Specular sets

Francesco Dolce

WORDS

KIEL | 17 September 2015

Joint work with

V. Berthé, C. De Felice, V. Delecroix, J. Leroy, D. Perrin, C. Reutenauer, G. Rindone

Introduction

Generalization of links between Sturmian sets and Free groups to general objects : *Specular sets* and *Specular groups*.

Introduction of new concepts : *parity* of words (*odd* and *even* words), *mixed return words*.

Framework allowing to handle linear involutions (generalization of interval exchanges).

Adaptation of results holding for tree sets : *Maximal Bifix Decoding Theorem*, *Finite Index Basis Theorem*, *Return Theorem*.

Outline

Outline

Given an involution $\theta : A \to A$ (possibly with some fixed point), let us define

$$G_\theta = \langle a \in A \mid a \cdot \theta(a) = 1 \text{ for every } a \in A \rangle.$$

$G_\theta = \mathbb{Z}^i * (\mathbb{Z}/2\mathbb{Z})^j$ is a *specular group* of type (i, j), and $\mathrm{Card}\,(A) = 2i + j$ is its *symmetric rank*.

Example

Let $A = \{a, b, c, d\}$ and let θ be the involution which exchanges b, d and fixes a, c, i.e.,

$$G_\theta = \langle a, b, c, d \mid a^2 = c^2 = bd = db = 1 \rangle.$$

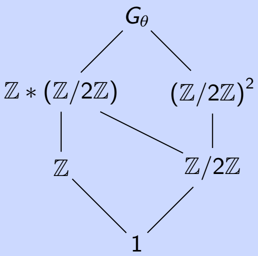
$G_\theta = \mathbb{Z} * (\mathbb{Z}/2\mathbb{Z})^2$ is a specular group of type $(1, 2)$ and symmetric rank 4.

Theorem

Any subgroup of a specular group is specular.

Example

Let $G_\theta = \mathbb{Z} * (\mathbb{Z}/2\mathbb{Z})^2$, then one has



A word is θ-*reduced* if it has no factor of the form $a\theta(a)$ for $a \in A$.

Any element of a specular group is represented by a unique reduced word.

Example

Let θ be the involution on the alphabet $\{a, b, c, d\}$ that fixes a, c and exchanges b, d.

The θ-reduction of the word $daaacbd$ is dac.

A word is *θ-reduced* if it has no factor of the form $a\theta(a)$ for $a \in A$.

Any element of a specular group is represented by a unique reduced word.

Example

Let θ be the involution on the alphabet $\{a, b, c, d\}$ that fixes a, c and exchanges b, d.

The θ-reduction of the word $d\not{a}\not{a}ac\not{b}\not{d}$ is dac.

A subset of a group G is called *symmetric* if it is closed under taking inverses (under θ).

Example

The set $X = \{a, adc, b, cba, d\}$ is symmetric, for $\theta : b \leftrightarrow d$ fixing a, c.

A subset of a group G is called *symmetric* if it is closed under taking inverses (under θ).

Example

The set $X = \{a, adc, b, cba, d\}$ is symmetric, for $\theta : b \leftrightarrow d$ fixing a, c.

A set X in a specular group G is called a *monoidal basis* of G if :

- it is symmetric ;
- the monoid that it generates is G ;
- any product $x_1 x_2 \cdots x_m$ such that $x_k x_{k+1} \neq 1$ for every k is distinct of 1.

Example

The alphabet A is a monoidal basis of G_θ.

The *symmetric rank* of a specular group is the cardinality of any monoidal basis.

Outline

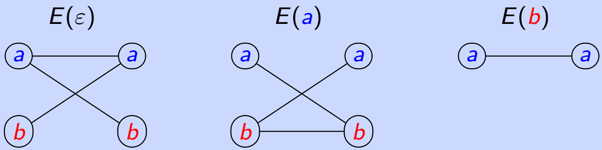
Let S be a factorial over an alphabet A.
The *extension graph* of a word $w \in S$ is the undirected bipartite graph $G(w)$ with vertices the disjoint union of

$$L(w) = \{a \in A \mid aw \in S\} \quad \text{and} \quad R(w) = \{a \in A \mid wa \in S\},$$

and edges the pairs $E(w) = \{(a, b) \in A \times A \mid awb \in S\}$.

Example

The *Fibonacci set* is the set of factors of the Fibonacci word, i.e. the fixed point $\varphi^\omega(a)$ of the morphism $\varphi : a \mapsto ab, b \mapsto a$.



Indeed one has $S = \{\varepsilon,\ a, b,\ aa, ab, ba,\ aab, aba, baa, bab, \ldots\}$.

A biextendable set S is called a *tree set* of *characteristic* c if for any nonempty $w \in S$, the graph $E(w)$ is a tree (acyclic and connected) and if $E(\varepsilon)$ is a union of c trees.

Example

The Fibonacci set is a tree set of characteristic 1.

A biextendable set S is called a *tree set* of *characteristic* c if for any nonempty $w \in S$, the graph $E(w)$ is a tree (acyclic and connected) and if $E(\varepsilon)$ is a union of c trees.

Example

The Fibonacci set is a tree set of characteristic 1.

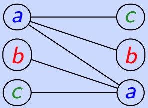
Proposition [Berthé, De Felice, D., Leroy, Perrin, Reutenauer, Rindone (2014, 2015)]

Factors of an Arnoux-Rauzy word and regular interval exchange sets are both uniformly recurrent tree sets of characteristic 1.

Example

The Tribonacci set is a tree set of characteristic 1.

$E(\varepsilon)$



A *specular set* on an alphabet A (w.r.t. an involution θ) is a

- biextendable and
- symmetric set
- of θ-reduced words
- which is a tree set of characteristic 2.

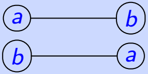
A *specular set* on an alphabet A (w.r.t. an involution θ) is a

- biextendable and
- symmetric set
- of θ-reduced words
- which is a tree set of characteristic 2.

Example

Let $A = \{a, b\}$ and θ be the identity on A. The set of factors of $(ab)^\omega$ is a specular set.

$E(\varepsilon)$



Proposition [J. Cassaigne (1997)]

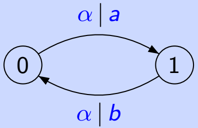
The factor complexity of a specular set is given by $p_0 = 1$ and $p_n = n\,(\mathrm{Card}\,(A) - 2) + 2$.

A *doubling transducer* is a transducer with set of states $Q = \{0, 1\}$ on the input alphabet Σ and the output alphabet A such that :

1. the input automaton is a group automaton, that is, every letter of Σ acts on Q as a permutation,
2. the output labels of the edges are all distinct.

Example

$\Sigma = \{\alpha\}$
$A = \{a, b\}$



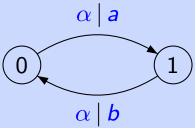
A *doubling transducer* is a transducer with set of states $Q = \{0, 1\}$ on the input alphabet Σ and the output alphabet A such that :

1. the input automaton is a group automaton, that is, every letter of Σ acts on Q as a permutation,
2. the output labels of the edges are all distinct.

A *doubling map* is a pair $\delta = (\delta_0, \delta_1)$, where $\delta_0, \delta_1 : \Sigma^* \to A^*$ are two maps such that $\delta_i(u) = v$ is the path starting at the state i with input label u and output label v.

Example

$\Sigma = \{\alpha\}$
$A = \{a, b\}$



$\delta_0(\alpha^\omega) = (ab)^\omega$
$\delta_1(\alpha^\omega) = (ba)^\omega$

A *doubling transducer* is a transducer with set of states $Q = \{0, 1\}$ on the input alphabet Σ and the output alphabet A such that :

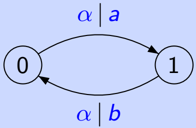
1. the input automaton is a group automaton, that is, every letter of Σ acts on Q as a permutation,
2. the output labels of the edges are all distinct.

A *doubling map* is a pair $\delta = (\delta_0, \delta_1)$, where $\delta_0, \delta_1 : \Sigma^* \to A^*$ are two maps such that $\delta_i(u) = v$ is the path starting at the state i with input label u and output label v.

The *image* of a set T by a doubling map is the set $\delta(T) = \delta_0(T) \cup \delta_1(T)$.

Example

$\Sigma = \{\alpha\}$
$A = \{a, b\}$



$\delta_0(\alpha^\omega) = (ab)^\omega$
$\delta_1(\alpha^\omega) = (ba)^\omega$

$\delta(\alpha^\omega) = (ab)^\omega \cup (ba)^\omega$

Proposition

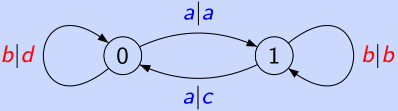
The image of a tree set of characteristic 1 closed under reversal by a doubling map is a specular set.

Example

Two possible *doublings* of the Fibonacci set are :

- the set of factors of the two infinite sequences $abaababa\cdots$ and $cdccdcdc\cdots$,

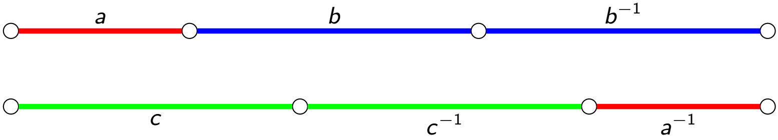
- the set of factors of the two infinite sequences $abcabcda\cdots$ and $cdacdabc\cdots$.



Both are specular sets. Their factor complexity is $2n+2$.

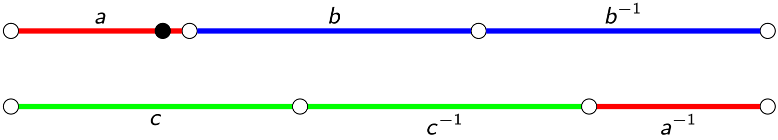
Theorem

The natural coding of a linear involution without connections is a specular set.



Theorem

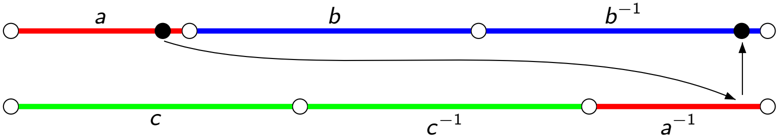
The natural coding of a linear involution without connections is a specular set.



$$\Sigma_T(z) = a$$

Theorem

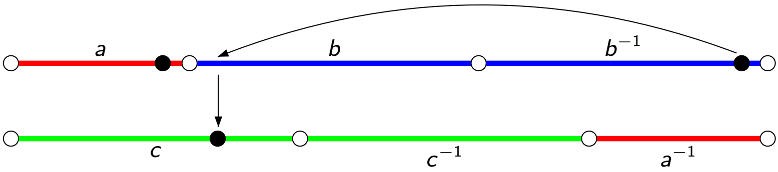
The natural coding of a linear involution without connections is a specular set.



$$\Sigma_T(z) = a\, b^{-1}$$

Theorem

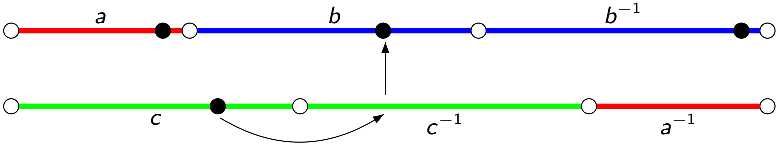
The natural coding of a linear involution without connections is a specular set.



$\Sigma_T(z) = a\, b^{-1}\, c$

Theorem

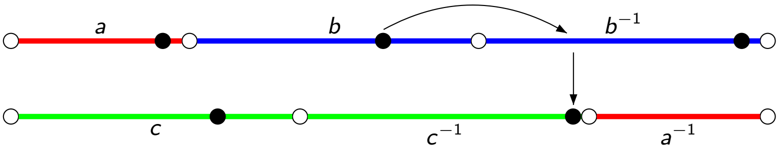
The natural coding of a linear involution without connections is a specular set.



$\Sigma_T(z) = a\, b^{-1}\, c\, b$

Theorem

The natural coding of a linear involution without connections is a specular set.

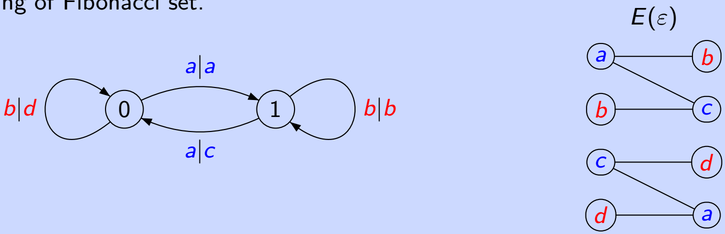


$$\Sigma_T(z) = a\, b^{-1}\, c\, b\, c^{-1} \cdots$$

A letter is said to be *even* if its two occurences (as a element of $L(\varepsilon)$ and of $R(\varepsilon)$) appear in the same tree of $E(\varepsilon)$. Otherwise it is said to be *odd*.

Example

Doubling of Fibonacci set.



The letters b, d are even, while the letters a, c are odd.

A word is said to be *even* if it has an even number of odd letters. Otherwise it is said to be *odd*.

Outline

A set $X \subset A^+$ of nonempty words over an alphabet A is a *bifix code* if it does not contain any proper prefix or suffix of its elements.

Example

- $\{aa, ab, ba\}$
- $\{aa, ab, bba, bbb\}$
- $\{ac, bcc, bcbca\}$

A set $X \subset A^+$ of nonempty words over an alphabet A is a *bifix code* if it does not contain any proper prefix or suffix of its elements.

Example

- $\{aa, ab, ba\}$
- $\{aa, ab, bba, bbb\}$
- $\{ac, bcc, bcbca\}$

A bifix code $X \subset S$ is *S-maximal* if it is not properly contained in a bifix code $Y \subset S$.

Example

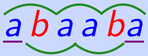
Let S be the Fibonacci set. The set $X = \{aa, ab, ba\}$ is an S-maximal bifix code. It is not an A^*-maximal bifix code, indeed $X \subset Y = X \cup \{bb\}$.

A *parse* of a word w with respect to a bifix code X is a triple (q, x, p) with $w = qxp$ and such that q has no suffix in X, $x \in X^*$ and p has no prefix in X.

Example

Let $X = \{aa, ab, ba\}$ and $w = abaaba$. The two possible parses of w are

- $(\varepsilon, ab\,aa\,ba, \varepsilon)$,
- $(a, ba\,ab, a)$.

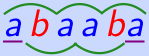


A *parse* of a word w with respect to a bifix code X is a triple (q, x, p) with $w = qxp$ and such that q has no suffix in X, $x \in X^*$ and p has no prefix in X.

Example

Let $X = \{aa, ab, ba\}$ and $w = abaaba$. The two possible parses of w are

- $(\varepsilon, ab\,aa\,ba, \varepsilon)$,
- $(a, ba\,ab, a)$.



The S-*degree* of X is the maximal number of parses with respect to X of a word of S.

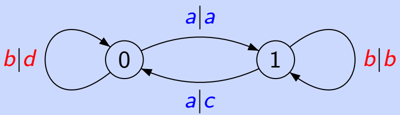
Example

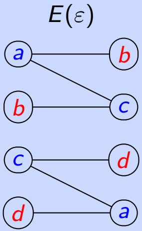
- For the Fibonacci set S, the set $X = \{aa, ab, ba\}$ has S-degree 2
- The set $X = S \cap A^n$ has S-degree n.

The set of even words in a specular set S has the form $X^* \cap S$, where $X \subset S$ is a bifix code called the *even code*.
The set X is the set of even words without a nonempty even prefix (or suffix).

Example

Doubling of Fibonacci set.





The even code is $X = \{abc, ac, b, ca, cda, d\}$.

Proposition

The even code of a recurrent specular set S is an S-maximal bifix code of S-degree 2.

Let S be a factorial set and X be a finite S-maximal bifix code.
A *coding morphism* for X is a morphism $f : B^* \to A^*$ which maps bijectively an alphabet B onto X.

The set $f^{-1}(S)$ is called a *maximal bifix decoding* of S.

Let S be a factorial set and X be a finite S-maximal bifix code.
A *coding morphism* for X is a morphism $f : B^* \to A^*$ which maps bijectively an alphabet B onto X.

The set $f^{-1}(S)$ is called a *maximal bifix decoding* of S.

Maximal Bifix Decoding Theorem

The decoding of a uniformly recurrent specular set by the even code is a union of two uniformly recurrent tree sets of characteristic 1.

Example

The set $S = \text{Fac}\left((ab)^\omega\right)$ is a specular set. Its even code is $X = \{ab, ba\}$.
Let us consider the coding morphism for X

$$f : \begin{cases} u \mapsto ab \\ v \mapsto ba \end{cases}$$

Then, $f^{-1}(S) = \text{Fac}\left(u^\omega\right) \cup \text{Fac}\left(v^\omega\right)$.

Finite Index Basis Theorem

Let S be a uniformly recurrent specular set and $X \subset S$ a finite symmetric bifix code. X is an S-maximal bifix code of S-degree d if and only if it is a monoidal basis of a subgroup of index d.

Example

- $S \cap A^n$.
- The even code is a monoidal basis of a subgroup of index 2 of G_θ called the *even subgroup*.

Finite Index Basis Theorem

Let S be a uniformly recurrent specular set and $X \subset S$ a finite symmetric bifix code.
X is an S-maximal bifix code of S-degree d if and only if it is a monoidal basis of a subgroup of index d.

Example

- $S \cap A^n$.
- The even code is a monoidal basis of a subgroup of index 2 of G_θ called the *even subgroup*.

The Finite Index Basis Theorem has also a converse.

Theorem

Let S be a recurrent and symmetric set of reduced words having factor complexity $p_n = n\,(\mathrm{Card}\,(A) - 2) + 2$.
If $S \cap A^n$ is a monoidal basis of the subgroup $\langle A^n \rangle$ for all $n \geq 1$, then S is a specular set.

Let S be a factorial set of words and $x \in S$.
A (*right*) *return word* to x in S is a nonempty word u such that $xu \in S \cap A^* x$, but has no internal factor equal to x.

We denote by $\mathcal{R}_S(w)$ the set of return words to x in S.

Example

Let S be the Fibonacci set. One has $\mathcal{R}_S(aa) = \{baa, babaa\}$.

$$\varphi(a)^\omega = ab\underline{aa}bab\underline{aa}baababaababaab\underline{aa}babaabaab\cdots$$

Remark. A recurrent set S is uniformly recurrent if and only if the set $\mathcal{R}_S(w)$ is finite for every $w \in S$.

Theorem [Balková, Palentová, Steiner (2008)]

Let S be a uniformly recurrent tree set of characteristic 1.
For every $w \in S$, the set $\mathcal{R}_S(w)$ has exactly $\text{Card}\,(A)$ elements.

Theorem [Balková, Palentová, Steiner (2008)]

Let S be a uniformly recurrent tree set of characteristic 1.
For every $w \in S$, the set $\mathcal{R}_S(w)$ has exactly $\mathrm{Card}\,(A)$ elements.

Theorem [Berthé, De Felice, D., Leroy, Perrin, Reutenauer, Rindone (2015)]

Let S be a uniformly recurrent tree set of characteristic 1.
For every $w \in S$, the set $\mathcal{R}_S(w)$ is a (tame) basis of the free group on A.

Theorem [Balková, Palentová, Steiner (2008)]

Let S be a uniformly recurrent tree set of characteristic 1.
For every $w \in S$, the set $\mathcal{R}_S(w)$ has exactly $\mathrm{Card}\,(A)$ elements.

Theorem [Berthé, De Felice, D., Leroy, Perrin, Reutenauer, Rindone (2015)]

Let S be a uniformly recurrent tree set of characteristic 1.
For every $w \in S$, the set $\mathcal{R}_S(w)$ is a (tame) basis of the free group on A.

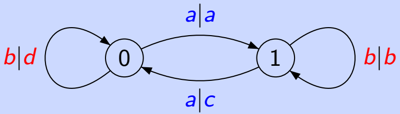
Return Theorem

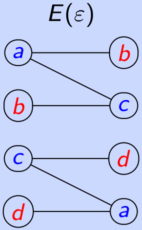
Let S be a uniformly recurrent specular set on the alphabet A.
For any $w \in S$, the set $\mathcal{R}_S(w)$ is a basis of the even subgroup.

In particular, $\mathrm{Card}\,(\mathcal{R}_S(x)) = \mathrm{Card}\,(A) - 1$.

Example

Let $G_\theta = \langle a, b, c, d \mid a^2 = c^2 = bd = 1 \rangle$ and S be the doubling of the Fibonacci set :





The even code is $X = \{abc, ac, b, ca, cda, d\}$,
while $\mathcal{R}_S(a) = \{bca, bcda, cda\}$.

Then, $\langle \mathcal{R}_S(a) \rangle = \langle X \rangle$, indeed :

$$\begin{cases} cda = cda & ca = (b)^{-1}(bca) \\ abc = (cda)^{-1} & ac = (ca)^{-1} \\ b = (bcda)(abc) & d = b^{-1} \end{cases}$$

Conclusions

Quick summary for those who fell asleep (wake up : it's lunch time!)

- Introduction of specular groups and specular sets.
- Generalization within these sets of results holding for tree sets.

Conclusions

Quick summary for those who fell asleep (wake up : it's lunch time!)

- Introduction of specular groups and specular sets.
- Generalization within these sets of results holding for tree sets.

Further research directions

- Investigation about recurrence (uniformly recurrence and tree condition, bifix decoding, . . .).
- Interesting connection with G-full (or G-rich) words.
- Generalization towards larger classes of groups (virtually free).


Danke !