Abelian bordered factors and periodicity

Émilie Charlier

(Joint work with Tero Harju, Svetlana Puzynina and Luca Zamboni)

Discrete Mathematics Day - Liège - 2015, January 8th

Bordered words and periodicity

- A finite word is bordered if it has a proper prefix equal to a suffix.
- For example, *aababbaaab* is bordered.
- If a word is not bordered, we say it is unbordered or border-free.
- For example, $10^n = 100\cdots0$ are unbordered.
- A periodic infinite word: $w = uv^\omega = uvvv\cdots$.
- A purely periodic infinite word: $w = v^\omega = vvv\cdots$.

Theorem (Ehrenfeucht-Silberger 1979)

An infinite word is purely periodic iff every sufficiently long factor is bordered.

We study both abelian and weak-abelian versions of this result.

Abelian bordered words

- Finite words u and v are abelian equivalent if $|u|_a = |v|_a$ for all letters a.
- A finite word is abelian bordered if it has a proper prefix abelian equivalent to a suffix.
- For example, *abababbaabb* is abelian bordered.
- If a word is not abelian bordered, we say it is abelian unbordered or abelian border-free.
- For example, *abababbaabbb* is abelian unbordered.

Abelian borders in binary words [Christodoulakis-Christou-Crochemore-Iliopoulos 2014]

On the number of abelian bordered words [Rampersad-Rigo-Salimov 2013]

Link with abelian periodicity ?

- An abelian periodic infinite word: $w = uv_1 v_2 v_3 \cdots$ where the v_i's are all abelian equivalent.
- The Thue-Morse word

$$w = 0110100110010110\cdots$$

abelian periodic with period 2. (It is the fixed point of the morphism $0 \mapsto 01$, $1 \mapsto 10$.)

Finitely many abelian unbordered factors	$\overset{?}{\longleftrightarrow}$ relation	Abelian periodicity

Abelian periodicity is not a sufficient condition for having finitely many abelian unbordered factors.

For example, the Thue-Morse word $0110100110010110 \cdots$ is

- abelian periodic with period 2
- has infinitely many abelian unbordered factors: $0p1$ where p is a palindrome. E.g. take $p = \mu^{2n}(0)$.

We even have:

Proposition

If a uniformly recurrent aperiodic word contains infinitely many palindromes, then it admits an infinite number of abelian unbordered factors.

Having finitely many abelian unbordered factors is not a sufficient condition for periodicity

Proposition

There exists an infinite aperiodic word w and constants C, D such that every factor v of w with $|v| \geq C$ has an abelian border of length at most D.

For example, any aperiodic infinite word

$$w \in \{010100110011, 0101001100110011\}^\omega$$

satisfies the condition with $C = 15$ and $D = 14$.

Is it a sufficient condition for abelian periodicity?

Open question 1

Let w be an infinite word with finitely many abelian unbordered factors. Does it follow that w is abelian periodic?

We do not know. But we are able to answer the question in a weak abelian setting.

Weak abelian periodicity

- The frequency $\rho_b(w)$ of a letter b in a finite word w is defined as $\rho_b(w) = \frac{|w|_b}{|w|}$.
- A weakly abelian periodic (WAP) infinite word: $w = uv_1v_2v_3\cdots$ where the v_i's have the same letter frequencies.
- w is called bounded weakly abelian periodic, if it is WAP with bounded lengths of blocks: $\exists C\ \forall i\ |v_i| \leq C$.

On certain sequences of lattice points [Gerver-Ramsey 1979]
Words that do not contain consecutive factors with equal frequencies of letters [Krainev 1980]

Weak abelian periodicity of infinite words [Avgustinovich-Puzynina 2013]

Geometric interpretation of WAP

Aim: $w = a_1 a_2 \cdots \in \Sigma^\omega \mapsto$ graph $G_w \subseteq \mathbb{Z}^{|\Sigma|}$

- In the binary case:

$$0 \to \text{move along } \mathbf{v}_0, \text{ e.g. } \mathbf{v}_0 = (1, -1)$$
$$1 \to \text{move along } \mathbf{v}_1, \text{ e.g. } \mathbf{v}_1 = (1, 1)$$

  Start at the origin $(x_0, y_0) = (0, 0)$.

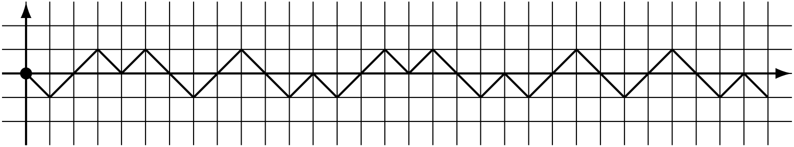
  At step $n > 0$: $(x_n, y_n) = (x_{n-1}, y_{n-1}) + \mathbf{v}_{a_n}$.

  Then (x_{n-1}, y_{n-1}) and (x_n, y_n) are connected by a line segment.

- For a k-letter alphabet one can consider a similar graph in $\mathbb{Z}^k$.

WAP infinite words correspond to a graph with infinitely many integer points on a line with a rational slope.

Bounded WAP infinite words correspond to a graph with infinitely many integer points on a line with a rational slope with bounded gaps.



The graph of the Thue-Morse word with $\mathbf{v}_0 = (1, -1)$, $\mathbf{v}_1 = (1, 1)$.

The paper-folding word:

$$0010011000110110001001110011011 0\cdots$$

The paper-folding word:

$$0010011000110110001001110011011 0 \cdots$$

It can be defined as a Toeplitz word with pattern 0?1?:

$$0?1?0?1?0?1?0?1?0?1?0?1?0?1?0?1? \ldots$$

The paper-folding word:

$$0010011000110110001001110011011 0\cdots$$

It can be defined as a Toeplitz word with pattern 0?1?:

$$001?011?001?011?001?011?001?011?\ldots$$

The paper-folding word:

$$0010011000110110001001110011011 0\cdots$$

It can be defined as a Toeplitz word with pattern 0?1?:

$$0010011?0011011?0010011?0011011?\ldots$$

The paper-folding word:

$$0010011000110110001001110011011 \cdots$$

It can be defined as a Toeplitz word with pattern 0?1?:

$$0010011000110111?0010011100110111? \ldots$$

The paper-folding word:

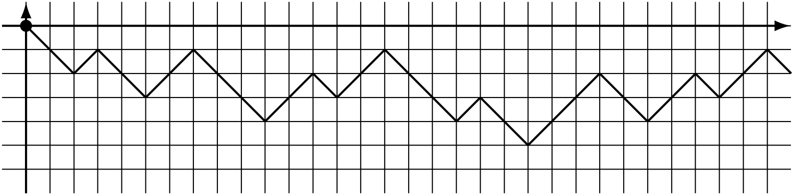
$$00100110001101100010011100110110\cdots$$

It can be defined as a Toeplitz word with pattern 0?1?:

$$00100110001101100010011100110110\ldots$$

The paper-folding word:

$$001001100011011000100111001101100 \cdots$$



The graph of the paperfolding word with $\mathbf{v}_0 = (1, -1)$, $\mathbf{v}_1 = (1, 1)$.
It is WAP along the line $y = -1$ (and actually along any line $y = C$, $C = -1, -2, \dots$).

Weak abelian borders

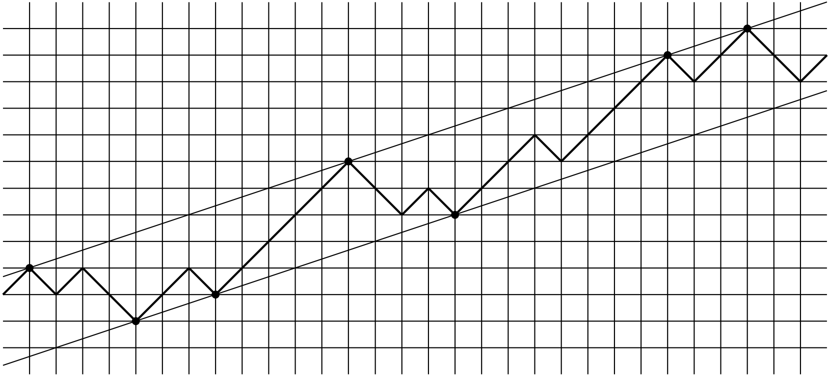
- A finite word w is weakly abelian bordered if it has a proper prefix and a suffix with the same letter frequencies.
- For example, *abaabbaabb* is weakly abelian bordered while *abaabbaabbb* isn't.

Finitely many weakly abelian unbordered factors $\overset{?}{\longleftrightarrow}$ Weak abelian periodicity

relation

Bounded width

If the graph of a word lies between two lines, then we say the word is of bounded width.



Bounded width is equivalent to balance

An infinite word w is K-balanced if for each letter a and two factors u and v of equal length, the inequality $||u|_a - |v|_a| \leq K$ holds. It is balanced if it is K-balanced for some K.

Lemma

An infinite word is balanced iff it is of bounded width.

Main theorem: Having finitely many weakly abelian unbordered factors implies WAP

Theorem

Let w be an infinite binary word. If there exists a constant C such that every factor v of w with $|v| \geq C$ is weakly abelian bordered, then w is bounded WAP.
Moreover, its graph lies between two rational lines and has points on each of these two lines with bounded gaps.

Is the converse true?

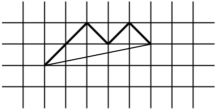
Open question 2

Let w be a bounded WAP infinite binary word such that its graph lies between two rational lines and has points on each of these two lines with bounded gaps.
Does w necessarily have only finitely many weakly abelian unbordered factors?

True for $\rho_0 = \rho_1 = \frac{1}{2}$.

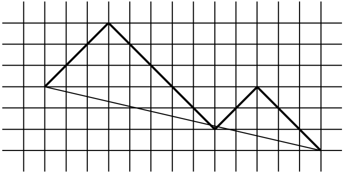
Lemma

Let w be an infinite binary word, and i, j be integers, $i < j$. If $g_w(k) > \frac{g_w(j)-g_w(i)}{j-i}k + \frac{g_w(i)j-g_w(j)i}{j-i}$ for each $i < k < j$, then the factor $w[i+1..j]$ is weakly abelian unbordered.



Weakly abelian unbordered factor 11010.

However, the reciprocal does not hold:



Weakly abelian unbordered factor 1110000011000.

Lemma

Suppose w have finitely many weakly abelian unbordered factors. Then w is of bounded width.

We derive a consequence in the abelian setting:

Proposition

Let w be an infinite word having only finitely many abelian unbordered factors. Then w has bounded abelian factor complexity.

Rational letter frequencies

Lemma

Let w have finitely many weakly abelian unbordered factors. Then the letter frequencies are rational.

Sketch of the proof

We know that balance implies the existence of letter frequencies.

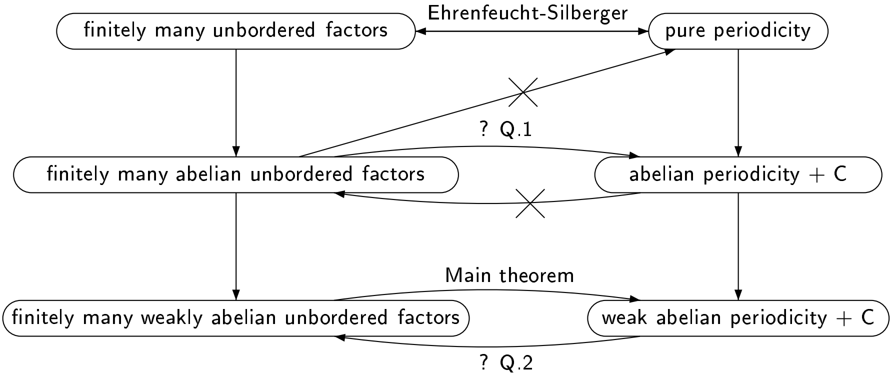
What we have to show is that they are rational.

Since w is of bounded width, i.e., there exist α, β', β'' such that the graph of w lies between two lines: $\alpha x + \beta' \leq g_w(x) \leq \alpha x + \beta''$.

Here α is related to the letter frequencies as follows: $\rho_0 = \frac{1-\alpha}{2}$ and $\rho_1 = \frac{1+\alpha}{2}$.

So, we need to prove that α is rational.

Results and open questions



C: The graph lies between two rational lines and has points on each of these two lines with bounded gaps.