# Regular subsequences in finite words

Andrew Ryzhikov

LIGM, Université Paris-Est

Words and Complexity Workshop
Lyon, 21 February 2018

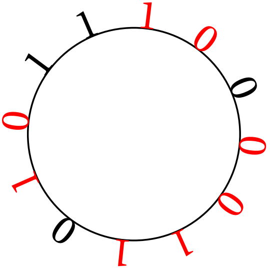
# Definitions: linear words

A *linear word* (or just a word) is a finite sequence of symbols over some alphabet. Ex.: 01110011.

A *subsequence* of a linear word $w = a_1 \ldots a_n$ is a word $w' = a_{i_1} \ldots a_{i_m}$ with $i_1 \leq \ldots \leq i_m$. Ex.: 0001 is a subsequence of 01110011.

# Definitions: circular words

A *circular word* is a class of equivalence of linear words under rotations.
A linear word is a *subsequence* of a circular word if it is a subsequence of some linear word from the corresponding equivalence class (some its *linear representation*).




10001110 is a subsequence.

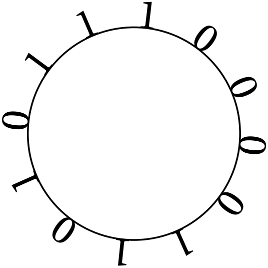
# Definitions: palindromes and antipalindromes

A word $w = a_1 \dots a_n$ is a *palindrome* if $a_i = a_{n-i}$ for any $1 \leq i \leq \frac{n}{2}$.
Ex.: 1000001.
A binary word $w = a_1 \dots a_n$ is an *antipalindrome* if $a_i \neq a_{n-i}$ for any $1 \leq i \leq \frac{n}{2}$. Ex.: 10001110.
The *reversal* $w^R$ of a word $w = a_1 \dots a_n$ is the word $a_n \dots a_1$. Ex.: $000111^R = 111000$.

# Weak antipalindromic conjecture

**Conjecture (Lyngsø and Pedersen, 1999)**

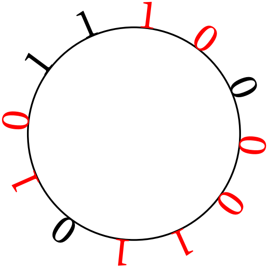
Every binary circular word of length $n$ divisible by $6$ with equal number of zeros and ones has an antipalindromic subsequence of length at least $\frac{2}{3}n$.



# Weak antipalindromic conjecture

**Conjecture (Lyngsø and Pedersen, 1999)**

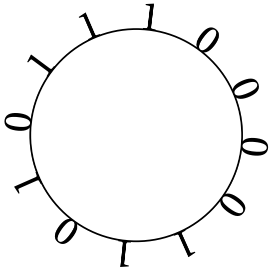
Every binary circular word of length $n$ divisible by $6$ with equal number of zeros and ones has an antipalindromic subsequence of length at least $\frac{2}{3}n$.



# Strong antipalindromic conjecture

## Conjecture (Brevier, Preissmann and Sebő)

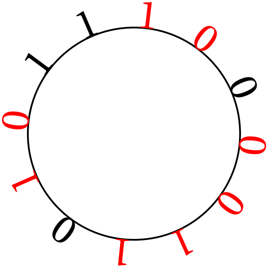
Let $w$ be a binary circular word of length $n$ divisible by $6$ with equal number of zeros and ones. Then $w$ can be partitioned into two linear words $w_1, w_2$ of equal length, $w = w_1 w_2$, having subsequences $s_1, s_2$ such that $s_1 s_2$ is an antipalindrome and $|s_1 s_2| = \frac{2}{3}|w|$.



# Strong antipalindromic conjecture

## Conjecture (Brevier, Preissmann and Sebő)

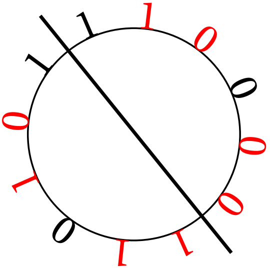
Let $w$ be a binary circular word of length $n$ divisible by 6 with equal number of zeros and ones. Then $w$ can be partitioned into two linear words $w_1, w_2$ of equal length, $w = w_1 w_2$, having subsequences $s_1, s_2$ such that $s_1 s_2$ is an antipalindrome and $|s_1 s_2| = \frac{2}{3}|w|$.



# Strong antipalindromic conjecture

## Conjecture (Brevier, Preissmann and Sebő)

Let $w$ be a binary circular word of length $n$ divisible by 6 with equal number of zeros and ones. Then $w$ can be partitioned into two linear words $w_1, w_2$ of equal length, $w = w_1 w_2$, having subsequences $s_1, s_2$ such that $s_1 s_2$ is an antipalindrome and $|s_1 s_2| = \frac{2}{3}|w|$.

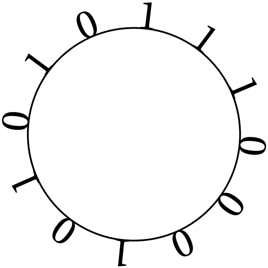


Checked up to $n = 30$. The word $w = 0^i 1^i (01)^i 1^i 0^i$ shows tightness.

# Weak palindromic conjecture

## Conjecture

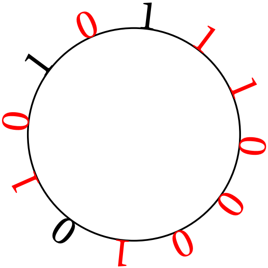
Every binary circular word of length $n$ has a palindromic subsequence of length at least $\frac{3}{4}n$.



# Weak palindromic conjecture

## Conjecture

Every binary circular word of length $n$ has a palindromic subsequence of length at least $\frac{3}{4}n$.

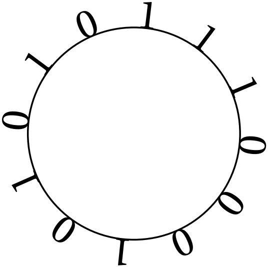


Checked up to $n = 30$. The word $0^i(10)^i1^i$ shows tightness.

# Strong palindromic conjecture

## Conjecture

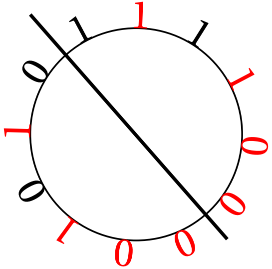
Let $w$ be a binary circular word of length $n$ divisible by $6$. Then $w$ can be partitioned into two linear words $w_1, w_2$ of equal length, $w = w_1 w_2$, having subsequences $s_1, s_2$ such that $s_1 = s_2^R$ (that is, $s_1 s_2$ is a palindrome) and $|s_1 s_2| = \frac{2}{3}|w|$.



# Strong palindromic conjecture

## Conjecture

Let $w$ be a binary circular word of length $n$ divisible by 6. Then $w$ can be partitioned into two linear words $w_1, w_2$ of equal length, $w = w_1 w_2$, having subsequences $s_1, s_2$ such that $s_1 = s_2^R$ (that is, $s_1 s_2$ is a palindrome) and $|s_1 s_2| = \frac{2}{3}|w|$.

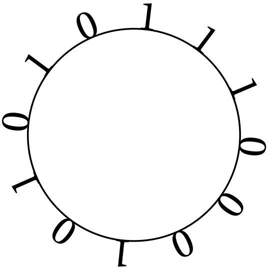


Checked up to $n = 30$. The word $0^{2i}(10)^i 1^{2i}$ shows tightness.

# Two cuts conjecture

## Conjecture

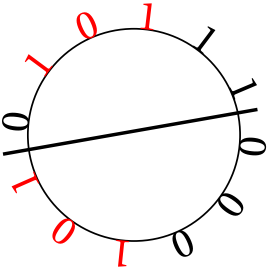
For any word $w$ of length $n$ divisible by 4 and any its linear representation $w = w_1 w_2 w_3 w_4$, the maximum of the lengths of longest palindromic subsequences of $w_1 w_2 | w_3 w_4$ and $w_2 w_3 | w_4 w_1$ is at least $\frac{1}{2} n$.



# Two cuts conjecture

## Conjecture

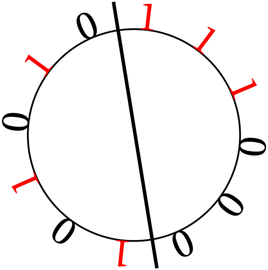
For any word $w$ of length $n$ divisible by 4 and any its linear representation $w = w_1 w_2 w_3 w_4$, the maximum of the lengths of longest palindromic subsequences of $w_1 w_2 | w_3 w_4$ and $w_2 w_3 | w_4 w_1$ is at least $\frac{1}{2} n$.



# Two cuts conjecture

## Conjecture

For any word $w$ of length $n$ divisible by 4 and any its linear representation $w = w_1 w_2 w_3 w_4$, the maximum of the lengths of longest palindromic subsequences of $w_1 w_2 | w_3 w_4$ and $w_2 w_3 | w_4 w_1$ is at least $\frac{1}{2} n$.



Checked up to $n = 30$. The word $0^i (10)^i 1^i$ shows tightness.

# Linear words

## Proposition

*Every binary word of length $n$ without three consecutive equal letters has a palindromic subsequence of length at least $\frac{2}{3}(n-2)$. The same is true for an antipalindromic subsequence.*

001001011011

# Linear words

## Proposition

*Every binary word of length $n$ without three consecutive equal letters has a palindromic subsequence of length at least $\frac{2}{3}(n-2)$. The same is true for an antipalindromic subsequence.*

001001011011

# Linear words

## Conjecture

Let $w$ be a word of length $n$ over an alphabet of size $k$, $k \geq 2$. If $w$ has no two consecutive equal letters, then it has a palindromic subsequence of length at least $\frac{1}{k-1}(n-1)$.

0101010101
01010121212
010101212123232 3

# Linear words

## Conjecture

Let $w$ be a word of length $n$ over an alphabet of size $k$, $k \geq 2$. If $w$ has no two consecutive equal letters, then it has a palindromic subsequence of length at least $\frac{1}{k-1}(n-1)$.

0101010101
01010121212
0101012121232323

Proved for $k = 2, 3$. Checked up to $n = 21$ for $k = 4$ and $n = 18$ for $k = 5$. The word which is a concatenation of the word $(a_1 a_2)^i$ and words $(a_{\ell+1} a_\ell)^{i-1} a_{\ell+1}$ for $1 < \ell < k-1$ shows tightness.

# Linear words

## Conjecture

Let $w$ be a word of length $n$ over an alphabet of size $k$, $k \geq 2$. If $w$ has no two consecutive equal letters, then it has a palindromic subsequence of length at least $\frac{1}{k-1}(n-1)$.

0101010101
01010121212
01010121212323

Proved for $k = 2, 3$. Checked up to $n = 21$ for $k = 4$ and $n = 18$ for $k = 5$. The word which is a concatenation of the word $(a_1 a_2)^i$ and words $(a_{\ell+1} a_\ell)^{i-1} a_{\ell+1}$ for $1 < \ell < k-1$ shows tightness.

# Future work

- Structure of extremal examples (there are more of them).
- Circular case for non-binary alphabet.
- Non-binary antipalindromes?
- Relations between the conjectures.
- Any other questions about the combinatorics of subsequences (almost nothing is known).

# Future work

- Structure of extremal examples (there are more of them).
- Circular case for non-binary alphabet.
- Non-binary antipalindromes?
- Relations between the conjectures.
- Any other questions about the combinatorics of subsequences (almost nothing is known).

# Future work

- Structure of extremal examples (there are more of them).
- Circular case for non-binary alphabet.
- Non-binary antipalindromes?
- Relations between the conjectures.
- Any other questions about the combinatorics of subsequences (almost nothing is known).

# Future work

- Structure of extremal examples (there are more of them).
- Circular case for non-binary alphabet.
- Non-binary antipalindromes?
- Relations between the conjectures.
- Any other questions about the combinatorics of subsequences (almost nothing is known).

# Future work

- Structure of extremal examples (there are more of them).
- Circular case for non-binary alphabet.
- Non-binary antipalindromes?
- Relations between the conjectures.
- Any other questions about the combinatorics of subsequences (almost nothing is known).

Thank you! Questions?
(Or answers?)