A new approach to the 2-regularity of the ℓ-abelian complexity of 2-automatic sequences

Aline Parreau*
LIRIS
University of Lyon, CNRS
Lyon, France
aline.parreau@univ-lyon1.fr

Michel Rigo
Department of Mathematics
University of Liege
Liege, Belgium
M.Rigo@ulg.ac.be

Eric Rowland†
Department of Mathematics
University of Liege
Liege, Belgium
erowland@ulg.ac.be

Élise Vandomme
Department of Mathematics
University of Liege
Liege, Belgium
Institut Fourier
University of Grenoble
Grenoble, France
E.Vandomme@ulg.ac.be

January 16, 2015

Abstract

We prove that a sequence satisfying a certain symmetry property is 2-regular in the sense of Allouche and Shallit, i.e., the $\mathbb{Z}$-module generated by its 2-kernel is finitely generated. We apply this theorem to develop a general approach for studying the ℓ-abelian complexity of 2-automatic sequences. In particular, we prove that the period-doubling word and the Thue–Morse word have 2-abelian complexity sequences that are 2-regular. Along the way, we also prove that the 2-block codings of these two words have 1-abelian complexity sequences that are 2-regular.

1 Introduction

This paper is about some structural properties of integer sequences that occur naturally in combinatorics on words. Since the fundamental work of Cobham [8], the so-called automatic sequences have been extensively studied. We refer the reader to [3] for basic

*This work has been done when this author was an FNRS post-doctoral fellow at the University of Liege.
†BeIPD-COFUND post-doctoral fellow at the University of Liege.

definitions and properties. These infinite words over a finite alphabet can be obtained by iterating a prolongable morphism of constant length to get an infinite word (and then, an extra letter-to-letter morphism, also called coding, may be applied). As a fundamental example, the *Thue–Morse word* $\mathbf{t} = \sigma^\omega(0) = 0110100110010110\cdots$ is a fixed point of the morphism σ over the free monoid $\{0,1\}^*$ defined by $\sigma(0) = 01$, $\sigma(1) = 10$. Similarly, the *period-doubling word* $\mathbf{p} = \psi^\omega(0) = 0100010101000100010\cdots$ is a fixed point of the morphism ψ over $\{0,1\}^*$ defined by $\psi(0) = 01$, $\psi(1) = 00$. We will discuss again these two examples of 2-automatic sequences.

Since an infinite word is just a sequence over $\mathbb{N}$ taking values in a finite alphabet, we use the terms 'infinite word' and 'sequence' interchangeably.

Let $k \geq 2$ be an integer. One characterization of k-automatic sequences is that their k-kernels are finite; see [9] or [3, Section 6.6].

Definition 1. The *k-kernel* of a sequence $\mathbf{s} = s(n)_{n\geq 0}$ is the set

$$\mathcal{K}_k(\mathbf{s}) = \{s(k^i n + j)_{n\geq 0} : i \geq 0 \text{ and } 0 \leq j < k^i\}.$$

For instance, the 2-kernel $\mathcal{K}_2(\mathbf{t})$ of the Thue–Morse word contains exactly two elements, namely $\mathbf{t}$ and $\sigma^\omega(1)$.

A natural generalization of automatic sequences to sequences on an infinite alphabet is given by the notion of k-regular sequences. We will restrict ourselves to sequences taking integer values only.

Definition 2. Let $k \geq 2$ be an integer. A sequence $\mathbf{s} = s(n)_{n\geq 0} \in \mathbb{Z}^{\mathbb{N}}$ is *k-regular* if $\langle \mathcal{K}_k(\mathbf{s}) \rangle$ is a finitely-generated $\mathbb{Z}$-module, i.e., there exist a finite number of sequences $t_1(n)_{n\geq 0}, \ldots, t_\ell(n)_{n\geq 0}$ such that every sequence in the k-kernel $\mathcal{K}_k(\mathbf{s})$ is a $\mathbb{Z}$-linear combination of the t_r's. Otherwise stated, for all $i \geq 0$ and for all $j \in \{0, \ldots, k^i - 1\}$, there exist integers $c_1, \ldots, c_\ell$ such that

$$\forall n \geq 0, \quad s(k^i n + j) = \sum_{r=1}^{\ell} c_r\, t_r(n).$$

There are many natural examples of k-regular sequences [1, 2]. There is a convenient matrix representation for k-regular sequences which leads to an efficient algorithm for computing the values of such a sequence (and many related quantities). See also [4, Chapter 5] for connections with rational series. In particular, a sequence taking finitely many values is k-regular if and only if it is k-automatic. The k-regularity of a sequence provides us with structural information about how the different terms are related to each other.

A classical measure of complexity of an infinite word $\mathbf{x}$ is its *factor complexity* $\mathcal{P}^{(\infty)}_{\mathbf{x}} : \mathbb{N} \to \mathbb{N}$ which maps n to the number of distinct factors of length n occurring in $\mathbf{x}$. It is well known that a k-automatic sequence $\mathbf{x}$ has a k-regular factor complexity function and the sequence $(\mathcal{P}^{(\infty)}_{\mathbf{x}}(n+1) - \mathcal{P}^{(\infty)}_{\mathbf{x}}(n))_{n\geq 0}$ is k-automatic. See [6, 7] for a proof and relevant extensions. As an example, again for the Thue–Morse word, we have

$$\mathcal{P}^{(\infty)}_{\mathbf{t}}(2n+1) = 2\mathcal{P}^{(\infty)}_{\mathbf{t}}(n+1) \text{ and } \mathcal{P}^{(\infty)}_{\mathbf{t}}(2n) = \mathcal{P}^{(\infty)}_{\mathbf{t}}(n+1) + \mathcal{P}^{(\infty)}_{\mathbf{t}}(n)$$

for all $n \geq 2$. See also [10] where a formula was obtained for the factor complexity of fixed points of some uniform morphisms.

Recently there has been a renewal of interest in abelian notions arising in combinatorics on words (e.g., avoiding abelian or ℓ-abelian patterns, abelian bordered words, etc.). For instance, two finite words u and v are *abelian equivalent* if one is obtained by permuting the letters of the other one, i.e., the two words share the same Parikh vector, $\Psi(u) = \Psi(v)$. Since the Thue–Morse word is an infinite concatenation of factors 01 and 10, this word is *abelian periodic* of period 2. The *abelian complexity* of an infinite word $\mathbf{x}$ is a function $\mathcal{P}^{(1)}_{\mathbf{x}} : \mathbb{N} \to \mathbb{N}$ which maps n to the number of distinct factors of length n occurring in $\mathbf{x}$, counted up to abelian equivalence. Madill and Rampersad [15] provided the first example of regularity in this setting: the abelian complexity of the paper-folding word (which is another typical example of an automatic sequence) is unbounded and 2-regular.

Let $\ell \geq 1$ be an integer. Based on [12] the notions of abelian equivalence and thus abelian complexity were recently extended to ℓ-abelian equivalence and ℓ-abelian complexity [13].

Definition 3. Let u, v be two finite words. We let $|u|_v$ denote the number of occurrences of the factor v in u. Two finite words x and y are *ℓ-abelian equivalent* if $|x|_v = |y|_v$ for all words v of length $|v| \leq \ell$.

As an example, the words 011010011 and 001101101 are 2-abelian equivalent but not 3-abelian equivalent (the factor 010 occurs in the first word but not in the second one). Hence one can define the function $\mathcal{P}^{(\ell)}_{\mathbf{x}} : \mathbb{N} \to \mathbb{N}$ which maps n to the number of distinct factors of length n occurring in the infinite word $\mathbf{x}$, counted up to ℓ-abelian equivalence. That is, we count ℓ-abelian equivalence classes partitioning the set of factors $\mathrm{Fac}_{\mathbf{x}}(n)$ of length n occurring in $\mathbf{x}$. In particular, for any infinite word $\mathbf{x}$, we have for all $n \geq 0$

$$\mathcal{P}^{(1)}_{\mathbf{x}}(n) \leq \cdots \leq \mathcal{P}^{(\ell)}_{\mathbf{x}}(n) \leq \mathcal{P}^{(\ell+1)}_{\mathbf{x}}(n) \leq \cdots \leq \mathcal{P}^{(\infty)}_{\mathbf{x}}(n).$$

In this paper, we show that both the period-doubling word and the Thue–Morse word have 2-abelian complexity sequences which are 2-regular. The computations and arguments leading to these results permit us to exhibit some similarities between the two cases and a quite general scheme that we hope can be used again to prove additional regularity results. Indeed, one conjectures that *any k-automatic sequence has an ℓ-abelian complexity function that is k-regular.*

We mention some other papers containing related work. In [14], the authors studied the asymptotic behavior of $\mathcal{P}^{(\ell)}_{\mathbf{t}}(n)$ and also derived some recurrence relations[1] showing that the abelian complexity $\mathcal{P}^{(1)}_{\mathbf{p}}(n)_{n\geq 0}$ of the period-doubling word $\mathbf{p}$ is 2-regular. In [5], the abelian complexity of the fixed point $\mathbf{v}$ of the non-uniform morphism $0 \mapsto 012, 1 \mapsto 02, 2 \mapsto 1$ is studied and the authors obtain results similar to those discussed in this paper. Even though the authors of [5] are not directly interested in the k-regularity of

[1] It seems that there is some subtle error in the relation for $\mathcal{P}^{(1)}_{\mathbf{p}}(4n+2)$ proposed in [14, Lemma 6]. Correct relations are given by [5, Proposition 2] and could also be obtained by Theorem 4 and Proposition 47.

$\mathcal{P}^{(1)}_{\mathbf{v}}(n)_{n\geq 0}$, they derive recurrence relations. From these relations, following the approach described in this paper, one can possibly prove some regularity result. In particular, the result of replacing in $\mathbf{v}$ all 2's by 0's leads back to the period-doubling word. Hence, Blanchet-Sadri et al. also proved some other relations about the abelian complexity of $\mathbf{p}$.

Given the first few terms of a sequence, one can easily conjecture the potential k-regularity of this sequence by exhibiting relations that should be satisfied; see [2, Section 6] for such a "predictive" algorithm that recognizes regularity. Of course, in such an algorithm, a finite examination does not lead to a proof of the k-regularity of a sequence. The first few terms of the 2-abelian complexity $\mathcal{P}^{(2)}_{\mathbf{t}}(n)_{n\geq 0}$ of the Thue–Morse word are

$$1, 2, 4, 6, 8, 6, 8, 10, 8, 6, 8, 8, 10, 10, 10, 8, 8, 6, 8, 10, 10, 8, 10, 12, 12, 10, 12, 12, \ldots.$$

The second and last authors of this paper conjectured the 2-regularity of the sequence $\mathcal{P}^{(2)}_{\mathbf{t}}(n)_{n\geq 0}$ (and proved some recurrence relations for this sequence) [17]. Recently, after hearing a talk given by the last author during the *Representing Streams II* meeting in January 2014, Greinecker proved the recurrence relations needed to prove the 2-regularity of this sequence [11]. Hopefully, the two approaches are complementary: in this paper, we prove 2-regularity without exhibiting the explicit recurrence relations.

Let us now describe the content and organization of this paper.

In Section 2 we prove Theorem 4, which establishes the 2-regularity of a large family of sequences satisfying a recurrence relation with a parameter c and 2^{ℓ_0} initial conditions. The form of the recurrence implies that sequences in this family exhibit a reflection symmetry in the values taken over each interval $[2^\ell, 2^{\ell+1})$ for $\ell \geq \ell_0$. For the special case of the Thue–Morse word, a similar property is shown in [11]. Computer experiments suggest that many 2-abelian complexity functions satisfy such a reflection property.

Theorem 4. *Let $\ell_0 \geq 0$ and $c \in \mathbb{Z}$. Suppose $s(n)_{n\geq 0}$ is a sequence such that, for all $\ell \geq \ell_0$ and for all r such that $0 \leq r \leq 2^\ell - 1$, we have*

$$s(2^\ell + r) = \begin{cases} s(r) + c & \text{if } r \leq 2^{\ell-1} \\ s(2^{\ell+1} - r) & \text{if } r > 2^{\ell-1}. \end{cases} \tag{1}$$

Then $s(n)_{n\geq 0}$ is 2-regular.

The recurrence satisfied by $s(n)$ in Theorem 4 reads words from left to right, i.e., starting with the most significant digit. Our proof of this theorem will express sequences in the 2-kernel of $s(n)_{n\geq 0}$ as in Definition 2, starting with the least significant digit.

From Equation (1) one can get some information about the asymptotic behavior of the sequence $s(n)_{n\geq 0}$. We have $s(n) = O(\log n)$, and moreover

$$s\left(\tfrac{4^{\ell+1}-1}{3}\right) = s(4^\ell + \cdots + 4^1 + 4^0) = \left(\ell - \left\lfloor \tfrac{\ell_0-1}{2} \right\rfloor\right) c + s\left(\tfrac{4^{\lfloor(\ell_0+1)/2\rfloor}-1}{3}\right)$$

for $\ell \geq \lfloor \frac{\ell_0-1}{2} \rfloor$. At the same time, there are many subsequences of $s(n)_{n\geq 0}$ which are constant; for example, $s(2^\ell) = c$ for $\ell \geq \ell_0$.

Example 5. As an illustration of the reflection property described in Theorem 4, we consider the abelian complexity of the 2-block coding of the period-doubling word $\mathbf{p}$. (The recurrence satisfied by this sequence is given in Theorem 21.) Some values of this sequence are depicted in Figures 1 and 2.

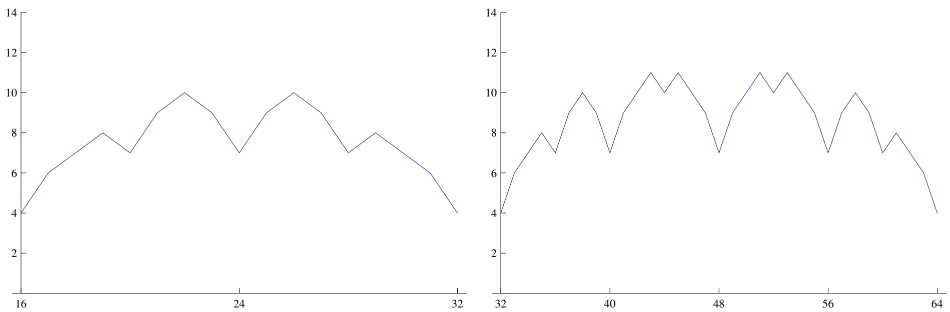

Figure 1: The abelian complexity of block$(\mathbf{p}, 2)$ on the intervals $[16, 32]$ and $[32, 64]$.

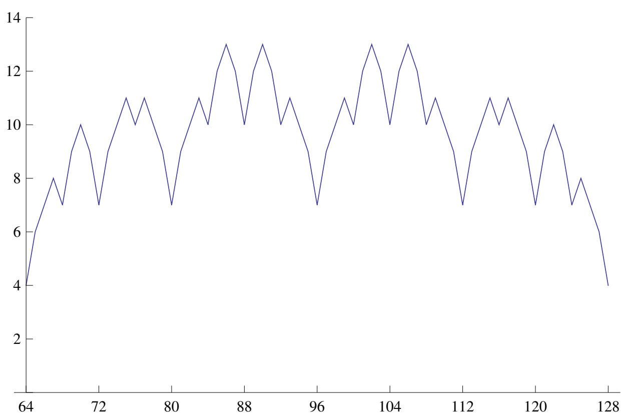

Figure 2: The abelian complexity of block$(\mathbf{p}, 2)$ on the interval $[64, 128]$.

In Section 3, we collect some general results and definitions about words and k-regular sequences (in particular stability properties of the set of k-regular sequences under sum and product) that are needed in the other parts of this paper.

In Section 4, we study the abelian complexity of the 2-block coding $\mathbf{x} = \text{block}(\mathbf{p}, 2)$ of the period-doubling word $\mathbf{p}$. In particular, we consider the difference $\Delta_0(n)$ between the maximal and minimal numbers of 0's occurring in factors of length n in block$(\mathbf{p}, 2)$. We prove that the sequences $\Delta_0(n)_{n\geq 0}$ and $\mathcal{P}_{\mathbf{x}}^{(1)}(n)_{n\geq 0}$ are 2-regular. In Section 5, we study the 2-abelian complexity of $\mathbf{p}$. We show that the 2-regularity of $\Delta_0(n)_{n\geq 0}$ and $\mathcal{P}_{\mathbf{x}}^{(1)}(n)_{n\geq 0}$ implies the 2-regularity of $\mathcal{P}_{\mathbf{p}}^{(2)}(n)$.

Sections 6 and 7 share some similarities with Sections 4 and 5. The reader will see that the strategy used to prove the 2-regularity of $\mathcal{P}_{\mathbf{p}}^{(2)}(n)$ can also be applied to the Thue–Morse word. Nevertheless, some differences do not permit us to treat the two cases within a completely unified framework.

In Section 6, we study the abelian complexity of the 2-block coding $\mathbf{y} = \mathrm{block}(\mathbf{t}, 2)$ of the Thue–Morse word $\mathbf{t}$. We define $\Delta_{12}(n)$ to be the difference between the maximal total and minimal total numbers of 1's and 2's occurring in factors of length n in $\mathrm{block}(\mathbf{t}, 2)$. It turns out that $\Delta_{12}(n) + 1 = \mathcal{P}^{(1)}_{\mathbf{p}}(n)$ and our results can thus be related to [5] and [14]. We prove that $\Delta_{12}(n)_{n\geq 0}$ and $\mathcal{P}^{(1)}_{\mathbf{y}}(n)_{n\geq 0}$ are 2-regular. In Section 7, we show that the 2-regularity of $\mathcal{P}^{(2)}_{\mathbf{t}}(n)$ follows from the 2-regularity of $\Delta_{12}(n)_{n\geq 0}$ and $\mathcal{P}^{(1)}_{\mathbf{y}}(n)_{n\geq 0}$.

Finally, in Section 8 we suggest a direction for future work.

Contents

2 Sequences satisfying a reflection symmetry

The aim of this section is to prove Theorem 4 stated in the introduction. Before proving it in generality, we first examine the sequence satisfying the recurrence for $\ell_0 = 0$ and $c = 1$. It will turn out that the general solution can be expressed naturally in terms of this sequence.

Let $A(0) = 0$. For each $\ell \geq 0$ and $0 \leq r \leq 2^\ell - 1$, let

$$A(2^\ell + r) = \begin{cases} A(r) + 1 & \text{if } r \leq 2^{\ell-1} \\ A(2^{\ell+1} - r) & \text{if } r > 2^{\ell-1}. \end{cases} \tag{2}$$

The sequence $A(n)_{n\geq 0}$ is

$$0,1,1,2,1,2,2,2,1,2,2,3,2,3,2,2,\ldots$$

and appears as [16, A007302]. Allouche and Shallit [2, Example 12] identified this sequence as an example of a regular sequence. We include a proof here.

Proposition 6. *For all $n \geq 0$ we have*

$$\begin{aligned}
A(2n) &= A(n)\\
A(8n+1) &= A(4n+1)\\
A(8n+3) &= A(2n+1)+1\\
A(8n+5) &= A(2n+1)+1\\
A(8n+7) &= A(4n+3).
\end{aligned}$$

In particular, $A(n)_{n\geq 0}$ is 2-regular.

Proof. This proof is typical of many of the proofs throughout the paper. We work by induction on n. The case $n=0$ can be checked easily using the first few values of the sequence $A(n)_{n\geq 0}$. Therefore, let $n \geq 1$ and assume that the recurrence holds for all values less than n. Write $n = 2^\ell + r$ with $\ell \geq 0$ and $0 \leq r \leq 2^\ell - 1$.

First let us address the equation $A(2n) = A(n)$. If $0 \leq r \leq 2^{\ell-1}$, then

$$\begin{aligned}
A(2n) &= A(2^{\ell+1}+2r) &&\\
&= A(2r)+1 && \text{(by Equation (2))}\\
&= A(r)+1 && \text{(by induction hypothesis)}\\
&= A(2^\ell + r) && \text{(by Equation (2))}\\
&= A(n).
\end{aligned}$$

On the other hand, if $2^{\ell-1} < r < 2^\ell$, then

$$\begin{aligned}
A(2n) &= A(2^{\ell+1}+2r) &&\\
&= A(2^{\ell+2}-2r) && \text{(by Equation (2))}\\
&= A(2^{\ell+1}-r) && \text{(by induction hypothesis)}\\
&= A(2^\ell + r) && \text{(by Equation (2))}\\
&= A(n).
\end{aligned}$$

Next we consider $A(8n+1) = A(4n+1)$. If $0 \leq r \leq 2^{\ell-1}-1$, then

$$\begin{aligned}
A(8n+1) &= A(2^{\ell+3}+8r+1) &&\\
&= A(8r+1)+1 && \text{(by Equation (2))}\\
&= A(4r+1)+1 && \text{(by induction hypothesis)}\\
&= A(2^{\ell+2}+4r+1) && \text{(by Equation (2))}\\
&= A(4n+1).
\end{aligned}$$

If $2^{\ell-1} \leq r < 2^{\ell}$, then

$$\begin{aligned}
A(8n+1) &= A(2^{\ell+3}+8r+1) \\
&= A(2^{\ell+4}-8r-1) && \text{(by Equation (2))} \\
&= A(2^{\ell+4}-8r-8+7) \\
&= A(2^{\ell+3}-4r-4+3) && \text{(by induction hypothesis)} \\
&= A(2^{\ell+3}-(4r+1)) \\
&= A(2^{\ell+2}+4r+1) && \text{(by Equation (2))} \\
&= A(4n+1).
\end{aligned}$$

The equations for $A(8n+3)$, $A(8n+5)$ and $A(8n+7)$ are handled similarly. □

Now we prove Theorem 4. We show that for general $\ell_0 \geq 0$, a sequence $s(n)_{n\geq 0}$ satisfying the recurrence can be written in terms of $A(n)_{n\geq 0}$.

Proof of Theorem 4. There are 2^{ℓ_0} initial conditions for the recurrence, namely $s(0), \ldots, s(2^{\ell_0}-1)$. We claim that most of the 2^{ℓ_0+2} subsequences of the form $s(2^{\ell_0+2}n+i)_{n\geq 0}$ depend on only one of the initial conditions $s(j)$; each of these subsequences is essentially $A(n)_{n\geq 0}$, $A(4n+1)_{n\geq 0}$, $A(2n+1)_{n\geq 0}$, or $A(4n+3)_{n\geq 0}$. Furthermore, each of the remaining subsequences is equal to $s(2^{\ell_0}n+j)+c$ for some j. More precisely, for $0 \leq i \leq 2^{\ell_0+2}-1$ and $n \geq 0$ we have the identity

$$s(2^{\ell_0+2}n+i) = \begin{cases} c\,A(n)+s(0) & \text{if } i=0 \\ c\,A(4n+1)-c+s(i) & \text{if } 1 \leq i \leq 2^{\ell_0}-1 \\ c\,A(4n+1)+s(0) & \text{if } i=2^{\ell_0} \\ s(2^{\ell_0}n+i-2^{\ell_0})+c & \text{if } 2^{\ell_0}+1 \leq i \leq 2^{\ell_0}+2^{\ell_0-1}-1 \\ c\,A(2n+1)+s(|i-2^{\ell_0+1}|) & \text{if } 2^{\ell_0}+2^{\ell_0-1} \leq i \leq 2^{\ell_0+1}+2^{\ell_0-1} \\ s(2^{\ell_0}n+i-2^{\ell_0+1})+c & \text{if } 2^{\ell_0+1}+2^{\ell_0-1}+1 \leq i \leq 2^{\ell_0+1}+2^{\ell_0}-1 \\ c\,A(4n+3)+s(0) & \text{if } i=2^{\ell_0+1}+2^{\ell_0} \\ c\,A(4n+3)-c+s(2^{\ell_0+2}-i) & \text{if } 2^{\ell_0+1}+2^{\ell_0}+1 \leq i \leq 2^{\ell_0+2}-1. \end{cases}$$

(Note the symmetry among the eight cases, which reflects the symmetry $s(2^{\ell}+r) = s(2^{\ell+1}-r)$ of the recurrence for $r > 2^{\ell-1}$.) It will follow from this identity that the $\mathbb{Z}$-module generated by the 2-kernel of $s(n)_{n\geq 0}$ is generated by the sequences $s(2^{\ell}n+j)_{n\geq 0}$ for $0 \leq \ell \leq \ell_0+1$ and $0 \leq j \leq 2^{\ell}-1$, $A(n)_{n\geq 0}$, $A(4n+1)_{n\geq 0}$, $A(2n+1)_{n\geq 0}$, $A(4n+3)_{n\geq 0}$, and the constant 1 sequence. In particular, this module is finitely generated.

We prove the identity by induction on n. Recall that for all $\ell \geq \ell_0$ and for all r such that $0 \leq r \leq 2^{\ell}-1$, we have Equation (1), i.e.,

$$s(2^{\ell}+r) = \begin{cases} s(r)+c & \text{if } r \leq 2^{\ell-1} \\ s(2^{\ell+1}-r) & \text{if } r > 2^{\ell-1}. \end{cases}$$

For $n = 0$, one uses $A(1) = 1$ and $A(3) = 2$ to verify that all eight cases of the identity hold. Inductively, let $n \geq 1$, and assume the identity is true for all $n' < n$. Write $n = 2^\ell + r$ with $\ell \geq 0$ and $0 \leq r \leq 2^\ell - 1$.

First we consider the case $0 \leq r \leq 2^{\ell-1} - 1$. For all $i \in \{0, \dots 2^{\ell_0+2} - 1\}$, we have $2^{\ell_0+2}r + i \leq 2^{(\ell_0+2+\ell)-1} - 1$, so

$$\begin{aligned} s(2^{\ell_0+2}n + i) &= s(2^{\ell_0+2+\ell} + (2^{\ell_0+2}r + i)) \\ &= s(2^{\ell_0+2}r + i) + c && \text{(by Equation (1))}. \end{aligned}$$

If $1 \leq i \leq 2^{\ell_0} - 1$, then the induction hypothesis now gives

$$\begin{aligned} s(2^{\ell_0+2}n + i) &= s(2^{\ell_0+2}r + i) + c \\ &= c\,A(4r + 1) + s(i) \\ &= c\left(A(2^{\ell+2} + 4r + 1) - 1\right) + s(i) \\ &= c\,A(4n + 1) - c + s(i), \end{aligned}$$

where we have used $A(2^{\ell+2} + 4r + 1) = A(4r + 1) + 1$ from the recurrence for $A(n)$, since $4r + 1 \leq 2^{(\ell+2)-1}$. The other seven intervals for i are verified similarly; in each case one applies the induction hypothesis to $s(2^{\ell_0+2}r + i) + c$ and then uses the recurrence for either $A(n)$ or $s(n)$ to raise an argument in r to an argument in n.

It remains to consider $2^{\ell-1} \leq r \leq 2^\ell - 1$. First we address the case $i = 0$. If $r = 2^{\ell-1}$ then

$$\begin{aligned} s(2^{\ell_0+2}n + i) &= s(2^{\ell_0+2+\ell} + 2^{\ell_0+2+\ell-1}) \\ &= s(2^{\ell_0+2+\ell-1}) + c && \text{(by Equation (1))} \\ &= c\,A(2^{\ell-1}) + s(0) + c && \text{(by inductive hypothesis)} \\ &= c\left(A(2^\ell + 2^{\ell-1}) - 1\right) + s(0) + c && \text{(by Equation (2))} \\ &= c\,A(n) + s(0) \end{aligned}$$

as desired. Alternatively, if $2^{\ell-1} < r \leq 2^\ell - 1$ then $2^{\ell_0+2}r > 2^{(\ell_0+2+\ell)-1}$, so

$$\begin{aligned} s(2^{\ell_0+2}n + i) &= s(2^{\ell_0+2+\ell} + 2^{\ell_0+2}r) \\ &= s(2^{\ell_0+2+\ell+1} - 2^{\ell_0+2}r) && \text{(by Equation (1))} \\ &= s(2^{\ell_0+2}(2^{\ell+1} - r) + 0) \\ &= c\,A(2^{\ell+1} - r) + s(0) && \text{(by inductive hypothesis)} \\ &= c\,A(2^\ell + r) + s(0) && \text{(by Equation (2))} \\ &= c\,A(n) + s(0). \end{aligned}$$

Therefore it remains to consider $2^{\ell-1} \leq r \leq 2^\ell - 1$ for $1 \leq i \leq 2^{\ell_0+2} - 1$. In this range we have $2^{\ell_0+2}r + i > 2^{(\ell_0+2+\ell)-1}$, so

$$\begin{aligned} s(2^{\ell_0+2}n + i) &= s(2^{\ell_0+2+\ell} + (2^{\ell_0+2}r + i)) \\ &= s(2^{\ell_0+2+\ell+1} - 2^{\ell_0+2}r - i) && \text{(by Equation (1))} \\ &= s(2^{\ell_0+2}n' + i'), \end{aligned}$$

where $n' = 2^{\ell+1} - r - 1$ and $i' = 2^{\ell_0+2} - i$. We prove the identity for the seven intervals for i using the same steps we have already used several times; we have just applied the recurrence for $s(n)$, so next we use the induction hypothesis, followed by the recurrence for $A(n)$ or $s(n)$, depending on which term appears. For the first interval, if $1 \le i \le 2^{\ell_0} - 1$, then $2^{\ell_0+1} + 2^{\ell_0} + 1 \le i' \le 2^{\ell_0+2} - 1$, so

$$\begin{aligned}
s(2^{\ell_0+2}n + i) &= s(2^{\ell_0+2}n' + i') \\
&= c\,A(4n' + 3) - c + s(2^{\ell_0+2} - i') && \text{(by inductive hypothesis)} \\
&= c\,A(2^{\ell+3} - (4r + 1)) - c + s(i) \\
&= c\,A(2^{\ell+2} + 4r + 1) - c + s(i) && \text{(by Equation (2))} \\
&= c\,A(4n + 1) - c + s(i).
\end{aligned}$$

The proofs for the remaining six intervals are routine at this point, so we omit the steps here. □

Example 7. In Section 4, we will use Theorem 4 with $\ell_0 = 2$ to conclude that $\Delta_0(n)_{n\ge0}$ and $\mathcal{P}^{(1)}_{\mathbf{x}}(n)_{n\ge0}$ are 2-regular for the period-doubling word. For $\ell_0 = 2$ the value of $s(16n + i)$ is

$$s(16n+i) = \begin{cases} c\,A(n) + s(0) & \text{if } i = 0 \\ c\,A(4n+1) - c + s(i) & \text{if } 1 \le i \le 3 \\ c\,A(4n+1) + s(0) & \text{if } i = 4 \\ s(4n+1) + c & \text{if } i = 5 \\ c\,A(2n+1) + s(|i-8|) & \text{if } 6 \le i \le 10 \\ s(4n+3) + c & \text{if } i = 11 \\ c\,A(4n+3) + s(0) & \text{if } i = 12 \\ c\,A(4n+3) - c + s(16-i) & \text{if } 13 \le i \le 15. \end{cases}$$

In Section 6, we will use Theorem 4 with $\ell_0 = 1$ to conclude that $\Delta_{12}(n)_{n\ge0}$ is 2-regular for the Thue–Morse word.

3 About regular sequences and words

We will often make use of the following composition theorem for a function F defined piecewise on several k-automatic sets.

Lemma 8. *Let $k \ge 2$. Let $P_1, \dots, P_\ell : \mathbb{N} \to \{0, 1\}$ be unary predicates that are k-automatic. Let $f_1, \dots, f_\ell$ be k-regular functions. The function $F : \mathbb{N} \to \mathbb{N}$ defined by*

$$F(n) = \sum_{i=1}^{\ell} f_i(n)\,P_i(n)$$

is k-regular.

Proof. It is a direct consequence of [1, Theorem 2.5]: if $s(n)_{n\geq 0}$ and $t(n)_{n\geq 0}$ are k-regular, then $(s(n)+t(n))_{n\geq 0}$ and $(s(n)t(n))_{n\geq 0}$ are both k-regular sequences. Recall that k-automatic sequences are special cases of k-regular sequences. □

Note that if, for each n, there is exactly one i such that $P_i(n) = 1$, then we can write

$$F(n) = \begin{cases} f_1(n) & \text{if } P_1(n) = 1 \\ f_2(n) & \text{if } P_2(n) = 1 \\ \vdots & \qquad \vdots \\ f_\ell(n) & \text{if } P_\ell(n) = 1. \end{cases}$$

This is the setting in which we will apply Lemma 8.

We will also make use of the following classical results.

Lemma 9. *[1, Theorem 2.3] Let $k \geq 2$ be an integer. A sequence taking finitely many values is k-regular if and only if it is k-automatic.*

Lemma 10. *[1, Corollary 2.4] Let $k, m \geq 2$ be integers. If a sequence $s(n)_{n\geq 0}$ is k-regular, then $(s(n) \bmod m)_{n\geq 0}$ is k-automatic.*

Lemma 11. *Let $k \geq 2$ be an integer. Let $s(n)_{n\geq 0}$ be a sequence. The sequence $s(n)_{n\geq 0}$ is k-regular if and only if $s(n+1)_{n\geq 0}$ is k-regular.*

Proof. It is a direct consequence of two results stated in [1], namely Theorem 2.6 and its following remark. □

Let us now give some definitions about combinatorics on words.

Definition 12. If a word w starts with the letter a, then $a^{-1}w$ denotes the word obtained from w by deleting its first letter. Similarly, if a word w ends with the letter a, then wa^{-1} denotes the word obtained from w by deleting its last letter. As usual, we let $|w|$ denote the length of the finite word w. If a is a letter, we let $|w|_a$ denote the number of occurrences of a in w. If $w = w_0 \cdots w_{\ell-1}$, then we let $w^{\mathrm{R}} = w_{\ell-1} \cdots w_0$ denote the reversal of w. Our convention is that we index letters in an infinite word beginning with 0.

Since we are interested in ℓ-abelian complexity, it is natural to consider the following operation that permits us to compare factors of length ℓ occurring in an infinite word. Indeed, if two finite words are ℓ-abelian equivalent, then their ℓ-block codings are abelian equivalent (but the converse does not hold).

Definition 13. Let $\ell \geq 1$. The *ℓ-block coding* of the word $\mathbf{w} = w_0w_1w_2\cdots$ over the alphabet A is the word

$$\operatorname{block}(\mathbf{w}, \ell) = (w_0 \cdots w_{\ell-1})\,(w_1 \cdots w_\ell)\,(w_2 \cdots w_{\ell+1}) \cdots (w_j \cdots w_{j+\ell-1}) \cdots$$

over the alphabet A^ℓ. If $A = \{0, \dots, r-1\}$, then it is convenient to identify A^ℓ with the set $\{0, \dots, r^\ell - 1\}$ and each word $w_0 \cdots w_{\ell-1}$ of length ℓ is thus replaced with the integer obtained by reading the word in base r, i.e.,

$$\sum_{i=0}^{\ell-1} w_i\, r^{\ell-1-i}.$$

One can also define accordingly the ℓ-block coding of a finite word u of length at least ℓ. The resulting word $\mathrm{block}(u, \ell)$ has length $|u| - \ell + 1$.

Example 14. The 2-block codings of 011010011 and 001101101 are respectively 13212013 and 01321321, which are abelian equivalent.

Lemma 15. *[13, Lemma 2.3] Let $\ell \geq 1$. Two finite words u and v of length at least $\ell - 1$ are ℓ-abelian equivalent if and only if they share the same prefix (resp. suffix) of length $\ell - 1$ and the words* $\mathrm{block}(u, \ell)$ *and* $\mathrm{block}(v, \ell)$ *are abelian equivalent.*

It is well known that the ℓ-block coding of a k-automatic sequence is again a k-automatic sequence [8]. (Note that the operation of ℓ-block compression that one also encounters in the literature is not the same as the ℓ-block coding given in Definition 13.)

Example 16. For the period-doubling word $\mathbf{p}$, the 2-block coding is given by

$$\mathrm{block}(\mathbf{p}, 2) = \phi^\omega(1) = 120012121200120012001212120012 12\cdots$$

where ϕ is the morphism over $\{0, 1, 2\}^*$ defined by $\phi : 0 \mapsto 12, 1 \mapsto 12, 2 \mapsto 00$.

Example 17. For the Thue–Morse word $\mathbf{t}$, the 2-block coding is given by

$$\mathrm{block}(\mathbf{t}, 2) = \nu^\omega(1) = 132120132012132120121320\cdots$$

where ν is the morphism over $\{0, 1, 2, 3\}^*$ defined by $\nu : 0 \mapsto 12, 1 \mapsto 13,\ 2 \mapsto 20, 3 \mapsto 21$.

4 Abelian complexity of $\mathrm{block}(\mathbf{p}, 2)$

We let $\mathbf{x}$ denote $\mathrm{block}(\mathbf{p}, 2) = 12001212120012001200121212001212\cdots$, the 2-block coding of $\mathbf{p}$, introduced in Example 16. We consider in this section the abelian complexity of $\mathbf{x}$ and then, in Section 5, we compare $\mathcal{P}^{(1)}_{\mathbf{x}}(n)$ with $\mathcal{P}^{(2)}_{\mathbf{p}}(n)$.

Definition 18. We will make use of functions related to the number of 0's in the factors of $\mathbf{x}$ of a given length. Let $n \in \mathbb{N}$. We let $\max_0(n)$ (resp. $\min_0(n)$) denote the maximum (resp. minimum) number of 0's in a factor of $\mathbf{x}$ of length n. Let $\Delta_0(n) = \max_0(n) - \min_0(n)$ be the difference between these two values.

Each of the $\Delta_0(n) + 1$ integers in the interval $[\min_0(n), \max_0(n)]$ is attained as the number of 0's in some factor of $\mathbf{x}$ of length n, since when we slide a window of length n along $\mathbf{x}$ from a factor with $\min_0(n)$ zeros to a factor with $\max_0(n)$ zeros, the number of 0's changes by at most 1 per step.

Lemma 19. *If n is even, then $\max_0(n)$, $\min_0(n)$ and $\Delta_0(n)$ are even.*

Proof. Suppose a factor $w = w_1 \cdots w_{2n}$ of $\mathbf{x}$ of even length $2n$ has an odd number n_0 of zeros. Since $\phi(0) = \phi(1) = 12$ and $\phi(2) = 00$, the factor w starts or ends with 0. Without loss of generality, assume it starts with $w_1 = 0$. Then its last letter must be $w_{2n} = 1$. The words $0w_1 \cdots w_{2n-1}$ and $w_2 \cdots w_{2n}2$ are two factors of length $2n$ with respectively $n_0 + 1$ and $n_0 - 1$ zeros. Hence, these two factors have even numbers of zeros which are respectively greater than and less than n_0. The conclusion follows. $\square$

We give two related proofs of the 2-regularity of the sequence $\mathcal{P}_{\mathbf{x}}^{(1)}(n)_{n\geq 0}$. The first uses the following proposition, which we prove in Section 4.1, together with the fact that $\Delta_0(n)_{n\geq 0}$ is 2-regular and the two sequences $(\Delta_0(n) \bmod 2)_{n\geq 0}$ and $(\min_0(n) \bmod 2)_{n\geq 0}$ are 2-automatic (see Section 4.2, Corollary 26). Then the 2-regularity of the sequence $\mathcal{P}_{\mathbf{x}}^{(1)}(n)_{n\geq 0}$ will follow from Lemma 8.

Proposition 20. *For $n \in \mathbb{N}$,*

$$\mathcal{P}_{\mathbf{x}}^{(1)}(n) = \begin{cases} \frac{3}{2}\Delta_0(n) + \frac{3}{2} & \text{if } \Delta_0(n) \text{ is odd} \\ \frac{3}{2}\Delta_0(n) + 1 & \text{if } \Delta_0(n) \text{ and } n - \min_0(n) \text{ are even} \\ \frac{3}{2}\Delta_0(n) + 2 & \text{if } \Delta_0(n) \text{ and } n - \min_0(n) + 1 \text{ are even.} \end{cases}$$

In the second proof, we prove in Section 4.3 the following theorem, which allows us to apply our general result expressed by Theorem 4.

Theorem 21. *Let $\ell \geq 2$ and r such that $0 \leq r < 2^\ell - 1$. We have*

$$\mathcal{P}_{\mathbf{x}}^{(1)}(2^\ell + r) = \begin{cases} \mathcal{P}_{\mathbf{x}}^{(1)}(r) + 3 & \text{if } r \leq 2^{\ell-1} \\ \mathcal{P}_{\mathbf{x}}^{(1)}(2^{\ell+1} - r) & \text{if } r > 2^{\ell-1}. \end{cases}$$

In particular, the sequence $\mathcal{P}_{\mathbf{x}}^{(1)}(n)_{n\geq 0}$ is 2-regular.

From Theorem 21 we see that $\mathcal{P}_{\mathbf{x}}^{(1)}(2^\ell) = \mathcal{P}_{\mathbf{x}}^{(1)}(0) + 3 = 4$ for all $\ell \geq 2$. Additionally, one can check that $\mathcal{P}_{\mathbf{x}}^{(1)}(2^1) = 4$.

4.1 Proof of Proposition 20

First we mention some properties of factors of the word $\mathbf{x}$.

Lemma 22. *The set of factors of $\mathbf{x}$ of length 2 is $\mathrm{Fac}_{\mathbf{x}}(2) = \{00, 01, 12, 20, 21\}$.*

Proof. It is easy to check that these five words are factors. To prove that they are the only ones, it is enough to check that for any element u in $\{00, 01, 12, 20, 21\}$ the three factors of length 2 of $\phi(u)$ are in $\{00, 01, 12, 20, 21\}$. $\square$

Lemma 23. *If w is a factor of $\mathbf{x}$ then $\big||w|_1 - |w|_2\big| \leq 1$. In particular, the letters 1 and 2 alternate in the sequence obtained from $\mathbf{x}$ after erasing the 0's.*

Proof. Let w be a factor of $\mathbf{x}$. There are two cases to consider.

If w can be de-substituted (that is, $w = \phi(v)$ for some v), then $|w|_1 = |w|_2$ since $|\phi(i)|_1 = |\phi(i)|_2$ for all $i \in \{0, 1, 2\}$.

If w cannot be de-substituted, then either w has even length and occurs at an odd index in $\mathbf{x}$, or w has odd length. If w has odd length, then deleting either the first or last letter results in a word that can be de-substituted, so $\big||w|_1 - |w|_2\big| \leq 1$. If w has even length and occurs at an odd index, then its first letter is 0 or 2 and its last letter is 0 or 1; deleting the first and last letters results in a word that can be de-substituted, so $\big||w|_1 - |w|_2\big| \leq 1$.

Finally, observe that if for all factors of a word u, the numbers of two letters x and y differ by at most 1, then x and y alternate in u. □

Lemma 24. *Let τ be the morphism defined by $\tau : 0 \mapsto 0, 1 \mapsto 2, 2 \mapsto 1$. If w is a factor of $\mathbf{x}$, then $\tau(w)^{\mathrm{R}}$ is also a factor of $\mathbf{x}$.*

Proof. We first prove by induction that

$$\tau(\phi(2u1))^{\mathrm{R}} = \phi(\tau(12u)^{\mathrm{R}})$$

for every factor of the form $2u1$ of $\mathbf{x}$.

One checks that this is true for 21 and 2001. If $2u1$ is a factor not equal to 21 nor 2001, then u must contain a 2 and we can write $2u1 = 2u'12u''1$ where $2u'1$ and $2u''1$ are factors of $\mathbf{x}$. By the induction hypothesis we have

$$\begin{aligned}
\tau(\phi(2u1))^{\mathrm{R}} &= \tau(\phi(2u'12u''1))^{\mathrm{R}} \\
&= \tau(\phi(2u''1))^{\mathrm{R}}\tau(\phi(2u'1))^{\mathrm{R}} \\
&= \phi(\tau(12u'')^{\mathrm{R}})\phi(\tau(12u')^{\mathrm{R}}) \\
&= \phi(\tau(12u'12u'')^{\mathrm{R}}) \\
&= \phi(\tau(12u)^{\mathrm{R}}).
\end{aligned}$$

We now prove the lemma by induction on the length of w. One can check by hand that the lemma is true for w of length at most 15. Assume the lemma is true for every factor of length at most $n \geq 15$, and let w be a factor of length $n+1$. Then w is a factor of $\phi(v)$ for some factor v of $\mathbf{x}$ with $\frac{n+1}{2} \leq |v| \leq \frac{n+3}{2}$.

Since all factors of length 4 contain a 1 and a 2, there exists a factor u such that v is a factor of $2u1$ and $|2u1| \leq \frac{n+3}{2} + 6$. In particular, w is a factor of $\phi(2u1)$ and $\tau(w)^{\mathrm{R}}$ is a factor of $\tau(\phi(2u1))^{\mathrm{R}}$. To obtain the conclusion, we just need to show that $\tau(\phi(2u1))^{\mathrm{R}}$ is a factor of $\mathbf{x}$.

As by Lemma 22, a 2 is always preceded by a 1 in $\mathbf{x}$, the word $12u$ is a factor of $\mathbf{x}$ and it has length $|12u| \leq \frac{n+3}{2} + 6 \leq n$. By induction hypothesis, $\tau(12u)^{\mathrm{R}}$ is a factor of $\mathbf{x}$. Hence $\phi(\tau(12u)^{\mathrm{R}})$ is also a factor. Finally, using the previous result, $\tau(\phi(2u1))^{\mathrm{R}} = \phi(\tau(12u)^{\mathrm{R}})$ is a factor of $\mathbf{x}$. □

We can now express $\mathcal{P}_{\mathbf{x}}^{(1)}$ in terms of Δ_0.

Proof of Proposition 20. Let w be a factor of $\mathbf{x}$ of length $|w| = n$.

If $|w|-|w|_0 = |w|_1+|w|_2$ is even, it follows from Lemma 23 that $|w|_1 = |w|_2$. Therefore every factor of length n containing exactly $|w|_0$ zeros is abelian-equivalent to w, so the pair $(n, |w|_0)$ determines a unique abelian equivalence class of factors.

If $|w| - |w|_0$ is odd, then by Lemma 23 either $|w|_1 = |w|_2 + 1$ or $|w|_2 = |w|_1 + 1$. By Lemma 24, there is another factor, $v = \tau(w)^{\mathrm{R}}$, of length n, with $|v|_0 = |w|_0$ and $|v|_1 - |v|_2 = |w|_2 - |w|_1$. Therefore both possibilities occur, so the number of abelian equivalence classes corresponding to a pair $(n, |w|_0)$ is 2.

There are $\Delta_0(n)+1$ possible values for the number of 0's in a factor of length n. Since each value occurs for some factor, we have

$$\begin{aligned}\mathcal{P}^{(1)}_{\mathbf{x}}(n) &= \sum_{i=\min_0(n)}^{\max_0(n)} \begin{cases} 1 & \text{if } n-i \text{ is even} \\ 2 & \text{if } n-i \text{ is odd} \end{cases} \\ &= \sum_{j=n-\max_0(n)}^{n-\min_0(n)} \begin{cases} 1 & \text{if } j \text{ is even} \\ 2 & \text{if } j \text{ is odd.} \end{cases}\end{aligned}$$

Therefore $\mathcal{P}^{(1)}_{\mathbf{x}}(n) = \frac{3}{2}\Delta_0(n) + c(n)$, where $c(n)$ depends only on the parities of $\Delta_0(n)$ and $n - \min_0(n)$; computing four explicit values allows one to determine the values of $c(n)$ and obtain the equation claimed for $\mathcal{P}^{(1)}_{\mathbf{x}}(n)$. □

4.2 $\Delta_0(n)_{n\geq 0}$ is 2-regular, $(\min_0(n) \bmod 2)_{n\geq 0}$ is 2-automatic

In this section, we prove the following result.

Proposition 25. *Let $\ell \geq 2$ and r such that $0 \leq r < 2^\ell$. We have*

$$\Delta_0(2^\ell + r) = \begin{cases} \Delta_0(r) + 2 & \text{if } r \leq 2^{\ell-1} \\ \Delta_0(2^{\ell+1} - r) & \text{if } r > 2^{\ell-1}. \end{cases}$$

Moreover,

$$\min_0(2^\ell + r) \equiv \begin{cases} \min_0(r) \pmod 2 & \text{if } r \leq 2^{\ell-1} \\ \min_0(2^{\ell+1} - r) + \Delta_0(2^{\ell+1} - r) \pmod 2 & \text{if } r > 2^{\ell-1}. \end{cases}$$

Before giving the proof, we prove a corollary. The 2-regularity of $\mathcal{P}^{(1)}_{\mathbf{x}}(n)_{n\geq 0}$ follows from Proposition 20 and Corollary 26.

Corollary 26. *The following statements are true.*

- *The sequence $\Delta_0(n)_{n\geq 0}$ is 2-regular.*
- *The sequence $(\Delta_0(n) \bmod 2)_{n\geq 0}$ is 2-automatic.*
- *The sequence $(\min_0(n) \bmod 2)_{n\geq 0}$ is 2-automatic.*

Proof. The first assertion is a direct consequence of Proposition 25 and Theorem 4. Note that one can obtain explicit relations satisfied by $\Delta_0(n)_{n\geq 0}$ from Example 7. The second assertion follows from Lemma 10.

For the last assertion, for $i \in \{0,\dots,31\}$ we prove that, modulo 2,

$$\min\nolimits_0(32n+i) \equiv \begin{cases} \min\nolimits_0(8n+1) & \text{if } i \in \{1,5,9,17,25\} \\ \min\nolimits_0(8n+3) & \text{if } i = 11 \\ \min\nolimits_0(8n+5) & \text{if } i = 21 \\ \min\nolimits_0(8n+7) & \text{if } i \in \{7,15,23,27,31\} \\ 0 & \text{otherwise} \end{cases}$$

and

$$\Delta_0(32n+i) \equiv \begin{cases} \Delta_0(8n+1) & \text{if } i \in \{1,5,9,17,25\} \\ \Delta_0(8n+3) & \text{if } i = 11 \\ \Delta_0(8n+5) & \text{if } i = 21 \\ \Delta_0(8n+7) & \text{if } i \in \{7,15,23,27,31\} \\ 0 & \text{otherwise.} \end{cases}$$

By Lemma 19, we already know that $\min_0(2n) \equiv \Delta_0(2n) \equiv 0 \pmod 2$ for any $n \in \mathbb{N}$. Hence the relations above are true for i even. We prove the other relations by induction on n. They are true for $n = 0$. Let $n > 0$ and assume the relations are satisfied for all n' such that $0 \leq n' < n$. We can write $n = 2^\ell + r$ with $\ell \geq 0$ and $0 \leq r < 2^\ell$. Let $i \in \{1,\dots,31\}$ be odd.

Assume first that $r < 2^{\ell-1}$. We have $32n + i = 2^{\ell+5} + 32r + i$ and $32r + i < 2^{\ell+4}$.

$$\begin{aligned} \min\nolimits_0(32n+i) &\equiv \min\nolimits_0(32r+i) && \text{(Proposition 25)} \\ &\equiv \min\nolimits_0(8r+j) && \text{(induction)} \\ &\equiv \min\nolimits_0(2^{\ell+3}+8r+j) && \text{(Proposition 25)} \\ &\equiv \min\nolimits_0(8n+j) \pmod 2 \end{aligned}$$

for some $j \in \{0,\dots,7\}$ according to the relations. A similar reasoning holds for the Δ_0 relations.

Assume now that $r \geq 2^{\ell-1}$. Since $32r + i > 2^{\ell+4}$, we have

$$\begin{aligned} \min\nolimits_0(32n+i) &\equiv \min\nolimits_0(2^{\ell+6}-32r-i) + \Delta_0(2^{\ell+6}-32r-i) && \text{(Proposition 25)} \\ &\equiv \min\nolimits_0(32n'+j) + \Delta_0(32n'+j) \pmod 2 \end{aligned}$$

with $j = 32 - i$ and $n' = 2^{\ell+1} - r - 1$. If $i \in \{3,13,19,29\}$, then $j \in \{3,13,19,29\}$. By the induction hypothesis, $\min_0(32n'+j) \equiv \Delta_0(32n'+j) \equiv 0 \pmod 2$ and we are done.

For the remaining cases, $i, j \notin \{3,13,19,29\}$. As $\min_0$ and Δ_0 satisfy the same recurrence relations, by the induction hypothesis, there exists $k \in \{1,3,5,7\}$ such that

$$\begin{aligned}
\min_0(32n+i) &\equiv \min_0(8n'+k) + \Delta_0(8n'+k) \\
&\equiv \min_0(2^{\ell+4} - (8r+8-k)) + \Delta_0(2^{\ell+4} - (8r+8-k)) \\
&\equiv \min_0(2^{\ell+3} + (8r+8-k)) \qquad \text{(Proposition 25)} \\
&\equiv \min_0(8n + (8-k)) \pmod 2.
\end{aligned}$$

Observe that the value of $8-k$ is the value given in the relation for i. This concludes the proof of the $\min_0$ relations. A similar argument works for the Δ_0 relations. □

We break the proof of Proposition 25 into three parts, covered by Lemmas 27, 29 and 31. We first deal with powers of 2.

Lemma 27. *Let $\ell \in \mathbb{N}$, $\ell \geq 1$. We have $\mathcal{P}^{(1)}_{\mathbf{x}}(2^\ell) = 4$,*

$$\Delta_0(2^\ell) = 2,\ \max_0(2^{\ell+1}) = 2^\ell - \min_0(2^\ell) \text{ and } \min_0(2^{\ell+1}) = 2^\ell - \max_0(2^\ell).$$

Proof. Recall that $\Psi(w) = (|w|_0, |w|_1, |w|_2)$ is the Parikh vector of w. We show by induction that

$$\begin{aligned}
&\{\Psi(w) : w \text{ factor of } \mathbf{x} \text{ with } |w| = 2^\ell\} \\
&\qquad = \{P_\ell + (0,0,0), P_\ell + (-2,1,1), P_\ell + (-1,1,0), P_\ell + (-1,0,1)\}
\end{aligned}$$

and that

$$\Psi(\phi^\ell(0)) = \begin{cases} P_\ell & \text{if } \ell \text{ is even} \\ P_\ell + (-2,1,1) & \text{if } \ell \text{ is odd} \end{cases}$$

$$\Psi(\phi^\ell(2)) = \begin{cases} P_\ell + (-2,1,1) & \text{if } \ell \text{ is even} \\ P_\ell & \text{if } \ell \text{ is odd,} \end{cases}$$

where $P_\ell = (\frac{2^\ell+4}{3}, \frac{2^\ell-2}{3}, \frac{2^\ell-2}{3})$ if ℓ is odd and $P_\ell = (\frac{2^\ell+2}{3}, \frac{2^\ell-1}{3}, \frac{2^\ell-1}{3})$ if ℓ is even. Since Parikh vectors of factors of length 2^ℓ can take exactly four values, the conclusion is immediate.

The result is true for $\ell \in \{1,2\}$. Let $\ell > 2$ and assume the result holds for $\ell - 1$. Let w be a factor of length 2^ℓ.

If w can be de-substituted, then $w = \phi(v)$ for some factor v of length $2^{\ell-1}$, and $\Psi(w) = (2|v|_2, |v|_0 + |v|_1, |v|_0 + |v|_1)$. Using the induction hypothesis, it is easy to check that $\Psi(w) = P_\ell$ or $\Psi(w) = P_\ell + (-2,1,1)$ and that the equalities for $\Psi(\phi^\ell(0)), \Psi(\phi^\ell(2))$ are satisfied.

If w cannot be de-substituted, then w occurs at an odd index in $\mathbf{x}$ and w is of the form

$$0^{-1}\phi(v)0, \quad 1^{-1}\phi(v)1, \quad 0^{-1}\phi(v)1 \quad \text{or} \quad 1^{-1}\phi(v)0$$

for some factor v of length $2^{\ell-1}$. If w is of one of the first two forms, then $\Psi(w) = \Psi(\phi(v))$ and $\Psi(w) = P_\ell$ or $\Psi(w) = P_\ell + (-2,1,1)$ (as in the previous case).

If $w = 0^{-1}\phi(v)1$, then w can also be written as $w = 0\phi(u)2^{-1}$ for some factor u of length $2^{\ell-1}$. So both Parikh vectors $\Psi(\phi(v))$ and $\Psi(\phi(u))$ belong to $\{P_\ell, P_\ell + (-2,1,1)\}$. Since by construction $\phi(v)$ has two more zeros than $\phi(u)$, we obtain $\Psi(\phi(v)) = P_\ell$ and $\Psi(\phi(u)) = P_\ell + (-2,1,1)$. Thus $\Psi(w) = \Psi(\phi(v)) + (-1,1,0) = P_\ell + (-1,1,0)$.

Similarly, if $w = 1^{-1}\phi(v)0$, then $\Psi(w) = P_\ell + (-1,0,1)$.

To conclude the proof, we just need to show that these four cases actually occur for all ℓ. Since $\{\Psi(\phi^\ell(0)), \Psi(\phi^\ell(2))\} = \{P_\ell, P_\ell + (-2,1,1)\}$, consider all factors of length 2^ℓ occurring between two consecutive occurrences of $\Psi(\phi^\ell(0))$ and $\Psi(\phi^\ell(2))$. By continuity[2], of the number of 0's, one of these factors must have a Parikh vector equal to $P_\ell + (-1,1,0)$ or $P_\ell + (-1,0,1)$. Using Lemma 24, we obtain that w is a factor of length 2^ℓ with $\Psi(w) = P_\ell + (-1,1,0)$ if and only if $\tau(w)^{\mathrm{R}}$ is a factor of length 2^ℓ with $\Psi(w) = P_\ell + (-1,0,1)$. So all four values actually occur. □

To show Lemmas 29 and 31, we first prove the following technical result.

Lemma 28. *Let u be a factor of $\mathbf{x}$ of length $n \geq 1$. Let $\max_2(n)$ (resp. $\min_2(n)$) denote the maximum (resp. minimum) of $\{|w|_2 : w$ factor of $\mathbf{x}$ of length $n\}$. We have $|u|_2 = \max_2(n)$ if and only if $|\phi(u)|_0 = \max_0(2n)$, and $|u|_2 = \min_2(n)$ if and only if $|\phi(u)|_0 = \min_0(2n)$.*

Proof. For the first assertion, assume that $|u|_2 = \max_2(n)$ and suppose that $|\phi(u)|_0 < \max_0(2n)$. Note that $|\phi(u)|_0 = 2|u|_2$ by definition of ϕ. Let v be a factor of length $2n$ such that $|v|_0 = \max_0(2n)$, which is even by Lemma 19. In addition, we can assume that v starts with 00. Indeed, if it is not the case, then either v starts with 01 and ends with 0, or v is of the form $t00s$ where t does not contain any zero. In the first case, we can consider the word $0v0^{-1}$ that starts with 00 and has $\max_0(2n)$ zeros. In the second case, we can consider the word $00sw$ for some w with $|w| = |t|$. This factor has also $\max_0(2n)$ zeros. Therefore v can be de-substituted. So $v = \phi(z)$ and $|z|_2 = \frac{1}{2}|v|_0 > |u|_2$, which is a contradiction.

For the other direction, assume $|\phi(u)|_0 = \max_0(2n)$ and suppose $|u|_2$ does not maximize the number of 2's. Then there exists a factor v of length n such that $|v|_2 = \max_2(n)$. Hence,

$$|\phi(v)|_0 = 2|v|_2 > 2|u|_2 = |\phi(u)|_0 = \max_0(2n),$$

which is a contradiction. Similar arguments hold for the second assertion. □

Lemma 29. *If $\ell \geq 2$ and $0 \leq r \leq 2^{\ell-1}$, then*

$$\max_0(2^\ell + r) = \max_0(2^\ell) + \max_0(r),$$
$$\min_0(2^\ell + r) = \min_0(2^\ell) + \min_0(r).$$

Proof. We work by induction on ℓ. One checks the case $\ell = 2$. Let $\ell > 2$ and assume the statements are true for $\ell - 1$. Let r such that $0 \leq r \leq 2^{\ell-1}$.

Assume first that r is even. We shall exhibit a factor of length $2^\ell + r$ that has $\max_0(2^\ell) + \max_0(r)$ zeros and maximizes the number of 0's. By the induction hypothesis,

[2] We mean by *continuity* that the number of 0's is varying by at most 1 between two factors of the same length starting at consecutive indexes.

the result is true for $2^{\ell-1}+r/2$. So there exists a factor u of length $2^{\ell-1}+r/2$ with $\min_0(2^{\ell-1}+r/2) = \min_0(2^{\ell-1})+\min_0(r/2)$ zeros. In addition, we can assume that u maximizes the number of 2's. Indeed, since $|u|_0 = \min_0(2^{\ell-1}+r/2)$, $|u|_1+|u|_2$ is maximal among all factors of length $2^{\ell-1}+r/2$. If the number of 1 and 2 in u is even, then $|u|_2 = |u|_1$ is maximal. Otherwise, either $|u|_2 = |u|_1+1$ and $|u|_2$ is maximal, or $|u|_2 = |u|_1-1$ and u does not maximize the number of 2's. In the last case, by Lemma 24, we can consider the factor $\tau(u)^{\mathrm{R}}$ which satisfies $|\tau(u)^{\mathrm{R}}|_0 = |u|_0$ and $|\tau(u)^{\mathrm{R}}|_2 = |u|_1$. Hence, $\tau(u)^{\mathrm{R}}$ minimizes the number of 0's and maximizes the number of 2's.

Let us write $u = vw$ with $|v| = 2^{\ell-1}$ and $|w| = r/2$. Then, as $|v|_0+|w|_0 = |u|_0 = \min_0(2^{\ell-1})+\min_0(r/2)$, the words v and w minimize the number of 0's for words of their respective lengths. The word v maximizes also the number of 2's for factors of length $2^{\ell-1}$ because $|v|$ and $|v|_0 = \min_0(2^{\ell-1})$ are even by Lemma 27 and so is $|v|_1+|v|_2$. Since u maximizes the number of 2's and $|v|_2 = |v|_1$, the word w also maximizes the number of 2's. Hence, by Lemma 28, $\phi(u)$, $\phi(v)$ and $\phi(w)$ maximize the number of 0's for words of their respective lengths. Thus,

$$\max_0(2^\ell+r) = |\phi(u)|_0 = |\phi(v)|_0+|\phi(w)|_0 = \max_0(2^\ell)+\max_0(r).$$

If r is odd, we still have $0 \le r-1 \le r+1 \le 2^{\ell-1}$ and we can use the previous results:

$$\begin{aligned}\max_0(2^\ell+r-1) &= \max_0(2^\ell)+\max_0(r-1),\\ \max_0(2^\ell+r+1) &= \max_0(2^\ell)+\max_0(r+1).\end{aligned}$$

Note that $\max_0$ is even for even values and can only grow by 0 or 1. So there are two cases to consider: either $\max_0(2^\ell+r+1) = \max_0(2^\ell+r-1)$ or $\max_0(2^\ell+r+1) = \max_0(2^\ell+r-1)+2$.

If the two maxima are equal, then $\max_0(r+1) = \max_0(r-1)$, $\max_0(2^\ell+r) = \max_0(2^\ell+r-1)$ and $\max_0(r) = \max_0(r-1)$, and we are done. Otherwise, the two maxima differ by 2, and then $\max_0(r+1) = \max_0(r-1)+2$, $\max_0(2^\ell+r) = \max_0(2^\ell+r-1)+1$ and $\max_0(r) = \max_0(r-1)+1$, and we are done.

A similar proof shows that $\min_0(2^\ell+r) = \min_0(2^\ell)+\min_0(r)$. □

Lemma 31 will follow directly from the following lemma.

Lemma 30. *If $\ell \ge 2$ and $2^{\ell-1} \le r \le 2^\ell$, then*

$$\begin{aligned}\max_0(2^{\ell+1}) &= \max_0(2^\ell+r)+\min_0(2^\ell-r),\\ \min_0(2^{\ell+1}) &= \min_0(2^\ell+r)+\max_0(2^\ell-r).\end{aligned}$$

Moreover, there is a factor of length $2^{\ell+1}$ maximizing (resp. minimizing) the number of 0's such that the prefix of length $2^\ell+r$ also maximizes (resp. minimizes) the number of 0's. In addition, the first equality $\max_0(2^{\ell+1}) = \max_0(2^\ell+r)+\min_0(2^\ell-r)$ holds even if $\ell = 1$.

Proof. We proceed by induction on ℓ. One checks that the results are true for $\ell = 2$ and, for the first equality, for $\ell = 1$. Let $\ell > 2$ and assume both equalities hold for $\ell - 1$. Let r such that $2^{\ell-1} \leq r \leq 2^\ell$.

Assume first that r is even. By the induction hypothesis, there exists a factor $u = vw$ of length 2^ℓ such that

$$|u|_0 = \min_0(2^\ell) = \min_0(2^{\ell-1} + r/2) + \max_0(2^{\ell-1} - r/2),$$

$|v| = 2^{\ell-1} + r/2$ and v minimizes the number of 0's. Hence, $|v|_0 = \min_0(2^{\ell-1} + r/2)$ and $|w|_0 = \max_0(2^{\ell-1} - r/2)$.

Observe that u maximizes the number of 2's as $|u|$ and $|u|_0 = \min_0(2^\ell)$ are even. In addition, we can assume that v also maximizes the number of 2's. Indeed, if v is of even length, $|v|_0 = \min_0(2^{\ell-1} + r/2)$ implies $|v|_2$ is maximal. If v is of odd length and v does not maximize the number of 2's, then it ends with 1. Thus, v is followed by a 2. In particular, v occurs at an even index in $\mathbf{x}$. So is u and $u12$ or $u00$ is a factor of $\mathbf{x}$. If $u12$ is a factor, then consider, instead of u, $u' = z^{-1}u1$ where z denotes the first letter of u. In that case, the prefix of length $2^{\ell-1}+r/2$ of u' is $z^{-1}v2$. It still minimizes the number of 0's and now maximizes the number of 2's. Assume now that $u00$ is a factor. Observe that $\mathbf{x}$ is the fixed point of ϕ. So it is also the fixed point of ϕ^2. Therefore, $\mathbf{x}$ is a concatenation of blocks of length 4 of the form $\phi^2(0) = \phi^2(1) = 1200$ and $\phi^2(2) = 1212$. Since $u00$ is a factor of $\mathbf{x}$, the only extension of this factor is $12u00$ as $|u| = 2^\ell \equiv 0 \pmod 4$. Consider then $u' = 2u2^{-1}$.

Since $|u|_1 = |u|_2$ and $|v|_2 \geq |v|_1$, $|w|_1 \geq |w|_2$. Thus, as $|w|_0 = \max_0(2^{\ell-1} - r/2)$, w minimizes the number of 2's. By Lemma 28, we obtain $|\phi(u)|_0 = \max_0(2^{\ell+1})$, $|\phi(v)|_0 = \max_0(2^\ell + r)$, $|\phi(w)|_0 = \min_0(2^\ell - r)$. So

$$\begin{aligned}\max_0(2^{\ell+1}) = |\phi(u)|_0 &= |\phi(v)|_0 + |\phi(w)|_0 \\ &= \max_0(2^\ell + r) + \min_0(2^\ell - r).\end{aligned}$$

We can show similarly that $\min_0(2^{\ell+1}) = \min_0(2^\ell + r) + \max_0(2^\ell - r)$. Note that in this case, we can assume that the factor u with $|u|_0 = \max_0(2^\ell)$, given by the induction hypothesis, starts with 00 as in the proof of Lemma 28.

Assume now that r is odd. Then $2^{\ell-1} \leq r - 1 < r + 1 \leq 2^\ell$ and we can apply the previous result:

$$\begin{aligned}\max_0(2^{\ell+1}) &= \max_0(2^\ell + r - 1) + \min_0(2^\ell - r + 1) \\ &= \max_0(2^\ell + r + 1) + \min_0(2^\ell - r - 1).\end{aligned}$$

Since $\max_0$ is even for even values and can only grow by 0 or 1, there are two cases to consider: either $\max_0(2^\ell+r-1) = \max_0(2^\ell+r+1)$ or $\max_0(2^\ell+r-1)+2 = \max_0(2^\ell+r+1)$.

If the two maxima are equal, then $\min_0(2^\ell - r + 1) = \min_0(2^\ell - r - 1) = \min_0(2^\ell - r)$ and $\max_0(2^\ell + r) = \max_0(2^\ell + r - 1)$, and we are done. Otherwise, the two maxima differ by 2, and then $\min_0(2^\ell-r+1)-2 = \min_0(2^\ell-r-1)$. So $\max_0(2^\ell+r) = \max_0(2^\ell+r-1)+1$ and $\min_0(2^\ell - r) = \min_0(2^\ell - r + 1) - 1$, and we are done. Using similar argument, we can conclude that $\min_0(2^{\ell+1}) = \min_0(2^\ell + r) + \max_0(2^\ell - r)$.

For the construction of the factors, one can construct them using the factors $\phi(u)$ and $\phi(u')$ given for $r-1$ and $r+1$ in the previous construction. We consider the same two cases as before.

If the maxima are equal, then $\max_0(2^\ell+r) = \max_0(2^\ell+r-1)$. By construction, $\phi(u)$ has a prefix $\phi(v)$ of length $2^\ell+r-1$, maximizing the number of 0's. The letter z following the prefix $\phi(v)$ in $\phi(u)$ is not a 0. Otherwise, $\phi(v)0$ would be a factor of length $2^\ell+r$ with $\max_0(2^\ell+r)+1$ zeros, which is a contradiction. Hence, $\phi(v)z$ is a prefix of length $2^\ell+r$ of $\phi(u)$ that maximizes the number of 0's.

If $\max_0(2^\ell+r-1)+2 = \max_0(2^\ell+r+1)$, then $\max_0(2^\ell+r) = \max_0(2^\ell+r+1)-1$. By construction, $\phi(u')$ has a prefix $\phi(v')$ of length $2^\ell+r+1$, maximizing the number of 0's. This prefix must end with 0. Otherwise, deleting the last letter of $\phi(v')$ would give a factor of length $2^\ell+r$ with $\max_0(2^\ell+r+1) = \max_0(2^\ell+r)+1$ zeros, which is a contradiction. Hence, $\phi(v')0^{-1}$ is a prefix of length $2^\ell+r$ of $\phi(u')$ that maximizes the number of 0's.

A similar construction yields a factor of length $2^{\ell+1}$ minimizing the number of 0's such that the prefix of length $2^\ell+r$ also minimizes the number of 0's. $\square$

The previous lemma permits us to reformulate some relations between the two sequences $\max_0(n)_{n\geq 0}$ and $\min_0(n)_{n\geq 0}$.

Lemma 31. *If $\ell \geq 2$ and $2^{\ell-1} \leq r \leq 2^\ell$, then*

$$\begin{aligned}\max\nolimits_0(2^\ell+r) &= 2^\ell - \min\nolimits_0(2^{\ell+1}-r),\\ \min\nolimits_0(2^\ell+r) &= 2^\ell - \max\nolimits_0(2^{\ell+1}-r).\end{aligned}$$

The first equality holds even if $\ell = 1$.

Proof. One can check the first equality for $\ell = 1$. Let $\ell \geq 2$ and r such that $2^{\ell-1} \leq r \leq 2^\ell$. From the previous lemma, we have

$$\max\nolimits_0(2^\ell+r) = \max\nolimits_0(2^{\ell+1}) - \min\nolimits_0(2^\ell-r).$$

Note that, by Lemma 27, we have $\max_0(2^{\ell+1}) = 2^\ell - \min_0(2^\ell)$. Moreover, by Lemma 29, since $0 \leq 2^\ell - r \leq 2^\ell$, we get

$$\min\nolimits_0(2^\ell) + \min\nolimits_0(2^\ell-r) = \min\nolimits_0(2^\ell+2^\ell-r).$$

Since similar relations hold when exchanging $\min_0$ and $\max_0$, the conclusion follows. $\square$

The proof of Proposition 25 about the reflection relation satisfied by $\Delta_0(n)$ and the recurrence relation of $\min_0(n)$ is now immediate.

Proof of Proposition 25. Let $\ell \geq 2$. For r such that $0 \leq r \leq 2^{\ell-1}$, subtracting the two relations provided by Lemma 29 gives $\Delta_0(2^\ell+r) = \Delta_0(2^\ell) + \Delta_0(r)$ and we can conclude using the first relation given in Lemma 27, $\Delta_0(2^\ell) = 2$. Furthermore, $\min_0(2^\ell+r) \equiv \min_0(2^\ell) + \min_0(r) \pmod 2$ by Lemma 29. The expression for $\min_0(2^\ell+r)$ follows since $\min_0(2^\ell) \equiv 0 \pmod 2$ by Lemma 27.

For $2^{\ell-1} < r < 2^\ell$, subtracting the two relations provided by Lemma 31 permits us to conclude the proof of the expression claimed for $\Delta_0(2^\ell + r)$. Moreover, using Lemma 31, we get

$$\begin{aligned}\min\nolimits_0(2^\ell + r) &\equiv \max\nolimits_0(2^{\ell+1} - r) \pmod 2\\ &\equiv \min\nolimits_0(2^{\ell+1} - r) + \Delta_0(2^{\ell+1} - r) \pmod 2.\square\end{aligned}$$

4.3 Another proof of the 2-regularity of $\mathcal{P}^{(1)}_{\mathbf{x}}(n)_{n\geq 0}$

In this section we prove the 2-regularity of the abelian complexity $\mathcal{P}^{(1)}_{\mathbf{x}}(n)_{n\geq 0}$ in a second way, by proving Theorem 21. The proof makes use of Propositions 20 and 25.

Proof of Theorem 21. If $2^{\ell-1} \leq r \leq 2^\ell$, since all the conditions in Proposition 20 are equivalent whether considering $2^\ell + r$ or $2^{\ell+1} - r$, we have

$$\mathcal{P}^{(1)}_{\mathbf{x}}(2^\ell + r) = \mathcal{P}^{(1)}_{\mathbf{x}}(2^{\ell+1} - r).$$

Assume now that $0 \leq r \leq 2^{\ell-1}$. If $\Delta_0(2^\ell+r)$ is odd, $\Delta_0(r)$ is also odd by Proposition 25. By Proposition 20, we have $\mathcal{P}^{(1)}_{\mathbf{x}}(2^\ell+r) = \frac{3}{2}(\Delta_0(2^\ell+r)+1)$ and $\mathcal{P}^{(1)}_{\mathbf{x}}(r) = \frac{3}{2}(\Delta_0(r)+1)$. By Proposition 25, we have $\Delta_0(2^\ell + r) = \Delta_0(r) + 2$. Putting these three equalities together, we get $\mathcal{P}^{(1)}_{\mathbf{x}}(2^\ell + r) = \mathcal{P}^{(1)}_{\mathbf{x}}(r) + 3$.

The other cases can be done similarly. If $\Delta_0(2^\ell+r)$ and $2^\ell+r-\min_0(2^\ell+r)$ are even, then $\Delta_0(r)$ and $r - \min_0(r)$ are even and

$$\begin{aligned}\mathcal{P}^{(1)}_{\mathbf{x}}(2^\ell + r) &= \tfrac{3}{2}\Delta_0(2^\ell + r) + 1 && \text{(by Proposition 20)}\\ &= \tfrac{3}{2}(\Delta_0(r) + 2) + 1 && \text{(by Proposition 25)}\\ &= \mathcal{P}^{(1)}_{\mathbf{x}}(r) + 3 && \text{(by Proposition 20).}\end{aligned}$$

If $\Delta_0(2^\ell+r)$ is even and $2^\ell+r-\min_0(2^\ell+r)$ is odd, then $\Delta_0(r)$ is even and $r-\min_0(r)$ is odd. Then

$$\begin{aligned}\mathcal{P}^{(1)}_{\mathbf{x}}(2^\ell + r) &= \tfrac{3}{2}\Delta_0(2^\ell + r) + 2 && \text{(by Proposition 20)}\\ &= \tfrac{3}{2}(\Delta_0(r) + 2) + 2 && \text{(by Proposition 25)}\\ &= \mathcal{P}^{(1)}_{\mathbf{x}}(r) + 3 && \text{(by Proposition 20).}\square\end{aligned}$$

One can prove the following result in a manner similar to the proof of Theorem 4. There may be simpler recurrences, but these relations exhibit the same symmetry as in Theorem 4.

Theorem 32. *The abelian complexity sequence $\mathcal{P}^{(1)}_{\mathbf{x}}(n)_{n\geq 0}$ of the 2-block coding of the* *period-doubling word satisfies the following relations.*

$$\begin{aligned}
\mathcal{P}_{\mathbf{x}}^{(1)}(8n) &= \mathcal{P}_{\mathbf{x}}^{(1)}(2n)\\
4\mathcal{P}_{\mathbf{x}}^{(1)}(8n+1) &= -2\mathcal{P}_{\mathbf{x}}^{(1)}(2n+1)+7\mathcal{P}_{\mathbf{x}}^{(1)}(4n+1)-2\mathcal{P}_{\mathbf{x}}^{(1)}(4n+2)+\mathcal{P}_{\mathbf{x}}^{(1)}(4n+3)\\
4\mathcal{P}_{\mathbf{x}}^{(1)}(8n+2) &= -6\mathcal{P}_{\mathbf{x}}^{(1)}(2n+1)+9\mathcal{P}_{\mathbf{x}}^{(1)}(4n+1)-2\mathcal{P}_{\mathbf{x}}^{(1)}(4n+2)+3\mathcal{P}_{\mathbf{x}}^{(1)}(4n+3)\\
4\mathcal{P}_{\mathbf{x}}^{(1)}(8n+3) &= -6\mathcal{P}_{\mathbf{x}}^{(1)}(2n+1)+5\mathcal{P}_{\mathbf{x}}^{(1)}(4n+1)+2\mathcal{P}_{\mathbf{x}}^{(1)}(4n+2)+3\mathcal{P}_{\mathbf{x}}^{(1)}(4n+3)\\
\mathcal{P}_{\mathbf{x}}^{(1)}(8n+4) &= \mathcal{P}_{\mathbf{x}}^{(1)}(4n+2)\\
4\mathcal{P}_{\mathbf{x}}^{(1)}(8n+5) &= -6\mathcal{P}_{\mathbf{x}}^{(1)}(2n+1)+3\mathcal{P}_{\mathbf{x}}^{(1)}(4n+1)+2\mathcal{P}_{\mathbf{x}}^{(1)}(4n+2)+5\mathcal{P}_{\mathbf{x}}^{(1)}(4n+3)\\
4\mathcal{P}_{\mathbf{x}}^{(1)}(8n+6) &= -6\mathcal{P}_{\mathbf{x}}^{(1)}(2n+1)+3\mathcal{P}_{\mathbf{x}}^{(1)}(4n+1)-2\mathcal{P}_{\mathbf{x}}^{(1)}(4n+2)+9\mathcal{P}_{\mathbf{x}}^{(1)}(4n+3)\\
4\mathcal{P}_{\mathbf{x}}^{(1)}(8n+7) &= -2\mathcal{P}_{\mathbf{x}}^{(1)}(2n+1)+\mathcal{P}_{\mathbf{x}}^{(1)}(4n+1)-2\mathcal{P}_{\mathbf{x}}^{(1)}(4n+2)+7\mathcal{P}_{\mathbf{x}}^{(1)}(4n+3)
\end{aligned}$$

5 2-abelian complexity of the period-doubling word

To prove the 2-regularity of $\mathcal{P}_{\mathbf{p}}^{(2)}(n)_{n\geq 0}$, the aim of this section is to express the 2-abelian complexity $\mathcal{P}_{\mathbf{p}}^{(2)}$ in terms of the 1-abelian complexity $\mathcal{P}_{\mathbf{x}}^{(1)}$ and the following additional 2-regular functions.

Definition 33. We define the *max-jump* function $\mathrm{MJ}_0 : \mathbb{N} \to \{0,1\}$ by $\mathrm{MJ}_0(0) = 0$ and, for $n \geq 1$,

$$\mathrm{MJ}_0(n) = \begin{cases} 1 & \text{if } \max_0(n) > \max_0(n-1)\\ 0 & \text{otherwise,}\end{cases}$$

i.e., $\mathrm{MJ}_0(n) = 1$ when the function $\max_0$ increases. Similarly, let $\mathrm{mj}_0 : \mathbb{N} \to \{0,1\}$ be the *min-jump* function defined by

$$\mathrm{mj}_0(n) = \begin{cases} 1 & \text{if } \min_0(n+1) > \min_0(n)\\ 0 & \text{otherwise.}\end{cases}$$

Since $\max_0(n)$ and $\min_0(n)$ are non-decreasing, we can write

$$\begin{aligned}
\mathrm{MJ}_0(n+1) &= \max_0(n+1) - \max_0(n),\\
\mathrm{mj}_0(n) &= \min_0(n+1) - \min_0(n).
\end{aligned}$$

The relationship between these sequences and $\mathcal{P}_{\mathbf{p}}^{(2)}$ and $\mathcal{P}_{\mathbf{x}}^{(1)}$ is stated in the following result.

Proposition 34. *Let $n \geq 1$ be an integer. Then*

$$\mathcal{P}_{\mathbf{p}}^{(2)}(n+1) - \mathcal{P}_{\mathbf{x}}^{(1)}(n) = \begin{cases} 0 & \text{if } n \text{ is odd}\\ \frac{\Delta_0(n)}{2} + 1 - \mathrm{MJ}_0(n) - \mathrm{mj}_0(n) & \text{if } n \text{ is even.}\end{cases}$$

We require several preliminary results.

Proposition 35. *Let u and v be factors of $\mathbf{p}$ of length n. Let u' and v' be the 2-block codings of u and v. The factors u and v are 2-abelian equivalent if and only if u' and v' are abelian equivalent and either u' and v' both start with 2 or none of them start with 2.*

Proof. By Lemma 15, u and v are 2-abelian equivalent if and only if they start with the same letter and have the same number of factors 00, 01 and 10. The number of 00 (respectively 01 and 10) in u is exactly the number of 0 (resp. 1 and 2) in u'. Moreover, u starts with 0 (resp. by 1) if and only if u' starts with 0 or 1 (resp. by 2). Therefore, u and v are 2-abelian equivalent if and only if u' and v' are abelian equivalent and both start with 2 or none of them start with 2. □

To compute $\mathcal{P}_{\mathbf{p}}^{(2)}$, we will use the abelian complexity of $\mathbf{x} = \text{block}(\mathbf{p}, 2)$, $\mathcal{P}_{\mathbf{x}}^{(1)}$, and study when an abelian equivalence class of $\mathbf{x}$ splits into two 2-abelian equivalence classes of $\mathbf{p}$, or in other words, study when two abelian equivalent factors of $\mathbf{x}$ can start, respectively, with 2 and with 0 or 1. If the class does not split, we say that it leads to only one class.

Lemma 36. *Let $\mathcal{X}$ be an abelian equivalence class of factors of length n of $\mathbf{x}$. If the number of 1's in an element of $\mathcal{X}$ differs from the number of 2's, then $\mathcal{X}$ leads to only one 2-abelian equivalence class of $\mathbf{p}$.*

Proof. It is enough to prove that if an element of $\mathcal{X}$ starts with 2, all the other elements of $\mathcal{X}$ start with 2. If u starts with 2, then all the elements of $\mathcal{X}$ have more 2's than 1's. But any factor with more 2's than 1's starts with a 2. □

Corollary 37. *If n is odd, $\mathcal{P}_{\mathbf{p}}^{(2)}(n+1) = \mathcal{P}_{\mathbf{x}}^{(1)}(n)$.*

Proof. Let $\mathcal{X}$ be an abelian equivalence class of factors of odd length n. If no element of $\mathcal{X}$ starts with a 2, $\mathcal{X}$ leads to only one 2-abelian equivalence class of factors of $\mathbf{p}$. So assume that there is a factor u in $\mathcal{X}$ starting with 2. Since n is odd, we can write $u = 2\phi(u')$. Then the number of 0's in u is even and there is a different number of 2's than 1's. By Lemma 36, $\mathcal{X}$ again leads to a unique 2-abelian equivalence class of $\mathbf{p}$. □

Corollary 38. *Let $\mathcal{X}$ be an abelian equivalence class of factors of $\mathbf{x}$ of even length n with an odd number of zeros. Then $\mathcal{X}$ leads to only one 2-abelian equivalence class of $\mathbf{p}$.*

Proof. Factors in $\mathcal{X}$ have an odd number of 1's and 2's counted together, so the number of 1's and the number of 2's are different and we can apply Lemma 36. □

Thus, an abelian equivalence class $\mathcal{X}$ of factors of length n of $\mathbf{x}$ can possibly lead to two 2-abelian equivalence classes of factors of length $n+1$ of $\mathbf{p}$ only if n is even and if there are an even number of zeros in $\mathcal{X}$. In most cases $\mathcal{X}$ will indeed lead to two different equivalence classes. The exceptions are identified by the following lemma.

Lemma 39. *Let n be a positive even integer and n_0 such that $\min_0(n) \leq n_0 \leq \max_0(n)$. Let $\mathcal{X}$ be an abelian equivalence class of factors of $\mathbf{x}$ of length n with exactly n_0 zeros.*

- *We have* $n_0 = \max_0(n)$ *and* $\mathrm{MJ}_0(n) = 1$ *if and only if every factor* u *in* $\mathcal{X}$ *can be written as* $u = 00u'00$.

- *We have* $n_0 = \min_0(n)$ *and* $\mathrm{mj}_0(n) = 1$ *if and only if every factor* u *in* $\mathcal{X}$ *is preceded and followed only by* 00.

Proof. We start by proving the first part of the lemma. Assume that all the elements of $\mathcal{X}$ have the form $00u'00$. In particular, n_0 is even. If $n_0 \neq \max_0(n)$, it means that there is a factor v of length n with $n_0 + 1$ zeros. Indeed, sliding a window of length n from a word of $\mathcal{X}$ to a factor with $\max_0(n)$ zeros gives factors with all possibilities between n_0 and $\max_0(n)$ for the number of zeros. Since $|v|_0$ is odd and n is even, we must have $v = 0\phi(v')1$ or $v = 2\phi(v')0$. But then $0^{-1}v2$ or $1v0^{-1}$ is an element of $\mathcal{X}$ not of the form $00u'00$, a contradiction. Hence $n_0 = \max_0(n)$. If $\mathrm{MJ}_0(n) = 0$, then $\max_0(n-1) = n_0$ and there is a factor v of odd length $n-1$ with even number n_0 of 0's. We must have $v = 2\phi(v')$ or $v = \phi(v')1$ but then $1v$ or $v2$ is an element of $\mathcal{X}$ not of the form $00u'00$, a contradiction and $\mathrm{MJ}_0(n) = 1$.

For the other direction, assume that $n_0 = \max_0(n)$ and $\mathrm{MJ}_0(n) = 1$. In particular, $\max_0(n-1) = n_0 - 1$. Assume there exists a factor u of $\mathcal{X}$ not of the form $u = 00u'00$. Since u has even length and even number of 0's, we must have $u = 01u'20$ or u has its first or last letter y not equal to 0. In the first case, $v = 001u'$ has length $n-1$ and n_0 zeros, a contradiction. In the second case, removing the letter y leads also to a factor of length $n-1$ with n_0 zeros.

The second part of the lemma is similar. Assume first that all the elements of $\mathcal{X}$ are preceded and followed by 00. In particular, n_0 is even. If $n_0 \neq \min_0(n)$, there is a factor v of length n with $n_0 - 1$ zeros. Since $|v|_0$ is odd but n is even, we must have $v = 0\phi(v')1$ or $v = 2\phi(v')0$ but then $0v1^{-1}$ or $2^{-1}v0$ is an element of $\mathcal{X}$ that starts or ends with 00 and so is preceded or followed by 12, a contradiction. Hence we have $n_0 = \min_0(n)$. If $\mathrm{mj}_0(n) = 0$, then $\min_0(n+1) = n_0$ and there is a factor v of odd length $n+1$ with even number n_0 of 0's. We must have $v = 2\phi(v')$ or $v = \phi(v')1$ but then $\phi(v')$ is an element of $\mathcal{X}$ without a 00 preceding or following it.

For the other direction, assume that $n_0 = \min_0(n)$ and $\mathrm{mj}_0(n) = 1$. In particular $\min_0(n+1) = n_0 + 1$. If there exists a factor u of $\mathcal{X}$ such that $1u$, $2u$, $u1$ or $u2$ is a factor, then $\min_0(n+1) \leq n_0$, a contradiction. Hence all the factors u of $\mathcal{X}$ can only be extended by $0u0$. Finally, note that $u \in \mathcal{X}$ cannot occur in $\mathbf{x}$ at odd index. In other words, any $u \in \mathcal{X}$ can be de-substituted. Indeed, if it is not the case, then u is of the form $0\phi(u')0$, $0\phi(u')1$, $2\phi(u')0$ or $2\phi(u')1$. If u is of the first form, then $\phi(u')001$ is a factor of length $n+1$ with only n_0 zeros, which is a contradiction. Otherwise, u is of one of the last three forms. Then either $u2$ or $1u$ is a factor of $\mathbf{x}$, which is not possible. So the only extension of u as a factor of $\mathbf{x}$ is $00u00$. $\square$

Lemma 40. *Let* n *be a positive even integer and* n_0 *even such that* $\min_0(n) \leq n_0 \leq \max_0(n)$. *Let* $\mathcal{X}$ *be an abelian equivalence class of factors of* $\mathbf{x}$ *of length* n *with* n_0 *zeros. The class* $\mathcal{X}$ *leads to only one* 2*-abelian equivalence class of* $\mathbf{p}$ *if and only if* $n_0 = \min_0(n)$ *and* $\mathrm{mj}_0(n) = 1$ *or* $n_0 = \max_0(n)$ *and* $\mathrm{MJ}_0(n) = 1$. *Otherwise,* $\mathcal{X}$ *splits into two classes.*

Proof. The factors in $\mathbf{x}$ of length $n = 2$ are $00, 01, 12, 21, 20$. The two classes to consider are $\mathcal{X}_1 = \{00\}$, which leads to one class, and $\mathcal{X}_2 = \{12, 21\}$, which splits into two classes. Since $\mathrm{MJ}_0(2) = 1$ and $\mathrm{mj}_0(2) = 0$, the proposition is true.

Hence let $n \geq 4$ even. If $n_0 = \min_0(n)$ and $\mathrm{mj}_0(n) = 1$, then by Lemma 39, all the elements of $\mathcal{X}$ are preceded by 00. In particular, they all start with 1 and $\mathcal{X}$ leads to only one 2-abelian equivalence class. Similarly, if $n_0 = \max_0(n)$ and $\mathrm{MJ}_0(n) = 1$, then by Lemma 39, all the elements of $\mathcal{X}$ start with 0 and we have only one class.

Assume now that $\mathcal{X}$ leads to only one class. If an element u of $\mathcal{X}$ starts with 2, we have $u = 2\phi(u')1$ since n and n_0 are even. Then $1u1^{-1}$ is an element of $\mathcal{X}$ starting with 1 and $\mathcal{X}$ splits into two classes. Hence every element u of $\mathcal{X}$ starts with 0 or 1. Assume there exists a factor u in $\mathcal{X}$ that starts with a 1. Then $u = 12\phi(u')$ and u cannot be followed by a 1 since otherwise $1^{-1}u1$ would be an element of $\mathcal{X}$ starting with 2. Hence u is always followed by 00 and so ends with 12. Similarly, it can only be preceded by 00. Hence all the factors in $\mathcal{X}$ starting with a 1 are preceded and followed by 00. In particular, if a factor in $\mathcal{X}$ starts with 1 and occurs in $\mathbf{x}$ at index i, then the two factors starting at indices $i-1$ and $i+1$ in $\mathbf{x}$ have $n_0 + 1$ zeros. Assume now there exists a factor u in $\mathcal{X}$ starting with a 0. Then, u can be de-substituted. Otherwise, as n and n_0 are even, u is of the form $0\phi(u')0$ where $\phi(u')$ ends with 12. Thus $2\phi(u')2^{-1}$ is an element of $\mathcal{X}$ starting with 2, which is a contradiction. Hence u starts with 00. If u ends with 12, then again, $2u2^{-1}$ is an element of $\mathcal{X}$ starting with 2. Hence $u = 00\phi(u')00$ and all elements of $\mathcal{X}$ starting with 0 start and end with 00. In particular, if a factor in $\mathcal{X}$ starts with 0 and occurs in $\mathbf{x}$ at index i, then the two factors starting at indices $i-1$ and $i+1$ in $\mathbf{x}$ have $n_0 - 1$ zeros.

If no elements of $\mathcal{X}$ start with 1 or no elements start with 0, we are done by Lemma 39. Otherwise, since one can show that $\mathbf{x}$ is uniformly recurrent[3], we can assume that there exist a factor $u \in \mathcal{X}$ that starts with 0 and occurs at index i in $\mathbf{x}$, and a factor $v \in \mathcal{X}$ that starts with 1 and occurs at index $i + \ell$ in $\mathbf{x}$, such that any factor w_s of length n occurring at index $i + s$ in $\mathbf{x}$ does not belong to $\mathcal{X}$ for $0 < s < \ell$. Then w_1 has $n_0 - 1$ zeros whereas $w_{\ell-1}$ has $n_0 + 1$ zeros. But there is no factor w_s with n_0 zeros. This is a contradiction since the number of 0's changes by at most one between two factors of the same length starting at consecutive indexes. $\square$

Proof of Proposition 34. The case n odd is given by Corollary 37. Assume now that n is even. Then by Lemma 19, $\min_0(n)$ and $\max_0(n)$ are even, and therefore $\Delta_0(n)$ is even as well. Let $\mathcal{X}$ be an abelian equivalence class of factors of $\mathbf{x}$ of length n. Let n_0 be the number of 0's in the elements of $\mathcal{X}$. There are exactly $\frac{\Delta_0(n)}{2}$ odd values of n_0 and $\frac{\Delta_0(n)}{2} + 1$ even values. By Corollary 38, if n_0 is odd, $\mathcal{X}$ leads to one 2-abelian equivalence class of $\mathbf{p}$. By Lemma 40, $\mathcal{X}$ splits into two classes except for $n_0 = \min_0(n)$ if $\mathrm{mj}_0(n) = 1$ and for $n_0 = \max_0(n)$ if $\mathrm{MJ}_0(n) = 1$. Hence there are in total $\frac{\Delta_0(n)}{2} + 1 - \mathrm{MJ}_0(n) - \mathrm{mj}_0(n)$ cases

[3] A word is *uniformly recurrent* if every factor occurs infinitely often and, for each factor, there is a constant c such that two consecutive occurrences of the factor occur within c of each other. To prove that $\mathbf{x}$ is uniformly recurrent, it is enough to observe that ϕ is primitive since for each letter $y \in \{0, 1, 2\}$, $\phi^3(y)$ contains all the letters.

where $\mathcal{X}$ leads to two 2-abelian equivalence classes of $\mathbf{p}$ instead of one and this is exactly the difference between $\mathcal{P}_{\mathbf{p}}^{(2)}(n+1)$ and $\mathcal{P}_{\mathbf{x}}^{(1)}(n)$. $\square$

Corollary 41. *The sequence $\mathcal{P}_{\mathbf{p}}^{(2)}(n)_{n\geq 0}$ is 2-regular.*

Proof. We can make use of Lemma 8. Thanks to Proposition 34, $\mathcal{P}_{\mathbf{p}}^{(2)}(n+1)$ can be expressed as a combination of $\mathcal{P}_{\mathbf{x}}^{(1)}(n)$, $\Delta_0(n)$, $\mathrm{MJ}_0(n)$, $\mathrm{mj}_0(n)$ using the predicate (n mod 2). Note that the predicate (n mod 2) is trivially 2-automatic.

We proved the 2-regularity of $\mathcal{P}_{\mathbf{x}}^{(1)}(n)_{n\geq 0}$ and of $\Delta_0(n)_{n\geq 0}$ in Section 4. Observe that

$$\mathrm{MJ}_0(n+1) = \max_0(n+1) - \max_0(n) = \min_0(n+1) + \Delta_0(n+1) - \min_0(n) - \Delta_0(n).$$

Since $\mathrm{MJ}_0(n+1)$ can only take the values 0 and 1, the latter relation can also be expressed using $(\min_0(n) \bmod 2)_{n\geq 0}$ and $(\Delta_0(n) \bmod 2)_{n\geq 0}$. These latter sequences are 2-regular by Corollary 26. By Lemma 11, $\mathrm{MJ}_0(n+1)_{n\geq 0}$ is thus a combination of four 2-regular sequences. Applying again Lemma 11, $\mathrm{MJ}_0(n)_{n\geq 0}$ is also 2-regular. We can show similarly that $\mathrm{mj}_0(n)_{n\geq 0}$ is 2-regular. In fact, both sequences $\mathrm{MJ}_0(n)_{n\geq 0}$ and $\mathrm{mj}_0(n)_{n\geq 0}$ are 2-automatic since they only take values 0 and 1. Thus, all the functions in the expression for $\mathcal{P}_{\mathbf{p}}^{(2)}(n+1)$ are 2-regular.

Finally, as $\mathcal{P}_{\mathbf{p}}^{(2)}(n+1)_{n\geq 0}$ is 2-regular, $\mathcal{P}_{\mathbf{p}}^{(2)}(n)_{n\geq 0}$ is 2-regular by Lemma 11. $\square$

6 Abelian complexity of $\mathrm{block}(\mathbf{t}, 2)$

In this section, we turn our attention to the Thue–Morse word $\mathbf{t}$. Let $\mathbf{y}$ denote

$$\mathrm{block}(\mathbf{t}, 2) = 132120132012132120121320\cdots,$$

the 2-block coding of $\mathbf{t}$ introduced in Example 17. Recall that $\mathbf{y}$ is a fixed point of the morphism ν defined by $\nu : 0 \mapsto 12, 1 \mapsto 13, 2 \mapsto 20, 3 \mapsto 21$. The approach here is similar to that of the period-doubling word: we consider in this section the abelian complexity of $\mathbf{y}$, and then we compare $\mathcal{P}_{\mathbf{y}}^{(1)}(n)$ with $\mathcal{P}_{\mathbf{t}}^{(2)}(n)$ in Section 7.

Our study of the period-doubling word in Sections 4 and 5 made substantial use of counting 0's in factors of $\mathbf{x}$. Alternatively, we could have counted the total number of 1's and 2's in factors of $\mathbf{x}$, since this is equivalent information and since the letters 1 and 2 alternate in $\mathbf{x}$.

For the Thue–Morse word, the appropriate statistic for factors of $\mathbf{y}$ is the total number of 1's and 2's (or, equivalently, the total number of 0's and 3's). We will show in Lemma 45 that the letters 1 and 2 alternate in $\mathbf{y}$. Therefore, for $n \in \mathbb{N}$ we set

$$\begin{aligned}
\max_{12}(n) &:= \max\{|u|_1 + |u|_2 : u \text{ is a factor of } \mathbf{y} \text{ with } |u| = n\},\\
\min_{12}(n) &:= \min\{|u|_1 + |u|_2 : u \text{ is a factor of } \mathbf{y} \text{ with } |u| = n\},\\
\Delta_{12}(n) &:= \max_{12}(n) - \min_{12}(n).
\end{aligned}$$

Remark 42. Note that $g(\mathbf{y})$ is exactly the period-doubling word $\mathbf{p}$, where g is the coding defined by $g(0) = 1$, $g(1) = 0$, $g(2) = 0$ and $g(3) = 1$. In particular, $\Delta_{12}(n) + 1$ is the abelian complexity function of the period-doubling word. This function was also studied in [5, 14]. Here we obtain relations of the same type as the relations in Theorem 4.

The fact that $\mathcal{P}^{(1)}_{\mathbf{y}}(n)_{n\geq 0}$ is 2-regular will follow from the next statement.

Proposition 43. *Let $n \in \mathbb{N}$. We have*

$$\mathcal{P}^{(1)}_{\mathbf{y}}(n) = \begin{cases} 2\Delta_{12}(n) + 2 & \text{if } n \text{ is odd} \\ \frac{5}{2}\Delta_{12}(n) + \frac{5}{2} & \text{if } n \text{ and } \Delta_{12}(n)+1 \text{ are even} \\ \frac{5}{2}\Delta_{12}(n) + 4 & \text{if } n,\ \Delta_{12}(n) \text{ and } \min_{12}(n)+1 \text{ are even} \\ \frac{5}{2}\Delta_{12}(n) + 1 & \text{if } n,\ \Delta_{12}(n) \text{ and } \min_{12}(n) \text{ are even.} \end{cases} \tag{3}$$

To be able to apply the composition result given by Lemma 8 to the expression of $\mathcal{P}^{(1)}_{\mathbf{y}}$ derived in Proposition 43, we have therefore to prove that

- the sequence $\Delta_{12}(n)_{n\geq 0}$ is 2-regular and
- the predicates occurring in (3) are 2-automatic.

Section 6.1 is dedicated to the proof of Proposition 43. In Section 6.2, we give a proof of the two previous items. In particular, we show that $\Delta_{12}(n)_{n\geq 0}$ satisfies a reflection symmetry. This permits us to express recurrence relations for $\mathcal{P}^{(1)}_{\mathbf{y}}$ at the end of Section 6.2.

6.1 Proof of Proposition 43

We first need three technical lemmas about factors of $\mathbf{y} = \mathrm{block}(\mathbf{t}, 2)$.

Lemma 44. *The set of factors of $\mathbf{y}$ of length 2 is $\mathrm{Fac}_{\mathbf{y}}(2) = \{01, 12, 13, 20, 21, 32\}$.*

Proof. It is easy to check that these six words are factors. To prove that they are the only ones, it is enough to check that for any element u in $\{01, 12, 13, 20, 21, 32\}$ the three factors of length 2 of $\nu(u)$ are still in $\{01, 12, 13, 20, 21, 32\}$. □

The following lemma has already been observed in [14, Lemma 10].

Lemma 45. *If w is a factor of $\mathbf{y}$, then $\big||w|_1 - |w|_2\big| \leq 1$ and $\big||w|_0 - |w|_3\big| \leq 1$. In particular, the letters 1 and 2 (respectively 0 and 3) alternate in $\mathbf{y}$.*

Proof. First note that if for all factors of a word u, the numbers of two letters x and y differ by at most 1, then x and y alternate in u. Furthermore, if the first or the last occurrence of one of these letters is x, then $|u|_x \geq |u|_y$. If both the first and the last occurrences are x, then $|u|_x = |u|_y + 1$.

We prove the result by induction on the length ℓ of the factor. The result is true for factors of length $\ell = 1$. Let w be a factor of length $\ell > 1$ and assume the result holds for factors of length smaller than ℓ. If w can be de-substituted as $w = \nu(w')$, we have

$$\begin{aligned}
|w|_0 &= |w'|_2, \\
|w|_1 &= |w'|_0 + |w'|_1 + |w'|_3, \\
|w|_2 &= |w'|_0 + |w'|_2 + |w'|_3, \\
|w|_3 &= |w'|_1.
\end{aligned}$$

Using the induction hypothesis, we have

$$\big||w|_1 - |w|_2\big| = \big||w|_0 - |w|_3\big| = \big||w'|_1 - |w'|_2\big| \leq 1.$$

If w cannot be de-substituted and has odd length, we have

$$w \in \left\{1^{-1}\nu(w'),\, 2^{-1}\nu(w'),\, \nu(w')1,\, \nu(w')2\right\}$$

for some factor w' with $|w'| < \ell$. Assume that $w = 1^{-1}\nu(w')$. Then as before $\big||w|_0 - |w|_3\big| = \big||w'|_1 - |w'|_2\big| \leq 1$. For the numbers of 1 and 2, w' starts with 0 or 1. Since by Lemma 44 a 0 is always followed by a 1, w' starts either with 01 or with 1. In both cases, since 1 and 2 alternate, we have $|w'|_1 \geq |w'|_2$ and thus

$$\big||w|_1 - |w|_2\big| = \big||w'|_1 - |w'|_2 - 1\big| \leq 1.$$

The same reasoning can be done for $w = 2^{-1}\nu(w')$. If $w = \nu(w')1$, then we clearly have $\big||w|_0 - |w|_3\big| \leq 1$ using the result on $\nu(w')$. By Lemma 44, the factor $\nu(w')$ must end either with 0 or 2. So w' ends with 0 or 2 as well. Since a 0 is always preceded by a 2, we necessarily have $|w'|_2 \geq |w'|_1$ and

$$\big||w|_1 - |w|_2\big| = \big||w'|_1 - |w'|_2 + 1\big| \leq 1.$$

The same reasoning applies to $w = \nu(w')2$.

If w cannot be de-substituted and has even length, then we have

$$w \in \left\{1^{-1}\nu(w')1,\, 1^{-1}\nu(w')2,\, 2^{-1}\nu(w')1,\, 2^{-1}\nu(w')2\right\}$$

for some factor w' with $|w'| < \ell$. If the same letter is removed and added to $\nu(w')$, then the result is clearly true. Otherwise, assume that $w = 1^{-1}\nu(w')2$ (the same reasoning holds for the last case). It is clear that $\big||w|_0 - |w|_3\big| \leq 1$ using the result on $\nu(w')$. For the numbers of 1 and 2, as before, w' starts with 01 or 1 and ends with 13 or 1. Hence we have $|w'|_1 = |w'|_2 + 1$ and then

$$\big||w|_1 - |w|_2\big| = \big||w'|_1 - |w'|_2 - 2\big| \leq 1. \qquad \square$$

Lemma 46. *Let τ, τ' be the morphisms respectively defined by*

$$\tau : \begin{cases} 0 \mapsto 0 \\ 1 \mapsto 2 \\ 2 \mapsto 1 \\ 3 \mapsto 3 \end{cases} \qquad \textit{and} \qquad \tau' : \begin{cases} 0 \mapsto 3 \\ 1 \mapsto 1 \\ 2 \mapsto 2 \\ 3 \mapsto 0 \end{cases} .$$

If w is a factor of $\mathbf{y}$, then $\tau'(w)^{\mathrm{R}}$, $\tau(w)^{\mathrm{R}}$ and $\tau'(\tau(w))$ are also factors of $\mathbf{y}$.

Proof. We prove the lemma for $\tau'(w)^{\mathrm{R}}$ and $\tau(w)^{\mathrm{R}}$ since $\tau'(\tau(w)) = \tau'(\tau(w)^{\mathrm{R}})^{\mathrm{R}}$.

We first prove by induction that for any factor u starting with the letter x and ending with the letter y,

$$\tau'(\nu(u))^{\mathrm{R}} = a^{-1}\nu(\tau(u)^{\mathrm{R}})b \tag{4}$$

where $a = 1$ (respectively $a = 2$, $b = 1$, $b = 2$) if and only if $y \in \{0, 2\}$ (resp. $y \in \{1, 3\}$, $x \in \{0, 1\}$, $x \in \{2, 3\}$). Note that $a^{-1}\nu(\tau(u)^{\mathrm{R}})b$ is well defined. Indeed, if $y \in \{0, 2\}$, then $\tau(u)^{\mathrm{R}}$ starts with 0 or 1 and thus $\nu(\tau(u)^{\mathrm{R}})$ starts with $a = 1$. The same holds with $y \in \{1, 3\}$.

The relation (4) is true for u of length 1. We have for example

$$\tau'(\nu(0))^{\mathrm{R}} = 21 = 1^{-1}\nu(0)1 = 1^{-1}\nu(\tau(0)^{\mathrm{R}})1$$

and

$$\tau'(\nu(1))^{\mathrm{R}} = 01 = 2^{-1}\nu(2)1 = 2^{-1}\nu(\tau(1)^{\mathrm{R}})1.$$

Let $u = u'yx$ be a factor with at least two letters x and y. Assume the conclusion holds for words of length at most $|u| - 1$. By the induction hypothesis, we have $\tau'(\nu(u'y))^{\mathrm{R}} = a^{-1}\nu(\tau(u'y)^{\mathrm{R}})b$ and $\tau'(\nu(x))^{\mathrm{R}} = c^{-1}\nu(\tau(x)^{\mathrm{R}})d$ with appropriate a, b, c, d. Since yx is a factor, one can check using Lemma 44 that $a = d$. Indeed, if $y \in \{0, 2\}$, then $x \in \{0, 1\}$. So $a = 1$ and $d = 1$. Similarly, if $y \in \{1, 3\}$, then $x \in \{2, 3\}$. Hence, $a = 2$ and $d = 2$. Thus, we have

$$\begin{aligned} \tau'(\nu(u))^{\mathrm{R}} &= \tau'(\nu(u'yx)^{\mathrm{R}}) \\ &= \tau'(\nu(x))^{\mathrm{R}}\tau'(\nu(u'y))^{\mathrm{R}} \\ &= c^{-1}\nu(\tau(x)^{\mathrm{R}})da^{-1}\nu(\tau(u'y)^{\mathrm{R}})b \\ &= c^{-1}\nu(\tau(u'yx)^{\mathrm{R}})b \\ &= c^{-1}\nu(\tau(u)^{\mathrm{R}})b. \end{aligned}$$

We can similarly prove by induction that for any factor u starting with the letter x and ending with the letter y,

$$\tau(\nu(u))^{\mathrm{R}} = a^{-1}\nu(\tau'(u)^{\mathrm{R}})b$$

where $a = 1$ (respectively $a = 2$, $b = 1$, $b = 2$) if and only if $y \in \{1, 3\}$ (resp. $y \in \{0, 2\}$, $x \in \{2, 3\}$, $x \in \{0, 1\}$).

We now prove the lemma (for τ and τ' together) by induction on the length of w. One can check by hand that the lemma is true for w of length at most 4. Assume the lemma is true for any factor of length at most $n \geq 4$, and let w be a factor of length $n+1$. There exist some factors s, t and v such that $swt = \nu(v)$, $0 \leq |t| \leq 1$ and $1 \leq |s| \leq 2$. Then we have $|v| \leq \frac{n+4}{2} \leq n$. By the induction hypothesis, $\tau(v)^{\mathrm{R}}$ is a factor of $\mathbf{y}$. Hence $\nu(\tau(v)^{\mathrm{R}})$ is also a factor of $\mathbf{y}$. Using the previous result, $\tau'(\nu(v))^{\mathrm{R}} = a^{-1}\nu(\tau(v)^{\mathrm{R}})b$ for some letters a and b. But we also have $\tau'(\nu(v))^{\mathrm{R}} = \tau'(t)^{\mathrm{R}}\tau'(w)^{\mathrm{R}}\tau'(s)^{\mathrm{R}}$ and since s has at least one letter, $\tau'(w)^{\mathrm{R}}$ is a factor of $\nu(\tau(v)^{\mathrm{R}})$. Hence it is a factor of $\mathbf{y}$. We do the same proof for $\tau(w)^{\mathrm{R}}$. □

We are now ready to prove the relationship between $\mathcal{P}^{(1)}_{\mathbf{y}}(n)$ and $\Delta_{12}(n)$.

Proof of Proposition 43. Let u be a factor of length n of $\mathbf{y}$. Let $n_{12} = |u|_1 + |u|_2$ and $n_{03} = |u|_0 + |u|_3$.

Assume first that n is odd. If n_{12} is even, then there are the same number of 1's and 2's in u by Lemma 45. Since n_{13} is odd, if $|u|_0 = |u|_3+1$ (resp. $|u|_3 = |u|_0+1$), then $\tau'(u)^{\mathrm{R}}$ is a factor by Lemma 46 and $|\tau'(u)^{\mathrm{R}}|_3 = |\tau'(u)^{\mathrm{R}}|_0 + 1$ (resp. $|\tau'(u)^{\mathrm{R}}|_0 = |\tau'(u)^{\mathrm{R}}|_3 + 1$). In either case, $\tau'(u)^{\mathrm{R}}$ still has n_{12} ones and twos. Hence there are exactly two abelian equivalence classes for fixed n odd and n_{12} even. We can do the same reasoning if n_{12} is odd. Finally, there are $\Delta_{12}(n) + 1$ possible values for n_{12} and thus $2(\Delta_{12}(n) + 1)$ abelian equivalence classes for a fixed odd n.

Assume now that n is even. If both n_{12} and n_{03} are even, then u necessarily has the same number of 1's as 2's and the same number of 0's as 3's, and thus there is only one abelian equivalence class. Hence assume that n_{12} and n_{03} are odd. We have $(|u|_0 - |u|_3, |u|_1 - |u|_2) \in \{-1, 1\}^2$. By Lemma 46, the four factors u, $\tau'(u)^{\mathrm{R}}$, $\tau(u)^{\mathrm{R}}$ and $\tau'(\tau(u))$ realize the four possibilities for $(|u|_0 - |u|_3, |u|_1 - |u|_2)$. Hence if n_{12} and n_{03} are both odd, there are four abelian equivalence classes.

Now, we just have to count pairs (n, n_{12}) with n and n_{12} even. If $\Delta_{12}(n)$ is odd, there are exactly $(\Delta_{12}(n) + 1)/2$ such pairs. So there are

$$1 \cdot (\Delta_{12}(n) + 1)/2 + 4 \cdot (\Delta_{12}(n) + 1)/2 = \frac{5}{2}(\Delta_{12}(n) + 1)$$

abelian classes for this value of n. If $\Delta_{12}(n)$ is even and $\min_{12}(n)$ is odd, there are exactly $\Delta_{12}(n)/2$ even values for n_{12}, and so there are

$$1 \cdot \Delta_{12}(n)/2 + 4 \cdot (\Delta_{12}(n)/2 + 1) = \frac{5}{2}\Delta_{12}(n) + 4$$

abelian classes. Finally, if $\Delta_{12}(n)$ is even and $\min_{12}(n)$ is even, there are $\Delta_{12}(n)/2 + 1$ even values for n_{12}, and so there are

$$1 \cdot (\Delta_{12}(n)/2 + 1) + 4 \cdot \Delta_{12}(n)/2 = \frac{5}{2}\Delta_{12}(n) + 1$$

abelian classes. □

6.2 $\Delta_{12}(n)_{n\geq 0}$ is 2-regular, $(\min_{12}(n) \bmod 2)_{n\geq 0}$ is 2-automatic

In this section, we prove the following result.

Proposition 47. *Let $\ell \geq 1$ and r such that $0 \leq r < 2^\ell$. We have*

$$\Delta_{12}(2^\ell + r) = \begin{cases} \Delta_{12}(r) + 1 & \text{if } r \leq 2^{\ell-1} \\ \Delta_{12}(2^{\ell+1} - r) & \text{if } r > 2^{\ell-1}. \end{cases}$$

Moreover,

$$\min_{12}(2^\ell + r) \equiv \begin{cases} \min_{12}(r) + \ell \pmod 2 & \text{if } r \leq 2^{\ell-1} \\ \min_{12}(2^{\ell+1} - r) + \Delta_{12}(2^{\ell+1} - r) \pmod 2 & \text{if } r > 2^{\ell-1}. \end{cases}$$

Note that those latter relations have a form similar to (but slightly different from) the assumptions of Theorem 4. Before giving the proof, we prove a corollary. The 2-regularity of $\mathcal{P}^{(1)}_{\mathbf{y}}(n)_{n\geq 0}$ follows from Proposition 43 and Corollary 48.

Corollary 48. *The following statements are true.*

- *The sequence $\Delta_{12}(n)_{n\geq 0}$ is 2-regular.*
- *The sequence $(\Delta_{12}(n) \bmod 2)_{n\geq 0}$ is 2-automatic.*
- *The sequence $(\min_{12}(n) \bmod 2)_{n\geq 0}$ is 2-automatic.*

Proof. The first assertion is a direct consequence of Proposition 47 and Theorem 4. The second assertion follows from Lemma 10.

To prove the last assertion, we prove by induction that, modulo 2,

$$\min_{12}(16n + i) \equiv \begin{cases} \min_{12}(4n) & \text{if } i = 0 \\ \min_{12}(4n+1) & \text{if } i \in \{1,4,5\} \\ \min_{12}(4n+1) + 1 & \text{if } i \in \{2,3\} \\ \min_{12}(4n+2) & \text{if } i \in \{6,8,9\} \\ \min_{12}(4n+2) + 1 & \text{if } i \in \{7,10\} \\ \min_{12}(4n+3) & \text{if } i \in \{12,13,15\} \\ \min_{12}(4n+3) + 1 & \text{if } i \in \{11,14\} \end{cases}$$

and

$$\Delta_{12}(16n + i) \equiv \begin{cases} \Delta_{12}(4n) & \text{if } i = 0 \\ \Delta_{12}(4n+1) & \text{if } i \in \{1,2,4\} \\ \Delta_{12}(4n+1) + 1 & \text{if } i \in \{3,5\} \\ \Delta_{12}(4n+2) & \text{if } i = 8 \\ \Delta_{12}(4n+2) + 1 & \text{if } i \in \{6,7,9,10\} \\ \Delta_{12}(4n+3) & \text{if } i \in \{12,14,15\} \\ \Delta_{12}(4n+3) + 1 & \text{if } i \in \{11,13\}. \end{cases}$$

The relations are true for $n = 0$. Let $n > 0$ and assume they are true for $n' < n$. We can write $n = 2^\ell + r$ with $\ell \geq 0$ and $0 \leq r < 2^\ell$. Let $i \in \{0, \ldots, 15\}$. We consider two cases.

Assume first that $r < 2^{\ell-1}$. We have $16n + i = 2^{\ell+4} + 16r + i$ and $16r + i < 2^{\ell+3}$.

$$\begin{aligned}
\min_{12}(16n+i) &\equiv \min_{12}(16r+i) + \ell + 4 && \text{(Proposition 47)}\\
&\equiv \min_{12}(4r+j) + \delta + \ell + 4 && \text{(induction)}\\
&\equiv \min_{12}(2^{\ell+2} + 4r + j) + \delta && \text{(Proposition 47)}\\
&\equiv \min_{12}(4n+j) + \delta \pmod 2
\end{aligned}$$

for some $j \in \{0, \ldots, 3\}$ and $\delta \in \{0, 1\}$ according to the relations. A similar reasoning holds for the Δ_{12} relations.

Assume now that $r \geq 2^{\ell-1}$ and $i \neq 0$. Setting $i' = 16 - i$ and $n' = 2^{\ell+1} - r - 1$, we obtain $16n' + i' = 2^{\ell+5} - 16r - i$. It follows that, by Proposition 47,

$$\begin{aligned}
\min_{12}(16n+i) &\equiv \min_{12}(2^{\ell+5} - 16r - i) + \Delta_{12}(2^{\ell+5} - 16r - i)\\
&\equiv \min_{12}(16n'+i') + \Delta_{12}(16n'+i')\\
&\equiv \min_{12}(4n'+k) + \delta + \Delta_{12}(4n'+k') + \delta' && \text{(induction)}
\end{aligned}$$

for some $k, k' \in \{0, \ldots, 3\}$ and $\delta, \delta' \in \{0, 1\}$ according to the relations. Note that we have $k = k'$, so

$$\begin{aligned}
\min_{12}(16n+i) &\equiv \min_{12}(4n'+k) + \delta + \Delta_{12}(4n'+k) + \delta'\\
&\equiv \min_{12}(2^{\ell+3} - (4r+4-k)) + \delta + \Delta_{12}(2^{\ell+3} - (4r+4-k)) + \delta'\\
&\equiv \min_{12}(2^{\ell+2} + (4r+4-k)) + \delta + \delta' && \text{(Proposition 47)}\\
&\equiv \min_{12}(4n + (4-k)) + \delta + \delta' \pmod 2.
\end{aligned}$$

Table 1 gives the values of i', k, δ and δ' for all the values of $i \neq 0$. Observe that the values of $4 - k$ and $(\delta + \delta' \bmod 2)$ are the values given in the relation for i. To conclude the proof, consider the case $i = 0$. We have

$$\begin{aligned}
\min_{12}(16n) &\equiv \min_{12}(16(2^{\ell+1} - r)) + \Delta_{12}(16(2^{\ell+1} - r)) && \text{(Proposition 47)}\\
&\equiv \min_{12}(4(2^{\ell+1} - r)) + \Delta_{12}(4(2^{\ell+1} - r)) && \text{(induction)}\\
&\equiv \min_{12}(4n) \pmod 2 && \text{(Proposition 47).}
\end{aligned}$$

A similar reasoning works for the Δ_{12} relations. □

Proposition 47 is a direct consequence of Lemmas 49, 52 and 54 given in this section.

Lemma 49. *Let $\ell \in \mathbb{N}$, $\ell \geq 1$. We have* $\Delta_{12}(2^\ell) = 1$, $\min_{12}(2^\ell) \equiv \ell \pmod 2$,

$$\min_{12}(2^\ell) + \max_{12}(2^{\ell+1}) = 2^{\ell+1} \text{ and } \max_{12}(2^\ell) + \min_{12}(2^{\ell+1}) = 2^{\ell+1}.$$

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
i'	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
k	3	3	3	3	3	2	2	2	2	2	1	1	1	1	1
δ	0	1	0	0	1	1	0	0	1	0	0	0	1	1	0
δ'	0	0	1	0	1	1	1	0	1	1	1	0	1	0	0

Table 1: The corresponding values of $i' = 16 - i$, k, δ and δ'.

Proof. Let $\ell \geq 1$, $A_\ell = \frac{2^{\ell+1}+(-1)^\ell}{3}$ and $B_\ell = \frac{2^{\ell+1}+2(-1)^{\ell+1}}{3}$. The sequences

$$(A_\ell)_{\ell\geq 1} = (1, 3, 5, 11, 21, \ldots) \text{ and } (B_\ell)_{\ell\geq 1} = (2, 2, 6, 10, 22, \ldots)$$

are integer sequences and both satisfy the recurrence relation $X_{\ell+1} = 2^{\ell+1} - X_\ell$. Moreover we have $A_\ell = B_\ell + 1$ for even ℓ and $B_\ell = A_\ell + 1$ for odd ℓ. Note that $|\nu^\ell(1)|_1 + |\nu^\ell(1)|_2 = A_\ell$ and $|\nu^\ell(0)|_1 + |\nu^\ell(0)|_2 = B_\ell$.

We show by induction that

$$\left\{|w|_1 + |w|_2 : w \text{ factor of } \mathbf{y} \text{ with } |w| = 2^\ell\right\} = \{A_\ell, B_\ell\}.$$

Note that this result will imply the lemma and that we already have A_ℓ and B_ℓ in the set.

It is easy to check the result for $\ell = 1$. Assume the result is true for $\ell \geq 1$. Let w be a factor of $\mathbf{y}$ of length $2^{\ell+1}$. If w can be de-substituted, then $w = \nu(u)$ and $|w|_1 + |w|_2 = 2|u|_0 + |u|_1 + |u|_2 + 2|u|_3$ as in the proof of Lemma 45. Hence $|w|_1 + |w|_2 = 2|u| - (|u|_1 + |u|_2) = 2^{\ell+1} - (|u|_1 + |u|_2)$. Using the recurrence relation for A_ℓ and B_ℓ and since $|u|_1 + |u|_2 \in \{A_\ell, B_\ell\}$, we have $|w|_1 + |w|_2 \in \{A_{\ell+1}, B_{\ell+1}\}$. If w cannot be de-substituted, then we can write $w = a^{-1}\nu(u)b$ for some letters $a, b \in \{1, 2\}$ and $|\nu(u)| = 2^{\ell+1}$. So $|w|_1 + |w|_2 = |\nu(u)|_1 + |\nu(u)|_2$. Since we already proved that $|\nu(u)|_1 + |\nu(u)|_2 \in \{A_{\ell+1}, B_{\ell+1}\}$, we are done.

To prove the second assertion of the lemma, observe that $\min_{12}(2^\ell) = A_\ell$ if ℓ is odd and $\min_{12}(2^\ell) = B_\ell$ if ℓ is even. Furthermore, A_ℓ is always odd whereas B_ℓ is always even. $\square$

In order to prove Lemmas 52 and 54, we first need some technical results.

Lemma 50. *Let u be a factor of $\mathbf{y}$ of length n. We have $|u|_1 + |u|_2 = \max_{12}(n)$ if and only if $|\nu(u)|_1 + |\nu(u)|_2 = \min_{12}(2n)$, and $|u|_1 + |u|_2 = \min_{12}(n)$ if and only if $|\nu(u)|_1 + |\nu(u)|_2 = \max_{12}(2n)$.*

Proof. Recall that $|\nu(u)|_1 + |\nu(u)|_2 = 2n - (|u|_1 + |u|_2)$. Assume that $|u|_1 + |u|_2 = \max_{12}(n)$ and that $|\nu(u)|_1 + |\nu(u)|_2 = x > \min_{12}(2n)$. Thus $x = 2n - \max_{12}(n)$. There exists a factor w of length $2n$ with $x - 1$ ones and twos. We can assume that w can be de-substituted. Otherwise, we can write w as $w = a^{-1}\nu(v)b$ for some $a, b \in \{1, 2\}$. Thus $\nu(v)$ has the same length as w and the same number of 1's and 2's. So we can assume $w = \nu(v)$. Then $|v|_1 + |v|_2 = 2n - (x - 1) = \max_{12}(n) + 1$, a contradiction.

For the other direction, assume that $|u|_1 + |u|_2 = x < \max_{12}(n)$ and that $|\nu(u)|_1 + |\nu(u)|_2 = \min_{12}(2n)$. Thus $x = n - \min_{12}(n)$. As before, there exists a factor v of length n with $x + 1$ ones and twos. Then $\nu(v)$ has $\min_{12}(n) - 1$ ones and twos, a contradiction.

The second part of the lemma is similar. $\square$

Lemma 51. *Let n be an odd integer. Then we have*

$$\begin{aligned}\min_{12}(n) &= \min_{12}(n+1) - 1,\\ \max_{12}(n) &= \max_{12}(n-1) + 1.\end{aligned}$$

Proof. Let u be a factor of even length $n+1$ minimizing the number of 1's and 2's. Then either u starts with 1 or 2, or ends with 1 or 2. Indeed, if u can be de-substituted, then it starts with 1 or 2. Otherwise, its last letter is the beginning of an image of ν and thus is 1 or 2. Removing this letter, we get a word of length n with $\min_{12}(n+1) - 1$ ones and twos. Since the function $\min_{12}$ increases by 0 or 1 from n to $n+1$, we have $\min_{12}(n) = \min_{12}(n+1) - 1$.

For the second equality, consider a factor u of even length $n-1$ with $\max_{12}(n-1)$ ones and twos. There exist two letters a and b such that aub is a factor. Then, as before, since aub has even length, a or b must be a 1 or a 2. Then au or ub is a factor of length n with $\max_{12}(n-1) + 1$ ones and twos and we conclude as before. $\square$

Lemma 52. *If $\ell \geq 1$ and $0 \leq r \leq 2^{\ell-1}$, then*

$$\begin{aligned}\max_{12}(2^\ell + r) &= \max_{12}(2^\ell) + \max_{12}(r)\\ \min_{12}(2^\ell + r) &= \min_{12}(2^\ell) + \min_{12}(r).\end{aligned}$$

Proof. We prove the two results together by induction on ℓ. One checks the case $\ell = 1$. Let $\ell > 1$ and assume the result is true for $\ell - 1$. Let r such that $0 \leq r \leq 2^{\ell-1}$.

Assume first that r is even. By the induction hypothesis, there exists a factor u of length $2^{\ell-1} + r/2$ such that

$$|u|_1 + |u|_2 = \min_{12}(2^{\ell-1} + r/2) = \min_{12}(2^{\ell-1}) + \min_{12}(r/2).$$

We can write $u = vw$ with v of length $2^{\ell-1}$ and w of length $r/2$. Both the words v and w must minimize the number of 1's and 2's for their respective lengths. By Lemma 50, $\nu(u) = \nu(v)\nu(w)$ maximizes the number of 1's and 2's and so do $\nu(v)$ and $\nu(w)$. Thus, $\max_{12}(2^\ell + r) = |\nu(u)|_1 + |\nu(u)|_2$ and

$$\max_{12}(2^\ell + r) = |\nu(v)|_1 + |\nu(v)|_2 + |\nu(w)|_1 + |\nu(w)|_2 = \max_{12}(2^\ell) + \max_{12}(r).$$

A similar proof shows that $\min_{12}(2^\ell + r) = \min_{12}(2^\ell) + \min_{12}(r)$.

Assume now that r is odd. We still have $0 \leq r - 1 < r + 1 \leq 2^{\ell-1}$. Hence we can apply the previous result to obtain $\max_{12}(2^\ell + r - 1) = \max_{12}(2^\ell) + \max_{12}(r-1)$. By Lemma 51,

$$\begin{aligned}\max_{12}(2^\ell + r) &= \max_{12}(2^\ell + r - 1) + 1\\ &= \max_{12}(2^\ell) + \max_{12}(r-1) + 1\\ &= \max_{12}(2^\ell) + \max_{12}(r).\end{aligned}$$

For the $\min_{12}$ equality, a similar argument holds (using the previous result for $r+1$). $\square$

Lemma 53. *If $\ell \geq 1$ and $2^{\ell-1} \leq r \leq 2^\ell$, then*

$$\begin{aligned}\max_{12}(2^{\ell+1}) &= \max_{12}(2^\ell + r) + \min_{12}(2^\ell - r)\\ \min_{12}(2^{\ell+1}) &= \min_{12}(2^\ell + r) + \max_{12}(2^\ell - r).\end{aligned}$$

Moreover, there is a factor of length $2^{\ell+1}$ maximizing (resp. minimizing) the number of 1's and 2's such that the prefix of length $2^\ell + r$ also maximizes (resp. minimizes) the number of 1's and 2's.

Proof. We proceed by induction on ℓ. The result is true for $\ell = 1$ since the only non-trivial case is $r = 1$. Then $\max_{12}(4) = \max_{12}(3) + \min_{12}(1)$ and $\min_{12}(4) = \min_{12}(3) + \max_{12}(1)$ and the factors 2120 and 0132 satisfy the claim.

Let $\ell > 1$ and assume the result is true for $\ell - 1$. Let r such that $2^{\ell-1} \leq r \leq 2^\ell$. Assume first that r is even. Then $2^{\ell-2} \leq r/2 \leq 2^{\ell-1}$. By the induction hypothesis, there is a factor u of length 2^ℓ minimizing the number of 1's and 2's such that the prefix v of length $2^{\ell-1} + r/2$ minimizes the number of 1's and 2's. Thus we can write $u = vw$ and $|v|_1 + |v|_2 = \min_{12}(2^{\ell-1} + r/2)$ and necessarily $|w|_1 + |w|_2 = \max_{12}(2^{\ell-1} - r/2)$. By Lemma 50, $\nu(u)$ and $\nu(v)$ maximize the number of 1's and 2's and $\nu(w)$ minimizes the number of 1's and 2's. So we can conclude the result. A similar proof shows the other relation. If r is odd, then we still have $2^{\ell-1} \leq r - 1 \leq 2^\ell$ since $\ell > 1$. Thus we can use the previous result and together with Lemma 51, we have

$$\begin{aligned}\max_{12}(2^{\ell+1}) &= \max_{12}(2^\ell + r - 1) + \min_{12}(2^\ell - r + 1)\\ &= \max_{12}(2^\ell + r) - 1 + \min_{12}(2^\ell - r) + 1\\ &= \max_{12}(2^\ell + r) + \min_{12}(2^\ell - r).\end{aligned}$$

Similarly, using the fact that $r + 1 \leq 2^\ell$,

$$\begin{aligned}\min_{12}(2^{\ell+1}) &= \min_{12}(2^\ell + r + 1) + \max_{12}(2^\ell - r - 1)\\ &= \min_{12}(2^\ell + r) + 1 + \max_{12}(2^\ell - r) - 1\\ &= \min_{12}(2^\ell + r) + \max_{12}(2^\ell - r).\end{aligned}$$

For the construction of the factors, one can construct them using the factor $\nu(u)$ maximizing the number of 1's and 2's given for $r - 1$ and the factor $\nu(u')$ minimizing the number of 1's and 2's given for $r + 1$ in the previous construction. Since r is odd, the letter between the prefix $\nu(v)$ of length $2^\ell + r - 1$ and $2^\ell + r$ of $\nu(u)$ is 1 or 2. Since the prefix of length $2^\ell + r - 1$ of $\nu(u)$ maximizes the number of 1's and 2's, so does the prefix of length $2^\ell + r$ of $\nu(u)$. For $\min_{12}$, consider $\nu(u')$. There exist letters a and b such that $w = a^{-1}\nu(u')b$ is still a factor. We must have $a, b \in \{1, 2\}$. Then the prefix of length $2^\ell + r$ of w minimizes the number of 1's and 2's. $\square$

The previous lemma permits us to reformulate some relations between the two sequences $\max_{12}(n)_{n\geq 0}$ and $\min_{12}(n)_{n\geq 0}$.

Lemma 54. *If $\ell \geq 1$ and $2^{\ell-1} \leq r \leq 2^\ell$, then*

$$\begin{aligned}\max\nolimits_{12}(2^\ell + r) &= 2^{\ell+1} - \min\nolimits_{12}(2^{\ell+1} - r)\\ \min\nolimits_{12}(2^\ell + r) &= 2^{\ell+1} - \max\nolimits_{12}(2^{\ell+1} - r).\end{aligned}$$

Proof. From the previous lemma, we have

$$\max\nolimits_{12}(2^\ell + r) = \max\nolimits_{12}(2^{\ell+1}) - \min\nolimits_{12}(2^\ell - r).$$

By Lemma 49, we have $\max_{12}(2^{\ell+1}) = 2^{\ell+1} - \min_{12}(2^\ell)$. Moreover, by Lemma 52, since $0 \leq 2^\ell - r \leq 2^{\ell-1}$, we get

$$\min\nolimits_{12}(2^\ell - r) = \min\nolimits_{12}(2^\ell + 2^\ell - r) - \min\nolimits_{12}(2^\ell).$$

Similar relations hold when changing $\max_{12}$ to $\min_{12}$. □

The proof of Proposition 47 about the reflection relation satisfied by $\Delta_{12}(n)$ and the recurrence relation of $\min_{12}(n)$ is now immediate.

Proof of Proposition 47. If $\ell \geq 1$ and $0 \leq r \leq 2^{\ell-1}$, then subtracting the two relations provided by Lemma 52 gives

$$\Delta_{12}(2^\ell + r) = \Delta_{12}(\ell) + \Delta_{12}(r)$$

and we can conclude using the first relation given in Lemma 49, $\Delta_{12}(2^\ell) = 1$. By Lemma 52, $\min_{12}(2^\ell+r) \equiv \min_{12}(2^\ell)+\min_{12}(r) \pmod 2$. The expression for $\min_{12}(2^\ell+r)$ follows since $\min_{12}(2^\ell) \equiv \ell \pmod 2$ by Lemma 49.

If $\ell \geq 1$ and $2^{\ell-1} < r < 2^\ell$, then subtracting the two relations provided by Lemma 54 permits us to conclude the proof of the expression claimed for $\Delta_{12}(2^\ell + r)$. Moreover, using Lemma 54, we get

$$\begin{aligned}\min\nolimits_{12}(2^\ell + r) &\equiv \max\nolimits_{12}(2^{\ell+1} - r) \pmod 2\\ &\equiv \min\nolimits_{12}(2^{\ell+1} - r) + \Delta_{12}(2^{\ell+1} - r) \pmod 2.\end{aligned}$$ □

Using Propositions 43 and 47, we can express recurrence relations for $\mathcal{P}_{\mathbf{y}}^{(1)}$ as we did for the proof of Theorem 21.

Theorem 55. *Let $\ell \geq 2$ and r such that $0 \leq r < 2^\ell$. For $r \leq 2^{\ell-1}$, we have*

$$\mathcal{P}_{\mathbf{y}}^{(1)}(2^\ell + r) = \begin{cases} \mathcal{P}_{\mathbf{y}}^{(1)}(r) + 2 & \text{if } r \text{ is odd}\\ \mathcal{P}_{\mathbf{y}}^{(1)}(r) + 1 & \text{if } (r,\ \Delta_{12}(2^\ell + r) \text{ and } \min_{12}(2^\ell + r) \text{ are even})\\ & \text{or } (r \text{ and } \Delta_{12}(2^\ell + r) + 1 \text{ are even}\\ & \quad \text{and } \min_{12}(2^\ell + r) \equiv \ell + 1 \pmod 2)\\ \mathcal{P}_{\mathbf{y}}^{(1)}(r) + 4 & \text{otherwise.}\end{cases}$$

For $r > 2^{\ell-1}$, we have $\mathcal{P}_{\mathbf{y}}^{(1)}(2^\ell + r) = \mathcal{P}_{\mathbf{y}}^{(1)}(2^{\ell+1} - r)$.

7 2-abelian complexity of the Thue–Morse word

The aim of this section is to express, in Theorem 56, $\mathcal{P}^{(2)}_{\mathbf{t}}(n+1)$ in terms of $\mathcal{P}^{(1)}_{\mathbf{y}}(n)$, $\Delta_{12}(n)$, $(\min_{12}(n) \bmod 2)$ and two new functions $\mathrm{MJ}_{03}(n)$ and $\mathrm{mj}_{03}(n)$ that are defined analogously to $\mathrm{MJ}_0(n)$ and $\mathrm{mj}_0(n)$ of Section 5. Let

$$\begin{aligned}\max_{03}(n) &:= \max\{|u|_0 + |u|_3 \,:\, u \text{ is a factor of } \mathbf{y} \text{ with } |u| = n\},\\ \min_{03}(n) &:= \min\{|u|_0 + |u|_3 \,:\, u \text{ is a factor of } \mathbf{y} \text{ with } |u| = n\},\end{aligned}$$

and let

$$\mathrm{MJ}_{03}(n) := \begin{cases} 1 & \text{if } \max_{03}(n) > \max_{03}(n-1) \\ 0 & \text{otherwise,} \end{cases}$$

$$\mathrm{mj}_{03}(n) := \begin{cases} 1 & \text{if } \min_{03}(n+1) > \min_{03}(n) \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 56. *For n odd, we have*

$$\mathcal{P}^{(2)}_{\mathbf{t}}(n+1) - \mathcal{P}^{(1)}_{\mathbf{y}}(n) = \\ \begin{cases} \Delta_{12}(n) + 2 - 2\,\mathrm{MJ}_{03}(n) - 2\,\mathrm{mj}_{03}(n) & \textit{if } \min_{12}(n) \textit{ and } \Delta_{12}(n) \textit{ are even} \\ \Delta_{12}(n) + 1 - 2\,\mathrm{MJ}_{03}(n) & \textit{if } \min_{12}(n) \textit{ and } \Delta_{12}(n)+1 \textit{ are even} \\ \Delta_{12}(n) + 1 - 2\,\mathrm{mj}_{03}(n) & \textit{if } \min_{12}(n) \textit{ and } \Delta_{12}(n) \textit{ are odd} \\ \Delta_{12}(n) & \textit{if } \min_{12}(n)+1 \textit{ and } \Delta_{12}(n) \textit{ are even.} \end{cases}$$

For n even, we have

$$\mathcal{P}^{(2)}_{\mathbf{t}}(n+1) - \mathcal{P}^{(1)}_{\mathbf{y}}(n) = \begin{cases} \frac{1}{2}\Delta_{12}(n) + 1 & \textit{if } \min_{12}(n) \textit{ and } \Delta_{12}(n) \textit{ are even} \\ \frac{1}{2}\Delta_{12}(n) & \textit{if } \min_{12}(n)+1 \textit{ and } \Delta_{12}(n) \textit{ are even} \\ \frac{1}{2}\Delta_{12}(n) + \frac{1}{2} & \textit{if } \Delta_{12}(n) \textit{ is odd.} \end{cases}$$

As in Section 5, we study when an abelian equivalence class of $\mathbf{y} = \mathrm{block}(\mathbf{t}, 2)$ splits into two 2-abelian equivalence classes of $\mathbf{t}$. We have similar propositions.

Proposition 57. *Let u and v be factors of $\mathbf{t}$ of length n. Let u' and v' be the 2-block codings of u and v. The factors u and v are 2-abelian equivalent if and only if u' and v' (of length $n-1$) are abelian equivalent and either u' and v' both have first letter in $\{0,1\}$ or both have first letter in $\{2,3\}$.*

Let $\mathcal{X}$ be an abelian equivalence class of factors of $\mathbf{y}$ of length n. For a letter a, let n_a denote the number of a's in each element of $\mathcal{X}$ and let $n_{12} = n_1 + n_2$, $n_{03} = n_0 + n_3$.

Lemma 58. *If n_{12} is odd, then $\mathcal{X}$ leads to a unique 2-abelian equivalence class of $\mathbf{t}$.*

Proof. Assume that $n_1 > n_2$ (the other case is similar). Then a word of $\mathcal{X}$ cannot start with 2 since the letters 1 and 2 alternate in $\mathbf{y}$ by Lemma 45. It cannot start with 3 neither since $n_1 > n_2$ and a 3 is always followed by 2 by Lemma 44. Hence it starts with 0 or 1. Thus $\mathcal{X}$ leads to a unique 2-abelian equivalence class. □

Lemma 59. *If n and n_{12} are even, then $\mathcal{X}$ splits into two 2-abelian equivalence classes of $\mathbf{t}$.*

Proof. If n and n_{12} are even, then n_{03} is also even and thus $n_1 = n_2$ and $n_0 = n_3$. Let u be an element of $\mathcal{X}$. Then $u' = \tau'(\tau(u))$ is also an element of $\mathcal{X}$. Moreover, the first letter of u is in $\{0, 1\}$ if and only if the first letter of u' is in $\{2, 3\}$. Hence $\mathcal{X}$ splits into two 2-abelian equivalence classes. □

So the last and hardest case happens when n is odd and n_{12} is even, i.e., when n and n_{03} are odd. The MJ_{03} and mj_{03} functions permit us to handle this case.

Lemma 60. *Let n and n_{03} are odd. Let $a \in \{0, 3\}$ (resp. $b \in \{0, 3\}$) be the letter in majority (resp. in minority) in factors in $\mathcal{X}$, among $\{0, 3\}$.*

- *We have $n_{03} = \max_{03}(n)$ and $\mathrm{MJ}_{03}(n) = 1$ if and only if every factor in $\mathcal{X}$ starts and ends with a.*
- *We have $n_{03} = \min_{03}(n)$ and $\mathrm{mj}_{03}(n) = 1$ if and only if every factor in $\mathcal{X}$ is preceded and followed by b.*

Proof. Assume that $a = 0$ and $b = 3$ (the other case is symmetric). We first prove the statement for the maximum. Assume that all the factors in $\mathcal{X}$ start and end with 0. If $n_{03} < \max_{03}(n)$, by continuity of the number of 0's and 3's and since $\mathbf{y}$ is uniformly recurrent, there exists a factor yuz such that the factor yu (resp. uz) is of length n with n_{03} (resp. $n_{03} + 1$) zeros and threes. We necessarily have $z \in \{0, 3\}$ and u is not finishing with a letter in $\{0, 3\}$. Since yu has n_{03} zeros and threes, yu or $\tau'(yu)^{\mathrm{R}}$ is an element of $\mathcal{X}$ that is either not finishing or not starting with 0, a contradiction. Hence we have $n_{03} = \max_{03}(n)$. Assume now that $\max_{03}(n - 1) = n_{03}$. There exists a factor u of even length $n - 1$ with n_{03} zeros. Without loss of generality, we can assume that u has more 0's than 3's (otherwise one can consider $\tau'(u)^{\mathrm{R}}$ by Lemma 46). Since u has even length, either u occurs at an even index in $\mathbf{y}$ and is always followed by 1 or 2, or u occurs at an odd index in $\mathbf{y}$ and is always preceded by 1 or 2. In other words, there is a factor of the form yu or uy with $y \in \{1, 2\}$. Then yu or uy is an element of $\mathcal{X}$ with the first or last letter different from 0, a contradiction.

For the other direction, assume that $n_{03} = \max_{03}(n)$ and $\mathrm{MJ}_{03}(n) = 1$. Let u be a factor in $\mathcal{X}$. If $u = xu'$ or $u = u'x$ with $x \neq 0$, then u' has length $n - 1$ and n_{03} zeros and threes. Thus $\mathrm{MJ}_{03}(n) = 0$, a contradiction.

The second statement is proved in the same way. Assume that all the factors in $\mathcal{X}$ are preceded and followed by 3. If $n_{03} > \min_{03}(n)$, by continuity of the number of 0's and 3's and since $\mathbf{y}$ is uniformly recurrent, there exists a factor yuz such that the factor yu (resp. uz) is of length n with n_{03} (resp. $n_{03} - 1$) zeros and threes. We necessarily have

$z \in \{1,2\}$. Then as before yu or $\tau'(yu)^{\mathrm{R}}$ is and element of $\mathcal{X}$ that is either not always followed or not always preceded by 3, a contradiction. Hence we have $n_{03} = \min_{03}(n)$. Assume now that $\min_{03}(n+1) = n_{03}$. There exists a factor u of even length $n+1$ with n_{03} zeros. Without loss of generality, we can assume that u has more 0's than 3's (otherwise one can consider $\tau'(u)^{\mathrm{R}}$ by Lemma 46). Since u has even length, either u occurs at an even index and starts with 1 or 2 or u occurs at an odd index and ends with 1 or 2. In other words, $u = yu'$ or $u = u'y$ with $y \in \{1,2\}$ and u' is an element of $\mathcal{X}$ preceded or followed by a letter different from 3, a contradiction.

For the other direction, assume that $n_{03} = \min_{03}(n)$ and $\mathrm{mj}_{03}(n) = 1$. Let u be a factor in $\mathcal{X}$. If $u' = ux$ or $u' = xu$ is a factor with $x \in \{1,2\}$, then u' has length $n+1$ and n_{03} zeros and threes. So $\mathrm{mj}_{03}(n) = 0$, which is a contradiction. Observe also that it is impossible to have $0u$ or $u0$ as factors of $\mathbf{y}$ since $|u|_0 > |u|_3$ by assumption and the letters 0 and 3 alternate in $\mathbf{y}$ by Lemma 45. The conclusion is immediate. □

Lemma 61. *If n is odd and n_{12} is even, then $\mathcal{X}$ leads to only one 2-abelian equivalence class of $\mathbf{t}$ if and only if $n_{03} = \min_{03}(n)$ and $\mathrm{mj}_{03}(n) = 1$, or $n_{03} = \max_{03}(n)$ and $\mathrm{MJ}_{03}(n) = 1$. Otherwise, $\mathcal{X}$ splits into two classes.*

Proof. If n is odd and n_{12} is even, then n_{03} is even. Assume that $n_0 > n_3$ (the other case is symmetric). If $n_{03} = \min_{03}(n)$ and $\mathrm{mj}_{03}(n) = 1$ then, by Lemma 60, all the factors in $\mathcal{X}$ start with 0, and so $\mathcal{X}$ leads to only one class. If $n_{03} = \max_{03}(n)$ and $\mathrm{MJ}_{03}(n) = 1$, then all the factors in $\mathcal{X}$ are preceded and followed by 3. In particular, they all start with 2 and again $\mathcal{X}$ leads to only one class.

For the other direction, suppose that $\mathcal{X}$ leads to only one class. All the factors in $\mathcal{X}$ must start either with a letter in $\{0,1\}$ or with a letter in $\{2,3\}$. Assume first that all the elements of $\mathcal{X}$ start with 0 or 1. Let u be a factor in $\mathcal{X}$. If the first letter of u is 1, it must start with 120 since u has more 0's than 3's. Thus u is always preceded by 2. It cannot end with 1 (since $n_1 = n_2$). So it must end with 0 or 2. If $u = 120u'2$, then $2120u'$ is an element of $\mathcal{X}$ starting with 2, which is a contradiction. If $u = 120u'0$ then $u1$ is a factor of $\mathbf{y}$. So $20u'01$ is an element of $\mathcal{X}$ starting with 2, a contradiction. Hence u cannot start with 1 and thus starts with 0. Observe that, if u does not end with 0, then $\tau(u)^{\mathrm{R}}$ is still an element of $\mathcal{X}$ by Lemma 46 and $\tau(u)^{\mathrm{R}}$ does not start with 0, a contradiction. Hence all the factors in $\mathcal{X}$ start and end with 0. By Lemma 60, we have $n_{03} = \max_{03}(n)$ and $\mathrm{MJ}_{03}(n) = 1$.

Assume now that all the elements of $\mathcal{X}$ start with 2 or 3. Since $n_0 > n_3$, they all start with 2. Moreover, as $n_1 = n_2$, they must end with 0 or 1. If $u \in \mathcal{X}$ ends with 0, then $\tau'(u)^{\mathrm{R}} \in \mathcal{X}$ starts with 3 by Lemma 46, a contradiction. So all factors in $\mathcal{X}$ end with 1. Let $u = 2u'1$ be an element of $\mathcal{X}$. By Lemma 44, the only possible extensions of u as a factor of length $n+1$ of $\mathbf{y}$ are $1u$, $3u$, $u2$ and $u3$. If $1u$ is a factor of $\mathbf{y}$, then $12u' \in \mathcal{X}$ starts with 1, which is a contradiction. If $u2$ is factor of $\mathbf{y}$, then $\tau(u'12)^{\mathrm{R}} \in \mathcal{X}$ starts with 1, a contradiction. Hence all the factors in $\mathcal{X}$ are preceded and followed by 3 in $\mathbf{y}$. By Lemma 60, this means that $n_{03} = \min_{03}(n)$ and $\mathrm{mj}_{03}(n) = 1$. □

We are now ready to prove Theorem 56.

Proof of Theorem 56. The difference between $\mathcal{P}_{\mathbf{t}}^{(2)}(n+1)$ and $\mathcal{P}_{\mathbf{y}}^{(1)}(n)$ is the number of abelian equivalence classes of factors of length n of $\mathbf{y}$ that split into two 2-abelian equivalence classes of factors of length $n+1$ of $\mathbf{t}$.

For even n, by Lemmas 58 and 59, it happens when n_{12} is even. The number of even values of $n_{12} \in \{\min_{12}(n), \ldots, \max_{12}(n)\}$ is

$$\begin{cases} \frac{1}{2}\Delta_{12}(n)+1 & \text{if } \min_{12}(n) \text{ and } \Delta_{12}(n) \text{ are even} \\ \frac{1}{2}\Delta_{12}(n) & \text{if } \min_{12}(n)+1 \text{ and } \Delta_{12}(n) \text{ are even} \\ \frac{1}{2}\Delta_{12}(n)+\frac{1}{2} & \text{if } \Delta_{12}(n) \text{ is odd,} \end{cases}$$

which leads to the result.

For odd n, by Lemmas 58 and 61, it happens when n_{12} is even, except if $n_{03} = \min_{03}(n)$ and $\mathrm{mj}_{03}(n) = 1$ or $n_{03} = \max_{03}(n)$ and $\mathrm{MJ}_{03}(n) = 1$. The number of such cases is

$$\begin{cases} \frac{\Delta_{12}(n)}{2}+1-\mathrm{MJ}_{03}(n)-\mathrm{mj}_{03}(n) & \text{if } \min_{12}(n) \text{ and } \Delta_{12}(n) \text{ are even} \\ \frac{\Delta_{12}(n)+1}{2}-\mathrm{MJ}_{03}(n) & \text{if } \min_{12}(n) \text{ and } \Delta_{12}(n)+1 \text{ are even} \\ \frac{\Delta_{12}(n)+1}{2}-\mathrm{mj}_{03}(n) & \text{if } \min_{12}(n) \text{ and } \Delta_{12}(n) \text{ are odd} \\ \frac{\Delta_{12}(n)}{2} & \text{if } \min_{12}(n)+1 \text{ and } \Delta_{12}(n) \text{ are even.} \end{cases}$$

Indeed, consider for example the case that $\min_{12}(n)$ and $\Delta_{12}(n)$ are even. First, there are $\frac{\Delta_{12}(n)}{2}+1$ even values of n_{12}. Second, since $\min_{12}(n)$ is even and n is odd, we have $\max_{03}(n) = n - \min_{12}(n)$ odd. Since $\Delta_{12}(n)$ is even, $\max_{12}(n)$ is also even and $\min_{03}(n)$ is odd.

If n is such that $\mathrm{mj}_{03}(n) = 1$ (resp. $\mathrm{MJ}_{03}(n) = 1$) then the case $n_{03} = \min_{03}(n)$ and $\mathrm{mj}_{03}(n) = 1$ (resp. $n_{03} = \max_{03}(n)$ and $\mathrm{MJ}_{03}(n) = 1$) indeed happens. So we have to remove 1, i.e., $\mathrm{mj}_{03}(n)$ or $\mathrm{MJ}_{03}(n)$ for each case.

As another example, consider the case that $\min_{12}(n)$ and $\Delta_{12}(n)$ are odd. Then $\max_{03}(n)$ is even and $\min_{03}(n)$ is odd. There are $\frac{\Delta_{12}(n)+1}{2}$ even values of n_{12}. We cannot have $n_{03} = \max_{03}(n)$ (for parity reasons) and thus we never have $n_{03} = \max_{03}(n)$ and $\mathrm{MJ}_{03}(n) = 1$. But the case $n_{03} = \min_{03}(n)$ happens and thus we have to remove one case when $\mathrm{mj}_{03}(n) = 1$.

Finally, observe that to each pair (n, n_{12}), with n odd and n_{12} even, correspond two abelian equivalence classes of $\mathbf{y}$ (see the proof of Proposition 43). Each of these classes splits into two 2-abelian equivalence classes. Hence multiplying by 2 the number of pairs (n, n_{12}), with n odd and n_{12} even, gives the result claimed for n odd. □

Corollary 62. *The sequence $\mathcal{P}_{\mathbf{t}}^{(2)}(n)_{n\geq 0}$ is 2-regular.*

Proof. We can make use of Lemma 8. Thanks to Theorem 56, $\mathcal{P}_{\mathbf{t}}^{(2)}(n+1)$ can be expressed as a combination of $\mathcal{P}_{\mathbf{y}}^{(1)}(n)$, $\Delta_{12}(n)$, $\mathrm{MJ}_{03}(n)$, $\mathrm{mj}_{03}(n)$ using the predicates ($n \bmod 2$), ($\Delta_{12}(n) \bmod 2$) and ($\min_{12}(n) \bmod 2$).

The sequences $\mathcal{P}_{\mathbf{y}}^{(1)}(n)_{n\geq 0}$ and $\Delta_{12}(n)_{n\geq 0}$ are 2-regular from Section 6. Note that we have $\mathrm{MJ}_{03}(n+1) = \min_{12}(n) - \min_{12}(n+1) + 1$ and

$$\begin{aligned}\mathrm{mj}_{03}(n) &= \max{}_{12}(n) - \max{}_{12}(n+1) + 1\\ &= \min{}_{12}(n) - \min{}_{12}(n+1) + \Delta_{12}(n) - \Delta_{12}(n+1) + 1.\end{aligned}$$

As $\mathrm{MJ}_{03}(n+1)$ and $\mathrm{mj}_{03}(n)$ can only take the values 0 and 1, these relations can also be expressed using $(\min_{12}(n) \bmod 2)_{n\geq 0}$ and $(\Delta_{12}(n) \bmod 2)_{n\geq 0}$. Since these two latter sequences are 2-regular, the sequences $(\min_{12}(n+1) \bmod 2)_{n\geq 0}$ and $(\Delta_{12}(n+1) \bmod 2)_{n\geq 0}$ are 2-regular by Lemma 11 and so are $\mathrm{MJ}_{03}(n+1)_{n\geq 0}$ and $\mathrm{mj}_{03}(n)_{n\geq 0}$ by Lemma 8. Thus, $\mathrm{MJ}_{03}(n)_{n\geq 0}$ is 2-regular by Lemma 11.

Since all the functions (resp. all the predicates) occurring in the statement of Theorem 56 are 2-regular (resp. 2-automatic), the composition given in Lemma 8 implies that the sequence $\mathcal{P}_{\mathbf{t}}^{(2)}(n+1)_{n\geq 0}$ is 2-regular. Hence, by Lemma 11, $\mathcal{P}_{\mathbf{t}}^{(2)}(n)_{n\geq 0}$ is 2-regular. □

8 Conclusions

The two examples treated in this paper, namely the 2-abelian complexity of the period-doubling word and the Thue–Morse word, suggest that a general framework to study the ℓ-abelian complexity of k-automatic sequences may exist. As an example, we consider the 3-block coding of the period-doubling word,

$$\mathbf{z} = \mathrm{block}(\mathbf{p}, 3) = 240125252401240124\cdots.$$

The abelian complexity $\mathcal{P}_{\mathbf{z}}^{(1)}(n)_{n\geq 0} = (1, 5, 5, 8, 6, 10, 19, 11, \ldots)$ seems to satisfy, for $\ell \geq 4$, the following relations (which are quite similar to what we have discussed so far)

$$\mathcal{P}_{\mathbf{z}}^{(1)}(2^\ell + r) = \begin{cases} \mathcal{P}_{\mathbf{z}}^{(1)}(r) + 5 & \text{if } r \leq 2^{\ell-1} \text{ and } r \text{ even} \\ \mathcal{P}_{\mathbf{z}}^{(1)}(r) + 7 & \text{if } r \leq 2^{\ell-1} \text{ and } r \text{ odd} \\ \mathcal{P}_{\mathbf{z}}^{(1)}(2^{\ell+1} - r) & \text{if } r > 2^{\ell-1}. \end{cases}$$

Then, the next step would be to relate $\mathcal{P}_{\mathbf{p}}^{(3)}$ with $\mathcal{P}_{\mathbf{z}}^{(1)}$ (and try to extend the developments from Section 5).

Acknowledgments

We thank Jeffrey Shallit for some motivating discussions we had about this problem at some early stage of development.

References

[1] Jean-Paul Allouche and Jeffrey Shallit, The ring of k-regular sequences, *Theoretical Computer Science* **98** (1992), 163–197.

[2] Jean-Paul Allouche and Jeffrey Shallit, The ring of k-regular sequences II, *Theoretical Computer Science* **307** (2003), 3–29.

[3] Jean-Paul Allouche and Jeffrey Shallit, *Automatic sequences. Theory, applications, generalizations*, Cambridge University Press, Cambridge, 2003.

[4] Jean Berstel and Christophe Reutenauer, Noncommutative rational series with applications, *Encyclopedia of Mathematics and its Applications*, **137**, Cambridge University Press, Cambridge, 2011.

[5] Francine Blanchet-Sadri, James D. Currie, Narad Rampersad and Nathan Fox, Abelian complexity of fixed point of morphism $0 \mapsto 012, 1 \mapsto 02, 2 \mapsto 1$, *INTEGERS* **14** (2014) #A11.

[6] Carpi, Arturo and D'Alonzo, Valerio, On factors of synchronized sequences, *Theoretical Computer Science*, **411** (2010), 3932–3937.

[7] Émilie Charlier, Narad Rampersad and Jeffrey Shallit, Enumeration and decidable properties of automatic sequences, *Internat. J. Found. Comput. Sci.* **23** (2012), no. 5, 1035–1066.

[8] Alan Cobham, Uniform tag sequences, *Math. Systems Theory* **6** (1972), 186–192.

[9] Samuel Eilenberg, *Automata, Languages and Machines*, Vol. A, Academic Press, New York, 1974.

[10] Anna Frid, On uniform DOL words, STACS'98, *Lecture Notes in Computer Science*, **1373** (1998), 544–554.

[11] Florian Greinecker, On the 2-abelian complexity of Thue–Morse subwords, `arXiv:1404.3906`.

[12] Juhani Karhumäki, Generalized Parikh mappings and homomorphisms, *Information and Control* **47** (1980), 155–165.

[13] Juhani Karhumäki, Aleksi Saarela and Luca Q. Zamboni, On a generalization of Abelian equivalence and complexity of infinite words, *J. Combin. Theory Ser. A* **120** (2013), no. 8, 2189–2206.

[14] Juhani Karhumäki, Aleksi Saarela and Luca Q. Zamboni, Variations of the Morse–Hedlund theorem for k-abelian equivalence, DLT 2014, *Lecture Notes in Computer Science* **8633** (2014), 203–214.

[15] Blake Madill and Narad Rampersad, The abelian complexity of the paperfolding word, *Discrete Math.* **313** (2013), no. 7, 831–838.

[16] The OEIS Foundation, The On-Line Encyclopedia of Integer Sequences, `http://oeis.org`.

[17] Michel Rigo and Élise Vandomme, 2-abelian complexity of the Thue–Morse sequence, `http://hdl.handle.net/2268/135841`, 2012.