2-FOLD AND 3-FOLD MIXING: WHY 3-DOT-TYPE COUNTEREXAMPLES ARE IMPOSSIBLE IN ONE DIMENSION

THIERRY DE LA RUE

ABSTRACT. V.A. Rohlin asked in 1949 whether 2-fold mixing implies 3-fold mixing for a stationary process $(\xi_i)_{i \in \mathbb{Z}}$, and the question remains open today. In 1978, F. Ledrappier exhibited a counterexample to the 2-fold mixing implies 3-fold mixing problem, the so-called *3-dot system*, but in the context of stationary random fields indexed by $\mathbb{Z}^2$.

In this work, we first present an attempt to adapt Ledrappier's construction to the one-dimensional case, which finally leads to a stationary process which is 2-fold but not 3-fold mixing *conditionally to the σ-algebra generated by some factor process*. Then, using arguments coming from the theory of joinings, we will give some strong obstacles proving that Ledrappier's counterexample can not be fully adapted to one-dimensional stationary processes.

1. INTRODUCTION: ROHLIN'S MULTIFOLD MIXING PROBLEM AND LEDRAPPIER'S TWO-DIMENSIONAL COUNTEREXAMPLE

The following work is based on two recent results concerning Rohlin's multifold mixing problem which are contained in [17] and [19]. It seemed to me interesting to put these results together and show them in a different light, emphasizing mainly on the underlying ideas rather than on technical details.

The object of our study is a stochastic process, that is to say a family $\xi = (\xi_i)_{i \in \mathbb{Z}}$ of random variables indexed by the set of integers, and we will always assume that these random variables take their values in a finite alphabet $\mathbb{A}$. If two integers $i \leq j$ are given, we will denote by ξ_i^j the finite sequence $(\xi_i, \xi_{i+1}, \dots, \xi_j)$. Obvious generalization of this notation to the case where $i = -\infty$ or $j = +\infty$ will also be used.

We are more particularly interested in the case where the stochastic process is *stationary*, which means that the probability of observing a given cylindrical event E (*i.e.* an event depending only on finitely many coordinates) at any position $i \in \mathbb{Z}$ does not depend on i:

$$\text{(1)} \qquad \forall \ell \geq 0, \ \forall E \subset \mathbb{A}^{\ell+1}, \ \forall i \in \mathbb{Z}, \ \mathbb{P}\left(\xi_i^{i+\ell} \in E \right) \ = \ \mathbb{P}\left(\xi_0^\ell \in E \right).$$

Another way to characterize the stationarity of the process is to say that its distribution is invariant by the coordinate shift: Let $T : \mathbb{A}^{\mathbb{Z}} \to \mathbb{A}^{\mathbb{Z}}$ be the transformation defined by $T(\xi) = \tilde{\xi}$, where for all $i \in \mathbb{Z}$, $\tilde{\xi}_i := \xi_{i+1}$. Then the stochastic process ξ is stationary if and only if the distribution of $T(\xi)$ is the same as the distribution of ξ.

The stochastic process ξ is said to be *mixing* if, considering two windows of arbitrarily large size ℓ, what happens in one window is asymptotically independent of what happens in the second window when the distance between them tends to infinity:

$$\text{(2)} \quad \forall \ell \geq 0, \ \forall E_1, E_2 \subset \mathbb{A}^{\ell+1}, \ \mathbb{P}\left(\xi_0^\ell \in E_1, \xi_p^{p+\ell} \in E_2 \right) - \mathbb{P}\left(\xi_0^\ell \in E_1 \right) \mathbb{P}\left(\xi_p^{p+\ell} \in E_2 \right) \xrightarrow[p \to \infty]{} 0.$$

1.1. **Rohlin's question.** In 1949, V.A. Rohlin [14] proposed a strengthening of the previous definition involving more than two windows: ξ is said to be *3-fold mixing* if

$$\text{(3)} \quad \forall \ell \geq 0, \ \forall E_1, E_2, E_3 \subset \mathbb{A}^{\ell+1},$$

$$\mathbb{P}\left(\xi_0^\ell \in E_1, \xi_p^{p+\ell} \in E_2, \xi_{p+q}^{p+q+\ell} \in E_3 \right) - \mathbb{P}\left(\xi_0^\ell \in E_1 \right) \mathbb{P}\left(\xi_p^{p+\ell} \in E_2 \right) \mathbb{P}\left(\xi_{p+q}^{p+q+\ell} \in E_3 \right) \xrightarrow[p,q \to \infty]{} 0.$$

A straightforward generalization to k windows naturally gives rise to the property of being *k-fold mixing*. To avoid any confusion, we will henceforth call the classical mixing property defined by (2): *2-fold mixing*[1].

Rohlin asked whether any stationary process which is 2-fold mixing is also 3-fold mixing. This question is still open today, but a large number of mathematical works have been devoted to the subject. Many of these works show that 2-fold mixing implies 3-fold mixing for special classes of stationary processes (see *e.g.* [13] and [22] for Gaussian processes, [7] for processes with singular spectrum, [9] and [20] for finite-rank processes).

1.2. **Ledrappier's counterexample in 2 dimensions: the 3-dot system.**
In the opposite direction, Ledrappier [12] produced in 1978 a counterexample showing that in the case of stationary processes indexed by $\mathbb{Z}^2$ (we should rather speak of stationary *random fields* in this context), 2-fold mixing does not necessarily imply 3-fold mixing. Here is a description of his example: Consider

$$G := \left\{ (\xi_{i,j}) \in \{0,1\}^{\mathbb{Z}^2} : \forall (i,j),\ \xi_{i,j} + \xi_{i+1,j} + \xi_{i,j+1} = 0 \mod 2 \right\}.$$

Let us describe a probability law μ on G by the way we pick a random element in G: First, use independent unbiased coin tosses to choose the $\xi_{i,0}$ on the horizontal axis (one coin toss for each $i \in \mathbb{Z}$: these random variables are independent). Now, note that the "3-dot rule" $\xi_{i,j} + \xi_{i+1,j} + \xi_{i,j+1} = 0 \mod 2$ for each (i,j) completely determines the coordinates $\xi_{i,j}$ on the upper-half plane $j \geq 0$. It remains to choose the $\xi_{i,j}$ for $j < 0$. For this, observe that we have yet no constraint on $\xi_{0,-1}$. We choose it with an unbiased coin toss, and then the entire line $\xi_{-1,j}$ is completely determined by the 3-dot rule. To complete the whole plane, we just have to pick each of the $\xi_{0,j}$ $(j < -1)$ with a coin toss, and then inductively fill each horizontal line with the 3-dot rule.

0	0	1	0	1	1	1	0	1
1	1	1	0	0	1	0	1	1
1	0	1	0	0	0	1	1	0
1	0	0	1.	1	1	1	0	1
1	0	0	0.	1.	0	1	0	0
0	1	1	1	1	0	0	1	1
0	0	1	0	1	0	0	0	1
1	1	1	0	0	1	1	1	1

FIGURE 1. Generation of a random configuration in G. First, use independent coin tosses to choose the values of the shaded cells, then apply the 3-dot rule to complete the others: Three adjacent cells disposed as the three dotted ones must contain an even number of 1's.

The addition mod 2 on each coordinate turns G into a compact Abelian group. We let the reader check that the probability law μ defined above on G is invariant by addition of an arbitrary element of G, thus μ is the unique normalized Haar measure on G. Since any shift of coordinates in $\mathbb{Z}^2$ is an automorphism of the group G, such a shift preserves μ. Hence μ turns $(\xi_{i,j})$ into a stationary random field.

The definition of k-fold mixing for a stationary random field is formally the same as in the case of processes, except that a window is no longer an interval on the line but a square in the plane: $\{(i_0 + i, j_0 + j) : 0 \leq i, j \leq \ell\}$ for some $(i_0, j_0) \in \mathbb{Z}^2$ and some $\ell \geq 0$. Let us sketch a geometric argument showing why the 2-fold mixing property holds for $(\xi_{i,j})$. Starting with the cells on the

---

[1] We must point out that in Rohlin's article, the definition of k-fold mixing originally involved $k + 1$ windows, thus the classical mixing property was called *1-fold mixing*. However it seems that the convention we adopt here is used by most authors, and we find it more coherent when translated in the language of multifold self-joinings (see section 3.1).

horizontal axis and the lower-half vertical axis filled with independent coin tosses , we observe that, when filling the other cells using the 3-dot rule,

- the region $R_1 := \{(i,j) : i < 0, \ 0 < j < -i\}$ only depends on the cells $(i,0)$, for $i < 0$;
- the region $R_2 := \{(i,j) : j < 0, \ 0 < i < -j\}$ only depends on the cells $(0,j)$, for $j < 0$;
- the region $R_3 := \{(i,j) : 0 < i, \ 0 < j\}$ only depends on the cells $(i,0)$, for $i \geq 0$. (See Figure 2.)

These three regions are therefore independent. Now, if we take two windows of size ℓ, and if the distance between them is large enough ("large enough" depending on ℓ), it is always possible to shift the coordinates in such a way that each of the shifted windows entirely lies in one of these three regions, and not both in the same region. The two shifted windows are then independent, and since μ is preserved by coordinate shift, this means that the two windows we started with are also independent.
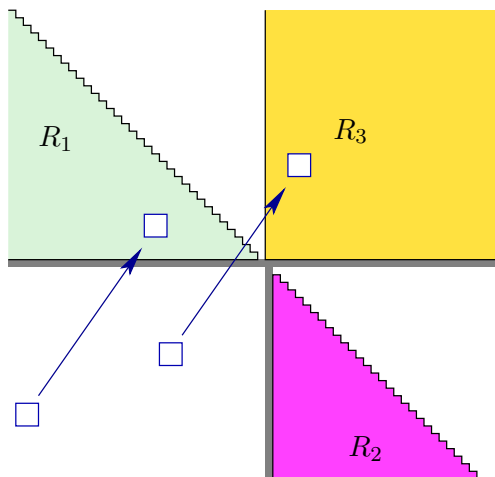

FIGURE 2. 2-fold mixing for the 3-dot system: If the distance between them is large enough, the two square windows can be shifted in such a way that one lies in one of the three colored regions, and the other one in another, independent, region.

It remains to see why Ledrappier's example is not 3-fold mixing. For this, apply the 3-dot rule from corner (i,j), from corner $(i+1,j)$ and from corner $(i,j+1)$, then add the three equalities (see Figure 3). In the sum, the random variables $\xi_{i+1,j}$, $\xi_{i+1,j+1}$ and $\xi_{i,j+1}$ are counted twice, thus they vanish since we work modulo 2. We get the following equality, which could be called the *scale-2 3-dot rule*:

$$(4) \qquad \xi_{i,j} + \xi_{i+2,j} + \xi_{i,j+2} \ = \ 0 \mod 2.$$

A straightforward induction then shows that for any $n \geq 0$, the *scale-2^n 3-dot rule* holds:

$$(5) \qquad \xi_{i,j} + \xi_{i+2^n,j} + \xi_{i,j+2^n} \ = \ 0 \mod 2.$$

But this shows that the three windows of size 1 $\{(i,j)\}$, $\{(i+2^n,j)\}$ and $\{(i,j+2^n)\}$ are always far from being independent, although the distance between them can be made arbitrarily large. Hence the random field ξ is not 3-fold mixing.

2. ATTEMPT TO CONSTRUCT A 3-DOT-TYPE ONE-DIMENSIONAL PROCESS

2.1. Block construction of a 2-fold but not 3-fold mixing process.
In this section we describe a naive attempt to mimic the 3-dot construction on a one-dimensional process. Our process will take its values in the same alphabet $\mathbb{A} = \{0,1\}$ as for Ledrappier's example, and we
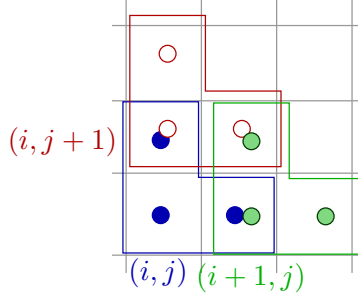
FIGURE 3. Applying the 3-dot rule from three different corners (i, j), $(i + 1, j)$ and $(i, j + 1)$, and adding the three equalities gives the scale-2 3-dot rule.

start by randomly picking the two random variables ξ_0 and ξ_1 with two independent unbiased coin tosses : ξ_0 and ξ_1 are independent, and each one is equal to 1 with probability $1/2$. Then, we set

$$\xi_2 \; := \; \xi_0 + \xi_1 \mod 2.$$

Each random variable ξ_i is called a 0-block, and the triple (ξ_0, ξ_1, ξ_2) is called a *1-block*. We pick the second 1-block (ξ_3, ξ_4, ξ_5) in the same way as the first one, but independently. The third 1-block (ξ_6, ξ_7, ξ_8) is now set to be the pointwise sum of the first two 1-blocks:

$$\xi_6 \; := \; \xi_0 + \xi_3 \mod 2,$$
$$\xi_7 \; := \; \xi_1 + \xi_4 \mod 2,$$
$$\xi_8 \; := \; \xi_2 + \xi_5 \mod 2.$$

Observe that this third 1-block follows the same distribution as the first two: ξ_6 and ξ_7 are two independent Bernoulli random variables with parameter $1/2$, and ξ_8 is the sum mod 2 of these variables. Note also that the third 1-block is independent of the first one, independent of the second one, but of course not independent of the first two together. The 9-tuple $(\xi_0, \xi_1, \ldots, \xi_8)$ is called a *2-block*.

We can repeat this procedure inductively to construct k-blocks for each $k \geq 0$: Suppose that for some k we already have constructed the first k-block, which is the 3^k-tuple $(\xi_0, \ldots, \xi_{3^k-1})$. Then, choose the second k-block $(\xi_{3^k}, \ldots, \xi_{2 \times 3^k-1})$ with the same probability distribution, but independently of the first one, and set the third k-block $(\xi_{2 \times 3^k}, \ldots, \xi_{3^{k+1}-1})$ to be the pointwise sum of the first two k-blocks:

$$(6) \qquad\qquad \xi_{2 \times 3^k + j} \; := \; \xi_j + \xi_{3^k + j} \mod 2 \; (0 \leq j \leq 3^k - 1).$$
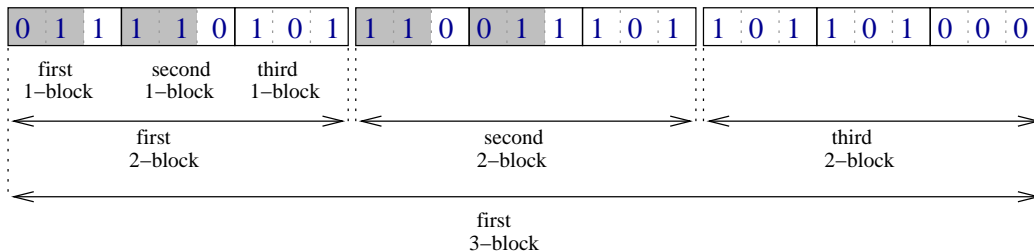

FIGURE 4. Block construction of a stochastic process: The shaded coordinates are given by independent coin tosses . The non-shaded coordinates are computed from the shaded ones by 3-dot-type rules.

This inductive procedure gives the construction of a stochastic one-dimensional process $(\xi_i)_{i \geq 0}$. (This construction can easily be extended to a process indexed by $\mathbb{Z}$: Set the negative coordinates independently of the nonnegative ones by a similar symmetric construction.) Let us sketch the

proof that our process is 2-fold mixing. For this, we use the two following facts, whose verification is left to the reader:

- Two different k-blocks are always independent.
- Call a k-*overlapping* the concatenation of two consecutive $(k-1)$-blocks lying in two different k-blocks. Any k-overlapping is independent of any concatenation of two consecutive $(k-1)$-blocks lying in other k-blocks.

Now, take two windows of fixed size ℓ, and let k be an integer such that $\ell \leq 3^{k-1}$. Then, if the distance between the two windows is greater than 3^k, either they lie in two different k-blocks, or at least one of them lie in a k-overlapping. In both cases the two windows are independent.

However, the stochastic process ξ is clearly not 3-fold mixing, since for any $k \geq 0$, we have

$$\xi_0 + \xi_{3^k} + \xi_{2 \times 3^k} = 0 \mod 2.$$

This, of course, does not make ξ a counterexample to Rohlin's question: The process we have just constructed is not a stationary one! Indeed, the pattern '111' for example can not be seen in the sequence $\xi_0\xi_1\xi_2$, but it can occur in the sequence $\xi_1\xi_2\xi_3$ with probability $1/8$.

2.2. **How to make the construction stationary.** The example described in the preceding section can be turned into a stationary process by applying some trick which is presented here. The process is still inductively constructed with k-blocks which follow the same distribution as before. The difference consists in the way k-blocks are extended to $(k+1)$-blocks. Observe that a k-block lying in a given $(k+1)$-block can have three positions, which will be denoted by '0' (the first k-block in the $(k+1)$-block), '1' (the second one) and '2' (the third one). We are going to define the increasing family of k-blocks ($k \geq 0$) containing the coordinate ξ_0 by using a sequence $S = (S_k)_{k \geq 0}$ of independent, uniformly distributed random variables, taking their values in $\{0, 1, 2\}$.

We start the construction by picking the first 0-block ξ_0 in the usual way, with a coin toss. Now, we have to decide whether this 0-block is in the first, second or third position in the 1-block. This is done by using the first random variable S_0. Next, we complete the 1-block by tossing a coin for the first missing variable, and setting the last one to be the sum mod 2 of the two others. The extension from the k-block to the $(k+1)$-block containing ξ_0 goes on in a similar way: Once we have determined the k-block, we use the random variable S_k to decide whether this k-block is in the first, second or third position in the $(k+1)$-block. Then the first missing k-block is chosen independently, and the last one is set to be the pointwise sum of the two other k-blocks.
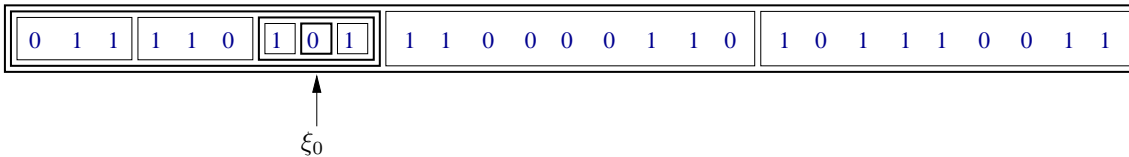

FIGURE 5. Beginning of the construction with the skeleton sequence $S_0 = 1$, $S_1 = 2$, and $S_2 = 0$.

The embedding of k-blocks in $k+1$-blocks is called the *skeleton* of the process, and the i.i.d. sequence $(S_k)_{k\geq 0}$ coding this embedding is the *skeleton sequence*. Since almost every realization of the skeleton sequence contains infinitely many 1's, the preceding procedure applied for all $k \geq 0$ gives rise to k-blocks extending arbitrarily far away from 0 on both sides with probability one. This defines the whole process $\xi = (\xi_i)_{i \in \mathbb{Z}}$.

Let us see how the skeleton sequence evolves when a coordinate shift is applied to the process ξ. It is not difficult to convince oneself that a shift of one coordinate to the left corresponds to the addition of '1' on the 3-adic number defined by the sequence (S_k). (Write the sequence from right to left, and see it as a "number" written in base 3 with infinitely many digits, S_0 being the unitary digit; then add '1' to the sequence as you would do it for an ordinary number: Add '1' to S_0, and if S_0 reaches 3, then set $S_0 = 0$ and add '1' to S_1, and so on.) Observe that the distribution of the

skeleton sequence is the same after this addition of '1', hence the distribution of the skeleton is invariant under the action of the coordinate shift. But once the skeleton is fixed, the distribution of the process is entirely determined by giving the distribution of k-blocks for every $k \geq 0$, which is the distribution described in the preceding section. Therefore the whole distribution of the process is invariant under the coordinate shift, and the process we get now is stationary.
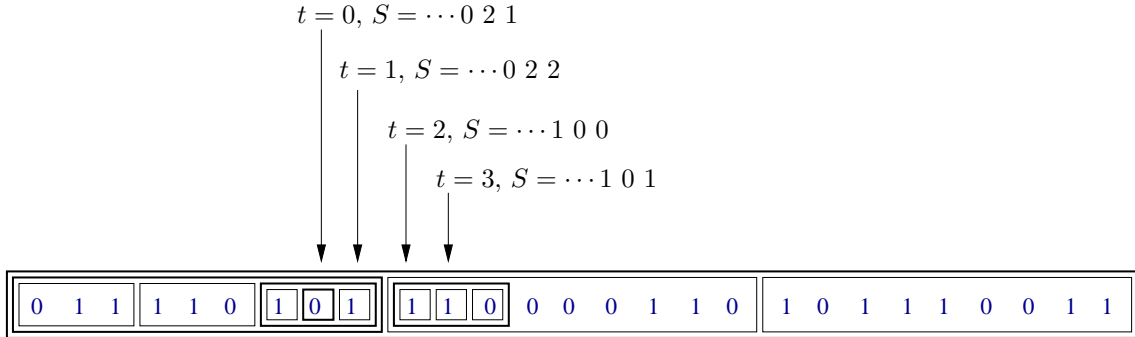
$$t = 0,\ S = \cdots 0\ 2\ 1$$
$$t = 1,\ S = \cdots 0\ 2\ 2$$
$$t = 2,\ S = \cdots 1\ 0\ 0$$
$$t = 3,\ S = \cdots 1\ 0\ 1$$

| 0 1 1 | 1 1 0 | 1 0 1 | 1 1 0 | 0 0 0 1 1 0 | 1 0 1 1 1 0 0 1 1 |

FIGURE 6. Action of the coordinate shift on the skeleton sequence. The arrows denote the position of ξ_0 at successive times during the iteration of the shift.

Unfortunately, making the process stationary has a cost: We have lost the 2-fold mixing property! Indeed, if for example we look at the rightmost coordinate S_0 of the skeleton sequence, we see that any realization of the process ξ gives rise for S_0 to the periodic sequence

$$\cdots 012012012\cdots$$

So, the process we get when we only observe S_0 is periodic. But if ξ was 2-fold mixing, then every *factor* of ξ (that is to say, every stationary process which can be seen as a function of ξ, such as the process generated by S_0 for example[2]) would also be 2-fold mixing.

2.3. **A relative counterexample to Rohlin's question.** The stationary process generated by the whole skeleton sequence S is well-known in ergodic theory, and is called the *3-adic odometer*. (Be careful: S does not take its values in a finite alphabet, it has infinitely many coordinates taking their values in $\{0, 1, 2\}$.) This process, which has appeared as a factor of ξ in our new construction, is far from being mixing, since each of its coordinates is periodic. However, we can notice some interesting facts regarding the 2-fold and 3-fold mixing properties of ξ. Namely, once the skeleton is fixed (that is to say, conditionally to the σ-algebra generated by S), the mixing properties of ξ are similar to those of the non-stationary process constructed in Section 2.1. Thus, the process ξ is 2-fold, but not 3-fold mixing *relatively to the factor σ-algebra generated by S*. (More details on relative k-fold mixing can be found in [17].)

It is a common idea in abstract ergodic theory to say that the study of stationary processes relatively to their factor σ-algebras gives rise to similar results as in the absolute study (one of the best examples of this fact is Thouvenot's relative version of Ornstein's isomorphism theorem [21]; another example is the proof of Proposition 3.2 presented below). Therefore, the process that we have just constructed could make us think that a one-dimensional counterexample to Rohlin's question should exist. However, we are going to show in the next Section that, if such a process exists, it must be of a different nature than this one, or than Ledrappier's counterexample in $\mathbb{Z}^2$.

3. OBSTRUCTION TO THE CONSTRUCTION OF A 3-DOT-TYPE ONE-DIMENSIONAL COUNTEREXAMPLE

3.1. **Multifold mixing and self-joinings.** We need now to present a powerful tool which has been introduced in ergodic theory by Furstenberg [4]: The notion of self-joining of a stationary process. Let $\xi = (\xi_i)_{i \in \mathbb{Z}}$ be a stationary process taking its values in the alphabet A, and denote

---

[2]We leave as an exercise for the reader the verification of the fact that the skeleton sequence is indeed a function of ξ.

by μ its probability distribution on $\mathbb{A}^{\mathbb{Z}}$. Take ξ' another process defined on the same probability space, taking its values in the same alphabet $\mathbb{A}$, and following the same distribution μ. Then we can consider the joint process (ξ, ξ'), taking its values in the Cartesian square $\mathbb{A} \times \mathbb{A}$. If this joint process is still stationary, then we say that its distribution λ on $\mathbb{A}^{\mathbb{Z}} \times \mathbb{A}^{\mathbb{Z}} \approx (\mathbb{A} \times \mathbb{A})^{\mathbb{Z}}$ is a *2-fold self-joining of* ξ. In other words, a 2-fold self-joining of ξ is a probability distribution on $\mathbb{A}^{\mathbb{Z}} \times \mathbb{A}^{\mathbb{Z}}$ whose marginals are both equal to μ, and which is invariant under the coordinate shift.

Let us see some simple examples of such self-joinings. The first idea is to take the two processes ξ and ξ' independent of each other. Then we get $\mu \otimes \mu$ as our first example of a 2-fold self-joining. Another very simple example is obtained by taking $\xi' = \xi$, and we denote by Δ_0 ("diagonal measure") the 2-fold self-joining of ξ concentrated on the diagonal of $\mathbb{A}^{\mathbb{Z}} \times \mathbb{A}^{\mathbb{Z}}$. This can be generalized by considering the case where ξ' is equal to a shifted copy of ξ: We fix some $p \in \mathbb{Z}$, and we set $\xi'_i := \xi_{i+p}$ for each $i \in \mathbb{Z}$. We denote by Δ_p the shifted diagonal measure obtained in this way.

The set $J_2(\xi)$ of all 2-fold self-joinings of ξ is endowed with the metrizable topology defined by the following distance:

$$d(\lambda_1, \lambda_2) := \sum_{n \geq 0} \sum_{n' \geq 0} \frac{1}{2^{n+n'}} \left| \lambda_1(\xi \in C_n, \xi' \in C_{n'}) - \lambda_2(\xi \in C_n, \xi' \in C_{n'}) \right|,$$

where $(C_n)_{n \geq 0}$ is the countable collection af all cylinder sets in $\mathbb{A}^{\mathbb{Z}}$. This topology (which is nothing else than the weak topology restricted to the set of 2-fold self-joinings of ξ) turns $J_2(\xi)$ into a compact metrizable topological space. The link with the 2-fold mixing property is now straightforward: The stationary process ξ is 2-fold mixing if and only if the sequence (Δ_p) of shifted diagonal measures converges in $J_2(\xi)$ to the product measure $\mu \otimes \mu$ as $p \to +\infty$.

To translate the 3-fold mixing property into the language of joinings, we have to generalize the notion of self-joining to the case where 3 processes ξ, ξ' and ξ'' with the same distribution μ are involved. This naturally leads to the definition of a *3-fold self-joining of* ξ. (We can of course define an r-fold self-joining of ξ for any $r \geq 2$, but for our purpose the cases $r = 2$ and $r = 3$ will suffice.) The set $J_3(\xi)$ is also turned into a compact metrizable space when endowed with the restriction of the weak topology. Particularly simple and interesting elements of $J_3(\xi)$ are again the product measure $\mu \otimes \mu \otimes \mu$ and the shifted diagonal measures $\Delta_{p,q}$, $p, q \in \mathbb{Z}$, the latter denoting the distribution of the triple (ξ, ξ', ξ'') when for all $i \in \mathbb{Z}$,

$$(7) \qquad \xi'_i = \xi_{i+p} \quad \text{and} \quad \xi''_i = \xi_{i+p+q}.$$

The process ξ is 3-fold mixing if and only if the following convergence holds in $J_3(\xi)$:

$$(8) \qquad \Delta_{p,q} \xrightarrow[p,q \to +\infty]{} \mu \otimes \mu \otimes \mu.$$

Now, let us assume that ξ is a 2-fold mixing stationary process which is not 3-fold mixing. Then, since $\Delta_{p,q}$ does not converge to the product measure, we can find a subsequence Δ_{p_n, q_n} converging to some 3-fold self-joining $\lambda \neq \mu \otimes \mu \otimes \mu$. But the 2-fold mixing property of ξ tells us that, under λ, the 3 processes ξ, ξ' and ξ'' have to be pairwise independent. Hence, we get the following conclusion:

Proposition 3.1. *If ξ is a 2-fold mixing stationary process which is not 3-fold mixing, then ξ has a 3-fold self-joining $\lambda \neq \mu \otimes \mu \otimes \mu$ with pairwise independent coordinates.*

3.2. **Restriction to zero-entropy processes.** The natural question now is whether stationary processes satisfying the conclusion of Proposition 3.1 can exist. But, without extra requirements, it is easy to find examples of such pairwise independent self-joinings which are not the product measure: Let ξ consist of i.i.d. random variables $(\xi_p)_{p \in \mathbb{Z}}$, taking their values in $\{0, 1\}$, each one with probability $1/2$. Take an independent copy $(\xi'_p)_{p \in \mathbb{Z}}$ of this process and set

$$\xi''_p := \xi_p + \xi'_p \mod 2.$$

Then the three processes ξ, ξ' and ξ'' have the same distribution, are pairwise independent, but the 3-fold self-joining of ξ we get in this way is not the product measure.

However, this process is not a counterexample to Rohlin's question. Indeed, all its coordinates being independent, ξ is of course k-fold mixing for any $k \geq 2$. It was Thouvenot who explained that this kind of Bernoulli shift situation could be avoided when one studies Rohlin's question, because it is always possible to restrict the analysis to the case of zero-entropy stationary processes. A stationary process $\xi = (\xi_i)_{i \in \mathbb{Z}}$ taking its values in a finite alphabet has zero entropy if and only if ξ_0 is measurable with respect to the *past, i.e.* with respect to the σ-algebra $\sigma(\xi_i, \ i < 0)$. (Main definitions and results concerning entropy in ergodic theory can be found *e.g.* in [2, 5, 8].)

Proposition 3.2. *If there exists a stationary process ξ which is 2-fold, but not 3-fold mixing, then one can find such a counterexample in the class of zero-entropy stationary processes.*

Proof. One of the main ingredient to prove this result is the so-called *Pinsker σ-algebra* (or *tail field*) of the process, which is the σ-algebra

$$\Pi(\xi) \ := \ \bigcap_{p \in \mathbb{Z}} \sigma(\xi_{-\infty}^p) \ = \ \bigcap_{p \in \mathbb{Z}} \sigma(\xi_p^{+\infty}).$$

The Pinsker σ-algebra of ξ is invariant by the coordinate shift T: $(\xi_n) \mapsto (\tilde{\xi}_n)$, where $\tilde{\xi}_n := \xi_{n+1}$. Therefore, if we choose any $\Pi(\xi)$-measurable random variable ζ_0, taking its value in a finite alphabet $\mathbb{B}$, and if we set for all $i \in \mathbb{Z}$

$$\zeta_i \ := \ \zeta_0 \circ T^i,$$

then the whole stationary process $\zeta = (\zeta_i)_{i \in \mathbb{Z}}$ is $\Pi(\xi)$-measurable. What is remarkable is that such a process always has zero entropy, and that any stationary process ζ with zero entropy which is a factor of ξ is automatically $\Pi(\xi)$-measurable (see *e.g.* [5], Theorem 18.6). Besides, Krieger's finite generator theorem ensures that it is always possible to find such a factor process ζ, taking its values in an alphabet $\mathbb{B}$ containing only two letters, and generating the whole Pinsker σ-algebra (nice proofs of Krieger's theorem can be found in [5], [8], or [16]). We henceforth fix such a process ζ with

$$\Pi(\xi) = \sigma(\zeta_i, i \in \mathbb{Z}).$$

As a factor of ξ, ζ automatically inherits the 2-fold mixing property. It remains to show that if ζ is 3-fold mixing, then so is ξ.

For this, let us fix three cylinder events E_1, E_2 and E_3, which are measurable with respect to ξ_0^ℓ for some $\ell \geq 0$. We have to compute the limit, as p and q go to $+\infty$, of the quantity

$$\mathbb{P}\left(\xi_0^\ell \in E_1, \xi_p^{p+\ell} \in E_2, \xi_{p+q}^{p+q+\ell} \in E_3\right)$$

$$= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_1} \ \big| \ \sigma(\xi_p^{+\infty})\right] \mathbb{1}_{\xi_p^{p+\ell} \in E_2} \mathbb{1}_{\xi_{p+q}^{p+q+\ell} \in E_3}\right]$$

$$= \mathbb{E}\left[\left(\mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_1} \ \big| \ \sigma(\xi_p^{+\infty})\right] - \mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_1} \ \big| \ \Pi(\xi)\right]\right) \mathbb{1}_{\xi_p^{p+\ell} \in E_2} \mathbb{1}_{\xi_{p+q}^{p+q+\ell} \in E_3}\right]$$

$$+ \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_1} \ \big| \ \Pi(\xi)\right] \mathbb{1}_{\xi_p^{p+\ell} \in E_2} \mathbb{1}_{\xi_{p+q}^{p+q+\ell} \in E_3}\right].$$

The martingale convergence theorem gives

$$\mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_1} \ \big| \ \sigma(\xi_p^{+\infty})\right] \ \xrightarrow[p \to +\infty]{L^1} \ \mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_1} \ \big| \ \Pi(\xi)\right],$$

hence the first term can be bounded by ε if p is large enough. We are left with

$$\mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_1} \ \big| \ \Pi(\xi)\right] \mathbb{1}_{\xi_p^{p+\ell} \in E_2} \mathbb{1}_{\xi_{p+q}^{p+q+\ell} \in E_3}\right]$$

$$= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\xi_{-p}^{-p+\ell} \in E_1} \ \big| \ \Pi(\xi)\right] \mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_2} \ \big| \ \sigma(\xi_q^{+\infty})\right] \mathbb{1}_{\xi_q^{q+\ell} \in E_3}\right],$$

and again the martingale convergence theorem allows us to replace $\mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_2} \mid \sigma(\xi_q^{+\infty})\right]$ with $\mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_2} \mid \Pi(\xi)\right]$ if q is large enough. We thus have proven

$$(9) \quad \mathbb{P}\left(\xi_0^\ell \in E_1, \xi_p^{p+\ell} \in E_2, \xi_{p+q}^{p+q+\ell} \in E_3\right)$$
$$- \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_1} \mid \Pi(\xi)\right] \ \mathbb{E}\left[\mathbb{1}_{\xi_p^{p+\ell} \in E_2} \mid \Pi(\xi)\right] \ \mathbb{E}\left[\mathbb{1}_{\xi_{p+q}^{p+q+\ell} \in E_3} \mid \Pi(\xi)\right]\right] \xrightarrow[p,q\to+\infty]{} 0.$$

Note that in this equation, the expectation can also be written as

$$\mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_1} \mid \Pi(\xi)\right] \ \mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_2} \mid \Pi(\xi)\right] \circ T^p \ \mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_3} \mid \Pi(\xi)\right] \circ T^{p+q}\right].$$

Now, if the process ζ generating $\Pi(\xi)$ is 3-fold mixing, we get that this expectation converges, as p and q go to $+\infty$, to the product

$$\mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_1} \mid \Pi(\xi)\right]\right] \ \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_2} \mid \Pi(\xi)\right]\right] \ \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\xi_0^\ell \in E_3} \mid \Pi(\xi)\right]\right]$$
$$= \mathbb{P}\left(\xi_0^\ell \in E_1\right) \mathbb{P}\left(\xi_0^\ell \in E_2\right) \mathbb{P}\left(\xi_0^\ell \in E_3\right),$$

which means that ξ is also 3-fold mixing. □

Let us make some comment about this proof. The main argument is a nice illustration of the principle that the behaviour of a stationary process conditionally to one of its factors gives rise to similar results as in the non-conditioned case. The key notion here is the so-called K-*property*: We say that the stationary process ξ has the K-property if its Pinsker σ-algebra $\Pi(\xi)$ is trivial. It is easy to show that this K-property implies 3-fold mixing: Just write the preceding proof until (9) in the case where $\Pi(\xi)$ is trivial, and you get the result. (In fact, the same argument gives that the K-property implies k-fold mixing for any k). Now, when a factor ζ of ξ is given, we can also define the K-property of ξ *relatively to* ζ (see *e.g.* [18]), and check that ξ always has the K-property relatively to its Pinsker σ-algebra: This comes from the fact that any factor of ξ with zero entropy is $\Pi(\xi)$-measurable. But this in turn gives that ξ is 3-fold mixing *relatively to* $\Pi(\xi)$: This is more or less what (9) says. The end of the proof consists in checking that, if ζ is 3-fold mixing, and if ξ is 3-fold mixing relatively to ζ, then ξ is 3-fold mixing.

Now that we have reduced Rohlin's problem to the case of zero-entropy processes, we can ask the question of pairwise independent self-joinings in this zero-entropy class.

Question 3.3. *Does there exist a zero-entropy, 2-fold mixing stationary process ξ, and a 3-fold self-joining λ of ξ for which the coordinates are pairwise independent but which is different from the product measure?*

Note also that if the assumption of 2-fold mixing is dropped, we can again find some counterexample: Take a periodic stationary process ξ taking its values in $\{0, 1, 2\}$, where ξ_0 is uniformly distributed in the alphabet, and $\xi_{i+1} = \xi_i + 1 \mod 3$ for all $i \in \mathbb{Z}$. Take an independent copy ξ' of ξ, and set

$$\xi_i'' := 2\xi_i' - \xi_i \mod 3.$$

Then ξ, ξ' and ξ'' share the same zero-entropy distribution, they are pairwise independent but the 3-fold self-joining they define is not the product measure.

3.3. **3-dot-type pairwise independent self-joinings.**
Let us consider now the two-dimensional example of Ledrappier from the point of view of self-joinings. The definition of the shifted diagonal 3-fold self-joining $\Delta_{p,q}$ is formally the same as in (7), but in this case p and q are both elements of $\mathbb{Z}^2$. Particularly interesting is the sequence Δ_{p_n,q_n}, where for all $n \geq 0$,

$$p_n := (0, 2^n) \quad \text{and} \quad q_n := (2^n, 0).$$

From the relation (5), we see that the sequence Δ_{p_n,q_n} converges as $n \to +\infty$ to the 3-fold self-joining λ of ξ under which the three coordinates ξ, ξ' and ξ'' are pairwise independent (because ξ is 2-fold mixing), but which is not the product measure since, for all $i \in \mathbb{Z}^2$,

$$(10) \qquad\qquad \xi_i'' = \xi_i + \xi_i' \mod 2.$$

Observe also that, for the non-stationary process constructed in Section 2.1, the sequence of 3-fold self-joinings $\Delta_{3^n, 2 \times 3^n}$ also converges to a joining with pairwise independent marginals which, by (6), satisfies a relation similar to (10). Since this kind of self-joining seems to appear naturally when one tries to construct examples which are 2-fold but not 3-fold mixing, we introduce the (slightly more general) following definition.

Definition 3.4. *Let ξ be a stationary process taking its values in a finite alphabet $\mathbb{A}$. We call 3-dot-type self-joining a 3-fold self-joining λ of ξ which has pairwise independent marginals, and for which there exists a map $f :\ \mathbb{A} \times \mathbb{A} \longrightarrow \mathbb{A}$ such that*

$$(11) \qquad\qquad \forall i \in \mathbb{Z}, \ \xi_i'' \ = \ f(\xi_i, \xi_i') \quad (\lambda\text{-a.s.}).$$

The result that we shall prove now shows that there is no hope to find a 2-fold but not 3-fold zero-entropy stationary process which admits a 3-dot-type self-joining: The only 3-dot-type self-joinings which can be seen in zero entropy are those arising from *periodic* (therefore non 2-fold mixing) processes, like in the example presented at the end of Section 3.2.

Theorem 3.5. *If the stationary process ξ admits a 3-dot-type self-joining, then*
- *either ξ is a periodic process,*
- *or ξ has entropy at least $\log 2$.*

The proof of the theorem is based on the following lemma.

Lemma 3.6. *Let X, Y and Z be 3 random variables, sharing the same distribution on the finite alphabet $\mathbb{A}$. Assume that these random variables are pairwise independent, and that there exists a map $f :\ \mathbb{A} \times \mathbb{A} \longrightarrow \mathbb{A}$ such that*

$$(12) \qquad\qquad\qquad Z \ = \ f(X, Y) \quad (\text{a.s.}).$$

Then their common distribution is the uniform distribution on a subset of $\mathbb{A}$

Proof. Taking a subset of $\mathbb{A}$ instead of $\mathbb{A}$ if necessary, we can assume that each letter of $\mathbb{A}$ is seen with positive probability. Let us fix some $y \in \mathbb{A}$, and condition with respect to the event $(Y = y)$. Since X, Y and Z are pairwise independent, we get for any $x, z \in \mathbb{A}$

$$\mathbb{P}(X = x | Y = y) \ = \ \mathbb{P}(X = x) \quad \text{and} \quad \mathbb{P}(Z = z | Y = y) \ = \ \mathbb{P}(Z = z).$$

But knowing $(Y = y)$, we have the equivalence

$$(Z = z) \iff (f(X, y) = z).$$

Since X and Z can take the same number of values, we deduce that for any $z \in \mathbb{A}$, there exists a unique x such that

$$f(x, y) \ = \ z,$$

and that furthemore, this x satisfies

$$(13) \qquad\qquad\qquad \mathbb{P}(X = x) \ = \ \mathbb{P}(Z = z).$$

Finally, observe that we can condition on $(Y = y)$ for any $y \in \mathbb{A}$, and therefore that (13) holds for any x and z in $\mathbb{A}$ for which we can find a $y \in \mathbb{A}$ satisfying $f(x, y) = z$. But since X and Z are independent, this is true for *any* $x, z \in \mathbb{A}$. $\qquad\qquad\square$

Proof of Theorem 3.5. Fix some $m \geq 1$, and apply Lemma 3.6 with $X := \xi_0^{m-1}$, $Y := {\xi'}_0^{m-1}$ and $Z := {\xi''}_0^{m-1}$ under the 3-dot-type self-joining of ξ: We get that ξ_0^{m-1} is uniformly distributed on the subset of sequences in $\mathbb{A}^m$ which are seen with positive probability. Denote by p_m the number of such sequences. The same argument applied with $m + 1$ in place of m gives that each of the p_{m+1} possible sequences of length $m + 1$ is seen with probability $1/p_{m+1}$. Therefore, any possible sequence of length m has exactly $a_m := p_{m+1}/p_m$ different ways to extend to some possible sequence of length $m+1$, and each of the possible extensions has a conditional probability $1/a_m$. Now, observe that since ξ is stationary, the number of possible extensions of a sequence of length $m + 1$ can not be greater than the number of possible extensions of the sequence of length m obtained by removing the first letter, which gives the inequality $a_{m+1} \leq a_m$. Therefore, there are two cases:

- Either $a_m = 1$ for m large enough, and in this case the process ξ is periodic;
- Or a_m is always greater than or equal to 2, and then the entropy of ξ is at least $\log 2$.

□

3.4. Difference between dimension one and two.
It seems interesting to analyze the result proved in the preceding section, to see precisely which obstacle prevents us from constructing a 3-dot-type counterexample in dimension one, although it is allowed in dimension two. At which point does the one-dimensional argument developped above stop applying to the case of dimension two?

First, note that the reduction of Rohlin's question to zero-entropy processes is still valid in the case of two-dimensional random fields: We can still consider in this case the Pinsker σ-algebra $\Pi(\xi)$, with respect to which every zero-entropy factor of ξ is measurable, and prove that if ξ is 2-fold but not 3-fold mixing, then this property comes from $\Pi(\xi)$. (For some presentations of the Pinsker σ-algebra in the multidimensional case, we refer the reader to [1, 6, 10, 11]. For the generalization of Krieger's finite generator theorem to $\mathbb{Z}^d$-actions, see *e.g.* [15, 3].) As far as Ledrappier's 3-dot example is concerned, there is no need to take a factor since this stationary random field already has zero entropy.

Next, Lemma 3.6 gives that if ξ is a finite-valued stationary random field indexed by $\mathbb{Z}^2$ admitting a 3-dot-type self joining, then the distribution of ξ on any finite window is uniform on all the configurations which are seen with positive probability. And indeed, we can easily check that this is the case for Ledrappier's construction (for example, any allowed configuration in a rectangular $\ell_1 \times \ell_2$ window has probability $2^{-(\ell_1 + \ell_2 - 1)}$). And this seems to be the point of the argument where there is a difference between $\mathbb{Z}$ and $\mathbb{Z}^2$: Although this uniform-probability property implies periodicity for one-dimensional zero-entropy stationary processes, this is no more true for two-dimensional stationary random fields, as we can see with Ledrappier's example.

We can also observe that in the one-dimensional case, if the stationarity property is dropped the argument fails at the same point: The relative counterexample constructed in Section 2.2, that is to say the process conditioned on the skeleton sequence, also has the uniform-probability property (which means in the non-stationary setting that for any $i \leq j$, all the possible configurations for the sequence ξ_i^j have the same probability). However, this process has zero entropy: There exists almost surely some k such that $S_k = 2$, therefore ξ_0 lies in the third k-block in its $(k+1)$-block. Then ξ_0 can be computed from the first two k-blocks, which are measurable with respect to the past of the process.

4. FURTHER QUESTIONS

Many important results have already been presented around Rohlin's multifold mixing problem. Most of them consider some special category of stationary processes (*e.g.* finite-rank processes, or processes with singular spectrum), and prove that in this category a pairwise-independent 3-fold self-joining has to be the product measure.

We hope that the work presented here can be the beginning of a slightly different approach: Consider a stationary process which admits some kind of pairwise-independent 3-fold self-joining which is not the product measure, and see which other properties on the process this assumption entails. In this direction, the most natural generalization of the present study should be the following situation: Take a stationary process ξ which admits a pairwise-independent 3-fold self-joining λ, and assume that the three coordinates ξ, ξ' and ξ'' of this self-joining satisfy

$$(14) \qquad \xi'' = \varphi(\xi, \xi') \quad \lambda\text{-a.s.}$$

for some measurable function $\varphi : \mathbb{A}^{\mathbb{Z}} \times \mathbb{A}^{\mathbb{Z}} \to \mathbb{A}^{\mathbb{Z}}$. Then what can be said on the process ξ? Can it have zero entropy and be 2-fold mixing?

Since this general question seems to be quite difficult, some more restricted classes of pairwise independent self-joinings could be considered first, satisfying (14) with some regularity assumption on φ. For example, what happens if under λ, the coordinate ξ_0'' is a function of finitely many coordinates of ξ and ξ'?

References

1. J.P. Conze, *Entropie d'un groupe abélien de transformations*, Z. Wahrscheinlichkeitstheorie verw. Geb. **25** (1972), 11–30.
2. I.P. Cornfeld, S.V. Fomin, and Ya.G. Sinai, *Ergodic theory*, Springer-Verlag, 1982.
3. Alexandre I. Danilenko and Kyewon K. Park, *Generators and Bernoullian factors for amenable actions and cocycles on their orbits*, Ergodic Theory Dynam. Systems **22** (2002), no. 6, 1715–1745.
4. H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in diophantine approximation*, Math. Systems Theory **1** (1967), 1–49.
5. E. Glasner, *Ergodic theory via joinings*, Mathematical Surveys and Monographs, vol. 101, American Mathematical Society, Providence, RI, 2003.
6. E. Glasner, J-P. Thouvenot, and B. Weiss, *Entropy theory without a past*, Ergod. Th. & Dynam. Sys. **20** (2000), 1355–1370.
7. B. Host, *Mixing of all orders and pairwise independent joinings of systems with singular spectrum*, Israel J. Math. **76** (1991), no. 3, 289–298.
8. S.A. Kalikow, *Outline of ergodic theory*, available at
   http://www.math.uvic.ca/faculty/aquas/kalikow/kalikow.html.
9. _____, *Twofold mixing implies threefold mixing for rank one transformations*, Ergod. Th. & Dynam. Sys. **4** (1984), 237–259.
10. B. Kamiński, *The theory of invariant partitions for $\mathscr{Z}^d$-actions*, Bull. Acad. Pol. Sci., Ser. Sci. Math. **29** (1981), 349–362.
11. _____, *Decreasing nets of σ-algebras and their applications to ergodic theory*, Tôhoku Mathematical Journal **43** (1991), 263–274.
12. F. Ledrappier, *Un champ markovien peut être d'entropie nulle et mélangeant*, C. R. Acad. Sci. Paris Sér. A-B **287** (1978), no. 7, A561–A563.
13. V.P. Leonov, *The use of the characteristic functional and semi-invariants in the ergodic theory of stationary processes*, Soviet Math. Dokl. **1** (1960), 878–881.
14. V.A. Rohlin, *On endomorphisms of compact commutative groups*, Izvestiya Akad. Nauk SSSR. Ser. Mat. **13** (1949), 329–340.
15. A. Rosenthal, *Finite uniform generators for ergodic, finite entropy, free actions of amenable groups*, Probab. Theory Related Fields **77** (1988), no. 2, 147–166.
16. D.J. Rudolph, *Fundamentals of measurable dynamics – Ergodic theory on Lebesgue spaces*, Oxford University Press, 1990.
17. T. de la Rue, *An extension which is relatively twofold mixing but not threefold mixing*, Colloquium Mathematicum **101** (2004), 271–277.
18. _____, *Sur les processus quasi-markoviens et certains de leurs facteurs*, 2005, pp. 215–230.
19. _____, *An introduction to joinings in ergodic theory*, Discrete and Continuous Dynamical Systems **15** (2006), no. 1, 121–142.
20. V.V. Ryzhikov, *Joinings and multiple mixing of finite rank actions*, Funct. Anal. Appl. **27** (1993), no. 2, 128–140.
21. J.P. Thouvenot, *Quelques propriétés des systèmes dynamiques qui se décomposent en un produit de deux systèmes dont l'un est un schéma de Bernoulli*, Israel Journal of Mathematics **21** (1975), 177–207.
22. H. Totoki, *The mixing property of Gaussian flows*, Memoirs of the Faculty of Science, Kyushu University, Ser. A **18** (1964), no. 2, 136–139.

Laboratoire de Mathématiques Raphaël Salem, UMR 6085 CNRS – Université de Rouen, Avenue de l'Université, B.P. 12, F76801 Saint-Étienne-du-Rouvray Cedex

E-mail address: thierry.de-la-rue@univ-rouen.fr