# Variations of the Morse-Hedlund Theorem for $k$-Abelian Equivalence

Juhani Karhumäki Aleksi Saarela Luca Q. Zamboni

Department of Mathematics and Statistics
and FUNDIM Centre,
University of Turku, Finland

DLT 2014
Ekaterinburg
29.8.2014

# Outline


1. Background
2. Morse–Hedlund
3. Period-Doubling Word
4. Thue–Morse Word
5. Complexity Gap
6. Other Results

# $k$-abelian equivalence

Let $k \geq 1$. Words $u, v$ are *$k$-abelian equivalent* if

- $|u|_t = |v|_t$ for all words $t$ such that $|t| \leq k$

or, equivalently, if

- $\mathrm{pref}_{k-1}(u) = \mathrm{pref}_{k-1}(v)$ and $|u|_t = |v|_t$ for all words $t$ of length $k$.

The $k$-abelian equivalence class of $u$ is denoted by $[u]_k$.

**Example**

$$[001011]_2 = \{001011, 010011, 001101, 011001\}$$

- 1-abelian equivalence is abelian equivalence.
- "$\infty$-abelian" equivalence is equality.

# Factor complexity

Let $\mathcal{F}_n(w)$ be the set of factors of $w$ of length $n$.

The *factor complexity* of $w$ is the function

$$\mathcal{P}_w : \mathbb{N}_1 \to \mathbb{N}_1, \mathcal{P}_w(n) = \#\mathcal{F}_n(w).$$

The *k-abelian complexity* of $w$ is the function

$$\mathcal{P}_w^k : \mathbb{N}_1 \to \mathbb{N}_1, \mathcal{P}_w^k(n) = \#\left\{[u]_k \mid u \in \mathcal{F}_n(w)\right\}.$$

# Outline


1. Background
2. Morse–Hedlund
3. Period-Doubling Word
4. Thue–Morse Word
5. Complexity Gap
6. Other Results

# Equality

## Theorem (Morse and Hedlund)

- $\mathcal{P}_w(n) < n + 1$ *for some* $n \Leftrightarrow w$ *ult. per.* $\Leftrightarrow \mathcal{P}_w(n) = O(1)$
- $\mathcal{P}_w(n) = n + 1$ *for all* $n \Leftrightarrow w$ *Sturmian*

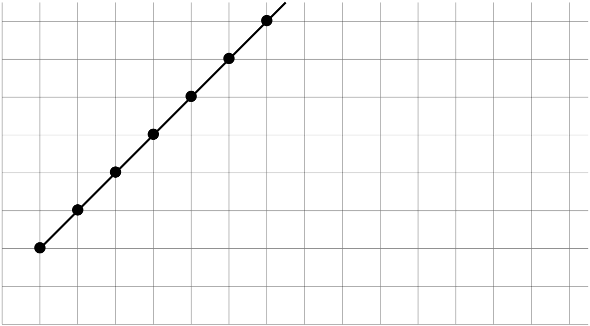


Figure: The factor complexity of Sturmian words.

# Abelian equivalence

## Theorem (Coven and Hedlund)

- $\mathcal{P}^1_w(n) < 2$ *for some* $n \Leftrightarrow w$ *per.* $\Rightarrow w$ *ult. per.* $\Rightarrow \mathcal{P}^1_w(n) = O(1)$
- $\mathcal{P}^1_w(n) = 2$ *for all* $n$ *and* $w$ *aper.* $\Leftrightarrow w$ *Sturmian*

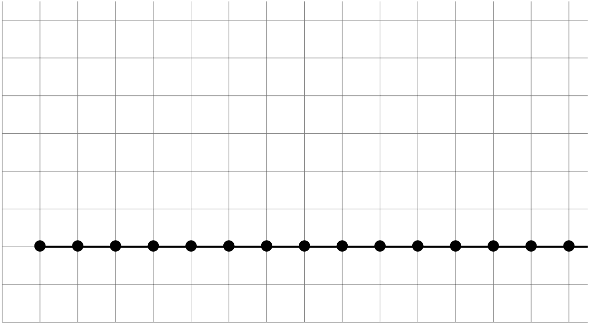


Figure: The abelian complexity of Sturmian words.

# $k$-abelian equivalence

**Theorem (Karhumäki, Saarela and Zamboni)**

- $\mathcal{P}_w^k(n) < \min(2k, n+1)$ *for some* $n$ $\Rightarrow$ *w ult. per.* $\Rightarrow$ $\mathcal{P}_w^k(n) = O(1)$
- $\mathcal{P}_w^k(n) = \min(2k, n+1)$ *for all* $n$ *and* $w$ *aper.* $\Leftrightarrow$ *w Sturmian*

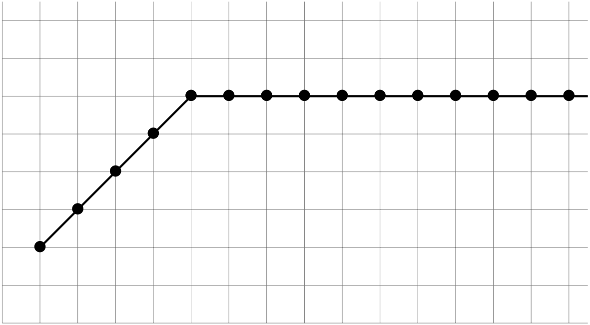


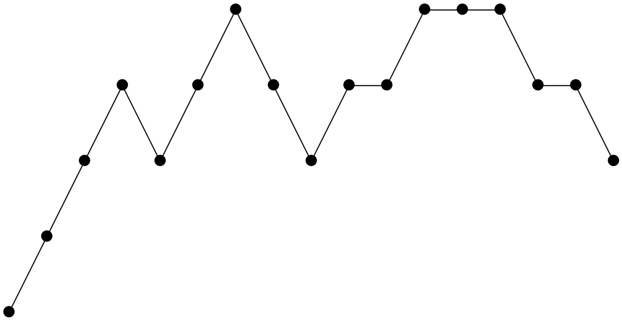
Figure: The 3-abelian complexity of Sturmian words.

# Gap theorems

- For ordinary factor complexity, there is a gap between bounded complexity and complexity $n + 1$.
- For abelian or $k$-abelian complexity, there is no such gap, as will be seen later.

# Upper and lower complexities

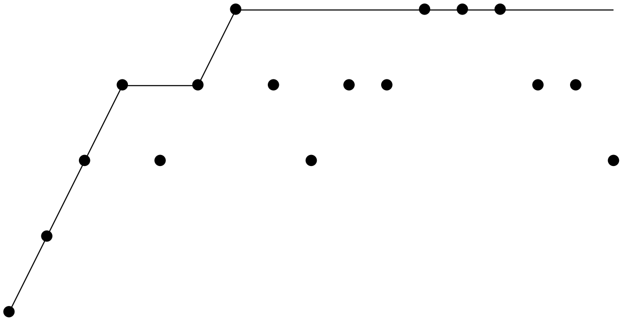
$\mathcal{P}_w^k$ can decrease



# Upper and lower complexities

$\mathcal{P}_w^k$ can decrease, but the following functions are increasing:

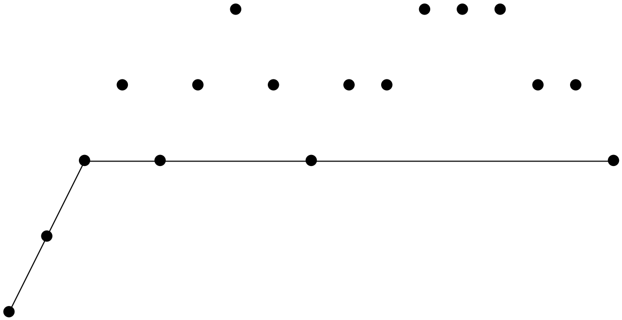
- *Upper k-abelian complexity* $\mathcal{U}_w^k(n) = \max_{m \leq n} \mathcal{P}_w^k(m)$.



# Upper and lower complexities

$\mathcal{P}_w^k$ can decrease, but the following functions are increasing:

- *Upper k-abelian complexity* $\mathcal{U}_w^k(n) = \max_{m \leq n} \mathcal{P}_w^k(m)$.
- *Lower k-abelian complexity* $\mathcal{L}_w^k(n) = \min_{m \geq n} \mathcal{P}_w^k(m)$.



# Outline

1 Background

2 Morse–Hedlund

3 Period-Doubling Word

4 Thue–Morse Word

5 Complexity Gap

6 Other Results

# Period-doubling word

Let $\sigma$ be the morphism defined by $\sigma(0) = 01, \sigma(1) = 00$. Let

$$S = 0100010101000100010001010100010 1\cdots$$

be the *period-doubling word*, which is the fixed point of $\sigma$.

# Recursion formulas

(Proved independently by Blanchet-Sadri, Currie, Rampersad, and Fox.)

**Lemma**

*For* $n \geq 1$, $\mathcal{P}_S^1(2n) = \mathcal{P}_S^1(n)$ *and* $\mathcal{P}_S^1(4n \pm 1) = \mathcal{P}_S^1(n) + 1$.

**Proof.**

Based on the formula

$$\mathcal{P}_w^1(n) = \max\{|u|_1 \mid u \in \mathcal{F}_n(w)\} - \min\{|u|_1 \mid u \in \mathcal{F}_n(w)\} + 1$$

for binary words $w \in \{0,1\}^\omega$. □

# Complexity

## Theorem

*For $n \geq 1$ and $m \geq 0$,*

$$\mathcal{P}_S^1(n) = O(\log n), \quad \mathcal{P}_S^1((2 \cdot 4^m + 1)/3) = m + 2, \quad \mathcal{P}_S^1(2^m) = 2.$$

The abelian complexity of $S$ has a logarithmic upper bound and a constant lower bound. These bounds are the best possible increasing bounds.

## Corollary

$\mathcal{U}_S^1(n) = \Theta(\log n)$ *and* $\mathcal{L}_S^1(n) = 2$.

# Outline


1. Background
2. Morse–Hedlund
3. Period-Doubling Word
4. Thue–Morse Word
5. Complexity Gap
6. Other Results

# Thue–Morse word

Let $\tau$ be the Thue–Morse morphism defined by $\tau(0) = 01, \tau(1) = 10$. Let

$$T = 01101001100101101001011001101001 \cdots$$

be the Thue–Morse word, which is a fixed point of $\tau$.

The first values of $\mathcal{P}_T^2$ are

$$2, 4, 6, 8, 6, 8, 10, 8, 6, 8, 8, 10, 10, 10, 8, 8, 6, 8, 10, 10.$$

# Complexity

## Lemma

*For* $n \geq 2$, $\mathcal{P}_S^1(n-1) \leq \mathcal{P}_T^2(n) \leq 4\mathcal{P}_S^1(n-1)$.

## Theorem

*For* $n \geq 1$ *and* $m \geq 1$,

$$\mathcal{P}_T^2(n) = O(\log n), \quad \mathcal{P}_T^2((2 \cdot 4^m + 4)/3) = \Theta(m), \quad \mathcal{P}_T^2(2^m + 1) = 6.$$

## Corollary

*Let* $k \geq 2$. *Then* $\mathcal{U}_T^k(n) = \Theta(\log n)$ *and* $\mathcal{L}_T^k(n) = \Theta(1)$.

# Outline

1 Background

2 Morse–Hedlund

3 Period-Doubling Word

4 Thue–Morse Word

5 Complexity Gap

6 Other Results

# Complexity gap

Questions:

- Does there exist an increasing unbounded function $f : \mathbb{N}_1 \to \mathbb{N}_1$ such that for every infinite word $w$ either $\mathcal{P}_w^k = O(1)$ or $\mathcal{P}_w^k = \Omega(f)$?
- Does there exist an increasing unbounded function $f : \mathbb{N}_1 \to \mathbb{N}_1$ such that for every infinite word $w$ either $\mathcal{P}_w^k = O(1)$ or $\mathcal{P}_w^k \neq O(f)$?

Answers:

- The first question has already been answered negatively.
- The answer to the second question is also negative, even if only uniformly recurrent words are considered.

# Construction

Let $n_1, n_2, \ldots$ be a sequence of integers greater than 1.

$$
\begin{aligned}
U_0 &= \varepsilon \\
U_1 &= (U_0 0)^{n_1 - 1} U_0 = 0^{n_1 - 1} \\
U_2 &= (U_1 1)^{n_2 - 1} U_1 \\
U_3 &= (U_2 0)^{n_3 - 1} U_2 \\
U_4 &= (U_3 1)^{n_4 - 1} U_2 \\
&\vdots
\end{aligned}
$$

Let $U$ be the limit.

The faster the sequence $n_1, n_2, \ldots$ grows, the slower $\mathcal{P}_U^k(n)$ grows.

# Construction

## Example

If $n_i = i + 1$ for all $i$, then

$$U = 01010001010001010001010 1\ldots$$

```
0 0 0 0 0 0 0 0 0 0 0 0
 1 1   1 1   1 1   1 1
   0     0     0
                     1
```

# Result

## Theorem

*For every increasing unbounded function* $f : \mathbb{N}_1 \to \mathbb{N}_1$ *there is a uniformly recurrent word* $w \in \{0,1\}^\omega$ *such that* $\mathcal{P}_w^k(n) = O(f(n))$ *but* $\mathcal{P}_w^k(n) \neq O(1)$.

Thus there are no gaps in the $k$-abelian case!

# Outline

1 Background

2 Morse–Hedlund

3 Period-Doubling Word

4 Thue–Morse Word

5 Complexity Gap

6 Other Results

# Other results

There are words having bounded $k$-abelian complexity but linear $(k+1)$-abelian complexity. These words can even be assumed to be $k$-abelian periodic, meaning that they are of the form $u_1 u_2 \cdots$, where $u_1, u_2, \ldots$ are $k$-abelian equivalent.

**Theorem**

*For every $k \geq 1$ there is a $k$-abelian periodic word $w$ such that $\mathcal{P}_w^{k+1}(n) = \Theta(n)$.*

**Proof.**

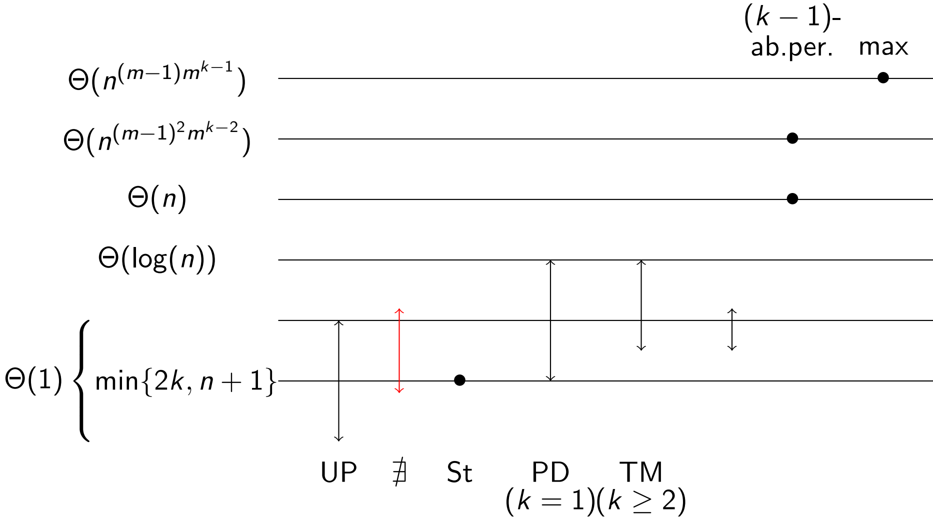
Let $w \in \{0,1\}^\omega$ be a word with linear abelian complexity (e.g., the Champernowne word) and let $h$ be the morphism defined by $h(0) = 0^{k+1}10^{k-1}1$, $h(1) = 0^k10^k1$. Then the word $h(w)$ is $k$-abelian periodic and $\mathcal{P}_{h(w)}^{k+1}((2k+2)n) = \Theta(\mathcal{P}_w^1(n)) = \Theta(n)$. If $m$ is the size of the alphabet, then $\mathcal{P}_{h(w)}^{k+1}(n+1) \leq m\mathcal{P}_{h(w)}^{k+1}(n)$ for all $n$, so the claim follows. □

# Other results

**Lemma (Cassaigne, Karhumäki and Saarela, preprint)**

*For every $k \geq 1$ there is a $k$-abelian periodic word $w$ such that $\mathcal{P}_w^{k+1}(n) = \Theta(n^{(m-1)^2 m^{k-1}})$, where $m$ is the size of the alphabet.*

# Summary


$(k-1)$-
ab.per.
max
$\Theta(n^{(m-1)m^{k-1}})$
$\Theta(n^{(m-1)^2 m^{k-2}})$
$\Theta(n)$
$\Theta(\log(n))$
$\Theta(1)$
$\min\{2k, n+1\}$
UP
$\nexists$
St
PD
TM
$(k=1)(k\geq 2)$

# Thank You!