

# Zero Krengel entropy does not kill Poisson entropy

Élise Janvresse and Thierry de la Rue

Laboratoire de Mathématiques Raphaël Salem, UMR 6085 CNRS, Université de Rouen, Avenue de l'Université, B.P. 12, F76801  
Saint-Étienne-du-Rouvray Cedex, France. E-mail: [Elise.Janvresse@univ-rouen.fr](mailto:Elise.Janvresse@univ-rouen.fr); [Thierry.de-la-Rue@univ-rouen.fr](mailto:Thierry.de-la-Rue@univ-rouen.fr)

Received 28 June 2010; revised 30 September 2010; accepted 7 October 2010

---

**Abstract.** We prove that the notions of Krengel entropy and Poisson entropy for infinite-measure-preserving transformations do not always coincide: We construct a conservative infinite-measure-preserving transformation with zero Krengel entropy (the induced transformation on a set of measure 1 is the Von Neumann–Kakutani odometer), but whose associated Poisson suspension has positive entropy.

**Résumé.** Nous prouvons que les notions d'entropie de Krengel et d'entropie de Poisson pour les transformations préservant une mesure infinie ne coïncident pas toujours : nous construisons une transformation conservative préservant une mesure infinie qui a une entropie de Krengel nulle (la transformation induite sur un ensemble de mesure 1 est l'odomètre de Von Neumann–Kakutani), mais dont la suspension de Poisson a une entropie strictement positive.

MSC: 37A05; 37A35; 37A40; 28D20

Keywords: Krengel entropy; Poisson suspension; Infinite-measure-preserving transformation;  $\bar{d}$ -distance.

---

## 1. Introduction

### 1.1. Entropy for infinite-measure-preserving transformations

There exist several notions of entropy for infinite transformations, which elegantly generalize Kolmogorov's entropy of a probability-preserving transformation. Krengel [4] comes down to the finite-measure case by considering the entropy of the induced transformation on a set of finite measure: The *Krengel entropy* of a conservative measure-preserving transformation  $(X, \mathcal{B}, \mu, T)$  is defined as

$$h_{\text{Kr}}(X, \mathcal{B}, \mu, T) := \sup_{A \in \mathcal{F}_+} \mu(A)h(A, \mathcal{B} \cap A, \mu_A, T_A),$$

where  $\mathcal{F}_+$  is the collection of sets in  $\mathcal{B}$  with finite positive measure,  $\mu_A$  is the normalized probability measure on  $A$  obtained by restricting  $\mu$  to  $\mathcal{B} \cap A$ , and  $T_A : A \rightarrow A$  is the induced map on  $A$ . Recall that this map is defined by

$$T_A(x) := T^{r_A(x)}(x),$$

where  $r_A(x) := \min\{k \geq 1 : T^k(x) \in A\}$  is the *first-return-time map* associated to  $A$ . As soon as  $T$  is not purely periodic, Krengel proved that

$$h_{\text{Kr}}(X, \mathcal{B}, \mu, T) = \mu(A)h(A, \mathcal{B} \cap A, \mu_A, T_A),$$

where  $A$  is any finite-measure *sweep-out* set (i.e., a set such that  $\bigcup_{n \geq 0} T^{-n}A = X$ ).

The *Parry entropy* of an infinite-measure-preserving transformation  $T$  has been defined in [7] as the supremum of the conditional entropy of  $\mathcal{C}$  with respect to  $T^{-1}\mathcal{C}$ , for all  $\sigma$ -finite sub- $\sigma$ -algebras  $\mathcal{C}$  such that  $T^{-1}\mathcal{C} \subset \mathcal{C}$ .

Recall now that to each infinite-measure-preserving transformation  $T$  we can associate a probability-preserving transformation  $T_*$  called its *Poisson suspension*, and which can be described as follows (we refer to [8] for details): We consider a Poisson process on  $X$  with intensity  $\mu$ , which we can consider as a random collection of particles. These particles are distributed over  $X$  in such a way that, denoting by  $N_B$  the random number of particles in any finite-measure set  $B$ , for any finite collection of pairwise disjoint, finite-measure sets  $B_1, \dots, B_n \in \mathcal{B}$ , the random variables  $N_{B_1}, \dots, N_{B_n}$  are independent, and follow Poisson distributions with respective parameters  $\mu(B_1), \dots, \mu(B_n)$ . Then  $T_*$  is defined on the canonical space of this Poisson process, and it consists in moving individually each of these particle according to the transformation  $T$  on  $X$ . The *Poisson entropy* of an infinite-measure-preserving transformation was defined by Roy [8] as the Kolmogorov entropy of its Poisson suspension.

Relations between these notions of entropy are studied in [3]: On large classes of transformations (e.g., quasi-finite transformations, rank-one transformations), it is proved that Poisson entropy is equal to Krengel entropy and to Parry entropy. Moreover, in any case, Parry entropy is dominated by both Krengel and Poisson entropy.

It was asked in [3] whether, for any conservative measure-preserving transformation, these three definitions always coincide. The purpose of the present paper is to show that the answer is negative, by constructing a counterexample.

**Theorem 1.1.** *There exists a conservative infinite-measure-preserving transformation with zero Krengel entropy (hence zero Parry entropy), but whose associated Poisson suspension has positive entropy.*

## 2. Construction

### 2.1. Von Neumann–Kakutani odometer

The transformation  $T$  is constructed as a tower over the Von Neumann–Kakutani odometer  $S$ . Let us recall the construction of the latter by cutting and stacking (see Fig. 1). We start with the interval  $A := [0, 1]$ . The first step consists in cutting  $A$  into two sub-intervals  $A_1$  and  $A \setminus A_1$  of measure  $1/2$ , and stacking  $A \setminus A_1$  over  $A_1$ . We get a tower of height 2 which we call Tower 1. After step  $n$ ,  $A$  has been cut into  $2^n$  sub-intervals which are stacked to get Tower  $n$ . This means that at this step each point of  $A$  (except those lying on the top of the tower) is mapped by  $S$  to the point of  $A$  lying above it. We construct Tower  $n + 1$  by cutting Tower  $n$  into two equal parts. We call  $A_{n+1}$  the left half of the top interval of Tower  $n$  and we stack the right part of the tower over  $A_{n+1}$ , thus dividing by 2 the measure of the set where  $S$  is not yet defined. Repeating this procedure defines the Von Neumann–Kakutani odometer which preserves the Lebesgue measure on  $A$ . It is well known that the odometer  $S$  is ergodic and has zero entropy.

### 2.2. Construction of $T$

$T$  is constructed on  $\mathbb{R}_+$  in such a way that the induced transformation  $T_A$  coincides with the odometer previously defined.  $T$  is completely defined (up to isomorphism) by giving for each point  $x \in A$  the first return time  $r_A(x)$  to  $A$ .

We fix an increasing sequence of integers  $(M_n)_{n \geq 0}$ , with  $M_n \rightarrow \infty$ . For any  $n \geq 1$ , we choose a large enough integer  $k_n$  (to be precised later). We define the first return time to  $A$  so that its restriction to  $A_n$  is uniformly distributed on  $\{M_n, M_n + 1, \dots, M_n + k_n - 1\}$  for any  $n$ .

We will see in Section 4 that by choosing  $k_n$  large enough, the entropy of  $T_*$  is positive.

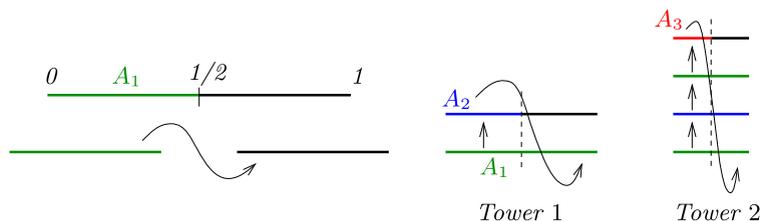


Fig. 1. First steps in the construction of the Von Neumann–Kakutani odometer by cutting and stacking.

### 3. Poisson approximation lemma

The purpose of this section is to prove the key lemma. This lemma roughly states that when  $k_n$  is large enough, it is almost impossible in the Poisson suspension to keep track individually of the particles when they leave  $A_n$  if we only have access to the number of particles in  $A$ . For this, we compare two processes: The first one is simply an i.i.d. sequence of Poisson random variables, whereas the second one modelizes particles leaving  $A_n$  and coming back to  $A$ . The comparison between the two processes uses the notion of  $\bar{d}$ -distance, of which we recall some properties.

#### 3.1. The $\bar{d}$ -distance

The  $\bar{d}$ -distance between two stationary processes has been introduced by Ornstein for the proof of the isomorphism theorem of Bernoulli shifts. We refer to [6] or [9] for the properties of this distance which we use later and which we recall here.

Let  $\xi$  and  $\zeta$  be two stationary processes taking values in a countable alphabet  $\mathbb{B}$ . For any integers  $p < q$ , we denote by  $\xi|_p^q$  the finite sequence  $(\xi_p, \xi_{p+1}, \dots, \xi_q)$ .

For  $L \geq 1$ , let  $J_L(\xi, \zeta)$  be the set of all joinings of  $\xi|_1^L$  and  $\zeta|_1^L$ , that is probability distributions on  $\mathbb{B}^L \times \mathbb{B}^L$  whose marginals are the distributions of  $\xi|_1^L$  and  $\zeta|_1^L$ . We first define  $\bar{d}_L$  for any  $L \geq 1$ , by

$$\bar{d}_L(\xi, \zeta) := \min_{\lambda \in J_L(\xi, \zeta)} \mathbb{E}_\lambda[d_L(\xi|_1^L, \zeta|_1^L)],$$

where  $d_L$  is the Hamming distance between sequences of length  $L$ :

$$d_L(x_1 \cdots x_L, z_1 \cdots z_L) := \frac{1}{L} \sum_1^L \mathbb{1}_{x_i \neq z_i}.$$

Then, the  $\bar{d}$ -distance between  $\xi$  and  $\zeta$  is defined by

$$\bar{d}(\xi, \zeta) := \sup_{L \geq 1} \bar{d}_L(\xi, \zeta).$$

It can be shown that  $\bar{d}(\xi, \zeta)$  is also the minimum of  $\lambda(\xi_0 \neq \zeta_0)$  when  $\lambda$  ranges over all stationary joinings of  $\xi$  and  $\zeta$ .

The two key properties of the  $\bar{d}$ -distance that we shall use are: On the one hand, the fact that entropy of processes close in  $\bar{d}$ -distance can be compared (Lemma 3.1). On the other hand, a practical tool to estimate the  $\bar{d}$ -distance between processes using conditional distributions on the past: If, for all large enough  $n$ ,

$$\sum_{b \in \mathbb{B}} |\mathbb{P}(\xi_0 = b | \xi|_{-n}^{-1}) - \mathbb{P}(\zeta_0 = b | \zeta|_{-n}^{-1})| < \varepsilon \tag{1}$$

for all past  $\xi|_{-n}^{-1}$  outside a set of measure  $\varepsilon$  and all past  $\zeta|_{-n}^{-1}$  outside a set of measure  $\varepsilon$ , then  $\bar{d}(\xi, \zeta) \leq 3\varepsilon$ . Moreover, the same conclusion holds if we replace in (1) the conditional distributions with respect to the past by conditional distributions with respect to finer  $\sigma$ -algebras.

The fact that entropy is a continuous function of processes taking values in a given finite alphabet, when these processes are topologized with the  $\bar{d}$ -distance, is well known (see, e.g., [9], p. 100). This is however no longer true if we consider processes taking countably many values. In this case, entropy is only a lower semicontinuous function of the process.

**Lemma 3.1.** *Let  $\xi$  be a stationary process taking values in a countable alphabet  $\mathbb{B}$  which has finite entropy. For any  $\varepsilon > 0$ , there exists  $\delta > 0$  such that any stationary process  $\zeta$  taking values in the same alphabet  $\mathbb{B}$  with  $\bar{d}(\xi, \zeta) < \delta$  satisfies  $h(\zeta) > h(\xi) - \varepsilon$ .*

**Proof.** Without loss of generality we can assume  $\mathbb{B} = \mathbb{Z}^+$ . For any integer  $\ell > 0$  we define the process  $(\xi \wedge \ell)$  taking values in the finite alphabet  $\{0, 1, \dots, \ell\}$  by

$$(\xi \wedge \ell)_x = \begin{cases} \xi_x & \text{if } \xi_x < \ell, \\ \ell & \text{otherwise.} \end{cases}$$

The  $\sigma$ -algebra generated by the process  $(\xi \wedge \ell)$  increases to the  $\sigma$ -algebra generated by  $\xi$ . Hence we can choose  $\ell$  large enough so that  $h(\xi \wedge \ell) > h(\xi) - \varepsilon/2$ . By continuity of the entropy in the case of a finite alphabet, we can find  $\delta > 0$  such that any process taking values in  $\{0, 1, \dots, \ell\}$  at  $\bar{d}$ -distance at most  $\delta$  from  $(\xi \wedge \ell)$  has entropy at least  $h(\xi) - \varepsilon$ . Now, if  $\bar{d}(\xi, \zeta) < \delta$ , then  $\bar{d}((\xi \wedge \ell), (\zeta \wedge \ell)) < \delta$  (where  $(\zeta \wedge \ell)$  is defined from  $\zeta$  in a similar way), hence  $h(\zeta) \geq h(\zeta \wedge \ell) \geq h(\xi) - \varepsilon$ .  $\square$

### 3.2. Comparison between connected and disconnected processes

Let  $\mathbb{A}$  be a finite alphabet,  $P^B$  and  $P^W$  be two probability measures on  $\mathbb{A}$  and  $\lambda_{\mathbb{A} \times \mathbb{A}}$  be a joining of  $P^B$  and  $P^W$ . Let  $\delta$  be some fixed positive real number. We define two processes  $\xi$  (the *disconnected process*) and  $\zeta$  (the *connected process*) on  $(\mathbb{N}^{\mathbb{A}})^{\mathbb{Z}}$ , and estimate the  $\bar{d}$ -distance between them.

The process  $\xi$  is constructed from two independent sequences of i.i.d. random variables distributed according to the Poisson distribution of parameter  $\delta/2$ , which can be interpreted as numbers of black and white particles lying on each site of  $\mathbb{Z}$ . Then to each black (resp., white) particle we randomly and independently associate a label picked in  $\mathbb{A}$  according to  $P^B$  (resp.,  $P^W$ ). For any  $x \in \mathbb{Z}$  and any labels  $a, b \in \mathbb{A}$ ,  $\xi_x^B(a)$  (resp.,  $\xi_x^W(b)$ ) is the total number of black (resp., white) particles labelled by  $a$  (resp.,  $b$ ) at position  $x$ . In other words, the process  $\xi$  associates in an i.i.d. way to each site  $x \in \mathbb{Z}$  a finite sequence  $\xi_x = (\xi_x^B(a), \xi_x^W(b))_{a,b \in \mathbb{A}}$  of independent random variables respectively distributed according to the Poisson distribution of parameter  $P^B(a)\delta/2$  and  $P^W(b)\delta/2$ . The process  $\xi$  is called the disconnected process because its two components  $\xi^B$  and  $\xi^W$  are independent.

Let  $k$  and  $M$  be two integers. The process  $\zeta$  is also constructed from black and white particles on  $\mathbb{Z}$ , but which are no longer independent. This is why we call it the connected process. The number of black particles at each site is given by a sequence of i.i.d. random variables distributed according to the Poisson distribution of parameter  $\delta/2$ . For each black particle at position  $x \in \mathbb{Z}$ , we first pick a random integer  $j$ , uniformly in  $\{M, M+1, \dots, M+k-1\}$  and independently of all other particles. Then we link the black particle to a white particle that we put at position  $x+j$ . For each such couple of black and white particles, a couple of labels is picked in  $\mathbb{A} \times \mathbb{A}$  with probability  $\lambda_{\mathbb{A} \times \mathbb{A}}$ : The first label is associated to the black particle and the second label to the white one. Then, for any  $x \in \mathbb{Z}$ ,  $\zeta_x = (\zeta_x^B(a), \zeta_x^W(b))_{a,b \in \mathbb{A}}$  denote the number of black particle labelled by  $a$  and the number of white particles labelled by  $b$  at position  $x$ .

**Lemma 3.2.** *For any  $\varepsilon > 0$ , if  $k$  is large enough,  $\bar{d}(\xi, \zeta) < \varepsilon$ .*

**Proof (Simple case).** We first prove the lemma in the case where  $\mathbb{A}$  is reduced to a singleton. Since all particles have the same label, we just forget it and simply count the number of black and white particles on each site.

We now prove that if  $k$  is large enough, for any  $n \geq M+k$ ,

$$\sum_{j, \ell \in \mathbb{N}} \left| P(\zeta_0^B = j, \zeta_0^W = \ell | \zeta|_{[-n]}^{-1}) - \frac{e^{-\delta/2}(\delta/2)^j}{j!} \frac{e^{-\delta/2}(\delta/2)^\ell}{\ell!} \right| < \varepsilon,$$

with probability  $1 - \varepsilon$  on  $\zeta|_{[-n]}^{-1}$ . Since  $\zeta_0^B$  is Poisson distributed with parameter  $\delta/2$ , and independent from  $\zeta|_{[-n]}^{-1}$ , and since  $\zeta_0^W$  is independent from  $\zeta_0^B$  conditionally to the past, it is enough to prove that if  $k$  is large enough, for any  $n \geq M+k$ ,

$$\sum_{\ell \in \mathbb{N}} \left| P(\zeta_0^W = \ell | \zeta|_{[-n]}^{-1}) - \frac{e^{-\delta/2}(\delta/2)^\ell}{\ell!} \right| < \varepsilon,$$

with probability  $1 - \varepsilon$  on  $\zeta|_{[-n]}^{-1}$ .

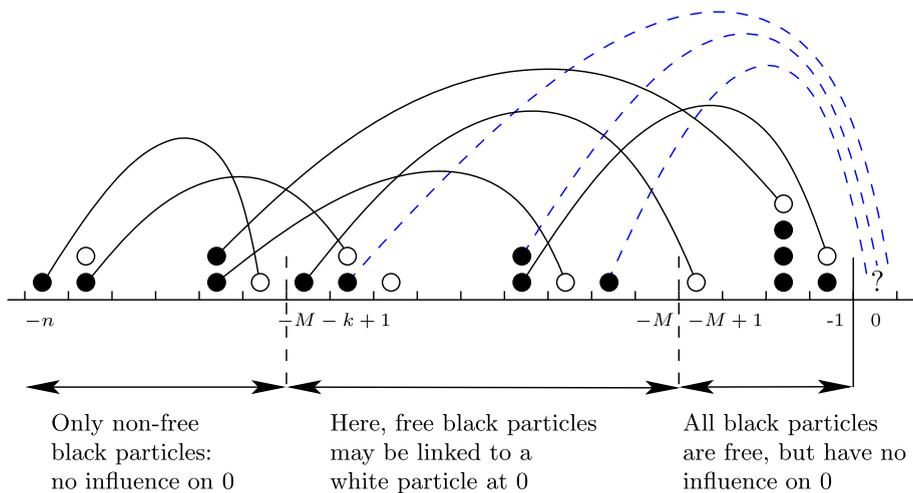


Fig. 2. The enriched past.

In fact we rather condition with respect to an enriched past (see Fig. 2): Formally, the *enriched past* is the  $\sigma$ -algebra generated by the number of black and white particles on each site  $x \in \{-n, \dots, -1\}$ , and the knowledge of all links existing between a black particle and a white particle both lying in  $\{-n, \dots, -1\}$ .

Conditionally to the enriched past, we can distinguish two kinds of black particles between  $-n$  and  $-1$ : Those which are linked to a white particle lying on the left of 0, and those, called *free particles*, whose white particle's position is unknown. Observe that only free particles may have some influence on  $\zeta_0^W$ . Hence, black particles lying on the left of site  $-(M+k-1)$  have no influence on  $\zeta_0^W$  since they are not free. Black particles lying on the right of site  $-M$  are free but nevertheless have no influence on  $\zeta_0^W$  since a black particle is linked to a white particle at distance at least  $M$ .

So it remains to study the influence on  $\zeta_0^W$  of free black particles at sites between  $-(M+k-1)$  and  $-M$ . For  $1 \leq j \leq k$  let us denote by  $F_j$  the number of free particles at site  $-(M+k)+j$ . Any black particle at site  $-(M+k)+j$  has probability  $j/k$  to be free. Therefore,  $F_j$  follows the Poisson distribution with parameter  $\frac{\delta j}{2k}$ .

Fix  $1 \leq j \leq k$ . Assume there is a free particle at site  $-(M+k)+j$ . Since there are  $j$  possible positions for its white particle, the latter has probability  $1/j$  to lie at site 0. Hence, conditionally to our enriched past, the number of white particles at site 0 can be written as

$$\sum_{j=1}^k \sum_{\ell=1}^{F_j} B_\ell^j,$$

where  $(B_\ell^j)$  are independent Bernoulli random variables with respective parameter  $1/j$ . The law of such a sum of independent Bernoulli variables is close to a Poisson distribution of parameter  $\delta/2$  as soon as the sum of the parameters is close to  $\delta/2$  and all parameters are small enough (see [2], Theorem 23.2, p. 312). Therefore we can choose a large enough integer  $J$ , and  $\varepsilon_1$  small enough so that, if  $Z$  is a sum of independent Bernoulli random variables, each with parameter less than  $1/J$  and such that the sum of parameters is within  $\varepsilon_1$  of  $\delta/2$ , then

$$\sum_{\ell \geq 0} \left| \mathbb{P}(Z = \ell) - \exp(-\delta/2) \frac{(\delta/2)^\ell}{\ell!} \right| < \varepsilon.$$

We have now to avoid bad configurations, that is configurations of the enriched past which have free particles close to  $-(M+k-1)$  giving rise to Bernoulli with large parameters, and configurations such that the sum of the parameters is not close enough to  $\delta/2$ .

*Control of the parameters' size.* We compute the probability that no free particle lie between  $-(M+k)+1$  and  $-(M+k)+J$ :

$$\mathbb{P}(F_j = 0, j = 1, \dots, J) = \prod_{j=1}^J \exp\left(-\frac{\delta j}{2k}\right) = \exp\left(-\frac{\delta J(J+1)}{4k}\right).$$

Under this condition,  $\zeta_0^W$  is (conditionally to the enriched past) the sum of independent Bernoulli variables with parameters smaller than  $1/J$ . If  $k$  is large enough, this happens with probability larger than  $1 - \varepsilon/2$ .

*Control of the parameters' sum.* Since the  $F_j$  are independent and Poisson distributed with parameter  $\frac{\delta j}{2k}$ , the expected value of the sum  $S := \sum_{j=1}^k j^{-1} F_j$  of the parameters is  $\delta/2$ , and its variance is

$$\text{var } S = \sum_{j=1}^k \frac{1}{j^2} \frac{\delta j}{2k} = \frac{\delta}{2k} \sum_{j=1}^k \frac{1}{j}.$$

Hence, if  $k$  is large enough,

$$\mathbb{P}(|S - \delta/2| < \varepsilon_1) > 1 - \varepsilon/2.$$

Putting things together, we have proved that with probability larger than  $1 - \varepsilon$  on the enriched past, the conditional distribution of the number  $\zeta_0^W$  of white particles at site 0 satisfies

$$\sum_{\ell \geq 0} \left| \mathbb{P}(\zeta_0^W = \ell | \text{enriched past}) - \exp(-\delta/2) \frac{(\delta/2)^\ell}{\ell!} \right| < \varepsilon.$$

This proves Lemma 3.2 when  $\mathbb{A}$  is reduced to a singleton. □

**Proof (General case).** We consider the family of independent processes  $\zeta^{a,b} = (\zeta^{B,a,b}, \zeta^{W,a,b})$ ,  $(a, b) \in \mathbb{A} \times \mathbb{A}$ , which counts the number of black and white  $\zeta$  particles at  $x$  belonging to a pair of black and white particles respectively labelled by  $a$  and  $b$ . Then  $\zeta^{a,b}$  is a simple-case  $\zeta$  process, for which the expected number of black particles per site is  $\delta \lambda_{\mathbb{A} \times \mathbb{A}}(a, b)/2$ . From the proof in the simple case, we know that as soon as  $k$  is large enough, the  $\bar{d}$ -distance between  $\zeta^{a,b}$  and  $\xi^{a,b}$  is smaller than  $\varepsilon/|\mathbb{A}|^2$ , where  $\xi^{a,b} = (\xi^{B,a,b}, \xi^{W,a,b})$  is composed of two i.i.d. sequences of Poisson random variables of parameter  $\delta \lambda_{\mathbb{A} \times \mathbb{A}}(a, b)/2$ . We can recover  $\zeta_x^B(a)$  and  $\zeta_x^W(b)$  by

$$\zeta_x^B(a) = \sum_{b \in \mathbb{A}} \zeta_x^{B,a,b} \quad \text{and} \quad \zeta_x^W(b) = \sum_{a \in \mathbb{A}} \zeta_x^{W,a,b}.$$

On the other hand,  $\sum_{b \in \mathbb{A}} \xi^{B,a,b}$  (resp.,  $\sum_{a \in \mathbb{A}} \xi^{W,a,b}$ ) has the same distribution as  $\xi^B(a)$  (resp.,  $\xi^W(b)$ ): It is an i.i.d. sequence of Poisson random variables of parameter  $P^B(a)\delta/2$  (resp.,  $P^W(b)\delta/2$ ). Summing over  $a$  and  $b$ , it follows that  $\bar{d}(\zeta, \xi) < \varepsilon$ . □

#### 4. Positive Poisson entropy

We denote by  $\xi^{(\infty)}$  the stationary process living in the Poisson suspension of our transformation  $T$ , defined by

$$\xi_x^{(\infty)} := \text{number of particles in } A \text{ at time } x.$$

The purpose of this section is to show that the entropy of the process  $\xi^{(\infty)}$  is positive as soon as the  $k_n$ 's are chosen large enough. This will be proved by showing that the  $\bar{d}$ -distance between  $\xi^{(\infty)}$  and an i.i.d. sequence  $\xi^{(0)}$  of random Poisson variables with parameter 1 can be made as small as we want. By Lemma 3.1, this will be enough to conclude.

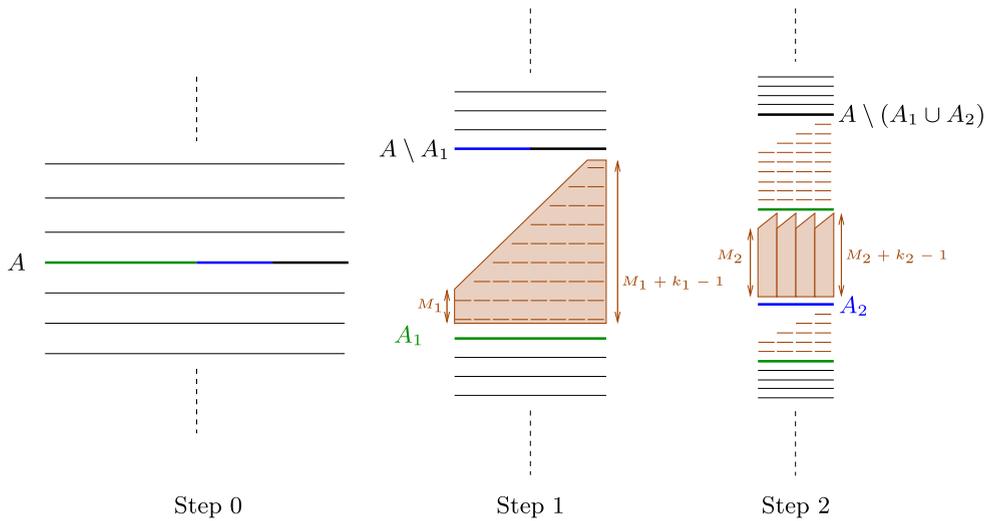


Fig. 3. First steps in the construction of the return time to  $A$ .

Our strategy is the following: As one goes along in the construction of the return time to  $A$ , we define a sequence  $(T^{(n)})$  of infinite-measure-preserving transformations acting on  $\mathbb{R}_+$ . Then we consider the process  $\xi^{(n)}$  living in the Poisson suspension of  $T^{(n)}$ :

$$\xi_x^{(n)} := \text{number of particles in } A \text{ at time } x \text{ for the Poisson suspension over } T^{(n)}.$$

The transformations  $T^{(n)}$  will be constructed so that the process  $\xi^{(n)}$  converge in  $\bar{d}$ -distance to  $\xi$ .

The transformation  $T^{(0)}$  is constructed by stacking infinitely many pairwise disjoint intervals of length 1,  $A$  being one of them, into a doubly infinite tower, each interval being mapped onto the one just above. Therefore, as mentioned previously,  $\xi^{(0)}$  is an i.i.d. sequence of Poisson variables with parameter 1. At step  $n$  of the construction, we define the return time to  $A$  on the subset  $A_n$ . This return time to  $A$  will be the same for all transformations  $T^{(m)}$ ,  $m \geq n$ , and for the final transformation  $T$ . By choosing the return time adequately, we will make sure that

$$\bar{d}(\xi^{(n-1)}, \xi^{(n)}) < 2^{-n} \varepsilon,$$

so that for all  $n$

$$\bar{d}(\xi^{(0)}, \xi^{(n)}) < \varepsilon. \tag{2}$$

Let us describe the first step. Recall that Tower 1 is of height 2, with basis  $A_1$  (see Fig. 1). We cut  $A_1$  into  $k_1$  equal subintervals, and we define on  $A_1$  the return time to  $A$  to be  $M_1 + j - 1$  on the  $j$ th subinterval. We insert Tower 1 into a doubly infinite tower of intervals of length  $1/2$  and add spacers between  $A_1$  and its image by the odometer  $S$ : We insert  $M_1 + j - 2$  spacers of width  $1/(2k_1)$  between the  $j$ th subinterval of  $A_1$  and its image by  $S$ . The transformation  $T^{(1)}$  is defined by mapping each point to the point right above it (see Fig. 3).

The process  $\xi^{(1)}$  counts the number of particles in  $A$  at time  $x$  for the Poisson suspension over  $T^{(1)}$ : We interpret points in  $A_1$  as black particles and points in  $A \setminus A_1$  as white particles. For the suspension over  $T^{(0)}$ , black and white particles are independent, whereas for the suspension over  $T^{(1)}$ , they are linked through the return time to  $A$  of the point corresponding to the black particle. Hence, a direct application of Lemma 3.2 in the simple case with no alphabet gives that  $\bar{d}(\xi^{(0)}, \xi^{(1)}) < \varepsilon/2$  when  $k_1$  is large enough.

Suppose the return time to  $A$  has already been defined on all  $A_i$ ,  $1 \leq i \leq n - 1$ . Consider Tower  $n$ . The return time to  $A$  has already been defined on all rungs but the roof and  $A_n$  (which is the rung of level  $2^{n-1}$ ). Let  $R_{n-1}$  be the maximum value of the already defined return time to  $A$  on  $A_1 \cup \dots \cup A_{n-1}$ . We consider the finite alphabet  $\mathbb{A} := \{1, \dots, R_{n-1}\}^{2^{n-1}-1}$ . For each point  $y$  in  $A_n$ , let  $h^B(y) \in \mathbb{A}$  be the sequence of the return times to  $A$  when we climb the first half of Tower  $n$  before reaching  $y$ , in other words  $h^B(y)$  is the sequence of the last  $2^{n-1} - 1$  return

times to  $A$ . Let  $h^W(y) \in \mathbb{A}$  be the sequence of the return times to  $A$  when we climb the second half of Tower  $n$  starting from  $Sy$ . We denote by  $r(y)$  the return time to  $A$ , which is to be defined at this step. We want  $r(y)$  to be independent of  $h^B(y)$  and  $h^W(y)$ . To this end, we consider the finite partition of  $A_n$  generated by  $h^B$  and  $h^W$ . Each atom of this partition is cut into  $k_n$  equal pieces, and we define  $r$  to be  $M_n + j - 1$  on the  $j$ th piece of each atom. Here is how we define the transformation  $T^{(n)}$ : we insert Tower  $n$  into a doubly infinite tower of intervals of length  $1/2^n$  and insert as many spacers as we need between the rungs of Tower  $n$  to achieve the already defined return time to  $A$ . The transformation  $T^{(n)}$  maps each point to the point right above it.

Let us turn to the estimation of  $\bar{d}(\xi^{(n-1)}, \xi^{(n)})$  for  $n \geq 2$ . We want to apply Lemma 3.2: Black particles are points in  $A_n$  and white particles are points in  $SA_n$ . To each point  $y$  in  $A_n$ , we associate the label  $h^B(y) \in \mathbb{A}$ , and to each point  $Sy$  in  $SA_n$ , we attach the label  $h^W(y) \in \mathbb{A}$ . Let the process  $\xi$  (resp.,  $\zeta$ ) count the numbers of black and white particles together with their label in the suspension over  $T^{(n-1)}$  (resp.,  $T^{(n)}$ ). These processes are exactly of the form studied in Lemma 3.2.

Now, observe that we can recover  $\xi^{(n-1)}$  and  $\xi^{(n)}$  from  $\xi$  and  $\zeta$ :

$$\xi_0^{(n-1)} = \sum_{h \in \mathbb{A}} \xi_0^B(h) + \xi_{h_1}^B(h) + \dots + \xi_{h_1+\dots+h_{2^{n-1}-1}}^B(h) + \sum_{h \in \mathbb{A}} \xi_0^W(h) + \xi_{-h_1}^W(h) + \dots + \xi_{-(h_1+\dots+h_{2^{n-1}-1})}^W(h)$$

and

$$\xi_0^{(n)} = \sum_{h \in \mathbb{A}} \zeta_0^B(h) + \zeta_{h_1}^B(h) + \dots + \zeta_{h_1+\dots+h_{2^{n-1}-1}}^B(h) + \sum_{h \in \mathbb{A}} \zeta_0^W(h) + \zeta_{-h_1}^W(h) + \dots + \zeta_{-(h_1+\dots+h_{2^{n-1}-1})}^W(h).$$

It follows that, if  $\xi$  and  $\zeta$  coincide on  $\{-R_{n-1}(2^{n-1} - 1), \dots, R_{n-1}(2^{n-1} - 1)\}$ , then  $\xi_0^{(n-1)} = \xi_0^{(n)}$ . Let  $\lambda$  be a stationary joining of  $\xi$  and  $\zeta$  achieving the  $\bar{d}$ -distance: We have  $\bar{d}(\xi, \zeta) = \lambda(\xi_0 \neq \zeta_0)$ . Then

$$\bar{d}(\xi^{(n-1)}, \xi^{(n)}) \leq \lambda(\xi_0^{(n-1)} \neq \xi_0^{(n)}) \leq 2^n R_{n-1} \lambda(\xi_0 \neq \zeta_0) = 2^n R_{n-1} \bar{d}(\xi, \zeta).$$

By Lemma 3.2,  $\bar{d}(\xi, \zeta)$  can be made arbitrarily small by choosing  $k_n$  large enough. Hence we can assure that  $\bar{d}(\xi^{(n-1)}, \xi^{(n)}) < 2^{-n} \varepsilon$ .

Finally, note that since  $M_n$  is increasing, the return time to  $A$  on the roof of Tower  $n$  (the union of  $A_i$ ,  $i > n$ ) will be larger than  $M_{n+1}$ . Hence, for any  $L > 0$ , if  $n$  is large enough so that  $M_{n+1} > L$ , the distribution of  $\xi^{(n)}|_0^{L-1}$  coincides with the distribution of  $\xi^{(\infty)}|_0^{L-1}$ . Therefore,

$$\bar{d}_L(\xi^{(\infty)}, \xi^{(0)}) = \bar{d}_L(\xi^{(n)}, \xi^{(0)}) \leq \bar{d}(\xi^{(n)}, \xi^{(0)}) < \varepsilon,$$

which implies

$$\bar{d}(\xi^{(\infty)}, \xi^{(0)}) \leq \varepsilon.$$

### 5. Comments and open questions

In view of previously known results on the subject, some comments on the infinite-measure-preserving transformation  $T$  constructed in Section 2 may be made.

First, although its construction is derived from the standard cutting-and-stacking procedure used to build the most elementary rank-one system (the Von Neumann–Kakutani odometer), the transformation  $T$  is not even of finite rank. Indeed, Proposition 10.1 in [3] shows that for finite-rank systems, both the Poisson and Krengel entropy vanish.

Second, it was also proved in [3] that Krengel and Poisson entropies coincide for *quasi-finite* transformations, namely transformations for which there exists a sweep-out set  $A$  of measure 1 such that the return-time partition of  $A$  has finite entropy. There exist only few examples of transformation for which the non quasi-finiteness has been established: an unpublished example constructed by Ornstein has been mentioned by Krengel in [5], and the only published example which we are aware of is a rank-one system, given by Aaronson and Park in [1]. Our construction thus provides a new example of a non-quasi-finite transformation.

After having proved that the different notions of entropy for infinite-measure-preserving transformations do not always coincide, a natural question is to ask whether they are always ordered in the same way: Is it true that Poisson entropy always dominates Krengel entropy? Can we at least decide whether zero Poisson entropy implies zero Krengel entropy? And what about similar questions regarding the comparison between Parry entropy and Poisson entropy? It may be worth recalling here that the equality of Parry entropy and Krengel entropy in the quasi-finite case was proved by Parry in 1969 [7], but that the question whether they always coincide is, as far as we know, still open.

## Acknowledgements

The construction of the transformation  $T$  in Section 2 has been inspired by a private communication of Benjamin Weiss concerning the Shannon–McMillan–Breiman theorem.

We are also much indebted to Emmanuel Roy for stimulating conversations on the subject.

## References

- [1] J. Aaronson and K. K. Park. Predictability, entropy and information of infinite transformations. *Fund. Math.* **206** (2009) 1–21. [MR2576257](#)
- [2] P. Billingsley. *Probability and Measure*, 2nd edition. Wiley, New York, 1986. [MR0830424](#)
- [3] É. Janvresse, T. Meyerovitch, E. Roy and T. de la Rue. Poisson suspensions and entropy for infinite transformations. *Trans. Amer. Math. Soc.* **362** (2010) 3069–3094. [MR2592946](#)
- [4] U. Krengel. Entropy of conservative transformations. *Z. Wahrsch. Verw. Gebiete* **7** (1967) 161–181. [MR0218522](#)
- [5] U. Krengel. On certain analogous difficulties in the investigation of flows in a probability space and of transformations in an infinite measure space. In *Functional Analysis (Proc. Sympos., Monterey, CA, 1969)* 75–91. Academic Press, New York, 1969. [MR0265558](#)
- [6] D. S. Ornstein. *Ergodic Theory, Randomness, and Dynamical Systems*. Yale Univ. Press, New Haven, CT, 1974. [MR0447525](#)
- [7] W. Parry. *Entropy and Generators in Ergodic Theory*. W. A. Benjamin, New York, 1969. [MR0262464](#)
- [8] E. Roy. Mesures de poisson, infinie divisibilité et propriétés ergodiques. Ph.D. thesis, 2005.
- [9] P. C. Shields. *The Ergodic Theory of Discrete Sample Paths. Graduate Studies in Mathematics* **13**. Amer. Math. Soc. Providence, RI, 1996. [MR1400225](#)