sigma_2^*$, where A denotes the set of states of $\mathcal{B}$ and

$$\tau(x) = \begin{cases} 1, & \text{if } x = b; \\ 0, & \text{if } x = c \text{ or } x = d; \\ \varepsilon, & \text{otherwise,} \end{cases}$$

the automaton $\mathcal{B}_\tau$ generates the sequence $S \bmod 2$, where S is the increasing sequence of elements of $\mathcal{K}_1$.

The 2-uniform morphism corresponding to $\mathcal{B}$ is $\sigma\colon A^* \to A^*$ defined by

$$a \mapsto ab, \quad b \mapsto cb, \quad c \mapsto de, \quad d \mapsto dg, \quad e \mapsto fb, \quad f \mapsto ge, \quad g \mapsto gg.$$

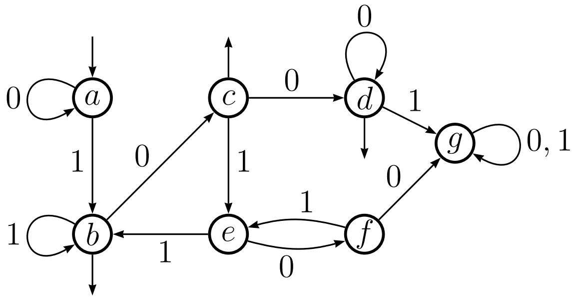

Figure 2: A finite deterministic automaton $\mathcal{B}$ accepting $0^*\mathrm{rep}_2(\mathcal{K}_1)$.

By the above reasoning, it is clear that $\tau(\sigma^\omega(a)) = S \bmod 2$. Let us denote $B = \{a, b, c, d, e, f, g, h\}$. Using the techniques described in the proof of Theorem 7.6.1 and Theorem 7.7.4 in [1], we obtain a coding $\nu\colon B^* \to \Sigma_2^*$ and a non-erasing morphism $\mu\colon B^* \to B^*$ such that $\nu(\mu^\omega(a)) = \mathcal{S}(\tau(\sigma^\omega(a)))$, where $\mathcal{S}$ is the shift function deleting the first element of the infinite word. The morphism μ is defined by

$$a \mapsto abcdbeb, \quad b \mapsto cdb, \quad c \mapsto fgb, \quad d \mapsto eb, \quad e \mapsto fh,$$
$$f \mapsto f, \quad g \mapsto gbeb, \quad h \mapsto hcdb$$

and

$$\nu(x) = \begin{cases} 1, & \text{if } x = b \text{ or } x = h; \\ 0, & \text{otherwise.} \end{cases}$$

Our goal is to show that $\nu(\mu^\omega(a))$ is the infinite Fibonacci word. For this purpose, let us first simplify the morphism μ. Since $\mu(fgb) = fgbebcdb = \mu(cdb)$, we conclude that $\nu(\mu^i(fgb)) = \nu(\mu^i(cdb))$ for every $i \geq 0$ and, consequently, we may set $\mu(c) = cdb$ without changing $\nu(\mu^\omega(a))$. Similarly, $\mu(fh) = fhcdb = \mu(eb)$ and therefore $\nu(\mu^i(e)) = \nu(\mu^i(d))$ for $i \geq 0$. Thus, we may set $e = d$ and replace the morphism μ by a simpler morphism on a four-letter alphabet $\{a, b, c, d\}$:

$$a \mapsto abcdbdb, \quad b \mapsto cdb, \quad c \mapsto cdb, \quad d \mapsto db.$$

Note that b and c have a different role with respect to the coding, i.e., $\nu(b) \neq \nu(c)$. Since b is always preceded by d except in the very beginning, we finally redefine the morphism $\mu\colon \{a, b, c, d\}^* \to \{a, b, c, d\}^*$ by

$$a \mapsto abcdbdbc, \quad b \mapsto db, \quad c \mapsto cdb, \quad d \mapsto dbc.$$

Hence, the sequence obtained by reducing S modulo 2 and omitting the first element can be obtained as the image of a coding ν of the fixed point $\mu^\omega(a)$.

Let us next modify the morphism generating the Fibonacci word. First, note that we may replace φ by φ^2, since clearly $\lim_{n\to\infty} \varphi^n(0) = \lim_{n\to\infty}(\varphi^2)^n(0)$. Since $\varphi^2(0) = 010$ and $\varphi^2(1) = 01$, we notice that there are two types of zeros in the Fibonacci word: those followed by 0 will be denoted by c and those followed by 1 will be denoted by d. Let us also replace every 1 by b. Hence, we have $\varphi^\omega(0) = \nu(\phi^\omega(d))$, where ν is the coding defined above and $\phi\colon \{b, c, d\}^* \to \{b, c, d\}^*$ is a morphism such that

$$b \mapsto db, \quad c \mapsto dbc, \quad d \mapsto dbc.$$

We denote $(f_n)_{n\ge 0} = \phi^\omega(d) = dbcdb\cdots$ and $(s_n)_{n\ge 0} = \mu^\omega(a) = abcdbdbcdb\cdots$. In order to prove the result of Kimberling, we have to show that $\nu(\phi^\omega(d)) = \nu(\mu^\omega(a))$. Since $\nu(f_0) = \nu(d) = 0 = \nu(a) = \nu(s_0)$, it suffices to show that $f_n = s_n$ for all $n \ge 1$. We do this by induction.

First observe that if $s_n = f_n$ for all $n = 1, 2, \ldots, k$, then

$$|\mu(s_0 \cdots s_k)| = |\phi(f_0 \cdots f_k)| + 5. \tag{4}$$

This holds because $|\mu(x)|_y = |\phi(x)|_y$ for every x and y in $\{b, c, d\}$ and $|\mu(s_0)| = |\mu(a)| = |\phi(f_0)| + 5$. Here $|w|_y$ denotes the number of letters y occurring in the word w.

Now assume that $s_n = f_n$ for $1 \le n \le l$ and l is such that $\phi(f_0 \cdots f_k) = f_0 f_1 \cdots f_l$ for some $k > 1$ satisfying $f_k = b$. This implies that $\phi(f_0 \cdots f_k) = u f_{l-1} f_l = udb$ and, by (4) and by the assumption, we have

$$\mu(s_0 \cdots s_k) = udb s_{l+1} s_{l+2} s_{l+3} s_{l+4} s_{l+5} = udb.dbc.db, \tag{5}$$

where $s_{l+4}s_{l+5} = \mu(s_k) = \mu(b)$ and $s_{l+1}s_{l+2}s_{l+3} = \mu(s_{k-1}) = \mu(d)$, since $s_k = b$ must be preceded by d if $k > 1$. We have two possibilities, either $f_{k+1}f_{k+2} = db$ or $f_{k+1}f_{k+2}f_{k+3} = cdb$.

If $f_{k+1}f_{k+2} = db$, then $\phi(f_0 \cdots f_{k+2}) = udb\phi(f_{k+1})\phi(f_{k+2}) = udb.dbc.db$ and, by comparing this to (5), we conclude that the claim $s_n = f_n$ holds for $1 \le n \le l + 5$.

Assume next that $f_{k+1}f_{k+2}f_{k+3} = cdb$. Now $f_1 \cdots f_{k+3} = s_1 \cdots s_{k+3}$, since we must have $k + 3 \le l$. Hence, we obtain

$$\begin{aligned} \phi(f_0 \cdots f_{k+3}) &= udb.dbc.dbc.db, \\ \mu(s_0 \cdots s_{k+3}) &= udb.dbc.db.cdb.dbc.db, \end{aligned}$$

which implies that $s_n = f_n$ for $1 \le n \le l + 8$.

Since in the first case $f_{k+2} = b$ and in the second case $f_{k+3} = b$, we may proceed by induction. This concludes the proof, since the claim clearly holds for small values of $n \ge 1$. $\square$

5 Multiplicatively Independent Case

In this section our aim is to show that $F^\omega(I) \subseteq \mathbb{N}$ given in Definition 3 is not recognizable in any base $k \ge 2$ provided that $\sum_{i=1}^{r} k_i^{-1} < 1$ and that there are at least two multiplicatively independent coefficients k_i. For the proof, we introduce the following notation. Let $X = \{x_0 < x_1 < x_2 < \cdots\}$ be an infinite ordered subset of $\mathbb{N}$. Then we denote

$$R_X = \limsup_{i\to\infty} \frac{x_{i+1}}{x_i} \text{ and } D_X = \limsup_{i\to\infty} (x_{i+1} - x_i).$$

In order to prove that a set is not k-recognizable for any base $k \ge 2$, we use the following result from [6], see also Eilenberg's book [7, Chapter V, Theorem 5.4].

Theorem 14 (Gap Theorem). *Let $k \ge 2$. If X is a k-recognizable infinite subset of $\mathbb{N}$, then either $R_X > 1$ or $D_X < \infty$.*

Note that $D_X < \infty$ means that X is *syndetic*, i.e., there exists a constant C such that the gap $x_{i+1} - x_i$ between any two consecutive elements x_i, x_{i+1} in X is bounded by C. Let us first show that if $\sum_{i=1}^{r} k_i^{-1} < 1$, then the set $F^\omega(I)$ given in Definition 3 contains arbitrarily large gaps.

Theorem 15. *Let $X = F^\omega(I)$ be a self-generating subset of $\mathbb{N}$ given in Definition 3. If $\sum_{i=1}^{r} k_i^{-1} < 1$, then X is not syndetic.*

Proof. Let $n \geq 1$ and $K = k_1 k_2 \cdots k_r$. Let $g = g_1 \circ g_2 \circ \cdots \circ g_n$ be a composite function, where g_j belongs to $G = \{\varphi_1, \varphi_2, \ldots, \varphi_r\}$ for every $j = 1, 2, \ldots, n$ and $g_j = \varphi_i$ for exactly n_i integers $j \in \{1, \ldots, n\}$. Note that $n_1 + n_2 \cdots + n_r = n$. By definition, we have $g(x) = k_1^{n_1} k_2^{n_2} \cdots k_r^{n_r} x + c_g$, where c_g is some constant depending on g. Since $k_1^{n_1} k_2^{n_2} \cdots k_r^{n_r}$ divides K^n, we get

$$\#\{g(x) \bmod K^n \mid x \in \mathbb{Z}\} = k_1^{n-n_1} k_2^{n-n_2} \cdots k_r^{n-n_r}.$$

Recall that $F = G \cup \{\varphi_0\}$, where φ_0 denotes the identity function. The set $F^n(I)$ contains exactly the integers obtained by at most n applications of maps in G. For any interval of integers $[\![N, N + K^n - 1]\!]$ where $N > \max F^n(I)$, the elements of X belonging to this interval have been obtained by applying at least $n+1$ maps. Hence, in the interval $[\![N, N + K^n - 1]\!]$ there can be at most $k_1^{n-n_1} k_2^{n-n_2} \cdots k_r^{n-n_r}$ integers $x \in X$ such that the last n maps which produce x correspond to the composite function g, i.e., such that there exists $y \in X$ satisfying $g(y) = x$. For fixed numbers n_i, $i = 1, 2, \ldots, r$, there are $n!/(n_1! n_2! \cdots n_r!)$ functions g of the type described above. Thus, the number of integers in $X \cap [\![N, N + K^n - 1]\!]$ for any large enough N is at most

$$\sum_{n_1, n_2, \ldots, n_r} \left(\frac{n!}{n_1! n_2! \cdots n_r!} \right) k_1^{n-n_1} k_2^{n-n_2} \cdots k_r^{n-n_r} = K^n \left(\frac{1}{k_1} + \frac{1}{k_2} + \cdots + \frac{1}{k_r} \right)^n$$

where the sum is over $n_1, n_2, \ldots, n_r \geq 0$ satisfying $n_1 + n_2 + \cdots + n_r = n$.

Hence, the biggest gap $x_{i+1} - x_i$ between two consecutive elements $x_i, x_{i+1} \in X$ in the interval $[\![N, N + K^n - 1]\!]$ is at least

$$d(n) = \frac{K^n}{K^n \left(\frac{1}{k_1} + \frac{1}{k_2} + \cdots + \frac{1}{k_r} \right)^n} = \left(\frac{1}{k_1} + \frac{1}{k_2} + \cdots + \frac{1}{k_r} \right)^{-n}.$$

Since $\sum_{i=1}^{r} k_i^{-1} < 1$, the function $d(n)$ tends to infinity as n tends to infinity. This means that there are arbitrarily large gaps in X. In other words, the self-generating set X is not syndetic. $\square$

Before showing that $R_X = 1$ let us first recall the density property of multiplicatively independent integers. A set S is *dense* in an interval $I \subseteq \mathbb{R}$ if every subinterval of I contains an element of S.

Theorem 16. *If $k, \ell \geq 2$ are multiplicatively independent, $\{k^p/\ell^q \mid p, q \geq 0\}$ is dense in $[0, \infty)$.*

This is a consequence of Kronecker's theorem, which states that for any irrational number θ the sequence $(\{n\theta\})_{n\geq 0}$ is dense in the interval $[0,1)$. Here $\{x\}$ denotes the fractional part of the real number x. The proof of Kronecker's theorem as well as the proof of Theorem 16 can be found in [1, Section 2.5] or [11]. As an easy consequence of the previous theorem, we obtain the following result.

Corollary 17. *Let $\alpha > 0$ and β be two real numbers. If k and ℓ are multiplicatively independent, then the set $\{(\alpha k^p + \beta)/\ell^q \mid p, q \geq 0\}$ is dense in $[0, \infty)$.*

Proof. We show how to get arbitrarily close to any positive real number x. Let $\epsilon > 0$. By Theorem 16, there exist integers p and q such that

$$\left|\frac{x}{\alpha} - \frac{k^p}{\ell^q}\right| < \frac{\epsilon}{2\alpha} \quad \text{and} \quad \left|\frac{\beta}{\ell^q}\right| < \frac{\epsilon}{2}.$$

Hence, it follows that

$$\left|x - \frac{\alpha k^p + \beta}{\ell^q}\right| \leq \left|x - \frac{\alpha k^p}{\ell^q}\right| + \left|\frac{\beta}{\ell^q}\right| < \frac{\epsilon}{2\alpha}\alpha + \frac{\epsilon}{2} = \epsilon.$$

□

Let us next consider the ratio R_X of a self-generating set X.

Theorem 18. *For any self-generating set $X = F^\omega(I) \subseteq \mathbb{N}$ given in Definition 3 where k_i and k_j are multiplicatively independent for some i and j, we have $R_X = 1$.*

Proof. Without loss of generality, we may assume that $F = \{\varphi_0, \varphi_1, \varphi_2\}$, where $\varphi_1 \colon n \mapsto k_1 n + \ell_1$, $\varphi_2 \colon n \mapsto k_2 n + \ell_2$, and k_1 and k_2 are multiplicatively independent. Namely, for $F \subseteq F'$, it is obvious that $F^\omega(I) \subseteq F'^\omega(I)$ and consequently, $R_{F^\omega(I)} = 1$ implies $R_{F'^\omega(I)} = 1$. By Lemma 7, we may also assume that ℓ_1 and ℓ_2 are non-negative.

Let $a \in X$ be a positive integer and set $X_n := X \cap [\varphi_1^{n-1}(a), \varphi_1^n(a)]$ for all $n > 0$. Note that $\cup_{n\in\mathbb{N}} X_n = X \cap [a, \infty)$. Recall that $X = \{x_0 < x_1 < x_2 < \cdots\}$ and define

$$r_n := \max\left\{\frac{x_{i+1} - x_i}{x_i} \,\middle|\, x_{i+1}, x_i \in X_n\right\}.$$

Note that, for all x and for $j = 1, 2$, if we set $b_j := \ell_j/(k_j - 1)$, then we have

$$\varphi_j^n(x) = k_j^n x + \ell_j \sum_{i=0}^{n-1} k_j^i = (x + b_j)\, k_j^n - b_j. \tag{6}$$

Let $m \geq 0$ and x_i, x_{i+1} be two consecutive elements belonging to the set X_m. By Corollary 17, there exist infinitely many positive integers p and q such that $\frac{\varphi_2^p(a)}{k_1^q}$ is equal to

$$\frac{(a + b_2)k_2^p - b_2}{k_1^q} \in \left[x_{i+1} + b_1 - \frac{3}{4}(x_{i+1} - x_i), x_i + b_1 + \frac{3}{4}(x_{i+1} - x_i)\right].$$

Therefore $\varphi_2^p(a)$ is an element of X belonging to the interval

$$[c,d] := \left[k_1^q(x_{i+1}+b_1) - \frac{3}{4}k_1^q(x_{i+1}-x_i), k_1^q(x_i+b_1) + \frac{3}{4}k_1^q(x_{i+1}-x_i)\right],$$

which is a sub-interval[2] of the interval $[\varphi_1^q(x_i), \varphi_1^q(x_{i+1})]$. In other words, we have

$$\varphi_1^q(x_i) < c < \varphi_2^p(a) < d < \varphi_1^q(x_{i+1}).$$

Hence, for all $t > q$, the difference $x_{j+1} - x_j$ of any two consecutive elements x_j, x_{j+1} of X in the interval $[\varphi_1^t(x_i), \varphi_1^t(x_{i+1})]$ is at most

$$\begin{aligned}&\max\{\varphi_1^{t-q}(\varphi_1^q(x_{i+1})) - \varphi_1^{t-q}(\varphi_2^p(a)), \varphi_1^{t-q}(\varphi_2^p(a)) - \varphi_1^{t-q}(\varphi_1^q(x_i))\}\\ &\quad \le \max\{\varphi_1^t(x_{i+1}) - \varphi_1^{t-q}(c), \varphi_1^{t-q}(d) - \varphi_1^t(x_i)\} = \frac{3}{4}k_1^t(x_{i+1}-x_i) + b_1 k_1^{t-q}.\end{aligned}$$

Thus, the ratio $(x_{j+1}-x_j)/x_j$ is at most

$$\frac{3\,k_1^t(x_{i+1}-x_i)}{4\,\varphi_1^t(x_i)} + \frac{b_1 k_1^{t-q}}{\varphi_1^t(x_i)} = \frac{3\,k_1^t(x_{i+1}-x_i)}{4\,\varphi_1^t(x_i)} + \frac{1}{k_1^q}\frac{b_1 k_1^t}{(x_i+b_1)k_1^t - b_1}. \tag{7}$$

The latter term in this sum can be taken as small as possible for q and t large enough ($1/k_1^q$ tends to 0 and the other factor tends to the constant $b_1/(x_i+b_1)$). In particular, for q and t large enough, we have

$$\frac{b_1 k_1^{t-q}}{\varphi_1^t(x_i)} < \frac{x_{i+1}-x_i}{12 x_i}.$$

Moreover, we have

$$\frac{3\,k_1^t(x_{i+1}-x_i)}{4\,\varphi_1^t(x_i)} = \frac{3\,(x_{i+1}-x_i)}{4\,(x_i+b_1-b_1/k_1^t)} < \frac{3\,(x_{i+1}-x_i)}{4\,x_i} < \frac{10\,(x_{i+1}-x_i)}{12\,x_i}.$$

Thus, by (7), we obtain

$$\frac{x_{j+1}-x_j}{x_j} < \frac{11\,(x_{i+1}-x_i)}{12\,x_i}. \tag{8}$$

Since the above holds for any consecutive elements x_i and x_{i+1} in X_m and there are only finitely many such pairs, we conclude that there exists an integer N_1 such that (8) holds for any consecutive elements $x_j, x_{j+1} \in X_n$ where $n \ge N_1$. Hence, we obtain $r_n < \frac{11}{12} r_m$ for every $n \ge N_1$. Moreover, by repeating this procedure, we conclude that there exists an integer N_k such that

$$r_n < \left(\frac{11}{12}\right)^k r_m$$

for every $n \ge N_k$. This implies that $\limsup_{n\to\infty} r_n = 0$ and, consequently,

$$R_X = 1 + \limsup_{n\to\infty} r_n = 1.$$

□

[2] $c - \varphi_1^q(x_i) = \frac{1}{4}k_1^q(x_{i+1}-x_i) + b_1$ and $\varphi_1^q(x_{i+1}) - d = \frac{1}{4}k_1^q(x_{i+1}-x_i) - b_1$ which is positive for large enough q.

Our main result is a straightforward consequence of the previous theorems.

Theorem 19. *Let $X = F^\omega(I) \subseteq \mathbb{N}$ be given in Definition 3. If $\sum_{t=1}^{r} k_t^{-1} < 1$ and there exist i, j such that k_i and k_j are multiplicatively independent, then $F^\omega(I)$ is not k-recognizable for any integer base $k \geq 2$.*

Proof. Let $X = F^\omega(I)$ satisfy the assumptions of the theorem. By Theorem 15, we have $D_X = \infty$ and, by Theorem 18, we have $R_X = 1$. Thus, Theorem 14 implies that X is not k-recognizable for any $k \geq 2$. $\square$

As a corollary, we have solved the conjecture of Allouche, Shallit and Skordev [2].

Corollary 20. *Let $F = \{\varphi_0, n \mapsto k_1 n + \ell_1, n \mapsto k_2 n + \ell_2\}$, where k_1 and k_2 are multiplicatively independent. Then any infinite self-generating set $F^\omega(I)$ given in Definition 3 is not k-recognizable for any $k \geq 2$.*

Proof. This follows directly from Theorem 19. Namely, if k_1 and k_2 are multiplicatively independent, then $k_1 \geq 2$ and $k_2 \geq 3$ and $k_1^{-1} + k_2^{-1} \leq 1/2 + 1/3 = 5/6 < 1$. $\square$

The condition $\sum_{t=1}^{r} k_t^{-1} < 1$ is not needed in a very special case of self-generating sets where $\ell_i = 0$ for every $i = 1, 2, \ldots, r$. This situation is related to so-called y-smooth numbers. An integer is *y-smooth* if it has no prime factors greater than y. For more on smooth numbers, see, e.g., [10].

Theorem 21. *Let $X = F^\omega(I)$ be given in Definition 3. If $\ell_i = 0$ for every $i = 1, 2, \ldots, r$ and there exist i, j such that k_i and k_j are multiplicatively independent, then $F^\omega(I)$ is not k-recognizable for any integer base $k \geq 2$. In particular, for $y \geq 3$, the set of y-smooth numbers is not k-recognizable for any $k \geq 2$.*

Proof. Assume that $\varphi_i \colon n \mapsto k_i n$ for $i = 1, 2, \ldots, r$ and denote $X = F^\omega(I)$. Let $x \geq 2$ be an integer and consider $n \in X \cap [0, x]$. By the definition of X, the integer n must be of the form $k_1^{e_1} \cdots k_r^{e_r} a$, where $a \in I$. Since the exponent e_i is at most $\log_2(x)$ for every $i = 1, 2, \ldots, r$, the number of integers in $X \cap [0, x]$ is at most $(1 + \log_2(x))^r |I| = O(\log^r(x))$. It follows that $x/|X \cap [0, x]|$ tends to infinity when x tends to infinity. This implies that $F^\omega(I)$ cannot be syndetic, i.e., $D_X = \infty$. If there are two multiplicatively independent constants k_1 and k_2, then $R_X = 1$ by Theorem 18. Hence, by Theorem 14, the self-generating set X is not k-recognizable for any $k \geq 2$. The second claim follows, since the set of y-smooth numbers can be represented as a self-generating set $F^\omega(I)$, where $I = \{1\}$ and $\varphi_i \colon n \mapsto p_i n$ for $i = 1, 2, \ldots, r$. Here p_i is the ith smallest prime and p_r is the largest prime less than or equal to y. $\square$

6 Acknowledgments

We thank the anonymous referees of the MFCS version of this paper for suggesting improvements in the presentation of this paper. In particular, it is one of the referees who pointed out a possible connection with smooth numbers.

References

[1] J.-P. Allouche, J. Shallit, *Automatic Sequences: Theory, Applications, Generalizations*, Cambridge University Press, 2003.

[2] J.-P. Allouche, J. Shallit, G. Skordev, Self-generating sets, integers with missing blocks, and substitutions, *Discrete Math.* **292** (2005), 1–15.

[3] V. Bruyère, G. Hansel, C. Michaux, R. Villemaire, Logic and p-recognizable sets of integers, *Bull. Belg. Math. Soc.* **1** (1994), 191–238.

[4] A. Cobham, On the Hartmanis-Stearns problem for a class of tag machines, in *IEEE Conference Record of 1968 Ninth Annual Symposium on Switching and Automata Theory*, 1968, pp. 51–60. Also appeared as IBM Research Technical Report **RC-2178**, August 23, 1968.

[5] A. Cobham, On the base-dependence of sets of numbers recognizable by finite automata, *Math. Systems Theory* **3** (1969), 186–192.

[6] A. Cobham, Uniform tag sequences, *Math. Systems Theory* **6** (1972), 164–192.

[7] S. Eilenberg, *Automata, Languages, and Machines*, Vol. A., Pure and Applied Mathematics, Vol. **58**, Academic Press, 1974.

[8] C. Frougny, Representations of numbers and finite automata, *Math. Systems Theory* **25** (1992), 37–60.

[9] D. Garth, A. Gouge, Affinely self-generating sets and morphisms, *J. Integer Seq.* **10** (2007), Article 07.1.5.

[10] A. Granville, Smooth numbers: computational number theory and beyond, in J. P. Buhler, P. Stevenhagen, eds., *Algorithmic number theory: lattices, number fields, curves and cryptography*, Math. Sci. Res. Inst. Publ. **44**, Cambridge University Press, 2008, pp. 267–323.

[11] G. H. Hardy, E. M. Wright, *Introduction to the Theory of Numbers*, Oxford University Press, 1985.

[12] T. Kärki, A. Lacroix, M. Rigo, On the recognizability of self-generating sets, in R. Královič, D. Niwiński, eds., *Proceedings of the 34st International Symposium on Mathematical Foundations of Computer Science, Bratislava, August 24 - 28, 2009*, Lecture Notes in Comput. Sci. **5734** (2009), 525–536.

[13] C. Kimberling, A self-generating set and the golden mean, *J. Integer Seq.* **3** (2000), Article 00.2.8.

[14] C. Kimberling, Affinely recursive sets and orderings of languages, *Discrete Math.* **274** (2004), 147–159.

[15] P. B. A. Lecomte, M. Rigo, Numeration systems on a regular language, *Theory Comput. Syst.* **34** (2001), 27–44.

[16] M. Lothaire, *Algebraic Combinatorics on Words*, Encyclopedia of Mathematics and its Applications, **90**, Cambridge University Press, 2002.

2000 *Mathematics Subject Classification*: Primary 68Q45; Secondary 68R15, 11B85.
Keywords: self-generating set, recognizability, numeration systems, multiplicatively independent integers, Fibonacci word.

(Concerned with sequences A000045, A000201, A001950, A003754, A003849, A052499.)

Received November 16 2009; revised version received January 21 2010. Published in *Journal of Integer Sequences*, January 27 2010.

Return to Journal of Integer Sequences home page.