A generalized definition of the scale of an automorphism

Stéphane Laurent

May 22, 2016

Abstract

Given an adic transformation S on the path space of a Bratteli graph and an automorphism T of a Lebesgue, we define a filtration whose tail σ-field is the $(T \times S)$-invariant σ-field. Considering the standardness property of this filtration provides a generalization of the scale of an automorphism as defined in [5].

Contents

1 Introduction

For a given adic transformation S acting on the set of infinite paths Γ of a Bratteli graph, one can straightforwardly associate a filtration on Γ whose tail σ-field is the S-invariant σ-field. When, in addition, an invertible measure-preserving transformation T is given, we show how to define a filtration whose tail σ-field is the $(T \times S)$-invariant σ-field. This construction was done in [5] in the case when S is an ordinary odometer, and considering the standardness property of this filtration yields a generalization of the scale of an automorphism as defined in [5].

2 Filtration associated to an adic transformation

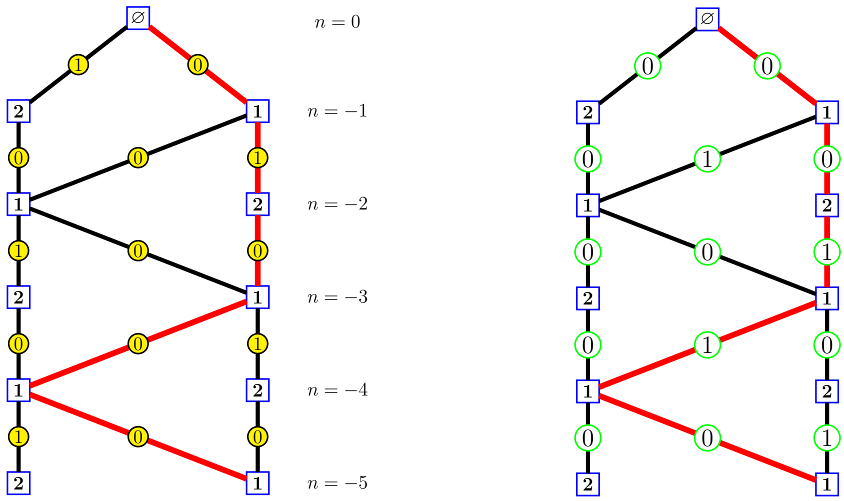
Consider an adic transformation S acting on the set of infinite paths Γ of a Bratteli graph, preserving a probability measure μ on Γ. In this section we introduce the sequence of measurable partitions and the corresponding filtration on the probability space Γ.

We consider that the root level of the graph is graded by the index $n = 0$ and the subsequent levels are graded by $n = -1$, $n = -2$, $\ldots$. The example of the golden graph is shown on Figure 1.

The usual labels on the arcs of a Bratteli graph, such as the one shown on Figure 1(a), provide, for each vertex v_n at a level n, an ordering of the arcs between v_n and the vertices connected to v_n at level $n-1$. The labels shown on Figure 1(b) are obtained by considering the other direction: they provide, for each vertex v_{n-1} at a level $n-1$, an ordering of the arcs between v_{n-1} and the vertices at level n connected to v_{n-1}. After a choice of such labels, we denote by $\boxed{\epsilon_n(\gamma)}$ the label of the edge connecting $v_{n-1}(\gamma)$ to $v_n(\gamma)$.

A path $\gamma \in \Gamma$ is a sequence of arcs $\gamma = (\gamma_0, \gamma_{-1}, \ldots)$, where γ_n connects a vertex at level n to a vertex at level $n-1$. Note that γ is determined by $(\gamma_{-1}, \gamma_{-2}, \ldots)$ when there is a unique arc between each vertex at level $n = -1$ and the root vertex at level $n = 0$.

We denote by $v_n(\gamma)$ the vertex at level n through which passes a path γ. The dimension of a vertex v, that is to say the number of paths connecting v to the root vertex, is denoted by $\dim(v)$, and we denote by $\boxed{d_n(\gamma) = \dim\big(v_n(\gamma)\big)}$ the dimension of the vertex at level n through which passes the path γ.



(a) Usual labels on the arcs

(b) Order labels ϵ_n on the arcs

Figure 1: Golden graph

2.1 The sequence of measurable partitions ζ

There is an increasing sequence of measurable partitions $(\zeta_n)_{n \leq 0}$ on (Γ, μ), defined by

$$\boxed{\gamma \overset{\zeta_n}{\sim} \gamma' \iff \gamma_k = \gamma'_k \quad \text{for every } k \leq n}.$$

The partition ζ_0 is the partition into singletons. The partition ζ_{-1} is also the partition into singletons when there is a unique arc between each vertex at level $n = -1$ and the root vertex at level $n = 0$.

The ζ_n-equivalence class $\zeta_n(\gamma)$ consists of $d_n(\gamma)$ elements. These elements are ordered: there is one element in $\zeta_n(\gamma)$, denoted by $\bar{\gamma}_n$, such that

$$\boxed{\zeta_n(\gamma) = \{\bar{\gamma}_n, S\bar{\gamma}_n, \ldots, S^{d_n(\gamma)-1}\bar{\gamma}_n\}}.$$

We consider $\bar{\gamma}_n$ as the ζ_n-representative of γ.

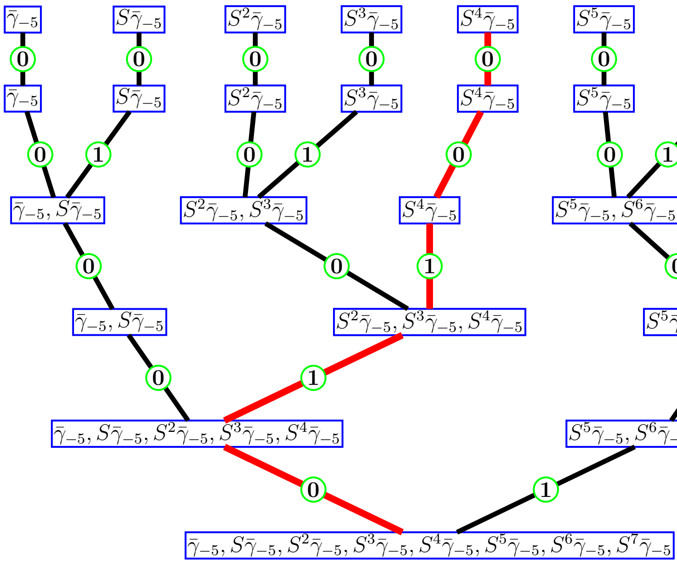


Figure 2: A ζ_{-5}-equivalence class

The increasing property of $(\zeta_n)_{n\leq 1}$ provides a structure on the ζ_n-equivalence classes: a ζ_n-equivalence class is a union of ζ_{n+1}-equivalences classes. This is illustrated on Figure 2 for the case of the golden graph. The labels on the edges of the tree shown on this figure are the labels ϵ_n also shown on Figure 1(b). The label $\epsilon_n(\gamma)$ between level n and level $n+1$ indicates the location of $\zeta_{n+1}(\gamma)$ as a block of $\zeta_n(\gamma)$.

Thus, $\boxed{\bar\gamma_n = S^{-k_n(\gamma)}\gamma}$ where the nonnegative integer $k_n(\gamma)$ is a function of $(\gamma_0, \gamma_{-1}, \ldots, \gamma_n)$, and

$$\zeta_n(\gamma) = \{S^{-k_n(\gamma)}\gamma, S^{-k_n(\gamma)+1}\gamma, \ldots, S^{-k_n(\gamma)+d_n(\gamma)-1}\gamma\}.$$

More precisely, knowing the vertex $v_n(\gamma)$, the integer $k_n(\gamma)$ is a one-to-one function of $\big(\epsilon_{n+1}(\gamma), \ldots, \epsilon_0(\gamma)\big)$.

The adic transformation on the path space of a Bratteli graph is a representation of a cutting-and-stacking construction. The cutting-and-stacking construction corresponding to the adic transformation on the golden graph is shown on Figure 3. The ζ_n-equivalence class $\zeta_n(\gamma)$ of γ is shown by the blue points on this figure. The ζ_n-representative $\bar\gamma_n$ of γ corresponds to the point in the base of the tower. For the example shown on Figure 3, one has $k_0(\gamma) = k_{-1}(\gamma) = k_{-2}(\gamma) = 0$ and $k_{-3}(\gamma) = 1$.

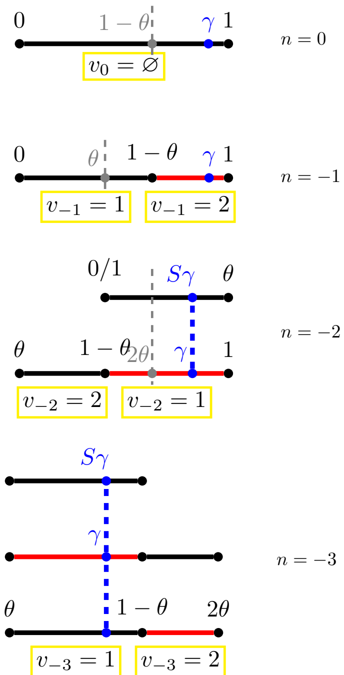


Figure 3: Golden towers

Lemma 1. *For almost $\gamma \in \Gamma$, $k_n(\gamma) \to \infty$ and $-k_n(\gamma) + d_n(\gamma) \to \infty$.*

Proof. The integer $k_n(\gamma)$ increases as n decreases to $-\infty$. If $k_n(\gamma) \to j < \infty$, that is to say $k_n(\gamma) = j$ for n small enough, then $S^{-j}\gamma$ belongs to the set of minimal paths of Γ, and this set has measure 0. Thus the set where $k_n(\gamma) \to j$ has measure 0 for every j, therefore the set where $k_n(\gamma) \not\to \infty$ has measure 0 by countable additivity. In the same way, the set where $-k_n(\gamma) + d_n(\gamma) \not\to \infty$ has measure 0 because the set of maximal paths of Γ has measure 0. □

Note that the sequence $\big(v_n(\gamma), \epsilon_n(\gamma)\big)_{n\leq n_0}$ determines the path γ truncated at level n_0. In other words it determines the ζ_n-equivalence class $\zeta_n(\gamma)$. We will come back to the above points in the next section, in the language of σ-fields.

2.2 The filtration $\mathcal{F}$

One gets a filtration $(\mathcal{F}_n)_{n\leq 0}$ by defining the σ-field $\mathcal{F}_n$ as the one generated by the measurable partition ζ_n. Here, a path $\gamma \in \Gamma$ is considered as the actual point taken at random in the probability space. Thus the σ-field $\mathcal{F}_n$ is $\boxed{\mathcal{F}_n = \sigma(R_n)}$ where we denote by R_n the random variable whose value at γ is the ζ_n-representative $\bar{\gamma}_n$ of γ. The random variable R_0 is a path taken at random according to μ. Note that $\sigma(R_n) \supset \sigma(R_{n-1})$ because $\bar{\gamma}_n$ determines the path γ truncated at n.

In the previous section, we introduced the integer $k_n(\gamma)$ such that $\bar{\gamma}_n = S^{-k_n(\gamma)}\gamma$. Here we consider k_n as random variable but we use the notation K_n instead of k_n. Thus $\boxed{R_n = S^{-K_n}R_0}$.

We also introduced the notations v_n and ϵ_n in the previous section. Here v_n and ϵ_n are random variables, and we use the notation V_n instead of v_n.

Thus the filtration $\mathcal{F}$ is generated by the stochastic process $(V_n, \epsilon_n)_{n\leq 0}$:

$$\boxed{\mathcal{F}_n = \sigma(V_m, \epsilon_m; m \leq n)}.$$

The random variable ϵ_{n+1} is a "novation" from $\mathcal{F}_n$ to $\mathcal{F}_{n+1}$, that is to say $\mathcal{F}_{n+1} = \mathcal{F}_n \vee \sigma(\epsilon_{n+1})$, since V_{n+1} is a function of V_n and ϵ_{n+1}.

Lemma 2. *The random variable ϵ_{n+1} is conditionally independent of $\mathcal{F}_n$ given V_n.*

Proof. Given $\mathcal{F}_n$, the random variable ϵ_{n+1} is the label of an arc connecting the vertex V_n to a vertex at level $n+1$. $\square$

Conditionally to V_n, the random integer K_n is a one-to-one function of $(\epsilon_{n+1}, \ldots, \epsilon_0)$, and it has the uniform distribution on $\{0, \ldots, \dim(V_n) - 1\}$. This is the *centrality* property of μ. Because of this property, the conditional law of V_{n+1} given V_n is given by

$$\Pr(V_{n+1} = v_{n+1} \mid V_n = v_n) = m(v_n, v_{n+1}) \frac{\dim(v_{n+1})}{\dim(v_n)}$$

where $m(v_n, v_{n+1})$ is the number of edges connecting v_n and v_{n+1}.

Observe that $R_{n+1} = S^{K_n - K_{n+1}} R_n$ and, conditionally to V_n, the nonnegative integer $K_n - K_{n+1}$ is a one-one function of ϵ_{n+1}. Thus, conditionally to V_n, the random integer K_n has always an expression of the form $K_n = \sum_{k=0}^{n+1} f_k(\epsilon_k)$.

2.3 Example: the golden graph

The random integer K_n has a very convenient expression for the case of the adic transformation on the golden graph with our choice of the labels shown on Figure 1(b):

$$K_n = \epsilon_{n+1} f_{n+1} + \ldots + \epsilon_{-1} f_{-1} (+\epsilon_0 f_0).$$

where $f_0 = 0$, $f_{-1} = f_{-2} = 1$, $f_{-3} = 2$, $\ldots$ are the Fibonacci numbers. Here its expression does not depend on V_n, but its distribution does.

There is no multiple edges in this graph, therefore the filtration $\mathcal{F}$ is generated by the stochastic process $(V_n)_{n\leq 0}$. Denoting by ϕ the golden number, the law of $(V_n)_{n\leq 0}$ is given by:

- $V_0 = \varnothing$;
- for $n \leq -1$, V_n takes the value 1 or 2 and $\Pr(V_n = 2) = \frac{f_n}{f_n + \phi f_{n-1}} =: p_n$;
- The transition matrix from V_n to V_{n+1} is

$V_n \backslash V_{n+1}$	1	2
1	f_n/f_{n-1}	f_{n+1}/f_{n-1}
2	1	0

2.4 Example: the Chacon graph

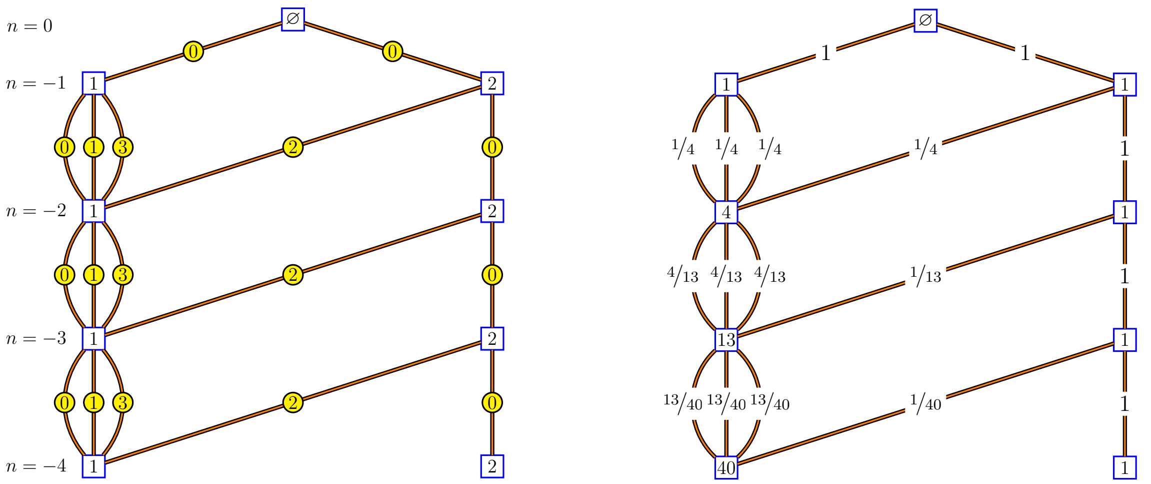
As another example, consider the Chacon graph shown on Figure 4 which also shows our choice of the labels on the arcs. For this example, $K_n = \sum_{k=0}^{n+1} f_k(\epsilon_k)$, with

$$\begin{cases} f_n(0) = 0 \\ f_n(1) = h_n \\ f_n(2) = 2h_n \\ f_n(3) = 2h_n + 1 \end{cases}.$$

where $h_n = \frac{3^{|n|}-1}{2}$ is the dimension of the vertex 1 at level n.

The law of the process $(V_n, \epsilon_n)_{n \leq 0}$ is given by:

- $V_0 = \varnothing$;
- for $n \leq -1$, V_n takes the value 1 or 2, and $\Pr(V_n = 2) = 1/3^{|n|}$;
- $\epsilon_n = 0$ and $V_n = 2$ if $V_{n-1} = 2$;
- conditionally to $V_{n-1} = 1$, the label ϵ_n of the edge between V_{n-1} and V_n equals 2 with probability $1/h_{n-1}$, or equals a value in $\{0, 1, 3\}$ with probability h_n/h_{n-1}.



(a) Chacon graph (b) Dimensions and transition probabilities

Figure 4: Chacon graph

Later, in order to prove Proposition 2, we will use the following property of the Chacon transformation, which is easy to see with the help of Figure 5. Let $I = \{\gamma \mid v_{-1}(\gamma) = 2\}$ be the set of infinite paths which pass through the vertex 2 at level $n = -1$. If $V_n = 1$, then

$$(\mathbf{1}_{R_n \in I}, \mathbf{1}_{SR_n \in I}, \ldots, \mathbf{1}_{S^{h_n-1}R_n \in I}) = c_0 c_1 \ldots c_{h_n-1} \tag{1}$$

where $c = (0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 1, 0 \ldots)$ is the limit of the words w_k obtained by initially setting $w_0 = 0$ and recursively setting $w_{k+1} = w_k w_k 1 w_k$. We call c the *Chacon word*.

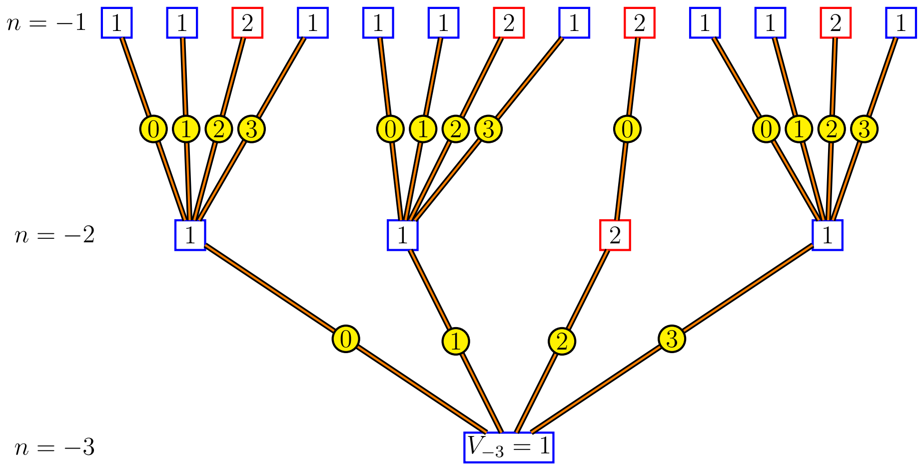


Figure 5: The process (V_n, ϵ_n) for the Chacon graph

3 Filtration associated to an adic transformation and an automorphism

Now, in addition to the adic transformation S, let T be an invertible measure-preserving transformation on a Lebesgue space $(\mathcal{X}, \nu)$.

We will define a filtration $\mathcal{G}$ locally isomorphic to $\mathcal{F}$, whose tail σ-field $\mathcal{G}_{-\infty}$ is the invariant σ-field of the product transformation $T \times S$.

3.1 The sequence of measurable partitions ξ

Here one defines an increasing sequence of measurable partitions $(\xi_n)_{n\leq 0}$ locally isomorphic to the elementary sequence $(\zeta_n)_{n\leq 0}$ associated to S.

For two paths γ, γ' in the same S-orbit, denote by $k(\gamma,\gamma')$ the integer such that $\gamma' = S^{k(\gamma,\gamma')}\gamma$. Thus $k(\gamma,\gamma') = k_n(\gamma) - k_n(\gamma')$ when $\gamma \overset{\zeta_n}{\approx} \gamma'$. Then define the measurable partition ξ_n by

$$\boxed{(x,\gamma) \overset{\xi_n}{\approx} (x',\gamma') \iff \gamma \overset{\zeta_n}{\approx} \gamma' \quad \text{and } x' = T^{k(\gamma,\gamma')}x}.$$

That is, the ξ_n-equivalence class of (x,γ) is

$$\boxed{\xi_n(x,\gamma) = \Big\{(\bar{x}_n,\bar{\gamma}_n),(T\bar{x}_n,S\bar{\gamma}_n),\ldots,(T^{d_n(\gamma)-1}\bar{x}_n,S^{d_n(\gamma)-1}\bar{\gamma}_n)\Big\}}$$

where $\bar{x}_n = T^{-k_n(\gamma)}x$ and, as already seen, $\bar{\gamma}_n = S^{-k_n(\gamma)}\gamma$ is the ζ_n-representative of γ. It is clear that $\xi_n \preceq \xi_{n+1}$. We consider $(\bar{x}_n,\bar{\gamma}_n)$ as the ξ_n-representative of (x,γ).

Remark 1. It is known that the set-theoretic intersection $\cap_{n\leq 0}\zeta_n$ is the orbital partition of S. As we will see (Proposition 1), the measurable hull of the tail partition $\cap\xi_n$ is the invariant σ-field of $T\times S$. But I do not know whether $\cap\xi_n$ is the orbital partition of $T \times S$.

3.2 The filtration $\mathcal{G}$

We take a random variable X_0 distributed on $\mathcal{X}$ according to ν and we set $\boxed{X_n = T^{-K_n}X_0}$, similarly to $R_n = S^{-K_n}R_0$.

Thus $(X_n,R_n)(x,\gamma) = (\bar{x}_n,\bar{\gamma}_n)$, and setting $\boxed{\mathcal{G}_n = \sigma(X_n,R_n)}$ then $\mathcal{G}_n = \sigma(\xi_n)$ is the σ-field corresponding to the measurable partition ξ_n.

Note that $X_n \sim \nu$ for every $n \leq 0$ because K_n is independent of X_0.

The filtration $\mathcal{G} = (\mathcal{G}_n)_{n\leq 0}$ is generated by the stochastic process $(X_n, V_n, \epsilon_n)_{n\leq 0}$:

$$\boxed{\mathcal{G}_n = \sigma(X_m, V_m, \epsilon_m; m \leq n)}.$$

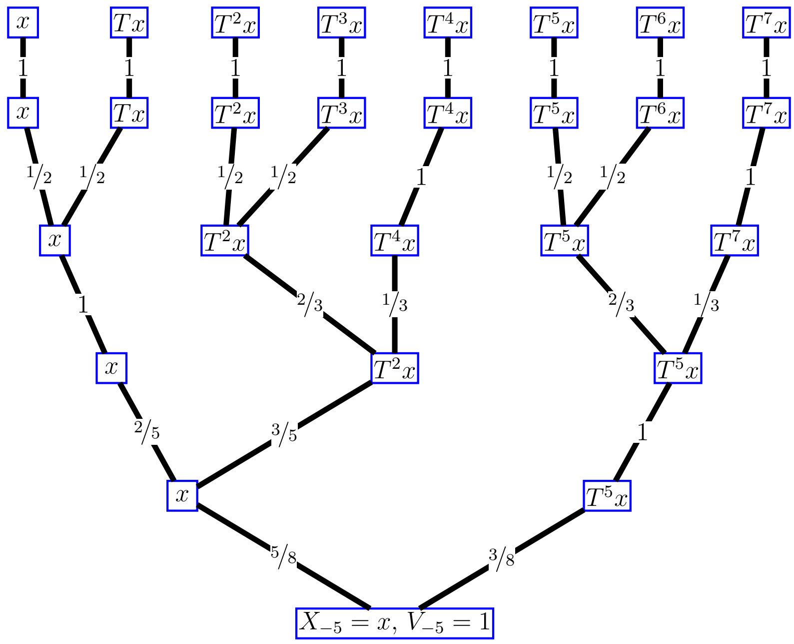


Figure 6: The process (X_n, V_n) for the golden graph

Lemma 3. *The following properties hold:*

1. *The random integer K_n is conditionally independent of $\mathcal{G}_n$ given V_n (equivalently, $(\epsilon_{n+1}, \ldots, \epsilon_0)$ is conditionally independent of $\mathcal{G}_n$ given V_n because K_n is a one-to-one function of $(\epsilon_{n+1}, \ldots, \epsilon_0)$ given V_n).*

2. *ϵ_{n+1} is a "novation" from $\mathcal{G}_n$ to $\mathcal{G}_{n+1}$, that is to say $\mathcal{G}_{n+1} = \mathcal{G}_n \vee \sigma(\epsilon_{n+1})$, and the filtration $\mathcal{F}$ is immersed in $\mathcal{G}$.*

Proof. Given $\mathcal{G}_n$, the random integer K_n corresponds to a path connecting the vertex V_n to the root vertex. That shows the conditional independence. As seen before, $K_n - K_{n+1}$ is, conditionally to V_n, a one-to-one function of ϵ_{n+1}. Since $X_{n+1} = T^{K_n - K_{n+1}} X_n$, that shows the equality $\mathcal{G}_{n+1} = \mathcal{G}_n \vee \sigma(\epsilon_{n+1})$. The immersion stems from the fact that ϵ_{n+1} is also a novation from $\mathcal{F}_n$ to $\mathcal{F}_{n+1}$. $\square$

As a consequence, the process $(X_n, V_n, \epsilon_n)_{n\leq 0}$ is Markovian.

Lemma 4. *$X_n \perp\!\!\!\perp \mathcal{F}_n$ for every $n \leq 0$.*

Proof. Take $f \in L^1$. Then

$$\mathbb{E}\Big[f(X_n) \mid \mathcal{F}_n\Big] = \mathbb{E}\Big[f(T^{-K_n} X_0) \mid \mathcal{F}_n\Big].$$

Since K_n is $\mathcal{F}_0$-measurable and $X_0 \perp\!\!\!\perp \mathcal{F}_0$,

$$\mathbb{E}\Big[f(T^{-K_n} X_0) \mid \mathcal{F}_n\Big] = \mathbb{E}\Big[h(K_n) \mid \mathcal{F}_n\Big]$$

where $h(k) = \mathbb{E}\Big[f(T^{-k} X_0)\Big]$. But $h(k) = \mathbb{E}\Big[f(X_0)\Big] = \mathbb{E}\Big[f(X_n)\Big]$ since T preserves the law of X_0. $\square$

Therefore the law of the process $(X_n, V_n, \epsilon_n)_{n\leq 0}$ can be described as follows:

- $(V_n, \epsilon_n)_{n\leq 0}$ is a path taken at random in Γ according to μ;
- $X_n \perp\!\!\!\perp (V_n, \epsilon_n)$;
- ϵ_{n+1} is conditionally independent of $\mathcal{G}_n$ given V_n;
- $X_{n+1} = T^{K_n - K_{n+1}} X_n$.

Proposition 1. *The tail σ-field $\mathcal{G}_{-\infty}$ is degenerate if and only if $T \times S$ is ergodic. More precisely, $\mathcal{G}_{-\infty}$ equals the $(T \times S)$-invariant σ-field.*

Proof. Denote by $\mathcal{I}$ the $(T \times S)$-invariant σ-field. Since the pair (X_0, R_0) generates $\mathcal{G}_0$, the degeneracy of $\mathcal{G}_{-\infty}$ is equivalent to the L^1-convergence of $\mathbb{E}\Big[f(X_0, R_0) \mid \mathcal{G}_n\Big]$ to $\mathbb{E}\Big[f(X_0, R_0) \mid \mathcal{I}\Big]$ for every bounded measurabe function f.

Recall that $X_0 = T^{K_n} X_n$ and $R_0 = S^{K_n} R_n$. Conditionally to $\mathcal{G}_n$, the random integer K_n has the uniform distribution on $\Big\{0, \ldots, \dim(V_n) - 1\Big\}$, therefore

$$\begin{aligned}\mathbb{E}\Big[f(X_0, R_0) \mid \mathcal{G}_n\Big] &= \frac{1}{\dim(V_n)} \sum_{k=0}^{\dim(V_n)-1} f\Big(T^k X_n, S^k R_n\Big) \\ &= \frac{1}{\dim(V_n)} \sum_{k=0}^{\dim(V_n)-1} f\Big(T^k T^{-K_n} X_0, S^k S^{-K_n} R_0\Big).\end{aligned}$$

Now, write

$$\sum_{k=0}^{\dim(V_n)-1} f\Big(T^k T^{-K_n} X_0, S^k S^{-K_n} R_0\Big) = \sum_{M=1}^{\dim(V_n)} \left(\sum_{k=0}^{\dim(V_n)-1} f\Big(T^k T^{-M} X_0, S^k S^{-M} R_0\Big) \right) \mathbf{1}_{K_n = M}.$$

and denote by $E(f \mid \mathcal{I})$ the conditional expectation of f given $\mathcal{I}$.

Let $\epsilon > 0$. By the ergodic theorem, for every integer N large enough and for every pair of random variables $(U, V) \sim \nu \otimes \mu$, the average $\frac{1}{N}\left(\sum_{k=0}^{N-1} f\Big(T^k U, S^k V\Big)\right)$ is ϵ-close in $L^2(\nu \otimes \mu)$ to $E(f \mid \mathcal{I})(U, V)$. For n large enough, one can apply this fact to $U = T^{-M} X_0$ and $V = S^{-M} R_0$ and $N = \dim(V_n)$, and one gets that the average

$$\frac{1}{\dim(V_n)} \sum_{k=0}^{\dim(V_n)-1} f\Big(T^k T^{-M} X_0, S^k S^{-M} R_0\Big)$$

is ϵ-close in $L^2(\nu \otimes \mu)$ to $E(f \mid \mathcal{I})(T^{-M} X_0, S^{-M} R_0) = E(f \mid \mathcal{I})(X_0, R_0)$.

Finally, using the Cauchy-Schwarz inequality,

$$\mathbb{E}\Big[\Big|\mathbb{E}\Big[f(X_0, R_0) \mid \mathcal{G}_n\Big] - E(f \mid \mathcal{I})(X_0, R_0)\Big|\Big] \leq \epsilon,$$

and the proof is over. $\square$

Remark 2. For people who deal with the filtration $\mathcal{G}$ on an abstract probability space, the equality is $\mathcal{G}_{-\infty} = (X_0, R_0)^{-1}(\mathcal{I})$, where $\mathcal{I}$ is the $(T \times S)$-invariant σ-field.

3.3 An application of the tail σ-field

As an application of Proposition 1, we provide in Proposition 2 a weighted ergodic average. We will use the following equality, seen in the proof of Proposition 1:

$$\mathbb{E}\Big[f(X_0, R_0) \mid \mathcal{G}_n\Big] = \frac{1}{\dim(V_n)} \sum_{k=0}^{\dim(V_n)-1} f\Big(T^k X_n, S^k R_n\Big). \tag{2}$$

Proposition 2. *Let c be the Chacon word defined below equality* (1)*. Define the weights*

$$a_{n,k} = \begin{cases} \frac{3}{h_n} c_k & \text{if } 0 \leq k \leq h_n - 1 \\ 0 & \text{if } k \geq h_n \end{cases}$$

Let T be an invertible measure-preserving transformation of a Lebesgue space $(\mathcal{X}, \nu)$. For $g \in L^1(\nu)$, define the weighted average

$$S_n(g)(x) = \frac{3}{h_n} \sum_{k=0}^{h_n-1} c_k g(T^k x) = \sum_{k=0}^{\infty} a_{n,k} g(T^k x).$$

Then $S_n(g) \to E(g \mid \mathcal{I})$ in probability, where $\mathcal{I}$ is the T-invariant σ-field.

We give two lemmas before proving the previous proposition.

Lemma 5. *On the same probability space, let (A_n) and (B_n) be two sequences of random variables and (E_n) be a sequence of events. Set $Y_n = A_n \mathbf{1}_{E_n} + B_n \mathbf{1}_{E_n^c}$. Assume that*

- $A_n \perp\!\!\!\perp E_n$;
- $\Pr(E_n) \to p > 0$;
- $Y_n \to Y > 0$ *in probability.*

Then $A_n \to Y$ in probability.

Proof. This follows from

$$\Pr\Big(|A_n - Y| > \epsilon\Big) = \frac{\Pr\Big(|A_n - Y| > \epsilon, E_n\Big)}{\Pr(E_n)} = \frac{\Pr\Big(|Y_n - Y| > \epsilon, E_n\Big)}{\Pr(E_n)} \leq \frac{\Pr\Big(|Y_n - Y| > \epsilon\Big)}{\Pr(E_n)}.$$

□

Lemma 6. *Let S and T be two invertible measure-preserving transformations. If S is weakly mixing, then the $(T \times S)$-invariant σ-field is the product of the T-invariant σ-field times the trivial σ-field.*

Proof. It is well known that the product of an ergodic transformation times a weakly mixing transformation is ergodic. The lemma, which generalizes this result, can be proved by looking at the space of ergodic components. □

Proof of Proposition 2. Let S be the Chacon adic transformation, which is known to be weakly mixing. Apply equality (2) to the function $f(x, \gamma) = g(x)\mathbf{1}_{x \in I}$ where $I = \{\gamma \mid v_{-1}(\gamma) = 2\}$ was introduced before equality (1). This gives

$$\begin{aligned}
\mathbb{E}\Big[g(X_0)\mathbf{1}_{R_0\in I} \mid \mathcal{G}_n\Big] &= \frac{1}{\dim(V_n)}\sum_{k=0}^{\dim(V_n)-1} g(T^k X_n)\mathbf{1}_{S^k R_n\in I} \\
&= \left(\frac{1}{h_n}\sum_{k=0}^{h_n-1} g(T^k X_n)\mathbf{1}_{S^k R_n\in I}\right)\mathbf{1}_{V_n=1} + \left(g(X_n)\mathbf{1}_{R_n\in I}\right)\mathbf{1}_{V_n=2} \\
&= \underbrace{\left(\frac{1}{h_n}\sum_{k=0}^{h_n-1} c_k g(T^k X_n)\right)}_{=:A_n}\mathbf{1}_{V_n=1} + \underbrace{\left(g(X_n)\mathbf{1}_{R_n\in I}\right)}_{=:B_n}\mathbf{1}_{V_n=2},
\end{aligned}$$

where the last equality comes from equality (1).

By Proposition 1, $\mathbb{E}\Big[g(X_0)\mathbf{1}_{R_0\in I} \mid \mathcal{G}_n\Big] \to \mathbb{E}\Big[g(X_0)\mathbf{1}_{R_0\in I} \mid \mathcal{G}_{-\infty}\Big]$. We know that $\mathcal{G}_{-\infty} = \mathcal{I}\otimes\{\varnothing,\Gamma\}$ by Lemma 6, therefore

$$\mathbb{E}\Big[g(X_0)\mathbf{1}_{R_0\in I} \mid \mathcal{G}_{-\infty}\Big] = \Pr(R_0\in I)\mathbb{E}\Big[g(X_0) \mid \mathcal{I}\otimes\{\varnothing,\Gamma\}\Big] = \frac{1}{3}E(g\mid\mathcal{I}).$$

Since $X_n \perp\!\!\!\perp V_n$ (Lemma 4) and $\Pr(V_n=1)\to 1$, one gets $A_n \to \frac{1}{3}E(g\mid\mathcal{I})$ by virtue of Lemma 5. $\square$

4 The scale of an automorphism

In the case when S is the usual adic transformation isomorphic to the (r_n)-ary odometer, the filtration $\mathcal{G}$ is the one introduced by Laurent in [5], whose standardness provides an equivalent definition of the first definition of the scale of an automorphism originally introduced by Vershik in [7]. This definition is the following one.

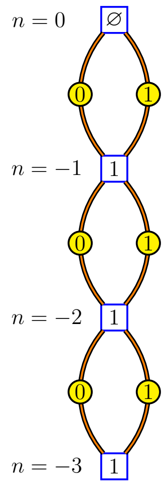


Figure 7: The Bratteli graph of the dyadic odometer

Definition 1. Let $\mathcal{G}$ be the filtration of the previous section in the case when S is the (r_n)-ary odometer. The sequence (r_n) belongs to the scale of T if $\mathcal{G}$ is standard.

In this case, it is shown in [5] that the tail σ-field $\mathcal{G}_{-\infty}$ is degenerate if and only if $T^{\prod_{k=n+1}^{0} r_k}$ is ergodic for every $n\le 0$. This is equivalent to the ergodicity of the product of T with the (r_n)-ary odometer, as expected in view of Proposition 1.

The filtration $\mathcal{G}$ cannot be standard when $\mathcal{G}_{-\infty}$ is not degenerate. A more general definition is proposed in [5] to deal with this situation: say that (r_n) belongs to the scale of T if $\mathcal{G}$ is conditionally standard given $\mathcal{G}_{-\infty}$. But this generalization of standardness has not been studied yet in the literature.

Therefore, the definition of the scale can be generalized as follows.

Definition 2. Let $\mathcal{G}$ be the filtration of the previous section associated to an adic transformation S and an automorphism T. Say that S is in the scale of T if $\mathcal{G}$ is standard, or, more generally, if $\mathcal{G}$ is conditionally standard given $\mathcal{G}_{-\infty}$.

In the case when S is the (r_n)-ary odometer and T is a Bernoulli automorphism, standardness of $\mathcal{G}$ has been characterized in terms of the asymptotic behavior or the sequence (r_n) (see [1, 3, 5]).

5 Adic split-words processes

We use the notations of the previous section. Let P be a finite or countable partition of $\mathcal{X}$. The elements of P are labelled by the letters of an alphabet A. For $x \in \mathcal{X}$ we denote by $P(x) \in A$ the label of the block to which x belongs.

Define the random word W_n by

$$W_n = P(X_n)P(TX_n)\dots P(T^{\dim(V_n)-1}X_n).$$

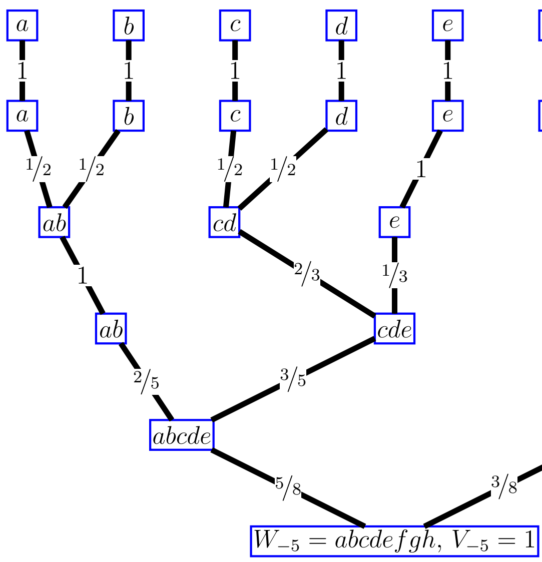


Figure 8: The process (W_n, V_n) for the golden graph

Lemma 7. *The filtration generated by $(W_n, V_n, \epsilon_n)_{n\leq 0}$ is immersed in $\mathcal{G}$. It equals $\mathcal{G}$ when P is a generating partition of T.*

Proof. Denote by $\mathcal{G}'$ this filtration. The immersion follows from the fact that the random variable ϵ_{n+1} is a novation of $\mathcal{G}'_n$ to $\mathcal{G}'_{n+1}$ and its conditional law given $\mathcal{G}'_n$ is the same as given $\mathcal{G}_n$ (Lemma 3).

To show that $\mathcal{G}' = \mathcal{G}$ when P is generating, it suffices to show that X_0 is measurable with respect to $\mathcal{G}'_0$. Since $X_n = T^{-K_n}X_0$, this follows from Lemma 1. $\square$

The filtration generated by $(W_n, V_n, \epsilon_n)_{n\leq 0}$ is also generated by $(W_n, \epsilon_n)_{n\leq 0}$ when two different vertices at each level n have different dimensions, because the length of W_n is $\dim(V_n)$.

The space of trajectories of a split-word process can be displayed on a Bratteli graph. Figure 9 shows the case when S is the dyadic odometer, T is the Bernoulli shift on $\{a, b\}^{\mathbb{Z}}$ and

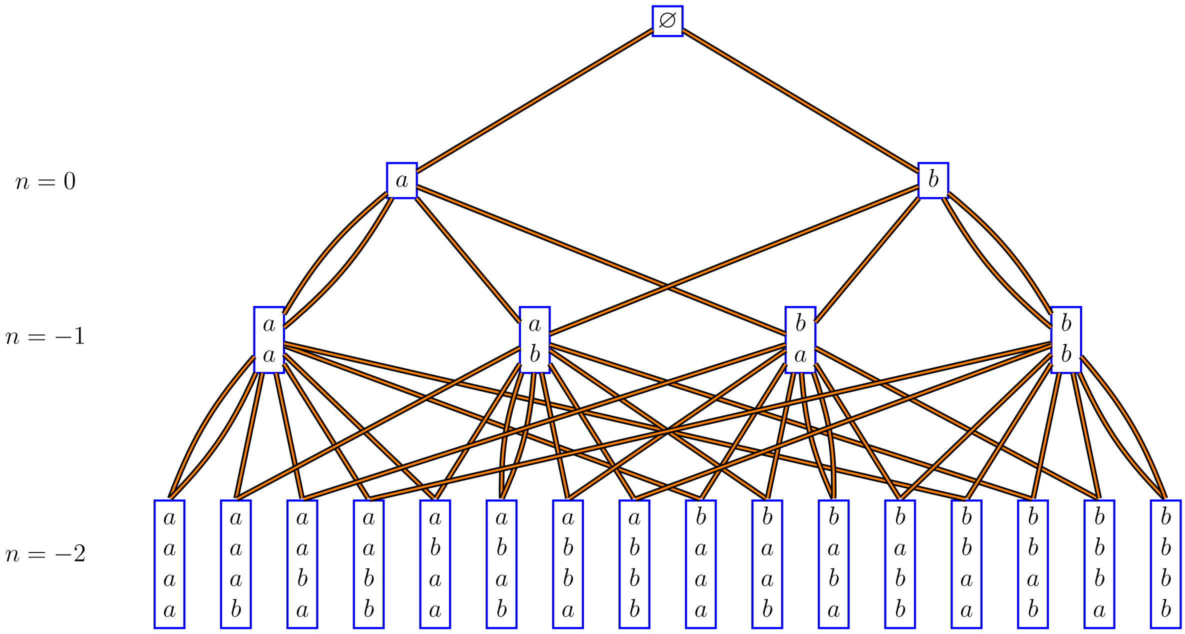


Figure 9: (W_n, ϵ_n) when S is the dyadic odometer and T is the Bernoulli 2-shift

P is the usual generating partition according to the central coordiinate. This graph is called the *graph of ordered pairs* by Vershik.

Note that $W_{-1} = W_0$ when there is a unique edge between the root vertex and each vertex at level -1 (that is to say when ϵ_0 takes only one value). In the general situation, W_0 is a function of W_{-1} and ϵ_0. Therefore one can set $W_0 = \varnothing$ without changing the filtration. An example is shown on Figure 10 for the case when S is the golden adic transformation, T is the Chacon transformation, and P is the generating partition $\{[0, 2/3[, [2/3, 1[\}$.

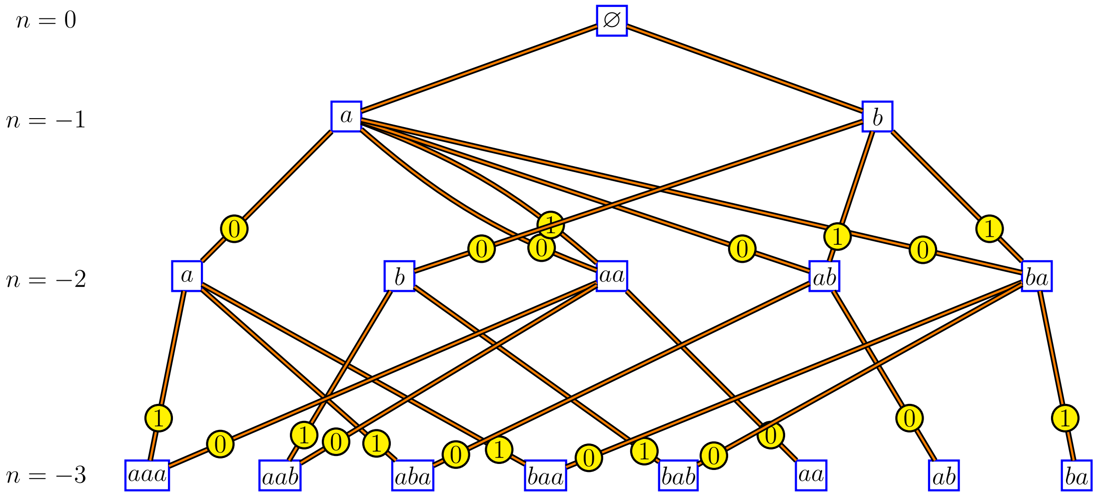


Figure 10: (W_n, ϵ_n) when S is the golden adic transformation and T is the Chacon transformation

Question 1. What is the adic transformation on these graphs? Since the tail σ-field is the $(T \times S)$-invariant σ-field, one could expect that the adic transformation is isomorphic to $T \times S$. I believe it is true.

6 Example: the golden graph and the golden rotation

The filtration $\mathcal{F}$ was introduced in Section 2.3 in the case when S is the adic transformation on the golden graph.

This adic transformation is isomorphic to the golden rotation on S^1, with angle $\theta = \frac{1}{1+\phi} = \frac{2}{3+\sqrt{5}}$.

For a given transformation T, the law of the stochastic process $(X_n, V_n, \epsilon_n)_{n\leq 0}$ generating the filtration $\mathcal{G}$ is given by

$$X_{n+1} = T^{\epsilon_{n+1} f_{n+1}} X_n.$$

The product of a rotation with itself is not ergodic. Hence, in the case when T is the golden rotation, the tail σ-field $\mathcal{G}_{-\infty}$ is not degenerate. For this case, one can see that X_n almost surely goes to a random variable $X_{-\infty}$ as $n \to -\infty$. Indeed, first observe that

$$X_{n+1} = X_n \text{ or } X_{n+1} = X_n + \theta f_{n+1}.$$

But the distance between θf_n and 0 in S^1 is less than $1/f_{n-1}$, because of the inequality

$$|\theta f_n - f_{n+2}| \leq \frac{1}{f_{n-1}},$$

coming from the well-known results about continued fraction (the continued fraction expansion of θ is $[0, 2, 1, 1, \ldots]$). Therefore $|X_{n+1} - X_n| \leq 1/f_n$, and $X_n \to X_{-\infty}$ because $1/f_n$ is the general term of a convergent series. Of course $X_{-\infty}$ has the uniform distribution on S^1, like each X_n.

Question 2. Does $\mathcal{G}_{-\infty} = \sigma(X_{-\infty})$?

Question 3. Conditionally to $X_{-\infty}$, is the filtration $\mathcal{G}$ isomorphic to $\mathcal{F}$?

7 Example: the Chacon graph and Bernoulli automorphisms

7.1 The filtration $\mathcal{G}$

In the case when S is the Chacon adic transformation, the filtration $\mathcal{F}$ was introduced in Section 2.4.

The law of the process $(X_n, V_n, \epsilon_n)_{n\leq 0}$ generating the filtration $\mathcal{G}$ is given by (see Figure 11, where a red box indicates $V_n = 2$):

- $(V_n, \epsilon_n)_{n\leq 0}$ has the law given in Section 2.4;
- $X_n \sim \nu$ is independent of (V_n, ϵ_n);
- $X_{n+1} = T^{f_{n+1}(\epsilon_{n+1})} X_n$ (the f_n are given in Section 2.4).

In this case, a split-word process $(W_n, V_n, \epsilon_n)_{n\leq 0}$ (Section 5) has the following dynamics (see Figure 12). The length of the word W_n is h_n if $V_n = 1$, and length 1 if $V_n = 2$. When $V_n = 1$, we consider W_n as the concatenation of four subwords with respective lengths h_{n+1}, h_{n+1}, 1 and h_{n+1}. Then W_{n+1} is the subword of W_n selected by the value of ϵ_{n+1}.

The Chacon transformation S is known to be weakly mixing. This implies that $T \times S$ is ergodic whenever T is ergodic, therefore the filtration $\mathcal{G}$ has a degenerate tail σ-field $\mathcal{G}_{-\infty}$ whenever T is ergodic (Proposition 1).

We will prove the following result.

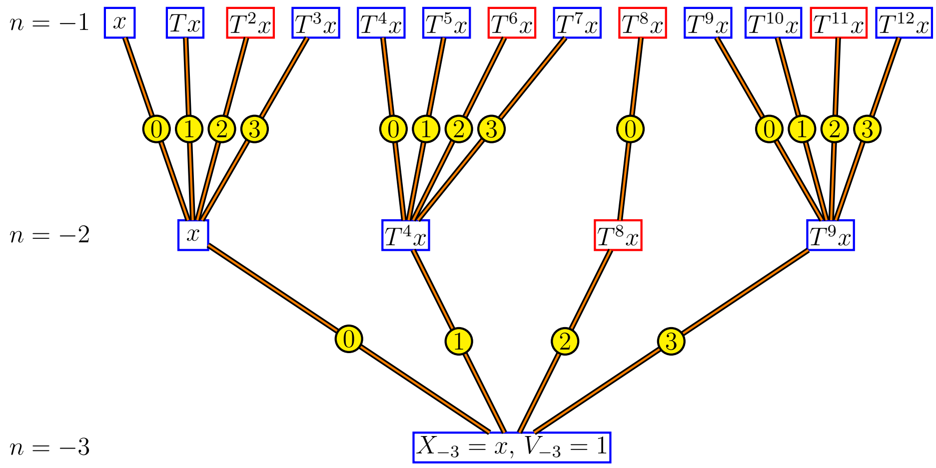


Figure 11: (X_n, V_n, ϵ_n)

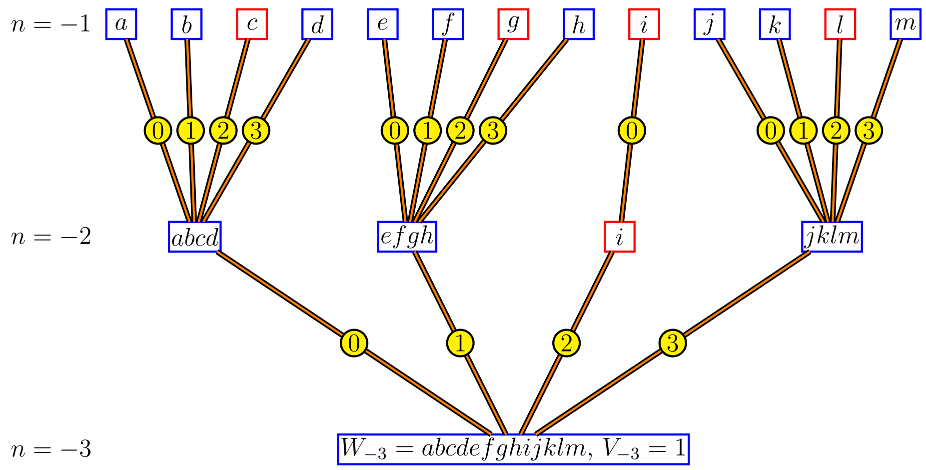


Figure 12: (W_n, V_n, ϵ_n)

Proposition 3. *When T is a Bernoulli automorphism, the filtration $\mathcal{G}$ is not standard.*

Hereafter we consider that T is a Bernoulli shift on a a countable alphabet A and that $(W_n, V_n, \epsilon_n)_{n\leq 0}$ is the split-word process obtained with the partition of $A^{\mathbb{Z}}$ according to the central coordinate. Thus W_n is a word made of i.i.d. letters on A.

7.2 The process $(\widetilde{W}_n, \widetilde{V}_n, \widetilde{\epsilon}_n)$

Proposition 3 will be derived from the non-standardness of the filtration $\widetilde{\mathcal{G}}$ generated by the following process $(\widetilde{W}_n, \widetilde{V}_n, \widetilde{\epsilon}_n)_{n\leq 0}$ (see Figure 13), which is asymptotically the same as the process (W_n, V_n, ϵ_n), in a sense that will be made precise in Section 7.3.

For every $n \leq -1$, $\widetilde{V}_n$ is a constant random variable equal to 1, and $\widetilde{V}_0 = \varnothing$. The process $(\widetilde{\epsilon}_n)_{n\leq -1}$ is a sequence of independent random variables having the uniform law on $\{0, 1, 3\}$, and $\widetilde{\epsilon}_0$ is a constant random variable equal to 0. The random variable $\widetilde{W}_n$ is a random word of length h_n on A, made of i.i.d. letters, and it is independent of $\widetilde{\epsilon}_n$. The random variable $\widetilde{\epsilon}_{n+1}$ is independent of $\widetilde{GG}_n$, and $\widetilde{W}_{n+1}$ is the subword of $\widetilde{W}_n$ selected by the value of $\widetilde{\epsilon}_{n+1}$, in the same way that W_{n+1} is the subword of W_n selected by the value of ϵ_{n+1}.

Thus, $(\widetilde{\mathcal{G}}_n)_{n\leq -1}$ is a 3-adic filtration: $\widetilde{\mathcal{G}}_{n+1} = \widetilde{\mathcal{G}}_n \vee \sigma(\widetilde{\epsilon}_{n+1})$ and $\widetilde{\epsilon}_{n+1}$ is independent of $\widetilde{\mathcal{G}}_n$ and has a uniform distribution on a set with three elements.

We will see in Section 7.3 that the tail σ-field $\widetilde{\mathcal{G}}_{-\infty}$ is degenerate as a consequence of the first point of Theorem 7.1.

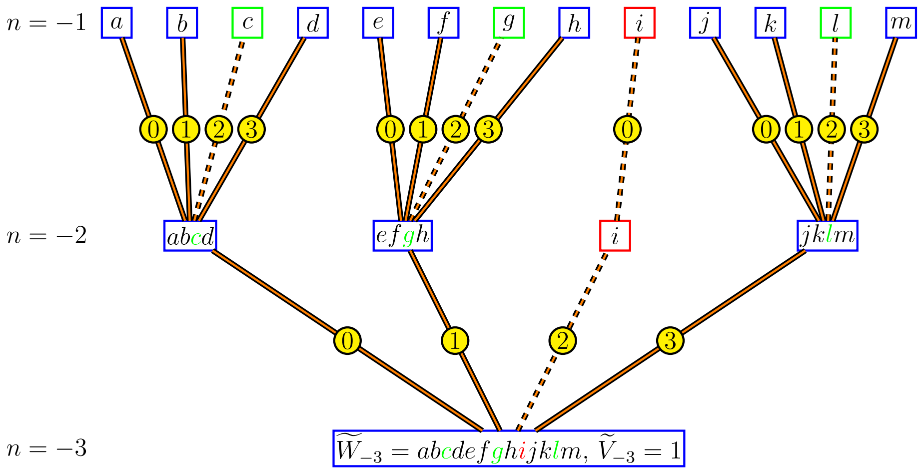


Figure 13: $(\widetilde{W}_n, \widetilde{V}_n, \widetilde{\epsilon}_n)$

Proposition 4. *The filtration $\widetilde{\mathcal{G}}$ is not standard.*

Proof. Let $\widetilde{Y}_n$ be the word obtained by deleting the letters of $\widetilde{W}_n$ shown in color on Figure 13. These are the letters of $\widetilde{W}_n$ which have a zero probability to be selected by the conditional law $\mathcal{L}(\widetilde{W}_{-1} \mid \widetilde{W}_n)$. The positions of these letters are given by the 1's in the Chacon word introduced after equality (1). The process $(\widetilde{Y}_n, \widetilde{\epsilon}_n)_{n \leq -1}$ is a 3-adic split-word process with i.i.d. letters on A. It is known that the filtration it generates is not standard (see [1, 3, 5]). This filtration is immersed in $\widetilde{\mathcal{G}}$, therefore $\widetilde{\mathcal{G}}$ is not standard. $\square$

7.3 Closely joinable processes

In this section we provide Theorem 7.1 which is the tool with the help of which we will derive Proposition 3 (non-standardness of $\mathcal{G}$) from Proposition 4 (non-standardness of $\widetilde{\mathcal{G}}$).

When two random vectors $(X_n)_{0 \leq n \leq n_0}$ and $(Y_n)_{0 \leq n \leq n_0}$ are defined on the same probability space and the filtrations they generate are jointly immersed, we will write and say that $\left\{ \begin{smallmatrix} (X_n)_{0 \leq n \leq n_0} \\ (Y_n)_{0 \leq n \leq n_0} \end{smallmatrix} \right.$ is a *synchronous joining*. That means, after introducing the σ-fields $\mathcal{B}_n = \sigma(X_0, \ldots, X_n)$ and $\mathcal{C}_n = \sigma(Y_0, \ldots, Y_n)$, that $\mathcal{B}_{n+1} \perp\!\!\!\perp_{\mathcal{B}_n} \mathcal{B}_n \vee \mathcal{C}_n$ and $\mathcal{C}_{n+1} \perp\!\!\!\perp_{\mathcal{C}_n} \mathcal{B}_n \vee \mathcal{C}_n$.

When $\left\{ \begin{smallmatrix} (X'_n)_{0 \leq n \leq n_0} \\ (Y'_n)_{0 \leq n \leq n_0} \end{smallmatrix} \right.$ is a *synchronous joining* in the situation when $(X'_n)_{0 \leq n \leq n_0}$ is a copy of a random vector $(X_n)_{0 \leq n \leq n_0}$ and $(Y'_n)_{0 \leq n \leq n_0}$ is a copy of a random vector $(Y_n)_{0 \leq n \leq n_0}$, we also say that $\left\{ \begin{smallmatrix} (X'_n)_{0 \leq n \leq n_0} \\ (Y'_n)_{0 \leq n \leq n_0} \end{smallmatrix} \right.$ is a *synchronous joining of $(X_n)_{0 \leq n \leq n_0}$ and $(Y_n)_{0 \leq n \leq n_0}$*.

Definition 3. Let $(X_n)_{n \leq 0}$ and $(Y_n)_{n \leq 0}$ be two stochastic processes. We say that they are *closely joinable* if for every $\delta > 0$ <u>and for every $M \leq 0$</u>, there exists <u>$N_\delta \leq M$</u> such that for every $n_0 \leq N_\delta$ there exists a synchronous joining $\left\{ \begin{smallmatrix} (X'_n)_{n_0 \leq n \leq N_\delta} \\ (Y'_n)_{n_0 \leq n \leq N_\delta} \end{smallmatrix} \right.$ of $(X_n)_{n_0 \leq n \leq N_\delta}$ and $(Y_n)_{n_0 \leq n \leq N_\delta}$ such that

$$\Pr(X'_{N_\delta} = Y'_{N_\delta}) > 1 - \delta.$$

Theorem 7.1. *Let $(X_n)_{n \leq 0}$ and $(Y_n)_{n \leq 0}$ be two Markovian stochastic processes, and denote by $\mathcal{F}$ and $\mathcal{G}$ the filtrations they respectively generate. Assume $(X_n)_{n \leq 0}$ and $(Y_n)_{n \leq 0}$ are closely joinable.*

1. *The σ-fields $\mathcal{F}_{-\infty}$ and $\mathcal{G}_{-\infty}$ are equal.*
2. *The filtration $\mathcal{F}$ is I-cosy if and only if the filtration $\mathcal{G}$ is I-cosy.*

We firstly prove the first point of this theorem.

Proof of 1 in Theorem 7.1. To establish the claim, it suffices to show that the conditional law $\mathcal{L}(X_M \mid \mathcal{F}_{-\infty})$ is $\mathcal{G}_{-\infty}$-measurable random variable for every integer $M \leq 0$. To do so, we will prove that the conditional expectation $\mathbb{E}\Big[f(X_M) \mid \mathcal{F}_{-\infty}\Big]$ can be made as L^1-close as desired to a $\mathcal{G}_{-\infty}$-measurable random variable for any Borelian function f taking its values in $[0,1]$.

We set $F_M = f(X_M)$. Let $\delta > 0$ and $N = N_\delta \leq M$ the integer provided by the joinability assumption. One has $\mathbb{E}[F_M \mid \mathcal{F}_N\Big] = g(X_N)$ for a certain Borelian function g taking its values in $[0,1]$. Take $n_0 \leq N$ small enough in order that

$$\mathbb{E}\left[\Big|\mathbb{E}[F_M \mid \mathcal{F}_{n_0}] - \mathbb{E}[F_M \mid \mathcal{F}_{-\infty}]\Big|\right] \leq \delta \quad \text{and} \quad \mathbb{E}\left[\Big|\mathbb{E}\Big[g(Y_N) \mid \mathcal{G}_{n_0}\Big] - \mathbb{E}\Big[g(Y_N) \mid \mathcal{G}_{-\infty}\Big]\Big|\right] \leq \delta,$$

which is possibly by virtue of the theorem on reverse martingale convergence.

Now, take the joining provided by the joinability assumption and set $F'_M = f(X'_M)$. One has $\mathbb{E}\Big[\big|g(X'_N) - g(Y'_N)\big|\Big] \leq \delta$. Therefore

$$\mathbb{E}\left[\Big|\mathbb{E}\Big[g(X'_N) \mid \sigma(X'_{n_0}, Y'_{n_0})\Big] - \mathbb{E}\Big[g(Y'_N) \mid \sigma(X'_{n_0}, Y'_{n_0})\Big]\Big|\right] \leq \delta$$

because of the contractivity of the conditional expectation. By immersion,

$$\mathbb{E}\Big[g(X'_N) \mid \sigma(X'_{n_0}, Y'_{n_0})\Big] = \mathbb{E}\Big[g(X'_N) \mid \sigma(X'_{n_0})\Big] = \mathbb{E}\Big[F'_M \mid \sigma(X'_{n_0})\Big]$$

and

$$\mathbb{E}\Big[g(Y'_N) \mid \sigma(X'_{n_0}, Y'_{n_0})\Big] = \mathbb{E}\Big[g(Y'_N) \mid \sigma(Y'_{n_0})\Big].$$

On the other hand,

$$\begin{aligned}\mathbb{E}\left[\Big|\mathbb{E}\Big[F'_M \mid \sigma(X'_{n_0})\Big] - \mathbb{E}\Big[g(Y'_N) \mid \sigma(Y'_{n_0})\Big]\Big|\right] &= \mathbb{E}\left[\Big|\mathbb{E}\Big[F_M \mid \sigma(X_{n_0})\Big] - \mathbb{E}\Big[g(Y_N) \mid \sigma(Y_{n_0})\Big]\Big|\right] \\ &= \mathbb{E}\left[\Big|\mathbb{E}\Big[F_M \mid \mathcal{F}_{n_0}\Big] - \mathbb{E}\Big[g(Y_N) \mid \mathcal{G}_{n_0}\Big]\Big|\right]\end{aligned}$$

By combining the previous equalities and inequalities,

$$\mathbb{E}\left[\Big|\mathbb{E}[F_M \mid \mathcal{F}_{-\infty}\Big] - \mathbb{E}\Big[g(Y_N) \mid \mathcal{G}_{-\infty}\Big]\Big|\right] \leq 3\delta,$$

and the proof is over. □

The second point of the theorem will be proved with the help of the following lemma.

Lemma 8. *Let* $\left\{\begin{smallmatrix} (X_n)_{0\leq n\leq n_0} \\ (Y_n)_{0\leq n\leq n_0}\end{smallmatrix}\right.$ *be a synchronous joining of two random vectors* $(X_n)_{0\leq n\leq n_0}$ *and* $(Y_n)_{0\leq n\leq n_0}$. *One assumes that a synchronous joining* $\left\{\begin{smallmatrix} (X'_n)_{0\leq n\leq n_0} \\ (X''_n)_{0\leq n\leq n_0}\end{smallmatrix}\right.$ *of two copies of* $(X_n)_{0\leq n\leq n_0}$ *is given on some probability space. Then, on an enlargement of this probability space, there exists a synchronous joining* $\left\{\begin{smallmatrix} (Y'_n)_{0\leq n\leq n_0} \\ (Y''_n)_{0\leq n\leq n_0}\end{smallmatrix}\right.$ *of two copies of* $(Y_n)_{0\leq n\leq n_0}$ *such that* $\left\{\begin{smallmatrix} (X'_n)_{0\leq n\leq n_0} \\ (Y'_n)_{0\leq n\leq n_0}\end{smallmatrix}\right.$ *and* $\left\{\begin{smallmatrix} (X''_n)_{0\leq n\leq n_0} \\ (Y''_n)_{0\leq n\leq n_0}\end{smallmatrix}\right.$ *are two copies of* $\left\{\begin{smallmatrix} (X_n)_{0\leq n\leq n_0} \\ (Y_n)_{0\leq n\leq n_0}\end{smallmatrix}\right.$.

Moreover, if $X'_0 \perp\!\!\!\perp_{\sigma(X'_0)\cap\sigma(X''_0)} X''_0$, *then* $Y'_0 \perp\!\!\!\perp_{\sigma(X'_0)\cap\sigma(X''_0)} Y''_0$.

Proof. In the proof, we will use the three following elementary facts about conditional independence:

(i) If $U \perp\!\!\!\perp \mathcal{A} \supset \mathcal{B}$, then $X \perp\!\!\!\perp_{\mathcal{B}} \mathcal{A}$ for any $\sigma(\mathcal{B}, U)$-measurable r.v. X.

(ii) If $U \perp\!\!\!\perp \sigma(\mathcal{B}, X)$ then $X \perp\!\!\!\perp_{\mathcal{B}} U$ (hence $X \perp\!\!\!\perp_{\mathcal{B}} \sigma(\mathcal{B}, U)$).

(iii) If $\mathcal{B} \subset \mathcal{A}$, then the two conditional independences $Y \perp\!\!\!\perp_{\mathcal{B}\vee\sigma(X)} \mathcal{A}$ and $X \perp\!\!\!\perp_{\mathcal{B}} \mathcal{A}$ imply $Y \perp\!\!\!\perp_{\mathcal{B}} \mathcal{A}$.

On the probability space of $\left\{ \begin{smallmatrix} (X_n)_{0\le n\le n_0} \\ (Y_n)_{0\le n\le n_0} \end{smallmatrix} \right.$, one can assume there exist some random variables $U_0, \ldots, U_{n_0}$ such that

- $U_n \perp\!\!\!\perp \Big((X_0, U_0), \ldots, (X_{n-1}, U_{n-1}), X_n\Big)$;
- $Y_n = f_n\Big((X_0, Y_0), \ldots, (X_{n-1}, Y_{n-1}), X_n, U_n\Big)$ for some Borelian functions f_n.

We denote by $(\mathcal{E}'_0, \ldots, \mathcal{E}'_{n_0})$ the filtration generated by $(X'_n)_{0\le n\le n_0}$ and by $(\mathcal{E}''_0, \ldots, \mathcal{E}''_{n_0})$ the filtration generated by $(X''_n)_{0\le n\le n_0}$.

On the probability space of $\left\{ \begin{smallmatrix} (X'_n)_{0\le n\le n_0} \\ (X''_n)_{0\le n\le n_0} \end{smallmatrix} \right.$, one can assume there are two copies $\mathbf{U}' = (U'_0, \ldots, U'_{n_0})$ and $\mathbf{U}'' = (U''_0, \ldots, U''_{n_0})$ of $(U_0, \ldots, U_{n_0})$ such that $\mathbf{U}' \perp\!\!\!\perp \mathbf{U}''$ and $(\mathbf{U}', \mathbf{U}'') \perp\!\!\!\perp \mathcal{E}'_{n_0} \vee \mathcal{E}''_{n_0}$.

We set $Y'_0 = f_0(X'_0, U'_0)$ and $Y''_0 = f_0(X''_0, U''_0)$. Then it is not difficult to check the last point of the lemma: $Y'_0 \perp\!\!\!\perp_{\sigma(X'_0)\cap\sigma(X''_0)} Y''_0$ if $X'_0 \perp\!\!\!\perp_{\sigma(X'_0)\cap\sigma(X''_0)} X''_0$.

Now we recursively set

$$Y'_n = f_n\Big((X'_0, Y'_0), \ldots, (X'_{n-1}, Y'_{n-1}), X'_n, U'_n\Big)$$

and

$$Y''_n = f_n\Big((X''_0, Y''_0), \ldots, (X''_{n-1}, Y''_{n-1}), X''_n, U''_n\Big).$$

In this way, it is clear that $\left\{ \begin{smallmatrix} (X'_n)_{0\le n\le n_0} \\ (Y'_n)_{0\le n\le n_0} \end{smallmatrix} \right.$ and $\left\{ \begin{smallmatrix} (X''_n)_{0\le n\le n_0} \\ (Y''_n)_{0\le n\le n_0} \end{smallmatrix} \right.$ are two copies of $\left\{ \begin{smallmatrix} (X_n)_{0\le n\le n_0} \\ (Y_n)_{0\le n\le n_0} \end{smallmatrix} \right.$.

We define the filtrations $(\mathcal{G}'_0, \ldots, \mathcal{G}'_{n_0})$ and $(\mathcal{G}''_0, \ldots, \mathcal{G}''_{n_0})$ by setting

$$\mathcal{G}'_n = \sigma\Big((X'_0, Y'_0), \ldots, (X'_n, Y'_n)\Big) \quad \text{and} \quad \mathcal{G}''_n = \sigma\Big((X''_0, Y''_0), \ldots, (X''_n, Y''_n)\Big),$$

and the filtration $(\mathcal{H}_0, \ldots, \mathcal{H}_{n_0})$ by setting

$$\mathcal{H}_n = (\mathcal{E}'_n \vee \mathcal{E}''_n) \vee \sigma\Big((U'_0, U''_0), \ldots, (U'_n, U''_n)\Big) \supset \mathcal{G}'_n \vee \mathcal{G}''_n.$$

By property *(ii)*,

$$X'_{n+1} \perp\!\!\!\perp_{\mathcal{E}'_n\vee\mathcal{E}''_n} \mathcal{H}_n \quad \text{and} \quad X''_{n+1} \perp\!\!\!\perp_{\mathcal{E}'_n\vee\mathcal{E}''_n} \mathcal{H}_n$$

and by the immersion property,

$$X'_{n+1} \perp\!\!\!\perp_{\mathcal{E}'_n} \mathcal{H}_n \quad \text{and} \quad X''_{n+1} \perp\!\!\!\perp_{\mathcal{E}''_n} \mathcal{H}_n \tag{3}$$

Therefore

$$X'_{n+1} \perp\!\!\!\perp_{\mathcal{G}'_n} \mathcal{H}_n \quad \text{and} \quad X''_{n+1} \perp\!\!\!\perp_{\mathcal{G}''_n} \mathcal{H}_n \tag{4}$$

because $\mathcal{E}'_n \subset \mathcal{G}'_n \subset \mathcal{H}_n$ and $\mathcal{E}''_n \subset \mathcal{G}''_n \subset \mathcal{H}_n$.

By property *(i)*,

$$Y'_{n+1} \perp\!\!\!\perp_{\mathcal{G}'_n\vee\sigma(X'_{n+1})} \mathcal{H}_n \quad \text{and} \quad Y''_{n+1} \perp\!\!\!\perp_{\mathcal{G}''_n\vee\sigma(X''_{n+1})} \mathcal{H}_n,$$

and by (4) and property *(iii)*,

$$Y'_{n+1} \perp\!\!\!\perp_{\mathcal{G}'_n} \mathcal{H}_n \quad \text{and} \quad Y''_{n+1} \perp\!\!\!\perp_{\mathcal{G}''_n} \mathcal{H}_n$$

and by the immersion property

$$Y'_{n+1} \perp\!\!\!\perp_{\sigma(Y'_0,\ldots,Y'_n)} \mathcal{H}_n \quad \text{and} \quad Y''_{n+1} \perp\!\!\!\perp_{\sigma(Y''_0,\ldots,Y''_n)} \mathcal{H}_n.$$

Thus, we have proved that the four filtrations generated by $(X'_n)_{0\le n\le n_0}$, $(X''_n)_{0\le n\le n_0}$, $(Y'_n)_{0\le n\le n_0}$, and $(Y''_n)_{0\le n\le n_0}$ are jointly immersed in the filtration $(\mathcal{H}_n)_{0\le n\le n_0}$. $\square$

Lemma 9. *Let $(Y_n)_{n\le 0}$ be a Markov process. For two integers $n_0 \le N < 0$, let* $\left\{\begin{smallmatrix} (Y'_n)_{n_0\le n\le N} \\ (Y''_n)_{n_0\le n\le N}\end{smallmatrix}\right.$ *be a synchronous joining of two copies of $(Y_n)_{n_0\le n\le N}$ such that $\Pr(Y'_N \neq Y''_N) < \delta$. Then it is possible to extend this joining to a synchronous joining* $\left\{\begin{smallmatrix} (Y'_n)_{n_0\le n\le 0} \\ (Y''_n)_{n_0\le n\le 0}\end{smallmatrix}\right.$ *of $(Y_n)_{n_0\le n\le 0}$ such that* $\Pr(Y'_N \neq Y''_N, \ldots, Y'_0 \neq Y''_0) < \delta$.

Proof. One can assume that $Y_{n+1} = f_n(Y_n, U_{n+1})$ where $(U_{N+1}, \ldots, U_0)$ is a vector of independent $\mathcal{U}(0,1)$ random variables such that $U_{n+1} \perp\!\!\!\perp (Y_N, U_{N+1}, \ldots, U_n)$. On the other hand, on the probability space of $\left\{\begin{smallmatrix} (Y'_n)_{n_0\le n\le N} \\ (Y''_n)_{n_0\le n\le N}\end{smallmatrix}\right.$, one can assume there is a vector $(\bar{U}_{N+1}, \ldots, \bar{U}_0)$ of independent $\mathcal{U}(0,1)$ random variables which is independent of $(Y'_n, Y''_n)_{n_0\le n\le N}$. Then we extend the joining by recursively setting $Y'_{n+1} = f_n(Y'_n, \bar{U}_{n+1})$ and $Y''_{n+1} = f_n(Y''_n, \bar{U}_{n+1})$. □

Lemma 10. *Let $\mathcal{G}$ be the filtration generated by a Markov process $(Y_n)_{n\le 0}$. Assume that for any integer $M \le 0$ and any real number $\delta > 0$, there exists two integers $n_0 \le N_\delta \le M$ and a synchronous joining* $\left\{\begin{smallmatrix} (Y'_n)_{n_0\le n\le N_\delta} \\ (Y''_n)_{n_0\le n\le N_\delta}\end{smallmatrix}\right.$ *of two copies of $(Y_n)_{n_0\le n\le N_\delta}$ such that $Y'_{n_0} \perp\!\!\!\perp Y''_{n_0}$ and $\Pr(Y'_{N_\delta} \neq Y''_{N_\delta}) < \delta$. Then $\mathcal{G}$ is I-cosy.*

Proof. In order for $\mathcal{G}$ to be I-cosy, it suffices that $(Y_M, \ldots, Y_0)$ satisfies the I-cosiness criterion for every $M \le 0$ (see [3]). By Lemma 9, the synchronous joining $\left\{\begin{smallmatrix} (Y'_n)_{n_0\le n\le N_\delta} \\ (Y''_n)_{n_0\le n\le N_\delta}\end{smallmatrix}\right.$ given by the assumption can be extended to a synchronous joining $\left\{\begin{smallmatrix} (Y'_n)_{n_0\le n\le 0} \\ (Y''_n)_{n_0\le n\le 0}\end{smallmatrix}\right.$ of two copies of $(Y_n)_{n_0\le n\le 0}$ satisfying $\Pr(Y'_{N_\delta} \neq Y''_{N_\delta}, \ldots, Y'_0 \neq Y''_0) < \delta$.

Now, we will construct two copies $(Y^*_n)_{n\le 0}$ and $(Y^{**}_n)_{n\le 0}$ of $(Y_n)_{n\le 0}$, independent up to n_0 and such that $\left\{\begin{smallmatrix} (Y^*_n)_{n_0\le n\le 0} \\ (Y^{**}_n)_{n_0\le n\le 0}\end{smallmatrix}\right.$ is a copy of $\left\{\begin{smallmatrix} (Y'_n)_{n_0\le n\le 0} \\ (Y''_n)_{n_0\le n\le 0}\end{smallmatrix}\right.$ One can assume that

$$(Y'_{n+1}, Y''_{n+1}) = f_n(Y'_{n_0}, Y''_{n_0}, \ldots, Y'_n, Y''_n, U_{n+1})$$

where $U_{n+1} \perp\!\!\!\perp (Y'_{n_0}, Y''_{n_0}, U_{n_0}, \ldots, Y'_n, Y''_n, U_n)$ is $\mathcal{U}(0,1)$. Then, given two independent copies $(Y^*_n)_{n\le n_0}$ and $(Y^{**}_n)_{n\le n_0}$ of $(Y_n)_{n\le n_0}$, we recursively set

$$(Y^*_{n+1}, Y^{**}_{n+1}) = f_n(Y^*_{n_0}, Y^{**}_{n_0}, \ldots, Y^*_n, Y^{**}_n, \bar{U}_{n+1})$$

where $(\bar{U}_{n_0+1}, \ldots, \bar{U}_0) \perp\!\!\!\perp (Y^*_n, Y^{**}_n)_{n\le n_0}$ is a vector of independent $\mathcal{U}(0,1)$ random variables. In this way, for $n \ge n_0$, one has

$$\mathcal{L}(Y^*_{n+1} \mid Y^*_m, Y^{**}_m; m \le n) = \mathcal{L}(Y^*_{n+1} \mid Y^*_{n_0}, Y^{**}_{n_0}, \ldots, Y^*_n, Y^{**}_n),$$

and $\mathcal{L}(Y^*_{n+1} \mid Y^*_{n_0}, Y^{**}_{n_0}, \ldots, Y^*_n, Y^{**}_n) = \mathcal{L}(Y^*_{n+1} \mid Y^*_n)$ because $\left\{\begin{smallmatrix} (Y^*_n)_{n_0\le n\le 0} \\ (Y^{**}_n)_{n_0\le n\le 0}\end{smallmatrix}\right.$ is a copy of $\left\{\begin{smallmatrix} (Y'_n)_{n_0\le n\le 0} \\ (Y''_n)_{n_0\le n\le 0}\end{smallmatrix}\right.$. Thus the two filtrations generated by $(Y^*_n)_{n\le 0}$ and $(Y^{**}_n)_{n\le 0}$ are jointly immersed, and that shows that $(Y_M, \ldots, Y_0)$ satisfies the I-cosiness criterion. □

Now we prove the second point of Theorem 7.1.

Proof of 2 in Theorem 7.1. Assume $\mathcal{F}$ is I-cosy. Take an integer $M \le 0$. To prove the claim, it suffices to show that $(Y_M, \ldots, Y_0)$ satisfies the I-cosiness criterion with respect to $\mathcal{G}$ (see [3]).

Let $\delta > 0$ and take the integer $N_\delta \le M$ provided by the joinability assumption (Definition 3).

The random variable X_{N_δ} satisfies the I-cosiness criterion with respect to $\mathcal{F}$. Thus, one has two jointly immersed copies $(\mathcal{F}'_n)_{n\le N_\delta}$ and $(\mathcal{F}''_n)_{n\le N_\delta}$ of the filtration $(\mathcal{F}_n)_{n\le N_\delta}$, independent up to an integer $n_0 \le N_\delta$ and such that $\Pr(X'_{N_\delta} \neq X''_{N_\delta}) < \delta$.

Thanks to the Markov property, $\begin{cases} (X'_n)_{n_0 \le n \le N_\delta} \\ (X''_n)_{n_0 \le n \le N_\delta} \end{cases}$ is a synchronous joining of two copies of $(X_n)_{n_0 \le n \le N_\delta}$.

Now, one has the joining $\begin{cases} (X_n)_{n_0 \le n \le N_\delta} \\ (Y_n)_{n_0 \le n \le N_\delta} \end{cases}$ provided by the joinability assumption.

By Lemma 8, one has, on an enlargement of the probability space of $\begin{cases} (X'_n)_{n_0 \le n \le N_\delta} \\ (X''_n)_{n_0 \le n \le N_\delta} \end{cases}$, two copies $\begin{cases} (X'_n)_{n_0 \le n \le N_\delta} \\ (Y'_n)_{n_0 \le n \le N_\delta} \end{cases}$ and $\begin{cases} (X''_n)_{n_0 \le n \le N_\delta} \\ (Y''_n)_{n_0 \le n \le N_\delta} \end{cases}$ of $\begin{cases} (X_n)_{n_0 \le n \le N_\delta} \\ (Y_n)_{n_0 \le n \le N_\delta} \end{cases}$ such that $\begin{cases} (Y'_n)_{n_0 \le n \le N_\delta} \\ (Y''_n)_{n_0 \le n \le N_\delta} \end{cases}$ is a synchronous joining of $(Y_n)_{n_0 \le n \le N_\delta}$ and $Y'_{n_0} \perp\!\!\!\perp Y''_{n_0}$. Clearly, $\Pr(Y'_{N_\delta} \neq Y''_{N_\delta}) < 3\delta$. The result follows from Lemma 10. $\square$

7.4 Proof of Proposition 3

Proposition 3 follows from Proposition 4, Theorem 7.1 and the following lemma.

Lemma 11. *The Markov processes* $(W_n, V_n, \epsilon_n)_{n \le 0}$ *and* $(\widetilde{W}_n, \widetilde{V}_n, \widetilde{\epsilon}_n)_{n \le 0}$ *are closely joinable.*

Proof. We firstly construct a joining of $(V_n, \epsilon_n)_{n_0 \le n \le 0}$ and $(\widetilde{V}_n, \widetilde{\epsilon}_n)_{n_0 \le n \le 0}$.

For $i \in \{0, 1, 3\}$,

$$\Pr(V_n = 1, \epsilon_n = i) = \frac{3^{|n|} - 1}{3^{|n|+1}} < \frac{1}{3}.$$

Therefore, one can construct a joining of $(V_{n_0}, \epsilon_{n_0})$ and $(\widetilde{V}_{n_0}, \widetilde{\epsilon}_{n_0})$ such that

$$\{V_{n_0} = 1, \epsilon_{n_0} = i\} \subset \{\widetilde{V}_{n_0} = 1, \widetilde{\epsilon}_{n_0} = i\}$$

for $i \in \{0, 1, 3\}$.

Now, take some independent random variables U_{n_0+1}, ..., U_{-1} having the $\mathcal{U}(0,1)$ distribution, and such that the vector $(U_{n_0+1}, \ldots, U_{-1})$ is independent of $(V_{n_0}, \epsilon_{n_0})$ and $(\widetilde{V}_{n_0}, \widetilde{\epsilon}_{n_0})$.

We recursively construct $(V_{n_0+1}, \epsilon_{n_0+1})$, ..., (V_{-1}, ϵ_{-1}) and $(\widetilde{V}_{n_0+1}, \widetilde{\epsilon}_{n_0+1})$, ..., $(\widetilde{V}_{-1}, \widetilde{\epsilon}_{-1})$ as follows.

Once the construction is done up to time n, we set $\epsilon_{n+1} = f_n(V_n, U_{n+1})$ and $\widetilde{\epsilon}_{n+1} = g_n(U_{n+1})$ where the functions f_n and g_n are defined as follows. We simply set $f_n(2, u) = 0$. For $i \in \{0, 1, 3\}$, the function $f_n(1, \cdot)$ is such that $f_n(1, u) = i$ for $u \in J_i$ where J_i is the set having Lebesgue measure $|J_i| = \frac{h_{n+1}}{h_n} < \frac{1}{3}$. We take a set $J'_i \supset J_i$ such that $|J'_i| = \frac{1}{3}$ and set $g_n(u) = i$ for $u \in J'_i$. Thus, on the event $\{V_n = 1\}$, one has $\widetilde{\epsilon}_{n+1} = i$ if $\epsilon_{n+1} = i$.

With this joining,

$$\Pr\Big((V_{n_0+1}, \epsilon_{n_0+1}) = (\widetilde{V}_{n_0+1}, \widetilde{\epsilon}_{n_0+1}), \ldots, (V_N, \epsilon_N) = (\widetilde{V}_N, \widetilde{\epsilon}_N) \mid V_{n_0} = 1\Big) \ge \prod_{n=n_0}^{N-1} \frac{h_{n+1}}{h_n}.$$

Now, we construct a synchronous joining $\begin{cases} (W_n, V_n, \epsilon_n)_{n_0+1 \le n \le N} \\ (\widetilde{W}_n, \widetilde{V}_n, \widetilde{\epsilon}_n)_{n_0+1 \le n \le N} \end{cases}$ as follows. To initialize the joining, we write $W_{n_0} = f(V_{n_0}, U)$ where U is a $\mathcal{U}(0,1)$ random variable independent of V_{n_0}, and we set $\widetilde{W}_{n_0} = f(\widetilde{V}_{n_0}, U) = f(1, U)$. Then we construct $\begin{cases} (W_n, V_n, \epsilon_n)_{n_0+1 \le n \le N} \\ (\widetilde{W}_n, \widetilde{V}_n, \widetilde{\epsilon}_n)_{n_0+1 \le n \le N} \end{cases}$ with the joining $\begin{cases} (V_n, \epsilon_n)_{n_0+1 \le n \le N} \\ (\widetilde{V}_n, \widetilde{\epsilon}_n)_{n_0+1 \le n \le N} \end{cases}$ we previously constructed. In this way,

$$\begin{aligned} &\Pr\Big((W_{n_0+1}, V_{n_0+1}, \epsilon_{n_0+1}) = (\widetilde{W}_{n_0+1}, \widetilde{V}_{n_0+1}, \widetilde{\epsilon}_{n_0+1}), \ldots, (W_N, V_N, \epsilon_N) = (\widetilde{W}_N, \widetilde{V}_N, \widetilde{\epsilon}_N) \mid V_{n_0} = 1\Big) \\ &= \Pr\Big((V_{n_0+1}, \epsilon_{n_0+1}) = (\widetilde{V}_{n_0+1}, \widetilde{\epsilon}_{n_0+1}), \ldots, (V_N, \epsilon_N) = (\widetilde{V}_N, \widetilde{\epsilon}_N) \mid V_{n_0} = 1\Big) \end{aligned}$$

The lemma follows because the product $\prod_{n=n_0}^{N-1} \frac{h_{n+1}}{h_n}$ is divergent as $n_0 \to -\infty$ and $\Pr(V_{n_0} = 1) \to 1$. $\square$

References

[1] Ceillier, G.: The filtration of the split-words process. Probability Theory and Related Fields 153, Issue 1–2, 269–292 (2012).

[2] Janvresse, E., Laurent, S., de la Rue, T.: Standardness of monotonic Markov filtrations.

[3] Laurent, S.: On standardness and I-cosiness. Séminaire de Probabilités XLIII, Springer Lecture Notes in Mathematics 2006, 127–186 (2010).

[4] Laurent, S.: On Vershikian and I-cosy random variables and filtrations. Teoriya Veroyatnostei i ee Primeneniya 55 (2010), 104–132. Also published in: Theory Probab. Appl. 55 (2011), 54–76.

[5] Laurent, S.: Vershik's Intermediate Level Standardness Criterion and the Scale of an Automorphism. Séminaire de Probabilités XLV, Springer Lecture Notes in Mathematics 2078, 123–139 (2013).

[6] Laurent, S.: Uniform entropy scalings of filtrations. `https://hal.archives-ouvertes.fr/hal-01006337`

[7] Vershik, A.M.: Four definitions of the scale of an automorphism. Funktsional'nyi Analiz i Ego Prilozheniya, 7:3, 1–17 (1973). English translation: Functional Analysis and Its Applications, 7:3, 169–181 (1973)

[8] Vershik, A.M.: The theory of decreasing sequences of measurable partitions (in Russian). Algebra i Analiz, 6:4, 1–68 (1994). English translation: St. Petersburg Mathematical Journal, 6:4, 705–761 (1995)