



**HAL**  
open science

# When $T$ is an irrational rotation, $[T, Id]$ and $[T, T^{-1}]$ are Bernoulli: explicit isomorphisms

Christophe Leuridan

► **To cite this version:**

Christophe Leuridan. When  $T$  is an irrational rotation,  $[T, Id]$  and  $[T, T^{-1}]$  are Bernoulli: explicit isomorphisms. 2023. hal-03933493v1

**HAL Id: hal-03933493**

**<https://hal.science/hal-03933493v1>**

Preprint submitted on 10 Jan 2023 (v1), last revised 25 Mar 2023 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# When $T$ is an irrational rotation, $[T, \text{Id}]$ and $[T, T^{-1}]$ are Bernoulli: explicit isomorphisms.

Christophe Leuridan

January 10, 2023

## Abstract

Let  $\theta$  be an irrational real number. The map  $T_\theta : y \mapsto (y + \theta) \bmod 1$  from the unit interval  $\mathbf{I} = [0, 1[$  (endowed with the Lebesgue measure) to itself is ergodic. In 2002, Rudolph and Hoffman showed in [7] the measure-preserving map  $[T_\theta, \text{Id}]$  is isomorphic to a one-sided dyadic Bernoulli shift. Their proof is not constructive. A few years before, Parry [12] had provided an explicit isomorphism under the assumption that  $\theta$  is extremely well approached by the rational numbers, namely

$$\inf_{q \geq 1} q^4 4^{q^2} \text{dist}(\theta, q^{-1}\mathbf{Z}) = 0.$$

Whether the explicit map considered by Parry is an isomorphism or not in the general case was still an open question. In [10] we relaxed Parry's condition into

$$\inf_{q \geq 1} q^4 \text{dist}(\theta, q^{-1}\mathbf{Z}) = 0.$$

In the present paper, we remove the condition by showing that the explicit map considered by Parry is always an isomorphism. With a few adaptations, the same method works with  $[T, T^{-1}]$ .

MSC Classification : 37A05, 60J05.

Keywords : irrational rotations, skew products, Bernoulli shifts, dyadic filtrations, head-and-tail filtrations, constructive Markov chains, coupling from the past.

## 1 Introduction

In this section, we present general properties of Meilijson's skew products  $[T, \text{Id}]$  and  $[T, T^{-1}]$  before focusing on the case where  $T$  is an irrational rotation.

### 1.1 Meilijson's skew products $[T, \text{Id}]$ and $[T, T^{-1}]$

Let  $X = \{0, 1\}^{\mathbf{Z}_+}$ , endowed with the product sigma-field  $\mathcal{X}$  and the product measure

$$\mu := \bigotimes_{n \in \mathbf{Z}_+} (\delta_0 + \delta_1)/2.$$

The shift operator  $S : X \rightarrow X$ , defined by

$$S(x)(i) := x(i + 1) \text{ for every } x \in X \text{ and } i \in \mathbf{Z}_+,$$

preserves the measure  $\mu$ .

Given any automorphism  $T$  of a probability space  $(Y, \mathcal{Y}, \nu)$  one defines the map  $[T, \text{Id}]$  from  $X \times Y$  to itself by

$$[T, \text{Id}](x, y) := (S(x), T^{x(0)}(y)).$$

On the first component, the map  $[T, \text{Id}]$  simply applies the shift  $S$ . If we totally ignore the first component, the map  $[T, \text{Id}]$  applies  $T$  or  $\text{Id}$  at random on the second one. The map  $[T, T^{-1}]$  can be defined with the same formula, just by replacing  $\{0, 1\}^{\mathbb{Z}^+}$  with  $\{-1, 1\}^{\mathbb{Z}^+}$ .

One checks that the transformations  $[T, \text{Id}]$  and  $[T, T^{-1}]$  preserve the measure  $\mu \otimes \nu$  and are two-to-one: if  $(x, y)$  is chosen randomly according to the distribution  $\mu \otimes \nu$ , and if one knows  $[T, \text{Id}](x, y)$  (or  $[T, T^{-1}]$ ), the only information missing to recover  $(x, y)$  is the value of  $x(0)$ , which is uniformly distributed on  $\{0, 1\}$  and independent of  $[T, \text{Id}](x, y)$  (of  $[T, T^{-1}]$ ).

Meilijson's theorem [11] ensures that the natural extension of  $[T, \text{Id}]$  is a  $K$ -automorphism whenever  $T$  is ergodic and the natural extension of  $[T, T^{-1}]$  is a  $K$ -automorphism whenever  $T$  is totally ergodic (i.e. all positive powers of  $T$  are ergodic). Actually, the ergodicity of  $T$  only is sufficient (and necessary) to guarantee that the endomorphism  $[T, \text{Id}]$  is exact, so its natural extension is a  $K$ -automorphism. And the ergodicity of  $T^2$  only is sufficient (and necessary) to guarantee that the endomorphism  $[T, T^{-1}]$  is exact, so its natural extension is a  $K$ -automorphism. See theorem 1 in [9].

Hence, when  $T$  (respectively  $T^2$ ) is ergodic, a natural question arises: is the endomorphism  $[T, \text{Id}]$  (respectively  $[T, T^{-1}]$ ) isomorphic to the Bernoulli shift  $S$ ? The answer depends on the automorphism  $T$  considered.

- An adaptation of techniques and ideas introduced by Vershik [14, 15], Heicklen and Hoffman [5] shows that when  $T$  is a two-sided Bernoulli shift, and more generally when  $T$  has positive entropy, the endomorphism  $[T, T^{-1}]$  cannot be Bernoulli since the standardness - a weaker property - fails. This argument can be adapted to the endomorphism  $[T, \text{Id}]$ . More details are given in [9].
- In 2000, Hoffman constructed in [6] a zero-entropy transformation  $T$  such that the  $[T, \text{Id}]$  endomorphism is nonstandard. The proof is detailed in [9].
- In 2002, Rudolph and Hoffman showed in [7] that when  $T$  is an irrational rotation, the endomorphism  $[T, \text{Id}]$  is isomorphic to the dyadic Bernoulli shift  $S$ . Their proof is not constructive.

Actually, non-trivial examples of explicit isomorphisms between measure preserving maps are quite rare. Yet, an independent generating partition providing an explicit isomorphism between  $[T, \text{Id}]$  and the dyadic one-sided Bernoulli shift was given in 1996 by Parry in [12] when the rotation  $T$  is extremely well approximated by rational ones.

## 1.2 Parry's partition

From now on, we fix an irrational number  $\theta$ . We call  $\mathbf{I} = [0, 1[$  the unit interval,  $\nu$  the uniform distribution on  $\mathbf{I}$  and  $T_\theta : \mathbf{I} \rightarrow \mathbf{I}$  the translation of  $\theta$  modulo 1. This transformation preserves the measure  $\nu$  and can be viewed as an irrational rotation on

the circle  $\mathbf{R}/\mathbf{Z}$ . However, we choose to work with the unit interval  $\mathbf{I} = [0, 1[$  to avoid ambiguity in the definition of the sub-intervals.

Since the transformation  $T_\theta$  depends only on the equivalence class of  $\theta$  in  $\mathbf{R}/\mathbf{Z}$ , we may - and we shall - assume that  $0 < \theta < 1$ . Hence for every  $y \in \mathbf{I}$ ,

$$T_\theta(y) = y + \theta - \mathbf{1}_{[1-\theta, 1[}(y).$$

The map  $T_\theta$  is bijective with inverse  $T_{1-\theta} = T_{-\theta}$ , given by

$$T_{-\theta}(y) = y - \theta + \mathbf{1}_{[0, \theta[}(y).$$

Since  $T_{1-\theta} = T_{-\theta}$  is isomorphic to  $T_\theta$ , we may - and we shall - assume that  $0 < \theta < 1/2$ .

In [12], Parry introduced the partition  $\alpha_\theta = \{A_0^\theta, A_1^\theta\}$  on  $X \times \mathbf{I}$  defined by

$$A_0^\theta := \{x \in X : x(0) = 0\} \times [0, \theta[ \cup \{x \in X : x(0) = 1\} \times [0, 1 - \theta[,$$

$$A_1^\theta := \{x \in X : x(0) = 0\} \times [\theta, 1[ \cup \{x \in X : x(0) = 1\} \times [1 - \theta, 1[.$$

One checks that the partitions  $([T_\theta, \text{Id}]^{-n} \alpha_\theta)_{n \geq 0}$  are independent.

For every  $(x, y) \in X \times \mathbf{I}$ , denote by  $\alpha_\theta(x, y) = \mathbf{1}_{A_1^\theta}(x, y)$  the index of the only block containing  $(x, y)$  in the partition  $\alpha_\theta$ . By construction, the ‘ $\alpha_\theta$ -name’ map  $\Phi_\theta : X \times \mathbf{I} \rightarrow X$  defined by

$$\Phi_\theta(x, y) := ((\alpha_\theta \circ S_\theta^n)(x, y))_{n \geq 0}.$$

is also a factor map which sends the dynamical system  $(X \times \mathbf{I}, \mathcal{X} \otimes \mathcal{B}(\mathbf{I}), \mu \otimes \nu, [T_\theta, \text{Id}])$  on the Bernoulli shift  $(X, \mathcal{X}, \mu, S)$ .

Under the assumption

$$\inf_{q \geq 1} q^4 4^{q^2} \text{dist}(\theta, q^{-1}\mathbf{Z}) = 0,$$

Parry shows that for  $\mu \otimes \nu$ -almost every  $(x, y) \in X \times \mathbf{I}$ , the knowledge of  $\Phi_\theta(x, y)$  is sufficient to recover  $(x, y)$ , so the factor map  $\Phi_\theta$  is an isomorphism.

Our purpose is to remove Parry’s assumption.

### 1.3 Probabilistic reformulation

Since the transformation  $T_\theta$  preserves the uniform measure on  $[0, 1[$ , we can consider a stationary Markov chain  $((\xi_n, Y_n))_{n \in \mathbf{Z}}$  on  $\{0, 1\} \times [0, 1[$  such that for every  $n \in \mathbf{Z}$ ,

- $(\xi_n, Y_n)$  is uniform on  $\{0, 1\} \times [0, 1[$ ;
- $\xi_n$  is independent of  $\mathcal{F}_{n-1}^{\xi, Y} := \sigma((\xi_k, Y_k)_{k \leq n-1})$ ;
- $Y_n = T_{-\theta}^{\xi_n}(Y_{n-1}) = (Y_{n-1} - \xi_n \theta) \bmod 1$ .

For every real number  $r$ , denote by  $\bar{r} = r + \mathbf{Z}$  the equivalence class of  $r$  modulo  $\mathbf{Z}$ . The process  $(\bar{Y}_n)_{n \in \mathbf{Z}}$  is as a random walk on  $\mathbf{R}/\mathbf{Z}$ . The steps  $(\bar{Y}_n - \bar{Y}_{n-1})_{n \in \mathbf{Z}} = (-\xi_n \theta)$  are uniformly distributed on  $\{\bar{0}, -\bar{\theta}\}$ . Since each random variable  $\xi_n$  can be recovered from the knowledge of  $Y_n$  and  $Y_{n-1}$ , the natural filtration  $(\mathcal{F}_n^{\xi, Y})_{n \in \mathbf{Z}}$  of the Markov chain  $((\xi_n, Y_n))_{n \in \mathbf{Z}}$  coincides with the natural filtration  $(\mathcal{F}_n^Y)_{n \in \mathbf{Z}}$  of the process  $(Y_n)_{n \in \mathbf{Z}}$ .

The Markov chain  $((\xi_n, Y_n))_{n \in \mathbf{Z}}$  thus defined is closely related to the transformation  $[T_\theta, \text{Id}]$  since for every  $n \in \mathbf{Z}$ ,

$$[T_\theta, \text{Id}]((\xi_{n-k})_{k \geq 0}, Y_n) = ((\xi_{n-(k+1)})_{k \geq 0}, T_\theta^{\xi_n}(Y_n)) = ((\xi_{n-1-k})_{k \geq 0}, Y_{n-1}).$$

Knowing the sequence  $(\xi_n)_{n \in \mathbf{Z}}$  only is not sufficient to recover the positions  $(Y_n)_{n \in \mathbf{Z}}$ . Indeed, one can check that  $Y_0$  is independent of the whole sequence  $(\xi_n)_{n \in \mathbf{Z}}$ .

Reformulating Parry's method, we introduce the sequence  $(\eta_n)_{n \in \mathbf{Z}}$  defined by

$$\begin{aligned} \eta_n &:= (\xi_n + \mathbf{1}_{[\theta, 1[}(Y_{n-1})) \bmod 2. \\ &= \mathbf{1}_{[\xi_n=0; Y_{n-1} \geq \theta]} + \mathbf{1}_{[\xi_n=1; Y_{n-1} < \theta]}. \end{aligned}$$

By construction,  $\eta_n$  is an  $\mathcal{F}_n^{\xi, Y}$ -measurable Bernoulli random variable; and  $\eta_n$  is uniform and independent of  $\mathcal{F}_{n-1}^{\xi, Y}$  since

$$\mathbf{E}[\eta_n | \mathcal{F}_{n-1}^{\xi, Y}] = \frac{1}{2} \mathbf{1}_{[Y_{n-1} \geq \theta]} + \frac{1}{2} \mathbf{1}_{[Y_{n-1} < \theta]} = \frac{1}{2}.$$

Moreover, since  $Y_n = T_{-\theta}^{\xi_n}(Y_{n-1})$ , we have also

$$\begin{aligned} \eta_n &= \mathbf{1}_{[\xi_n=0; Y_n \geq \theta]} + \mathbf{1}_{[\xi_n=1; Y_n \geq 1-\theta]} \\ &= \mathbf{1}_{A_1^\theta}((\xi_{n-k})_{k \geq 0}, Y_n) \\ &= \mathbf{1}_{A_1^\theta}([T_\theta, \text{Id}]^{-n}((\xi_{-k})_{k \geq 0}, Y_0)) \text{ if } n \leq 0. \end{aligned}$$

Hence, the  $\alpha_\theta$ -name of  $((\xi_{-i})_{i \geq 0}, Y_0)$  is the sequence  $(\eta_{-i})_{i \geq 0}$ .

Thus, proving that the partition  $\alpha_\theta$  is generating is equivalent to prove that almost surely, the random variable  $Y_0$  is completely determined by the knowledge of  $(\eta_k)_{k \leq 0}$ . By stationarity, it means that the filtration  $(\mathcal{F}_n^{\xi, Y})_{n \in \mathbf{Z}} = (\mathcal{F}_n^Y)_{n \in \mathbf{Z}}$  is generated by the sequence  $(\eta_n)_{n \in \mathbf{Z}}$ .

Actually, we show that this property holds without any additional assumption on  $\theta$ .

**Theorem 1** *For every  $n \in \mathbf{Z}$ , the random variable  $Y_n$  can be almost surely recovered from the knowledge of  $(\eta_k)_{k \leq n}$ . Hence the filtration  $(\mathcal{F}_n^{\xi, Y})_{n \in \mathbf{Z}} = (\mathcal{F}_n^Y)_{n \in \mathbf{Z}}$  is generated by the sequence  $(\eta_n)_{n \in \mathbf{Z}}$  of independent uniform Bernoulli random variables. Equivalently, Parry's partition  $\alpha_\theta$  is an independent generator of  $[T_\theta, \text{Id}]$ .*

We refer the reader to subsection 5 for the similar statement about  $[T_\theta, T_\theta^{-1}]$ .

At the moment, let us explain the interest why the sequence  $(\eta_n)_{n \in \mathbf{Z}}$  gives more information than the sequence  $(\xi_n)_{n \in \mathbf{Z}}$ . Define two maps  $f_0$  and  $f_1$  from the interval  $\mathbf{I} = [0, 1[$  to itself by

$$f_0(y) := y - \theta \mathbf{1}_{[\theta, 1[}(y), \quad \text{and } f_1(y) := y + (1 - \theta) \mathbf{1}_{[0, \theta]}(y).$$

The union of the graphs of  $f_0$  and  $f_1$  is also the union of the graphs of  $\text{Id}$  and  $T_{-\theta}$  (see figure 1). Therefore, applying at random  $f_0$  or  $f_1$  has the same effect as applying at random  $\text{Id}$  or  $T_{-\theta}$ . Actually, a direct computation distinguishing four cases, according that  $\xi_n$  equals 0 or 1, and that  $Y_{n-1} < \theta$  or  $Y_{n-1} \geq \theta$  yields

$$Y_n = f_{\eta_n}(Y_{n-1}).$$

The great difference with the initial recursion relation  $Y_n = T_{-\theta}^{\xi_n}(Y_{n-1})$  is that the maps  $f_0$  and  $f_1$  are not injective, unlike  $T_{-\theta}^0 = \text{Id}$  and  $T_{-\theta}^1 = T_{-\theta}$ .

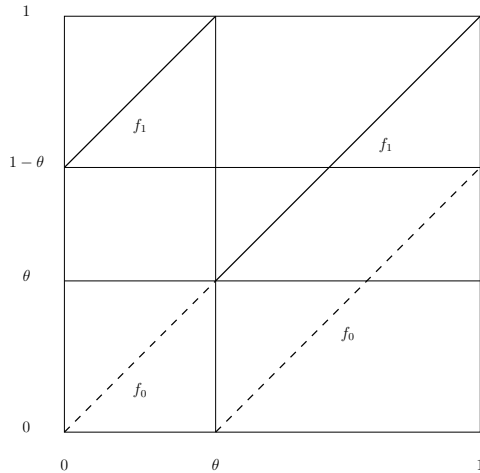


Figure 1: Graphs of  $f_0$  and  $f_1$

## 1.4 Intermediate and related results

For every  $n \geq 1$ , denote by  $F_n$  the random map from  $\mathbf{I} = [0, 1[$  to itself defined by  $F_n := f_{\eta_n} \circ \dots \circ f_{\eta_1}$ . By convention,  $F_0$  is the identity map from  $\mathbf{I}$  to  $\mathbf{I}$ .

The sequence  $(x_n)_{n \geq 0}$  defined by  $x_n := n\theta - \lfloor n\theta \rfloor$  plays an important role in our study. Indeed, a recursion yields the following statement.

**Lemma 2** *Let  $n \geq 0$ . The points  $(x_i)_{0 \leq i \leq n}$  split  $\mathbf{I}$  into  $n + 1$  semi-open intervals and, on each of those intervals, the random map  $F_n$  coincides with  $T_{-k\theta}$  for some integer  $k \in [0, n]$ .*

We would like to say that when  $n$  is large, the random map  $F_n$  is close to a random constant map, so the random variable  $Y_n = F_n(Y_0)$  is almost a function of  $(\eta_1, \dots, \eta_n)$  and has a weak dependence on  $Y_0$ . The next result seems to be a good step in this direction.

**Proposition 3** *The Lebesgue measure of the range  $R_n := F_n(\mathbf{I})$  goes to 0 as  $n$  goes to infinity.*

When proving Proposition 3, we will provide an explicit upper bound of  $|R_n|$  involving the continued fraction expansion of  $\theta$ .

Yet, getting Proposition 3 is far from proving Theorem 1. Although the range  $R_n$  can be written as the union of at most  $n + 1$  intervals (and actually much less than  $n + 1$ ), these intervals may be spread out in the unit interval, and the evolution is difficult to study. The only simple observation is that the Lebesgue measure can only decrease. Indeed, to deduce  $R_n$  from  $R_{n-1}$ , we split  $R_{n-1}$  at  $\theta$ , and we apply translation  $T_{-\theta}$  to  $R_{n-1} \cap [\theta, 1[$  or to  $R_{n-1} \cap [0, \theta[$ , according that  $\eta_n = 0$  or  $\eta_n = 1$ .

To prove Theorem 1, we use a coupling argument: given  $(y, y') \in \mathbf{I}^2$ , we study the processes  $((F_n(y), F_n(y'))_{n \geq 0}$ . These processes are Markov chains on  $\mathbf{I}^2$  whose transition probabilities are given by

$$K((y, y'), \cdot) = \frac{1}{2} (\delta_{(f_0(y), f_0(y'))} + \delta_{(f_1(y), f_1(y'))}).$$

See figure 2 in Subsection 2.2.

Set  $B_\theta = [0, \theta[ \times ]\theta, 1[ \cup ]\theta, 1[ \times ]0, \theta[$ . The key step in the proof of Theorem 1 is to show that the set  $B_\theta$  is visited more and more rarely.

**Proposition 4** *For every  $(y, y') \in \mathbf{I}^2$ ,*

$$\frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}_{B_\theta}(F_k(y), F_k(y')) \rightarrow 0 \text{ almost surely as } n \rightarrow +\infty.$$

For every real number  $\delta$ , we denote by  $\|\delta\|$  the shortest distance from  $\delta$  to  $\mathbf{Z}$ :

$$\|\delta\| := \text{dist}(\delta, \mathbf{Z}) = \min(\delta - \lfloor \delta \rfloor, 1 - \delta + \lfloor \delta \rfloor)$$

By construction,  $0 \leq \|\delta\| \leq 1/2$ .

The next result is similar and can be proved with the help of proposition 4.

**Proposition 5** *For every  $(y, y') \in \mathbf{I}^2$ ,*

$$\frac{1}{n} \sum_{k=0}^{n-1} \|F_k(y') - F_k(y)\| \rightarrow 0 \text{ in probability as } n \rightarrow +\infty.$$

Yet, Theorem 1 follows directly from proposition 4: see subsection 1.5 below.

## 1.5 Deriving Theorem 1 from Proposition 4

Let us explain the final arguments once Proposition 4 is established.

First, we prove that if  $\pi$  is an invariant probability with regard to kernel  $K$  then  $\pi$  is carried by  $D$ , the diagonal of  $\mathbf{I}^2$ . Indeed, by stationarity, one has for every  $n \geq 1$

$$\int_{\mathbf{I}^2} \mathbf{1}_{B_\theta}((y, y')) d\pi(y, y') = \int_{\mathbf{I}^2} \mathbf{E} \left[ \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}_{B_\theta}((F_k(y), F_k(y'))) \right] d\pi(y, y').$$

In the right-hand side, the fonction integrated with regard to  $\pi$  takes values in  $[0, 1]$  and goes to 0 everywhere as  $n$  goes to infinity. Hence Lebesgue dominated convergence applies. Passing to the limit, we get  $\pi(B_\theta) = 0$ . Since all future orbits under the action of  $T_{-\theta}$  are dense in  $\mathbf{I}$ , every Markov chain with transition kernel  $K$  starting from anywhere in  $D^c$  reaches the set  $B_\theta$  with positive probability (precise bounds on the number of steps will be given in section 2). Thus  $\pi(D^c) = 0$ .

Next, by enlarging if necessary the probability space on which the Markov chain  $Y := (Y_n)_{n \in \mathbf{Z}}$  is defined, we can construct a copy  $Y' := (Y'_n)_{n \in \mathbf{Z}}$  such that  $Y$  and  $Y'$  are i.i.d. conditionally on  $\eta := (\eta_n)_{n \in \mathbf{Z}}$ . Since the process  $((Y_n, Y'_n))_{n \in \mathbf{Z}}$  thus obtained is a stationary Markov chain with transition kernel  $K$ , it lives almost surely on  $D$ , so the processes  $Y = Y'$  almost surely. But

$$\mathcal{L}((Y, Y')|\eta) = \mathcal{L}(Y|\eta) \otimes \mathcal{L}(Y|\eta),$$

hence the conditional law  $\mathcal{L}(Y|\eta)$  is almost surely a Dirac mass, so  $Y$  is almost surely a deterministic function of the process  $\eta$ .

Remind that for every  $n \in \mathbf{Z}$ , the random variable  $\eta_n$  is independent of  $\mathcal{F}_{n-1}^{\xi, Y}$ . Given  $n \in \mathbf{Z}$ , the sequence  $(\eta_m)_{k \geq n+1}$  is independent of  $\mathcal{F}_n^{\xi, Y}$  and therefore of the sub- $\sigma$ -field  $\sigma(Y_n) \vee \sigma((\eta_m)_{k \leq n})$ . Hence one has almost surely,

$$\delta_{Y_n} = \mathcal{L}(Y_n | (\eta_k)_{k \in \mathbf{Z}}) = \mathcal{L}(Y_n | (\eta_k)_{k \leq n}),$$

so the random variable  $Y_n$  is almost surely a measurable function of the process  $(\eta_m)_{k \leq n}$  only. Theorem 1 follows.

## 1.6 Plan of the paper

In section 2, we introduce the main tools and Lemmas involved to prove Propositions 3, 4 and 5: continued fractions expansions, three gaps Theorem and bounds for hitting and return times under the action of  $T_{-\theta}$ .

In section 3, we prove Proposition 3.

In section 4, we prove Propositions 4 and 5.

In section 5, we explain how the proof must be adapted to get a constructive proof that  $[T_\theta, T_\theta^{-1}]$  is Bernoulli.

In section 6, we view still open related questions.

## 2 Tools

In this section, we present the techniques and preliminary results involved to prove Propositions 3, 4 and 5. Our results heavily rely on the three gaps Theorem.

### 2.1 Three gaps theorem and consequences

We consider again the sequence  $(x_n)_{n \geq 0}$  defined by  $x_n := n\theta - \lfloor n\theta \rfloor$ . The well-known three gaps theorem states that for every  $n \geq 1$ , the intervals defined by the subdivision  $(x_k)_{0 \leq k \leq n-1}$  have at most three different lengths. We shall use a very precise statement of the three gaps theorem, involving the continued fraction expansion  $\theta = [0; a_1, a_2, \dots]$ . Let us recall some classical facts. A more complete exposition can be found in [8].

Set  $a_0 = \lfloor \theta \rfloor = 0$  and define two sequences  $(p_k)_{k \geq -2}$  and  $(q_k)_{k \geq -2}$  of integers by  $p_{-2} = 0$ ,  $q_{-2} = 1$ ,  $p_{-1} = 1$ ,  $q_{-1} = 0$ , and for every  $k \geq 0$ ,

$$p_k = a_k p_{k-1} + p_{k-2}, \quad q_k = a_k q_{k-1} + q_{k-2}.$$

Moreover,  $p_k$  and  $q_k$  are relatively prime since  $p_k q_{k-1} - q_k p_{k-1} = (-1)^{k-1}$ .

The fractions  $(p_k/q_k)_{k \geq 0}$  are the convergents of  $\theta$ . The following inequalities hold

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \dots < \theta < \dots < \frac{p_3}{q_3} < \frac{p_1}{q_1}.$$

In particular, the difference  $\theta - p_k/q_k$  has the same sign as  $(-1)^k$  and

$$\frac{a_{k+2}}{q_k q_{k+2}} = \left| \frac{p_{k+2}}{q_{k+2}} - \frac{p_k}{q_k} \right| < \left| \theta - \frac{p_k}{q_k} \right| < \left| \frac{p_{k+1}}{q_{k+1}} - \frac{p_k}{q_k} \right| = \frac{1}{q_k q_{k+1}}.$$

Hence the sequence  $(p_k/q_k)_{n \geq 0}$  converges to  $\theta$ .



For every  $k \geq -1$ , set  $\ell_k = (-1)^k(q_k\theta - p_k) = |q_k\theta - p_k|$ . In particular,  $\ell_{-1} = 1$ ,  $\ell_0 = \theta$  and  $\ell_1 = 1 - \lfloor 1/\theta \rfloor \theta \leq 1 - \theta$ . Since for every  $k \geq 0$ ,

$$\frac{1}{q_{k+2}} \leq \frac{a_{k+2}}{q_{k+2}} < |q_k\theta - p_k| < \frac{1}{q_{k+1}},$$

the sequence  $(\ell_k)_{k \geq -1}$  is decreasing. For every  $k \geq 1$ , we have  $\ell_k = -a_k\ell_{k-1} + \ell_{k-2}$ , so  $a_k = \lfloor \ell_{k-2}/\ell_{k-1} \rfloor$ .

We can now give a precise statement of the three gaps theorem. The formulation below is close to the formulation given by Alessandri and Berthé [3]

**Theorem 6** *Let  $n \geq 1$ . Set  $n = bq_k + q_{k-1} + r$  where  $k \geq 0$ ,  $b \in [1, a_{k+1}]$  and  $r \in [0, q_k]$  are integers. Then the points  $(x_i)_{0 \leq i \leq n-1}$  split  $\mathbf{I}$  into  $n$  intervals:*

- $n - q_k = (b - 1)q_k + q_{k-1} + r$  intervals with length  $\ell_k$ ;
- $q_k - r$  intervals with length  $\ell_{k-1} - (b - 1)\ell_k$ ;
- $r$  intervals with length  $\ell_{k-1} - b\ell_k$ .

**Remark 7** *Each positive integer  $n$  can be written as above. The decomposition is unique if we require that  $r \leq q_k - 1$ . In this case,*

- $k$  is the largest non-negative integer such that  $q_k + q_{k-1} \leq n$ , so  $q_k + q_{k-1} \leq n < q_{k+1} + q_k = (a_{k+1} + 1)q_k + q_{k-1}$ ,
- $b$  is the integer part of  $(n - q_{k-1})/q_k$ , so  $1 \leq b \leq a_{k+1}$ ,
- $r = (n - q_{k-1}) - bq_k$ , so  $0 \leq r \leq q_{k-1} - 1$ .

Under the assumptions of theorem 6, we have  $\ell_{k-1} - (b - 1)\ell_k = (\ell_{k-1} - b\ell_k) + \ell_k$ . Moreover,  $\ell_{k-1} - b\ell_k < \ell_k$  if and only if  $b = a_k$ . From these observations, we deduce the following consequences.

**Corollary 8** *Fix  $n \geq 1$ , and consider the  $n$  sub-intervals of  $\mathbf{I}$  defined by the subdivision  $(x_i)_{0 \leq i \leq n-1}$ .*

- *If  $bq_k + q_{k-1} \leq n < (b + 1)q_k + q_{k-1}$  with  $k \geq 0$  and  $b \in [1, a_{k+1}]$ , the greatest length is  $\ell_{k-1} - (b - 1)\ell_k$ .*
- *If  $q_k < n \leq q_{k+1}$ , the smallest length is  $\min\{||i\theta|| : i \in [1, n - 1]\} = \ell_k$ .*

We can now answer precisely these two questions: given a semi-open subinterval  $J$  of  $\mathbf{I}$ , what is the maximal number of iterations of  $T_{-\theta}$  to reach  $J$  from anywhere in  $\mathbf{I}$ ? what is the minimal number of iterations of  $T_{-\theta}$  necessary to return in  $J$  from anywhere in  $J$ ?

**Corollary 9** *Let  $J$  be a semi-open subinterval of  $\mathbf{I}$ , modulo 1, namely  $J = [\alpha, \beta[$  with  $0 \leq \alpha < \beta \leq 1$  (with length  $|J| = \beta - \alpha$ ), or  $J = [\alpha, 1[ \cup [0, \beta[$  with  $0 \leq \beta < \alpha \leq 1$  (with length  $|J| = 1 + \beta - \alpha$ ).*

- Let  $k \geq 0$  be the least integer such that  $\ell_k + \ell_{k+1} \leq |J|$ , and  $b \in [1, a_{k+1}]$  be the least integer such that  $\ell_{k-1} - (b-1)\ell_k \leq |J|$ . This integer is well-defined since  $\ell_{k-1} - (a_{k+1}-1)\ell_k = \ell_k + \ell_{k+1}$ . Then  $bq_k + q_{k-1} - 1$  iterations of  $T_{-\theta}$  are sufficient to reach  $J$  from anywhere in  $\mathbf{I}$ , and this bound is optimal.
- Let  $k \geq 0$  be the least integer such that  $\ell_k \leq |J|$ . For every  $y \in J$ , the first integer  $m(y) \geq 1$  such that  $T_{-\theta}^{m(y)}(y) \in J$  is at least equal to  $q_k$ .

Proof. We are searching for the least non-negative integer  $m$  such that

$$\bigcup_{k=0}^m T_{-\theta}^k(J) = \mathbf{I} \quad (1)$$

By rotation, one may assume that  $J = [0, |J|$ . In this case, for every  $k \geq 0$ ,

$$T_{-\theta}^k(J) = \begin{cases} [x_k, x_k + |J| & \text{if } x_k + |J| \leq 1, \\ [x_k, 1] \cup [0, x_k + |J| - 1] & \text{if } x_k + |J| > 1. \end{cases}$$

Therefore equality 1 holds if and only if the greatest length of the sub-intervals of  $\mathbf{I}$  defined by the subdivision  $(x_k)_{0 \leq k \leq m}$  is at most  $|J|$ . Thus, first item follows from corollary 8.

Next, given  $y \in J$ , we are searching for the least integer  $m(y) \geq 1$  such that  $T_{-\theta}^{m(y)}(y) \in J$ , which implies

$$\min_{0 \leq i < j \leq m(y)} \|x_j - x_i\| \leq \|m(y)\theta\| = \|y - T^{m(y)}(y)\| < |J|.$$

Denoting by  $k \geq 0$  the least integer such that  $\ell_k \leq |J|$ , we derive  $m(y) \geq q_k$  by corollary 8.  $\square$

## 2.2 A Markov chain on $\mathbf{I}^2$

In this section, we study the Markov chains on  $\mathbf{I}^2$  with kernel  $K$ . Our purpose is to give upper bounds on the entrance time in the set  $B_\theta = [0, \theta] \times [\theta, 1] \cup [\theta, 1] \times [0, \theta]$ .

The transition probabilities can be also written

$$K((y, y'), \cdot) = \frac{1}{2}(\delta_{(T_{-\theta}(y), y')} + \delta_{(y, T_{-\theta}(y'))}) \text{ if } (y, y') \in B_\theta,$$

$$K((y, y'), \cdot) = \frac{1}{2}(\delta_{(y, y')} + \delta_{(T_{-\theta}(y), T_{-\theta}(y'))}) \text{ if } (y, y') \in B_\theta^c.$$

Note that the transition probabilities from any  $(y, y') \in B_\theta^c$  preserve the difference  $y' - y$  modulo 1. See figure 2.

Since  $D$  is absorbing and has no intersection with  $B_\theta$ , the set  $B_\theta$  cannot be reached from  $D$ . But the set  $B_\theta$  can be reached from any  $(y, y') \in D^c$ . We set

$$t_{y, y'} := \inf\{n \geq 0 : (T_{-\theta}^n(y), T_{-\theta}^n(y')) \in B_\theta\}$$

$$T_{y, y'} = \inf\{n \geq 0 : (F_n(y), F_n(y')) \in B_\theta\}.$$

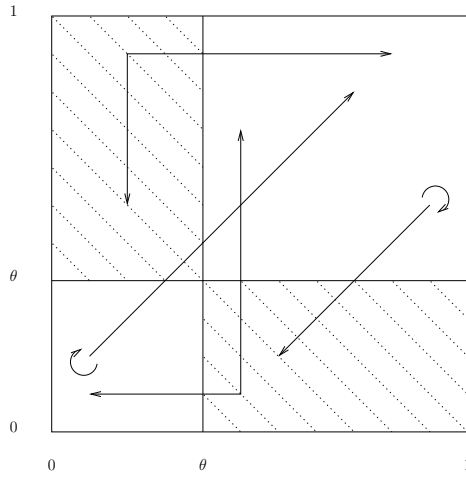


Figure 2: Probability transitions from different points  $(y, y')$ . The probability associated to each arrow is  $1/2$ . The set  $B_\theta$  is the union of the shaded rectangles. When  $(y, y') \in B_\theta$ , only one coordinate moves, so the transitions are parallel to the axis. When  $(y, y') \in B_\theta^c$ , the same translation  $\text{Id}$  or  $T_{-\theta}$  is applied to both coordinates, so the transitions are parallel to the diagonal.

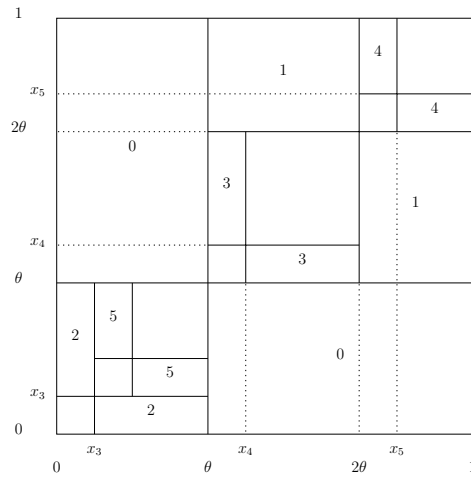


Figure 3: Entrance time in  $B_\theta$  from any point under the action of the map  $(y, y') \mapsto (T_{-\theta}(y), T_{-\theta}(y'))$ . Points far from the diagonal have low entrance time.

Note that we have always  $T_{y,y'} \geq t_{y,y'}$ , and that  $t_{y,y'}$  and  $T_{y,y'}$  are null whenever  $(y, y')$  is in  $B_\theta$ . Moreover,

$$T_{y,y'} + 1 = \inf\{n \geq 0 : F_n(y') - F_n(y) \not\equiv y' - y \pmod{1}\}.$$

Informally, we want to show that the times  $t_{y,y'}$  and  $T_{y,y'}$  are small when  $\|y' - y\|$  is far from 0, and are ‘often’ big when  $\|y' - y\|$  is close from 0. The restriction ‘often’ cannot be removed since  $t_{\theta-\varepsilon_1, \theta+\varepsilon_2} = 0$  for every  $\varepsilon_1 \in ]0, \theta[$  and  $\varepsilon_2 \in [0, 1 - \theta[$ . See figure 3.

Before, we prove a useful Lemma.

**Lemma 10** *Let  $y \in \mathbf{R}$ . There exists a unique sequence  $(\zeta_n^y)_{n \geq 1}$  of independent uniform Bernoulli random variables such that for every integer  $n \geq 0$ ,*

$$\forall n \geq 0, \quad F_{n+1}(y) = T_{-\theta}^{\zeta_n^y}(F_n(y)).$$

*Proof.* The uniqueness holds since  $T_{-\theta}(y') \neq y'$  for every  $y' \in \mathbf{I}$ . To prove the existence, we construct a sequence  $(\zeta_n^y)_{n \geq 1}$  of random variables by setting for every  $n \geq 1$ ,

$$\zeta_n^y := (\eta_n + \mathbf{1}_{[\theta, 1[}(F_{n-1}(y))) \pmod{2} = \mathbf{1}_{[\eta_n=0 ; F_{n-1}(y) \geq \theta]} + \mathbf{1}_{[\eta_n=1 ; F_{n-1}(y) < \theta]}.$$

For every  $n \geq 0$ , set  $\mathcal{F}_n := \sigma(\eta_1, \dots, \eta_n)$ , with the convention  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ .

By construction,  $\zeta_n^y$  is an  $\mathcal{F}_n$ -measurable Bernoulli random variable, which is uniform and independent of  $\mathcal{F}_{n-1}$  since

$$\mathbf{E}[\zeta_n^y | \mathcal{F}_{n-1}] = \frac{1}{2} \mathbf{1}_{[F_{n-1}(y) \geq \theta]} + \frac{1}{2} \mathbf{1}_{[F_{n-1}(y) < \theta]} = \frac{1}{2}.$$

Moreover, a computation distinguishing two cases whether  $F_{n-1}(y) < \theta$  or  $F_{n-1}(y) \geq \theta$  yields

$$F_n(y) = f_{\eta_n}(F_{n-1}(y)) = T_{-\theta}^{\zeta_n^y} F_{n-1}(y).$$

The proof is complete. □

**Lemma 11** *Assume that  $0 < \theta < 1/2$ . Let  $y$  and  $y'$  be two different points in  $\mathbf{I}$ .*

*The time  $t_{y,y'}$  is finite. More precisely, let  $k \geq 0$  be the least integer such that  $\ell_k + \ell_{k+1} \leq \|y' - y\|$ , and  $b \in [1, a_{k+1}]$  the least integer such that  $\ell_{k-1} - (b-1)\ell_k \leq \|y' - y\|$ . Then  $t_{y,y'} \leq bq_k + q_{k-1} - 2$  if  $\|y' - y\| \leq \theta$  and  $t_{y,y'} \leq q_1 - 1$  otherwise.*

*The random variable  $T_{y,y'}$  is binomial negative, with parameters  $t_{y,y'}$  and  $1/2$ . In particular,  $\mathbf{E}[T_{y,y'}] = 2t_{y,y'}$ .*

*Proof.* Let  $\delta = \|y' - y\|$ . Since  $y$  and  $y'$  play the same role, we may - and we do - assume that  $y' = T_\delta(y)$ .

We use the sequence  $(\zeta_n^y)_{n \geq 1}$  introduced in Lemma 10 and abbreviate the notation into  $(\zeta_n)_{n \geq 1}$ .

For every integer  $n \geq 0$ ,

$$\forall n \geq 0, \quad F_{n+1}(y) = T_{-\theta}^{\zeta_n}(F_n(y)).$$

For every integer  $n \in [0, T_{y,y'} - 1]$ , the two points  $F_n(y')$  and  $F_n(y)$  belong to the same interval  $[0, \theta[$  or  $[\theta, 1[$ , so  $F_{n+1}(y') = T_{-\theta}^{\zeta_n}(F_n(y')) \equiv F_n(y') - \zeta_n \theta \pmod{1}$ . By recursion, we get:

$$\forall n \geq 0, \quad F_n(y) = T_{-\theta}^{\zeta_1 + \dots + \zeta_n}(y)$$

and

$$\forall n \in [0, T_{y,y'}], \quad F_n(y') = T_{-\theta}^{\zeta_1 + \dots + \zeta_n}(y') = T_\delta(F_n(y)).$$

Hence

$$\begin{aligned} T_{y,y'} &= \inf\{n \geq 0 : (F_n(y), T_\delta(F_n(y))) \in B_\theta\} \\ &= \inf\{n \geq 0 : (T_{-\theta}^{\zeta_1 + \dots + \zeta_n}(y), T_\delta(T_{-\theta}^{\zeta_1 + \dots + \zeta_n}(y))) \in B_\theta\} \\ &= \inf\{n \geq 0 : \zeta_1 + \dots + \zeta_n = t_{y,y'}\} \end{aligned}$$

Since  $(\zeta_n)_{n \geq 1}$  is a sequence of independent uniform Bernoulli random variables, the random variable  $T_{y,y'}$  is negative binomial, with parameters  $t_{y,y'}$  and  $1/2$ .

If  $\delta \leq \theta$ , then for every  $x \in \mathbf{I}$ ,

$$(x, T_\delta(x)) \in B_\theta \iff x \in [\theta - \delta, \theta[ \cup [1 - \delta, 1[.$$

Thus  $t_{y,y'} = \inf\{n \geq 0 : T_{-\theta}^n(y) \in [\theta - \delta, \theta[ \cup [1 - \delta, 1[ \}$ . By corollary 9, at most  $bq_k + q_{k-1} - 1$  iterations of the map  $T_{-\theta}$  are sufficient to reach each one of the intervals  $[\theta - \delta, \theta[$  or  $[1 - \delta, 1[$  from  $y$ . But these two intervals are disjoint since  $0 < \theta - \delta < \theta < 1/2 < 1 - \delta < 1$ . Hence  $t_{y,y'} \leq bq_k + q_{k-1} - 2$ .

If  $\delta > \theta$ , then for every  $x \in \mathbf{I}$ ,

$$(x, T_\delta(x)) \in B_\theta \iff x \in [0, \theta[ \cup [1 - \delta, 1 - \delta + \theta[.$$

Thus  $T = \inf\{n \geq 0 : F_n(y) \in [0, \theta[ \cup [1 - \delta, 1 - \delta + \theta[ \}$ . By corollary 9 applied with  $k = 1$  and  $b = 1$ , at most  $q_1$  iterations of the map  $T_{-\theta}$  are sufficient to reach each one of the intervals  $[0, \theta[$  or  $[1 - \delta, 1 - \delta + \theta[$  from  $y$ . But these two intervals are disjoint since  $0 < \theta < 1/2 < 1 - \delta < 1 - \delta + \theta < 1$ . Hence  $t_{y,y'} \leq q_1 - 1$ .

The proof is complete.  $\square$

**Corollary 12** *Let  $y$  and  $y'$  be in  $\mathbf{I}$  and  $k \geq 1$  be an integer such that  $\ell_{k-1} \leq \|y' - y\|$ . Then  $t_{y,y'} \leq q_k + q_{k-1} - 2$ .*

*Proof.* By Lemma 11, if  $\theta \leq \|y' - y\|$ , then  $t_{y,y'} \leq q_1 - 1 \leq q_k + q_{k-1} - 2$ .

From now on, we focus on the case where  $\|y' - y\| \leq \theta$ . Observe that

$$\ell_k + \ell_{k+1} \leq a_{k+1}\ell_k + \ell_{k+1} = \ell_{k-1} \leq \|y' - y\|.$$

If  $\|y' - y\| < \ell_{k-1} + \ell_k$ , then Lemma 11 applied to the integers  $k$  and  $b = 1$  yields  $t_{y,y'} \leq q_k + q_{k-1} - 2$ .

Otherwise,  $k \geq 2$  since  $\ell_{k-1} < \ell_{k-1} + \ell_k \leq \|y' - y\| \leq \theta = \ell_0$ . Furthermore, the least integer  $k' \geq 1$  such that  $\ell_{k'} + \ell_{k'+1} \leq \|y' - y\|$  is at most  $k - 1$  and Lemma 11 applies to the integers  $k' \geq 1$  and some  $b \in [1, a_{k'+1}]$ , so

$$t_{y,y'} \leq bq_{k'} + q_{k'-1} - 2 \leq a_{k'+1}q_{k'} + q_{k'-1} - 2 = q_{k'+1} - 2 \leq q_k - 2.$$

Hence, in all cases, we get  $t_{y,y'} \leq q_k + q_{k-1} - 2$ .  $\square$

### 2.3 A time-changed symmetric random walk

For every real number  $x$ , denote by  $\bar{x} = x + \mathbf{Z}$  its equivalence class in  $\mathbf{R}/\mathbf{Z}$ . Then Lemma 11 and the strong Markov property show that the process  $(\overline{F_n(y') - F_n(y)})_{n \geq 0}$  is a time-changed symmetric random walk on  $\mathbf{R}/\mathbf{Z}$  with steps  $\pm\bar{\theta}$ .

**Lemma 13** *For every  $n \geq 0$ , set  $\mathcal{F}_n = \sigma(\eta_1, \dots, \eta_n)$ , with the convention  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ . Consider the non-decreasing process  $(A_n)_{n \geq 0}$  defined by*

$$A_n := \sum_{k=0}^{n-1} \mathbf{1}_{B_\theta}(F_k(y), F_k(y')), \text{ with the convention } A_0 = 0,$$

and set  $A_\infty := \lim_n A_n$ . Consider its inverse  $(N_a)_{a \geq 0}$  defined by ,

$$N_a := \inf\{n \geq 0 : A_n \geq a\}, \text{ with the convention } \inf \emptyset = +\infty.$$

On the event  $[N_a < +\infty] = [a \leq A_\infty]$ , set

$$\Delta N_a := N_{a+1} - N_a \text{ and } W_a := \sum_{b=1}^a (2\zeta_{N_b}^y - 1),$$

with the notation of Lemma 10.

1. For every  $n \geq 0$ ,

$$F_n(y') - F_n(y) \equiv y' - y + \theta W_{A_n} \pmod{1}.$$

2. The random times  $(N_a - 1)_{a \geq 1}$  are stopping times in the filtration  $(\mathcal{F}_n)_{n \geq 0}$ .
3. The process  $(W_a)_{a \geq 0}$  is a simple symmetric random walk on  $\mathbf{Z}$ , with possibly finite life-time  $A_\infty$ . More precisely, for every positive integer  $a \geq 0$ , conditionally on the event  $[A_\infty \geq a]$ , the increments  $W_1 - W_0, \dots, W_a - W_{a-1}$  are independent and uniform on  $\{-1, 1\}$ .
4. One has  $A_\infty = A_{T_{y,y'}^D}$  where  $T_{y,y'}^D := \inf\{n \geq 0 : F_n(y) = F_n(y')\}$ . If  $y' - y \in \mathbf{Z} + \theta\mathbf{Z}$ , then  $A_\infty$  is almost surely finite. Otherwise,  $A_\infty$  is infinite.
5. Assume that  $y' - y \notin \mathbf{Z} + \theta\mathbf{Z}$ . Conditionally on the process  $((F_{N_a}(y), F_{N_a}(y')))_{a \geq 0}$ , the random variables  $(\Delta N_a)_{a \geq 0}$  are independent and each  $\Delta N_a - 1$  is negative binomial with parameters  $t_{F_{N_a}(y), F_{N_a}(y')}$  and  $1/2$ .

*Proof.* We use the sequence  $(\zeta_n)_{n \geq 1} := (\zeta_n^y)_{n \geq 1}$  introduced in Lemma 10, and define  $(\zeta'_n)_{n \geq 1} := (\zeta_n^{y'})_{n \geq 1}$  in the same way. By construction,

$$F_n(y) = f_{\eta_n}(F_{n-1}(y)) \equiv F_{n-1}(y) - \zeta_n \theta \pmod{1}.$$

Then

$$F_n(y') - F_n(y) \equiv F_{n-1}(y') - F_{n-1}(y) + (\zeta_n - \zeta'_n)\theta \pmod{1}.$$

A recursion yields, for every  $n \geq 0$ ,

$$F_n(y') - F_n(y) \equiv y' - y + M_n \theta \pmod{1}, \text{ where } M_n := \sum_{k=1}^n (\zeta_k - \zeta'_k),$$

with the convention  $M_0 = 0$ . For every  $k \geq 1$ ,  $\zeta'_k = 1 - \zeta_k$  if  $(F_{k-1}(y), F_{k-1}(y')) \in B_\theta$  whereas  $\zeta'_k = \zeta_k$  if  $(F_{k-1}(y), F_{k-1}(y')) \in B_\theta^c$ , so

$$\begin{aligned}\zeta_k - \zeta'_k &= \mathbf{1}_{B_\theta}(F_{k-1}(y), F_{k-1}(y')) \times (2\zeta_k - 1) \\ &= (A_k - A_{k-1}) \times (2\zeta_k - 1).\end{aligned}$$

By definition, for every  $a \geq 0$ , on the event  $[N_a < +\infty]$ , the process  $(A_n)_{n \geq 0}$  remains constant and equal to  $a$  during the time interval  $[N_a, N_{a+1} - 1]$ . Hence

$$M_n = \sum_{k=1}^n (M_k - M_{k-1}) = \sum_{b=1}^{A_n} (M_{N_b} - M_{N_b-1}) = \sum_{b=1}^{A_n} (2\zeta_{N_b} - 1) = W_{A_n}.$$

Item 1 follows.

The process  $(A_n)_{n \geq 0}$  defined by

$$A_n = \sum_{k=0}^{n-1} \mathbf{1}_{B_\theta}(F_k(y), F_k(y'))$$

is  $(\mathcal{F}_n)_{n \geq 0}$ -predictable: the random variable  $A_0 = 0$  is constant and for every  $n \geq 1$ , the random variable  $A_n$  is  $\mathcal{F}_{n-1}$ -measurable.

For every  $a \geq 0$ , the random variable  $N_a$  is a stopping time in the filtration  $(\mathcal{F}_n)_{n \geq 0}$ . If  $a \geq 1$ , the random variable  $N_a - 1$  is still a stopping time since for every  $n \geq 0$ ,

$$[N_a - 1 \leq n] = [N_a \leq n + 1] = [A_{n+1} \geq a] \in \mathcal{F}_n.$$

Item 2 follows.

Let  $a \geq 0$ . The event  $[N_a < +\infty] = [N_a - 1 < +\infty]$  belongs to  $\mathcal{F}_{N_a-1}$ . On this event, the sequence  $(\tilde{\eta}_n)_{n \geq 0} := (\eta_{N_a+n})_{n \geq 0}$  is an i.i.d. sequence of uniform Bernoulli random variables and is independent of  $\mathcal{F}_{N_a-1}$ . Therefore,

- the Bernoulli random variable  $\zeta_{N_a} = (\eta_{N_a} + \mathbf{1}_{[0, \theta[}(F_{N_a-1}(y))) \bmod 2$  is uniform, independent of  $\mathcal{F}_{N_a-1}$  and  $\mathcal{F}_{N_a}$ -measurable. Thus the random variable  $W_a - W_{a-1} = 2\zeta_{N_a} - 1$  is uniform on  $\{-1, 1\}$ , independent of  $\mathcal{F}_{N_a-1}$  and  $\mathcal{F}_{N_a}$ -measurable. Item 3 follows.

- the random variable  $(Y, Y') := (F_{N_a}(y), F_{N_a}(y'))$  is  $\mathcal{F}_{N_a}$ -measurable.
- the process  $(\tilde{F}_n)_{n \geq 0}$  defined by  $\tilde{F}_n := f_{\tilde{\eta}_n} \circ \dots \circ f_{\tilde{\eta}_1}$  is independent of  $\mathcal{F}_{N_a}$  and has the same law as  $(F_n)_{n \geq 0}$ .

Observe that with obvious notations,

$$\Delta N_a - 1 = (N_{a+1} - 1) - N_a = \tilde{T}_{Y, Y'} := \inf \{n \geq 0 : (\tilde{F}_n(Y), \tilde{F}_n(Y')) \in B_\theta\}.$$

Hence, conditionally on  $\mathcal{F}_{N_a}$  and on the event  $[N_a < +\infty] \cap [Y \neq Y']$ , the random variable  $\Delta N_a - 1$  is almost surely finite and negative binomial with parameters  $t_{Y, Y'}$  and  $1/2$ . Item 5 follows. Moreover, for every  $a \geq 0$ , we have almost surely

$$N_{a+1} < +\infty \iff (N_a < +\infty \text{ and } F_{N_a}(y) \neq F_{N_a}(y')).$$

Hence, almost surely,

$$A_\infty = \inf \{a \geq 0 : N_{a+1} = +\infty\} = \inf \{a \geq 0 : N_a < +\infty ; F_{N_a}(y) = F_{N_a}(y')\}.$$

During each time interval  $[N_a, N_{a+1} - 1]$ , the process  $(F_n(y') - F_n(y))$  remains constant modulo 1 and the process  $(A_n)_{n \geq 0}$  remains constant and equal to  $a$ . Thus

$$A_\infty = A_{T_{y,y'}^D}, \text{ where } T_{y,y'}^D := \inf\{n \geq 0 : F_n(y) = F_n(y')\}.$$

Since for every  $n \geq 0$ ,  $F_n(y') - F_n(y) \equiv y' - y + \theta W_{A_n} \pmod{1}$ , the difference  $F_n(y') - F_n(y)$  remains in the coset  $y' - y + \theta \mathbf{Z} + \mathbf{Z}$ . Therefore, the time  $T_{y,y'}^D$  cannot be finite unless  $y' - y \in \theta \mathbf{Z} + \mathbf{Z}$ . Conversely, if  $y' - y \in \theta \mathbf{Z} + \mathbf{Z}$ , then  $T_{y,y'}^D$  is almost surely finite by recurrence of the symmetric simple random walk on  $\mathbf{Z}$ . Item 4 follows.  $\square$

By density of the subgroup  $\theta \mathbf{Z} + \mathbf{Z}$  in  $\mathbf{R}$  and by recurrence of the symmetric simple random walk on  $\mathbf{Z}$ , we derive the following consequence.

**Corollary 14** *Let  $y$  and  $y'$  be two points in  $\mathbf{I}$  such that  $y' - y \notin \mathbf{Z} + \theta \mathbf{Z}$ . Then, almost surely,*

$$\liminf_{n \rightarrow +\infty} \|F_n(y') - F_n(y)\| = 0 \text{ and } \limsup_{n \rightarrow +\infty} \|F_n(y') - F_n(y)\| = 1/2.$$

*Proof.* We use the notations and the results of Lemma 13. The recurrence of the symmetric simple random walk on  $\mathbf{Z}$  and the density of the subgroup  $\theta \mathbf{Z} + \mathbf{Z}$  in  $\mathbf{R}$  yields

$$\inf_{n \geq 0} \|F_n(y') - F_n(y)\| = \inf_{a \geq 0} \|y' - y + \theta W_a\| = \inf_{w \in \mathbf{Z}} \|y' - y + \theta w\| = 0 \text{ a.s..}$$

Since the quantities  $\|F_n(y') - F_n(y)\|$  are all positive, the greatest lower bound of the sequence  $(\|F_n(y') - F_n(y)\|)_{n \geq 0}$  is not achieved and is therefore a lower limit. We argue in the same way with the least upper bound.  $\square$

To prove that the Cesàro averages of the sequence  $(\|F_n(y') - F_n(y)\|)_{n \geq 0}$  tend to 0, we observe that the time-change in the process  $(W_{A_n})_{n \geq 0}$  slows down the random walk much more when the difference is  $\|y' - y\|$  is small. Indeed, Corollary 12 shows that if  $\|y' - y\|$  is not close to 0, then the process  $((F_n(y), F_n(y'))_{n \geq 0})$  reaches  $B_\theta$  quickly, and the difference  $F_n(y') - F_n(y)$  changes immediately after.

Conversely, if  $\|y' - y\|$  is small, we hope that it will take a long time for the process  $((F_n(y), F_n(y'))_{n \geq n_0})$  to reach  $B_\theta$ . During this time, the difference  $F_n(y') - F_n(y)$  remains unchanged. Yet, we may be unlucky, so the times  $t_{y,y'}$  and  $T_{y,y'}$  may be small. Fortunately, the distance  $\|F_n(y') - F_n(y)\|$  is close to  $\theta$  at time  $N_1 = T_{y,y'} + 1$  so it will change soon. With probability 1/2, we will get  $W_2 = 0$ , so  $\|F_{N_2}(y') - F_{N_2}(y)\| = \|y' - y\|$  will be small, and this time the distance  $\|F_n(y') - F_n(y)\|$  will remain unchanged during a long time. The role of the next Lemma is to provide a lower bound of this time.

**Lemma 15** *Let  $y$  and  $y'$  be two distinct points in  $\mathbf{I}$  such that  $(y, y') \in [0, 1 - \theta]^2 \cup [\theta, 1]^2$  and  $\|y' - y\| < \ell_2$ . Let  $k \geq 3$  be the integer such that  $\ell_k \leq \|y' - y\| < \ell_{k-1}$ . Keep the notations of Lemma 13. Then*

$$\text{On the event } [W_2 = 0], \quad \Delta N_2 - 1 \geq t_{F_{N_2}(y), F_{N_2}(y')} \geq q_k + 1 - q_2 - q_1.$$

Before proving Lemma 15, it is important to note that all probability transitions  $K((y, y'), \cdot)$  give full measure to the set  $[0, 1 - \theta]^2 \cup [\theta, 1]^2$  (see figure 4). Hence, the additional hypothesis  $(y, y') \in [0, 1 - \theta]^2 \cup [\theta, 1]^2$  will not be a true restriction. Actually, this assumption could be removed by changing the lower bound into  $q_k - q_2 - q_1$ .



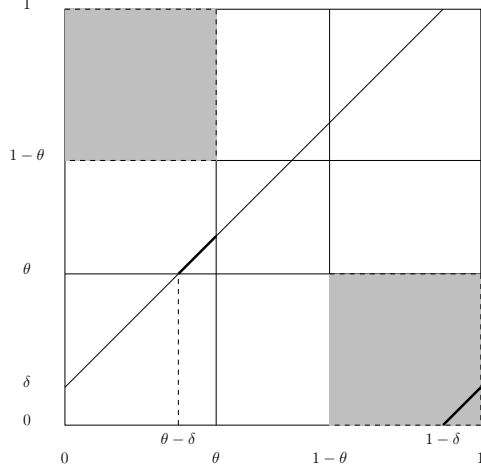


Figure 4: The squares  $[0, \theta[ \times [1 - \theta, 1[$  and  $[1 - \theta, 1[ \times [0, \theta[$  cannot be reached from anywhere via the kernel  $K$ . If  $0 < \delta < \theta$ , then  $B_\theta \cap \{(y, T_\delta(y)); y \in \mathbf{I}\}$  is the union of two ‘semi-open segments’, defined by  $\theta - \delta \leq y < \theta$  and  $1 - \delta \leq y < 1$ . The second semi-open segment cannot be reached from anywhere via the kernel  $K$  and is the image of the first one by the map  $(y, y') \mapsto (T_{-\theta}(y), T_{-\theta}(y'))$ .

Proof. The inequality  $\Delta N_2 - 1 \geq t_{F_{N_2}(y), F_{N_2}(y')} \geq 0$  follows from item 5 of Lemma 13.

The assumption  $(y, y') \in [0, 1 - \theta]^2 \cup [\theta, 1]^2$  ensures that  $(F_n(y), F_n(y'))$  belongs to  $[0, 1 - \theta]^2 \cup [\theta, 1]^2$  not only for every  $n \geq 1$ , but also when  $n = 0$ .

Set  $\delta = \|y' - y\|$ . By symmetry, we may assume that  $y' - y \equiv \delta \pmod{1}$ , i.e.  $y' = T_\delta(y)$ . Hence, for every  $a \geq 0$ ,  $F_{N_{a+1}-1}(y') - F_{N_{a+1}-1}(y) \equiv \delta + \theta W_a \pmod{1}$ .

On the event  $[W_2 = 0]$ , we have

$$F_{N_3-1}(y') - F_{N_3-1}(y) \equiv F_{N_1-1}(y') - F_{N_1-1}(y) \equiv \delta \pmod{1}.$$

Since the points  $(F_{N_1-1}(y), F_{N_1-1}(y'))$  and  $(F_{N_3-1}(y), F_{N_3-1}(y'))$  belong at the same time to  $B_\theta$  and to  $[0, 1 - \theta]^2 \cup [\theta, 1]^2$ , we deduce that  $F_{N_1-1}(y)$  and  $F_{N_3-1}(y)$  belong to  $[\theta - \delta, \theta[$ . But  $F_{N_3-1}(y)$  is obtained from  $F_{N_1-1}(y)$  by applying the transformation  $T_{-\theta}$  exactly  $\zeta_{N_1}^y + t_{F_{N_1}(y), F_{N_1}(y')} + \zeta_{N_2}^y + t_{F_{N_2}(y), F_{N_2}(y')}$  times.

On the event  $[W_2 = 0]$ , we have  $\zeta_{N_1}^y + \zeta_{N_2}^y = 1$ , so corollary 9 yields

$$t_{F_{N_1}(y), F_{N_1}(y')} + t_{F_{N_2}(y), F_{N_2}(y')} + 1 \geq q_k.$$

But  $F_{N_1}(y') - F_{N_1}(y) \equiv \delta \pm \theta \pmod{1}$ , so

$$\|F_{N_1}(y') - F_{N_1}(y)\| \geq \theta - \delta > \ell_0 - \ell_k \geq \ell_1.$$

Corollary 12 yields  $t_{F_{N_1}(y), F_{N_1}(y')} \leq q_2 + q_1 - 2$ . Thus

$$t_{F_{N_2}(y), F_{N_2}(y')} + q_2 + q_1 - 1 \geq q_k.$$

The desired inequality follows.  $\square$

## 2.4 Martingales and symmetric simple random walk on $\mathbb{Z}$

In this subsection, we establish three more results that which help us in the proof of Proposition 4. The first of them is a kind of law of large numbers.

**Lemma 16** *Let  $(\mathcal{G}_n)_{n \geq 0}$  be a filtration of  $(\Omega, \mathcal{A}, \mathbf{P})$  and  $(Z_n)_{n \geq 1}$  be a sequence of real random variables, bounded in  $L^2(\mathbf{P})$ , and adapted to the filtration  $(\mathcal{G}_n)_{n \geq 1}$ . Then*

$$\frac{1}{n} \sum_{k=1}^n (Z_k - \mathbf{E}[Z_k | \mathcal{G}_{n-1}]) \rightarrow 0 \text{ almost surely as } n \rightarrow +\infty.$$

Proof. For every  $n \geq 0$ , let

$$M_n = \sum_{k=1}^n \frac{1}{k} (Z_k - \mathbf{E}[Z_k | \mathcal{G}_{n-1}]).$$

By construction, the sequence  $(M_n)_{n \geq 0}$  is a square-integrable martingale in  $(\mathcal{G})_{n \geq 0}$ . The increments  $M_k - M_{k-1}$  are pairwise orthogonal in  $L^2(\mathbf{P})$  and

$$\sum_{k \geq 1} \|M_k - M_{k-1}\|_2^2 = \sum_{k \geq 1} \frac{1}{k^2} \|Z_k - \mathbf{E}[Z_k | \mathcal{G}_{n-1}]\|_2^2 \leq \sum_{k \geq 1} \frac{1}{k^2} \|Z_k\|_2^2 < +\infty.$$

Hence the martingale  $(M_n)_{n \geq 0}$  is bounded in  $L^2(\mathbf{P})$  and it converges almost surely and in  $L^2(\mathbf{P})$  to some real random variable  $M_\infty$ .

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^n (Z_k - \mathbf{E}[Z_k | \mathcal{G}_{n-1}]) &= \frac{1}{n} \sum_{k=1}^n k(M_k - M_{k-1}) \\ &= \frac{1}{n} \left( nM_n + \sum_{k=1}^{n-1} kM_k - M_0 - \sum_{k=1}^{n-1} (k+1)M_k \right) \\ &= M_n - \frac{1}{n} \sum_{k=1}^{n-1} M_k. \end{aligned}$$

The conclusion follows from Cesàro Lemma.  $\square$

The second result is a refinement of gambler's ruin Theorem [13].

**Lemma 17** *Let  $(W_n)_{n \geq 0}$  be a symmetric simple random walk on  $\mathbf{Z}$ , i.e.  $W_0 = 0$  and  $(W_n - W_{n-1})_{n \geq 1}$  is an i.i.d. sequence of Rademacher random variables. Fix two non-negative integers  $a$  and  $b$  and set*

$$\tau_{-a,b} = \inf\{n \geq 0 : W_n = \{-a, b\}\}.$$

Then  $\tau_{-a,b} < +\infty$  almost surely. Moreover,

- $\mathbf{P}[W_{\tau_{-a,b}} = -a] = b/(a+b)$  and  $\mathbf{P}[W_{\tau_{-a,b}} = b] = a/(a+b)$ .
- $\mathbf{E}[\tau_{-a,b}] = ab$ .
- For every  $\lambda \geq 0$ ,  $\mathbf{E}[(\cosh \lambda)^{-\tau_{-a,b}}] = \cosh\left(\frac{-a+b}{2}\lambda\right) / \cosh\left(\frac{a+b}{2}\lambda\right)$ .

Proof. One checks that the processes  $(W_n)_{n \geq 0}$ ,  $(W_n^2 - n)_{n \geq 0}$  and  $(e^{\pm \lambda W_n} (\cosh \lambda)^{-n})_{n \geq 0}$  are martingales, and that the martingales

$$(W_n)_{n \geq 0}, \quad \left( \left( W_n - \frac{b-a}{2} \right)^2 - n \right)_{n \geq 0}, \quad \left( \cosh \left( \lambda \left( W_n - \frac{b-a}{2} \right) \right) (\cosh \lambda)^{-n} \right)_{n \geq 0}$$

are uniformly bounded on the time interval  $[0, \tau_{-a,b}]$ . Hence optional stopping time Theorem applies. Writing that the expectations at time  $\tau_{-a,b}$  are equal to the expectations at time 0 yields items 1, 2, 3.  $\square$

The third result states an equidistribution modulo 1.

**Lemma 18** *Let  $(W_n)_{n \geq 0}$  be a symmetric simple random walk on  $\mathbf{Z}$  and  $\theta$  be any irrational number. Then for every continuous 1-periodic function  $f : \mathbf{R} \rightarrow \mathbf{C}$ ,*

$$\frac{1}{n} \sum_{k=0}^{n-1} f(\theta W_k) \rightarrow \int_0^1 f(x) dx \text{ almost surely as } n \rightarrow +\infty.$$

*Proof.* Since the trigonometric polynomials form a dense subspace of the vector space of all continuous 1-periodic functions, we need only to prove the result when  $f$  is the function  $e_m : x \mapsto e^{i2\pi m x}$  for some  $m \in \mathbf{Z}$ . The case where  $m = 0$  is trivial, so we fix  $m \neq 0$ .

For every  $n \geq 1$ , let

$$S_n := \sum_{k=0}^{n-1} e^{i2\pi m \theta W_k}.$$

Then

$$|S_n|^2 = S_n \overline{S_n} = \sum_{k=0}^{n-1} \sum_{\ell=0}^{n-1} e^{i2\pi m \theta (W_k - W_\ell)}$$

Let  $z = \mathbf{E}[e^{i2\pi m \theta (W_1 - W_0)}] = \cos(2\pi m \theta)$ . Then  $z \in ]-1, 1[$  since  $m \neq 0$  and  $\theta$  is irrational. By independence equidistribution of the increments of the random walk  $W$ , we derive

$$\begin{aligned} \mathbf{E}[|S_n|^2] &= \sum_{k=0}^{n-1} \sum_{\ell=0}^{n-1} z^{|k-\ell|} = n + 2 \sum_{d=1}^{n-1} (n-d) z^d \\ &\leq n \left( 1 + 2 \sum_{d=1}^{n-1} |z|^d \right) \leq n \frac{1+|z|}{1-|z|}. \end{aligned}$$

Thus

$$\sum_{n=1}^{+\infty} \left\| \frac{1}{n^3} S_{n^3} \right\|_2 \leq \sum_{n=1}^{+\infty} \left( \frac{1+|z|}{1-|z|} \right)^{-1/2} n^{-3/2} < +\infty.$$

The series  $\sum_n S_{n^3}/n^3$  is absolutely convergent in  $L^2(\mathbf{P})$ , so it converges almost surely and in  $L^2(\mathbf{P})$ . Therefore,  $S_{n^3}/n^3 \rightarrow 0$  almost surely as  $n \rightarrow +\infty$ . To deduce that  $S_n/n \rightarrow 0$  almost surely as  $n \rightarrow +\infty$ , it suffices to set  $r_n = \lfloor n^{1/3} \rfloor$  and to note that

$$|S_n - S_{r_n^3}| \leq n - r_n^3 \leq 3n^{2/3}(n^{1/3} - r_n) \leq 3n^{2/3}.$$

### 3 Proof of Proposition 3

We begin with the key Lemma. This lemma (and Fubini's Theorem) ensures that when  $n$  is large, the random map  $F_n$  is close to be  $\theta$ -periodic, namely that  $F_n$  and  $F_n \circ T_\theta$  coincide on a set of Lebesgue measure close to 1. View figure 5.

**Lemma 19** *Let  $y \in \mathbf{I}$  and set  $y' = T_\theta(y) = (y + \theta) \bmod 1$ . Given  $k \geq 1$ , we have*

$$\forall n \geq 2(q_k + q_{k-1})q_k^2, \quad \mathbf{P}[F_n(y') \neq F_n(y)] \leq \frac{2}{q_k}.$$

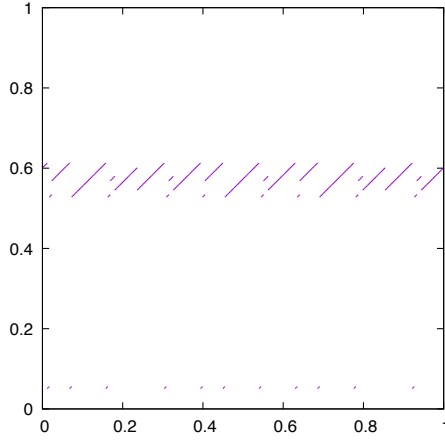


Figure 5: A simulation of  $F_n$ , for  $n = 1000$  and  $\theta = (\sqrt{5} - 1)/2$

Proof. We use the notations and the results of Lemma 13. For every  $n \geq 0$ ,

$$F_n(y') - F_n(y) \equiv \theta(1 + W_{A_n}) \pmod{1},$$

where  $(A_n)_{n \geq 0}$  is a time-change given by

$$A_n = \sum_{k=0}^{n-1} \mathbf{1}_{B_\theta}(F_k(y), F_k(y')).$$

and  $(W_n)_{0 \leq n \leq A_\infty}$  is a simple symmetric random walk on  $\mathbf{Z}$ , with lifetime  $A_\infty$ .

Call  $(N_a)_{a \geq 0}$  the inverse of  $(A_n)_{n \geq 0}$ : for every integer  $a \geq 0$ ,

$$N_a := \inf\{n \geq 0 : A_n \geq a\}.$$

For every  $d \in \mathbf{Z}$ , set

$$\sigma_d := \inf\{t \geq 0 : W_t = d\},$$

$$\tau_d := \inf\{n \geq 0 : W_{A_n} = d\} = \inf\{n \geq 0 : F_n(y') - F_n(y) = (1 + d)\theta\}.$$

Then  $\tau_d = N_{\sigma_d}$ , with the conventions  $\inf \emptyset = +\infty$  and  $N_\infty = \infty$ .

In particular, since  $y' = T_\theta(y)$ , lemma 13 yields  $A_\infty = A_{T_{y,y'}^D}$  where

$$T_{y,y'}^D = \inf\{n \geq 0 : F_n(y) = F_n(y')\} = \inf\{n \geq 0 : W_{A_n} = -1\} = N_{\sigma_{-1}} = \tau_{-1}.$$

By adapting item 5 to the present situation, we see that, conditionally on the process  $Z := ((F_{N_a}(y), F_{N_a}(y')))_{0 \leq a \leq \sigma_{-1}-1}$ , the random variables  $(\Delta N_a)_{0 \leq a \leq \sigma_{-1}}$  are independent and each random variable  $\Delta N_a$  is negative binomial with parameters  $t_{F_{N_a}(y), F_{N_a}(y')}$  and  $1/2$ . Since

$$\min(\tau_{-1}, \tau_{q_k-1}) = \min(N_{\sigma_{-1}}, N_{\sigma_{q_k-1}}) = \sum_{a=0}^{\min(\sigma_{-1}, \sigma_{q_k-1})-1} \Delta N_a,$$

and since  $\min(\tau_{-1}, \tau_{q_k-1})$  is a measurable function of  $Z$ , we derive

$$\mathbf{E}[\min(\tau_{-1}, \tau_{q_k-1}) | Z] = \sum_{a=0}^{\min(\sigma_{-1}, \sigma_{q_k-1})-1} (2t_{F_{N_a}(y), F_{N_a}(y')} + 1).$$

Next, while  $0 \leq n < \min(\tau_0, \tau_{q_k})$ , we have  $1 + W_{A_n} \in [1, q_k - 1]$ , so Corollary 8 yields

$$\|F_n(y') - F_n(y)\| \geq \min\{\|i\theta\| : i \in [1, q_k - 1]\} = \ell_{k-1},$$

and Corollary 12 yields  $t_{F_n(y), F_n(y')} \leq q_k + q_{k-1} - 2$ . Putting things together, we get

$$\mathbf{E}[\min(\tau_{-1}, \tau_{q_k-1}) | Z] \leq 2(q_k + q_{k-1}) \min(\sigma_{-1}, \sigma_{q_k-1})$$

Taking expectations and applying the gambler's ruin theorem (see Lemma 17 or [13], chapter 10, section 14, subsection 4) yields

$$\mathbf{E}[\min(\tau_{-1}, \tau_{q_k-1})] \leq 2(q_k + q_{k-1})q_k.$$

Now, let  $n \geq 1$ . Since the processes  $((F_n(y))_{n \geq 0})$  and  $((F_n(y'))_{n \geq 0})$  coincide from the instant  $\tau_{-1}$ , we have the inclusions

$$[F_n(y') \neq F_n(y)] \subset [n < \tau_{-1}] \subset [n < \min(\tau_{-1}, \tau_{q_k-1})] \cup [\tau_{q_k-1} < \tau_{-1}],$$

so Markov inequality and the gambler's ruin theorem (see Lemma 17 or [13], chapter 10, section 14, subsection 4) yield

$$\begin{aligned} \mathbf{P}[F_n(y') \neq F_n(y)] &\leq \mathbf{P}[n < \min(\tau_{-1}, \tau_{q_k-1})] + \mathbf{P}[\sigma_{q_k-1} < \sigma_{-1}] \\ &\leq \frac{2(q_k + q_{k-1})q_k}{n} + \frac{1}{q_k}. \end{aligned}$$

The statement of Lemma 19 follows. □

We now derive an upper bound for the Lebesgue measure of the range  $F_n(\mathbf{I})$ .

**Corollary 20** *Given  $k \geq 1$ , we have*

$$\forall n \geq 2(q_k + q_{k-1})q_k^2, \quad \mathbf{E}[|F_n(\mathbf{I})|] \leq \frac{2}{q_k}.$$

*Proof.* Let  $E = \{y \in \mathbf{I} : F_n(T_\theta(y)) \neq F_n(y)\}$ . Since  $F_n \circ T_\theta - F_n$  is piecewise constant on  $\mathbf{I}$  and right-continuous, the set  $E$  is a finite union of semi-open intervals. Moreover, by Lemma 19 and Fubini's theorem,  $\mathbf{E}[|E|] \leq 2/q_k$ .

Let us show that  $F_n(\mathbf{I}) \subset F_n(E)$  (the reverse inclusion is trivial). Fix  $y \in E$ . We claim that there exists  $k \geq 0$  such that  $F_n(T_\theta^{k+1}(y)) \neq F_n(T_\theta^k(y))$ . Otherwise, the function  $F_n$  would be constant along the future  $T_\theta$ -orbit of  $y$ , hence constant on  $\mathbf{I}$  by right-continuity and by density. But  $F_n$  cannot be constant since it is a piecewise translation. Hence,  $m(y) := \min\{k \geq 0 : F_n(T_\theta^{k+1}(y)) \neq F_n(T_\theta^k(y))\}$  is well-defined. By construction,  $F(y) = F(T_\theta^{m(y)}(y))$  and  $T_\theta^{m(y)}(y) \in E$ .

Putting things together, we get

$$|F_n(\mathbf{I})| = |F_n(E)| \leq |E| \leq \frac{2}{q_k}$$

since  $F_n$  is a piecewise translation. □

## 4 Proof of Propositions 4 and 5

### 4.1 Proof of Proposition 4

Let  $y, y'$  be in  $\mathbf{I}$  and  $\delta = y' - y$ . We use again the notations and the results contained in Lemma 13.

If  $\delta \in \mathbf{Z} + \theta\mathbf{Z}$ , then almost surely,  $F_n(y') - F_n(y)$  is null for all large enough  $n$ , so the conclusion holds.

We now focus on the difficult case, when  $\delta \notin \mathbf{Z} + \theta\mathbf{Z}$ . First, we prove the following Lemma.

**Lemma 21** *Assume that  $\delta \notin \mathbf{Z} + \theta\mathbf{Z}$ , keep the notations of Lemma 13.*

1. Consider the function  $g$  from  $[0, 1/2]$  to  $[0, +\infty]$  defined by  $g(0) = +\infty$  and

$$g(r) := 1 + (q_k + 1 - q_2 - q_1)_+ \text{ whenever } \ell_k \leq r < \ell_{k-1}.$$

*This function  $g$  is non-increasing on  $[0, 1/2]$  and bounded below by 1. Moreover, the integral of  $g$  on  $[0, 1/2]$  is infinite.*

2. Moreover,

$$\forall a \geq 3, \quad \Delta N_a \geq g(\|\theta W_a\|) \mathbf{1}_{[W_a=W_{a-2}]}$$

*Proof.* The sequence  $(\ell_k)_{k \geq 0}$  is decreasing and the sequence  $(q_k)_{k \geq 0}$  is non-decreasing. The monotonicity of  $g$  and  $h$  and the lower bounds follow. On each interval  $[\ell_k, \ell_{k-1}[$ ,  $g(r) \geq q_k - q_3$ . thus

$$\int_0^{\ell_2} g(r) dr \geq \sum_{k=3}^{+\infty} q_k (\ell_{k-1} - \ell_k) - q_3 \ell_2.$$

But  $q_k \ell_{k-1} \geq 1/2$  for every  $k \geq 0$  and  $\ell_{k-1} \geq \frac{1+\sqrt{5}}{2} \ell_k$  for at least one  $k$  among two consecutive  $k$  (since  $\ell_{k-1} \geq \ell_k + \ell_{k+1}$  for every  $k \geq 0$ ). Item 1 follows.

Now, assume that  $a \geq 3$ . Then  $N_{a-2} \geq a - 2 \geq 1$ , so the point  $(Y, Y') := (F_{N_{a-2}}(y), F_{N_{a-2}}(y'))$  belongs to  $[0, 1 - \theta]^2 \cup [\theta, 1]^2$ . The sequence  $(\eta_{N_{a-2}+n})_{n \geq 0}$  is independent of  $(Y, Y')$  and has the same law as  $(\eta_n)_{n \geq 0}$ . Hence, we may apply Lemma 15 to  $(Y, Y')$  and the sequence  $(\eta_{N_{a-2}+n})_{n \geq 0}$  instead of  $(y, y')$  and  $(\eta_n)_{n \geq 0}$ . We get

$$\Delta N_a \geq (1 + (q_k + 1 - q_1 - q_2)_+) \mathbf{1}_{[W_a=W_{a-2}]} = g(\|\delta + \theta W_a\|) \mathbf{1}_{[W_a=W_{a-2}]}$$

Item 2 follows. □

We now prove Proposition 4.

On each time interval  $[N_a, N_{a+1} - 1]$ , we have  $A_n = a$ , so

$$\frac{a}{N_{a+1}} < \frac{A_n}{n} \leq \frac{a}{N_a}.$$

Hence we only need to show that  $N_a/a \rightarrow +\infty$  almost surely as  $a \rightarrow +\infty$ . To do this, we observe that for every  $a \geq 3$ ,

$$\begin{aligned} \frac{N_a}{a} &= \frac{1}{a} \sum_{b=0}^{a-1} \Delta N_b \geq \frac{1}{a} \sum_{b=3}^{a-1} \mathbf{1}_{[W_b=W_{b-2}]} g(\|\delta + \theta W_{b-2}\|) \\ &= \frac{1}{a} \sum_{k=2}^{a-2} \mathbf{1}_{[W_{k+1}=W_{k-1}]} g(\|\delta + \theta W_{k-1}\|). \end{aligned}$$

Therefore, it is sufficient to prove that

$$\frac{1}{a} \sum_{k=1}^a \mathbf{1}_{[W_{k+1}=W_{k-1}]} g(|\delta + \theta W_{k-1}|) \rightarrow +\infty \text{ almost surely as } a \rightarrow +\infty.$$

The positive function  $x \mapsto g(|\delta + x|)$  can be written as the limit of a non-decreasing sequence of continuous 1-periodic functions. Given such functions, the sequence of their integrals on  $[0, 1]$  tends to infinity by Beppo-Levi's Lemma.

Thus, it suffices to prove that for every continuous 1-periodic function  $f : \mathbf{R} \rightarrow \mathbf{R}$ ,

$$\frac{1}{a} \sum_{k=1}^a \mathbf{1}_{[W_{k+1}=W_{k-1}]} f(\theta W_{k-1}) \rightarrow \frac{1}{2} \int_0^1 f(x) dx \text{ almost surely as } a \rightarrow +\infty.$$

Since for every  $k \geq 1$ ,  $\mathbf{P}[W_{k+1} = W_{k-1} | \mathcal{F}_{N_k}] = 1/2$ , Lemma 16 applied to the random variables  $Z_k := \mathbf{1}_{[W_{k+1}=W_{k-1}]} f(\theta W_{k+1}) = \mathbf{1}_{[W_{k+1}=W_{k-1}]} f(\theta W_{k-1})$  and to the filtration  $(\mathcal{G}_k)_{k \geq 0} := (\mathcal{F}_{N_k})_{k \geq 0}$  yields

$$\frac{1}{a} \sum_{k=1}^a \left( \mathbf{1}_{[W_{k+1}=W_{k-1}]} f(\theta W_{k-1}) - \frac{1}{2} f(\theta W_{k-1}) \right) \rightarrow 0 \text{ almost surely as } n \rightarrow +\infty.$$

But Lemma 18 yields

$$\frac{1}{a} \sum_{k=1}^a f(\theta W_{k-1}) \rightarrow \int_0^1 f(x) dx \text{ almost surely as } a \rightarrow +\infty.$$

Hence the conclusion follows.  $\square$

## 4.2 Proof of Proposition 5

We now deduce Proposition 5 from Proposition 4, Lemma 11 and corollary 12.

Let  $(y, y') \in \mathbf{I}^2$ . Then, by Fubini's theorem

$$\mathbf{E} \left[ \frac{1}{n} \sum_{k=0}^{n-1} \|F_k(y') - F_k(y)\| \right] = \int_0^{1/2} \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{P}[\|F_k(y') - F_k(y)\| \geq r] dr.$$

By the Lebesgue dominated convergence theorem, it is sufficient to check that for every  $r \in ]0, 1/2[$ ,

$$\frac{1}{n} \sum_{k=0}^{n-1} \mathbf{P}[\|F_k(y') - F_k(y)\| \geq r] \rightarrow 0 \text{ almost surely as } n \rightarrow +\infty.$$

Since the sequence  $(\ell_j)_{j \geq 0}$  tends to 0, one only needs to consider the case where  $r = \ell_j$ , with  $k \geq 0$ .

Observe that when  $(\zeta_n)_{n \geq 1}$  is any sequence of independent Bernoulli random variables with parameter  $1/2$ , and  $t$  is a positive integer, the random variable

$$\rho_t = \inf\{n \geq 0 : \zeta_1 + \dots + \zeta_n = t\}$$

is negative binomial with parameters  $t$  and  $1/2$ , and we have then

$$\mathbf{P}[\rho_t \leq 2t - 1] = \mathbf{P}[\zeta_1 + \dots + \zeta_{2t-1} \geq t] = \frac{1}{2},$$

since the binomial random variable  $\zeta_1 + \dots + \zeta_{2t-1}$  is symmetric with regard to  $t - 1/2$ .

For every  $k \geq 0$ , set

$$\tilde{T}_k := \inf\{n \geq 0 : (F_{k+n}(y), F_{k+n}(y')) \in B_\theta\}$$

By Lemma 11 and Markov property, conditionally on  $\mathcal{F}_k$ , the random variable  $\tilde{T}_k$  is negative binomial with parameters  $t_{(F_k(y), F_k(y'))}$  and  $1/2$ . Hence

$$\mathbf{P}[\tilde{T}_k \leq 2t_{(F_k(y), F_k(y'))} - 1 | \mathcal{F}_k] = \frac{1}{2}.$$

Set  $n_j := q_{j+1} - q_j - 2$ . On the event  $[|F_k(y') - F_k(y)| \geq \ell_j]$ , we have  $t_{(F_k(y), F_k(y'))} \leq n_j$  by corollary 12, so  $\mathbf{P}[\tilde{T}_k \leq 2n_j - 1 | \mathcal{F}_k] \geq 1/2$ . Hence

$$\begin{aligned} \sum_{i=k}^{k+2n_j-1} \mathbf{P}[(F_i(y), F_i(y')) \in B_\theta] &\geq \mathbf{P}[\tilde{T}_k \leq 2n_j - 1] \\ &\geq \mathbf{P}[\tilde{T}_k \leq 2n_j - 1 ; |F_k(y') - F_k(y)| \geq \ell_j] \\ &\geq \frac{1}{2} \mathbf{P}[|F_k(y') - F_k(y)| \geq \ell_j] \end{aligned}$$

Since any non-negative integer belong to at most  $2n_j$  intervals of the form  $[k, k + 2n_j - 1]$ , summation over all integers  $k \in [0, n - 1]$  yields

$$\sum_{k=0}^{n-1} \mathbf{P}[|F_k(y') - F_k(y)| \geq \ell_j] \leq 4n_j \sum_{i=0}^{n+2n_j-2} \mathbf{P}[(F_i(y), F_i(y')) \in B_\theta]$$

Dividing by  $n + 2n_j - 1$  and letting  $n$  go to infinity yields the result by Proposition 4.

## 5 Constructive proof that $[T_\theta, T_\theta^{-1}]$ is Bernoulli

The strategy used for  $[T_\theta, \text{Id}]$  also works for  $[T_\theta, T_\theta^{-1}]$  with some adaptations, which are often slight but at some places more substantial. Let us explain them.

We define  $[T_\theta, T_\theta^{-1}]$  on the space  $\{-1, 1\}^{\mathbf{Z}^+} \times \mathbf{I}$  with the same formula used to define  $[T_\theta, \text{Id}]$  on the space  $\{0, 1\}^{\mathbf{Z}^+} \times \mathbf{I}$ . In the probabilistic reformulation, we work with an i.i.d. sequence  $(\xi_n)_{n \in \mathbf{Z}}$  of Rademacher random variables. The process  $(Y_n)_{n \in \mathbf{Z}}$  is a Markov chain on  $\mathbf{I}$  whose evolution is governed by  $(\xi_n)_{n \in \mathbf{Z}}$  through the recursion relation

$$\forall n \in \mathbf{Z}, \quad Y_n = (Y_{n-1} - \xi_n \theta) \pmod{1},$$

where  $\xi_n$  is uniform on  $\{-1, 1\}$  and independent of  $\mathcal{F}_{n-1}^{\xi, Y}$ .

For every  $n \in \mathbf{Z}$ , the random variable

$$\eta_n := (\mathbf{1}_{[0, \theta[}(Y_{n-1}) - \mathbf{1}_{[\theta, 1[}(Y_{n-1})) \xi_n.$$

is also uniform on  $\{-1, 1\}$  and independent of  $\mathcal{F}_{n-1}^{\xi, Y}$ , and the recursion relation can be written alternatively

$$\forall n \in \mathbf{Z}, \quad Y_n = f_{\eta_n}(Y_{n-1}),$$

where the maps  $f_{-1}$  and  $f_1$  from  $\mathbf{I}$  to  $\mathbf{I}$  are defined by

$$f_h(y) := \left( y - h\theta(\mathbf{1}_{[0, \theta[}(y) - \mathbf{1}_{[\theta, 1[}(y)) \right) \pmod{1}.$$



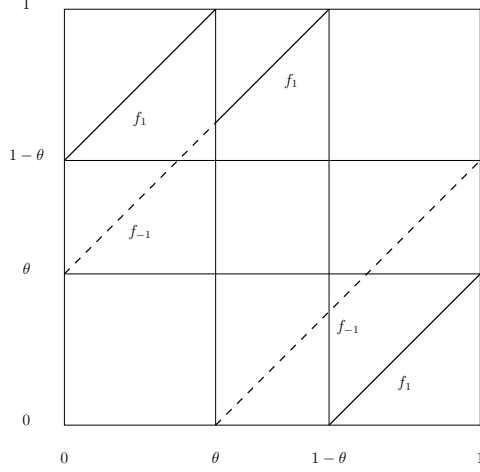


Figure 6: Graphs of  $f_{-1}$  and  $f_1$

We want to show that for every  $n \in \mathbf{Z}$ , the random variable  $Y_n$  can be almost surely recovered from the knowledge of  $(\eta_k)_{k \leq n}$ . To do this, we introduce the random maps  $F_n := f_{\eta_n} \circ \dots \circ f_{\eta_1}$  and the  $\sigma$ -fields  $\mathcal{F}_n = \sigma(\eta_1, \dots, \eta_n)$  for  $n \geq 0$ . Let  $B_\theta := [0, \theta[ \times ]\theta, 1[ \cup ]\theta, 1[ \times [0, \theta[$  and  $D$  be the diagonal of  $\mathbf{I}^2$ . The global strategy, based on a coupling argument is unchanged:

1. For every  $(y, y') \in \mathbf{I}^2$ , the process  $((F_n(y), F_n(y'))_{n \geq 0}$  is a Markov chain on  $\mathbf{I}^2$  in the filtration  $(\mathcal{F}_n)_{n \geq 0}$ , whose transition probabilities (see figure 7) are given by

$$K((y, y'), \cdot) = \frac{1}{2} (\delta_{(f_{-1}(y), f_{-1}(y'))} + \delta_{(f_1(y), f_1(y'))}).$$

2. The process  $(A_n)_{n \geq 0}$  defined by

$$A_n := \sum_{k=0}^{n-1} \mathbf{1}_{B_\theta}(F_k(y), F_k(y'))$$

is predictable in the filtration  $(\mathcal{F}_n)_{n \geq 0}$ .

3. Given  $(y, y') \in \mathbf{I}^2$ , the process  $(\overline{(F_n(y') - F_n(y))})_{n \geq 0}$  with values in  $\mathbf{R}/\mathbf{Z}$  is still a time-changed symmetric random walk, making steps  $\pm 2\bar{\theta}$ , with probability 1/2 each. More precisely, for every  $n \geq 0$ ,

$$F_n(y') - F_n(y) \equiv y' - y + 2\theta W_{A_n} \pmod{1},$$

where  $(W_a)_{a \geq 0}$  is a simple symmetric random walk on  $\mathbf{Z}$ , independent of  $(A_n)_{n \geq 0}$ .

4. The process  $((F_n(y), F_n(y'))_{n \geq 0}$  visits  $B_\theta$  more and more rarely:  $A_n/n \rightarrow 0$  almost surely as  $n \rightarrow +\infty$ .
5. The final argument given in subsection 1.5 (which uses that invariant probability measures for  $K$  are carried by  $B_\theta^c$ , and therefore by  $D$ ) is unchanged.

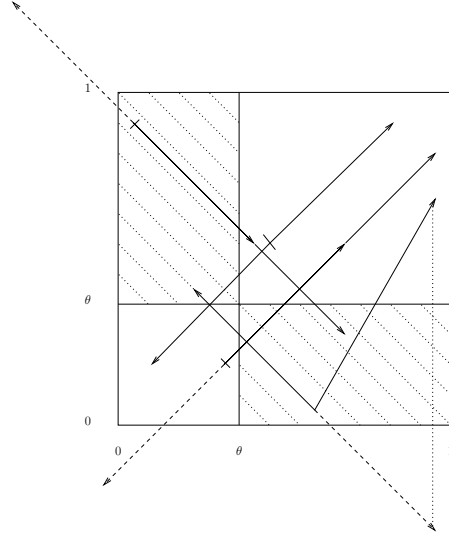


Figure 7: Probability transitions from different points  $(y, y')$ . The probability associated to each arrow is  $1/2$ . The set  $B_\theta$  is the union of the shaded rectangles. When  $(y, y') \in B_\theta^c$ , the same rotation  $T_{\pm\theta}$  is applied to both coordinates. When  $(y, y') \in B_\theta$ , rotation  $T_\theta$  is applied to one coordinate whereas rotation  $T_{-\theta}$  is applied to the other one.

Yet, the laws of the reaching times of the set  $B_\theta$  from any point  $(y, y') \in D^c$  are very different in our new situation, and this changes notably the proof of step 3. The main changes occur in Lemmas 15 and 21 which must be replaced by Lemmas 22 and 23 below. The purpose is the same, namely exhibiting situations that occur sufficiently often where the process  $((F_n(y), F_n(y'))_{n \geq 0})$  spends a lot of time outside of  $B_\theta$  (and therefore close to the diagonal  $D$ ), but the assumptions and the lower bounds for the reaching time of  $B_\theta$  are different.

For every  $a \geq 0$ , we set  $N_a = \inf\{n \geq 0 : A_n \geq a\}$  and  $\Delta N_a = N_{a+1} - N_a$ . By construction, the random times  $(N_a - 1)_{a \geq 1}$  are exactly the visit times of  $B_\theta$  by the process  $((F_n(y), F_n(y'))_{n \geq 0})$ , so they are stopping-time in the filtration  $(\mathcal{F}_n)_{n \geq 0}$ .

We now relate the distribution of  $N_2 - N_1$  given  $W_1 = -1$  with the distribution of stopping times of the form  $\tau_{-1,b} := \inf\{n \geq 0 : W_n \in \{-1, b\}\}$  with  $b \geq 0$ .

Remind that  $\ell_{-1} = 1$ ,  $\ell_0 = \theta$ ,  $\ell_1 = 1 - \lfloor 1/\theta \rfloor \theta < \min(\theta, 1 - 2\theta)$  and  $q_1 = a_1 \geq 2$  since we assumed that  $0 < \theta < 1/2$ . Hence the condition  $\delta \in ]0, \min(\theta, 1 - 2\theta)[$  below holds whenever  $0 < \delta < \ell_1$ .

**Lemma 22** *Let  $(y, y') \in \mathbf{I}^2$ . Assume that  $(y, y') \in B_\theta$  (so  $N_1 = 1$ ) and  $y' = T_{2\theta+\delta}(y)$ . for some  $\delta \in ]0, \min(\theta, 1 - 2\theta)[$ . Then*

1.  $(T_\theta(y), T_{-\theta}(y')) \notin B_\theta$ . Hence on the event  $[W_1 = -1]$ ,  $(F_1(y), F_1(y')) \notin B_\theta$ .
2. If  $\ell_k \leq \delta < \ell_{k-1}$  with  $k \geq 2$ , then for every  $\lambda > 0$ ,

$$\mathbf{E} \left[ \frac{1 - (\cosh \lambda)^{-\Delta N_1}}{1 - (\cosh \lambda)^{-1}} \middle| W_1 = -1 \right] \geq \mathbf{E} \left[ \frac{1 - (\cosh \lambda)^{-1 - \tau_{-1, q_k - 2}}}{1 - (\cosh \lambda)^{-1}} \right] = \frac{\tanh\left(\frac{q_k - 1}{2} \lambda\right)}{\tanh(\lambda/2)}.$$

Proof. By assumption, we have  $(y, y') \in [0, \theta[ \times ]\theta, 1[$  or  $(y, y') \in ]\theta, 1[ \times ]0, \theta[$ . Since

$y' = (y + 2\theta + \delta) \bmod 1$ , we get (see figure 8)

$$\begin{cases} 0 \leq y < \min(\theta, 1 - 2\theta - \delta) \\ 2\theta + \delta \leq y' < \min(3\theta + \delta, 1) \end{cases} \quad \text{or} \quad \begin{cases} \max(\theta, 1 - 2\theta - \delta) \leq y < 1 - \theta - \delta \\ \max(3\theta + \delta - 1, 0) \leq y' < \theta \end{cases}$$

Thus

$$\begin{cases} \theta \leq T_\theta(y) < \min(2\theta, 1 - \theta - \delta) \\ \theta + \delta \leq T_{-\theta}(y') < \min(2\theta + \delta, 1 - \theta) \end{cases} \quad \text{or} \quad \begin{cases} \max(2\theta, 1 - \theta - \delta) \leq T_\theta(y) < 1 - \delta \\ \max(2\theta + \delta, 1 - \theta) \leq T_{-\theta}(y') < 1 \end{cases}$$

In both cases,  $(T_\theta(y), T_{-\theta}(y')) \notin B_\theta$ . This shows item 1.

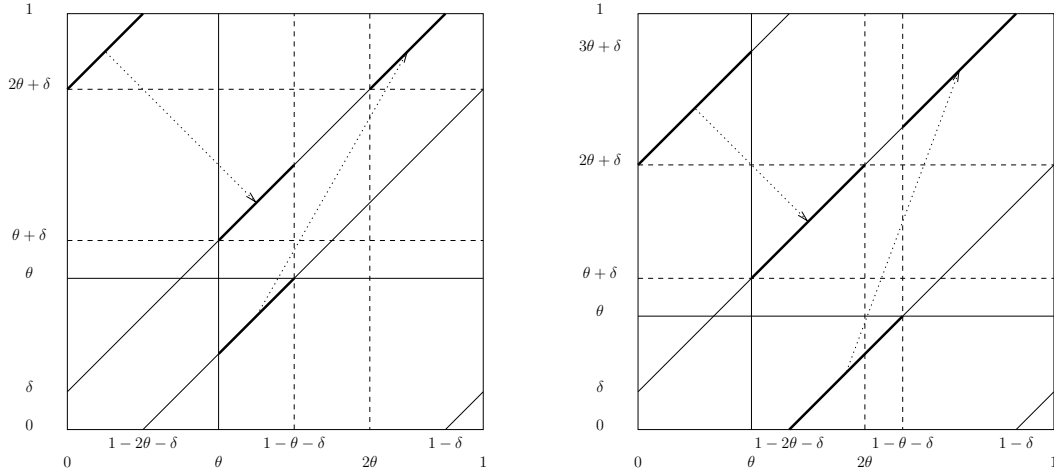


Figure 8: Possible locations of  $(y, y')$  and  $(T_\theta(y), T_{-\theta}(y'))$  when  $(y, y') \in B_\theta$  and  $y' = T_{2\theta+\delta}(y)$  with  $0 < \delta < \min(\theta, 1 - 2\theta)$ , according that  $3\theta + \delta \geq 1$  (left) or  $3\theta + \delta \leq 1$  (right). Under these assumptions, the couple  $(T_\theta(y), T_{-\theta}(y'))$  is not in  $B_\theta$ .

Let us prove item 2. Let  $a$  (respectively  $b$ ) be the number of iterations of the Cartesian product map  $T_{-\theta} \times T_{-\theta}$  (respectively  $T_\theta \times T_\theta$ ) necessary to reach  $B_\theta$  from  $(T_\theta(y), T_{-\theta}(y'))$ . By item 1, we have  $a \geq 1$  and  $b \geq 1$ . But, we have also

$$a = \min\{k \geq 0 : T_\theta^{-k}(T_\theta(y)) \in [\theta - \delta, \theta] \cup [1 - \delta, 1]\},$$

$$b = \min\{k \geq 0 : T_\theta^k(T_\theta(y)) \in [\theta - \delta, \theta] \cup [1 - \delta, 1]\}.$$

Since  $T_\theta([1 - \delta, 1]) = [\theta - \delta, \theta]$ , corollary 9 yields  $a + b \geq q_k - 1$ .

On the event  $[W_1 = -1]$ , the random variable  $\Delta N_1 - 1 = (N_2 - 1) - N_1$  is the hitting time of  $\{-a, b\}$  by  $(W_{1+n} - W_1)_{n \geq 1}$ . This process is independent of  $W_1$  and is a symmetric simple random walk, like  $W$ . Hence the distribution of  $\Delta N_1 - 1$  under  $\mathbf{P}[\cdot | W_1 = -1]$  is the distribution on  $\tau_{-a,b} := \inf\{n \geq 1 : W_n \in \{-a, b\}\}$  under  $\mathbf{P}$ . Thus

gambler's ruin Theorem (Proposition 17) yields for every  $\lambda > 0$ ,

$$\begin{aligned}
\mathbf{E}[(\cosh \lambda)^{-\Delta N_1} | W_1 = -1] &= \mathbf{E}[(\cosh \lambda)^{-\tau_{-a,b}}] \\
&= \frac{\cosh\left(\frac{-a+b}{2}\lambda\right)}{\cosh\left(\frac{a+b}{2}\lambda\right)} \\
&\leq \frac{\cosh\left(\left(\frac{a+b}{2}-1\right)\lambda\right)}{\cosh\left(\frac{a+b}{2}\lambda\right)} \\
&= \mathbf{E}[(\cosh \lambda)^{-\tau_{-1,a+b-1}}] \\
&\leq \mathbf{E}[(\cosh \lambda)^{-\tau_{-1,q_k-2}}] \text{ since } \tau_{-1,a+b-1} \geq \tau_{-1,q_k-2} \\
&= \frac{\cosh\left(\frac{q_k-3}{2}\lambda\right)}{\cosh\left(\frac{q_k-1}{2}\lambda\right)}.
\end{aligned}$$

Hence

$$\begin{aligned}
\mathbf{E}\left[\frac{1 - (\cosh \lambda)^{-\Delta N_1}}{1 - (\cosh \lambda)^{-1}} \middle| W_1 = -1\right] &\geq \mathbf{E}\left[\frac{1 - (\cosh \lambda)^{-1-\tau_{-1,q_k-2}}}{1 - (\cosh \lambda)^{-1}}\right] \\
&= \mathbf{E}\left[\frac{\cosh \lambda - (\cosh \lambda)^{-\tau_{-1,q_k-2}}}{\cosh \lambda - 1}\right] \\
&= \frac{\cosh \lambda \cosh\left(\frac{q_k-1}{2}\lambda\right) - \cosh\left(\frac{q_k-3}{2}\lambda\right)}{(\cosh \lambda - 1) \cosh\left(\frac{q_k-1}{2}\lambda\right)} \\
&= \frac{\sinh \lambda \sinh\left(\frac{q_k-1}{2}\lambda\right)}{2 \sinh^2(\lambda/2) \cosh\left(\frac{q_k-1}{2}\lambda\right)} \\
&= \frac{\tanh\left(\frac{q_k-1}{2}\lambda\right)}{\tanh(\lambda/2)}.
\end{aligned}$$

Item 2 follows. □

Now, let us explain how we use Lemma 22.

By symmetry, the roles of  $y$  and  $y'$  in can be switched provided the event  $[W_1 = -1]$  is replaced with the event  $[W_1 = 1]$ .

For every  $a \geq 1$ , we introduce a sign  $\varepsilon_a$  which indicates the shortest path in the circle  $\mathbf{R}/\mathbf{Z}$  to go from  $\overline{Y_{N_a}}$  to  $\overline{Y'_{N_a}}$ . More precisely, we set

$$\begin{aligned}
\varepsilon_a &= 1 & \text{if } Y'_{N_a} - Y_{N_a} \in ]0, 1/2[ + \mathbf{Z}, \\
\varepsilon_a &= -1 & \text{if } Y'_{N_a} - Y_{N_a} \in ]1/2, 1[ + \mathbf{Z} \\
\varepsilon_a &= 0 & \text{if } Y'_{N_a} - Y_{N_a} \in (1/2)\mathbf{Z}.
\end{aligned}$$

Given  $\lambda > 0$ , call  $h_\lambda$  the function from  $[0, 1/2]$  to  $[0, +\infty]$  defined by  $h_\lambda(0) = +\infty$ ,  $h_\lambda(r) = 1$  if  $r \geq \ell_1$  and

$$h_\lambda(r) := \frac{\tanh\left(\frac{q_k-1}{2}\lambda\right)}{\tanh\left(\frac{1}{2}\lambda\right)} \text{ whenever } \ell_k \leq r < \ell_{k-1} \text{ with } k \geq 2.$$

As  $\lambda \rightarrow 0+$ ,  $h_\lambda(r)$  increases to  $h(r)$ , where  $h(0) = +\infty$ ,  $h(r) = 1$  if  $r \geq \ell_1$  and  $h(r) := q_k - 1$  whenever  $\ell_k \leq r < \ell_{k-1}$  with  $k \geq 2$ . We see that the integral of  $h$  over  $[0, 1/2]$  is infinite in the same way as we did for the function  $g$  in Lemma 21.

Hence, strong Markov property at time  $N_a$  and Lemma 22 yield the next Lemma, which replaces Lemma 21.

**Lemma 23** For every  $a \geq 1$  and  $\lambda > 0$ ,

$$\begin{aligned} \mathbf{E}\left[\frac{1 - (\cosh \lambda)^{\Delta N_a}}{1 - \cosh \lambda} \middle| \mathcal{F}_{N_a}\right] &\geq h_\lambda(\|Y'_{N_a} - Y_{N_a}\|)\mathbf{1}_{[W_a - W_{a-1} = -\varepsilon_a]} \\ &\geq h_\lambda(\|Y'_{N_{a-1}} - Y_{N_{a-1}} - 2\theta\|)\mathbf{1}_{[W_a - W_{a-1} = -\varepsilon_a]}. \end{aligned}$$

Therefore,

$$\mathbf{E}\left[\frac{1 - (\cosh \lambda)^{\Delta N_a}}{1 - \cosh \lambda} \middle| \mathcal{F}_{N_{a-1}}\right] \geq \frac{1}{2}h_\lambda(\|Y'_{N_{a-1}} - Y_{N_{a-1}} - 2\theta\|).$$

We now use Lemma 23. The inequalities above may be false for  $a = 0$ , but this has no influence on the Cesàro limits below. Observing that for every  $\lambda > 0$ ,

$$\Delta N_a \geq \frac{1 - (\cosh \lambda)^{\Delta N_a}}{1 - \cosh \lambda}$$

and applying Lemma 16, we get almost surely

$$\begin{aligned} \liminf_{a \rightarrow +\infty} \frac{N_a}{a} &= \liminf_{a \rightarrow +\infty} \frac{1}{a} \sum_{b=0}^{a-1} \Delta N_b \\ &\geq \liminf_{a \rightarrow +\infty} \frac{1}{a} \sum_{b=0}^{a-1} \frac{1 - (\cosh \lambda)^{\Delta N_b}}{1 - \cosh \lambda} \\ &= \liminf_{a \rightarrow +\infty} \frac{1}{a} \sum_{b=0}^{a-1} \mathbf{E}\left[\frac{1 - (\cosh \lambda)^{\Delta N_b}}{1 - \cosh \lambda} \middle| \mathcal{F}_{N_{a-1}}\right] \\ &\geq \liminf_{a \rightarrow +\infty} \frac{1}{a} \sum_{b=0}^{a-1} h_\lambda(\|y' - y + 2\theta W_{a-1} - 2\theta\|). \end{aligned}$$

The function  $x \mapsto h_\lambda(\|x + y' - y - 2\theta\|)$  can be written as the limit of an increasing sequence of non-negative continuous 1-periodic functions. Applying Lemma 18 and Beppo-Levi's Lemma yields almost surely

$$\begin{aligned} \liminf_{a \rightarrow +\infty} \frac{1}{a} \sum_{b=0}^{a-1} h_\lambda(\|y' - y + 2\theta W_{a-1} - 2\theta\|) &\geq \int_0^1 h_\lambda(\|x + y' - y - 2\theta\|) dx \\ &= 2 \int_0^{1/2} h_\lambda(r) dr. \end{aligned}$$

Hence

$$\liminf_{a \rightarrow +\infty} \frac{N_a}{a} \geq 2 \int_0^{1/2} h_\lambda(r) dr \quad \text{a.s.}$$

Letting  $\lambda$  go to 0 and applying Beppo-Levi's Lemma again yields

$$\liminf_{a \rightarrow +\infty} \frac{N_a}{a} \geq +\infty \quad \text{a.s.}$$

We derive that  $A_n/n \rightarrow 0$  almost surely in the same way as we did for  $[T_\theta, \text{Id}]$ .

## 6 Related results and open questions

In this section, we view still open related questions and a few variants of what we have done for  $[T_\theta, \text{Id}]$  and  $[T_\theta, T_\theta^{-1}]$ .

### 6.1 Almost sure continuity of the isomorphism and its inverse

We come back to the transformation  $[T_\theta, \text{Id}]$  defined on  $X \times \mathbf{I}$ , where  $X = \{0, 1\}^{\mathbf{Z}^+}$  and  $\mathbf{I} = [0, 1[$ . We endow  $X$  with the product topology,  $\mathbf{I}$  with the topology deriving from the metric defined by  $d(y, y') := \|y' - y\|$ , and  $X \times \mathbf{I}$  with the product topology. Since the map  $T_\theta$  is continuous on  $\mathbf{I}$ , the map  $[T_\theta, \text{Id}]$  is continuous on  $X \times \mathbf{I}$ .

The partition of  $X \times \mathbf{I}$  introduced by Parry is  $\alpha_\theta := \{A_0^\theta, A_1^\theta\}$ , where

$$A_0^\theta := \{x \in X : x(0) = 0\} \times [0, \theta[ \cup \{x \in X : x(0) = 1\} \times [0, 1 - \theta[,$$

$$A_1^\theta := \{x \in X : x(0) = 0\} \times [\theta, 1[ \cup \{x \in X : x(0) = 1\} \times [1 - \theta, 1[.$$

For every  $(x, y) \in X \times \mathbf{I}$ , denote by  $\alpha_\theta(x, y)$  the index of the only block containing  $(x, y)$  in the partition  $\alpha_\theta$ . The ‘ $\alpha_\theta$ -name’ map is  $\Phi_\theta : X \times \mathbf{I} \rightarrow X$  defined by

$$\Phi_\theta(x, y) := ((\alpha_\theta \circ [T_\theta, \text{Id}]^n)(x, y))_{n \geq 0}.$$

It transforms the dynamical system  $(X \times [0, 1[, \mathcal{B}(X) \otimes \mathcal{P}([0, 1]), \mu \otimes \nu, [T_\theta, \text{Id}])$  into  $(X, \mathcal{B}(X), \mu, S)$ .

By construction, for every  $(x, y) \in X \times \mathbf{I}$ ,

$$\begin{aligned} \alpha_\theta(x, y) = \mathbf{1}_{A_1^\theta}(x, y) &= (\mathbf{1}_{[(1-x_0)\theta+x_0(1-\theta), 1[}(y)) \pmod 2 \\ &= (x_0 + \mathbf{1}_{[\theta, 1[}(T_\theta^{x(0)}(y))) \pmod 2. \end{aligned}$$

Since  $\nu(\{0, \theta, 1 - \theta\}) = 0$ , the map  $\alpha_\theta : X \times \mathbf{I} \rightarrow \{0, 1\}$  is continuous  $\mu \otimes \nu$ -almost everywhere. And since  $[T_\theta, \text{Id}]$  is continuous and preserves  $\mu \otimes \nu$ , we deduce that the morphism  $\Phi_\theta$  is continuous  $\mu \otimes \nu$ -almost everywhere.

We proved that the partition  $\alpha_\theta$  is generating, so  $\Phi_\theta$  is an isomorphism. Thouvenot asked me whether  $\Phi_\theta^{-1}$  is continuous  $\mu$ -almost everywhere or not. Answering this question is difficult since we do not have simple formulas to recover  $(x, y) \in X \times \mathbf{I}$  from  $\Phi_\theta(x, y)$ .

Once again, we reformulate our problem in probabilistic terms. Remind that

$$\Phi_\theta((\xi_{-i})_{i \geq 0}, Y_0) = (\eta_{-i})_{i \geq 0}.$$

We know that  $Y_0$  is almost surely a function of  $(\eta_{-i})_{i \geq 0}$ . To prove that  $\Phi_\theta^{-1}$  is continuous  $\mu$ -almost everywhere, we just have to prove that this function is almost surely continuous, thanks to the recursion relations

$$\begin{aligned} \xi_{-i} &= (\eta_{-i} + \mathbf{1}_{[\theta, 1[}(Y_{-i-1})) \pmod 2 \\ &= (\eta_{-i} + \mathbf{1}_{[\theta, 1[}(T_\theta^{\xi_{-i} + \dots + \xi_0}(Y_0))) \pmod 2. \end{aligned}$$

For every  $n \geq 0$ , we know that

$$Y_0 = f_{\eta_0} \circ \dots \circ f_{\eta_{-n+1}}(Y_{n+1}),$$

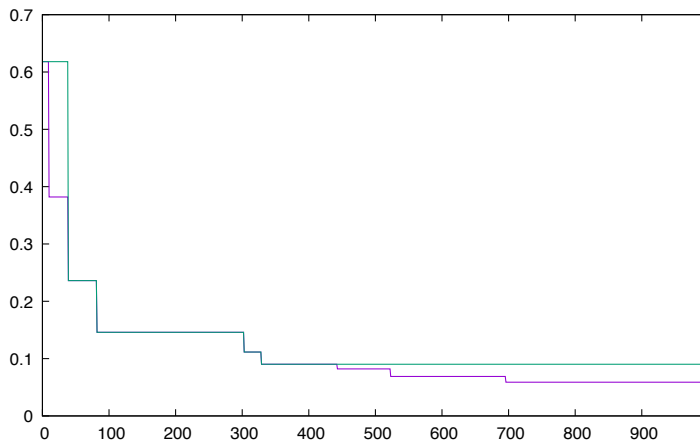


Figure 9: Diameter (for the metric induced by the usual metric on  $\mathbf{R}$ ) and Lebesgue measure of  $P_n(\mathbf{I})$ , for different values of  $n$  and for  $\theta = (\sqrt{5} - 1)/2$ . When  $P_n(\mathbf{I})$  is an interval, these two quantities coincide. Otherwise, the diameter is larger.

that  $(\eta_0, \dots, \eta_{-n+1})$  is independent of  $\mathcal{F}_{-n}^{\xi, Y}$  therefore of  $Y_{-n}$  and  $Y_{-n}$  is uniform on  $\mathbf{I}$ .

Given  $n \geq 0$ , the random map  $P_n = f_{\eta_0} \circ \dots \circ f_{\eta_{-n+1}}$  has the same law as  $F_n = f_{\eta_n} \circ \dots \circ f_{\eta_1}$ . Yet, there is a great different between the processes  $(P_n)_{n \geq 0}$  and  $(F_n)_{n \geq 0}$ . When we couple in the past, the range  $P_n(\mathbf{I})$  can only decrease or stay equal as  $n$  increases. But when we couple in the future, no such inclusion holds; we only know that the Lebesgue measures  $|F_n(\mathbf{I})|$  can only decrease or stay unchanged and we proved (Proposition 3) that  $|F_n(\mathbf{I})|$  goes to 0 almost surely as  $n$  goes to infinity.

Since for each  $n \geq 0$ , the random variables  $|P_n(\mathbf{I})|$  and  $|F_n(\mathbf{I})|$  have the same distribution, and since the sequence  $(P_n(\mathbf{I}))_{n \geq 0}$  is non-decreasing, we deduce that  $|P_n(\mathbf{I})|$  also goes to 0 almost surely as  $n$  goes to infinity. These facts and simulations (view figure 9) lead us to formulate the following conjecture.

**Conjecture 24** *The diameter of  $P_n(\mathbf{I})$  (for the distance  $d$ ) goes almost surely to 0 as  $n$  goes to infinity.*

Observe that the conclusion will be false if we replace  $P_n(\mathbf{I})$  by  $F_n(\mathbf{I})$ .

Assume that this conjecture is true. Then, given  $\delta \in ]0, 1/2[$ , we can find almost surely an integer  $N \geq 1$  such that the diameter of  $P_n(\mathbf{I})$  is at most  $\delta$ ; then, the knowledge of  $\eta_0, \dots, \eta_{-N+1}$  only is sufficient to provide a random variable  $\tilde{Y}_0$  such that  $\|Y_0 - \tilde{Y}_0\| \leq \delta$ . Hence,  $Y_0$  can be recovered with probability 1 as an almost everywhere continuous function of  $(\eta_{-i})_{i \geq 0}$ , giving a positive answer to Thouvenot's question.

Yet, proving this conjecture requires new ideas and looks difficult.

## 6.2 Changing the partition

To define the random sequence  $(\eta_n)_{n \in \mathbf{Z}}$ , we split the interval  $\mathbf{I} = [0, 1[$  at  $\theta$ . An advantage of the (non-trivial) partition  $\iota = \{[0, \theta[, [\theta, 1[\}$  is that for every  $n \geq 0$ , the partition

$$\bigvee_{k=0}^n T_{\theta}^k \iota$$

comprises only  $n + 1$  intervals given by the subdivision  $((k\theta) \bmod 1)_{0 \leq k \leq n}$ , which is the least number possible. This facilitates the study of the random maps  $F_n = f_{\eta_n} \circ \dots \circ f_{\eta_1}$ .

Yet, if the definition of the sequence  $(\eta_n)_{n \in \mathbf{Z}}$  is modified only by a modification of the splitting point -  $\alpha$  instead of  $\theta$ , say - a large part of our preceding analysis remains true, namely Lemma 13 and its variant for  $[T_\theta, T_\theta^{-1}]$ . Yet, the set  $B_\theta$  must be replaced by  $B_\alpha$ , and estimating the reaching time of  $B_\alpha$  from any point in  $\mathbf{I}^2$  are less simple.

A remarkable phenomenon occurs when we study  $[T_\theta, T_\theta^{-1}]$  and split  $\mathbf{I}$  at  $1/2$ , namely when we set

$$\eta_n := (\mathbf{1}_{[0, 1/2[}(Y_{n-1}) - \mathbf{1}_{[1/2, 1[}(Y_{n-1}))\xi_n.$$

Indeed, replacing the Markov chain  $((\xi_n, Y_n)_{n \in \mathbf{Z}}$  by  $((-\xi_n, (1 - Y_n) \bmod 1)_{n \in \mathbf{Z}}$  preserves its law and modifies the sequence  $(\eta_n)_{n \in \mathbf{Z}}$  only on a null set, namely on the negligible event  $[\exists n \in \mathbb{N} : Y_n \in \{0, 1/2\}]$ . Hence, the knowledge of  $(\eta_n)_{n \in \mathbf{Z}}$  is not sufficient to recover almost surely the Markov chain  $((\xi_n, Y_n)_{n \in \mathbf{Z}}$ . By symmetry, at least one bit of information is lost.

We complete this observation with a coupling argument to prove that exactly one bit of information is lost. Given  $y$  and  $y'$  in  $\mathbf{I}$ , look at the Markov chain  $((F_n(y), F_n(y'))_{n \geq 0}$ , where  $F_n = f_{\eta_n} \circ \dots \circ f_{\eta_1}$ .

- On the event  $[(F_n(y), F_n(y')) \in B_{1/2}^c]$ , we have

$$\|F_{n+1}(y') - F_{n+1}(y)\| = \|F_n(y') - F_n(y)\| \leq \|F_n(y') + F_n(y)\|.$$

- On the event  $[(F_n(y), F_n(y')) \in B_{1/2}]$ , we have

$$\|F_{n+1}(y') + F_{n+1}(y)\| = \|F_n(y') + F_n(y)\| \leq \|F_n(y') - F_n(y)\|.$$

As a result, the quantity  $\min(\|F_n(y') + F_n(y)\|, \|F_n(y') - F_n(y)\|)$  can only decrease or stay unchanged, so it tends to 0 since  $\liminf_{n \rightarrow +\infty} \|F_n(y') - F_n(y)\| = 0$ .

Therefore, if  $Y' := (Y'_n)_{n \in \mathbf{Z}}$  is a copy of  $Y := (Y_n)_{n \in \mathbf{Z}}$  such that  $Y$  and  $Y'$  are i.i.d. conditionally on  $\eta := (\eta_n)_{n \in \mathbf{Z}}$ , the process  $((Y_n, Y'_n)_{n \in \mathbf{Z}}$  thus obtained is a stationary Markov chain with transition kernel  $K$ , so it lives almost surely on  $D \cup D'$ , where  $D'$  is the anti-diagonal

$$D' = \{(y, y') \in \mathbf{I}^2 : y + y' \in \mathbb{Z}\} = \{(0, 0)\} \cup \{(y, 1 - y) : y \in ]0, 1[ \}.$$

Moreover,  $D'$  is an absorbing set, like  $D$ . Since

$$\mathcal{L}((Y, Y')|\eta) = \mathcal{L}(Y|\eta) \otimes \mathcal{L}(Y|\eta),$$

we derive that the conditional law  $\mathcal{L}(Y|\eta)$  is almost surely carried by two points, which gives the desired conclusion.

This symmetry is an exceptional case. Intuitively, we may expect that no information is lost in the other cases, namely when the splitting point is different of  $1/2$ , or when we work with  $[T_\theta, \text{Id}]$ .

### 6.3 Working with more general skew products

So far, we worked only with  $[T_\theta, \text{Id}]$  and  $[T_\theta, T_\theta^{-1}]$ . A natural generalization of these two transformations is  $[T_\alpha, T_\beta]$  where  $\alpha$  and  $\beta$  are real numbers such that  $\beta - \alpha$  is irrational. The map  $[T_\alpha, T_\beta]$  can be defined on the product  $\{\alpha, \beta\}^{\mathbf{Z}^+} \times \mathbf{I}$  by

$$[T_\alpha, T_\beta]((x_n)_{n \geq 0}, y) := ((x_{n+1})_{n \geq 0}, T_{x_0}(y)).$$



Let  $(Y_n)_{n \in \mathbf{Z}}$  be a Markov chain  $(Y_n)_{n \in \mathbf{Z}}$  governed by the recursion relation

$$Y_n = T_{-\xi_n}(Y_{n-1}) \text{ where } \xi_n \text{ is independent of } \mathcal{F}_{n-1}^{\xi, Y} \text{ and uniform on } \{\alpha, \beta\}.$$

Let  $\theta = \beta - \alpha$ . Define  $f_0$  and  $f_1$  from  $\mathbf{I}$  to  $\mathbf{I}$  by

$$f_0(y) := \mathbf{1}_{[0, \theta[}(y)T_{-\alpha}(y) - \mathbf{1}_{[\theta, 1[}(y)T_{-\beta}(y),$$

$$f_1(y) := \mathbf{1}_{[0, \theta[}(y)T_{-\beta}(y) - \mathbf{1}_{[\theta, 1[}(y)T_{-\alpha}(y).$$

We may write alternatively the recursion  $Y_n = f_{\eta_n}(Y_{n-1})$  by adapting the definition of  $(\eta_n)_{n \in \mathbf{Z}}$  accordingly.

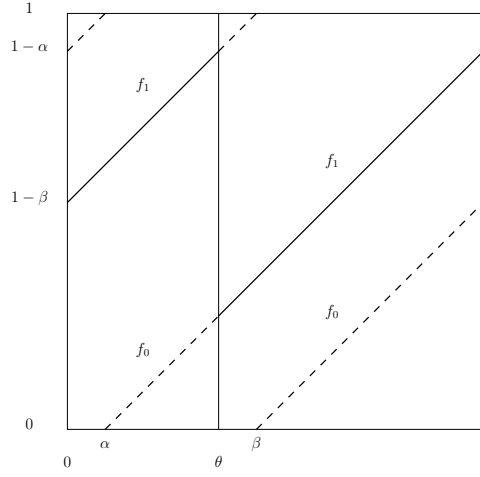


Figure 10: The modified maps  $f_0$  and  $f_1$ .

Once again, we study the Markov chain  $((F_n(y), F_n(y'))_{n \geq 0})$ , where  $(F_n)_{n \geq 0}$  is the sequence of random maps defined by  $F_n = f_{\eta_n} \circ \dots \circ f_{\eta_1}$ . Given  $y$  and  $y'$  in  $\mathbf{I}$ , the difference  $F_n(y') - F_n(y)$  is still (modulo 1)  $\theta$  times a time-changed symmetric simple random walk on  $\mathbf{Z}$ , where the time change is given by the past sojourn time in  $B_\theta$  of the process  $((F_n(y), F_n(y'))_{n \geq 0})$ .

But this time, when the Markov chain  $((F_n(y), F_n(y'))_{n \geq 0})$  is in  $B_\theta^c$ , the steps it makes are uniform on  $\{(-\alpha, -\alpha), (-\beta, -\beta)\}$ . In this general context, estimating the reaching time of  $B_\theta$  from any point in  $\mathbf{I}^2$  is much harder.

## Acknowledgments

I thank Jean-Paul Thouvenot for drawing my attention on this question and Jean-Brossard for his careful reading and fruitful comments. I also thank Vincent Beffara for nice simulations, namely figures 5 and 9.

## References

- [1] R.L. Adler, P.C. Shields, *Skew products of Bernoulli shifts with rotations*. Israel Journal of Mathematics **12**, 215–222 (1972).

- [2] R.L. Adler, P.C. Shields, *Skew products of Bernoulli shifts with rotations*. Israel Journal of Mathematics **19**, 228–236 (1974).
- [3] P. Alessandri, V. Berthé, *Three distance theorems and combinatorics on words*, Enseign. Math. (2) **44**, no. 1-2, 103–132 (1998).  
<https://www.theoremoftheday.org/Docs/3dAlessandriBerthe.pdf>
- [4] J. Feldman, D. Rudolph, *Standardness of sequences of  $\sigma$ -fields given by certain endomorphisms*. Fundamenta Mathematicae. **157**, 175–189 (1998).
- [5] D. Hecklen, C Hoffman,  *$[T, T^{-1}]$  is not standard*. Ergodic Theory and Dynamical Systems **18**, no. 4, 875–878 (1998).
- [6] C. Hoffman, *A zero entropy  $T$  such that the  $[T, \text{Id}]$  endomorphism is nonstandard*. Proceedings of the American Mathematical Society **128**, no. 1, 183–188 (2000).
- [7] C. Hoffman, D. Rudolph *Uniform endomorphisms which are isomorphic to a Bernoulli shift*. Annals of Mathematics, Second Series, Vol. 156, No. 1 (Jul., 2002), pp. 79–101 (2002).
- [8] A. Khintchine, *Metrische Kettenbruchprobleme*. Compositio Mathematica, Tome 1, 361–382 (1935).
- [9] C. Leuridan, *Filtrations associated to some two-to-one transformations*, to be published in Séminaire de Probabilités LI.
- [10] C. Leuridan, *Bernoulliness of  $[T, \text{Id}]$  when  $T$  is an irrational rotation: towards an explicit isomorphism*, Ergodic Theory and dynamical Systems **41**-7, 2110–2135 (2021).
- [11] I. Meilijson, *Mixing properties of a class of skew-products*. Israel J. Math. **19**, 266–270 (1974).
- [12] W. Parry, *Automorphisms of the Bernoulli endomorphism and a class of skew-products*. Ergodic Theory and Dynamical Systems **16**, 519–529 (1996).
- [13] Resnick, S. I., *A probability path*, Reprint of the 2005 Edition, Birkhäuser.
- [14] A.M. Vershik, *Decreasing sequences of measurable partitions, and their applications*. Dokl. Akad. Nauk SSSR, **193**, 748–751 (1970). English translation: Soviet mathematics - Doklady, **11**, 1007–1011 (1970).
- [15] A.M. Vershik, *The theory of decreasing sequences of measurable partitions*. Algebra i Analiz **6**:4, 1–68 (1994) (in Russian). English translation: St. Petersburg Mathematical Journal **6**:4, 705–761 (1995).

**Christophe Leuridan**

Université Grenoble Alpes, CNRS, IF, 38000 Grenoble, FRANCE