On the group of a rational maximal bifix code

Alfredo Costa, Centre for Mathematics, University of Coimbra joint work with Jorge Almeida, Revekka Kyriakoglou & Dominique Perrin

17e journées montoises d'informatique théorique — 2018

 \mathbb{C}

Centre for Mathematics University of Coimbra

• <u>Relativization</u> of maximal bifix codes:

filtration

$$X = Z \cap F$$

of a maximal bifix code Z by a (uniformly) recurrent set F

- Relativization of descriptors of maximal bifix codes:
 - $d(Z) \rightsquigarrow d_F(X)$
 - $G(Z) \rightsquigarrow G_F(X)$
- We give necessary and sufficient conditions for

 $G(Z) \cong G_F(X)$

• Methodological novelty: the use of free profinite monoids

Alfredo Costa

▲日▼ ▲圖▼ ▲目▼ ▲目▼ ■ ●のの⊙

The syntactic monoid M(L) of a language L is the transition monoid of the minimal automaton of L

We denote the map $A^* \to M(L)$ by η_L

e.g.,
$$L = \{aa, ab, ba\}^{*}$$





M(L):

$^*ab^2a$	ab^2a^2	ab^2
a^2b^2a	$a^2b^2a^2$	$^{*a^{2}b^{2}}$
b^2a	$^{*b^{2}a^{2}}$	b^2
	*0	

Green's relations

- $u \mathcal{R} v \Leftrightarrow uM = vM$
- $u \mathcal{L} v \Leftrightarrow Mu = Mv$
- $u \leq_{\mathcal{J}} v \Leftrightarrow MuM \subseteq MvM$
- In a finite monoid,

 $\mathcal{J} = \mathcal{R} \circ \mathcal{L} = \mathcal{L} \circ \mathcal{R} = \mathcal{R} \lor \mathcal{L}$

All monoids we consider have this property.

For such monoids, if a *J*-class contains an idempotent, then its *H*-classes containing idempotents are isomorphic subgroups.

The abstract semigroup thus defined is the Schützenberger group of the \mathcal{J} -class.





A B F A B F

Green's relations

- $u \mathcal{R} v \Leftrightarrow uM = vM$
- $u \mathcal{L} v \Leftrightarrow Mu = Mv$
- $u \leq_{\mathcal{J}} v \Leftrightarrow MuM \subseteq MvM$
- In a finite monoid,

 $\mathcal{J} = \mathcal{R} \circ \mathcal{L} = \mathcal{L} \circ \mathcal{R} = \mathcal{R} \lor \mathcal{L}$

All monoids we consider have this property.

For such monoids, if a *J*-class contains an idempotent, then its *H*-classes containing idempotents are isomorphic subgroups.

The abstract semigroup thus defined is the Schützenberger group of the \mathcal{J} -class.

Alfredo Costa

*1				
$*a^{2}$	*ab			
ab^2a	a^2b^3			
a	a^2b			
ab^2a^2	ab^2			
a^2b^2a	ab^3			
ab^3a	a^2b^2			
*ba	ba^2b			
$b^{3}a^{2}$	$^{*}b^{3}$			
ba^2	b			
b^3a	ba^2b^2			
b^2a	b^2a^2b			
b^2a^2	b^2			

・ロト ・行下・ ・ヨト ・ ヨト

A bifix code X of A^* is maximal if

$$X \subseteq Y$$
 and Y is bifix $\implies X = Y$

Motivation

- Perhaps the most studied and "tractable" class of codes.
- Important case: if $M(X^*)$ is a group, then X is a maximal bifix code, called a group code.

A bifix code X is *F*-maximal if $X \subseteq F$ and

 $X \subseteq Y \subseteq F$ and Y is bifix $\implies X = Y$

Theorem (Berstel & De Felice & Perrin & Reutenauer & Rindone; 2012)

If Z is maximal bifix, then $X = Z \cap F$ is...

- ... <u>F-maximal bifix</u> if F is recurrent
- ... <u>finite</u> if F is uniformly recurrent

If X is a <u>rational</u> maximal bifix code, then the degree of X is the rank d(X) of the minimum ideal J(X) of $M(X^*)$.

Theorem (The five authors of the 2012 paper + Dolce & Leroy; 2015)

If Z is a maximal bifix code and F is a tree set, then $X = Z \cap F$ is a basis of a subgroup of index d(Z) of the free group FG(A).

・ロト ・行下・ ・ヨト ・ ヨト

Let X be a <u>rational</u> maximal bifix code.

The *F*-minimum \mathcal{J} -class of $M(X^*)$, denoted $J_F(X)$, is the \mathcal{J} -class of $M(X^*)$ that is \mathcal{J} -minimum among the \mathcal{J} -classes intersecting $\eta_{X^*}(F)$.

The *F*-degree of X is the rank $d_F(X)$ of $J_F(X)$.

The group of Z, denoted G(Z), is the Schützenberger group of J(Z).

The *F*-group of *X*, denoted $G_F(X)$, is the Schützenberger group of $J_F(X)$.

- G(Z) is a permutation group of degree d(Z)
- $G_F(X)$ is a permutation group of degree $d_F(X)$

We want to relate G(Z) with $G_F(X)$...

Theorem (Berstel & De Felice & Perrin & Reutenauer & Rindone; 2012)

If Z is a group code and F is Sturmian, then $G(Z) \simeq G_F(X)$ and $d(Z) = d_F(X)$.

(the theorem does not hold if we just replace "Sturmian" by "uniformly recurrent", or "group code" by "maximal bifix code")

Alfredo Costa

The group of Z, denoted G(Z), is the Schützenberger group of J(Z).

The *F*-group of *X*, denoted $G_F(X)$, is the Schützenberger group of $J_F(X)$.

- G(Z) is a permutation group of degree d(Z)
- $G_F(X)$ is a permutation group of degree $d_F(X)$

We want to relate G(Z) with $G_F(X)$...

Theorem (Berstel & De Felice & Perrin & Reutenauer & Rindone; 2012)

If Z is a group code and F is Sturmian, then $G(Z) \simeq G_F(X)$ and $d(Z) = d_F(X)$.

(the theorem does not hold if we just replace "Sturmian" by "uniformly recurrent", or "group code" by "maximal bifix code")

Alfredo Costa



- $Z = \{aa, ab, ba, bb\}$
- *F* = "Fibonacci set"

Minimum automaton of X^* , where $X = Z \cap F$:



	*	*1	
	a^{*a^2}	*ab a^2b	,
	*ba ba^2	ba ² l	6
			_
$^*ab^2a$	ab	$^{2}a^{2}$	ab^2
a^2b^2a	$*a^{2}l$	$b^{2}a^{2}$	$^{*a^{2}b^{2}}$
b^2a	*b2	$^{2}a^{2}$	b^2

*0

イロト 不得下 イヨト イヨト

$$G(Z)\simeq G_F(X)\simeq \mathbb{Z}/2\mathbb{Z}$$

3

The profinite completion of A^*

If x, y are distinct elements of A^* , then there is a finite quotient $M = A^*/\sim$ separating x and y (i.e. $x \not\sim y$).

Let r(x, y) be the smallest possible cardinal for M.

$$d(x,y)=2^{-r(x,y)}$$

For this metric, let $\widehat{A^*}$ be the metric completion of A^* .

 $\widehat{A^*}$ is a topological monoid, with a profinite topology.

More: $\widehat{A^*}$ is the free profinite monoid generated by A.

If F is recurrent, then there is a \mathcal{J} -minimum \mathcal{J} -class contained in \overline{F} , denoted J(F).

The group G(F) is the (topological!) Schützenberger group of J(F).

▲日▼ ▲冊▼ ▲目▼ ▲目▼ 目 ろの⊙

Theorem (Almeida & Costa; 2017)

If F is a tree set, then G(F) is a free profinite group of rank |A|.

(the proof uses the "Return Theorem")

More precisely:

if F is a tree set, then $\pi \mid : G(F) \to \widehat{FG(A)}$ is an isomorphism:



More generally, $\pi \mid : G(F) \to \widehat{FG(A)}$ is onto if F is connected.

▲圖 ▶ ▲ 臣 ▶ → 臣 ▶ …

We say that a code X is:

• F-charged if

$$\hat{\eta}_{X^*}(G(F))=G(X)$$



A B F A B F

weakly F-charged if

 $\hat{\eta}_{X^*}(G(F))=G_F(X)$

э

We say that a code X is:

• F-charged if

$$\hat{\eta}_{X^*}(G(F))=G(X)$$



A B F A B F

weakly F-charged if

 $\hat{\eta}_{X^*}(G(F))=G_F(X)$

э

Necessary and sufficient conditions for $G(Z) \simeq G_F(Z \cap F)$

Theorem (Almeida & Costa & Kyriakoglou & Perrin; 2018)

Let:

- F recurrent
- Z rational maximal bifix code

Suppose also that $X = Z \cap F$ is rational.

The following conditions are equivalent:

- Z is F-charged
- $d_F(X) = d(Z)$, $G_F(X) \simeq G(Z)$ and X is weakly F-charged

Additionally: if F is uniformly recurrent, then the equality $d_F(X) = d(Z)$ is redundant in the second condition.

- **(())) (()))**

Necessary and sufficient conditions for $G(Z) \simeq G_F(Z \cap F)$

Theorem (Almeida & Costa & Kyriakoglou & Perrin; 2018)

Let:

- F recurrent
- Z rational maximal bifix code

Suppose also that $X = Z \cap F$ is rational.

The following conditions are equivalent:

- Z is F-charged
- $d_F(X) = d(Z)$, $G_F(X) \simeq G(Z)$ and X is weakly F-charged

Additionally: if F is uniformly recurrent, then the equality $d_F(X) = d(Z)$ is redundant in the second condition.





When $F = F_{\varphi}$ is described by a primitive substitution φ , we have an algorithm to decide if X is (weakly) *F*-charged: this is done via a profinite presentation for $G(F_{\varphi})$, obtained by Almeida & Costa (2013).

Example

• F_{τ} : the Prouhet-Thue-Morse set, where

 $\tau: a \mapsto ab, b \mapsto ba$

• Z: group code over $\{a, b\}$ generating the stabilizer of 1 via

$$a \mapsto (123), \quad b \mapsto (345)$$

- $G(Z) = A_5$
- Z is F_{τ} -charged

•
$$Z = \{aa, ab, ba\} \cup b^2(a^+b)^*b$$

•
$$F = A^* \setminus A^*ab(b^2)^*aA^*$$

Z is maximal bifix, but not a group code *F* is recurrent, but not uniformly

Z is F-charged and $G(Z) \simeq G_{F}(X) \simeq S_{3}$

*1				
*a2	*ab			
ab^2a	$a^{2}b^{3}$			
a	a^2b			
ab^2a^2	ab^2			
a^2b^2a	ab^3			
ab^3a	a^2b^2			
*ba	ba^2b			
$b^{3}a^{2}$	$*b^{3}$			
ba^2	b			
b^3a	ba^2b^2			
b^2a	b^2a^2b			
$b^{2}a^{2}$	b^2			

	Ϋ́	
	Ь	
	Т	
$*b^4a$	$b^{2}a^{2}b^{2}$	$b^{4}a^{2}b$
$b^{3}a^{2}$	*63	$b^{3}ab$
$b^{4}a^{2}$	b^4	b^4ab
b^3a	$b^{2}a^{2}b^{3}$	$b^{3}a^{2}b$
b^2a	$b^{3}a^{2}b^{2}$	$b^{2}a^{2}b$
$b^{2}a^{2}$	b^2	b^2ab
*a ²	*ab4	ab
ab^2a	$a^{2}b^{3}$	ab^2a^2b
4	$a^{2}b^{4}$	a^2b
ab^2a^2	ab^2	ab^2ab
a^2b^2a	ab ³	ab^3ab
$ab^{3}a$	$a^{2}b^{2}$	a^2b^2ab
ba	ba^2b^4	ba^2b
ba^2b^2a	*bab ³	$*bab^{3}ab$
ba ²	bab4	bab
$bab^{3}a$	ba^2b^2	ba^2b^2ab
bab^2a	ba^2b^3	bab^2a^2b
bab^2a^2	bab^2	bab^2ab

Arbite Arbite<								
	a^2ba	*abab	a^2bab^3	$abab^2a$	$abab^2a^2b$	$abab^2a^2$	abab ²	$abab^2ab$
مالي <th< td=""><td>aba</td><td>$a^{2}bab$</td><td>abab³</td><td>a^2bab^2a</td><td>abab³ab</td><td>abab³a</td><td>a^2bab^2</td><td>a²bab²ab</td></th<>	aba	$a^{2}bab$	abab ³	a^2bab^2a	abab ³ ab	abab ³ a	a^2bab^2	a ² bab ² ab
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	*baba	ba^2bab	*babab ³	ba^2bab^2a	*babab ³ ab	babab ³ a	ba ² bab ²	ba ² bab ² ab
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	ba"ba	babab	ba"bab"	babab" a	babab"a"b	babab"a"	babab"	babab"ab
$\frac{1}{2}$ $\frac{1}$	b^3a^2ba	*b ³ abab	$b^3a^2bab^3$	*b ³ abab ² a	$b^3abab^2a^2b$	b ³ abab ² a ²	b ³ abab ²	b ³ abab ² ab
$h^{2} \phi^{2} h^{2} h^{2} = h^{2} \phi^{2} h^{2} h^$	b^3aba	b^3a^2bab	b ³ abab ³	$b^3a^2bab^2a$	b ³ abab ³ ab	b ³ abab ³ a	$b^3a^2bab^2$	$b^3a^2bab^2ab$
$\frac{bah^2aba}{ba} \frac{ba^2k^2abab}{ba^2abab^2} \frac{bah^2abab^2a}{ba^2k^2abab^2a} \frac{ba^2k^2abab^2a^2}{ba^2k^2abab^2a^2} \frac{ba^2k^2abab^2}{ba^2k^2abab^2} ba^2k^2ab$	ba ² b ² aba	*bab ³ abab	$ba^2b^2abab^3$	bab ³ abab ² a	$ba^2b^2abab^3ab$	ba ² b ² abab ³ a	bab ³ abab ²	bab^3abab^2ab
ab^2aba ab^2a^2bab $*ab^2abab^3$ $*ab^2a^2bab^2a$ ab^2abab^3ab ab^2abab^3a $ab^2a^2bab^2$ $ab^2a^2bab^2ab$	bab ³ aba	ba^2b^2abab	bab ³ abab ³	ba ² b ² abab ² a	$ba^2b^2abab^2a^2b$	$ba^2b^2abab^2a^2$	$ba^2b^2abab^2$	$ba^2b^2abab^2ab$
-1 ² - ² h = 1 ² -1+1 = 1 ² - ² h = 3 = 1 ² -1+2 ² = 1 ² -1+2 ² - ² h = 1 ² -1+2 ² - ² = 1 ² -1+2 ² = 1 ² -1+2 ² -1	ab^2aba	ab^2a^2bab	*ab ² abab ³	*ab ² a ² bab ² a	ab ² abab ³ ab	ab ² abab ³ a	$ab^2a^2bab^2$	$ab^2a^2bab^2ab$
	ab^2a^2ba	ab^2abab	$ab^2a^2bab^3$	ab^2abab^2a	ab ² abab ² a ² b	$ab^2abab^2a^2$	ab^2abab^2	ab^2abab^2ab
$b^{2}aba$ $b^{2}a^{2}bab$ $b^{2}abab^{3}$ $b^{2}a^{2}bab^{2}a$ $b^{2}abab^{3}ab$ $*b^{2}abab^{3}a$ $b^{2}a^{2}bab^{2}$ $b^{2}a^{2}bab^{2}ab$	b^2aba	b^2a^2bab	b ² abab ³	$b^2a^2bab^2a$	b ² abab ³ ab	*b ² abab ³ a	$b^2a^2bab^2$	b ² a ² bab ² ab
b^2a^2ba b^2abab $b^2a^2bab^3$ b^2abab^2a $b^2abab^2a^2b$ $b^2abab^2a^2$ $*b^2abab^2$ $*b^2abab^2ab$	b^2a^2ba	b^2abab	$b^2 a^2 b a b^3$	b^2abab^2a	$b^2abab^2a^2b$	$b^2abab^2a^2$	$*b^2abab^2$	$*b^2abab^2ab$
$a^{2}b^{2}aba$ $ab^{3}abab$ $a^{2}b^{2}abab^{3}$ $ab^{3}abab^{2}a$ $a^{2}b^{2}abab^{3}ab$ $*a^{2}b^{2}abab^{3}a$ $*ab^{3}abab^{2}$ $ab^{3}abab^{2}ab$	a^2b^2aba	ab ³ abab	$a^2b^2abab^3$	ab^3abab^2a	a ² b ² abab ³ ab	$*a^2b^2abab^3a$	*ab ³ abab ²	ab^3abab^2ab
$ab^{3}aba \ a^{2}b^{2}abab \ ab^{3}abab^{3} \ a^{2}b^{2}abab^{2}a \ a^{2}b^{2}abab^{2}a^{2}b \ a^{2}b^{2}abab^{2}a^{2} \ a^{2}b^{2}abab^{2}a^{2}$	ab^3aba	a^2b^2abab	ab^3abab^3	$a^2b^2abab^2a$	$a^2b^2abab^2a^2b$	$a^2b^2abab^2a^2$	$a^2b^2abab^2$	$a^2b^2abab^2ab$
$bab^2aba \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	bab ² aba	bab^2a^2bab	bab^2abab^3	bab ² a ² bab ² a	bab ² abab ³ ab	bab ² abab ³ a	$bab^2a^2bab^2$	$bab^2a^2bab^2ab$
bab^2a^2ba bab^2abab $bab^2a^2bab^3$ bab^2abab^2a $bab^2abab^2a^2b$ $bab^2abab^2a^2$ $*bab^2abab^2$ $*bab^2abab^2a^2$	bab^2a^2ba	bab^2abab	$bab^2a^2bab^3$	bab ² abab ² a	$bab^2abab^2a^2b$	$bab^2abab^2a^2$	bab^2abab^2	bab^2abab^2ab

*abab ² aba	abab ² a ² ba	abab ² abab	abab ² abab ²	abab ² abab ² a	abab ² abab ² ab
a ² bab ² aba	*abab ³ aba	*a ² bab ² abab	a ² bab ² abab ²	$a^2bab^2abab^2a$	a ² bab ² abab ² ab
babab ² aba	$*babab^2a^2ba$	babab ² abab	*babab ² abab ²	$babab^2abab^2a$	$*babab^2abab^2ab$
b ² abab ² aba	$b^2abab^2a^2ba$	*b ² abab ² abab	b ² abab ² abab ²	*b ² abab ² abab ² a	b ² abab ² abab ² ab
bab^2abab^2aba	$bab^2abab^2a^2ba$	*bab ² abab ² abab	$bab^2abab^2abab^2$	$bab^2abab^2abab^2a$	$bab^2abab^2abab^2ab$
ab ² abab ² aba	$ab^2abab^2a^2ba$	ab^2abab^2abab	ab ² abab ² abab ²	$ab^2abab^2abab^2a$	$ab^2abab^2abab^2ab$

Image: A matrix of the second seco

э

- - ≣ →

∃ ⊳.

- $Z = \{aa, ab, ba\} \cup b^2(a^+b)^*b$
- F = "Fibonacci set"
 - Z is maximal bifix, but not a group code
 - F is Sturmian
 - Z is not F-charged and $G(Z) \not\simeq G_{\mathbf{F}}(X)$

*1				
$*a^{2}$	*ab			
ab^2a	a^2b^3			
a	a^2b			
ab^2a^2	ab^2			
a^2b^2a	ab^3			
ab^3a	a^2b^2			
*ba	ba^2b			
$b^{3}a^{2}$	$^{*}b^{3}$			
ba^2	b			
b^3a	ba^2b^2			
b^2a	b^2a^2b			
b^2a^2	b^2			

	3		
	$*a^2$	*ab	
	a	a^2b	
	*ba	ba^2l	5
	ba^2	b	
$^*ab^2a$	ab	$^{2}a^{2}$	ab^2
a^2b^2a	$*a^{2}$	$b^{2}a^{2}$	$^{*a^{2}b^{2}}$
b^2a	*b	a^2a^2	b^2
	Я	°0	

イロト イロト イヨト イヨト

3

We gave the first relevant "external" applications of the profinite Schützenberger group G(F) of a uniformly recurrent set.

• the statement

 $G(Z) \cong G_F(Z \cap F)$ if F is connected and Z is group code

uses no "profinite jargon"

- the definition of "*F*-charged" gives a comprehensive framework to improve the latter
- the profinite group G(F) serves as a sort of universal cover for F-groups
- "profinite stuff" facilitates synthetic statements
- advantages of the profinite monoid for proofs:
 - more "conceptual" proofs (diagram style)
 - enhanced combinatorics ("pseudowords" can be idempotent!)

イロト イポト イヨト イヨト

We gave the first relevant "external" applications of the profinite Schützenberger group G(F) of a uniformly recurrent set.

• the statement

 $G(Z) \cong G_F(Z \cap F)$ if F is connected and Z is group code

uses no "profinite jargon"

- the definition of "*F*-charged" gives a comprehensive framework to improve the latter
- the profinite group G(F) serves as a sort of universal cover for F-groups
- "profinite stuff" facilitates synthetic statements
- advantages of the profinite monoid for proofs:
 - more "conceptual" proofs (diagram style)
 - enhanced combinatorics ("pseudowords" can be idempotent!)

イロト イポト イヨト イヨト

We gave the first relevant "external" applications of the profinite Schützenberger group G(F) of a uniformly recurrent set.

• the statement

 $G(Z) \cong G_F(Z \cap F)$ if F is connected and Z is group code

uses no "profinite jargon"

- the definition of "*F*-charged" gives a comprehensive framework to improve the latter
- the profinite group G(F) serves as a sort of universal cover for F-groups
- "profinite stuff" facilitates synthetic statements
- advantages of the profinite monoid for proofs:
 - more "conceptual" proofs (diagram style)
 - enhanced combinatorics ("pseudowords" can be idempotent!)

ヘロト 不得下 不可下 不可下

We gave the first relevant "external" applications of the profinite Schützenberger group G(F) of a uniformly recurrent set.

• the statement

 $G(Z) \cong G_F(Z \cap F)$ if F is connected and Z is group code

uses no "profinite jargon"

- the definition of "*F*-charged" gives a comprehensive framework to improve the latter
- the profinite group G(F) serves as a sort of universal cover for F-groups
- "profinite stuff" facilitates synthetic statements
- advantages of the profinite monoid for proofs:
 - more "conceptual" proofs (diagram style)
 - enhanced combinatorics ("pseudowords" can be idempotent!)

ヘロト 不得下 不可下 不可下

We gave the first relevant "external" applications of the profinite Schützenberger group G(F) of a uniformly recurrent set.

• the statement

 $G(Z) \cong G_F(Z \cap F)$ if F is connected and Z is group code

uses no "profinite jargon"

- the definition of "*F*-charged" gives a comprehensive framework to improve the latter
- the profinite group G(F) serves as a sort of universal cover for F-groups
- "profinite stuff" facilitates synthetic statements
- advantages of the profinite monoid for proofs:
 - more "conceptual" proofs (diagram style)
 - enhanced combinatorics ("pseudowords" can be idempotent!)

(ロト (四) (ヨト (ヨト) ヨー ショう

We gave the first relevant "external" applications of the profinite Schützenberger group G(F) of a uniformly recurrent set.

• the statement

 $G(Z) \cong G_F(Z \cap F)$ if F is connected and Z is group code

uses no "profinite jargon"

- the definition of "*F*-charged" gives a comprehensive framework to improve the latter
- the profinite group G(F) serves as a sort of universal cover for F-groups
- "profinite stuff" facilitates synthetic statements
- advantages of the profinite monoid for proofs:
 - more "conceptual" proofs (diagram style)
 - enhanced combinatorics ("pseudowords" can be idempotent!)

(ロト (四) (ヨト (ヨト) ヨー ショう

We gave the first relevant "external" applications of the profinite Schützenberger group G(F) of a uniformly recurrent set.

• the statement

 $G(Z) \cong G_F(Z \cap F)$ if F is connected and Z is group code

uses no "profinite jargon"

- the definition of "*F*-charged" gives a comprehensive framework to improve the latter
- the profinite group G(F) serves as a sort of universal cover for F-groups
- "profinite stuff" facilitates synthetic statements
- advantages of the profinite monoid for proofs:
 - more "conceptual" proofs (diagram style)
 - enhanced combinatorics ("pseudowords" can be idempotent!)

ロト 本得下 オヨト オヨト ヨー わらぐ