Abelian complexity function of the Tribonacci word

Ondřej Turek

Nuclear Physics Institute
Academy of Sciences of the Czech Republic
250 68 Řez, Czech Republic

and

Bogolyubov Laboratory of Theoretical Physics
Joint Institute for Nuclear Research
141980 Dubna, Russia

email: turek@theor.jinr.ru

May 22, 2014

Abstract

According to a result of Richomme, Saari and Zamboni, the abelian complexity of the Tribonacci word satisfies $\rho^{ab}(n) \in \{3, 4, 5, 6, 7\}$ for each $n \in \mathbb{N}$. In this paper we derive a formula for evaluating the function $\rho^{ab}(n)$ explicitly. The formula uses the representation $n = \sum_{j=0}^{k} \delta_j T_j$, where $(T_j)_{j=0}^{\infty} = (1, 2, 4, 7, 13, \ldots)$ are Tribonacci numbers and $\delta_j \in \{0, 1\}$ for every $j = 0, 1, \ldots, k$. We demonstrate that $\rho^{ab}(n) = (A \circ \tau_0 \circ \tau_1 \circ \cdots \circ \tau_k)(0)$, where $\tau_0$ and $\tau_1$ are certain endomorphisms of $\{0, 1, \ldots, 67\}$ and $A$ is a map $\{1, \ldots, 67\} \rightarrow \{3, 4, 5, 6, 7\}$. The evaluation of $\rho^{ab}(n)$ needs $O(\log n)$ operations, and thus it is fast even for large values of $n$. In addition, the result implies a characterization of those $n$ for which $\rho^{ab}(n) = m$ in terms of the existence of walks in a certain graph, which de facto solves an open problem proposed by Richomme et al.

1 Introduction

Abelian complexity of a word $u$ is a function $\mathbb{N} \rightarrow \mathbb{N}$ that counts the number of pairwise non-abelian-equivalent factors of $u$ of length $n$. The notion was introduced by Richomme, Saari and Zamboni in 2009 [1], and since then it has been extensively studied [2, 3, 4, 5, 6, 7, 8, 9]. One of the first papers on the subject [2] was devoted to the Tribonacci word $t$, which is the fixed point of the substitution $0 \mapsto 01, 1 \mapsto 02, 2 \mapsto 0$. The authors of [2], Richomme, Saari and
Zamboni, showed that \( \rho_{ab}^t(n) \in \{3, 4, 5, 6, 7\} \) for all \( n \). They also characterized numbers \( n \) for which \( \rho_{ab}^t(n) = 3 \), and proposed the following open problem: For each \( m \in \{4, 5, 6, 7\} \), characterize those \( n \) for which \( \rho_{ab}^t(n) = m \).

Explicit characterization of \( \rho_{ab}^u(n) \) of a given infinite word \( u \) is generally a difficult task, particularly in case of words defined over alphabets consisting of more than two letters. It is known that any Sturmian (and thus binary) word \( u \) satisfies \( \rho_{ab}^u(n) = 2 \) for all \( n \in \mathbb{N} \) [10]. Concerning non-Sturmian binary words, formulas for \( \rho_{ab}^u(n) \) have been derived for the Thue–Morse word [1] and for quadratic Parry words [6]. Also, recently the abelian complexity of the paperfolding word \( f \) has been explored [9]; it has been shown that \( \rho_{ab}^f \) is 2-
regular and can be evaluated using a finite set of recurrent relations.

Much less results of this type are known for words over \( m \)-letter alphabes for \( m \geq 3 \). A recurrent word over a ternary alphabet with constant abelian complexity equal to 3 for all \( n \in \mathbb{N} \) has been already constructed [1], but it seems there is no other example to date of a recurrent \( m \)-ary word with \( m \geq 3 \) whose abelian complexity function has been precisely determined. In particular, the problem of precise characterization of the abelian complexity \( \rho_{ab}^t(n) \) of the Tribonacci word \( t \), which is a ternary word, has remained open since 2009.

In this paper we propose a formula for evaluating the function \( \rho_{ab}^t(n) \), together with a method that can be applied to certain other words as well. Our approach relies on the technique of abelian co-decomposition. The technique has been introduced in [8], where it served as a tool for proving that the abelian complexity of the Tribonacci word attains each value in \( \{4, 5, 6\} \) infinitely often.

## 2 Preliminaries

Let us consider a set \( A = \{0, 1, 2, \ldots, m-1\} \) (alphabet) consisting of \( m \) symbols (letters) \( 0, 1, \ldots, m-1 \). Concatenations of letters from \( A \) are called words. Let \( A^* \) denote the free monoid of all finite words over \( A \) including the empty word \( \varepsilon \).

The length of a \( w = w_0w_1w_2 \cdots w_{n-1} \in A^* \) is the number of its letters, \( |w| = n \); the length of the empty word is defined to be 0. The symbol \( |w|_\ell \) for \( \ell \in A \) and \( w \in A^* \) denotes the number of occurrences of the letter \( \ell \) in the word \( w \).

Infinite sequences of letters are called infinite words. The set of all infinite words over \( A \) is denoted by \( A^\infty \). A finite word \( w \) is a factor of a (finite or infinite) word \( u \) if there exists a finite word \( x \) and a (finite or infinite, respectively) word \( y \) such that \( u = xwy \). The word \( w \) is called a prefix of \( u \) if \( x = \varepsilon \), and a suffix of \( u \) if \( y = \varepsilon \).

An infinite word \( u \) is called recurrent if every factor of \( u \) occurs infinitely many times in \( u \).

The Parikh vector of \( w \) is the \( m \)-tuple \( \Psi(w) = (|w|_0, |w|_1, \ldots, |w|_{m-1}) \); note that \( |w|_0 + |w|_1 + \cdots + |w|_{m-1} = |w| \). For any given infinite word \( u \), let \( \mathcal{P}_u(n) \) denote the set of all Parikh vectors corresponding to factors of \( u \) having the length \( n \), i.e.,

\[
\mathcal{P}_u(n) = \{ \Psi(w) \mid w \text{ is a factor of } u, |w| = n \}.
\]
The abelian complexity of a word $u$ is the function $\rho_{ab}^u : \mathbb{N} \to \mathbb{N}$ defined as

$$\rho_{ab}^u(n) = \# P_u(n),$$

where $\#$ denotes the cardinality.

In [8] we introduced the relative Parikh vector, defined for any factor $w$ of $u$ of length $n$ as

$$\Psi_{rel}^u(w) = \Psi(w) - \Psi(u[n]),$$

where $u[n]$ is the prefix of $u$ of length $n$. Since the subtrahend $\Psi(u[n])$ is independent of $w$, the set of relative Parikh vectors corresponding to the length $n$,

$$P_{rel}^u(n) := \{ \Psi_{rel}^u(w) \mid w \text{ is a factor of } u, |w| = n \},$$

has the same cardinality as $P_u(n)$. Hence we obtain, with regard to (1),

$$\rho_{ab}^u(n) = \# P_{rel}^u(n).$$

An infinite word $u$ is said to be $c$-balanced if for every $\ell \in A$ and for every pair of factors $v$, $w$ of $u$ such that $|v| = |w|$, it holds $\| |v| - |w| \| \leq c$. If $u$ is a $c$-balanced word, the components of relative Parikh vectors are bounded by $c$ [8]. Therefore, the set of all relative Parikh vectors $\bigcup_{n \in \mathbb{N}} P_{rel}^u(n)$ is finite for any $c$-balanced word $u$.

A simple Parry word is defined over the alphabet $A = \{0, 1, \ldots, m-1\}$ as the fixed point of a substitution

$$\varphi_u : 0 \mapsto 0^{\alpha_0}1$$
$$1 \mapsto 0^{\alpha_1}2$$
$$\vdots$$
$$m-2 \mapsto 0^{\alpha_{m-2}}(m-1)$$
$$m-1 \mapsto 0^{\alpha_{m-1}}$$

with $\alpha_i$ satisfying conditions $\alpha_0 \geq 1$ and $\alpha_\ell \leq \alpha_0$ for all $\ell \in A$; see [11, 12].

Let $U_j = |\varphi_u^j(0)|$ for every $j \in \mathbb{N}_0$. Any $n \in \mathbb{N}$ can be represented as a sum

$$n = \sum_{j=0}^{k} d_j U_j$$

with integer coefficients $d_j$. If coefficients $d_j$ are obtained by the greedy algorithm, the sequence $d_k d_{k-1} \cdots d_1 d_0$ is called normal $U$-representation of $n$ [13] and denoted

$$\langle n \rangle_U = d_k d_{k-1} \cdots d_1 d_0.$$\hspace{1cm} (5)

In can be shown that coefficients in (5) satisfy $d_j \in \{0, 1, \ldots, \alpha_0\}$ for all $j = 0, 1, \ldots, k$.

It is well known (cf. [14, 12]) that a prefix of a simple Parry word $u$ of length $n$ takes the form

$$u[n] = (\varphi_u^k(0))^{d_k} (\varphi_u^{k-1}(0))^{d_{k-1}} \cdots (\varphi_u(0))^{d_1} 0^{d_0},$$

3
where \( d_j \) are the coefficients from the normal \( U \)-representation (5).

In this paper we consider a special case of a simple Parry word, namely the Tribonacci word \( t \) that is defined over the alphabet \( A = \{0, 1, 2\} \) as the fixed point of the substitution

\[
\varphi_t : \begin{align*}
0 & \mapsto 01 \\
1 & \mapsto 02 \\
2 & \mapsto 0
\end{align*}
\]

i.e.,

\[
t = \lim_{k \to \infty} \varphi_t^k(0) = 01020100120102010201020102010201020102010201 \cdots.
\]

From now on denote \( T_j = |\varphi_t^j(0)| \) for every \( j \in \mathbb{N}_0 \). The numbers \( T_j \) are Tribonacci numbers for all \( j \in \mathbb{N}_0 \), because \( T_0 = 1, T_1 = 2, T_2 = 4 \) are Tribonacci numbers and the relation \( \varphi_t^j(0) = \varphi_t^{j-1}(0) \varphi_t^{j-2}(0) \varphi_t^{j-3}(0) \) (hence \( T_j = T_{j-1} + T_{j-2} + T_{j-3} \)) holds for every \( j \geq 3 \). Note, however, that \((T_0, T_1, T_2, T_3, \ldots) = (1, 2, 4, 7, \ldots)\) whereas the standard Tribonacci sequence reads \((0, 1, 1, 2, 4, 7, \ldots)\).

Any \( n \in \mathbb{N} \) can be represented as a sum of Tribonacci numbers with binary coefficients,

\[
n = \sum_{j=0}^{k} \delta_j T_j \quad \text{for} \quad \delta_j \in \{0, 1\}, \quad k \in \mathbb{N}_0.
\]

Similarly as above, if coefficients \( \delta_j \in \{0, 1\} \) are obtained by the greedy algorithm, we write

\[
\langle n \rangle_T = \delta_k \delta_{k-1} \cdots \delta_1 \delta_0.
\]

### 3 Abelian co-decomposition

Abelian co-decomposition has been developed as a tool for calculating \( \rho_u^{ab}(n) \) of recurrent words [8]. The matter of the method consists in using the expansion \( \langle n \rangle_U \) for transforming the set of all factors of length \( n \) into a certain set \( Z_u(n) \) of pairs of factors. In this section we summarize main facts and add illustrative examples. More details can be found in [8].

For any two factors \( v, w \) of \( u \) such that \( \Psi(v) = \Psi(w) \), we define

\[
P \left( \begin{array}{c} v \\ w \end{array} \right) = \{ \Psi(s) - \Psi(r) \mid r \text{ is a prefix of } v, s \text{ is a prefix of } w, |s| = |r| \}.
\]

**Example 3.1.** Let \( v = 0102 \), \( w = 1020 \). Then

\[
P \left( \begin{array}{c} v \\ w \end{array} \right) = \{ \Psi(1) - \Psi(0), \Psi(10) - \Psi(01), \Psi(102) - \Psi(010), \Psi(1020) - \Psi(0102) \}
\]

\[= \{ (-1, 1, 0), (0, 0, 0), (-1, 0, 1) \}. \]

Consider factors \( v, w \) of \( u \) such that \( \Psi(v) = \Psi(w) \). Let

\[
v = z_0 z_1 z_2 \cdots z_h \\
w = \tilde{z}_0 \tilde{z}_1 \tilde{z}_2 \cdots \tilde{z}_h
\]

(10)
for non-empty factors $z_0, z_1, \ldots, z_h$ and $\tilde{z}_0, \tilde{z}_1, \ldots, \tilde{z}_h$ satisfying $\Psi(\tilde{z}_j) = \Psi(z_j)$ for all $j \in \{0, 1, \ldots, h\}$. The set of ordered pairs

$$\text{Dec} \left( \begin{array}{c} v \\ w \end{array} \right) = \left\{ \left( \begin{array}{c} z_0 \\ \tilde{z}_0 \end{array} \right), \left( \begin{array}{c} z_1 \\ \tilde{z}_1 \end{array} \right), \ldots, \left( \begin{array}{c} z_h \\ \tilde{z}_h \end{array} \right) \right\}$$

(11)

is called abelian co-decomposition of the ordered pair $\left( \begin{array}{c} v \\ w \end{array} \right)$. An abelian co-decomposition (11) exists for any $v, w$ such that $\Psi(v) = \Psi(w)$, because one can take e.g. $\text{Dec} \left( \begin{array}{c} v \\ w \end{array} \right) = \left\{ \left( \begin{array}{c} v \\ w \end{array} \right) \right\}$. The decomposition (10) is in general not unique, but it can be made unique by an additional requirement. Here we will adopt, throughout the whole paper, the following convention: The number $h$ in equation (10) is chosen maximal possible. This requirement ensures uniqueness of $\text{Dec} \left( \begin{array}{c} v \\ w \end{array} \right)$.

Example 3.2. For $v = 0102$, $w = 1020$ we have

$$\begin{array}{ccl}
v &=& \begin{array}{cc}
z_1 & z_2 \\
01 & 02 \\
\tilde{z}_1 & \tilde{z}_2 \\
\end{array} \\
w &=& \begin{array}{cc}
10 & 20 \\
\tilde{z}_1 & \tilde{z}_2 \\
\end{array}
\end{array}$$

hence

$$\text{Dec} \left( \begin{array}{c} v \\ w \end{array} \right) = \left\{ \left( \begin{array}{c} 01 \\ 10 \end{array} \right), \left( \begin{array}{c} 02 \\ 20 \end{array} \right) \right\}.$$ 

For a given fixed point $u$ of (3) and for any $n \in \mathbb{N}$, we set (cf. [8, Def. 3.7 and Prop. 4.8])

$$Z_u(n) = \text{Dec} \left( \begin{array}{c} \varphi_{\eta}^{k+R}(0) \\ u_{[n]}^{-1}\varphi_{\eta}^{k+R}(0)u_{[n]} \end{array} \right),$$

(12)

where $R = m - 1 + \min\{\ell \geq 1 \mid \alpha \ell \geq 1\}$ and $k$ is any integer such that $n \leq U_k$. (Recall that $u_{[n]}$ denotes the prefix of $u$ of length $n$.)

The set $Z_u(n)$ defined in this way allows to determine the set of relative Parikh vectors corresponding to the number $n$. According to [8, Prop. 3.8], it holds

$$P_{u}^{\text{rel}}(n) = \bigcup_{\left( \begin{array}{c}
\tilde{z} \\
\end{array} \right) \in Z_u(n)} P \left( \begin{array}{c}
\tilde{z} \\
\end{array} \right)$$

(13)

for any $n \in \mathbb{N}$. Consequently, if $Z_u(n)$ is known, one can calculate $\rho_u^{ab}(n)$ using formula

$$\rho_u^{ab}(n) = \# \bigcup_{\left( \begin{array}{c}
\tilde{z} \\
\end{array} \right) \in Z_u(n)} P \left( \begin{array}{c}
\tilde{z} \\
\end{array} \right),$$

(14)

which follows immediately from equations (2) and (13).
Example 3.3. Let us calculate $Z_t(1)$. We have $t[1] = 0$. Since $1 \leq T_0 = 1$ and $R = 3 - 1 + \min\{1\} = 3$, we shall use formula (12) with $k + R = 0 + 3 = 3$. Therefore, due to (12),

$$Z_t(1) = \text{Dec}\left(\frac{\varphi_t^3(0)}{0^{-1}\varphi_t^3(0)0}\right).$$

We calculate

$$\varphi_t^3(0) = 01 \ 02 \ 01 \ 0$$
$$0^{-1}\varphi_t^3(0)0 = 10 \ 20 \ 10 \ 0$$

whence we obtain

$$Z_t(1) = \left\{\left(\begin{array}{c} 0 \\ 0 \end{array}\right), \left(\begin{array}{c} 01 \\ 10 \end{array}\right), \left(\begin{array}{c} 02 \\ 20 \end{array}\right)\right\}.$$  

It is important that $Z_u(n)$ does not need to be calculated from definition (using (12)) for each $n \in \mathbb{N}$. One can take advantage of the following fact (cf. [8, Prop. 5.1]). Let $u$ be the fixed point of (3) and $\langle N \rangle_U = \langle n \rangle_U \langle q \rangle_U$, i.e.,

$$\langle N \rangle_U = d_\ell \cdots d_{k+1}d_k \cdots d_0,$$
$$\langle n \rangle_U = d_\ell \cdots d_{k+1}, \quad \langle q \rangle_U = d_k \cdots d_0. \quad (15)$$

If $\varphi_u^{k+1}(\tilde{z})$ has the prefix $u[q]$ for all $(\tilde{z}) \in Z_u(n)$, then

$$Z_u(N) = \bigcup_{(\tilde{z}) \in Z_u(n)} \text{Dec}\left(\frac{\varphi_u^{k+1}(z)}{u^{-1}\varphi_u^{k+1}(\tilde{z})u[q]}\right). \quad (16)$$

When exploring the Tribonacci word, we will use a corollary of (16) that is obtained straightforwardly for $k = 0$ and $q = 0$ or $q = 1$.

Corollary 3.4. It holds

$$\langle N \rangle_T = \langle n \rangle_T 0 \quad \Rightarrow \quad Z_t(N) = \bigcup_{(\tilde{z}) \in Z_t(n)} \text{Dec}\left(\frac{\varphi_t(z)}{\varphi_t(\tilde{z})}\right), \quad (17)$$
$$\langle N \rangle_T = \langle n \rangle_T 1 \quad \Rightarrow \quad Z_t(N) = \bigcup_{(\tilde{z}) \in Z_t(n)} \text{Dec}\left(\frac{\varphi_t(0)}{0^{-1}\varphi_t(\tilde{z})0}\right). \quad (18)$$

For further convenience, we introduce the following shorthand. Let $z, \tilde{z}$ be factors of $t$ satisfying $\Psi(z) = \Psi(\tilde{z})$, and $\zeta$ stand for $(\frac{z}{\tilde{z}})$. Then we denote

$$D_0(\zeta) = \text{Dec}\left(\frac{\varphi_t(z)}{\varphi_t(\tilde{z})}\right), \quad D_1(\zeta) = \text{Dec}\left(\frac{\varphi_t(z)}{0^{-1}\varphi_t(\tilde{z})0}\right). \quad (19)$$
Recall that numbers $\delta_i$ in the representation (8) attain values 0 and 1 only. The statement of Corollary 3.4 can be thus formulated as

$$\langle N \rangle_T = \langle n \rangle_T \delta \Rightarrow Z_t(N) = \bigcup_{\zeta \in Z_t(n)} D_\delta(\zeta).$$

(20)

4 Structure of sets $Z(n)$

From now on we will deal with the Tribonacci word. Therefore, we will simplify the notation by dropping the subscript $t$ from symbols $\rho_t^{ab}$, $P_t^{rel}$, $\phi_t$, $Z_t$.

In this section we will show that there exist sets $Z_1, Z_2, \ldots, Z_M$ such that for any $n \in \mathbb{N}$, the set $Z(n)$ is equal to $Z_j$ for a certain $j \in \{1, \ldots, M\}$.

**Observation 4.1.** Let an $N \in \mathbb{N}$ have the representation $\langle N \rangle_T = 1 \delta_{k-1} \cdots \delta_1 \delta_0$. Let $n$ be the number with the representation $\langle n \rangle_T = 1 \delta_{k-1} \cdots \delta_1$, i.e., $\langle N \rangle_T = \langle n \rangle_T \delta_0$. Then

$$Z(N) = \bigcup_{\zeta \in Z(n)} D_0(\zeta) \quad \text{or} \quad Z(N) = \bigcup_{\zeta \in Z(n)} D_1(\zeta).$$

(21)

Observation 4.1 is just a trivial reformulation of equation (20). The main result of this section follows.

**Theorem 4.2.** There exist a 56-element set

$$Z_{\text{super}} = \{\zeta_1, \zeta_2, \zeta_3, \ldots, \zeta_{56}\}$$

and 277 its subsets $Z_1, Z_2, \ldots, Z_{277} \subset Z_{\text{super}}$ such that

$$\forall n \in \mathbb{N} \left( \exists j \in \{1, 2, \ldots, 277\} \right) \left( Z(n) = Z_j \right).$$

(22)

**Proof.** We begin the search for $Z_j$ having the property (22) by exploring $Z(n)$ for $n$ having a 1-digit representation, i.e., $\langle n \rangle_T = \delta_0$. Trivially, there is one single positive number having such representation, namely $n = 1$. We know from Example 3.3 that

$$Z(1) = \{\zeta_1, \zeta_2, \zeta_3\}$$

(23)

for

$$\zeta_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \zeta_2 = \begin{pmatrix} 01 \\ 10 \end{pmatrix}, \quad \zeta_3 = \begin{pmatrix} 02 \\ 20 \end{pmatrix}. $$

From now on we denote the set $Z(1)$ by $Z_1$.

Let us proceed to exploring $Z(N)$ for numbers $N$ having 2-digit representations, $\langle N \rangle_T = 1 \delta_0$. We apply Observation 4.1 for $k = 1$. Writing $\langle N \rangle_T = 1 \delta_0$ in the form $\langle n \rangle_T \delta_0$ implies $\langle n \rangle_T = 1$, hence $n = 1$. For such $N$ and $n$, formula (21) reads

$$Z(N) = \bigcup_{\zeta \in Z_1} D_0(\zeta) \quad \text{or} \quad Z(N) = \bigcup_{\zeta \in Z_1} D_1(\zeta).$$

7
Since $Z_1 = \{\zeta_1, \zeta_2, \zeta_3\}$, we need to find $D_0(\zeta_j)$ and $D_1(\zeta_j)$ for $j = 1, 2, 3$. It is an easy task. We start with $D_0(\zeta_j)$. With regard to (19), we calculate
\[
\varphi(0) = 01 \quad \varphi(01) = 0102 \quad \varphi(02) = 010 \quad \varphi(0) = 01 \quad \varphi(10) = 0201 \quad \varphi(20) = 011
\]
Hence
\[
D_0(\zeta_1) = \{\zeta_1, \zeta_4\}, \quad D_0(\zeta_2) = \{\zeta_1, \zeta_5\}, \quad D_0(\zeta_3) = \{\zeta_1, \zeta_6\}, \tag{24}
\]
where $\zeta_1 = \left(\begin{array}{c} 0 \\ 0 \end{array}\right)$ has been defined above and
\[
\zeta_4 = \left(\begin{array}{c} 1 \\ 1 \end{array}\right), \quad \zeta_5 = \left(\begin{array}{c} 102 \\ 201 \end{array}\right), \quad \zeta_6 = \left(\begin{array}{c} 10 \\ 01 \end{array}\right).
\]
A similar calculation leads to sets $D_1(\zeta_j)$ for $j = 1, 2, 3$, see (19). We have
\[
\varphi(0) = 01, \quad \varphi(01) = 0102, \quad \varphi(02) = 010, \quad 0^{-1}\varphi(00) = 10, \quad 0^{-1}\varphi(01) = 2010, \quad 0^{-1}\varphi(02) = 0110
\]
hence
\[
D_1(\zeta_1) = \{\zeta_2\}, \quad D_1(\zeta_2) = \{\zeta_7\}, \quad D_1(\zeta_3) = \{\zeta_1, \zeta_4\} \tag{25}
\]
for $\zeta_1, \zeta_2, \zeta_4$ defined above and
\[
\zeta_7 = \left(\begin{array}{c} 0102 \\ 2010 \end{array}\right).
\]
To sum up, if $\langle N \rangle_T = 1\delta_0$, then
\[
Z(N) = \{\zeta_1, \zeta_4, \zeta_5, \zeta_6\} := Z_2 \quad \text{or} \quad Z(N) = \{\zeta_1, \zeta_2, \zeta_4, \zeta_7\} := Z_3. \tag{26}
\]
We will list sets $Z_j$ in Table 1. The elements $\zeta_i$ together with $D_0(\zeta_i)$ and $D_1(\zeta_i)$ will be listed in Table 2.

Now we can apply Observation 4.1 again, this time for $k = 2$, to explore $Z(N)$ for numbers $N$ having 3-digit representations, $\langle N \rangle_T = 1\delta_1\delta_0$. Any such $N$ can be written in the form $\langle n \rangle_T\delta_0$ for $\langle n \rangle_T = 1\delta_1$. The representation of $n$ has two digits, thus $Z(n) = Z_2$ or $Z(n) = Z_3$, see the previous step. The application of equation (21) for $Z(n) = Z_2$ and $Z(n) = Z_3$ requires the knowledge of $D_0(\zeta_j)$ and $D_1(\zeta_j)$ for $j = 1, \ldots, 7$. Sets $D_0(\zeta_j)$ and $D_1(\zeta_j)$ for $j = 1, 2, 3$ are already known, see equations (24), (25). Concerning $j = 4, \ldots, 7$, a short calculation gives
\[
D_0(\zeta_4) = \{\zeta_1, \zeta_8\}, \quad D_0(\zeta_5) = \{\zeta_1, \zeta_9\}, \quad D_0(\zeta_6) = \{\zeta_1, \zeta_{10}\}, \quad D_0(\zeta_7) = \{\zeta_1, \zeta_6, \zeta_{11}\},
\]
\[
D_1(\zeta_4) = \{\zeta_3\}, \quad D_1(\zeta_5) = \{\zeta_1, \zeta_{10}\}, \quad D_1(\zeta_6) = \{\zeta_{12}\}, \quad D_1(\zeta_7) = \{\zeta_1, \zeta_4, \zeta_8\} \tag{27}
\]
for
\[
\zeta_8 = \left(\begin{array}{c} 2 \\ 2 \end{array}\right), \quad \zeta_9 = \left(\begin{array}{c} 2010 \\ 0102 \end{array}\right), \quad \zeta_{10} = \left(\begin{array}{c} 201 \\ 102 \end{array}\right), \quad \zeta_{11} = \left(\begin{array}{c} 20 \\ 02 \end{array}\right), \quad \zeta_{12} = \left(\begin{array}{c} 0201 \\ 1020 \end{array}\right).
\]
If we substitute $Z(n) = Z_2$ and $Z(n) = Z_3$ into equation (21) and use the known structure of $Z_2, Z_3$ (cf. (26)) together with equations (24), (25) and (27), we get

$$Z(n) = Z_2 \quad \Rightarrow \quad Z(N) = \{\zeta_1, \zeta_4, \zeta_8, \zeta_9, \zeta_{10}\} =: Z_6,$$

$$Z(n) = Z_3 \quad \Rightarrow \quad Z(N) = \{\zeta_1, \zeta_4, \zeta_5, \zeta_6, \zeta_8, \zeta_{11}\} =: Z_7.$$

The calculation continues in the same way for $k = 3$. For each $j \in \{4, 5, 6, 7\}$, we substitute the set $Z_j$ into (21) for $Z(n)$. Evaluation of $Z(N)$ requires sets $D_0(\zeta_j)$ and $D_1(\zeta_j)$ for $j = 1, \ldots, 12$. We draw them from previous results (for $j = 1, \ldots, 7$), or calculate them now (for $j = 8, \ldots, 12$). In this way we find eight sets $Z(N)$, namely

$$Z(n) = Z_4 \quad \Rightarrow \quad Z(N) = Z_7 \quad \text{or} \quad Z(N) = Z_{10},$$

$$Z(n) = Z_5 \quad \Rightarrow \quad Z(N) = Z_8 \quad \text{or} \quad Z(N) = Z_{11},$$

$$Z(n) = Z_6 \quad \Rightarrow \quad Z(N) = Z_9 \quad \text{or} \quad Z(N) = Z_{12},$$

$$Z(n) = Z_7 \quad \Rightarrow \quad Z(N) = Z_5 \quad \text{or} \quad Z(N) = Z_7.$$

Their structure is shown in Table 1. Note that the eight sets $Z(N)$ listed above are of two types:

- three of them ($Z_7$ occurring 2 times and $Z_5$) have been found in previous steps;
- five of them ($Z_8, Z_9, Z_{10}, Z_{11}, Z_{12}$) are “new” – they appear in the calculation for the first time now.

| $j$ | $\{i \mid \zeta_i \in Z_j\}$ |
|-----|-------------------------------|
| 1   | $1 ; 2 ; 3$                   |
| 2   | $1 ; 4 ; 5 ; 6$               |
| 3   | $1 ; 2 ; 4 ; 7$               |
| 4   | $1 ; 4 ; 8 ; 9 ; 10$          |
| 5   | $1 ; 4 ; 5 ; 6 ; 8 ; 11$      |
| 6   | $1 ; 2 ; 3 ; 10 ; 12$         |
| 7   | $1 ; 2 ; 3 ; 4 ; 7 ; 8$       |
| 8   | $1 ; 2 ; 4 ; 8 ; 9 ; 10$      |
| 9   | $1 ; 4 ; 5 ; 6 ; 7 ; 13$      |
| 10  | $1 ; 2 ; 3 ; 14 ; 15$         |
| 11  | $1 ; 2 ; 3 ; 10 ; 12 ; 16$    |
| 12  | $1 ; 2 ; 4 ; 7 ; 15 ; 17$     |
| ... |                               |
| 277 | $1 ; 2 ; 4 ; 7 ; 15 ; 17 ; 22 ; 23 ; 24 ; 36 ; 43 ; 50$ |

Table 1: Structure of sets $Z_j$. 
Table 2: Elements of $\mathcal{Z}(n)$. The table shows also the structure of sets $D_0(\zeta_j), D_1(\zeta_j), \text{and } P(\zeta_j)$ (see Sect. 5).

| $j$ | $\zeta_j$ | $\{i \mid \zeta_i \in D_0(\zeta_j)\}$ | $\{i \mid \zeta_i \in D_1(\zeta_j)\}$ | $\{i \mid \psi_i \in P(\zeta_j)\}$ |
|-----|-----------|--------------------------------------|--------------------------------------|----------------------------------|
| 1   | $\left(\frac{0}{0}\right)$ | 1:4 | 2 | 0 |
| 2   | $\left(\frac{01}{10}\right)$ | 1:5 | 7 | 0:1 |
| 3   | $\left(\frac{02}{20}\right)$ | 1:6 | 1:4 | 0:2 |
| 4   | $\left(\frac{1}{1}\right)$ | 1:8 | 3 | 0 |
| 5   | $\left(\frac{102}{201}\right)$ | 1:9 | 1:10 | 0:3 |
| 6   | $\left(\frac{10}{01}\right)$ | 1:10 | 12 | 0:4 |
| 7   | $\left(\frac{0102}{2010}\right)$ | 1:6:11 | 1:4:8 | 0:2:3 |
| 8   | $\left(\frac{2}{2}\right)$ | 1 | 1 | 0 |
| 9   | $\left(\frac{2010}{0102}\right)$ | 1:2:3 | 14 | 0:5:6 |
| 10  | $\left(\frac{201}{102}\right)$ | 1:7 | 15 | 0:6 |
| 11  | $\left(\frac{20}{20}\right)$ | 1:2 | 16 | 0:5 |
| 12  | $\left(\frac{0201}{1020}\right)$ | 1:13 | 17 | 0:1:6 |
| ... |  |  |  |  |
| 56  | $\left(\frac{10010201020}{02010201001}\right)$ | 1:25:26:52 | 1:27:28:29 | 0:3:4 |

Since sets $\mathcal{Z}_5$ and $\mathcal{Z}_7$ have been already explored, they are no more interesting at this moment and may be put aside. In the next step, when exploring numbers $N$ with representations having 5 digits ($k = 4$), we will use formula (21) only with $\mathcal{Z}(n) = \mathcal{Z}_j$ for such $\mathcal{Z}_j$ that promise new results, i.e., $j = 8, 9, 10, 11, 12$. Henceforth we will proceed similarly – we will always put aside those sets $\mathcal{Z}_j$ that reappear after having been explored earlier, and queue the “new” ones for further use in (21).

The progress of the calculation is illustrated with Table 3. We see that when $k$ reaches the value 22, no new set $\mathcal{Z}_j$ is found. In other words, all sets $\mathcal{Z}(N)$ obtained by formula (21) for $k = 22$ have been already found (and explored) earlier. The search is then completed. We conclude: there exist altogether 277 sets $\mathcal{Z}_1, \ldots, \mathcal{Z}_{277}$ such that for any $n \in \mathbb{N}$, it holds $\mathcal{Z}(n) = \mathcal{Z}_j$ for a certain $j \in \{1, \ldots, 277\}$.

Since every $\mathcal{Z}_j$ consists of elements $\zeta_i$ for $i = 1, \ldots, 56$ (note that new elements $\zeta_i$ stop appearing already at $k = 17$), the set $\mathcal{Z}_{\text{super}} := \{\zeta_1, \zeta_2, \zeta_3, \ldots, \zeta_{55}, \zeta_{56}\}$ obviously satisfies $\mathcal{Z}_j \subset \mathcal{Z}_{\text{super}}$ for all $j = 1, \ldots, 277$.

The search for sets $\mathcal{Z}_j$ and elements $\zeta_i$ can be in principle completely performed using pen and paper, but since the procedure is a little lengthy and cumbersome, it is better to use a computer, which also helps to avoid mistakes.
### Table 3: Progress of the calculation. Sets $Z_j$ expressed in terms of $\zeta_i$ can be found in Tab. 1. Elements $\zeta_i$ are listed in Tab. 2.

| $k$ | new $Z_j$ found | new $\zeta_i$ found | $k$ | new $Z_j$ found | new $\zeta_i$ found |
|-----|-----------------|---------------------|-----|-----------------|---------------------|
| 0   | $Z_1$           | $\zeta_1, \zeta_2, \zeta_3$ | 15  | $Z_{201}, Z_{221}$ | $\zeta_{54}, \zeta_{55}$ |
| 1   | $Z_2, Z_3$     | $\zeta_4, \zeta_5, \zeta_6, \zeta_7$ | 16  | $Z_{222}, Z_{245}$ | $\zeta_{56}$ |
| 2   | $Z_4, Z_5, Z_6, Z_7$ | $\zeta_8, \zeta_9, \zeta_{10}, \zeta_{11}, \zeta_{12}$ | 17  | $Z_{246}, Z_{260}$ | none |
| 3   | $Z_8, Z_9, Z_{10}, Z_{11}, Z_{12}$ | $\zeta_{13}, \zeta_{14}, \zeta_{15}, \zeta_{16}, \zeta_{17}$ | 18  | $Z_{261}, Z_{271}$ | none |
| 4   | $Z_{13}, Z_{21}$ | $\zeta_{18}, \zeta_{24}$ | 19  | $Z_{272}, Z_{273}$ | none |
| 5   | $Z_{22}, Z_{32}$ | $\zeta_{25}, \zeta_{30}$ | 20  | $Z_{274}, Z_{275}$ | none |
| ... |                 |                     | 21  | $Z_{276}, Z_{277}$ | none |
| 14  | $Z_{179}, Z_{200}$ | $\zeta_{52}, \zeta_{53}$ | 22  | none             | none |

5 Range of $\rho^{ab}(n)$

For any $n \in \mathbb{N}$, formula $\rho^{ab}(n) = \# \bigcup_{\zeta \in Z(n)} P(\zeta)$ (eq. (14)) allows to determine $\rho^{ab}(n)$ if the set $Z(n)$ is known. At this moment we do not know $Z(n)$ explicitly, thus we cannot use (14) as it is. Nevertheless, since $Z(n) = Z_j$ for a certain $j \in \{1, \ldots, 277\}$, we are already able to restrict the range of $\rho^{ab}(n)$.

At first we find $P(\zeta_i)$ for $i = 1, \ldots, 56$. If formula (9) is applied on $\zeta_1$, it gives

$$P(\zeta_1) = P \binom{0}{0} = \{\Psi(s) - \Psi(r) \mid r \text{ is a prefix of } 0, \text{ s is a prefix of } 0, |s| = |r|\}$$

$$= \{(0, 0, 0)\}.$$

Similarly, for $\zeta_2$ and $\zeta_3$ we get

$$P(\zeta_2) = P \binom{01}{10} = \{\Psi(s) - \Psi(r) \mid r \text{ is a prefix of } 01, \text{ s is a prefix of } 10, |s| = |r|\}$$

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

$$P(\zeta_3) = P \binom{02}{20} = \{\Psi(s) - \Psi(r) \mid r \text{ is a prefix of } 02, \text{ s is a prefix of } 20, |s| = |r|\}$$

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

Performing the same calculation for other values of $i$, we find that for all $i = 1, \ldots, 56$, it holds $P(\zeta_i) \subset \{\psi_0, \psi_1, \ldots, \psi_8\}$, where

- $\psi_0^s = (0, 0, 0)$;
- $\psi_0^r = (-1, 1, 0)$;
- $\psi_0^r = (-1, 0, 1)$;
- $\psi_1^s = (0, -1, 1)$;
- $\psi_1^r = (1, -1, 0)$;
- $\psi_1^r = (1, 0, -1)$;
- $\psi_4^s = (0, 1, -1)$;
- $\psi_4^r = (-1, 2, -1)$;
- $\psi_5^r = (-1, -1, 2)$.

The structure of sets $P(\zeta_j)$ in terms of $\psi_0^s, \ldots, \psi_8^s$ is partly described in Table 2.
Table 4: Map $C$ (partial list for $j = 1, \ldots, 42$).

| $j$ | $C(j)$ | $j$ | $C(j)$ | $j$ | $C(j)$ | $j$ | $C(j)$ | $j$ | $C(j)$ |
|-----|--------|-----|--------|-----|--------|-----|--------|-----|--------|
| 1   | 3      | 6   | 4      | 11  | 4      | 16  | 4      | 21  | 4      |
| 2   | 3      | 7   | 4      | 12  | 4      | 17  | 4      | 22  | 5      |
| 3   | 4      | 8   | 4      | 13  | 4      | 18  | 4      | 23  | 5      |
| 4   | 3      | 9   | 4      | 14  | 4      | 19  | 4      | 24  | 3      |
| 5   | 4      | 10  | 3      | 15  | 3      | 20  | 4      | 25  | 4      |
| 6   | 4      | 11  | 4      | 16  | 4      | 21  | 4      | 26  | 4      |

Since $Z(n) = Z_j$ implies $\rho^{ab}(n) = \# \bigcup_{\zeta \in Z_j} P(\zeta)$, we now calculate the quantities

$$C(j) = \# \bigcup_{\zeta \in Z_j} P(\zeta)$$

for $j = 1, \ldots, 277$. It is a straightforward task, consisting in combining data from Tables 1 and 2.

1. Table 1 shows the structure of $Z_j$ in terms of $\zeta_i$.

2. Table 2 contains $P(\zeta_i)$ for $\zeta_i$.

For example, for $j = 1$ we obtain

$$\bigcup_{\zeta \in Z_1} P(\zeta) = \{\psi_0^r, \psi_1^r, \psi_2^r\}$$

where we used data from Table 1 at (a) and data from Table 2 at (b). Hence $C(1) = \# \bigcup_{\zeta \in Z_1} P(\zeta) = 3$. We proceed similarly for $j = 2, \ldots, 277$. The values $C(j)$ obtained by the calculation are listed in Table 4. The list is only partial by reason of saving space, however, in the full list we could see that $C(j)$ \in \{3, 4, 5, 6, 7\} for all $j \in \{1, \ldots, 277\}$. Consequently, $\rho^{ab}(n) \in \{3, 4, 5, 6, 7\}$ for all $n \in \mathbb{N}$.

Moreover, we can show similarly that for any $j = 1, \ldots, 277$, the absolute values of the elements of the difference vector $\psi^r_i - \psi^r_h$ for $\psi^r_i, \psi^r_h \in \bigcup_{\zeta \in Z_j} P(\zeta)$ do not exceed 2. This fact implies that the Tribonacci word is 2-balanced. However, this bound as well as the result $\rho^{ab}(n) \in \{3, 4, 5, 6, 7\}$ are not new; they have been derived already in [1].

6 Evaluation of $\rho^{ab}(n)$

Let us introduce maps (endomorphisms) $\sigma_0, \sigma_1$ on $\{0, 1, \ldots, 277\}$ defined for $j = 1, \ldots, 277$ as

$$\sigma_0(j) = \ell \Leftrightarrow \bigcup_{\zeta \in Z_j} D_0(\zeta) = Z_\ell \quad \text{and} \quad \sigma_1(j) = \ell \Leftrightarrow \bigcup_{\zeta \in Z_j} D_1(\zeta) = Z_\ell$$

(30)
Values $\sigma_0(j)$ and $\sigma_1(j)$ for $j = 1, \ldots, 277$ follow from calculations carried out in the proof of Theorem 4.2. For example, it holds $\bigcup_{\zeta \in \mathbb{Z}} D_0(\zeta) = \mathbb{Z}$ and $\bigcup_{\zeta \in \mathbb{Z}} D_0(\zeta) = \mathbb{Z}_3$, hence $\sigma_0(1) = 2$ and $\sigma_1(1) = 3$. The values can be tabulated, see Table 5.

Maps $\sigma_0$ and $\sigma_1$ play an important role, as they allow to determine $\mathcal{Z}(n)$ for any $n \in \mathbb{N}$ using the following proposition.

**Proposition 6.1.** If $\langle n \rangle_T = \delta_k \delta_{k-1} \cdots \delta_1 \delta_0$, then

$$\mathcal{Z}(n) = \mathcal{Z}_j \quad \text{for} \quad j = (\sigma_{\delta_0} \circ \sigma_{\delta_1} \circ \cdots \circ \sigma_{\delta_{h-1}} \circ \sigma_{\delta_h})(0).$$  \hfill (32)

**Proof.** At first we prove the statement by induction on $k$ on the assumption that $\delta_k = 1$.

I. Let $k = 0$. Then $\langle n \rangle_T = 1$, i.e., $n = 1$. We have $\sigma_{\delta_0}(0) = \sigma_1(0) = 1$, thus formula (32) gives $\mathcal{Z}(1) = \mathcal{Z}_1$. Indeed, this result holds true by definition of $\mathcal{Z}_1$.

II. Let $\langle n \rangle_T = \delta_k \delta_{k-1} \cdots \delta_1 \delta_0$ have $k + 1$ digits. We assume that formula (32) is valid for $\langle q \rangle_T = \delta_k \delta_{k-1} \cdots \delta_1$ (because $\langle q \rangle_T$ has only $k$ digits), hence $\mathcal{Z}(q) = \mathcal{Z}_\ell$ for $\ell = (\sigma_{\delta_0} \circ \cdots \circ \sigma_{\delta_h})(0)$.

Due to equation (20) it holds $\mathcal{Z}(n) = \bigcup_{\zeta \in \mathbb{Z}_T} D_0(\zeta)$. Consequently, with regard to the definition of $\sigma_0$ and $\sigma_1$ (cf. (30)), we have $\mathcal{Z}(n) = \mathcal{Z}_j$ for $j = \sigma_{\delta_h}(\ell)$.

Combining these two facts, we obtain

$$\mathcal{Z}(n) = \mathcal{Z}_j \quad \text{for} \quad j = (\sigma_{\delta_0} \circ \sigma_{\delta_1} \circ \cdots \circ \sigma_{\delta_{h-1}} \circ \sigma_{\delta_h})(0).$$

Finally, let $\delta_k = 0$, i.e., $\langle n \rangle_T = 00 \cdots 0 \delta_h \delta_{h-1} \cdots \delta_0$ for an $h < k$ and $\delta_h = 1$. The initial block of zeros can be omitted, i.e., we can write $\langle n \rangle_T = \delta_h \delta_{h-1} \cdots \delta_0$ for $\delta_h = 1$. Then, due to previous considerations,

$$\mathcal{Z}(n) = \mathcal{Z}_j \quad \text{for} \quad j = (\sigma_{\delta_0} \circ \cdots \circ \sigma_{\delta_{h-1}} \circ \sigma_{\delta_h})(0).$$

Since $(\sigma_0 \circ \cdots \circ \sigma_0)(0) = 0$, it holds also $j = (\sigma_{\delta_0} \circ \cdots \circ \sigma_{\delta_{h-1}} \circ \sigma_{\delta_h} \circ \sigma_{\delta_{h+1}} \cdots \circ \sigma_{\delta_k})(0)$.

\[ \boxend \]
Now we recall that if $Z(n) = Z_j$, then $\rho^{ab}(n) = C(j)$, where $C$ is the map introduced in Section 5. It follows from equations (14) and (29). Combining this fact with equation (32), we get

**Theorem 6.2.** If $\langle n \rangle_T = \delta_k \delta_{k-1} \cdots \delta_1 \delta_0$, then

$$\rho^{ab}(n) = C(j) \quad \text{for} \quad j = (\sigma_{\delta_0} \circ \sigma_{\delta_1} \circ \cdots \circ \sigma_{\delta_{k-1}} \circ \sigma_{\delta_k})(0). \quad (33)$$

Theorem (6.2) together with Tables 5 and 4 allows to determine $\rho^{ab}(n)$ for any $n \in \mathbb{N}$. However, tables 5 and 4 have rather an illustrative function. Their structure is temporary and they will be improved (condensed) in Section 7. For this reason these two tables are currently shown only partially.

The number of steps needed for evaluating $\rho^{ab}(n)$ is proportional to the value $k$, i.e., to the number of digits in the representation $\langle n \rangle_T$. Numbers $T_j$ grow roughly exponentially ($T_j \approx 1.84^j$), thus $k \sim \log n$. To sum up, the evaluation of $\rho^{ab}(n)$ using formula (33) needs $\mathcal{O}(\log n)$ operations, i.e., it is fast even for large values of $n$.

**Example 6.3.** Let us calculate $\rho^{ab}(n)$ for $n = 2013$. Since

$$(T_j)_{j=0}^{\infty} = (1, 2, 4, 7, 13, 24, 44, 81, 149, 274, 504, 927, 1705, 3136, \ldots),$$

we have $\langle 2013 \rangle_T = 1001000101011$. Applying the endomorphisms $\sigma_0$, $\sigma_1$ step by step, we obtain

$$0 \overset{\sigma_0}{\rightarrow} 1 \overset{\sigma_3}{\rightarrow} 2 \overset{\sigma_3}{\rightarrow} 4 \overset{\sigma_3}{\rightarrow} 10 \overset{\sigma_3}{\rightarrow} 15 \overset{\sigma_3}{\rightarrow} 24 \overset{\sigma_3}{\rightarrow} 7 \overset{\sigma_3}{\rightarrow} 7 \overset{\sigma_3}{\rightarrow} 5 \overset{\sigma_3}{\rightarrow} 11 \overset{\sigma_3}{\rightarrow} 16 \overset{\sigma_3}{\rightarrow} 19 \overset{\sigma_3}{\rightarrow} 32,$$

thus $(\sigma_1 \circ \sigma_1 \circ \sigma_0 \circ \sigma_1 \circ \sigma_0 \circ \sigma_1 \circ \sigma_0 \circ \sigma_0 \circ \sigma_0 \circ \sigma_0 \circ \sigma_0 \circ \sigma_1)(0) = 32$. Hence

$\rho^{ab}(2013) = C(32) = 4$.

### 7 Graph of $\rho^{ab}$. Simplified formula for $\rho^{ab}(n)$

The result of Theorem 6.2 can be interpreted in terms of walks in a certain oriented graph. Consider an oriented graph $\Gamma = (V, E)$ such that

- $V = \{v_0, v_1, \ldots, v_{277}\}$ is the set of vertices;
- $E = E_0 \cup E_1$ is the set of oriented edges;
- $e = (v_j, v_\ell) \in E_0$ iff $\sigma_0(j) = \ell$, and $e = (v_j, v_\ell) \in E_1$ iff $\sigma_1(j) = \ell$.

Note that each vertex has exactly one outgoing edge belonging to $E_0$ and exactly one outgoing edge belonging to $E_1$. The vertices and edges of graph $\Gamma$ are labeled as follows.

- **Edges:** Each $e \in E_0$ is labeled by 0; each $e \in E_1$ is labeled by 1.
- **Vertices:** For all $j = 1, \ldots, 277$, $v_j \in V$ is labeled by the number $C(j)$.
The interpretation of the graph $\Gamma$ follows straightforwardly from Theorem 6.2. Let $\langle n \rangle_T = \delta_k \delta_{k-1} \cdots \delta_0$ for $\delta_i \in \{0, 1\}$. Consider a walk in $\Gamma$ consisting of $k + 1$ edges, defined by the conditions

i. the start vertex is $v_0$;

ii. the edges in the walk are chosen so that their labels form the sequence $(\delta_k, \delta_{k-1}, \ldots, \delta_0)$.

Then the label of the end vertex of the walk is equal to $\rho^{ab}(n)$.

**Reduction of $\Gamma$**

In the rest of the section we will show that the number of vertices of $\Gamma$ can be substantially reduced. It will allow us to shorten the tables needed for the use of Theorem 6.2, and thus to simplify the calculation of $\rho^{ab}(n)$.

It is easy to see that if certain $j, \ell \in \{1, \ldots, 277\}$ ($j \neq \ell$) satisfy

$$\sigma_0(j) = \sigma_0(\ell), \quad \sigma_1(j) = \sigma_1(\ell) \quad \text{and} \quad C(j) = C(\ell).$$

then vertices $v_j$ and $v_\ell$ can be merged into a single vertex with label $C(j)$. That is, we

1. remove vertex $v_\ell$ and both its outgoing edges;

2. replace each edge $(v_h, v_\ell) \in E_\delta$ (where $\delta = 0$ or $\delta = 1$) by a new edge $(v_h, v_j) \in E_\delta$ (i.e., the replaced edge and the replacing edge have identical labels, 0 or 1).

Then the walk starting at $v_0$ and consisting of edges with labels $\delta_k, \ldots, \delta_0$ in the original graph, and the walk starting at $v_0$ and consisting of edges with labels $\delta_k, \ldots, \delta_0$ in the new graph, obviously both traverse vertices with exactly the same sequence of labels. In particular, their end vertices have identical labels. Consequently, despite the new graph has a smaller number of vertices and a smaller number of edges than the original graph, it has the same interpretation and can serve for determining $\rho^{ab}(n)$ using the representation $\langle n \rangle_T = \delta_k \cdots \delta_0$ in the same way.

For example, in $\Gamma$ it holds (see Tables 4 and 5)

$$\sigma_0(i) = 14, \quad \sigma_1(i) = 19, \quad C(i) = 4 \quad \text{for all} \quad i = 9, 16, 27, 37, 38, 67, 68, 123.$$

Therefore, all vertices $v_9, v_{16}, v_{27}, v_{37}, v_{38}, v_{67}, v_{68}, v_{123}$ can be merged together, which reduces the number of vertices of $\Gamma$ by 7.

Let the merging procedure be applied on the graph $\Gamma$ wherever possible. After finishing the first run of the procedure, we repeat it again, time after time, until no more mergeable vertices exist. In this way one obtains a new, reduced, graph $\Gamma_{\text{red}}$. It turns out that $\Gamma_{\text{red}}$ has just 68 vertices; let us renumber them by $w_0, w_1, \ldots, w_{67}$.
Graph $\Gamma_{\text{red}}$ inherits the main qualitative properties of $\Gamma$. In particular, each vertex has exactly one outgoing edge labeled with 0 and exactly one outgoing edge labeled with 1. This fact allows us to define endomorphisms $\tau_0, \tau_1$ of $\{0, 1, \ldots, 67\}$ by the relations

$$
\tau_0(j) = \ell \quad \text{iff the outgoing edge from } w_j \text{ labeled with 0 ends in } w_\ell;
$$

$$
\tau_1(j) = \ell \quad \text{iff the outgoing edge from } w_j \text{ labeled with 1 ends in } w_\ell.
$$

We also define a map $A$ by

$$
A(j) = \text{label of the vertex } w_j, \text{ inherited from the graph } \Gamma.
$$

Values of maps $\tau_0, \tau_1$ and $A$ are tabulated in Table 6. It is obvious from the construction of $\Gamma_{\text{red}}$ that Theorem 6.2 can be reformulated in terms of $\tau_0, \tau_1$ and $A$.

**Theorem 7.1.** If $\langle n \rangle_T = \delta_k \delta_{k-1} \cdots \delta_1 \delta_0$, then

$$
\rho^{ab}(n) = \left(A \circ \tau_{\delta_k} \circ \tau_{\delta_{k-1}} \circ \cdots \circ \tau_{\delta_1} \circ \tau_{\delta_0}\right)(0).
$$

(34)
Equation (34) supplemented by Table 6 represents the main result of our paper. Formula (34) is technically easier to be applied than (33), because a smaller table is used.

**Example 7.2.** Let us calculate $\rho_{ab}(n)$ for $n = 2013$ using formula (34). Since $\langle 2013 \rangle_T = 1001000101011$ (see Ex. 6.3), we obtain, using Table 6,

\[0 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6 \rightarrow 7 \rightarrow 8 \rightarrow 9 \rightarrow 10.\]

Hence $\rho_{ab}(2013) = A(10) = 4.$

**Characterization of $n$ such that $\rho_{ab}(n) = m \in \{4, 5, 6, 7\}$**

Now we can return to the open problem proposed by Richomme, Saari and Zamboni [2] concerning the characterization of numbers $n$ for which $\rho_{ab}(n) = m$, $m \in \{4, 5, 6, 7\}$.

The graph $\Gamma_{\text{red}}$ gives a certain answer. Namely, the set \{\(n | \rho_{ab}(n) = m\)\} contains such \(\langle n \rangle_T = \delta_k \delta_{k-1} \cdots \delta_1 \delta_0\) for which the walk from $v_0$ along edges labeled with $\delta_k, \ldots, \delta_1, \delta_0$ (in this order) ends in any of the vertices labeled with the value $m$.

### 8 Conclusions

The first question that naturally arises at this point is whether the approach can give results for other words. It is easy to see that the procedure can be straightforwardly applied to $m$-bonacci words, which are fixed points of substitutions $0 \mapsto 01, 1 \mapsto 02, \ldots, m - 2 \mapsto 0(m - 1), m - 1 \mapsto 0$ for $m \geq 2$. To explore an $m$-bonacci word for any $m \geq 2$ in this way, it suffices to change just the constant $R$ in Example 3.3 from the value 3 to $m$. On the other hand, the cardinality of $\mathbb{Z}_{\text{super}}$ quickly grows with $m$ and the method ceases to be efficient.

The method can be also easily adapted for application to other simple Parry words. The modifications that are needed are again rather small. Further, despite we have considered simple Parry words in this paper, the abelian co-decomposition method allows to examine non-simple Parry words as well, cf. [8]. To sum up, the approach can work for both simple and non-simple Parry words, but in practice it will more likely work well in cases when the image of the abelian complexity function is a set of low cardinality. Nevertheless, it can give new results for various words for which other methods fail.

Another potentially interesting question is whether this method (possibly after a certain improvement) can be used for dealing with a word that depends on a parameter, i.e., whether one can explore a parametric family of words en bloc. Consider for instance the $m$-bonacci word for a general $m \geq 2$. We are convinced that the procedure would work with a parameter as well, although the calculation would be of course intricate and lengthy.

The procedure can also give as a by-product the optimal balance bound of the examined word. Consequently, the method can be regarded not only as
a tool for evaluating the abelian complexity, but also as a tool for exploring balance properties of words. In particular, it is possible that this approach can lead to the optimal balance bound for the $m$-bonacci word for any $m$. Recall that the optimal bound for the $m$-bonacci word is not known yet, despite an upper bound has been already determined [15].

References

[1] G. Richomme, K. Saari, L. Q. Zamboni, Abelian complexity in minimal subshifts, *J. London Math. Soc.* **83**(1) (2011) 79–95.

[2] G. Richomme, K. Saari, L. Q. Zamboni, Balance and Abelian complexity of the Tribonacci word, *Adv. Appl. Math.* **45** (2010) 212–231.

[3] J. Currie, N. Rampersad, Recurrent words with constant Abelian complexity, *Adv. Appl. Math.* **47** (2011), 116–124.

[4] J. Cassaigne, G. Richomme, K. Saari, L. Q. Zamboni, Avoiding Abelian powers in binary words with bounded Abelian complexity, *Int. J. Found. Comp. S.* **22** (2011) 905–920.

[5] O. Turek, Balances and Abelian complexity of a certain class of infinite ternary words, *RAIRO: Theoret. Informatics Appl.* **44** (2010) 313–337.

[6] L. Balková, K. Brinda, O. Turek, Abelian complexity of infinite words associated with quadratic Parry numbers, *Theor. Comput. Sci.* **412** (2011) 6252–6260.

[7] H. Ardal, T. Brown, V. Jungić, J. Sahasrabudhe, On abelian and additive complexity in infinite words, *Integers* **12** (2012) #A21.

[8] O. Turek, Abelian complexity and abelian co-decomposition, *Theor. Comput. Sci.* **469** (2013) 77–91.

[9] B. Madill, N. Rampersad, The abelian complexity of the paperfolding word, *Discrete Math.* **313** (2013), 831–838.

[10] E. M. Coven and G. A. Hedlund, Sequences with minimal block growth, *Math. Syst. Theory* **7** (1973) 138–153.

[11] W. Parry, On the $\beta$-expansions of real numbers, *Acta Math. Acad. Sci. H.* **11** (1960) 401–416.

[12] S. Fabre, Substitutions et $\beta$-systèmes de numération, *Theor. Comput. Sci.* **137** (1995) 219–236.

[13] M. Lothaire, *Algebraic Combinatorics on Words*, volume 90 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2002.
[14] J.-M. Dumont and A. Thomas. Systèmes de numération et fonctions fractales relatifs aux substitutions. *Theor. Comput. Sci.* **65** (1989) 153–169.

[15] K. Břinda, E. Pelantová, O. Turek, Balances of $m$-bonacci words, *Fundam. Inform.*, to appear.