The reflection complexity of sequences over finite alphabets

Jean-Paul Allouche John M. Campbell Jeffrey Shallit
Manon Stipulanti

Abstract

In combinatorics on words, the well-studied factor complexity function $\rho_{\mathbf{x}}$ of a sequence $\mathbf{x}$ over a finite alphabet counts, for any nonnegative integer n, the number of distinct length-n factors of $\mathbf{x}$. In this paper, we introduce the *reflection complexity* function $r_{\mathbf{x}}$ to enumerate the factors occurring in a sequence $\mathbf{x}$, up to reversing the order of symbols in a word. We introduce and prove results on $r_{\mathbf{x}}$ regarding its growth properties and relationship with other complexity functions. We prove that if $\mathbf{x}$ is k-automatic, then $r_{\mathbf{x}}$ is computably k-regular, and we use the software `Walnut` to evaluate the reflection complexity of automatic sequences, such as the Thue–Morse sequence. We prove a Morse–Hedlund-type result characterizing eventually periodic sequences in terms of their reflection complexity, and we deduce a characterization of Sturmian sequences. Furthermore, we investigate the reflection complexity of episturmian, $(s+1)$-dimensional billiard, and Rote sequences. There are still many unanswered questions about this measure.

Keywords: reflection complexity, factor complexity, reversal, automatic sequence, Sturmian sequence, episturmian sequence, Rote sequence, Morse-Hedlund theorem, `Walnut`.
MSC: 05A05, 11B85, 68R15

1 Introduction

The discipline of combinatorics on words continues to be burgeoning as a relatively new and interdisciplinary area of mathematics. In this regard, the significance of combinatorics on words within disciplines such as theoretical computer science leads us to explore variants and generalizations of fundamental objects and constructions involved within the field. If $\mathbf{x}$ is an infinite sequence over a finite alphabet (see Section 1.1 for precise definitions), natural problems that arise in the combinatorial study of $\mathbf{x}$ and in the context of computer science-based problems concern the behavior of factors of $\mathbf{x}$. Writing $\mathbb{N} = \{1, 2, \ldots\}$ and $\mathbb{N}_0 = \{0, 1, \ldots\}$, we are led to consider the *factor complexity function* $\rho_{\mathbf{x}} \colon \mathbb{N}_0 \to \mathbb{N}$ which maps $n \geq 0$ to the number of distinct factors of $\mathbf{x}$ of length n. (The term *factor* refers to any contiguous block occurring in $\mathbf{x}$.)

Variations on this definition can be considered as a measure of how "complicated" a sequence is. For example, the *abelian complexity function* of $\mathbf{x}$ counts the number of factors of $\mathbf{x}$ of a given length and up to permutation of symbols, so that two factors $u = u(1)u(2)\cdots u(m)$ and $v = v(1)v(2)\cdots v(n)$ are *abelian equivalent* if $m = n$ and if there exists a bijection $\sigma \colon \{1, 2, \ldots, m\} \to \{1, 2, \ldots, m\}$ such that $u(1)u(2)\cdots u(m) = v(\sigma(1))v(\sigma(2))\cdots v(\sigma(m))$. Similarly, the *cyclic complexity function* $c_{\mathbf{x}}$ introduced in 2017 [29] is equal to the number of length-n factors of $\mathbf{x}$, up to equivalence under cyclic permutations. The abelian and cyclic complexity functions lead us to introduce a *reflection complexity function* on sequences via an invariant property involving permutations of symbols, by analogy with the abelian and cyclic complexity functions.

In addition to the factor, abelian, and cyclic complexity functions indicated above, there have been many different complexity functions on sequences that have been previously introduced. In this regard, we highlight the following "complexities": additive complexity [9], arithmetical complexity [12], gapped binomial complexity [77], k-abelian complexity [55], k-binomial complexity [76], Kolmogorov complexity [56], Lempel–Ziv complexity [58], Lie complexity [17], linear complexity (see the survey by Niederreiter [69]), maximal pattern complexity [54], maximum order complexity [44], (initial) (non-)repetitive complexity [25, 65], opacity complexity [8], open and closed complexity [71], palindrome complexity [4], periodicity complexity [64], privileged complexity [73], relational factor complexity [32], span and leftmost complexity [30], string attractor profile complexity (implicitly defined in [81]; also see [30]), and window complexity [31]. Also see the references in the surveys in [1, 46, 47, 48]. The complexity function $r_{\mathbf{x}}$ defined below does not seem to have been previously studied, but may be thought of as natural in terms of its relationships with automatic sequences such as the Thue–Morse sequence. To begin with, we require the equivalence relation $\sim_r$ defined below.

Definition 1. Let m, n be nonnegative integers. Given a finite word $u = u(1)u(2)\cdots u(m)$, its *reversal* is the word $u^R = u(m)u(m-1)\cdots u(1)$, that is, $u^R(i) = u(m+1-i)$ for all $i \in \{1, \ldots, m\}$. A *palindrome* is a word that is equal to its reversal. Two finite words u and v are *reflectively equivalent* if $u = v$ or $u = v^R$. We denote this equivalence relation by $u \sim_r v$.

Example 2. Over the alphabet $\{\mathtt{a}, \mathtt{b}, \ldots, \mathtt{z}\}$, the English word `reward` is reflectively equivalent to `drawer`, while `deed`, `kayak`, and `level` are palindromes.

Definition 3. Let $\mathbf{x}$ be a sequence. The *reflection complexity function* $r_{\mathbf{x}} \colon \mathbb{N}_0 \to \mathbb{N}$ of $\mathbf{x}$ maps every $n \geq 0$ to the number of distinct length-n factors of $\mathbf{x}$, up to equivalence by $\sim_r$.

Example 4. Let

$$\mathbf{t} = 011010011001011010010110011010011 \cdots \tag{1}$$

denote the Thue–Morse sequence, where the nth term in (1) for $n \geq 1$ is defined as the number of ones, modulo 2, in the base-2 expansion of n. The initial terms of the integer sequence $(r_{\mathbf{t}}(n))_{n \geq 0}$ are such that

$$(r_{\mathbf{t}}(n))_{n \geq 0} = 1, 2, 3, 4, 6, 6, 10, 10, 13, 12, 16, 16, 20, 20, 22, \ldots. \tag{2}$$

We see that $r_{\mathbf{t}}(2) = 3$, for example, since there are 3 length-2 factors of $\mathbf{t}$, up to reflection complexity, i.e., the factors 00 and 11 and one member of the equivalence class $\{01, 10\}$, with respect to $\sim_r$.

The integer sequence in (2) is not currently included in the On-Line Encyclopedia of Integer Sequences [85], which suggests that our notion of "reflection complexity" is new. See also the work of Krawchuk and Rampersad in [57], which introduced the notion of *cyclic/reversal complexity* for sequences. The evaluation of reflection complexity functions is closely related to the work of Rampersad and Shallit [74], who investigated sequences $\mathbf{x}$ such that all sufficiently long factors w have the property that that w^R is not a factor of $\mathbf{x}$. Also, the evaluation of reflection complexities for sequences is related to the enumeration of palindromes contained in sequences; see, for example, Fici and Zamboni [49].

This paper is organized as follows. In Section 1.1, we introduce the notation and definitions needed for the paper. In Section 2, we give general results on the reflection complexity. In particular, we investigate growth properties and relationships with other complexity functions. In Sections 3, 4, and 5, respectively, we investigate reflection complexities for eventually periodic sequences, Sturmian sequences and generalizations, and reversal-closed and rich sequences. Next, in Section 6, we focus on classical automatic sequences and, with the use of `Walnut`, we prove that the reflection complexity function for automatic sequences is a regular sequence. We also study the reflection complexity for famous automatic sequences such as the Thue–Morse sequence. Finally, some further research directions are considered in Section 7.

1.1 Preliminaries

Generalities. For a general reference on words, we cite [59]. An *alphabet* is a finite set of elements called *letters*. A *word* on an alphabet A is a finite sequence of letters from A. The *length* of a word, denoted between vertical bars, is the number of its letters (counting multiplicities). The *empty word* is the only 0-length word, denoted as ε. For all $n \geq 0$, we let A^n denote the set of all length-n words over A. We let A^* denote the set of words over A, including the empty word and equipped with the concatenation operation. In order to distinguish finite words and infinite sequences, we write the latter in bold. Except for complexity functions, we start indexing words and sequences at 1. A *factor* of a word or a sequence is any of its (finite and contiguous) subblocks. A *prefix* (resp., *s*uffix) is any starting (resp., ending) factor. Given a word w, its nth term is written $w(n)$ for $1 \leq n \leq |w|$. The factor starting at position n and ending at position m with $1 \leq m \leq n \leq |w|$ is written $w[m..n]$. A factor u of a word w over A is *right* (resp., *left*) special if ua and ub (resp., au and bu) are factors of w for some distinct letters $a, b \in A$. We extend these notations to (finite) words as well. A sequence $\mathbf{x}$ is *reversal-closed* if, for any factor w of $\mathbf{x}$, the word w^R is also a factor of $\mathbf{x}$. A sequence $\mathbf{x}$ is *eventually periodic* if there exist finite words u, v, with v nonempty, such that $\mathbf{x} = uv^\omega$ where $v^\omega = vvv\cdots$ denotes the infinite concatenation of the word v. A sequence that is not eventually periodic is said to be *aperiodic*. A sequence is said to be *recurrent* if every factor occurs infinitely many times; it is *uniformly recurrent* if each factor occurs with bounded gaps, i.e., for all factors w, there is some length $m = m(w)$ such that w occurs in every length-m block.

Morphisms. Let A and B be finite alphabets. A *morphism* $f \colon A^* \to B^*$ is a map satisfying $f(uv) = f(u)f(v)$ for all $u, v \in A^*$. In particular, $f(\varepsilon) = \varepsilon$, and f is entirely determined by the images of the letters in A. For an integer $k \geq 2$, a morphism is *k-uniform* if it maps each letter to a length-k word. A 1-uniform morphism is called a *coding*. A sequence $\mathbf{x}$ is *morphic* if there exist a morphism $f\colon A^* \to A^*$, a coding $g\colon A^* \to B^*$, and a letter $a \in A$ such that $\mathbf{x} = g(f^\omega(a))$, where $f^\omega(a) = \lim_{n\to\infty} f^n(a)$. We let $E\colon \{0,1\}^* \to \{0,1\}^*$ be the *exchange morphism* defined by $E(0) = 1$ and $E(1) = 0$. We naturally extend E on sequences.

Numeration systems. Let $U = (U(n))_{n\geq 0}$ be an increasing sequence of integers with $U(0) = 1$. Any integer n can be decomposed, not necessarily uniquely, as $n = \sum_{i=0}^{t} c(i)\,U(i)$ with non-negative integer coefficients $c(i)$. The word $c(t)\cdots c(0) \in \mathbb{N}^*$ is a *U-representation* of n. If this representation is computed greedily, then for all $j < t$ we have $\sum_{i=0}^{j} c(i)\,U(i) < U(j+1)$ and $\mathrm{rep}_U(n) = c(t)\cdots c(0)$ is said to be the *greedy* U-representation of n. By convention, the greedy representation of 0 is the empty word ε, and the greedy representation of $n > 0$ starts with a non-zero digit. For any $c_t \cdots c_0 \in \mathbb{N}^*$, we let $\mathrm{val}_U(c(t)\cdots c(0))$ denote the integer $\sum_{i=0}^{t} c(i)\,U(i)$. A sequence U satisfying all the above conditions defines a *positional numeration system.*

Automatic and regular sequences. For the case of integer base numeration systems, a classical reference on automatic sequences is [7], while [75, 82] treat the case of more exotic numeration systems.

Let $U = (U(n))_{n\geq 0}$ be an positional numeration system. A sequence $\mathbf{x}$ is *U-automatic* if there exists a deterministic finite automaton with output (DFAO) $\mathcal{A}$ such that, for all $n \geq 0$, the nth term $\mathbf{x}(n)$ of $\mathbf{x}$ is given by the output $\mathcal{A}(\mathrm{rep}_U(n))$ of $\mathcal{A}$. In particular, if U is the sequence of consecutive powers of an integer $k \geq 2$, then $\mathbf{x}$ is said to be *k-automatic.*

It is known that a sequence is k-automatic if and only if it is the image, under a coding, of a fixed point of a k-uniform morphism [7].

A generalization of automatic sequences to infinite alphabets is the following [7, 75, 82]. Let $U = (U(n))_{n\geq 0}$ be an positional numeration system. A sequence $\mathbf{x}$ is *U-regular* if there exist a column vector λ, a row vector γ and matrix-valued morphism μ such that $\mathbf{x}(n) = \lambda\mu(\mathrm{rep}_U(n))\gamma$. Such a system of matrices forms a *linear representation* of $\mathbf{x}$.

Another definition of k-regular sequences is the following one [7]. Consider a sequence $\mathbf{x}$ and an integer $k \geq 2$. The k-kernel of $\mathbf{x}$ is the set of subsequences of the form $(\mathbf{x}(k^e n + r))_{n\geq 0}$ where $r \in \{1, 2, \ldots, k^e\}$. Equivalently, a sequence is k-regular if the $\mathbb{Z}$-module generated by its k-kernel is finitely generated. A sequence is then k-automatic if and only if its k-kernel is finite [7].

Sturmian sequences. A sequence $\mathbf{x}$ is *Sturmian* if its factor complexity function satisfies $\rho_{\mathbf{x}}(n) = n + 1$ (see, e.g., [7, 60]). Sturmian sequences have minimal factor complexity among all non-eventually periodic sequences, as proved by Morse and Hedlund [66].

Theorem 5 ([66])**.** *Let* $\mathbf{x}$ *be a sequence and let* ℓ *be the number of distinct letters occurring in* $\mathbf{x}$*. The following properties are equivalent.*

(a) The sequence $\mathbf{x}$ *is eventually periodic.*

(b) We have $\rho_{\mathbf{x}}(n) = \rho_{\mathbf{x}}(n + 1)$ *for some* $n \geq 0$*.*

(c) We have $\rho_{\mathbf{x}}(n) < n + \ell - 1$ *for some* $n \geq 1$*.*

(d) The factor complexity $\rho_{\mathbf{x}}$ *is bounded.*

Remark 6. This theorem implies in particular that $\rho_{\mathbf{x}}(n)$ is either bounded or it satisfies $\rho_{\mathbf{x}}(n) \geq n + \ell$ for all n. Thus the minimal factor complexity among all non-eventually periodic sequences is $\rho_{\mathbf{x}}(n) = n + 1$ for all n, i.e., the complexity of Sturmian sequences. Actually there is another "growth gap" for $\rho_{\mathbf{x}}$. Recall that a sequence $\mathbf{x}$ is called *quasi-Sturmian* if there exists a constant C such that, for n large enough, one has $\rho_{\mathbf{x}}(n) = n + C$ (see [28]; also see [34]). It is known that if $\mathbf{x}$ is neither eventually periodic nor quasi-Sturmian, then $\rho_{\mathbf{x}}(n) - n$ tends to infinity (this result is due to Coven [37, Lemma 1.3]; the first author [3, Theorem 3, p. 23] attributed the result to J. Cassaigne, because he first learned it from him). Thus we have:

(a) either $\rho_{\mathbf{x}}(n)$ is bounded, which happens if and only if $\mathbf{x}$ is eventually periodic;

(b) or else $\rho_{\mathbf{x}}(n) = n + C$ for some constant C and all n large enough, which means that $\mathbf{x}$ is quasi-Sturmian;

(c) or else $\rho_{\mathbf{x}}(n) - n$ tends to infinity.

One more point (once explained to the first author by J. Berstel) is that if $\rho_{\mathbf{x}}(n) = n + 1$ for n large enough, then $\rho_{\mathbf{x}}(n) = n + 1$ for all n. Namely, let n_0 be the least integer n for which $\rho_{\mathbf{x}}(n) = n + 1$, and suppose that $n_0 > 1$. Hence $\rho_{\mathbf{x}}(n_0 - 1) \neq n_0$. The sequence $\mathbf{x}$ cannot be eventually periodic, since its factor complexity is not bounded. Thus, one has $\rho_{\mathbf{x}}(n) \geq n + 1$ for all

n. Hence, in particular, $\rho_{\mathbf{x}}(n_0 - 1) \geq n_0$. Thus $\rho_{\mathbf{x}}(n_0 - 1) > n_0$. Since $\rho_{\mathbf{x}}$ is non-decreasing, we have that $n_0 < \rho_{\mathbf{x}}(n_0 - 1) \leq \rho_{\mathbf{x}}(n_0) = n_0 + 1$. This gives $\rho_{\mathbf{x}}(n_0 - 1) = n_0 + 1 = \rho_{\mathbf{x}}(n_0)$, which is impossible since $\mathbf{x}$ is not eventually periodic. In other words, in the second item above, if $C = 1$, then $\mathbf{x}$ is Sturmian.

2 General results

Given a sequence $\mathbf{x}$, we can decompose its factor complexity function $\rho_{\mathbf{x}}$ and its reflection complexity function $r_{\mathbf{x}}$ by using the following functions:

(a) The number $\mathrm{Unr}_{\mathbf{x}}(n)$ of "unreflected" length-n factors w of $\mathbf{x}$ such that w^R is not a factor of $\mathbf{x}$;

(b) The number $\mathrm{Ref}_{\mathbf{x}}(n)$ of "reflected" length-n factors w of $\mathbf{x}$ such that w^R is also a factor of $\mathbf{x}$; and

(c) The number $\mathrm{Pal}_{\mathbf{x}}(n)$ of length-n palindrome factors w of $\mathbf{x}$ (i.e., the *palindrome complexity* function of $\mathbf{x}$, see [4]).

In particular, we have

$$\begin{aligned} \rho_{\mathbf{x}}(n) &= \mathrm{Unr}_{\mathbf{x}}(n) + \mathrm{Ref}_{\mathbf{x}}(n), \\ r_{\mathbf{x}}(n) &= \mathrm{Unr}_{\mathbf{x}}(n) + \frac{1}{2}(\mathrm{Ref}_{\mathbf{x}}(n) - \mathrm{Pal}_{\mathbf{x}}(n)) + \mathrm{Pal}_{\mathbf{x}}(n). \end{aligned} \tag{3}$$

Example 7. Let $\mathbf{f} = 01101010\cdots$ denote the Fibonacci sequence, which is the fixed point of $0 \mapsto 01, 1 \mapsto 0$. Its length-5 factors are $u_1 = 01001$, $u_2 = 10010$, $u_3 = 00101$, $u_4 = 01010$, $u_5 = 10100$, and $u_6 = 00100$. Observe that u_4 is an unreflected factor (Type 1 above), u_1, u_2, u_3, u_5 are reflected (Type 2), and u_6 is a palindrome (Types 2 and 3). We obtain $r_{\mathbf{f}}(6) = 1 + \frac{1}{2}(5 - 1) + 1 = 4$.

There is a rich interplay among the complexity functions $\mathrm{Unr}_{\mathbf{x}}$, $\mathrm{Ref}_{\mathbf{x}}$, $\rho_{\mathbf{x}}$, $r_{\mathbf{x}}$, and $\mathrm{Pal}_{\mathbf{x}}$, which motivates the study of the combinations of these functions indicated in Equalities (3). This is illustrated below.

Lemma 8. *For a sequence* $\mathbf{x}$ *and for all* $n \geq 1$*, we have*

$$\rho_{\mathbf{x}}(n) - r_{\mathbf{x}}(n) = \frac{1}{2}(\mathrm{Ref}_{\mathbf{x}}(n) - \mathrm{Pal}_{\mathbf{x}}(n))$$

and

$$2r_{\mathbf{x}}(n) - \rho_{\mathbf{x}}(n) = \mathrm{Unr}_{\mathbf{x}}(n) + \mathrm{Pal}_{\mathbf{x}}(n).$$

Proof. Immediate from Equalities (3). □

This lemma implies the following bounds on the ratio $\frac{r}{\rho}$.

Theorem 9. *For any sequence* $\mathbf{x}$ *and for all* $n \geq 0$*, we have*

$$\frac{1}{2}\rho_{\mathbf{x}}(n) \leq \frac{1}{2}(\rho_{\mathbf{x}}(n) + \mathrm{Pal}_{\mathbf{x}}(n)) \leq r_{\mathbf{x}}(n) \leq \rho_{\mathbf{x}}(n).$$

Furthermore, the equality cases are as follows.

(a) We have $r_{\mathbf{x}}(n) = \rho_{\mathbf{x}}(n)$ *if and only if every reflected length-n factor of* $\mathbf{x}$ *is a palindrome.*

(b) We have $r_{\mathbf{x}}(n) = \frac{1}{2}(\rho_{\mathbf{x}}(n) + \mathrm{Pal}_{\mathbf{x}}(n))$ *if and only if* $\mathbf{x}$ *has no unreflected length-n factors. In particular, if the sequence* $\mathbf{x}$ *is reversal-closed, we have* $r_{\mathbf{x}} = \frac{1}{2}(\rho_{\mathbf{x}} + \mathrm{Pal}_{\mathbf{x}})$.

(c) We have $r_{\mathbf{x}}(n) = \frac{1}{2}\rho_{\mathbf{x}}(n)$ *if and only if* $\mathbf{x}$ *has no palindrome of length n and each of its length-n factors is reflected.*

Proof. The inequalities and the equality cases are immediate consequences of Lemma 8. □

Remark 10. It is known that if a sequence $\mathbf{x}$ is reversal-closed, then $\mathbf{x}$ is recurrent: it suffices to adapt the proof of [40, Proposition 1, p. 176], as indicated in [22]. Also note that if a sequence $\mathbf{x}$ is uniformly recurrent and contains infinitely many distinct palindromes, then $\mathbf{x}$ is reversal-closed [15, Theorem 3.2].

One can say more for uniformly recurrent sequences. The following dichotomy holds.

Theorem 11. *Let* $\mathbf{x}$ *be a uniformly recurrent sequence. Then either it is reversal-closed, or else it has no long reflected factors (which implies that* $\mathbf{x}$ *has no long palindromes). In other words,*

(a) either $\rho_{\mathbf{x}} = \mathrm{Ref}_{\mathbf{x}}$*, which implies the equality* $r_{\mathbf{x}} = \frac{1}{2}(\rho_{\mathbf{x}} + \mathrm{Pal}_{\mathbf{x}})$*;*

(b) or else there exists n_0 such that $\rho_{\mathbf{x}}(n) = \mathrm{Unr}_{\mathbf{x}}(n)$ for all $n \geq n_0$, which implies $r_{\mathbf{x}}(n) = \rho_{\mathbf{x}}(n)$ for all $n \geq n_0$.

Proof. If $\mathbf{x}$ is reversal-closed, then $r_{\mathbf{x}} = \frac{1}{2}(\rho_{\mathbf{x}} + \mathrm{Pal}_{\mathbf{x}})$ from Theorem 9(b) above. Now suppose that $\mathbf{x}$ has an unreflected factor w. Since $\mathbf{x}$ is uniformly recurrent, any long enough factor of $\mathbf{x}$ contains w as a factor, which implies that this long factor itself is unreflected. This exactly says that $\mathrm{Ref}_{\mathbf{x}}(n) = 0$ for n large enough (and in particular $\mathrm{Pal}_{\mathbf{x}}(n) = 0$ for n large enough). This implies from Equalities (3) that, for n large enough, $\rho_{\mathbf{x}}(n) = \mathrm{Unr}_{\mathbf{x}}(n)$, and so $r_{\mathbf{x}}(n) = \rho_{\mathbf{x}}(n)$. □

Now we exhibit sequences with particular behaviors of their reflection complexity.

Example 12. It is possible to construct an aperiodic automatic sequence $\mathbf{x}$ such that $r_{\mathbf{x}}(n) = \rho_{\mathbf{x}}(n)$ and $\mathrm{Pal}_{\mathbf{x}}(n) > 0$ for all n. An example of such a sequence is given by a fixed point of the morphism $0 \mapsto 01$, $1 \mapsto 23$, $2 \mapsto 45$, $3 \mapsto 23$, $4 \mapsto 44$, and $5 \mapsto 44$. This sequence has no reflected factors except palindromes, and there is exactly one palindrome of each length > 1.

Example 13. Consider the sequence $\mathbf{x}$ on $\{0,1\}$ whose nth prefix x_n is given recursively as follows: $x_0 = 01$ and $x_{n+1} = x_n 01 x_n^R$ for all $n \geq 0$. See [19, Section 3] or [15, Example 3.1]. The sequence $\mathbf{x}$ is uniformly recurrent, reversal-closed, and 2-automatic and accepted by a DFAO of 6 states (see, e.g., [6]), and contains only a finite number of palindromes. Furthermore, for all sufficiently large n, we have $r_{\mathbf{x}}(n) = \dfrac{1}{2}\rho_{\mathbf{x}}(n)$.

Example 14. It is also possible to construct an aperiodic automatic sequence where the only palindromes are of length 1, but there are reflected factors of each length > 1. In this regard, we let g_n be the prefix of length $2^n - 2$ of $(012)^\omega$. Then an example of an automatic sequence satisfying the desired properties is

$$3\, g_1\, 4\ 5\, g_2^R\, 6\ 3\, g_3\, 4\ 5\, g_4^R\, 6\ 3\, g_5\, 4\ 5\, g_6^R\, 6 \cdots,$$

where [74, Theorem 1] is required (observe that we intertwine the sequences $(3456)^\omega$ and $g_1 g_2^R g_3 g_4^R \cdots$ to build our sequence).

Example 15. There is an automatic sequence $\mathbf{x}$ over the alphabet $\{0,1\}$ such that $\mathrm{Ref}_{\mathbf{x}}(n) = \Omega(n)$ and such that $\mathrm{Unr}_{\mathbf{x}}(n) = \Omega(n)$. Namely, consider the image under the coding $0,1,2 \mapsto 0$ and $3,4 \mapsto 1$ of the fixed point, starting with 0, of the morphism $0 \mapsto 01$, $1 \mapsto 23$, $2 \mapsto 32$, $3 \mapsto 42$, and $4 \mapsto 43$.

Example 16. We also provide a construction of an automatic sequence $\mathbf{x}$ such that $r_{\mathbf{x}}(n+1) < r_{\mathbf{x}}(n)$ for all odd $n \geq 3$. In particular, let $\mathbf{x}$ denote the sequence given by applying the coding $a, b, d \mapsto 1$ and $c \mapsto 0$ to the fixed point, starting with a, of the morphism defined by $a \mapsto ab$, $b \mapsto cd$, $c \mapsto cd$, and $d \mapsto bb$. This gives us sequence [85, `A039982`] in the OEIS. Computing the reflection complexity of $\mathbf{x}$ (e.g., using `Walnut`) gives that

$$r_{\mathbf{x}}(n) = \begin{cases} n+1, & \text{for odd } n \geq 1; \\ n-1, & \text{for even } n \geq 4. \end{cases}$$

Actually we even have, for this sequence $\mathbf{x}$, that $r_{\mathbf{x}}(n+1) = r_{\mathbf{x}}(n) - 1$ for all odd $n \geq 3$.

With extra hypotheses on a sequence $\mathbf{x}$, we can give more precise results in comparing the respective growths of reflection and factor complexities. We will need Theorem 17 below. Note that Part (b) of this theorem was originally stated for uniformly recurrent sequences: see [14, Theorem 1.2]. However, its proof only requires the sequences to be recurrent (see [15, p. 449] and also [23, Footnote, p. 493]). Furthermore we have seen that a reversal-closed sequence must be recurrent (see Remark 10). Thus we can state the theorem as follows (also see Theorem 11).

Theorem 17.

(a) Let $\mathbf{x}$ be a uniformly recurrent sequence. If $\mathbf{x}$ is not closed under reversal, then $\mathrm{Pal}(n) = 0$ *for n large enough (actually one even has* $\mathrm{Ref}_{\mathbf{x}}(n) = 0$ *for n large enough).*

(b) Let $\mathbf{x}$ be a reversal-closed sequence. For all $n \geq 0$, we have

$$\mathrm{Pal}_{\mathbf{x}}(n+1) + \mathrm{Pal}_{\mathbf{x}}(n) \leq \rho_{\mathbf{x}}(n+1) - \rho_{\mathbf{x}}(n) + 2.$$

Remark 18. There exist sequences that are uniformly recurrent, reversal-closed, and have no long palindromes (see [19]; also see Example 13 above).

We deduce the following theorem from Theorem 17.

Theorem 19. *Let* $\mathbf{x}$ *be a reversal-closed sequence. For all* $n \geq 0$*, we have*

$$\frac{1}{2}\rho_{\mathbf{x}}(n) \leq r_{\mathbf{x}}(n) \leq \frac{1}{2}\rho_{\mathbf{x}}(n+1) + 1.$$

Proof. Using Theorem 9(a) and (b), Remark 10 and Theorem 17(b), we have

$$\frac{1}{2}\rho_{\mathbf{x}}(n) \leq r_{\mathbf{x}}(n) = \frac{1}{2}(\rho_{\mathbf{x}}(n) + \mathrm{Pal}_{\mathbf{x}}(n)) \leq \frac{1}{2}(\rho_{\mathbf{x}}(n) + \rho_{\mathbf{x}}(n+1) - \rho_{\mathbf{x}}(n) + 2)$$

for all $n \geq 0$. Thus

$$\frac{1}{2}\rho_{\mathbf{x}}(n) \leq r_{\mathbf{x}}(n) \leq \frac{1}{2}\rho_{\mathbf{x}}(n+1) + 1$$

for all $n \geq 0$, as desired. □

On the other hand, we can use a result of [4] to obtain the following theorem.

Theorem 20. *Let* $\mathbf{x}$ *be a non-eventually periodic and reversal-closed sequence. For all* $n \geq 1$*, we have*

$$\frac{1}{2}\rho_{\mathbf{x}}(n) \leq r_{\mathbf{x}}(n) < \frac{1}{2}\rho_{\mathbf{x}}(n) + \frac{8}{n}\rho_{\mathbf{x}}\left(n + \left\lfloor\frac{n}{4}\right\rfloor\right).$$

Proof. Given a non-eventually periodic sequence $\mathbf{x}$, we have from [4, Theorem 12] the inequality

$$\mathrm{Pal}_{\mathbf{x}}(n) < \frac{16}{n}\rho_{\mathbf{x}}\left(n + \left\lfloor\frac{n}{4}\right\rfloor\right)$$

for all $n \geq 1$. The statement follows from this and Theorem 9(a) and (b). □

Corollary 21. *Let* $\mathbf{x}$ *be a non-eventually periodic and reversal-closed sequence. If its factor complexity satisfies* $\rho_{\mathbf{x}}(n+1) \sim \rho_{\mathbf{x}}(n)$ *or* $\frac{\rho_{\mathbf{x}}(2n)}{\rho_{\mathbf{x}}(n)} = o(n)$*, then*

$$r_{\mathbf{x}}(n) \sim \frac{1}{2}\rho_{\mathbf{x}}(n)$$

when n *tends to infinity. In particular, this equivalence holds if* $\mathbf{x}$ *is non-eventually periodic, reversal-closed and morphic.*

Proof. Let $\mathbf{x}$ be a non-eventually periodic and reversal-closed sequence. If $\rho_{\mathbf{x}}(n+1) \sim \rho_{\mathbf{x}}(n)$, then, from Theorem 19, we obtain that $r_{\mathbf{x}}(n) \sim \frac{1}{2}\rho_{\mathbf{x}}(n)$ when n tends to infinity. Now, if $\frac{\rho_{\mathbf{x}}(2n)}{\rho_{\mathbf{x}}(n)} = o(n)$, we obtain, from Theorems 9 and 20, and using the fact that $\rho_{\mathbf{x}}$ is non-decreasing,

$$\frac{1}{2}\rho_{\mathbf{x}}(n) \leq r_{\mathbf{x}}(n) < \frac{1}{2}\rho_{\mathbf{x}}(n) + \frac{8}{n}\rho_{\mathbf{x}}(2n) = \frac{1}{2}\rho_{\mathbf{x}}(n) + o(\rho_{\mathbf{x}}(n)),$$

which is enough.

Now suppose that, in addition, the sequence $\mathbf{x}$ is morphic. We know that either $\rho_{\mathbf{x}}(n) = \Theta(n^2)$ or $\rho_{\mathbf{x}}(n) = O(n^{3/2})$ (see [41]). In the first case, then $\frac{\rho_{\mathbf{x}}(2n)}{\rho_{\mathbf{x}}(n)}$ is bounded, and hence $o(n)$. If $\rho_{\mathbf{x}}(n) = O(n^{3/2})$, since $\mathbf{x}$ is not eventually periodic (hence $\rho_{\mathbf{x}}(n) \geq n+1$), we have

$$\frac{\rho_{\mathbf{x}}(2n)}{\rho_{\mathbf{x}}(n)} \leq C\frac{n^{3/2}}{n+1} = o(n).$$

This finishes the proof. $\square$

The upper bound in Theorem 20 raises questions as to growth properties of the function $r_{\mathbf{x}}$ more generally, apart from the case where the set of factors of $\mathbf{x}$ satisfies the hypotheses of Theorem 20. This leads us toward the growth property in Theorem 22 below.

Theorem 22. *Let $\mathbf{x}$ be a sequence. Then $r_{\mathbf{x}}(n) \leq r_{\mathbf{x}}(n+2)$ for all $n \geq 0$.*

Proof. The result is clear for $n = 0$, so assume $n > 0$ in what follows. Let c be a letter not in the alphabet of $\mathbf{x}$, and define $\mathbf{y} = c\mathbf{x}$. Then $r_{\mathbf{y}}(n) = r_{\mathbf{x}}(n) + 1$ for all $n > 0$, since $\mathbf{y}$ has exactly one additional factor for each length $n \geq 1$; namely, the prefix of length n. Thus, it suffices to prove the claim for $\mathbf{y}$ instead of $\mathbf{x}$.

With each length-n factor w of $\mathbf{y}$ associate a set S_w of length-$(n+2)$ factors of $\mathbf{y}$, as follows: If w is the length-n prefix of $\mathbf{y}$, then $S_w := \{w'\}$, where w' is the prefix of length $n+2$ of $\mathbf{y}$. We call such a factor *exceptional.* Otherwise, define $S_w := \{z \in \mathrm{Fac}(\mathbf{y}) \,:\, z = awb \text{ for some letters } a, b\}$. Note that the sets S_w, over all length-n factors of $\mathbf{y}$, are pairwise disjoint, and cover all the length-$(n+2)$ factors of $\mathbf{y}$.

For a factor w of $\mathbf{y}$, define $[w]_1 = 1$ if w is a palindrome, and 0 otherwise. Similarly, $[w]_2 = 1$ if w^R is not a factor of $\mathbf{y}$ and 0 otherwise. Finally, define $[w]_3 = 1$ if w^R is also a factor of $\mathbf{y}$ but w is not a palindrome, and 0 otherwise.

Notice that these three cases are disjoint and subsume all possibilities for factors of $\mathbf{y}$ (also recall the decomposition at the beginning of the section). We can extend this notation to sets by defining $[S]_i = \sum_{w \in S} [w]_i$ for $i \in \{1, 2, 3\}$. Define $[w] = [w]_1 + [w]_2 + [w]_3/2$ and similarly for $[S]$. From Equalities (3), we know that

$$r_{\mathbf{y}}(n) = \sum_{\substack{|w|=n \\ w \in \mathrm{Fac}(\mathbf{y})}} [w]$$

while

$$r_{\mathbf{y}}(n+2) = \sum_{\substack{|w|=n \\ w \in \mathrm{Fac}(\mathbf{y})}} [S_w].$$

Therefore, to show the desired inequality $r_{\mathbf{y}}(n) \leq r_{\mathbf{y}}(n+2)$, it suffices to show that $[w] \leq [S_w]$ for all length-n factors w of $\mathbf{y}$.

Suppose w is exceptional. Recall that w starts with c, which appears nowhere else in $\mathbf{y}$. Then $[w]_1 = [w]_3 = 0$, but $[w]_2 = 1$. And $S_w = \{wab\}$, so $[S_w]_1 = [S_w]_3 = 0$, but $[S_w]_2 = 1$. Therefore $[w] \leq [S_w]$.

Now suppose w is not exceptional. There are three cases to consider.

Case 1: If $[w]_1 = 1$, then w is a palindrome. Consider a factor $awb \in S_w$. If it is a palindrome, then $[awb]_1 = 1$, so $[w] \leq [awb]$. If awb is not a palindrome, then $awb \neq (awb)^R = bw^Ra = bwa$. Thus $a \neq b$. If bwa is not a factor of $\mathbf{y}$, then $[awb]_2 = 1$, so $[w] \leq [awb]$. If bwa is a factor of $\mathbf{y}$, then $bwa \in S_w$ and $[awb]_3 + [bwa]_3 = 2$, so in all cases $[w] \leq [awb]$. Thus $[w] \leq [S_w]$.

Case 2: If $[w]_2 = 1$, then w^R is not a factor of $\mathbf{y}$. Consider a factor $awb \in S_w$. Then $(awb)^R = bw^Ra$, so $(awb)^R$ cannot be a factor of y either. Hence $[awb]_2 = 1$, $[w] \leq [awb]$, and hence $[w] \leq [S_w]$.

Case 3: If $[w]_3 = 1$, then w^R is a factor of $\mathbf{y}$, but w is not a palindrome. Consider a factor $awb \in S_w$. If awb is a palindrome, then $awb = (awb)^R = bw^Ra$, so w^R would be a palindrome, a contradiction. So awb is not a palindrome and $[awb]_1 = 0$. If $(awb)^R = bw^Ra$ is a factor of $\mathbf{y}$, then $[awb]_3 = 1$, so $[w] \leq [awb]$. If $(awb)^R$ is not a factor of $\mathbf{y}$, then $[awb]_2 = 1$, so $[w] < [awb]$. Thus $[w] \leq [S_w]$.

This completes the proof. □

Remark 23. Another formulation of Theorem 22 above is that the sequence $(r_{\mathbf{x}}(n+1) + r_{\mathbf{x}}(n))_{n\geq 0}$ is non-decreasing.

Numerical experiments concerning the growth of the reflection complexity have led us to formulate Conjectures 24–27 below. We leave these conjectures as open problems.

Conjecture 24. *Let* $\mathbf{x}$ *be a sequence. Then* $r_{\mathbf{x}}(n) = r_{\mathbf{x}}(n+2)$ *for some* n *if and only if* $\mathbf{x}$ *is eventually periodic.*

Note that one direction is true. We have even more: namely, if the sequence $\mathbf{x}$ is eventually periodic, then $r_{\mathbf{x}}(n) = r_{\mathbf{x}}(n+2)$, for all n large enough (see Theorem 32 below).

Conjecture 25. *Let* $\mathbf{x}$ *be a sequence. Then* $r_{\mathbf{x}}(n) - 1 \leq r_{\mathbf{x}}(n+1)$ *for all* $n \geq 0$.

Note that we can have equality for infinitely many values of n—see Example 16 above.

Conjecture 26. *Let* $\mathbf{x}$ *be a sequence of at most linear factor complexity. Then* $r_{\mathbf{x}}(n+1) - r_{\mathbf{x}}(n)$ *is bounded for all* $n \geq 0$. *Hence, in particular, if* $\mathbf{x}$ *is (generalized) automatic, so is* $(r_{\mathbf{x}}(n+1) - r_{\mathbf{x}}(n))_{n\geq 0}$.

It can be shown that Conjecture 26 holds for the Thue–Morse, period-doubling, Golay–Shapiro, second-bit, paperfolding, Stewart choral, Baum-Sweet, Chacon, and Mephisto-Waltz sequences. (Also see Corollary 30.)

Conjecture 27. *Let* $\mathbf{x}$ *be a sequence. If the limit* $\lim_{n\to\infty} \frac{r_{\mathbf{x}}(n)}{\rho_{\mathbf{x}}(n)}$ *exists, then it is either equal to* $\frac{1}{2}$ *or to* 1.

Actually we prove below weaker forms of Conjectures 25 and 26 for reversal-closed sequences, and weaker forms of Conjectures 24 and 25 for sequences without long palindromes. Also we can prove that Conjecture 27 holds for primitive morphic sequence.

Theorem 28. *Let* $\mathbf{x}$ *be a reversal-closed sequence. Then, for all* $n \geq 0$, *we have* $r_{\mathbf{x}}(n) - 1 \leq r_{\mathbf{x}}(n+1)$. *If, in addition,* $\mathbf{x}$ *has at most linear factor complexity, then* $r_{\mathbf{x}}(n+1) - r_{\mathbf{x}}(n)$ *is bounded.*

Proof. Let $\mathbf{x}$ be a reversal-closed sequence. By Remark 10, $\mathbf{x}$ is recurrent. Thus, putting together Theorems 17(b) and 9(b) gives, for all $n \geq 0$,

$$2(r_{\mathbf{x}}(n+1) - r_{\mathbf{x}}(n)) = \rho_{\mathbf{x}}(n+1) + \mathrm{Pal}_{\mathbf{x}}(n+1) - \rho_{\mathbf{x}}(n) - \mathrm{Pal}_{\mathbf{x}}(n). \quad (4)$$

Equality (4) implies $2(r_{\mathbf{x}}(n+1) - r_{\mathbf{x}}(n)) \geq \mathrm{Pal}_{\mathbf{x}}(n+1) - 2$, hence we obtain $r_{\mathbf{x}}(n+1) - r_{\mathbf{x}}(n) \geq -1$ as desired.

If, in addition, the factor complexity of $\mathbf{x}$ is at most linear, then Equality (4) gives

$$2|r_{\mathbf{x}}(n+1) - r_{\mathbf{x}}(n)| \leq |\rho_{\mathbf{x}}(n+1) - \rho_{\mathbf{x}}(n)| + \mathrm{Pal}_{\mathbf{x}}(n+1) + \mathrm{Pal}_{\mathbf{x}}(n).$$

But $|\rho_{\mathbf{x}}(n+1) - \rho_{\mathbf{x}}(n)|$ is bounded (see [27]) and $\mathrm{Pal}_{\mathbf{x}}$ is also bounded (see [4, Theorem 12] or use Theorem 17(b) above). □

Theorem 29. *Let $n_0 \geq 0$ be an integer and let $\mathbf{x}$ be a sequence with no palindrome of length $\geq n_0$. Then $(r_{\mathbf{x}}(n))_{n\geq 0}$ is eventually non-decreasing: $r_{\mathbf{x}}(n) \leq r_{\mathbf{x}}(n+1)$ for $n \geq n_0$. Furthermore, if $r_{\mathbf{x}}(n+2) = r_{\mathbf{x}}(n)$ for some $n \geq n_0$, then the sequence $\mathbf{x}$ is eventually periodic.*

Proof. By combining the assumption and the second equality of Lemma 8, we have that

$$r_{\mathbf{x}}(n) = \frac{1}{2}(\rho_{\mathbf{x}}(n) + \mathrm{Unr}_{\mathbf{x}}(n)) \quad (5)$$

for $n \geq n_0$. Since both $(\rho_{\mathbf{x}}(n))_{n\geq 0}$ and $(\mathrm{Unr}_{\mathbf{x}}(n))_{n\geq 0}$ are non-decreasing, we see that $(r_{\mathbf{x}}(n))_{n\geq n_0}$ is non-decreasing, which gives that $r_{\mathbf{x}}(n) \leq r_{\mathbf{x}}(n+1)$ for $n \geq n_0$. This shows the first part of the statement. For the second part, if we have $r_{\mathbf{x}}(n+2) = r_{\mathbf{x}}(n)$ for some $n \geq n_0$, then Equality (5) implies that $\rho_{\mathbf{x}}(n+2) + \mathrm{Unr}_{\mathbf{x}}(n+2) = \rho_{\mathbf{x}}(n) + \mathrm{Unr}_{\mathbf{x}}(n)$. Hence

$$\rho_{\mathbf{x}}(n+2) + \mathrm{Unr}_{\mathbf{x}}(n+2) = \rho_{\mathbf{x}}(n+1) + \mathrm{Unr}_{\mathbf{x}}(n+1) = \rho_{\mathbf{x}}(n) + \mathrm{Unr}_{\mathbf{x}}(n).$$

Hence $\rho_{\mathbf{x}}(n+1) = \rho_{\mathbf{x}}(n)$, which implies that $\mathbf{x}$ is eventually periodic from Theorem 5, as desired. □

Actually, Theorems 11, 28 and 29 imply the following corollary.

Corollary 30. *Conjectures 25 and 26 hold if $\mathbf{x}$ is uniformly recurrent.*

Proof. Let $\mathbf{x}$ be a uniformly recurrent sequence. Then, from Theorem 11, we have that $\mathbf{x}$ is either reversal-closed and $\rho_{\mathbf{x}} = \mathrm{Ref}_{\mathbf{x}}$, so that $r_{\mathbf{x}} = \frac{1}{2}(\rho_{\mathbf{x}} + \mathrm{Pal}_{\mathbf{x}})$, or else that there exists n_0 such that $\rho_{\mathbf{x}}(n) = \mathrm{Unr}_{\mathbf{x}}(n)$ for all $n \geq n_0$, which implies $r_{\mathbf{x}}(n) = \rho_{\mathbf{x}}(n)$ for all $n \geq n_0$.

In the first case, the claim is proved by using Theorem 28, and that, if, in addition, $\mathbf{x}$ is automatic, then both sequences $(\rho_{\mathbf{x}}(n+1) - \rho_{\mathbf{x}}(n))_{n \geq 0}$ and $(\mathrm{Pal}_{\mathbf{x}}(n))_{n \geq 0}$ are automatic (see [26, Theorem 4.3] and [26, Theorem 4.8] respectively). Hence the sequence $(\mathrm{Pal}_{\mathbf{x}}(n+1) - \mathrm{Pal}_{\mathbf{x}}(n))_{n \geq 0}$ is also automatic. The proofs extend easily to generalized automatic sequences.

In the second case, inspired by the proof of Theorem 29, we note that there is an integer n_0 such that $r_{\mathbf{x}}(n) = \rho_{\mathbf{x}}(n) = \mathrm{Unr}_{\mathbf{x}}(n)$ for all $n \geq n_0$. But $\mathrm{Unr}_{\mathbf{x}}$ is clearly non-decreasing, which gives the property of Conjecture 25. If, in addition, $\mathbf{x}$ has at most linear complexity, we know that $\rho_{\mathbf{x}}(n+1) - \rho_{\mathbf{x}}(n)$ is bounded (see [27]), hence $r_{\mathbf{x}}(n+1) - r_{\mathbf{x}}(n)$ is bounded for n large enough, hence for all n. The (generalized) automatic property is proved by using, as above, that $(\rho_{\mathbf{x}}(n+1) - \rho_{\mathbf{x}}(n))_{n \geq 0}$ is (generalized) automatic. □

In the same vein, Theorem11 and Corollary 21 imply the following corollary.

Corollary 31. *Conjecture 27 holds for non-eventually periodic primitive morphic sequences.*

Proof. Let $\mathbf{x}$ be a primitive morphic sequence. We know that $\mathbf{x}$ is uniformly recurrent. Thus, from Theorem 11, $\mathbf{x}$ is either reversal-closed, or else it has no long palindromes. If $\mathbf{x}$ is reversal-closed, then, by Corollary 21, we have $r_{\mathbf{x}}(n) \sim \frac{1}{2}\rho_{\mathbf{x}}(n)$. Otherwise, $\mathbf{x}$ has no long palindromes, then, still from Theorem 11, we have that $r_{\mathbf{x}}(n) = \rho_{\mathbf{x}}(n)$ for n large enough. □

3 Eventually periodic sequences

We can characterize eventually periodic sequences (i.e., sequences that are periodic from some index on) in terms of their reflection complexity.

Theorem 32. *A sequence $\mathbf{x}$ is eventually periodic if and only if both sequences $(r_{\mathbf{x}}(2n))_{n \geq 0}$ and $(r_{\mathbf{x}}(2n+1))_{n \geq 0}$ are eventually constant.*

Proof. From Theorem 22 both sequences $(r_{\mathbf{x}}(2n))_{n \geq 0}$ and $(r_{\mathbf{x}}(2n+1))_{n \geq 0}$ are non-decreasing. Also, from the inequality $\frac{1}{2}\rho_{\mathbf{x}}(n) \leq r_{\mathbf{x}}(n) \leq \rho_{\mathbf{x}}(n)$ in

Theorem 9, and the fact that the sequence $(\rho_{\mathbf{x}}(n))_{n\geq 0}$ is non-decreasing, we have that either the three integer sequences $(r_{\mathbf{x}}(2n))_{n\geq 0}$, $(r_{\mathbf{x}}(2n+1))_{n\geq 0}$, and $(\rho_{\mathbf{x}}(n))_{n\geq 0}$ are all bounded, or else none of them is. Furthermore, we know that $(\rho_{\mathbf{x}}(n))_{n\geq 0}$ is bounded if and only if the sequence $\mathbf{x}$ is eventually periodic (Theorem 5(d) above). Hence, we have two cases depending on the periodicity of $\mathbf{x}$.

(a) If $\mathbf{x}$ is eventually periodic, then $(\rho_{\mathbf{x}}(n))_{n\geq 0}$ is bounded, so $(r_{\mathbf{x}}(2n))_{n\geq 0}$ and $(r_{\mathbf{x}}(2n+1))_{n\geq 0}$ are eventually constant.

(b) If $\mathbf{x}$ is not eventually periodic, its factor complexity is not bounded: thus both sequences $(r_{\mathbf{x}}(2n))_{n\geq 0}$ and $(r_{\mathbf{x}}(2n+1))_{n\geq 0}$ tend to infinity.

This ends the proof. □

Remark 33. If $\mathbf{x}$ is eventually periodic, the eventual values of $(r_{\mathbf{x}}(2n))_{n\geq 0}$ and $(r_{\mathbf{x}}(2n+1))_{n\geq 0}$ can be either equal or distinct, as seen from the examples of the sequences $(01)^{\omega}$ and $(011)^{\omega}$.

Remark 34. Theorem 35, Corollary 36 and Remark 38 below give more precise results for the growth of the reflection complexity of eventually periodic and non-eventually periodic sequences.

4 Sturmian sequences and generalizations

In this section, we study Sturmian sequences as well as some generalizations.

4.1 Sturmian sequences

First we state the following result, which notably characterizes Sturmian sequences in terms of their reflection complexity.

Theorem 35. *Let $\mathbf{x}$ be a non-eventually periodic sequence on a finite alphabet.*

(a) For all $n \geq 1$, we have $r_{\mathbf{x}}(n) \geq 1 + \lfloor \frac{n+1}{2} \rfloor$;

(b) We have $r_{\mathbf{x}}(n) = 1 + \lfloor \frac{n+1}{2} \rfloor$ if and only if $\mathbf{x}$ is Sturmian.

Proof. For each integer $n \geq 1$, let $\mathcal{S}_n$ be the permutation group on n elements. Let σ_n be the permutation defined by

$$\sigma_n := \begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ n & n-1 & \dots & 2 & 1 \end{pmatrix}$$

and G_n be the subgroup of $\mathcal{S}_n$ generated by σ_n, i.e., the group $\{\sigma_n, \mathrm{id}_n\}$. The number of distinct orbits of $\{1, 2, \dots, n\}$ under G_n is equal to $n/2$ if n is even, and to $(n+1)/2$ if n is odd, which can be written $\lfloor (n+1)/2 \rfloor$ in both cases. Thus, applying [33, Theorem 1] proves the first item of the theorem and the implication $\Longrightarrow$ of the second item.

To prove the last assertion, suppose that $\mathbf{x}$ is a Sturmian sequence. We know that any Sturmian sequence is reversal-closed (see [43, Theorem 4, p. 77], where reversals are called mirror images). Furthermore, it is proved in [43, Theorem 5, p. 77] that a sequence is Sturmian if and only if it has one palindrome of all even lengths and two palindromes of all odd lengths. Now, from Theorem 9(b) we have that

$$r_{\mathbf{x}}(n) = \frac{1}{2}(\rho_{\mathbf{x}}(n) + \mathrm{Pal}_{\mathbf{x}}(n)) = \begin{cases} \frac{n+2}{2} = 1 + \lfloor \frac{n+1}{2} \rfloor, & \text{if } n \text{ even;} \\ \frac{n+3}{2} = 1 + \lfloor \frac{n+1}{2} \rfloor, & \text{if } n \text{ odd.} \end{cases}$$

This ends the proof. □

With regard to the above referenced work of Charlier et al. [33]; also see the related and recent work by Luchinin and Puzynina [61].

The following is an analog of the Morse–Hedlund theorem (which is recalled in Theorem 5 above).

Corollary 36. *A sequence $\mathbf{x}$ is eventually periodic if and only if there exists $n \geq 1$ such that $r_{\mathbf{x}}(n) \leq \lfloor \frac{n+1}{2} \rfloor$. Furthermore both sequences $(r_{\mathbf{x}}(2n))_n$ and $(r_{\mathbf{x}}(2n+1))_n$ are then eventually constant.*

Proof. Let $\mathbf{x}$ be a sequence. Contraposing Property (a) of Theorem 35, we obtain that if $r_{\mathbf{x}}(n) \leq \lfloor \frac{n+1}{2} \rfloor$ for some n, then $\mathbf{x}$ must be eventually periodic. Conversely, if $\mathbf{x}$ is eventually periodic, it has a bounded number of factors, hence there exists some integer n for which the inequality of the statement is true. The last assertion is Theorem 32 above. □

Remark 37. Actually, it is possible to prove the first part of the proof of Corollary 36 without using Theorem 35 in the case where *the integer n is even*. Namely, Theorem 9 implies that $\rho_{\mathbf{x}}(n)/2 \leq r_{\mathbf{x}}(n)$ for all n. So that, if $r_{\mathbf{x}}(n_0) \leq \lfloor \frac{n_0+1}{2} \rfloor$ for some n_0, then $\rho_{\mathbf{x}}(n_0) \leq 2\lfloor \frac{n_0+1}{2} \rfloor = n_0$ if n_0 is even. Then we apply the Morse–Hedlund theorem (Theorem 5).

Remark 38. As in Remark 6, there is another "growth gap". Namely, an easy consequence of Theorem 9 above is that, for any sequence $\mathbf{x}$ that is neither eventually periodic, nor quasi-Sturmian, one has $r_{\mathbf{x}}(n) - \frac{n}{2} \to +\infty$ when $n \to +\infty$ (recall the definition of quasi-Sturmian in Remark 6; and use the fact that, for a sequence $\mathbf{x}$ that is neither eventually periodic nor quasi-Sturmian, one has that $\rho_{\mathbf{x}}(n) - n \to +\infty$). So that we can have the following possibilities for a sequence $\mathbf{x}$:

(a) $r_{\mathbf{x}}(n)$ is bounded, which happens if and only if $\mathbf{x}$ is eventually periodic. Then both sequences $(r_{\mathbf{x}}(2n))_n$ and $(r_{\mathbf{x}}(2n+1))_n$ are eventually constant;

(b) $r_{\mathbf{x}}(n) = 1 + \lfloor \frac{n+1}{2} \rfloor$ for all n, which happens if and only if $\mathbf{x}$ is Sturmian. Note that $r_{\mathbf{x}}(n) - \frac{n}{2} > \frac{1}{2}$ for any non-eventually periodic sequence, and that $r_{\mathbf{x}}(n) - \frac{n}{2} \in \{1, \frac{3}{2}\}$ for any Sturmian sequence;

(c) $r_{\mathbf{x}}(n) - \frac{n}{2}$ is bounded, and $\mathbf{x}$ is not Sturmian. This implies that $\rho_{\mathbf{x}}(n) - n$ is bounded (use Theorem 9), hence that $\mathbf{x}$ is quasi-Sturmian;

(d) $r_{\mathbf{x}}(n) - \frac{n}{2}$ is not bounded and $\mathbf{x}$ is quasi-Sturmian.

(e) $r_{\mathbf{x}}(n) - \frac{n}{2}$ tends to infinity. This is the case where $\mathbf{x}$ is neither eventually periodic nor quasi-Sturmian.

Note that both behaviors $r_{\mathbf{x}}(n) - \frac{n}{2}$ bounded (Item (c)) or $r_{\mathbf{x}}(n) - \frac{n}{2}$ not bounded (Item (d)) are possible for quasi-Sturmian sequences. Namely, if we start from the binary Fibonacci sequence $\mathbf{f}$ (fixed point of the morphism $0 \to 01$, $1 \to 0$), and apply two particular morphisms, then (using `Walnut`), we have for the corresponding quasi-Sturmian sequences:

(a) the image of $\mathbf{f}$ under the morphism $0 \to 0101$, $1 \to 1111$ is reversal closed and has arbitrarily long palindromes; its reflection complexity is given by $r(n) - \frac{n}{2} = \frac{9}{2}$ for $n \geq 7$ odd and $r(n) - \frac{n}{2} = 3$ for $n \geq 8$ even;

(b) the image of $\mathbf{f}$ under the morphism $0 \to 01101$, $1 \to 10100$ is not reversal-closed, and has no large palindrome; its complexity has the property that $r(n) = n + 9$ for $n \geq 11$.

It is also worth noting that in [38, p. 133], it is indicated that, for any Sturmian sequence $\mathbf{x}$ with values in $\{0, 1\}$, the image of $\mathbf{x}$ by the morphism f defined by $f(0) = 011001$, $f(1) = 001011$ is a quasi-Sturmian sequence without long palindromes (the authors call such a sequence a *non-palindromic sequence*, using the terminology of [53]). More precise results on the reflection complexity of quasi-Sturmian sequences are given in the next subsection.

4.2 Quasi-Sturmian sequences

Recall that any quasi-Sturmian sequence $\mathbf{x}$ can be written as $\mathbf{x} = uf(\mathbf{z})$, where u is any word on any alphabet, $\mathbf{z}$ is a (necessarily binary) Sturmian sequence, and f an aperiodic morphism from $\{0, 1\}$ to any alphabet (see [28, 37, 72]). We state the following theorem.

Theorem 39. *Let* $\mathbf{x} = yf(\mathbf{z})$ *be a quasi-Sturmian sequence, where* y *is any word,* $\mathbf{z}$ *is a Sturmian sequence, and* f *is an aperiodic morphism from* $\{0, 1\}$ *to any alphabet. Then*

(a) either $f(\mathbf{z})$ *is reversal-closed and* $r_{\mathbf{x}} = \frac{n}{2} + O(1)$*;*

(b) or else $f(\mathbf{z})$ *is not reversal-closed and* $r_{\mathbf{x}} = n + O(1)$*.*

Proof. First we note that, clearly, $r_{\mathbf{x}} = r_{f(\mathbf{z})}(n) + O(1)$, hence it suffices to prove both statements for $r_{f(\mathbf{z})}$ instead of $r_{\mathbf{x}}$. Then we note that $\mathbf{z}$ and hence $f(\mathbf{z})$ are both uniformly recurrent. Thus we can apply Theorem 11 to $f(\mathbf{z})$. We now consider two cases.

(a) If $f(\mathbf{z})$ is reversal-closed, then, $r_{f(\mathbf{z})} = \frac{1}{2}(\rho_{f(\mathbf{z})} + \mathrm{Pal}_{f(\mathbf{z})})$. Since $f(\mathbf{z})$ is reversal-closed and quasi-Sturmian, we have from Theorem 17(b) that $\mathrm{Pal}_{f(\mathbf{z})} \leq 3$. Hence $r_{f(\mathbf{z})}(n) = \frac{1}{2}\rho_{f(\mathbf{z})}(n) + O(1)$ as desired.

(b) If $f(\mathbf{z})$ is not reversal-closed, then, $r_{f(\mathbf{z})}(n) = \rho_{f(\mathbf{z})}(n)$ for n large enough. By assumption, we have $\rho_{f(\mathbf{z})}(n) = n + C$ for some constant C and for n large enough, so we have $r_{f(\mathbf{z})}(n) = n + O(1)$ as desired.

This ends the proof. □

Remark 40. One can compare the first part of Theorem 39 above with Corollary 21.

4.3 Episturmian sequences

Among several generalizations of Sturmian sequences, episturmian sequences have in particular the property–sometimes even taken as part of their definition–to be reversal-closed. Furthermore, their palindrome complexity has been studied. See the survey [50]; also see the survey [18]. We develop here for these sequences a theorem similar to Theorem 35 above.

Definition 41. Let A be a finite alphabet with cardinality ℓ. A sequence $\mathbf{x}$ over A is *episturmian* if it is reversal-closed and has at most one left special factor of each length. An episturmian sequence $\mathbf{x}$ is ℓ-strict if it has exactly one left special factor of each length and for which every left special factor u of $\mathbf{x}$ has ℓ distinct left extensions in $\mathbf{x}$.

We compute the reflection complexity of episturmian sequences as follows. (Recall that the factor complexity of an ℓ-strict episturmian sequence is given by $\rho_{\mathbf{x}}(n) = (\ell - 1)n + 1$.)

Theorem 42. *Let $\mathbf{x}$ be an ℓ-strict episturmian sequence. Then, for all $n \geq 0$,*

$$r_{\mathbf{x}}(n) = (\ell - 1) \left\lfloor \frac{n+1}{2} \right\rfloor + 1.$$

Proof. Let $\mathbf{x}$ be an ℓ-strict episturmian sequence. The case $n = 0$ is true. Assume that $n \geq 1$. Then by [42, Theorem 7], we have $\rho_{\mathbf{x}}(n) = (\ell - 1)n + 1$. We also know that

$$\mathrm{Pal}_{\mathbf{x}}(n) = \begin{cases} 1, & \text{if } n \text{ is even;} \\ \ell, & \text{if } n \text{ is odd.} \end{cases}$$

Using these together with Theorem 9(b), we deduce the desired result. □

Example 43. For the Tribonacci sequence $\mathbf{tr}$, which is the fixed point of the morphism $0 \mapsto 01, 1 \mapsto 02, 2 \mapsto 0$, we have $r_{\mathbf{tr}}(n) = 2\lfloor (n+1)/2 \rfloor + 1$ for all $n \geq 0$.

4.4 Billiard sequences on a hypercube

Since one interpretation of Sturmian sequences is the binary coding of irrational trajectories on a square billiard table, one can turn to irrational trajectories on a hypercube. The following result can be found in [16]: the first item is in [16, Corollary 1.6] and the second is the main theorem of that paper (which proves a conjecture due to Tamura; note the unexpected symmetry between s and n, where $(s+1)$ is the dimension of the hypercube).

Theorem 44 ([16]). *Let* $\mathbf{x}$ *be an irrational billiard sequence on an* $(s+1)$*-dimensional hypercube.*

(a) The sequence $\mathbf{x}$ *is reversal-closed.*

(b) The factor complexity of $\mathbf{x}$ *is given by*

$$\rho_{\mathbf{x}}(n) = \sum_{k=0}^{\min(s,n)} k! \binom{s}{k} \binom{n}{k}.$$

Remark 45. Note that, if $\mathbf{x}$ is an irrational billiard sequence on an $(s+1)$-dimensional hypercube, the previous result implies that $\rho_{\mathbf{x}}(n) = \Theta(n^s)$. In particular for $s+1=2$, we obtain $\rho_{\mathbf{x}}(n) = n+1$ (which gives back Sturmian sequences), and for $s+1=3$, we obtain $\rho_{\mathbf{x}}(n) = n^2+n+1$, which had been conjectured by Rauzy and proved in [10].

Corollary 46. *Let* $\mathbf{x}$ *be an irrational billiard sequence on an hypercube of dimension* $(s+1)$*. Then its reflection complexity has the property that*

$$r_{\mathbf{x}}(n) \sim \frac{1}{2} \sum_{k=0}^{\min(s,n)} k! \binom{s}{k} \binom{n}{k}$$

when n *tends to infinity. In particular,* $r_{\mathbf{x}}(n) = \Theta(n^s)$ *when* n *tends to infinity.*

Proof. Use Theorem 44 above with Corollary 21. □

4.5 Complementation-symmetric Rote sequences

So-called complementation-symmetric Rote sequences, which were defined and studied in [78], are related to Sturmian sequences as stated below in Theorem 48. In this section, after recalling their definition, we study their reflection complexity.

Definition 47. Let $\mathbf{x}$ be a binary sequence. Then $\mathbf{x}$ is called a *Rote sequence* if its factor complexity satisfies $\rho_{\mathbf{x}}(n) = 2n$ for all $n \geq 1$. The sequence $\mathbf{x}$ is said to be *complementation-symmetric* if its set of factors is closed under the exchange morphism, i.e., if w is a factor of $\mathbf{x}$, so is $E(w)$.

We consider the mapping $\Delta\colon \{0,1\}^+ \to \{0,1\}^*$ defined as follows: $\Delta(a) = a$ for all $a \in \{0,1\}$ and for $n \geq 1$, $\Delta(v(0)v(1)\cdots v(n)) = u(0)\, u(1) \cdots\, u(n-1)$ with $u(i) = (v(i+1)-v(i)) \bmod 2$ for all $i \in \{0,\dots,n-1\}$. There is a natural extension of Δ to sequences: if $\mathbf{x} = (x(n))_{n\geq 0}$ is a binary sequence, then $\Delta(\mathbf{x})$ is the sequence whose nth letter is defined by $(x(n+1) - x(n)) \bmod 2$ for all $n \geq 0$. Observe that $\Delta(\mathbf{x})$ is the sequence of first differences of $\mathbf{x}$, taken modulo 2.

Theorem 48 ([78]). *A binary sequence* $\mathbf{x}$ *is a complementation-symmetric Rote sequence if and only if* $\Delta(\mathbf{x})$ *is Sturmian.*

In fact, with each Sturmian sequence $\mathbf{s}$, there are two associated complementation-symmetric Rote sequences $\mathbf{x}$ and $\mathbf{x}'$ with $\mathbf{x}' = E(\mathbf{x})$. The factors in $\mathbf{s}$ and its corresponding Rote sequences are closely related as shown below.

Proposition 49 ([78]; also see [62, Proposition 2] or [63, Lemma 2.7]). *Let* $\mathbf{s}$ *be a Sturmian sequence and let* $\mathbf{x}$ *be the complementation-symmetric Rote sequence such that* $\mathbf{s} = \Delta(\mathbf{x})$. *Then* u *is a factor of* $\mathbf{s}$ *if and only if both words* v, v' *such that* $u = \Delta(v) = \Delta(v')$ *are factors of* $\mathbf{x}$. *Furthermore, for every* $n \geq 0$, x *occurs at position* n *in* $\mathbf{s}$ *if and only if* v *or* v' *occurs at position* n *in* $\mathbf{x}$.

Lemma 50. *A complementation-symmetric Rote sequence is reversal-closed.*

Proof. Let $\mathbf{x}$ be a complementation-symmetric Rote sequence. Let $\mathbf{s}$ be the Sturmian sequence corresponding to $\mathbf{x}$, i.e., $\mathbf{s} = \Delta(\mathbf{x})$ given by Theorem 48. Consider a factor v of $\mathbf{x}$. Write $u = \Delta(v)$. Since $\mathbf{s}$ is reversal-closed, the word u^R is also a factor of $\mathbf{s}$. Let w and w' be the binary words such that $u^R = \Delta(w) = \Delta(w')$ and $w' = E(w)$. By Proposition 49, both w and w' are factors of $\mathbf{x}$. Now observe that we have either $v^R = w$ or $v^R = w'$. This ends the proof. □

We compute the reflection complexity of Rote sequences as follows.

Theorem 51. *Let* $\mathbf{x}$ *be a complementation-symmetric Rote sequence. Then its reflection complexity satisfies* $r_{\mathbf{x}}(n) = n + 1$ *for all* $n \geq 0$.

Proof. Let $\mathbf{x}$ be a complementation-symmetric Rote sequence. We clearly have $r_{\mathbf{x}}(0) = 1$. Now, for $n \geq 1$, [4, Theorem 8] states that $\mathrm{Pal}_{\mathbf{x}}(n) = 2$. We finish the proof using Lemma 50 and Theorems 9(b). □

5 Reversal-closed and rich sequences

Proposition 52. *Let* $\mathbf{x}$ *be a reversal-closed sequence. Then we have* $r_{\mathbf{x}}(n+1) + r_{\mathbf{x}}(n) \le \rho_{\mathbf{x}}(n+1) + 1$ *for all* $n \ge 0$.

Proof. It is enough to combine Remark 10 and Theorems 17(b) and 9(b). □

Rich sequences have several equivalent definitions. It is known that a word w contain at most $|w|+1$ many palindromic factors [42]. A sequence is called *rich* if each factor contains the maximal number of palindromic factors.

Theorem 53. *Let* $\mathbf{x}$ *be a reversal-closed sequence. Then* $\mathbf{x}$ *is rich if and only if* $r_{\mathbf{x}}(n+1) + r_{\mathbf{x}}(n) = \rho_{\mathbf{x}}(n+1) + 1$ *for all* $n \ge 0$.

Proof. From [24, Theorem 1.1], the sequence $\mathbf{x}$ is rich if and only if the inequality in Theorem 17(b) is an equality. The result then follows from Theorem 9(b). □

Those among binary quasi-Sturmian sequences that are coding of rotations are rich, see [20, Thm. 19].

Corollary 54. *Let* $\mathbf{x}$ *be a binary reversal-closed quasi-Sturmian sequence. There exists a constant* C *such that* $r_{\mathbf{x}}(n+1) + r_{\mathbf{x}}(n) = n + C$ *for* n *large enough.*

Proof. Let C' be a constant such that $\rho_{\mathbf{x}}(n) = n + C'$ for n large enough. It is enough to choose $C = C' + 2$. □

Using Theorem 9(b) and [80, Cor. 2.27 and 2.29], it is possible to bound the reflection complexity of rich sequences as follows.

Proposition 55. *Let* $\mathbf{x}$ *be a rich sequence over an alphabet of* q *letters and write* $\delta = \frac{2}{3(\log 3 - \log 2)}$. *Then* $r_{\mathbf{x}}(n) \le \frac{nq}{2}(2q^2n)^{\delta \log n}(1 + nq^3(2q^2n)^{\delta \log n})$ *for all* $n \ge 1$.

Other sequences have a reflection complexity satisfying the equality of Theorem 53. For instance, it is the case of complementation–symmetric sequences, sequences canonically associated with some specific Parry numbers, and sequences coding particular interval exchange transformations. For more details, see [14, Section 3].

6 Automatic sequences

In this section, we study the reflection complexity of automatic sequences. First, in a positional numeration system U having an adder, we show that if a sequence is U-automatic, then its reflection complexity is a U-regular sequence. Furthermore we show how to effectively compute a linear representation for the sequence, making use of the free software `Walnut` [68, 83]. Next, we explore the reflection complexity of some famous automatic sequences, namely the Thue–Morse, the period-doubling, generalized paperfolding, generalized Golay–Shapiro, and the Baum-Sweet sequences.

6.1 Reflection complexity is computably regular

We now show that the reflection complexity of any automatic sequence is regular.

Theorem 56. *Let $U = (U(n))_{n\geq 0}$ be a positional numeration system such that there is an adder, i.e., addition is recognizable by an automaton reading U-representations, and let $\mathbf{x}$ be a U-automatic sequence. Then $(r_{\mathbf{x}}(n))_{n\geq 0}$ is a U-regular sequence. Furthermore, a linear representation for $(r_{\mathbf{x}}(n))_{n\geq 0}$ is computable from the DFAO for $\mathbf{x}$.*

Proof. Here is a sketch of the proof before we give the details: We create a first-order logical formula asserting that the factor $\mathbf{x}[i..i+n-1]$ is the first occurrence of this factor, or its reversal. Then the number of such i is precisely the reflection complexity at n. From this, we can create a linear representation for the number of such i.

Now some more details. We define the following logical formulas:

$$\begin{aligned} \textsc{FactorEq}(i,j,n) &:= \forall t\ (t<n) \implies \mathbf{x}[i+t] = \mathbf{x}[j+t] \\ \textsc{FactorRevEq}(i,j,n) &:= \forall t\ (t<n) \implies \mathbf{x}[i+t] = \mathbf{x}[(j+n)-(t+1)] \end{aligned} \tag{6}$$

$$\begin{aligned} \textsc{RefComp}(i,n) := \forall j\ (j<i) \implies &((\neg\,\textsc{FactorEq}(i,j,n)) \\ &\wedge\ (\neg\,\textsc{FactorRevEq}(i,j,n))). \end{aligned}$$

Now we use the fundamental result on Büchi arithmetic to translate each of these formulas to their corresponding automata accepting the base-k representation of those pairs (i,n) making the formula true. Next, we use a basic result to convert the automaton for RefComp to the corresponding linear representation computing the reflection complexity. □

Once we have a linear representation for the reflection complexity, we can easily compute it for a given n. Furthermore, we can compare it to a guessed formula, provided that this formula can also be expressed as a linear representation (see [83]). In the next section we carry this out in detail for a number of famous sequences.

6.2 The Thue–Morse and period-doubling sequences

We can compute a linear representation for the reflection complexity $r_{\mathbf{t}}(n)$ of the 2-automatic Thue–Morse sequence, using the same approach as in the preceding section. Here we use the following `Walnut` code:

```
def factoreq_tm "At (t<n) => T[i+t]=T[j+t]"::
def factorreveq_tm "At (t<n) => T[i+t]=T[(j+n)-(t+1)]"::
def rc_tm n "Aj (j<i) => ((~$factoreq_tm(i,j,n))
   & (~$factorreveq_tm(i,j,n)))"::
```

This generates a linear representation of rank 66, which can be minimized to the following.

$$v = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad w = \begin{bmatrix} 1 & 2 & 3 & 4 & 6 & 6 & 10 & 10 & 13 \end{bmatrix}^T$$

$$\mu(0) = \frac{1}{33}\begin{bmatrix} 33 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 33 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 33 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 33 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 33 \\ 0 & 0 & -26 & 0 & 0 & 23 & 10 & -10 & 36 \\ 0 & 0 & -57 & 33 & 0 & 6 & -6 & 6 & 51 \\ 0 & 0 & -79 & 33 & 33 & -5 & -28 & 28 & 51 \\ 0 & 0 & -72 & 0 & 33 & 18 & -18 & 18 & 54 \end{bmatrix}, \quad \mu(1) = \frac{1}{33}\begin{bmatrix} 0 & 33 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 33 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 33 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 33 & 0 \\ 0 & 0 & -24 & 0 & 0 & 39 & -6 & 6 & 18 \\ 0 & 0 & -40 & 0 & 0 & 43 & -10 & 10 & 30 \\ 0 & 0 & -78 & 33 & 33 & 3 & -36 & 36 & 42 \\ 0 & 0 & -86 & 33 & 33 & 5 & -38 & 38 & 48 \\ 0 & 0 & -72 & 0 & 0 & 51 & -18 & 18 & 54 \end{bmatrix}. \tag{7}$$

Recall that Brlek [21], de Luca and Varricchio [39], and Avgustinovich [11] independently gave a simple recurrence for the number of length-n factors of $\mathbf{t}$, namely $\rho_{\mathbf{t}}(2n) = \rho_{\mathbf{t}}(n) + \rho_{\mathbf{t}}(n+1)$ and $\rho_{\mathbf{t}}(2n+1) = 2\rho_{\mathbf{t}}(n+1)$ for $n \geq 2$. As it turns out, there is a simple relationship between $r_{\mathbf{t}}$ and $\rho_{\mathbf{t}}$.

Theorem 57. *Let* $\mathbf{t}$ *be the Thue–Morse sequence.*

(a) For all $n \geq 0$*, we have* $r_{\mathbf{t}}(2n+1) = \rho_{\mathbf{t}}(n+1)$.

(b) For all $n \geq 2$*, we have*

$$r_{\mathbf{t}}(2n) = \begin{cases} \rho_{\mathbf{t}}(n+1) + 1, & \textit{if } \exists m \geq 0 \textit{ with } 3 \cdot 4^{m-1} + 1 \leq n \leq 4^m; \\ \rho_{\mathbf{t}}(n+1), & \textit{otherwise.} \end{cases}$$

(c) *There is an automaton of* 14 *states computing the first difference* $r_{\mathtt{t}}(n+1) - r_{\mathtt{t}}(n)$.

Proof. We prove each item separately.

(a) Above in Equalities (7) we computed a linear representation for $r_{\mathtt{t}}(n)$. From this linear representation we can easily compute one for $r_{\mathtt{t}}(2n+1)$ merely by replacing w with $\mu(1)w$. (Indeed, base-2 representations of integers $2n+1$ all end with 1.)

Next, we can compute a linear representation for $\rho_{\mathtt{t}}(n+1)$ using the following `Walnut` command.

```
def sc_tm_offset n "Aj (j<i) => ~$factoreq_tm(i,j,n+1)":
```

This creates a linear representation of rank 6.

Finally, we use a block matrix construction to compute a linear representation for the difference $r_{\mathtt{t}}(2n+1) - \rho_{\mathtt{t}}(n+1)$ and minimize it; the result is the 0 representation. This computation gives a rigorous proof of item (a).

(b) This identity can be proven in a similar way. We form the linear representation for

$$r_{\mathtt{t}}(2n) - \rho_{\mathtt{t}}(n+1) - [\exists m\colon 3 \cdot 4^{m-1} + 1 \le n \le 4^m],$$

where the last term uses the Iverson bracket. We then minimize the result and obtain the 0 representation.

(c) We can compute a linear representation for the first difference $r_{\mathtt{t}}(n+1) - r_{\mathtt{t}}(n)$, and then use the "semigroup trick" [83, §4.11] to prove that the difference is bounded and find the automaton for it. It is displayed in Figure 1.

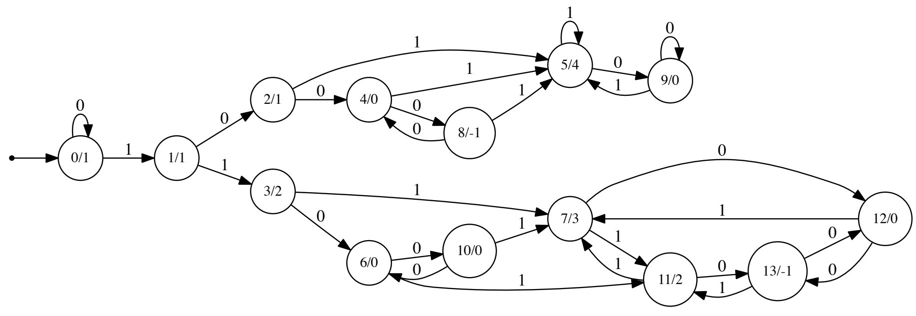

Figure 1: Automaton computing $r_{\mathbf{t}}(n+1)-r_{\mathbf{t}}(n)$ where $\mathbf{t}$ is the Thue–Morse sequence.

These computations rigorously prove the three items of the claim. □

The *period-doubling sequence* $(d(n))_{n\geq 0}$ is a natural companion of the Thue–Morse sequence. Recall that $\mathbf{t}$ is the fixed point, starting with 0, of the morphism defined by $0 \mapsto 01$ and $1 \mapsto 10$. We similarly define $\mathbf{p}$ as the fixed point of the morphism $0 \mapsto 01$ and $1 \mapsto 00$. This gives us that $\mathbf{p}$ is 2-automatic as well. By defining $d(n)$ as the highest power of 2, modulo 2, dividing $n+1$, the sequence $\mathbf{p}$ can be equivalently defined as the sequence $(d(n))_{n\geq 0}$. The close relationship between $\mathbf{t}$ and $\mathbf{p}$ is captured by the identity $\rho_{\mathbf{p}}(n) = \frac{\rho_{\mathbf{t}}(n+1)}{2}$ for all n. We may devise a close analogue of Theorem 57 for the reflection complexity of $\mathbf{p}$, again with the use of `Walnut`. Explicitly, it can be shown that: For all $n \geq 0$, we have $r_{\mathbf{p}}(2n+1) = \rho_{\mathbf{p}}(n)+1$, and, for all $n \geq 2$, we have

$$r_{\mathbf{p}}(2n) = \begin{cases} \rho_{\mathbf{p}}(n+1)-1, & \text{if } \exists m \geq 0 \text{ with } 3\cdot 2^{m-1} \leq n \leq 2^{m+1}-1; \\ \rho_{\mathbf{p}}(n+1)-2, & \text{otherwise,} \end{cases}$$

and we may similarly devise an analogue of part (c) of Theorem 57. Observe that

$$\liminf_{n\to\infty} \frac{r_{\mathbf{p}}(n)}{n} = \frac{3}{4} \quad \text{and} \quad \limsup_{n\to\infty} \frac{r_{\mathbf{p}}(n)}{n} = \frac{5}{6},$$

and similarly for the reflection complexity of $\mathbf{t}$.

6.3 The generalized paperfolding sequences

A paperfolding sequence $\mathbf{p_f}$ is a binary sequence $p_1p_2p_3\cdots$ specified by a sequence of binary unfolding instructions $f_0f_1f_2\cdots$, as the limit of the sequences $\mathbf{p}_{f_0f_1f_2\cdots}$, defined as follows:

$$\mathbf{p}_\varepsilon = \varepsilon \quad \text{and} \quad \mathbf{p}_{f_0\cdots f_{i+1}} = \mathbf{p}_{f_0\cdots f_i}\ f_{i+1}\ E(\mathbf{p}^R_{f_0\cdots f_i}) \ \text{ for all } i \geq 0$$

where E is the exchange morphism. For example, if $\mathbf{f} = 000\cdots$, we get the simplest paperfolding sequence

$$\mathbf{p} = 0010011000110110001001110011011\cdots .$$

Note that a paperfolding sequence is 2-automatc if and only if the sequence of unfolding instructions is eventually periodic [7, Theorem 6.5.4].

Allouche [2], and later, Baake [13] proved that no paperfolding sequence contains a palindrome of length > 13. In fact, even more is true as shown below.

Proposition 58. *No paperfolding sequence contains a reflected factor of length* > 13.

Proof. It suffices to show that no paperfolding sequence contains a reflected factor of length 14. For if this holds, but there is a longer reflected factor x, we could write $x = yz$ where $|y| = 14$. Then $x^R = z^R y^R$, so y would be a reflected factor of length 14, a contradiction.

Now, by a known result on the appearance function of paperfolding sequences [83, Theorem 12.2.1], we know that every length-14 factor of a paperfolding sequence $\mathbf{p_f}$ appears in a prefix of length 109, which is in turn specified by the first 7 unfolding instructions. We can then simply examine each of the 56 length-14 factors of these 128 (finite) words and verify that no factor is reflected. □

We can now prove the following result.

Theorem 59. *Let* $\mathbf{p_f}$ *be a paperfolding sequence. Then*

(a) *For all* $n \geq 13$, *we have* $r_{\mathbf{f}}(n) = \rho_{\mathbf{f}}(n) = 4n$.

(b) *The reflection complexity of every paperfolding sequence is the same, and takes the values* $2, 3, 6, 7, 12, 15, 22, 24, 32, 36, 42, 46$ *for* $1 \leq n \leq 12$.

Proof. We prove each item separately.

(a) For $n \geq 14$, the result follows from combining the results of Allouche [2] and Proposition 58. For $n = 13$, we can verify the claim by explicit enumeration.

(b) The result for $n \geq 13$ follows from (a). For $n < 13$ the result can be verified by enumeration of all length-109 prefixes of paperfolding sequences specified by instructions of length 7.

This ends the proof. □

6.4 The generalized Golay–Shapiro sequences

A generalized Golay–Shapiro sequence $\mathbf{g}$ is defined by taking the running sum, modulo 2, of a paperfolding sequence $\mathbf{p_f}$. The famous Golay–Shapiro sequence (also called the Rudin–Shapiro sequence) [51, 52, 79, 84] corresponds to the case of unfolding instructions $0(01)^\omega$ [4, Def. 6]. Note that the 2-automaticity of a generalized Golay–Shapiro sequence follows from that of its corresponding generalized paperfolding sequence.

We can prove the analogue of Proposition 58.

Proposition 60. *No generalized Golay–Shapiro sequence contains a reflected factor of length* > 14.

Proof. As above, it suffices to show that no Golay–Shapiro sequence contains a reflected factor of length 15.

Now, by a known result on the recurrence function of generalized Golay–Shapiro sequences [5, Proposition 4.1], we know that every length-15 factor of a paperfolding sequence $\mathbf{p_f}$ appears in a prefix of length 2408, which is in turn specified by the first 12 unfolding instructions. We can then simply examine each of the 60 length-15 factors of these 4096 (finite) words and verify that no factor is reflected. □

We can now prove the following result.

Theorem 61. *Let* $\mathbf{g}$ *be a generalized Golay–Shapiro sequence.*

(a) For all $n \geq 15$*, we have* $r_{\mathbf{g}}(n) = \rho_{\mathbf{g}}(n) = 8n - 8$.

(b) *The reflection complexity of every generalized Golay–Shapiro sequence is the same, and takes the values* 2*,* 3*,* 6*,* 10*,* 14*,* 22*,* 30*,* 42*,* 48*,* 62*,* 72*,* 83*,* 92*,* 103 *for* $1 \le n \le 14$*.*

Proof. We prove each item separately.

(a) For $n \ge 15$, the result follows from combining the results of Allouche and Bousquet-Melou [5] and Proposition 60.

(b) The result for $n \ge 15$ follows from (a). For $n < 15$ the result can be verified by enumeration of all length-2408 prefixes of paperfolding sequences specified by instructions of length 12.

This ends the proof. □

6.5 The Baum–Sweet sequence

Let the *Baum–Sweet sequence*

$$\mathbf{b} = 11011001010010011001000001001001010010 01 \cdots$$

be defined by $\mathbf{b}(0) = 1$ and for $n \ge 1$, $b(n)$ is 1 if the base-2 expansion of n contains no block of successive zeros of odd length and 0 otherwise. It is 2-automatic as well. The factor complexity function for $\mathbf{b}$ is such that

$$(\rho_{\mathbf{b}}(n))_{n\ge 0} = 1, 2, 4, 7, 13, 17, 21, 27, 33, 38, 45, 52, 59, 65, 70, \ldots \qquad (8)$$

and the reflection complexity function for $\mathbf{b}$ starts with

$$(r_{\mathbf{b}}(n))_{n\ge 0} = 1, 2, 3, 5, 8, 11, 13, 17, 21, 25, 30, 35, 40, 46, 50, 56, \ldots. \qquad (9)$$

We can again compute a linear representation for $r_{\mathbf{b}}(n)$ using the following `Walnut` code:

```
def factoreq_bs "At (t<n) => BS[i+t]=BS[j+t]"::
def factorreveq_bs "At (t<n) => BS[i+t]=BS[(j+n)-(t+1)]"::
def rc_bs n "Aj (j<i) => ((~$factoreq_bs(i,j,n))
   & (~$factorreveq_bs(i,j,n)))"::
```

This gives us a linear representation of rank 90. From this linear representation, a computation proves the following result.

Corollary 62. *Let* $\mathbf{b}$ *be the Baum-Sweet sequence. Then the first difference of the sequence* $r_{\mathbf{b}}(n)$ *is* 2*-automatic, over the alphabet* $\{1, 2, \ldots, 8\}$*.*

7 Further directions

We conclude the paper by considering some further research directions to pursue in relation to reflection complexities of sequences and by raising some open problems.

We encourage further explorations of the evaluation of $r_{\mathbf{x}}$ for sequences $\mathbf{x}$ for which properties of $\mathrm{Pal}_{\mathbf{x}}$ and/or $\rho_{\mathbf{x}}$ are known, especially if `Walnut` cannot be used directly in the investigation of $r_{\mathbf{x}}$. For example, by letting the *Chacon sequence* $\mathbf{c}$ be the fixed point of the morphism $0 \mapsto 0010$ and $1 \mapsto 1$, it is known that $\mathrm{Pal}_{\mathbf{c}}(n) = 0$ for all $n \geq 13$. Also, its factor complexity satisfies $\rho_{\mathbf{c}}(n) = 2n - 1$ for $n \geq 2$ [45]. We have

$$(r_{\mathbf{c}}(n))_{n\geq 0} = 1, 2, 2, 4, 4, 6, 7, 10, 11, 14, 16, 20, 23, 25, 27, 29, 31, 33, \ldots.$$

This sequence is not automatic in a given so-called addable numeration system (where there is an adder). Therefore, we cannot use `Walnut`, in this case. However, an inductive argument can be applied to prove that $r_{\mathbf{c}}(n) = \rho_{\mathbf{c}}(n)$ for all $n \geq 13$.

Question 63. To what extent can the reflection complexity be used to discriminate between different families of sequences, by analogy with our characterizations of Sturmian and eventually periodic sequences?

The complexity function $\mathrm{Unr}_{\mathbf{x}}$ defined above may be of interest in its own right, as is the case with the "reflection-free" complexity function enumerating factors such that the reversal of every sufficiently large factor is not a factor.

Question 64. How can Theorem 57 be generalized with the use of standard generalizations of the Thue–Morse sequence?

For example, if we let

$$\mathbf{t3} = 011212201220200112202001200\cdots$$

denote the generalized Thue–Morse sequence for which the nth term $\mathbf{t3}(n)$ is equal to the number of 1's, modulo 3, in the base-2 expansion of n, it can be shown that $r_{\mathbf{t3}}(n) = \rho_{\mathbf{t3}}(n)$ for all $n \geq 3$, and it appears that a similar property holds for the cases given by taking the number of 1's modulo $\ell > 4$.

Question 65. What is the reflection complexity of the Thue–Morse sequence over polynomial extractions, with regard to the work of Yossi Moshe [67]?

Question 66. How could the upper bound in Theorem 20 be improved? If $r_{\mathbf{x}}(n)$ is of the form $\Omega(n)$, then how could this be improved?

Question 67. How does the reflection complexity compare with other complexity functions, as in the complexity functions listed in Section 1?

This leads us to ask about the respective growths of the complexities listed in Section 1, in particular for morphic sequences. In this direction, recall that the factor complexity of a morphic sequence is either $\Theta(1)$, $\Theta(n)$, $\Theta(n \log \log n)$, $\Theta(n \log n)$ or $\Theta(n^2)$, see [70] (more details can be found, e.g., in [35]; also see [41]). As an illustration with a result that has not been already cited above, a comparison between growths for the factor complexity and the Lempel-Ziv complexity can be found in [36]. We end with an easy result for the growth of the reflection complexity in the case of morphic sequences.

Proposition 68. *The reflection complexity of a morphic sequence is either* $\Theta(1)$, $\Theta(n)$, $\Theta(n \log \log n)$, $\Theta(n \log n)$ *or* $\Theta(n^2)$.

Proof. Use the inequalities in Theorem 9: for any sequence $\mathbf{x}$ and for all $n \geq 0$, we have $\frac{1}{2}\rho_{\mathbf{x}}(n) \leq r_{\mathbf{x}}(n) \leq \rho_{\mathbf{x}}(n)$. □

Acknowledgments

We thank Boris Adamczewski for discussions. John Campbell is grateful to acknowledge support from a Killam Postdoctoral Fellowship from the Killam Trusts. Manon Stipulanti is an FNRS Research Associate supported by the Research grant 1.C.104.24F. The research of Jeffrey Shallit is supported by NSERC grant 2018-04118.

References

[1] Jean-Paul Allouche. Sur la complexité des suites infinies. *Bull. Belg. Math. Soc.—Simon Stevin*, 1(2):133–143, 1994.

[2] Jean-Paul Allouche. Schrödinger operators with Rudin-Shapiro potentials are not palindromic. *J. Math. Phys.*, 38(4):1843–1848, 1997.

[3] Jean-Paul Allouche. Nouveaux résultats de transcendance de réels à développement non aléatoire. *Gaz. Math.*, 84:19–34, 2000.

[4] Jean-Paul Allouche, Michael Baake, Julien Cassaigne, and David Damanik. Palindrome complexity. *Theoret. Comput. Sci.*, 292(1):9–31, 2003.

[5] Jean-Paul Allouche and Mireille Bousquet-Mélou. Facteurs des suites de Rudin-Shapiro généralisées. *Bull. Belg. Math. Soc.*, 1:145–164, 1994.

[6] Jean-Paul Allouche and Jeffrey Shallit. Generalized perturbed symmetry. *European J. Combin.*, 19:401–411, 1998.

[7] Jean-Paul Allouche and Jeffrey Shallit. *Automatic Sequences. Theory, Applications, Generalizations.* Cambridge: Cambridge University Press, 2003.

[8] Jean-Paul Allouche and Jia-Yan Yao. Opacity complexity of automatic sequences. the general case, 2024. Preprint available at `http://arxiv.org/abs/2404.13601`.

[9] Hayri Ardal, Tom Brown, Veselin Jungić, and Julian Sahasrabudhe. On abelian and additive complexity in infinite words. *Integers*, 12(5):795–804, a21, 2012.

[10] Pierre Arnoux, Christian Mauduit, Iekata Shiokawa, and Jun-ichi Tamura. Complexity of sequences defined by billiard in the cube. *Bull. Soc. Math. France*, 122:1–12, 1994.

[11] Sergeĭ V. Avgustinovich. The number of different subwords of given length in the Morse-Hedlund sequence. *Sibirsk. Zh. Issled. Oper.*, 1(2):3–7, 103, 1994.

[12] Sergeĭ V. Avgustinovich, Dmitriĭ G. Fon-Der-Flaass, and Anna E. Frid. Arithmetical complexity of infinite words. In *Words, languages & combinatorics, III (Kyoto, 2000)*, pages 51–62. World Sci. Publ., River Edge, NJ, 2003.

[13] Michael Baake. A note on palindromicity. *Lett. Math. Phys.*, 49(3):217–227, 1999.

[14] Peter Baláži, Zuzana Masáková, and Edita Pelantová. Factor versus palindromic complexity of uniformly recurrent infinite words. *Theoret. Comput. Sci.*, 380(3):266–275, 2007.

[15] L'ubomíra Balková, Edita Pelantová, and Štěpán Starosta. Sturmian jungle (or garden?) on multiliteral alphabets. *RAIRO Theor. Inform. Appl.*, 44(4), 2010.

[16] Yuliy Baryshnikov. Complexity of trajectories in rectangular billiards. *Comm. Math. Phys.*, 174:43–56, 1995.

[17] Jason P. Bell and Jeffrey Shallit. Lie complexity of words. *Theoret. Comput. Sci.*, 927:98–108, 2022.

[18] Jean Berstel. Sturmian and episturmian words (a survey of some recent results). In *Algebraic informatics*, volume 4728 of *Lecture Notes in Comput. Sci.*, pages 23–47. Springer, Berlin, 2007.

[19] Jean Berstel, Luc Boasson, Olivier Carton, and Isabelle Fagnot. Infinite words without palindrome, 2009. Preprint available at `https://arxiv.org/abs/0903.2382`.

[20] Alexandre Blondin-Massé, Srecko Brlek, Sébastien Labbé, and Laurent Vuillon. Palindromic complexity of codings of rotations. *Theoret. Comput. Sci.*, 412(46):6455–6463, 2011.

[21] Srećko Brlek. Enumeration of factors in the Thue-Morse word. *Discrete Appl. Math.*, 24(1-3):83–96, 1989.

[22] Srećko Brlek and Annie Ladouceur. A note on differentiable palindromes. *Theoret. Comput. Sci.*, 302:167–178, 2003.

[23] Srećko Brlek and Christophe Reutenauer. Complexity and palindromic defect of infinite words. *Theoret. Comput. Sci.*, 412:493–497, 2011.

[24] Michelangelo Bucci, Alessandro De Luca, Amy Glen, and Luca Q. Zamboni. A connection between palindromic and factor complexity using return words. *Adv. in Appl. Math.*, 42(1):60–74, 2009.

[25] Yann Bugeaud and Dong Han Kim. A new complexity function, repetitions in Sturmian words, and irrationality exponents of Sturmian numbers. *Trans. Amer. Math. Soc.*, 371:3281–3308, 2019.

[26] Arturo Carpi and Valerio D'Alonzo. On factors of synchronized sequences. *Theoret. Comput. Sci.*, 411(44-46):3932–3937, 2010.

[27] Julien Cassaigne. Special factors of sequences with linear subword complexity. In *Developments in language theory, II (Magdeburg, 1995)*, pages 25–34. World Sci. Publ., River Edge, NJ, 1996.

[28] Julien Cassaigne. Sequences with grouped factors. In Symeon Bozapalidis, editor, *Proceedings of the 3rd International Conference Developments in Language Theory, DLT 1997, Thessaloniki, Greece, July 20-23, 1997*, pages 211–222. Aristotle University of Thessaloniki, 1997.

[29] Julien Cassaigne, Gabriele Fici, Marinella Sciortino, and Luca Q. Zamboni. Cyclic complexity of words. *J. Comb. Theory, Ser. A*, 145:36–56, 2017.

[30] Julien Cassaigne, France Gheeraert, Antonio Restivo, Giuseppe Romana, Marinella Sciortino, and Manon Stipulanti. New string attractor-based complexities for infinite words, 2023. Preprint available at `https://orbi.uliege.be/handle/2268/309672`.

[31] Julien Cassaigne, Idrissa Kaboré, and Théodore Tapsoba. On a new notion of complexity on infinite words. *Acta Univ. Sapientiae Math.*, 2:127–136, 2010.

[32] Julien Cassaigne, Tomi Kärki, and Luca Q. Zamboni. Relationally periodic sequences and subword complexity. In *Developments in Language Theory. 12th International Conference, DLT 2008, Kyoto, Japan, September 16–19, 2008. Proceedings*, pages 196–205. Berlin: Springer, 2008.

[33] Émilie Charlier, Svetlana Puzynina, and Luca Q. Zamboni. On a group theoretic generalization of the Morse-Hedlund theorem. *Proc. Amer. Math. Soc.*, 145(8):3381–3394, 2017.

[34] A. L. Chernyat'ev. Words with a minimal growth function. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.*, 6:42–44, 78–79, 2008.

[35] Célia Cisternino. *Fonction de complexité en facteurs et un théorème de PANSIOT*. Mémoire de Master, Université de Liège, Belgique, 2018.

[36] Sorin Constantinescu and Lucian Ilie. The Lempel-Ziv complexity of fixed points of morphisms. *SIAM J. Discrete Math.*, 21:466–481, 2007.

[37] Ethan M. Coven. Sequences with minimal block growth. II. *Math. Syst. Theory*, 8:376–382, 1975.

[38] David Damanik and Daniel Lenz. Uniform spectral properties of one-dimensional quasicrystals. IV. Quasi-Sturmian potentials. *J. Anal. Math.*, 90:115–139, 2003.

[39] Aldo de Luca and Stefano Varricchio. Some combinatorial properties of the Thue-Morse sequence and a problem in semigroups. *Theoret. Comput. Sci.*, 63(3):333–348, 1989.

[40] F. Michel Dekking. On the structure of self-generating sequences. In *Séminaire de Théorie des Nombres, Bordeaux, 1980–1981*. Univ. Bordeaux I, Talence, 1981. Exp. No. 31, 6 pp.

[41] Rostislav Deviatov. On subword complexity of morphic sequences. In *Computer science—theory and applications*, volume 5010 of *Lecture Notes in Comput. Sci.*, pages 146–157. Springer, Berlin, 2008.

[42] Xavier Droubay, Jacques Justin, and Giuseppe Pirillo. Episturmian words and some constructions of de Luca and Rauzy. *Theoret. Comput. Sci.*, 255(1-2):539–553, 2001.

[43] Xavier Droubay and Giuseppe Pirillo. Palindromes and Sturmian words. *Theoret. Comput. Sci.*, 223(1-2):73–85, 1999.

[44] Diane Erdmann and Sean Murphy. An approximate distribution for the maximum order complexity. *Des. Codes Cryptography*, 10(3):325–339, 1997.

[45] Sébastien Ferenczi. Les transformations de chacon : combinatoire, structure géométrique, lien avec les systèmes de complexité $2n+1$. *Bull. Soc. Math. Fr.*, 123(2):271–292, 1995.

[46] Sébastien Ferenczi. Complexity of sequences and dynamical systems. *Discrete Math.*, 206(1-3):145–154, 1999.

[47] Sébastien Ferenczi and Zoltán Kása. Complexity for finite factors of infinite sequences. *Theoret. Comput. Sci.*, 218(1):177–195, 1999.

[48] Gabriele Fici and Svetlana Puzynina. Abelian combinatorics on words: a survey. *Comput. Sci. Rev.*, 47:21, 2023. Id/No 100532.

[49] Gabriele Fici and Luca Q. Zamboni. On the least number of palindromes contained in an infinite word. *Theoret. Comput. Sci.*, 481:1–8, 2013.

[50] Amy Glen and Jacques Justin. Episturmian words: a survey. *RAIRO Theor. Inform. Appl.*, 43(3):403–442, 2009.

[51] Marcel J. E. Golay. Multi-slit spectrometry. *J. Optical Society Amer.*, 39:437–444, 1949.

[52] Marcel J. E. Golay. Static multislit spectrometry and its application to the panoramic display of infrared spectra. *J. Optical Society Amer.*, 41:468–472, 1951.

[53] A. Hof, O. Knill, and B. Simon. Singular continuous spectrum for palindromic Schrödinger operators. *Comm. Math. Phys.*, 174:149–159, 1995.

[54] Teturo Kamae and Luca Zamboni. Sequence entropy and the maximal pattern complexity of infinite words. *Ergodic Theory Dyn. Syst.*, 22(4):1191–1199, 2002.

[55] Juhani Karhumaki, Aleksi Saarela, and Luca Q. Zamboni. On a generalization of abelian equivalence and complexity of infinite words. *J. Comb. Theory, Ser. A*, 120(8):2189–2206, 2013.

[56] Andrey Kolmogorov. On tables of random numbers. *Sankhyā, Ser. A*, 25:369–376, 1963.

[57] Colin Krawchuk and Narad Rampersad. Cyclic complexity of some infinite words and generalizations. *Integers*, 18A:paper a12, 13 pp., 2018.

[58] Abraham Lempel and Jacob Ziv. On the complexity of finite sequences. *IEEE Trans. Inf. Theory*, 22:75–81, 1976.

[59] M. Lothaire. *Combinatorics on words*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1997.

[60] M. Lothaire. *Algebraic combinatorics on words*, volume 90 of *Encycl. Math. Appl.* Cambridge: Cambridge University Press, 2002.

[61] Sergey Luchinin and Svetlana Puzynina. Sequences of symmetry groups of infinite words. *Discrete Math.*, 346(1):Paper No. 113171, 15, 2023.

[62] Kateřina Medková. Return words and derivated sequences to Rote sequences, 2018. Available at https://www.labri.fr/perso/vdelecro/jm2018/215749.pdf.

[63] Kateřina Medková, Edita Pelantová, and Laurent Vuillon. Derived sequences of complementary symmetric Rote sequences. *RAIRO Theor. Inform. Appl.*, 53:125–151, 2019.

[64] Filippo Mignosi and Antonio Restivo. A new complexity function for words based on periodicity. *Internat. J. Algebra Comput.*, 23:963–987, 2013.

[65] T. K. Subrahmonian Moothathu. Eulerian entropy and non-repetitive subword complexity. *Theoret. Comput. Sci.*, 420:80–88, 2012.

[66] Marston Morse and Gustav A. Hedlund. Symbolic Dynamics. *Amer. J. Math.*, 60:815–866, 1938.

[67] Yossi Moshe. On the subword complexity of Thue-Morse polynomial extractions. *Theoret. Comput. Sci.*, 389(1-2):318–329, 2007.

[68] Hamoon Mousavi. Automatic theorem proving in Walnut, 2016. Preprint available at https://arxiv.org/abs/1603.06017.

[69] Harald Niederreiter. Linear complexity and related complexity measures for sequences. In *Progress in Cryptology – INDOCRYPT 2003. 4th International Conference on Cryptology in India, New Delhi, India, December 8–10, 2003. Proceedings*, pages 1–17. Berlin: Springer, 2003.

[70] Jean-Jacques Pansiot. Complexité des facteurs des mots infinis engendrés par morphismes itérés. In *Automata, languages and programming (Antwerp, 1984)*, volume 172 of *Lecture Notes in Comput. Sci.*, pages 380–389. Springer, Berlin, 1984.

[71] Olga Parshina and Mickaël Postic. Open and closed complexity of infinite words, 2020. Preprint available at https://arxiv.org/abs/2005.06254.

[72] Michael E. Paul. Minimal symbolic flows having minimal block growth. *Math. Systems Theory*, 8:309–315, 1974/75.

[73] Jarkko Peltomäki. Introducing privileged words: privileged complexity of Sturmian words. *Theoret. Comput. Sci.*, 500:57–67, 2013.

[74] Narad Rampersad and Jeffrey Shallit. Words avoiding reversed subwords. *J. Comb. Math. Comb. Comput.*, 54:157–164, 2005.

[75] Michel Rigo and Arnaud Maes. More on generalized automatic sequences. *J. Autom. Lang. Comb.*, 7(3):351–376, 2002.

[76] Michel Rigo and Pavel Salimov. Another generalization of abelian equivalence: binomial complexity of infinite words. *Theoret. Comput. Sci.*, 601:47–57, 2015.

[77] Michel Rigo, Manon Stipulanti, and Whiteland Markus. Gapped binomial complexities in sequences. In *IEEE International Symposium on Information Theory (ISIT), Tapai, Taiwan*, pages 1294–1299. IEEE, 2023.

[78] Günter Rote. Sequences with subword complexity $2n$. *J. Number Theory*, 46(2):196–213, 1994.

[79] Walter Rudin. Some theorems on Fourier coefficients. *Proc. Amer. Math. Soc.*, 10:855–859, 1959.

[80] Josef Rukavicka. Upper bound for palindromic and factor complexity of rich words. *RAIRO Theor. Inform. Appl.*, 55:Paper No. 1, 15 pp., 2021.

[81] Luke Schaeffer and Jeffrey Shallit. String attractors for automatic sequences, 2021. Preprint available at `https://arxiv.org/abs/2012.06840`.

[82] Jeffrey Shallit. A generalization of automatic sequences. *Theoret. Comput. Sci.*, 61(1):1–16, 1988.

[83] Jeffrey Shallit. *The Logical Approach to Automatic Sequences—Exploring Combinatorics on Words with* `Walnut`, volume 482 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2023.

[84] Harold S. Shapiro. Extremal problems for polynomials and power series. Master's thesis, MIT, 1952.

[85] Neil J. A. Sloane. The On-Line Encyclopedia of Integer Sequences. Available at https://oeis.org.

Jean-Paul Allouche
CNRS, IMJ-PRG
Sorbonne, 4 Place Jussieu
75252 Paris Cedex 05, France
jean-paul.allouche@imj-prg.fr

John M. Campbell
Department of Mathematics and Statistics, Dalhousie University
Halifax, 6299 South St, NS B3H 4R2, Canada
jmaxwellcampbell@gmail.com

Jeffrey Shallit
School of Computer Science, University of Waterloo
Waterloo, ON N2L 3G1, Canada
shallit@uwaterloo.ca

Manon Stipulanti
Department of Mathematics, University of Liège
4000 Liège, Allée de la Découverte 12, Belgium
m.stipulanti@uliege.be