On abelian versions of Critical Factorization Theorem

S. Avgustinovich, J. Karhumäki, S. Puzynina

University of Turku, Finland
Sobolev Institute of Mathematics, Novosibirsk, Russia
Periodicity

\[ \Sigma - \text{alphabet} \]
\[ \Sigma^* - \text{finite words over } \Sigma \]
\[ \Sigma^\omega - \text{right infinite words over } \Sigma \]
\[ \Sigma^\mathbb{Z} - \text{biinfinite words over } \Sigma \]

- A finite word \( w \) is **periodic**, if there exists an integer \( d \) \( 1 \leq d \leq |w| \), such that \( w_i = w_{i+d} \) for all \( 1 \leq i \leq |w| - d \).

- A biinfinite word \( w \) is **periodic**, if there exists a finite word \( v \) such that \( w = v^\mathbb{Z} \).

- A right infinite word \( w \) is **ultimately periodic**, if for some finite words \( u \) and \( v \) it holds \( w = uv^\omega \);

- A right infinite word \( w \) is **purely periodic** (or briefly **periodic**) if \( u = \epsilon \).
Critical Factorization Theorem for finite words

Critical Factorization Theorem: local periods vs global period

Y. Césari, M. Vincent (1978); J. Duval (1979)

- \( w = uv, \ |u| = p \)
- \( z \): the shortest suffix of \( w_1u \) and prefix of \( vw_2 \) for suitable \( w_1 \) and \( w_2 \)
- \( z \) is the shortest repetition word in \( w \) centered at position \( p \)
- if \( w_1 = w_2 = \epsilon \), then we have a "proper" square, otherwise we have "virtual" square

- CFT: in every word \( w \) there is a position \( p \) where the shortest repetition word \( z \) is as long as the global period \( d \) of \( w \). The position \( p \) is called critical.

- \( |z| \) is local period at position \( p \), then global period is maximum of local periods
Example

$w = abbab$

period: 3
Example

\[ w = a.bbab \]

period: 3
position 1, local period 3 (bba)
Example

\[ w = ab.bab \]

period: 3
position 2, local period 1 (b)
Example

\( w = abb.ab \)

period: 3
position 3, local period 3 (abb)
Example

\( w = abba.b \)

period: 3
position 4, local period 2 (ba)
Example

$w = abbab$

period: 3
critical positions: 1, 3
Consequences for infinite words

Theorem

A biinfinite word \( w \) is periodic if and only if there exists an integer \( l \) such that \( w \) has at every position a centered square with period at most \( l \).

Theorem

A right infinite word \( w \) is periodic if and only if there exists an integer \( l \) such that \( w \) has at every position a (virtual) centered square with period at most \( l \).

Theorem

A right infinite word \( w = w_1w_2 \ldots \) is ultimately periodic if and only if there exists an integer \( k \) such that for every \( j \geq k \) there exists a suffix of \( w_1 \ldots w_j \) which is also a prefix of \( w_{j+1}w_{j+2} \ldots \), i.e., at every position a proper centered square.
F. Mignosi, A. Restivo, S. Salemi (1998): local period as powers immediately to the left from each position

\[ \text{pref}_m(w) \] – a prefix of \( w \) of length \( m \)

For a rational number \( k \geq 1 \), \( w \) is a \( k \)-th power, if there exists a word \( v \) such that \( w = v^k \), where \( v^k \) denotes the word \( v'v'' \) with \( v' = v^{\lfloor k \rfloor} \) and \( v'' = \text{pref}_{|v|}(k - \lfloor k \rfloor)v \).

For a real number \( \rho \geq 1 \), \( w \) contains a repetition of order \( \rho \), if it contains as a factor a \( k \)-th power with \( k \geq \rho \).

\( w \) has a \( \rho \)-suffix, if \( w \) contains a repetition of order \( \rho \) as its suffix.
A right-infinite word $w$ is ultimately periodic if and only if there exists $n_0$ such that for every $n \geq n_0$ the word $\text{pref}_n(w)$ has a $\varphi^2$-suffix, where $\varphi = (1 + \sqrt{5})/2$.

The bound is optimal: for any real number $\varepsilon > 0$ there exists a natural number $n_0 > 0$ such that for any $n \geq n_0$ the prefix of length $n$ of the infinite Fibonacci word has a $(\varphi^2 - \varepsilon)$-suffix.
The document describes a concept in combinatorics on words, specifically focusing on powers with bounded periods. The authors cited are J. Karhumäki, A. Lepistö, and W. Plandowski (2002) who studied left powers with bounded periods.

Mathematically, a real number $\rho \geq 1$ and an integer $p \geq 1$ are considered. An infinite word $w$ is said to be $(\rho, p)$-repetitive if there exists an integer $n_0$ such that each prefix of $w$ of length at least $n_0$ ends with a repetition of order $\rho$ of a word of length at most $p$.

Moreover, the document states that if $w$ is $(\rho, p)$-repetitive, then it is also $(\rho', p)$-repetitive for any $p' \geq p$, and $(\rho', p)$-repetitive for any $\rho' \leq \rho$. This is a property that links different $(\rho, p)$-repetitive words.
J. Karhumäki, A. Lepistö, W. Plandowski (2002)

Connection between \((\rho, p)\)-repetitive and ultimately periodic words:

- there exist non-ultimately periodic \((2, 5)\)-repetitive words, and all such words are described.
- the pair \((2, 5)\) is optimal in the sense that \((\rho, p)\)-repetitive word with \(\rho > 2\) and \(p = 5\) or \(\rho = 2\) and \(p = 4\) is ultimately periodic.
Abelian powers

\[ u = u_1 u_2 \ldots u_n \]  
finite word over alphabet \( \Sigma \)

\[ a \in \Sigma, \ |u|_a \]  
the number of occurrences of the letter \( a \) in \( u \)

Two finite words \( u \) and \( v \) are abelian equivalent if \( |u|_a = |v|_a \) for all \( a \in \Sigma \).

An abelian \( k \)-power is a non-empty word of the form \( u = v_1 v_2 \ldots v_k \) where the words \( v_i \) are pairwise abelian equivalent.

\[ |v_1| = \cdots = |v_k| \]  
length of abelian period of \( u \).
A biinfinite word $w$ is $(k, l)$-abelian central repetitive, if it has at every position a centered abelian $2k$-power with length of period at most $l$ as its factor, i. e., for every $i$ there exists $l' \leq l$ such that $w_{i-l'k+1} \ldots w_{i+l'k}$ is an abelian $2k$-power.
Basic notions

(k, l)-ACR word

A biinfinite word $w$ is $(k, l)$-abelian central repetitive, if it has at every position a centered abelian $2k$-power with length of period at most $l$ as its factor,

i. e., for every $i$ there exists $l' \leq l$ such that $w_{i-l'k+1} \ldots w_{i+l'k}$ is an abelian $2k$-power.

(k, l)-ARR and ALR word

$w$ is $(k, l)$-abelian right (resp. left) repetitive, if it has an abelian $k$-power immediately to the right (resp. left) from every position with length of period at most $l$,

i. e., for every $i$ the word $w_{i+1} \ldots w_{i+l'k}$ (resp. $w_{i-l'k+1} \ldots w_i$) is an abelian $k$-power for some $l' \leq l$. 

S. Avgustinovich, J. Karhumäki, S. Puzynina

On abelian versions of Critical Factorization Theorem
Problem

\((k, l)\)-ACR, ALR, ARR properties

relations

⇔

periodicity
Some simple properties

- a periodic biinfinite word with the period $l$ is $(k, l)$-ACR, ALR and ARR for every $k$
- a $(k, l)$-ACR word is also $(k, l)$-ARR and ALR
- a word having at every position a centered (right, left) abelian square of fixed length is periodic
Theorem

For every integer $k$, there exists a biinfinite non-periodic $(k, l)$-ACR word with $l = 2(k + 1)^2$.

Compare!

For usual powers bounded squares imply periodicity.
Let $\Sigma = \{a_1, \ldots, a_n\}$ be an alphabet. Denote by $\mu_\Sigma$ the generalized Thue-Morse morphism:

$$
\mu_\Sigma(a_1) = a_1a_2\ldots a_n,
$$

$$
\mu_\Sigma(a_2) = a_2a_3\ldots a_na_1,
$$

$$
\vdots
$$

$$
\mu_\Sigma(a_n) = a_na_1\ldots a_{n-1}.
$$

Notice that $\mu = \mu_{\{0,1\}}$ is the usual Thue-Morse morphism, and its fixed point is the Thue-Morse word.
A biinfinite word $w$ is \((k, l)\)-arithmetically centered, if it contains at every position $i$ an arithmetical progression of one letter of length at least $2k + 1$ with common difference at most $l$ centered at this position, i.e., for every $i$ there exists $l' \leq l$ such that

$$
\omega_i = \omega_{i+l'} = \omega_{i-l'} = \omega_{i+2l'} = \omega_{i-2l'} = \cdots = \omega_{i+kl'} = \omega_{i-kl'}.
$$

**Lemma**

Let $w$ be a \((k, l)\)-arithmetically centered word over an alphabet $\Sigma$. Then $\mu_\Sigma(w)$ is \((k, |\Sigma|l)\)-ACR word.
Example: A family of non-periodic (1, 8)-ACR words.

First we build a (1, 4)-arithmetically centered word:

\[ w_i = \begin{cases} 
1, & \text{if } i \text{ is not divisible by } 2, \\
?, & \text{if } i = 4j, j \in \mathbb{Z}, \\
0, & \text{otherwise}.
\end{cases} \quad (1) \]

Instead of each ? one can use either 0 or 1.

Applying the Thue-Morse morphism

\[ \varphi(0) = 01, \varphi(1) = 10 \]

to this sequence, we obtain (1, 8)-ACR word.

The bound is optimal: if \( w \) is a biinfinite non-periodic (1, \( l \))-ACR word, then \( l \geq 8 \).
Example: $(1, 8)$-ACR words

\[ \cdots 010?010?010?010? \cdots \]

apply $\mu = \mu_{\{0,1\}}$

\[ \downarrow \]

\[ \cdots 011001\mu(?)011001\mu(?)011001\mu(?)011001\mu(?) \cdots \]
Example: $(1, 8)$-ACR words

\[ \ldots 010?010?010?010? \ldots \]

apply $\mu = \mu_{\{0,1\}}$

\[ \downarrow \]

\[ \ldots 011001\mu(?)01100.1\mu(?)011001\mu(?)011001\mu(?) \ldots \]
Example: \((1, 8)\)-ACR words

\[ \ldots 010?010?010?010? \ldots \]

apply \(\mu = \mu_{\{0,1\}}\)

\[ \downarrow \]

\[ \ldots 011001\mu(?)01100.1\mu(?)011001\mu(?)011001\mu(?) \ldots \]
Example: \((1, 8)\)-ACR words

apply \(\mu = \mu_{\{0,1\}}\)

\[
\downarrow
\]

\[
\ldots 010?010?010?010?\ldots
\]

\[
\ldots 011001\mu(?)011001\mu(?)011001\mu(?)011001\mu(?)\ldots
\]
Example: \((1, 8)\)-ACR words

\[ \ldots 010?010?010?010? \ldots \]

apply \(\mu = \mu\{0,1\}\)

\[ \Downarrow \]

\[ \ldots 011001\mu(?)011001\mu(?)011.001\mu(?)011001\mu(?) \ldots \]
Example: \((1, 8)\)-ACR words

\[
\ldots 010?010?010?010? \ldots
\]

apply \(\mu = \mu\{0,1\}\)

\[
\downarrow
\]

\[
\ldots 011001\mu(?)011001\mu(?)011001\mu(?)011001\mu(?) \ldots
\]
Example: $(1, 8)$-ACR words

\[
\ldots 010?010?010?010?\ldots 
\]

if \( ? = 0 \), \( \mu(?) = 01 \)

apply \( \mu = \mu\{0,1\} \)

\[
\downarrow
\]

\[
\ldots 011001\mu(?)011001.0.1011001\mu(?)011001\mu(?)\ldots
\]

S. Avgustinovich, J. Karhumäki, S. Puzynina

On abelian versions of Critical Factorization Theorem
Example: (1, 8)-ACR words

\[ \ldots 010?010?010?010? \ldots \]

if \( ? = 1 \), \( \mu(?) = 10 \)

apply \( \mu = \mu_{\{0,1\}} \)

\[ \downarrow \]

\[ \ldots 011001\mu(?)011001.0011001\mu(?)011001\mu(?) \ldots \]
Theorem

*There exists a biinfinite non-periodic word \( w \), such that for every integer \( k \) there exists \( l \) such that \( w \) is \((k, l)\)-ACR word.*
If in the previous theorem we replace quantifiers “there exists $l$” and “for every $k$”, then we obtain an opposite result:

**Theorem**

*Let $w$ be a biinfinite word. If there exists an integer $l$ such that for every $k$ the word $w$ is $(k, l)$-ACR, then $w$ is periodic.*
All statements can be reformulated for right infinite words:

- periodicity $\rightarrow$ ultimate periodicity
- suppose that all conditions for abelian repetitions hold starting from some position

For example, $w \in \Sigma^\omega$ is $(k, l)$-ACR, if there exists $i_0$ such that for every $i \geq i_0$ there is an abelian central $2k$-power with length of period at most $l$ at position $i$. 
For every $k$ there exist non-periodic biinfinite $(k, 2(k + 1)^2)$-ARR (ALR) words.

Moreover, there exists a non-periodic biinfinite word, such that for every $k$ there exists $l$ for which $w$ is a $(k, l)$-ARR (ALR) word.
Right and left abelian squares

Biinfinite $(2, 2)$-ARR binary words.

$0^\mathbb{Z}, (01)^\mathbb{Z}, (1100)^-\omega(10)\omega, 0^-\omega(1100)\omega, (0)^-\omega(0011)^*(01)\omega, (0)^-\omega(1100)^*(10)\omega$ (up to replacing 0 and 1)
Right and left abelian squares

Biinfinite $(2, 2)$-ARR binary words.

$0^\mathbb{Z}$, $(01)^\mathbb{Z}$, $(1100)^{-\omega}(10)^\omega$, $0^{-\omega}(1100)^\omega$, $(0)^{-\omega}(0011)^*(01)^\omega$, $(0)^{-\omega}(1100)^*(10)^\omega$ (up to replacing 0 and 1)

Some of these words are non-periodic, but they have some periodicity structure similar to ultimately periodic right infinite words.
Biinfinite \((2, 2)\)-ARR binary words.

\[
0^\mathbb{Z}, (01)^\mathbb{Z}, (1100)^-\omega(10)^\omega, 0^-\omega(1100)^\omega, (0)^-\omega(0011)^*(01)^\omega, (0)^-\omega(1100)^*(10)^\omega \text{ (up to replacing 0 and 1)}
\]

Some of these words are non-periodic, but they have some periodicity structure similar to ultimately periodic right infinite words.

We say that a biinfinite word \(w\) is \textbf{ultimately periodic}, if there exist finite words \(u, v_1, v_2\), such that \(w = v_1^-\omega uv_2^\omega\).
Biinfinite \((2, 2)\)-ARR binary words.

\[0^\mathbb{Z}, (01)^\mathbb{Z}, (1100)^-\omega (10)^\omega, 0^-\omega (1100)^\omega, (0)^-\omega (0011)^* (01)^\omega, (0)^-\omega (1100)^* (10)^\omega \text{ (up to replacing 0 and 1)}\]

Some of these words are non-periodic, but they have some periodicity structure similar to ultimately periodic right infinite words.

We say that a biinfinite word \(w\) is **ultimately periodic**, if there exist finite words \(u\), \(v_1\), \(v_2\), such that \(w = v_1^-\omega uv_2^\omega\).

Non-periodic biinfinite \((2, 3)\)-ARR binary words.

\[\supseteq \{01, 011\}^\mathbb{Z}\]
The bound distinguishing periodic and non-periodic words

- for $l = 1$ there exist only periodic words, namely, words of the form $a^\mathbb{Z}$, where $a \in \Sigma$
- for $l = 2$ there exist only periodic and ultimately periodic words
- for $l \geq 3$ there exist non-periodic words
Sturmian words

Let $s$ be a Sturmian word. Then

- $s$ is not $(k, l)$-ACR for every pair $(k, l)$ (follows from properties of sturmian words)
- for every $k$ there exists $l$ such that $s$ is $(k, l)$-ARR (theorem G. Richomme, K. Saari and L. Zamboni, 2009)
Sturmian and Thue-Morse words

Sturmian words

Let $s$ be a Sturmian word. Then

- $s$ is not $(k, l)$-ACR for every pair $(k, l)$ (follows from properties of sturmian words)
- for every $k$ there exists $l$ such that $s$ is $(k, l)$-ARR (theorem G. Richomme, K. Saari and L. Zamboni, 2009)

Thue-Morse word

- Thue-Morse word is $(1, 14)$-ACR and not $(k, l)$-ACR for every $k > 1$ or $l < 14$;
- Thue-Morse word is $(2, 5)$-ARR and not $(k, l)$-ARR for every $k > 2$ or $l < 5$. 

S. Avgustinovich, J. Karhumäki, S. Puzynina

On abelian versions of Critical Factorization Theorem
Sturmian and Thue-Morse words

**Sturmian words**

Let $s$ be a Sturmian word. Then

- $s$ is not $(k, l)$-ACR for every pair $(k, l)$ (follows from properties of Sturmian words)
- For every $k$ there exists $l$ such that $s$ is $(k, l)$-ARR (theorem G. Richomme, K. Saari and L. Zamboni, 2009)

**Thue-Morse word**

- Thue-Morse word is $(1, 14)$-ACR and not $(k, l)$-ACR for every $k > 1$ or $l < 14$;
- Thue-Morse word is $(2, 5)$-ARR and not $(k, l)$-ARR for every $k > 2$ or $l < 5$.

Compare!
Open problems

Open problem 1.

For any $k$, find $l$ such that there exists a non-periodic $(k, l)$-ACR (ARR) word, and all $(k, l - 1)$-ACR (ARR) words are periodic.

Open problem 2.

Let $k$ be an integer. Does there exist a non-periodic word $w$ and integers $l_1, l_2$, such that $w$ contains a central abelian $2^k$-power of length $l_1$ or $l_2$ at every position?

Open problem 3.

Find other natural conditions for abelian powers enforcing periodicity.

S. Avgustinovich, J. Karhumäki, S. Puzynina

On abelian versions of Critical Factorization Theorem
Open problems

Open problem 1.
For any $k$, find $l$ such that there exists a non-periodic $(k, l)$-ACR (ARR) word, and all $(k, l-1)$-ACR (ARR) words are periodic.

Open problem 2.
Let $k$ be an integer. Does there exist a non-periodic word $w$ and integers $l_1, l_2$, such that $w$ contains a central abelian $2k$-power of length $l_1$ or $l_2$ at every position?
Open problems

Open problem 1.
For any $k$, find $l$ such that there exists a non-periodic $(k, l)$-ACR (ARR) word, and all $(k, l - 1)$-ACR (ARR) words are periodic.

Open problem 2.
Let $k$ be an integer. Does there exist a non-periodic word $w$ and integers $l_1, l_2$, such that $w$ contains a central abelian $2k$-power of length $l_1$ or $l_2$ at every position?

Open problem 3.
Find other natural conditions for abelian powers enforcing periodicity.
Thank you!