Abelian Complexity of Thue-Morse Word over a Ternary Alphabet

Idrissa Kaboré(✉) and Boucaré Kientéga

UFR–Sciences et Techniques, Université NAZI BONI,
01 BP 1091 Bobo-Dioulasso 01, Burkina Faso
ikaborei@yahoo.fr, boucarekientaga@yahoo.fr

Abstract. In this paper, we study the Thue-Morse word on a ternary alphabet. We establish some properties on special factors of this word and prove that it is 2-balanced. Moreover, we determine its Abelian complexity function.

Keywords: Infinite word · Factor · Morphism · Abelian complexity

Mathematics Subject Classification: 68R15 · 11B85

1 Introduction

Abelian complexity is a combinatorial notion used in the study of infinite words. It counts the number of Parikh vectors of given length in a word. The study of Abelian complexity was developed recently [6,7,9,16,17]. In particular, the Abelian complexity of some words and some classes of words have been studied [4,8,10,12,13,21,23].

The Thue-Morse word $\mathbf{t}_2$ on the binary alphabet $\{0,\ 1\}$ is the infinite word generated by the morphism μ_2 defined by $\mu(0) = 01$, $\mu(1) = 10$. The study of this word goes back to the beginning of the twentieth century with the works of Thue [19,20]. It was extensively studied during the last three decades [1–3,14]. In [17] the authors have determined its Abelian complexity: for all $n \geq 1$, $\rho^{ab}_{\mathbf{t}_2}(n) = 2$ if n is odd and $\rho^{ab}_{\mathbf{t}_2}(n) = 3$ otherwise. The Thue-Morse word can be naturally generalized over an alphabet $\mathcal{A}_q$ of size $q \geq 3$. More precisely, it is, on the alphabet $\mathcal{A}_q = \{0,\ 1,\ ...,\ q-1\}$, the infinite word $\mathbf{t}_q$ generated by the morphism μ_q defined by: $\mu_q(k) = k(k+1)...(k+q-1)$, where the letters are expressed modulo q. A study of this word has been done in [18]. In this paper, we are interested in the study of the Abelian complexity of the Thue-Morse word over the alphabet $\mathcal{A}_3 = \{0,\ 1,\ 2\}$. More exactly, it is the word $\mathbf{t}_3$ generated by the morphism μ_3 defined by $\mu_3(0) = 012, \mu_3(1) = 120$ and $\mu_3(2) = 201$.

The paper is organized as follows. After some definitions and notations, we recall in Sect. 2 some useful results. In Sect. 3, we establish some combinatorial properties of the word $\mathbf{t}_3$. We determine, in particular, its triprolongable factors, then we show that it is 2-balanced. Lastly, in Sect. 4, we determine the Abelian complexity function of $\mathbf{t}_3$.

S. Brlek et al. (Eds.): WORDS 2017, LNCS 10432, pp. 132–143, 2017.
DOI: 10.1007/978-3-319-66396-8_13

2 Definitions and Notations

Let $\mathcal{A}$ be a finite alphabet. The set of finite words over $\mathcal{A}$ is noted $\mathcal{A}^*$ and ε represents the empty word. The set of non-empty finite words over $\mathcal{A}$ is denoted by $\mathcal{A}^+$. For all $u \in \mathcal{A}^*$, $|u|$ designates the length of u and the number of occurrences of a letter a in u is denoted $|u|_a$. A word u of length n formed by repeating a single letter x is denoted x^n.

An infinite word is a sequence of letters of $\mathcal{A}$, indexed by $\mathbb{N}$. We denote by $\mathcal{A}^\omega$ the set of infinite words on $\mathcal{A}$. The set of finite or infinite words on $\mathcal{A}$ is denoted $\mathcal{A}^\infty$.

Let u be a finite or infinite word and v a finite word on $\mathcal{A}$. The word v is called factor of u if there exists $u_1 \in \mathcal{A}^*$ and $u_2 \in \mathcal{A}^\infty$ such that $u = u_1 v u_2$. The factor v is called prefix (resp. suffix) if u_1 (resp. u_2) is the empty word. The set of the prefixes (resp. the suffixes) of u is denoted $pref(u)$ (resp. $suff(u)$).

Let u be an infinite word. The set of factors of length n of u is denoted $F_n(u)$. The set of all the factors of u is denoted $F(u)$.

Let v be a factor of u and a be a letter of $\mathcal{A}$. We say that v is right (resp. left) prolongable by a, if va (resp. av) is also a factor of u. The word va (resp. av) is called a right (resp. left) extension of v in u. The factor v is said to be right (resp. left) special it admits at least two right (resp. left) extensions. If v is both right special and left special, it is called bispecial.

An infinite word u is said to be recurrent if any factor of u appears infinitely often. It is said to be uniformly recurrent if for any natural n, it exists a natural n_0 such that any factor of length n_0 contains all the factors length n of u.

A morphism on $\mathcal{A}^*$ is a map $f : \mathcal{A}^* \to \mathcal{A}^*$ such that $f(uv) = f(u)f(v)$, for all $u, v \in \mathcal{A}^*$. A morphism f is said to be primitive if it exists a positive integer n such that, for all letter a in $\mathcal{A}$, $f^n(a)$ contains all the letters of $\mathcal{A}$. It is k-uniform, if $|f(a)| = k$ for all a in $\mathcal{A}$. A morphism f is said to be prolongable on a letter a if $f(a) = aw$ where $w \in \mathcal{A}^+$, and $f^n(w)$ is non empty for any natural n. A morphism f defined on an alphabet $A = \{a_1, a_2, ..., a_d\}$ is said to be left (resp. right) marked, if the first (resp. last) letters of $f(a_i)$ and $f(a_j)$ are different, for all $i \neq j$. If f is both left marked and right marked, it is said marked. An infinite word u is generated by a morphism f if there exists a letter a such that the words a, $f(a)$, ..., $f^n(a)$, ... are longer and longer prefixes of u. We note $u = f^\omega(a)$. An infinite word generated by a morphism is called purely morphic word. Let u be an infinite purely morphic word and w, a factor of u verifying

$$|w| \geq max\{|f(a)| : a \in \mathcal{A}\}.$$

Then w can be decomposed in the form

$$p_0 f(a_1) f(a_2) ... f(a_n) s_{n+1},$$

where

- $n \geq 0$, $a_0,\ a_1,\ ...,\ a_{n+1} \in \mathcal{A}$;
- p_0 is a suffix of $f(a_0)$ and s_{n+1} is a prefix of $f(a_{n+1})$.

This decomposition is called *synchronization* [5].

Let u be an infinite word on an alphabet $\mathcal{A}_q = \{a_0, a_1, ..., a_{q-1}\}$ and v, a factor of u. The Parikh vector of v is the q-uplet $\psi(w) = (|v|_{a_0}, |v|_{a_1}, ..., |v|_{a_{q-1}})$. We denote by $\Psi_n(u)$, the set of the Parikh vectors of the factors of length n of u:

$$\Psi_n(u) = \{\psi(v) : v \in F_n(u)\}.$$

The Abelian complexity of u is the application of $\mathbb{N}$ to $\mathbb{N}$ defined by: $\rho^{ab}(n) = card(\Psi_n(u))$. Let θ be a natural. An infinite word u is said to be θ-balanced if for any letter a of $\mathcal{A}$ and any couple (v, w) of factors of u with the same length, one has $||v|_a - |w|_a| \leq \theta$.

Let u be an infinite word and v, a factor of u of length n. We denote by $u_{[n]}$ the prefix of u of length n. The relative Parikh vector [22] of v is:

$$\psi^{rel}(v) = \psi(v) - \psi(u_{[n]}).$$

The set of the relative Parikh vectors of the factors of u of length n will be simply denoted:

$$\Psi_n^{rel}(u) = \{\psi^{rel}(v) : v \in F_n(u)\}.$$

This set has the same cardinal as $\Psi_n(u)$. So,

$$\rho^{ab}(n) = card(\Psi_n^{rel}(u)).$$

If u is θ-balanced, then all the components of relative Parikh vector are bounded by θ [23].

Let us consider the alphabet $\mathcal{A}_3 = \{0, 1, 2\}$. The Thue-Morse word over $\mathcal{A}_3$ is the infinite word $\mathbf{t}_3$ generated by the morphism μ_3 defined by $\mu_3(0) = 012$, $\mu_3(1) = 120$, $\mu_3(2) = 201$:

$$\mathbf{t}_3 = \lim_{n \longrightarrow +\infty} \mu_3^{(n)}(0) = 012120201120201012201012120120201012201012120012...$$

Theorem 2.1 *[11]. Let f be a primitive morphism, prolongable on a letter a. Then, the infinite word $f^\omega(a)$, generated by f on a, is uniformly recurrent.*

The morphism μ_3 being primitive and prolongable on 0, the word $\mathbf{t}_3 = \mu_3^\omega(0)$ is uniformly recurrent.

In the following, we consider the alphabet $\mathcal{A}_3 = \{0, 1, 2\}$.

3 Triprolongable Factors and Balance

In this section, we establish some combinatorial properties of $\mathbf{t}_3$, then we show that it is 2-balanced.

Recall the following useful lemma called synchronization lemma applied to the morphism μ_3.

Lemma 3.1 *Let u be a factor of $\mathbf{t}_3$. Then, there exist some factors v, δ_1 and δ_2 of $\mathbf{t}_3$ such that $u = \delta_1\mu_3(v)\delta_2$ with $|\delta_1|$, $|\delta_2| \leq 2$. This decomposition is unique if $|u| \geq 7$.*

Proposition 3.1 *Let u be a factor of* $\mathbf{t}_3$*. Then, u is right (resp. left) triprolongable if and only if $\mu_3(u)$ is right (resp. left) triprolongable.*

Proof: Let u be a factor of $\mathbf{t}_3$, right triprolongable. Then, for any $i \in \mathcal{A}_3$, ui is in $\mathbf{t}_3$. Therefore, $\mu_3(u)i$ is in $\mathbf{t}_3$, since $\mu_3(i)$ begins with i.

Conversely, let u be a factor of $\mathbf{t}_3$ such that $\mu_3(u)$ is right triprolongable with $|u| \geq 2$ (the case $|u| \leq 1$ is evident). Then, $\mu_3(u)i$ is in $\mathbf{t}_3$, for all $i \in \mathcal{A}_3$. So, the factor $\mu_3(u)i$ ends by the first letter of the image of $\mu_3(i)$, for all $i \in \mathcal{A}_3$; we use here the unicity in the Lemma 3.1 since $|\mu_3(u)i| \geq 7$. So, the factors $\mu_3(u)012$, $\mu_3(u)120$ and $\mu_3(u)201$ are in $\mathbf{t}_3$. These three factors can be written respectively $\mu_3(u0)$, $\mu_3(u1)$ and $\mu_3(u2)$. This proves that u is right triprolongable in $\mathbf{t}_3$. We proceed in the same way for the factors which are left triprolongable. ■

For the following, we denote by $BST(\mathbf{t}_3)$, the set of the factors of $\mathbf{t}_3$ which are both right triprolongable left triprolongable.

As a consequence of Proposition 3.1, a factor u is in $BST(\mathbf{t}_3)$ if and only if $\mu_3(u)$ is in $BST(\mathbf{t}_3)$.

Proposition 3.2. *Let u be an element of $BST(\mathbf{t}_3)$. If $|u| \geq 3$, it exists u' in $BST(\mathbf{t}_3)$ such that $u = \mu_3(u')$.*

Proof: Let u in $BST(\mathbf{t}_3)$ such that $|u| \geq 3$. One verifies manually the proposition for the case $3 \leq |u| \leq 6$. Now suppose $|u| \geq 7$. Then, the factor u can be written in a unique way in the form $u = \delta_1\mu_3(u')\delta_2$, where u', δ_1 and δ_2 are factors of $\mathbf{t}_3$. Let us verify that factors δ_1 and δ_2 are empty. As u is right triprolongable, the factors $\delta_2 0$, $\delta_2 1$ and $\delta_2 2$ are in $\mathbf{t}_3$. So, one of the words $\delta_2 i$, contains the square of a letter. This is impossible because the image of no letter does contain a square. In the same way, we show that δ_1 is empty. Hence, $u = \mu_3(u')$. By Proposition 3.1, u' is in $BST(\mathbf{t}_3)$. ■

Theorem 3.1. *The set $BST(\mathbf{t}_3)$ is given by:*

$$BST(\mathbf{t}_3) = \bigcup_{n \geq 0} \{\mu_3^n(0),\ \mu_3^n(1),\ \mu_3^n(2),\ \mu_3^n(01),\ \mu_3^n(12),\ \mu_3^n(20)\} \cup \{\varepsilon\}.$$

Proof: Let u be an element of $BST(\mathbf{t}_3)$ with length at least 3. By Proposition 3.3, it exists u' in $BST(\mathbf{t}_3)$ such that $u = \mu_3(u')$. Hence, to obtain the set $BST(\mathbf{t}_3)$, it suffices to find its elements of length at most 2, since the others can be obtained by applying successively μ_3. These factors are $0,\ 1,\ 2,\ 01,\ 12$ and 20. ■

Corollary 3.1 *Let u be a factor of* $\mathbf{t}_3$ *which is right triprolongable. If $|u| = 3^k$ or $|u| = 2 \times 3^k$, $k \geq 0$, then u is left triprolongable.*

Proof: Let u be a factor of $\mathbf{t}_3$, right triprolongable and verifying $|u| = 3^k$, $k \geq 1$. Then, u can be decomposed in the form $\delta_1\mu_3(v)\delta_2$ where v, δ_1, $\delta_2 \in F(\mathbf{t}_3)$. The factor u being right triprolongable, δ_2 is the empty word. So, $u = \delta_1\mu_3(v)$. We know that $|\delta_1| \leq 2$ and $|\mu_3(v)|$ is multiple of 3. The factor u being of length 3^k then δ_1 is the empty word. Hence, $u = \mu_3(v)$ where v is a right triprolongable factor of length 3^{k-1}. By the same process, the factor v can be written $v = \mu_3(v')$, where v' is a right triprolongable factor of length 3^{k-2}. In a successive way, we succeed in $u = \mu_3^k(i)$, $i \in \mathcal{A}_3$. With Theorem 3.1, we conclude that u is left triprolongable. We proceed in the same way for the factors of length 2×3^k. ■

Proposition 3.3. *Let u be in $BST(\mathbf{t}_3)$. Then, it exists a unique letter i in $\mathcal{A}_3$ such that iu (resp. ui) is right (resp. left) triprolongable.*

Proof: Let us construct the set $F(\mathbf{t}_3) \cap (\mathcal{A}_3 v \mathcal{A}_3)$ where $v \in BST(\mathbf{t}_3)$. We give those for which $|v| \leq 3$, and by induction we show that for those of upper length, the extensions respect the unicity. We have, for $i \in \mathcal{A}_3$:

$F(\mathbf{t}_3) \cap (\mathcal{A}_3 i \mathcal{A}_3) = \{0i1,\ 1i1,\ 2i0,\ 2i1,\ 2i2\}$;
$F(\mathbf{t}_3) \cap (\mathcal{A}_3 i(i+1) \mathcal{A}_3) = \{0i(i+1)2,\ 1i(i+1)2,\ 2i(i+1)0,\ 2i(i+1)1,\ 2i(i+1)2\}$;
$F(\mathbf{t}_3) \cap (\mathcal{A}_3 \mu_3(i) \mathcal{A}_3) = \{0\mu_3(i)1,\ 1\mu_3(i)0,\ 1\mu_3(i)1,\ 1\mu_3(i)2,\ 2\mu_3(i)0,\ 2\mu_3(i)1\}$;

where $i + 1$ is taken modulo 3.

Let us take a factor $v = \mu_3^n(0)$ and suppose that the set $F(\mathbf{t}_3) \cap \mathcal{A}_3 \mu_3^n(0)$ contains a single right triprolongable factor. Even if it means changing letter, let us take $0\mu_3^n(0)$ this factor. So, By Proposition 3.1, $2\mu_3^{n+1}(0)$ is a right triprolongable factor of $\mathbf{t}_3$. Let us verify that it is the only one. Suppose $0\mu_3^{n+1}(0)$ is right triprolongable. Then, $1\mu_3^n(0)$ is right triprolongable. This contradicts the recursion hypothesis. We proceed in the same way for the other factors. ■

Proposition 3.4. *Let u be a factor of $\mathbf{t}_3$, right triprolongable. If u is left special, then it is left triprolongable.*

Proof: Let u be a factor of $\mathbf{t}_3$, right triprolongable and left special. Then, u can be written in the form $\delta_1 \mu_3(v_1) \delta_2$. As the factor u is right triprolongable, δ_2 is empty by Proposition 3.2. Furthermore, as u is left special, δ_1 is empty; because otherwise, δ_1 would be proper suffix of the image of some letter and by this fact u would be extended on left in a unique way. So, u can be synchronized in the form $u = \mu_3(v_1)$, where v_1 is a factor of $\mathbf{t}_3$. Since the morphism μ_3 is marked, then v_1 is left special. Moreover, it is right triprolongable by Proposition 3.1. Thus, v_1 can be synchronized in the form $v_1 = \mu_3(v_2)$, $v_2 \in F(\mathbf{t}_3)$. In a successive way, we succeed in $u = \mu_3^k(v_k)$ with $k \geq 0$ and v_k a right triprolongable factor, left special and of length at most 2. Therefore, v_k is left triprolongable by Theorem 3.1. ■

Proposition 3.5. *For all positive natural n, $\mathbf{t}_3$ admits exactly three right (resp. left) triprolongable factors of length n.*

Proof: It is known that 0, 1 and 2 are the right triprolongable factors of length 1. Let us show that any right triprolongable factor of length n is suffix of a unique right triprolongable factor of length $n + 1$.

Let w be a right triprolongable factor of length n. If w admits a unique extension a on left, then aw is a right triprolongable factor since $\mathbf{t}_3$ is recurrent. If it admits at least two left extensions, then w is in $BST(\mathbf{t})$ by Proposition 3.4 and only one of its left extensions is right triprolongable by Proposition 3.3.

Thus, the number of right triprolongable factors of length $n+1$ is equal to the number of right triprolongable factors of length n in $\mathbf{t}_3$. In the same way, we treat the case of the left triprolongable factors. ■

The following remark is a consequence of Proposition 3.5.

Remark 3.1. *Let u be a right (resp. left) triprolongable factor of $\mathbf{t}_3$. If the length of u is $3k$, $k \geq 1$, then it exists a right (resp. left) triprolongable factor v of $\mathbf{t}_3$ such that $u = \mu_3(v)$.*

Proposition 3.6. *For all positive natural n, the right (resp. left) triprolongable factors of length n begin (resp. end) with different letters.*

Proof: We proceed by induction on n. Suppose all the right triprolongable factors of $\mathbf{t}_3$ of length at most n begin with different letters. Let u_1 and u_2 be two factors of $\mathbf{t}_3$, right triprolongable of length n. We distinguish the two following cases.

Case 1: n is multiple of 3. Then, it exists some factors v_1 and v_2 of $\mathbf{t}_3$ such that $u_1 = \mu_3(v_1)$ and $u_2 = \mu_3(v_2)$. Suppose there exists a letter a of $\mathcal{A}_3$ such that au_1 and au_2 are right triprolongable. Even if it means changing letter, let us take $a = 0$. Thus, the factors $120u_1$ and $120u_2$ are right triprolongable in $\mathbf{t}_3$. These factors can be written respectively $\mu_3(1v_1)$ and $\mu_3(1v_2)$. By Proposition 3.1, $1v_1$ and $1v_2$ are right triprolongable factors. This fact contradicts the hypothesis of induction since $1v_1$ and $1v_2$ are of length lower than n.

Case 2: $n-1$ is multiple of 3. Then, there exist some factors v_1 and v_2 of $\mathbf{t}_3$, right triprolongable such that $u_1 = i\mu_3(v_1)$ and $u_2 = j\mu_3(v_2)$. As i and j are suffix of images of letters, they have each a unique left extension. Since they are different by hypothesis, the extensions are different.

Case 3: $n-2$ is multiple of 3. We proceed similarly like previous cases. ■

Theorem 3.2. *The word $\mathbf{t}_3$ is 2-balanced.*

Proof: Let u_1 and u_2 be two factors of length $n \geq 7$ of $\mathbf{t}_3$. Then, u_1 and u_2 can be synchronized in a unique way in the forms $u_1 = \delta_1\mu_3(v_1)\delta_2$ and $u_2 = \delta_1^{'}\mu_3(v_2)\delta_2^{'}$, $v_1,\ v_2,\ \delta_1,\ \delta_2,\ \delta_1',\ \delta_2' \in F(\mathbf{t}_3)$. Let us put $\alpha_i = |\delta_1|_i + |\delta_2|_i$, $\beta_i = |\delta_1'|_i + |\delta_2'|_i$, for all $i \in \mathcal{A}_3$. Consider the following cases.

Case 1: n is multiple of 3. Then, u_1 (resp. u_2) can be written uniquely in the form $\mu_3(v)$, $ij\mu_3(v)k$ or $i\mu_3(v)jk$, with $i,\ j,\ k \in \mathcal{A}_3$ and $v \in F(\mathbf{t}_3)$. Consider the different forms taken by u_1 and u_2.

- Suppose $u_1 = \mu_3(v_1)$ and $u_2 = \mu_3(v_2)$. Then, we have $\psi(u_1) = \psi(u_2)$ and we have:

$$||u_1|_i - |u_2|_i| = 0,$$

  for any letter i.

- Suppose $u_1 = \mu_3(v_1)$ and $u_2 = i\mu_3(v_2)jk$. Write u_1 in the form $\mu_3(v_1')\mu_3(a)$, $a \in \mathcal{A}_3$. Thus, we have $|v_1'| = |v_2|$. As the letters have the same number of occurrences in image of each letter, we have $|\mu_3(a)|_i = 1$, for all $i \in \mathcal{A}_3$. Moreover, $\beta_i \leq 2$, for all $i \in \mathcal{A}_3$ since jk is the prefix of the image of some letter. Thus,
$$||u_1|_i - |u_2|_i| = ||\mu_3(a)|_i - \beta_i| \leq 1.$$
- Suppose $u_1 = ij\mu_3(v_1)k$ and $u_2 = l\mu_3(v_2)mn$, $i,\ j,\ k,\ l,\ m,\ n \in \mathcal{A}_3$. As previously, one verifies that $\alpha_i, \beta_i \leq 2$, for all $i \in \mathcal{A}_3$. Thus,
$$||u_1|_i - |u_2|_i| = |\alpha_i - \beta_i| \leq 2.$$
By taking $u_1 = 101212$ and $u_2 = 010120$, we observe that the bound 2 is reached.

Case 2: $n-1$ is multiple of 3. Then, u_1 (resp. u_2) is of the form $i\mu_3(v)$, $\mu_3(v)k$ or $ij\mu_3(v)kl$, $i,\ j,\ k,\ l \in \mathcal{A}_3$, $v \in F(\mathbf{t}_3)$.

- Suppose $u_1 = i\mu_3(v_1)$ and $u_2 = \mu_3(v_2)j$. Then, we have $|v_1| = |v_2|$. So $|\alpha_i - \beta_i| \leq 1$, for all $i \in \mathcal{A}_3$.

- Suppose $u_1 = ij\mu_3(v_1)kl$ and $u_2 = i'j'\mu_3(v_2)k'l'$, where ij and $i'j'$ (resp. kl and $k'l'$) are suffix (resp. prefix) of images of letters. Then, note that $(i,\ k) \neq (j,\ l)$ and $(i',\ k') \neq (j',\ l')$. Thus, $|v_1| = |v_2|$. By analogy with the previous case, one verifies that $\alpha_i,\ \beta_i \leq 2$. Therefore,
$$||u_1|_i - |u_2|_i| = |\alpha_i - \beta_i| \leq 2,$$
for all $i \in \mathcal{A}_3$. By taking $u_1 = 01\mu_3(12)01$ and $u_2 = 20\mu_3(01)20$ we observe that the bound 2 is reached.

- Suppose $u_1 = i'\mu_3(v_1)$ and $u_2 = ij\mu_3(v_2)kl$. Then, we write u_1 in the form $i'\mu_3(v_1')\mu_3(a)$, $a \in \mathcal{A}_3$ and $v' \in F(\mathbf{t}_3)$. Thus, $|v_1'| = |v_2|$ and $\alpha_i,\ \beta_i \leq 2$. So
$$||u_1|_i - |u_2|_i| = |\alpha_i - \beta_i| \leq 2.$$

Case 3: $n-2$ is multiple of 3. Suppose u_1 (resp. u_2) can be written in the form $ij\mu_3(v_1)$, $i\mu_3(v_1)k$ or $\mu_3(v_1)kl$ (resp. $ij\mu_3(v_2)$, $i\mu_3(v_2)k$ or $\mu_3(v_2)kl$). Then, we have $|v_1| = |v_2|$. In a similar way as previous cases, one verifies that $|\alpha_i - \beta_i| \leq 2$, for $i \in \mathcal{A}_3$. ■

4 Abelian Complexity

In this section we give an explicit formula of the Abelian complexity function ρ^{ab} of $\mathbf{t}_3$. We show that the sequence $(\rho^{ab}(n))_{n\geq 2}$ of the word $\mathbf{t}_3$ is 3-periodic.

Proposition 4.1. *For all $k \geq 1$, $\rho^{ab}(3k) = 7$.*

Proof: Let u be a factor of $\mathbf{t}_3$ of length $3k$, $k \geq 1$. Then, u synchronizes in the form $\mu_3(v)$, $i\mu_3(v)jk$ or $ij\mu_3(v)k$ with $i,\, j,\, k \in \mathcal{A}_3$, $ij,\, jk \in \{01,\, 12,\, 20\}$ and $v \in F(\mathbf{t}_3)$. As u is chosen arbitrary one verifies that these three forms are taken by u. As the prefix $\mathbf{t}_{3[3k]}$ begins with the image of some letter, it is in the form $\mu_3(v)$. For the sequel, we note $\mathbf{t}_{3[3k]} = \mu_3(v_1)$. We have three cases to discuss.

Case 1: The factor u is in the form $\mu_3(v_2)$. Then, $|v_1| = |v_2|$ and so $\psi^{rel}(u) = (0,\, 0,\, 0)$.

Case 2: The factor u is in the form $i\mu_3(v_2)jk$. Then, we have:

$$\psi(u) = (|v_2| + |ijk|_0,\; |v_2| + |ijk|_1,\; |v_2| + |ijk|_2).$$

Let us show that the set of the values taken by ijk is

$$\{001,\, 012,\, 020,\, 101,\, 112,\, 120,\, 201,\, 212,\, 220\}.$$

By Proposition 3.5, for any integer $k \geq 1$, $\mathbf{t}_3$ possesses exactly 3 right triprolongeable factors of length $3k$. Let us denote by R_1, R_2 and R_3 the right triprolongeable factors of length $3k-3$. As these factors begin with different letters, we can suppose, even if it means changing the indexes, that $0R_1$, $1R_2$ and $2R_3$ are the right triprolongeable factors of $\mathbf{t}_3$ of length $3k-2$. Therefore, the words $0R_101$, $0R_112$, $0R_120$, $1R_201$, $1R_212$, $1R_220$, $2R_301$, $2R_312$ and $2R_320$ are factors of $\mathbf{t}_3$ of length $3k$. Hence, ijk browses the announced set. So, $\psi(ijk)$ takes all the values of the following set

$$\{(2,\, 1,\, 0),\; (1,\, 1,\, 1),\; (2,\, 0,\, 1),\; (1,\, 2,\, 0),\; (0,\, 2,\, 1),\; (0,\, 1,\, 2),\; (1,\, 0,\, 2)\}.$$

Write the prefix $\mathbf{t}_{3[n]}$ in the form $\mu_3(v_1')\mu_3(l)$, $l \in \mathcal{A}_3$. Then, $|v_1'| = |v_2|$ and $\psi(\mu_3(l)) = (1,\, 1,\, 1)$. Thus, for all the factors u of length $3k$, $\psi^{rel}(u) = \psi(ijk) - \psi(\mu_3(l))$ takes all the values of the set

$$\{(1,\, 0,\, -1), (0,\, 0,\, 0),\; (1,\, -1,\, 0),\; (0,\, 1,\, -1),\; (-1,\, 1,\, 0),\; (-1,\, 0,\, 1),\; (0,\, -1,\, 1)\}.$$

Case 3: The factor u is in the form $ij\mu_3(v_2)k$. Then

$$\psi(u) = (|v_2| + |ijk|_0, |v_2| + |ijk|_1, |v_2| + |ijk|_2).$$

By proceeding in a similar way as in the case 2 and by using the left triprolongable factors, we verify that the set of values taken by ijk is

$$\{010,\, 011,\, 012,\, 120,\, 121,\, 122,\, 200,\, 201,\, 202\}.$$

Consequently, for all the factors u satisfying these conditions, $\psi^{rel}(u)$ takes all the values of the set

$$\{(1,\, 0,\, -1),\; (0,\, 0,\, 0),\; (1,\, -1,\, 0),\; (0,\, 1,\, -1), (-1,\, 1,\, 0),\; (-1,\, 0,\, 1),\; (0,\, -1,\, 1)\}.$$

After all, we have:

$$\Psi_n^{rel}(\mathbf{t}_3) = \{(1,0,-1),(0,0,0),(1,-1,0),(0,1,-1),(-1,1,0),(-1,0,1),(0,-1,1)\}.$$

■

Proposition 4.2. *For all $k \geq 1$, $\rho^{ab}(3k+1) = 6$.*

Proof: Let u be a factor of $\mathbf{t}_3$ of length $3k+1$, $k \geq 1$. Then, u synchronizes in the form $i\mu_3(v)$, $\mu_3(v)j$ or $ij\mu_3(v)kl$, $i, j, k, l \in \mathcal{A}_3$, $ij, kl \in \{01, 12, 20\}$ and $v \in F(\mathbf{t}_3)$. The prefix $\mathbf{t}_{3[3k+1]}$ is in the form $\mu_3(v_1)i$, $i \in \mathcal{A}_3$. We have:

Case 1: $i = 0$. Then, $\mathbf{t}_{3[3k+1]} = \mu_3(v_1)0$. Let us determine $\Psi^{rel}_{3k+1}(\mathbf{t}_3)$.

- Let v_2 be a factor of $\mathbf{t}_3$ such that $u = i\mu_3(v_2)$. By using the left triprolongable factors of length $3k$, we verify that the values taken by i are 0, 1 and 2. Consequently, $\psi^{rel}(u)$ takes all the values of $\{(0,0,0), (-1,1,0), (-1,0,1)\}$. In the same way, we verify that if $u = \mu_3(v_2)j$, $\psi^{rel}(u)$ browses all the elements of the set $\{(0,0,0), (-1,1,0), (-1,0,1)\}$.

- Let u be a factor of $\mathbf{t}_3$ with the form $u = ij\mu_3(v_2)kl$. We write $\mathbf{t}_{3[3k+1]}$ in the form $\mu_3(v_1')\mu_3(m)0$, $m \in \mathcal{A}_3$. It is known that each factor of the form $ij\mu_3(v_2)kl$ is the left extension of a factor of the form $j\mu_3(v_2)kl$ whose the set of values taken by jkl is

  $$\{001, 012, 020, 101, 112, 120, 201, 212, 220\}.$$

  Thus, those taken by $ijkl$ is

  $$\{2001, 2012, 2020, 0101, 0112, 0120, 1201, 1212, 1220\}.$$

  So, $\psi^{rel}(u)$ browses all the elements of the set

  $$\{(0,-1,1), (-1,1,0), (0,0,0), (-2,1,1), (-1,0,1), (0,1,-1)\}.$$

  Finally, we get

  $$\Psi^{rel}_{3k+1}(\mathbf{t}_3) = \{(0,-1,1), (-1,1,0), (0,0,0), (-2,1,1), (-1,0,1), (0,1,-1)\}.$$

Case 2: $i = 1$. Then, $\mathbf{t}_{3[3k+1]} = \mu_3(v_1)1$. Consider the different forms of u.

- Let u be a factor of $\mathbf{t}_3$ of the form $u = i\mu_3(v_2)$. As in the case 1, we verify that i takes the values 0, 1 and 2. Therefore, $\psi^{rel}(u)$ browses all the elements of $\{(0,0,0), (1,-1,0), (0,-1,1)\}$.

- Let u be a factor of $\mathbf{t}_3$ of the form $u = ij\mu_3(v_2)kl$. Then, we write $\mathbf{t}_{3[3k+1]}$ in the form $\mathbf{t}_{3[3k+1]} = \mu_3(v_1')\mu_3(m)1$, $m \in \mathcal{A}_3$. By proceeding in a similar way as in the case 1, we verify that the set of values taken by $ijkl$ is

  $$\{2001, 2012, 2020, 0101, 0112, 0120, 1201, 1212, 1220\}.$$

  Thus, $\psi^{rel}(u)$ browses all of the elements of the set

  $$\{(1,0,-1), (1,-1,0), (0,0,0), (-1,0,1), (0,-1,1), (1,-2,1)\}.$$

After all, we have:

$$\Psi^{rel}_{3k+1}(\mathbf{t}_3) = \{(1, 0, -1),\ (1, -1, 0),\ (0, 0, 0),\ (-1, 0, 1),\ (0, -1, 1),\ (1, -2, 1)\}.$$

Case 3: $i = 2$. Then, $\mathbf{t}_{3[3k+1]} = \mu_3(v_1)2$. By proceeding in a similar way as in the previous case we get:

$$\Psi^{rel}_{3k+1}(\mathbf{t}_3) = \{(1, 1, -2), (0, 1, -1), (1, 0, -1), (-1, 1, 0), (0, 0, 0), (1, -1, 0)\}. \blacksquare$$

Proposition 4.3. *For all $k \geq 1$, $\rho^{ab}(3k+2) = 6$.*

Proof: Let u be a factor of length $3k+2$ of $\mathbf{t}_3$, $k \geq 1$. Then, u can be written in the form $i\mu_3(v_2)j$, $ij\mu_3(v_2)$ or $\mu_3(v_2)kl$, $i, j, k, l \in \mathcal{A}_3$, $v_2 \in F(\mathbf{t}_3)$. Otherwise, the prefix $\mathbf{t}_{3[3k+2]}$ is in the form $\mu_3(v_1)ij$.

Case 1: $ij = 01$. Then, $\mathbf{t}_{3[3k+1]} = \mu_3(v_1)01$. Let us determine the set $\Psi^{rel}_{3k+2}(\mathbf{t}_3)$.

- Let u be a factor of $\mathbf{t}_3$ of the form $i\mu_3(v_2)j$. Then, v_1 and v_2 have the same length. So, $\psi^{rel}(u) = \psi(ij) - \psi(01)$. With right triprolongable factors of length $k-1$, we verify that the set of values taken by ij is $\{00, 01, 02, 10, 11, 12, 20, 21, 22\}$. So, $\psi^{rel}(u)$ takes all the values of the set

$$\{(1, -1, 0),\ (0, 0, 0),\ (0, -1, 1),\ (-1, 1, 0),\ (-1, 0, 1),\ (-1, -1, 2)\}.$$

- Let u be a factor of $\mathbf{t}_3$ of the form $ij\mu_3(v_2)$. Then, v_1 and v_2 have the same length. So, $\psi^{rel}(u) = \psi(ij) - \psi(01)$. The factor ij is the suffix of the image of a letter. It takes the values 01, 12 and 20. Thus, $\psi^{rel}(u)$ takes all the values of the set

$$\{(0, 0, 0),\ (0, -1, 1),\ (-1, 0, 1)\}.$$

In a same way, we verify that if u has the form $\mu_3(v_2)kl$, $\psi^{rel}(u)$ takes all the values of the set $\{(0, 0, 0),\ (0, -1, 1),\ (-1, 0, 1)\}$.
Finally, we get:

$$\Psi^{rel}_{3k+2}(\mathbf{t}_3) = \{(1, -1, 0), (0, 0, 0), (0, -1, 1), (-1, 1, 0), (-1, 0, 1), (-1, -1, 2)\}$$

Case 2: $ij = 12$. Then, $\mathbf{t}_{3[3k+2]} = \mu_3(v_1)12$. Let us determine the set $\Psi^{rel}_{3k+2}(\mathbf{t}_3)$.

- Let u be a factor of $\mathbf{t}_3$ of the form $u = i\mu_3(v_2)j$. As in the previous case the set of values taken by ij is $\{00, 01, 02, 10, 11, 12, 20, 21, 22\}$. Therefore, $\psi^{rel}(u)$ takes all the values of the set

$$\{(2, -1, -1),\ (1, 0, -1),\ (0, 1, -1),\ (0, 0, 0),\ (0, -1, 1),\ (1, -1, 0)\}.$$

- Let u be a factor of $\mathbf{t}_3$ of the form $u = ij\mu_3(v_2)$. Then, we show as in the case 1 that $\psi^{rel}(u)$ takes all the values of the set $\{(1, 0, -1),\ (1, -1, 0),\ (0, 0, 0)\}$.
After all, we have

$$\Psi^{rel}_{3k+2}(\mathbf{t}_3) = \{(2, -1, -1), (1, 0, -1), (0, 1, -1), (0, 0, 0), (0, -1, 1), (1, -1, 0)\}.$$

Case 3: $ij = 20$. Then, $\mathbf{t}_{3[3k+2]} = \mu_3(v_1)20$. As in the previous cases we verify that:

$$\Psi^{rel}_{3k+2}(\mathbf{t}_3) = \{(1, 0, -1), (0, 1, -1), (0, 0, 0), (-1, 2, -1), (-1, 1, 0), (-1, 0, 1)\} \blacksquare$$

Theorem 4.1. *The Abelian complexity function of* $\mathbf{t}_3$ *is given by:*

$$\rho^{ab}_{\mathbf{t}_3}(n) = \begin{cases} 1 & si\ \ n = 0 \\ 3 & si\ \ n = 1 \\ 7 & si\ \ n = 3k,\ \ k \geq 1 \\ 6 & sinon \end{cases}$$

Proof: The result follows from Propositions 4.1, 4.2 and 4.3. ■

The ternary Thue-Morse word $\mathbf{t}_3$ is 3-automatic. Its Abelian complexity is the eventually periodic word $(\rho^{ab}(n))_{n\geq 0} = 136(766)^{\omega}$. Thus, we note that the word $\mathbf{t}_3$ responds to the conjecture of Parreau, Rigo, Rowland and Vandomme: *Any k-automatic word admits a l-Abelian complexity function which is k-automatic.* The reader can find this conjecture and more information on the concepts in [15].

References

1. Allouche, J.P., Arnold, A., Berstel, J., Brlek, S., Jockush, W., Plouffe, S., Sagan, B.E.: A relative of the thue-morse sequence. Disc. Math. **139**, 455–461 (1995)
2. Allouche, J.P., Peyrière, J., Wen, Z.-X., Wen, Z.-Y.: Hankel determinants of the thue-morse sequense. Ann. Inst. Fourier **48**(1), 1–27 (1998)
3. Allouche, J.P., Shallit, J.: The Ubiquitous Prouhet-Thue-Morse Sequence. In: Ding, C., Helleseth, T., Niederreiter, H. (eds.) Sequences and their Applications. Discrete Mathematics and Theoretical Computer Science, pp. 1–16. Springer, London (1999)
4. Balkovà, L., Břinda, K., Turek, O.: Abelian complexity of infinite words associated with non-simple parry number. Theor. Comput. Sci. **412**, 6252–6260 (2011)
5. Cassaigne, J.: Facteurs spéciaux. Bull. Belg. Math. **4**, 67–88 (1997)
6. Cassaigne, J., Kaboré, I.: Abelian complexity and frequencies of letters in infinite words. Int. J. Found. Comput. Sci. **27**(05), 631–649 (2016)
7. Cassaigne, J., Richomme, G., Saari, K., Zamboni, L.Q.: Avoiding abelian powers in binary wordswith bounded abelian complexity. Int. J. Found. Comput. Sci. **22**(2011), 905–920 (2011)
8. Chen, J., Lü, X., Wu, W.: On the k-abelian complexity of the cantor sequence. arXiv: 1703.04063 (2017)
9. Coven, E.M., Hedlund, G.A.: Sequences with minimal block growth. Math. Syst. Theo. **7**, 138–153 (1973)
10. Curie, J., Rampersad, N.: Recurrent words with constant abelian complexity. Adv. Appl. Math. **47**, 116–124 (2011)
11. Gottschalk, W.H.: Substitution on minimal sets. Trans. Amer. Math. Soc. **109**, 467–491 (1963)
12. Lü, X., Chen, J., Wen, Z., Wu, W.: On the abelian complexity of the rudin-shapiro sequence. J. Math. Anal. Appl. **451**, 822–838 (2017)

13. Madill, B., Rampersad, N.: The abelian complexity of the paperfolding word. Disc. Math. **313**, 831–838 (2013)
14. Massé, A.B., Brlek, S., Garon, A., Labbé, S.: Combinatorial properties of f-palindromes in the thue-morse sequence. Pure Math. Appl. **19**, 39–52 (2008)
15. Parreau, A., Rigo, M., Rowland, E., Vandomme, E.: A new approach to the $2-$regularity of the $l-$abelian complexity of $2-$automatic sequences. Electron. J. Combin. **22**, 1–27 (2015)
16. Richomme, G., Saari, K., Zamboni, L.Q.: Balance and abelian complexity of the tribonacci word. Adv. Appl. Math. **45**, 212–231 (2010)
17. Richomme, G., Saari, K., Zamboni, L.Q.: Abelian complexity in minimal subshifts. J. London Math. Soc. **83**, 79–95 (2011)
18. Štarosta, S.: Generalised thue-morse word and palindromic richness. Kybernetika **48**(3), 361–370 (2012)
19. Thue, A.: Über unendliche zeichenreihen. Norske Vid. Selsk. Skr. Mat. Nat. Kl. **7**, 1–22 (1906)
20. Thue, A.: Über die gegenseilige lage gleicher teile gewisser zeichenreihen. Norske Vid. Selsk. Skr. Mat. Nat. Kl. **1**, 139–158 (1912)
21. Turek, O.: Balance and abelian complexity of a certain class of infinite ternary words. RAIRO Theor. Inf. Appl. **44**, 313–337 (2010)
22. Turek, O.: Abelian complexity and abelian co-decomposition. Theor. Comput. Sci. **469**, 77–91 (2013)
23. Turek, O.: Abelian complexity of the tribonacci word. J. Integer Seq. **18**, 212–231 (2015)