More symmetries occurring in an infinite word

Palindromic and G-palindromic defect

E. Pelantová[1] **Š. Starosta**[2]

[1]Department of Mathematics
Faculty of Nuclear Sciences and Physical Engineering
Czech Technical University in Prague

[2]Department of Applied Mathematics
Faculty of Information Technology
Czech Technical University in Prague

September 30, 2012 MELA 2012

Outline

Outline

Combinatorics on words

	example
alphabet $\mathcal{A}$	$\{0,1\}$
infinite word $\mathbf{u} = (u_i)_{i=0}^{+\infty}$, $u_i \in \mathcal{A}$	$011010\ldots$
factor $w = u_k u_{k+1} \ldots u_{k+n-1} \in \mathcal{A}^*$	01, 11, 10

language of u is the set of its factors, denoted $\mathcal{L}(\mathbf{u})$

Reversal mapping and its fixed points

reversal mapping R

$$R(w_0 w_1 \dots w_n) = w_n \dots w_1 w_0$$

palindrome $w = R(w)$

examples: $0, 00, 010, \varepsilon$

Reversal mapping and its fixed points

reversal mapping R

$$R(w_0 w_1 \dots w_n) = w_n \dots w_1 w_0$$

palindrome $w = R(w)$

examples: $0, 00, 010, \varepsilon$

Rauzy graphs (1/2)

Rauzy graph of order $n \in \mathbb{N}$

- is a subgraph of n-dimensional De Bruijn graph;
- represents factors of an infinite word up to the length $n+1$.

vertices = factors of length n

there is an edge e from w to v if e is a factor of length $n+1$ and there exist letters a and b such that $e = wa = bv$

Rauzy graphs (1/2)

Rauzy graph of order $n \in \mathbb{N}$

- is a subgraph of n-dimensional De Bruijn graph;
- represents factors of an infinite word up to the length $n+1$.

vertices = factors of length n

there is an edge e from w to v if e is a factor of length $n+1$ and there exist letters a and b such that $e = wa = bv$

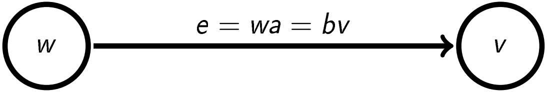
Rauzy graphs (1/2)

Rauzy graph of order $n \in \mathbb{N}$

- is a subgraph of n-dimensional De Bruijn graph;
- represents factors of an infinite word up to the length $n + 1$.

vertices = factors of length n

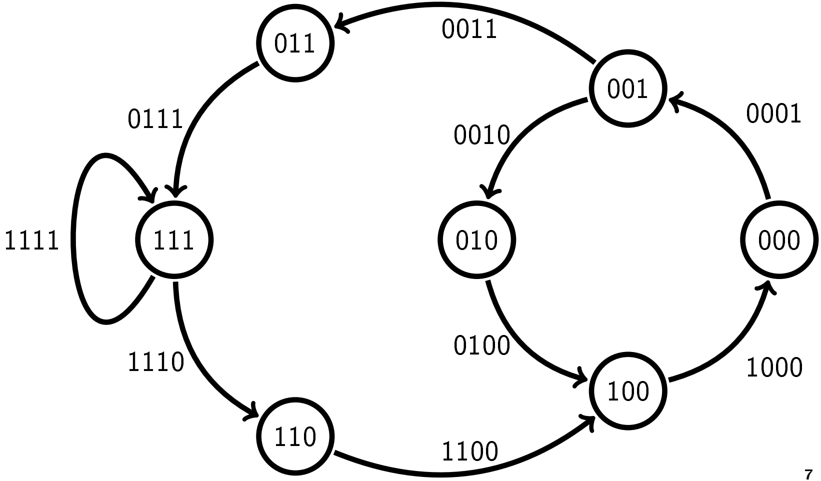
there is an edge e from w to v if e is a factor of length $n + 1$ and there exist letters a and b such that $e = wa = bv$



Rauzy graphs (2/2)

Language of an infinite word 0001000100011100010001000 11...

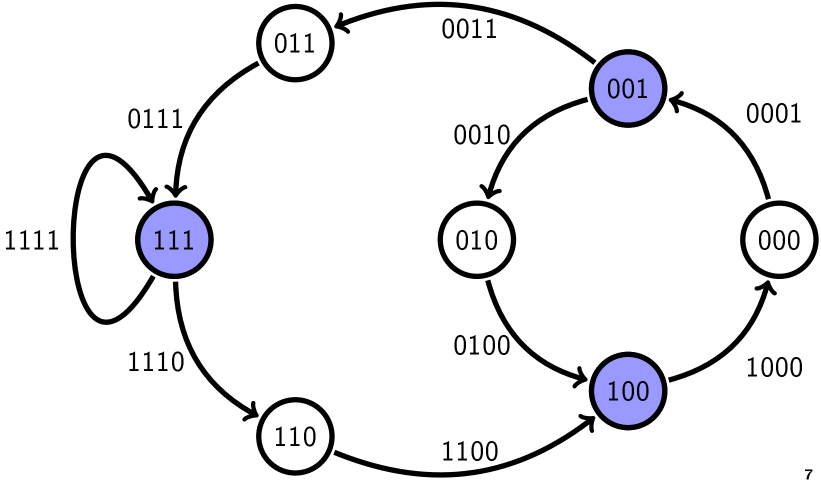
Rauzy graph of order 3



Rauzy graphs (2/2)

Language of an infinite word 0001000100011100010001000110...

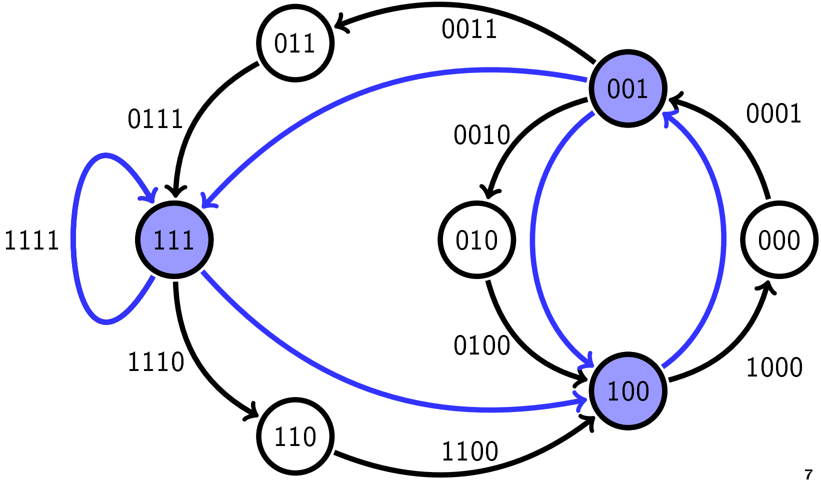
transformation of Rauzy graph to reduced Rauzy graph of order 3



Rauzy graphs (2/2)

Language of an infinite word 0001000100011100010001000 11 . . .

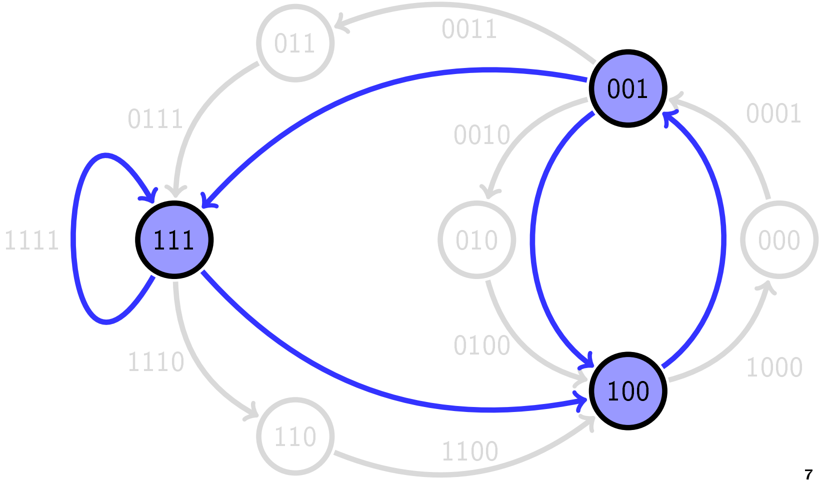
transformation of Rauzy graph to reduced Rauzy graph of order 3



Rauzy graphs (2/2)

Language of an infinite word 00010001000111000100010001 1...

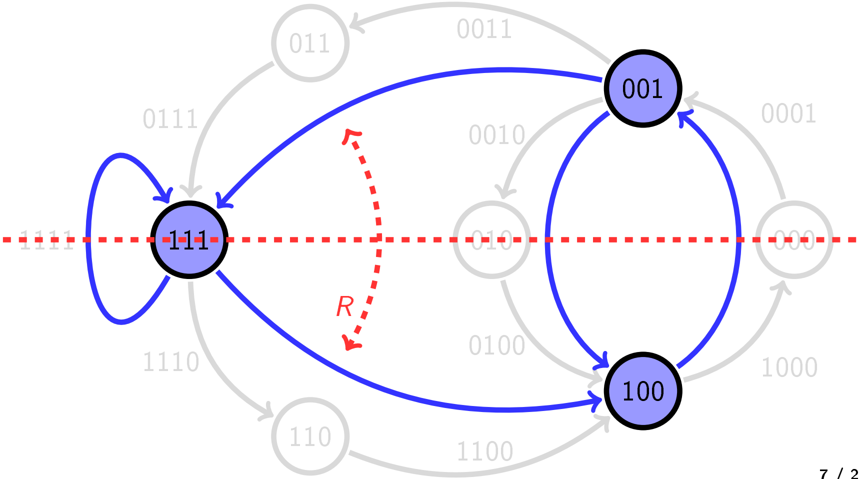
reduced Rauzy graph of order 3



Rauzy graphs (2/2)

Language of an infinite word 0001000100011100010001000011 . . .

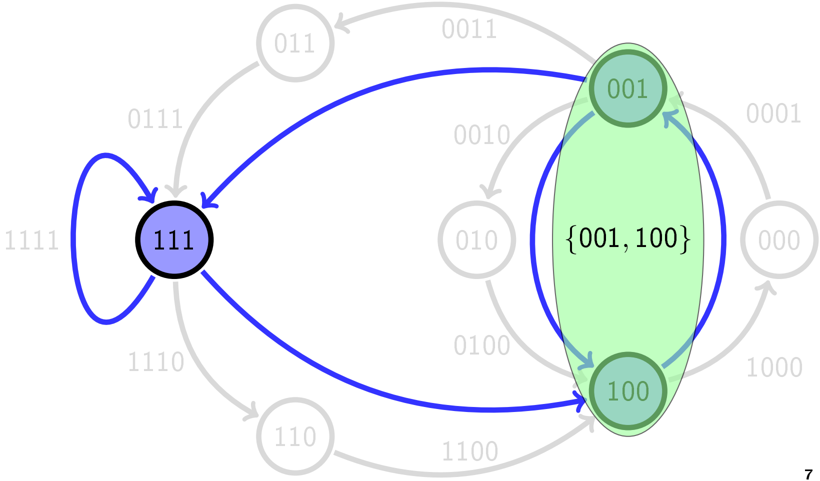
reduced Rauzy graph of order 3, the language is closed under R



Rauzy graphs (2/2)

Language of an infinite word 0001000100011100010001000 11 . . .

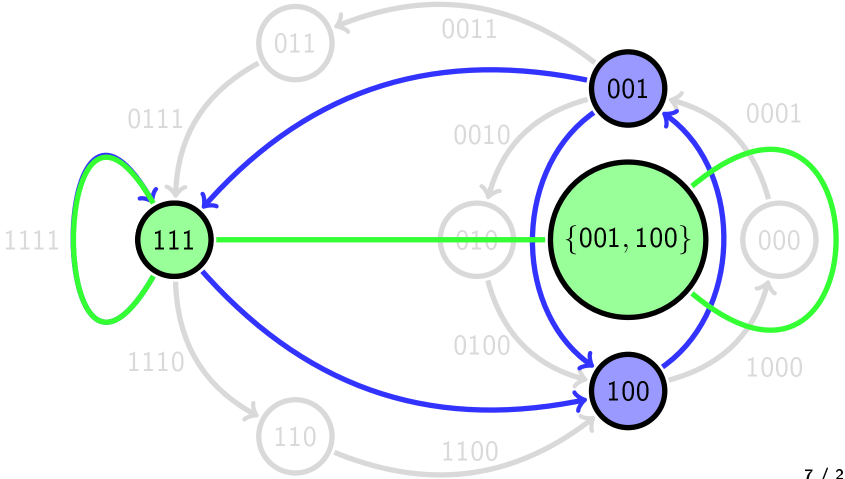
transformation of reduced to super reduced Rauzy graph of order 3



Rauzy graphs (2/2)

Language of an infinite word 00010001000111000100010001 1...

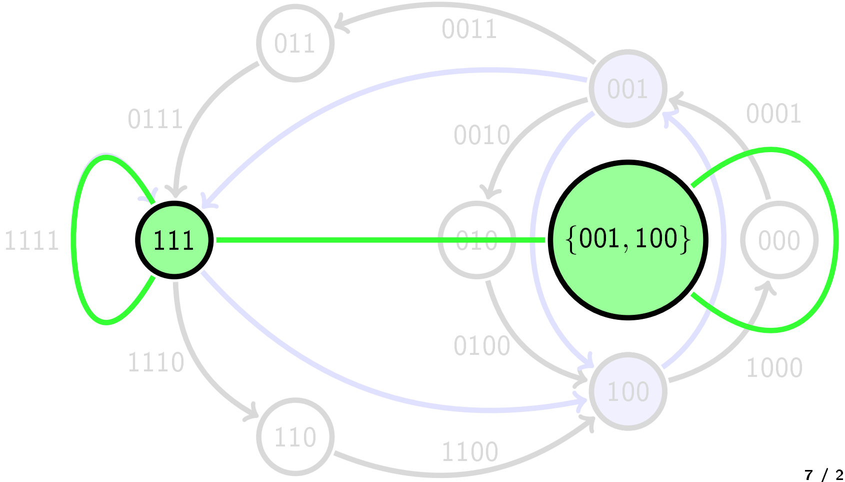
transformation of reduced to super reduced Rauzy graph of order 3



Rauzy graphs (2/2)

Language of an infinite word 0001000100011100010001000011...

super reduced Rauzy graph of order 3



Rauzy graphs - applications

Rauzy graph of order n of $\mathcal{L}(\mathbf{u})$ can be used to determine:

- factors up to the length $n + 1$
- symmetries (palindromic complexity)
- special / bispecial factors (factor complexity)
- ...

Return words cannot be determined.

Evolution of Rauzy graphs.

Rauzy graphs - applications

Rauzy graph of order n of $\mathcal{L}(\mathbf{u})$ can be used to determine:

- factors up to the length $n+1$
- symmetries (palindromic complexity)
- special / bispecial factors (factor complexity)
- ...

Return words cannot be determined.

Evolution of Rauzy graphs.

Rauzy graphs - applications

Rauzy graph of order n of $\mathcal{L}(\mathbf{u})$ can be used to determine:

- factors up to the length $n + 1$
- symmetries (palindromic complexity)
- special / bispecial factors (factor complexity)
- ...

Return words cannot be determined.

Evolution of Rauzy graphs.

Outline

Words with finite palindromic defect

By $\Gamma_n(\mathbf{u})$ we denote the super reduced Rauzy graph of order n of an infinite word $\mathbf{u}$.

Definition

We say that an infinite word $\mathbf{u}$ has finite palindromic defect (or is almost rich/full) if $\mathcal{L}(\mathbf{u})$ is invariant under R and there exists $N \in \mathbb{N}$ such for each $n \geq N$ the following holds

- if e is a loop in $\Gamma_n(\mathbf{u})$, then e represents palindrome;
- the graph obtained from $\Gamma_n(\mathbf{u})$ by removing loops is a tree.

If $N = 0$, then we say that $\mathbf{u}$ has palindromic defect 0 (or is rich/full).

Examples: episturmian words, words coding symmetric interval exchange transformation, ...

Words with finite palindromic defect

By $\Gamma_n(\mathbf{u})$ we denote the super reduced Rauzy graph of order n of an infinite word $\mathbf{u}$.

Definition

We say that an infinite word $\mathbf{u}$ has finite palindromic defect (or is almost rich/full) if $\mathcal{L}(\mathbf{u})$ is invariant under R and there exists $N \in \mathbb{N}$ such for each $n \geq N$ the following holds

- if e is a loop in $\Gamma_n(\mathbf{u})$, then e represents palindrome;
- the graph obtained from $\Gamma_n(\mathbf{u})$ by removing loops is a tree.

If $N = 0$, then we say that $\mathbf{u}$ has palindromic defect 0 (or is rich/full).

Examples: episturmian words, words coding symmetric interval exchange transformation, ...

Words with finite palindromic defect

By $\Gamma_n(\mathbf{u})$ we denote the super reduced Rauzy graph of order n of an infinite word $\mathbf{u}$.

Definition

We say that an infinite word $\mathbf{u}$ has finite palindromic defect (or is almost rich/full) if $\mathcal{L}(\mathbf{u})$ is invariant under R and there exists $N \in \mathbb{N}$ such for each $n \geq N$ the following holds

- if e is a loop in $\Gamma_n(\mathbf{u})$, then e represents palindrome;
- the graph obtained from $\Gamma_n(\mathbf{u})$ by removing loops is a tree.

If $N = 0$, then we say that $\mathbf{u}$ has palindromic defect 0 (or is rich/full).

Examples: episturmian words, words coding symmetric interval exchange transformation, ...

Characterizations of words with defect 0

For an infinite word $\mathbf{u}$ with language invariant under R the following statements are equivalent:

1. $\mathbf{u}$ has palindromic defect 0;
2. the longest palindromic suffix of any factor $w \in \mathcal{L}(\mathbf{u})$ is unioccurrent in w;
3. any complete return word of any palindromic factor of $\mathbf{u}$ is a palindrome;
4. for any factor w of $\mathbf{u}$, every factor of $\mathbf{u}$ that contains w only as its prefix and $R(w)$ only as its suffix is a palindrome;
5. for each n the following equality holds

$$\mathcal{C}(n+1) - \mathcal{C}(n) + 2 = \mathcal{P}(n) + \mathcal{P}(n+1).$$

[Droubay et al. 2001, Glen et al. 2009, Bucci et al. 2009]

Outline

Θ-palindromes

let $\Theta : \mathcal{A}^* \mapsto \mathcal{A}^*$ be an involutive antimorphism, i.e., $\Theta^2 = \mathrm{Id}$ and $\Theta(wv) = \Theta(v)\Theta(w)$ for all $w, v \in \mathcal{A}^*$

Θ-**palindrome** $w = \Theta(w)$

[Kari et al.; Anne et al., 2005; de Luca et al., 2006; ...]

Example: Watson-Crick complementarity $A \leftrightarrow T, G \leftrightarrow C$

Θ-palindromes: AT, $AATT$, $AGCT$

Θ-palindromes

let $\Theta : \mathcal{A}^* \mapsto \mathcal{A}^*$ be an involutive antimorphism, i.e., $\Theta^2 = \mathrm{Id}$ and $\Theta(wv) = \Theta(v)\Theta(w)$ for all $w, v \in \mathcal{A}^*$

Θ-**palindrome** $w = \Theta(w)$

[Kari et al.; Anne et al., 2005; de Luca et al., 2006; ...]

Example: Watson-Crick complementarity $A \leftrightarrow T, G \leftrightarrow C$

Θ-palindromes: AT, $AATT$, $AGCT$

Θ-palindromes

let $\Theta : \mathcal{A}^* \mapsto \mathcal{A}^*$ be an involutive antimorphism, i.e., $\Theta^2 = \mathrm{Id}$ and $\Theta(wv) = \Theta(v)\Theta(w)$ for all $w, v \in \mathcal{A}^*$

Θ-**palindrome** $w = \Theta(w)$

[Kari et al.; Anne et al., 2005; de Luca et al., 2006; ...]

Example: Watson-Crick complementarity $A \leftrightarrow T, G \leftrightarrow C$

Θ-palindromes: AT, $AATT$, $AGCT$

More symmetries

Let $\mathbf{u}$ be an infinite word over $\mathcal{A}$.

Let G be a **finite** group consisting of morphisms and antimorphisms over $\mathcal{A}$ such that $\mathcal{L}(\mathbf{u})$ is invariant under all elements of G.

In this context, invariance under R is the same as invariance under all elements of the group $\{\mathrm{Id}, R\}$.

$[w] = \{\nu(w) \mid \nu \in G\}$

More symmetries

Let $\mathbf{u}$ be an infinite word over $\mathcal{A}$.

Let G be a **finite** group consisting of morphisms and antimorphisms over $\mathcal{A}$ such that $\mathcal{L}(\mathbf{u})$ is invariant under all elements of G.

In this context, invariance under R is the same as invariance under all elements of the group $\{\mathrm{Id}, R\}$.

$[w] = \{\nu(w) \mid \nu \in G\}$

More symmetries

Let $\mathbf{u}$ be an infinite word over $\mathcal{A}$.

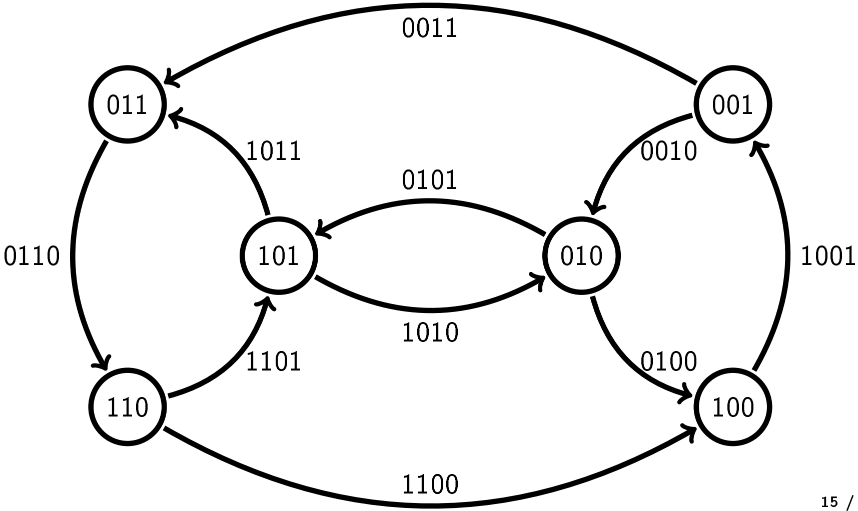
Let G be a **finite** group consisting of morphisms and antimorphisms over $\mathcal{A}$ such that $\mathcal{L}(\mathbf{u})$ is invariant under all elements of G.

In this context, invariance under R is the same as invariance under all elements of the group $\{\mathrm{Id}, R\}$.

$[w] = \{\nu(w) \mid \nu \in G\}$

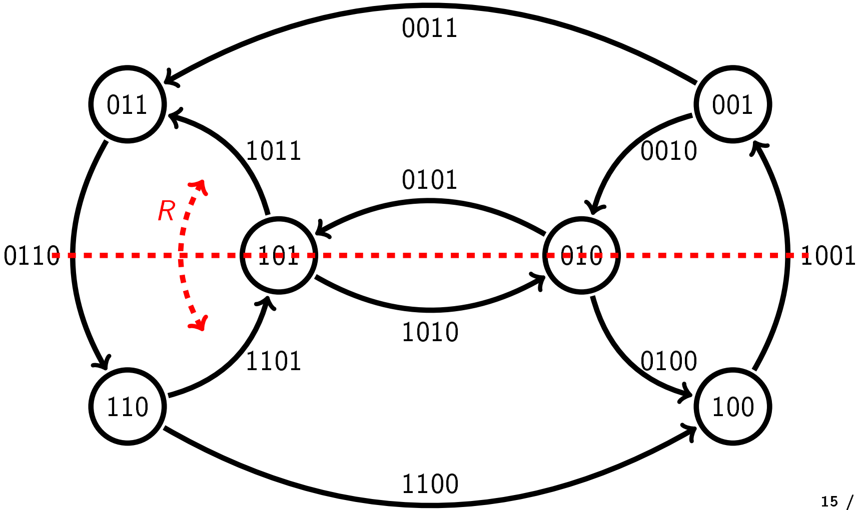
Graph of symmetries of the Thue-Morse word

the Thue-Morse word is a fixed point of the morphism
$0 \mapsto 01, 1 \mapsto 10$



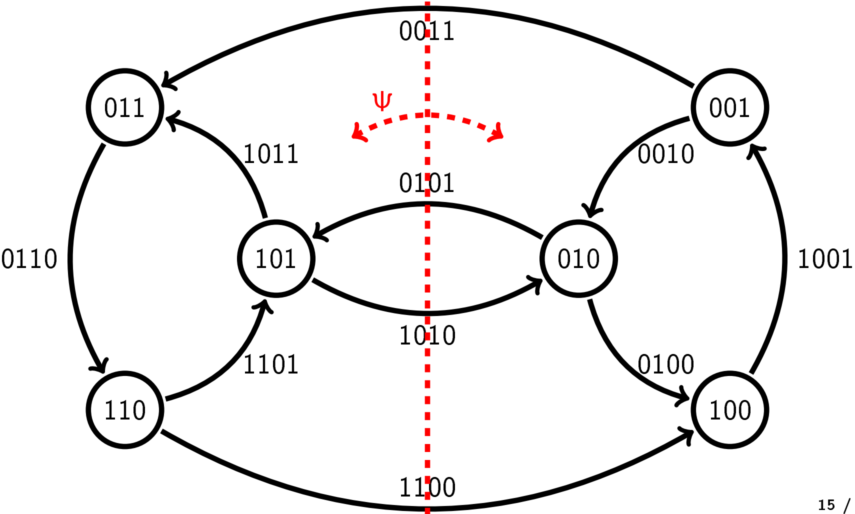
Graph of symmetries of the Thue-Morse word

the Thue-Morse word is a fixed point of the morphism
$0 \mapsto 01, 1 \mapsto 10$



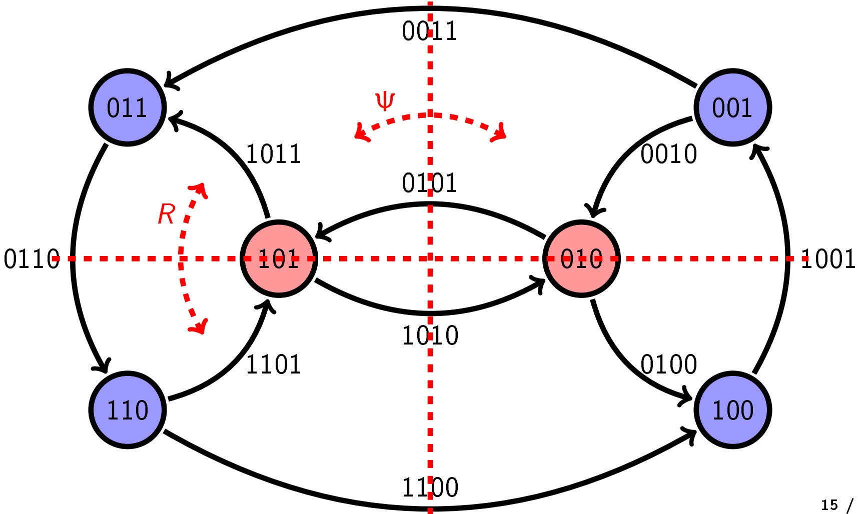
Graph of symmetries of the Thue-Morse word

the Thue-Morse word is a fixed point of the morphism $0 \mapsto 01, 1 \mapsto 10$, its language is invariant under $\Psi : 0 \mapsto 1, 1 \mapsto 0$



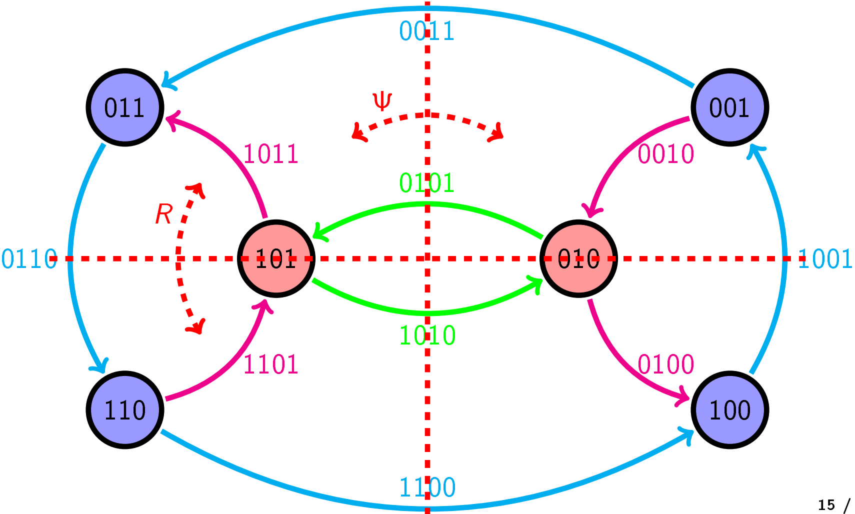
Graph of symmetries of the Thue-Morse word

the Thue-Morse word is a fixed point of the morphism
$0 \mapsto 01, 1 \mapsto 10$, its language is invariant under $\Psi : 0 \mapsto 1, 1 \mapsto 0$



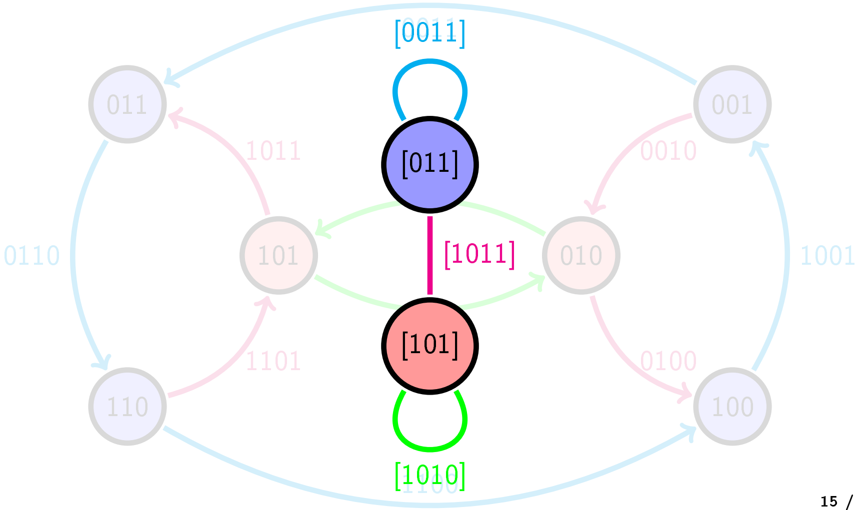
Graph of symmetries of the Thue-Morse word

the Thue-Morse word is a fixed point of the morphism $0 \mapsto 01, 1 \mapsto 10$, its language is invariant under $\Psi : 0 \mapsto 1, 1 \mapsto 0$



Graph of symmetries of the Thue-Morse word

the Thue-Morse word is a fixed point of the morphism
$0 \mapsto 01, 1 \mapsto 10$



Graph of symmetries - definition

The **directed graph of symmetries** of the word $\mathbf{u}$ of order n, denoted is $\overrightarrow{\Gamma}_n(\mathbf{u})$, is the graph $(V, \overrightarrow{E})$ such that

$$V = \left\{ [w] \;\middle|\; w \in \mathcal{L}(\mathbf{u}), |w| = n, w \text{ is special} \right\}$$

and an edge $e \in \overrightarrow{E} \subset \mathcal{L}(\mathbf{u})$ starts in a vertex $[w]$ and ends in a vertex $[v]$, if

- the prefix of e of length n belongs to $[w]$,
- the suffix of e of length n belongs to $[v]$,
- e has exactly two occurrences of special factors of length n.

The **graph of symmetries** of the word $\mathbf{u}$ of order n, denoted $\Gamma_n(\mathbf{u})$, is the graph (V, E) with the same set of vertices as $\overrightarrow{\Gamma}_n(\mathbf{u})$ and for any $e \in \mathcal{L}(\mathbf{u})$ we have

$$[e] \in E \quad \iff \quad e \in \overrightarrow{E}.$$

Graph of symmetries - definition

The **directed graph of symmetries** of the word $\mathbf{u}$ of order n, denoted is $\overrightarrow{\Gamma}_n(\mathbf{u})$, is the graph $(V, \overrightarrow{E})$ such that

$$V = \{[w] \mid w \in \mathcal{L}(\mathbf{u}), |w| = n, w \text{ is special } \}$$

and an edge $e \in \overrightarrow{E} \subset \mathcal{L}(\mathbf{u})$ starts in a vertex $[w]$ and ends in a vertex $[v]$, if

- the prefix of e of length n belongs to $[w]$,
- the suffix of e of length n belongs to $[v]$,
- e has exactly two occurrences of special factors of length n.

The **graph of symmetries** of the word $\mathbf{u}$ of order n, denoted $\Gamma_n(\mathbf{u})$, is the graph (V, E) with the same set of vertices as $\overrightarrow{\Gamma}_n(\mathbf{u})$ and for any $e \in \mathcal{L}(\mathbf{u})$ we have

$$[e] \in E \iff e \in \overrightarrow{E}.$$

Words having finite G-defect

Definition

Let $G \subset AM(\mathcal{A}^*)$ be a finite group containing at least one antimorphism. We say that an infinite word $\mathbf{u}$ has finite G-defect (or almost G-rich) if $\mathcal{L}(\mathbf{u})$ is invariant under all elements of G and there exists $N \in \mathbb{N}$ such that for each $n > N$ the following holds:

- if $[e]$ is a loop in $\Gamma_n(\mathbf{u})$, then e is a Θ-palindrome for some involutive antimorphism $\Theta \in G$;
- the graph obtained from $\Gamma_n(\mathbf{u})$ by removing loops is a tree.

If $N = 0$, we say that $\mathbf{u}$ has G-defect 0 (or G-rich).

The Thue-Morse word has $\{\mathrm{Id}, R, \Psi, R\Psi\}$-defect 0.

Words having finite G-defect

Definition

Let $G \subset AM(\mathcal{A}^*)$ be a finite group containing at least one antimorphism. We say that an infinite word $\mathbf{u}$ has finite G-defect (or almost G-rich) if $\mathcal{L}(\mathbf{u})$ is invariant under all elements of G and there exists $N \in \mathbb{N}$ such that for each $n > N$ the following holds:

- if $[e]$ is a loop in $\Gamma_n(\mathbf{u})$, then e is a Θ-palindrome for some involutive antimorphism $\Theta \in G$;
- the graph obtained from $\Gamma_n(\mathbf{u})$ by removing loops is a tree.

If $N = 0$, we say that $\mathbf{u}$ has G-defect 0 (or G-rich).

The Thue-Morse word has $\{\mathrm{Id}, R, \Psi, R\Psi\}$-defect 0.

Words having finite G-defect

Definition

Let $G \subset AM(\mathcal{A}^*)$ be a finite group containing at least one antimorphism. We say that an infinite word $\mathbf{u}$ has finite G-defect (or almost G-rich) if $\mathcal{L}(\mathbf{u})$ is invariant under all elements of G and there exists $N \in \mathbb{N}$ such that for each $n > N$ the following holds:

- if $[e]$ is a loop in $\Gamma_n(\mathbf{u})$, then e is a Θ-palindrome for some involutive antimorphism $\Theta \in G$;
- the graph obtained from $\Gamma_n(\mathbf{u})$ by removing loops is a tree.

If $N = 0$, we say that $\mathbf{u}$ has G-defect 0 (or G-rich).

The Thue-Morse word has $\{\mathrm{Id}, R, \Psi, R\Psi\}$-defect 0.

Characterizations of words with G-defect 0

As in the case of palindromic defect, such words are fully saturated by generalized palindromes.

Theorem

Let $\mathbf{u}$ be an infinite word with language closed under all elements of G. The following conditions are equivalent:

1. $\mathbf{u}$ *has G-defect* 0*;*
2. *for all $v \in \mathcal{L}(\mathbf{u})$ the G-longest palindromic suffix of v is G-unioccurrent in v or the last letter of v is G-unioccurrent in v;*
3. *for all $w \in \mathcal{L}(\mathbf{u})$ every complete G-return word of $[w]$ is a G-palindrome;*

Characterizations of words with finite G-defect

Theorem

Let $\mathbf{u}$ *be a* ***uniformly recurrent*** *infinite word with language closed under all elements of* G*. The following conditions are equivalent:*

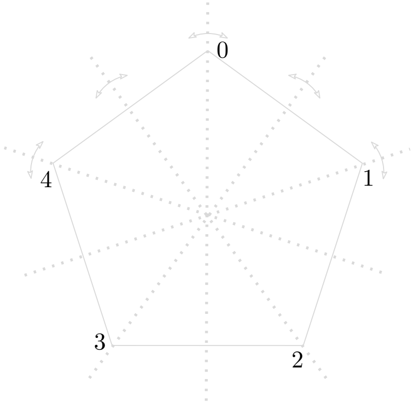
1. $\mathbf{u}$ *has finite* G*-defect;*
2. *there exists an integer* N *such that for all* $v \in \mathcal{L}(\mathbf{u})$, $|v| > N$, *the* G*-longest palindromic suffix of* v *is* G*-unioccurrent in* v *or the last letter of* v *is* G*-unioccurrent in* v*;*
3. *there exists an integer* N *such that for all* $w \in \mathcal{L}(\mathbf{u})$ *of length greater than* N *every complete* G*-return word of* $[w]$ *is a* G*-palindrome;*
4. *there exists an integer* N *such that*

$$\Delta\mathcal{C}(n) + \#G = \sum_{\Theta \in G^{(2)}} \Big(\mathcal{P}_\Theta(n) + \mathcal{P}_\Theta(n+1) \Big) \qquad \text{for all } n \geq N;$$

G-defect and Coxeter groups

If there exists a word with **finite** G-defect, then G is a Coxeter group.

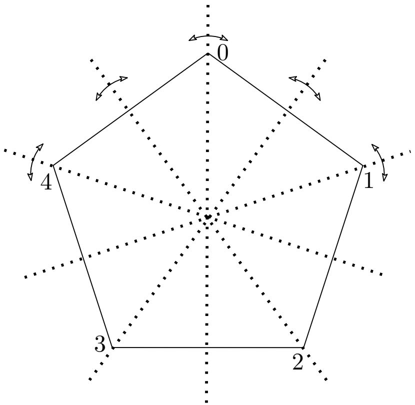
Every so-called generalized Thue-Morse word [Allouche et al., 2000] has zero G-defect where G is isomorphic to a dihedral group.



G-defect and Coxeter groups

If there exists a word with **finite** G-defect, then G is a Coxeter group.

Every so-called generalized Thue-Morse word [Allouche et al., 2000] has zero G-defect where G is isomorphic to a dihedral group.



Outline

Open questions

1. Given (finite Coxeter group) G, is there a G-rich word? Can it be constructed as a fixed point of a morphism?
2. Is an almost G_1-rich word related to a G_2-rich word if G_1 is isomorphic to G_2.
3. Conjecture: the G-defect of an aperiodic fixed point of a primitive non-injective morphism 0 or $+\infty$.
4. Given n, how many G-rich words of length n exist?

Open questions

1. Given (finite Coxeter group) G, is there a G-rich word? Can it be constructed as a fixed point of a morphism?
2. Is an almost G_1-rich word related to a G_2-rich word if G_1 is isomorphic to G_2.
3. Conjecture: the G-defect of an aperiodic fixed point of a primitive non-injective morphism 0 or $+\infty$.
4. Given n, how many G-rich words of length n exist?

Open questions

1. Given (finite Coxeter group) G, is there a G-rich word? Can it be constructed as a fixed point of a morphism?
2. Is an almost G_1-rich word related to a G_2-rich word if G_1 is isomorphic to G_2.
3. Conjecture: the G-defect of an aperiodic fixed point of a primitive non-injective morphism 0 or $+\infty$.
4. Given n, how many G-rich words of length n exist?

Open questions

1. Given (finite Coxeter group) G, is there a G-rich word? Can it be constructed as a fixed point of a morphism?
2. Is an almost G_1-rich word related to a G_2-rich word if G_1 is isomorphic to G_2.
3. Conjecture: the G-defect of an aperiodic fixed point of a primitive non-injective morphism 0 or $+\infty$.
4. Given n, how many G-rich words of length n exist?

Open questions

1. Given (finite Coxeter group) G, is there a G-rich word? Can it be constructed as a fixed point of a morphism?
2. Is an almost G_1-rich word related to a G_2-rich word if G_1 is isomorphic to G_2.
3. Conjecture: the G-defect of an aperiodic fixed point of a primitive non-injective morphism 0 or $+\infty$.
4. Given n, how many G-rich words of length n exist?

Thank you for your attention.