Automaticity and Parikh-collinear Morphisms

Michel Rigo[0000−0001−7463−8507], Manon Stipulanti[0000−0002−2805−2465], and Markus A. Whiteland⋆[0000−0002−6006−9902]

Department of Mathematics, University of Liège, Liège, Belgium
{m.rigo,m.stipulanti,mwhiteland}@uliege.be

Abstract. Parikh-collinear morphisms have recently received a lot of attention. They are defined by the property that the Parikh vectors of the images of letters are collinear. We first show that any fixed point of such a morphism is automatic. Consequently, we get under some mild technical assumption that the abelian complexity of a binary fixed point of a Parikh-collinear morphism is also automatic, and we discuss a generalization to arbitrary alphabets. Then, we consider the abelian complexity function of the fixed point of the Parikh-collinear morphism $0 \mapsto 010011$, $1 \mapsto 1001$. This 5-automatic sequence is shown to be aperiodic, answering a question of Salo and Sportiello.

Keywords: Automatic sequences · Morphic words · Abelian complexity · Automated theorem proving · Walnut

1 Introduction

Let us briefly introduce the main concept of this paper. Details and precise definitions are given in Section 2. Let A be a finite alphabet. A morphism $f\colon A^* \to B^*$ is *Parikh-collinear* if the Parikh vectors $\Psi(f(a))$, $a \in A$, are collinear (or pairwise $\mathbb{Z}$-linearly dependent).

Example 1. The morphism $f\colon \{0,1\}^* \to \{0,1\}^*$ with $0 \mapsto 010011$, $1 \mapsto 1001$ is Parikh-collinear, as is the morphism $g\colon \{0,1,2\}^* \to \{0,1,2\}^*$ defined by $0 \mapsto 012$, $1 \mapsto 102201$, $2 \mapsto \varepsilon$. Any *Parikh-constant* morphism [26] (i.e., the Parikh vectors of images of letters are equal) is Parikh-collinear.

Parikh-collinear morphisms have received some attention in recent years. Cassaigne *et al.* characterized Parikh-collinear morphisms as those morphisms that map all words to words with bounded abelian complexity [8]. These morphisms also provide infinite words with interesting properties with respect to the so-called k-binomial equivalence $\sim_k$. Two words $u, v \in A^*$ are *k-binomially equivalent* if $\binom{u}{x} = \binom{v}{x}$, for all $x \in A^*$ with $|x| \leq k$. Recall that a binomial coefficient $\binom{u}{x}$ counts the number of times x occurs as a subword of u. The *k-binomial complexity function* of an infinite word $\mathbf{x}$ introduced in [26] is defined as $\mathsf{b}^{(k)}_{\mathbf{x}}\colon \mathbb{N} \to \mathbb{N}$, $n \mapsto \#(\mathcal{L}_n(\mathbf{x})/\sim_k)$, i.e., length-$n$ factors in $\mathbf{x}$ are counted up

⋆ Corresponding author

to k-binomial equivalence. (Here $\mathtt{b}_{\mathbf{x}}^{(1)}$ is the usual abelian complexity function [14].) For a survey on abelian properties of words, see [15]. In a recent work[1], we showed that a morphism is Parikh-collinear iff it maps all words with bounded k-binomial complexity to words with bounded $(k+1)$-binomial complexity (for all k) [27]. Thus any fixed point of a Parikh-collinear morphism has a bounded k-binomial complexity for all k (and thus a bounded abelian complexity).

In our computer experiments and research presentations, the first few values of these bounded complexities on various examples suggested that the abelian complexity of a fixed point of a Parikh-collinear morphism might be ultimately periodic. This question was asked independently by Ville Salo when the third author visited Turku University and by Andrea Sportiello when the second author gave a presentation at "Journées Combinatoires de Bordeaux" 2023.

Our contributions. Even though Parikh-collinear morphisms are generally non-uniform, we show in Section 3 that their fixed points are k-automatic for $k = \sum_{b\in A} |f(b)|_b$. The result itself, in fact, can be considered folklore. The (constructive) proof given here was inspired by [7,9] but, as pointed out by the referees, it can be seen as a consequence of [1, Thm. 2.2 or 4.2], the former of which is itself a reformulation of a result of Dekking [12] (we note however, that the statements speak of non-erasing morphisms). It is well known that there exist infinite sequences that are the fixed points of non-uniform morphisms, but not k-automatic for any k, and that every k-automatic sequence is the image of a fixed point of a non-uniform morphism [3]. A recent preprint [18] completely characterizes those uniformly recurrent (i.e., every factor occurs infinitely often and with bounded gaps) morphic words that are automatic.

Making use of Büchi's theorem and first-order logic, we prove in Section 4 that under some mild assumptions the abelian complexity of a binary fixed point of a Parikh-collinear morphism is automatic. This result supports the expectation of the abelian complexity of a k-automatic sequence to exhibit regular behavior in base-k (cf. Rigo's conjecture [22], named after the first author of this paper). We however recall that abelian properties in general cannot be handled with such a formalism (for instance, Schaeffer showed that the set of occurrences of abelian squares in the paperfolding word is not k-automatic for any k [28]).

Coming back to Salo and Sportiello's question; a positive answer to it would suggest that our second main result is trivial in the sense that any ultimately periodic sequence is automatic over any base. In Section 5, we propose an answer to their question by considering the abelian complexity of the fixed point $\mathbf{w} = 0100111001\cdots$ of the morphism $f\colon \{0,1\}^* \to \{0,1\}^*$ given in Example 1: we show that its abelian complexity is aperiodic. We provide two proofs of this result: we make use of, on the one hand, classical combinatorial arguments, and on the other, the software `Walnut` [21,30] to illustrate, in this case, that the study of abelian complexities is amenable to the first-order logic formalism in practice.

Finally, we explain in Section 6 how the mild assumptions considered here may be alleviated and the result generalized to arbitrary alphabets.

[1] A long version is available at https://doi.org/10.48550/arXiv.2201.04603.

2 Preliminaries

We recall some basics needed for the paper. More can be found in [2,24,25]. We let A^* (resp., $A^{\mathbb{N}}$) denote the set of finite (resp., infinite) words over A equipped with concatenation. Infinite words are written in bold unless otherwise stated. We let ε denote the empty word. The length of the word w is denoted by $|w|$ and the number of occurrences of a letter a in w is denoted by $|w|_a$. The *Parikh vector* of a word $w \in A^*$ is defined as the vector $\Psi(w) = (|w|_a)_{a\in A} \in \mathbb{N}^A$. An infinite word is *ultimately periodic* if it can be written as $uvvv\cdots$ where $u, v \in A^*$ and $v \neq \varepsilon$. If it is not the case, it is said to be *aperiodic*. For an infinite word $\mathbf{x}$ and an integer $n \geq 0$, we let $\mathcal{L}(\mathbf{x})$ and $\mathcal{L}_n(\mathbf{x})$ respectively denote the set of factors of $\mathbf{x}$ and that of length-n factors of $\mathbf{x}$.

A *morphism* is a map $f\colon A^* \to B^*$, where A, B are alphabets, such that $f(xy) = f(x)f(y)$ for all $x, y \in A^*$. The morphism f is *prolongable* on the letter $a \in A$ if $f(a) = ax$ for some $x \in A^*$ and $\lim_{n\to\infty} |f^n(x)| = \infty$. We let $f^\omega(a) := \lim_{n\to\infty} f^n(a)$ denote the fixed point of f starting with a; an infinite word $\mathbf{x}$ is called *pure morphic* if $\mathbf{x} = f^\omega(a)$ for some such f and a. An infinite word is *morphic* if it can be written as $g(f^\omega(a))$, where g and f are morphisms such that f is prolongable on a. For a given morphism $f\colon A^* \to A^*$, a letter $a \in A$ is called *mortal* if $f^n(a) = \varepsilon$ for some $n \geq 1$. If a is not mortal, we call it *immortal*. For an integer $k \geq 1$, a morphism $f\colon A^* \to B^*$ is *k-uniform* if $|f(a)| = k$ for all letters $a \in A$. A 1-uniform morphism is called a *coding*.

For a morphism $f\colon A^* \to A^*$ let $M_f \in \mathbb{N}^{A\times A}$ denote the associated matrix defined by $(M_f)_{a,b} = |f(b)|_a$ for $a, b \in A$. Then we have $\Psi(f(w)) = M_f\Psi(w)$ for all words $w \in A^*$. A morphism f is *primitive* if the corresponding matrix M_f is primitive, that is, there exists a power of M_f having only positive entries.

Two words $u, v \in A^*$ are *abelian equivalent* if they are obtained as permutations of each other, and we write $u \sim_1 v$. The latter relation is called the *abelian equivalence* already introduced by Erdős [14]. The *abelian complexity function* of an infinite word $\mathbf{x}$ is defined as $\mathsf{a}_{\mathbf{x}}\colon \mathbb{N} \to \mathbb{N}$, $n \mapsto \#(\mathcal{L}_n(\mathbf{x})/{\sim_1})$.

Introduced by Cobham [11], automatic words have several equivalent definitions. See [2] for a comprehensive presentation. Let $k \geq 2$ be an integer. An infinite word $\mathbf{x}$ is *k-automatic* if it is the image, under a coding, of a fixed point of a k-uniform morphism; they form a subclass of morphic words. They can also be characterized by means of first-order logic. Consider the structure $\langle \mathbb{N}, +, V_k\rangle$, where $V_k(0) := 1$ and, for all $n \geq 1$, $V_k(n)$ is the largest power of k dividing n. A set $X \subseteq \mathbb{N}^d$ is *k-definable* if it can be defined by a first-order formula with d free variables within $\langle \mathbb{N}, +, V_k\rangle$. As a consequence of a theorem of Büchi [5], an infinite word $\mathbf{x}$ is k-automatic iff for every letter a, the set of positions where a occurs in $\mathbf{x}$ is k-definable. Again, we refer the reader to [2,4,10,25].

3 Automaticity of Parikh-collinear Fixed Points

We focus on Parikh-collinear morphisms f that are prolongable on some letter. For any mortal letter $m \in A$ of f we have that $f(m) = \varepsilon$. Indeed, if $f(b) \neq \varepsilon$

for a letter $b \in A$ and f is prolongable on a, then $f(b)$ contains an occurrence of a by Parikh-collinearity. Therefore b cannot be mortal. We shall always assume that the underlying alphabet is minimal. More precisely, we assume that each letter of A appears in $f(a)$ for any immortal letter a. Again, Parikh-collinearity implies that the minimal alphabet is well-defined. For an immortal letter $a \in A$, for each $b \in A$ there exists $r_b \in \mathbb{Q}$ such that $\Psi(f(b)) = r_b \Psi(f(a))$.

Lemma 2. *Let $f\colon A^* \to A^*$ be Parikh-collinear and $a \in A$ be immortal. Then $\Psi(f(a))$ is an eigenvector of M_f associated with the eigenvalue $\sum_{b\in A} |f(b)|_b$.*

Proof. For any word $w \in A^*$, we have

$$M_f\Psi(w) = \sum_{b\in A} |w|_b M_f \Psi(b) = \sum_{b\in A} |w|_b \Psi(f(b)) = \sum_{b\in A} (|w|_b r_b) \cdot \Psi(f(a)).$$

With the choice $w = f(a)$, we find that $\Psi(f(a))$ is an eigenvector of M_f associated with the eigenvalue $\sum_{b\in A} |f(a)|_b r_b = \sum_{b\in A} |f(b)|_b$. □

When speaking of *the eigenvalue* of a Parikh-collinear morphism f, we mean the eigenvalue $\sum_{b\in A} |f(b)|_b$ of M_f. As M_f has rank 1, the only other eigenvalue is 0 (with multiplicity $\#A - 1$).

Remark 3. If f is prolongable on a letter a, then the eigenvalue k of f is at least 2. Indeed, $f(a)$ must contain at least two occurrences of immortal letters (the first letter a and another one, say b). If $b = a$ then $k \geq |f(a)|_a \geq 2$, otherwise $|f(a)|_a$, $|f(b)|_b \geq 1$ by Parikh-collinearity and again $k \geq 2$.

In what follows, for an infinite word $\mathbf{x}$, a letter $a \in A$ is called *left deterministic* (resp., *right deterministic*) if it is always preceded (resp., followed) by a unique letter $b \in A$ in $\mathbf{x}$. In particular, the first letter of $\mathbf{x}$ is not left deterministic.

Lemma 4. *Let $\mathbf{x} \in A^{\mathbb{N}}$ be a fixed point of the morphism $f\colon A^* \to A^*$. Assume further that, for distinct letters $a_1, \ldots, a_\ell$, such that $a_1 \cdots a_\ell \in \mathcal{L}(\mathbf{x})$, a_i is right deterministic for $i < \ell$, and a_j is left deterministic for $j \geq 2$. Factorize $f(a_1 \cdots a_\ell) = u_1 \cdots u_\ell$, with $u_i \in A^*$. Then $g(\mathbf{x}) = \mathbf{x}$, where g is the morphism defined by $a_i \mapsto u_i$ for all $i \in \{1, \ldots, \ell\}$, and $c \mapsto f(c)$ for all other $c \in A$.*

Proof. Writing $w = a_1 \cdots a_\ell$, we may factorize $\mathbf{x} = x_0 w x_1 \cdots w x_n \cdots$, where $x_i \in (A \setminus \{a_1, \ldots, a_\ell\})^*$ (and if w appears only a finite number n of times, then $x_n \in (A \setminus \{a_1, \ldots, a_\ell\})^{\mathbb{N}}$). But we now have $g(w) = f(w)$ and $g(x_i) = f(x_i)$ for all i by construction, whence $g(\mathbf{x}) = f(\mathbf{x}) = \mathbf{x}$. □

Notice that if g above is prolongable on the first letter a of $\mathbf{x}$, then $\mathbf{x} = g^\omega(a)$.

The reader may wish to consult Example 6 for illustrations of the constructions provided the proof the following theorem.

Theorem 5. *Let $f\colon A^* \to A^*$ be a Parikh-collinear morphism prolongable on a letter $a \in A$. Then $f^\omega(a)$ is k-automatic for the eigenvalue k of f.*

Proof. Let us write $\mathbf{x} := f^\omega(a)$. Recall from Lemma 2 that $\Psi(f(b)) = r_b\Psi(f(a))$ is either the zero vector or an eigenvector of the adjacency matrix M_f associated with the eigenvalue k. It therefore follows that $|f^2(b)| = k \cdot |f(b)|$ for each $b \in A$.

As an intermediate step, we construct a pure morphic word $\mathbf{y}$ which can be mapped, by a coding, to $\mathbf{x}$. To this end, for each immortal letter $b \in A$ define the letters $\widehat{b}_1, \ldots, \widehat{b}_{|f(b)|}$ and let $B = \{\widehat{b}_i \colon b \in A \text{ is immortal}, i = 1, \ldots, |f(b)|\}$. Let $\tau\colon B^* \to A^*$ be the coding defined by $\tau(\widehat{b}_i) = c$ if the ith letter of $f(b)$ equals c. Set, for each $c \in A$, the word $\mathfrak{w}_c := \widehat{c}_1 \cdots \widehat{c}_{|f(c)|}$ if c is immortal, otherwise set $\mathfrak{w}_c = \varepsilon$. It is now evident that $\tau(\mathfrak{w}_c) = f(c)$. Define then the morphism $\varphi\colon B^* \to B^*$ as follows: for each immortal $b \in A$ and $i \in \{1, \ldots, |f(b)|\}$, we set $\varphi(\widehat{b}_i) = \mathfrak{w}_{\tau(\widehat{b}_i)}$. Notice now that $\tau(\varphi(\widehat{b}_i)) = \tau(\mathfrak{w}_{\tau(\widehat{b}_i)}) = f(\tau(\widehat{b}_i))$, and so $\tau \circ \varphi = f \circ \tau$ as morphisms $B^* \to A^*$. Moreover $\tau \circ \varphi^n = f \circ \tau \circ \varphi^{n-1} = \ldots = f^n \circ \tau$. Since a is immortal we get $\tau(\varphi^n(\widehat{a}_1)) = f^n(a)$ for all $n \geq 0$, yielding $\tau(\varphi^\omega(\widehat{a}_1)) = \mathbf{x}$. We set $\mathbf{y} = \varphi^\omega(\widehat{a}_1)$.

Next we define a k-uniform morphism $g\colon B^* \to B^*$ for which $g^\omega(\widehat{a}_1) = \mathbf{y}$ (recall that $k \geq 2$ by Remark 3). This implies that $\mathbf{x}$ is k-automatic, as then $\mathbf{x} = \tau(g^\omega(\widehat{a}_1))$. Fix an immortal letter $b \in A$ (we will proceed iteratively for each of them). The letters $\widehat{b}_i$, with $i \in \{1, \ldots, |f(b)|\}$, satisfy the assumptions of Lemma 4 in $\mathbf{y}$. Notice that $\tau(\varphi(\mathfrak{w}_b)) = \tau(\varphi^2(\widehat{b}_1)) = f^2(\tau(\widehat{b}_1)) = f^2(b)$, whence $|\varphi(\mathfrak{w}_b)| = k \cdot |f(b)|$. Factorize $\varphi(\mathfrak{w}_b) = u_1 \cdots u_{|\mathfrak{w}_b|}$, each of the words u_i having length k. By Lemma 4, we have $\mathbf{y} = g_b(\mathbf{y})$, where $g_b\colon B^* \to B^*$ is defined by $g_b(\widehat{b}_i) = u_i$ and $g_b(\mathfrak{a}) = \varphi(\mathfrak{a})$ for $\mathfrak{a} \in B \setminus \{\widehat{b}_i \colon i = 1, \ldots, |f(b)|\}$ (note that $|g_b(\widehat{b}_i)| = k$ for each $i \in \{1, \ldots, |f(b)|\}$). We may repeat this operation on g_b (sequentially) for all the other immortal letters $c \in A$, and the resulting morphism g is k-uniform for which $g(\mathbf{y}) = \mathbf{y}$. Clearly g is prolongable on $\widehat{a}_1$, which suffices for the claim. $\square$

Example 6. Let f be defined by $0 \mapsto 012$; $1 \mapsto 112002$; $2 \mapsto \varepsilon$, and let $\mathbf{x} = f^\omega(0)$. Here $k = 3$. We thus have $B = \{\widehat{0}_i, \widehat{1}_j \colon i = 1, 2, 3,\ j = 1, \ldots, 6\}$, and τ is defined by $\widehat{0}_1, \widehat{1}_4, \widehat{1}_5 \mapsto 0$; $\widehat{0}_2, \widehat{1}_1, \widehat{1}_2 \mapsto 1$; $\widehat{0}_3, \widehat{1}_3, \widehat{1}_6 \mapsto 2$. We then define φ by

$$\widehat{0}_1, \widehat{1}_4, \widehat{1}_5 \mapsto \mathfrak{w}_0 = \widehat{0}_1\widehat{0}_2\widehat{0}_3;\ \widehat{0}_2, \widehat{1}_1, \widehat{1}_2 \mapsto \mathfrak{w}_1 = \widehat{1}_1\widehat{1}_2\widehat{1}_3\widehat{1}_4\widehat{1}_5\widehat{1}_6;\ \widehat{0}_3, \widehat{1}_3, \widehat{1}_6 \mapsto \varepsilon.$$

Factorizing $\varphi(\mathfrak{w}_0)$ and $\varphi(\mathfrak{w}_1)$, respectively, as

$$\begin{aligned}
\varphi(\mathfrak{w}_0) &= \widehat{0}_1\widehat{0}_2\widehat{0}_3 \cdot \widehat{1}_1\widehat{1}_2\widehat{1}_3 \cdot \widehat{1}_4\widehat{1}_5\widehat{1}_6\\
\varphi(\mathfrak{w}_1) &= \widehat{1}_1\widehat{1}_2\widehat{1}_3 \cdot \widehat{1}_4\widehat{1}_5\widehat{1}_6 \cdot \widehat{1}_1\widehat{1}_2\widehat{1}_3 \cdot \widehat{1}_4\widehat{1}_5\widehat{1}_6 \cdot \widehat{0}_1\widehat{0}_2\widehat{0}_3 \cdot \widehat{0}_1\widehat{0}_2\widehat{0}_3,
\end{aligned}$$

we define g by $\widehat{0}_1, \widehat{1}_5, \widehat{1}_6 \mapsto \widehat{0}_1\widehat{0}_2\widehat{0}_3$; $\widehat{0}_2, \widehat{1}_1, \widehat{1}_3 \mapsto \widehat{1}_1\widehat{1}_2\widehat{1}_3$; and $\widehat{0}_3, \widehat{1}_2, \widehat{1}_4 \mapsto \widehat{1}_4\widehat{1}_5\widehat{1}_6$, which gives $\tau(g^\omega(\widehat{0}_1)) = \mathbf{x}$.

One notes that there are redundant letters (i.e., they have equal images under both τ and $g \circ \tau$); we find a simpler morphism h by identifying them: $0 \mapsto 012$; $1 \mapsto 134$; $2 \mapsto 506$; $3 \mapsto 506$; $4 \mapsto 134$; $5 \mapsto 506$; $6 \mapsto 012$, with which $\tau'(h^\omega(0)) = \mathbf{x}$, where τ' is defined by $0, 5 \mapsto 0$; $1, 3 \mapsto 1$; $2, 4, 6 \mapsto 2$.

4 Automaticity of the Abelian Complexity

In this section, we consider a Parikh-collinear morphism $f\colon A \to A^*$ prolongable on a letter $a \in A$ and its fixed point $\mathbf{x} = f^\omega(a)$. We set $k = \sum_{b\in A} |f(b)|_b$ to be the eigenvalue of f. For all $n \geq 0$, we let $\mathrm{pref}_n(\mathbf{x})$ be the length-n prefix of $\mathbf{x}$. The corresponding *cutting set* is defined by

$$\mathsf{CS}_{f,a} := \{|f(\mathrm{pref}_n(\mathbf{x}))| \colon\ n \geq 0\}. \tag{1}$$

This set simply provides the indices where blocks $f(b)$, with $b \in A$, start in a factorization of $\mathbf{x}$ of the form $f(x_0)f(x_1)f(x_2)\cdots$.

To help the reader, we now start a running example throughout the section.

Running Example 7. Consider $f\colon 0 \mapsto 010011,\ 1 \mapsto 1001$ and $\mathbf{w} = f^\omega(0)$. The first five elements in $\mathsf{CS}_{f,0}$ are 0,6, 10, 16, and 22.

Given any integer i, we look for two consecutive integers: the next and previous elements found in C around i.

Lemma 8. *Let $C = \{0 = c_0 < c_1 < c_2 < \cdots\}$ be an infinite k-definable subset of $\mathbb{N}$. The functions* $\mathrm{ne} : \mathbb{N} \to \mathbb{N}$ *mapping i to the least element in C greater than or equal to i and* $\mathrm{pr} : \mathbb{N} \to \mathbb{N}$ *mapping i to the greatest element in C less than i, are k-definable. (We set* $\mathrm{pr}(0) = 0$.*)*

Proof. Since C is k-definable by some formula φ_C, i.e., $\varphi_C(j)$ holds iff $j \in C$. The functions of the statement are then defined by

$$\begin{aligned}
\mathrm{ne}(i) = j &\equiv\ \varphi_C(j) \wedge (i \leq j) \wedge (\forall k)(\varphi_C(k) \wedge i \leq k) \to (j \leq k),\\
\mathrm{pr}(i) = j &\equiv\ \varphi_C(j) \wedge (j < i) \wedge (\forall k)(\varphi_C(k) \wedge k < i) \to (k \leq j).
\end{aligned} \qquad \square$$

It is easy to see that the abelian complexity of a binary word $\mathbf{x}$, fixed point of a (non-erasing) Parikh-collinear morphism, is given by

$$\mathsf{a}_{\mathbf{x}}(n) = \frac{1}{r+r'}\left(\max_{x\in\mathcal{L}_n(\mathbf{x})}(r'|x|_1 - r|x|_0) - \min_{x\in\mathcal{L}_n(\mathbf{x})}(r'|x|_1 - r|x|_0)\right) + 1, \tag{2}$$

where $|f(a)|_1 = \frac{r}{q}|f(a)|$ and $|f(a)|_0 = \frac{r'}{q}|f(a)|$ for both $a \in \{0,1\}$. Notice that we have $r + r' = q$ and $r'|f(a)|_1 = r|f(a)|_0$.

Observe that, since f is Parikh-collinear, each full block $f(b)$ occurring in a factor x has no contribution to the value of $r'|x|_1 - r|x|_0$ in the above formula. This observation is at the core of our reasoning. A similar strategy can be found in [16] where for factors of the Thue–Morse word, one can disregard full images.

Let x be a length-n factor of $\mathbf{x}$. Then there exist two letters $b, b' \in A$, a proper suffix s of $f(b)$, a factor u, and a proper prefix p of $f(b')$ such that $x = sf(u)p$. Due to the previous observation what matters to compute the abelian classes is therefore the total contribution of both p and s. Note that there are only finitely many such proper prefixes and suffixes, which is enough to be encoded into a formula. In addition, empty prefixes or suffixes have no contribution.

Running Example 9. The word $11|1001|010011|010 = 11f(10)010$ is a factor of length 15 of $\mathbf{w}$ occurring at position 4. For this morphism f, we have $r = r' = 1$ and $q = 2$ with the notation of Eq. (2). We have $\mathrm{pr}(4) = 0$ and $\mathrm{ne}(4) = 6$. The prefix 11 has a contribution of 2 to $r'|x|_1 - r|x|_0$ and the suffix 010 a contribution of -1. In Table 1, the contribution (symbolized by c) of each suffix (symbolized by s) and each prefix (symbolized by p) of $f(0)$ and $f(1)$ is given.

s	c	s	c	p	c	p	c
1	1	1	1	01001	-1	100	-1
11	2	01	0	0100	-2	10	0
011	1	001	-1	010	-1	1	1
0011	0			01	0		
10011	1			0	-1		

Table 1. Contributions (c) to $r'|\cdot|_1 - r|\cdot|_0$ of the suffixes (s) and prefixes (p) of $f(0)$ and $f(1)$ respectively.

For the sake of presentation, we give the next result for binary words under a mild assumption on $\mathsf{CS}_{f,a}$. In Section 6 we discuss how to generalize it.

Theorem 10. *Let $\mathbf{x} = f^\omega(0) \in \{0,1\}^{\mathbb{N}}$ be a binary fixed point of a Parikh-collinear morphism. If $\mathsf{CS}_{f,a}$ is k-automatic then $\mathsf{a}_{\mathbf{x}}(n)$ is k-automatic.*

Proof. We can assume that $|f(0)| \neq |f(1)|$. Because otherwise f is Parikh-constant and Guo *et al.* [17, Thm. 3] showed that the Parikh-constant image of a k-automatic sequence has k-automatic abelian complexity.

Let $x = sf(u)p$ be the factor of length n occurring in position i where, as usual, s is a proper suffix of $f(b)$, u is a factor, and p is a proper prefix of $f(b')$ for some letters b, b'. Let r, r' be the constants as in Eq. (2).

If s is non-empty, i.e., if $i \neq \mathrm{ne}(i)$, the letter b is uniquely determined by $\mathrm{ne}(i) - \mathrm{pr}(i)$. This difference is equal to $|f(b)|$ and by assumption, distinct letters have images with distinct length. The length of s is given by $\mathrm{ne}(i) - i$. Consequently, $i, \mathrm{ne}(i), \mathrm{pr}(i)$ determine if $s = \varepsilon$ or, a unique suffix s with a specific contribution to $r'|x|_1 - r|x|_0$.

Similarly, the length of p is zero if $i + n - 1 = \mathrm{ne}(i + n - 1)$, i.e., $i + n - 1$ belongs to $\mathsf{CS}_{f,a}$. If p is non-empty, the letter b' is uniquely determined by $\mathrm{ne}(i+n-1) - \mathrm{pr}(i+n-1)$. Otherwise, the length of p is given by $i+n-1-\mathrm{pr}(i+n-1)$. Consequently, $i+n-1, \mathrm{ne}(i+n-1), \mathrm{pr}(i+n-1)$ determine if $p = \varepsilon$ or, a unique prefix p with a specific contribution to $r'|x|_1 - r|x|_0$. Since there are finitely many prefixes and suffixes, we may define a function $\mathrm{contr} : (i, n) \mapsto r'|x|_1 - r|x|_0$ where x is the length-n factor occurring at position i in $\mathbf{x}$. It is a finite disjunction of terms of the form

$$(\mathrm{ne}(i) - \mathrm{pr}(i) = z_1 \wedge \mathrm{ne}(i) - i = z_2 \wedge \mathrm{ne}(i+n-1) - \mathrm{pr}(i+n-1) = z_3$$
$$\wedge\, i+n-1-\mathrm{pr}(i+n-1) = z_4) \rightarrow \mathrm{contr}(i,n) = \lambda,$$

where λ depends on the 4-tuple (z_1, z_2, z_3, z_4). The complete formula should also take into account the test for empty p or s. We illustrate this on our running example below.

The conclusion follows: the maximum in Eq. (2) is defined as a function of n by $\max_{x\in\mathcal{L}_n(\mathbf{x})}(r'|x|_1 - r|x|_0) = M \equiv (\exists i)(\mathrm{contr}(i,n) = M) \wedge (\forall j)(\mathrm{contr}(i,n) \leq M)$. One proceeds similarly with the minimum and we thus see that $a_{\mathbf{x}}(n)$ is k-definable. Simply recall that multiplication by a constant is definable in Presburger arithmetic and by Büchi's theorem, the claim follows. □

Running Example 11. With $i = 4$ and $n = 15$, we find $\mathrm{pr}(i) = 0$, $\mathrm{ne}(i) = 6$, so $z_1 = 6 = |f(0)|$, and $i+n-1 = 18$, $\mathrm{pr}(18) = 15$, $\mathrm{ne}(18) = 21$, so $z_3 = 6 = |f(0)|$. Since $z_2 = \mathrm{ne}(i)-i = 2$, we know that we have the suffix of $f(0)$ of length 2 with contribution 2 (recall Table 1). Similarly, since $z_4 = i+n-1-\mathrm{pr}(i+n-1) = 3$, we known that we have the prefix of $f(0)$ of length 3 with contribution -1. So the 4-tuple $(6,2,6,3)$ is associated with the total contribution $\frac{3-(-1)}{2} - 1 = 1$.

We now show on our running example that $\mathsf{CS}_{f,0}$ is k-definable.

Running Example 12. The factor 0100 occurs in $\mathbf{w}$ only in position corresponding to prefixes of blocks $f(0)$. Since $\mathbf{w}$ is k-automatic for $k = 5$, there are two k-definable unary relations $\varphi_0(i)$ and $\varphi_1(i)$ which are true iff 0 and 1 respectively occur in $\mathbf{w}$ at position i. So the index i is the starting position of a block $f(0)$ iff $\Psi_0(i) \equiv \varphi_0(i) \wedge \varphi_1(i+1) \wedge \varphi_0(i+2) \wedge \varphi_0(i+3)$ holds. In a similar way, every block $f(1)$ is preceded in $\mathbf{w}$ by a letter 1, so the factor 1100 occurs in position $i-1$ iff an occurrence of $f(1)$ starts in position i. Therefore the index i is the starting position of a block $f(1)$ iff $\Psi_1(i) \equiv \varphi_1(i-1) \wedge \varphi_1(i) \wedge \varphi_0(i+1) \wedge \varphi_0(i+2)$ holds. From this, we deduce that $\mathsf{CS}_{f,0}$ is k-definable by $\{i\colon \Psi_0(i) \vee \Psi_1(i)\}$ and we can therefore apply the above theorem.

5 The Fixed Point of $0 \mapsto 010011$, $1 \mapsto 1001$

We consider the pure morphic word $\mathbf{w} = f^\omega(0)$, where $f\colon 0 \mapsto 010011$, $1 \mapsto 1001$ (see Running Example 7). We give two different proofs of the following result.

Proposition 13. *The abelian complexity function* $\mathsf{a}_{\mathbf{w}}$ *of* $\mathbf{w}$ *is aperiodic.*

5.1 Proving Aperiodicity the Old-fashioned Way

It is straightforward to see that $\mathbf{w}$ is aperiodic and uniformly recurrent. Furthermore, the frequency of 0 exists and equals $1/2$ as this is the case for the images of both letters. Hence, for each $n \geq 0$, there exist length-n factors u, v for which $|u|_0 > n/2$ and $|v|_0 < n/2$ (see, for example, [23, Lem. 4.3]). The next lemma shows that $\mathsf{a}_{\mathbf{w}}$ is bounded by 4.

Lemma 14. *We have* $\mathsf{a}_{\mathbf{w}}(2n+1) \leq 4$ *and* $\mathsf{a}_{\mathbf{w}}(2n) = 3$ *for all* $n \geq 0$*. More precisely, for a factor* $x \in \mathcal{L}(\mathbf{w})$*, we have* $\lfloor |x|/2 \rfloor - 1 \leq |x|_0 \leq \lceil |x|/2 \rceil + 1$*.*

Proof. The claim can be verified for factors of length at most 4 straightforwardly. For a factor x of length $n \geq 5$ we may write $x = sf(u)p$, where s is a proper suffix and p is a proper prefix of the image of a letter. Notice that $|f(u)|_0 = |f(u)|/2$, hence $|x|_0 = (|x| - |ps|)/2 + |ps|_0$. Notice here that $|ps|$ and $|x|$ have the same parity, as $|f(u)|$ has even length. By inspecting all combinations of p and s (with total length odd and even separately), we find the more precise claim.

To see that $\mathsf{a}_{\mathbf{w}}(2n) = 3$, all three values n and $n \pm 1$, for the number of 0s, are attained, as was asserted in the beginning of this section. □

The proof of the following lemma is a tedious exercise using properties established above. We omit the proof due to space constraints.

Lemma 15. *The abelian complexity function of* $\mathbf{w}$ *satisfies the following.*

1. *We have* $\mathsf{a}_{\mathbf{w}}(5n) = \mathsf{a}_{\mathbf{w}}(n)$ *for all* $n \geq 0$*.*
2. *For each* $j \in \{1,2,3,4\}$ *and for all* $n \geq 1$*, we have* $\mathsf{a}_{\mathbf{w}}(5n+j) \geq 3$*.*

Proof sketch. 1. The claim essentially follows from two correspondences: We have $\min_{u \in \mathcal{L}_n(\mathbf{w})}\{|u|_0\} < \lfloor n/2 \rfloor$ iff $\max_{x \in \mathcal{L}_{5n}(\mathbf{w})}\{|x|_0\} > \lceil 5n/2 \rceil$. Similarly, $\max_{u \in \mathcal{L}_n(\mathbf{w})}\{|u|_0\} > \lceil n/2 \rceil$ iff $\min_{x \in \mathcal{L}_n(\mathbf{w})}\{|x|_0\} < \lfloor 5n/2 \rfloor$. For the proof one takes a factor $x = sf(u)p$, and computes bounds on $|u|$. Inspecting all possibilities (utilizing Lemma 14) gives the claimed properties.

2. When $5n+j$ is even, Lemma 14 shows that $\mathsf{a}_{\mathbf{w}}(5n+j) = 3$. For each of the cases $j \in \{1,2,3,4\}$ and for each n for which $5n+j$ is odd, we exhibit a factor x of length $5n+j$ for which $|x|_0 \notin \frac{5n+j\pm 1}{2}$. For example, let $u \in \mathcal{L}(\mathbf{w})$ be of odd length $n-1$ such that $1u0 \in \mathcal{L}(\mathbf{w})$ and $|u|_0 = \lfloor \frac{n-1}{2} \rfloor$. Such a factor always exists: there exist two factors v, v' and an index i such that $|v|_0 = \frac{n-2}{2}$, $|v'|_0 = \frac{n}{2}$, v begins at position i in $\mathbf{w}$, and v' begins at position $i+1$ in $\mathbf{w}$. The only possibility is that v begins with 1 and v' ends with 0; u may be then taken to be the overlap of v and v'. We have $|f(u)| = 4 \cdot |u| + 2 \cdot |u|_0 = 5n-6$ since n is even. Take $x = 001f(u)0100$. Then $|x| = 5n+1$ and $|x|_0 = \frac{|f(u)|}{2} + 5 = \frac{5n-6}{2} + 5 = \lceil \frac{5n+1}{2} \rceil + 1$. □

Proof of Proposition 13. We show that $\mathsf{a}_{\mathbf{w}}(n) = 2$ iff $n = 5^m$ for some $m \geq 0$. The value 2 is unattainable for all even n by Lemma 14. Let n be odd and write $n = 5^m \cdot k$ with $\gcd(k,5) = 1$; then $\mathsf{a}_{\mathbf{w}}(5^m \cdot k) = \mathsf{a}_{\mathbf{w}}(k)$ by repeated application of Lemma 15(1). Write $k = 5\ell + j$ for some $1 \leq j \leq 4$ and $\ell \geq 0$. If $\ell = 0$ and $j = 1$ (so $n = 5^m$) we find $\mathsf{a}_{\mathbf{w}}(n) = 2$, and if $j \geq 2$ we find $\mathsf{a}_{\mathbf{w}}(n) = 3$ by inspection. For $\ell \geq 1$ Lemma 15(2) shows $\mathsf{a}_{\mathbf{x}}(n) \geq 3$. □

5.2 Proof of Proposition 13 via `Walnut` in a `Shell`

Originally designed by Mousavi [21], `Walnut` is a free and publicly available software that allows to prove theorems and properties for the particular family of automatic words through the lens of first-order logic; see Shallit's recent book [30] for a comprehensive presentation. We shall use `Walnut` to prove Proposition 13.

The procedure associated to Theorem 5 (complemented with identifying redundant letters and renaming) gives the 5-uniform morphism $g\colon 0 \mapsto 01023$, $1 \mapsto 14501$, $2 \mapsto 10102$, $3 \mapsto 31450$, $4 \mapsto 45010$, and $5 \mapsto 10231$, and the coding $\tau\colon 0,2,5 \mapsto 0$; $1,3,4 \mapsto 1$, for which $\mathbf{w} = \tau(g^\omega(0))$, and which can be used to define $\mathbf{w}$ in `Walnut` conveniently by:

```
morphism g "0->01023 1->14501 2->10102 3->31450 4->45010 5->10231";
morphism tau "0->0 1->1 2->0 3->1 4->1 5->0";
promote G g;
image W tau G;
```

Following the procedure described in Section 4, we construct a DFAO generating $\mathtt{a}_{\mathbf{w}}$. We first recall basic syntax for `Walnut`. The letters `A` and `E` are abbreviations for $\forall$ ("for all") and $\exists$ ("there exists") respectively. The letters `?msd_5` indicates that an expression is to be evaluated using base-5 representations. The symbol `@` specifies the value of an automatic sequence. The symbols `&`, `|`, `~`, and `=>` are logical "and" , "or", negation and implication respectively.

As in Eq. (1), we define the cutting set (details to understand the formula can be found in Running Example 12):

```
def CS "?msd_5 (W[i]=@0 & W[i+1]=@1 & W[i+2]=@0 & W[i+3]=@0)
                  | (i>0 & W[i-1]=@1 & W[i]=@1 & W[i+1]=@0 & W[i+2]=@0)";
```

As in Lemma 8, we define the maps ne and pr:

```
   def ne "?msd_5 k >= i & $CS(k) & (A j (j >= i & $CS(j)) => j>=k)";
   def pr "?msd_5 k < i & $CS(k) & (A j (j < i & $CS(j)) => j<=k)";
```

Next we describe the contribution of factors of $\mathbf{w}$ starting at position i for all $i \geq 0$. Using the notation from Section 4, for the length-n factor x of $\mathbf{w}$ starting at position i, we write $x = sf(u)p$ where $b, b' \in \{0,1\}$, s is a proper suffix of $f(b)$, u is a word, and p is a proper prefix of $f(b')$. We know that x only contributes to the abelian complexity $\mathtt{a}_{\mathbf{w}}$ via s and p. Based on Table 1, the following two binary predicates correspond to integer pairs (c, i) such that the contribution from the suffix s of $f(b)$ (resp., prefix p of $f(b')$) starting at position i (resp., ending at position $i-1$) equals $c-2$. We notably shift true value by 2 as the variable domain in `Walnut` is $\mathbb{N}$.

```
def suffContr "?msd_5 (c=1 & ($ne(i,i+3) & $pr(i,i-1)))
   | (c=2 & ($ne(i,i) | $ne(i,i+4)) | ($ne(i,i+2) & $pr(i,i-2)))
   | (c=3 & ($ne(i,i+1) | ($ne(i,i+3) & $pr(i,i-3)) | $ne(i,i+5)))
   | (c=4 & ($ne(i,i+2) & $pr(i,i-4)))";
def prefContr "?msd_5 (c=0 & ($pr(i,i-4) & $ne(i,i+2)))
 | (c=1 & ($pr(i,i-3) | ($pr(i,i-1) & $ne(i,i+5)) | $pr(i,i-5)))
 | (c=2 & ($ne(i,i) | $pr(i,i-2))) | (c=3 & ($pr(i,i-1) & $ne(i,i+3)))";
```

The ternary predicate $\mathtt{contr}(c, i, n)$ is satisfied when the prefix contribution at position i and the suffix contribution at position $i+n$ total up to $c-4$.

```
def contr "?msd_5 Ed,e $suffContr(d,i) & $prefContr(e,i+n) & (e+d=c)";
```

Here recall that $d-2$ (resp., $e-2$) is the contribution of p at position i (resp., s ending at position $i+n-1$). (In the end we will call on this predicate only for $n > 4$, so one does not need to worry what happens when, e.g., $n = 0$.)

The following predicates accept the pairs (c, n), where $c-4$ is the max. (resp., min.) contribution for the prefixes and suffixes p, s for length-n factors.

```
def MaxContr "?msd_5 (Ei $contr(c,i,n)) & (Ai,d $contr(d,i,n) => d<=c)";
```

```
def MinContr "?msd_5 (Ei $contr(c,i,n)) & (Ai,d $contr(d,i,n) => d>=c)";
```

We may now define the abelian complexity function as a binary predicate $\texttt{abComp}(a, n)$. We have $\mathtt{a_w}(n) = a$ iff $2 \cdot a = \max(|u|_1 - |u|_0) - \min(|u|_1 - |u|_0) + 2$. Adding 4 to both min and max simultaneously does not change the sum, so we may use `MaxContr` and `MinContr` in their places (interpreted as functions).

```
def abComp "?msd_5 (n=0 & a=1) | (n=1 & a=2) | ((n>=2 & n<=4) & a =3)
      | (n>4 & (Ec,d $MaxContr(c,n) & $MinContr(d,n)  & 2*a = c-d + 2))";
```

Next we generate $\mathtt{a_w}$ as a 5-automatic word, so we define the following predicates, for $z = 1, 2, 3, 4$, which recognize the base-5 representations of integers n such that $\mathtt{a_w}(n) = z$:

```
def abCompz "?msd_5 $abComp(z,n)";
```

To express $\mathtt{a_w}$ as an automatic sequence, we combine the predicates `abCompz` into one; in `Walnut`, this is done with the command

```
combine abCompW abComp1=1 abComp2=2 abComp3=3 abComp4=4;
```

`Walnut` then returns the 9-state deterministic finite automaton with output reading base-5 representations generating $\mathtt{a_w}$ in Fig. 1.

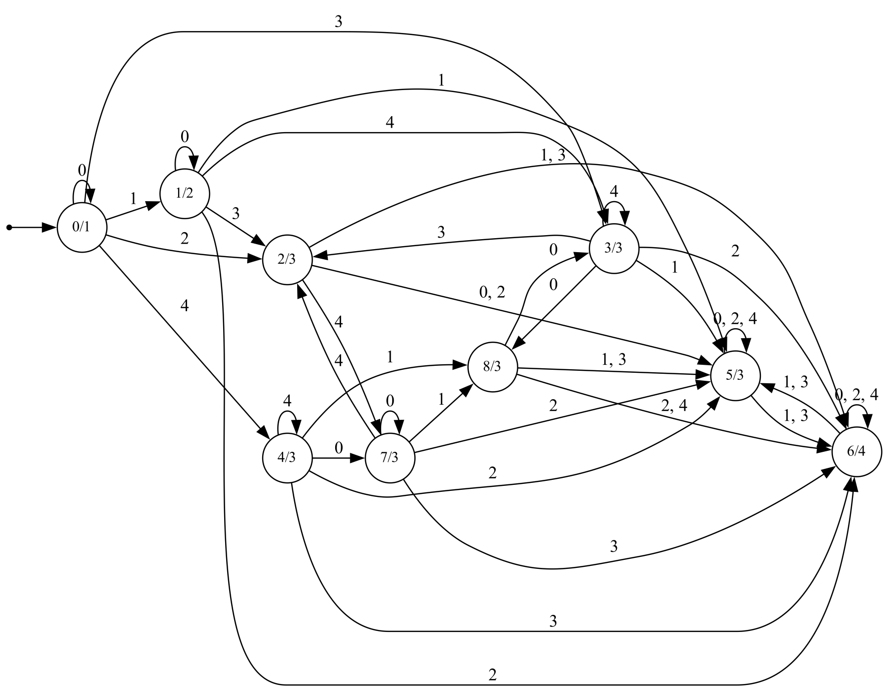

Fig. 1. The minimal deterministic finite automaton with output reading base-5 representations generating the abelian complexity $\mathtt{a_w}$.

We now have an automatic way of proving Proposition 13. One may perform the following query about the aperiodicity of $\mathtt{a_w}$, for which `Walnut` replies `TRUE`:

```
eval isAper "?msd_5 ~(Ei,p (p>0) & Aj ((j>=i)
                              => (abCompW[j] = abCompW[j+p])))";
```

6 Final remarks

Our first main result Theorem 5 is constructive, given the morphism f and letter $a \in A$ for which $\mathbf{x} = f^{\omega}(a)$. Our second main result Theorem 10 holds for those binary $\mathbf{x}$ for which the set $\mathsf{CS}_{f,a}$ is k-definable. We briefly sketch a plan for relaxing the assumptions of Theorem 10 to obtain: *Assuming f above is non-erasing, $\mathsf{a}_{\mathbf{x}}$ is k-automatic (for the eigenvalue k of f), and a DFAO defining it can be effectively constructed.* The main idea is the following: we show that, for the sequence $(p_n)_{n\geq 0}$ of length-n prefixes $p_n = \mathrm{pref}_n(\mathbf{x})$ and any letter $b \in A$, the sequence $(|p_n|_b)_{n\geq 0}$ is k-*synchronized*, namely, there is an automaton which accepts the tuples $(\mathrm{rep}_k(n), \mathrm{rep}_k(|p_n|_b))$ accordingly padded (where $\mathrm{rep}_k(n)$ denote the base-k expansion of n). See again [30, §10] for an excellent introduction. We may then invoke the result and methods of Shallit [29] to conclude.

As mentioned above, with k the eigenvalue of f, Theorem 5 shows that $\mathbf{x}$ is effectively k-automatic. Since f is assumed non-erasing, it is primitive. Therefore, by a result of Mossé [20,19], there exists a constant L such that, from any position i of $\mathbf{x}$, one can determine the indices $\mathrm{pr}(i)$ and $\mathrm{ne}(i)$ of $\mathsf{CS}_{f,a}$ by inspecting the factor $\mathbf{x}[i-L...i+L]$. Moreover, given f and a, the constant L is effectively computable by [13]. Therefore, $\mathsf{CS}_{f,a}$ is effectively k-definable.

For a prefix p_n of the form $f(x_n)t_n$ where x_n is a prefix of $\mathbf{x}$ such that $\mathrm{pr}(n) = |f(x_n)|$, we have that $|f(x_n)|_b = \frac{r}{q}|f(x)|$, where $r = |f(a)|_b$ and $q = |f(a)|$. Hence $q|p_n|_b = r|f(x_n)| + q|t_n|_b$. Define the function $F(n) = |p_n|_b$. Then

$$y = F(n) \equiv \exists m, z \colon (\mathrm{pr}(n) = m) \wedge (q \cdot (y - z) = r \cdot m) \wedge (|\mathbf{x}[m...n]|_b = z).$$

Recall that r and q are constants, and notice that $|\mathbf{x}[\mathrm{pr}(n)...n]|$ attains finitely many values, whence the last check $(|\mathbf{x}[m...n]|_b = z)$ can be expressed by a first-order logical formula with indexing into $\mathbf{x}$. Hence the function $F(n)$ is defined by a first-order logical formula for which [30, Thm. 10.2.3] applies, and is therefore synchronized.

We remark that Mossé's recognizability result referred to in the above has recently been extended to deal with erasing morphisms [6]. One could hope that the results therein are useful to obtain the automaticity of the abelian complexity of any infinite word generated by a Parikh-collinear morphism.

Acknowledgments M. Stipulanti and M. Whiteland are supported by the FNRS Research grants 1.B.397.20F and 1.B.466.21F, respectively. We thank Julien Leroy for fruitful discussions on morphic words and pointing out useful references. We thank A. Sportiello and V. Salo for asking the question leading to this paper. We also thank the reviewers for their suggestions.

References

1. Allouche, J.P., Dekking, M., Queffélec, M.: Hidden automatic sequences. Comb. Theory **1**(#20) (2021). https://doi.org/10.5070/C61055386
2. Allouche, J.P., Shallit, J.: Automatic sequences: Theory, applications, generalizations. Cambridge University Press, Cambridge (2003). https://doi.org/10.1017/CBO9780511546563
3. Allouche, J.P., Shallit, J.: Automatic sequences are also non-uniformly morphic. In: Discrete mathematics and applications, Springer Optim. Appl., vol. 165, pp. 1–6. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-55857-4_1
4. Bruyère, V., Hansel, G., Michaux, C., Villemaire, R.: Logic and p-recognizable sets of integers. Bull. Belg. Math. Soc. Simon Stevin **1**(2), 191–238 (1994), http://projecteuclid.org/euclid.bbms/1103408547, journées Montoises (Mons, 1992)
5. Büchi, J.R.: Weak second-order arithmetic and finite automata. Z. Math. Logik Grundlagen Math. **6**, 66–92 (1960). https://doi.org/10.1002/malq.19600060105
6. Béal, M.P., Perrin, D., Restivo, A.: Recognizability of morphisms. Ergodic Theory and Dynamical Systems p. 1–25 (2023). https://doi.org/10.1017/etds.2022.109
7. Cassaigne, J., Nicolas, F.: Quelques propriétés des mots substitutifs. Bull. Belg. Math. Soc. Simon Stevin **10**(suppl.), 661–676 (2003), http://projecteuclid.org/euclid.bbms/1074791324
8. Cassaigne, J., Richomme, G., Saari, K., Zamboni, L.Q.: Avoiding Abelian powers in binary words with bounded Abelian complexity. Int. J. Found. Comput. S. **22**(4), 905–920 (2011). https://doi.org/10.1142/S0129054111008489
9. Charlier, É., Leroy, J., Rigo, M.: Asymptotic properties of free monoid morphisms. Linear Algebra and its Applications **500**, 119–148 (2016). https://doi.org/10.1016/j.laa.2016.02.030
10. Charlier, É., Rampersad, N., Shallit, J.: Enumeration and decidable properties of automatic sequences. Internat. J. Found. Comput. Sci. **23**(5), 1035–1066 (2012). https://doi.org/10.1142/S0129054112400448
11. Cobham, A.: Uniform tag sequences. Math. Systems Theory **6**, 164–192 (1972). https://doi.org/10.1007/BF01706087
12. Dekking, F.M.: The spectrum of dynamical systems arising from substitutions of constant length. Z. Wahrscheinlichkeitstheor. Verw. Geb. **41**, 221–239 (1978). https://doi.org/10.1007/BF00534241
13. Durand, F., Leroy, J.: The constant of recognizability is computable for primitive morphisms. Journal of Integer Sequences **20**(#17.4.5) (2017), https://cs.uwaterloo.ca/journals/JIS/VOL20/Leroy/leroy4.html
14. Erdős, P.: Some unsolved problems. Michigan Math. J. **4**, 291–300 (1958)
15. Fici, G., Puzynina, S.: Abelian combinatorics on words: A survey. Comput. Sci. Rev. **47**, Paper No. 100532 (2023). https://doi.org/10.1016/j.cosrev.2022.100532
16. Goč, D., Rampersad, N., Rigo, M., Salimov, P.: On the number of abelian bordered words (with an example of automatic theorem-proving). Internat. J. Found. Comput. Sci. **25**(8), 1097–1110 (2014). https://doi.org/10.1142/S0129054114400267
17. Guo, Y.J., Lü, X.T., Wen, Z.X.: On the boundary sequence of an automatic sequence. Discrete Math. **345**(1) (2022). https://doi.org/10.1016/j.disc.2021.112632
18. Krawczyk, E., Müllner, C.: Automaticity of uniformly recurrent substitutive sequences (2023). https://doi.org/10.48550/arXiv.2111.13134
19. Mossé, B.: Reconnaissabilité des substitutions et complexité des suites automatiques. Bulletin de la Société Mathématique de France **124**(2), 329–346 (1996). https://doi.org/10.24033/bsmf.2283

20. Mossé, B.: Puissance de mots et reconnaissabilité des points fixes d'une substitution. Theoretical Computer Science **99**(2), 327–334 (1992). https://doi.org/10.1016/0304-3975(92)90357-L
21. Mousavi, H.: Automatic theorem proving in Walnut (2016). https://doi.org/10.48550/ARXIV.1603.06017
22. Parreau, A., Rigo, M., Rowland, E., Vandomme, É.: A new approach to the 2-regularity of the ℓ-abelian complexity of 2-automatic sequences. The Electronic Journal of Combinatorics **22**(#P1.27) (2022). https://doi.org/10.37236/4478
23. Puzynina, S., Whiteland, M.A.: Abelian closures of infinite binary words. Journal of Combinatorial Theory, Series A **185**, 105524 (2022). https://doi.org/10.1016/j.jcta.2021.105524
24. Rigo, M.: Formal languages, automata and numeration systems. Vol. 1. ISTE, London; John Wiley & Sons, Inc., Hoboken, NJ (2014), introduction to combinatorics on words, With a foreword by Valérie Bethé
25. Rigo, M.: Formal languages, automata and numeration systems. Vol. 2. Networks and Telecommunications Series, ISTE, London; John Wiley & Sons, Inc., Hoboken, NJ (2014), applications to recognizability and decidability, With a foreword by Valérie Berthé
26. Rigo, M., Salimov, P.: Another generalization of abelian equivalence: binomial complexity of infinite words. Theor. Comput. Sci. **601**, 47–57 (2015). https://doi.org/10.1016/j.tcs.2015.07.025
27. Rigo, M., Stipulanti, M., Whiteland, M.A.: Binomial complexities and Parikh-collinear morphisms. In: Diekert, V., Volkov, M.V. (eds.) Developments in Language Theory - 26th International Conference, DLT 2022, Tampa, FL, USA, May 9-13, 2022, Proceedings. Lecture Notes in Computer Science, vol. 13257, pp. 251–262. Springer (2022). https://doi.org/10.1007/978-3-031-05578-2_20
28. Schaeffer, L.: Deciding Properties of Automatic Sequences. Master's thesis, Univ. of Waterloo (2013), https://uwspace.uwaterloo.ca/handle/10012/7899
29. Shallit, J.: Abelian complexity and synchronization. INTEGERS: Electronic Journal of Combinatorial Number Theory (#A.36) (2021), http://math.colgate.edu/~integers/v36/v36.pdf
30. Shallit, J.: The Logical Approach to Automatic Sequences: Exploring Combinatorics on Words with Walnut. London Mathematical Society Lecture Note Series, Cambridge University Press (2022). https://doi.org/10.1017/9781108775267