Codes in symbolic systems

Marie-Pierre Béal and Dominique Perrin

Institut Gaspard-Monge
Université de Marne-la-Vallée et CNRS

Developments in Language Theory Conference, 2005

Symbolic systems

Definitions

Any subshift on an alphabet A is a closed and shift invariant subset of $A^{\mathbb{Z}}$.

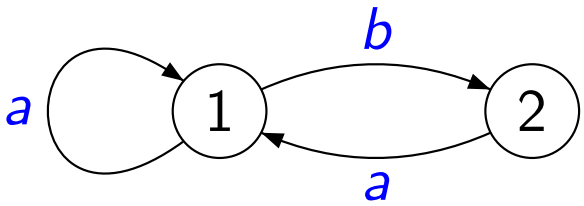
Any sofic shift is defined by a finite automaton (or a regular set of forbidden words).

A subshift of finite type is defined by a finite set of of forbidden blocks.

An irreducible sofic shift is defined by a finite automaton with a strongly connected graph.

The golden mean system

The golden mean system

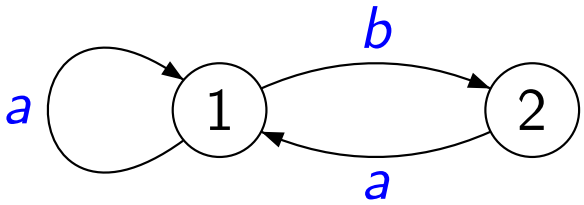


is a shift of finite type.

The factor bb is forbidden.

The golden mean system

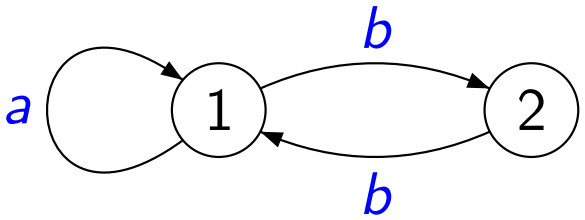
The golden mean system



is a shift of finite type.

The factor bb is forbidden.

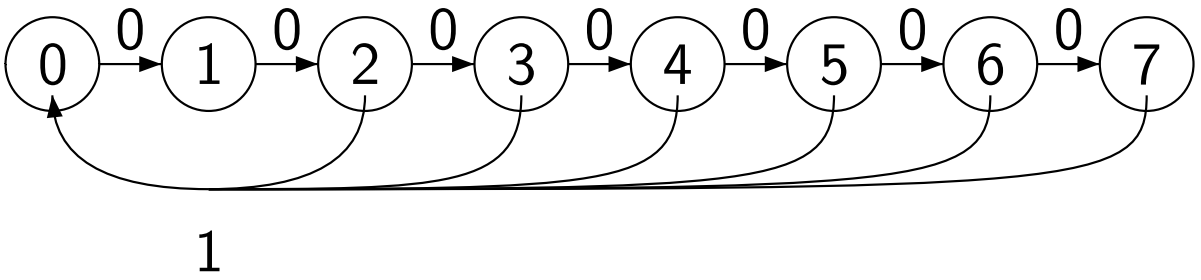
The even system



is a sofic shift.

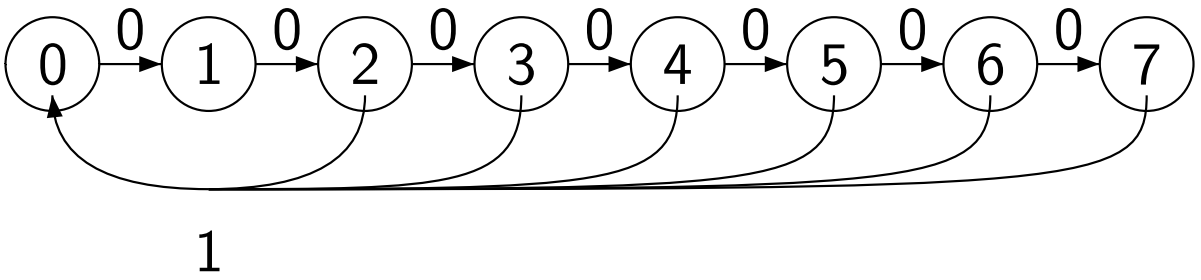
Factors $a(bb)^n ba$ are forbidden.

The [2,7] constraint



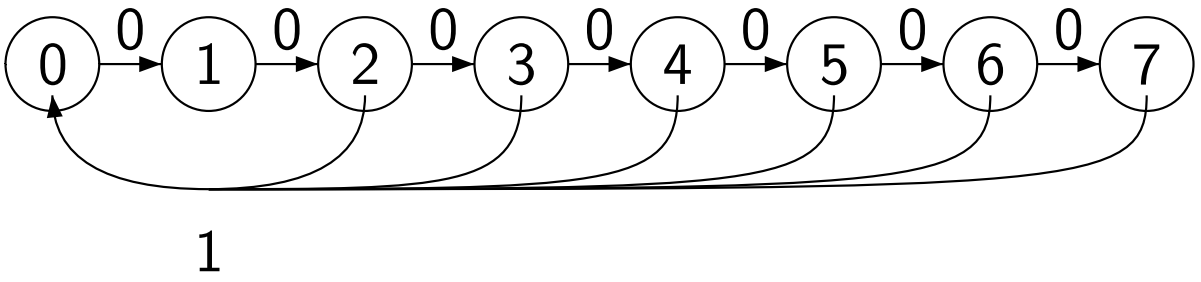
- A least 2 and at most 7 consecutives symbols 0.

The [2,7] constraint



- A least 2 and at most 7 consecutives symbols 0.
- The channel is defined by a finite list of forbidden blocks $\mathcal{F} = \{11, 101, 00000000\}$.

The [2,7] constraint



- A least 2 and at most 7 consecutives symbols 0.
- The channel is defined by a finite list of forbidden blocks $\mathcal{F} = \{11, 101, 00000000\}$.
- It is a shift of finite type.

Entropy

The entropy of a shift S is

$$h(S) = \lim_{n \to \infty} \frac{1}{n} \log u_n,$$

where u_n = number of admissible blocks of length n.

$$h(S) = -\log \rho_S.$$

- For the full shift on k symbols, $\rho_S = 1/k$.

Entropy

The entropy of a shift S is

$$h(S) = \lim_{n \to \infty} \frac{1}{n} \log u_n,$$

where u_n = number of admissible blocks of length n.

$$h(S) = -\log \rho_S.$$

- For the full shift on k symbols, $\rho_S = 1/k$.
- For the golden mean system, ρ_S is the inverse of the golden mean, i.e. such that $\rho_S^2 + \rho_S = 1$.

Zeta function

Let s_n be the number of points of a shift S of period dividing n:

$$s_n = \mathrm{Card}\{x \in S \mid \sigma^n(x) = x\}.$$

The zeta function of S is

$$\zeta(z) = \exp \sum_{n=1}^{\infty} \frac{s_n}{n} z^n.$$

Zeta function

- For an irreducible subshift of finite type realized by a minimal deterministic automaton $\mathcal{A}$ (its Shannon cover),

$$\zeta(z) = \det(I - Mz)^{-1},$$

where M is the adjacency matrix of $\mathcal{A}$ (Manning, 1971).

Zeta function

- For an irreducible subshift of finite type realized by a minimal deterministic automaton $\mathcal{A}$ (its Shannon cover),

$$\zeta(z) = \det(I - Mz)^{-1},$$

  where M is the adjacency matrix of $\mathcal{A}$ (Manning, 1971).
- For the golden mean system, $\zeta(z) = \frac{1}{1-z-z^2}$.

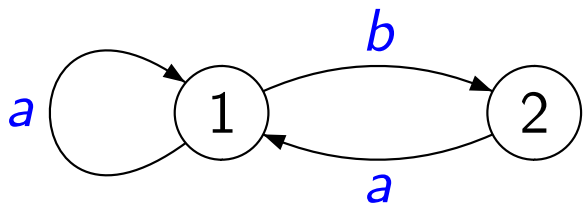
Zeta function

- For an irreducible subshift of finite type realized by a minimal deterministic automaton $\mathcal{A}$ (its Shannon cover),

$$\zeta(z) = \det(I - Mz)^{-1},$$

where M is the adjacency matrix of $\mathcal{A}$ (Manning, 1971).

- For the golden mean system, $\zeta(z) = \frac{1}{1-z-z^2}$.



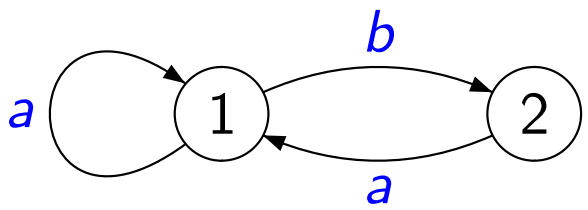
Zeta function

- For an irreducible subshift of finite type realized by a minimal deterministic automaton $\mathcal{A}$ (its Shannon cover),

$$\zeta(z) = \det(I - Mz)^{-1},$$

where M is the adjacency matrix of $\mathcal{A}$ (Manning, 1971).

- For the golden mean system, $\zeta(z) = \frac{1}{1-z-z^2}$.



$$M = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Codes

Let S be a subshift on A. Let Fact(S) denotes the factors of S. A set $X \subset \text{Fact}(S)$ is

- a code if X^* is unambiguous.

Codes

Let S be a subshift on A. Let $\mathrm{Fact}(S)$ denotes the factors of S. A set $X \subset \mathrm{Fact}(S)$ is

- a code if X^* is unambiguous.
- complete in S if $X \subset \mathrm{Fact}(S) \subset \mathrm{Fact}(X^*)$

Codes

Let S be a subshift on A. Let $\mathrm{Fact}(S)$ denotes the factors of S. A set $X \subset \mathrm{Fact}(S)$ is

- a code if X^* is unambiguous.
- complete in S if $X \subset \mathrm{Fact}(S) \subset \mathrm{Fact}(X^*)$
- caveat: complete $\not\Rightarrow$ maximal: $X = \{ab\}$ is complete in $(ab)^{\mathbb{Z}} \cup (ba)^{\mathbb{Z}}$ although $X \subset \{ab, ba\}$ which is a code.

Codes

Let S be a subshift on A. Let $\mathrm{Fact}(S)$ denotes the factors of S. A set $X \subset \mathrm{Fact}(S)$ is

- a code if X^* is unambiguous.
- complete in S if $X \subset \mathrm{Fact}(S) \subset \mathrm{Fact}(X^*)$
- caveat: complete $\not\Rightarrow$ maximal: $X = \{ab\}$ is complete in $(ab)^{\mathbb{Z}} \cup (ba)^{\mathbb{Z}}$ although $X \subset \{ab, ba\}$ which is a code.
- For example, $X = \{aa, ab, ba\}$ is a complete code in the golden mean system.

Codes

Let S be a subshift on A. Let $\mathrm{Fact}(S)$ denotes the factors of S. A set $X \subset \mathrm{Fact}(S)$ is

- a code if X^* is unambiguous.
- complete in S if $X \subset \mathrm{Fact}(S) \subset \mathrm{Fact}(X^*)$
- caveat: complete $\not\Rightarrow$ maximal: $X = \{ab\}$ is complete in $(ab)^{\mathbb{Z}} \cup (ba)^{\mathbb{Z}}$ although $X \subset \{ab, ba\}$ which is a code.
- For example, $X = \{aa, ab, ba\}$ is a complete code in the golden mean system.
- Note that we do not require $X^* \subset \mathrm{Fact}(S)$ as in Restivo (1990).

Codes

Let S be a subshift on A. Let $\mathrm{Fact}(S)$ denotes the factors of S. A set $X \subset \mathrm{Fact}(S)$ is

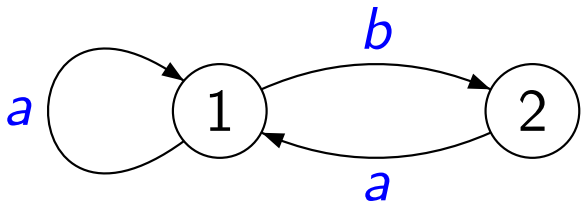
- a code if X^* is unambiguous.
- complete in S if $X \subset \mathrm{Fact}(S) \subset \mathrm{Fact}(X^*)$
- caveat: complete $\not\Rightarrow$ maximal: $X = \{ab\}$ is complete in $(ab)^{\mathbb{Z}} \cup (ba)^{\mathbb{Z}}$ although $X \subset \{ab, ba\}$ which is a code.
- For example, $X = \{aa, ab, ba\}$ is a complete code in the golden mean system.
- Note that we do not require $X^* \subset \mathrm{Fact}(S)$ as in Restivo (1990).
- Our definition captures also the codes of paths of Reutenauer (1986).

Matrix representations

In the sequel, S is an irreducible sofic shift recognized by its Shannon cover $\mathcal{A} = (Q, E)$. For $u \in A^*$ and $p, q \in Q$.

$$\mu(u)_{pq} = \begin{cases} u & \text{if } p \cdot u = q \\ 0 & \text{otherwise} \end{cases}$$

For the golden mean



$$\mu(A = \{a, b\}) = \begin{bmatrix} a & b \\ a & 0 \end{bmatrix}.$$

Series of matrices

Let π be the morphism obtained from μ through an assignment of positive real values to the elements of A.
For a set X, we denote

$$f_X(z) = \sum_{x \in X} \pi(x) z^{|x|}.$$

For the golden mean, if $\pi(a) = p, \pi(b) = q$,

$$\mu(A) = \begin{bmatrix} a & b \\ a & 0 \end{bmatrix}, \quad f_A(z) = \begin{bmatrix} pz & qz \\ pz & 0 \end{bmatrix}.$$

Series of matrices

- If X is a code, one has

$$f_{X^*}(z) = (I - f_X(z))^{-1}.$$

Series of matrices

- If X is a code, one has

$$f_{X^*}(z) = (I - f_X(z))^{-1}.$$

- Its radius of convergence $\rho(f_{X^*}(z))$ is the positive root of minimal modulus of $\det(I - f_X(z))$ (Perron-Frobenius).

Series of matrices

- If X is a code, one has

$$f_{X^*}(z) = (I - f_X(z))^{-1}.$$

- Its radius of convergence $\rho(f_{X^*}(z))$ is the positive root of minimal modulus of $\det(I - f_X(z))$ (Perron-Frobenius).
- For the golden mean,

$$\det(I - f_A(z)) = \begin{vmatrix} 1 - pz & -qz \\ -pz & 1 \end{vmatrix} = 1 - pz - pqz^2,$$

Series of matrices

- If X is a code, one has

$$f_{X^*}(z) = (I - f_X(z))^{-1}.$$

- Its radius of convergence $\rho(f_{X^*}(z))$ is the positive root of minimal modulus of $\det(I - f_X(z))$ (Perron-Frobenius).
- For the golden mean,

$$\det(I - f_A(z)) = \begin{vmatrix} 1 - pz & -qz \\ -pz & 1 \end{vmatrix} = 1 - pz - pqz^2,$$

- If $p + pq = 1$, $\rho(f_{X^*}(z)) = 1$.

Admissible assigments

Definition (A generalization of Bernoulli distributions)

An assignment π is admissible if $z = 1$ is the positive root of minimal modulus of $\det(I - f_A(z))$.

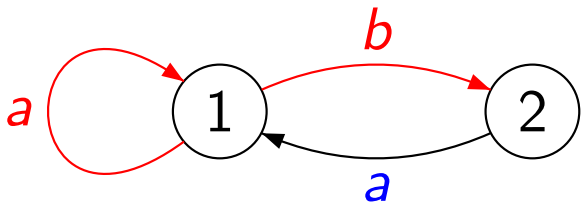
Let $A = \{a, b\}$ and let $\pi(a) = p, \pi(b) = q$.

- For the full shift, π is admissible if $p + q = 1$.
- For the golden mean, π is admissible if $p + pq = 1$.
- The uniform assignment $\pi(a) = \rho_S$ for all $a \in A$ is admissible.
- If S is of finite type and π is uniform,

$$\zeta_S(z) = det(I - f_A(z/\rho_S))^{-1}.$$

Example

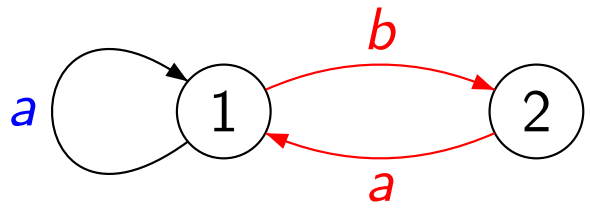
Consider the code $X = \{ab, ba\}$ in the golden mean system.



$$f_X(z) = \begin{bmatrix} pqz^2 & pqz^2 \\ 0 & pqz^2 \end{bmatrix}.$$

Example

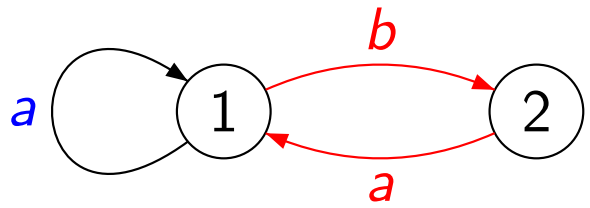
Consider the code $X = \{ab, ba\}$ in the golden mean system.



$$f_X(z) = \begin{bmatrix} pqz^2 & pqz^2 \\ 0 & pqz^2 \end{bmatrix}.$$

Example

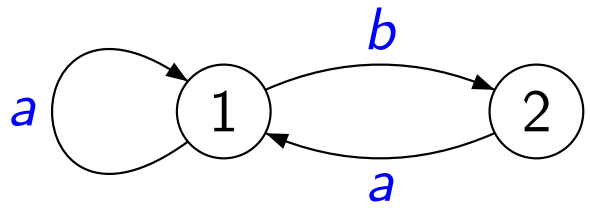
Consider the code $X = \{ab, ba\}$ in the golden mean system.



$$f_X(z) = \begin{bmatrix} pqz^2 & pqz^2 \\ 0 & pqz^2 \end{bmatrix}.$$

Example

Consider the code $X = \{ab, ba\}$ in the golden mean system.



$$f_X(z) = \begin{bmatrix} pqz^2 & pqz^2 \\ 0 & pqz^2 \end{bmatrix}.$$

Since,

$$\det(I - f_X(1)) = \begin{vmatrix} 1-pq & -pq \\ 0 & 1-pq \end{vmatrix} = 1 - pq + p^2q^2,$$

$$\det(I - f_X(1)) \geq 0 \iff pq(1-pq) \leq 1.$$

True if π is admissible, $p + pq = 1$.

Kraft's inequality

Theorem

Let S be an irreducible sofic shift. If $X \subset \mathrm{Fact}(S)$ is a code, then $\det(I - f_X(1)) \geq 0$ for any admissible assignment π.

Kraft's inequality

Theorem

Let S be an irreducible sofic shift. If $X \subset \mathrm{Fact}(S)$ is a code, then $\det(I - f_X(1)) \geq 0$ for any admissible assignment π.

Let $p_X(z) = 1 - \det(I - f_X(z/\rho_S))$.

Kraft's inequality

Theorem

Let S be an irreducible sofic shift. If $X \subset \mathrm{Fact}(S)$ is a code, then $\det(I - f_X(1)) \geq 0$ for any admissible assignment π.

Let $p_X(z) = 1 - \det(I - f_X(z/\rho_S))$.

Corollary

Let S be an irreducible sofic shift. If $X \subset \mathrm{Fact}(S)$ is a code, then $p_X(\rho_S) \leq 1$.

Kraft's inequality

When S is the full shift on k symbols with the uniform assignment, one recovers

Theorem (Kraft)

Let X be a code on A with Card $A = k$. *Let* $p_X(z) = \sum_{x \in X} z^{|x|}$, *we have*

$$\sum_{x \in X} k^{-|x|} \leq 1.$$

Example

Let $X = \{a, ab, bb\}$, $\pi(a) = \pi(b) = 1/2$,

$$\sum_{x \in X} (1/2)^{-|x|} = 1/2 + 1/4 + 1/4 = 1.$$

Kraft's inequality

Theorem

Let S be an irreducible sofic shift. If $X \subset \mathrm{Fact}(S)$ is a code, then $\det(I - f_X(1)) \geq 0$ for any admissible assignment π.

Proof.

- π is admissible implies $\rho(f_{A^*}(z)) = 1$. Thus $\rho(f_{X^*}(z)) \geq 1$.

Kraft's inequality

Theorem

Let S be an irreducible sofic shift. If $X \subset \mathrm{Fact}(S)$ is a code, then $\det(I - f_X(1)) \geq 0$ for any admissible assignment π.

Proof.

- π is admissible implies $\rho(f_{A^*}(z)) = 1$. Thus $\rho(f_{X^*}(z)) \geq 1$.
- X is a code implies $f_{X^*}(z) = (I - f_X(z))^{-1}$.

Kraft's inequality

Theorem

Let S be an irreducible sofic shift. If $X \subset \mathrm{Fact}(S)$ is a code, then $\det(I - f_X(1)) \geq 0$ for any admissible assignment π.

Proof.

- π is admissible implies $\rho(f_{A^*}(z)) = 1$. Thus $\rho(f_{X^*}(z)) \geq 1$.
- X is a code implies $f_{X^*}(z) = (I - f_X(z))^{-1}$.
- Thus for $0 \leq z < 1$, $\det(I - f_X(z)) \neq 0$.

Kraft's inequality

Theorem

Let S be an irreducible sofic shift. If $X \subset \mathrm{Fact}(S)$ is a code, then $\det(I - f_X(1)) \geq 0$ for any admissible assignment π.

Proof.

- π is admissible implies $\rho(f_{A^*}(z)) = 1$. Thus $\rho(f_{X^*}(z)) \geq 1$.
- X is a code implies $f_{X^*}(z) = (I - f_X(z))^{-1}$.
- Thus for $0 \leq z < 1$, $\det(I - f_X(z)) \neq 0$.
- Hence $\det(I - f_X(1)) \geq 0$.

The equality case

Theorem

Let S be an irreducible sofic shift, π be an admissible assignment and $X \subset \mathrm{Fact}(S)$ be a regular code. The code X is complete in S if and only if $\det(I - f_X(1)) = 0$.

The equality case

Theorem

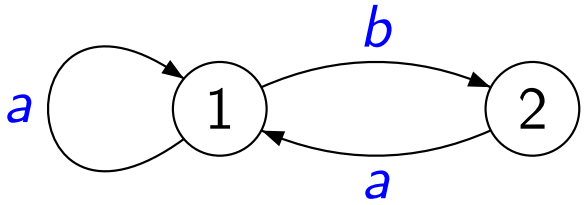
Let S be an irreducible sofic shift, π be an admissible assignment and $X \subset \mathrm{Fact}(S)$ be a regular code. The code X is complete in S if and only if $\det(I - f_X(1)) = 0$.

Corollary

Let S be an irreducible sofic shift and let $X \subset \mathrm{Fact}(S)$ be a regular code. The code X is complete in S if and only if $p_X(\rho_S) = 1$.

Example

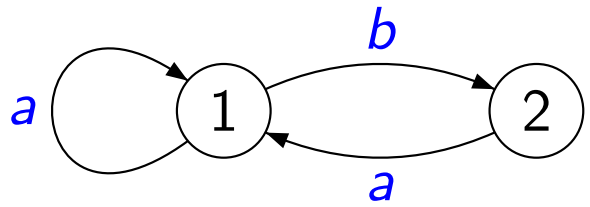
Consider the code $X = \{aa, ab, ba\}$ in the golden mean system.



$$f_X(z) = \begin{bmatrix} (p^2 + pq)z^2 & pqz^2 \\ p^2z^2 & pqz^2 \end{bmatrix}.$$

Example

Consider the code $X = \{aa, ab, ba\}$ in the golden mean system.



$$f_X(z) = \begin{bmatrix} (p^2+pq)z^2 & pqz^2 \\ p^2z^2 & pqz^2 \end{bmatrix}.$$

$$\begin{aligned} \det(I - f_X(1)) &= 1 - p^2 - 2pq + p^2q^2 = (1-p-pq)(1+p-pq) \\ &= 0. \end{aligned}$$

The code X is complete in the golden mean system.

The equality case

Theorem

Let S be an irreducible sofic shift, π be an admissible assignment and $X \subset \mathrm{Fact}(S)$ be a regular code. The code X is complete in S if and only if $\det(I - f_X(1)) = 0$.

Proof.

- Let X a regular code complete in S, $\mu(A^*)$ is contained in a finite union of two-sided residuals of entries of $\mu(X^*)$. This shows that $\rho(f_{X^*}(z)) = 1$, whence $\det(I - f_X(1)) = 0$.

The equality case

Theorem

Let S be an irreducible sofic shift, π be an admissible assignment and $X \subset \mathrm{Fact}(S)$ be a regular code. The code X is complete in S if and only if $\det(I - f_X(1)) = 0$.

Proof.

- Let X a regular code complete in S, $\mu(A^*)$ is contained in a finite union of two-sided residuals of entries of $\mu(X^*)$. This shows that $\rho(f_{X^*}(z)) = 1$, whence $\det(I - f_X(1)) = 0$.
- Conversely, if X is not complete, then $\rho(f_{X^*}(z)) > 1$ whence $\det(I - f_X(1)) > 0$.

A factorization problem

We denote by α the morphism obtained from μ by taking the commutative image of the elements. For any finite code X, we define in $\mathbb{Z}[A]$ the polynomial $p(X) = \det(I - \alpha(X))$.

Theorem

Let S be an irreducible sofic shift and X is a finite code complete in S. We have $p(A)$ divides $p(X)$.

The above result solves a problem raised by Reutenauer (1986), who proved the same result under an additional hypothesis.

A factorization problem

Theorem

Let S be an irreducible sofic shift and X is a finite code complete in S. We have $p(A)$ divides $p(X)$.

Example

For the golden mean and $X = \{aa, ab, ba\}$, whe have

- $p(A) = 1 - a - ab$
- $p(X) = 1 - aa - 2ab + a^2b^2 = (1 + a - ab)(1 - a - ab)$.

Proof of the main result

Theorem

Let S be an irreducible sofic shift and X is a finite code complete in S. We have $p(A)$ divides $p(X)$.

We have

- an analytic proof,
- a combinatorial proof.

An analytic proof

The proof is reduced to the case of an edge shift (the labels of all edges are distinct). In this case, $p(A)$ is irreducible in $\mathbb{Z}[A]$ (Reutenauer, 1986).
We select a letter $a \in A$ and write

$$p(A) = -aq + r$$

where q, r are polynomials in $\mathbb{Z}[A - a]$.
Then

$$p(X)q^n = p(A)s' + t'$$

for some $n \geq 0$ and some $s' \in \mathbb{Z}[A]$, $t' \in \mathbb{Z}[A - a]$.
There exists a ball $B((\rho_S)_{b \in A-a}, \epsilon)$ such that the assignment $\pi(b) = x_b$ for $x \in B$ with $\pi(a) = r(x)/q(x)$ is admissible.

An analytic proof

Then $p(X) = p(A) = 0$ for any such assignment π and thus t' vanishes on B whence $t' = 0$. We use here previous Theorem.

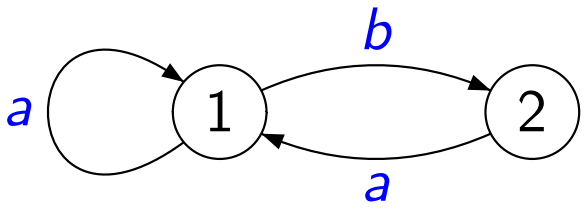
Since $p(A)$ is irreducible, this forces $p(A)/p(X)$.

A combinatorial proof

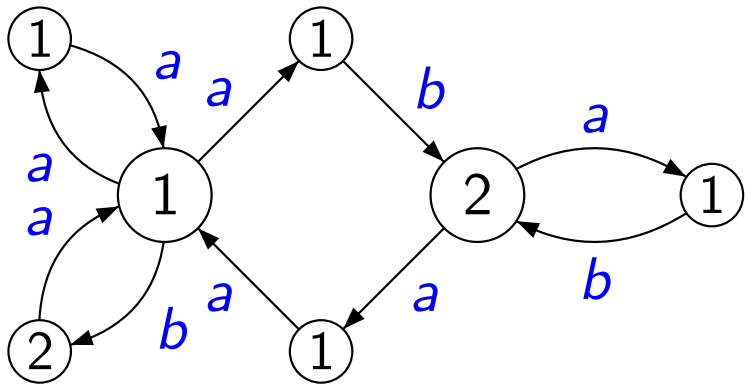
We build an automaton $\mathcal{B}$ induced from X in S.

Example

Consider the code $X = \{aa, ab, ba\}$ in S



Automaton $\mathcal{A}$



Automaton $\mathcal{B}$

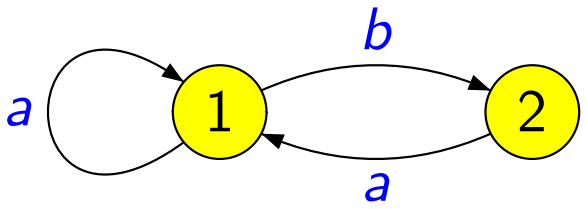
Since X is a complete code, $\mathcal{B}$ is unambiguous and recognizes S.

A combinatorial proof

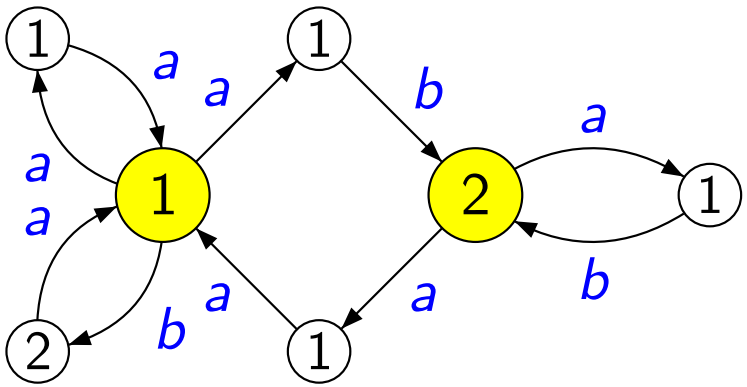
We build an automaton $\mathcal{B}$ induced from X in S.

Example

Consider the code $X = \{aa, ab, ba\}$ in S



Automaton $\mathcal{A}$



Automaton $\mathcal{B}$

Since X is a complete code, $\mathcal{B}$ is unambiguous and recognizes S.

A combinatorial proof

By construction,

$$\begin{aligned} p(X) &= \det(I - \mu_{\mathcal{B}}(A)) \\ p(A) &= \det(I - \mu_{\mathcal{A}}(A)). \end{aligned}$$

Lemma

If $\mathcal{B}$ is an unambiguous and irreducible automaton representing S, $\det(I - \mu_{\mathcal{A}}(A))$ divides $\det(I - \mu_{\mathcal{B}}(A))$.

Proof.

- Let $\mathcal{C}$ be an automaton. The degree of a word u in $\mathcal{C}$ is the number of paths labelled u in $\mathcal{C}$.

A combinatorial proof

By construction,

$$\begin{aligned} p(X) &= \det(I - \mu_{\mathcal{B}}(A)) \\ p(A) &= \det(I - \mu_{\mathcal{A}}(A)). \end{aligned}$$

Lemma

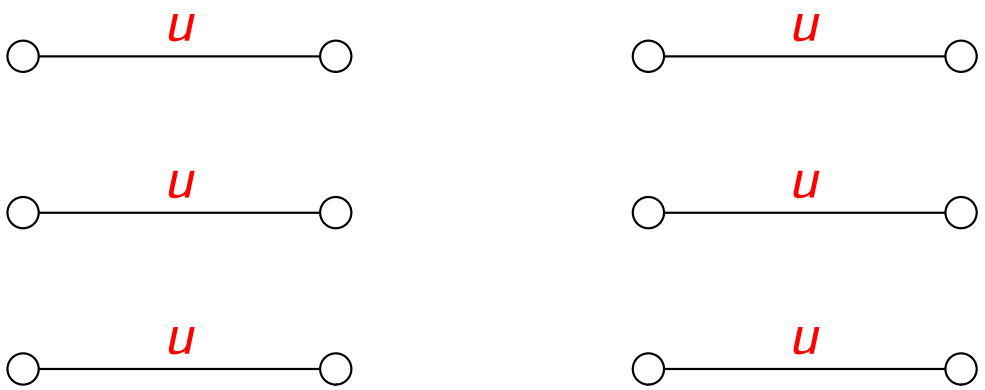
If $\mathcal{B}$ is an unambiguous and irreducible automaton representing S, $\det(I - \mu_{\mathcal{A}}(A))$ divides $\det(I - \mu_{\mathcal{B}}(A))$.

Proof.

- Let $\mathcal{C}$ be an automaton. The degree of a word u in $\mathcal{C}$ is the number of paths labelled u in $\mathcal{C}$.
- One can choose a word u which has a nonnull and minimal degree d for $\mathcal{A}$ and k for $\mathcal{B}$.

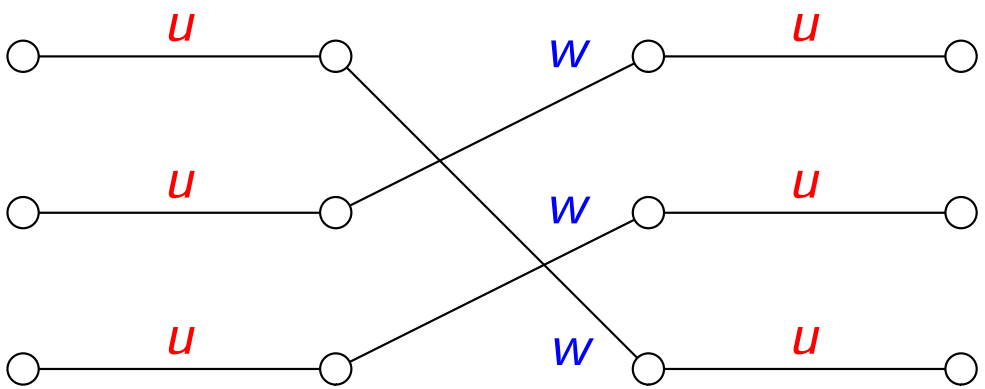
A combinatorial proof

- Since the automaton is unambiguous, for any word u of minimal nonnull degree k and any word w such that uwu has a nonnull degree, uwu has degree k.



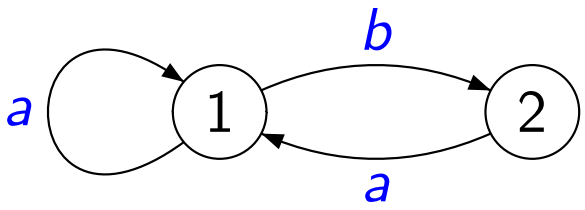
A combinatorial proof

- Since the automaton is unambiguous, for any word u of minimal nonnull degree k and any word w such that uwu has a nonnull degree, uwu has degree k.

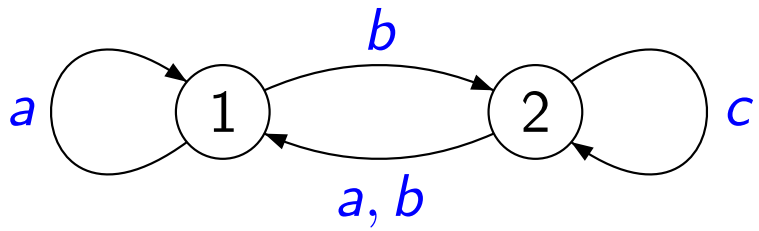


A combinatorial proof

We first assume that S has almost finite type: its Shannon cover $\mathcal{A}$ has no distinct left-infinite paths with the same label ending in a same state.



almost finite type shift



not almost finite type

If S has almost finite type, $\mathcal{A}$ has a word of degree $d = 1$.

A combinatorial proof

- Hence, the two $\mathbb{N}$-automata below define the same series.

$$s_{\mathcal{B}} = \langle d \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \mu_{\mathcal{B}}(u), \quad \mu_{\mathcal{B}}, \quad \mu_{\mathcal{B}}(u) \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \rangle,$$

$$s_{\mathcal{A}} = \langle k \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \mu_{\mathcal{A}}(u), \quad \mu_{\mathcal{A}}, \quad \mu_{\mathcal{A}}(u) \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \rangle$$

A combinatorial proof

- Hence, the two $\mathbb{N}$-automata below define the same series.

$$s_{\mathcal{B}} = \langle d \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \mu_{\mathcal{B}}(u), \quad \mu_{\mathcal{B}}, \quad \mu_{\mathcal{B}}(u) \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \rangle,$$

$$s_{\mathcal{A}} = \langle k \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \mu_{\mathcal{A}}(u), \quad \mu_{\mathcal{A}}, \quad \mu_{\mathcal{A}}(u) \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \rangle$$

- The Schützenberger reductions of the automata $s_{\mathcal{B}}$ and $s_{\mathcal{A}}$ over a field are similar.

A combinatorial proof

- Hence, the two $\mathbb{N}$-automata below define the same series.

$$s_{\mathcal{B}} = \langle d \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \mu_{\mathcal{B}}(u), \quad \mu_{\mathcal{B}}, \quad \mu_{\mathcal{B}}(u) \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \rangle,$$

$$s_{\mathcal{A}} = \langle k \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \mu_{\mathcal{A}}(u), \quad \mu_{\mathcal{A}}, \quad \mu_{\mathcal{A}}(u) \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \rangle$$

- The Schützenberger reductions of the automata $s_{\mathcal{B}}$ and $s_{\mathcal{A}}$ over a field are similar.
- Since $s_{\mathcal{A}}$ has degree 1, it is already reduced.

A combinatorial proof

- Hence, the two $\mathbb{N}$-automata below define the same series.

$$s_{\mathcal{B}} = \langle d \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \mu_{\mathcal{B}}(u), \quad \mu_{\mathcal{B}}, \quad \mu_{\mathcal{B}}(u) \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \rangle,$$

$$s_{\mathcal{A}} = \langle k \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \mu_{\mathcal{A}}(u), \quad \mu_{\mathcal{A}}, \quad \mu_{\mathcal{A}}(u) \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \rangle$$

- The Schützenberger reductions of the automata $s_{\mathcal{B}}$ and $s_{\mathcal{A}}$ over a field are similar.
- Since $s_{\mathcal{A}}$ has degree 1, it is already reduced.
- The reduction of $s_{\mathcal{B}}$ is obtained through a left and a right reduction.

A combinatorial proof

- Each one is a conjugacy of automata: $\langle I, \mu, T\rangle \xleftarrow{U} \langle J, \mu', F\rangle$. For any word w,

$$JU = I, \;\; U\mu(w) = \mu'(w)U, \text{ and } F = UT,$$

where the rows of U are a basis of $\langle I\mu(A^*)\rangle$ and $\mu'(w)$ is the transformation $\mu(w)$ in this basis.

A combinatorial proof

- Each one is a conjugacy of automata: $\langle I, \mu, T\rangle \xleftarrow{U} \langle J, \mu', F\rangle$. For any word w,

$$JU = I, \;\; U\mu(w) = \mu'(w)U, \text{ and } F = UT,$$

  where the rows of U are a basis of $\langle I\mu(A^*)\rangle$ and $\mu'(w)$ is the transformation $\mu(w)$ in this basis.
- We also have $U\mu(A) = \mu'(A)U$.

A combinatorial proof

- Each one is a conjugacy of automata: $\langle I, \mu, T\rangle \xleftarrow{U} \langle J, \mu', F\rangle$. For any word w,

$$JU = I, \;\; U\mu(w) = \mu'(w)U, \text{ and } F = UT,$$

  where the rows of U are a basis of $\langle I\mu(A^*)\rangle$ and $\mu'(w)$ is the transformation $\mu(w)$ in this basis.
- We also have $U\mu(A) = \mu'(A)U$.
- It follows that the Jordan form of $\mu'(A)$ is a principal submatrix of $\mu(A)$.

A combinatorial proof

- Each one is a conjugacy of automata: $\langle I, \mu, T\rangle \xleftarrow{U} \langle J, \mu', F\rangle$. For any word w,

$$JU = I, \;\; U\mu(w) = \mu'(w)U, \text{ and } F = UT,$$

  where the rows of U are a basis of $\langle I\mu(A^*)\rangle$ and $\mu'(w)$ is the transformation $\mu(w)$ in this basis.
- We also have $U\mu(A) = \mu'(A)U$.
- It follows that the Jordan form of $\mu'(A)$ is a principal submatrix of $\mu(A)$.
- Hence $\det(I - \mu_{\mathcal{A}}(A))$ divides $\det(I - \mu_{\mathcal{B}}(A))$.

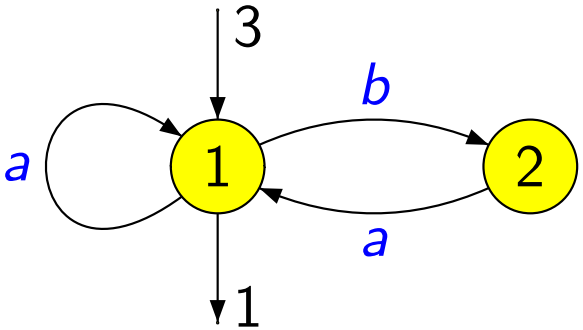
A combinatorial proof

- Each one is a conjugacy of automata: $\langle I, \mu, T\rangle \xleftarrow{U} \langle J, \mu', F\rangle$. For any word w,

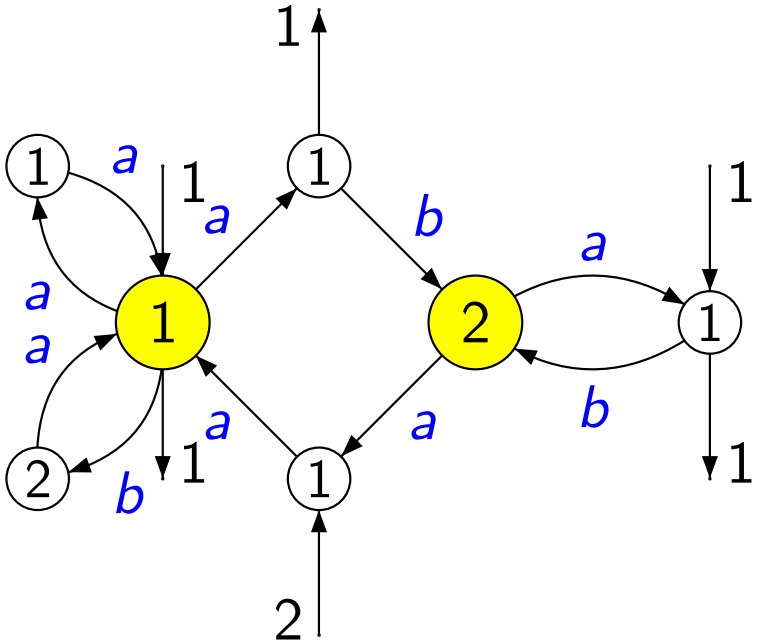
$$JU = I, \;\; U\mu(w) = \mu'(w)U, \text{ and } F = UT,$$

  where the rows of U are a basis of $\langle I\mu(A^*)\rangle$ and $\mu'(w)$ is the transformation $\mu(w)$ in this basis.
- We also have $U\mu(A) = \mu'(A)U$.
- It follows that the Jordan form of $\mu'(A)$ is a principal submatrix of $\mu(A)$.
- Hence $\det(I - \mu_{\mathcal{A}}(A))$ divides $\det(I - \mu_{\mathcal{B}}(A))$.
- Extension to all sofic shifts by a specialization argument.

The equivalent $\mathbb{N}$-automata



Automaton $\mathcal{A}$



Automaton $\mathcal{B}$

A variant of the combinatorial proof

From Nasu [1985],

- Let $M = \sum_{a \in A} \mu_{\mathcal{A}}(a)$ and $M' = \sum_{a \in A} \mu_{\mathcal{B}}(a)$. Let $h(S) = \log \lambda$.

A variant of the combinatorial proof

From Nasu [1985],

- Let $M = \sum_{a \in A} \mu_{\mathcal{A}}(a)$ and $M' = \sum_{a \in A} \mu_{\mathcal{B}}(a)$. Let $h(S) = \log \lambda$.
- Let $\mathbf{u}$ (resp. $\mathbf{v}$) be a positive left (resp. right) eigenvector of M for λ. Let $\mathbf{u}'$ (resp. $\mathbf{v}'$) be a positive left (resp. right) eigenvector of M' for λ with $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}' \cdot \mathbf{v}'$.

A variant of the combinatorial proof

From Nasu [1985],

- Let $M = \sum_{a \in A} \mu_{\mathcal{A}}(a)$ and $M' = \sum_{a \in A} \mu_{\mathcal{B}}(a)$. Let $h(S) = \log \lambda$.
- Let $\mathbf{u}$ (resp. $\mathbf{v}$) be a positive left (resp. right) eigenvector of M for λ. Let $\mathbf{u}'$ (resp. $\mathbf{v}'$) be a positive left (resp. right) eigenvector of M' for λ with $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}' \cdot \mathbf{v}'$.
- Then $s_{\mathcal{A}} = \langle \mathbf{u}, \mu_{\mathcal{A}}, \mathbf{v} \rangle$ and $s_{\mathcal{B}} = \langle \mathbf{u}', \mu_{\mathcal{B}}, \mathbf{v}' \rangle$ are equivalent $\mathbb{R}$-automata.

A variant of the combinatorial proof

From Nasu [1985],

- Let $M = \sum_{a \in A} \mu_{\mathcal{A}}(a)$ and $M' = \sum_{a \in A} \mu_{\mathcal{B}}(a)$. Let $h(S) = \log \lambda$.
- Let $\mathbf{u}$ (resp. $\mathbf{v}$) be a positive left (resp. right) eigenvector of M for λ. Let $\mathbf{u}'$ (resp. $\mathbf{v}'$) be a positive left (resp. right) eigenvector of M' for λ with $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}' \cdot \mathbf{v}'$.
- Then $s_{\mathcal{A}} = \langle \mathbf{u}, \mu_{\mathcal{A}}, \mathbf{v} \rangle$ and $s_{\mathcal{B}} = \langle \mathbf{u}', \mu_{\mathcal{B}}, \mathbf{v}' \rangle$ are equivalent $\mathbb{R}$-automata.
- same end.

A variant of the combinatorial proof

Indeed, let us assume that $\lambda = 1$. We have, for $s = s_{\mathcal{A}}$ or $s_{\mathcal{B}}$, a right (and left) invariance property:

$$\sum_{w \in A^k} \langle s, xw \rangle = \langle s, x \rangle.$$

and an ergodic property:

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \sum_{w \in A^i} \langle s, xwy \rangle = \langle s, x \rangle \langle s, y \rangle.$$

The equivalence of the series follows then from the inequalities

$$\langle s_{\mathcal{B}}, w \rangle \leq K \langle s_{\mathcal{A}}, w \rangle.$$

$$\langle s_{\mathcal{A}}, w \rangle \leq K' \langle s_{\mathcal{B}}, w \rangle.$$

The noncommutative factorization theorem

Theorem (Reutenauer 1984)

For any finite complete code $X \subset A^*$ there exists two polynomials $L, R \in \mathbb{Z}\langle A \rangle$ such that $X - 1 = L(A - 1)R$.

The noncommutative factorization theorem

Conjecture (Schützenberger)

For any finite complete code $X \subset A^*$ there exists two polynomials $L, R \in \mathbb{N}\langle A \rangle$ such that $X - 1 = L(A - 1)R$.

References

- M.-P. Béal and D. Perrin. Codes and sofic constraints. *The Art of Theory*, *Theoret. Comput. Sci.*, 340(2):381-393, 2005
- M. Nasu. An invariant for bounded-to-one factor maps between transitive sofic shifts. *Ergod. Th. & Dynamic. Sys.*, 5:89-105, 1985
- A. Restivo. Codes and local constraints. *Theoret. Comput. Sci.*, 72:55-64 1990
- C. Reutenauer. Ensembles libres de chemins dans un graphe. *Bull. Soc. Math. France*, 114:135-152, 1986