Densities of rational languages by example

Dominique Perrin

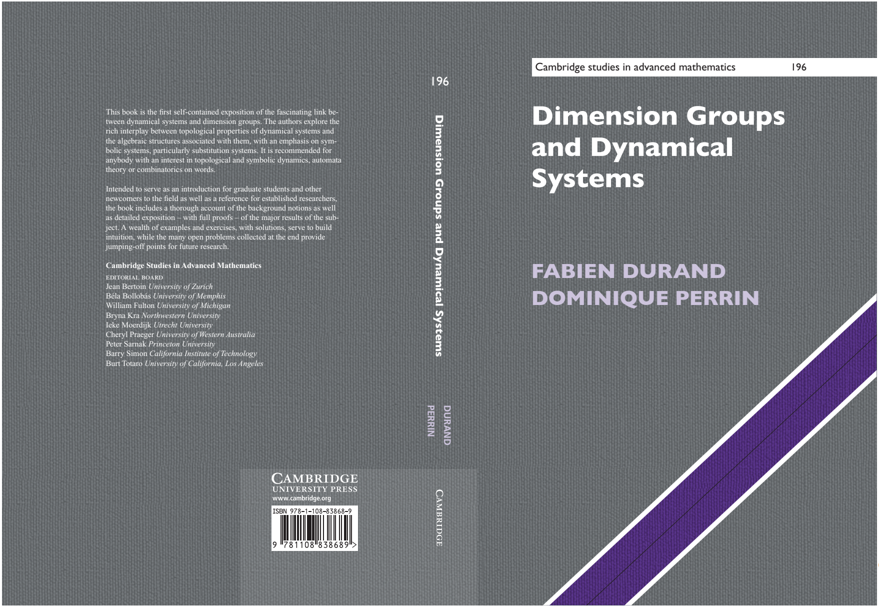
Université Gustave Eiffel
Laboratoire d'informatique Gaspard-Monge UMR 8049

SAMSA 2025

Outline

- Part 1. Computing invariant probability measures
  - Invariant probability mesures
  - Substitution shifts
  - Invariant measure on the Fibonacci shift
  - Invariant measure on the Thue-Morse shift
- Part 2. Computing in finite monoids
  - Green's relations
  - The $\mathcal{J}$-class $J_X(M)$
  - Parity of a in the Fibonacci shift
  - The Schützenberger representation

recommended reading :))



Stochastic processes

Let $\mu\colon A^* \to [0,1]$ be such that $\mu(\varepsilon) = 1$ and

$$\mu(w) = \sum_{a \in A} \mu(wa)$$

for every $w \in A^*$. Thus, we can interpret $\pi(wa)/\pi(w)$ as the probability of seeing the letter a after the word w.
Such a map μ is called a stochastic process on A^*.
For $L \subset A^*$, we denote $\mu(L) = \sum_{w \in L} \mu(w)$.
A simple example is a Bernoulli process, defined by a morphism $\mu\colon A^* \to [0,1]$ such that $\sum_{a \in A} \mu(a) = 1$. Equivalently, $\mu(wa)/\mu(w)$ does not depend on w.
If μ is a uniform Bernoulli process, that is if $\mu(a) = 1/\operatorname{card}(A)$, then

$$\mu(w) = \frac{1}{\operatorname{card}(A)^{|w|}}$$

Probability measures

A Borel probability measure on a topological space X is a map μ defined on the family of Borel sets of X such that $\mu(X) = 1$ and

$$\mu(\cup_{n\geq 0} U_i) = \sum_{n\geq 0} \mu(U_n)$$

for every family of paiwise disjoint Borel sets U_n.
Let $[w] = \{x \in A^{\mathbb{Z}} \mid x_{[0,|w|)} = w\}$ be the cylinder defined by the word w. Given a stochastic process μ, there is a unique Borel probability measure μ on $A^{\mathbb{Z}}$ such that $\mu([w]) = \mu(w)$ for every $w \in A^*$.
The support of μ is the set

$$X = \{x \in A^{\mathbb{Z}} \mid \mu(w) > 0 \text{ for every } w \in \mathcal{L}(x)\}.$$

It is a closed subset and $\mu(X) = 1$. Thus μ is a Borel probability measure on X.

Prefix codes

A prefix code on A is a set $C \subset A^*$ such that no word in C is a proper prefix of another word in C. A suffix code is the reversal of a prefix code.
For $X \subset A^{\mathbb{Z}}$, a prefix code $C \subset \mathcal{L}(X)$ is X-maximal if it is not properly included in a prefix code $C' \subset \mathcal{L}(X)$.
If μ is a stochastic process, one has $\mu(C) \leq 1$ for every prefix code C because the cylinders $[w]$ for $w \in C$ are disjoint.
Let X be the support of μ. If C is a finite X-maximal prefix code, then $\mu(X) = 1$ because $X = \cup_{c \in C}[c]$. Moreover, the average length of C

$$\lambda(C) = \sum_{c \in C} |c|\mu(c)$$

is equal to $\mu(P)$, where P is the set of proper prefixes of the words of C.

Invariant measures

A measure μ on $A^{\mathbb{Z}}$ is invariant if $\mu(S^{-1}U) = \mu(U)$ for every Borel set U, where S denotes the shift transformation.
The measure μ is invariant if the associated stochatic process satisfies

$$\mu(w) = \sum_{a \in A} \mu(aw)$$

for every $w \in A^*$.
The support of an invariant measure is closed and invariant. Thus, it is a shift space. Conversely, for every shift space X, there exists an invariant measure supported by X.
A Bernoulli measure is invariant.

Ergodic measures

An invariant measure μ is ergodic if every invariant Borel set has measure 0 or 1. As an equivalent condition, μ is ergodic if and only if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(U \cap S^{-i} V) = \mu(U)\mu(V)$$

for every pair U, V of Borel sets.
Every shift space has ergodic measures. If there is a unique invariant measure, it is ergodic. The shift is said to be uniquely ergodic.
A Bernoulli measure is ergodic.

Substitution shifts

Let $\sigma: A^* \to A^*$ be a substitution. The shift space $\mathrm{X}(\sigma)$ is the set of sequences $x \in A^{\mathbb{Z}}$ such that all the blocks of x appear in some $\sigma^n(a)$ for $a \in A$ and $n \geq 0$.
The substitution σ is primitive if for every $a \in A$, there is $n \geq 1$ such that every letter $b \in A$ appears in $\sigma^n(a)$.

Theorem (Michel)

Every primitive substitution shift is uniquely ergodic.

Computation of the unique invariant measure

The composition matrix of $\sigma \colon A^* \to A^*$ is the $A \times A$-matrix

$$M(\sigma)_{a,b} = |\sigma(b)|_a.$$

Proposition

If σ is primitive, and μ is the unique invariant measure, then

$$\mu([a])_{a \in A}$$

is a right Perron eigenvector of $M(\sigma)$.

Relation to average length

Let $\sigma\colon A^* \to A^*$ be a primitive substitution. Let μ be the unique invariant probability distribution on $\mathsf{X}(\sigma)$. The average length of σ

$$\lambda(\sigma) = \sum_{a\in A} |\sigma(a)|\mu(a)$$

is equal to the Perron eigenvalue ρ of $M(\sigma)$. Indeed,

$$\lambda(\sigma) = \sum_{a\in A} |\sigma(a)|\mu(a) = \sum_{a,b\in A} |\sigma(a)|_b\mu(a) = \rho\sum_{b\in A}\mu(b) = \rho.$$

Recognizability of substitutions

Let $\sigma \colon A^* \to B^*$ be a substitution. Let X be a shift space on the alphabet A and let Y be the closure under the shift of $\sigma(X)$.
The substitution σ is recognizable in X if for every $y \in Y$ there is exactly one pair (x, k) with $x \in X$ and $0 \le k < |\sigma(x_0)|$ such that

$$y = S^k(\sigma(x)).$$

The following result is well known.

Theorem (Mossé)

Every primitive aperiodic substitution $\sigma \colon A^* \to A^*$ is recognizable in $\mathrm{X}(\sigma)$.

Consequences of recognizability

If $\sigma\colon A^* \to B^*$ is recognizable in X, then it is a homeomorphism from X onto Y.
Therefore, by Kac's formula, if $\sigma\colon A^* \to A^*$ is primitive and aperiodic, one has

$$\mu(\sigma(U)) = \mu(U)/\lambda$$

for every Borel set U, where λ is the Perron eigenvalue of $M(\sigma)$.
Thus, we have an enlightening interpretation of the fact that $(\mu(a))_{a\in A}$ is a left eigenvector of $M(\sigma)$: there is a partition of $\mathrm{X}(\sigma)$ is clopen sets $S^k\sigma([a])$ for $a \in A$ and $0 \le k < |\sigma(a)|$). Therefore

$$1 = \sum_{a\in A} |\sigma(a)|\mu(\sigma(a)) = \sum_{a\in A} |\sigma(a)|\mu(a)/\lambda$$

The k-th higher block shift

Let X be a shift space on A. Let $u \mapsto \langle u \rangle$ be a bijection from the set $\mathcal{L}_k(X)$ of blocks of length k of X onto an alphabet A_k. The k-th higher block shift $X^{(k)}$ is the image of X under the map γ_k defined by $y = \gamma_k(x)$ if

$$y_n = \langle x_n x_{n+1} \cdots x_{n+k-1} \rangle \quad (n \in \mathbb{Z})$$

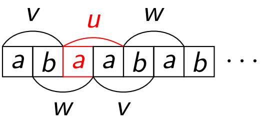
For $X = \mathrm{X}(\sigma)$, one has $X^{(k)} = \mathrm{X}(\sigma_k)$ where σ_k is the k-th higher block presentation of a non-erasing substitution σ.
Let $u \in \mathcal{L}_k(\sigma)$ and let a be the first letter of u. Set $s = |\sigma(a)|$. If $\sigma(u) = b_1 b_2 \cdots b_\ell$ with $b_i \in A$, then

$$\sigma_k(\langle u \rangle) = \langle b_1 b_2 \cdots b_k \rangle \langle b_2 b_3 \cdots b_{k+1} \rangle \cdots \langle b_s \cdots b_{s+k-1} \rangle.$$

The vector $\mu(u)_{u \in \mathcal{L}_k(X)}$ is a right Perron eigenvector of $M(\sigma_k)$.

Example

Let $\sigma\colon a \mapsto ab, b \mapsto a$ be the Fibonacci substitution. Set $u = \langle aa \rangle$, $v = \langle ab \rangle$, $w = \langle ba \rangle$. Then $\sigma_2\colon u \mapsto vw, v \mapsto vw, w \mapsto u$ generates $\mathsf{X}(\sigma)^{(2)}$.



The invariant measure on the Fibonacci shift

Let $\sigma\colon a \mapsto ab, b \mapsto a$ be the Fibonacci substitution and let $X = X(\sigma)$ be the Fibonacci shift. Then

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

Its eigenvalues are the roots $\lambda = (1 + \sqrt{5})/2$ and $\hat{\lambda} = (1 - \sqrt{5})/2$ of $z^2 = z + 1$. Then $\begin{bmatrix} \lambda^{-1} & \lambda^{-2} \end{bmatrix}^t$ is a right eigenvector for the eigenvalue λ. Thus $\mu(a) = \lambda^{-1}$ and $\mu(b) = \lambda^{-2}$.

Let us compute $\mu(u)$ for $u \in \mathcal{L}_2(X)$. Set $u = \langle aa \rangle$, $v = \langle ab \rangle$, $w = \langle ba \rangle$. Then

$$\sigma_2 \colon u \mapsto vw, v \mapsto vw, w \mapsto u$$

and thus

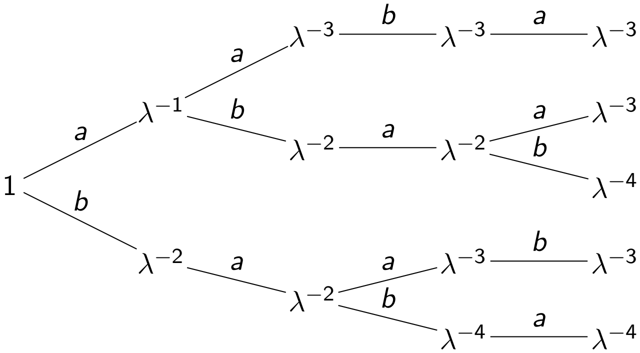
$$M(\sigma_2) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

has the eigenvector

$$\begin{bmatrix} \lambda^{-3} & \lambda^{-2} & \lambda^{-2} \end{bmatrix}^t$$

Thus $\mu(aa) = \lambda^{-3}$, $\mu(ab) = \lambda^{-2}$ and $\mu(ba) = \lambda^{-2}$.

The invariant probability measure on the Fibonacci shift



The Thue-Morse shift

Let $\sigma\colon a \mapsto ab, b \mapsto ba$ be the Thue-Morse substitution. The matrix

$$M(\sigma) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

has $\begin{bmatrix} 1/2 & 1/2 \end{bmatrix}$ as eigenvector for the eigenvalue 2. Set $u = \langle aa \rangle$, $v = \langle ab \rangle$, $w = \langle ba \rangle$, $t = \langle bb \rangle$. We find

$$\sigma_2\colon u \mapsto vw, v \mapsto vt, w \mapsto wu, t \mapsto wv$$

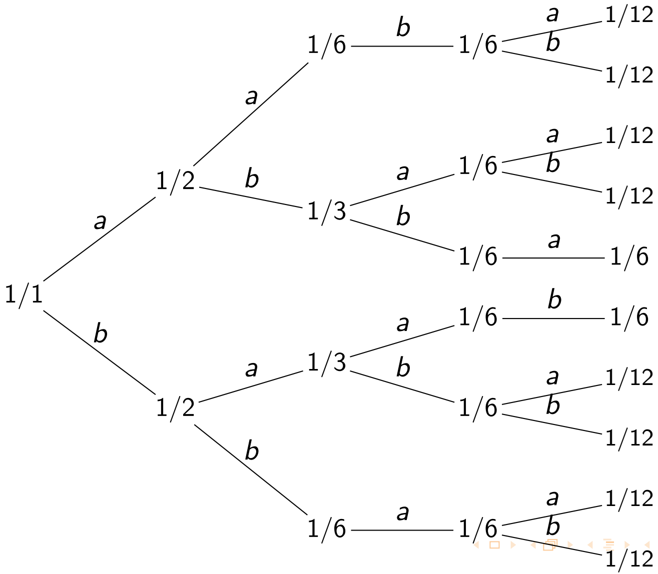
and thus

$$M(\sigma_2) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

with right eigenvector

$$\begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{bmatrix}^t$$

The invariant probability measure on the TM shift



Computing in finite monoids

Recall the Green relations in a monoid M.

- $m\mathcal{R}n \Leftrightarrow mM = nM \Leftrightarrow m, n$ generate the same right ideal
- $m\mathcal{L}n \Leftrightarrow Mm = Mn \Leftrightarrow m, n$ generate the same left ideal.
- $m\mathcal{J}n \Leftrightarrow MmM = MnM \Leftrightarrow m, n$ generate the same ideal.
- $m\mathcal{H}n \Leftrightarrow m\mathcal{R}n$ and $m\mathcal{L}n$.

When M is a monoid of partial mappings from a set Q to itself, the Green relations have natural interpretations. The kernel of m is the equivalence relation on Q defined by $p \equiv q$ if $pm = qm$. Likewise the image $\mathrm{Im}(m)$ of m is the set of $q \in Q$ of the form pm for some $p \in Q$. If $m\mathcal{R}n$, then m and n have the same kernel. Symmetrically, if $m\mathcal{L}n$, then m and n have the same image. Finally, if $m\mathcal{J}n$, then m, n have the same rank (where *rank* means the cardinality of the image).

A $\mathcal{J}$-class J is regular if it contains an idempotent. We have

- All $\mathcal{H}$-classes contained in J have the same number of elements.
- Each $\mathcal{H}$-class containing an idempotent is a group and there is one in each $\mathcal{R}$-class and each $\mathcal{L}$-class.
- All groups in J are isomorphic to the Schützenberger group of J.

When M is a group, it is a single $\mathcal{H}$-class.

The $\mathcal{J}$-class $J_X(M)$

Let X be a shift space on A and let $\varphi\colon A^* \to M$ be a morphism onto a finite monoid M. Let $K_X(M)$ be the intersection of all two-sided ideals I of M such that $I \cap \varphi(\mathcal{L}(X)) \neq \varnothing$. Let $J_X(M)$ be the $\mathcal{J}$-class

$$J_X(M) = \{m \in M \mid MmM = K_X(M)\}.$$

Proposition

Let X be an irreducible shift space on A and let $\varphi\colon A^* \to M$ be a morphism onto a finite monoid M. Then

1. *$K_X(M)$ is an ideal of M which meets $\varphi(\mathcal{L}(X))$.*
2. *$J_X(M) = \{m \in K_X(M) \mid MmM \cap \varphi(\mathcal{L}(X)) \neq \emptyset\}$.*
3. *$J_X(M)$ is either the minimal ideal $K(M)$ of M, or the unique 0-minimal ideal in the quotient of M by the largest ideal of M which does not meet $\varphi(\mathcal{L}(X))$.*

X-degree of an automaton

When the monoid M is the monoid of transitions of a deterministic automaton $\mathcal{A}$, the $\mathcal{J}$-class $J_X(M)$ has a simple definition in terms of ranks of the mappings. The minimal rank of the elements of $\varphi(\mathcal{L}(X))$ as partial mappings is called the X-degree of the automaton, denoted $d_X(\mathcal{A})$.

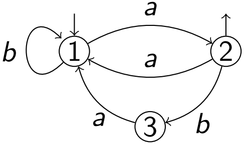
The X-degree of an automaton is computable provided $J_X(M)$ is computable, using the following statement.

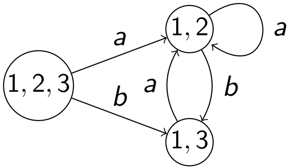
Proposition

Let $\mathcal{A}$ be a deterministic automaton and let $M = \varphi(A^*)$ be the transition monoid of $\mathcal{A}$. Let X be an irreducible shift space. The $\mathcal{J}$-class $J_X(M)$ contains all elements of $\varphi(\mathcal{L}(X))$ of rank $d_X(\mathcal{A})$.

The parity of aa in the Fibonacci shift

Let $\mathcal{A}$ be the automaton represented below on the left. Let X be the Fibonacci shift. Inside $\mathcal{L}(X)$, the automaton $\mathcal{A}$ recognizes (with $i = 1$ and $t = 2$) the blocks of X with an even number of aa.





$1, 2$	$1, 3$
a, a^2	ab, a^2b
ba, ba^2	b, bab

The action on subsets shown in the middle shows that $d_X(\mathcal{A}) = 2$. The $\mathcal{J}$-class $J_X(M)$ is represented on the right.

The Schützenberger representation

Let M be a finite monoid and let J be a regular $\mathcal{J}$-class. Let Λ be the set of $\mathcal{H}$-classes of J in the same $\mathcal{R}$-class R. We have an action of M on Λ defined by $H \cdot m = Hm$ if $Hm \subset J$ and $\emptyset$ otherwise.
Let $e \in J$ be an idempotent of R and let G be its $\mathcal{H}$-class. A system of coordinates of G is a family $(r_H, r'_H)_{H\in\Lambda}$ of pairs of elements of M such that for every $H \in \Lambda$,

$$er_H \in H, er_H r'_H = e$$

with $r_G = r'_G = e$. Set $H * m = er_H m r'_H$. The map

$$\lambda(m)_{H,K} = \begin{cases} H * m & \text{if } H \cdot m = K \\ 0 & \text{otherwise} \end{cases}$$

is a morphism from M to the monoid of $\Lambda \times \Lambda$-matrices with elements in $G \cup \{0\}$, called the Schützenberger representation of M on J.

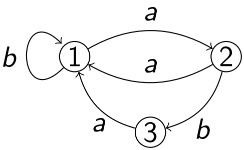
Computation in the transition monoid of an automaton

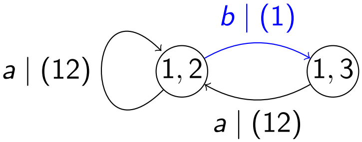
When M is the transition monoid of a deterministic automaton $\mathcal{A}$, the following simplifications occur:

1. The set Λ can be identified with the set $\mathcal{I}$ of images of minimal cardinality (equal to $d_X(\mathcal{A})$) of words in $\mathcal{L}(X)$.
2. One can choose a system of coordinates of R such that $H * \varphi(a) = e$ for every edge in a spanning tree of the graph with edges $H \xrightarrow{a} H \cdot \varphi(a)$.
3. For every $m \in M$ such that $G \cdot m = G$, the permutation $G * m$ is the restriction of m to the image of e.

Example

Let $\mathcal{A}$ be the automaton represented below on the left. Let μ be the invariant probability measure on the Fibonacci shift.





1, 2	1, 3
a	ab
ba	b

The action on minimal images is shown in the middle. Then

$$\lambda(a) = \begin{bmatrix} (12) & 0 \\ (12) & 0 \end{bmatrix} \quad \lambda(b) = \begin{bmatrix} 0 & (1) \\ 0 & 0 \end{bmatrix}$$

is the Schützenberger representation relative to $e = \varphi(a^2)$ with $r_{13} = \varphi(b)$.

An automaton of X-degree 3

Let X be the Fibonacci shift. Consider the X-maximal prefix code C represented below with the states of the minimal automaton of C^* indicated.

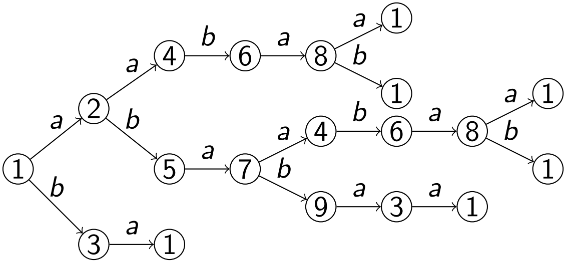


Figure: A prefix code of X-degree 3

It is not bifix because $aabaa, abaabaa \in C$.

The X-minimal rank is 3 because the image of aa is $\{1, 2, 4\}$ and the action on the minimal images is indicated below.

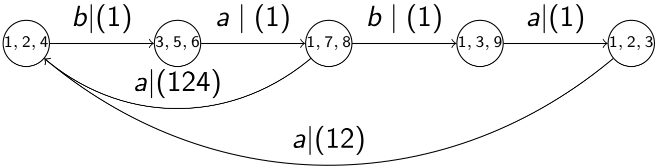
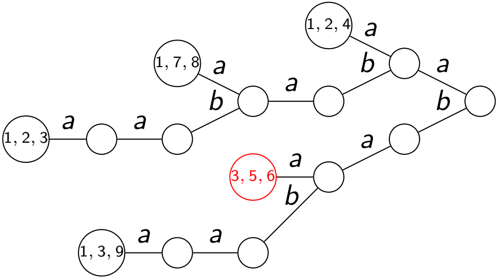


Figure: A prefix code of X-degree 3

The group is transitive because baa defines the permutation (124).

Computation of the density of C^*



Let G be the above X-maximal suffix code. One has $J_X(M) = \varphi(A^*G) \cap \mathcal{L}(X)$ and

$$\bigcup_{m \in J_X(M) \cap \varphi(C^*)} \varphi^{-1}(Mm) = A^*(G \setminus \{aab\}) \cap \mathcal{L}(X).$$

Therefore

$$\delta_\mu(C^*) = \frac{1-\lambda^{-3}}{3} \equiv .$$

Computation of $\lambda(C)$

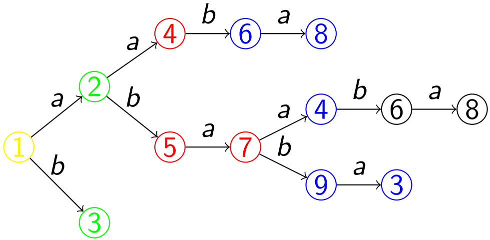


Figure: The set P of prefixes of C.

The yellow, green, red and blue sets are X-maximal suffix codes. Therefore,

$$\lambda(C) = 4 + \mu(abaab) + \mu(abaaba) = 4 + 2\lambda^{-3}.$$