



HAL
open science

Joinings in Ergodic Theory

Thierry de La Rue

► **To cite this version:**

Thierry de La Rue. Joinings in Ergodic Theory. Encyclopedia of Complexity and Systems Science, 2020, 978-3-642-27737-5. 10.1007/978-3-642-27737-5_300-2 . hal-02469083

HAL Id: hal-02469083

<https://hal.science/hal-02469083v1>

Submitted on 6 Feb 2020

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.

Joinings in Ergodic Theory

Thierry de la Rue
Laboratoire de Mathématiques Raphaël Salem,
CNRS – Université de Rouen
Avenue de l'Université
F76801 Saint Étienne du Rouvray.
email: Thierry.de-la-Rue@univ-rouen.fr

Outline

Glossary	2
1 Definition of the subject	4
2 Introduction	4
3 Joinings of two or more dynamical systems	6
3.1 The set of joinings	6
3.2 From disjointness to isomorphy	7
3.2.1 Disjointness	7
3.2.2 Joinings and isomorphism	8
3.3 Joinings and factors	10
3.4 Markov intertwinings and composition of joinings	12
4 Self-joinings	13
4.1 Self-joinings and commuting transformations	13
4.2 Minimal self-joinings	14
4.3 Simple systems	16
4.4 Relative properties with respect to a factor	17
5 Some applications and future directions	18
5.1 Filtering problems	18
5.2 Joinings proofs of Ornstein's and Krieger's theorems	18
5.3 Joinings and Rohlin's multifold mixing question	19
5.3.1 Pairwise-independent joinings	20
5.3.2 Host's and Ryzhikov's theorems	21
5.4 Joinings and multiple ergodic averages	21
5.5 Joinings and conjectures in number theory	23
5.6 Future Directions	25

Glossary

Disjoint measure-preserving systems

The two measure-preserving dynamical systems (X, \mathcal{A}, μ, T) and (Y, \mathcal{B}, ν, S) are said to be *disjoint* if their only joining is the product measure $\mu \otimes \nu$.

Joining

Let I be a finite or countable set, and for each $i \in I$, let $(X_i, \mathcal{A}_i, \mu_i, T_i)$ be a measure-preserving dynamical system. A *joining* of these systems is a probability measure on the Cartesian product $\prod_{i \in I} X_i$, which has the μ_i 's as marginals, and which is invariant under the product transformation $\bigotimes_{i \in I} T_i$.

Marginal of a probability measure on a product space

Let λ be a probability measure on the Cartesian product of a finite or countable collection of measurable spaces $(\prod_{i \in I} X_i, \bigotimes_{i \in I} \mathcal{A}_i)$, and let $J = j_1, \dots, j_k$ be a finite subset of I . The *k-fold marginal* of λ on X_{j_1}, \dots, X_{j_k} is the probability measure μ defined by:

$$\forall A_1 \in \mathcal{A}_{j_1}, \dots, A_k \in \mathcal{A}_{j_k}, \quad \mu(A_1 \times \dots \times A_k) := \lambda \left(A_1 \times \dots \times A_k \times \prod_{i \in I \setminus J} X_i \right).$$

Markov intertwining

Let (X, \mathcal{A}, μ, T) and (Y, \mathcal{B}, ν, S) be two measure-preserving dynamical systems. We call *Markov intertwining* of T and S any operator $P : L^2(X, \mu) \rightarrow L^2(Y, \nu)$ enjoying the following properties:

- $PU_T = U_S P$, where U_T and U_S are the unitary operators on $L^2(X, \mu)$ and $L^2(Y, \nu)$ associated respectively to T and S (i.e. $U_T f(x) = f(Tx)$, and $U_S g(y) = g(Sy)$).
- $P\mathbb{1}_X = \mathbb{1}_Y$,
- $f \geq 0$ implies $Pf \geq 0$, and $g \geq 0$ implies $P^*g \geq 0$, where P^* is the adjoint operator of P .

Minimal self-joinings

Let $k \geq 2$ be an integer. The ergodic measure-preserving dynamical system T has *k-fold minimal self-joinings* if, for any ergodic joining λ of k copies of T , we can partition the set $\{1, \dots, k\}$ of coordinates into subsets J_1, \dots, J_ℓ such that

1. for j_1 and j_2 belonging to the same J_i , the marginal of λ on the coordinates j_1 and j_2 is supported on the graph of T^n for some integer n (depending on j_1 and j_2);

2. for $j_1 \in J_1, \dots, j_\ell \in J_\ell$, the coordinates j_1, \dots, j_ℓ are independent.

We say that T has *minimal self-joinings* if T has k -fold minimal self-joinings for every $k \geq 2$.

Off-diagonal self-joinings

Let (X, \mathcal{A}, μ, T) be a measure-preserving dynamical system, and S be an invertible measure-preserving transformation of (X, \mathcal{A}, μ) commuting with T . Then the probability measure Δ_S defined on $X \times X$ by

$$\Delta_S(A \times B) := \mu(A \cap S^{-1}B) \quad (1)$$

is a 2-fold self-joining of T supported on the graph of S . We call it an *off-diagonal self-joining* of T .

Process in a measure-preserving dynamical systems

Let (X, \mathcal{A}, μ, T) be a measure-preserving dynamical system, and let $(E, \mathcal{B}(E))$ be a measurable space (which may be a finite or countable set, or \mathbb{R}^d , or $\mathbb{C}^d \dots$). For any E -valued random variable ξ defined on the probability space (X, \mathcal{A}, μ) , we can consider the stochastic process $(\xi_i)_{i \in \mathbb{Z}}$ defined by

$$\xi_i := \xi \circ T^i.$$

Since T preserves the probability measure μ , $(\xi_i)_{i \in \mathbb{Z}}$ is a *stationary process*: For any ℓ and n , the distribution of (ξ_0, \dots, ξ_ℓ) is the same as the probability distribution of $(\xi_n, \dots, \xi_{n+\ell})$.

Self-joining

Let T be a measure-preserving dynamical system. A *self-joining* of T is a joining of a family $(X_i, \mathcal{A}_i, \mu_i, T_i)_{i \in I}$ of systems where each T_i is a copy of T . If I is finite and has cardinal k , we speak of a *k-fold self-joining* of T .

Simplicity

For $k \geq 2$, we say that the ergodic measure-preserving dynamical system T is *k-fold simple* if, for any ergodic joining λ of k copies of T , we can partition the set $\{1, \dots, k\}$ of coordinates into subsets J_1, \dots, J_ℓ such that

1. for j_1 and j_2 belonging to the same J_i , the marginal of λ on the coordinates j_1 and j_2 is supported on the graph of some $S \in C(T)$ (depending on j_1 and j_2);
2. for $j_1 \in J_1, \dots, j_\ell \in J_\ell$, the coordinates j_1, \dots, j_ℓ are independent.

We say that T is *simple* if T is k -fold simple for every $k \geq 2$.

1 Definition of the subject

The word *joining* can be considered as the counterpart in ergodic theory of the notion of *coupling* in probability theory (see *e.g.* [63]): Given two or more processes defined on different spaces, what are the possibilities of embedding them together in the same space? There always exists the solution of making them independent of each other, but interesting cases arise when we can do this in other ways. The notion of joining originates in ergodic theory from pioneering works of Furstenberg [18], who introduced the fundamental notion of *disjointness*, and Rudolph, who laid the basis of joining theory in his article on minimal self-joinings [52]. It has today become an essential tool in the classification of measure-preserving dynamical systems and in the study of their intrinsic properties.

2 Introduction

A central question in ergodic theory is to tell when two measure-preserving dynamical systems are essentially the same, *i.e.* when they are *isomorphic*. When this is not the case, a finer analysis consists in asking what these two systems could share in common: For example, do there exist stationary processes which can be observed in both systems? This latter question can also be asked in the following equivalent way: Do these two systems have a *common factor*? The arithmetical flavour of this question is not fortuitous: There are deep analogies between the arithmetic of integers and the classification of measure-preserving dynamical systems, and these analogies were at the starting point of the study of joinings in ergodic theory.

In the seminal paper [18] which introduced the concept of joinings in ergodic theory, Furstenberg observed that two operations can be done with dynamical systems: We can consider the *product* of two dynamical systems, and we can also take a *factor* of a given system. Like the multiplication of integers, the product of dynamical systems is commutative, associative, it possesses a neutral element (the trivial single-point system), and the systems S and T are both factors of their product $S \times T$. It was then natural to introduce the property for two measure-preserving systems to be *relatively prime*. As far as integers are concerned, there are two equivalent ways of characterizing the relative primeness: First, the integers a and b are *relatively prime* if their unique positive common factor is 1. Second, a and b are *relatively prime* if, each time both a and b are factors of an integer c , their product ab is also a factor of c . It is a well-known theorem in number theory that these two properties are equivalent, but this was not clear for their analog in ergodic theory. Furstenberg reckoned that the second way of defining relative primeness was the most interesting property in ergodic theory, and called it *disjointness* of measure-preserving systems (we will discuss precisely in Section 3.2 what the correct analog is in the setting of ergodic theory). He also asked whether the non-existence of a non-trivial common factor between two systems was equivalent to their disjointness. He was able to

prove that disjointness implies the impossibility of a non-trivial common factor, but not the converse. And in fact, the converse turns out to be false: In 1979, Rudolph exhibited a counterexample in his paper introducing the important notion of *minimal self-joinings*. The relationships between disjointness and the lack of common factor will be presented in details in Section 3.3.

Given two measure-preserving dynamical systems S and T , the study of their disjointness naturally leads one to consider all the possible ways these two systems can be both seen as factors of a third system. As we shall see, this is precisely the study of their *joinings*. The concept of joining turns out to be related with many important questions in ergodic theory, and a large number of deep results can be stated and proved inside the theory of joinings. For example, the fact that the dynamical systems S and T are isomorphic is equivalent to the existence of a special joining between S and T , and this can be used to give a joining proof of Krieger's finite generator theorem, as well as Ornstein's isomorphism theorem (see Section 5.2). As it already appears in Furstenberg's article, joinings provide a powerful tool in the classification of measure-preserving dynamical systems: Many classes of systems can be characterized in terms of their disjointness with other systems. Joinings are also strongly connected with difficult questions arising in the study of almost everywhere convergence of non-conventional averages (see Section 5.4).

Amazingly, a situation in which the study of joinings leads to most interesting results consists in considering two or more *identical* systems. We then speak of the *self-joinings* of the dynamical system T . Again, the study of self-joinings is closely related to many ergodic properties of the system: its mixing properties, the structure of its factors, the transformations which commute with T , and so on. . . We already mentioned minimal self-joinings, and we will see in Section 4.2 how this property may be used to get many interesting examples, such as a transformation with no root, or a process with no non-trivial factor. In the same section we will also discuss a very interesting generalization of minimal self-joinings: the property of being *simple*.

The range of applications of joinings in ergodic theory is very large; only some of them will be given in Section 5: The use of joinings in proving Krieger's and Ornstein's theorems, the links between joinings and some questions of pointwise convergence, and the strong connections between the study of self-joinings and Rohlin's famous question on multifold mixing, which has been opened since 1949 [50].

In parallel with the concept of joinings of measure-preserving systems, Furstenberg also introduced in [18] the notion of *topological joinings*, concerning topological dynamical systems (that is, systems given by a continuous transformation of a compact metric space). In the present article we have restricted ourselves to the measure-preserving setting. In addition to Furstenberg's paper, applications of topological joinings can be found for example in [21, 38, 67].

3 Joinings of two or more dynamical systems

In the following, we are given a finite or countable family $(X_i, \mathcal{A}_i, \mu_i, T_i)_{i \in I}$ of measure-preserving dynamical systems: T_i is an invertible measure-preserving transformation of the standard Borel probability space $(X_i, \mathcal{A}_i, \mu_i)$. When it is not ambiguous, we shall often use the symbol T_i to denote both the transformation and the system.

A *joining* λ of the T_i 's (see the definition in the Glossary) defines a new measure-preserving dynamical system: The product transformation

$$\bigotimes_{i \in I} T_i : (x_i)_{i \in I} \mapsto (T_i x_i)_{i \in I}$$

acting on the Cartesian product $\prod_{i \in I} X_i$, and preserving the probability measure λ . We will denote this big system by $(\bigotimes_{i \in I} T_i)_\lambda$. Since all marginals of λ are given by the original probabilities μ_i , observing only the coordinate i in the big system is the same as observing only the system T_i . Thus, each system T_i is a factor of $(\bigotimes_{i \in I} T_i)_\lambda$, via the homomorphism π_i which maps any point in the Cartesian product to its i -th coordinate.

Conversely, if we are given a measure-preserving dynamical system $(Z, \mathcal{C}, \rho, R)$ admitting each T_i as a factor via some homomorphism $\varphi_i : Z \rightarrow X_i$, then we can construct the map $\varphi : Z \rightarrow \prod_{i \in I} X_i$ sending z to $(\varphi_i(z))_{i \in I}$. We can easily check that the image of the probability measure ρ is then a joining of the T_i 's.

Therefore, studying the joinings of a family of measure-preserving dynamical system amounts to study all the possible ways these systems can be together seen as factors in another big system.

3.1 The set of joinings

The set of all joinings of the T_i 's will be denoted by $J(T_i, i \in I)$. Before anything else, we have to observe that this set is never empty. Indeed, whatever the systems are, the product measure $\bigotimes_{i \in I} \mu_i$ always belongs to this set. Note also that any convex combination of joinings is a joining: $J(T_i, i \in I)$ is a convex set.

The set of joinings is turned into a compact metrizable space, equipped with the topology defined by the following notion of convergence: $\lambda_n \xrightarrow[n \rightarrow \infty]{} \lambda$ if and only if, for all family of measurable subsets $(A_i)_{i \in I} \in \prod_{i \in I} \mathcal{A}_i$, finitely many of them being different from X_i , we have

$$\lambda_n \left(\prod_{i \in I} A_i \right) \xrightarrow[n \rightarrow \infty]{} \lambda \left(\prod_{i \in I} A_i \right). \quad (2)$$

We can easily construct a distance defining this topology by observing that it is enough to check (2) when each of the A_i 's is chosen in some countable algebra \mathcal{C}_i generating the σ -algebra \mathcal{A}_i . We can also point out that, when the X_i 's are themselves compact metric spaces, this topology on the set of joinings is nothing but the restriction to $J(T_i, i \in I)$ of the usual weak* topology.

It is particularly interesting to study *ergodic* joinings of the T_i 's, whose set will be denoted by $J_e(T_i, i \in I)$. Since any factor of an ergodic system is itself ergodic, a necessary condition for $J_e(T_i, i \in I)$ not to be empty is that all the T_i 's be themselves ergodic. Conversely, if all the T_i 's are ergodic, we can prove by considering the ergodic decomposition of the product measure $\bigotimes_{i \in I} \mu_i$ that ergodic joinings do exist: Any ergodic measure appearing in the ergodic decomposition of some joining has to be itself a joining. This result can also be stated in the following way:

Proposition 3.1. *If all the T_i 's are ergodic, the set of their ergodic joinings is the set of extremal points in the compact convex set $J(T_i, i \in I)$.*

3.2 From disjointness to isomorphy

In this section, as in many others in this article, we are focusing on the case where our family of dynamical systems is reduced to two of them. We will then rather call them S and T , standing for (Y, \mathcal{B}, ν, S) and (X, \mathcal{A}, μ, T) . We are interested here in two extremal cases for the set of joinings $J(T, S)$. The first one occurs when the two systems are as far as possible from each other: They have nothing to share in common, and therefore their set of joinings is reduced to the singleton $\{\mu \otimes \nu\}$: This is called the *disjointness* of S and T . The second one arises when the two systems are isomorphic, and we will see how this property shows through $J(T, S)$.

3.2.1 Disjointness

Many situations where disjointness arises were already given by Furstenberg in [18]. Particularly interesting is the fact that classes of dynamical systems can be characterized through disjointness properties. We list here some of the main examples of disjoint classes of measure-preserving systems.

Theorem 3.2.

1. T is ergodic if and only if it is disjoint from every identity map.
2. T is weakly mixing if and only if it is disjoint from any rotation on the circle.
3. T has zero entropy if and only if it is disjoint from any Bernoulli shift.
4. T is a K -system if and only if it is disjoint from any zero-entropy system.

The first result is the easiest, but is quite important, in particular when it is stated in the following form: If λ is a joining of T and S , with T ergodic, and if λ is invariant by $T \times \text{Id}$, then $\lambda = \mu \otimes \nu$.

The second, third and fourth results were originally proved by Furstenberg. They can also be seen as corollaries of the theorems presented in Section 3.3, linking the non-disjointness property with the existence of a particular factor.

Both the first and the second results can be derived from the next theorem, giving a general spectral condition in which disjointness arises. The proof of this theorem can be found in [64]. It is a direct consequence of the fact that, if f and g are square-integrable functions in a given dynamical system, and if their spectral measures are mutually singular, then f and g are orthogonal in L^2 .

Theorem 3.3. *If the reduced maximum spectral types of T and S are mutually singular, then T and S are disjoint.*

As we already said in the introduction, disjointness was recognized by Furstenberg as the most pertinent way to define the analog of the arithmetic property “ a and b are relatively prime” in the context of measure-preserving dynamical systems. We must however point out that the statement

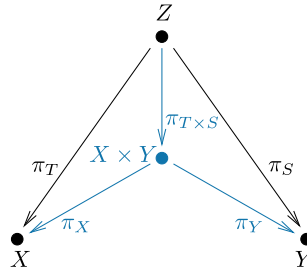
- (i) S and T are disjoint

is, in general, strictly stronger than the straightforward translation of the arithmetic property:

- (ii) Each time both S and T appear as factors in a third dynamical system, then their product $S \times T$ also appears as a factor in this system.

Indeed, contrary to the situation in ordinary arithmetic, there exist non-trivial dynamical systems T which are isomorphic to $T \times T$: For example, this is the case when T is the product of countably many copies of a single non-trivial system. Now, if T is such a system and if we take $S = T$, then S and T do not satisfy statement (i): A non-trivial system is never disjoint from itself, as we will see in the next Section. However they obviously satisfy the statement (ii).

A correct translation of the arithmetic property is the following: S and T are disjoint if and only if, each time T and S appear as factors in some dynamical system through the respective homomorphisms π_T and π_S , $T \times S$ also appears as a factor through a homomorphism $\pi_{T \times S}$ such that $\pi_X \circ \pi_{T \times S} = \pi_T$ and $\pi_Y \circ \pi_{T \times S} = \pi_S$, where π_X and π_Y are the projections on the coordinates in the Cartesian product $X \times Y$ (see the diagram below).



3.2.2 Joinings and isomorphism

We first introduce some notations: For any probability measure λ on a measurable space, let ‘ $A \stackrel{\lambda}{=} B$ ’ stand for ‘ $\lambda(A \Delta B) = 0$ ’. Similarly, if \mathcal{C} and \mathcal{D} are

σ -algebras of measurable sets, we write ' $\mathcal{C} \stackrel{\lambda}{\subset} \mathcal{D}$ ' if, for any $C \in \mathcal{C}$, we can find some $D \in \mathcal{D}$ such that $C \stackrel{\lambda}{=} D$, and by ' $\mathcal{C} \stackrel{\lambda}{=} \mathcal{D}$ ' we naturally mean that both $\mathcal{C} \stackrel{\lambda}{\subset} \mathcal{D}$ and $\mathcal{D} \stackrel{\lambda}{\subset} \mathcal{C}$ hold.

Let us assume now that our two systems S and T are isomorphic: This means that we can find some measurable one-to-one map $\varphi : X \rightarrow Y$, with $T(\mu) = \nu$, and $\varphi \circ T = S \circ \varphi$. With such a φ , we construct the measurable map $\psi : X \rightarrow X \times Y$ by setting

$$\psi(x) := \left((x, \varphi(x)) \right).$$

Let Δ_φ be the image measure of μ by ψ . This measure is supported on the graph of φ , and is also characterized by

$$\forall A \in \mathcal{A}, \forall B \in \mathcal{B}, \Delta_\varphi(A \times B) = \mu(A \cap \varphi^{-1}B). \quad (3)$$

We can easily check that, φ being an isomorphism of T and S , Δ_φ is a joining of T and S . And this joining satisfies very special properties:

- For any measurable $A \subset X$, $A \times Y \stackrel{\Delta_\varphi}{=} X \times \varphi(A)$;
- Conversely, for any measurable $B \subset Y$, $X \times B \stackrel{\Delta_\varphi}{=} \varphi^{-1}(B) \times Y$;

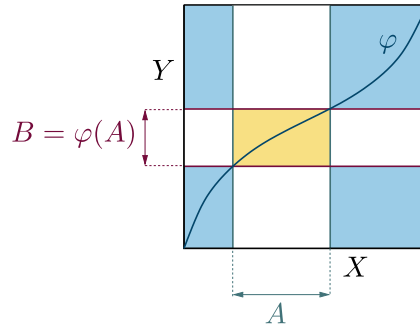


Figure 1: The joining Δ_φ identifies the sets $A \times Y$ and $X \times \varphi(A)$.

Thus, in the case where S and T are isomorphic, we can find a special joining of S and T , which is supported on the graph of an isomorphism, and which identifies the two σ -algebras generated by the two coordinates. What is remarkable is that the converse is true: The existence of an isomorphism between S and T is characterized by the existence of such a joining, and we have the following theorem:

Theorem 3.4. *The measure-preserving dynamical systems S and T are isomorphic if and only if there exists a joining λ of S and T such that*

$$\{X, \emptyset\} \otimes \mathcal{B} \stackrel{\lambda}{=} \mathcal{A} \otimes \{Y, \emptyset\}. \quad (4)$$

When λ is such a joining, it is supported on the graph of an isomorphism of T and S , and both systems are isomorphic to the joint system $(T \otimes S)_\lambda$.

This theorem finds nice applications in the proof of classical isomorphism results. For example, it can be used to prove that two discrete-spectrum systems which are spectrally isomorphic are isomorphic (see [64] or [55]). We will also see in Section 5.2 how it can be applied in the proofs of Krieger's and Ornstein's deep theorems.

Consider now the case where T and S are no longer isomorphic, but where S is only a factor of T . Then we have a factor map $\pi : X \rightarrow Y$ which has the same properties as an isomorphism φ , except that it is not one-to-one (π is only onto). The measure Δ_π , constructed in the same way as Δ_φ , is still a joining supported on the graph of π , but it does not identify the two σ -algebras generated by the two coordinates any more: Instead of Condition (4), Δ_π only satisfies the weaker one:

$$\{X, \emptyset\} \otimes \mathcal{B} \stackrel{\Delta_\pi}{\subset} \mathcal{A} \otimes \{Y, \emptyset\}. \quad (5)$$

The existence of a joining satisfying (5) is a criterion for S being a factor of T .

For more details on the results stated in this section, we refer the reader to [55].

3.3 Joinings and factors

The purpose of this section is to investigate the relationships between the disjointness of two systems S and T , and the lack of a common factor. The crucial fact which was pointed out by Furstenberg is that the existence of a common factor enables one to construct a very special joining of S and T : The *relatively independent joining over this factor*.

Let us assume that our systems S and T share a common factor $(Z, \mathcal{C}, \rho, R)$, which means that we have measurable onto maps $\pi_X : X \rightarrow Z$ and $\pi_Y : Y \rightarrow Z$, respectively sending μ and ν to ρ , and satisfying $\pi_X \circ T = R \circ \pi_X$ and $\pi_Y \circ S = R \circ \pi_Y$. We can then consider the joinings supported on their graphs $\Delta_{\pi_X} \in J(T, R)$ and $\Delta_{\pi_Y} \in J(S, R)$, as defined in the preceding section. Next, we construct a joining λ of the 3 systems S , T and R . Heuristically, λ is the probability distribution of the triple (x, y, z) when we first pick z according to the probability distribution ρ , then x and y according to their conditionnal distribution knowing z in the respective joinings Δ_{π_X} and Δ_{π_Y} , but independently of each other. More precisely, λ is defined by setting, for all $A \in \mathcal{A}$, $B \in \mathcal{B}$ and $C \in \mathcal{C}$

$$\lambda(A \times B \times C) := \int_C \mathbb{E}_{\Delta_{\pi_X}}[\mathbb{1}_{x \in A} | z] \mathbb{E}_{\Delta_{\pi_Y}}[\mathbb{1}_{y \in B} | z] d\rho(z). \quad (6)$$

Observe that the 2-fold marginals of λ on $X \times Z$ and $Y \times Z$ are respectively Δ_{π_X} and Δ_{π_Y} , which means that we have $z = \pi_X(x) = \pi_Y(y)$ λ -almost surely. In other words, we have identified in the two systems T and S the projections on their common factor R . The 2-fold marginal of λ on $X \times Y$ is itself a joining

of T and S , which we call the *relatively independent joining* over the common factor R . This joining will be denoted by $\mu \otimes_R \nu$. (Be careful: The projections π_X and π_Y are hidden in this notation, but we have to know them to define this joining.) From (6), we immediately get the formula defining $\mu \otimes_R \nu$:

$$\forall A \in \mathcal{A}, \forall B \in \mathcal{B},$$

$$\mu \otimes_R \nu(A \times B) := \int_Z \mathbb{E}_{\Delta_{\pi_X}}[\mathbb{1}_{x \in A} | z] \mathbb{E}_{\Delta_{\pi_Y}}[\mathbb{1}_{y \in B} | z] d\rho(z). \quad (7)$$

This definition of the relatively independent joining over a common factor can easily be extended to a finite or countable family of systems sharing the same common factor.

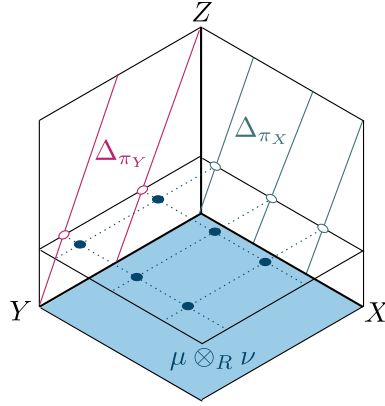


Figure 2: The relatively independent joining $\mu \otimes_R \nu$ and its disintegration over z .

Note that $\mu \otimes_R \nu$ coincides with the product measure $\mu \otimes \nu$ if and only if the common factor is the trivial one-point system. We therefore get the following result:

Theorem 3.5. *If S and T have a non-trivial common factor, then these systems are not disjoint.*

As we already said in the introduction, Rudolph exhibited in [52] a counterexample showing that the converse is not true. There exists however an important result, which was published in [26, 40] allowing us to derive some information on factors from the non-disjointness of two systems.

Theorem 3.6. *If T and S are not disjoint, then S has a non-trivial common factor with some joining of a countable family of copies of T .*

This result leads to the introduction of a special class of factors when some dynamical system T is given: For any other dynamical system S , call T -factor of S any common factor of S with a joining of countably many copies of T . If

$(Z, \mathcal{C}, \rho, R)$ is a T -factor of S and $\pi : Y \rightarrow Z$ is a factor map, we say that the σ -algebra $\pi^{-1}(\mathcal{C})$ is a T -factor σ -algebra of S . Another way to state Theorem 3.6 is then the following: If S and T are not disjoint, then S has a non-trivial T -factor. In fact, an even more precise result can be derived from the proof of Theorem 3.6: For any joining λ of S and T , for any bounded measurable function f on X , the factor σ -algebra of S generated by the function $\mathbb{E}_\lambda[f(x)|y]$ is a T -factor σ -algebra of S .

With the notion of T -factor, Theorem 3.6 has been extended in [42] in the following way, showing the existence of a special T -factor σ -algebra of S comprising anything in S which could lead to a non-trivial joining between T and S .

Theorem 3.7. *Given two measure-preserving dynamical systems (X, \mathcal{A}, μ, T) and (Y, \mathcal{B}, ν, S) , there always exists a maximum T -factor σ -algebra of S , denoted by \mathcal{F}_T .*

Under any joining λ of T and S , the σ -algebras $\mathcal{A} \otimes \{\emptyset, Y\}$ and $\{\emptyset, X\} \otimes \mathcal{B}$ are independent conditionally to the σ -algebra $\{\emptyset, X\} \otimes \mathcal{F}_T$.

Theorem 3.6 gives a powerful tool to prove some important disjointness results, such as those stated in Theorem 3.2. These results involve properties of dynamical systems which are stable under the operations of taking joinings and factors. We will call these properties *stable properties*. This is *e.g.* the case of the zero-entropy property: We know that any factor of a zero-entropy system still has zero entropy, and that any joining of zero-entropy systems also has zero entropy. In other words, T has zero entropy implies that any T -factor has zero entropy. But the property of S being a K -system is precisely characterized by the fact that any non-trivial factor of S has positive entropy. Hence a K -system S cannot have a non-trivial T -factor if T has zero entropy, and is therefore disjoint from T . The converse is a consequence of Theorem 3.5: If S is not a K -system, then it possesses a non-trivial zero-entropy factor, and therefore there exists some zero-entropy system from which it is not disjoint.

The same argument also applies to the disjointness of discrete-spectrum systems with weakly mixing systems, since discrete spectrum is a stable property, and weakly mixing systems are characterized by the fact that they do not have any discrete-spectrum factor.

3.4 Markov intertwining and composition of joinings

There is another way of defining joinings of two measure-preserving dynamical systems involving operators on L^2 spaces, mainly put to light by Ryzhikov (see [60]): Observe that for any joining $\lambda \in J(T, S)$, we can consider the operator $P_\lambda : L^2(X, \mu) \rightarrow L^2(Y, \nu)$ defined by

$$P_\lambda(f) := \mathbb{E}_\lambda[f(x)|y].$$

It is easily checked that P_λ is a *Markov intertwining* of T and S . Conversely, given any Markov intertwining P of T and S , it can be shown that the measure

λ_P defined on $X \times Y$ by

$$\lambda_P(A \times B) := \langle P\mathbb{1}_A, \mathbb{1}_B \rangle_{L^2(Y, \nu)}$$

is a joining of T and S .

This reformulation of the notion of joining is useful when joinings are studied in connection with spectral properties of the transformations (see *e.g.* [28]). It also provides us with a convenient setting to introduce the *composition* of joinings: If we are given three dynamical systems (X, \mathcal{A}, μ, T) , (Y, \mathcal{B}, ν, S) and $(Z, \mathcal{C}, \rho, R)$, a joining $\lambda \in J(T, S)$ and a joining $\lambda' \in J(S, R)$, the composition of the Markov intertwining P_λ and $P_{\lambda'}$ is easily seen to give a third Markov intertwining, which itself corresponds to a joining of T and R denoted by $\lambda \circ \lambda'$. When $R = S = T$, *i.e.* when we are speaking of *2-fold self-joinings* of a single system T (*cf.* next section), this operation turns $J(T, T) = J^2(T)$ into a semigroup. Ahn and Lemańczyk [1] have shown that the subset $J_e^2(T)$ of ergodic 2-fold self-joinings is a sub-semigroup if and only if T is semisimple (see Section 4.3).

4 Self-joinings

We now turn to the case where the measure-preserving dynamical systems we want to join together are all copies of a single system T . For $k \geq 2$, any joining of k copies of T is called a *k-fold self-joining of T* . We denote by $J^k(T)$ the set of all k -fold self-joinings of T , and by $J_e^k(T)$ the subset of *ergodic k-fold self-joinings*.

4.1 Self-joinings and commuting transformations

As soon as T is not the trivial single-point system, T is never disjoint from itself: Since T is obviously isomorphic to itself, we can always find a 2-fold self-joining of T which is not the product measure by considering self-joinings supported on graphs of isomorphisms (see Section 3.2.2). The simplest of them is obtained by taking the identity map as an isomorphism, and we get that $J^2(T)$ always contains the *diagonal measure* $\Delta_0 := \Delta_{\text{Id}}$.

In general, an isomorphism of T with itself is an invertible measure-preserving transformation S of (X, \mathcal{A}, μ) which commutes with T . We call *commutant* of T the set of all such transformations (it is a subgroup of the group of automorphisms of (X, \mathcal{A}, μ)), and denote it by $C(T)$. It always contains, at least, all the powers T^n , $n \in \mathbb{Z}$.

Each element S of $C(T)$ gives rise to a 2-fold self-joining Δ_S supported on the graph of S . Such self-joinings are called *off-diagonal self-joinings*. They also belong to $J_e^k(T)$ if T is ergodic.

It follows that properties of the commutant of an ergodic T can be seen in its ergodic joinings. As an example of application, we can cite Ryzhikov's proof of

King's *weak closure theorem* for rank-one transformations¹. Rank-one measure-preserving transformations form a very important class of zero-entropy, ergodic measure-preserving transformations. They have many remarkable properties, among which is the fact that their commutant is reduced to the weak limits of powers of T . In other words, if T is rank one, for any $S \in C(T)$ there exists a subsequence of integers (n_k) such that,

$$\forall A \in \mathcal{A}, \mu(T^{-n_k} A \Delta S^{-1} A) \xrightarrow[k \rightarrow \infty]{} 0. \quad (8)$$

King proved this result in 1986 [36], using a very intricate coding argument. Observing that (8) was equivalent to the convergence, in $J_e^2(T)$, of $\Delta_{T^{n_k}}$ to Δ_S , Ryzhikov showed in [58] that King's theorem could be seen as a consequence of the following general result concerning 2-fold self-joinings of rank-one systems:

Theorem 4.1. *Let T be a rank-one measure-preserving transformation, and $\lambda \in J_e^2(T)$. Then there exist $t \geq 1/2$, a subsequence of integers (n_k) and another 2-fold self-joining λ' of T such that*

$$\Delta_{T^{n_k}} \xrightarrow[k \rightarrow \infty]{} t\lambda + (1-t)\lambda'.$$

4.2 Minimal self-joinings

For any measure-preserving dynamical system T , the set of 2-fold self-joinings of T contains at least the product measure $\mu \otimes \mu$, the off-diagonal joinings Δ_{T^n} for each $n \in \mathbb{Z}$, and any convex combination of these. Rudolph [52] discovered in 1979 that we can find systems for which there are no other 2-fold self-joinings than these obvious ones. When this is the case, we say that T has *2-fold-minimal self-joinings*, or for short: $T \in \text{MSJ}(2)$. It can be shown (see e.g. [53]) that, as soon as the underlying probability space is not atomic (which we henceforth assume), 2-fold minimal self-joinings implies that T is weakly mixing, and therefore that $\mu \otimes \mu$ and Δ_{T^n} , $n \in \mathbb{Z}$, are *ergodic* 2-fold self-joinings of T . That is why 2-fold minimal self-joinings are often defined by the following:

$$T \in \text{MSJ}(2) \iff J_e^2(T) = \{\mu \otimes \mu\} \cup \{\Delta_{T^n}, n \in \mathbb{Z}\}. \quad (9)$$

Systems with 2-fold minimal self-joinings have very interesting properties. First, since for any S in $C(T)$, Δ_S belongs to $J_e^2(T)$, we immediately see that the commutant of T is reduced to the powers of T . In particular, it is impossible to find a square root of T , i.e. a measure-preserving S such that $S \circ S = T$. Second, the existence of a non-trivial factor σ -algebra of T would lead, via the relatively independent self-joining over this factor, to some ergodic 2-fold self-joining of T which is not in the list prescribed by (9). Therefore, any factor σ -algebra of a system with 2-fold minimal self-joinings must be either the trivial σ -algebra $\{\emptyset, X\}$ or the whole σ -algebra \mathcal{A} . This has the remarkable consequence that

¹An introduction to finite-rank transformations can be found e.g. in [43]; we also refer the reader to the quite complete survey [14].

if ξ is any random variable on the underlying probability space which is not almost-surely constant, then the process $(\xi \circ T^n)_{n \in \mathbb{Z}}$ always generates the whole σ -algebra \mathcal{A} . This also implies that T has zero entropy, since positive-entropy systems have many non-trivial factors.

The notion of 2-fold minimal self-joinings extends for any integer $k \geq 2$ to *k-fold minimal self-joinings*, which roughly means that there are no other k -fold ergodic self-joinings than the “obvious” ones: Those for which the k coordinates are either independent or just translated by some power of T (see the glossary for a more precise definition). We denote in this case: $T \in \text{MSJ}(k)$. If T has k -fold minimal self-joinings for all $k \geq 2$, we simply say that T has *minimal self-joinings*.

Rudolph’s construction of a system with 2-fold minimal self-joinings [52] was inspired by a famous work of Ornstein [46], giving the first example of a transformation with no roots. It turned out that Ornstein’s example is a mixing rank-one system, and all mixing rank-one systems were later proved by King [37] to have 2-fold minimal self-joinings. This can also be viewed as a consequence of Ryzhikov’s Theorem 4.1. Indeed, in the language of joinings, the mixing property of T translates as follows:

$$T \text{ is mixing} \iff \Delta_{T^n} \xrightarrow{|n| \rightarrow \infty} \mu \otimes \mu. \quad (10)$$

Therefore, if in Theorem 4.1 we further assume that T is mixing, then either the sequence (n_k) we get in the conclusion is bounded, and then λ is some Δ_{T^n} , or it is unbounded and then $\lambda = \mu \otimes \mu$.

$T \in \text{MSJ}(k)$ obviously implies $T \in \text{MSJ}(k')$ for any $2 \leq k' \leq k$, but the converse is not known. The question whether 2-fold minimal self-joinings implies k -fold minimal self-joinings for all k is related to the important open problem of *pairwise independent joinings* (see Section 5.3.1). But the latter problem is solved for some special classes of systems, in particular in the category of mixing rank-one transformations. It follows that, if T is mixing and rank one, then T has minimal self-joinings.

In 1980, Del Junco, Rahe and Swanson proved that Chacon’s transformation also has minimal self-joinings [9]. This well-known transformation is also a rank-one system, but it is not mixing (it had been introduced by R.V. Chacon in 1969 [6] as the first explicit example of a weakly mixing transformation which is not mixing). For another example of a transformation with 2-fold minimal self-joinings, constructed as an exchange map on three intervals, we refer to [7].

The existence of a transformation with minimal self-joinings has been used by Rudolph as a wonderful tool to construct a large variety of striking counterexamples, such as

- a transformation T which has no roots, while T^2 has roots of any order,
- a transformation with a cubic root but no square root,
- two measure-preserving dynamical systems which are weakly isomorphic (each one is a factor of the other) but not isomorphic. . .

Let us now sketch the argument showing that we can find two systems with no common factor but which are not disjoint: We start with a system T with minimal self-joinings. Consider the direct product of T with an independent copy T' of itself, and take the *symmetric factor* S of $T \otimes T'$, that is to say the factor we get if we only look at the non-ordered pair of coordinates $\{x, x'\}$ in the Cartesian product. Then S is surely not disjoint from T , since the pair $\{x, x'\}$ is not independent of x . However, if S and T had a non-trivial common factor, then this factor should be isomorphic to T itself (because T has minimal self-joinings). Therefore we could find in the direct product $T \otimes T'$ a third copy \tilde{T} of T , which is measurable with respect to the symmetric factor. In particular, \tilde{T} is invariant by the flip map $(x, x') \mapsto (x', x)$, and this prevents \tilde{T} from being measurable with respect to only one coordinate. Then, since $T \in \text{MSJ}(3)$, the systems T , T' and \tilde{T} have no choice but being independent. But this contradicts the fact that \tilde{T} is measurable with respect to the σ -algebra generated by x and x' . Hence, T and S have no non-trivial common factor.

We can also cite the example given by Glasner and Weiss [27] of a pair of horocycle transformations which have no nontrivial common factor, yet are not disjoint. Their construction relies on the deep work by Ratner [49], which describes the structure of joinings of horocycle flows.

4.3 Simple systems

An important generalization of 2-fold minimal self-joinings has been proposed by William A. Veech in 1982 [66]. We say that the measure-preserving dynamical system T is *2-fold simple* if it has no other ergodic 2-fold self-joinings than the product measure $\{\mu \otimes \mu\}$ and joinings supported on the graph of a transformation $S \in C(T)$. (The difference with $\text{MSJ}(2)$ lies in the fact that $C(T)$ may contain other transformations than the powers of T .) It turns out that simple systems may have non-trivial factors, but the structure of these factors can be explicitly described: They are always associated with some compact subgroup of $C(T)$. More precisely, if K is a compact subgroup of $C(T)$, we can consider the factor σ -algebra

$$\mathcal{F}_K := \{A \in \mathcal{A} : \forall S \in K, A = S(A)\},$$

and the corresponding factor transformation $T|_{\mathcal{F}_K}$ (called a *group factor*). Then Veech proved the following theorem concerning the structure of factors of a 2-fold simple system.

Theorem 4.2. *If the dynamical system T is 2-fold simple, and if $\mathcal{F} \subset \mathcal{A}$ is a non-trivial factor σ -algebra of T , then there exists a compact subgroup K of the group $C(T)$ such that $\mathcal{F} = \mathcal{F}_K$.*

There is a natural generalisation of Veech's property to the case of k -fold self-joinings, which has been introduced by Del Junco and Rudolph in 1987 [10] (see the precise definition of simple systems in the glossary). In their work, important results concerning the structure of factors and joinings of simple systems are proved. In particular, they are able to completely describe the

structure of the ergodic joinings between a given simple system and any ergodic system (see also [64] and [24]). Recall that, for any $r \geq 1$, the *symmetric factor* of $T^{\otimes r}$ is the system we get if we observe the r coordinates of the point in X^r and forget their order. This is a special case of group factor, associated to the subgroup of $C(T^{\otimes r})$ consisting of all permutations of the coordinates. We denote this symmetric factor by $T^{(r)}$.

Theorem 4.3. *Let T be a simple system and S an ergodic system. Assume that λ is an ergodic joining of T and S which is different from the product measure. Then there exists a compact subgroup K of $C(T)$ and an integer $r \geq 1$ such that*

- $(T|_{\mathcal{F}_K})^{(r)}$ is a factor of S ,
- λ is the projection on $X \times Y$ of the relatively independent joining of $T^{\otimes r}$ and S over their common factor $(T|_{\mathcal{F}_K})^{(r)}$.

If we further assume that the second system is also simple, then in the conclusion we can take $r = 1$. In other words, ergodic joinings of simple systems S and T are either the product measure or relatively independent joinings over a common group factor. This leads to the following corollary:

Theorem 4.4. *Simple systems without non-trivial common factor are disjoint.*

As for minimal self-joining, it is not known in general whether 2-fold simplicity implies k -fold simplicity for all k . This question is studied in [25], where sufficient spectral conditions are given for this implication to hold. It is also proved that any 3-fold simple weakly mixing transformation is simple of all order.

4.4 Relative properties with respect to a factor

In fact, Veech also introduced a weaker, “relativised”, version of the 2-fold simplicity. If $\mathcal{F} \subset \mathcal{A}$ is a non-trivial factor σ -algebra of T , let us denote by $J^2(T, \mathcal{F})$ the set of 2-fold self-joinings of T which are “constructed over \mathcal{F} ”, which means that their restriction to the product σ -algebra $\mathcal{F} \otimes \mathcal{F}$ coincides with the diagonal measure. (The relatively-independent joining over \mathcal{F} is the canonical example of such a joining.) For the conclusion of Theorem 4.2 to hold, it is enough to assume only that the ergodic elements of $J^2(T, \mathcal{F})$ be supported on the graph of a transformation $S \in C(T)$. This is an important situation where the study of $J^2(T, \mathcal{F})$ gives strong informations on the way \mathcal{F} is embedded in the whole system T , or, in other words, on the relative properties of T with respect to the factor $T|_{\mathcal{F}}$. A simple example of such a relative property is the *relative weak mixing with respect to \mathcal{F}* , which is characterized by the ergodicity of the relatively-independent joining over \mathcal{F} (recall that weak-mixing is itself characterized by the ergodicity of the direct product $T \otimes T$).

For more details on this subject, we refer the reader to [41]. We also wish to mention the generalization of simplicity called *semisimplicity* proposed by Del Junco, Lemańczyk and Mentzen in [8], which is precisely characterized by the

fact that, for any $\lambda \in J_e^2(T)$, the system $(T \otimes T)_\lambda$ is a relatively weakly mixing extension of T .

5 Some applications and future directions

5.1 Filtering problems

Filtering problems were one of the motivations presented by Furstenberg for the introduction of the disjointness property in [18], and he considered in particular the following situation. Suppose we are given two real-valued stationary processes (X_n) and (Y_n) , with their joint distribution also stationary. We can interpret (X_n) as a signal, perturbed by a noise (Y_n) . Under which condition can we recover the original signal (X_n) from the observation of $(X_n + Y_n)$? If this is possible, we say that the sequence (X_n, Y_n) admits a *perfect filter*. Furstenberg proved that a perfect filter exists if the two processes (X_n) and (Y_n) are integrable, and if the two measure-preserving dynamical systems constructed as the shift of the two processes are disjoint. Furstenberg also observed that the integrability assumption can be removed if a stronger disjointness property is satisfied: A perfect filter exists if the system T generated by (X_n) is *doubly disjoint* from the system S generated by (Y_n) , in the sense that T is disjoint from any ergodic self-joining of S . Several generalizations have been studied (see [20, 4]), but the question whether the integrability assumption of the processes can be removed remained open for a long time. Finally, a short and clever argument by Garbit [22] proved in 2011 that indeed the only assumption of disjointness is enough to ensure the existence of a perfect filter.

5.2 Joinings proofs of Ornstein's and Krieger's theorems

We have already seen that joinings could be used to prove isomorphisms between systems. This fact found a nice application in the proofs of two major theorems in ergodic theory: Ornstein's isomorphism theorem [45], stating that two Bernoulli shifts with the same entropy are isomorphic, and Krieger's finite generator theorem [39], which says that any dynamical system with finite entropy is isomorphic to the shift transformation on a finite-valued stationary process. The idea of this joining approach to the proofs of Krieger's and Ornstein's theorems was originally due to Burton and Rothstein, who circulated a preliminary report on the subject which was never published [5]. The first published and fully detailed exposition of these proofs can be found in Rudolph's book [53] (see also in Glasner's book [24]).

In fact, Ornstein's theorem goes far more beyond the isomorphism of two given Bernoulli shifts: It also gives a powerful tool for showing that a specific dynamical system is isomorphic to a Bernoulli shift. In particular, Ornstein introduced the property for an ergodic stationary process to be *finitely determined*. We shall not give here the precise definition of this property (for a complete exposition of Ornstein's theory, we refer the reader to [47]), but sim-

ply point out that Bernoulli shifts and mixing Markov chains are examples of finitely determined processes. Rudolph's argument to show Ornstein's theorem via joinings makes use of Theorem 3.4, and of the topology of $J(T, S)$.

Theorem 5.1 (Ornstein's isomorphism theorem). *Let T and S be two ergodic dynamical systems with the same entropy, and both generated by finitely determined stationary processes. Then the set of joinings of T and S which are supported on graphs of isomorphisms forms a dense G_δ in $\overline{J_e(T, S)}$.*

Krieger's theorem is not as easily stated in terms of joinings, because it does not refer to the isomorphism of two specific systems, but rather to the isomorphism of one given system with some other system which has to be found. We have therefore to introduce a larger set of joinings: Given an integer n , we denote by Y_n the set of double-sided sequences taking values in $\{1, \dots, n\}$. We consider on Y_n the shift transformation S , but we do not determine yet the invariant measure. Now, for a specific measure-preserving dynamical system T , consider the set $J(n, T)$ of all possible joinings of T with some system (Y_n, ν, S) , when ν ranges over all possible shift-invariant probability measures on Y_n . $J(n, T)$ can also be equipped with a topology which turns it into a compact convex metric space, and as soon as T is ergodic, the set $J_e(n, T)$ of ergodic elements of $J(n, T)$ is not empty. In this setting, Krieger's theorem can be stated as follows:

Theorem 5.2 (Krieger's finite generator theorem). *Let T be an ergodic dynamical system with entropy $h(T) < \log_2 n$. Then the set of $\lambda \in J(n, T)$ which are supported on graphs of isomorphisms between T and some system (Y_n, ν, S) forms a dense G_δ in $\overline{J_e(n, T)}$.*

Since any system of the form (Y_n, ν, S) obviously has an n -valued generating process, we obtain as a corollary that T itself is generated by an n -valued process.

As another nice example of how joinings can be used to describe the isomorphisms between two transformations, we can also cite the paper [16] by Foreman, Rudolph and Weiss. With a clever construction of a family of transformations for which they can completely describe the joinings between T and its inverse, they are able to prove that the isomorphism relation between measure preserving transformations of the unit interval is not Borel.

5.3 Joinings and Rohlin's multifold mixing question

We have already seen that the property of T being mixing could be expressed in terms of 2-fold self-joinings of T (see (10)). Rohlin proposed in 1949 [50] a generalization of this property, called *multifold mixing*: The measure-preserving transformation T is said to be *k-fold mixing* if $\forall A_1, A_2, \dots, A_k \in \mathcal{A}$,

$$\lim_{n_2, n_3, \dots, n_k \rightarrow \infty} \mu \left(A_1 \cap T^{-n_2} A_2 \cap \dots \cap T^{-(n_2 + \dots + n_k)} A_k \right) = \prod_{i=1}^k \mu(A_i).$$

Again, this definition can easily be translated into the language of joinings: T is k -fold mixing when the sequence $(\Delta_{T^{n_2}, \dots, T^{n_2 + \dots + n_k}})$ converges in $J^k(T)$ to $\mu^{\otimes k}$ as n_2, \dots, n_k go to infinity, where $(\Delta_{T^{n_2}, \dots, T^{n_2 + \dots + n_k}})$ is the obvious generalization of Δ_{T^n} to the case of k -fold self-joinings. The classical notion of mixing corresponds in this setting to 2-fold mixing. (We must point out that Rohlin's original definition of k -fold mixing involved $k+1$ sets, thus the classical mixing property was called *1-fold mixing*. However it seems that the convention we adopt here is now used by most authors, and we find it more coherent when translated in the language of multifold self-joinings.)

Obviously, 3-fold mixing is stronger than 2-fold mixing, and Rohlin asked in his article whether the converse is true. This question is still open today, even though many important works dealt with it and supplied partial answers. Most of these works directly involve self-joinings via the argument exposed in the following section.

5.3.1 Pairwise-independent joinings

Let T be a 2-fold mixing dynamical system. If T is not 3-fold mixing, $(\Delta_{T^n, T^{n+m}})$ does not converge to the product measure as n and m go to ∞ . By compactness of $J^3(T)$, we can find subsequences (n_k) and (m_k) such that $(\Delta_{T^{n_k}, T^{n_k+m_k}})$ converges to a cluster point $\lambda \neq \mu \otimes \mu \otimes \mu$. However, by 2-fold mixing, the 3 coordinates must be pairwise independent under λ . We therefore get a 3-fold self-joining λ with the unusual property that λ has pairwise independent coordinates, but λ is not the product measure.

A system is said to be *pairwise independently determined* (PID) if for all $k \geq 3$, the only pairwise independent k -fold self-joining of the system is the product measure. Thus, in the class of PID systems, 2-fold mixing implies 3-fold mixing.

In fact, non-PID systems are easy to find (see *e.g.* [55]), but the examples we know so far are either periodic transformations (which can not be counterexamples to Rohlin's question since they are not mixing), or transformations with positive entropy. However, using an argument provided by Thouvenot, we can prove that, if there exists a 2-fold mixing T which is not 3-fold mixing, then we can find such a T in the category of zero-entropy dynamical systems (see *e.g.* [54]). Therefore, a negative answer to the following question would solve Rohlin's multifold mixing problem:

Question 5.3. *Does there exist a zero-entropy, weakly mixing dynamical system T which is not PID?*

The problem of (non) existence of such pairwise-independent joinings is also related to the question whether MSJ(2) implies MSJ(3), or whether 2-fold simplicity implies 3-fold simplicity. Indeed, any counter-example to one of these implication would necessarily be of zero entropy, and would possess a pairwise-independent 3-fold self-joining which is not the product measure. Moreover, it is noticed in [24, Corollary 12.22] that if we find a system which is MSJ(2) but

not MSJ(3), then this system would also be 2-fold mixing without being 3-fold mixing.

Question 5.3 has been answered by Host and Ryzhikov for some special classes of zero-entropy dynamical systems.

5.3.2 Host's and Ryzhikov's theorems

The following theorem, proved in 1991 by Host [30] (see also [44, 24]), establishes a spectacular connection between the spectral properties of a finite family of dynamical systems and the non-existence of a pairwise-independent, non-independent joining:

Theorem 5.4 (Host's theorem on singular spectrum). *Let $(X_i, \mathcal{A}_i, \mu_i, T_i)_{1 \leq i \leq r}$ be a finite family of measure-preserving dynamical systems with purely singular spectrum. Then any pairwise-independent joining $\lambda \in J(T_1, \dots, T_r)$ is the product measure $\mu_1 \otimes \dots \otimes \mu_r$.*

Corollary 5.5. *If a dynamical system with singular spectrum is 2-fold mixing, then it is k -fold mixing for any $k \geq 2$.*

The multifold-mixing problem for rank-one measure-preserving systems was solved in 1984 by Kalikow [35], using arguments which do not involve the theory of joinings. In 1993, Ryzhikov [59] extended Kalikow's result to finite-rank systems, by giving a negative answer to Question 5.3 in the category of finite-rank mixing systems:

Theorem 5.6 (Ryzhikov's theorem for finite rank systems). *All finite-rank mixing transformations are PID.*

Corollary 5.7. *If a finite-rank transformation is 2-fold mixing, then it is k -fold mixing for any $k \geq 2$.*

5.4 Joinings and multiple ergodic averages

Important researches in ergodic theory are devoted to the convergence of so-called *multiple ergodic averages*, that is to say expressions of the form

$$\frac{1}{n} \sum_{k=0}^{n-1} f_1(T_1^k x) f_2(T_2^k x) \cdots f_d(T_d^k x), \quad (11)$$

where $d \geq 2$, T_1, \dots, T_d are d measure-preserving transformations of the probability space (X, \mathcal{A}, μ) and f_1, \dots, f_d are bounded measurable functions. A case of special interest is when $T_i = T^i$ ($i = 1, \dots, d$) for some measure-preserving transformation T , which is strongly connected to Furstenberg's ergodic proof of Szemerédi's Theorem, and more generally when T_1, \dots, T_d are *commuting* transformations.

The study of convergence in L^2 of (11) has a long history, and has led to the emergence of deep and successful ideas. We can cite in particular the work of

Host and Kra [31] who established the L^2 -convergence for successive powers of a single transformation, introducing and describing the structure of a sequence of factors for which joinings play a central role. The first proof of the convergence in L^2 of (11) in the case of d commuting transformations has been given by Tao in [62]. While Tao's original proof relies on a purely combinatorial argument and does not make use of joinings, we would like here to emphasize the alternative proof established by Austin [2], which strongly relies on the theory of joinings, and how they are involved in the description of the structure of the factors of a dynamical system. For an even more joining oriented version of Austin's proof, see [56].

As far as pointwise convergence of multiple ergodic averages is concerned, still more questions remain open, and they certainly also involve the theory of joinings. As a simple example, we present the relationships between the study of pointwise convergence of (11) when $d = 2$, that is, we consider two non necessarily commuting transformations S and T and we study the sequence

$$\left(\frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) g(S^k x) \right)_{n>0}. \quad (12)$$

It turns out that disjointness of T and S is a sufficient condition for the almost-sure convergence to hold. Indeed, let us first consider the case where T and S are defined on *a priori* different spaces (X, \mathcal{A}, μ) and (Y, \mathcal{B}, ν) respectively, and consider the ergodic average in the product

$$\frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) g(S^k y), \quad (13)$$

which can be viewed as the integral of the function $f \otimes g$ with respect to the empirical distribution

$$\delta_n(x, y) := \frac{1}{n} \sum_{k=0}^{n-1} \delta_{(T^k x, S^k y)}.$$

We can always assume that T and S are continuous transformations of compact metric spaces (indeed, any measure-preserving dynamical system is isomorphic to such a transformation on a compact metric space: see e.g. [19]). Then the set of probability measures on $X \times Y$ equipped with the topology of weak convergence is metric compact. Now, here is the crucial point where joinings appear: If T and S are ergodic, we can easily find subsets $X_0 \subset X$ and $Y_0 \subset Y$ with $\mu(X_0) = \nu(Y_0) = 1$, such that for all $(x, y) \in X_0 \times Y_0$, any cluster point of the sequence $(\delta_n(x, y))_{n>0}$ is automatically a joining of T and S . (We just have to pick x and y among the “good points” for the ergodic theorem in their respective spaces.) When T and S are disjoint, the only cluster point of the sequence $(\delta_n(x, y))$ is therefore $\mu \otimes \nu$. This ensures that, for continuous f and g , (13) converges to the product of the integrals of f and g as soon as (x, y) is picked in $X_0 \times Y_0$. The subspace of continuous functions being dense in L^2 , the

classical ergodic maximal inequality (see [23]) ensures that, for any f and g in $L^2(\mu)$, (13) converges for any (x, y) in a rectangle of full measure $X_0 \times Y_0$.

Coming back to the original question where the spaces on which T and S act are identified, we observe that with probability one, x belongs both to X_0 and Y_0 , and therefore the sequence (12) converges.

The existence of a rectangle of full measure in which the sequence of empirical distributions $(\delta_n(x, y))_{n>0}$ always converges to some joining has been studied in [42] as a natural generalization of the notion of disjointness. This property was called *weak disjointness* of S and T , and it is indeed strictly weaker than disjointness, since there are examples of transformations which are weakly disjoint from themselves.

Recent works use joinings to study the pointwise convergence of multiple ergodic averages. We mention in particular [32], where the case of successive powers of an ergodic distal transformation is treated, and [12] where this result is extended to the case of d commuting transformations generating a distal action of \mathbb{Z}^d , using some of Austin's ideas.

Also, let us observe that a strong connection with Question 5.3 on pairwise independent joinings has been established in [29]. Indeed, using a result of Bourgain according to which the almost-sure convergence of (11) holds when $d = 2$ and T_1, T_2 are powers of a single transformation, the authors prove that if T is PID and weakly mixing, then (11) converge almost surely for $T_i = T^i$, $i = 1, \dots, d$.

5.5 Joinings and conjectures in number theory

Joinings have recently proved to be a relevant tool in the study of a conjecture by Sarnak [61], which lies at the interface between number theory and dynamical systems. Sarnak conjecture deals with the famous arithmetic Möbius function μ , which is defined for each integer $n \geq 0$ by

$$\mu(n) := \begin{cases} 0 & \text{if there exists a prime } p \text{ such that } p^2|n, \\ (-1)^k & \text{if } n \text{ is the product of } k \text{ distinct prime numbers } (k \geq 0). \end{cases}$$

A famous heuristic called the *Möbius randomness principle* states that the patterns of symbols $-1, 1$ and 0 in the sequence μ behaves so chaotically that μ has no correlation with any *reasonably simple* sequence ξ . Sarnak proposed a precise interpretation of this principle in the context of dynamical systems:

Conjecture 5.8 (Sarnak's conjecture). *For any topological dynamical system (X, T) with zero topological entropy, any continuous real-valued function f defined on X and any $x \in X$, we have*

$$\frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) \mu(k) \xrightarrow[n \rightarrow \infty]{} 0. \quad (14)$$

Sarnak pointed out that this conjecture was supported by an older conjecture of Chowla which says that $\boldsymbol{\mu}$ has no autocorrelation of any order. More precisely, it can be stated as follows:

Conjecture 5.9 (Chowla's conjecture). *For any integers $r \geq 0$, $i_0, i_1, \dots, i_r \geq 0$ with at least one odd i_j , we have*

$$\frac{1}{n} \sum_{k=0}^{n-1} \boldsymbol{\mu}(k)^{i_0} \boldsymbol{\mu}(k+1)^{i_1} \cdots \boldsymbol{\mu}(k+r)^{i_r} \xrightarrow{n \rightarrow \infty} 0.$$

As pointed out by Sarnak in [61], that the Chowla conjecture implies the Sarnak conjecture can be proved via ergodic theory arguments. Indeed, an interpretation of the Chowla conjecture is that $\boldsymbol{\mu}$ is the product of $\boldsymbol{\mu}^2$ (the characteristic function of the set of square-free numbers) with a sequence $\boldsymbol{\pi} \in \{-1, 1\}^{\mathbb{N}}$ behaving as a typical output of an infinite sequence of balanced coin tosses. Now, with the assumptions of the statement of the Sarnak conjecture, the sequence $f(T^k x) \boldsymbol{\mu}^2(k)$ is produced by a zero-entropy system, whereas the sequence $\boldsymbol{\pi}$ comes from a K -system. Hence (14) can be viewed as a consequence of the disjointness between K -systems and systems of zero entropy. (See *e.g.* [13] for details.)

Another reason why joinings are used in the study of Sarnak conjecture comes from the following criterion introduced by Bourgain, Sarnak and Ziegler [3], which provides a sufficient condition for a bounded sequence (a_n) to be orthogonal to any bounded *multiplicative* function ν . Recall that the Möbius function is itself multiplicative, which means that $\boldsymbol{\mu}(mn) = \boldsymbol{\mu}(m)\boldsymbol{\mu}(n)$ whenever m, n are coprime integers.

Lemma 5.10 (Criterion of Bourgain, Sarnak and Ziegler). *Assume that (a_n) is a bounded sequence of complex numbers, such that*

$$\limsup_{\substack{r, s \rightarrow \infty, \\ r, s \text{ different primes}}} \limsup_{n \rightarrow \infty} \left| \frac{1}{n} \sum_{k=0}^{n-1} a_{kr} \bar{a}_{ks} \right| = 0. \quad (15)$$

Then, for each bounded multiplicative function $\nu : \mathbb{N} \rightarrow \mathbb{C}$, we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} a_k \cdot \nu(k) = 0.$$

To apply the above lemma in the context of the Sarnak conjecture, we have to consider the sequence $a_k = f(T^k x)$, hence (15) becomes an assumption on this kind of multiple ergodic average

$$\frac{1}{n} \sum_{k=0}^{n-1} f((T^r)^k x) \overline{f((T^s)^k x)}.$$

Now, let us assume for simplification that the topological dynamical system (X, T) has a unique invariant probability measure μ . Then arguments similar to those exposed in Section 5.4 lead to considering joinings of different prime powers T^r and T^s of the measure-preserving system (X, T, μ) . For example, as explained in [3], disjointness of different prime powers of T implies the validity of the Sarnak conjecture for the system. But Lemma 5.10 can also be applied in a larger class of systems, which is defined below and for which we can control all possible joinings of different prime powers of T .

The ergodic measure preserving system (X, μ, T) is said to have *Asymptotically Orthogonal Powers* (AOP) if for each given f and g in $L^2(\mu)$ with $\int_X f d\mu = \int_X g d\mu = 0$, we have

$$\lim_{\substack{r, s \rightarrow \infty, \\ r, s \text{ different primes}}} \sup_{\kappa \in J_e(T^r, T^s)} \left| \int_{X \times X} f \otimes g d\kappa \right| = 0.$$

Surprisingly, the class of AOP systems includes examples where all positive powers of T are isomorphic. Examples of system with AOP and consequences of this property are given in [15] and references therein. In particular, the Sarnak conjecture holds for any uniquely ergodic model of an AOP system, with uniform convergence of (14) with respect to x .

5.6 Future Directions

A lot of important open questions in ergodic theory involve joinings, and we already have cited several of them: joinings are a natural tool when we want to deal with some problems of pointwise convergence involving several transformations (see Section 5.4). It can therefore be assumed that they will play an important role in future progress on pointwise convergence of multiple ergodic averages. Their use is also fundamental in the study of Rohlin's question on multifold mixing. As far as this latter problem is concerned, we may mention a recent approach to Question 5.3: start with a transformation for which some special pairwise-independent self-joining exists, and see what this assumption entails. In particular, we can ask under which conditions there exists a pairwise-independent 3-fold self-joining of T under which the third coordinate is a function of the two others. It has already been proved in [33] that if this function is sufficiently regular (continuous for some topology), then T is periodic or has positive entropy. And there are strong evidences leading to the conjecture that, when T is weakly mixing, such a situation can only arise when T is a Bernoulli shift of entropy $\log n$ for some integer $n \geq 2$. A question in the same spirit was raised by Ryzhikov, who asked in [57] under which conditions we can find a factor of the direct product $T \times T$ which is independent of both coordinates.

There is also a lot of work to do with joinings in order to understand the structure of factors of some dynamical systems, and how different classes of systems are related. An example of such a work is given in the class of Gaussian dynamical systems, *i.e.* dynamical systems constructed from the shift on a

stationary Gaussian process: For some of them (which are called *GAG*, from the french *Gaussien à Autocouplages Gaussiens*), it can be proved that any ergodic self-joining is itself a Gaussian system (see [65, 40]), and this gives a complete description of the structure of their factors. This kind of analysis is expected to be applicable to other classes of dynamical systems. In particular, Gaussian joinings find a nice generalization in the notion of *infinitely divisible joinings*, studied by Roy in [51]). These ID joinings concern a wider class of dynamical systems of probabilistic origin, among which we can also find Poisson suspensions. The counterpart of Gaussian joinings in this latter class are *Poisson joinings*, which have been introduced by Derriennic, Frączek, Lemańczyk and Parreau in [11]. As far as Poisson suspensions are concerned, the analog of the GAG property in the Gaussian class can also be considered, and examples of Poisson suspensions for which the only ergodic self-joinings are Poisson joinings have been given in [48] and [34]. In [11], a general joining property is described: T satisfies the *ELF property* (from the french: *Ergodicité des Limites Faibles*) if any joining which can be obtained as a limit of off-diagonal joinings $\Delta_{T^{n_k}}$ is automatically ergodic. It turns out that this property is satisfied by any system arising from an infinitely divisible stationary process (see [11, 51]). It is proved in [11] that the ELF property implies disjointness with any system which is 2-fold simple and weakly mixing but not mixing. The ELF property is expected to give a useful tool to prove disjointness between dynamical systems of probabilistic origin and other classes of systems (see *e.g.* [17] in the case of \mathbb{R} -action for disjointness between ELF systems and a class of special flows over irrational rotations).

References

- [1] Y.-H. Ahn and M. Lemańczyk, *An algebraic property of joinings*, Proc. Amer. Math. Soc. **131** (2003), no. 6, 1711–1716 (electronic).
- [2] Tim Austin, *Multiple recurrence and the structure of probability-preserving systems*, arXiv:1006.0491, 2010.
- [3] J. Bourgain, P. Sarnak, and T. Ziegler, *Disjointness of Möbius from horocycle flows*, From Fourier analysis and number theory to Radon transforms and geometry, Dev. Math., vol. 28, Springer, New York, 2013, pp. 67–83.
- [4] W. Bułatek, M. Lemańczyk, and E. Lesigne, *On the filtering problem for stationary random \mathbb{Z}^2 -fields*, IEEE Trans. Inform. Theory **51** (2005), no. 10, 3586–3593.
- [5] R. Burton and A. Rothstein, *Isomorphism theorems in ergodic theory*, Technical report, Oregon State University, 1977.
- [6] R.V. Chacon, *Weakly mixing transformations which are not strongly mixing*, Proc. Amer. Math. Soc. **22** (1969), 559–562.

- [7] A. del Junco, *A family of counterexamples in ergodic theory*, Israel J. Math. **44** (1983), no. 2, 160–188.
- [8] A. del Junco, M. Lemańczyk, and M.K. Mentzen, *Semisimplicity, joinings and group extensions*, Studia Math. **112** (1995), no. 2, 141–164.
- [9] A. del Junco, M. Rahe, and L. Swanson, *Chacon’s automorphism has minimal self-joinings*, J. Analyse Math. **37** (1980), 276–284.
- [10] A. del Junco and D.J. Rudolph, *On ergodic actions whose self-joinings are graphs*, Ergodic Theory Dynam. Systems **7** (1987), no. 4, 531–557.
- [11] Y. Derriennic, K. Frączek, M. Lemańczyk, and F. Parreau, *Ergodic automorphisms whose weak closure of off-diagonal measures consists of ergodic self-joinings*, To appear in Colloquium Mathematicum.
- [12] S. Donoso and W. Sun, *Pointwise convergence of some multiple ergodic averages*, arXiv:1609.02529, 2016.
- [13] H. El Abdalaoui, J. Kułaga-Przymus, M. Lemańczyk, and T. de la Rue, *The Chowla and the Sarnak conjectures from ergodic theory point of view*, Discrete and Continuous Dynamical Systems **37** (2017), no. 6, 2899–2944.
- [14] S. Ferenczi, *Systems of finite rank*, Colloq. Math. **73** (1997), no. 1, 35–65.
- [15] S. Ferenczi, J. Kułaga-Przymus, and M. Lemańczyk, *Sarnak’s conjecture – what’s new*, arXiv:1710.04039, 2017.
- [16] Matthew Foreman, Daniel J. Rudolph, and Benjamin Weiss, *The conjugacy problem in ergodic theory.*, Ann. Math. (2) **173** (2011), no. 3, 1529–1586 (English).
- [17] K. Frączek and M. Lemańczyk, *A class of special flows over irrational rotations which is disjoint from mixing flows*, Ergodic Theory Dynam. Systems **24** (2004), no. 4, 1083–1095.
- [18] H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Math. Systems Theory **1** (1967), 1–49.
- [19] ———, *Recurrence in ergodic theory and combinatorial number theory*, Princeton University Press, Princeton, N.J., 1981, M. B. Porter Lectures.
- [20] H. Furstenberg, Y. Peres, and B. Weiss, *Perfect filtering and double disjointness*, Ann. Inst. H. Poincaré Probab. Statist. **31** (1995), no. 3, 453–465.
- [21] Harry Furstenberg, Harvey B. Keynes, and Leonard Shapiro, *Prime flows in topological dynamics.*, Isr. J. Math. **14** (1973), 26–38 (English).
- [22] Rodolphe Garbit, *A note on Furstenberg’s filtering problem*, Israel J. Math. **182** (2011), 333–336.

- [23] A.M. Garsia, *Topics in almost everywhere convergence*, Lectures in Advanced Mathematics, vol. 4, Markham Publishing Co., Chicago, Ill., 1970.
- [24] E. Glasner, *Ergodic theory via joinings*, Mathematical Surveys and Monographs, vol. 101, American Mathematical Society, Providence, RI, 2003.
- [25] E. Glasner, B. Host, and D.J. Rudolph, *Simple systems and their higher order self-joinings*, Israel J. Math. **78** (1992), no. 1, 131–142.
- [26] E. Glasner, J.-P. Thouvenot, and B. Weiss, *Entropy theory without a past*, Ergodic Theory Dynam. Systems **20** (2000), no. 5, 1355–1370.
- [27] S. Glasner and B. Weiss, *Minimal transformations with no common factor need not be disjoint*, Israel J. Math. **45** (1983), no. 1, 1–8.
- [28] G.R. Goodson, *Joining properties of ergodic dynamical systems having simple spectrum*, Sankhyā Ser. A **62** (2000), no. 3, 307–317, Ergodic theory and harmonic analysis (Mumbai, 1999).
- [29] Y. Gutman, W. Huang, S. Shao, and X. Ye, *Almost sure convergence of the multiple ergodic average for certain weakly mixing systems*, Acta Math. Sinica, English Series **34** (2018), no. 1, 79–90.
- [30] B. Host, *Mixing of all orders and pairwise independent joinings of systems with singular spectrum*, Israel J. Math. **76** (1991), no. 3, 289–298.
- [31] Bernard Host and Bryna Kra, *Nonconventional ergodic averages and nilmanifolds*, Ann. of Math. (2) **161** (2005), no. 1, 397–488.
- [32] W. Huang, S. Shao, and X. Ye, *Pointwise convergence of multiple ergodic averages and strictly ergodic models*, arXiv:1406.5930, 2014.
- [33] É. Janvresse and T. de la Rue, *On a class of pairwise-independent joinings*, Preprint, arXiv:0704.3358v2 [math.PR], 2007.
- [34] É. Janvresse, E. Roy, and T. de la Rue, *Poisson suspensions and Sushis*, Ann. Scient. Éc. Norm. Sup. **50** (2017), no. 6, 1301–1334.
- [35] S.A. Kalikow, *Twofold mixing implies threefold mixing for rank one transformations*, Ergodic Theory Dynam. Systems **4** (1984), no. 2, 237–259.
- [36] J. King, *The commutant is the weak closure of the powers, for rank-1 transformations*, Ergodic Theory Dynam. Systems **6** (1986), no. 3, 363–384.
- [37] ———, *Joining-rank and the structure of finite rank mixing transformations*, J. Analyse Math. **51** (1988), 182–227.
- [38] Jonathan L. King, *A map with topological minimal self-joinings in the sense of del Junco.*, Ergodic Theory Dyn. Syst. **10** (1990), no. 4, 745–761 (English).

- [39] W. Krieger, *On entropy and generators of measure-preserving transformations*, Trans. Amer. Math. Soc. **149** (1970), 453–464.
- [40] M. Lemańczyk, F. Parreau, and J.-P. Thouvenot, *Gaussian automorphisms whose ergodic self-joinings are Gaussian*, Fund. Math. **164** (2000), no. 3, 253–293.
- [41] M. Lemańczyk, J.-P. Thouvenot, and B. Weiss, *Relative discrete spectrum and joinings*, Monatsh. Math. **137** (2002), no. 1, 57–75.
- [42] E. Lesigne, B. Rittaud, and T. de la Rue, *Weak disjointness of measure-preserving dynamical systems*, Ergodic Theory Dynam. Systems **23** (2003), no. 4, 1173–1198.
- [43] M.G. Nadkarni, *Basic ergodic theory*, second ed., Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], Birkhäuser Verlag, Basel, 1998.
- [44] ———, *Spectral theory of dynamical systems*, Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], Birkhäuser Verlag, Basel, 1998.
- [45] D.S. Ornstein, *Bernoulli shifts with the same entropy are isomorphic*, Advances in Math. **4** (1970), 337–352 (1970).
- [46] ———, *On the root problem in ergodic theory*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory (Berkeley, Calif.), Univ. California Press, 1972, pp. 347–356.
- [47] ———, *Ergodic theory, randomness, and dynamical systems*, Yale University Press, New Haven, Conn., 1974, James K. Whittemore Lectures in Mathematics given at Yale University, Yale Mathematical Monographs, No. 5.
- [48] F. Parreau and E. Roy, *Poisson suspensions with a minimal set of self-joinings*, Preprint, 2007.
- [49] Marina Ratner, *Horocycle flows, joinings and rigidity of products*, Ann. of Math. (2) **118** (1983), no. 2, 277–313.
- [50] V.A. Rohlin, *On endomorphisms of compact commutative groups*, Izvestiya Akad. Nauk SSSR. Ser. Mat. **13** (1949), 329–340.
- [51] E. Roy, *Poisson suspensions and infinite ergodic theory*, Preprint, 2007.
- [52] D.J. Rudolph, *An example of a measure preserving map with minimal self-joinings, and applications*, J. Analyse Math. **35** (1979), 97–122.

- [53] ———, *Fundamentals of measurable dynamics*, Oxford Science Publications, The Clarendon Press Oxford University Press, New York, 1990, Ergodic theory on Lebesgue spaces.
- [54] T. de la Rue, *2-fold and 3-fold mixing: why 3-dot-type counterexamples are impossible in one dimension*, Bull. Braz. Math. Soc. (N.S.) **37** (2006), no. 4, 503–521.
- [55] ———, *An introduction to joinings in ergodic theory*, Discrete Contin. Dyn. Syst. **15** (2006), no. 1, 121–142.
- [56] ———, *Notes on Austin’s multiple ergodic theorem*, hal-00400975, 2009.
- [57] V. V. Ryzhikov, *Stochastic wreath products and joinings of dynamical systems*, Mat. Zametki **52** (1992), no. 3, 130–140, 160.
- [58] V.V. Ryzhikov, *Mixing, rank and minimal self-joining of actions with invariant measure*, Mat. Sb. **183** (1992), no. 3, 133–160.
- [59] ———, *Joinings and multiple mixing of the actions of finite rank*, Funktsional. Anal. i Prilozhen. **27** (1993), no. 2, 63–78, 96.
- [60] ———, *Joinings, wreath products, factors and mixing properties of dynamical systems*, Izv. Ross. Akad. Nauk Ser. Mat. **57** (1993), no. 1, 102–128.
- [61] Peter Sarnak, *Three lectures on the Möbius function, randomness and dynamics*.
- [62] Terence Tao, *Norm convergence of multiple ergodic averages for commuting transformations*, Ergodic Theory Dynam. Systems **28** (2008), no. 2, 657–688.
- [63] H. Thorisson, *Coupling, stationarity, and regeneration*, Probability and its Applications (New York), Springer-Verlag, New York, 2000.
- [64] J.-P. Thouvenot, *Some properties and applications of joinings in ergodic theory*, Ergodic theory and its connections with harmonic analysis (Alexandria, 1993), London Math. Soc. Lecture Note Ser., vol. 205, Cambridge Univ. Press, Cambridge, 1995, pp. 207–235.
- [65] Jean-Paul Thouvenot, *The metrical structure of some Gaussian processes*, Proceedings of the conference on ergodic theory and related topics, II (Georgenthal, 1986) (Leipzig), Teubner-Texte Math., vol. 94, Teubner, 1987, pp. 195–198.
- [66] W.A. Veech, *A criterion for a process to be prime*, Monatsh. Math. **94** (1982), no. 4, 335–341.

- [67] B. Weiss, *Multiple recurrence and doubly minimal systems.*, Topological dynamics and applications. A volume in honor of Robert Ellis. Proceedings of a conference in honor of the retirement of Robert Ellis, Minneapolis, MN, USA, April 5–6, 1995, Providence, RI: American Mathematical Society, 1998, pp. 189–196 (English).