On the Recognizability of Self-Generating Sets

Tomi Kärki*, Anne Lacroix, Michel Rigo

University of Liège, Institute of Mathematics, Grand Traverse 12 (B 37), B-4000 Liège, Belgium. {T.Karki,A.Lacroix,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 $\varphi \in F$. In particular, solving a conjecture of Allouche, Shallit and Skordev, we show 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 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 [4], 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 introduced the so-called *self-generating* sets of integers [10]. 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 $\varphi_i : n \mapsto k_i\, n + \ell_i$ 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

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

* supported by Osk. Huttunen Foundation.

Example 1. In [10], for $G = \{n \mapsto 2n, n \mapsto 4n-1\}$ and $I = \{1\}$, it is shown 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. 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 $\mathcal{K}_0 \cup \mathcal{K}_1$.

These self-generating sets are also called *affinely recursive* in [11] 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 [12].

In [2] a general framework for self-generating sets is considered. The k-ary representations of the elements in a self-generating set 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). These 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, 13]. 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 1. 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 notation once and for all.

Definition 1. *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*

$$F = G \cup \{\varphi_0\} = \{\varphi_0, \varphi_1, \ldots, \varphi_r\} \text{ where } \varphi_0 : n \mapsto n$$

and $\varphi_i : n \mapsto k_i n + \ell_i$ with $k_i, \ell_i \in \mathbb{Z}$ and $2 \leq k_1 \leq k_2 \leq \cdots \leq k_r$. Having identity function 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.*

The content of the paper is the following.

1. If we add to F an extra map $\psi : n \mapsto n+\ell$ with $\ell \neq 0$, then the corresponding self-generating set $F^\omega(I)$ is ultimately periodic and therefore k-recognizable for all $k \geq 2$.
2. If all the 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. Let us note that the case where the constants k_i are powers of a fixed base is considered in [8].
3. 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 1 in the affirmative.

The techniques rely on a classical gap theorem; see Theorem 3. We study differences and ratios of consecutive elements in the considered self-generating set.

2 Ultimately Periodic Self-Generating Sets

Theorem 1. *If we add to F in Definition 1 an extra map $\psi : n \mapsto n + \ell$ with $\ell \neq 0$, then the corresponding self-generating set $F^\omega(I)$ is ultimately periodic of period ℓ.*

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

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

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

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

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

Remark 1. In Definition 1 and in what follows, we always assume that all the 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 $F^\omega(\{0\}) = \mathbb{N}$.

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 assume that all the additive constants ℓ_i are non-negative.

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

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

Since $k_i \geq 2$, the constants $\widehat{\ell}_i$ are non-negative for every i.

We show by induction on the number of applied maps that x belongs to $F^\omega(I)$ if and only if $x-y$ belongs to $\widehat{F}^\omega(I-y)$. First, for any $x \in I$, it is obvious that $x-y$ belongs to $I - y$. Recall that $F = G \cup \{\varphi_0\}$ and assume now that $x \in G^m(I)$ for some $m \geq 1$. Otherwise stated, x is obtained by applying m maps in $\{\varphi_1, \ldots, \varphi_r\}$. Therefore there exist $z \in G^{m-1}(I)$ and $i \in \{1, \ldots, r\}$ such that $x = \varphi_i(z)$. By induction hypothesis, $z-y$ belongs to $\widehat{G}^{m-1}(I-y)$ where $\widehat{G} = \widehat{F} \setminus \{\varphi_0\}$. Then 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 $\widehat{G}^m(I - y)$. Assume now that $x - y \in \widehat{G}^m(I - y)$ for some $m \geq 1$. There exist $z \in \widehat{G}^{m-1}(I - y)$ and $i \in \{1, \ldots, r\}$ such that $x - y = \widehat{\varphi}_i(z)$. Then $x = k_i(z + y) + \ell_i = \varphi_i(z + y)$ and by induction hypothesis $z + y$ belongs to $G^{m-1}(I)$. This concludes the proof.

Example 2. Consider the set $X = \mathcal{K}_1$ given in Example 1 and generated from $\{1\}$ by the maps $n \mapsto 2n$ and $n \mapsto 4n - 1$. Applying the constructions 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 $X - 1$.

3 Multiplicatively Dependent Case

In this section, we assume that the multiplicative coefficients k_i appearing in Definition 1 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 2. *Let F given in Definition 1 be such that the multiplicative coefficients $k_1, \ldots, k_r$ are all pairwise multiplicatively dependent. The corresponding self-generating set $X = F^\omega(I)$ is k-recognizable if k_i is a power of k for every $i = 1, 2, \ldots, r$.*

We first sketch a proof relying on Frougny's normalization theorem.

Proof. Assume that $X \subseteq \mathbb{N}$ and that the maps are of the kind $\varphi_i : n \mapsto k^{e_i}\, n + \ell_i$ with $e_1 \geq 1$ for all $i \in \{1, \ldots, r\}$. Let $n = \varphi_{i_m}(\varphi_{i_{m-1}}(\cdots \varphi_{i_1}(a) \cdots))$ 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 $I \cup \{0, \ell_1, \ldots, \ell_r\} \subset \mathbb{Z}$. One can notice that $\mathrm{val}_k(w) = n$. Apply Proposition 7.1.4 in [13] (see also [7]) and Theorem 4.3.6 in [1] to the language $I\{0^{e_1-1}\ell_1, \ldots, 0^{e_r-1}\ell_r\}^*$ to get the regular language $\mathrm{rep}_k(F^\omega(I))$.

We give below another proof which is independent of Frougny's normalization theorem. It describes a way to build an automaton recognizing the k-ary representations of $F^\omega(I)$. We denote by $\mathbb{Z}_{\geq 0}$ (resp., $\mathbb{Z}_{\leq 0}$) the set of non-negative (resp., non-positive) integers.

Remark 2. The set of non-negative elements (resp., the set of absolute values of non-positive elements) in $F^\omega(I)$ can be obtained from a finite set of non-negative elements. Let $m_\ell = \max\{|\ell_i| \mid i = 1, 2, \ldots, r\}$ and denote by M_ℓ the interval of integers $[\![-m_\ell, m_\ell]\!]$. Define $I_j := F^j(I) \cap M_\ell$ 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_ℓ, then $\varphi_i(n) \notin M_\ell$ 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_ℓ are exactly the ones in I_J. We set $I^+ := (I_J \cup I) \cap \mathbb{Z}_{\geq 0}$ and $I^- := (I_J \cup I) \cap \mathbb{Z}_{\leq 0}$. By the property above, we conclude that

$$F^\omega(I) = (F^\omega(I^+) \cap \mathbb{Z}_{\geq 0}) \cup (F^\omega(I^-) \cap \mathbb{Z}_{\leq 0}).$$

Hence, $F^\omega(I^+) \cap \mathbb{Z}_{\geq 0}$ is obtained from I^+ by considering only non-negative images of the maps φ_i. Let $\overline{F} = \{\varphi_0, \overline{\varphi}_1, \overline{\varphi}_2, \ldots, \overline{\varphi}_r\}$, where $\overline{\varphi}_i : n \mapsto k_i\, n - \ell_i$ for $i = 1, 2, \ldots, r$. Then we have $F^\omega(I^-) \cap \mathbb{Z}_{\leq 0} = -(\overline{F}(-I^-) \cap \mathbb{Z}_{\geq 0})$. Otherwise stated, the negation of the elements in $F^\omega(I^-) \cap \mathbb{Z}_{\leq 0}$ are obtained from $-I^-$ by considering only non-negative images of the maps $\overline{\varphi_i}$.

Proof. From the previous remark, we may assume without loss of generality that I and X are subsets of $\mathbb{N}$. Let k be an integer such that all the coefficients k_i are powers of k. Note that $\mathrm{val}_k^{-1}(n)$ contains all the representations of the integer n in base k over Σ_k^*, including those with leading zeros. We define a non-deterministic finite automaton $\mathcal{A} = (Q, \{q_0\}, \Sigma_k, \Delta, T)$ accepting the reversal of the elements in $\mathrm{val}_k^{-1}(F^\omega(I))$, so we may allow leading zeros in front of the most significant digit. The transition relation Δ is a finite subset of $Q \times \Sigma_k^* \times Q$. If (p, w, q) belongs to Δ, we write $p \xrightarrow{w} q$. An input $x \in \Sigma_k^*$ is accepted if and only if there is a sequence of states $q_0, q_1, \ldots, q_i$ such that $q_i \in T$, x can be factored as $u_1 \cdots u_i$ and $(q_0, u_1, q_1), (q_1, u_2, q_2), \ldots (q_{i-1}, u_i, q_i) \in \Delta$.

Let M be the maximal element in $I \cup \{k_1, \ldots, k_r, |\ell_1|, \ldots, |\ell_r|\}$ and $m = |\mathrm{rep}_k(M)|$. Define $Q = \{q_0\} \cup (\{-1, 0, +1\} \times \Sigma_k^{m+1})$. A state $q = (c, x) \in Q \setminus \{q_0\}$ is final if and only if $c = 0$ and $x \in \mathrm{val}_k^{-1}(I)$. From the initial state q_0, we have all the transitions

$$q_0 \xrightarrow{w} (0, \widetilde{w})$$

where $w \in \Sigma_k^{m+1}$ and $\widetilde{w}$ is the reversal of w. Recall that entries are read in $\mathcal{A}$ the least significant digit first, that is from right to left. This explains why we consider the reversals in the encoding. From each state $Q \setminus \{q_0\}$ there are transitions corresponding to the maps φ_i, $i = 1, 2, \ldots, r$. The idea is to guess the sequence of maps $(\varphi_{i_1}, \varphi_{i_2}, \ldots, \varphi_{i_m})$ that was used to obtain the integer corresponding to the input belonging to $\mathrm{val}_k^{-1}(F^\omega(I))$ and apply the inverses of these maps in reversed order to get back the representation of one of the initial values in I. The first component of a state $q = (c, x_m x_{m-1} \cdots x_0)$ corresponds to a carry bit and the second component represents the last $m+1$ digits of a number, x_0 being the least significant one. We show how to simulate the multiplications and additions in the successive applications of the affine functions φ_i using only the carry bit c and the digits $x_m x_{m-1} \cdots x_0$.

Consider first a state $p = (0, x_m x_{m-1} \cdots x_0)$ and φ_i-transitions, where $\varphi_i : n \mapsto k_i\, n + \ell_i$ and $\ell_i \geq 0$. For the inverse of φ_i, we want to subtract ℓ_i and then divide by $k_i = k^t$ for some positive integer t. We do this using the classical paper-and-pencil method as illustrated in Figure 1(a), where $\mathrm{val}_k(y_m y_{m-1} \cdots y_0) = \ell_i$

$$\begin{array}{rr} & x x_m x_{m-1} \cdots x_1 x_0 \\ - & y_m y_{m-1} \cdots y_1 y_0 \\ \hline & z_m z_{m-1} \cdots z_1 z_0 \end{array}$$

(a) subtraction

$$\begin{array}{rr} & x_m x_{m-1} \cdots x_1 x_0 \\ + & y_m y_{m-1} \cdots y_1 y_0 \\ \hline & z z_m z_{m-1} \cdots z_1 z_0 \end{array}$$

(b) addition

Fig. 1. The paper-and-pencil subtraction and addition.

and $x = 1$, if $\mathrm{val}_k(x_m x_{m-1} \cdots x_0) < \ell_i$ and $x = 0$, otherwise. Note that, by the definition of m, we have $y_m = 0$. Hence, a "carry" bit x might be needed only if $x_m = 0$. Multiplying an integer n by k^t corresponds to adding t zeros at the end of the k-ary representation of n. Hence, if φ_i is the correct guess, $z_m z_{m-1} \cdots z_0$ should have at least t zeros as suffix. If this is not the case, we choose to have no φ_i-transitions starting from p. If $x = 0$, then φ_i-transitions are of the form

$$p \xrightarrow{w} (0, \widetilde{w} z_m \cdots z_t), \tag{3}$$

where w is any word over Σ_k of length t. If $x = 1$, then we have two cases depending on the form of $w \in \Sigma_k^t$:

1. If $w = 0^t$, then the transition is

$$p \xrightarrow{w} \left(-1, (k-1)^t z_m \cdots z_t\right), \tag{4}$$

   where the first component -1 indicates that a carry was needed in a "previous" subtraction and it must be borrowed from the first non-zero digit of the input that will be read in the future.
2. Otherwise $\widetilde{w} = vu0^s$, where $s < t$, $u \in \{1, 2, \ldots, k-1\}$ and $v \in \Sigma_k^{t-s-1}$, then the transition is

$$p \xrightarrow{w} (0, v(u-1)(k-1)^s z_m \cdots z_t). \tag{5}$$

   Here the carry $x = 1$ was borrowed from u and no carry is postponed to future calculations.

Consider next a state $p = (0, x_m x_{m-1} \cdots x_0)$ and φ_i-transitions, where $\varphi_i : n \mapsto k_i\, n + \ell_i$ and $\ell_i < 0$. Instead of subtraction, we consider now addition by the paper-and-pencil method where $\text{val}_k(y_m y_{m-1} \cdots y_0) = |\ell_i|$. This is illustrated in Figure 1(b). Note that since $y_m = 0$ by the definition of m, a carry $z = 1$ can occur only if $x_m = k-1$. As above, the φ_i-transitions exist only if the last t digits of $z_m z_{m-1} \cdots z_0$ are zeros. This holds also for any transition considered in the sequel. If $z = 0$, then we have the transitions of the form (3). If $z = 1$ we have again two cases depending on the digits of $w \in \Sigma_k^t$:

1. If $w = (k-1)^t$, then the carry is shifted to future calculations.

$$p \xrightarrow{w} \left(+1, 0^t z_m \cdots z_t\right). \tag{6}$$

2. If $\widetilde{w} = vu(k-1)^s$, where $s < t$, $u \in \{0, 1, \ldots, k-2\}$ and $v \in \Sigma_k^{t-s-1}$, then

$$p \xrightarrow{w} (0, v(u+1)0^s z_m \cdots z_t). \tag{7}$$

   Here the carry is added to the digit u and no carry is postponed to future calculations.

Secondly, consider a state of the form $p = (-1, x_m x_{m-1} \cdots x_0)$ and assume that $\ell_i \geq 0$. The carry component -1 means that we have borrowed a carry in a subtraction and after the subtraction we have read only zeros, which have been turned into digits $k-1$. Otherwise, if non-zero digits were read, there would be no longer a carry -1. Hence, we can be sure that $x_m = k-1$ in Figure 1(a), and consequently, we have $x = 0$, since $y_m = 0$. This is important, since it means that no "new" carry is borrowed. Again, assume that $z_{t-1} \cdots z_0 = 0^t$. If $w = 0^t$, then the transition is of type (4). Otherwise, the transitions are of type (5).

If $\ell_i < 0$, then we perform addition as in Figure 1(b) and assume $z_{t-1} \cdots z_0 = 0^t$. If $z = 0$ and $w = 0^t$, no new carries occur and again the transition is of type (4). If $z = 0$ and $w \neq 0^t$, then the transitions are of type (5). If $z = 1$, then this positive carry and the negative carry borrowed in a previous calculation annihilate each other. Hence, the transition is of type (3).

Finally, consider the states of the form $p = (+1, x_m x_{m-1} \cdots x_0)$. The carry component $+1$ means that we obtained a carry in an addition and after the addition we have read only digits $k-1$, which have been turned into zeros. Otherwise, if a digit $u \neq k-1$ were read, it would have been turned into $u+1$ and there would be no longer a carry $+1$. Hence, we conclude that $x_m = 0$ in Figures 1(a) and 1(b). If $\ell_i \geq 0$, then consider Figure 1(a) and assume $z_{t-1} \cdots z_0 = 0^t$. If $x = 0$ and $w = (k-1)^t$, then the transition is of type (6). If $x = 0$ and $w \neq (k-1)^t$, then the transitions are of type (7). If $x = 1$, then the negative and positive carry annihilate each other and the transitions are of type (3). Assume now that $\ell_i < 0$ and $z_{t-1} \cdots z_0 = 0^t$. In Figure 1(b) no new carry $z = 1$ can occur, since both $x_m = 0$ and $y_m = 0$. Hence, we have only two cases. If $w = (k-1)^t$, then the transition is of type (6). Otherwise, it is of the type (7).

If $n = \varphi_{i_m}(\varphi_{i_{m-1}}(\cdots \varphi_{i_1}(a) \cdots))$ for some $a \in I$, then using the above transitions and k-ary representations, we are able to correctly simulate the calculation $n \mapsto \varphi_{i_m}^{-1}(n) \mapsto \varphi_{i_{m-1}}^{-1}(\varphi_{i_m}^{-1}(n)) \mapsto \cdots \mapsto a$ as long as the k-ary representation of n given as input contains enough leading zeros. However, we may fix this by replacing the set of final states T by an enlarged set T'. A state $q' \in Q$ belongs to T' if there exists a path with label 0^t, $t \geq 0$, from q' to some state $q \in T$. Hence, with the modified final states the automaton $\mathcal{A}$ accepts all the reversals of the words in $\mathrm{val}_k^{-1}(F^\omega(I))$. On the other hand, it cannot accept any other word. Namely, for any word w accepted by $\mathcal{A}$ there is a sequence $(\varphi_{i_1}, \varphi_{i_2}, \ldots, \varphi_{i_m})$ such that (1) holds for $n = \mathrm{val}_k(w)$. It is well-known that any non-deterministic finite automaton can be turned into a DFA, e.g., by the subset construction. Hence, $F^\omega(I)$ is k-recognizable.

Remark 3. 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 4 described below can be applied, then $F^\omega(I)$ contains arbitrarily large gaps.

4 Multiplicatively Independent Case

In this section, our aim is to show that $F^\omega(I) \subseteq \mathbb{N}$ given in Definition 1 is not recognizable in any base $k \geq 2$ 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 $k \geq 2$, we use the following result from [5], see also Eilenberg's book [6, Chapter V, Theorem 5.4].

Theorem 3 (Gap Theorem). *Let $k \geq 2$. 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 1 contains arbitrarily large gaps.

Theorem 4. *Let $X = F^\omega(I)$ be a self-generating subset of $\mathbb{N}$ given in Definition 1. 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 $\{\varphi_1, \varphi_2, \ldots, \varphi_r\}$ 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 only on g. Since $k_1^{n_1} k_2^{n_2} \cdots k_r^{n_r}$ divides K^n, we get

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

The set $F^n(I)$ contains exactly the integers obtained by at most n applications of maps in F. For any interval of integers $[\![N, N + K^n - 1]\!]$ where $N > \max F^n(I)$, the elements in 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.

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 if every subinterval of I contains an element of S.

Theorem 5. *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 5 can be found in [1, Section 2.5] or [9]. As an easy consequence of the previous theorem, we obtain the following result.

Corollary 1. *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 5, there exists 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 6. *For any self-generating set $X = F^\omega(I)$ given in Definition 1 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 $\varphi_1 : n \mapsto k_1 n + \ell_1$, $\varphi_2 : n \mapsto k_2 n + \ell_2$, 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 1, we may also assume that ℓ_1 and ℓ_2 are non-negative. Moreover, with Remark 2, we may consider that both I and X are subsets of $\mathbb{N}$.

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

$$\varphi_j^n(x) = k_j^n x + \ell_j \sum_{i=0}^{n-1} k_j^i = (x + b_j)\, k_j^n - b_j. \tag{8}$$

Let $m \geq 0$ and x_i, x_{i+1} be two consecutive elements belonging to the set X_m. By Corollary 1, there exist infinitely many positive integers p and q such that

$$\frac{\varphi_2^p(a)}{k_1^q} = \frac{(a + b_2)k_2^p - b_2}{k_1^q} \in \left[x_{i+1} + b_1 - \frac{3}{4}(x_{i+1} - x_i), x_i + b_1 + \frac{3}{4}(x_{i+1} - x_i) \right].$$

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[1] 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

$$\frac{3}{4}\frac{k_1^t(x_{i+1}-x_i)}{\varphi_1^t(x_i)} + \frac{b_1 k_1^{t-q}}{\varphi_1^t(x_i)} = \frac{3}{4}\frac{k_1^t(x_{i+1}-x_i)}{\varphi_1^t(x_i)} + \frac{1}{k_1^q}\frac{b_1 k_1^t}{(x_i+b_1)k_1^t - b_1}. \tag{9}$$

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 a 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

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

Thus, by (9), we obtain

$$\frac{x_{j+1}-x_j}{x_j} < \frac{11\,(x_{i+1}-x_i)}{12\,x_i}. \tag{10}$$

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 (10) 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,

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

[1] $c - \varphi_1^q(x_i) = \frac{1}{4}k_1^q(x_{i+1}-x_i) + b_1$ 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 7. *Let $X = F^\omega(I)$ be given in Definition 1. 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 4, we have $D_X = \infty$ and, by Theorem 6, we have $R_X = 1$. Thus, Theorem 3 implies that X is not k-recognizable for any k.

As a corollary, we have solved the conjecture presented in [2].

Corollary 2. *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 1 is not k-recognizable for any $k \geq 2$.*

Proof. This follows directly from Theorem 7. Namely, if k_1 and k_2 are multiplicatively independent, then $k_1 \geq 2$ and $k_2 \geq 3$ and $k_1^{-1} + k_2^{-1} \leq 1/2 + 1/3 = 5/6 < 1$.

References

1. J.-P. Allouche, J. Shallit, *Automatic Sequences: Theory, Applications, Generalizations*, Cambridge University Press, Cambridge, (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 base-dependence of sets of numbers recognizable by finite automata, *Math. Systems Theory* **3** (1969), 186–192.
5. A. Cobham, Uniform tag sequences, *Math. Systems Theory* **6** (1972), 186–192.
6. S. Eilenberg, *Automata, languages, and machines*, Vol. A., Pure and Applied Mathematics, Vol. **58**, Academic Press, New York, (1974).
7. C. Frougny, Representations of numbers and finite automata, *Math. Systems Theory* **25** (1992), 37–60.
8. D. Garth, A. Gouge, Affinely self-generating sets and morphisms, *J. Integer Seq.* **10** (2007), Article 07.1.5.
9. G.H. Hardy, E.M. Wright, *Introduction to the Theory of Numbers*, Oxford Univ. Press, (1985).
10. C. Kimberling, A self-generating set and the golden mean, *J. Integer Seq.* **3** (2000), Article 00.2.8.
11. C. Kimberling, Affinely recursive sets and orderings of languages, *Discrete Math.* **274** (2004), 147–159.
12. P.B.A. Lecomte, M. Rigo, Numeration systems on a regular language, *Theory Comput. Syst.* **34** (2001), 27–44.
13. M. Lothaire, *Algebraic Combinatorics on Words*, Encyclopedia of Mathematics and its Applications, **90**, Cambridge University Press, Cambridge, (2002).