

# Abelian complexity of infinite words

**Karel Břinda**

karel.brinda@jfifi.cvut.cz

**T**heoretical **I**nformatics **G**roup  
Faculty of Nuclear Sciences and Physical Engineering  
Czech Technical University in Prague

Analytic and algebraic methods in physics VII  
Prague, March 19, 2011

# Preliminaries

## Definition

**Abelian complexity** of an infinite word  $\mathbf{u}$  is a function

$$\mathcal{AC}(n) = \#\{\Psi(w) \mid w \in \mathcal{L}_n(\mathbf{u})\},$$

where  $\Psi(w) = (|w|_a)_{a \in \mathcal{A}}$  is **Parikh vector** of  $w$ .

## Definition

We define **balance function** of an infinite word  $\mathbf{u}$  as

$$B(n) = \max_{a \in \mathcal{A}} \max_{v, w \in \mathcal{L}_n(\mathbf{u})} \{|v|_a - |w|_a\}.$$

The word  $\mathbf{u}$  is  **$C$ -balanced** if  $B(n)$  is bounded by some  $C$ . We say that  $\mathbf{u}$  is **balanced** if it is 1-balanced.

## Example

$$\mathbf{u} = (0010)^\omega = 001000100010 \dots$$

Factors of length 3:  $\mathcal{L}_3(\mathbf{u}) = \{001, 010, 100, 000\}$ ,  $\mathcal{C}(3) = 4$

Parikh vectors:  $(2, 1), (2, 1), (2, 1), (3, 0)$ ,  $\mathcal{AC}(3) = 2$

Factors of length 4:  $\mathcal{L}_4(\mathbf{u}) = \{0010, 0100, 1000, 0001\}$ ,  $\mathcal{C}(4) = 4$

Parikh vectors:  $(3, 1), (3, 1), (3, 1), (3, 1)$ ,  $\mathcal{AC}(4) = 1$

# Properties of $\mathcal{AC}(n)$ and $B(n)$

- $1 \leq \mathcal{AC}(n) \leq \binom{n+d-1}{d-1}, \quad 0 \leq B(n) \leq n$
- $\mathcal{AC}(n)$  is bounded  $\Leftrightarrow B(n)$  is bounded
- $\mathcal{AC}(p) = 1$  for some  $p \Leftrightarrow \mathbf{u}$  is periodic of period  $p$

# Incidence matrix vs. balance function

## Theorem (Adamczewski 2003)

*Let  $\mathbf{u}$  be a fixed point of a primitive substitution  $\varphi$ . Let us denote the second biggest (in absolute value) eigenvalue of the incidence matrix of  $\varphi$  by  $\theta_2$ .  $|\theta_2| < 1$  if and only if  $B(n)$  is bounded.*

# Words over binary alphabets

# Words over binary alphabets

## Theorem

*Let  $\mathbf{u}$  be an infinite word over a binary alphabet. Then  $AC(n) = B(n) + 1$  for all  $n \in \mathbb{N}$ .*

# Sturmian words

## Definition

An infinite aperiodic word  $\mathbf{u}$  is **Sturmian**, if for all  $n \in \mathbb{N}$  it holds

$$\mathcal{C}(n) = n + 1.$$

(Necessarily  $\#\mathcal{A} = 2$ .)

## Theorem (Coven, Handlund 1973)

*Let  $\mathbf{u}$  be an infinite word over a binary alphabet. Then  $\mathbf{u}$  is Sturmian if and only if  $\mathbf{u}$  is aperiodic and  $\mathcal{AC}(n) = 2$  for all  $n \in \mathbb{N}$ .*

# Thue-Morse word

**Thue-Morse word**  $\text{TM}_0$  is a fixed point of the substitution

$$\varphi : \begin{cases} 0 & \rightarrow & 01 \\ 1 & \rightarrow & 10 \end{cases}$$

## Example (Thue-Morse)

$$\text{TM}_0 = 011010011001011010010110011010011001011001 \dots$$

## Theorem (Richomme, Saari, Zamboni 2009)

*The Abelian complexity of  $\text{TM}_0$  satisfies*

$$\mathcal{AC}(n) = \begin{cases} 2 & \text{for } n \text{ odd,} \\ 3 & \text{for } n \text{ even.} \end{cases}$$

# Quadratic non-simple Parry

$\mathbf{u}_\beta$  – **word associated with a quadratic non-simple Parry number  $\beta$**   
 – is the fixed point of the substitution

$$\varphi : \begin{cases} 0 & \rightarrow & 0^p 1 \\ 1 & \rightarrow & 0^q 1 \end{cases}, \text{ where } p > q \geq 1$$

## Theorem (Balková, B., Turek 2011)

Let  $(U_n)_{n=0}^\infty$  be a sequence such that  $U_n = |\varphi^n(0)|$ . Then

$$AC(n) = 1 + U_k - \sum_{j=1}^k (d_j + e_j) U_{j-1},$$

where

- $k$  is "enough big",
- $\langle n \rangle_U = (d_k, d_{k-1}, \dots, d_1, d_0)$ ,
- $\langle U_{k+1} - n \rangle_U = (e_k, e_{k-1}, \dots, e_1, e_0)$ .

# Quadratic simple Parry

$\mathbf{u}_\beta$  – **word associated with a quadratic simple Parry number**  $\beta$  – is the fixed point of the substitution

$$\varphi : \begin{cases} 0 & \rightarrow & 0^p 1 \\ 1 & \rightarrow & 0^q \end{cases}, \text{ where } p \geq q \geq 1$$

# Quadratic simple Parry

## Theorem (Balková, B., Turek 2011)

Let  $q > 1$ , and let  $J \in \mathbb{N}_0$  satisfy  $U_J \leq n < U_{J+1}$ , where  $(U_n)_{n=0}^\infty$  is a sequence such that  $U_n = |\varphi^n(0)|$ .

- If  $J$  is even: put  $N := \frac{J}{2}$ , and put  $M := \frac{J}{2}$  if  $|v^{(\frac{J}{2})}| \leq n$  and  $M := \frac{J}{2} - 1$  otherwise.
- If  $J$  is odd: put  $M := \frac{J-1}{2}$ , and put  $N := \frac{J+1}{2}$  if  $|w^{(\frac{J+1}{2})}| \leq n$  and  $N := \frac{J-1}{2}$  otherwise.

Then  $\mathcal{AC}(n) = 2 + (1, 0) \left( (q-1) \left[ (M_\varphi + I)^{-1} (M_\varphi^{2N} - I) - (M - N + 1) M_\varphi^{2N} \right] + \sum_{i=0}^J (c_i - d_i) M_\varphi^i \right) \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ , where

$$\begin{aligned} \langle n - |w^{(N)}| \rangle_U &= \langle n - 1 - (q-1) \sum_{i=0}^{N-1} U_{2i+1} \rangle_U &= (c_J, c_{J-1}, \dots, c_1, c_0), \\ \langle n - |v^{(M)}| \rangle_U &= \langle n - 1 - (q-1) \sum_{i=0}^M U_{2i} \rangle_U &= (d_J, d_{J-1}, \dots, d_1, d_0). \end{aligned}$$

# Words over multiliteral alphabets

# Tribonacci word

**Tribonacci word**  $\mathbf{t}$  is the fixed point of the substitution

$$\varphi : \begin{cases} 0 & \rightarrow & 01 \\ 1 & \rightarrow & 02 \\ 2 & \rightarrow & 0 \end{cases}$$

## Example (Tribonacci)

$\mathbf{t} = 0102010010201010201001020102010010201010201001 \dots$

# Properties of Tribonacci word

## Theorem (Richomme, Saari, Zamboni 2009)

*The Tribonacci word  $\mathfrak{t}$  is 2-balanced.*

## Theorem (Richomme, Saari, Zamboni 2009)

*For the Tribonacci word  $\mathfrak{t}$ , the Abelian complexity  $\mathcal{AC}(n) \in \{3, 4, 5, 6, 7\}$  for all  $n \in \mathbb{N}$ . All these values are attained and, moreover, 3 and 7 are attained infinitely many times.*

# K-bonacci word

**K-bonacci word** is the fixed point of the substitution

$$\varphi : \left\{ \begin{array}{l} 0 \rightarrow 01 \\ 1 \rightarrow 02 \\ \vdots \\ (K-2) \rightarrow 0(K-1) \\ (K-1) \rightarrow 0 \end{array} \right.$$

## Example (5-bonacci)

0102010301020104010201030102010010201030102010401020...

# Cubic Pisot

$\mathbf{u}_\beta$  – **word associated with a cubic Pisot number**  $\beta$  – is the fixed point of the substitution

$$\varphi : \begin{cases} 0 & \rightarrow & 0^p 1 \\ 1 & \rightarrow & 2 \\ 2 & \rightarrow & 0^{p-1} 1 \end{cases}, \text{ where } p > 1$$

## Theorem (Turek 2010)

*The word  $\mathbf{u}_\beta$  is 3-balanced. Its Abelian complexity satisfies  $\mathcal{AC}(n) \in \{3, 4, 5, 6, 7\}$  for all  $n \in \mathbb{N}$ . Moreover,  $\mathcal{AC}(n) = 7$  is attained for infinitely many  $n$ .*

# Conclusion

# Open questions

- **Tribonacci**

- Does  $\mathcal{AC}(n)$  attain the values 4, 5, 6 infinitely many times?
- Characterize those  $n$  for which  $\mathcal{AC}(n) = m$ , where  $m \in \{4, 5, 6, 7\}$ .

- **$K$ -bonacci**

- Prove or disprove that it is  $(K - 1)$ -balanced.
- Everything about  $\mathcal{AC}(n)$ .

# Summary of the presentation

- Defined and shown the connection between
  - Abelian complexity  $\mathcal{AC}(n)$
  - Balance function  $B(n)$
- Binary words
  - Sturmian
  - Thue-Morse
  - Simple Parry
  - Non-simple Parry
- Words over multiliteral alphabets
  - Tribonacci
  - $K$ -bonacci
  - Cubic Pisot
- Open questions

**Thank you for your attention!**

# Reference



B. Adamczewski.

Balances for fixed points of primitive substitutions.

*THEORETICAL COMPUTER SCIENCE*, 307(1):47–75, SEP 26 2003.

3rd Conference on WORDS, PALERMO, ITALY, SEP, 2001.



Ľ. Balková, K. Břinda, and O. Turek.

Abelian Complexity of Infinite Words Associated with Quadratic Parry Numbers.

*ArXiv e-prints*, Jan. 2011.



G. Richomme, K. Saari, and L. Q. Zamponi.

Abelian Complexity in Minimal Subshifts.

*ArXiv e-prints*, Nov. 2009.



G. Richomme, K. Saari, and L. Q. Zamponi.

Balance and Abelian complexity of the Tribonacci word.

*ADVANCES IN APPLIED MATHEMATICS*, 45(2):212–231, AUG 2010.



O. Turek.

Balances and Abelian Complexity of a Certain Class of Infinite Ternary Words.

*ArXiv e-prints*, Mar. 2010.