Remerciements

Je voudrais d'abord remercier Thierry de la Rue et Emmanuel Roy qui m'ont accompagné dans toutes les étapes de mon projet de thèse, dès mon master 2, dans le choix de mon sujet, et ensuite en tant que directeur et co-directeur de ma thèse. Emmanuel pour m'avoir mis en contact avec Thierry, pour avoir maintenu ton intérêt pour mon travail malgré mon départ à Rouen et pour tes conseils, tes idées et tes visites. Thierry pour ta patience, ton suivi et les discussions qui ont permis à notre travail d'aller de l'avant, tout en étant intéressant et diversifié, même lorsqu'on était tous confinés.

Je veux exprimer ma gratitude pour Benjamin Weiss et Christophe Leuridan qui ont pris le temps de regarder mon travail en grands détails en étant rapporteurs de ma thèse. En particulier, je remercie Christophe Leuridan pour m'avoir accueilli à Grenoble à deux reprises.

Je suis aussi reconnaissant envers Tim Austin, Élise Janveresse et Irène Marcovici pour avoir accepté de participer à mon jury de thèse.

Mes remerciements vont aussi à Olivier Guibé et Yves Coudène qui ont participé à mon CSI, ainsi que pour l'invitation à Jussieu de la part de Yves Coudène. Lors de cette visite, j'ai aussi pu rencontrer Jean-Paul Thouvenot que je remercie pour nos échanges qui ont permis des avancées importantes dans mon travail.

J'ai toujours apprécié l'ambiance de travail au LMRS, et j'en remercie tous les membres qui participent à entretenir cet environnement convivial.

Je remercie tout particulièrement les doctorants du laboratoire, avec qui j'ai passé de très bons moments tous les midis, et lors de nos sorties: Anas, Angèle, Audrey, Cécile, Gauthier, Liam, Mirellla, Nessim, Samuel, Silvio et Victor. Je suis très reconnaissant envers Mirella, Cécile, Gauthier et Victor qui étaient là dès mes premières semaines a Rouen, et pour toutes nos aventures, à Rouen et ailleurs. Je garde aussi de très bons souvenir des moments passés au badminton avec Mirella, Victor, puis Nicolas (que je ne peux que compter parmi les doctorants), qui ont étés essentiels pour me changer les idées. Je n'oublierai pas non plus Adrien, Elisa et Rania qui se sont joints à nos sorties avec humour et bonne humeur.

J'en profite aussi pour rappeler ma gratitude pour tous eux qui ont contribué, depuis que j'étais au collège, pour que ma scolarité se déroule le plus normalement possible.

Enfin, je remercie ma famille, mon père, ma mère, mon frère et ma sœur qui étaient là pour moi bien avant qu'il soit question de partir en thèse, et qui m'ont permis de toujours aborder les problèmes sereinement. Je remercie particulièrement ma mère qui m'a aidé à m'orienter dans un monde qu'elle connaît bien, et mon père pour sa bienveillance (et les saucissons).

Contents

Introduction

Ergodic theory is the abstract framework for the study of measurable dynamical systems: we take a set X equipped with a σ-algebra $\mathscr{A}$ and a measure μ on $\mathscr{A}$. Our basic assumption is that $(X, \mathscr{A})$ is a standard Borel space and μ is a Borel measure. A transformation T acts on this measure space, which is a measurable, invertible map with measurable inverse. We assume that T preserves the measure: for all $A \in \mathscr{A}$, $\mu(T^{-1}A) = \mu(A)$. In other words, the image measure of μ by T, denoted by $T_*\mu$, is equal to μ. The quadruple $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ is what we call a *dynamical system*. One of the goals of ergodic theory is to give a classification of those systems. In particular, we have a "hierarchy" on dynamical systems: if $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ and $\mathbf{Y} := (Y, \mathscr{C}, \nu, S)$ are dynamical systems, we say that $\mathbf{Y}$ is a factor of $\mathbf{X}$ if there is a map $\pi : X \to Y$ that preserves the measure, i.e. $\pi_*\mu = \nu$, and such that

$$\pi \circ T = S \circ \pi, \ \ \mu\text{-almost surely.}$$

In that case we say that π is a *factor map*. If, in addition, π is bijective outside of sets of measure 0, we say that π is an isomorphism. We can also view this factor map through the σ-algebra $\mathscr{B} := \pi^{-1}\mathscr{C} \subset \mathscr{A}$ that it generates. One can note that the image of $\mathscr{B}$ under T, i.e. the σ-algebra $T^{-1}\mathscr{B} := \{T^{-1}B \,;\, B \in \mathscr{B}\}$, is equal to $\mathscr{B}$. Such a σ-algebra is called a *factor σ-algebra*.

A well known result from ergodic theory tells us that, for any other system $\mathbf{Z} := (Z, \mathscr{D}, \rho, R)$ and any factor map $\widetilde{\pi} : \mathbf{X} \to \mathbf{Z}$ such that $\widetilde{\pi}^{-1}\mathscr{D} = \mathscr{B}$ mod μ, there exists an isomorphism $\varphi : \mathbf{Z} \to \mathbf{Y}$ such that $\pi = \varphi \circ \widetilde{\pi}$ (see [**?**, Chapter 2, Section 2] or [**?**, Section 2.1]). This means that, by knowing $\mathscr{B}$, we can describe $\mathbf{Y}$ up to isomorphism.

Moreover, while the internal structure of $\mathscr{B}$ informs us on the behavior of the factor system $\mathbf{Y}$, obtaining a complete understanding of $\mathscr{B}$ inside of $\mathbf{X}$ (and consequently of π) requires examining how $\mathscr{B}$ is embedded within $\mathscr{A}$. Here, it is helpful to change our perspective: instead of considering $\mathscr{B}$ as a factor of $\mathscr{A}$, we view $\mathscr{A}$ as an *extension* of $\mathscr{B}$. For convenience, we usually say that the pair $(\mathscr{A}, \mathscr{B})$ is an extension, which we denote by $\mathscr{A} \to \mathscr{B}$, or even by $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ if we want to specify the factor map. The purpose in studying $\mathscr{A} \to \mathscr{B}$ is to determine the relative structure of $\mathscr{A}$ over $\mathscr{B}$. Such extensions constitute an important area of study in ergodic theory and have been considered in many different contexts (see [**?**], [**?**], [**?**], [**?**] or [**?**]). They are also the focus of the first two chapters of my thesis.

An important concept that emerges from our work on extensions is the notion of *confined extensions*, which relies on the theory of joinings, initiated by Furstenberg [**?**]. On $\mathbf{X}$, an extension $\mathscr{A} \to \mathscr{B}$ is confined if, for any joining of $\mathbf{X}$ with another system $\mathbf{Z} := (Z, \mathscr{D}, \rho, R)$ such that $\mathscr{D}$ is independent of $\mathscr{B}$, $\mathscr{D}$ is also independent of $\mathscr{A}$, and the said joining is simply the product joining.

Confined extensions and non-standard dynamical filtrations

As the main part of Chapter **??**, we introduce confined extensions and study their basic properties. We give examples of such extensions appearing in diverse contexts (compact extensions, T, T^{-1} transformations, flow extensions). We also find many ergodic properties of dynamical systems that can be lifted through confined extensions, i.e. properties $\mathcal{P}$ such that, if $\mathcal{P}$ is true on $\mathscr{B}$ and the extension $\mathscr{A} \to \mathscr{B}$ is confined, then $\mathcal{P}$ is true on $\mathscr{A}$. Actually, it turns out that confinement is close to the notion of *stability* introduced by Robinson in [**?**] specifically to get such lifting results. But confinement is easier to manipulate, applicable without any ergodicity assumption and invariant under isomorphism.

In the rest of Chapter **??**, we question whether there is a general structure that can describe, for an extension $\mathscr{A} \to \mathscr{B}$, how we rebuild $\mathscr{A}$, starting from $\mathscr{B}$. Initially, our work is guided by the "static case", i.e. the study of a probability space $(X, \mathscr{A}, \mu)$, where $(X, \mathscr{A})$ is a standard Borel space and μ a Borel measure, with a sub-σ-algebra $\mathscr{B} \subset \mathscr{A}$. In that setup, it is known (see [**?**, Proposition 3.25]) that, up to embedding $\mathscr{A}$ in a larger σ-algebra $\tilde{\mathscr{A}}$, there is a σ-algebra $\mathscr{C}$ independent of $\mathscr{B}$ such that

$$\mathscr{A} \subset \mathscr{B} \vee \mathscr{C} \mod \mu. \tag{1}$$

Such a σ-algebra $\mathscr{C}$ is called a *super-innovation*. We want to know if an equivalent result could be true in our dynamical case. Namely, we wonder whether, for any extension $\mathscr{A} \to \mathscr{B}$ on a system $\mathbf{X}$, up to embedding $\mathbf{X}$ in a larger system $\tilde{\mathbf{X}} := (\tilde{X}, \tilde{\mathscr{A}}, \tilde{\mu}, \tilde{T})$, there exists a factor σ-algebra $\mathscr{C}$ on $\tilde{\mathbf{X}}$ independent from $\mathscr{B}$ for which (**??**) would be true. The key difference here is that $\mathscr{C}$ is assumed to be a *factor* σ-algebra. However, it follows relatively easily from the definitions that a confined extension cannot have such a "dynamical" super-innovation (see Proposition **??**). This is actually our initial motivation for introducing confined extensions.

Then, we keep wondering if more general structures could describe extensions in the dynamical case. We take inspiration from the theory of dynamical filtrations introduced in [**?**], [**?**]. A dynamical filtration on a system $\mathbf{X}$ is an increasing sequence $(\mathscr{F}_n)_{n\leq 0}$ of factor σ-algebras. In [**?**], [**?**], Lanthier and de la Rue identified a significant class of dynamical filtration: *standard dynamical filtrations*. Although that notion was mainly introduced as a way to characterize the asymptotic behavior of dynamical filtrations, we are interested in understanding the local structures it imposes on a finite number of factors. Specifically, we view an extension $\mathscr{A} \to \mathscr{B}$ as a filtration with only two steps, and copy the definition of standard filtrations to get the definition of a *standard extension* (see Definition **??**). For an extension $\mathscr{A} \to \mathscr{B}$, being standard is a weaker property than admitting a super-innovation, as it only requires that $\mathscr{A} \to \mathscr{B}$ can be immersed into an extension that admits a super-innovation. Then, our goal is to figure out if there exist non-standard extensions: first, we note that there are standard confined extensions (see Lemma **??** and Theorem **??**), so this question requires additional work. At that point, we turn our attention towards T, T^{-1} transformations, which were part of our examples of confined extensions. Ultimately, we manage to show, with additional assumptions on T, that the T, T^{-1} transformation gives us a *non-standard extension*.

Our work on extensions highlights the complexity of the possible structures than can arise, making, in turn, the study of dynamical filtrations more difficult. Indeed, in the static case (i.e. when $T = \mathrm{Id}$), the study of filtrations of the form $(\mathscr{F}_n)_{n\leq 0}$ is mainly an asymptotic problem. However, in general, for a dynamical filtration $(\mathscr{F}_n)_{n\leq 0}$, it is necessary to understand both the asymptotic behavior of the filtration and the local structure of each extension $\mathscr{F}_{n+1} \to \mathscr{F}_n$. In particular, from the existence of non-standard extensions, we deduce a negative answer to a question left open in [**?**] regarding the characterization of standard dynamical filtrations (see Proposition **??**).

Throughout Chapter **??**, we define confined extensions, give interesting properties and show that they are a useful tool in the classification of extensions. Therefore, we want to see in which contexts that behavior can appear, and we start doing so in Section **??**, by using extensions that are well known in the literature. Next, we want to develop more intricate examples, to better understand the type of structures that can yield confinement.

Confined Poisson extensions

The work from Chapter **??** concerns a new kind of confined extensions, built using Poisson suspensions.

Poisson suspensions bridge two areas of ergodic theory: "classical" ergodic theory, where the invariant measure μ is a probability measure (i.e. $\mu(X) = 1$), and infinite ergodic theory, where μ is only σ-finite, i.e. there are sets $(X_n)_{n\geq 0}$ of finite measure such that $X = \bigcup_{n\geq 0} X_n$, but $\mu(X) = \infty$.

This is done using *Poisson point processes*: take a σ-finite measure space $(X, \mathscr{A}, \mu)$ and consider the set X^* of counting measures on X of the form $\sum_{i\in\mathbb{N}} \delta_{x_i}$, with $(x_i)_{i\in\mathbb{N}} \in X^{\mathbb{N}}$. The measurable sets on X^* are generated by the maps $\omega \in X^* \mapsto \omega(A) \in \mathbb{N}$, with $A \in \mathscr{A}$. This emphasizes that we do not keep track of the position of each individual point, but simply look at how they are distributed on X. Finally, the Poisson point process on $(X, \mathscr{A}, \mu)$ is given by the probability measure μ^* characterized by the following: for any disjoint sets $A_1, ..., A_n \in \mathscr{A}$, the random variables $\omega(A_1), ..., \omega(A_n)$ are independent Poisson random variables of respective parameter $\mu(A_i)$, for $i \in [\![1, n]\!]$.

If we now take an infinite measure dynamical system $\mathbf{X} := (X, \mathscr{A}, \mu, T)$, we have the corresponding probability space $(X^*, \mathscr{A}^*, \mu^*)$ and T induces a natural transformation: $T_* : \omega \mapsto T_*\omega = \omega(T^{-1}(\cdot))$. This yields the dynamical system $\mathbf{X}^* := (X^*, \mathscr{A}^*, \mu^*, T_*)$, and it is the *Poisson suspension* over $\mathbf{X}$. So we have an infinite measure system, $\mathbf{X}$, and a probability measure system, $\mathbf{X}^*$, whose properties are intertwined.

Now, if we have a factor map $\pi : \mathbf{Z} \to \mathbf{X}$ between two infinite measure systems, it yields a factor map

$$\begin{array}{rccl} \pi_* : & Z^* & \longrightarrow & X^* \\ & \omega & \longmapsto & \pi_*\omega = \omega(\pi^{-1}(\cdot)) \end{array},$$

between the associated Poisson suspensions. The corresponding extension $\mathbf{Z}^* \xrightarrow{\pi_*} \mathbf{X}^*$ is a *Poisson extension*. We are interested in a particular type of Poisson extensions: the case where $\mathbf{Z}$ is a compact extension of $\mathbf{X}$. Specifically, we take a compact group G with Haar measure m_G, and a measurable map $\varphi : X \to G$, which is called a cocycle. Then, on the measure space $(X \times G, \mathscr{A} \otimes \mathcal{B}(G), \mu \otimes m_G)$, we define the transformation

$$\begin{array}{rccl} T_\varphi : & X \times G & \longrightarrow & X \times G \\ & (x, g) & \longmapsto & (Tx, g \cdot \varphi(x)) \end{array}.$$

The system $\mathbf{Z} := (X \times G, \mathscr{A} \otimes \mathcal{B}(G), \mu \otimes m_G, T_\varphi)$ is a compact extension of $\mathbf{X}$. The Poisson extensions we study are those of the form

$$((X \times G)^*, (\mu \otimes m_G)^*, (T_\varphi)_*) \xrightarrow{\pi_*} (X^*, \mu^*, T_*).$$

We are able to get two results. First, we look at the case where φ acts as the identity map (i.e. $T_\varphi = T \times \mathrm{Id}_G$), and show that if all the Cartesian products $\mathbf{X}^{\otimes k}$, $k \geq 1$, are ergodic, the Poisson extension $\mathbf{Z}^* \xrightarrow{\pi_*} \mathbf{X}^*$ is confined (see Theorem **??**). However, in that case, the infinite measure system $\mathbf{Z}$ is not ergodic (even though the Poisson suspension $\mathbf{Z}^*$ is). Our second result concerns a case where $\mathbf{Z}$ is ergodic. Specifically, we show that, if all the Cartesian products $\mathbf{Z}^{\otimes k}$, $k \geq 1$, are ergodic, then $\mathbf{Z}^* \xrightarrow{\pi_*} \mathbf{X}^*$ is confined (see Theorem **??**). Moreover, we point out that, although the underlying extension $\mathbf{Z} \xrightarrow{\pi} \mathbf{X}$ is compact, the Poisson extension is not (see Lemma **??**).

To conclude the chapter, we provide a concrete example of an infinite measure system $\mathbf{X}$ along with a compact extension $\mathbf{Z}$ of $\mathbf{X}$ such that all Cartesian products $\mathbf{Z}^{\otimes k}$ for $k \geq 1$ are ergodic, as detailed in Theorem **??**. This shows that Theorem **??** is not void.

In conclusion, the work from Chapter **??** is a continuation of Chapter **??**, since it explores a new and original setup in which confinement appears, which adds to the examples of confined extensions from Section **??**.

Overall, Chapters **??** and **??** give us tools to better study the structure of extensions $\mathscr{A} \to \mathscr{B}$ given by a pair of factor σ-algebras. In the next chapter, we look at a situation in which there appear infinite sequences $(\mathscr{F}_n)_{n \leq 0}$ of factor σ-algebras (i.e. dynamical filtrations) where the structure of each extension $\mathscr{F}_n \to \mathscr{F}_{n-1}$ is known, and our concerns turn to the asymptotic properties of those filtrations.

A class of dynamical filtrations: weak Pinsker filtrations

For any given dynamical system $\mathbf{X} := (X, \mathscr{A}, \mu, T)$, there are several factor σ-algebras that can emerge naturally. In Chapter **??**, we look at one area of study in ergodic theory in which such factors appear: the Kolmogorov-Sinaï entropy.

This entropy was introduced by Kolmogorov and Sinaï in 1958 as an isomorphism invariant that measures the unpredictability of a dynamical system (see [**?**], [**?**], [**?**]). It is structured as follows: for any finite set A, for a random variable $\xi_0 : X \to A$, the entropy $h_\mu(\xi, T)$ is a real number assigned to the T-process $\xi = (\xi_n)_{n\in\mathbb{Z}} := (\xi_0 \circ T^n)_{n\in\mathbb{Z}}$ that indicates the average information missing to determine ξ_0 when we know the value of the past process $\xi_{]\infty,0[}$. This means that if $h_\mu(\xi, T) = 0$, by knowing the past values of ξ, we can determine its future: the process becomes deterministic. On the other hand, if $h_\mu(\xi, T) > 0$, we know that some part of the process cannot be predicted. Then the entropy of a factor σ-algebra $\mathscr{B}$ is the maximal possible entropy obtained from a $\mathscr{B}$-measurable process:

$$h_\mu(\mathscr{B}, T) := \sup\{h_\mu(\xi, T)\,;\ \xi_0 \text{ a } \mathscr{B}\text{-measurable random variable}\}.$$

In that context, a natural factor to consider is $\Pi_{\mathbf{X}}$, the *Pinsker factor* of $\mathbf{X}$: it is generated by the processes of $\mathbf{X}$ that have entropy 0. Equivalently, $\Pi_{\mathbf{X}}$ can be defined as the largest factor of $\mathbf{X}$ that has entropy 0. Questions regarding the relative structure of $\mathbf{X}$ over $\Pi_{\mathbf{X}}$, i.e. the structure of the extension $\mathbf{X} \to \Pi_{\mathbf{X}}$ (i.e. $\mathscr{A} \to \Pi_{\mathbf{X}}$), have played an important role in the development of entropy in ergodic theory ever since its inception. At the center of those questions was a particular class of systems: *K-systems*. They are the systems for which the Pinsker factor is trivial, meaning that they are entirely non-deterministic. Initially, Kolmogorov conjectured that K-systems were all isomorphic to Bernoulli shifts, i.e. systems generated by an i.i.d. process. Subsequently, Pinsker proved that any K-factor of $\mathbf{X}$ is automatically independent of the Pinsker factor, and formulated an additional conjecture, known as the "Pinsker conjecture", which suggests that, on any system $\mathbf{X} := (X, \mathscr{A}, \mu, T)$, there exists a K-factor $\mathscr{C}$ (necessarily independent of $\Pi_{\mathbf{X}}$) such that $\mathbf{X} = \Pi_{\mathbf{X}} \vee \mathscr{C}$. Using the vocabulary developed in Chapter **??**, we can rephrase this conjecture by saying that $\mathbf{X} \to \Pi_{\mathbf{X}}$ is an extension of product-type (see Definition **??**).

However, both Kolomogorov's and Pinsker's conjectures turned out to be false, highlighting the complex structure of positive entropy systems. In 1970, Ornstein solved the classification of Bernoulli shifts by showing a remarkable result [**?**], [**?**]: Bernoulli shifts with the same entropy are isomorphic. But, using the tools he developed, he was also able to build a non-Bernoulli K-system [**?**], contradicting Kolmogorov's conjecture. Following his work, many other examples were built, showing that the class of non-Bernoulli K-systems is quite broad. Ornstein also further developed his arguments to build a counterexample to Pinsker's conjecture [**?**], and he even managed to get a mixing counterexample [**?**]. Therefore, in general, we have no description of the extension $\mathbf{X} \to \Pi_{\mathbf{X}}$ given by a system over its Pinsker factor.

In 1977, Thouvenot proposed to consider systems with a new structure, which he called the *weak Pinsker property*: for every $\varepsilon > 0$, there exist a factor Π_ε of entropy ε and a Bernoulli factor $\mathcal{B}$ independent of Π_ε such that

$$\mathscr{A} = \Pi_\varepsilon \vee \mathcal{B} \mod \mu.$$

And four decades later, in 2018, Austin [**?**] proved that every ergodic system satisfies that property. This means that, on any ergodic system $\mathbf{X}$, we have factors over which the relative structure of $\mathbf{X}$ is simple, but unlike the Pinsker factor, those factors still contain some randomness. The approach we suggest is to iterate the weak Pinsker property, to obtain a new object: a *weak Pinsker filtration*. Specifically, we start by fixing a decreasing sequence $(\varepsilon_n)_{n\geq 0}$ that goes to 0, then Austin's theorem tells us that there exists a sequence of factor σ-algebras $(\mathscr{F}_n)_{n\leq 0}$ such that $\mathscr{F}_0 = \mathscr{A}$, for every $n \leq -1$, $h_\mu(\mathscr{F}_n, T) = \varepsilon_{|n|}$, $\mathscr{F}_n \subset \mathscr{F}_{n+1}$ and there exist Bernoulli factor σ-algebras $(\mathscr{B}_n)_{n\leq 0}$ such that

$$\mathscr{B}_n \perp\!\!\!\perp \mathscr{F}_{n-1} \text{ and } \mathscr{F}_n = \mathscr{F}_{n-1} \vee \mathscr{B}_n.$$

Then, we want to study those filtrations to better understand the underlying system. To that end, weak Pinsker filtrations being dynamical filtrations, we use the framework introduced in [**?**], which we also explore in Section **??**. This is our main reason for taking an interest in dynamical filtrations.

The tail σ-algebra $\bigcap_{n\leq 0} \mathscr{F}_n$ of any weak Pinsker filtration is equal to the Pinsker factor $\Pi_{\mathbf{X}}$ of the underlying system. Moreover, the main classes of dynamical filtrations considered in [**?**] have a trivial tail σ-algebra, therefore our proposed approach would mainly aim at distinguishing various structures of K-systems based on the precise asymptotic properties of their weak Pinsker filtrations. For now, we do not have many concrete results in that direction, but this could be a new and interesting way to describe and classify non-Bernoulli K-systems.

If we are going to consider weak Pinsker filtrations as a tool to understand positive entropy systems, we need to answer some basic questions, the first of which would regard uniqueness: we want to know if, on a given system $\mathbf{X}$, it is possible to find weak Pinsker filtrations with different behaviors. First, we point out that on a system $\mathbf{X}$, for a given sequence $(\varepsilon_n)_{n\geq 0}$, the choice of a weak Pinsker filtration $(\mathscr{F}_n)_{n\leq n}$ such that $h_\mu(\mathscr{F}_n, T) = \varepsilon_{|n|}$ is not unique. The question we are interested in is whether all the weak Pinsker filtrations on $\mathbf{X}$ with the same entropy are isomorphic. It turns out to be a surprisingly intricate problem, to which we only give a partial answer, in the case were $\mathbf{X}$ is a Bernoulli shift. In that case, for any decreasing values of entropy, we can find a weak Pinsker filtration $(\mathscr{F}_n)_{n\leq 0}$ of product type, i.e. there exists a sequence $(\mathscr{B}_n)_{n\leq 0}$ of mutually independent factor σ-algebras such that

$$\forall n \leq 0,\ \mathscr{F}_n = \bigvee_{k\leq n} \mathscr{B}_k.$$

Therefore, on Bernoulli shifts, the question becomes: are all weak Pinsker filtrations of product-type ? In Theorem **??**, we give a partial answer by proving that, for any weak Pinsker filtration $(\mathscr{F}_n)_{n\leq 0}$ on a Bernoulli shift, there exists a sub-sequence $(\mathscr{F}_{n_k})_{k\leq 0}$ which is of product type.

Finally, to help understand weak Pinsker filtrations, we build two explicit examples of such filtrations using cellular automata. The first one on a Bernoulli shift (see Section **??**) and the second one on a non-Bernoulli K-system (see Section **??**). However, despite the simplicity of their construction, many open questions still surround those examples.

Chapter 1

Confined extensions and non-standard dynamical filtrations

1.1 Introduction

1.1.1 Motivations

The main focus of this chapter is the study of the relation of a measure theoretic dynamical system with one of its factors. Given a dynamical system $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ with T invertible, we consider its *factor σ-algebras*, i.e. the sub-σ-algebras $\mathscr{B} \subset \mathscr{A}$ that are T-invariant. If $\mathscr{B}$ is a factor σ-algebra, we also say that $\mathscr{A}$ is an *extension* of $\mathscr{B}$, or, in short, that the pair $(\mathscr{A}, \mathscr{B})$ is an extension, which we denote $\mathscr{A} \to \mathscr{B}$. Our purpose is to understand some of the ways in which a factor σ-algebra can sit in $\mathscr{A}$. This is a key question in ergodic theory and has been studied from various points of view like, for example, in [**?**], [**?**], [**?**], [**?**] or [**?**]. Here, our approach is largely inspired by the study of filtrations. In general, a filtration is an ordered family of σ-algebras, so we can view an extension as a filtration with only two steps. We use some vocabulary and notions from the theory of filtrations initiated by Vershik (see [**?**], [**?**]) and its adaptation to dynamical filtrations, i.e. filtrations made of factor σ-algebras on a dynamical system (see [**?**], [**?**]). In return, our study of extensions enables us to get new results on dynamical filtrations.

First consider what is left when we remove the transformation T, which corresponds to the case where $T = \mathrm{Id}$. Let $(X, \mathscr{A}, \mu)$ be a standard Borel space equipped with a Borel measure and take a countably generated sub-σ-algebra $\mathscr{B} \subset \mathscr{A}$. In [**?**, §4], Rokhlin gave a complete description of the possible configurations that arise when we consider such objects. His approach was based on a close study of the conditional measures $(\mu_x)_{x\in X}$ obtained by decomposing μ over $\mathscr{B}$. In particular, if all of those measures are continuous (i.e. $\forall x, x' \in X$, $\mu_x(\{x'\}) = 0$), then $\mathscr{B}$ has an independent complement: a σ-algebra $\mathscr{C} \subset \mathscr{A}$ such that

$$\mathscr{C} \perp\!\!\!\perp \mathscr{B} \text{ and } \mathscr{A} = \mathscr{B} \vee \mathscr{C}. \tag{1.1}$$

(See Section **??** for the notation). In the general case where the measures $(\mu_x)_{x\in X}$ have atoms, Rokhlin's description is precise, but written in intricate measure theoretical terms. We prefer the probabilistic formulation found in [**?**, Proposition 3.25]: up to embedding $\mathscr{A}$ in a larger σ-algebra $\tilde{\mathscr{A}}$, there is a σ-algebra $\mathscr{C} \subset \tilde{\mathscr{A}}$ such that

$$\mathscr{C} \perp\!\!\!\perp \mathscr{B} \text{ and } \mathscr{A} \subset \mathscr{B} \vee \mathscr{C}. \tag{1.2}$$

Such a $\mathscr{C}$ is called a *super-innovation*.

The study of $(X, \mathscr{A}, \mu)$ over $\mathscr{B}$ that we briefly described above is what we refer to as the «static case». Our purpose in this chapter is to study the «dynamical case», that arises when a measure preserving transformation T is given and $\mathscr{B}$ is T-invariant.

The first question we consider regarding the dynamical case is to compare it to the setup obtained in the static case. We wonder if, in general, there always exists a *dynamical* super-innovation from $\mathscr{B}$ to $\mathscr{A}$, i.e. up to embedding $\mathbf{X}$ in a larger system $\tilde{\mathbf{X}} := (\tilde{X}, \tilde{\mathscr{A}}, \tilde{\mu}, \tilde{T})$, a $\tilde{T}$-invariant σ-algebra $\mathscr{C}$ satisfying (**??**).

We give in Example **??** an example to highlight the distinction between (static) super-innovations and dynamical super-innovations. From now on, the term «super-innovation» will only be used to refer to dynamical super-innovations.

We give several examples of extensions with no dynamical super-innovations, which include $\mathscr{A} \longrightarrow \tau^{-1}\mathscr{A}$ from Example **??**, thus showing the first difference with the static case. To get those examples, we introduce the key notion of *confined extensions*: extensions $\mathscr{A} \to \mathscr{B}$ such that for any joining of $\mathbf{X}$ with a system $\mathbf{Z} := (Z, \mathscr{C}, \rho, R)$ such that $\mathscr{C}$ is independent of $\mathscr{B}$, we have that $\mathscr{C}$ is also independent of $\mathscr{A}$(see Definition **??**). This is quite close to the notions of *stability* and *GW-property* presented in [**?**], but it is easier to use, invariant under isomorphism and applicable in a more general context, without any ergodicity assumptions. Since stability and the GW-property implicitly require that the considered extension be *relatively uniquely ergodic* (see Definition **??**), comparing confinement to those properties leads us to prove that a confined extension is always isomorphic to a relatively uniquely ergodic extension. We dedicate most of this chapter to studying the properties of confined extensions and giving various examples. One property is of particular interest: confined extensions do not admit super-innovations (see Proposition **??**).

As we mentioned earlier, our work on extensions finds an application to the study of *dynamical filtrations*, which are filtrations of the form $\mathscr{F} := (\mathscr{F}_n)_{n\leq 0}$ such that each $\mathscr{F}_n$ is T-invariant. The basic setup to study those objects was introduced in [**?**], [**?**], and we present it in Section **??**. A significant class that arises in this setup is the class of *standard* dynamical filtrations (see Definition **??**). This notion of standardness is a translation to the dynamical case of the notion of standardness introduced by Vershik in [**?**] for a filtration on a probability space. When $T = \mathrm{Id}$, those two notions are equivalent.

The fact that a dynamical filtration $\mathscr{F}$ is standard imposes some structure on the extensions $\mathscr{F}_{n+1} \to \mathscr{F}_n$. To formalize that, we introduce the notion of *standard extension* (see Definition **??**), which is a weaker property than admitting a super-innovation. The definition of standard extensions is chosen so that, for a standard dynamical filtration $\mathscr{F}$, every extension $\mathscr{F}_{n+1} \to \mathscr{F}_n$ is standard. Although it is more difficult than finding confined extensions, we also manage to build a non-standard extension, further emphasizing the variety of structures that can arise in the dynamical case.

In the static case, there are several equivalent criteria to characterize standard filtrations, one of them being I-cosiness. This notion translates to the dynamical case (see [**?**, Definition 3.7] or Section **??**). Although it was shown in [**?**] that standard dynamical filtrations are I-cosy, the converse result was left as an open question. We see in Proposition **??** that the existence of non-standard extensions gives a negative answer. This is the initial motivation for the work we present in this chapter.

Outline of the chapter

In Sections **??** and **??**, we define the main properties that we want to study and we use compact extensions to give concrete examples. In particular, compact extensions give us many examples of confined extensions (see Theorem **??**, Proposition **??**), but they are all standard (see Lemma **??**). In Section **??**, we see that T, T^{-1} transformations give non-compact confined extensions (see Theorem **??**), and we show that, provided T has the so-called PID property, we get a non-standard extension (see Theorem **??**). The PID (*pairwise independently determined*) property was introduced by Del Junco and Rudolph in [**?**], and we recall it in Definition **??**.

Finally, we give in Section **??** the details of our application of the existence of non-standard extensions to the study of dynamical filtrations.

1.1.2 Notation

A *dynamical system* is a quadruple $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ such that $(X, \mathscr{A}, \mu)$ is a Lebesgue probability space, and T is an invertible measure-preserving transformation. We denote $\mathscr{P}(X)$ the set of probability measures on $(X, \mathscr{A})$ and $\mathscr{P}_T(X) \subset \mathscr{P}(X)$ the set of T-invariant probability measures.

Let $\mathscr{B}, \mathscr{C} \subset \mathscr{A}$ be sub-σ-algebras. We write $\mathscr{B} \subset \mathscr{C}$ mod μ, if for every $B \in \mathscr{B}$, there exists $C \in \mathscr{C}$ such that $\mu(B \Delta C) = 0$. Then, $\mathscr{B} = \mathscr{C}$ mod μ if $\mathscr{B} \subset \mathscr{C}$ mod μ and $\mathscr{C} \subset \mathscr{B}$ mod μ. We denote $\mathscr{B} \vee \mathscr{C}$ the smallest σ-algebra that contains $\mathscr{B}$ and $\mathscr{C}$. We say that $\mathscr{B}$ is a *factor* σ-algebra (or T-*invariant* σ-algebra) if $T^{-1}(\mathscr{B}) = \mathscr{B}$ mod μ. Let $\mathscr{B}, \mathscr{C}$ and $\mathscr{D}$ be sub-σ-algebras of $\mathscr{A}$. We say that $\mathscr{B}$ and $\mathscr{C}$ are *relatively independent over* $\mathscr{D}$ if for any $\mathscr{B}$-measurable bounded function B and $\mathscr{C}$-measurable bounded function C

$$\mathbb{E}[BC \mid \mathscr{D}] = \mathbb{E}[B \mid \mathscr{D}]\, \mathbb{E}[C \mid \mathscr{D}] \text{ almost surely.}$$

In this case, we write $\mathscr{B} \perp\!\!\!\perp_{\mathscr{D}} \mathscr{C}$. If $\mathscr{D}$ is trivial, $\mathscr{B}$ and $\mathscr{C}$ are independent, which we denote $\mathscr{B} \perp\!\!\!\perp \mathscr{C}$.

If we have two systems $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ and $\mathbf{Y} := (Y, \mathscr{B}, \nu, S)$, a *factor map* is a measurable map $\pi : X \longrightarrow Y$ such that $\pi_* \mu = \nu$ and $\pi \circ T = S \circ \pi$, μ-almost surely. If such a map exists, we say that $\mathbf{Y}$ is a *factor* of $\mathbf{X}$ and we denote $\sigma(\pi) := \pi^{-1}(\mathscr{B})$ the σ-algebra generated by π. Conversely, we also say that $\mathbf{X}$ is an extension of $\mathbf{Y}$ or that $\mathbf{Y}$ is embedded in $\mathbf{X}$. Moreover, if there exist invariant sets $X_0 \subset X$ and $Y_0 \subset Y$ of full measure such that $\pi : X_0 \longrightarrow Y_0$ is a bijection, then π is an *isomorphism* and we write $\mathbf{X} \cong \mathbf{Y}$.

For a given factor σ-algebra $\mathscr{B}$, in general, the quadruple $(X, \mathscr{B}, \mu, T)$ is not a dynamical system since $\mathscr{B}$ need not separate points on X, and in this case $(X, \mathscr{B}, \mu)$ is not a Lebesgue probability space. However, there exist a dynamical system $\mathbf{Y}$ and a factor map $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$ such that $\sigma(\pi) = \mathscr{B}$ mod μ. Moreover, although this representation is not unique, for a given factor $\mathscr{B}$, there is a canonical construction to get a system $\mathbf{X}/_{\mathscr{B}}$ and a factor map $p_{\mathscr{B}} : \mathbf{X} \longrightarrow \mathbf{X}/_{\mathscr{B}}$ such that $\sigma(p_{\mathscr{B}}) = \mathscr{B}$ mod μ (see [**?**, Chapter 2, Section 2]).

A *joining* of $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ and $\mathbf{Z} := (Z, \mathscr{C}, \rho, R)$ is a $(T \times R)$-invariant measure λ on $X \times Z$ whose marginals are μ and ρ. It yields the dynamical system

$$\mathbf{X} \times_\lambda \mathbf{Z} := (X \times Z, \mathscr{A} \otimes \mathscr{C}, \lambda, T \times R).$$

On this system, the coordinate projections are factor maps that project onto $\mathbf{X}$ and $\mathbf{Z}$ respectively. So $\mathbf{X} \times_\lambda \mathbf{Z}$ is an extension of $\mathbf{X}$ (or $\mathbf{Z}$) via the coordinate projection. If it is not necessary to specify the measure, we will simply write $\mathbf{X} \times \mathbf{Z}$. For the product joining, we will use the notation $\mathbf{X} \otimes \mathbf{Z} := \mathbf{X} \times_{\mu\otimes\rho} \mathbf{Z}$. For the n-fold product self-joining, we will write $\mathbf{X}^{\otimes n}$.

Let $\hat{\mathbf{X}}$ be system of which $\mathbf{X}$ is a factor, via a factor map $p_X : \hat{\mathbf{X}} \longrightarrow \mathbf{X}$. Any object defined on $\mathbf{X}$ has a copy on $\hat{\mathbf{X}}$:

Definition 1.1.1. *Let* $p_{\mathbf{X}} : (\hat{X}, \hat{\mathscr{A}}, \hat{\mu}, \hat{T}) \to (X, \mathscr{A}, \mu, T)$.

- *If* $\mathscr{B} \subset \mathscr{A}$ *is a sub-*σ*-algebra, we call* $p_{\mathbf{X}}^{-1}(\mathscr{B})$ *the* copy of $\mathscr{B}$ on $\hat{\mathbf{X}}$.
- *If* $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$ *is a factor map, we say that* $\pi \circ p_{\mathbf{X}}$ *is the* copy of π on $\hat{\mathbf{X}}$.
- *The copy of* $\mathscr{A}$ *on* $\hat{\mathbf{X}}$ *will also be called the* copy of $\mathbf{X}$ on $\hat{\mathbf{X}}$.

When there is no confusion, we will still denote those copies $\mathscr{B}$*,* π*, and* $\mathbf{X}$*. We will also say that those copies are* embedded *in* $\hat{\mathbf{X}}$.

When $\hat{\mathbf{X}}$ is a self-joining of $\mathbf{X}$, all objects defined on $\mathbf{X}$ will have multiple copies on $\hat{\mathbf{X}}$: in this case we will add a number to identify each copy. For example, on $\mathbf{X}^{\times n}$, we will denote $\mathscr{B}_k := p_k^{-1}(\mathscr{B})$, where p_k is the projection on the k-th coordinate.

Assume that $\mathbf{X}$ and $\mathbf{Z}$ have a common factor, i.e. there are $\mathscr{B} \subset \mathscr{A}$ and $\tilde{\mathscr{B}} \subset \mathscr{C}$ such that $\mathbf{X}/_{\mathscr{B}} \cong \mathbf{Z}/_{\tilde{\mathscr{B}}}$. Equivalently, $\mathbf{X}$ and $\mathbf{Z}$ have a common factor if there are a system $\mathbf{Y} := (Y, \mathscr{D}, \nu, S)$ and two factor maps $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$ and $\tilde{\pi} : \mathbf{Z} \longrightarrow \mathbf{Y}$. In this case, decompose μ and λ over π and π' respectively

$$\mu := \int_Y \mu_y d\nu(y) \ \text{ and } \ \lambda := \int_Y \lambda_y d\nu(y).$$

We define the relatively independent product of $\mathbf{X}$ and $\mathbf{Z}$ over this common factor from the joining

$$\mu \otimes_{\mathbf{Y}} \lambda := \int_Y \mu_y \otimes \lambda_y d\nu(y).$$

We will denote the resulting system $\mathbf{X} \otimes_{\mathbf{Y}} \mathbf{Z}$ or $\mathbf{X} \otimes_{(\mathscr{B},\tilde{\mathscr{B}})} \mathbf{Z}$. It has the following well-known property (see [**?**, Proposition 6.11]):

Lemma 1.1.2. *Let* $\mathscr{B}, \mathscr{C} \subset \mathscr{A}$ *be two factor* σ*-algebras. Then* $\mathscr{B}$ *and* $\mathscr{C}$ *are independent if and only if, in the relatively independent product of* $\mathbf{X}$ *over* $\mathscr{B}$*, the two copies of* $\mathscr{C}$ *are independent.*

Proof. Take a factor map $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$ such that $\sigma(\pi) = \mathscr{B}$. Let f a $\mathscr{C}$-measurable random variable and denote f_1 and f_2 it's copies in $\mathbf{X} \otimes_{\mathbf{Y}} \mathbf{X}$. Using the Cauchy-Schwartz inequality, we have

$$\int_{X\times X} f_1 f_2 d(\mu \otimes_{\mathbf{Y}} \mu) = \int_Y \left(\int_X f d\mu_y\right)^2 d\nu(y) \geq \left(\int_Y \int_X f d\mu_y d\nu(y)\right)^2$$
$$= \int_{X\times X} f_1 d(\mu \otimes_{\mathbf{Y}} \mu) \cdot \int_{X\times X} f_2 d(\mu \otimes_{\mathbf{Y}} \mu).$$

Therefore, using the equality condition, this yields that f_1 and f_2 are uncorrelated if and only if $y \mapsto \int f d\mu_y$ is ν-almost surely constant. Therefore, f_1 and f_2 are uncorrelated if and only if f is independent of $\mathscr{B}$. One can now prove our lemma with a straightforward reasoning.

Indeed, if $\mathscr{C}$ is independent of $\mathscr{B}$, the above argument is not even necessary, as the result follows from the construction of the relative product. Conversely, if the copies of $\mathscr{C}$ in $\mathbf{X} \otimes_{\mathbf{Y}} \mathbf{X}$ are independent, for any $\mathscr{C}$-measurable random variable f, the copies of f in $\mathbf{X} \otimes_{\mathbf{Y}} \mathbf{X}$are independent. The above argument tells us that f is independent of $\mathscr{B}$, and therefore $\mathscr{C}$ is independent of $\mathscr{B}$. □

1.2 Product type, standardness and super-innovations for extensions

Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a dynamical system. We call *extension* on $\mathbf{X}$ a pair of factor σ-algebras $\tilde{\mathscr{A}}$, $\mathscr{B} \subset \mathscr{A}$ such that $\mathscr{B} \subset \tilde{\mathscr{A}}$, and we denote it $\tilde{\mathscr{A}} \to \mathscr{B}$. To avoid introducing too many notations, we will usually take $\tilde{\mathscr{A}} = \mathscr{A}$.

For a given extension $\mathscr{A} \to \mathscr{B}$ where $\mathscr{A}$ is the full σ-algebra on $\mathbf{X}$, we know that there is a factor map $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$, unique up to isomorphism, such that $\sigma(\pi) = \mathscr{B}$ mod μ (see Section **??**). For such a factor map, we say that the extension $\mathscr{A} \to \mathscr{B}$ is given by π, and we note it $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$. This representation of extensions is useful in the more concrete cases, but for a general discussion, we find it more convenient to write extensions in terms of T-invariant σ-algebras.

We first need a notion of isomorphism:

Definition 1.2.1. *Let* $\mathbf{X}_1 := (X_1, \mathscr{A}_1, \mu_1, T_1)$ *and* $\mathbf{X}_2 := (X_2, \mathscr{A}_2, \mu_2, T_2)$ *be dynamical systems. Two extensions* $\mathscr{C} \to \mathscr{D}$ *and* $\mathscr{I} \to \mathscr{J}$ *on* $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *respectively are* isomorphic *if there exists an isomorphism* $\Phi : \mathbf{X}_1/_{\mathscr{C}} \longrightarrow \mathbf{X}_2/_{\mathscr{I}}$ *such that* $\Phi\mathscr{D} = \mathscr{J}$ *mod* μ_2.

In the case where the extensions are given by two factor maps $\pi_1 : \mathbf{X}_1 \longrightarrow \mathbf{Y}_1$ and $\pi_2 : \mathbf{X}_2 \longrightarrow \mathbf{Y}_2$, they are isomorphic if there are two isomorphisms $\varphi : \mathbf{X}_1 \longrightarrow \mathbf{X}_2$ and $\psi : \mathbf{Y}_1 \longrightarrow \mathbf{Y}_2$ such that the following diagram is commutative:

$$\begin{array}{ccc} \mathbf{X}_1 & \xrightarrow{\varphi} & \mathbf{X}_2 \\ \downarrow \pi_1 & \circlearrowleft & \downarrow \pi_2 \\ \mathbf{Y}_1 & \xrightarrow{\psi} & \mathbf{Y}_2 \end{array}$$

We write the following definitions in terms of extensions given by factor σ-algebras, but they can all be translated for extensions given by a factor map similarly to Definition **??**. We then recall the concept of immersion (see [**?**]), which, in the theory of filtrations, expresses the idea of a «sub-filtration»:

Definition 1.2.2. *Let $\hat{\mathbf{X}} := (\hat{X}, \hat{\mathscr{A}}, \hat{\mu}, \hat{T})$ be a dynamical system and $\mathscr{A} \to \mathscr{B}$ and $\mathscr{D} \to \mathscr{E}$ be extensions defined on $\hat{\mathbf{X}}$. We say that $\mathscr{A} \to \mathscr{B}$ is* immersed *in $\mathscr{D} \to \mathscr{E}$ if we have $\mathscr{A} \subset \mathscr{D}$, $\mathscr{B} \subset \mathscr{E}$ and*

$$\mathscr{A} \perp\!\!\!\perp_{\mathscr{B}} \mathscr{E}.$$

If $\mathscr{A} \to \mathscr{B}$ is an extension defined on a dynamical system $\hat{\mathbf{X}}_1 := (\hat{X}_1, \hat{\mathscr{A}_1}, \hat{\mu}_1, \hat{T}_1)$ and $\mathscr{D} \to \mathscr{E}$ is defined on another system $\hat{\mathbf{X}}_2 := (\hat{X}_2, \hat{\mathscr{A}_2}, \hat{\mu}_2, \hat{T}_2)$, we say that $\mathscr{A} \to \mathscr{B}$ is immersible *in $\mathscr{D} \to \mathscr{E}$ if it is isomorphic to an extension on $\hat{\mathbf{X}}_2$ immersed in $\mathscr{D} \to \mathscr{E}$.*

Definition 1.2.3. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a dynamical system and $\mathscr{A} \to \mathscr{B}$ an extension defined on $\mathbf{X}$. We say that $\mathscr{A} \to \mathscr{B}$ is of* product type *if there exists a factor σ-algebra $\mathscr{C} \subset \mathscr{A}$ such that $\mathscr{B} \perp\!\!\!\perp \mathscr{C}$ and $\mathscr{A} = \mathscr{B} \vee \mathscr{C}$ mod μ.*

We can finally define standard extensions:

Definition 1.2.4. *An extension is* standard *if it is immersible in a product type extension. More explicitly, an extension $\mathscr{A} \to \mathscr{B}$ defined on $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ is standard if $\mathbf{X}$ can be embedded in a system $\hat{\mathbf{X}}$ on which there is an extension $\tilde{\mathscr{B}} \to \mathscr{B}$ and a factor σ-algebra $\mathscr{C}$ such that $\mathscr{A} \perp\!\!\!\perp_{\mathscr{B}} \tilde{\mathscr{B}}$, $\mathscr{C}$ is independent of $\tilde{\mathscr{B}}$ and $\mathscr{A} \subset \tilde{\mathscr{B}} \vee \mathscr{C}$.*

For example, we show below that all compact extensions are standard.

Definition 1.2.5. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ and $\mathbf{Y} := (Y, \mathscr{B}, \nu, S)$ be dynamical systems. An extension $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is compact if, up to isomorphism, there exists a* *compact group G, equipped with its Haar measure m_G, and a measurable map $\varphi : Y \longrightarrow G$ such that $\mathbf{X} = (Y \times G, \nu \otimes m_G, S_\varphi)$, where S_φ is given by*

$$S_\varphi : (y, g) \longmapsto (Sy, g \cdot \varphi(y)).$$

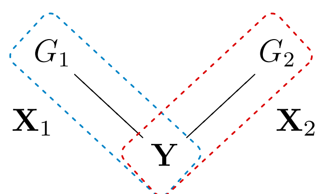
We denote it $\mathbf{Y} \ltimes_\varphi G := \mathbf{X}$.

Lemma 1.2.6. *A compact extension is standard.*

Proof. Let $\mathbf{X} := \mathbf{Y} \ltimes_\varphi G$ be a compact extension of $\mathbf{Y}$. Denote G_1 and G_2 two copies of G and consider $\mathbf{Z}$, the system on $(Y \times G_1 \times G_2, \nu \otimes m_G \otimes m_G)$ given by the transformation

$$(y, g_1, g_2) \mapsto (Sy, g_1 \cdot \varphi(y), g_2 \cdot \varphi(y)),$$

or, in short, $\mathbf{Z} := \mathbf{Y} \ltimes_{\varphi \times \varphi} (G_1 \otimes G_2)$. It is isomorphic to the 2-fold relative product of $\mathbf{X}$ over $\mathbf{Y}$ and can be viewed as in the following diagram



For $i = 1, 2$, denote $\mathbf{X}_i := \mathbf{Y} \ltimes_\varphi G_i$. The structure of the compact extension enables us to consider the factor map

$$\alpha : (y, g_1, g_2) \mapsto g_1 \cdot {g_2}^{-1},$$

which is independent of the coordinates (y, g_2) that generate $\mathbf{X}_2$, because of the invariance of the Haar measure. It is a factor map onto the identity map on G and satisfies

$$\sigma(y, g_1, g_2) = \sigma(y, g_2) \vee \sigma(\alpha).$$

This proves that the extension $\mathbf{Z} \longrightarrow \mathbf{X}_2$ is of product type. Finally, since the coordinates y, g_1 and g_2 are mutually independent, $\mathbf{X}_1$ and $\mathbf{X}_2$ are relatively independent over $\mathbf{Y}$. This means that $\mathbf{X}_1 \longrightarrow \mathbf{Y}$ is immersed in $\mathbf{Z} \longrightarrow \mathbf{X}_2$, and therefore it is standard. □

We also introduce an intermediate property between product type and standardness:

Definition 1.2.7. *Let* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *be a dynamical system and* $\mathscr{A} \to \mathscr{B}$ *an extension defined on* $\mathbf{X}$*. We say that* $\mathscr{A} \to \mathscr{B}$ admits a super-innovation *if there exists a system* $\hat{\mathbf{X}} := (\hat{X}, \hat{\mathscr{A}}, \hat{\mu}, \hat{T})$ *which extends* $\mathbf{X}$ *such that the extension* $\hat{\mathscr{A}} \to \mathscr{B}$ *is of product type, i.e. there is a factor* σ*-algebra* $\mathscr{C}$ *on* $\hat{\mathbf{X}}$ *independent of* $\mathscr{B}$ *such that* $\mathscr{A} \subset \hat{\mathscr{A}} = \mathscr{B} \vee \mathscr{C}$ *mod* $\hat{\mu}$*.*

An extension that admits a super-innovation is standard because, keeping the notations from the definition, we have that $\mathscr{A} \to \mathscr{B}$ is immersed in $\hat{\mathscr{A}} \to \mathscr{B}$ and $\hat{\mathscr{A}} \to \mathscr{B}$ is of product type.

Remark 1.2.8. For an extension given by a factor map $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$, we can rewrite the definition of standardness using super-innovations. First, we see that an extension given by π admits a super-innovation if there exists a system $\mathbf{Z}$ and a factor map $\varphi : \mathbf{Z} \otimes \mathbf{Y} \longrightarrow \mathbf{X}$ such that the following diagram is commutative:

$$\begin{array}{ccc} \mathbf{Z} \otimes \mathbf{Y} & \xrightarrow{\varphi} & \mathbf{X} \\ & {\scriptstyle p}\searrow & \downarrow{\scriptstyle \pi} \\ & & \mathbf{Y} \end{array} .$$

Next, $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is standard if there exists an extension $\tilde{\mathbf{Y}} \xrightarrow{\alpha} \mathbf{Y}$ such that $\mathbf{X} \otimes_{\mathbf{Y}} \tilde{\mathbf{Y}} \xrightarrow{\tilde{\pi}} \tilde{\mathbf{Y}}$ admits a super-innovation.

Indeed, if $\mathbf{X} \otimes_{\mathbf{Y}} \tilde{\mathbf{Y}} \xrightarrow{\tilde{\pi}} \tilde{\mathbf{Y}}$ has a super-innovation, we have a system $\mathbf{Z}$ such that $\mathbf{X} \otimes_{\mathbf{Y}} \tilde{\mathbf{Y}} \xrightarrow{\tilde{\pi}} \tilde{\mathbf{Y}}$ is immersible in $\mathbf{Z} \otimes \tilde{\mathbf{Y}} \xrightarrow{p} \tilde{\mathbf{Y}}$. Moreover, since, in $\mathbf{X} \otimes_{\mathbf{Y}} \tilde{\mathbf{Y}}$, we have that $\mathbf{X}$ is relatively independent of $\tilde{\mathbf{Y}}$ over $\mathbf{Y}$, we get that $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is immersible in $\mathbf{X} \otimes_{\mathbf{Y}} \tilde{\mathbf{Y}} \xrightarrow{\tilde{\pi}} \tilde{\mathbf{Y}}$. Therefore, $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is immersible in $\mathbf{Z} \otimes \tilde{\mathbf{Y}} \xrightarrow{p} \tilde{\mathbf{Y}}$, which means it is standard.

Conversely, if $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is standard, there are two systems $\mathbf{Z}$ and $\tilde{\mathbf{Y}}$ such that $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is immersible in $\mathbf{Z} \otimes \tilde{\mathbf{Y}} \xrightarrow{p} \tilde{\mathbf{Y}}$. This means we have two factor maps α and β and the following commutative diagram

$$\begin{array}{ccc} \mathbf{Z} \otimes \tilde{\mathbf{Y}} & \xrightarrow{p} & \tilde{\mathbf{Y}} \\ \downarrow{\scriptstyle \beta} & \circlearrowleft & \downarrow{\scriptstyle \alpha} \\ \mathbf{X} & \xrightarrow{\pi} & \mathbf{Y} \end{array} ,$$

in which $\mathbf{X}$ and $\tilde{\mathbf{Y}}$ are relatively independent over $\mathbf{Y}$. Therefore the product map $\beta \times p : Z \times \tilde{Y} \longrightarrow X \times \tilde{Y}$ is a factor map from $\mathbf{Z} \otimes \tilde{\mathbf{Y}}$ onto $\mathbf{X} \otimes_{\mathbf{Y}} \tilde{\mathbf{Y}}$ which sends $\tilde{\mathbf{Y}}$ onto $\tilde{\mathbf{Y}}$. This means that $\mathbf{Z}$ is a super-innovation for $\mathbf{X} \otimes_{\mathbf{Y}} \tilde{\mathbf{Y}} \xrightarrow{\tilde{\pi}} \tilde{\mathbf{Y}}$. $\square$

The following proposition deals with the static case.

Proposition 1.2.9. *Let* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *be a dynamical system and* $\mathscr{A} \to \mathscr{B}$ *an extension defined on* $\mathbf{X}$*. If* T *acts as the identity map on* $\mathscr{A}$*, i.e. for all* $A \in \mathscr{A}$*,* $T^{-1}A = A$ *mod* μ*, then* $\mathscr{A} \to \mathscr{B}$ *admits a super-innovation.*

Proof. Since our definition of "dynamical" super-innovation, when T acts as the identity map, is equivalent to the definition for the static case, the lemma follows from [**?**, Proposition 3.25]. □

Example 1.2.10. We give here an example to highlight the distinction between (static) super-innovations and dynamical super-innovations.

Take $(\varepsilon_n)_{n\in\mathbb{Z}}$ a sequence of independent coin tosses with 1 or -1 on each side of the coin, and $\mathscr{A}$ the associated σ-algebra. Consider the σ-algebra $\mathscr{B} := \tau^{-1}\mathscr{A} \subset \mathscr{A}$ generated by the cellular automaton

$$\begin{array}{cccc} \tau: & \{\pm 1\}^{\mathbb{Z}} & \longrightarrow & \{\pm 1\}^{\mathbb{Z}} \\ & (\varepsilon_n)_{n\in\mathbb{Z}} & \longmapsto & (\varepsilon_n \varepsilon_{n+1})_{n\in\mathbb{Z}} \end{array}.$$

From the study done in [**?**], we know that the σ-algebra generated by ε_0 gives a static super-innovation (it is even an independent complement) from $\tau^{-1}\mathscr{A}$ to $\mathscr{A}$. However, if we consider the dynamics given by the shift

$$T : (\varepsilon_n)_{n\in\mathbb{Z}} \mapsto (\varepsilon_{n+1})_{n\in\mathbb{Z}},$$

then $\tau^{-1}\mathscr{A}$ is a factor σ-algebra, but the σ-algebra generated by the projection $(\varepsilon_n)_{n\in\mathbb{Z}} \mapsto \varepsilon_0$ is not an invariant factor and therefore gives no information about the dynamical structure of $(\{\pm 1\}^{\mathbb{Z}}, \mathscr{A}, \mu, T)$ over $\tau^{-1}\mathscr{A}$. □

Remark 1.2.11. Super-innovations give an intermediate property between product type extensions and standardness. Let us give examples here to show that it is not equivalent to either of these properties. We can sum that up in the following diagram:

$$\boxed{\text{Product type}} \ \underset{\not\Leftarrow}{\Rightarrow} \ \boxed{\text{Admits a super-innovation}} \ \underset{\not\Leftarrow}{\Rightarrow} \ \boxed{\text{Standard}}$$

A standard extension with no super-innovation

As we have already mentioned, compact extensions are standard, but we will show that, in many cases, they do not admit a super-innovation. For a concrete example, consider the Anzai product given by the map

$$T : (x, y) \mapsto (x + \alpha, y + x), \text{ with } \alpha \in \mathbb{R}\backslash\mathbb{Q},$$

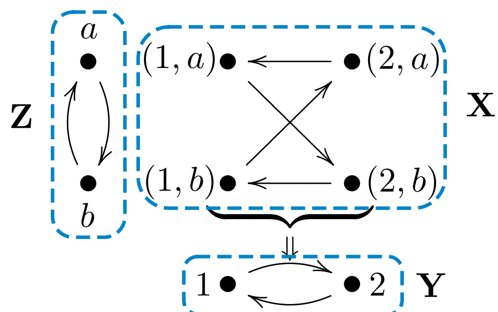
on the torus equipped with the Lebesgue measure. It is a compact, and therefore standard, extension of the rotation of angle α, but we will see that it has no super-innovation (Proposition **??** and Proposition **??**).

An extension which is not of product type but that admits a super-innovation

There already exist examples in the static case of extensions admitting super-innovations without being of product type, but here we build an *ergodic* example. We will denote $\mathbf{Z} := (\{a, b\}, \lambda, R)$ the ergodic two points system on $\{a, b\}$ (we could replace $\mathbf{Z}$ by any automorphism with no square root). Denote $\mathbf{Y} := (\{1, 2\}, \nu, S)$ the two points system on $\{1, 2\}$, and define $\mathbf{X}$ to be the system on the product space $(\{1, 2\} \times \{a, b\}, \nu \otimes \lambda)$ given by the map

$$T(i, z) := \begin{cases} (2, Rz) & \text{if } i = 1 \\ (1, z) & \text{if } i = 2 \end{cases}.$$

As we see on the following diagram, $\mathbf{X}$ is simply a cyclic four points system on $\{1, 2\} \times \{a, b\}$:



If $\mathscr{A}$ is the full σ-algebra on $\mathbf{X}$ and $\mathscr{B}$ the factor generated by the projection on $\{1, 2\}$, the extension $\mathscr{A} \to \mathscr{B}$ admits a super-innovation. Indeed, consider the transformation $\hat{R} : (z_1, z_2) \mapsto (z_2, Rz_1)$ and the system $\hat{\mathbf{Z}} := (Z \times Z, \lambda \otimes \lambda, \hat{R})$. Finally, define $\hat{\mathbf{X}}$ as the direct product of the two points system on $\{1, 2\}$, i.e. $\mathbf{Y}$, and $\hat{\mathbf{Z}}$, which extends $\mathbf{X}$ via the factor map

$$\pi : \begin{array}{ccc} \hat{\mathbf{X}} & \longrightarrow & \mathbf{X} \\ (i, z_1, z_2) & \longmapsto & (i, z_i) \end{array}.$$

An orbit on $\hat{\mathbf{X}}$ goes as follows

$$\underbrace{\begin{pmatrix}1\\a\\b\end{pmatrix} \xmapsto{S\times\hat{R}} \begin{pmatrix}2\\b\\b\end{pmatrix} \xmapsto{S\times\hat{R}} \begin{pmatrix}1\\b\\a\end{pmatrix} \xmapsto{S\times\hat{R}} \begin{pmatrix}2\\a\\a\end{pmatrix} \xmapsto{S\times\hat{R}} \begin{pmatrix}1\\a\\b\end{pmatrix}}$$

$$\pi \Big\Downarrow$$

$$\begin{pmatrix}1\\a\end{pmatrix} \xmapsto{T} \begin{pmatrix}2\\b\end{pmatrix} \xmapsto{T} \begin{pmatrix}1\\b\end{pmatrix} \xmapsto{T} \begin{pmatrix}2\\a\end{pmatrix} \xmapsto{T} \begin{pmatrix}1\\a\end{pmatrix}$$

It is then clear that $\hat{\mathbf{Z}}$ gives us the desired super-innovation.

However, the extension $\mathscr{A} \to \mathscr{B}$ is not of product type. Indeed, if it were, there would exist a system $\mathbf{W} := (W, \gamma, Q)$ and an isomorphism $\Phi : \mathbf{Y} \otimes \mathbf{W} \longrightarrow \mathbf{X}$ which sends $\mathbf{Y}$ onto $\mathbf{Y}$. In other words, there would exist two measure preserving bijections $\varphi_i : W \longrightarrow Z$ for $i = 1, 2$ such that

$$\Phi(i, w) = (i, \varphi_i(w)) \text{ almost surely.}$$

Then, the identity $\Phi \circ (S \times Q) = T \circ \Phi$ would become:

$$\varphi_2^{-1} \circ R \circ \varphi_1 = Q = \varphi_1^{-1} \circ \varphi_2.$$

This would give

$$R = \varphi_2 \circ \varphi_1^{-1} \circ \varphi_2 \circ \varphi_1^{-1} = \varphi \circ \varphi,$$

with $\varphi := \varphi_2 \circ \varphi_1^{-1}$. Since $\varphi \in \mathrm{Aut}(Z, \lambda)$, this would contradict the fact that R has no square root.

1.3 Confined extensions

In trying to build non-standard extensions, we first look for extensions with no super-innovations. The notion of *confined extension* that we introduce in this section follows that purpose, relying on the non-trivial joining properties associated to super-innovations. The link between confined extensions and super-innovations will be detailed in Section **??**.

1.3.1 Definitions, basic properties and examples

Definition 1.3.1. *Let* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *be a dynamical system and* $\mathscr{B}$ *be a factor* σ*-algebra. The extension* $\mathscr{A} \to \mathscr{B}$ *is said to be* confined *if it satisfies one of the following equivalent properties:*

(i) every 2*-fold self-joining of* $\mathbf{X}$ *in which the two copies of* $\mathscr{B}$ *are independent is the product joining;*

(ii) for every system $\mathbf{Z}$ *and every joining of* $\mathbf{X}$ *and* $\mathbf{Z}$ *in which the copies of* $\mathscr{B}$ *and* $\mathbf{Z}$ *are independent, the copies of* $\mathscr{A}$ *and* $\mathbf{Z}$ *are independent;*

(iii) for every $n \in \mathbb{N}^* \cup \{+\infty\}$*, every* n*-fold self-joining of* $\mathbf{X}$ *in which the* n *copies of* $\mathscr{B}$ *are mutually independent is the* n*-fold product joining.*

For a more explicit formulation, if the extension is given by a factor map $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$, with $\mathbf{Y} := (Y, \mathscr{C}, \nu, S)$, properties (i) and (ii) become:

(i) every $T \times T$-invariant measure λ on $X \times X$ such that $\lambda(\cdot \times X) = \lambda(X \times \cdot) = \mu$ and $(\pi \times \pi)_* \lambda = \nu \otimes \nu$ must be equal to $\mu \otimes \mu$;

(ii) for every system $\mathbf{Z} := (Z, \mathscr{D}, \rho, R)$, every $T \times R$-invariant measure λ on $X \times Z$ such that $\lambda(\cdot \times Z) = \mu$, $\lambda(X \times \cdot) = \rho$ and $(\pi \times \mathrm{Id}_Z)_* \lambda = \nu \otimes \rho$ must be equal to $\mu \otimes \rho$.

We will prove the equivalence of (i), (ii) and (iii) in Proposition **??**. Let us present here some examples of confined extensions:

Compact extensions. This is the most well known family of extensions, and it is therefore natural to start our study with them. We have a criterion for the confinement of compact extensions, which we state below. We thank Mariusz Lemańczyk for suggesting this criterion. Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$, $\mathbf{Y} := (Y, \mathscr{B}, \nu, S)$ and $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ be a compact extension, i.e. $\mathbf{X} = \mathbf{Y} \ltimes_\varphi G$, using notations from Definition **??**. Consider the ergodic decomposition of $\nu \otimes \nu$:

$$\nu \otimes \nu := \int \rho_\omega \, d\,\mathbb{P}(\omega).$$

This gives a (not necessarily ergodic) decomposition of $\nu \otimes \nu \otimes m_G \otimes m_G$:

$$\nu \otimes \nu \otimes m_G \otimes m_G = \int \rho_\omega \otimes m_G \otimes m_G \, d\,\mathbb{P}(\omega).$$

We can switch the coordinates from $(Y \times G) \times (Y \times G)$ to $Y \times Y \times G \times G$ so that $S_\varphi \times S_\varphi$ become

$$(S \times S)_{\varphi \times \varphi} : (y_1, y_2, g_1, g_2) \mapsto (Sy_1, Sy_2, g_1 \cdot \varphi(y_1), g_2 \cdot \varphi(y_2)).$$

Each measure $\rho_\omega \otimes m_G \otimes m_G$ is invariant under $(S \times S)_{\varphi \times \varphi}$.

Theorem 1.3.2. *The following are equivalent:*

(i) The extension $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ *is confined;*

(ii) The product extension $\mathbf{X} \otimes \mathbf{X} \xrightarrow{\pi \times \pi} \mathbf{Y} \otimes \mathbf{Y}$ *is relatively ergodic;*

(iii) For $\mathbb{P}$*-almost every* ω*, the measure* $\rho_\omega \otimes m_G \otimes m_G$ *is ergodic under* $(S \times S)_{\varphi \times \varphi}$.

In particular, weakly mixing compact extensions are confined.

In Section **??**, we will study the link between confinement and Robinson's notion of *stable extensions*. We can deduce the weakly mixing case of Theorem **??** from Robinson's work by combining [**?**, Corollary 3.8] and Lemma **??** (i). See Section **??** for a full proof of the theorem. Our theorem is easy to use for weakly mixing compact extensions, but for the non-weakly mixing case, our condition is more involved. In Section **??**, we give an application in the non-weakly mixing case with an Anzai skew-product (see Proposition **??**). We also give an example of a non-confined ergodic compact extension, illustrating that some condition is still necessary for compact extensions to be confined.

T, T^{-1} **transformations.** Using arguments from Lemańczyk and Lesigne [**?**], we show that, provided T^2 is ergodic, T, T^{-1} transformations yield confined extensions (see Theorem **??**). Moreover, if T is weakly mixing, the T, T^{-1} extension is confined and not compact (see Corollary **??**). Finally we show that, with more assumptions on T, we get an additional property: it is not *standard* (see Theorem **??**).

Flow extensions. A generic flow extension of a weakly mixing system is confined (see [**?**, Section 8] for the definitions). We do not discuss those examples in detail here, but Robinson showed in [**?**] that such extensions have the *GW-property* and we show below that this property yields confinement (see Proposition **??**).

Totally confined systems. We can build systems with a surprising property: they are a confined extension of any non-trivial factor. We call them *totally confined systems*. The existence of such systems with non-trivial factors follows from the work done in [**?**]: a system verifying the JP-property is totally confined. Building on similar arguments, it can be shown that any system whose reduced maximal spectral type σ is disjoint from $\sigma * \sigma$ is totally confined.

The key proposition is the following, where we prove the equivalence from Definition **??**:

Proposition 1.3.3. *Properties (i), (ii) and (iii) in Definition* **??** *are equivalent. Therefore, either of those properties can be used as a definition of confined extensions.*

Proof. **(i) ⇒ (ii).** Assume that the extension $\mathscr{A} \to \mathscr{B}$ satisfies property (i). Let $\mathbf{Z}$ be a dynamical system and let $\lambda \in \mathscr{P}(X \times Z)$ yield a joining of $\mathbf{X}$ and $\mathbf{Z}$, $\mathbf{X} \times_\lambda \mathbf{Z}$, for which $\mathscr{B}$ and $\mathbf{Z}$ are independent. The key ingredient of our proof is the relatively independent product and the Lemma **??**. We look at the system

$$(\mathbf{X} \times_\lambda \mathbf{Z}) \otimes_{\mathbf{Z}} (\mathbf{X} \times_\lambda \mathbf{Z}),$$

i.e. the relatively independent product of $\mathbf{X} \times_\lambda \mathbf{Z}$ over $\mathbf{Z}$. By taking the projection on $X \times X$, we get a self-joining γ of $\mathbf{X}$. Because of our assumption on λ and Lemma **??**, the copies of $\mathscr{B}$ in $\mathbf{X} \times_\gamma \mathbf{X}$ are independent. Therefore, by property (i), we have $\gamma = \mu \otimes \mu$. This means that on $(\mathbf{X} \times_\lambda \mathbf{Z}) \otimes_{\mathbf{Z}} (\mathbf{X} \times_\lambda \mathbf{Z})$, the copies of $\mathscr{A}$ are independent. However, using again Lemma **??**, this is only possible if $\mathscr{A}$ is independent of $\mathbf{Z}$ in $\mathbf{X} \times_\lambda \mathbf{Z}$. And this gives us property (ii).

(ii) ⇒ (iii). Assume that $\mathscr{A} \to \mathscr{B}$ satisfies property (ii). Let $n \in \mathbb{N}$ (if $n = +\infty$, we only need to show that finite families of copies of $\mathscr{A}$ are mutually independent) and let $\mathbf{Z} := \mathbf{X} \times \cdots \times \mathbf{X}$ be a n-fold joining of $\mathbf{X}$ for which the copies of $\mathscr{B}$ are independent. We show by induction on k that the family $(\mathscr{A}_1, ..., \mathscr{A}_k, \mathscr{B}_{k+1}, ..., \mathscr{B}_n)$ is mutually independent. The case $k = 0$ is simply our assumption on $\mathbf{Z}$. If the property is true for k, then $\mathscr{B}_{k+1}$ is independent of $(\mathscr{A}_1, ..., \mathscr{A}_k, \mathscr{B}_{k+2}, ..., \mathscr{B}_n)$, therefore, using (ii), we get that $\mathscr{A}_{k+1}$ is independent of $(\mathscr{A}_1, ..., \mathscr{A}_k, \mathscr{B}_{k+2}, ..., \mathscr{B}_n)$. Since the family $(\mathscr{A}_1, ..., \mathscr{A}_k, \mathscr{B}_{k+2}, ..., \mathscr{B}_n)$ is mutually independent, it implies that $(\mathscr{A}_1, ..., \mathscr{A}_k, \mathscr{A}_{k+1}, \mathscr{B}_{k+2}, ..., \mathscr{B}_n)$ is mutually independent. The case $k = n$ ends our proof.

(iii) ⇒ (i). Simply take $n = 2$.

□

Remark 1.3.4. To avoid complicating notations, we defined the notion of confined extension for an extension $\mathscr{A} \to \mathscr{B}$ defined on $\mathbf{X} := (X, \mathscr{A}, \mu, T)$, but this does not take into account extensions on $\mathbf{X}$ of the form $\mathscr{B} \to \mathscr{C}$, with $\mathscr{B} \neq \mathscr{A}$. One can still use the setup of Definition **??** because we know that there is a factor map $p_{\mathscr{B}} : \mathbf{X} \to \mathbf{X}/_{\mathscr{B}}$ such that $\sigma(p_{\mathscr{B}}) = \mathscr{B}$, and, since $\mathscr{C} \subset \mathscr{B}$, there is $\hat{\mathscr{C}}$ on $\mathbf{X}/_{\mathscr{B}}$ such that $p_{\mathscr{B}}^{-1}(\hat{\mathscr{C}}) = \mathscr{C}$ mod μ. Then, we say that $\mathscr{B} \to \mathscr{C}$ is confined if $\mathbf{X}/_{\mathscr{B}} \to \hat{\mathscr{C}}$ is confined. But this is equivalent to saying that $\mathscr{B} \to \mathscr{C}$ is confined if, for any self-joining of $\mathbf{X}$ on which the copies of $\mathscr{C}$ are independent, the copies of $\mathscr{B}$ are independent. This equivalence follows from the fact that any joining of $\mathbf{X}/_{\mathscr{B}}$ can be extended to a joining of $\mathbf{X}$, by taking the relative product over $\mathbf{X}/_{\mathscr{B}}$.

Let us state some simple manipulations possible with confined extensions:

Proposition 1.3.5. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a dynamical system. Let $\mathscr{C} \subset \mathscr{B} \subset \mathscr{A}$ be factor σ-algebras. The extension $\mathscr{A} \to \mathscr{C}$ is confined if and only if the extensions $\mathscr{A} \to \mathscr{B}$ and $\mathscr{B} \to \mathscr{C}$ are both confined.*

Proof. Assume that $\mathscr{A} \to \mathscr{B}$ and $\mathscr{B} \to \mathscr{C}$ are confined. Take $\mathbf{X} \times_\lambda \mathbf{X}$ a self-joining of $\mathbf{X}$ in which the copies of $\mathscr{C}$ are independent. Because $\mathscr{B} \to \mathscr{C}$ is confined, it follows that the copies of $\mathscr{B}$ are independent. Finally, since $\mathscr{A} \to \mathscr{B}$ is confined, the copies of $\mathscr{A}$ are independent, and λ is the product joining.

Conversely, assume that $\mathscr{A} \to \mathscr{C}$ is confined. Then $\mathscr{A} \to \mathscr{B}$ is confined because in any joining of $\mathbf{X}$ where the copies of $\mathscr{B}$ are independent, the copies of $\mathscr{C}$ are also independent (since $\mathscr{C} \subset \mathscr{B}$). For the case of $\mathscr{B} \to \mathscr{C}$, we use Remark **??**. We can take λ a self-joining of $\mathbf{X}$ on which the copies of $\mathscr{C}$ are independent. Because $\mathscr{A} \to \mathscr{C}$ is confined, λ is the product joining, and therefore the copies of $\mathscr{B}$ are independent. This proves that $\mathscr{B} \to \mathscr{C}$ is confined. □

Proposition 1.3.6. *Let $\hat{\mathbf{X}} := (\hat{X}, \hat{\mathscr{A}}, \hat{\mu}, \hat{T})$ be a dynamical system, $\mathscr{A} \to \mathscr{B}$ be a confined extension on $\hat{\mathbf{X}}$ and let $\mathscr{C} \subset \hat{\mathscr{A}}$ be a factor σ-algebra independent of $\mathscr{A}$. Then the extension $\mathscr{A} \vee \mathscr{C} \to \mathscr{B} \vee \mathscr{C}$ is confined.*

Proof. Let $\mathbf{Z}$ be a dynamical system and take a joining with $\mathbf{X}$ in which $\mathscr{B} \vee \mathscr{C}$ is independent of $\mathbf{Z}$. Therefore, $\mathscr{B}$, $\mathscr{C}$ and $\mathbf{Z}$ are mutually independent, so $\mathscr{B}$ is independent of $\mathscr{C} \vee \mathbf{Z}$. Using the definition of confined extensions, this implies that $\mathscr{A}$ is independent of $\mathscr{C} \vee \mathbf{Z}$. This means that $\mathscr{A} \vee \mathscr{C}$ is independent of $\mathbf{Z}$, which shows that the extension $\mathscr{A} \vee \mathscr{C} \to \mathscr{B} \vee \mathscr{C}$ is confined. □

Remark 1.3.7. Let us however note that some manipulations on confined extensions are not always true:

(i) For two confined extensions $\mathscr{A} \to \mathscr{B}$ and $\tilde{\mathscr{A}} \to \mathscr{B}$ over a same factor, we cannot conclude that the joint extension $\mathscr{A} \vee \tilde{\mathscr{A}} \to \mathscr{B}$ is confined. Therefore, in this case, there is no « largest confined extension of $\mathscr{B}$ », since such an extension would contain $\mathscr{A} \vee \tilde{\mathscr{A}}$. A counterexample, which relies on compact extensions, can be found in Example **??**.

(ii) Our first remark implies that the independence condition in Proposition **??** cannot be removed. Indeed, take $\mathscr{A}$, $\tilde{\mathscr{A}}$ and $\mathscr{B}$ such that $\mathscr{A} \to \mathscr{B}$ and $\tilde{\mathscr{A}} \to \mathscr{B}$ are confined but $\mathscr{A} \vee \tilde{\mathscr{A}} \to \mathscr{B}$ is not. We have

$$\mathscr{A} \vee \tilde{\mathscr{A}} \to \mathscr{B} \vee \tilde{\mathscr{A}} = \tilde{\mathscr{A}} \to \mathscr{B}.$$

By Proposition **??**, $\mathscr{A} \vee \tilde{\mathscr{A}} \to \mathscr{B} \vee \tilde{\mathscr{A}}$ cannot be confined, even though $\mathscr{A} \to \mathscr{B}$ is.

1.3.2 Confined extensions do not admit a super-innovation

We mentioned in the previous section that our initial interest in confined extensions stemmed from the fact that they do not admit super-innovations. Let us prove this here. We need the following (and basic) result:

Lemma 1.3.8. *Let (X, μ) be a probability space. Let $\mathscr{A}$, $\mathscr{B}$ and $\mathscr{C}$ be σ-algebras such that $\mathscr{B} \subset \mathscr{A}$, $\mathscr{A} \subset \mathscr{B} \vee \mathscr{C}$ mod μ and $\mathscr{C}$ is independent of $\mathscr{A}$. Then $\mathscr{A} = \mathscr{B}$ mod μ.*

Proof. Let A be a bounded real-valued $\mathscr{A}$-measurable random variable. Since $\mathscr{A} \subset \mathscr{B} \vee \mathscr{C}$, we have

$$A = \mathbb{E}[A \,|\, \mathscr{B} \vee \mathscr{C}].$$

To identify the right-hand term, take a $\mathscr{B}$-measurable bounded random variable B and a $\mathscr{C}$-measurable bounded random variable C. Since C is independent of $\mathscr{A}$ and $\mathscr{B} \subset \mathscr{A}$, we know that AB and C are independent, so

$$\begin{aligned} \mathbb{E}[ABC] &= \mathbb{E}[AB]\, \mathbb{E}[C] \\ &= \mathbb{E}\big[\mathbb{E}[A \,|\, \mathscr{B}]\, B\big]\, \mathbb{E}[C] \\ &= \mathbb{E}\big[\mathbb{E}[A \,|\, \mathscr{B}]\, BC\big]. \end{aligned}$$

So

$$A = \mathbb{E}[A \,|\, \mathscr{B} \vee \mathscr{C}] = \mathbb{E}[A \,|\, \mathscr{B}],$$

which means that A is $\mathscr{B}$-measurable. Therefore, $\mathscr{A} \subset \mathscr{B}$ mod μ. □

We can now show that super-innovations and confined extensions are incompatible:

Proposition 1.3.9. *Assume that $\mathscr{B}$ is a proper factor σ-algebra of the system $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ (i.e. we assume that $\mathscr{B} \neq \mathscr{A}$) and that the extension $\mathscr{A} \to \mathscr{B}$ admits a super-innovation. Then $\mathscr{A} \to \mathscr{B}$ is not confined.*

Proof. By assumption, $\mathscr{A} \to \mathscr{B}$ admits a super-innovation, so there exist a system $\mathbf{Z}$ and a joining $\mathbf{X} \times_\lambda \mathbf{Z}$ in which $\mathscr{B}$ and $\mathbf{Z}$ are independent and we have $\mathscr{A} \subset \mathscr{B} \vee \mathbf{Z} \mod \lambda$. However, if $\mathscr{A} \to \mathscr{B}$ were confined, $\mathscr{A}$ and $\mathbf{Z}$ would be independent. Now using Lemma **??**, we would get $\mathscr{A} = \mathscr{B} \mod \mu$, which contradicts our assumption.

□

Combining this with Proposition **??**, we get the following corollary, which is very useful when we want to show that an extension is not confined.

Corollary 1.3.9.1. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a dynamical system. If the extension $\mathscr{A} \to \mathscr{B}$ is confined, then for any factor σ-algebra $\tilde{\mathscr{A}}$ such that $\mathscr{B} \subsetneq \tilde{\mathscr{A}} \subsetneq \mathscr{A}$ mod μ, neither $\mathscr{A} \to \tilde{\mathscr{A}}$ nor $\tilde{\mathscr{A}} \to \mathscr{B}$ admit a super-innovation. In particular, there cannot be a factor in $\mathscr{A}$ independent of $\mathscr{B}$.*

Example 1.3.10. Let us use our corollary to illustrate Remark **??**. Take a system $\mathbf{Y}$, a measurable map $\varphi : \mathbf{Y} \longrightarrow G$ and consider the system $\mathbf{Y} \ltimes_\varphi G$ as defined in Definition **??**. Because of Theorem **??**, we can choose $\mathbf{Y} \ltimes_\varphi G$ so that the resulting compact extension is confined. As in the proof of Lemma **??**, take G_1 and G_2 two copies of G and consider $\mathbf{Z}$, the system on $(Y \times G_1 \times G_2, \nu \otimes m_G \otimes m_G)$ given by the transformation

$$(y, g_1, g_2) \mapsto (Sy, g_1 \cdot \varphi(y), g_2 \cdot \varphi(y)).$$

In this case, $\mathbf{Z} \longrightarrow \mathbf{Y}$ is the supremum of the compact extensions $\mathbf{Y} \ltimes_\varphi G_1 \longrightarrow \mathbf{Y}$ and $\mathbf{Y} \ltimes_\varphi G_2 \longrightarrow \mathbf{Y}$. However, the invariant map

$$\alpha : (y, g_1, g_2) \mapsto g_2 \cdot g_1^{-1}$$

is independent from $\mathbf{Y}$. Therefore, Corollary **??** tells us that $\mathbf{Z} \longrightarrow \mathbf{Y}$ is not confined.

This gives an example of a supremum of two confined extensions which is not confined.

1.3.3 Lifting results

In this section we list some properties of a dynamical system which are automatically lifted to any confined extension. Such results are not surprising. Indeed, we show in the next section that confined extensions resemble Robinson's *stable extensions* which he developed specifically to get lifting results [**?**].

Proposition 1.3.11. *Let $\mathscr{Z}$ be a family of dynamical systems. Let $\mathcal{P}$ be a property of a system that can be characterized as follows: a system $\mathbf{X}$ satisfies $\mathcal{P}$ if and only if, for every $\mathbf{Z} \in \mathscr{Z}$, $\mathbf{X}$ is disjoint from $\mathbf{Z}$. If $\mathbf{Y}$ satisfies $\mathcal{P}$ and the extension $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is confined, then $\mathbf{X}$ satisfies $\mathcal{P}$.*

Proof. It follows from Definition **??** (ii). Let $\mathcal{P}$ and $\mathscr{Z}$ be as in the proposition. Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ and $\mathbf{Y} := (Y, \mathscr{B}, \nu, S)$. Assume that $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is a confined extension and that $\mathbf{Y}$ satisfies $\mathcal{P}$. Let $\mathbf{Z} := (Z, \mathscr{C}, \rho, R) \in \mathscr{Z}$ and $\lambda \in \mathscr{P}(X \times Z)$ be a joining of $\mathbf{X}$ and $\mathbf{Z}$. We know that $(\pi \times \mathrm{Id}_Z)_* \lambda$ is a joining of $\mathbf{Y}$ and $\mathbf{Z}$, and since those systems are disjoint, it implies that $(\pi \times \mathrm{Id}_Z)_* \lambda = \nu \otimes \rho$. And, since $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is confined, Definition **??** (ii) implies that $\lambda = \mu \otimes \rho$. So $\mathbf{X}$ and $\mathbf{Z}$ are disjoint. This being true for every $\mathbf{Z} \in \mathscr{Z}$ implies that $\mathbf{X}$ satisfies $\mathcal{P}$. □

Using this proposition, we can prove that many properties are preserved under confined extensions:

1. Ergodicity: $\mathbf{X}$ is ergodic if and only if it is disjoint from every identity system (see [**?**, Theorem 6.26]).

2. Weak mixing: $\mathbf{X}$ is weakly mixing if and only if it is disjoint from every system with discrete spectrum (see [**?**, Theorem 6.27]).

3. Mild mixing: $\mathbf{X}$ is mildly mixing if and only if it is disjoint from every rigid system (see [**?**, Corollary 8.16]).

4. K-property: $\mathbf{X}$ is a K-automorphism if and only if it is disjoint from every 0-entropy system (see [**?**, Theorem 18.16]).

Remark 1.3.12. Conversely, one can use confinement to characterize disjointness from a family of systems. For example, the systems disjoint from all ergodic systems are the confined extensions of identity map systems (see [**?**, Theorem 3.1]).

We can also prove that other properties are preserved under confined extensions. In the following proposition, we prove this to be true for mixing of all orders:

Definition 1.3.13. *Let $n \geq 2$. A system $\mathbf{X}$ is n-mixing if for all measurable sets $A_1, ..., A_n \subset X$ we have*

$$\lim_{k_1,...,k_{n-1}\to\infty} \mu(A_1 \cap T^{-k_1}A_2 \cap \cdots \cap T^{-(k_1+\cdots+k_{n-1})}A_n) = \mu(A_1)\cdots\mu(A_n).$$

Proposition 1.3.14. *Let $n \geq 2$. If $\mathbf{Y}$ is n-mixing and the extension $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is confined, then $\mathbf{X}$ is n-mixing.*

Proof. Let $J_n(\mu) \subset \mathscr{P}(X^n)$ be the set of n-fold joinings of μ. We endow it with the topology given by: $\lambda_p \underset{p\to\infty}{\longrightarrow} \lambda$ if

$$\forall A_1, ..., A_n \in \mathscr{A}, \lambda_p(A_1 \times \cdots \times A_n) \underset{p\to\infty}{\longrightarrow} \lambda(A_1 \times \cdots \times A_n).$$

With this topology, $J_n(\mu)$ is a compact metrizable space (see [**?**]).

For $k := (k_1, ..., k_{n-1}) \in \mathbb{N}^{n-1}$, let us consider the off-diagonal joining on X^n:

$$\mu_k := \int_X \delta_x \otimes \delta_{T^{k_1}x} \otimes \cdots \otimes \delta_{T^{k_1+\cdots k_{n-1}}x} d\mu(x),$$

and similarly define ν_k on Y^n. We can re-write the definition of n-mixing as

$$\mu_k \underset{k\to\infty}{\longrightarrow} \mu^{\otimes n}.$$

We also define $\pi_n := \pi \times \cdots \times \pi$ and $T_n := T \times \cdots \times T$.

It suffices to check that the only limit point of $(\mu_k)_{k\in\mathbb{N}^{n-1}}$ as k goes to ∞ is $\mu^{\otimes k}$. Using the compactness of $J_n(\mu)$, fix a sequence $(k(i))_{i\in\mathbb{N}}$ on $\mathbb{N}^{n-1}$ such that, for every $\ell \in \{1, ..., n-1\}$

$$\lim_{i\to\infty} k_\ell(i) = +\infty,$$

and

$$\mu_{k(i)} \underset{i\to\infty}{\longrightarrow} \lambda,$$

for some measure λ on X^n. Clearly, λ is a T_n-invariant measure which projects to μ on each coordinate, and that means that it defines a n-fold joining of $\mathbf{X}$. Moreover, we have $(\pi_n)_*\mu_k = \nu_k$, which yields

$$(\pi_n)_*\lambda = \lim_{i\to\infty} \nu_{k(i)} = \nu^{\otimes n},$$

where we use the n-mixing property of $\mathbf{Y}$ to get the last equality. Finally, since the extension $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is confined, using property (iii) from Definition **??**, we must have $\lambda = \mu^{\otimes n}$.

Since this computation works for any converging subsequence of off-diagonal joinings, we have proved that $\mu_k \underset{k\to\infty}{\longrightarrow} \mu^{\otimes n}$, which means that $\mathbf{X}$ is n-mixing. □

Using again property (iii) from Definition **??**, we can easily prove that Del Junco and Rudolph's PID property is preserved under confined extensions. We first recall the definition:

Definition 1.3.15 (Del Junco and Rudolph [**?**])**.** *Let $\mathbf{X}$ be a dynamical system and $n \in \mathbb{N} \cup \{+\infty\}$. We say that $\mathbf{X}$ has the n-fold PID (pairwise independently determined) property if the only n-fold self-joining of $\mathbf{X}$ in which the copies of $\mathbf{X}$ are pairwise independent is the product joining.*

We then have:

Proposition 1.3.16. *If $\mathbf{Y}$ has the n-fold PID property and the extension $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is confined, then $\mathbf{X}$ has the n-fold PID property.*

Proof. It follows from property (iii) in Definition **??**. □

We can also see that confined extensions preserve the Kolmogorov-Sinaï entropy:

Proposition 1.3.17. *If $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is a confined extension, then $h(\mathbf{X}) = h(\mathbf{Y})$. Moreover, if $\mathbf{X}$ is a K-automorphism of finite entropy, the converse is true: the extension $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is confined if and only if $h(\mathbf{X}) = h(\mathbf{Y})$.*

The second part of the proposition was pointed out to us by Christophe Leuridan.

Proof. Assume that $h(\mathbf{X}) > h(\mathbf{Y})$. Using Thouvenot's relative version of Sinaï's theorem, we know there exists a Bernoulli factor of $\mathbf{X}$ with entropy $h(\mathbf{X}) - h(\mathbf{Y})$ which is independent of $\mathbf{Y}$ (we can get that result from [**?**, Proposition 2]). But we have seen in Corollary **??** that a confined extension of $\mathbf{Y}$ can have no non-trivial factors independent of $\mathbf{Y}$, therefore $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is not confined. This proves the first part of the proposition.

To prove the second part, we will use [**?**, Lemma 2]. It gives a relative version of the disjointness of K-automorphisms and 0-entropy systems: On a dynamical system $(Z, \mathscr{C}, \rho, R)$, take two R-invariant σ-algebras $\mathscr{A}$ and $\mathscr{B}$ such that $\mathscr{B} \subset \mathscr{A}$

and $h(\mathscr{A}, R) = h(\mathscr{B}, R) < \infty$. Next, take a third R-invariant σ-algebra $\mathscr{D}$ with finite entropy such that $(\mathscr{D}, R)$ has the K-property and $\mathscr{D}$ is independent of $\mathscr{B}$. Then, $\mathscr{D}$ is independent of $\mathscr{A}$.

Assume that $\mathbf{X}$ is a K-automorphism of finite entropy and that $h(\mathbf{X}) = h(\mathbf{Y})$. Let λ yield a self-joining $\mathbf{X}_1 \times_\lambda \mathbf{X}_2$ in which the copies of $\mathbf{Y}$ are independent. First, $\mathbf{Y}_2$ has the K-property and is independent of $\mathbf{Y}_1$. Moreover, $h(\mathbf{X}_1) = h(\mathbf{Y}_1)$, so [**?**, Lemma 2] tells us that $\mathbf{Y}_2$ is independent of $\mathbf{X}_1$. Similarly, $\mathbf{X}_1$ has the K-property and is independent of $\mathbf{Y}_2$. Since $h(\mathbf{X}_2) = h(\mathbf{Y}_2)$, [**?**, Lemma 2] tells us that $\mathbf{X}_1$is independent of $\mathbf{X}_2$. □

However, not all properties are preserved under confined extensions:

1. Rigidity: the system studied in Proposition **??** and Proposition **??** gives an example of a non-rigid confined extension of a rigid factor. We recall the definition of rigidity in Definition **??**.

2. The Bernoulli and loosely Bernoulli properties: we show in Theorem **??** that a T, T^{-1} transformation is a confined extension of its natural Bernoulli factor (provided T^2 is ergodic), but, when T is given by a Bernoulli shift, Kalikow proved in [**?**] that the T, T^{-1} transformation is not loosely Bernoulli (and therefore not Bernoulli either).

1.3.4 Confinement, stability and GW-property

In studying confined extensions, we found in the literature similar notions : stability and GW-property. The purpose of this section is to compare confinement to those properties.

A key notion of this section will be the following.

Definition 1.3.18. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ and $\mathbf{Y} := (Y, \mathscr{B}, \nu, S)$. We say that the extension given by a factor map $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$ is* relatively uniquely ergodic (RUE) *over ν if, for any T-invariant probability measure λ on X such that $\pi_*\lambda = \nu$, we have $\lambda = \mu$.*

Stable extensions

In [**?**], Robinson gives three notions of « stable extensions », which we recall here.

Definition 1.3.19. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ and $\mathbf{Y} := (Y, \mathscr{B}, \nu, S)$ be dynamical systems. Consider an extension given by a factor map $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$. We say that $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is:*

- stable *if* $\mathbf{X}$ *is ergodic and for every system* $\mathbf{Z} := (Z, \mathscr{C}, \rho, R)$ *such that* $\mathbf{X} \otimes \mathbf{Z}$ *is ergodic, the extension* $\mathbf{X} \otimes \mathbf{Z} \overset{\pi \times \mathrm{Id}}{\longrightarrow} \mathbf{Y} \otimes \mathbf{Z}$ *is relatively uniquely ergodic over* $\nu \otimes \rho$*.*

- n-fold self-stable *if* $\mathbf{Y}$ *is weakly mixing and the extension* $\mathbf{X}^{\otimes n} \overset{\pi_n}{\longrightarrow} \mathbf{Y}^{\otimes n}$ *is relatively uniquely ergodic over* $\nu^{\otimes n}$*. It is* self-stable *if it is* n*-fold self-stable, for every* n*.*

- weakly stable *if* $\mathbf{X}$ *is ergodic and for every system* $\mathbf{Z}$ *and every* ergodic *joining of* $\mathbf{X}$ *and* $\mathbf{Z}$ *for which* $\mathbf{Y}$ *and* $\mathbf{Z}$ *are independent,* $\mathbf{X}$ *and* $\mathbf{Z}$ *are also independent.*

Remarks. The main difference between weak stability and the first two definitions is that weak stability is an isomorphism invariant (see [**?**, Proposition 3.12]), while stability and self-stability depend on the model we consider.

Then, the distinction between stability and confinement lies mainly in the ergodicity and weak mixing assumptions in the definitions of stability. We discuss this more precisely in the next proposition.

Proposition 1.3.20. *Consider an extension given by a factor map* $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$ *and let* $n \geq 1$*. We have the following relations:*

(i) $\mathbf{X} \overset{\pi}{\longrightarrow} \mathbf{Y}$ *is* n*-fold self-stable if and only if* $\mathbf{Y}$ *is weakly mixing, and* $\mathbf{X} \overset{\pi}{\longrightarrow} \mathbf{Y}$ *is RUE and confined,*

(ii) if $\mathbf{Y}$ *is ergodic and* $\mathbf{X} \overset{\pi}{\longrightarrow} \mathbf{Y}$ *is confined, then* $\mathbf{X} \overset{\pi}{\longrightarrow} \mathbf{Y}$ *is weakly stable.*

*Since we have [**?**, Proposition 3.13.], we also get*

(iii) If $\mathbf{Y}$ *is ergodic and* $\mathbf{X} \overset{\pi}{\longrightarrow} \mathbf{Y}$ *is RUE and confined, then* $\mathbf{X} \overset{\pi}{\longrightarrow} \mathbf{Y}$ *is stable.*

Using those results, we can answer some questions left open by Robinson (see [**?**, §3.3]) on stable extensions:

1. 2-fold self-stability implies self-stability,

2. and self-stability implies stability.

Proving those result does not require the use of confined extensions, we could also have proven them directly using similar arguments to those used in Proposition **??**.

Proof of Proposition **??**. In this proof we say that an extension satisfying Definition **??** (iii) is n-confined.

We will use basic results from [**?**]:

- [**?**, Lemma 3.7, (v)] tells us that an n-fold self-stable extension is relatively uniquely ergodic.
- [**?**, Proposition 3.13] tells us that a relatively uniquely ergodic and weakly stable extension is stable.

Let us prove (i): assume that $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is n-fold self-stable. By definition, $\mathbf{Y}$ is weakly mixing. Then, [**?**, Lemma 3.7, (v)] gives the relative unique ergodicity. Finally, the n-fold self stability tells us that the only T^n-invariant measure which projects to $\nu^{\otimes n}$ is the product measure $\mu^{\otimes n}$, which means the extension $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is n-confined, and therefore confined. Conversely, assume that $\mathbf{Y}$ is weakly mixing and that $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is confined and RUE. We then know that $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is n-confined. To prove the n-fold self-stability: let $\lambda \in \mathscr{P}(X^n)$ be T^n-invariant and assume it projects to $\nu^{\otimes n}$. Using the RUE property of $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$, we get that λ is a n-fold joining of μ. Then the n-confinement of $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ implies that it is the product joining, that is $\lambda = \mu^{\otimes n}$.

We then prove (ii). Since $\mathbf{Y}$ is ergodic and $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ confined, using Proposition **??**, we get that $\mathbf{X}$ is ergodic. Then (ii) follows from the definitions.

We get (iii) by combining (ii) and [**?**, Proposition 3.13]. □

GW-property

The GW-property was introduced by Glasner and Weiss in [**?**] and named so by Robinson in [**?**]. Robinson defines this property on topological models: that is, X and Y are compact metric spaces and T, S and π are continuous maps. However, it will be more convenient here to define it in the more general setup of standard Borel spaces.

Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$, $\mathbf{Y} := (Y, \mathscr{B}, \nu, S)$ where, as in the rest of this chapter, $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ are standard Borel spaces. Let $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$ be a measurable factor map, that is, $\pi : (X, \mathscr{A}) \longrightarrow (Y, \mathscr{B})$ is a Borel map. We define $\mathscr{M}_\nu$ as the set of probability measures on X which project to ν under π:

$$\mathscr{M}_\nu := \{\gamma \in \mathscr{P}(X) \,|\, \pi_* \gamma = \nu\}.$$

Since $\mathscr{P}(X)$ equipped with the weak* topology, induced by bounded continuous (for a Polish topology on X) functions, is a Polish space (see [**?**, Theorem 17.23]),

$\mathscr{M}_\nu$ is a standard Borel space. Moreover, it is T_*-invariant, and therefore we can consider the measurable action of T_* on $\mathscr{M}_\nu$. Note that μ is a fixed point for T_*, and therefore δ_μ is a T_*-invariant measure on $\mathscr{M}_\nu$.

Definition 1.3.21. *Using the notations introduced above, we say that* $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ *has the GW-property if* $(\mathscr{M}_\nu, \delta_\mu, T_*)$ *is uniquely ergodic, i.e.* δ_μ *is the only* T_**-invariant measure on* $\mathscr{M}_\nu$*.*

Proposition 1.3.22. *The extension* $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ *has the GW-property if and only if it is relatively uniquely ergodic and confined.*

This equivalence relies on the canonical relation between quasifactors (see [**?**, Chapter 8]) and joinings. Similarly to the systems on $\mathscr{M}_\nu$ we consider here, a quasifactor of a system $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ is a dynamical system of the form $(\mathscr{P}(X), \rho, T_*)$ where ρ is T_*-invariant and $\int_{\mathscr{P}(X)} \tilde{\mu}\, d\rho(\tilde{\mu}) = \mu$.

Proof. The first part of the proof will be similar to [**?**, Proposition 2.1]. Assume that $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ has the GW-property. Let $\mathbf{Z} := (Z, \mathscr{C}, \rho, R)$ be a dynamical system and λ a measure which gives us a joining $\mathbf{X} \times_\lambda \mathbf{Z}$ where $\mathbf{Y}$ and $\mathbf{Z}$ are independent. We decompose λ over $\mathbf{Z}$:

$$\lambda = \int_Z \mu_z \otimes \delta_z \, d\rho(z).$$

We set $\varphi : z \mapsto \mu_z$ and $\gamma := \varphi_* \rho \in \mathscr{P}(\mathscr{P}(X))$. Since $\mathbf{Y}$ and $\mathbf{Z}$ are independent, we have

$$\nu \otimes \rho = (\pi \times \mathrm{Id})_* \lambda = \int_Z \pi_* \mu_z \otimes \delta_z \, d\rho(z),$$

and it follows that $\pi_* \mu_z = \nu$, ρ-almost surely. So γ is supported on $\mathscr{M}_\nu$. Moreover, using the invariance of λ and ρ, we have

$$\int_Z T_* \mu_z \otimes \delta_{Rz} \, d\rho(z) = (T \times R)_* \lambda = \lambda = \int_Z \mu_z \otimes \delta_z \, d\rho(z) = \int_Z \mu_{Rz} \otimes \delta_{Rz} \, d\rho(z),$$

so φ satisfies the equivariance condition: $\mu_{Rz} = T_* \mu_z$ ρ-almost surely. Therefore, γ is T_*-invariant. Finally, the GW-property implies $\gamma = \delta_\mu$, which in turn yields $\mu_z = \mu$, ρ-almost surely. This shows that $\lambda = \mu \otimes \rho$, so the extension is confined.

We now prove that $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is RUE: take a T-invariant measure λ such that $\pi_* \lambda = \nu$. Then $\lambda \in \mathscr{M}_\nu$ and it is a fixed point of T_*, therefore, δ_λ is a T_*-invariant measure, so, by the GW property, $\delta_\lambda = \delta_\mu$. Therefore, $\lambda = \mu$, and this proves that the extension is RUE.

Conversely, assume that $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is confined and RUE. Let ρ be a T_*-invariant measure on $\mathscr{M}_\nu$, and define the associated system: $\mathbf{Z} := (\mathscr{M}_\nu, \rho, T_*)$. We can then use the induced joining

$$\lambda := \int_{\mathscr{M}_\nu} \tilde{\mu} \otimes \delta_{\tilde{\mu}} \, d\rho(\tilde{\mu}) \in \mathscr{P}(X \times \mathscr{M}_\nu).$$

It projects to ρ on $\mathscr{M}_\nu$. Let us focus on the projection on X: $\int \tilde{\mu} \, d\rho(\tilde{\mu})$. Since ρ is supported on $\mathscr{M}_\nu$ and T_*-invariant, this measure projects to ν on Y, and is T-invariant. Now, using our relative unique ergodicity assumption, this means that $\int \tilde{\mu} \, d\rho(\tilde{\mu}) = \mu$. In conclusion, λ is a joining of $\mathbf{X}$ and $\mathbf{Z}$. Moreover, we have the computation

$$\begin{aligned} (\pi \times Id_{\mathbf{Z}})_* \lambda &= \int_{\mathscr{M}_\nu} \pi_* \tilde{\mu} \otimes \delta_{\tilde{\mu}} \, d\rho(\tilde{\mu}) \\ &= \int_{\mathscr{M}_\nu} \nu \otimes \delta_{\tilde{\mu}} \, d\rho(\tilde{\mu}) = \nu \otimes \rho. \end{aligned}$$

Then, since $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is confined, it follows that $\lambda = \mu \otimes \rho$. Given the construction of λ, it means that ρ-almost surely, $\tilde{\mu} = \mu$, which we can write as $\rho = \delta_\mu$. □

Relatively uniquely ergodic models of confined extensions

For a given extension $\mathscr{A} \to \mathscr{B}$ on $\mathbf{X} := (X, \mathscr{A}, \mu, T)$, we can find extensions defined on other dynamical systems that are isomorphic to $\mathscr{A} \to \mathscr{B}$. Moreover, as previously mentioned, once the system $\mathbf{X}$ is chosen, there exist a system $\mathbf{Y} := (Y, \mathscr{C}, \nu, S)$ and a factor map $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$ such that $\mathscr{B} = \pi^{-1}(\mathscr{C})$ mod μ, but π and $\mathbf{Y}$ are not unique. Once $\mathbf{X}$, $\mathbf{Y}$ and π are fixed, we say that $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is a *model* of $\mathscr{A} \to \mathscr{B}$. When studying properties invariant under isomorphism like confinement or standardness, the exact choice of the model of $\mathscr{A} \to \mathscr{B}$ has no impact on our results.

However, stability and the GW-property are not invariant under isomorphism, and are therefore specific to one model. The purpose of this section is to determine which of the differences between confinement and stability or the GW-property are solely due to a choice of the model. We state our result in the following theorem, where we see, in particular, that, up to the choice of the model, confinement and the GW-property are equivalent.

Theorem 1.3.23. *Let* $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ *be an extension and let* $n \geq 1$. *We have* *(i)* $\mathbf{Y}$ *is weakly mixing and* $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ *is confined if and only if* $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ *has a* n*-fold self-stable model.*

(ii) *If* $\mathbf{Y}$ *is ergodic and* $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ *is confined, then* $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ *has a stable model.*

(iii) $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ *is confined if and only if it has a model which has the GW-property.*

The key assumption on the choice of a model that emerges in Propositions **??** and **??** is relative unique ergodicity. Therefore, we want to show that

Proposition 1.3.24. *A confined extension is isomorphic to a relatively uniquely ergodic extension.*

We recall the definition of a relatively ergodic extension:

Definition 1.3.25. *Let* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$*. Denote* $\mathscr{I}_X := \{A \in \mathscr{A} \,|\, T^{-1}A = A \bmod \mu\}$*. An extension* $\mathscr{A} \to \mathscr{B}$ *on* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *is relatively ergodic if* $\mathscr{I}_X \subset \mathscr{B} \bmod \mu$*. Equivalently, if the extension is given by a factor map* $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$*, it is relatively ergodic if* $\pi^{-1}\mathscr{I}_Y = \mathscr{I}_X \bmod \mu$*.*

Our proof of Proposition **??** will be done in two steps: first we show in Lemma **??** that confined extensions are relatively ergodic, and then we show in Lemma **??** that a relatively ergodic extension admits a relatively uniquely ergodic model.

Lemma 1.3.26. *A confined extension is relatively ergodic.*

Proof. Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a dynamical system, $\mathscr{A} \to \mathscr{B}$ a confined extension and denote $\mathscr{I}_X$ the invariant factor of $\mathbf{X}$. We aim to show that $\mathscr{I}_X \subset \mathscr{B}$ mod μ. Set $\mathscr{I}_B := \mathscr{I}_X \cap \mathscr{B}$. Since T acts as the identity map on $\mathscr{I}_X$, the extension $\mathscr{I}_X \to \mathscr{I}_B$ admits a super-innovation (see Proposition **??**): there exist a probability space $(\Omega, \mathscr{E}, \mathbb{P})$, a measure preserving map $p : \Omega \longrightarrow \mathbf{X}/_{\mathscr{I}_X}$ and a σ-algebra $\mathscr{C} \subset \mathscr{E}$ independent of $p^{-1}(\mathscr{I}_B)$ such that $p^{-1}(\mathscr{I}_X) \subset p^{-1}(\mathscr{I}_B) \vee \mathscr{C}$. We now want to use $\mathscr{C}$ to get a super-innovation for the extension $\mathscr{B} \vee \mathscr{I}_X \to \mathscr{B}$.

Viewing $\mathbf{\Omega} := (\Omega, \mathscr{E}, \mathbb{P}, \mathrm{Id})$ as a dynamical system, we can set $\mathbf{Z} := (Z, \mathscr{D}, \rho, R)$ to be the relative product of $\mathbf{X}$ and $\mathbf{\Omega}$ over $\mathscr{I}_X$. Let $\tilde{\mathscr{I}_B}$, $\tilde{\mathscr{I}_X}$ and $\tilde{\mathscr{C}}$ be the respective copies of $p^{-1}(\mathscr{I}_B)$, $p^{-1}(\mathscr{I}_X)$ and $\mathscr{C}$ on $\mathbf{Z}$ obtained by taking the converse image of the projection on Ω. Also, denote $\bar{\mathscr{B}}$, $\bar{\mathscr{I}_B}$ and $\bar{\mathscr{I}_X}$ the copies of $\mathscr{B}$, $\mathscr{I}_B$

and $\mathscr{I}_X$ obtained using the projection map on X. Because $\mathscr{I}_B \subset \mathscr{I}_X$, we must have $\bar{\mathscr{I}}_B = \tilde{\mathscr{I}}_B$ mod ρ and $\bar{\mathscr{I}}_X = \tilde{\mathscr{I}}_X$ mod ρ. We then get, on $\mathbf{Z}$:

$$\bar{\mathscr{B}} \vee \bar{\mathscr{I}}_X \subset \bar{\mathscr{B}} \vee \tilde{\mathscr{C}} \mod \rho. \tag{1.3}$$

Moreover, $\tilde{\mathscr{C}}$ is independent of $\bar{\mathscr{B}}$: let C be a $\tilde{\mathscr{C}}$-measurable random variable, and consider $\mathbb{E}[C \,|\, \bar{\mathscr{B}}]$. Since $\bar{\mathscr{B}}$ is an invariant σ-algebra, we have

$$\mathbb{E}[C \,|\, \bar{\mathscr{B}}] \circ R = \mathbb{E}[C \circ R \,|\, \bar{\mathscr{B}}] = \mathbb{E}[C \,|\, \bar{\mathscr{B}}],$$

because, by construction, C is R-invariant. Therefore, $\mathbb{E}[C \,|\, \bar{\mathscr{B}}]$ is $\bar{\mathscr{I}}_B$-measurable, which means that $\mathbb{E}[C \,|\, \bar{\mathscr{B}}] = \mathbb{E}[C \,|\, \bar{\mathscr{I}}_B] = \mathbb{E}[C]$, so C is independent of $\bar{\mathscr{B}}$ under ρ. Combining this with (**??**) shows that the extension $\mathscr{B} \vee \mathscr{I}_X \to \mathscr{B}$ defined on $\mathbf{X}$ admits a super-innovation. However, by Proposition **??**, it is also confined, so Proposition **??** implies that it is a trivial extension: $\mathscr{B} \vee \mathscr{I}_X = \mathscr{B}$ mod μ and that yields $\mathscr{I}_X \subset \mathscr{B}$ mod μ. □

We continue our proof of Proposition **??** with the following lemma:

Lemma 1.3.27. *Any relatively ergodic extension $\mathscr{A} \to \mathscr{B}$ has a relatively uniquely ergodic model.*

The case when $\mathscr{B}$ (and therefore $\mathscr{A}$) is ergodic is already known: it is a result from Weiss ([**?**]). Weiss's result gives a stronger conclusion than ours since it gives a *topological* RUE model. In the non-ergodic case, we only get a model $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ where X and Y are standard Borel spaces and π is a Borel map. One could try to improve this result and build a model where $\mathbf{X}$ and $\mathbf{Y}$ are topological systems by making use of Weiss and Downarowicz's result from [**?**]. By improving the result from [**?**], one might also be able to get a model where π is continuous.

Proof. Let $\mathscr{A} \to \mathscr{B}$ be a relatively ergodic extension on $\mathbf{X} := (X, \mathscr{A}, \mu, T)$. Start by taking a system $\mathbf{Y} := (Y, \mathscr{C}, \nu, S)$ and a factor map $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$ such that $\mathscr{B} = \sigma(\pi)$ mod μ. In that case, the fact that $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is relatively ergodic means that $\mathscr{I}_X = \pi^{-1}(\mathscr{I}_Y)$ mod μ. Since $(X, \mathscr{A})$ and $(Y, \mathscr{C})$ are standard Borel spaces, we can assume that X and Y are respectively Borel subsets of some compact spaces $\tilde{X}$ and $\tilde{Y}$. Denote $\mathscr{T}_X$ and $\mathscr{T}_Y$ the induced topologies on X and Y and note that they satisfy $\mathscr{A} = \sigma(\mathscr{T}_X)$ and $\mathscr{C} = \sigma(\mathscr{T}_Y)$.

We recall that a point $x \in \tilde{X}$ is generic (for T) if there exists an invariant measure $\Phi_X(x)$ such that for every continuous bounded function $f : \tilde{X} \to \mathbb{R}$, we have

$$\frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k(x) \underset{n \to \infty}{\longrightarrow} \int_{\tilde{X}} f d\Phi_X(x). \tag{1.4}$$

Using Birkhoff's ergodic theorem and the fact that the space of continuous functions on $\tilde{X}$ is separable, we know that on $\mathbf{X}$ and $\mathbf{Y}$, almost every point is generic for an ergodic measure. Therefore, there are invariant measurable subsets $Y_0 \subset Y$ and $X_0 \subset \pi^{-1}Y_0 \subset X$ of full measure on which all points are generic for an ergodic measure and for every $x \in X_0$, $\pi \circ T(x) = S \circ \pi(x)$. Up to making X_0 slightly smaller, we can define the map $\Phi_X : X_0 \longrightarrow \mathscr{P}(X)$ that sends each point to the ergodic measure for which it is generic and for all $x \in X_0$, $\Phi_X(x)(X) = 1$. Finally, define Φ_X on $X \backslash X_0$ as any measurable map, which implies that $\Phi_X : X \longrightarrow \mathscr{P}(X)$ is measurable. We define $\Phi_Y : Y \longrightarrow \mathscr{P}(Y)$ similarly.

Those maps are measurable and generate the respective invariant factor σ-algebras of $\mathbf{X}$ and $\mathbf{Y}$. Indeed, let us show that $\sigma(\Phi_X) = \mathscr{I}_X$ mod μ: it is clear that $\sigma(\Phi_X) \subset \mathscr{I}_X$ mod μ, so we need to show the converse. This will follow from the equality, for any bounded measurable function f:

$$\int_X f d\Phi_X(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k x = \mathbb{E}[f \mid \mathscr{I}_X](x), \text{ for } \mu\text{-almost every } x. \quad (1.5)$$

We get (**??**) by first showing it for continuous functions, and then extending it to all bounded measurable functions. Once that is established, take f a T-invariant function and use (**??**) to get

$$\int_X f d\Phi_X(x) = \mathbb{E}[f \mid \mathscr{I}_X](x) = f(x), \text{ for } \mu\text{-almost every } x.$$

This shows that f is Φ_X-measurable, therefore completing the proof that $\sigma(\Phi_X) = \mathscr{I}_X$ mod μ.

Combining that with the relative ergodicity of $\mathscr{A} \to \mathscr{B}$, we have

$$\sigma(\Phi_X) = \mathscr{I}_X = \pi^{-1}(\mathscr{I}_Y) = \pi^{-1}\sigma(\Phi_Y) = \sigma(\Phi_Y \circ \pi) \text{ mod } \mu.$$

Moreover, the laws of Φ_X and Φ_Y, which we denote ξ_X and ξ_Y respectively, give the ergodic decompositions of μ and ν. Therefore, there exists a measurable map $\Gamma : (\mathscr{P}(Y), \xi_Y) \longrightarrow (\mathscr{P}(X), \xi_X)$ such that

$$\Phi_X = \Gamma \circ \Phi_Y \circ \pi \ \mu\text{-almost surely.}$$

Moreover, it is easy to verify that $\Phi_Y \circ \pi = \pi_* \circ \Phi_X$ on X_0. Finally, it yields that

$$\Gamma \circ \pi_* = \mathrm{Id}_{\mathscr{P}(X)} \ \xi_X\text{-almost surely.}$$

From that, we can use [**?**, Corollary 15.2] to deduce that there are two measurable sets $\Sigma_X \subset \mathscr{P}(X)$ and $\Sigma_Y \subset \mathscr{P}(Y)$ of full measure (for ξ_X and ξ_Y respectively) such that $\pi_* : \Sigma_X \longrightarrow \Sigma_Y$ is a measurable bijection (with $\pi_*^{-1} = \Gamma$). Finally, set $X_1 := \Phi_X^{-1}(\Sigma_X) \cap X_0$ and $Y_1 := \Phi_Y^{-1}(\Sigma_Y) \cap Y_0$ and define $\mathbf{X}_1$ (resp. $\mathbf{Y}_1$) as the restriction of $\mathbf{X}$ to X_1 (resp. $\mathbf{Y}$ to Y_1). This is well-defined because X_1 and Y_1 are invariant measurable sets that verify $\pi(X_1) \subset Y_1$. We know that $(X_1, \mathscr{A}_1) := (X_1, \mathscr{A} \cap X_1)$ and $(Y_1, \mathscr{C}_1) := (Y_1, \mathscr{C} \cap Y_1)$ are still standard Borel spaces because they are the restriction of a standard Borel space to a measurable subset.

Define $\Phi_{X_1} : X_1 \longrightarrow \mathscr{P}(X_1) \cup \{0\}$ by restricting $\Phi_X(x)$ to X_1. This is well-defined because, since $\Phi_X(x)$ is ergodic and X_1 is invariant, the restriction of $\Phi_X(x)$ to X_1 is either a probability measure or the null measure.

The extension $\mathbf{X}_1 \xrightarrow{\pi_1} \mathbf{Y}_1$ gives us the desired model. Let λ be a T-invariant probability measure on X_1 such that $\pi_*\lambda = \nu_1$ and set $\chi := (\Phi_{X_1})_*\lambda$. First, consider $\tilde{\lambda}(\cdot) := \lambda(\cdot \cap X_1)$ and integrate (**??**) over X with respect to λ: for any bounded $\mathscr{T}_X$-continuous f, using the dominated convergence theorem, we get

$$\int_{X_1} \int_X f d\Phi_X(x)\, d\lambda(x) = \lim_{n\to\infty} \int_{X_1} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(x) d\lambda(x) = \int_{X_1} f d\lambda = \int_X f\tilde{\lambda},$$

so

$$\tilde{\lambda} = \int_{X_1} \Phi_X(x) d\lambda(x).$$

So $\Phi_X(x)(X_1) = 1$ λ-almost surely and

$$\lambda = \int_{X_1} \Phi_{X_1}(x) d\lambda(x) = \int \rho \, d\chi(\rho).$$

By applying π_*, we get

$$\nu_1 = \int \pi_*\rho \, d\chi(\rho),$$

which yields $(\pi_*)_*\chi = \xi_{Y_1}$, by uniqueness of the ergodic decomposition of ν_1. Then, since π_* is a bijection, we get

$$\chi = (\pi_*^{-1})_*\xi_{Y_1} = \xi_{X_1},$$

and finally

$$\lambda = \int \rho \, d\chi(\rho) = \int \rho \, d\xi_{X_1} = \mu_1.$$

□

1.3.5 Compact extensions: a criterion for confinement

In this section, we prove the confinement criterion presented in Theorem **??** and give applications of this criterion.

A criterion for confinement of compact extensions

We recall our criterion from Theorem **??**. For a compact extension $\mathbf{X} := \mathbf{Y} \ltimes_\varphi G \longrightarrow \mathbf{Y}$, we want to show that the following are equivalent:

(i) The extension $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is confined;

(ii) The product extension $\mathbf{X} \otimes \mathbf{X} \xrightarrow{\pi\times\pi} \mathbf{Y} \otimes \mathbf{Y}$ is relatively ergodic;

(iii) Call $\int \rho_\omega \, d\,\mathbb{P}(\omega)$ the ergodic decomposition of $\nu \otimes \nu$; for $\mathbb{P}$-almost every ω, the measure $\rho_\omega \otimes m_G \otimes m_G$ is ergodic under $(S \times S)_{\varphi\times\varphi}$.

As we see in the proof below, the implications (i) $\Rightarrow$ (ii) and (ii) $\Rightarrow$ (iii) are general results for which the extension does not need to be compact. However, the implication (iii) $\Rightarrow$ (i) uses the compactness of the extension, because the main ingredient in our proof is Furstenberg's well known relative unique ergodicity result:

Lemma 1.3.28 (Furstenberg [**?**]). *Let $\mathbf{X} := (Y \times G, \nu \otimes m_G, S_\varphi)$ be the system introduced in Definition* **??** *and $\pi : (y, g) \mapsto y$. If $\mathbf{X}$ is ergodic, then, for any S_φ-invariant measure λ on $Y \times G$ which verifies $\pi_*\lambda = \nu$, we have $\lambda = \nu \otimes m_G$.*

See [**?**, Theorem 3.30] for a proof of the lemma.

Proof of Theorem **??**. **(i) $\Rightarrow$ (ii).** Assume that $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is a confined compact extension. From the definition (iii) of confined extensions, we know that the product extension $\mathbf{X} \otimes \mathbf{X} \xrightarrow{\pi\times\pi} \mathbf{Y} \otimes \mathbf{Y}$ is also confined. So it is relatively ergodic (see Lemma **??**).

(ii) $\Rightarrow$ (iii). It follows from the lemma (applied to Y^2 and G^2 instead of Y and G), which is true for any extension given by a Rokhlin cocycle:

Lemma 1.3.29. *Let $\mathbf{Y} := (Y, \mathscr{B}, \nu, S)$ be a dynamical system, (Z, ρ) a standard Borel space and $R_\bullet : Y \longrightarrow \mathrm{Aut}(Z, \rho)$ a measurable cocycle. Consider the Rokhlin extension $\mathbf{X}$ defined on $(X, \mu) := (Y \times Z, \nu \otimes \rho)$ by the skew product transformation*

$$T : (y, z) \mapsto (Sy, R_y(z)).$$

Consider

$$\nu = \int \widetilde{\nu}\, d\chi_\nu(\widetilde{\nu}),$$

the ergodic decomposition of ν*. If* $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ *is relatively ergodic, then the ergodic decomposition of* $\nu \otimes \rho$ *is*

$$\nu \otimes \rho = \int \widetilde{\nu} \otimes \rho\, d\chi_\nu(\widetilde{\nu}).$$

Proof. Let $\mathscr{I}_T$ be the σ-algebra of T-invariant sets on $Y \times Z$ and $\mathscr{I}_S$ the σ-algebra of S-invariant sets on Y. By assumption, $\pi^{-1}\mathscr{I}_S = \mathscr{I}_T \bmod \mu$, so if we take $f : Y \longrightarrow \mathbb{R}$ and $h : Z \longrightarrow \mathbb{R}$ bounded measurable functions, the independence of Y and Z gives

$$\begin{aligned}\mathbb{E}_{\nu\otimes\rho}[f(y)h(z) \,|\, \mathscr{I}_T] &= \mathbb{E}_{\nu\otimes\rho}[\mathbb{E}_{\nu\otimes\rho}[f(y)h(z) \,|\, \sigma(Y)] \,|\, \mathscr{I}_T] \\ &= \mathbb{E}_{\nu\otimes\rho}[f(y)\mathbb{E}_{\nu\otimes\rho}[h(z)] \,|\, \mathscr{I}_T] = \mathbb{E}_\nu[f(y) \,|\, \mathscr{I}_S] \cdot \mathbb{E}_\rho[h(z)].\end{aligned}$$

Since the ergodic decomposition is obtained by decomposing over the factor of invariant sets, this equality implies that

$$\int \widetilde{\nu} \otimes \rho\, d\chi_\nu(\widetilde{\nu}),$$

gives the ergodic decomposition of $\nu \otimes \rho$.

□

(iii) ⇒ (i). Assume that the condition (iii) holds. Let λ be a self-joining of $\mathbf{X}$ for which the copies of $\mathbf{Y}$ are independent, i.e. λ projects to $\nu \otimes \nu$ on $Y \times Y$. We consider the ergodic decomposition of λ

$$\lambda = \int \lambda_\omega\, d\,\mathbb{P}(\omega).$$

Then, since λ projects onto $\nu \otimes \nu$, we know that $\rho_\omega := (\pi \times \pi)_*\lambda_\omega$ gives an ergodic decomposition of $\nu \otimes \nu$. Therefore, by the uniqueness of the ergodic decomposition and our hypothesis, we get that, for $\mathbb{P}$-almost every ω, the measure $\rho_\omega \otimes m_{G\times G} = \rho_\omega \otimes m_G \otimes m_G$ is ergodic. Moreover, we can see $(X \times X, \rho_\omega \otimes m_{G\times G}, T \times T)$ as a compact extension of $(Y \times Y, \rho_\omega, S \times S)$ via the co-cycle $\psi : (y, y') \mapsto (\varphi(y), \varphi(y'))$ taking values in the compact group $G \times G$. Since $\rho_\omega \otimes m_{G\times G}$ is ergodic, Lemma **??** tells us that $\lambda_\omega = \rho_\omega \otimes m_{G\times G}$. Finally:

$$\lambda = \int \rho_\omega\, d\,\mathbb{P}(\omega) \otimes m_{G\times G} = \nu \otimes \nu \otimes m_G \otimes m_G = \mu \otimes \mu.$$

□

A non-confined ergodic compact extension

We give here a simple example of non-confined ergodic compact extension: take $\mathbf{Y} := (Y, \mathscr{B}, \nu, S)$ a weakly mixing dynamical system, α an irrationnal number and the constant cocycle

$$\begin{array}{rccc} \varphi: & Y & \longrightarrow & \mathbb{T} \\ & y & \mapsto & \alpha \end{array}$$

Then the associated compact extension $(Y \times \mathbb{T}, \nu \otimes m_{\mathbb{T}}, S_\varphi)$ defined by

$$S_\varphi : (y, z) \mapsto (Sy, z + \alpha)$$

is ergodic. Also, because the cocycle is constant, the extension is of product type, and therefore it is not confined.

A non-weakly mixing confined compact extension

Let us now turn our attention to an example illustrating our confinement criterion in the non-weakly mixing case. We consider the system $\mathbf{X}_\alpha$ on the two dimensional torus given by the following Anzai product:

$$T_\alpha : (x, y) \mapsto (x + \alpha, y + x),$$

as an extension of the system $\mathbf{Y}_\alpha$ given by the rotation

$$S_\alpha : x \mapsto x + \alpha,$$

with $\alpha \in \mathbb{R}$. We equip both systems with the Lebesgue measure.

Proposition 1.3.30. *The system $\mathbf{X}_\alpha$ is not weakly mixing but the extension $\mathbf{X}_\alpha \xrightarrow{\pi} \mathbf{Y}_\alpha$ is compact and confined.*

Proof of Proposition **??** *using Theorem* **??**. We know that $\mathbf{X}_\alpha$ is not weakly mixing because $\mathbf{Y}_\alpha$ is not. The extension $\mathbf{X}_\alpha \xrightarrow{\pi} \mathbf{Y}_\alpha$ is compact because $\mathbb{T}$ is compact. Let us show that it is confined.

Denote by ν the Lebesgue measure on the torus $\mathbb{T}$ and by $\mu := \nu \otimes \nu$ the Lebesgue measure on the 2-dimensional torus, $\mathbb{T}^2$. For $\omega \in \mathbb{T}$, define $\mu_\omega := \int \delta_x \otimes \delta_{x+\omega} d\nu(x)$. The measures $(\mu_\omega)_{\omega \in \mathbb{T}}$ give an ergodic decomposition of $\nu \otimes \nu$:

$$\nu \otimes \nu = \int \mu_\omega d\nu(\omega).$$

In light of Theorem **??**, we need to show that, for ν-almost every ω, the system $(\mathbb{T}^4, \mu_\omega \otimes \nu \otimes \nu, T_\alpha \times T_\alpha)$ is ergodic. This is isomorphic to the system on $(\mathbb{T}^3, \nu \otimes \nu \otimes \nu)$ given by

$$(x, y_1, y_2) \mapsto (x + \alpha, y_1 + x, y_2 + x + \omega), \tag{1.6}$$

which is a compact extension of $\mathbf{Y}_\alpha$ via the cocycle

$$\varphi : x \mapsto (x, x + \omega).$$

It is known (see [**?**, Theorem 3]) that this yields a non-ergodic system if and only if there exist $(n_1, n_2) \in \mathbb{Z}\backslash\{(0,0)\}$ and a measurable map $f : \mathbb{T} \longrightarrow \mathbb{U}$ such that

$$e^{2i\pi n_1 x} e^{2i\pi n_2 (x+\omega)} = f(x + \alpha)/f(x) \text{ for } \nu\text{-almost all } x.$$

By considering the Fourier series of such a function f, we see that this is only possible if there is $k \in \mathbb{Z}$ such that $n_2\omega - k\alpha \in \mathbb{Z}$.

Therefore, the system given by (**??**) is ergodic except for countably many $\omega \in \mathbb{T}$. Since ν is non-atomic, we conclude that for ν-almost every ω, the system $(\mathbb{T}^4, \mu_\omega \otimes \nu \otimes \nu, T_\alpha \times T_\alpha)$ is ergodic, and, from Theorem **??**, we know that $\mathbf{X}_\alpha \xrightarrow{\pi} \mathbf{Y}_\alpha$ is confined. □

Let us add the following result regarding this extension:

Definition 1.3.31. *A dynamical system $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ is rigid if there exists a sequence $(n_k)_{k\geq 0}$ that goes to ∞ such that, for every measurable set $A \subset X$, we have*

$$\lim_{k\to\infty} \mu(T^{n_k} A \,\Delta\, A) = 0.$$

For such a sequence, we say that $\mathbf{X}$ is (n_k)-rigid.

Proposition 1.3.32. *The factor $\mathbf{Y}_\alpha$ is rigid but the extension $\mathbf{X}_\alpha$ is not.*

It is known that $\mathbf{Y}_\alpha$ is rigid as an irrational rotation. The fact that $\mathbf{X}_\alpha$ is not rigid follows easily from [**?**, Theorem 6].

1.4 T, T^{-1} transformations

Let $\mathbf{Y} := (Y, \mathscr{B}, \nu, S)$ be a Bernoulli shift with $Y = \{-1, 1\}^{\mathbb{Z}}$, $\nu = \frac{1}{2}(\delta_{-1} + \delta_1)^{\otimes\mathbb{Z}}$ and S the shift on Y. Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a dynamical system. We introduce the system $\mathbf{Y} \ltimes \mathbf{X}$ defined on $(Y \times X, \nu \otimes \mu)$ by the transformation

$$\begin{array}{rccc} S \ltimes T : & Y \times X & \longrightarrow & Y \times X \\ & (y, x) & \mapsto & (Sy, T^{y(0)}x) \end{array}.$$

We call this map a T, T^{-1} transformation.

Most of our approach in Section **??** is based on the arguments given in [**?**], with the necessary adjustments to apply them to T, T^{-1} transformations and confined extensions. Then in Section **??**, we use new arguments to finally get a non-standard extension.

1.4.1 Confinement result for T, T^{-1} transformations

Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a dynamical system. In this section, we will prove that

Theorem 1.4.1. *If T^2 acts ergodically on $(X, \mathscr{A}, \mu)$, then $\pi : \mathbf{Y} \ltimes \mathbf{X} \longrightarrow \mathbf{Y}$ is a confined extension.*

As a consequence, we will get

Corollary 1.4.1.1. *If T is weakly mixing, then $\pi : \mathbf{Y} \ltimes \mathbf{X} \longrightarrow \mathbf{Y}$ is confined but not compact.*

Since all the confined extensions we built previously were compact, this gives us new examples.

Ergodic properties of the cocycle

We will prove the theorem using the framework proposed in [**?**, Section 6]. To do so, we first remark that the transformation $(S \ltimes T) \times (S \ltimes T)$ can easily be described using a cocycle. To see that, we define the $\mathbb{Z}^2$-action on $X \times X$ by

$$T_{(k_1,k_2)}(x_1, x_2) := (T^{k_1} x_1, T^{k_2} x_2),$$

and then $(S \ltimes T) \times (S \ltimes T)$ is isomorphic to

$$\begin{array}{cccc} S_{2,\varphi} : & Y^2 \times X^2 & \longrightarrow & Y^2 \times X^2 \\ & (y, x) & \mapsto & (S_2(y), T_{\varphi(y)}(x)) \end{array}$$

where $S_2 := S \times S$ and φ is the following cocycle

$$\begin{array}{cccc} \varphi : & Y^2 & \longrightarrow & \mathbb{Z}^2 \\ & (y_1, y_2) & \mapsto & (y_1(0), y_2(0)) \end{array}.$$

We then have to study the ergodic properties of the associated transformation

$$\begin{array}{cccc} \bar{S}_{2,\varphi} : & Y^2 \times \mathbb{Z}^2 & \longrightarrow & Y^2 \times \mathbb{Z}^2 \\ & (y, k) & \longmapsto & (S_2(y), k + \varphi(y)) \end{array}.$$

We will also note $\nu_2 := \nu \otimes \nu$.

On $\mathbb{Z}^2$, we introduce $Z_0 := \{k \in \mathbb{Z}^2 \,|\, k_1 - k_2 \text{ is even}\}$ and $Z_1 := \{k \in \mathbb{Z}^2 \,|\, k_1 - k_2 \text{ is odd}\}$. We point out that, for every $y \in Y^2$, we have

$$Z_0 + \varphi(y) = Z_0 \text{ and } Z_1 + \varphi(y) = Z_1. \tag{1.7}$$

Finally, for any subset $F \subset \mathbb{Z}^2$, we denote by λ_F the counting measure on F. We then have the following result, which can be viewed as a two dimensional version of the result in [**?**, §12]:

Lemma 1.4.2. *The ergodic components of* $(Y^2 \times \mathbb{Z}^2, \bar{S}_{2,\varphi}, \nu_2 \otimes \lambda_{\mathbb{Z}^2})$ *are* $\nu_2 \otimes \lambda_{Z_0}$ *and* $\nu_2 \otimes \lambda_{Z_1}$.

Proof. We need to show that $\nu_2 \otimes \lambda_{Z_0}$ and $\nu_2 \otimes \lambda_{Z_1}$ are ergodic. Let us consider the case of $\nu_2 \otimes \lambda_{Z_0}$.

We could adapt the arguments in [**?**, §12] to a 2-dimensional setting. We can also use the following map:

$$\begin{array}{rccc} \psi : & Y^2 \times Z_0 & \longrightarrow & {Z_0}^{\mathbb{Z}} \\ & (y,k) & \longmapsto & (k + \varphi_n(y))_{n \in \mathbb{Z}} \end{array},$$

with

$$\varphi_n(y) := \begin{cases} 0 & \text{if } n = 0 \\ \sum_{j=0}^{n-1} \varphi(S_2^j(y)) & \text{if } n > 0 \\ -\sum_{j=n}^{-1} \varphi(S_2^j(y)) & \text{if } n < 0 \end{cases}.$$

We chose the sequence $(\varphi_n)_{n \geq 0}$ because when we iterate $\bar{S}_{2,\varphi}$ we get

$$\forall n \in \mathbb{Z},\, \forall (y,k) \in Y \times Z_0,\, \bar{S}_{2,\varphi}^n(y,k) = (S_2^n y, k + \varphi_n(y)).$$

Therefore, the map ψ sends $\bar{S}_{2,\varphi}$ onto the shift on $Z_0^{\mathbb{Z}}$. Moreover, $\psi_*(\nu_2 \otimes \lambda_{Z_0})$ is the shift-invariant measure on $Z_0^{\mathbb{Z}}$ obtained by applying the symmetric random walk on the counting measure λ_{Z_0}. Therefore our system is isomorphic to the infinite measure preserving system associated to the symmetric random walk on Z_0, which is known to be a conservative and ergodic system (see [**?**, Theorem 4.5.3]). □

Corollary 1.4.2.1. *Define the map*

$$\begin{array}{rccc} H : & Y^2 & \longrightarrow & Y^2 \\ & y & \longmapsto & S_2^{N(y)}(y) \end{array},$$

where N *is the first return time to* $(0,0)$ *of the symmetric random walk associated to* $(\varphi_n)_{n \geq 0}$*. Then* H *is well-defined, measure preserving and ergodic on* (Y^2, ν_2).

Proof. The map H is simply isomorphic to the map induced by the system $(Y^2 \times Z_0, \nu_2 \otimes \lambda_{Z_0}, \bar{S}_{2,\varphi})$ on $Y^2 \times \{(0,0)\}$. Since we have seen that this system is conservative and ergodic, we get our corollary. □

Actually, a closer study of H would show that it is isomorphic to a Bernoulli shift, but this ergodicity result will be enough for our purposes.

The sequence $(\varphi_n)_{n\geq 0}$ and the return time N will be of use later because, as with $\bar{S}_{2,\varphi}$, when we iterate $S_{2,\varphi}$ we get

$$\forall n \in \mathbb{Z},\, \forall (y,x) \in Y \times X,\, S_{2,\varphi}^n(y,x) = (S_2^n y, T^{\varphi_n(y)}x).$$

And when we iterate precisely N times, the action on X reduces to the identity:

$$\forall (y,x) \in Y \times X,\, S_{2,\varphi}^{N(y)}(y,x) = (S_2^{N(y)} y, x).$$

Confinement for T, T^{-1} transformations

Using our previous section and the arguments taken from [**?**], we now prove Theorem **??**.

Lemańczyk and Lesigne gave a condition (see [**?**, Propostion 8]) for ergodic cocycles to yield stable extensions. Our proof here is a straightforward adaptation that takes into account the lack of ergodicity of the cocycle and shows that the extension is confined. We give a detailed proof for the sake of completeness.

Proof of Theorem **??**. Let λ be a $(S \ltimes T) \times (S \ltimes T)$-invariant self-joining of $\mathbf{Y} \ltimes \mathbf{X}$ whose projection on $Y \times Y$ is the product measure $\nu_2 = \nu \otimes \nu$. As we remarked previously λ being $(S \ltimes T) \times (S \ltimes T)$-invariant is, up to a permutation of the coordinates, equivalent to λ being $S_{2,\varphi}$-invariant.

Let us decompose λ over the projection on Y^2:

$$\lambda = \int_{Y^2} \delta_y \otimes \mu_y \, d\nu_2(y).$$

Since λ is $S_{2,\varphi}$-invariant, we get $\mu_{S_2(y)} = (T_{\varphi(y)})_* \mu_y$. Hence the map

$$F : (y,k) \mapsto (T_k^{-1})_* \mu_y$$

is $\bar{S}_{2,\varphi}$-invariant. So, by Lemma **??**, it is almost surely constant on $Y^2 \times Z_0$ and $Y^2 \times Z_1$. In particular, for ν_2-almost every y, we have $\mu_y = F(y,0) = \gamma_0$, where γ_0 is a probability measure on X^2. Therefore $\lambda = \nu_2 \otimes \gamma_0$. Moreover, since, for almost every y, $F(y,\cdot)$ is constant on Z_0, by taking $k=(2,0)$ and $k=(0,0)$, we get

$$((T^2\times \mathrm{Id})^{-1})_*\gamma_0=\gamma_0,$$

which we can also write as

$$(T^2\times \mathrm{Id})_*\gamma_0=\gamma_0.$$

Since both marginals of γ_0 are μ, it implies that γ_0 is a joining of (X,μ,T^2) and (X,μ,Id). However, all ergodic transformations are disjoint from the identity map, so γ_0 is a product measure, and more precisely $\gamma_0=\mu\otimes\mu$. This means that

$$\lambda=\nu\otimes\nu\otimes\mu\otimes\mu.$$

□

Remark 1.4.3. The ergodicity assumption on T^2 is necessary to get the result of the theorem. Indeed, for example, if $T^2=\mathrm{Id}$ we get $T=T^{-1}$, and then $S\ltimes T=S\times T$ cannot be confined, unless $\mathbf{X}$ is trivial (see Corollary **??**). In fact, for any transformation for which T^2 is not ergodic, the T,T^{-1} extension is not confined. Indeed, take f a non-trivial T^2-invariant function on X. Set $\xi:=(f,f\circ T)$ and $\tilde{\xi}(y,x):=\xi(x)$. Since $f\circ T^2=f$, we also get that $f\circ T^{-1}=f\circ T$. Using this, we check that

$$\tilde{\xi}\circ S\ltimes T(y,x)=(f(T^{y(0)}x),f(T^{y(0)+1}x))=(f(Tx),f(x)),$$

is $\tilde{\xi}$-measurable, meaning that $\sigma(\tilde{\xi})$ is invariant under the T,T^{-1} transformation. But, by construction, it is independent of the Y coordinate, which means (using again Corollary **??**) that the extension is not confined.

Our additional ergodicity assumption on T^2 is here to compensate the lack of ergodicity from the cocycle that arises from the fact that the random walk set by φ is not ergodic. Indeed, if we modify Y and take random variables uniformly distributed on $\{-1,0,1\}$, the random walk it generates on $\mathbb{Z}^2$ is ergodic, the cocycle is ergodic as well and we get the result of the theorem by only assuming that T is ergodic. This is the setup to which Lemańczyk and Lesigne applied their stability criterion.

Relative weak mixing of T,T^{-1} transformations

In this section, we prove Corollary **??** using the notion of relative weak mixing (see [**?**]).

Definition 1.4.4. *Let* $\mathbf{U}$ *and* $\mathbf{V}$ *be dynamical systems and* $\alpha : \mathbf{U} \longrightarrow \mathbf{V}$ *a factor map. The extension* $\mathbf{U} \xrightarrow{\alpha} \mathbf{V}$ *is relatively weakly mixing if the extension*

$$\mathbf{U} \otimes_{\mathbf{V}} \mathbf{U} \longrightarrow \mathbf{V}$$

is relatively ergodic.

It is well-known that relatively weakly mixing extensions cannot be compact (see [**?**, Chapter 6, Part 4]). It can also be seen from the construction we give in the proof of Lemma **??**.

Proof of Corollary **??**. Assume that T is weakly mixing. Therefore $(T \times T)^2$ acts ergodically on $(X \times X, \mathscr{A} \otimes \mathscr{A}, \mu \otimes \mu)$. One can check that the system, that we note $\mathbf{Y} \ltimes (\mathbf{X} \otimes \mathbf{X})$, given on $(Y \times X \times X, \nu \otimes \mu \otimes \mu)$ by the transformation

$$S \ltimes (T \times T) : \begin{array}{ccc} Y \times X \times X & \longrightarrow & Y \times X \times X \\ (y, x, x') & \mapsto & (Sy, T^{y(0)}x, T^{y(0)}x') \end{array},$$

is the relatively independent product of $\mathbf{Y} \ltimes \mathbf{X}$ by itself over $\mathbf{Y}$. Using Theorem **??**, we know that $\mathbf{Y} \ltimes (\mathbf{X} \otimes \mathbf{X}) \longrightarrow \mathbf{Y}$ is confined, and, by Lemma **??**, it is relatively ergodic. It follows that $\mathbf{Y} \ltimes \mathbf{X} \longrightarrow \mathbf{Y}$ is relatively weakly mixing, and therefore not compact. □

1.4.2 A non-standard T, T^{-1} extension

The goal of this section is to show the following result:

Theorem 1.4.5. *If* $\mathbf{X}$ *has the* 4*-fold PID property and* T^2 *acts ergodically on* $(X, \mathscr{A}, \mu)$*, then the extension* $\mathbf{Y} \ltimes \mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ *is not standard.*

We introduced the PID property in Definition **??**. It is known that there are systems satisfying this property. For example, it is shown in [**?**] that all finite rank mixing transformations have the PID property. We will use Ryzhikov's result from [**?**, §1, Section 2] which states that if $\mathbf{X}$ has the 4-fold PID property, any joining λ of $\mathbf{X}$ with any two systems $\mathbf{Z}_1$ and $\mathbf{Z}_2$ that is pairwise independent has to be the product joining. In fact, we will only need a simplified version, which we write here as a consequence of Lemma **??**:

Lemma 1.4.6. *Assume that* $\mathbf{X}$ *has the* 4*-fold PID property. Let* $\mathbf{Z}$ *be any dynamical system. Consider a joining* $\mathbf{X}_1 \times \mathbf{X}_2 \times \mathbf{Z}$ *where* $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *are copies of* $\mathbf{X}$*. If this joining is pairwise independent, then it is the product joining.*

Proof. Let λ be a pairwise independent joining on $X \times X \times Z$. Take $\mathbf{W}$ the system given by the relatively independent product of λ over $\mathbf{Z}$, and denote $\mathbf{X}'_1$, $\mathbf{X}'_2$, $\mathbf{X}''_1$ and $\mathbf{X}''_2$ the copies of $\mathbf{X}$ on $\mathbf{W}$. Using our assumption on λ, we know that $\mathbf{X}'_i$ and $\mathbf{X}''_j$ are relatively independent over $\mathbf{Z}$ and are both independent of $\mathbf{Z}$, so $\mathbf{X}'_i$ and $\mathbf{X}''_j$ are independent. Therefore the quadruplet $(\mathbf{X}'_1, \mathbf{X}'_2, \mathbf{X}''_1, \mathbf{X}''_2)$ is pairwise independent. Next, the 4-fold PID property tells us that this quadruplet is mutually independent. Therefore, $\mathbf{X}'_1 \vee \mathbf{X}'_2$ and $\mathbf{X}''_1 \vee \mathbf{X}''_2$ are independent. Finally, using Lemma **??** once more, we know that $\mathbf{X}_1 \vee \mathbf{X}_2$ and $\mathbf{Z}$ are independent, which implies that λ is the product joining. □

Proof of Theorem **??**. Let us take an extension $\tilde{\mathbf{Y}} \xrightarrow{\alpha} \mathbf{Y}$. Consider the system $\mathbf{W} := \tilde{\mathbf{Y}} \ltimes \mathbf{X} = (\tilde{Y} \times X, \tilde{\nu} \otimes \mu, Q)$ with Q being the map

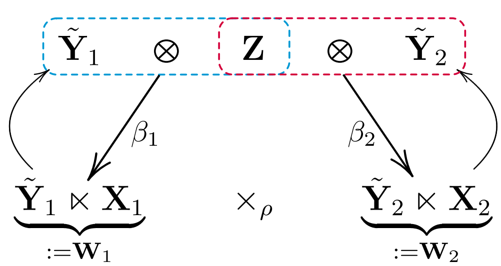
$$Q : (\tilde{y}, x) \mapsto (\tilde{S}\tilde{y}, T^{y(0)}x),$$

using the notation $y := \alpha(\tilde{y})$. One can check that $\mathbf{W}$ is the relatively independent product of $\mathbf{Y} \ltimes \mathbf{X}$ and $\tilde{\mathbf{Y}}$ over $\mathbf{Y}$. Following Remark **??**, we need to show that the extension $\tilde{\mathbf{Y}} \ltimes \mathbf{X} \xrightarrow{\tilde{\pi}} \tilde{\mathbf{Y}}$ admits no super-innovation.

Assume that $\mathbf{W} \xrightarrow{\tilde{\pi}} \tilde{\mathbf{Y}}$ has a super-innovation. Therefore there is a system $\mathbf{Z}$ and a factor map $\beta : \tilde{\mathbf{Y}} \otimes \mathbf{Z} \longrightarrow \mathbf{W}$ such that $\tilde{\pi} \circ \beta(\tilde{y}, z) = \tilde{y}$. We then use $\mathbf{Z}$ the get a self-joining of $\mathbf{W}$: we start with the product space $\tilde{\mathbf{Y}}_1 \otimes \tilde{\mathbf{Y}}_2 \otimes \mathbf{Z}$ and we get two copies of $\mathbf{W}$ by considering

$$\beta_1 : (\tilde{y}_1, \tilde{y}_2, z) \mapsto \beta(\tilde{y}_1, z) \;\text{ and }\; \beta_2 : (\tilde{y}_1, \tilde{y}_2, z) \mapsto \beta(\tilde{y}_2, z).$$

Formally, we consider the joining $\rho \in \mathscr{P}(\tilde{Y} \times X \times \tilde{Y} \times X)$ defined as the law of the factor map (β_1, β_2). Our goal is to show that ρ is the product joining. We represent the construction of ρ in the following diagram:



To make our notation explicit, we define $p_{\mathbf{X}_1}$, $p_{\mathbf{X}_2}$, $p_{\tilde{\mathbf{Y}}_1}$ and $p_{\tilde{\mathbf{Y}}_2}$ as the coordinate projections on $\tilde{Y} \times X \times \tilde{Y} \times X$. In the joining given by ρ, we have the two following properties:

(i) the copies of $\tilde{\mathbf{Y}}$, generated by $p_{\tilde{\mathbf{Y}}_1}$ and $p_{\tilde{\mathbf{Y}}_2}$, are independent,

(ii) we have the additional independence: for $i = 1, 2$ we get that $p_{\mathbf{X}_i}$ is independent of $(p_{\tilde{\mathbf{Y}}_1}, p_{\tilde{\mathbf{Y}}_2})$.

Indeed, (i) follows from the independence of $\mathbf{Z}$ and $\tilde{\mathbf{Y}}$ in the construction of ρ. To get (ii) (with $i = 1$ for example), first note that, by construction, $(p_{\mathbf{X}_1}, p_{\tilde{\mathbf{Y}}_1})$ is independent of $p_{\tilde{\mathbf{Y}}_2}$. Since $p_{\mathbf{X}_1}$ and $p_{\tilde{\mathbf{Y}}_1}$ are independent, this yields that $p_{\mathbf{X}_1}$ is independent of $(p_{\tilde{\mathbf{Y}}_1}, p_{\tilde{\mathbf{Y}}_2})$, which gives us (ii).

For the rest of this proof, we switch the coordinates and view ρ as a measure on $\tilde{Y} \times \tilde{Y} \times X \times X$. We note a point on this space as $(\tilde{y}_1, \tilde{y}_2, x_1, x_2)$ or $(\tilde{y}, x)$, for short. Moreover, we use the notation $y := \alpha \times \alpha(\tilde{y})$, and this enables us to define $\tilde{\varphi}$, $\tilde{\varphi}_n$ and $\tilde{N}$ as maps on $\tilde{Y} \times \tilde{Y}$ by setting

$$\tilde{\varphi}(\tilde{y}) := \varphi(y),\ \tilde{\varphi}_n(\tilde{y}) = \varphi_n(y)\ \text{ and }\ \tilde{N}(\tilde{y}) := N(y).$$

We also define $\hat{N}$ on $\tilde{Y} \times \tilde{Y} \times X \times X$ by setting $\hat{N}(\tilde{y}, x) := N(y)$. We now get back to our proof.

From (ii), we know that each $p_{\mathbf{X}_i}$ is independent of $(p_{\tilde{\mathbf{Y}}_1}, p_{\tilde{\mathbf{Y}}_2})$. Moreover, from Theorem **??**, we know that $p_{\mathbf{X}_1}$ and $p_{\mathbf{X}_2}$ are independent, so, under ρ, $(p_{\tilde{\mathbf{Y}}_1}, p_{\tilde{\mathbf{Y}}_2})$, $p_{\mathbf{X}_1}$ and $p_{\mathbf{X}_2}$ are pairwise independent. We then want to use Lemma **??** to get the mutual independence, but we do not have a transformation θ on $(\tilde{Y} \times \tilde{Y}, \tilde{\nu}_2)$ such that ρ is $(\theta \times T \times T)$-invariant. Therefore, our strategy below is, instead of considering $(p_{\tilde{\mathbf{Y}}_1}, p_{\tilde{\mathbf{Y}}_2})$, to only consider the conditional law of $(p_{\mathbf{X}_1}, p_{\mathbf{X}_2})$ given $(p_{\tilde{\mathbf{Y}}_1}, p_{\tilde{\mathbf{Y}}_2})$, which takes its values in $\mathscr{P}(X \times X)$. Our goal is then to find a suitable transformation on $\mathscr{P}(X \times X)$ and an invariant joining on $\mathscr{P}(X \times X) \times X \times X$ in order to finally apply Lemma **??**.

Since $p_{\tilde{\mathbf{Y}}_1}$ and $p_{\tilde{\mathbf{Y}}_2}$ are independent, ρ projects onto $\tilde{\nu}_2 := \tilde{\nu} \otimes \tilde{\nu}$ on $\tilde{Y}^2$. We decompose ρ over $\tilde{Y}^2$:

$$\rho = \int_{\tilde{Y}^2} \delta_{\tilde{y}} \otimes \mu_{\tilde{y}}\, d\tilde{\nu}_2(\tilde{y}).$$

where each $\mu_{\tilde{y}}$ is a probability measure on $X \times X$. We consider the following action on $\mathscr{P}(X \times X)$:

$$\forall k \in \mathbb{Z}^2,\ \theta_k : \gamma \mapsto (T_k)_* \gamma = (T^{k_1} \times T^{k_2})_* \gamma.$$

The fact that ρ is $Q \times Q$-invariant yields:

$$\int_{\tilde{Y}^2} \delta_{\tilde{S}_2(\tilde{y})} \otimes \mu_{\tilde{S}_2(\tilde{y})}\, d\tilde{\nu}_2(\tilde{y}) = \int_{\tilde{Y}^2} \delta_{\tilde{y}} \otimes \mu_{\tilde{y}}\, d\tilde{\nu}_2(\tilde{y})$$

$$= \rho = (Q \times Q)_* \rho = \int_{\tilde{Y}^2} \delta_{\tilde{S}_2(\tilde{y})} \otimes (T_{\tilde{\varphi}(\tilde{y})})_* \mu_{\tilde{y}}\, d\tilde{\nu}_2(\tilde{y}),$$

where $\tilde{\varphi}(\tilde{y}) = \varphi(y) = (y_1(0), y_2(0))$. So

$$\mu_{\tilde{S}_2(\tilde{y})} = (T_{\tilde{\varphi}(\tilde{y})})_* \mu_{\tilde{y}} = \theta_{\tilde{\varphi}(\tilde{y})} \mu_{\tilde{y}} \text{ almost surely.} \tag{1.8}$$

We now set the map $\mu_\bullet : (\tilde{y}, x) \mapsto \mu_{\tilde{y}}$ and consider the triplet $(\mu_\bullet, p_{\mathbf{X}_1}, p_{\mathbf{X}_2})$. It is invariant under

$$(Q \times Q)^{\hat{N}}.$$

Indeed:

$$(p_{\mathbf{X}_1}, p_{\mathbf{X}_2}) \circ (Q \times Q)^{\hat{N}(\tilde{y},x)}(\tilde{y}, x) = T_{\varphi_{N(y)}(y)}(x) = x,$$

and, using (**??**):

$$\mu_\bullet \circ (Q \times Q)^{\hat{N}(\tilde{y},x)}(\tilde{y}, x) = \theta_{\varphi_{N(y)}(y)} \mu_{\tilde{y}} = \mu_{\tilde{y}},$$

because $\varphi_{N(y)}(y) = 0$.

However, it follows from Corollary **??** that

$$(p_{\mathbf{Y}_1}, p_{\mathbf{Y}_2}) := (\alpha \circ p_{\tilde{\mathbf{Y}}_1}, \alpha \circ p_{\tilde{\mathbf{Y}}_2})$$

is ergodic under $(Q \times Q)^{\hat{N}}$. Therefore, since ergodic transformations are disjoint from any identity map, the factor map $(\mu_\bullet, p_{\mathbf{X}_1}, p_{\mathbf{X}_2})$ is independent of $(p_{\mathbf{Y}_1}, p_{\mathbf{Y}_2})$. Set $\hat{\rho} := (\mu_\bullet, p_{\mathbf{X}_1}, p_{\mathbf{X}_2})_* \rho$, which is a probability measure on $\mathscr{P}(X \times X) \times X \times X$. Also consider $\hat{\rho}_y$ the conditional law of $(\mu_\bullet, p_{\mathbf{X}_1}, p_{\mathbf{X}_2})$ given $(p_{\mathbf{Y}_1}, p_{\mathbf{Y}_2})$ under ρ. Using the $Q \times Q$-invariance of ρ and (**??**), we get that

$$\int_{Y^2} \delta_{S_2(y)} \otimes \hat{\rho}_{S_2(y)}\, d\nu_2(y) = \int_{Y^2} \delta_y \otimes \hat{\rho}_y\, d\nu_2(y) = (p_{\mathbf{Y}_1}, p_{\mathbf{Y}_2}, \mu_\bullet, p_{\mathbf{X}_1}, p_{\mathbf{X}_2})_* \rho$$

$$= (p_{\mathbf{Y}_1}, p_{\mathbf{Y}_2}, \mu_\bullet, p_{\mathbf{X}_1}, p_{\mathbf{X}_2})_* (Q \times Q)_* \rho = \int_{Y^2} \delta_{S_2(y)} \otimes (\theta_{\varphi(y)} \times T_{\varphi(y)})_* \hat{\rho}_y\, d\nu_2(y).$$

Therefore

$$\hat{\rho}_{S^2 y} = (\theta_{\varphi(y)} \times T_{\varphi(y)})_* \hat{\rho}_y \text{ almost surely.}$$

Moreover, because $(\mu_\bullet, p_{\mathbf{X}_1}, p_{\mathbf{X}_2})$ is independent of $(p_{\mathbf{Y}_1}, p_{\mathbf{Y}_2})$, we know that, ν_2-almost surely, $\hat{\rho}_y = \hat{\rho}$. This yields:

$$\hat{\rho} = (\theta_{\varphi(y)} \times T_{\varphi(y)})_* \hat{\rho} \text{ almost surely.}$$

In particular, since $\nu_2(\{\varphi(y) = (1,1)\}) > 0$, we get that $\hat{\rho}$ is $(\theta_{(1,1)} \times T \times T)$-invariant.

Let us study $\hat{\rho}$ more closely: using Theorem **??**, we know that, under ρ, $p_{\mathbf{X}_1}$ and $p_{\mathbf{X}_2}$ are independent. Moreover, using the property (ii) introduced earlier, we get that, for $i = 1, 2$, $p_{\mathbf{X}_i}$ is independent of $\mu_\bullet$. So $\hat{\rho}$ is a pairwise independent joining. We then use Ryzhikov's result from [**?**] as expressed in Lemma **??** and the 4-fold PID property of $\mathbf{X}$ to conclude that $\hat{\rho}$ is the product joining.

Therefore, $(p_{\mathbf{X}_1}, p_{\mathbf{X}_2})$ is independent of $\mu_\bullet$. However, since $\mu_\bullet = \mu_{\tilde{y}}$ is the conditional law of $(p_{\mathbf{X}_1}, p_{\mathbf{X}_2})$ given $(p_{\tilde{\mathbf{Y}}_1}, p_{\tilde{\mathbf{Y}}_2})$, the conditional law of $(p_{\mathbf{X}_1}, p_{\mathbf{X}_2})$ given $\mu_\bullet$ is also $\mu_\bullet$. As a result, $\mu_\bullet$ is constant mod ρ, which means that $(p_{\mathbf{X}_1}, p_{\mathbf{X}_2})$ is independent of $(p_{\tilde{\mathbf{Y}}_1}, p_{\tilde{\mathbf{Y}}_2})$.

In conclusion:

$$\rho = \tilde{\nu} \otimes \mu \otimes \tilde{\nu} \otimes \mu.$$

Using the notation introduced in the beginning of this proof, this means that β_1 and β_2 are independent. Since $\tilde{\mathbf{Y}}_1 \otimes \tilde{\mathbf{Y}}_2 \otimes \mathbf{Z}$ is the 2-fold relative product of $\tilde{\mathbf{Y}} \otimes \mathbf{Z}$ over $\mathbf{Z}$, Lemma **??** tells us that β is independent of $\mathbf{Z}$. Finally, Lemma **??** yields that β is $\tilde{\mathbf{Y}}$-measurable, implying that $\mathbf{W}$ must be $\tilde{\mathbf{Y}}$-measurable, which is absurd. So $\mathbf{W} \xrightarrow{\tilde{\pi}} \tilde{\mathbf{Y}}$ admits no super-innovation. □

Given our work in Section **??**, it would be natural to try to prove that $\tilde{\mathbf{Y}} \ltimes \mathbf{X} \xrightarrow{\tilde{\pi}} \tilde{\mathbf{Y}}$ has no super-innovation by showing that it is confined. This would show that $\mathbf{Y} \ltimes \mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ has a stronger property: it would be *hyper-confined*, as we define below

Definition 1.4.7. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a dynamical system. We say that $\mathscr{A} \to \mathscr{B}$ on $\mathbf{X}$ is* hyper-confined *if for every $\beta : \tilde{\mathbf{X}} \longrightarrow \mathbf{X}$ and for every extension $\tilde{\mathscr{B}} \to \mathscr{B}$ on $\tilde{\mathbf{X}}$ such that $\mathscr{A} \perp\!\!\!\perp_{\mathscr{B}} \tilde{\mathscr{B}}$, we have that $\mathscr{A} \vee \tilde{\mathscr{B}} \to \tilde{\mathscr{B}}$ is confined.*

Equivalently, an extension given by a factor map $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$ is hyper-confined *if, for every extension $\tilde{\mathbf{Y}} \xrightarrow{\alpha} \mathbf{Y}$, the extension $\mathbf{X} \otimes_{\mathbf{Y}} \tilde{\mathbf{Y}} \xrightarrow{\tilde{\pi}} \tilde{\mathbf{Y}}$ is confined.*

However, in trying to prove that the T, T^{-1} extension is hyper-confined, we get a similar setup to the proof we gave above, but with a self-joining of $\mathbf{W}$ that does not need to verify property (ii). Since we did not manage to complete the proof in that more general case, the following question remains open:

Question 1.4.8. *Is $\pi : \mathbf{Y} \ltimes \mathbf{X} \longrightarrow \mathbf{Y}$ hyper-confined ? More generally, is it possible to build a hyper-confined extension ?*

1.5 Application to non-standard dynamical filtrations

Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$. A *dynamical filtration* is a pair $(\mathscr{F}, T)$ such that $\mathscr{F} := (\mathscr{F}_n)_{n\leq 0}$ is a filtration in discrete negative time on $\mathscr{A}$ and each $\mathscr{F}_n$ is T-invariant. The theory of dynamical filtrations was initiated by Paul Lanthier in ([**?**], [**?**]). The definitions we give in Section **??** for extensions are based on the theory of filtrations we present here, therefore the process is very similar.

Definition 1.5.1. *Let $(\mathscr{F}, T_1)$ be a dynamical filtration on $\mathbf{X}_1 := (X_1, \mathscr{A}_1, \mu_1, T_1)$ and $(\mathscr{G}, T_2)$ a dynamical filtration on $\mathbf{X}_2 := (X_2, \mathscr{A}_2, \mu_2, T_2)$. We say that $(\mathscr{F}, T_1)$ and $(\mathscr{G}, T_2)$ are isomorphic if there is an isomorphism $\Phi : \mathbf{X}_1/_{\mathscr{F}_0} \to \mathbf{X}_2/_{\mathscr{G}_0}$ such that, for all $n \leq 0$, $\Phi(\mathscr{F}_n) = \mathscr{G}_n$ mod μ_2.*

If $\mathscr{F}$ and $\mathscr{G}$ are defined on the same system $(X, \mathscr{A}, \mu, T)$, we say that $(\mathscr{F}, T)$ is immersed in $(\mathscr{G}, T)$ if for every $n \leq 0$, $\mathscr{F}_n \subset \mathscr{G}_n$ and we have the following relative independence:

$$\mathscr{F}_{n+1} \perp\!\!\!\perp_{\mathscr{F}_n} \mathscr{G}_n.$$

In general, we say that $(\mathscr{F}, T_1)$ is immersible in $(\mathscr{G}, T_2)$ if there exists some dynamical filtration isomorphic to $(\mathscr{F}, T_1)$ immersed in $(\mathscr{G}, T_2)$.

We can then define our main classes of filtrations:

Definition 1.5.2. *Let $(\mathscr{F}, T)$ be a dynamical filtration on $\mathbf{X} := (X, \mathscr{A}, \mu, T)$. It is of product type if there is a sequence $(\mathscr{C}_n)_{n\leq 0}$ of mutually independent factor σ-algebras such that*

$$\forall n \leq 0,\ \mathscr{F}_n = \bigvee_{k\leq n} \mathscr{C}_k \ \text{mod}\ \mu.$$

It is standard if it is immersible in a product type dynamical filtration.

We chose our definitions to get the following properties:

Proposition 1.5.3. *We have*

1. *If $(\mathscr{F}, T)$ is of product type, then every extension $\mathscr{F}_{n+1} \to \mathscr{F}_n$ is of product type.*

2. *If* $(\mathscr{F}, T)$ *is standard, then every extension* $\mathscr{F}_{n+1} \to \mathscr{F}_n$ *is standard.*

Proof. It follows from the definitions. □

Below, we use this proposition to build a non-standard filtration using a non-standard extension.

In the static case (i.e. when $T = \mathrm{Id}$), the existence of super-innovations implies that the standardness of a filtration is an asymptotic property [**?**, Proposition 3.38]: $(\mathscr{F}_n)_{n\leq 0}$ is standard if and only if there is a $n_0 \leq 0$ such that $(\mathscr{F}_n)_{n\leq n_0}$ is standard. In the dynamical case, the existence of extensions without super-innovations puts that asymptotic property in question. Then, the existence of non-standard extensions shows that standardness is not an asymptotic property for dynamical filtrations (using Proposition **??**). This is the main guideline for what we do next.

One of the main goals in the study of dynamical filtration is to find a standardness criterion, and for that purpose, dynamical I-cosiness was introduced in [**?**]. It relies on the notion of *real time joinings* of filtrations: by that we mean a system $\mathbf{Z} := (Z, \mathscr{C}, \lambda, R)$ and a pair $((\mathscr{F}', R), (\mathscr{F}'', R))$ defined on $\mathbf{Z}$ such that both $(\mathscr{F}', R)$ and $(\mathscr{F}'', R)$ are isomorphic to $(\mathscr{F}, T)$ and immersed in $(\mathscr{F}' \vee \mathscr{F}'', R)$.

A dynamical filtration $\mathscr{F}$ is I-cosy if for every $\mathscr{F}_0$-measurable random variable ξ taking values in any compact metric space (E, d) and every $\delta > 0$ there exists some integer $n_0 \leq 0$ and a real time joining $((\mathscr{F}', R), (\mathscr{F}'', R))$ such that $\mathscr{F}'_{n_0} \perp\!\!\!\perp \mathscr{F}''_{n_0}$ and

$$\mathbb{E}[d(\xi', \xi'')] \leq \delta,$$

where ξ' and ξ'' are the respective copies of ξ in $\mathscr{F}'_0$ and $\mathscr{F}''_0$. In the static case, it is known that I-cosiness is equivalent to standardness (see [**?**, Theorem 4.9]). In the dynamical case that is of interest to us here, it was proved in [**?**] that standard dynamical filtrations are I-cosy, but the converse was left as an open question. The purpose of this section is to prove, using a non-standard extension, that the converse is not true: in the dynamical setting, I-cosiness is necessary but not sufficient for a dynamical filtration to be standard.

Proposition 1.5.4. *There exists a non-standard and* I*-cosy dynamical filtration.*

Proof. Let $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$ be a factor map yielding a non-standard extension (we know it exists from Theorem **??**). Take a sequence $(\mathbf{V}_n)_{n\leq -2}$ of non-trivial dynamical systems and set

$$\mathbf{Z} = (Z, \mathscr{C}, \rho, R) := \left(\bigotimes_{n\leq -2} \mathbf{V}_n \right) \otimes \mathbf{X}.$$

We consider the filtration defined by

$$\mathscr{F}_n := \begin{cases} \bigvee_{k\le n} \mathbf{V}_k & \text{if } n \le -2 \\ \bigvee_{k\le -2} \mathbf{V}_k \vee \mathbf{Y} & \text{if } n = -1 \\ \bigvee_{k\le -2} \mathbf{V}_k \vee \mathbf{X} & \text{if } n = 0 \end{cases}$$

Since, $\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ is not standard, it is easy to check that the extension

$$\bigotimes_{n\le -2} \mathbf{V}_n \otimes \mathbf{X} \xrightarrow{\tilde{\pi}} \bigotimes_{n\le -2} \mathbf{V}_n \otimes \mathbf{Y}$$

is not either. Therefore, Proposition **??** yields that $(\mathscr{F}, R)$ is not standard.

We will use and slightly adapt the argument used in [**?**] to show that product type filtrations are I-cosy to show that $\mathscr{F}$ is also I-cosy. Let ξ be a $\mathscr{F}_0$-measurable random variable taking values in a compact metric space (E, d) and $\delta > 0$. There exist $n_0 \le -2$ and a $\left(\bigvee_{n_0+1\le k\le -2} \mathbf{V}_k \vee \mathbf{X}\right)$-measurable $\tilde{\xi}$ such that

$$\mathbb{E}[d(\xi, \tilde{\xi})] \le \delta/2.$$

We now introduce our joining: we set the system

$$\mathbf{W} = (W, \mathscr{D}, \gamma, Q) := \bigotimes_{n\le n_0} \mathbf{V}'_n \otimes \bigotimes_{n\le n_0} \mathbf{V}''_n \otimes \bigotimes_{n_0+1\le n\le -2} \mathbf{V}_n \otimes \mathbf{X},$$

and denote $(\mathscr{F}', \mathscr{F}'')$ the copies of $\mathscr{F}$ on $\mathbf{W}$, namely

$$\mathscr{F}'_n := \begin{cases} \bigvee_{k\le n} \mathbf{V}'_k & \text{if } n \le n_0 \\ \mathscr{F}'_{n_0} \vee \bigvee_{n_0<k\le n} \mathbf{V}_k & \text{if } n_0 < n \le -2 \\ \mathscr{F}'_{-2} \vee \mathbf{Y} & \text{if } n = -1 \\ \mathscr{F}'_{-2} \vee \mathbf{X} & \text{if } n = 0 \end{cases},$$

and a similar definition for $\mathscr{F}''$. Clearly, $\mathscr{F}'_{n_0}$ and $\mathscr{F}''_{n_0}$ are independent and the copies $\tilde{\xi}'$ and $\tilde{\xi}''$ of ξ in the filtrations $\mathscr{F}'$ and $\mathscr{F}''$ coincide, so

$$\mathbb{E}[d(\xi', \xi'')] \le \mathbb{E}[d(\xi', \tilde{\xi}')] + \mathbb{E}[d(\xi'', \tilde{\xi}'')] \le \delta.$$

We now only need to check that $((\mathscr{F}', Q), (\mathscr{F}'', Q))$ is a real time joining, i.e. for every $n \le -1$:

$$\mathscr{F}'_{n+1} \perp\!\!\!\perp_{\mathscr{F}'_n} \mathscr{F}''_n \text{ and } \mathscr{F}''_{n+1} \perp\!\!\!\perp_{\mathscr{F}''_n} \mathscr{F}'_n.$$

For $n \le -2$ we get the relative independent condition like in the product type case. We now check that $\mathscr{F}'_0 \perp\!\!\!\perp_{\mathscr{F}'_{-1}} \mathscr{F}''_{-1}$, which reduces to $\mathbf{X}' \perp\!\!\!\perp_{\mathbf{Y}'} \mathbf{Y}''$. However, this is clearly true because $\mathbf{Y}' = \mathbf{Y}''$. □

We could have simplified this proof by showing that I-cosiness is an asymptotic property, and deducing that $\mathscr{F}$ is I-cosy since $(\mathscr{F}_n)_{n\leq -1}$ is of product type, and therefore I-cosy. Unfortunately, we have not been able show in general that I-cosiness is asymptotic.

Here we have exploited the strong structure of some specific extension to get a non-standard filtration. Therefore it is natural to ask:

Question 1.5.5. *Is there an I-cosy dynamical filtration such that each extension $\mathscr{F}_{n+1} \to \mathscr{F}_n$ is of product type, but which is still not standard ?*

Chapter 2

Confined Poisson extensions

2.1 Introduction

2.1.1 Motivations

This chapter investigates the concept of extensions of measure-preserving dynamical systems, specifically, extensions given by a factor map $\pi : (Z, \rho, R) \to (X, \mu, T)$. We mean that $\mathbf{Z} := (Z, \rho, R)$ and $\mathbf{X} := (X, \mu, T)$ are invertible measure preserving dynamical systems on standard Borel sets such that $\mathbf{X}$ is a factor of $\mathbf{Z}$ via π, and conversely, we also view $\mathbf{Z}$ as an extension of $\mathbf{X}$.

This chapter is a continuation of the work done in Chapter **??** (and published in [**?**]). There, we introduced the notion of confined extensions: they are extensions $(Z, \rho, R) \xrightarrow{\pi} (X, \mu, T)$ such that the only self-joining λ of (Z, ρ, R) in which the law of $\pi \times \pi$ is the product measure $\mu \otimes \mu$, is the product joining $\lambda = \rho \otimes \rho$ (see Definition **??**).

This notion was first of interest to us in the study of dynamical filtrations, which are filtrations defined on some dynamical system (X, μ, T) of the form $\mathscr{F} := (\mathscr{F}_n)_{n \leq 0}$ such that each $\mathscr{F}_n$ is T-invariant (see Section **??** for more details). But we also noticed other interesting results on confined extensions. For example, we listed properties $\mathcal{P}$ that are lifted through confined extensions, i.e. if (X, μ, T) satisfies $\mathcal{P}$ and $(Z, \rho, R) \xrightarrow{\pi} (X, \mu, T)$ is confined, then (Z, ρ, R) satisfies $\mathcal{P}$ (see Section **??**).

Since we noticed that confined extensions had many interesting properties, we look for examples in which that behavior arises. In Chapter **??**, we considered extensions well known in the literature, namely, compact extensions and T, T^{-1}-transformations. In both cases, we gave necessary and sufficient conditions for those extensions to be confined.

In this chapter, we give confinement results for a new kind of extension, in the setting of Poisson suspensions. Take (X, μ, T) a measure preserving dynamical system where μ is a σ-finite measure, but assume that $\mu(X) = \infty$. Consider the probability space (X^*, μ^*) where X^* is the set of counting measures of the form $\sum_{i\geq 1} \delta_{x_i}$, with $(x_i)_{i\geq 1} \in X^{\mathbb{N}}$, and μ^* the law of the Poisson point process of intensity μ. One can then define T_* on (X^*, μ^*) by applying T to each point of the point process. The resulting dynamical system (X^*, μ^*, T_*) is called the *Poisson suspension over* (X, μ, T). A factor map $\pi : (Z, \rho, R) \to (X, \mu, T)$ between infinite measure systems will then yield a factor map between the Poisson suspensions: $\pi_* : (Z^*, \rho^*, R_*) \to (X^*, \mu^*, T_*)$. The resulting extension is what we call a *Poisson extension*.

We will consider the case where $\mathbf{Z} := (X \times G, \mu \otimes m_G, T_\varphi)$ is the compact extension given by a cocycle $\varphi : X \to G$, with G a compact group. We recall that the map T_φ is defined as

$$T_\varphi(x, g) := (Tx, g \cdot \varphi(x)).$$

Our results concern the following Poisson extension:

$$((X \times G)^*, (\mu \otimes m_G)^*, (T_\varphi)_*) \xrightarrow{\pi_*} (X^*, \mu^*, T_*), \tag{2.1}$$

with $\pi : (x, g) \mapsto x$.

In Section **??**, we consider the case where $\varphi(x)$ acts as the identity map, for every $x \in X$. Using a splitting result from [**?**], we prove that in this case, if (X, μ, T) is of infinite ergodic index, the extension (**??**) is confined (see Theorem **??**).

In Section **??**, we deal with a more general cocycle φ. There, our argument will rely on the assumption that the compact extension $(X \times G, \mu \otimes m_G, T_\varphi)$ is of infinite ergodic index. In that case, we make use of Lemma **??**, which is a well know result from Furstenberg. Through some intricate manipulations, we manage to reduce our problem to a relative unique ergodicity problem for products of the extension $\mathbf{Z} \xrightarrow{\pi} \mathbf{X}$, so that we can use Furstenberg's lemma (i.e. Lemma **??**) to prove that (**??**) is confined (see Theorem **??**).

Since the argument developed in Section **??** requires a compact extension $(X \times G, \mu \otimes m_G, T_\varphi)$ of infinite ergodic index, in Section **??**, we give an example of such an extension, showing that Theorem **??** is not void.

2.1.2 Basic notions and notation in ergodic theory

A *dynamical system* is a quadruple $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ such that $(X, \mathscr{A})$ is a standard Borel space, μ is a Borel measure which is σ-finite, i.e. there exist measurable sets $(X_n)_{n\geq 1}$ such that $\mu(X_n) < \infty$ and $X = \bigcup_{n\geq 1} X_n$, and T is an invertible measure-preserving transformation. Throughout the chapter, we will often not specify the σ-algebra $\mathscr{A}$, and will write our dynamical systems as a triple of the form (X, μ, T).

If we have two systems $\mathbf{X} := (X, \mu, T)$ and $\mathbf{Z} := (Z, \rho, R)$, a *factor map* is a measurable map $\pi : Z \longrightarrow X$ such that $\pi_* \rho = \mu$ and $\pi \circ R = T \circ \pi$, ρ-almost surely. If such a map exists, we say that $\mathbf{X}$ is a *factor* of $\mathbf{Z}$. Conversely, we also say that $\mathbf{Z}$ is an extension of $\mathbf{X}$. Moreover, if there exist invariant sets $X_0 \subset X$ and $Z_0 \subset Z$ of full measure such that $\pi : Z_0 \longrightarrow X_0$ is a bijection, then π is an *isomorphism* and we write $\mathbf{Z} \cong \mathbf{X}$.

The system $\mathbf{X} := (X, \mu, T)$ is *ergodic* if $T^{-1}A = A$ implies that $\mu(A) = 0$ or $\mu(A^c) = 0$. It is *conservative* if there is no non-trivial set A such that the $\{T^n A\}_{n\in\mathbb{Z}}$ are disjoint. Let (X, μ, T) be a dynamical system with $\mu(X) = \infty$. If $(X, \mu, T)^{\otimes k}$ is conservative and ergodic, so are all the smaller exponents $(X, \mu, T)^{\otimes j}$, $j \leq k$. The *ergodic index* of (X, μ, T) is the largest integer $k \geq 0$ such that $(X, \mu, T)^{\otimes k}$ is conservative and ergodic. If $(X, \mu, T)^{\otimes k}$ is ergodic for every integer k, the ergodic index is infinite.

Let (X, μ, T) be a conservative system. For $A \subset X$ measurable such that $0 < \mu(A) < \infty$, we denote the restriction of μ to A by $\mu_{|A} := \frac{1}{\mu(A)}\mu(\cdot \cap A)$. Since the system is conservative, the return time $N_A(x) := \inf\{k \geq 1 \,|\, T^k \in A\}$ is almost surely finite, allowing us to define the induced transformation

$$T|_A : \begin{array}{ccc} A & \longrightarrow & A \\ x & \longmapsto & T^{N_A(x)}x \end{array} .$$

Lemma 2.1.1. *Let* $\mathbf{Z} := (Z, \rho, R)$ *and* $\mathbf{X} := (X, \mu, T)$ *be two σ-finite systems and* $\pi : \mathbf{Z} \longrightarrow \mathbf{X}$ *be a factor map. Then* $\mathbf{Z}$ *is conservative if and only if* $\mathbf{X}$ *is conservative.*

Proof. If $\mathbf{Z}$ is conservative, it is easy to see that $\mathbf{X}$ is as well. Conversely, assume that $\mathbf{X}$ is conservative. Let $A \subset Z$ be a measurable set such that $\rho(A) > 0$. We need to show that the sets $\{R^n A\}_{n\in\mathbb{Z}}$ are not disjoint. Because (X, μ) is σ-finite, there are measurable sets $(X_p)_{p\geq 0}$ of finite measure such that $X = \bigcup_{p\geq 0} X_p$. Therefore, we have

$$A = \bigcup_{p\geq 0} (A \cap \pi^{-1} X_p),$$

so there exists $p \geq 0$ such that $\rho(A \cap \pi^{-1} X_p) > 0$. Denote $C := X_p$ and $B := \pi^{-1} B$. Since $\mathbf{X}$ is conservative and $\mu(C) > 0$, the induced system $(C, \mu_{|C}, T|_C)$ is well-defined, and it follows that the induced system $(B, \rho_{|B}, R|_B)$ is well-defined. It is a probability preserving dynamical system, and $\rho_{|B}(A \cap B) > 0$, so, by the Poincaré recurrence theorem, there exists $n \geq 1$ such that $\rho_{|B}(R|_B{}^n(A \cap B) \cap (A \cap B)) > 0$. Moreover, if $N_B^{(n)}$ is the n-th return time on B, we have $R|_B{}^n = R^{N_B^{(n)}}$. Finally, there must be an integer k such that $R|_B{}^n(A \cap B \cap \{N_B^{(n)} = k\}) \cap (A \cap B)$ is not empty, but we have

$$R|_B{}^n(A \cap B \cap \{N_B^{(n)} = k\}) = R^k(A \cap B \cap \{N_B^{(n)} = k\}).$$

Therefore, $R^k(A \cap B) \cap (A \cap B)$ is not empty, which means that $R^k A \cap A$ is not empty. □

2.1.3 Joinings and confined extensions

Let $\mathbf{X} := (X, \mu, T)$ and $\mathbf{Y} := (Y, \nu, S)$ be two σ-finite measure preserving dynamical systems. A *joining* of $\mathbf{X}$ and $\mathbf{Y}$ is a $(T \times S)$-invariant measure λ on $X \times Y$ whose marginals are μ and ν (therefore the marginals have to be σ-finite). It yields the dynamical system

$$\mathbf{X} \times_\lambda \mathbf{Y} := (X \times Y, \lambda, T \times S).$$

On this system, the coordinate projections are factor maps that project onto $\mathbf{X}$ and $\mathbf{Y}$ respectively. If it is not necessary to specify the measure, we will simply write $\mathbf{X} \times \mathbf{Y}$. For the product joining, we will use the notation $\mathbf{X} \otimes \mathbf{Y} := \mathbf{X} \times_{\mu \otimes \nu} \mathbf{Y}$. For the n-fold product self-joining under the transformation $T^{\times n} := T \times \cdots \times T$, we will write $\mathbf{X}^{\otimes n}$.

When $\mathbf{X}$ and $\mathbf{Y}$ are probability measure preserving systems, there is at least one joining, the product joining $\mathbf{X} \otimes \mathbf{Y}$. However, if we have infinite measures, the product measure is not a joining because its marginals are not σ-finite. In fact, there exist pairs of systems for which there does not exist any joining. For example, Lemma **??** implies that there cannot exist a joining of a conservative and a non-conservative system.

We now recall the definition of *confined extensions*, which concerns only probability measure preserving dynamical systems.

Definition 2.1.2. *Let* $\mathbf{X} := (X, \mu, T)$ *and* $\mathbf{Y} := (Y, \nu, S)$ *be probability measure preserving dynamical systems, and* $\pi : \mathbf{X} \longrightarrow \mathbf{Y}$ *be a factor map. The extension*

$\mathbf{X} \xrightarrow{\pi} \mathbf{Y}$ *is said to be* confined *if it satisfies one of the following equivalent properties:*

(i) every 2*-fold self-joining of* $\mathbf{X}$ *in which the two copies of* π *are independent random variables is the product joining: i.e. the only measure* $\lambda \in \mathscr{P}(X \times X)$ *that is* $T \times T$*-invariant, with* $\lambda(\cdot \times X) = \lambda(X \times \cdot) = \mu$ *and* $(\pi \times \pi)_* \lambda = \nu \otimes \nu$*, is* $\lambda = \mu \otimes \mu$*;*

(ii) for every system $\mathbf{Z}$*, every joining of* $\mathbf{X}$ *and* $\mathbf{Z}$ *in which the copy of* π *and the projection on* $\mathbf{Z}$ *are independent random variables is the product joining;*

(iii) for every $n \in \mathbb{N}^* \cup \{+\infty\}$*, every* n*-fold self-joining of* $\mathbf{X}$ *in which the* n *copies of* π *are mutually independent random variables is the* n*-fold product joining.*

It was shown in Proposition **??** that the definitions (i), (ii), and (iii) are equivalent. In this chapter, we mainly use the definition (i). As we mentioned, this concerns only the case for probability measures. An adaptation to the infinite measure case would be more intricate, mainly because if we assume that a measure λ on $X \times X$ projects onto $\nu \otimes \nu$ on $Y \times Y$ and that ν is an infinite measure, then λ cannot be a joining of μ. That is because, in that case, both projections of λ on X would not be σ-finite.

The following lemma is useful when proving that an extension is confined:

Lemma 2.1.3. *Let* $\mathbf{Z} := (Z, \rho, R)$ *and* $\mathbf{X} := (X, \mu, T)$ *be probability measure preserving dynamical systems and* $\pi : \mathbf{Z} \longrightarrow \mathbf{X}$ *be factor map. If* $\mathbf{Z}$ *is ergodic and* $\mathbf{X}$ *is weakly mixing, then we can show that the extension* $\mathbf{Z} \xrightarrow{\pi} \mathbf{X}$ *is confined by verifying that for any* ergodic *self-joining* λ *of* $\mathbf{Z}$ *such that* $(\pi \times \pi)_* \lambda = \mu \otimes \mu$*, we have* $\lambda = \rho \otimes \rho$*.*

Proof. Assume that condition of the lemma is verified. Let λ be (not necessarily ergodic) self-joining of $\mathbf{Z}$ such that $(\pi \times \pi)_* \lambda = \mu \otimes \mu$. Take the ergodic decomposition of λ:

$$\lambda = \int \lambda_\omega \, d\mathbb{P}(\omega).$$

Now take the projection on the first coordinate

$$\rho = \lambda(\cdot \times Z) = \int \lambda_\omega(\cdot \times Z) \, d\mathbb{P}(\omega).$$

However, since ρ is ergodic and the measures $\lambda_\omega(\cdot \times Z)$ are R-invariant, we must have $\lambda_\omega(\cdot \times Z) = \rho$, $\mathbb{P}$-almost surely. Similarly, we get that $\lambda_\omega(Z \times \cdot) = \rho$, $\mathbb{P}$-almost surely. Moreover, we can also apply $\pi \times \pi$ to the ergodic decomposition of λ:

$$\mu \otimes \mu = (\pi \times \pi)_* \lambda = \int (\pi \times \pi)_* \lambda_\omega \, d\mathbb{P}(\omega).$$

Since $\mathbf{X}$ is weakly mixing, the measure $\mu \otimes \mu$ is ergodic. So, as before, we get that $(\pi \times \pi)_* \lambda_\omega = \mu \otimes \mu$, $\mathbb{P}$-almost surely. Therefore, for $\mathbb{P}$-almost every ω, the measure λ_ω verifies all the conditions for us to conclude that $\lambda_\omega = \rho \otimes \rho$. By integrating that, we get that $\lambda = \rho \otimes \rho$.

We have shown that $\xrightarrow{\pi} \mathbf{X}$ is confined. □

2.1.4 Compact extensions and relative unique ergodicity

Definition 2.1.4. *Let $(Z, \rho, R) \xrightarrow{\pi} (X, \mu, T)$ be an extension. It is* relatively uniquely ergodic *if the only R-invariant measure λ on Z such that $\pi_* \lambda = \mu$ is $\lambda = \rho$.*

Let $\mathbf{X} := (X, \mu, T)$ be a measure preserving dynamical system, G a compact group and $\varphi : X \longrightarrow G$ a cocycle. Let m_G denote the Haar probability measure on G. The compact extension of $\mathbf{X}$ given by φ is the system $\mathbf{Z}$ on $(X \times G, \mu \otimes m_G)$ given by the skew product

$$\begin{array}{rccc} T_\varphi : & X \times G & \longrightarrow & X \times G \\ & (x, g) & \longmapsto & (Tx, g \cdot \varphi(x)) \end{array}.$$

This is the most well known family of extensions. The only result we will need, is the following, due to Furstenberg:

Lemma 2.1.5 (Furstenberg [**?**])**.** *Let $\mathbf{X} := (X, \mu, T)$ be an ergodic measure preserving dynamical system where μ is a finite or σ-finite measure. Assume that the compact extension $\mathbf{Z} = (X \times G, \mu \otimes m_G, T_\varphi)$ is ergodic. Let λ be a σ-finite T_φ-invariant measure on $X \times G$ such that $\lambda(\cdot \times G) = \mu$. Then*

$$\lambda = \mu \otimes m_G.$$

This lemma is usually stated with μ a probability measure, but the infinite measure case is proven in the exact same way.

Furstenberg's lemma can be summarized by saying that an ergodic compact extension is relatively uniquely ergodic.

2.1.5 Poisson suspensions, splittings and extensions

Let $\mathbf{X} := (X, \mu, T)$ be a σ-finite measure preserving dynamical system. For convenience, we will assume that $X = \mathbb{R}^+$ and that μ is a locally finite measure, i.e. for any bounded set $B \subset \mathbb{R}^+$, we have $\mu(B) < \infty$. We define the set of counting measures on X by

$$X^* := \left\{ \text{locally finite measures of the form } \sum_{i\geq 1} \delta_{x_i} \right\}.$$

A point process is a probability measure on X^*. The Poisson point process of intensity μ, which we denote μ^*, is the point process characterized by the fact that, for $A_1, ..., A_n \subset X$ measurable disjoint subsets such that $0 < \mu(A_i) < \infty$, the random variables $\omega(A_1), ..., \omega(A_n)$, for $\omega \in X^*$, are independent Poisson random variables of respective parameter $\mu(A_i)$, for $i \in [\![1, n]\!]$.

On the probability space (X^*, μ^*), we define the transformation

$$T_* : \sum_{i\geq 1} \delta_{x_i} \mapsto T_* \left(\sum_{i\geq 1} \delta_{x_i} \right) = \sum_{i\geq 1} \delta_{Tx_i}.$$

The resulting dynamical system $\mathbf{X}^* := (X^*, \mu^*, T_*)$ is called the *Poisson suspension over* (X, μ, T).

It is well known that the Poisson suspension $\mathbf{X}^*$ is ergodic if and only if there is no T-invariant measurable subset $A \subset X$ such that $0 < \mu(A) < \infty$ (see [**?**]). Moreover, this implies that if $\mathbf{X}^*$ is ergodic, it is automatically weakly mixing. Also, note that it is not necessary that $\mathbf{X}$ is ergodic for $\mathbf{X}^*$ to be ergodic.

We use the notion of Poisson splittings from [**?**], but with different choices in the notation. A *splitting of order* n of the Poisson suspension (X^*, μ^*, T_*) is a family $\{\nu_i\}_{1\leq i\leq n}$ of T_*-invariant probability measures on X^* and λ a $T_*^{\times n}$-invariant joining of $\{\nu_i\}_{1\leq i\leq n}$ such that $\Sigma_*^{(n)}\lambda = \mu^*$, where

$$\begin{array}{cccc} \Sigma^{(n)}: & X^* \times \cdots \times X^* & \longrightarrow & X^* \\ & (\omega_1, ..., \omega_n) & \longmapsto & \omega_1 + \cdots + \omega_n \end{array}.$$

The splitting is said to be *ergodic* if λ is an ergodic joining. The splitting is a *Poisson splitting* if there exist $\{\mu_i\}_{1\leq i\leq n}$, σ-finite measures on X such that, for $i \in [\![1, n]\!]$, $\nu_i = \mu_i^*$, and λ is the product measure $\mu_1^* \otimes \cdots \otimes \mu_n^*$. With that notation, the result [**?**, Theorem 2.6] becomes

Theorem 2.1.6. *Let $\mathbf{X} := (X, \mu, T)$ be a σ-finite measure preserving dynamical system of infinite ergodic index. Then any ergodic splitting of the Poisson suspension (X^*, μ^*, T_*) is a Poisson splitting.*

Consider two σ-finite systems $\mathbf{Z} := (Z, \rho, R)$ and $\mathbf{X} := (X, \mu, T)$ and a factor map $\pi : \mathbf{Z} \longrightarrow \mathbf{X}$, which means that we have an extension $\mathbf{Z} \xrightarrow{\pi} \mathbf{X}$ of σ-finite systems. We can then define the map

$$\pi_* : \sum_{i \geq 1} \delta_{x_i} \mapsto \pi_* \left(\sum_{i \geq 1} \delta_{x_i} \right) = \sum_{i \geq 1} \delta_{\pi(x_i)}.$$

One can check that this yields a factor map from $\mathbf{Z}^*$ to $\mathbf{X}^*$, therefore we have defined an extension $\mathbf{Z}^* \xrightarrow{\pi_*} \mathbf{X}^*$ between Poisson suspensions. Such an extension is what we call a *Poisson extension*.

As we did in Corollary **??** for T, T^{-1} transformations, we make sure that the extensions we consider here are not compact, to make sure that we are looking at a new type of extension and getting original confinement results.

Lemma 2.1.7. *Let $\mathbf{Z} := (Z, \rho, R)$ and $\mathbf{X} := (X, \mu, T)$ be infinite σ-finite systems and $\pi : \mathbf{Z} \longrightarrow \mathbf{X}$ be a factor map. If $\mathbf{X}$ is conservative and ergodic, then $\mathbf{Z}^*$ is ergodic. Moreover, the extension $\mathbf{Z}^* \longrightarrow \mathbf{X}^*$ is relatively weakly mixing, and therefore it is not compact.*

The notion of relative weak mixing introduced in Definition **??** uses the relatively independent product, presented in Section **??**. Although the definition we gave was set in the context of probability measures, the construction of the relatively independent product also works in the exact same way for infinite measure systems, and we use it in the following proof.

Proof. As we mentioned earlier, a criterion for the ergodicity of $\mathbf{Z}^*$ was given in [**?**]. We need to show that the invariant sets of $\mathbf{Z}$ are of measure 0 or ∞. Let A be an R-invariant set such that $\rho(A) < \infty$.

Now take a set $B \subset X$ of finite measure. Using that A is R-invariant, we get

$$\begin{aligned} \rho(A \cap \pi^{-1}B) = \int \mathbb{1}_A (\mathbb{1}_B \circ \pi) d\rho &= \int \mathbb{1}_A (\mathbb{1}_B \circ T^j \circ \pi) d\rho \\ &= \frac{1}{n} \sum_{j=0}^{n-1} \int \mathbb{1}_A (\mathbb{1}_B \circ T^j \circ \pi) d\rho \\ &= \int \mathbb{1}_A \left(\frac{1}{n} \sum_{j=0}^{n-1} \mathbb{1}_B \circ T^j \circ \pi \right) d\rho. \end{aligned}$$

However, since $\mathbf{X}$ is conservative, ergodic and $\mu(X) = \infty$, we can use [?, Exercise 2.2.1] to see that

$$\frac{1}{n}\sum_{j=0}^{n-1} \mathbb{1}_B \circ T^j \underset{n\to\infty}{\longrightarrow} 0 \ \mu\text{-almost surely.}$$

Next, because $\rho(A) < \infty$, the dominated convergence theorem yields

$$\int \mathbb{1}_A \left(\frac{1}{n}\sum_{j=0}^{n-1} \mathbb{1}_B \circ T^j \circ \pi \right) d\rho \underset{n\to\infty}{\longrightarrow} 0,$$

meaning that $\rho(A \cap \pi^{-1}B) = 0$. Finally, using that μ is σ-finite, it follows that $\rho(A) = 0$.

To prove that the extension $\mathbf{Z}^* \longrightarrow \mathbf{X}^*$ is relatively weakly mixing, we need to look at the relative product $\mathbf{Z}^* \otimes_{\mathbf{X}^*} \mathbf{Z}^*$, and prove that it is ergodic. One can prove that $\mathbf{Z}^* \otimes_{\mathbf{X}^*} \mathbf{Z}^*$ is isomorphic to $(\mathbf{Z} \otimes_{\mathbf{X}} \mathbf{Z})^*$, the Poisson suspension on the relative product $\mathbf{Z} \otimes_{\mathbf{X}} \mathbf{Z}$, via the map

$$\psi : \sum_{k\geq 0} \delta_{\left(z_k^{(1)}, z_k^{(2)}\right)} \mapsto \left(\sum_{k\geq 0} \delta_{z_k^{(1)}}, \sum_{k\geq 0} \delta_{z_k^{(2)}} \right).$$

However, $(\mathbf{Z} \otimes_{\mathbf{X}} \mathbf{Z})^*$ is a Poisson extension of $\mathbf{X}^*$. Therefore, using the first part of the current lemma, we deduce that $\mathbf{Z}^* \otimes_{\mathbf{X}^*} \mathbf{Z}^*$ is ergodic, making $\mathbf{Z}^* \longrightarrow \mathbf{X}^*$ relatively weakly mixing. Then, as in the proof of Corollary **??**, we use [?, Chapter 6, Part 4] to conclude that $\mathbf{Z}^* \longrightarrow \mathbf{X}^*$ is not compact. □

2.2 A Poisson extension over a trivial cocycle

In this section, we study Poisson extensions over extensions of the form

$$\begin{array}{rccc} T \times \mathrm{Id} : & X \times K & \longrightarrow & X \times K \\ & (x, \kappa) & \longmapsto & (Tx, \kappa) \end{array},$$

on $(X \times K, \mu \otimes \rho)$, where K is a standard Borel space and ρ is a probability measure on K. We start in Section **??** by showing that if T has infinite ergodic index, the associated Poisson extension is confined. Then in Section **??**, we see that marked point processes enable us to write Poisson extensions through a Rokhlin cocycle, and we give an application in probability theory by giving an alternative proof of the De Finetti theorem (see Corollary **??**). Finally, in Section **??**, we give an example of a non-confined Poisson extension.

2.2.1 Confinement as a consequence of Poisson splittings

We derive the content of this section as a consequence of Theorem **??**. In [**?**], the authors proved Theorem **??** and gave an application of that splitting result (specifically, [**?**, Theorem 3.1]). Here, we note that it can be rephrased as a relative unique ergodicity result for the Poisson extension. In our notation, it becomes:

Theorem 2.2.1. *Let* $\mathbf{X} := (X, \mu, T)$ *be a* σ*-finite measure preserving dynamical system of infinite ergodic index, and* K *a standard Borel space. Let* λ *be an invariant marked point process over* μ^**, i.e. a* $(T \times \mathrm{Id})_*$*-invariant probability measure on* $(X \times K)^*$ *such that* $(\pi_*)_*\lambda = \mu^*$*. If* $(\lambda, (T \times \mathrm{Id})_*)$ *is ergodic, then there exists a probability measure* ρ *on* K *such that* $\lambda = (\mu \otimes \rho)^*$*.*

We deduce that the Poisson extension is confined:

Theorem 2.2.2. *Let* $\mathbf{X} := (X, \mu, T)$ *be a* σ*-finite measure preserving dynamical system of infinite ergodic index, and* (K, ρ) *a standard probability space. Then the Poisson extension*

$$((X \times K)^*, (\mu \otimes \rho)^*, (T \times \mathrm{Id})_*) \longrightarrow (X^*, \mu^*, T_*),$$

is confined.

Proof. Set $\mathbf{Z} := ((X \times K)^*, (\mu \otimes \rho)^*, (T \times \mathrm{Id})_*)$ and $\pi : (x, \kappa) \mapsto x$. Let λ be a 2-fold self joining of $\mathbf{Z}$ such that $(\pi_* \times \pi_*)_*\lambda = \mu^* \otimes \mu^*$. Since $\mathbf{Z}^*$ and $\mathbf{X}^*$ are ergodic, and even weakly mixing (see Section **??** or [**?**]), by Lemma **??**, one may assume that λ is ergodic. Note that λ is a probability measure on

$$(X \times K)^* \times (X \times K)^*.$$

Both marginals of λ on $(X \times K)^*$ yield a Poisson point process of intensity $\mu \otimes \rho$. We view the realization of both of those processes simultaneously on $X \times K$ and we tag the points coming from the first coordinate with a 1, and the points coming from the second coordinate with a 2. To do that formally, we define the map

$$\begin{array}{cccc} \Omega : & (X \times K)^* \times (X \times K)^* & \longrightarrow & (X \times K \times \{1,2\})^* \\ & (\omega_1, \omega_2) & \longmapsto & \omega_1 \otimes \delta_{\{1\}} + \omega_2 \otimes \delta_{\{2\}} \end{array},$$

so that

$$\Omega(\omega_1, \omega_2)(\cdot \times \{i\}) = \omega_i. \tag{2.2}$$

Consider $\eta := \Omega_* \lambda$. Since $\Omega \circ (T \times \mathrm{Id}_K)_* \times (T \times \mathrm{Id}_K)_* = (T \times \mathrm{Id}_{K \times \{1,2\}})_* \circ \Omega$, we know that η is $(T \times \mathrm{Id}_{K \times \{1,2\}})_*$-invariant. Moreover, because λ is assumed to be ergodic, the system $((X \times K \times \{1,2\})^*, \eta, (T \times \mathrm{Id}_{K \times \{1,2\}})_*)$ is ergodic. Finally, we need to look at the projection on X^* via $\tilde{\pi} : (x, \kappa, i) \mapsto x$. To that end, we note that $(\tilde{\pi}_*)_* \eta = (\tilde{\pi}_* \circ \Omega)_* \lambda$, and

$$\begin{aligned}\tilde{\pi}_* \circ \Omega(\omega_1, \omega_2) &= \tilde{\pi}_*(\omega_1 \otimes \delta_{\{1\}} + \omega_2 \otimes \delta_{\{2\}}) \\ &= \tilde{\pi}_*(\omega_1 \otimes \delta_{\{1\}}) + \tilde{\pi}_*(\omega_2 \otimes \delta_{\{2\}}) = \pi_* \omega_1 + \pi_* \omega_2.\end{aligned}$$

However, $(\pi_* \times \pi_*)_* \lambda = \mu^* \otimes \mu^*$, which means that, in the above notation, $\pi_* \omega_1$ and $\pi_* \omega_2$ are independent Poisson processes of intensity μ. It is known that the sum of such two independent Poisson processes is a Poisson process of intensity 2μ. Therefore, $(\tilde{\pi}_*)_* \eta = (\tilde{\pi}_* \circ \Omega)_* \lambda = (2\mu)^*$. Theorem **??** tells us that there exists $\chi \in \mathscr{P}(K \times \{1,2\})$ such that $\eta = (2\mu \otimes \chi)^*$.

Now we show that $\chi = \rho \otimes (\frac{1}{2}(\delta_{\{1\}} + \delta_{\{2\}}))$. Let $A \subset X$ such that $0 < \mu(A) < \infty$, $B \subset K$ and $i \in \{1,2\}$

$$\begin{aligned}e^{-2\mu(A)\chi(B \times \{i\})} &= \eta(\{\tilde{\omega}\,;\, \tilde{\omega}(A \times B \times \{i\}) = 0\}) \\ &= \lambda(\{(\omega_1, \omega_2)\,;\, \omega_i(A \times B) = 0\}) \text{ because of (??)} \\ &= (\mu \otimes \rho)^*(\{\omega\,;\, \omega(A \times B) = 0\}) = e^{-\mu(A)\rho(B)}.\end{aligned}$$

So, $\chi(B \times \{i\}) = \frac{1}{2}\rho(B)$. Therefore $\chi = \rho \otimes (\frac{1}{2}(\delta_{\{1\}} + \delta_{\{2\}}))$, so $\eta = (2\mu \otimes \rho \otimes (\frac{1}{2}(\delta_{\{1\}} + \delta_{\{2\}})))^*$. Finally, we get

$$\begin{aligned}\lambda = \Omega_*^{-1} \eta &= \Omega_*^{-1}(2\mu \otimes \rho \otimes (\frac{1}{2}(\delta_{\{1\}} + \delta_{\{2\}})))^* \\ &= (\mu \otimes \rho)^* \times (\mu \otimes \rho)^*.\end{aligned}$$

□

2.2.2 Marked Point processes

Let (X, μ) be a standard Borel space equipped with a σ-finite measure such that $\mu(X) = \infty$. Without loss of generality, we may assume that $X = \mathbb{R}_+$, thus enabling us to use the natural order on $\mathbb{R}^+$, but any other order could be used here. We way also assume that μ is the Lebesgue measure on $\mathbb{R}_+$ (by doing so, we ignore the case where μ has atoms, but for the rest of our work, that is not a problem). Up to a set of μ^*-measure 0 we may assume that the elements ω of $(\mathbb{R}_+)^*$ are locally finite measures with no multiplicity, i.e. such that for all

$x \in \mathbb{R}_+$, $\omega(\{x\}) \leq 1$. This allows us to define a sequence $(t_n)_{n\in\mathbb{N}}$ of measurable maps from $(\mathbb{R}_+)^*$ to $\mathbb{R}_+$ such that

$$\omega = \sum_{n\geq 1} \delta_{t_n(\omega)},$$

and

$$0 \leq t_1(\omega) < t_2(\omega) < \cdots .$$

Each $t_n(\omega)$ gives us the position of the n-th atom of the counting measure ω.

Now consider a Polish space K. We will call a *marked point process over* μ^* a probability measure λ on $(\mathbb{R}_+\times K)^*$ such that $(\pi_*)_*\lambda = \mu^*$, where $\pi : (x,\kappa) \mapsto x$. We already manipulated marked point processes in the previous section, we are simply giving them a name now. We can describe marked point processes as follows: define the map

$$f: \begin{array}{ccc} (\mathbb{R}_+)^* \times K^{\mathbb{N}} & \longrightarrow & (\mathbb{R}_+ \times K)^* \\ (\omega, (\kappa_n)_{n\geq 1}) & \longmapsto & \sum_{n\geq 1} \delta_{(t_n(\omega),\kappa_n)} \end{array}.$$

Since f is injective, we know from [**?**, Corollary 15.2] that $f((\mathbb{R}_+)^* \times K^{\mathbb{N}})$ is a Borel set and f^{-1} is measurable, and we can write it as

$$\Phi := f^{-1} : \tilde{\omega} \mapsto (\pi_*(\tilde{\omega}), (\kappa_n(\tilde{\omega}))_{n\geq 1}), \tag{2.3}$$

where $(\kappa_n(\tilde{\omega}))_{n\geq 1}$ is called the sequence of the *marks* of $\tilde{\omega}$. For a marked point process λ, $\lambda\big(f((\mathbb{R}_+)^* \times K^{\mathbb{N}})\big) = 1$, therefore, up to a set of measure 0, f is a bijection. Moreover, we have the following result, from [**?**, Lemma 6.4.VI]:

Proposition 2.2.3. *Let ρ be a probability measure on K. The Poisson process $(\mu\otimes\rho)^*$ is a marked point process over μ^* and $f_*(\mu^*\otimes\rho^{\otimes\mathbb{N}}) = (\mu\otimes\rho)^*$.*

In other words, if $\tilde{\omega} \in (\mathbb{R}_+ \times K)^*$ is distributed according to the Poisson process of intensity $\mu\otimes\rho$, the sequence of marks $(\kappa_n(\tilde{\omega}))_{n\geq 1}$ is i.i.d. of law ρ.

2.2.3 A $\mathfrak{S}(\mathbb{N})$-valued cocycle and its action on $K^{\mathbb{N}}$

In Section **??**, we saw that, assuming that $X = \mathbb{R}_+$, a point process on $(X\times K)^*$ can be represented on $X^* \times K^{\mathbb{N}}$, via the map introduced in (**??**):

$$\Phi: \begin{array}{ccc} (X\times K)^* & \longrightarrow & X^* \times K^{\mathbb{N}} \\ \tilde{\omega} & \longmapsto & (\pi_*(\tilde{\omega}), (\kappa_n(\tilde{\omega}))_{n\geq 1}) \end{array},$$

where $\kappa_n(\tilde{\omega})$ is the mark associated to the n-th point of $\tilde{\omega}$, once the points of $\tilde{\omega}$ are ordered according to their projection on X. Now we mean to determine the dynamic on $X^* \times K^{\mathbb{N}}$ that would correspond to $(T \times \mathrm{Id})^*$ on $(X \times K)^*$. To do this, we will need a tool to track how the points of $\omega \in X^*$ are reordered when T_* is applied.

The group $\mathfrak{S}(\mathbb{N})$ is the group of all permutations on $\mathbb{N}$, i.e. the bijections from $\mathbb{N}$ onto itself. Equipped with the metric

$$d(\sigma, \tau) := \sum_{n \in \mathbb{N}} \frac{1}{2^n} \mathbb{1}_{\sigma(n) \neq \tau(n)},$$

it is a Polish group, i.e. $(\mathfrak{S}(\mathbb{N}), d)$ is a complete separable metric space such that the map $(\sigma, \tau) \mapsto \sigma \circ \tau^{-1}$ is continuous. This group acts on $K^{\mathbb{N}}$ via the measurable action

$$(\sigma, (\kappa_n)_{n \geq 1}) \mapsto (\kappa_{\sigma^{-1}(n)})_{n \geq 1}. \tag{2.4}$$

We recall that, given $\omega \in X^*$, we denote by $(t_n(\omega))_{n \geq 1}$ the ordered sequence of the points of ω. To describe the action of T_* on $(t_n(\omega))_{n \geq 1}$, we define $\Psi(\omega)$ as the unique element of $\mathfrak{S}(\mathbb{N})$ such that for every $n \in \mathbb{N}$, $\Psi(\omega)(n)$ is the rank of the atom $T(t_n(\omega))$ in the counting measure $T_*\omega$. This define a cocycle

$$\Psi : X^* \longrightarrow \mathfrak{S}(\mathbb{N}),$$

so that

$$T(t_n(\omega)) = t_{\Psi(\omega)(n)}(T_*\omega).$$

We consider the skew-product define by the cocycle Ψ:

$$\begin{array}{cccc} (T_*)_\Psi : & X^* \times K^{\mathbb{N}} & \longrightarrow & X^* \times K^{\mathbb{N}} \\ & (\omega, (\kappa_n)_{n \geq 1}) & \longmapsto & (T_*\omega, (\kappa_{\Psi(\omega)^{-1}(n)})_{n \geq 1}) \end{array}.$$

Then we check that

$$\begin{aligned} \Phi \circ (T \times \mathrm{Id})_*(\widetilde{\omega}) &= (\pi_*(T \times \mathrm{Id})_*\widetilde{\omega}, (\kappa_n((T \times \mathrm{Id})_*\widetilde{\omega}))_{n \geq 1}) \\ &= (T_*\pi_*\widetilde{\omega}, (\kappa_{\Psi(\omega)^{-1}(n)}(\widetilde{\omega}))_{n \geq 1}) \\ &= (T_*)_\Psi(\pi_*\widetilde{\omega}, (\kappa_n(\widetilde{\omega}))_{n \geq 1}) \\ &= (T_*)_\Psi \circ \Phi(\widetilde{\omega}). \end{aligned} \tag{2.5}$$

Combined with Proposition **??**, this tells us that Φ is an isomorphism from

$$((X \times K)^*, (\mu \otimes \rho)^*, (T \times \mathrm{Id})_*) \longrightarrow (X^*, \mu^*, T_*),$$

to

$$(X^* \times K^{\mathbb{N}}, \mu^* \otimes \rho^{\otimes \mathbb{N}}, (T_*)_\Psi) \longrightarrow (X^*, \mu^*, T_*).$$

Through this isomorphism, Theorem **??** becomes

Theorem 2.2.4. *Let $\mathbf{X} := (X, \mu, T)$ be a σ-finite measure preserving dynamical system of infinite ergodic index, and K a standard Borel space. Let λ be a $(T_*)_\Psi$-invariant probability measure on $X^* \times K^{\mathbb{N}}$ such that $\lambda(\cdot \times K^{\mathbb{N}}) = \mu^*$. If $(\lambda, (T_*)_\Psi)$ is ergodic, then there exists a probability measure ρ on K such that $\lambda = \mu^* \otimes \rho^{\otimes \mathbb{N}}$.*

As an unexpected corollary, we get the following result, which is the De Finetti theorem written in the language of ergodic theory:

Corollary 2.2.4.1 (De Finetti, Hewitt-Savage)**.** *Let ρ_∞ be a $\mathfrak{S}(\mathbb{N})$-invariant (under the action defined by* (**??**)*) probability measure on $K^{\mathbb{N}}$ and ρ its marginal on the first coordinate. The action $(K^{\mathbb{N}}, \rho_\infty, \mathfrak{S}(\mathbb{N}))$ is ergodic if and only if $\rho_\infty = \rho^{\otimes \mathbb{N}}$.*

Proof. If $\rho_\infty = \rho^{\otimes \mathbb{N}}$, it is ergodic under the action of $\mathfrak{S}(\mathbb{N})$ for the same reason that it is ergodic under the shift. For completeness, we detail that argument. Let $A \subset K^{\mathbb{N}}$ be a measurable set such that, for every $\sigma \in \mathfrak{S}(\mathbb{N})$, $A = \sigma^{-1} A$ mod ρ_∞, where the action of σ is given by (**??**). Take $\delta > 0$: there exist $\ell \geq 1$ and a set B measurable with respect to $(\kappa_n)_{0 \leq n \leq \ell}$ such that $\rho_\infty(A \Delta B) \leq \delta$. Then, for any permutation $\sigma \in \mathfrak{S}(\mathbb{N})$ that sends the interval $[\![\ell+1, 2\ell]\!]$ onto $[\![1, \ell]\!]$, the set $\sigma^{-1} B$ is $(\kappa_n)_{\ell+1 \leq n \leq 2\ell}$-measurable and therefore independent of B. Finally, we get

$$\begin{aligned}\rho_\infty(A) &= \rho_\infty(\sigma^{-1} A \cap A) \leq \rho_\infty(\sigma^{-1} B \cap B) + 2\delta \\ &= \rho_\infty(\sigma^{-1} B) \rho_\infty(B) + 2\delta = \rho_\infty(B)^2 + 2\delta \leq \rho_\infty(A)^2 + 4\delta.\end{aligned}$$

And letting δ go to 0 yields $\rho_\infty(A) = \rho_\infty(A)^2$, meaning that $\rho_\infty(A) = 0$ or 1.

Assume that $(K^{\mathbb{N}}, \rho_\infty, \mathfrak{S}(\mathbb{N}))$ is ergodic. Let (X, μ, T) be a dynamical system of infinite ergodic index. Since ρ_∞ is $\mathfrak{S}$-invariant, it follows that $\mu^* \otimes \rho_\infty$ is $(T_*)_\Psi$-invariant. Then Theorem **??** tells us that the ergodic decomposition of $\mu^* \otimes \rho_\infty$ is of the form

$$\mu^* \otimes \rho_\infty = \int_\Gamma \mu^* \otimes \gamma^{\otimes \mathbb{N}} d\mathbb{P}(\gamma) = \mu^* \otimes \int_\Gamma \gamma^{\otimes \mathbb{N}} d\mathbb{P}(\gamma),$$

so $\rho_\infty = \int_\Gamma \gamma^{\otimes \mathbb{N}} d\mathbb{P}(\gamma)$. However, each measure $\gamma^{\otimes \mathbb{N}}$ is $\mathfrak{S}(\mathbb{N})$-invariant, and ρ_∞ is ergodic under $\mathfrak{S}(\mathbb{N})$. Therefore, there exists $\gamma \in \Gamma$ such that $\rho_\infty = \gamma^{\otimes \mathbb{N}}$. □

2.2.4 A non-confined Poisson extension

We give here an example of a non-trivial non-confined Poisson extension, to show that the infinite ergodic index assumption in Theorem **??** cannot be removed. Take $X := \mathbb{R}$, μ the Lebesgue measure on $\mathbb{R}$, and

$$T : x \mapsto x + 1.$$

The system (X, μ, T) is neither ergodic nor conservative and its ergodic index is 0, but the Poisson suspension (X^*, μ^*, T_*) is ergodic, since it is Bernoulli. We get the following:

Proposition 2.2.5. *Let (K, ρ) be a non-trivial standard probability space. The extension $((X \times K)^*, (\mu \otimes \rho)^*, (T \times \mathrm{Id})_*) \to (X^*, \mu^*, T_*)$ is not confined.*

Proof. We will make use of the setup presented in the previous section for the study of marked point processes. We make some slight adjustments since now $X = \mathbb{R}$ (instead of $\mathbb{R}_+$): define a sequence $(t_n(\omega))_{n\in\mathbb{Z}}$ such that

$$\omega = \sum_{n\in\mathbb{Z}} \delta_{t_n(\omega)},$$

and

$$\cdots t_{-1}(\omega) < t_0(\omega) < 0 \leq t_1(\omega) < t_2(\omega) \cdots .$$

We then also define a cocycle

$$\widetilde{\Psi} : X^* \longrightarrow \mathfrak{S}(\mathbb{Z}),$$

such that

$$T(t_n(\omega)) = t_{\widetilde{\Psi}(\omega)(n)}(T_*\omega).$$

In other words, $\widetilde{\Psi}(\omega)(n)$ is the rank of the atom $T(t_n(\omega))$ in $T_*\omega$. In our present case, the map $\widetilde{\Psi}$ can be described explicitly: denote the shift $S : k \mapsto k + 1$ and then one can check that

$$\widetilde{\Psi} : \omega \mapsto S^{\omega([-1,0[)}.$$

Indeed, the shift in the numbering of the atoms of ω is only affected by the atoms in $[-1, 0[$ as they go from being smaller than 0 to being greater than 0 once we apply T.

As in the previous section, we get an isomorphism in between the extensions

$$((X \times K)^*, (\mu \otimes \rho)^*, (T \times \mathrm{Id})_*) \xrightarrow{\pi_*} (X^*, \mu^*, T_*),$$

and

$$(X^* \times K^{\mathbb{Z}}, \mu^* \otimes \rho^{\otimes \mathbb{Z}}, (T_*)_{\widetilde{\Psi}}) \xrightarrow{\widetilde{\pi}} (X^*, \mu^*, T_*).$$

We prove our proposition by showing that the second extension is not confined. We need to build a non-product self-joining of $\mu^* \otimes \rho^{\otimes \mathbb{Z}}$ whose projection via $\widetilde{\pi} \times \widetilde{\pi}$ is $\mu^* \otimes \mu^*$. It will be more convenient to describe this joining as a measure on $X^* \times X^* \times K^{\mathbb{Z}} \times K^{\mathbb{Z}}$. We start with the marginal on $X^* \times X^*$ which has to be $\lambda(\cdot \times \cdot \times K^{\mathbb{Z}} \times K^{\mathbb{Z}}) = \mu^* \otimes \mu^*$, and for $(\omega_1, \omega_2) \in X^* \times X^*$, the conditional law $\lambda_{(\omega_1,\omega_2)}$ is as follows. The sequence of marks of ω_1, $(\kappa_n(\omega_1))_{n\in\mathbb{Z}}$ is chosen with probability $\rho^{\otimes \mathbb{Z}}$. For the choice of $\kappa_n(\omega_2)$, we distinguish two situations:

- If $\omega_1([t_n(\omega_2), t_{n+1}(\omega_2)[) \geq 1$, we set

  $$\ell := \min\{k \in \mathbb{Z} \,|\, t_k(\omega_1) \in [t_n(\omega_2), t_{n+1}(\omega_2)[\},$$

  and then we choose $\kappa_n(\omega_2) := \kappa_\ell(\omega_1)$.

- If $\omega_1([t_n(\omega_2), t_{n+1}(\omega_2)[) = 0$, we choose $\kappa_n(\omega_2)$ with law ρ, independently from all the other marks.

The construction of λ is concluded by taking

$$\lambda := \int \delta_{\omega_1} \otimes \delta_{\omega_2} \otimes \lambda_{(\omega_1,\omega_2)} \, d(\mu^* \otimes \mu^*)(\omega_1, \omega_2).$$

Note that our choices for $(\kappa_n(\omega_1))_{n\in\mathbb{Z}}$ and $(\kappa_n(\omega_2))_{n\in\mathbb{Z}}$ depend only on the relative positions of the points $\{t_n(\omega_1)\}_{n\in\mathbb{Z}}$ and $\{t_n(\omega_2)\}_{n\in\mathbb{Z}}$. Since those relative positions are preserved under application of $(T_* \times T_*)$, the measure λ is $(T_*)_{\widetilde{\Psi}} \times (T_*)_{\widetilde{\Psi}}$-invariant (up to a permutation of coordinates).

From our construction, it is clear that λ is not a product measure and that

$$\lambda(\cdot \times X^* \times \cdot \times K^{\mathbb{Z}}) = \mu^* \otimes \rho^{\otimes \mathbb{Z}}.$$

We are left with checking that $\lambda(X^* \times \cdot \times K^{\mathbb{Z}} \times \cdot) = \mu^* \otimes \rho^{\otimes \mathbb{Z}}$. Consider that ω_1, ω_2 and $(\kappa_n(\omega_2))_{n<n_0}$ are known and compute the law of $\kappa_{n_0}(\omega_2)$: if $\omega_1([t_n(\omega_2), t_{n+1}(\omega_2)[) = 0$, it is follows from our construction that the law of $\kappa_{n_0}(\omega_2)$ is ρ. If $\omega_1([t_n(\omega_2), t_{n+1}(\omega_2)[) \geq 1$, we have $\kappa_n(\omega_2) = \kappa_\ell(\omega_1)$ (see above for the definition of ℓ). One can check that $(\kappa_n(\omega_2))_{n<n_0}$ only informs us on (some of) the values of $(\kappa_n(\omega_1))_{n<\ell}$ and $\kappa_\ell(\omega_1)$ is independent of ω_1, ω_2 and $(\kappa_n(\omega_1))_{n<\ell}$. So, even with ω_1, ω_2 and $(\kappa_n(\omega_2))_{n<n_0}$ fixed, the law of $\kappa_\ell(\omega_1)$ is ρ, so the law of $\kappa_n(\omega_2)$ is also ρ.

To sum up, up to a permutation of coordinates, λ is a $(T_*)_{\widetilde{\Psi}} \times (T_*)_{\widetilde{\Psi}}$-invariant measure on $X^* \times K^{\mathbb{Z}} \times X^* \times K^{\mathbb{Z}}$, such that

$$\lambda(\cdot \times \cdot \times X^* \times K^{\mathbb{Z}}) = \mu^* \otimes \rho^{\otimes \mathbb{Z}} \text{ and } \lambda(X^* \times K^{\mathbb{Z}} \times \cdot \times \cdot) = \mu^* \otimes \rho^{\otimes \mathbb{Z}}.$$

So it a self-joining of $\mu^* \otimes \rho^{\otimes \mathbb{Z}}$. Moreover, it projects onto $\mu^* \otimes \mu^*$ without being equal to the product measure $\mu^* \otimes \rho^{\otimes \mathbb{Z}} \otimes \mu^* \otimes \rho^{\otimes \mathbb{Z}}$. This precisely means that the extension

$$(X^* \times K^{\mathbb{Z}}, \mu^* \otimes \rho^{\otimes \mathbb{Z}}, (T_*)_{\widetilde{\Psi}}) \longrightarrow (X^*, \mu^*, T_*),$$

is not confined. □

2.3 A Poisson suspension over a compact extension

In this section, we are interested in the Poisson extension over a compact extension, i.e. over the system $\mathbf{Z}$ given by

$$\begin{array}{rccc} T_\varphi : & X \times G & \longrightarrow & X \times G \\ & (x, g) & \longmapsto & (Tx, g \cdot \varphi(x)) \end{array},$$

for some compact group G and measurable cocycle $\varphi : X \to G$. Our goal will be to show that

Theorem 2.3.1. *Let $\mathbf{X} := (X, \mu, T)$ be a dynamical system of infinite ergodic index. If the compact extension $(X \times G, \mu \otimes m_G, T_\varphi)$ is also of infinite ergodic index, then the Poisson extension*

$$((X \times G)^*, (\mu \otimes m_G)^*, (T_\varphi)_*) \xrightarrow{\pi_*} (X^*, \mu^*, T_*)$$

is confined.

We start with the Sections **??** and **??**, where we introduce some useful notions and results from the literature. We then prove the main technical step in the proof of our theorem in Section **??**. We conclude in Section **??**.

2.3.1 Ergodicity of Cartesian products in spectral theory

We present briefly some results on the ergodicity of Cartesian products of σ-finite measure preserving dynamical systems.

We start by introducing some classic notions in spectral theory. Let (X, μ, T) be a σ-finite measure preserving dynamical system. Consider the space $L^2(X, \mu)$ and the Koopman operator

$$U_T : \begin{array}{ccc} L^2(X,\mu) & \longrightarrow & L^2(X,\mu) \\ f & \longmapsto & f \circ T \end{array}.$$

Denote $L^2_0(X, \mu) \subset L^2(X, \mu)$ the subspace orthogonal to the space of constant maps. For $f \in L^2(X, \mu)$, the spectral measure of f, σ_f, is define as the only measure on $\mathbb{U}$ such that, for every $n \in \mathbb{Z}$

$$\widehat{\sigma_f}(n) = \int_X f\, \overline{U_T^n f} d\mu.$$

There exists a finite measure σ^0_X on $\mathbb{U}$, unique up to equivalence, such that

- for every $f \in L^2_0(X, \mu)$, $\sigma_f \ll \sigma^0_X$,
- and for every finite measure σ such that $\sigma \ll \sigma^0_X$, there exists $f \in L^2_0(X, \mu)$ such that $\sigma = \sigma_f$.

It is the *restricted maximal spectral type* of (X, μ, T).

We define a L^∞-eigenvalue as $\lambda \in \mathbb{C}$ such that there exists $f \in L^\infty(X, \mu)$ such that

$$f \circ T = \lambda f.$$

Such a map f is called a L^∞-eigenfunction. Denote $e(T)$ the set of all L^∞-eigenvalues of (X, μ, T), which is a sub-group of $\mathbb{U}$, provided T is conservative and ergodic (see [**?**, Section 2.6]). The notion of L^∞-eigenvalues is mainly useful in the infinite measure case. Indeed, if $\mu(X) < \infty$, we have $L^\infty(X, \mu) \subset L^2(X, \mu)$, so L^∞-eigenfunctions are simply eigenvectors of the Koopman operator U_T. When T is ergodic, we also know that an eigenvector f of U_T is a L^∞-eigenfunction because $|f|$ is almost surely constant.

We will use the following ergodicity criterion, due to Keane (see [**?**, Section 2.7]):

Theorem 2.3.2. *Let $\mathbf{X} := (X, \mu, T)$ be a conservative and ergodic dynamical system and $\mathbf{Y} := (Y, \nu, S)$ be an ergodic probability measure preserving dynamical system. The Cartesian product $\mathbf{X} \otimes \mathbf{Y}$ is ergodic if and only if $\sigma^0_Y(e(T)) = 0$.*

We use that criterion to prove

Corollary 2.3.2.1. *Let* $\mathbf{X} := (X, \mu, T)$ *be a conservative and ergodic dynamical system and* $k \geq 1$. *If* $\mathbf{X}^{\otimes 2k}$ *is ergodic, the product system*

$$(X, \mu, T)^{\otimes k} \otimes (X^*, \mu^*, T_*)$$

is ergodic.

Proof. Let $k \geq 1$. Assume that $\mathbf{X}^{\otimes 2k}$ is ergodic. We show that $e(T^{\times k}) = \{1\}$. Take $\lambda \in e(T^{\times k})\backslash\{1\}$ and $f \in L^\infty(X^k, \mu^{\otimes k})$ the associated eigenfunction. Define the tensor function as

$$\begin{array}{rccc} f \otimes \overline{f}: & X^k \times X^k & \longrightarrow & \mathbb{C} \\ & (x_1, x_2) & \longmapsto & f(x_1)\overline{f(x_2)} \end{array}.$$

We have

$$(f \otimes \overline{f}) \circ T^{\times 2k} = (f \circ T^{\times k}) \otimes (\overline{f \circ T^{\times k}}) = \lambda\overline{\lambda} f \otimes \overline{f} = f \otimes \overline{f}.$$

Since $\mathbf{X}^{\otimes 2k}$ is ergodic, it yields that $f \otimes \overline{f}$ is constant, so f is constant, which implies that $\lambda = 1$. Therefore $e(T^{\times k}) = \{1\}$.

Moreover, since (X, μ, T) is ergodic, (X^*, μ^*, T_*) is as well, so $\sigma^0_{X^*}(\{1\}) = 0$. We have shown that $\sigma^0_{X^*}(e(T^{\times k})) = 0$, so Theorem **??** tells us that

$$(X, \mu, T)^{\otimes k} \otimes (X^*, \mu^*, T_*)$$

is ergodic. □

Remark 2.3.3. The result from Corollary **??** should be compared to the following result from Meyerovitch (see [**?**, Theorem 1.2]):

$$(X, \mu, T) \otimes (X^*, \mu^*, T_*) \text{ is ergodic if and only if } (X, \mu, T) \text{ is ergodic.}$$

This gives the result of Corollary **??** in the case $k = 1$, but with a weaker condition: we need (X, μ, T) to be ergodic, instead of $(X, \mu, T)^{\otimes 2}$. We conjecture that one could extend the result from Meyerovitch and get the conclusion of Corollary **??** under the weaker assumption that $(X, \mu, T)^{\otimes k}$ is ergodic.

2.3.2 Distinguishing points in a Poisson process

In this section, we use the sequence $(t_n(\omega))_{n\geq 1}$ introduced in Section **??**. Because of that, we need X to be $\mathbb{R}^+$ and the measure μ to be the Lebesgue measure. Some additional aspects of the structure of $(\mathbb{R}^+, \mu)$ will also be useful. The purpose of this section is to study the map

$$\begin{array}{cccc} \tilde{\Phi}_k : & X^k \times X^* & \longrightarrow & X^* \\ & (x_1, ..., x_k, \omega) & \longmapsto & \delta_{x_1} + \cdots + \delta_{x_k} + \omega \end{array}.$$

We view the points of $X^k \times X^*$ as a Poisson process for which the first k points are distinguished, so that we can track each of them individually. To avoid any multiplicity on the right-hand term, we will study this map on a smaller set $X^{(k)} \subset X^k \times X^*$, defined as

$$X^{(k)} := \{(x_1, ..., x_k, \omega) \in X^k \times X^* \,|\, x_1 < \cdots < x_k < t_1(\omega)\}.$$

From now on, Φ_k denotes the restriction of $\tilde{\Phi}_k$ to $X^{(k)}$. We start by computing the measure of $X^{(k)}$, using the fact t_1 follows an exponential law of parameter 1:

$$\begin{aligned} \mu^{\otimes k} \times \mu^*(X^{(k)}) &= \int_{X^*} \int_{(\mathbb{R}_+)^k} \mathbb{1}_{x_1 < \cdots < x_k < t_1(\omega)} d\mu(x_1) \cdots d\mu(x_k) d\mu^*(\omega) \\ &= \int_{\mathbb{R}_+} \int_{(\mathbb{R}_+)^k} \mathbb{1}_{x_1 < \cdots < x_k < t} d\mu(x_1) \cdots d\mu(x_k) e^{-t} dt \\ &= \int_{\mathbb{R}_+} \frac{t^k}{k!} e^{-t} dt = 1, \end{aligned}$$

the last equality being obtained through k successive integrations by parts. We complete this with the following result

Lemma 2.3.4. *Let $k \geq 1$. The map Φ_k sends $(\mu^{\otimes k} \otimes \mu^*)\big|_{X^{(k)}}$ onto μ^*. Therefore*

$$\Phi_k : (X^{(k)}, (\mu^{\otimes k} \otimes \mu^*)\big|_{X^{(k)}}) \longrightarrow (X^*, \mu^*),$$

is an isomorphism of probability spaces.

Proof. It is clear that Φ_k is a bijection whose inverse is

$$\omega \mapsto (t_1(\omega), ..., t_k(\omega), \omega - (\delta_{t_1(\omega)} + \cdots + \delta_{t_k(\omega)})).$$

We then need to prove that Φ_k is measure-preserving. We prove that result by induction on k. The case $k = 1$ can be found in [**?**, Proposition 6.1], but we give a proof for completeness. Denote Exp the law of an exponential variable of parameter 1.

To prove that $(\Phi_1)_*(\mu \otimes \mu^*)\big|_{X^{(1)}} = \mu^*$, we need to show that, if (x, ω) is chosen under $(\mu \otimes \mu^*)\big|_{X^{(1)}}$, the sequence

$$\big(x, t_1(\omega) - x, (t_{i+1}(\omega) - t_i(\omega))_{i\geq 1}\big)$$

is i.i.d. of law Exp. First, we know that $(t_{i+1} - t_i)_{i\geq 1}$ is i.i.d. of law Exp. It is also clear that $(x, t_1 - x)$ is independent from $(t_{i+1} - t_i)_{i\geq 1}$. Therefore, we now only have to compute the law of $(x, t_1 - x)$ under $(\mu \otimes \mu^*)\big|_{X^{(1)}}$. Let A, $B \subset \mathbb{R}_+$ be measurable sets. We have:

$$\begin{aligned}
\mu \otimes \mu^*(x < t_1, x \in A, t_1 - x \in B) &= \int_A \mu^*(x < t_1, t_1 - x \in B) d\mu(x) \\
&= \int_A \underbrace{\mu^*(x < t_1)}_{=e^{-x}} \underbrace{\mu^*(t_1 - x \in B \,|\, t_1 > x)}_{=\mu^*(t_1 \in B)} d\mu(x) \\
&= \int_A e^{-x} d\mu(x) \mu^*(t_1 \in B) = \mathrm{Exp}(A) \cdot \mathrm{Exp}(B),
\end{aligned}$$

where we use the fact that the law of t_1 is Exp, and the loss of memory property of Exp.

Let $k \geq 1$, and assume that the result is true for k. We start by noting that $\Phi_{k+1} = \Phi_1 \circ (\mathrm{Id} \times \Phi_k)$ and use the induction hypothesis to prove that

$$(\mathrm{Id} \times \Phi_k)_*(\mu^{\otimes k+1} \otimes \mu^*)\big|_{X^{(k+1)}} = (\mu \otimes \mu^*)\big|_{X^{(1)}}. \tag{2.6}$$

Indeed, for a measurable map, $F : X^{(1)} \to \mathbb{R}$, we have

$$\begin{aligned}
&\int_{X^{(k+1)}} F(x_1, \delta_{x_2} + \cdots + \delta_{k+1} + \omega) \, d\mu^{\otimes k+1}(x_1, ..., x_{k+1}) d\mu^*(\omega) \\
&\quad = \int_{\mathbb{R}_+} \int_{X^{(k)}} \mathbb{1}_{x_1 < x_2} F(x_1, \sum_{i=2}^{k+1} \delta_{x_i} + \omega) \, d\mu^{\otimes k}(x_2, ..., x_{k+1}) d\mu^*(\omega) d\mu(x_1) \\
&\quad = \int_{\mathbb{R}_+} \int_{X^*} \mathbb{1}_{x_1 < t_1(\omega)} F(x_1, \omega) d\mu^*(\omega) d\mu(x_1) \\
&\quad = \int_{X^{(1)}} F d(\mu \otimes \mu^*),
\end{aligned}$$

by the induction hypothesis and the fact that $x_2 = t_1(\delta_{x_2} + \cdots \delta_{k+1} + \omega)$. Therefore (**??**) is proven. We then combine it with the result for $k = 1$ to conclude that

$$(\Phi_{k+1})_*(\mu^{\otimes k+1} \otimes \mu^*)\big|_{X^{(k+1)}} = (\Phi_1)_*(\mu \otimes \mu^*)\big|_{X^{(1)}} = \mu^*.$$

□

We now want to study how Φ_k matches the dynamics on $X^{(k)}$ and X^*. We recall that we defined $\Psi : (\mathbb{R}_+)^* \to \mathfrak{S}(\mathbb{N})$ so that $T(t_n(\omega)) = t_{\Psi(\omega)(n)}(T_*\omega)$. We then iterate it to define

$$\Psi_p(\omega) := \Psi(T_*^{p-1}\omega) \circ \cdots \circ \Psi(\omega).$$

This iteration means that $\Psi_p(\omega)(n)$ is the rank of the atom $T^p(t_n(\omega))$ in the counting measure $T_*^p\omega$. Now consider

$$N^{(k)}(\omega) := \inf\{p \geq 1 \mid \Psi_p(\omega)(1) = 1, ..., \Psi_p(\omega)(k) = k\}.$$

This is the first time in which the first k points of ω are back to being the first k points of $T_*^p\omega$ and in their original order. If the random time $N^{(k)}$ is almost surely finite, we can define the automorphism $T^{N^{(k)}}$ on (X^*, μ^*) by

$$\left(T_*^{N^{(k)}}\right)(\omega) := T_*^{N^{(k)}(\omega)}(\omega).$$

We conclude this section with the following result:

Proposition 2.3.5. *Let* $\mathbf{X} := (X, \mu, T)$ *be a* σ*-finite measure preserving dynamical system. Assume that* T *has infinite ergodic index. Then, for any* $k \geq 1$*,* $N^{(k)}$ *is* μ^**-almost surely well-defined and* Φ_k *is an isomorphism between the systems*

$$(X^{(k)}, (\mu^{\otimes k} \otimes \mu^*)\big|_{X^{(k)}}, (T^{\times k} \times T_*)\big|_{X^{(k)}})$$

and

$$(X^*, \mu^*, T_*^{N^{(k)}}).$$

Proof. Let $k \geq 1$. Since $\mathbf{X} = (X, \mu, T)$ is of infinite ergodic index, the system $\mathbf{X}^{\otimes k} = (X^k, \mu^{\otimes k}, T^{\times k})$ is conservative and ergodic. Since μ^* is a probability measure, Lemma **??** tells us that the system $(X^k \times X^*, \mu^{\otimes k} \otimes \mu^*, T^{\times k} \times T_*)$ is also conservative. Therefore, the induced system

$$(X^{(k)}, (\mu^{\otimes k} \otimes \mu^*)\big|_{X^{(k)}}, (T^{\times k} \times T_*)\big|_{X^{(k)}})$$

is well-defined. Moreover, if $M^{(k)}$ is the first return time in $X^{(k)}$, then $M^{(k)}$ is $(\mu^{\otimes k} \otimes \mu^*)\big|_{X^{(k)}}$-almost surely finite. However, since we have

$$\tilde{\Phi}_k \circ (T^{\times k} \times T_*) = T_* \circ \tilde{\Phi}_k, \tag{2.7}$$

one can check that on $X^{(k)}$, we have

$$M^{(k)} = N^{(k)} \circ \Phi_k. \tag{2.8}$$

So, because Lemma **??** shows that $(\Phi_k)_*(\mu^{\otimes k} \otimes \mu^*)\big|_{X^{(k)}} = \mu^*$, we deduce that $N^{(k)}$ is μ^*-almost surely finite. Finally, by combining (**??**) and (**??**), one gets

$$\Phi_k \circ (T^{\times k} \times T_*)\big|_{X^{(k)}} = T_*^{N^{(k)}} \circ \Phi_k.$$

Since, by Lemma **??**, Φ_k is a bijection for which $(\Phi_k)_*(\mu^{\otimes k} \otimes \mu^*)\big|_{X^{(k)}} = \mu^*$, we have shown that it yields the desired isomorphism of dynamical systems. □

2.3.3 Relative unique ergodicity

As before, we assume that $X = \mathbb{R}^+$ and μ is the Lebesgue measure. The main step in proving Theorem **??** is the following relative unique ergodicity result:

Theorem 2.3.6. *Let $\mathbf{X} := (X, \mu, T)$ be a dynamical system of infinite ergodic index. If the compact extension $(X \times G, \mu \otimes m_G, T_\varphi)$ is also of infinite ergodic index, then the only $(T_\varphi)_*$-invariant measure $\rho \in \mathscr{P}((X \times G)^*)$ such that $(\pi_*)_*\rho = \mu^*$ is $\rho = (\mu \otimes m_G)^*$.*

As in Section **??**, we represent the Poisson extension $((X \times G)^*, (\mu \otimes m_G)^*, (T_\varphi)_*) \xrightarrow{\pi_*} (X^*, \mu^*, T_*)$ through a Rokhlin cocycle. We do this using the representation of $((X \times G)^*, (\mu \otimes m_G)^*)$ as a marked point process given in Proposition **??**.

We start by introducing the skew product group $G^{\mathbb{N}} \rtimes \mathfrak{S}(\mathbb{N})$ whose operation is defined by

$$((h_n)_{n\geq 1}, \tau) \cdot ((g_n)_{n\geq 1}, \sigma) = ((h_{\sigma(n)} \cdot g_n)_{n\geq 1}, \tau \circ \sigma).$$

This group acts on $(G^{\mathbb{N}}, m_G{}^{\otimes \mathbb{N}})$ via the maps

$$\chi_{(g_n)_{n\geq 1}, \sigma}\left((h_n)_{n\geq 1}\right) := \left(g_{\sigma^{-1}(n)} \cdot h_{\sigma^{-1}(n)}\right)_{n\geq 1}.$$

Then we define the cocycle from X^* to $G^{\mathbb{N}} \rtimes \mathfrak{S}(\mathbb{N})$ by

$$\overline{\varphi} : \omega \mapsto (\varphi(t_n(\omega)))_{n\geq 1}, \Psi(\omega)).$$

This cocycle induces the following transformation

$$\begin{array}{rccc} (T_*)_{\overline{\varphi}} : & X^* \times G^{\mathbb{N}} & \longrightarrow & X^* \times G^{\mathbb{N}} \\ & (\omega, (g_n)_{n\geq 1}) & \longmapsto & (T_*\omega, \chi_{\overline{\varphi}(\omega)}((g_n)_{n\geq 1})) \end{array}.$$

Then, for any $(T_\varphi)_*$-invariant measure ρ such that $(\pi_*)_*\rho = \mu^*$, an adaptation of the computation (**??**) show that the map Φ introduced in (**??**) gives an isomorphism between the extensions

$$((X \times G)^*, \rho, (T_\varphi)_*) \xrightarrow{\pi_*} (X^*, \mu^*, T_*),$$

and

$$(X^* \times G^{\mathbb{N}}, \Phi_*\rho, (T_*)_{\overline{\varphi}}) \longrightarrow (X^*, \mu^*, T_*).$$

Therefore, to prove Theorem **??**, we need to take a $(T_*)_{\overline{\varphi}}$-invariant measure λ such that $\lambda(\cdot \times G^{\mathbb{N}}) = \mu^*$ and show that $\lambda = \mu^* \otimes {m_G}^{\otimes \mathbb{N}}$. This is what we do below:

Proof of Theorem **??**. Let λ be a $(T_*)_{\overline{\varphi}}$-invariant measure on $X^* \times G^{\mathbb{N}}$ such that $\lambda(\cdot \times G^{\mathbb{N}}) = \mu^*$. Fix $k \geq 1$ and call λ_k the image of λ via p_k, the projection on $X^* \times G^k$. The main idea of this proof is to use Proposition **??** to distinguish the points $t_1(\omega), ..., t_k(\omega)$ since they determine the action of $(T_*)_{\overline{\varphi}}$ on $g_1, ..., g_k$ and then view $(t_1(\omega), g_1), ..., (t_k(\omega), g_k)$ as a compact extension of $t_1(\omega), ..., t_k(\omega)$ to which Furstenberg's relative unique ergodicity Lemma applies.

We start our argument by understanding better the dynamics on $g_1, ..., g_k$. Since (X, μ, T) has infinite ergodic index, Proposition **??** tells us that the random time

$$\tilde{N}^{(k)}(\omega, (g_n)_{n\geq 1}) := N^{(k)}(\omega)$$

is λ-almost surely finite. Now note that, by definition of $N^{(k)}$, we have

$$\begin{aligned} p_k \circ (T_*)_{\overline{\varphi}}^{\tilde{N}^{(k)}}&(\omega,(g_n)_{n\geq 1}) \\ &= (T_*^{N^{(k)}(\omega)}\omega, \varphi^{(N^{(k)}(\omega))}(t_1(\omega)) \cdot g_1, ..., \varphi^{(N^{(k)}(\omega))}(t_k(\omega)) \cdot g_k) \\ &= (T_*^{N^{(k)}(\omega)}\omega, \varphi_k(\omega) \cdot (g_1, ..., g_k)), \end{aligned}$$

where we define the cocycle $\varphi_k : X^* \to G^k$ by:

$$\varphi_k(\omega) := \varphi^{(N^{(k)}(\omega))}(t_1(\omega)), ..., \varphi^{(N^{(k)}(\omega))}(t_k(\omega)),$$

with

$$\varphi^{(p)}(x) := \varphi(T^{p-1}x) \cdots \varphi(x).$$

Therefore λ_k is invariant under the transformation

$$\begin{array}{cccc} (T_*^{N^{(k)}})_{\varphi_k} : & X^* \times G^k & \longrightarrow & X^* \times G^k \\ & (\omega, (g_1, ..., g_k)) & \longmapsto & ((T_*)^{N^{(k)}(\omega)}\omega, \varphi_k(\omega) \cdot (g_1, ..., g_k)) \end{array}.$$

This map yields a compact extension of $(T_*)^{N^{(k)}}$, but to apply Furstenberg's Lemma (i.e. Lemma **??**), we still have to prove that

$$(X^* \times G^k, \mu^* \otimes m_G^{\otimes k}, (T_*^{N^{(k)}})_{\varphi_k}) \tag{2.9}$$

is ergodic.

We recall that $M^{(k)}$ is defined as the return time on $X^{(k)}$ and that $M^{(k)} = N^{(k)} \circ \Phi_k$. Then, Proposition **??** tells us that $(X^* \times G^k, \mu^* \otimes m_G^{\otimes k}, (T_*^{N^{(k)}})_{\varphi_k})$ is isomorphic to

$$\left(X^{(k)} \times G^k, \left(\mu^{\otimes k} \otimes \mu^*\right)\big|_{X^{(k)}} \otimes m_G^{\otimes k}, \left((T^{\times k} \times T_*)\big|_{X^{(k)}}\right)_{\widehat{\varphi_k}}\right), \tag{2.10}$$

where $\widehat{\varphi_k}$ is the cocycle defined by

$$\widehat{\varphi_k} := \varphi_k \circ \Phi_k = (\varphi^{(M^{(k)})}(x_1), ..., \varphi^{(M^{(k)})}(x_k)).$$

However, it is straightforward to check that, up to a permutation of coordinates, (**??**) is an induced system of

$$(X^k \times G^k \times X^*, \mu^{\otimes k} \otimes m_G^{\otimes k} \otimes \mu^*, T_\varphi^{\times k} \times T_*),$$

which can be written as

$$(X \times G, \mu \otimes m_G, T_\varphi)^{\otimes k} \otimes (X^*, \mu^*, T_*). \tag{2.11}$$

However, this is a factor of

$$(X \times G, \mu \otimes m_G, T_\varphi)^{\otimes k} \otimes ((X \times G)^*, (\mu \otimes m_G)^*, (T_\varphi)_*). \tag{2.12}$$

But, since $(X \times G, \mu \otimes m_G, T_\varphi)$ is of infinite ergodic index, Corollary **??** applies and tells us that (**??**) is ergodic, and therefore (**??**) is as well. Since an induced system on an ergodic system is also ergodic, this yields that (**??**) is ergodic. In conclusion, Furstenberg's Lemma implies that $\lambda_k = \mu^* \otimes m_G^{\otimes k}$.

This being true for every $k \geq 1$, it follows that $\lambda = \mu^* \otimes m_G^{\otimes \mathbb{N}}$. □

2.3.4 Conclusion of the proof of Theorem ??

We now finish the proof of Theorem **??** by combining our relative unique ergodicity result (Theorem **??**) from the previous section with Theorem **??**. In our application of the splitting result (Theorem **??**), the fact that the marginals $\{\nu_i\}_{i \in [\![1,n]\!]}$ are Poisson measures will already be known, and the important part will be the fact that the associated joining λ is the product joining.

Proof of Theorem **??**. Let $\mathbf{X} := (X, \mu, T)$ be a dynamical system, and $\varphi : X \to G$ a cocycle such that the compact extension $\mathbf{Z} := (X \times G, \mu \otimes m_G, T_\varphi)$ has infinite ergodic index.

Let λ be a $(T_\varphi)_* \times (T_\varphi)_*$-invariant self-joining of $(\mu \otimes m_G)^*$ such that

$$(\pi_* \times \pi_*)_* \lambda = \mu^* \otimes \mu^*. \tag{2.13}$$

Since, as mentioned in Section **??**, (X^*, μ^*, T_*) is weakly mixing, Lemma **??** implies that, up to replacing λ with one of its ergodic components, we may assume that the system

$$((X \times G)^* \times (X \times G)^*, \lambda, (T_\varphi)_* \times (T_\varphi)_*),$$

is ergodic. Now set $\rho := \Sigma_* \lambda$. We then use (**??**) to compute

$$(\pi_*)_* \rho = (\pi_*)_* \Sigma_* \lambda = \Sigma_* (\pi_* \times \pi_*)_* \lambda = \Sigma_* (\mu^* \otimes \mu^*) = (2\mu)^*.$$

In other words, (**??**) means that the projection of ρ on X^* is the sum of to independent Poisson point processes of intensity μ, and the result of this sum is a Poisson point process of intensity 2μ. Now, since T_φ has infinite ergodic index, we can apply Theorem **??** to conclude that $\rho = (2\mu \otimes m_G)^*$.

Using again the fact that T_φ has infinite ergodic index, we can now deduce from Theorem **??** that λ is the product joining

$$\lambda = (\mu \otimes m_G)^* \times (\mu \otimes m_G)^*.$$

□

2.4 A compact extension of infinite ergodic index

The construction in Theorem **??** relies on a compact extension

$$(X \times G, \mu \otimes m_G, T_\varphi),$$

which is of infinite ergodic index. In this section, we build a compact extension that has that property, which shows that Theorem **??** is not void. We start with Section **??**, where we give a criterion for the ergodicity of compact extensions. Then, in Section **??**, we choose a suitable system (X, μ, T): the infinite Chacon transformation (see [**?**, Section 2]). Finally, in Section **??**, we describe our choice for the cocycle φ and prove that the resulting transformation T_φ is of infinite ergodic index.

2.4.1 Ergodic compact extensions

Let $\mathbf{X} := (X, \mu, T)$ be a measure preserving dynamical system, G a compact group and $\varphi : X \longrightarrow G$ a cocycle. We mean to study ergodic properties of the compact extension:

$$\begin{array}{rccc} T_\varphi : & X \times G & \longrightarrow & X \times G \\ & (x, g) & \longmapsto & (Tx, g \cdot \varphi(x)) \end{array}.$$

In this section, we prove the following lemma, for which a statement can be found in [**?**, Theorem 3] for the finite measure case and in [**?**, Lemma 1] for the infinite measure case. We give a proof inspired from [**?**] that works for both cases. This lemma relies on the notion of characters, for which we use the definition given by Rudin in [**?**]: the group of the characters of G, $\hat{G}$, is the set of all continuous maps $\chi : G \to \mathbb{U}$ such that, for all $g, h \in G$, $\chi(g \cdot h) = \chi(g)\chi(h)$.

Lemma 2.4.1. *Let $\mathbf{X} := (X, \mu, T)$ be an ergodic measure preserving dynamical system where μ is a finite or σ-finite infinite measure. Assume that G is an abelian group. The compact extension given by $\mathbf{Z} = (X \times G, \mu \otimes m_G, T_\varphi)$ is ergodic if and only if, for every character $\chi \in \hat{G} \backslash \{1\}$, there is no measurable function $f : X \longrightarrow \mathbb{U}$ such that*

$$\frac{f(Tx)}{f(x)} = \chi(\varphi(x)) \text{ almost everywhere.} \tag{2.14}$$

If there is a measurable map f such that (**??**) holds, we say that $\chi \circ \varphi$ is a co-boundary. Therefore, **Z** is ergodic if there is no character $\chi \in \hat{G} \backslash \{1\}$ for which $\chi \circ \varphi$ is a co-boundary. In our proof, we will use results on Fourier analysis on locally compact abelian groups from [**?**]. This is why we need to assume that G is abelian.

Proof. Assume that there is a character $\chi \neq 1$ and a map $f : X \longrightarrow \mathbb{U}$ that satisfies (**??**). Then define

$$h(x, g) := f(x) \cdot \chi(g)^{-1}.$$

One can simply check that

$$\begin{aligned} h \circ T_\varphi(x, g) &= h(Tx, g \cdot \varphi(x)) \\ &= f(Tx) \cdot \chi(g \cdot \varphi(x))^{-1} \\ &= f(x) \cdot \chi(\varphi(x)) \cdot \chi(\varphi(x))^{-1} \cdot \chi(g)^{-1} \\ &= f(x) \cdot \chi(g)^{-1} = h(x, g). \end{aligned}$$

Since $\chi \neq 1$, h is not constant, and therefore $\mathbf{Z}$ is not ergodic.

Conversely, if $\mathbf{Z}$ is not ergodic, there exists $h \in L^\infty(\mathbf{Z})$ not almost everywhere constant such that $h \circ T_\varphi = h$. For $\chi \in \hat{G}$, define

$$f_\chi(x) := \int_G h(x,g)\chi(g)dm_G(g).$$

We know that this integral is well-defined for almost every x because $h \in L^\infty(\mathbf{Z})$. Then we have, almost surely:

$$\begin{aligned} f_\chi(Tx) &= \int_G h(Tx,g)\chi(g)dm_G(g) \\ &= \int_G h(Tx, g\cdot\varphi(x))\chi(g\cdot\varphi(x))dm_G(g) \\ &= \int_G h(x,g)\chi(g)\chi(\varphi(x))dm_G(g) = \chi(\varphi(x))f_\chi(x). \end{aligned}$$

Now we simply need to find $\chi \neq 1$ such that f_χ is not almost everywhere equal to 0. First, notice that f_1 is T invariant, and therefore almost everywhere constant. Take $c \in \mathbb{C}$ such that $f_1 \equiv c$. We now argue by contradiction.

Assume that for every $\chi \neq 1$, $f_\chi(x) = 0$ for almost every x. Since G is compact, [**?**, Theorem 1.2.5] tells us that $\hat{G}$ is discrete, and therefore countable. From that, we deduce that for almost every x, we have

$$\forall \chi \neq 1, f_\chi(x) = 0.$$

Therefore, using the fact that, for $\chi \neq 1$, $\int_G \chi dm_G = 0$ and $f_1 \equiv c$, we get that, for almost every x:

$$\forall \chi \in \hat{G}, \int_G (h(x,g) - c)\chi(g)dm_G(g) = 0.$$

Then, using a result from [**?**, Section 1.7.3], we know that for those x, the map $g \mapsto h(x,g) - c$ is almost surely 0. Therefore, $h \equiv c$ almost everywhere, which contradicts our assumption on h.

This means that there exists $\chi \in \hat{G}\setminus\{1\}$ such that $\mu(\{f_\chi \neq 0\}) > 0$. Since the set $\{f_\chi \neq 0\}$ is T invariant, this yields that $\mu(\{f_\chi = 0\}) = 0$, by ergodicity of $\mathbf{X}$. Finally, define the function $f := f_\chi/|f_\chi|$, which takes its values in $\mathbb{U}$ and satisfies (**??**). □

In what follows we apply this lemma to a situation where the group G is $(\mathbb{Z}/2\mathbb{Z})^p$. So we need to know the characters of $(\mathbb{Z}/2\mathbb{Z})^p$:

Lemma 2.4.2. *The characters of* $\mathbb{Z}/2\mathbb{Z}^p$ *are of the form:*

$$\chi_I := \bigotimes_{i \notin I} \chi_1 \otimes \bigotimes_{i \in I} \chi_{-1} = \bigotimes_{i \in I} (-1)^{\bullet} : (\epsilon_1, ..., \epsilon_p) \mapsto (-1)^{\sum_{i \in I} \epsilon_i},$$

for $I \subset [\![1, p]\!]$.

Proof. We prove it by induction on p. For $p = 1$, take $\chi \in \widehat{\mathbb{Z}/2\mathbb{Z}}$ and note that, since characters are group morphisms, we have $\chi(0) = 1$. Next to determine $\chi(1)$, we note that $\chi(1)^2 = \chi(1+1) = \chi(0) = 1$. Therefore, $\chi(1) = 1$ or -1. We are left with two possible characters: $\chi \equiv 1$ or $\chi : \epsilon \mapsto (-1)^{\epsilon}$.

For $p \geq 1$, assume that the characters of $\mathbb{Z}/2\mathbb{Z}^p$ are as described in the lemma and take $\chi \in \widehat{\mathbb{Z}/2\mathbb{Z}^{p+1}}$. We have

$$\chi(\epsilon_1, ..., \epsilon_{p+1}) = \chi(\epsilon_1, ..., \epsilon_p, 0)\chi(0, ..., 0, \epsilon_{p+1}).$$

Finally, since $\epsilon_1, ..., \epsilon_p \mapsto \chi(\epsilon_1, ..., \epsilon_p, 0)$ is a character of $\mathbb{Z}/2\mathbb{Z}^p$ and $\epsilon_{p+1} \mapsto \chi(0, ..., 0, \epsilon_{p+1})$ is a character of $\mathbb{Z}/2\mathbb{Z}$, the induction hypothesis and the case for $p = 1$ end our proof. □

2.4.2 Description of the infinite Chacon transformation

Let $\mathbf{X} := (X, \mu, T)$ be the system given by the infinite Chacon transformation defined in [**?**, Section 2]. We chose this transformation because it is known that it has an infinite ergodic index (see [**?**, Theorem 2.2]), and because the rank one structure is convenient to define a suitable cocycle in Section **??**. Some other infinite measure preserving rank one transformations could be used here. For example the nearly finite Chacon transformation introduced in [**?**] has all the properties we require in this work. All the following constructions and proofs could be applied to that transformation.

As any rank one transformation, the infinite Chacon transformation can be defined as an increasing union of towers $(\mathcal{T}_n)_{n \geq 1}$. The tower $\mathcal{T}_n$ of order n, is composed of its levels $\{E_n^{(1)}, ..., E_n^{(h_n)}\}$ such that

$$\mathcal{T}_n = \bigsqcup_{k=1}^{h_n} E_n^{(k)}.$$

We say that h_n is the height of $\mathcal{T}_n$. The transformation T acts on $\mathcal{T}_n$ so that, for $k \in [\![1, h_n - 1]\!]$, we have

$$TE_n^{(k)} = E_n^{(k+1)}.$$

All levels of $\mathcal{T}_n$ have same measure under μ, and we denote it by $\mu_n := \mu(E_n^{(k)})$.

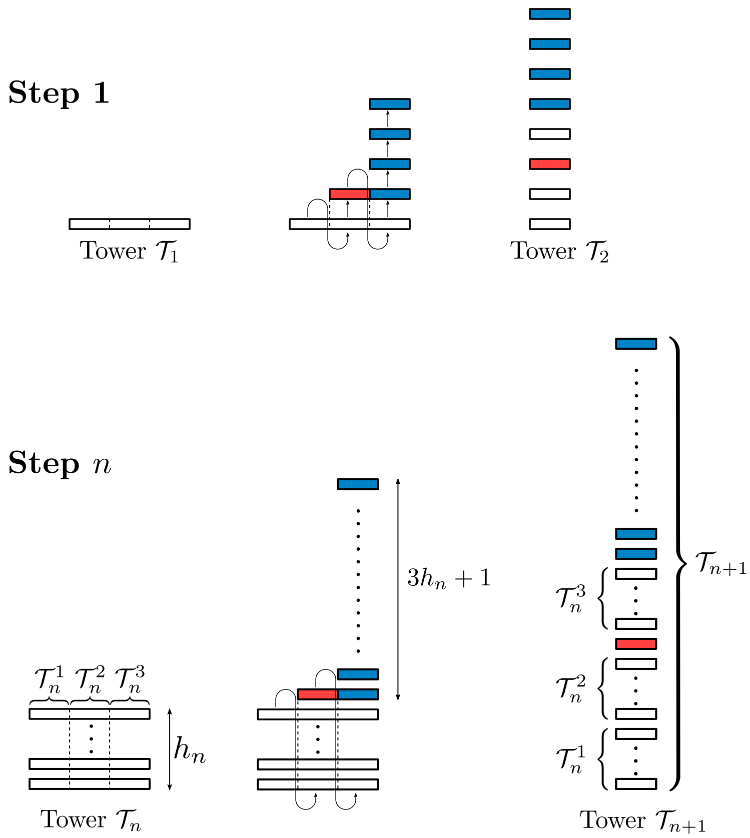


Figure 2.1: Construction of the infinite Chacon transformation

The construction of the sequence $(\mathcal{T}_n)_{n\geq 1}$ is done inductively. It will be done by taking intervals of $\mathbb{R}^+$ with the Lebesgue measure to be the levels of our towers. Start by taking the interval $[0, 1[$ to be $\mathcal{T}_1$. Now assume that the tower $\mathcal{T}_n$ has been built. The construction of $\mathcal{T}_{n+1}$ goes as follows.

Decompose $\mathcal{T}_n$ into three disjoint towers of equal measure $\mathcal{T}_n = \mathcal{T}_n^1 \sqcup \mathcal{T}_n^2 \sqcup \mathcal{T}_n^3$. Specifically, split each level of $\mathcal{T}_n$ into three intervals of length $\mu_n/3$, then put the left-most interval in $\mathcal{T}_n^1$, the middle one in $\mathcal{T}_n^2$ and the right one in $\mathcal{T}_n^3$. We will call *spacers* a collection of $3h_n + 2$ intervals of length $\mu_n/3$, disjoint from $\mathcal{T}_n$. We put a spacer on top of $\mathcal{T}_n^2$ and $3h_n + 1$ spacers on top of $\mathcal{T}_n^3$. Once the spacers are in place, we stack $\mathcal{T}_n^1$, $\mathcal{T}_n^2$ and $\mathcal{T}_n^3$ on top of each other, which yields $\mathcal{T}_{n+1}$. Therefore $\mathcal{T}_{n+1}$ is a tower of height $2(3h_n + 1)$ whose levels each have measure $\mu_n/3$, so $\mu(\mathcal{T}_{n+1}) \geq 2\mu(\mathcal{T}_n)$. Finally, for $k \in [\![1, h_{n+1} - 1]\!]$, define T on $E_{n+1}^{(k)}$ as the translation that sends $E_{n+1}^{(k)}$ to $E_{n+1}^{(k+1)}$ (which is possible because they are both intervals of the same length). The transformation is not yet defined on $E_{n+1}^{(h_{n+1})}$, that will be done in the next step of the construction.

We end the construction of (X, μ, T) by setting $X := \bigcup_{n\geq 1} \mathcal{T}_n$. Since $\mu(\mathcal{T}_{n+1}) \geq 2\mu(\mathcal{T}_n)$, we have $\mu(X) = \infty$.

2.4.3 Construction of the extension

Take $\mathbf{X} := (X, \mu, T)$ the system given by the infinite Chacon transformation introduced in Section **??**.

Fix $n_0 \in \mathbb{N}$ and $j_0 \in [\![1, h_{n_0}]\!]$. Set $A := E_{n_0}^{(j_0)}$, and consider the cocycle taking its values in $\mathbb{Z}/_{2\mathbb{Z}}$ (identified with $\{0, 1\}$):

$$\varphi := \mathbb{1}_A.$$

We study the system $\mathbf{Z}$ given on $(Z, \rho) := (X \times \{0, 1\}, \mu \otimes \mathscr{B}(1/2, 1/2))$ by the transformation

$$\begin{array}{cccc} T_\varphi : & X \times \{0,1\} & \longrightarrow & X \times \{0,1\} \\ & (x, \epsilon) & \longmapsto & (Tx, \epsilon + \varphi(x) \bmod 2) \end{array}.$$

This is a compact extension of $\mathbf{X}$.

Remark 2.4.3. By a simple induction, one can check that, for all $n \geq n_0$, there are 3^{n-n_0} levels in $\mathcal{T}_n$ that belong to A.

Theorem 2.4.4. *The system $\mathbf{Z} = (X \times \{0, 1\}, \rho, T_\varphi)$ is of infinite ergodic index, i.e. for every $p \geq 1$, $\mathbf{Z}^{\otimes p}$ is a conservative and ergodic system.*

It is known from the work in [**?**, Theorem 2.2] that $\mathbf{X}$ is of infinite ergodic index, therefore, our goal is to show that $\mathbf{Z}$ is as well. Let $p \geq 1$. Then $\mathbf{X}^{\otimes p}$ is a conservative and ergodic system, and $\mathbf{Z}^{\otimes p}$ is the compact extension given by the cocycle

$$\begin{array}{cccc} \varphi^{\times p} : & X^p & \longrightarrow & \{0,1\}^p \\ & (x_1, ..., x_p) & \mapsto & (\varphi(x_1), ..., \varphi(x_p)). \end{array}$$

By Lemma **??**, the characters of $\{0,1\}$ (identified to $\mathbb{Z}/_{2\mathbb{Z}}$) are

$$\chi_1 := 1 \text{ and } \chi_{-1} := (-1)^{\bullet}.$$

and the characters of $\{0,1\}^p$ are the tensor products of the form

$$\chi_I := \bigotimes_{i\notin I}\chi_1 \otimes \bigotimes_{i\in I}\chi_{-1} = \bigotimes_{i\in I}(-1)^{\bullet} : (\epsilon_1, ..., \epsilon_p) \mapsto (-1)^{\sum_{i\in I}\epsilon_i},$$

for $I \subset [\![1,p]\!]$. In particular, the character is entirely determined by the choice of the set I. Then, Lemma **??** tells us that $\mathbf{Z}^{\otimes p}$ is non-ergodic if and only if there exist $I \subset [\![1,p]\!]$ with $I \neq \varnothing$ and a map $f : X^p \longrightarrow \mathbb{U}$ such that

$$\frac{f(Tx_1, ..., Tx_p)}{f(x_1, ..., x_p)} = (-1)^{\sum_{i\in I}\varphi(x_i)} \text{ almost everywhere.} \tag{2.15}$$

Moreover, since the right-hand term can only take the values 1 and -1, the map that is equal to 1 when f is in $\{e^{i\theta};\ \theta \in [0,\pi[\}$ and -1 when f is in $\{e^{i\theta};\ \theta \in [\pi, 2\pi[\}$ satisfies the same equation as f. Therefore, we can simply look for maps $f : X^p \to \{\pm 1\}$ satisfying (**??**). However, Proposition **??** and Corollary **??** will show that such functions cannot exist, which will complete the proof of Theorem **??**.

We now turn our attention to the proofs of Proposition **??** and Corollary **??**. We start by giving some setup common to both of those results, and we will then conclude each proof separately in Sections **??** and **??**.

Denote $\nu_p := \mu^{\otimes p}$. We define $\mathcal{H}_n^p := \mathcal{T}_n \times \cdots \times \mathcal{T}_n$. We can decompose $\mathcal{H}_n^p$ as

$$\mathcal{H}_n^p = \bigsqcup_{k_1=1}^{h_n} \cdots \bigsqcup_{k_p=1}^{h_n} E_n^{(k_1)} \times \cdots \times E_n^{(k_p)}.$$

This gives a filtration on $X \times \cdots \times X$:

$$\mathscr{F}_n := \sigma\left(\{E_n^{(k_1)} \times \cdots \times E_n^{(k_p)}\}_{k_1,...,k_p \in [\![1,h_n]\!]}, (\mathcal{H}_n^p)^c\right).$$

Each $\mathscr{F}_n$ is not a σ-finite σ-algebra, because $\mu((\mathcal{H}_n^p)^c) = \infty$. So the conditional expectation $\mathbb{E}[\cdot \mid \mathscr{F}_n]$ is not well-defined. However, if we fix $N \in \mathbb{N}$ and consider the probability space $(\mathcal{H}_N^p, \frac{1}{\nu_p(\mathcal{H}_N^p)}\nu_p(\cdot \cap \mathcal{H}_N^p))$, for every $n \geq N$, $\mathscr{F}_n$ yields a finite partition of $\mathcal{H}_N^p$. Moreover, we can compute that, for any measurable function f:

$$\mathbb{E}_{\mathcal{H}_N^p}[f \mid \mathscr{F}_n] = \frac{1}{\mu_n^p}\int_{E_n^{(k_1)} \times \cdots \times E_n^{(k_p)}} f d\nu_p. \tag{2.16}$$

The important thing to note is that the right-hand term does not depend on N. Therefore we define the following, for $f : X^p \longrightarrow \mathbb{R}$:

$$\mathbb{E}[f \mid \mathscr{F}_n] := \begin{cases} \frac{1}{\mu_n^p} \int_{E_n^{(k_1)} \times \cdots \times E_n^{(k_p)}} f d\nu_p & \text{if } (x_1, ..., x_p) \in E_n^{(k_1)} \times \cdots \times E_n^{(k_p)} \\ 0 & \text{if } (x_1, ..., x_p) \notin \mathcal{H}_n^p \end{cases}.$$

Despite our choice of notation, this is not a true conditional expectation. However, since we have (**??**), we can conclude, using the fact that $X^p = \bigcup_{N \geq 1} \mathcal{H}_N^p$ and that $\bigvee_{n \geq 1} \mathscr{F}_n$ separates the points on X^p, that by the martingale convergence theorem, we have

$$\mathbb{E}[f \mid \mathscr{F}_n] \underset{n \to \infty}{\longrightarrow} f \text{ almost everywhere.}$$

Before we present the remaining details of the proof, we give a technical lemma:

Lemma 2.4.5. *Let $p \geq 1$. Let $i_1, ..., i_p \in \{1, 2, 3\}$. For almost every $(x_1, ..., x_p) \in X^p$, for every $M \geq 1$, there exits $n \geq M$ such that $(x_1, ..., x_p) \in \mathcal{T}_n^{i_1} \times \cdots \times \mathcal{T}_n^{i_p}$. In other words, for infinitely many $n \geq 1$, the points $x_1, ..., x_p$ belong to the thirds of $\mathcal{T}_n$ with respective indexes $i_1, ..., i_p$.*

Proof. We recall that $\mathcal{H}_n^p := \mathcal{T}_n \times \cdots \times \mathcal{T}_n$ and $\nu_p := \mu^{\otimes p}$. Let $M \geq 1$. Take $M' \geq M$, and note that

$$\begin{aligned} \nu_p\big(\forall n \geq M, &(x_1, ..., x_p) \in \mathcal{H}_n^p \backslash (\mathcal{T}_n^{i_1} \times \cdots \times \mathcal{T}_n^{i_p})\big) \\ &\leq \nu_p \left(\forall n \in [\![M, M']\!], (x_1, ..., x_p) \in \mathcal{H}_n^p \backslash (\mathcal{T}_n^{i_1} \times \cdots \times \mathcal{T}_n^{i_p})\right). \end{aligned}$$

A straightforward induction on M' shows that

$$\begin{aligned} \nu_p\big(\forall n \in [\![M, M']\!], (x_1, ..., x_p) \in \mathcal{H}_n^p \backslash (\mathcal{T}_n^{i_1} \times \cdots \times \mathcal{T}_n^{i_p})\big) \\ = \left(\frac{3^p - 1}{3^p}\right)^{(M' - M + 1)} \nu_p(\mathcal{H}_M^p) \underset{M' \to \infty}{\longrightarrow} 0. \end{aligned}$$

Therefore:

$$\nu_p \left(\forall n \geq M, (x_1, ..., x_p) \in \mathcal{H}_n^p \backslash (\mathcal{T}_n^{i_1} \times \cdots \times \mathcal{T}_n^{i_p})\right) = 0,$$

which implies

$$\nu_p \left(\exists M \geq 1, \forall n \geq M, (x_1, ..., x_p) \in \mathcal{H}_n^p \backslash (\mathcal{T}_n^{i_1} \times \cdots \times \mathcal{T}_n^{i_p})\right) = 0.$$

Combining this with the fact that $X^p = \bigcup_{n \geq 1} \mathcal{H}_n^p$ ends our proof. □

Case where $\#I$ is odd

Let us first give a rough sketch of the argument to prove that we cannot find f satisfying (**??**) when $\#I$ is odd. For such a function f, we study the evolution of its values along the orbit of a point $(x_1, ..., x_p)$. By (**??**), every time some coordinate of index $i \in I$ goes through A, it causes a change of sign for f. Now, if we choose n so that all the $\{x_i\}_{i \in [\![1,p]\!]}$ start in $\mathcal{T}_n^1$, to avoid hitting the spacers, after h_n applications of $T \times \cdots \times T$, each coordinate is then back in the level from which it started and its orbit has gone through each level of $\mathcal{T}_n$ exactly once. But we know from Remark **??** that each tower $\mathcal{T}_n$ has 3^{n-n_0} levels that are subsets of A. So, on the piece of the orbit that we consider, the sign of f changes $\#I \cdot 3^{n-n_0}$ times, and since $\#I$ is odd, this means that the sign of f changes. But, since each x_i is back on the level from which it started, if f is close enough to being constant on cells of the form $E_n^{(k_1)} \times \cdots \times E_n^{(k_p)}$, this yields a contradiction. We detail this argument in the following proposition:

Proposition 2.4.6. *Let $I \subset [\![1,p]\!]$ such that $\#I$ is odd. There is no measurable map $f : X^p \longrightarrow \{\pm 1\}$ that satisfies* (**??**)*.*

Proof. Suppose by contradiction that there exists $f : X^p \longrightarrow \{\pm 1\}$ that satisfies (**??**). Take $\delta > 0$ and $(x_1, ..., x_p) \in X^p$. Up to a set of measure 0, we may assume that there exists $N \geq 1$ such that $\forall n \geq N$, we have

$$|\mathbb{E}[f \mid \mathscr{F}_n](x_1, ..., x_p) - f(x_1, ..., x_p)| \leq \delta. \tag{2.17}$$

We know from Lemma **??**, that, up to another set of measure 0, we may assume that there is $n \geq N$ such that

$$(x_1, ..., x_p) \in \mathcal{T}_n^1 \times \cdots \times \mathcal{T}_n^1.$$

Let $(k_1, ..., k_p)$ such that for $i \in [\![1,p]\!]$, $x_i \in E_n^{(k_i)}$. Using the definition of $\mathbb{E}[f \mid \mathscr{F}_n]$, we denote $E := E_n^{(k_1)} \times \cdots \times E_n^{(k_p)}$, $x := (x_1, ..., x_p)$ and compute

$$\begin{aligned}
|\mathbb{E}[f \mid \mathscr{F}_n](x) - f(x)| &= \left| \frac{1}{\nu_p(E)} \int_E f \mathrm{d}\nu_p - f(x) \right| \\
&= \left| f(x) \frac{\nu_p(E \cap \{f = f(x)\}) - \nu_p(E \cap \{f = -f(x)\})}{\nu_p(E)} - f(x) \right| \\
&= \left| \frac{\nu_p(E) - \nu_p(E \cap \{f = -f(x)\}) - \nu_p(E \cap \{f = -f(x)\})}{\nu_p(E)} - 1 \right| \\
&= 2 \frac{\nu_p(E \cap \{f \neq f(x)\})}{\nu_p(E)},
\end{aligned}$$

where we use the facts that f takes only two possible values and $|f| = 1$. This shows that (**??**) implies

$$\nu_p((E_n^{(k_1)} \times \cdots \times E_n^{(k_p)}) \cap \{f \neq f(x_1, ..., x_p)\}) \leq \frac{\delta}{2}\mu_n^p. \tag{2.18}$$

Define

$$B := (E_n^{(k_1)} \times \cdots \times E_n^{(k_p)}) \cap (\mathcal{T}_n^1 \times \cdots \times \mathcal{T}_n^1),$$

and

$$C := (E_n^{(k_1)} \times \cdots \times E_n^{(k_p)}) \cap (\mathcal{T}_n^2 \times \cdots \times \mathcal{T}_n^2).$$

Since there is no spacer on top of $\mathcal{T}_n^1$, we get that $(T \times \cdots \times T)^{h_n} B = C$. However, since $\#I$ is odd, for any $(x'_1, .., x'_p) \in B$, Remark **??** and (**??**) imply that

$$f(T^{h_n}x'_1, ..., T^{h_n}x'_p) = (-1)^{\#I \cdot 3^{n-n_0}} f(x'_1, ..., x'_p) = -f(x'_1, ..., x'_p).$$

Therefore

$$\begin{aligned}\nu_p(C \cap \{f = -f(x_1, ..., x_p)\}) &= \nu_p(B \cap \{f = f(x_1, ..., x_p)\}) \\ &\geq \nu_p(B) - \frac{\delta}{2}\mu_n^p.\end{aligned}$$

But, by construction, $\nu_p(B) = \mu_n^p/3^p$. So

$$\begin{aligned}\nu_p(C \cap \{f = -f(x_1, ..., x_p)\}) &\geq \frac{\mu_n^p}{3^p} - \frac{\delta}{2}\mu_n^p \\ &= \left(\frac{1}{3^p} - \frac{\delta}{2}\right)\mu_n^p > \frac{\delta}{2}\mu_n^p,\end{aligned}$$

if $\delta < 1/3^p$. Since $C \subset E_n^{(k_1)} \times \cdots \times E_n^{(k_p)}$, combined with (**??**), this implies that $f(x_1, ..., x_p) = -f(x_1, ..., x_p)$. However, by definition, $f \neq 0$, so this is absurd. □

Case where $\#I$ is even

If $\#I$ is even, the argument given above no longer works: when the orbits of all the $\{x_i\}_{i \in [\![1,p]\!]}$ go through the levels of $\mathcal{T}_n$, it causes an even number of sign changes for f, which means that f remains unchanged. Here we need to make use of the placement of the spacers in the construction of T. We will fix $i_0 \in I$ and choose n so that all the points $\{x_i\}_{i \neq i_0}$ start in $\mathcal{T}_n^2$ so that they can hit the spacer. On the other hand, the choice of n enables us to assume that x_{i_0} will start in $\mathcal{T}_n^1$ so that its orbit can go through all the levels of $\mathcal{T}_n$ without hitting a spacer. Therefore, after $h_n + 1$ applications of $T \times \cdots \times T$, all the coordinates of index $i \neq i_0$ will be back in the level from which they started, but x_{i_0} will be one level higher, which will yield the equation we get in Proposition **??**. Then applying our reasoning from the odd case will give a contradiction, as stated in Corollary **??**.

Let us show the following:

Proposition 2.4.7. *Let $I \subset [\![1, p]\!]$ such that $I \neq \varnothing$ and $\#I$ is even. Assume that there exists $f : X^p \longrightarrow \{\pm 1\}$ that satisfies (**??**). Then, for ever $i \in I$, we have:*

$$\frac{f(x_1, ..., Tx_i, ..., x_p)}{f(x_1, ..., x_p)} = (-1)^{\varphi(x_i)} \text{ almost everywhere.} \tag{2.19}$$

Proof. Up to a permutation of coordinates, we can assume that $1 \in I$ and deal with the case $i = 1$. Define also $\tilde{\mathcal{T}}_n := \mathcal{T}_n \cap (X \backslash E_n^{(h_n)})$, i.e. we get $\tilde{\mathcal{T}}_n$ by removing the top level from $\mathcal{T}_n$.

Let $(x_1, ..., x_p) \in X^p$. Take a small $\delta > 0$. Up to a set of measure 0, we can assume that there exists $N \geq 1$ such that $\forall n \geq N$, we have

$$\begin{cases} |\mathbb{E}[f \mid \mathscr{F}_n](x_1, ..., x_p) - f(x_1, ..., x_p)| \leq \delta, \\ \\ |\mathbb{E}[f \mid \mathscr{F}_n](Tx_1, ..., x_p) - f(Tx_1, ..., x_p)| \leq \delta. \end{cases} \tag{2.20}$$

Also note that in the construction of a tower, the top level is obtained with a spacer, so if $x_1 \in \mathcal{T}_m$, then for every $n > m$, $x_1 \in \mathcal{T}_n \backslash E_n^{(h_n)} = \tilde{\mathcal{T}}_n$. Since $X = \bigcup_{n \geq 1} \mathcal{T}_n$, this means that if N is large enough, we can also assume that for every $n \geq N$, $x_1 \in \tilde{\mathcal{T}}_n$. Finally, using Lemma **??**, we know that up to another set of measure zero, we can find $n \geq N$ for which $(x_1, ..., x_p) \in \mathcal{T}_n^1 \times \mathcal{T}_n^2 \times \cdots \times \mathcal{T}_n^2$.

Set $k_1, ..., k_p$ such that $x_i \in E_n^{(k_i)}$ for every $i \in [\![1, p]\!]$. The fact that $x_1 \in \tilde{\mathcal{T}}_n$ implies that $k_1 \leq h_n - 1$. The same computation that gave (**??**) shows that (**??**) implies

$$\nu_p\big((E_n^{(k_1)} \times \cdots \times E_n^{(k_p)}) \cap \{f \neq f(x_1, ..., x_p)\}\big) \leq \frac{\delta}{2} \mu_n^p, \tag{2.21}$$

and

$$\nu_p\big((E_n^{(k_1+1)} \times E_n^{(k_2)} \times \cdots \times E_n^{(k_p)}) \cap \{f \neq f(Tx_1, ..., x_p)\}\big) \leq \frac{\delta}{2} \mu_n^p. \tag{2.22}$$

Denote

$$B := (E_n^{(k_1)} \times \cdots \times E_n^{(k_p)}) \cap (\mathcal{T}_n^1 \times \mathcal{T}_n^2 \times \cdots \times \mathcal{T}_n^2)$$

and

$$C := (E_n^{(k_1+1)} \times E_n^{(k_2)} \times \cdots \times E_n^{(k_p)}) \cap (\mathcal{T}_n^2 \times \mathcal{T}_n^3 \times \cdots \times \mathcal{T}_n^3).$$

One can check that, because there is a spacer on top of $\mathcal{T}_n^2$ and no spacer on top of $\mathcal{T}_n^1$, we get

$$(T \times \cdots \times T)^{h_n+1} B = C.$$

However, using the fact that $B \subset \tilde{\mathcal{T}}_n^1 \times \mathcal{T}_n^2 \times \cdots \times \mathcal{T}_n^2$, we can track the times f changes sign between B and C. The orbits of the $\{x_i\}_{i \neq 1}$ go through each level of $\mathcal{T}_n$, therefore contributing $(\#I - 1)3^{n-n_0}$ sign changes. The orbit of x_1 goes through every level of $\mathcal{T}_n$ after h_n applications of T, contributing 3^{n-n_0} sign changes, and one additional sign change from the $h_n + 1$-th application if $E_n^{(k_1)} \subset A$, or, equivalently, if $\varphi(x_1) = 1$. This means that for every $(x_1', ..., x_p') \in B$, we have

$$\begin{aligned} f(T^{h_n+1}x_1', ..., T^{h_n+1}x_p') &= (-1)^{\#I3^{n-n_0} + \mathbb{1}_{E_n^{(k_1)} \subset A}} f(x_1', ..., x_p') \\ &= (-1)^{\varphi(x_1)} f(x_1', ..., x_p') \text{ because } \#I \text{ is even.} \end{aligned}$$

So, we have

$$\begin{aligned} \nu_p(C \cap \{f = &(-1)^{\varphi(x_1)} f(x_1, ..., x_p)\}) = \nu_p(B \cap \{f = f(x_1, ..., x_p)\}) \\ &= \nu_p(B) - \nu_p(B \cap \{f \neq f(x_1, ..., x_p)\}) \\ &\geq \nu_p(B) - \nu_p\big((E_n^{(k_1)} \times \cdots \times E_n^{(k_p)}) \cap \{f \neq f(x_1, ..., x_p)\}\big) \\ &\geq \nu_p(B) - \frac{\delta}{2}\mu_n^p, \text{ because of (\textbf{??}).} \end{aligned}$$

Moreover, $\nu_p(B) = \mu_n^p / 3^p$, so

$$\begin{aligned} \nu_p\big((E_n^{k_1+1)} \times \cdots \times E_n^{(k_p)}) \cap &\{f = (-1)^{\varphi(x_1)} f(x_1, ..., x_p)\}\big) \\ &\geq \nu_p(C \cap \{f = (-1)^{\varphi(x_1)} f(x_1, ..., x_p)\}) \\ &\geq \left(\frac{1}{3^p} - \frac{\delta}{2}\right) \mu_n^p > \frac{\delta}{2}\mu_n^p, \end{aligned}$$

if δ is small enough ($\delta < 1/3^p$). However, this is only compatible with (**??**) if

$$(-1)^{\varphi(x_1)} f(x_1, ..., x_p) = f(Tx_1, ..., x_p).$$

□

Corollary 2.4.7.1. *Let $I \subset [\![1, p]\!]$ such that $I \neq \varnothing$ and $\#I$ is even. There is no measurable map $f : X^p \longrightarrow \{\pm 1\}$ that satisfies* (**??**).

Proof. Assume that there is a measurable map $f : X^p \longrightarrow \{\pm 1\}$ that satisfies (**??**). Up to a permutation of coordinates, we may assume that $1 \in I$. Fix $(x_2, ..., x_p) \in X^{p-1}$. Up to a set of measure 0 in our choice of $(x_2, .., x_p)$, we know from Proposition **??** that the map $g_{(x_2,...,x_p)} : x_1 \mapsto f(x_1, ..., x_p)$ satisfies

$$\frac{g_{(x_2,...,x_p)}(Tx_1)}{g_{(x_2,...,x_p)}(x_1)} = (-1)^{\varphi(x_1)} \text{ almost everywhere.}$$

However, Proposition **??** tells us that such a map cannot exist. □

Remark 2.4.8. Now that the proof of Theorem **??** is complete, we have an example of compact extension of an infinite measure preserving system that is of infinite ergodic index. But we only proved the infinite ergodic index of a specific extension, and we wonder if more general results can be found. In particular, a significant difference between finite and infinite ergodic theory is the fact that for probability preserving systems, an ergodic index greater or equal to 2 is automatically infinite. This is not true in the infinite measure case, but we could consider an intermediate situation: take an extension $\mathbf{Z} \xrightarrow{\pi} \mathbf{X}$ of σ-finite infinite measure systems and assume that $\mathbf{X}$ has an infinite ergodic index. Is it possible that $\mathbf{Z}$ have a finite ergodic index greater or equal to 2 ? In our example, proving that the ergodic index is at least 2 contains exactly as much difficulty as proving it is infinite, therefore suggesting that the answer could be negative.

Chapter 3

A class of dynamical filtrations: weak Pinsker filtrations

In 1958, Kolmogorov and Sinaï introduced the notion of entropy in ergodic theory: the Kolmogorov-Sinaï entropy (or KS-entropy). The same year, Kolmogorov introduced another important notion: K-systems. He defines a K-system as a dynamical system $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ on which there is a finite generator ξ whose tail σ-algebra $\bigcap_{n\geq 1} \sigma(\xi_{]-\infty,-n]})$ is trivial. There is an equivalent definition that is more intrinsic to the system: $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ is a K-system if, and only if, every non-trivial observable ξ_0 satisfies $h_\mu(\xi, T) > 0$ (a proof of this equivalence, and a more complete presentation of this notion can be found in [**?**]). It is also equivalent to assume that the Pinsker factor of the system is trivial, the Pinsker factor being the σ-algebra

$$\Pi_{\mathbf{X}} = \{A \in \mathscr{A} \mid h(\mathbb{1}_A, T) = 0\}.$$

The Pinsker factor is simply the largest factor of $\mathbf{X}$ that is of entropy 0. Therefore, a K-system has no non-trivial factor of entropy 0: it is entirely non-deterministic. For example, the most elementary K-systems are the Bernoulli shifts. They are K-systems because i.i.d. processes satisfy Kolmogorov's 0-1 law.

Entropy is an invariant that quantifies the "chaos" of a dynamical system, or more precisely its unpredictability, and many of the questions that arose after its discovery were aimed at understanding the structure of this "chaos". The first question, which Kolmogorov posed after proving that Bernoulli shifts are K-systems, was whether all K-systems are Bernoulli shifts, which would imply that these chaotic systems have a very simple structure. More general questions then emerged, and we will return to them in the following paragraphs.

The discovery of entropy first led to non-isomorphism results, particularly for Bernoulli shifts: two isomorphic Bernoulli shifts must have the same entropy. The converse of this result, shown by Ornstein ([**?**], [**?**]), is one of the most notable successes of the KS-entropy. But the ramifications of Ornstein's theory go far beyond Bernoulli shifts, and have had a profound impact on the evolution of ergodic theory. We will confine ourselves here to telling the story of the *weak Pinsker property*.

In the early 1960s, Pinsker, then working in Moscow with Kolmogorov, showed that any K-factor of $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ is independent of the Pinsker factor $\Pi_{\mathbf{X}}$ (see [**?**], but this reference is in Russian). Following this result, although the existence of any specific K-factor had not yet been proved, he issued a conjecture (later called the "Pinsker conjecture"): any system of non-zero entropy is isomorphic to the direct product of its Pinsker factor and a K-system. A few years later, Sinai published [**?**] which seemed to confirm this conjecture: he proved the existence of a factor of $\mathbf{X}$ isomorphic to a Bernoulli shift of the same entropy as $\mathbf{X}$. Given Pinsker's independence result, it would have been sufficient to prove that the factor constructed by Sinaï and the Pinsker factor generate the entire σ-algebra $\mathscr{A}$ to obtain a result even stronger than Pinsker's conjecture: $\mathbf{X}$ would then be isomorphic to the direct product of its Pinsker factor and a Bernoulli shift. This "strong Pinsker conjecture" would also have proved that any K-system is isomorphic to a Bernoulli shift.

But this conjecture turned out to be false: Ornstein published a first example of a non-Bernoulli K-system [**?**] which contradicts the strong Pinsker conjecture. Following that, many other counterexamples were built, and it turns out that the family of all K-systems is very broad, leaving little hope for a complete classification of those systems. Among all these counterexamples, we can find a construction by Ornstein [**?**] that can be used to contradict Pinsker's conjecture. Furthermore, he then refines this result by constructing a *mixing* system that does not verify Pinsker's conjecture [**?**]. Thus, all the conjectures formulated in the early years of the study of KS-entropy turned out to be wrong, revealing a wide variety of possible phenomena.

One of the ramifications of Ornstein's work can be found in the work of Thouvenot, who, starting in 1975, became interested in relatively Bernoulli systems and developed a "relative" version of Ornstein's theory. Following this work, in his 1977 paper [**?**], he introduced the *weak Pinsker property*: for any $\varepsilon > 0$, $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ is isomorphic to the direct product of a Bernoulli shift $\mathbf{B}$ and a system $\mathbf{X}_\varepsilon$ of entropy ε:

$$\mathbf{X} \cong \mathbf{X}_\varepsilon \otimes \mathbf{B}. \tag{3.1}$$

For four decades, however, it was unclear whether all systems verified the weak Pinsker property. But in 2018, Austin published a paper on the subject [**?**] in which he proved that all ergodic systems satisfy the weak Pinsker property.

We can then iterate this splitting operation: take $(\varepsilon_n)_{n\leq -1}$ an increasing sequence of positive numbers such that $\lim_{n\to-\infty}\varepsilon_n = 0$, and start by splitting $\mathbf{X}$ into

$$\mathbf{X} \cong \mathbf{X}_{\varepsilon_{-1}} \otimes \mathbf{B}_{-1},$$

then split $\mathbf{X}_{\varepsilon_{-1}}$ into

$$\mathbf{X}_{\varepsilon_{-1}} \cong \mathbf{X}_{\varepsilon_{-2}} \otimes \mathbf{B}_{-2},$$

and so on. This yields a sequence of systems $(\mathbf{X}_{\varepsilon_n})_{n\leq -1}$ such that, for every $n \leq -1$, $\mathbf{X}_{\varepsilon_n}$ is a factor of $\mathbf{X}_{\varepsilon_{n+1}}$. By composing the factor maps, it means that each $\mathbf{X}_{\varepsilon_n}$ is a factor of $\mathbf{X}$, and therefore generates a T-invariant σ-algebra $\mathscr{F}_n := \sigma(\mathbf{X}_{\varepsilon_n}) \subset \mathscr{A}$. Because of our iterating construction, we see that $\mathscr{F}_n \subset \mathscr{F}_{n+1}$, so the sequence $\mathscr{F} := (\mathscr{F}_n)_{n\leq 0}$ is a filtration. This is what we call a *weak Pinsker filtration on* $\mathbf{X}$ (see Definition **??**).

The purpose of this chapter is to discuss how weak Pinsker filtrations can be used as a tool to describe the structure of dynamical systems with positive entropy. In particular, in Section **??**, we have introduced some concepts from the theory of dynamical filtrations, and this is the framework we mean to use to study weak Pinsker filtrations. This framework is focused on the various possible structures of filtrations whose tail σ-algebra $\bigcap_{n\leq 0}\mathscr{F}_n$ is trivial, which is the type of weak Pinsker filtrations that appear on K-systems (see Theorem **??**). Therefore, the study of weak Pinsker filtrations we suggest would mainly be aimed at classifying K-systems, and in particular non-Bernoulli K-systems.

In Section **??**, we give an overview of the results and open questions that arise from the study of the properties of weak Pinsker filtrations, and their relation to the structure of the underlying dynamical system. One of those questions concerns the uniqueness, up to isomorphism, of weak Pinsker filtrations. In Section **??**, we give a partial answer to this uniqueness problem in the case of Bernoulli systems. That section is based on ideas suggested to us by Thouvenot. Finally, in Section **??**, we give explicit examples of weak Pinsker filtrations, in order to give a more concrete meaning to all of those abstract notions.

3.1 Introduction of weak Pinsker filtrations and related questions

In this section, we introduce the notion of weak Pinsker filtrations, the tools necessary to study them and state some of the main questions concerning those filtrations. Since weak Pinsker filtrations are dynamical filtrations, we will use the framework for classifying dynamical filtrations presented in Chapter **??** (specifically, see Section **??** for the complete definitions).

3.1.1 Reminders on KS-entropy and Ornstein's theory and its relative version

The notion of entropy first appeared in mathematics in 1948, introduced by Shannon in his foundational work on information theory [**?**]. It is defined as follows:

Definition 3.1.1 (Shannon entropy)**.** *Let* $(X, \mathscr{A}, \mu)$ *be a probability space and* $\xi : X \to A$ *a random variable, with* A *finite or countable. The Shannon entropy of* ξ *is*

$$H_\mu(\xi) := -\sum_{a \in A} \mu(\{\xi = a\}) \cdot \log \mu(\{\xi = a\}).$$

The number $H_\mu(\xi)$ gives the average amount of information given by the random variable ξ. If we have a probability measure ρ defined directly on A, we can also define the entropy of that measure

$$H(\rho) := -\sum_{a \in A} \rho(a) \cdot \log \rho(a).$$

In 1958, Kolmogorov and Sinaï used this entropy as a tool to help to describe quantitatively the behavior of measure preserving dynamical systems.

Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a dynamical system. To any random variable $\xi_0 : X \to A$, with A finite or countable, we associate $\xi : X \to A^{\mathbb{Z}}$ the corresponding T-process

$$\xi := (\xi_n)_{n \in \mathbb{Z}} := (\xi_0 \circ T^n)_{n \in \mathbb{Z}}.$$

Also, for $F \subset \mathbb{Z}$, set $\xi_F := (\xi_n)_{n \in F}$.

The Kolmogorov-Sinaï entropy (or KS-entropy) of a dynamical system is:

Definition 3.1.2 (Kolmogorov-Sinaï entropy)**.** *Let* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *be a dynamical system. For a random variable* $\xi_0 : X \to A$*, define*

$$h_\mu(\xi, T) := \lim_{n\to\infty} \frac{1}{n} H_\mu(\xi_{[\![0,n]\!]}).$$

For a T*-invariant* σ*-algebra* $\mathscr{B} \subset \mathscr{A}$*, define*

$$h_\mu(\mathscr{B}, T) := \sup\{h_\mu(\xi, T)\,;\ \xi_0 \text{ a } \mathscr{B}\text{-measurable random variable}\}.$$

Finally, set

$$h(\mathbf{X}) := h_\mu(\mathscr{A}, T).$$

The KS-entropy satisfies the following continuity result:

Lemma 3.1.3. *Let* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *be a dynamical system and a random variable* $\xi_0 : X \to A$*, with* A *finite. For* $\varepsilon > 0$*, there exists* $\delta > 0$ *such that, for any random variable* $\zeta_0 : X \to A$ *such that* $\mu(\zeta_0 \neq \xi_0) \leq \delta$*, we have*

$$|h_\mu(\xi, T) - h_\mu(\zeta, T)| \leq \varepsilon.$$

Proof. In this proof, we will use the conditional entropy: for $\chi_1 : X \to Y_1$ and $\chi_2 : X \to Y_2$ be two random variables, we define

$$H_\mu(\chi_1 \,|\, \chi_2) := \sum_{y_2 \in Y_2} \mu(\chi_2 = y_2) \sum_{y_1 \in Y_1} \varphi(\mu(\chi_1 = y_1 \,|\, \chi_2 = y_2)),$$

where $\varphi(x) = -x \cdot \log(x)$. We refer to [**?**, Chapter 2, Section 6] for the basic properties of this notion. Set $d := \mu(\zeta_0 \neq \xi_0)$ and note that [**?**, Theorem 6.2] states that

$$H_\mu(\xi_0 \,|\, \zeta_0) \leq \varphi(d) + \varphi(1-d) + d\log(\#A).$$

Now we compute:

$$\begin{aligned} h(\xi, T) &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\xi_{[0,n[}) \leq \lim_{n\to\infty} \frac{1}{n} H_\mu((\xi \vee \zeta)_{[0,n[}) \\ &\leq \lim_{n\to\infty} \frac{1}{n} \left(H_\mu(\zeta_{[0,n[}) + \sum_{j=0}^{n-1} H_\mu(\xi_j \,|\, \zeta_{[0,n[}) \right) \\ &\leq \lim_{n\to\infty} \frac{1}{n} \left(H_\mu(\zeta_{[0,n[}) + \sum_{j=0}^{n-1} H_\mu(\xi_j \,|\, \zeta_j) \right) \\ &\leq h_\mu(\zeta, T) + H_\mu(\xi_0 \,|\, \zeta_0) \\ &\leq h_\mu(\zeta, T) + \varphi(d) + \varphi(1-d) + d\log(\#A). \end{aligned}$$

And, since φ is continuous, there exists $\delta > 0$ such that, if $d \leq \delta$, we have

$$h_\mu(\xi, T) \leq h_\mu(\zeta, T) + \varepsilon.$$

By switching ξ and ζ and doing the same reasoning, we and the proof. □

It is useful to locate the deterministic aspects of a dynamical system. We do that by considering the Pinsker factor of a system, which is defined as, for any factor σ-algebra $\mathscr{B}$

$$\Pi_{\mathscr{B}} = \{A \in \mathscr{B} \mid h(\mathbb{1}_A, T) = 0\},$$

and $\Pi_{\mathbf{X}} := \Pi_{\mathscr{A}}$. We will use the following basic result, which can be found in [**?**, Theorem 14]:

Lemma 3.1.4. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a dynamical system and $\mathscr{B}$ and $\mathscr{C}$ be independent factor σ-algebras. We have*

$$\Pi_{\mathscr{B} \vee \mathscr{C}} = \Pi_{\mathscr{B}} \vee \Pi_{\mathscr{C}}.$$

To be able to compute the entropy of a system, the following result proves to be most useful.

Theorem 3.1.5 (Kolmogorov-Sinaï)**.** *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a dynamical system. Consider a random variable $\xi_0 : X \to A$ and the corresponding T-process $\xi := (\xi_0 \circ T^n)_{n \in \mathbb{Z}}$. Then we have*

$$h_\mu(\sigma(\xi), T) = h_\mu(\xi, T).$$

In particular, if ξ is a generator of $\mathscr{A}$ (i.e. $\mathscr{A} = \sigma(\xi)$ mod μ), then $h(\mathbf{X}) = h_\mu(\xi, T)$.

From the definition, one easily sees that KS-entropy is invariant under isomorphism of dynamical systems, which makes it a useful tool in the classification of measure preserving dynamical systems. The most remarkable classification results concern Bernoulli and relatively Bernoulli systems:

Definition 3.1.6 (Bernoulli and relatively Bernoulli)**.** *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a dynamical system.*

We say that $\mathbf{X}$ (or $\mathscr{A}$) is Bernoulli *if there exists a random variable $\xi_0 : X \to A$ such that the corresponding process $\xi := (\xi_0 \circ T^n)_{n \in \mathbb{Z}}$ is i.i.d. and generates $\mathscr{A}$, i.e. we have $\sigma(\xi) = \mathscr{A}$ mod μ.*

Let $\mathscr{B} \subset \mathscr{A}$ be a factor σ-algebra. We say that $\mathbf{X}$ (or $\mathscr{A}$) is relatively Bernoulli over $\mathscr{B}$ *if there is an i.i.d. process of the form $\xi := (\xi_0 \circ T^n)_{n \in \mathbb{Z}}$ such that $\sigma(\xi)$ is independent of $\mathscr{B}$ and $\mathscr{A} = \mathscr{B} \vee \sigma(\xi)$ mod μ.*

Those two definitions coincide when $\mathscr{B}$ is the trivial factor σ-algebra: $\mathbf{X}$ is relatively Bernoulli over $\{\varnothing, X\}$ if and only if $\mathbf{X}$ is Bernoulli.

Remark 3.1.7. We can consider another approach to define Bernoulli systems: take A a finite or countable set and ρ a probability measure on A. On the product probability space $(A^{\mathbb{Z}}, \rho^{\otimes \mathbb{Z}})$, consider the transformation

$$S : (a_n)_{n\in\mathbb{Z}} \mapsto (a_{n+1})_{n\in\mathbb{Z}}.$$

The map S is called the *shift* on $A^{\mathbb{Z}}$. One can easily check that $\rho^{\otimes\mathbb{Z}}$ is S-invariant. Therefore, this yields a measure preserving dynamical system

$$\mathbf{B} := (A^{\mathbb{Z}}, \rho^{\otimes\mathbb{Z}}, S), \tag{3.2}$$

which is called a *Bernoulli shift*. Then a system is Bernoulli if and only if it is isomorphic to a Bernoulli shift. Similarly, we can see that a system $\mathbf{X}$ is relatively Bernoulli over a factor σ-algebra $\mathscr{B}$ if and only if $\mathbf{X}$ is isomorphic to a system of the form $\mathbf{Y} \otimes \mathbf{B}$ via a factor map $\varphi : \mathbf{X} \longrightarrow \mathbf{Y} \times \mathbf{B}$ such that $\sigma(\pi_{\mathbf{Y}} \circ \varphi) = \mathscr{B}$ mod μ (where $\pi_{\mathbf{Y}}$ is the projection of $\mathbf{Y} \otimes \mathbf{B}$ onto $\mathbf{Y}$).

Using Theorem **??**, it is easy to compute the entropy of a Bernoulli system. Let ξ be an $i.i.d.$ process on $\mathbf{X}$ that generates $\mathscr{A}$. We then have

$$\begin{aligned} h(\mathbf{X}) = h_\mu(\mathscr{A}, T) = h_\mu(\xi, T) &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\xi_{[\![0,n]\!]}) \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} H_\mu(\xi_i) = H_\mu(\xi_0). \end{aligned}$$

In particular, if $\mathbf{X}$ is isomorphic to a system of the form (**??**), we get

$$h(\mathbf{X}) = h(\mathbf{B}) = -\sum_{a\in A} \rho(a) \log(\rho(a)).$$

Since, to be isomorphic, two systems need to have the same entropy, this computation enables us to get a non-isomorphism result between any two Bernoulli systems of different entropy. Remarkably, Ornstein proved that the converse is also true:

Theorem 3.1.8 (Ornstein [**?**], [**?**])**.** *If* $\mathbf{X}$ *and* $\mathbf{Y}$ *are Bernoulli systems such that* $h(\mathbf{X}) = h(\mathbf{Y})$*, then* $\mathbf{X} \cong \mathbf{Y}$.

This means that the KS-entropy gives a complete classification of Bernoulli systems. An outstanding result that emerged from Ornstein's theory was a criterion to characterize Bernoulli systems: finite determination. However, although this notion is useful in proving abstract results, when studying a given system, it is not easy to know whether or not it is finitely determined. Because of that, another criterion called *very weak Bernoullicity* was developed (see [**?**, Section 7]). This is the criterion we are interested in.

For the remainder of this section, we assume that the processes are defined on finite alphabets. We first need a technical definition. Given a finite alphabet A, an integer $\ell \geq 1$ and two words $\mathbf{a}, \mathbf{b} \in A^\ell$ of length ℓ on A, we define the normalized Hamming metric between $\mathbf{a}$ and $\mathbf{b}$ as:

$$d_\ell(\mathbf{a}, \mathbf{b}) := \frac{1}{\ell} \#\{i \in [\![1, \ell]\!] \,|\, a_i \neq b_i\},$$

where $\mathbf{a} = (a_1, ..., a_\ell)$ and $\mathbf{b} = (b_1, ..., b_\ell)$. We then consider the corresponding transportation metric on $\mathscr{P}(A^\ell)$:

$$\forall \mu, \nu \in \mathscr{P}(A^\ell),\ \bar{d}_\ell(\mu, \nu) := \inf \left\{ \int d_\ell(\mathbf{a}, \mathbf{b}) d\lambda(\mathbf{a}, \mathbf{b})\,;\ \lambda \text{ a coupling of } \mu \text{ and } \nu \right\}.$$

Then a process ξ is said to be very weak Bernoulli if, for some $\ell \geq 1$, the conditional law of $\xi_{[0,\ell[}$ given the past of ξ is close enough to the law of $\xi_{[0,\ell[}$ in the $\bar{d}_\ell$ metric. More formally, we state:

Definition 3.1.9 (Very weak Bernoulli)**.** *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be an ergodic dynamical system, equipped with a process ξ taking values in a finite alphabet. We say that ξ is very weak Bernoulli if, for every $\varepsilon > 0$, there exists $\ell \geq 1$ such that for every $m \geq 1$, we have*

$$\int \bar{d}_\ell \left(\nu_\ell(\cdot \,|\, \mathbf{a}_{[-m,0[}), \nu_\ell(\cdot) \right) d\nu(\mathbf{a}) \leq \varepsilon,$$

where ν is the law of ξ and, for $I \subset \mathbb{Z}$, $\nu_\ell(\cdot \,|\, \mathbf{a}_I)$ is the conditional law of $\xi_{[0,\ell[}$ given that ξ_I equals $\mathbf{a}_I$.

If $\mathscr{A} = \sigma(\xi)$, we say that $\mathbf{X}$ (or $\mathscr{A}$) is very weak Bernoulli.

The fact that very weak Bernoullicity characterizes Bernoulli systems can be stated as follows:

Theorem 3.1.10 (see [**?**], [**?**])**.** *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a dynamical system. A process ξ on $\mathbf{X}$ is very weak Bernoulli if and only if $\sigma(\xi)$ is Bernoulli.*

Following the work of Ornstein, Thouvenot studied relatively Bernoulli systems and adapted the definitions of finite determination and very weak Bernoullicity to get criteria that characterize relatively Bernoulli systems. Here we give his adaptation of very weak Bernoullicity:

Definition 3.1.11 (Relatively very weak Bernoulli)**.** *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be an ergodic dynamical system, equipped with two processes ξ and η with finite alphabets. We say that ξ is* relatively very weak Bernoulli over η *if, for every $\varepsilon > 0$, there exists $\ell \geq 1$ such that for every $m \geq 1$ and for all $k \geq 1$ large enough, we have*

$$\int \bar{d}_\ell \left(\nu_\ell(\cdot \,|\, \mathbf{a}_{[-m,0[}, \mathbf{b}_{[-k,k]}), \nu_\ell(\cdot \,|\, \mathbf{b}_{[-k,k]})\right) d\nu(\mathbf{a}, \mathbf{b}) \leq \varepsilon,$$

where ν is the law of (ξ, η) and, for $I, J \subset \mathbb{Z}$, $\nu_\ell(\cdot \,|\, \mathbf{a}_I, \mathbf{b}_J)$ is the conditional law of $\xi_{[0,\ell[}$ given that ξ_I equals $\mathbf{a}_I$ and that η_J equals $\mathbf{b}_J$.

If $\mathscr{A} = \sigma(\xi)$ and $\mathscr{B} = \sigma(\eta)$, we say that $\mathbf{X}$ (or $\mathscr{A}$) is relatively very weak Bernoulli over $\mathscr{B}$.

Many early results from Thouvenot's theory were stated for relatively finitely determined systems. But, we have the following:

Theorem 3.1.12 (see [**?**])**.** *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be an ergodic system and ξ and η be processes with finite alphabets defined on $\mathbf{X}$. Then ξ is relatively very weak Bernoulli over η if and only if it is relatively finitely determined over η.*

Although we have not explicitly defined relative finite determination here, this result is useful since it enables us to apply to relatively very weak Bernoulli processes results originally stated for relatively finitely determined processes. We give a summary of the results we will use:

Lemma 3.1.13. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a finite entropy dynamical system and $\mathscr{B}$ a factor σ-algebra. Let ξ and η be processes with finite alphabets defined on $\mathbf{X}$ such that $\mathscr{A} = \sigma(\xi)$ and $\mathscr{B} = \sigma(\eta)$. If ξ is relatively very weak Bernoulli over η, then*

(i) $\mathbf{X}$ is relatively Bernoulli over $\mathscr{B}$,

(ii) any process ρ on $\mathbf{X}$ is relatively very weak Bernoulli over η,

(iii) for any factor σ-algebra $\mathscr{C} \subset \mathscr{A}$, $\mathscr{B} \vee \mathscr{C}$ is relatively very weak Bernoulli over $\mathscr{B}$,

(iv) any factor σ-algebra $\mathscr{C} \subset \mathscr{A}$ that is independent from *$\mathscr{B}$ is Bernoulli.*

Proof. We prove the lemma mainly by referring to the literature. The statement (i) follows from [**?**, Proposition 5] and Theorem **??**. Then (ii) follows from [**?**, Proposition 4] and Theorem **??**, and (iii) follows from (ii). Let us prove (iv): take ρ a process on $\mathbf{X}$ such that $\mathscr{C} = \sigma(\rho)$ mod μ. From (ii), we know that ρ is relatively very weak Bernoulli over η. However, since $\mathscr{C}$ is independent of $\mathscr{B}$, ρ is independent of η. One can then notice that if we add this independence in the definition of relative very weak Bernoullicity, we end up with the fact that ρ is very weak Bernoulli. Finally, Theorem **??** tells us that $\mathscr{C} = \sigma(\rho)$ is Bernoulli. $\square$

We have just given many definitions and results concerning processes with finite alphabets, and the σ-algebras they generate. The following result from Krieger tells that it is applicable on any finite entropy system:

Theorem 3.1.14 (See [**?**]). *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be an ergodic dynamical system and $\mathscr{B} \subset \mathscr{A}$ be a factor σ-algebra. If $h_\mu(\mathscr{B}, T) < \infty$, there exists a finite alphabet A and a random variable $\xi_0 : X \to A$ such that*

$$\mathscr{B} = \sigma(\{\xi_0 \circ T^n\}_{n\in\mathbb{Z}}) \text{ mod } \mu.$$

We say that ξ is a finite generator of $\mathscr{B}$.

3.1.2 Positive entropy systems and weak Pinsker filtrations

In 2018, Austin proved the following:

Theorem 3.1.15 (Austin, 2018, [**?**]). *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be an ergodic dynamical system. For every $\varepsilon > 0$ there exists a factor σ-algebra $\mathscr{B}$ such that:*

- *$h_\mu(\mathscr{B}, T) \leq \varepsilon$,*
- *$\mathbf{X}$ is relatively Bernoulli over $\mathscr{B}$.*

In other words, $\mathbf{X}$ has the weak Pinsker property *(as in (**??**)).*

Definition 3.1.16. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a dynamical system and $\mathscr{F} := (\mathscr{F}_n)_{n\leq 0}$ a dynamical filtration on $\mathbf{X}$ such that $\mathscr{F}_0 = \mathscr{A}$. We say that $\mathscr{F}$ is a* weak Pinsker filtration *if*

- *for every $n \leq -1$, $\mathscr{F}_{n+1}$ is relatively Bernoulli over $\mathscr{F}_n$,*

• *and*

$$\lim_{n\to-\infty} h(\mathscr{F}_n, T) = 0.$$

Then, by iterating Austin's theorem, we see that we can obtain weak Pinsker filtrations on any ergodic system:

Proposition 3.1.17. *Let* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *be a dynamical system. If* $\mathbf{X}$ *is ergodic, there exists a weak Pinsker filtration on* $\mathbf{X}$*. More specifically, for every increasing sequence* $(\varepsilon_n)_{n\leq -1}$ *such that* $\lim_{n\to-\infty} \varepsilon_n = 0$*, there exists a weak Pinsker filtration* $(\mathscr{F}_n)_{n\leq 0}$ *such that* $\forall n \leq -1$*,* $h(\mathscr{F}_n, T) = \varepsilon_n$*.*

This simply tells us that weak Pinsker filtrations exist, but gives no explicit description. To start understanding those filtrations better, we can first link them to the Pinsker factor of the system:

Proposition 3.1.18. *Let* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *be a dynamical system and* $\mathscr{F} := (\mathscr{F}_n)_{n\leq 0}$ *a weak Pinsker filtration on* $\mathbf{X}$*. Then the tail* σ*-algebra* $\mathscr{F}_{-\infty} := \bigcap_{n\leq 0} \mathscr{F}_n$ *is the Pinsker factor of* $\mathbf{X}$*.*

Proof. Let $\mathscr{F} := (\mathscr{F}_n)_{n\leq 0}$ be a weak Pinsker filtration on $\mathbf{X}$. Since, for $n_0 \leq 0$, $\mathscr{F}_{-\infty} \subset \mathscr{F}_{n_0}$, it follows that $h(\mathscr{F}_{-\infty}, T) \leq h(\mathscr{F}_{n_0}, T)$. Then, by taking $n_0 \to -\infty$, this yields $h(\mathscr{F}_{-\infty}, T) = 0$. Therefore, $\mathscr{F}_{-\infty} \subset \Pi_{\mathbf{X}}$.

Conversely, let us show that, for every $n \leq 0$, $\Pi_{\mathbf{X}} \subset \mathscr{F}_n$. Since $\mathscr{F}$ is a weak Pinsker filtration, we can choose $\mathscr{B}_n \subset \mathscr{A}$ a Bernoulli factor σ-algebra such that

$$\mathscr{F}_n \perp\!\!\!\perp \mathscr{B}_n \text{ and } \mathscr{F}_n \vee \mathscr{B}_n = \mathscr{A} \bmod \mu.$$

Then we use Lemma **??**:

$$\Pi_{\mathbf{X}} = \Pi_{\mathscr{A}} = \Pi_{\mathscr{F}_n} \vee \Pi_{\mathscr{B}_n} = \Pi_{\mathscr{F}_n} \subset \mathscr{F}_n,$$

because, $\mathscr{B}_n$ being Bernoulli, its Pinsker factor is trivial. □

Weak Pinsker filtrations are dynamical filtrations, and in Chapter **??**, we introduced tools to classify dynamical filtrations, which we use here. For precise definitions, see Section **??**. While trying to connect the properties of a weak Pinsker filtration with the properties of the underlying system, we get the following simple results:

Theorem 3.1.19. *Let* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *be a dynamical system and* $\mathscr{F} := (\mathscr{F}_n)_{n\leq 0}$ *be a weak Pinsker filtration on* $\mathbf{X}$*. Then* *(i)* $\mathbf{X}$ *is a K-system if and only if* $\mathscr{F}$ *is kolmogorovian, i.e.* $\bigcap_{n\leq 0}\mathscr{F}_n = \{\varnothing, X\}$ *mod* μ*.*

(ii) If the filtration $\mathscr{F}$ *is of product-type, then* $\mathbf{X}$ *is Bernoulli.*

Proof. We know that a system is K if and only if its Pinsker factor is trivial. Then the equivalence in (i) follows from Proposition **??**.

We now prove (ii). Assume that $\mathscr{F}$ is a weak Pinsker filtration of product type. This means that there exists a sequence $(\mathscr{B}_n)_{n\leq 0}$ of mutually independent factor σ-algebra such that $\mathscr{F}_n = \bigvee_{k\leq n}\mathscr{B}_k$. Let $n \leq 0$. We know that $\mathscr{F}_n$ is relatively Bernoulli over $\mathscr{F}_{n-1}$ and that $\mathscr{B}_n$ is independent of $\mathscr{F}_{n-1}$. So, Lemma **??** tells us that $\mathscr{B}_n$ is Bernoulli. Therefore, we have $\mathscr{A} = \mathscr{F}_0 = \bigvee_{k\leq 0}\mathscr{B}_k$, which shows that we can write $\mathscr{A}$ as a product of mutually independent Bernoulli factors. Hence, $\mathscr{A}$ is Bernoulli. □

However, this result leaves many open questions. First, we can ask if the converse of (ii) is true, because we remark at the end of Section **??** that, on a Bernoulli shift, there is at least one weak Pinsker filtration of product type. Therefore the converse of (ii) is equivalent to the uniqueness problem given in Question **??**. Another area that is left open is to consider other properties from the theory of dynamical filtrations, like standardness or I-cosiness, and wonder what it implies of the system if a weak Pinsker filtrations has those properties:

Question 3.1.20. *What can we say about* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *if there is a weak Pinsker filtration* $\mathscr{F}$ *on* $\mathbf{X}$ *that is standard ? In that case, is* $\mathbf{X}$ *Bernoulli ? And if the weak Pinsker filtration is I-cosy ?*

Our hope is that answering those questions could give additional information on the structure of non-Bernoulli K-systems.

3.1.3 The uniqueness problem

Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be an ergodic dynamical system. As mentioned in Proposition **??**, the fact that every ergodic systems satisfies the weak Pinsker property implies that, for any given increasing sequence $(\varepsilon_n)_{n\leq -1}$ that goes to 0 in $-\infty$ such that $\varepsilon_{-1} \leq h(\mathbf{X})$, there exits a weak Pinsker filtration $\mathscr{F}$ on $\mathbf{X}$ such that $h(\mathscr{F}_n, T) = \varepsilon_n$. But this filtration is not unique, because in the splitting result given by the weak Pinsker property (**??**), the choice of the factor σ-algebra generated by $\mathbf{X}_\varepsilon$ is not unique. For example, take a system of the form

$$\mathbf{X} := \mathbf{Z} \otimes \mathbf{B}_1 \otimes \mathbf{B}_2,$$

where $\mathbf{Z}$ is a 0 entropy system and $\mathbf{B}_1$ and $\mathbf{B}_2$ are Bernoulli shifts of equal entropy. Note that $\mathbf{Z} \otimes \mathbf{B}_1$ and $\mathbf{Z} \otimes \mathbf{B}_2$ generate two different factor σ-algebras on $\mathbf{X}$. But they are both factors over which $\mathbf{X}$ is relatively Bernoulli, and they have the same entropy. However, we can notice in this example that $\mathbf{Z} \otimes \mathbf{B}_1$ and $\mathbf{Z} \otimes \mathbf{B}_2$ are isomorphic. This observation hints to a general result:

Theorem 3.1.21 (From Thouvenot in [**?**]). *Let* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *and* $\mathbf{Y} := (Y, \mathscr{B}, \nu, S)$ *be ergodic dynamical systems and* $\mathbf{B}$ *be a Bernoulli shift of finite entropy. If* $\mathbf{X} \otimes \mathbf{B}$ *and* $\mathbf{Y} \otimes \mathbf{B}$ *are isomorphic, then* $\mathbf{X}$ *and* $\mathbf{Y}$ *are isomorphic.*

Proof. This proof relies on the weak Pinsker property of $\mathbf{X}$ and $\mathbf{Y}$, and Lemma **??**. We also use many times that Bernoulli shifts with the same entropy are isomorphic.

Since $\mathbf{X} \otimes \mathbf{B}$ and $\mathbf{Y} \otimes \mathbf{B}$ are isomorphic, we have:

$$h(\mathbf{X}) = h(\mathbf{X} \otimes \mathbf{B}) - h(\mathbf{B}) = h(\mathbf{Y} \otimes \mathbf{B}) - h(\mathbf{B}) = h(\mathbf{Y}).$$

Set $a := h(\mathbf{X}) = h(\mathbf{Y})$. We can then apply the weak Pinsker property of $\mathbf{X}$ and $\mathbf{Y}$ to find two systems $\hat{\mathbf{X}}$, $\hat{\mathbf{Y}}$ and a Bernoulli shift $\hat{\mathbf{B}}$ such that

$$h(\hat{\mathbf{X}}) = h(\hat{\mathbf{Y}}) \leq a/3,$$

and

$$\mathbf{X} \cong \hat{\mathbf{X}} \otimes \hat{\mathbf{B}} \ \text{ and } \ \mathbf{Y} \cong \hat{\mathbf{Y}} \otimes \hat{\mathbf{B}}.$$

This implies

$$\hat{\mathbf{X}} \otimes (\hat{\mathbf{B}} \otimes \mathbf{B}) \cong \hat{\mathbf{Y}} \otimes (\hat{\mathbf{B}} \otimes \mathbf{B}).$$

In other words, there is a system $\mathbf{Z}$ and two factor maps $p_{\hat{\mathbf{X}}} : \mathbf{Z} \longrightarrow \hat{\mathbf{X}}$ and $p_{\hat{\mathbf{Y}}} : \mathbf{Z} \longrightarrow \hat{\mathbf{Y}}$ such that $\mathbf{Z}$ is relatively Bernoulli over $p_{\hat{\mathbf{X}}}$ and relatively Bernoulli over $p_{\hat{\mathbf{Y}}}$. But then, Lemma **??** tells us that the factor σ-algebra $\sigma(p_{\hat{\mathbf{X}}} \vee p_{\hat{\mathbf{Y}}})$ is relatively very weak Bernoulli over $p_{\hat{\mathbf{X}}}$ and relatively very weak Bernoulli over $p_{\hat{\mathbf{Y}}}$. Therefore, there exist a Bernoulli shift $\tilde{\mathbf{B}}$ and two factor maps $\varphi_1 : \mathbf{Z} \longrightarrow \tilde{\mathbf{B}}$ and $\varphi_2 : \mathbf{Z} \longrightarrow \tilde{\mathbf{B}}$ such that $\varphi_1 \perp\!\!\!\perp p_{\hat{\mathbf{X}}}$, $\varphi_2 \perp\!\!\!\perp p_{\hat{\mathbf{Y}}}$ and

$$\sigma(p_{\hat{\mathbf{X}}} \vee \varphi_1) = \sigma(p_{\hat{\mathbf{X}}} \vee p_{\hat{\mathbf{Y}}}) = \sigma(p_{\hat{\mathbf{Y}}} \vee \varphi_2).$$

This implies that

$$\hat{\mathbf{X}} \otimes \tilde{\mathbf{B}} \cong \hat{\mathbf{Y}} \otimes \tilde{\mathbf{B}}.$$

But, since we chose to have $h(\hat{\mathbf{X}}) = h(\hat{\mathbf{Y}}) \leq a/3$, we get

$$h(\tilde{\mathbf{B}}) \leq h(p_{\hat{\mathbf{X}}} \vee p_{\hat{\mathbf{Y}}}) \leq h(\hat{\mathbf{X}}) + h(\hat{\mathbf{Y}}) \leq 2a/3 \leq h(\hat{\mathbf{B}}).$$

Given a last Bernoulli shift $\overline{\mathbf{B}}$ of entropy $h(\hat{\mathbf{B}}) - h(\tilde{\mathbf{B}})$ we get $\hat{\mathbf{B}} \cong \tilde{\mathbf{B}} \otimes \overline{\mathbf{B}}$ and

$$\mathbf{X} \cong \hat{\mathbf{X}} \otimes \hat{\mathbf{B}} \cong \hat{\mathbf{X}} \otimes \tilde{\mathbf{B}} \otimes \overline{\mathbf{B}} \cong \hat{\mathbf{Y}} \otimes \tilde{\mathbf{B}} \otimes \overline{\mathbf{B}} \cong \hat{\mathbf{Y}} \otimes \hat{\mathbf{B}} \cong \mathbf{Y}.$$

□

As a consequence of this result, we see that if $\mathscr{F} := (\mathscr{F}_n)_{n\leq 0}$ and $\mathscr{G} := (\mathscr{G}_n)_{n\leq 0}$ are two weak Pinsker filtrations on $\mathbf{X}$ such that, for all $n \leq 0$, $h(\mathscr{F}_n, T) = h(\mathscr{G}_n, T)$, then we must have that, for each $n \leq 0$, $\mathbf{X}/_{\mathscr{F}_n} \cong \mathbf{X}/_{\mathscr{G}_n}$.

However, this only gives "local isomorphisms", and it does not necessarily mean that the filtrations $\mathscr{F}$ and $\mathscr{G}$ are isomorphic (according to the notion of isomorphism introduced in Definition **??**). Therefore, the following is still an open question:

Question 3.1.22. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be an ergodic dynamical system. Are all weak Pinsker filtrations on $\mathbf{X}$ with the same entropy isomorphic ?*

This question is what we call the *uniqueness problem*.

If $\mathbf{X}$ is a Bernoulli shift, and if we take a sequence $(\varepsilon_n)_{n\leq 0}$ such that $\varepsilon_0 = h(\mathbf{X})$, we can take Bernoulli shifts $(\mathbf{B}_n)_{n\leq 0}$ such that $h(\mathbf{B}_n) = \varepsilon_n - \varepsilon_{n-1}$, and define the system

$$\mathbf{B} := \bigotimes_{n\leq 0} \mathbf{B}_n.$$

It is a Bernoulli shift of entropy $\varepsilon_0 = h(\mathbf{X})$, so it is isomorphic to $\mathbf{X}$. Through this isomorphism, the factors of the form $\bigotimes_{k\leq n} \mathbf{B}_k$ generate a product type weak Pinsker filtration on $\mathbf{X}$. Therefore, in the case where $\mathbf{X}$ is a Bernoulli shift, the uniqueness problem becomes:

Question 3.1.23. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a Bernoulli shift. Are all weak Pinsker filtrations on $\mathbf{X}$ of product type ?*

3.2 Uniqueness problem on Bernoulli systems

In this section, we present our efforts to tackle Question **??**. The ideas developed here come from discussions with Jean-Paul Thouvenot, and we thank him for those insights. Specifically, we are going to show:

Theorem 3.2.1. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a Bernoulli system and let $\mathscr{F} := (\mathscr{F}_n)_{n\leq 0}$ be a weak Pinsker filtration. There exists some sub-sequence $(\mathscr{F}_{n_k})_{k\leq 0}$ which is a weak Pinsker filtration of product type.*

The fact that we are only able to describe the structure of a sub-sequence of $\mathscr{F}$, for now, seems to be significant. Indeed, we can compare that result to a well known result from Vershik about static filtrations on a probability space: any filtration whose tail σ-algebra $\bigcap_{n\leq 0}\mathscr{F}_n$ is trivial has a sub-sequence that is standard (see [**?**, Theorem 3]). However there are many examples of non-standard filtrations with trivial tail σ-algebra. Therefore, although the context of Vershik's result is very different, it emphasizes that Theorem **??** does not give a complete answer to Question **??**.

The main step in proving Theorem **??** is contained in the following proposition:

Proposition 3.2.2. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a Bernoulli system of finite entropy and $\mathcal{P}_0 : X \to A$ a finite generator of $\mathscr{A}$, i.e. a finite valued random variable such that $\mathscr{A} = \sigma(\{\mathcal{P}_0 \circ T^n\}_{n\in\mathbb{Z}})$. Let $\mathscr{H} \subset \mathscr{A}$ be a factor σ-algebra such that $\mathbf{X}$ is relatively Bernoulli over $\mathscr{H}$. For every $\varepsilon > 0$, there exists $\delta > 0$ such that, if $h_\mu(\mathscr{H}) \leq \delta$, there is a Bernoulli factor σ-algebra $\mathscr{B}$ such that*

(i) $\mathscr{B} \perp\!\!\!\perp \mathscr{H}$,

(ii) $\mathscr{A} = \mathscr{H} \vee \mathscr{B}$ *mod* μ,

(iii) and $\mathcal{P}_0 \preceq_\varepsilon \mathscr{B}$.

In this proposition, Krieger' theorem (Theorem **??**) ensures the existence of a finite generator $\mathcal{P}$ since $\mathbf{X}$ has finite entropy. The notation "$\mathcal{P}_0 \preceq_\varepsilon \mathscr{B}$", which we use many times below, means that there exists a $\mathscr{B}$-measurable random variable $\mathcal{Q}_0$ such that $\mu(\mathcal{P}_0 \neq \mathcal{Q}_0) \leq \varepsilon$.

The existence of a Bernoulli factor satisfying (i) and (ii) is simply the definition of relative Bernoullicity, the important part of this proposition is the ability to build a Bernoulli complement that satisfies (iii). Then iterating this result will yield Theorem **??** (see Section **??**).

3.2.1 The technical lemma

In this section, we tackle the main technical and constructive part of the proof of Proposition **??**. It is contained in Lemma **??**.

In Section **??**, we introduced the notion of very weak Bernoullicity, which gives a characterization of Bernoulli systems. Here, we use another equivalent notion: extremality, due to Thouvenot.

Definition 3.2.3 (See [**?**]). *Let* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *be an ergodic dynamical system and* $\xi := (\xi_0 \circ T^n)_{n\in\mathbb{Z}}$ *be a process where* ξ_0 *takes values in some finite alphabet* A*. We say that* ξ *is* extremal *if, for every* $\varepsilon > 0$*, there exist* $\delta > 0$ *and* $N \in \mathbb{N}$*, such that for every* $\ell \geq N$ *and every random variable* $\mathcal{Q} : X \to B$ *with* $\#B \leq 2^{\delta\ell}$*, there is a set* $B_0 \subset B$ *such that* $\mu(\mathcal{Q} \in B_0) \geq 1 - \varepsilon$ *and for* $b \in B_0$*, we have:*

$$\bar{d}_\ell(\nu_\ell(\cdot \,|\, b), \nu_\ell(\cdot)) \leq \varepsilon,$$

where ν_ℓ *is the law of* $\xi_{[0,\ell[}$ *and* $\nu_\ell(\cdot \,|\, b)$ *is the law of* $\xi_{[0,\ell[}$ *given that* $\mathcal{Q}$ *equals* b*.*

In [**?**, Theorem 6.4], it is shown that extremality is equivalent to very weak Bernoullicity (and hence to Bernoullicity). In particular, we will use the fact that any process defined on a Bernoulli system is extremal.

The proof of Lemma **??** uses many methods that are usual in Ornstein's theory of Bernoulli shifts (a presentation can be found in [**?**] or [**?**]). Therefore, we need to introduce some commonly used notions and results from that theory. The following combinatorial result is frequently used in Ornstein's theory:

Lemma 3.2.4 (Hall's marriage lemma [**?**]). *Let* E *and* F *be finite sets, and* $\{J_e\}_{e\in E}$ *be a family of subsets of* F*:* $\forall e \in E, J_e \subset F$*. There exists an injective map* $\psi : E \to F$ *such that* $\forall e \in E, \psi(e) \in J_e$ *if, and only if for every* $I \subset E$*, we have*

$$\#I \leq \#\bigcup_{e\in I} J_e.$$

The main way in which the entropy of the processes is used in our arguments comes from the Shannon-McMillan-Breiman Theorem (see [**?**, Theorem 13.1]):

Theorem 3.2.5. *Let* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *be an ergodic dynamical system and* $\xi_0 : X \to A$*. For* $\boldsymbol{a} \in A^{[0,n[}$*, define*

$$p_n(\boldsymbol{a}) := \mu(\xi_{[0,n[} = \boldsymbol{a}).$$

We have

$$\lim_{n\to\infty} -\frac{1}{n}\log(p_n(\xi_{[0,n[})) = h_\mu(\xi, T), \text{ } \mu\text{-almost surely.}$$

In particular, we also have the convergence in probability: for every $\varepsilon > 0$*, there exists* $N \geq 1$ *such that for every* $n \geq N$*, there exists a set* $\mathcal{A}_n \subset A^{[0,n[}$ *such that* $\mu(\xi_{[0,n[} \in \mathcal{A}_n) \geq 1 - \varepsilon$ *and for every* $\boldsymbol{a} \in \mathcal{A}_n$*,*

$$2^{-(h_\mu(\xi,T)+\varepsilon)n} \leq \mu(\xi_{[0,n[} = \boldsymbol{a}) \leq 2^{-(h_\mu(\xi,T)-\varepsilon)n}.$$

We also need to introduce another tool that is commonly used in Ornstein's theory: Rokhlin towers. On a dynamical system $\mathbf{X} := (X, \mathscr{A}, \mu, T)$, to get a tower of height n, we need a set F such that the sets $T^j F$, for $0 \leq j \leq n-1$ are disjoint. Then the family $\mathcal{T} := (F, TF, ..., T^{n-1}F)$ is what we call a Rokhlin tower, or, in short, a *tower*. However, we will also refer to the set $\bigsqcup_{j=0}^{n-1} T^j F$ as a tower. In particular, many times, we will write $\mu(\mathcal{T})$ for $\mu(\bigsqcup_{j=0}^{n-1} T^j F)$. The following result guaranties that Rokhlin of arbitrary height and total measure almost 1 exist under quite general conditions:

Proposition 3.2.6 (See [**?**]). *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be an ergodic dynamical system and ξ_0 a finite valued random variable. For all $n \geq 1$ and $\varepsilon > 0$, there exists a measurable set $F \subset X$ such that the sets $T^j F$, for $j \in [0, n[$, are disjoint, $\mu(\bigcup_{j=0}^{n-1} T^j F) \geq 1 - \varepsilon$ and $\mathcal{L}(\xi_0 \,|\, F) = \mathcal{L}(\xi_0)$.*

The set F is called the base of the tower $\mathcal{T}$ and the sets $T^j F$ are the levels. For any set $E \subset F$, the family

$$C_E := \{T^j E\}_{0 \leq j \leq n-1}$$

is a tower, and we say that it is a column of $\mathcal{T}$. If $\xi_0 : X \to A$ is a random variable, we will be interested in the columns defined by sets of the form $F_{\boldsymbol{a}} := F \cap \{\xi_{[0,n[} = \boldsymbol{a}\}$ with $\boldsymbol{a} \in A^{[0,n[}$. We say that $\boldsymbol{a}$ is the ξ-name of the column $C_{\boldsymbol{a}} := C_{F_{\boldsymbol{a}}}$. The columns $\{C_{\boldsymbol{a}}\}_{\boldsymbol{a} \in A^{[0,n[}}$ give a partition of the levels of $\mathcal{T}$. Now, conversely, assume that we have a partition of F given by sets $E_1, ..., E_p$, then the columns $C_{E_1}, ..., C_{E_p}$ give a partition of the levels of $\mathcal{T}$. If, moreover, we associate to each column C_{E_i} a name $\boldsymbol{a}^{(i)} \in A^{[0,n[}$ of length n, we can define a random variable ξ_0 on the levels of $\mathcal{T}$ so that, for every i, we have $C_{E_i} = C_{\boldsymbol{a}^{(i)}}$. We obtain this random variable simply by setting, for $i \in [\![1, p]\!], j \in [\![0, n[\![$

$$\xi_0 = \boldsymbol{a}_j^{(i)} \text{ on } T^j E_i.$$

This is the framework we will use to construct our random variables. We are now ready to turn our attention to the following:

Lemma 3.2.7. *Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a Bernoulli system of finite entropy and $\mathcal{P}_0 : X \to A$ a finite valued random variable such that $\mathscr{A} = \sigma(\{\mathcal{P}_0 \circ T^n\}_{n \in \mathbb{Z}})$ mod μ. Let $\mathcal{H}_0 : X \to H$ be a finite valued random variable such that $\mathbf{X}$ is relatively Bernoulli over $\sigma(\mathcal{H}) := \sigma(\{\mathcal{H}_0 \circ T^n\}_{n \in \mathbb{Z}})$, i.e. we can take a finite alphabet B and a B-valued i.i.d. process $\xi := (\xi_0 \circ T^n)_{n \in \mathbb{Z}}$ independent from $\mathcal{H}$ such that $\mathscr{A} = \sigma(\mathcal{H}) \vee \sigma(\xi)$ mod μ. For every $\varepsilon > 0$, there exists $\delta > 0$ such that, if $h_\mu(\mathcal{H}, T) \leq \delta$, for any $\alpha > 0$, there exists a process $\tilde{\xi}$ such that* *(i)* $\bar{d}_1(\mathcal{L}(\xi_0), \mathcal{L}(\tilde{\xi}_0)) \leq \alpha$,

(ii) $0 \leq h_\mu(\mathcal{H} \vee \xi, T) - h_\mu(\mathcal{H} \vee \tilde{\xi}, T) \leq \alpha$,

(iii) and $\mathcal{P}_0 \preceq_\varepsilon \sigma(\tilde{\xi})$.

The proof of the lemma being quite intricate, we start by giving a sketch of the proof. First, we will need a Rokhlin tower $\mathcal{T}_n$ of very large height n. This tower is then divided into the columns $C_{\boldsymbol{h}}$ (see (**??**)) generated by $\mathcal{H}$. Each of those columns is then divided into sub-columns $C_{\boldsymbol{h}}^{\boldsymbol{b}}$ (see (**??**)) generated by ξ. Because $\mathcal{H} \vee \xi$ generate $\mathscr{A}$, we can approach $\mathcal{P}_0$ by some random variable $\tilde{\mathcal{P}}_0$ depending on finitely many coordinates of $\mathcal{H} \vee \xi$. It enables us to associate to each $C_{\boldsymbol{h}}^{\boldsymbol{b}}$ a word $\tilde{\mathcal{P}}_{[0,n[}(\boldsymbol{h}, \boldsymbol{b})$ which gives a good approximation of $\mathcal{P}_0$ on the levels of $C_{\boldsymbol{h}}^{\boldsymbol{b}}$. We will define $\tilde{\xi}_0$ by giving $C_{\boldsymbol{h}}^{\boldsymbol{b}}$ a new $\tilde{\xi}$-name, to replace $\boldsymbol{b}$. Our goal is to choose those names so that we can get a good approximation of $\tilde{\mathcal{P}}_{[0,n[}(\boldsymbol{h}, \boldsymbol{b})$ by simply knowing the $\tilde{\xi}$-name of $C_{\boldsymbol{h}}^{\boldsymbol{b}}$, regardless of $\boldsymbol{h}$. To do that, we fix a column $C_{\boldsymbol{h}_0}$ and use it as a "model" for the other columns. Then the extremality of $\mathcal{P}$ comes into play: it tells us, for most choices of $\boldsymbol{h}$, the families $\{\tilde{\mathcal{P}}_{[0,n[}(\boldsymbol{h}_0, \boldsymbol{b})\}_{\boldsymbol{b} \in \mathcal{B}_n}$ and $\{\tilde{\mathcal{P}}_{[0,n[}(\boldsymbol{h}, \boldsymbol{b})\}_{\boldsymbol{b} \in \mathcal{B}_n}$ are quite similar. More specifically, we show that, for most $\boldsymbol{b}$, there are names $\tilde{\boldsymbol{b}}$ such that $d_n(\tilde{\mathcal{P}}_{[0,n[}(\boldsymbol{h}_0, \tilde{\boldsymbol{b}}), \tilde{\mathcal{P}}_{[0,n[}(\boldsymbol{h}, \boldsymbol{b}))$ is small. Those names are then suitable $\tilde{\xi}$-names for $C_{\boldsymbol{h}}^{\boldsymbol{b}}$. However, when we choose among those suitable names, we need to make sure that we are not giving the same name to too many columns, otherwise we might loose to much information, and we could not get (ii). This is done using Hall's marriage lemma.

Proof of Lemma **??**. In this proof, we will use many parameters, that will depend on each other. We start with a presentation of those parameters, and the order in which they are chosen.

- Let $\varepsilon > 0$. This parameter is chosen first, as it appears in the statement of the lemma. Then we choose $\delta > 0$ and $N \geq 1$, depending on ε via the extremality of $\mathcal{P}$, and require that $h_\mu(\mathcal{H}, T) < \delta$.

- Let $\alpha > 0$. This is another arbitrarily small parameter that appears in the statement of the lemma. It does not depend on ε nor δ.

- Next, we introduce $0 < \gamma < 1$, which needs to be small with respect to α and ε for (ii) and (iii) to hold.

- Then we take $\beta > 0$, which will be our most used parameter. It depends on all previous parameters, and throughout the proof, we will give many instances where it needs to be chosen small enough with respect to those parameters. With β fixed, we take $n_0 \geq 1$ to get $\mathcal{P}_0 \preceq_{\beta^2} (\mathcal{H} \vee \xi)_{[-n_0,n_0]}$.

- Finally, we choose an integer n, which will be the height of the Rokhlin tower. It is chosen larger than N. We also need it to be large enough for us to apply the Shannon-McMillan-Breiman theorem, as well as Birkhoff's ergodic theorem. As n will appear in several estimates, our choice of n also depends on ε, δ, γ, β and n_0.

Step 1: The setup of the tower

We apply the Shannon-McMillan-Breiman theorem (i.e. Theorem **??**) to know that, since n is large enough, there exist two sets $\mathcal{E}_n \subset H^{[0,n[}$ and $\mathcal{B}_n \subset B^{[0,n[}$ such that

$$\mu(\mathcal{H}_{[0,n[} \in \mathcal{E}_n) \geq 1 - \beta, \text{ and } \mu(\xi_{[0,n[} \in \mathcal{B}_n) \geq 1 - \beta, \tag{3.3}$$

and we have the following estimates:

$$\forall \boldsymbol{h} \in \mathcal{E}_n,\ 2^{-(h_\mu(\mathcal{H},T)+\beta)n} \leq \mu(\mathcal{H}_{[0,n[} = \boldsymbol{h}) \leq 2^{-(h_\mu(\mathcal{H},T)-\beta)n}, \tag{3.4}$$

$$\forall \boldsymbol{b} \in \mathcal{B}_n,\ 2^{-(h_\mu(\xi,T)+\beta)n} \leq \mu(\xi_{[0,n[} = \boldsymbol{b}) \leq 2^{-(h_\mu(\xi,T)-\beta)n}. \tag{3.5}$$

Throughout the remainder of the proof, we will want the elements of $\mathcal{B}_n$ and $\mathcal{E}_n$ to have some additional properties, and to get that we will need to restrict ourselves to subsets of $\mathcal{B}_n$ and $\mathcal{E}_n$. However, the measure of the complement of those subsets will always be controlled by a multiple of β, so, to simplify notation, we will still call those sets $\mathcal{B}_n$ and $\mathcal{E}_n$ and assume that (**??**) still holds.

Birkhoff's ergodic theorem gives us additional information on the elements of $\mathcal{B}_n$: for any sequence $\boldsymbol{b} \in B^{[0,n[}$ and any element $\boldsymbol{b'} \in B$, denote $f_n(\boldsymbol{b}, \boldsymbol{b'})$ the frequency at which the element $\boldsymbol{b'}$ appears in the sequence $\boldsymbol{b}$. This can also be defined as follows:

$$\forall x \in \{\xi_{[0,n[} = \boldsymbol{b}\},\ f_n(\boldsymbol{b}, \boldsymbol{b'}) := \frac{1}{n} \sum_{j=0}^{n-1} \mathbb{1}_{\{\xi_0 = \boldsymbol{b'}\}}(T^j x). \tag{3.6}$$

From this definition of f_n, by applying Birkhoff's ergodic theorem, provided n is large enough, we can choose $\mathcal{B}_n$ so that for every $\boldsymbol{b} \in \mathcal{B}_n$

$$\sum_{\boldsymbol{b'} \in B} |f_n(\boldsymbol{b}, \boldsymbol{b'}) - \mu(\xi_0 = \boldsymbol{b'})| \leq \beta. \tag{3.7}$$

Since $\mathcal{H}\vee\xi$ generates $\mathscr{A}$, we can find $n_0 \geq 1$ so that $\mathcal{P}_0 \preceq_{\beta^2} (\mathcal{H}\vee\xi)_{[-n_0,n_0]}$. This means that there exists a $(\mathcal{H}\vee\xi)_{[-n_0,n_0]}$-measurable random variable $\tilde{\mathcal{P}}_0$ such that $\mu(\tilde{\mathcal{P}}_0 \neq \mathcal{P}_0) \leq \beta^2$.

By making use of Proposition **??**, we can build a set G such that $F' := TG$ is disjoint from G and F' is the base of a tower $\mathcal{G}_n := \{T^j F'\}_{0\leq j\leq n-1}$ such that $\mu(\mathcal{G}_n) \geq 1-\beta$ and

$$\mathcal{L}((\mathcal{H}\vee\tilde{\mathcal{P}}\vee\xi)_{[0,n[} \,|\, F') = \mathcal{L}((\mathcal{H}\vee\tilde{\mathcal{P}}\vee\xi)_{[0,n[}). \tag{3.8}$$

The set G will be useful later to code the entrance of the tower. We slightly reduce the tower by setting $F := F' \cap \{\mathcal{H}_{[0,n[} \in \mathcal{E}_n\} \cap \{\xi_{[0,n[} \in \mathcal{B}_n\}$ and $\mathcal{T}_n := \{T^j F\}_{0\leq j\leq n-1}$. One can then use (**??**) with our previous estimates to see that $\mu(\mathcal{T}_n) \geq 1-3\beta$.

We then split $\mathcal{T}_n$ into $\mathcal{H}$-columns: for $\boldsymbol{h} \in \mathcal{E}_n$, we define

$$C_{\boldsymbol{h}} := \{T^j(F \cap \{\mathcal{H}_{[0,n[} = \boldsymbol{h}\})\}_{0\leq j\leq n-1}, \tag{3.9}$$

so that $\mathcal{T}_n = \bigsqcup_{\boldsymbol{h}\in\mathcal{E}_n} C_{\boldsymbol{h}}$ (we mean that the levels of $\mathcal{T}_n$ are disjoint unions of the levels of $C_{\boldsymbol{h}}$). For each $\boldsymbol{h} \in \mathcal{E}_n$, we say that $C_{\boldsymbol{h}}$ is the column of $\mathcal{H}_{[0,n[}$-name $\boldsymbol{h}$. We also denote by $F_{\boldsymbol{h}} := F \cap \{\mathcal{H}_{[0,n[} = \boldsymbol{h}\}$ the base of $C_{\boldsymbol{h}}$.

Step 2: Using the extremality of $\mathcal{P}$

We plan on modifying ξ into a process $\tilde{\xi}$ so that the joint law of $\mathcal{P}\vee\tilde{\xi}$ is almost the same in most of the columns $\{C_{\boldsymbol{h}}\}_{\boldsymbol{h}\in\mathcal{E}_n}$. We start by using the fact that $\mathbf{X}$ is Bernoulli to see that the law of $\mathcal{P}$ is almost the same on each column $C_{\boldsymbol{h}}$. Indeed, since $\mathbf{X}$ is Bernoulli, $\mathcal{P}$ is extremal, which means we can fix $\delta > 0$ and $N \geq 1$ as the numbers associated to $\varepsilon^3/4$ in the definition of extremality and assume that $h_\mu(\mathscr{H}, T) < \delta$. On the other hand, from (**??**), we deduce that

$$\#\mathcal{E}_n \leq 2^{(h_\mu(\mathcal{H},T)+\beta)n}.$$

Next we define the partition

$$\mathcal{Q} := \begin{cases} * & \text{on } \{\mathcal{H}_{[0,n[} \notin \mathcal{E}_n\} \cup \{\xi_{[0,n[} \notin \mathcal{B}_n\} \\ \mathcal{H}_{[0,n[} & \text{on } \{\mathcal{H}_{[0,n[} \in \mathcal{E}_n\} \cap \{\xi_{[0,n[} \in \mathcal{B}_n\} \end{cases}.$$

In particular, we know that $\mu(\mathcal{Q} = *) \leq 2\beta$. Moreover, the number of values taken by $\mathcal{Q}$ is bounded by

$$\#\mathcal{E}_n + 1 \leq 2^{(h_\mu(\mathcal{H},T)+\beta)n} + 1 \leq 2^{n\delta},$$

for β small enough. Therefore the extremality of $\mathcal{P}$ tells us that, since $n \geq N$, there exists a subset $\bar{\mathcal{E}}_n \subset \mathcal{E}_n$ such that

$$\mu(\mathcal{Q} \notin (\bar{\mathcal{E}}_n \cup \{*\})) \leq \varepsilon^3/4 + 2\beta \leq \varepsilon^3 \leq \varepsilon, \tag{3.10}$$

and for $\boldsymbol{h} \in \bar{\mathcal{E}}_n$, we have

$$\bar{d}_n(\mathcal{L}(\mathcal{P}_{[0,n[} \,|\, \mathcal{Q} = \boldsymbol{h}), \mathcal{L}(\mathcal{P}_{[0,n[})) \leq \varepsilon^3/4.$$

Moreover, since $\mu(\mathcal{P}_0 \neq \tilde{\mathcal{P}}_0) \leq \beta^2$, up to making $\mathcal{E}_n$ slightly smaller, we can assume that for every $\boldsymbol{h} \in \mathcal{E}_n$, we have

$$\mu(\tilde{\mathcal{P}}_0 \neq \mathcal{P}_0 \,|\, \mathcal{Q} = \boldsymbol{h}) \leq \beta. \tag{3.11}$$

Therefore, for β small enough with respect to ε^3, this yields:

$$\bar{d}_n(\mathcal{L}(\tilde{\mathcal{P}}_{[0,n[} \,|\, \mathcal{Q} = \boldsymbol{h}), \mathcal{L}(\tilde{\mathcal{P}}_{[0,n[})) \leq \varepsilon^3/3. \tag{3.12}$$

Step 3: Framework for the construction of $\tilde{\xi}_0$

We start the construction of $\tilde{\xi}$ by setting $\tilde{\xi}_0 := *$ on $G = T^{-1}F'$, where $*$ represents a symbol that does not belong to B. Later in the proof, this will allow us to detect the entrance into $\mathcal{T}_n$ from the value of the process $\tilde{\xi}$. Then define $\tilde{\xi}_0$ to take any value in B on the rest of $\mathcal{T}_n^c$. For $\boldsymbol{h} \in \mathcal{E}_n \backslash \bar{\mathcal{E}}_n$, on $C_{\boldsymbol{h}}$, we set $\tilde{\xi}_0 := \xi_0$. We are left with defining our new random variable $\tilde{\xi}_0$ on the columns $C_{\boldsymbol{h}}$, with $\boldsymbol{h} \in \bar{\mathcal{E}}_n$. We start by fixing $\boldsymbol{h}_0 \in \bar{\mathcal{E}}_n$, and the column $C_{\boldsymbol{h}_0}$ will serve as a "model" for the other columns.

Next we fix an $\boldsymbol{h} \in \bar{\mathcal{E}}_n$. We define sub-columns of $C_{\boldsymbol{h}}$: for $\boldsymbol{b} \in \mathcal{B}_n$,

$$C_{\boldsymbol{h}}^{\boldsymbol{b}} := \{T^j(F \cap \{\mathcal{H}_{[0,n[} = \boldsymbol{h}\} \cap \{\xi_{[0,n[} = \boldsymbol{b}\})\}_{0 \leq j \leq n-1}. \tag{3.13}$$

We say that $\boldsymbol{b}$ the ξ-name of $C_{\boldsymbol{h}}^{\boldsymbol{b}}$. Because of our definition of F and (**??**), the set $\mathcal{B}_n$ gives us exactly the ξ-names of all the sub-columns in $C_{\boldsymbol{h}}$. We will then give each sub-column $C_{\boldsymbol{h}}^{\boldsymbol{b}}$ a new word $\tilde{\boldsymbol{b}} \in \mathcal{B}_n$ and define the random variable $\tilde{\xi}_0$ on $C_{\boldsymbol{h}}^{\boldsymbol{b}}$ as the only variable such that $\tilde{\boldsymbol{b}}$ is the $\tilde{\xi}$-name of $C_{\boldsymbol{h}}^{\boldsymbol{b}}$. This means that to conclude the construction of $\tilde{\xi}_0$ on $C_{\boldsymbol{h}}$, we simply need to build a map $\varphi_{\boldsymbol{h}} : \mathcal{B}_n \longrightarrow \mathcal{B}_n$, and the properties we will obtain on $\tilde{\xi}$ will follow from our choice for $\varphi_{\boldsymbol{h}}$.

In order to give us some additional leeway, we use the parameter $\gamma > 0$ introduced at the start of the proof: we define $n_1 := \lfloor (1-\gamma)n \rfloor \leq n$, and for $\boldsymbol{b} \in \mathcal{B}_n$, we denote by $\boldsymbol{b}_{n_1} := \boldsymbol{b}_{[0,n_1[} \in B^{[0,n_1[}$ the truncated sub-sequence of $\boldsymbol{b}$ of length n_1. Conversely, for $\bar{\boldsymbol{b}} \in B^{[0,n_1[}$, define

$$B(\bar{\boldsymbol{b}}) := \{\boldsymbol{b} \in \mathcal{B}_n \,|\, \boldsymbol{b}_{n_1} = \bar{\boldsymbol{b}}\},$$

and

$$\mathcal{B}_{n_1} := \{\bar{\boldsymbol{b}} \in B^{[0,n_1[} \,|\, B(\bar{\boldsymbol{b}}) \neq \varnothing\}.$$

We will obtain the map $\varphi_{\boldsymbol{h}}$ by building an injective map $\psi_{\boldsymbol{h}} : \mathcal{B}_{n_1} \longrightarrow \mathcal{B}_n$ and setting $\varphi_{\boldsymbol{h}}(\boldsymbol{b}) := \psi_{\boldsymbol{h}}(\boldsymbol{b}_{n_1})$. We start by noting that, in our application of the Shannon-McMillan-Breiman Theorem, provided n is large enough, we can assume that the estimate (**??**) still holds when replacing n by n_1. More precisely, we mean that, up to making $\mathcal{B}_n$ smaller, we can assume that we have the estimate: for $\bar{\boldsymbol{b}} \in \mathcal{B}_{n_1}$

$$2^{-(h_\mu(\xi,T)+\beta)n_1} \leq \mu(\xi_{[0,n_1[} = \bar{\boldsymbol{b}}) \leq 2^{-(h_\mu(\xi,T)-\beta)n_1}. \tag{3.14}$$

Moreover, up to making $\mathcal{B}_n$ slightly smaller, we can assume that we also have, for $\bar{\boldsymbol{b}} \in \mathcal{B}_{n_1}$

$$\mu(\xi_{[0,n_1[} = \bar{\boldsymbol{b}}, \xi_{[0,n[} \in \mathcal{B}_n) \geq \frac{1}{2}\mu(\xi_{[0,n_1[} = \bar{\boldsymbol{b}}). \tag{3.15}$$

We do this by considering the set

$$\mathcal{C} := \{\bar{\boldsymbol{b}} \in \mathcal{B}_{n_1} \,|\, \mu(\xi_{[0,n_1[} = \bar{\boldsymbol{b}}, \xi_{[0,n[} \notin \mathcal{B}_n) \geq \frac{1}{2}\mu(\xi_{[0,n_1[} = \bar{\boldsymbol{b}})\}.$$

From the definition of $\mathcal{C}$, we get

$$\frac{1}{2}\mu(\xi_{[0,n_1[} \in \mathcal{C}) \leq \mu(\xi_{[0,n[} \notin \mathcal{B}_n) \leq \beta.$$

Then, we replace $\mathcal{B}_n$ with $\{\boldsymbol{b} \in \mathcal{B}_n \,|\, \boldsymbol{b}_{n_1} \notin \mathcal{C}\}$. Because the set we removed from $\mathcal{B}_n$ is measurable with respect to the truncated sequences of length n_1, for $\bar{\boldsymbol{b}} \notin \mathcal{C}$, this change does not affect the value of the left-hand term in (**??**). Using this and the definition of $\mathcal{C}$, we see that after replacing $\mathcal{B}_n$ with $\{\boldsymbol{b} \in \mathcal{B}_n \,|\, \boldsymbol{b}_{n_1} \notin \mathcal{C}\}$, we can indeed assume that (**??**) holds for every $\bar{\boldsymbol{b}} \in \mathcal{B}_{n_1}$.

Finally, putting (**??**) and (**??**) together, we get, for $\bar{\boldsymbol{b}} \in \mathcal{B}_{n_1}$

$$\mu(\xi_{[0,n_1[} = \bar{\boldsymbol{b}}, \xi_{[0,n[} \in \mathcal{B}_n) \geq \frac{1}{2}2^{-(h_\mu(\xi,T)+\beta)n_1} \geq 2^{-(h_\mu(\xi,T)+2\beta)n_1}, \tag{3.16}$$

for n_1 large enough. This will enable us to control the measure of the part of the truncated column over $\{\xi_{[0,n_1[} = \bar{\boldsymbol{b}}\}$ that is in $\mathcal{T}_n$.

Step 4: Estimates for Hall's marriage lemma

From (**??**) and (**??**), we can tell that

$$\begin{aligned}\#\mathcal{B}_{n_1} &\le 2^{(h_\mu(\xi,T)+\beta)n_1} \le 2^{(h_\mu(\xi,T)+\beta)(1-\gamma)n} \\ &\le 2^{(h_\mu(\xi,T)-\gamma h_\mu(\xi,T)+\beta)n} \le (1-\beta)2^{(h_\mu(\xi,T)-\beta)n} \le \#\mathcal{B}_n,\end{aligned}$$

for β small enough with respect to γ. This inequality is also clearly true from the definition of $\mathcal{B}_{n_1}$, but we include this computation, as a similar one will be essential later in the proof. That being said, this inequality means that it is possible to find an injective map from $\mathcal{B}_{n_1}$ to $\mathcal{B}_n$, but we want to be more specific about which injective map we choose. To that end, we will make use of Hall's marriage lemma. To do that, for each $\bar{\boldsymbol{b}} \in \mathcal{B}_{n_1}$, we need to specify which elements of $\mathcal{B}_n$ we consider suitable $\tilde{\xi}$-names for the columns $\{C_{\boldsymbol{h}}^{\boldsymbol{b}};\ \boldsymbol{b} \in B(\bar{\boldsymbol{b}})\}$.

We recall that n_0 is the integer chosen so that $\tilde{\mathcal{P}}_0$ is $(\mathcal{H}\vee\xi)_{[-n_0,n_0]}$-measurable. Define $L_n := [n_0, n_1 - n_0[\subset \mathbb{Z}$ and $\ell := n_1 - 2n_0$ the length of L_n. Because $\tilde{\mathcal{P}}_0$ is $(\mathcal{H}\vee\xi)_{[-n_0,n_0]}$-measurable, $\tilde{\mathcal{P}}_{L_n}$ is $(\mathcal{H}\vee\xi)_{[0,n_1[}$-measurable. So, for $\boldsymbol{h}$ fixed, for each $\bar{\boldsymbol{b}} \in \mathcal{B}_{n_1}$, on the set $\{\mathcal{H}_{[0,n[} = \boldsymbol{h}, \xi_{[0,n_1[} = \bar{\boldsymbol{b}}\}$, there can be only one value of $\tilde{\mathcal{P}}_{L_n}$, which we denote $\tilde{\mathcal{P}}_{L_n}(\boldsymbol{h}, \bar{\boldsymbol{b}})$.

For $\bar{\boldsymbol{b}} \in \mathcal{B}_{n_1}$, the suitable corresponding $\tilde{\xi}$-names will be the elements $\boldsymbol{b} \in \mathcal{B}_n$ for which $d_\ell(\tilde{\mathcal{P}}_{L_n}(\boldsymbol{h}_0, \boldsymbol{b}_{n_1}), \tilde{\mathcal{P}}_{L_n}(\boldsymbol{h}, \bar{\boldsymbol{b}})) \le \varepsilon$. More formally, we set

$$J_{\bar{\boldsymbol{b}}} := \{\boldsymbol{b} \in \mathcal{B}_n \,|\, d_\ell(\tilde{\mathcal{P}}_{L_n}(\boldsymbol{h}_0, \boldsymbol{b}_{n_1}), \tilde{\mathcal{P}}_{L_n}(\boldsymbol{h}, \bar{\boldsymbol{b}})) \le \varepsilon\},$$

and we want to build $\psi_{\boldsymbol{h}}$ so that we have

$$\psi_{\boldsymbol{h}}(\bar{\boldsymbol{b}}) \in J_{\bar{\boldsymbol{b}}}, \tag{3.17}$$

for as many $\bar{\boldsymbol{b}} \in \mathcal{B}_{n_1}$ as possible.

From (**??**), it follows that

$$\bar{d}_n(\mathcal{L}(\tilde{\mathcal{P}}_{[0,n[} \,|\, \mathcal{Q} = \boldsymbol{h}), \mathcal{L}(\tilde{\mathcal{P}}_{[0,n[} \,|\, \mathcal{Q} = \boldsymbol{h}_0)) \le 2\varepsilon^3/3.$$

Therefore:

$$\begin{aligned}&\bar{d}_\ell(\mathcal{L}(\tilde{\mathcal{P}}_{L_n} \,|\, \mathcal{Q} = \boldsymbol{h}), \mathcal{L}(\tilde{\mathcal{P}}_{L_n} \,|\, \mathcal{Q} = \boldsymbol{h}_0)) \\ &\qquad\le \frac{n}{n_1 - 2n_0}\bar{d}_n(\mathcal{L}(\tilde{\mathcal{P}}_{[0,n[} \,|\, \mathcal{Q} = \boldsymbol{h}), \mathcal{L}(\tilde{\mathcal{P}}_{[0,n[} \,|\, \mathcal{Q} = \boldsymbol{h}_0)) \\ &\qquad\le \frac{n}{(1-\gamma)n - n_0} 2\varepsilon^3/3 < \varepsilon^3,\end{aligned}$$

for n large enough. So there exists $\lambda \in \mathscr{P}(A^{L_n} \times A^{L_n})$ a coupling of $\mathcal{L}(\tilde{\mathcal{P}}_{L_n} \mid \mathcal{Q} = \boldsymbol{h})$ and $\mathcal{L}(\tilde{\mathcal{P}}_{L_n} \mid \mathcal{Q} = \boldsymbol{h}_0)$ such that

$$\int d_\ell(\boldsymbol{p}_1, \boldsymbol{p}_2) d\lambda(\boldsymbol{p}_1, \boldsymbol{p}_2) \leq \varepsilon^3.$$

Denote by λ_1 and λ_2 the marginals of λ, i.e. $\lambda_1 = \mathcal{L}(\tilde{\mathcal{P}}_{L_n} \mid \mathcal{Q} = \boldsymbol{h})$ and $\lambda_2 = \mathcal{L}(\tilde{\mathcal{P}}_{L_n} \mid \mathcal{Q} = \boldsymbol{h}_0)$. We are interested in the set $\mathcal{A}_\ell \subset A^{L_n}$ defined by

$$\mathcal{A}_\ell := \{\boldsymbol{p} \in A^{L_n}\, ;\, \lambda(d_\ell(\boldsymbol{p}_1, \boldsymbol{p}_2) \leq \varepsilon \,|\, \boldsymbol{p}_1 = \boldsymbol{p}) \geq 1 - \varepsilon\}.$$

The following gives an estimate on the measure of $\mathcal{A}_\ell$:

$$\begin{aligned}
\varepsilon^3 \geq \int d_\ell(\boldsymbol{p}_1, \boldsymbol{p}_2) d\lambda(\boldsymbol{p}_1, \boldsymbol{p}_2) &\geq \int_{\boldsymbol{p}_1 \notin \mathcal{A}_\ell} d_\ell(\boldsymbol{p}_1, \boldsymbol{p}_2) d\lambda(\boldsymbol{p}_1, \boldsymbol{p}_2) \\
&= \int_{\boldsymbol{p} \notin \mathcal{A}_\ell} \int d_\ell(\boldsymbol{p}_1, \boldsymbol{p}_2) d\lambda(\boldsymbol{p}_1, \boldsymbol{p}_2 \,|\, \boldsymbol{p}_1 = \boldsymbol{p}) d\lambda_1(\boldsymbol{p}) \\
&\geq \int_{\boldsymbol{p} \notin \mathcal{A}_\ell} \lambda(d_\ell(\boldsymbol{p}_1, \boldsymbol{p}_2) > \varepsilon \,|\, \boldsymbol{p}_1 = \boldsymbol{p}) \cdot \varepsilon d\lambda_1(\boldsymbol{p}) \\
&> \varepsilon^2 \mu(\tilde{\mathcal{P}}_{L_n} \notin \mathcal{A}_\ell \,|\, \mathcal{Q} = \boldsymbol{h}),
\end{aligned}$$

so $\mu(\tilde{\mathcal{P}}_{L_n} \notin \mathcal{A}_\ell \,|\, \mathcal{Q} = \boldsymbol{h}) < \varepsilon$. In other words, if we set

$$\bar{\mathcal{B}}_{n_1}(\boldsymbol{h}) := \{\bar{\boldsymbol{b}} \in \mathcal{B}_{n_1} \,|\, \tilde{\mathcal{P}}_{L_n}(\boldsymbol{h}, \bar{\boldsymbol{b}}) \in \mathcal{A}_\ell\},$$

we have $\mu(\xi_{[0,n_1[} \in \bar{\mathcal{B}}_{n_1}(\boldsymbol{h}) \,|\, \mathcal{Q} = \boldsymbol{h}) \geq 1 - \varepsilon$. The set $\bar{\mathcal{B}}_{n_1}(\boldsymbol{h})$ is the set on which we want (**??**) to hold. Hall's marriage lemma tells us that there exists an injective map $\psi_{\boldsymbol{h}} : \bar{\mathcal{B}}_{n_1}(\boldsymbol{h}) \to \mathcal{B}_n$ for which (**??**) is true if we have the following:

$$\forall I \subset \bar{\mathcal{B}}_{n_1}(\boldsymbol{h}),\ \#I \leq \# \bigcup_{\bar{\boldsymbol{b}} \in I} J_{\bar{\boldsymbol{b}}}. \tag{3.18}$$

Let $I \subset \bar{\mathcal{B}}_{n_1}(\boldsymbol{h})$. Consider $K := \bigcup_{\bar{\boldsymbol{b}} \in I} \{\tilde{\mathcal{P}}_{L_n}(\boldsymbol{h}, \bar{\boldsymbol{b}})\} \subset \mathcal{A}_\ell$ and note that

$$\bigcup_{\bar{\boldsymbol{b}} \in I} J_{\bar{\boldsymbol{b}}} = \{\boldsymbol{b} \in \mathcal{B}_n \,|\, d_\ell(\tilde{\mathcal{P}}_{L_n}(\boldsymbol{h}_0, \boldsymbol{b}_{n_1}), K) \leq \varepsilon\}.$$

Taking that into account, we have

$$\begin{aligned}
\#\bigcup_{\bar{\boldsymbol{b}}\in I} J_{\bar{\boldsymbol{b}}} &\geq 2^{(h_\mu(\xi,T)-\beta)n}\mu(d_\ell(\tilde{\mathcal{P}}_{L_n}(\boldsymbol{h_0},\xi_{[0,n_1[}),K)\leq\varepsilon,\xi_{[0,n[}\in\mathcal{B}_n)\\
&= 2^{(h_\mu(\xi,T)-\beta)n}\mu(d_\ell(\tilde{\mathcal{P}}_{L_n},K)\leq\varepsilon,\xi_{[0,n[}\in\mathcal{B}_n\,|\,\mathcal{H}_{[0,n[}=\boldsymbol{h}_0)\\
&= 2^{(h_\mu(\xi,T)-\beta)n}\mu(\xi_{[0,n[}\in\mathcal{B}_n)\mu(d_\ell(\tilde{\mathcal{P}}_{L_n},K)\leq\varepsilon\,|\,\mathcal{Q}=\boldsymbol{h}_0)\\
&\geq 2^{(h_\mu(\xi,T)-\beta)n}(1-\beta)\lambda(d_\ell(\boldsymbol{p}_1,\boldsymbol{p}_2)\leq\varepsilon,\boldsymbol{p}_1\in K) \qquad (3.19)\\
&\geq 2^{(h_\mu(\xi,T)-2\beta)n}\int_{\boldsymbol{p}\in K}\lambda(d_\ell(\boldsymbol{p}_1,\boldsymbol{p}_2)\leq\varepsilon\,|\,\boldsymbol{p}_1=\boldsymbol{p})d\lambda_1(\boldsymbol{p})\\
&\geq 2^{(h_\mu(\xi,T)-2\beta)n}(1-\varepsilon)\lambda_1(K),\ \text{ because } K\subset\mathcal{A}_\ell\\
&\geq 2^{(h_\mu(\xi,T)-3\beta)n}\mu(\tilde{\mathcal{P}}_{L_n}\in K\,|\,\mathcal{Q}=\boldsymbol{h}).
\end{aligned}$$

Moreover, using (**??**), we get

$$\begin{aligned}
\#I &\leq 2^{(h_\mu(\xi,T)+2\beta)n_1}\mu(\xi_{[0,n_1[}\in I,\xi_{[0,n[}\in\mathcal{B}_n)\\
&\leq 2^{(h_\mu(\xi,T)+2\beta)n_1}\mu(\xi_{[0,n_1[}\in I\,|\,\xi_{[0,n[}\in\mathcal{B}_n)\\
&= 2^{(h_\mu(\xi,T)+2\beta)n_1}\mu(\xi_{[0,n_1[}\in I\,|\,\mathcal{Q}=\boldsymbol{h}),\ \text{ because } \xi\perp\!\!\!\perp\mathcal{H}\\
&\leq 2^{(h_\mu(\xi,T)+2\beta)n_1}\mu(\tilde{\mathcal{P}}_{L_n}\in K\,|\,\mathcal{Q}=\boldsymbol{h}),
\end{aligned}$$

by definition of K. Together with (**??**), it yields

$$\begin{aligned}
\#I &\leq 2^{((1-\gamma)(h_\mu(\xi,T)+2\beta)-(h_\mu(\xi,T)-3\beta))n}\#\bigcup_{\bar{\boldsymbol{b}}\in I}J_{\bar{\boldsymbol{b}}}\\
&\leq 2^{(5\beta-\gamma h_\mu(\xi,T))n}\#\bigcup_{\bar{\boldsymbol{b}}\in I}J_{\bar{\boldsymbol{b}}}\leq\#\bigcup_{\bar{\boldsymbol{b}}\in I}J_{\bar{\boldsymbol{b}}},
\end{aligned}$$

for β small enough with respect to γ. Therefore there exists an injective map $\psi_{\boldsymbol{h}}:\bar{\mathcal{B}}_{n_1}(\boldsymbol{h})\to\mathcal{B}_n$ for which (**??**) holds. As we noted that $\#\mathcal{B}_{n_1}\leq\#\mathcal{B}_n$, $\psi_{\boldsymbol{h}}$ can then be extended to an injective map defined on $\mathcal{B}_{n_1}$ (still taking values in $\mathcal{B}_n$). We recall that, with $\psi_{\boldsymbol{h}}$ built, we set $\varphi_{\boldsymbol{h}}(\boldsymbol{b}):=\psi_{\boldsymbol{h}}(\boldsymbol{b}_{n_1})$.

As we announced at the start of our reasoning, we define $\tilde{\xi}_0$ on the levels of $C_{\boldsymbol{h}}$ so that the $\tilde{\xi}$-name of each sub-column $C_{\boldsymbol{h}}^{\boldsymbol{b}}$ is $\varphi_{\boldsymbol{h}}(\boldsymbol{b})=\psi_{\boldsymbol{h}}(\boldsymbol{b}_{n_1})$. Since this construction can be done with every $\boldsymbol{h}\in\bar{\mathcal{E}}_n$ (with the map $\psi_{\boldsymbol{h}}$ depending on $\boldsymbol{h}$), we have completed the construction of $\tilde{\xi}_0$. We now need to check that $\tilde{\xi}$ satisfies the conditions (i), (ii) and (iii) of our lemma.

Step 5: Proving that $\tilde{\xi}$ satsifies (i), (ii) and (iii)

We start by estimating the law of $\tilde{\xi}_0$. Since $\mu(\mathcal{T}_n) \geq 1 - 3\beta$, we have

$$\begin{aligned}\sum_{b\in B}|\mu(\tilde{\xi}_0 = b) - \mu(\xi_0 = b)| &\leq \sum_{b\in B}|\mu(\{\tilde{\xi}_0 = b\} \cap \mathcal{T}_n) - \mu(\xi_0 = b)\mu(\mathcal{T}_n)| + 6\beta \\ &\leq \sum_{b\in B}\sum_{\boldsymbol{h}\in\mathcal{E}_n, \boldsymbol{b}\in\mathcal{B}_n}|\mu(\{\tilde{\xi}_0 = b\} \cap C_{\boldsymbol{h}}^{\boldsymbol{b}}) - \mu(\xi_0 = b)\mu(C_{\boldsymbol{h}}^{\boldsymbol{b}})| + 6\beta.\end{aligned}$$

We recall that $f_n(\boldsymbol{b}, \boldsymbol{b'})$ is the frequency at which the element $\boldsymbol{b'}$ appears in the sequence $\boldsymbol{b}$ (see (**??**)). Moreover, one can see that, since $\varphi_{\boldsymbol{h}}(\boldsymbol{b})$ is the $\tilde{\xi}$-name of $C_{\boldsymbol{h}}^{\boldsymbol{b}}$ and all the levels of $C_{\boldsymbol{h}}^{\boldsymbol{b}}$ have the same measure, we have

$$\mu(\{\tilde{\xi}_0 = b\} \cap C_{\boldsymbol{h}}^{\boldsymbol{b}}) = \mu(C_{\boldsymbol{h}}^{\boldsymbol{b}}) \cdot f_n(\varphi_{\boldsymbol{h}}(\boldsymbol{b}), b).$$

Therefore, because $\varphi_{\boldsymbol{h}}$ takes values in $\mathcal{B}_n$, (**??**) yields:

$$\begin{aligned}\sum_{b\in B}|\mu(\tilde{\xi}_0 = b) - \mu(\xi_0 = b)| &\leq \sum_{b\in B}\sum_{\boldsymbol{h}\in\mathcal{E}_n, \boldsymbol{b}\in\mathcal{B}_n}\mu(C_{\boldsymbol{h}}^{\boldsymbol{b}})|f_n(\varphi_{\boldsymbol{h}}(\boldsymbol{b}), b) - \mu(\xi_0 = b)| + 6\beta \\ &\leq \beta\mu(\mathcal{T}_n) + 6\beta \leq 7\beta.\end{aligned}$$

This means that $\bar{d}_1(\mathcal{L}(\tilde{\xi}_0), \mathcal{L}(\xi_0)) \leq 7\beta \leq \alpha$, for β small enough.

We now turn our attention to the entropy of $\mathcal{H} \vee \tilde{\xi}$. The $\tilde{\xi}$-name of a column $C_{\boldsymbol{h}}^{\boldsymbol{b}}$ is $\psi_{\boldsymbol{h}}(\boldsymbol{b}_{n_1})$, and since $\psi_{\boldsymbol{h}}$ is invective, we can deduce $\boldsymbol{b}_{n_1}$ from the $\tilde{\xi}$-name of $C_{\boldsymbol{h}}^{\boldsymbol{b}}$. This means that, on the levels of the truncated tower $\mathcal{T}_{n_1} := \{T^j F\}_{0\leq j\leq n_1-1}$, ξ_0 is $(\mathcal{H} \vee \tilde{\xi})_{[-n_1,n[}$-measurable. Indeed, if x is in $\mathcal{T}_{n_1}$ and the sequence $(\mathcal{H} \vee \tilde{\xi})_{[-n_1,n[}(x)$ is known, the sequence $\tilde{\xi}_{[-n_1,0[}(x)$ must contain a "$*$", which indicates the moment the past orbit of x passes trough G before entering $\mathcal{T}_n$. So the position of "$*$" in $\tilde{\xi}_{[-n_1,0[}(x)$ tells us the index of the level of $\mathcal{T}_{n_1}$ the point x is on, which we call j_0. In other words, we mean that $T^{-j_0}x \in F$. Then, $(\mathcal{H} \vee \tilde{\xi})_{[-j_0,n-j_0[}$ gives the $(\mathcal{H} \vee \tilde{\xi})$-name of the column x is on, from which we deduce the truncated ξ-name of length n_1 of the column. Finally, the j_0-th letter of that name gives us $\xi_0(x)$.

Therefore, if we combine the previous paragraph with the fact that $\mu(\mathcal{T}_{n_1}) \geq (1 - \gamma)\mu(\mathcal{T}_n)$, there exists a $(\mathcal{H} \vee \tilde{\xi})_{[-n_1,n[}$-measurable random variable χ_0 such that $\mu(\chi_0 \neq \xi_0) \leq \beta + \gamma \leq 2\gamma$ (if $\beta \leq \gamma$). So, by applying Lemma **??**, for γ small enough with respect to α, we conclude that

$$h_\mu(\mathcal{H} \vee \xi, T) \leq h_\mu(\mathcal{H} \vee \chi, T) + \alpha \leq h_\mu(\mathcal{H} \vee \tilde{\xi}, T) + \alpha.$$

Because $\mathcal{H} \vee \xi$ generates $\mathscr{A}$, we also have the converse inequality

$$h_\mu(\mathcal{H} \vee \tilde{\xi}, T) \leq h_\mu(\mathcal{H} \vee \xi, T),$$

so we have proved that $\tilde{\xi}$ satisfies condition (ii) of our lemma.

We are now left with proving (iii). If we consider that $\tilde{\xi}_{[-n,n[}(x)$ is known, we deduce that, if the symbol "$*$" appears in $\tilde{\xi}_{[-n,0[}$, then x is in $\mathcal{T}_n$ and the position of "$*$" tells us the index j_0 of the level of $\mathcal{T}_n$ the point x is on. Then, using the notation introduced in our construction above, we can look at the random variable $\tilde{\mathcal{P}}_{j_0}(\boldsymbol{h}_0, \tilde{\xi}_{[-j_0,n_1-j_0[})$. It is $\tilde{\xi}_{[-n,n[}$-measurable and we are going to show that it satisfies

$$\mu(\tilde{\mathcal{P}}_{j_0}(\boldsymbol{h}_0, \tilde{\xi}_{[-j_0,n_1-j_0[}) \neq \mathcal{P}_0) \leq 5\varepsilon. \tag{3.20}$$

We start by looking at a column $C_{\boldsymbol{h}}$ for some $\boldsymbol{h} \in \bar{\mathcal{E}}_n$. We then split it into sub-columns $C_{\boldsymbol{h}}^{\boldsymbol{b}}$. If $\boldsymbol{b}_{n_1} \in \bar{\mathcal{B}}_{n_1}(\boldsymbol{h})$, we are going to use (**??**). First, we need to remember that if x is in $C_{\boldsymbol{h}}^{\boldsymbol{b}}$, then $\tilde{\xi}_{[-j_0,n-j_0[}(x)$ gives the $\tilde{\xi}$-name of the column $C_{\boldsymbol{h}}^{\boldsymbol{b}}$. But, by construction, that name is $\psi_{\boldsymbol{h}}(\boldsymbol{b}_{n_1})$, and, because we are looking at the case where $\boldsymbol{h} \in \bar{\mathcal{E}}_n$ and $\boldsymbol{b}_{n_1} \in \bar{\mathcal{B}}_{n_1}(\boldsymbol{h})$, (**??**) holds. So we have

$$d_\ell(\tilde{\mathcal{P}}_{L_n}(\boldsymbol{h}_0, \tilde{\xi}_{[-j_0,n_1-j_0[}), \tilde{\mathcal{P}}_{L_n}(\boldsymbol{h}, \boldsymbol{b}_{n_1})) \leq \varepsilon.$$

We recall that $L_n = [n_0, n_1 - n_0[$ and ℓ is its length. By definition of d_ℓ, we know that the number of levels $j_0 \in L_n$ on which we have $\tilde{\mathcal{P}}_{j_0}(\boldsymbol{h}_0, \tilde{\xi}_{[-j_0,n_1-j_0[}) = \tilde{\mathcal{P}}_{j_0}(\boldsymbol{h}, \boldsymbol{b}_{n_1})$ is greater than $(1-\varepsilon)\ell$. Moreover, by construction, for $j_0 \in L_n$, on the j_0-th level of $C_{\boldsymbol{h}}^{\boldsymbol{b}}$, we have $\tilde{\mathcal{P}}_0 = \tilde{\mathcal{P}}_{j_0}(\boldsymbol{h}, \boldsymbol{b}_{n_1})$. Finally, since $\ell = n_1 - 2n_0$, we have

$$\begin{aligned}
\mu(\tilde{\mathcal{P}}_{j_0}(\boldsymbol{h}_0, \tilde{\xi}_{[-j_0,n_1-j_0[}) \neq \tilde{\mathcal{P}}_0 \,|\, C_{\boldsymbol{h}}^{\boldsymbol{b}}) &\leq \frac{n - (1-\varepsilon)(n_1 - 2n_0)}{n} \\
&\leq \frac{n - (1-\varepsilon)(1-\gamma)n + 2n_0}{n} \\
&\leq \varepsilon + \gamma + \frac{2n_0}{n} \leq 2\varepsilon,
\end{aligned}$$

for $\gamma \leq \varepsilon/2$ and n large enough so that $2n_0/n \leq \varepsilon/2$. Moreover, the fact that $\boldsymbol{h} \in \bar{\mathcal{E}}_n$ implies that $\mu(\xi_{[0,n_1[} \in \bar{\mathcal{B}}_{n_1}(\boldsymbol{h}) \,|\, \mathcal{Q} = \boldsymbol{h}) \geq 1 - \varepsilon$, and, combining it with (**??**), we can see that

$$\mu\left(\bigcup_{\boldsymbol{b}_{n_1} \notin \bar{\mathcal{B}}_{n_1}(\boldsymbol{h})} C_{\boldsymbol{h}}^{\boldsymbol{b}} \,\middle|\, C_{\boldsymbol{h}}\right) \leq \varepsilon.$$

Therefore

$$\mu(\tilde{\mathcal{P}}_{j_0}(\boldsymbol{h}_0, \tilde{\xi}_{[-j_0, n_1-j_0[}) \neq \tilde{\mathcal{P}}_0 \,|\, C_{\boldsymbol{h}}) \leq 3\varepsilon.$$

Next

$$\begin{aligned}
\mu(\tilde{\mathcal{P}}_{j_0}(\boldsymbol{h}_0, \tilde{\xi}_{[-j_0, n_1-j_0[}) \neq \tilde{\mathcal{P}}_0) &\leq 3\varepsilon + \mu\left(\bigcup_{\boldsymbol{h}\notin\bar{\mathcal{E}}_n} C_{\boldsymbol{h}}\right) + \mu(\mathcal{T}_n^c) \\
&\leq 3\varepsilon + \varepsilon + \mu(\mathcal{T}_n^c) \;\text{ using } (\mathbf{??}) \text{ and } (\mathbf{??}) \\
&\leq 4\varepsilon + 3\beta.
\end{aligned}$$

Finally, $\tilde{\mathcal{P}}_0$ was chosen so that $\mu(\tilde{\mathcal{P}}_0 \neq \mathcal{P}_0) \leq \beta^2$, so for β small enough, we have proven (**??**), and therefore, up to replacing ε by $\varepsilon/5$, we have shown that

$$\mathcal{P}_0 \preceq_\varepsilon \sigma(\tilde{\xi}).$$

□

3.2.2 Application of the technical lemma

We are now left with proving Proposition **??** using Lemma **??**. This is done using some abstract results from Thouvenot [**?**, Proposition 2', Proposition 3]. We start by rewriting those results with our notation. We give a slight simplification, adapted to our setup.

First, [**?**, Proposition 2'] tells us that a process close enough to an i.i.d. process independent from $\mathscr{H}$ in law and entropy can be turned into an i.i.d. process independent from $\mathscr{H}$.

Proposition 3.2.8. *Let* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *be an ergodic system of finite entropy. Let* $\mathcal{H}$ *be a finite valued process defined on* $\mathbf{X}$ *and* ρ *be a probability measure on a finite alphabet* B*. For every* $\varepsilon > 0$*, there exist* $\alpha > 0$ *such that if a random variable* $\tilde{\xi}_0 : X \to B$ *satisfies*

(i) $\bar{d}_1(\rho, \mathcal{L}(\tilde{\xi}_0)) \leq \alpha$*,*

(ii) and $0 \leq h_\mu(\mathcal{H}, T) + H(\rho) - h_\mu(\mathcal{H} \vee \tilde{\xi}, T) \leq \alpha$*,*

then there exists a random variable ξ'_0 *of law* ρ *such that the process* $\xi' := (\xi'_0 \circ T^n)_{n\in\mathbb{Z}}$ *is i.i.d., independent from* $\mathcal{H}$ *and we have*

$$\mu(\tilde{\xi}_0 \neq \xi'_0) \leq \varepsilon.$$

Next, [**?**, Proposition 3] tells us that on a system that is relatively Bernoulli over a factor $\mathscr{H}$, any i.i.d. process independent from $\mathscr{H}$ with the right entropy can be turned into an independent complement of $\mathscr{H}$:

Proposition 3.2.9. *Let* $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ *be an ergodic system,* $\mathcal{H}$ *a finite valued process and* ξ *a finite valued i.i.d. process independent from* $\mathcal{H}$ *such that* $\mathscr{A} = \sigma(\mathcal{H}) \vee \sigma(\xi)$ *mod* μ*. For any* $\varepsilon > 0$ *and any i.i.d. process* ζ *independent from* $\mathcal{H}$ *such that* $h_\mu(\xi, T) = h_\mu(\zeta, T)$*, there exists* $\tilde{\zeta}_0$ *such that* $\mathcal{L}(\mathcal{H} \vee \tilde{\zeta}) = \mathcal{L}(\mathcal{H} \vee \zeta)$*,* $\mathscr{A} = \sigma(\mathcal{H}) \vee \sigma(\tilde{\zeta})$ *mod* μ *and*

$$\mu(\tilde{\zeta}_0 \neq \zeta_0) \leq \varepsilon.$$

We are now fully equipped to end the proof of Proposition **??**:

Proof of Proposition **??**. Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a Bernoulli shift of finite entropy and $\mathcal{P}_0 : X \to A$ a finite valued random variable such that $\mathscr{A} = \sigma(\{\mathcal{P}_0 \circ T^n\}_{n\in\mathbb{Z}})$. As we consider a factor σ-algebra $\mathscr{H}$ of $\mathbf{X}$, it has finite entropy, therefore there exists a finite valued random variable $\mathcal{H}_0 : X \to H$ such that the process $\mathcal{H} := (\mathcal{H}_0 \circ T^n)_{n\in\mathbb{Z}}$ generates $\mathscr{H}$. Lastly, we take an i.i.d. process ξ independent from $\mathscr{H}$ such that $\mathscr{A} = \mathscr{H} \vee \sigma(\xi)$ mod μ. Let $\varepsilon > 0$.

Now, Lemma **??** tells us that there is $\delta > 0$ for which, if $h_\mu(\mathscr{H}, T) \leq \delta$, then for any $\alpha > 0$, there is a random variable $\tilde{\xi}_0$ such that

(i) $\bar{d}_1(\mathcal{L}(\xi_0), \mathcal{L}(\tilde{\xi}_0)) \leq \alpha$,

(ii) $0 \leq h_\mu(\mathcal{H} \vee \xi, T) - h_\mu(\mathcal{H} \vee \tilde{\xi}, T) \leq \alpha$,

(iii) and $\mathcal{P}_0 \preceq_{\varepsilon/4} \sigma(\tilde{\xi})$.

Denote $\tilde{\mathcal{P}}_0$ a $\tilde{\xi}$-measurable random variable such that $\mu(\tilde{\mathcal{P}}_0 \neq \mathcal{P}_0) \leq \varepsilon/4$. We can find an integer $N \geq 1$ for which $\tilde{\mathcal{P}}_0 \preceq_{\varepsilon/4} \tilde{\xi}_{[-N,N]}$ and set $\varepsilon_1 := \varepsilon/(4(2N+1))$. If α is chosen small enough, then Proposition **??** tells us that there is a random variable ξ'_0 such that the process $(\xi'_0 \circ T^n)_{n\in\mathbb{Z}}$ is i.i.d., independent from $\mathcal{H}$ and we have $\mu(\xi'_0 \neq \tilde{\xi}_0) \leq \varepsilon_1$. Finally, Proposition **??** tells us that we can then find a random variable ξ''_0 for which the process $(\xi''_0 \circ T^n)_{n\in\mathbb{Z}}$ is still i.i.d., independent from $\mathcal{H}$, but we also have that $\mathscr{A} = \mathscr{H} \vee \sigma(\xi'')$ mod μ and $\mu(\xi'_0 \neq \tilde{\xi}_0) \leq \varepsilon_1$. So we have $\mu(\xi''_0 \neq \tilde{\xi}_0) \leq 2\varepsilon_1$.

Combining that with the fact that $\tilde{\mathcal{P}}_0 \preceq_{\varepsilon/4} \tilde{\xi}_{[-N,N]}$, we get that $\tilde{\mathcal{P}}_0 \preceq_{3\varepsilon/4} \xi''_{[-N,N]}$, so

$$\mathcal{P}_0 \preceq_\varepsilon \xi''_{[-N,N]}.$$

Setting $\mathscr{B} := \sigma(\xi'')$, we get the Bernoulli factor desired to prove our proposition. □

3.2.3 Proof of Theorem ??

In the previous section, we managed to conclude the proof of Proposition **??**. We now see how Theorem **??** follows from that proposition:

Proof of Theorem **??**. Let $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ be a Bernoulli system and $\mathscr{F} := (\mathscr{F}_n)_{n\leq 0}$ be a weak Pinsker filtration. Since $\mathscr{F}$ is a weak Pinsker filtration, if $(\mathscr{F}_n)_{\leq -1}$ is of product type, so is $\mathscr{F}$. Therefore, up to replacing $\mathbf{X}$ by the factor generated by $\mathscr{F}_{-1}$, we can assume that $\mathbf{X}$ has finite entropy. Thanks to Theorem **??**, this means that we can set a finite alphabet A and a random variable $\mathcal{P}_0 : X \to A$ such that the corresponding process $\mathcal{P} := (\mathcal{P}_0 \circ T^i)_{i\in\mathbb{Z}}$ generates $\mathscr{A}$, i.e. $\mathscr{A} = \sigma(\mathcal{P}) \mod \mu$. Let $(\varepsilon_k)_{k\geq 1}$ be a decreasing sequence of positive numbers such that $\lim_{k\to\infty} \varepsilon_k = 0$.

We need to build a strictly increasing sequence $(n_k)_{k\leq 0}$ such that $(\mathscr{F}_{n_k})_{k\leq 0}$ is of product type. We start by setting $n_0 = 0$. Since $\lim_{n\to-\infty} h_\mu(\mathscr{F}_n, T) = 0$, we can choose $n_{-1} \leq 0$ large enough (in absolute value), so that $h_\mu(\mathscr{F}_{n_{-1}}, T)$ is small enough for Proposition **??** to enable us to build a Bernoulli factor σ-algebra $\mathscr{B}_{n_{-1}}$ that is an independent complement of $\mathscr{F}_{n_{-1}}$ such that $\mathcal{P}_0 \preceq_{\varepsilon_1} \mathscr{B}_{n_{-1}}$.

Now take $k \leq -1$ and assume that we have built $(\mathscr{B}_{n_{-1}}, ..., \mathscr{B}_{n_k})$ such that they are mutually independent Bernoulli factors such that for $k \leq j \leq -1$, $\mathscr{B}_{n_j}$ is independent from $\mathscr{F}_{n_j}$, $\mathscr{F}_{n_{j+1}} = \mathscr{F}_{n_j} \vee \mathscr{B}_{n_j}$ and we have

$$\mathcal{P}_0 \preceq_{\varepsilon_{|k|}} \bigvee_{k\leq j\leq -1} \mathscr{B}_{n_j}. \tag{3.21}$$

By construction of the $\mathscr{B}_{n_j}$, we know that $\mathcal{P}$ is measurable with respect to $\mathscr{F}_{n_k} \vee \bigvee_{k\leq j\leq -1} \mathscr{B}_{n_j}$. Moreover, using again Theorem **??**, there is a random variable $\mathcal{P}_0^{(k)}$ such that the process $\mathcal{P}^{(k)} := (\mathcal{P}_0^{(k)} \circ T^i)_{i\in\mathbb{Z}}$ generates $\mathscr{F}_{n_k}$. So there exists an integer $N \geq 1$ such that

$$\mathcal{P}_0 \preceq_{\varepsilon_{|k|+1}/2} \mathcal{P}^{(k)}_{[-N,N]} \vee \bigvee_{k\leq j\leq -1} \mathscr{B}_{n_j}. \tag{3.22}$$

Then set $\tilde{\varepsilon} := \varepsilon_{|k|+1}/(2(2N+1)) > 0$. As we did above, we choose $n_{k-1} \leq n_k$ large enough in absolute value so that $h_\mu(\mathscr{F}_{n_{k-1}}, T)$ is small enough for us to apply Proposition **??** to find a Bernoulli factor $\mathscr{B}_{n_{k-1}} \subset \mathscr{F}_{n_k}$ such that $\mathscr{B}_{n_{k-1}} \perp\!\!\!\perp \mathscr{F}_{n_{k-1}}$, $\mathscr{F}_{n_k} = \mathscr{F}_{n_{k-1}} \vee \mathscr{B}_{n_{k-1}}$ and

$$\mathcal{P}_0^{(k)} \preceq_{\tilde{\varepsilon}} \mathscr{B}_{n_{k-1}}. \tag{3.23}$$

Putting (**??**) and (**??**) together, we get

$$\mathcal{P}_0 \preceq_{\varepsilon_{|k|+1}} \bigvee_{k-1 \le j \le -1} \mathscr{B}_{n_j}.$$

Iterating this for every $k \le -1$ ends our construction of $(n_k)_{k \le 0}$ and $(\mathscr{B}_{n_k})_{k \le -1}$. Therefore (**??**) holds for every $k \le -1$. It follows then that $\mathcal{P}_0$ is measurable with respect to

$$\bigvee_{j \le -1} \mathscr{B}_{n_j}.$$

Since the $\mathscr{B}_{n_j}$ are factor σ-algebras, the full process $\mathcal{P}$ is also $\bigvee_{j \le -1} \mathscr{B}_{n_j}$-measurable. Finally, $\mathcal{P}$ generates $\mathscr{A}$, so

$$\bigvee_{j \le -1} \mathscr{B}_{n_j} = \mathscr{A} = \mathscr{F}_0 \mod \mu.$$

Let $k \le -1$, and set $\mathscr{E}_1 := \bigvee_{j \le k-1} \mathscr{B}_{n_j}$ and $\mathscr{E}_2 := \bigvee_{k \le j \le -1} \mathscr{B}_{n_j}$. By construction, we have

$$\mathscr{E}_1 \subset \mathscr{F}_{n_k}, \; \mathscr{F}_{n_k} \perp\!\!\!\perp \mathscr{E}_2, \; \text{ and } \; \mathscr{F}_0 = \mathscr{E}_1 \vee \mathscr{E}_2.$$

We use this to see that if f is $\mathscr{F}_{n_k}$-measurable, we have

$$f = \mathbb{E}[f \,|\, \mathscr{F}_0] = \mathbb{E}[f \,|\, \mathscr{E}_1 \vee \mathscr{E}_2] = \mathbb{E}[f \,|\, \mathscr{E}_1],$$

which proves that

$$\mathscr{F}_{n_k} = \mathscr{E}_1 = \bigvee_{j \le k-1} \mathscr{B}_{n_j} \mod \mu.$$

□

3.3 Examples of weak Pinsker filtrations generated by a cellular automaton

Up to this point, we have discussed the existence and abstract properties of weak Pinsker filtrations. Now we want to give explicit examples to get a more concrete idea of what those objects can look like. We take inspiration from [**?**] and use cellular automata to generate our filtrations. We describe in the following paragraphs how this is done.

Let A be a finite alphabet. A cellular automaton (or, more precisely, a deterministic cellular automaton) $\tau : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ maps $A^{\mathbb{Z}}$ onto itself as follows: take $F \subset \mathbb{Z}$ finite, which we call a neighborhood, and a local map $\tau_0 : A^F \to A$. Then define

$$\tau : (a_n)_{n\in\mathbb{Z}} \mapsto (\tau_0((a_{n+k})_{k\in F}))_{n\in\mathbb{Z}}.$$

Here, we will only consider examples in which $F = \{0,1\}$. Therefore, our automata will be determined by a local map of the form $\tau_0 : A^{\{0,1\}} \to A$. One can note that, by construction, cellular automata commute with the shift transformation

$$S : (a_n)_{n\in\mathbb{Z}} \mapsto (a_{n+1})_{n\in\mathbb{Z}}.$$

So we can consider a dynamical system of the form $\mathbf{Y} := (A^{\mathbb{Z}}, \mathscr{B}, \nu, S)$ where ν is a S-invariant measure, and note that the σ-algebra $\sigma(\tau)$ generated by τ is a factor σ-algebra. We can do better and iterate τ to generate a filtration:

$$\text{for } n \leq 0,\ \mathscr{F}_n := \sigma(\tau^{|n|}).$$

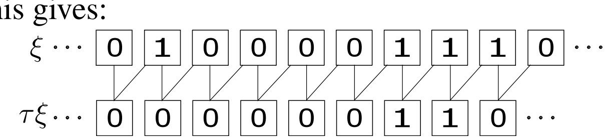
In that case, each $\mathscr{F}_n$ is a factor σ-algebra of $\mathbf{Y}$, and therefore $\mathscr{F} := (\mathscr{F}_n)_{n\leq 0}$ is a *dynamical filtration*. So, we see that cellular automata give a natural way to construct dynamical filtrations.

In fact, the theory of dynamical filtrations we presented in Section **??** was initiated in [**?**] in the setting of filtrations generated by cellular automata. However, the automata studied there preserve the product measure, and therefore the entropy of the associated factor σ-algebras $\mathscr{F}_n$ will be the same for every $n \leq 0$. This prevents the filtration from being weak Pinsker.

Here, we will consider a different automaton: take A a finite alphabet and assume that one element of A is labeled «0». Then define the following local map

$$\begin{array}{rccl} \tau_0 : & A^2 & \longrightarrow & A \\ & (\alpha_1, \alpha_2) & \mapsto & \begin{cases} \alpha_1 & \text{if } \alpha_1 = \alpha_2 \\ 0 & \text{otherwise.} \end{cases} \end{array} \tag{3.24}$$

The associated automaton will eliminate isolated elements, replacing them with 0, and a maximal string of the form $\alpha \cdots \alpha\alpha$ is replaced with $\alpha \cdots \alpha 0$. For example, if $A = \{0,1\}$, this gives:



Therefore, as we iterate the automaton, the proportion of «0» increases as all other elements are gradually replaced by «0». Heuristically, this indicates that the entropy of the factor σ-algebras $\sigma(\tau^{|n|})$ will go to zero as n goes to infinity. But to state this rigorously, one need to specify the system $\mathbf{Y} := (A^{\mathbb{Z}}, \mathscr{B}, \nu, S)$ on which we define $\mathscr{F}$. More accurately, it is the alphabet A and the measure ν that need to be specified. However, the entropy $h_\nu(\mathscr{F}_n)$ goes to 0 regardless of the choice of A and ν:

Proposition 3.3.1. *Let* $\mathbf{Y} := (A^{\mathbb{Z}}, \mathscr{B}, \nu, S)$*, where* ν *is a* S*-invariant measure and let* ξ *be the coordinate process on* $\mathbf{Y}$*. For every* $n \geq 1$*, we have*

$$h_\nu(\tau^n \xi, S) \leq \frac{\log(\#An^2)}{n}.$$

Proof. Let $B_n \subset A^n$ be the set of values taken by $(\tau^n \xi)_{[0,n[}$. We know that

$$H_\nu((\tau^n \xi)_{[0,n[}) \leq \log(\#B_n).$$

Because of the structure of τ, in $\tau^n \xi$, for $\alpha \neq 0$, any run of «α» is placed in between two runs of «0» of length at least $n+1$. Therefore, $(\tau^n \xi)_{[0,n[}$ is either a sequence of «0» or composed of one run of «α» (with $\alpha \neq 0$) in between runs of «0». So

$$\#B_n \leq 1 + (\#A - 1)n^2 \leq \#An^2.$$

In conclusion

$$h_\nu(\tau^n \xi, S) \leq \frac{1}{n} H_\nu((\tau^n \xi)_{[0,n[}) \leq \frac{\log(\#An^2)}{n}.$$

□

In Section **??**, we deal with the case where $\mathbf{Y}$ is a Bernoulli shift, and in Section **??**, we deal with the case where $\mathbf{Y}$ is Ornstein's example of a non-Bernoulli K-system from [**?**]. In both cases, by Proposition **??**, the entropy of the filtration generated by the cellular automaton goes to zero. Then we look at each example separately to show the more involved result: each $\mathscr{F}_{n+1}$ is relatively Bernoulli over $\mathscr{F}_n$. Therefore, we get two examples of weak Pinsker filtrations.

It is interesting to note that those two filtrations are very similar in their construction, but the filtration on Ornstein's K-system cannot be of product type (otherwise, the system would be Bernoulli), while we conjecture that the filtration on the Bernoulli shift is of product type. At least, we know from Theorem **??** that the latter has a sub-sequence that is of product type. It shows that there can be subtle differences in the asymptotic structure of weak Pinsker filtrations.

3.3.1 A cellular automaton on a Bernoulli shift

Here, we consider a Bernoulli shift $\mathbf{Y} := (A^{\mathbb{Z}}, \mathscr{B}, \nu, S)$ where ν is a product measure. To avoid unnecessarily complicated notations, we will also assume that $A = \{0,1\}$ and $\nu := (\frac{1}{2}(\delta_0 + \delta_1))^{\otimes \mathbb{Z}}$. Therefore, the local function (**??**) becomes:

$$\begin{array}{rccc} \tau_0 : & \{0,1\}^2 & \longrightarrow & \{0,1\} \\ & \alpha & \mapsto & \begin{cases} 1 & \text{si } \alpha = (1,1) \\ 0 & \text{otherwise.} \end{cases} \end{array}$$

And we study the corresponding automaton:

$$\begin{array}{rccc} \tau : & \{0,1\}^{\mathbb{Z}} & \longrightarrow & \{0,1\}^{\mathbb{Z}} \\ & (a_n)_{n\in\mathbb{Z}} & \mapsto & (\tau_0(a_n, a_{n+1}))_{n\in\mathbb{Z}} \end{array}$$

The automaton replaces an isolated «1» with a «0» and reduces sequences of «1» by replacing the final one by a «0».

Theorem 3.3.2. *On the system $\mathbf{Y} := (\{0,1\}^{\mathbb{Z}}, \mathscr{B}, \nu, S)$, the filtration given by $\mathscr{F} := (\sigma(\tau^{|n|}))_{n\leq 0}$ is a weak Pinsker filtration. That is, for every $n \leq -1$, $\mathscr{F}_{n+1}$ is relatively Bernoulli over $\mathscr{F}_n$ and we have*

$$h_\nu(\mathscr{F}_n) \underset{n\to-\infty}{\longrightarrow} 0. \tag{3.25}$$

The convergence of the entropy follows from Proposition **??**. However, when $\mathbf{Y}$ is a Bernoulli shift, we can compute a better bound, as stated in Proposition **??**. First, we give a simple technical lemma on Shannon entropy:

Lemma 3.3.3. *Let B be a finite set and $\varepsilon \in]0, e^{-1}[$. If ρ is a random variable taking values in B such that there exists $b_0 \in B$ satisfying $\nu(\rho = b_0) \geq 1 - \varepsilon$, then*

$$H_\nu(\rho) \leq \varepsilon(1 + \log(\#B) + \log(\varepsilon^{-1})).$$

Proof. Let $\varepsilon \in]0, e^{-1}[$. Using the concavity of $\varphi : x \mapsto -x\log(x)$, we see that

$\varphi(x) \leq 1 - x$ and so we get

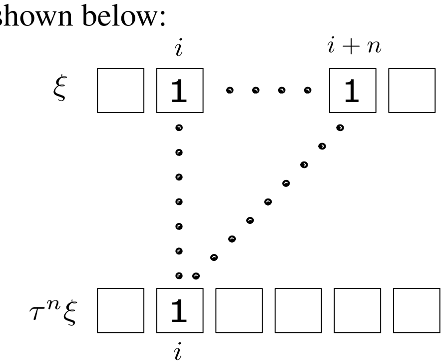
$$\begin{aligned}
H_\nu(\rho) &= \varphi(\nu(\rho = b_0)) + \sum_{b \neq b_0} \varphi(\nu(\rho = b)) \\
&\leq \varepsilon + (\#B - 1)\varphi\left(\frac{1}{\#B - 1} \sum_{b \neq b_0} \nu(\rho = b)\right) \\
&= \varepsilon + (\#B - 1)\varphi\left(\frac{\mu(\rho \neq b_0)}{\#B - 1}\right) \\
&= \varepsilon - \nu(\rho \neq b_0) \cdot \log\left(\frac{\nu(\rho \neq b_0)}{\#B - 1}\right) \\
&\leq \varepsilon + \nu(\rho \neq b_0)\log(\#B) + \varphi(\nu(\rho \neq b_0)) \\
&\leq \varepsilon(1 + \log(\#B) + \log(\varepsilon^{-1})),
\end{aligned}$$

because φ is increasing on $]0, e^{-1}[$. □

Proposition 3.3.4. *Let ξ denote the coordinate process on $\mathbf{Y}$. For every $n \geq 0$, we have*

$$h_\nu(\tau^n \xi, S) \leq 3\log(2)2^{-n/2}.$$

Proof. Let $n \geq 0$. One can see that $\tau^n\xi$ is 1 at i if and only if ξ is 1 over the entire segment $[i, i+n]$, as shown below:



We set $k_n := \lceil n/2 \rceil$, and we remark that

$$\nu(\{\exists i \in [0, k_n], (\tau^n\xi)_i = 1\}) \leq \nu(\{\xi_{[k_n, n]} = (1, ..., 1)\}) \leq 1/2^{n - k_n + 1} \leq 1/2^{n/2}.$$

Then, combining this with Lemma **??** we get

$$H_\nu((\tau^n\xi)_{[0,k_n]}) \leq 2^{-n/2}(1 + \log(2^{k_n+1}) + \log(2^{n/2})) \leq 2^{-n/2}3(k_n + 1)\log(2),$$

and we can conclude for the KS-entropy:

$$h_\nu(\tau^n\xi, S) \leq \frac{1}{k_n+1} H_\nu((\tau^n\xi)_{[0,k_n]}) \leq 3\log(2)2^{-n/2}.$$

□

In addition, we give the following simple lemma on conditional independence:

Lemma 3.3.5. *Let $(X, \mathscr{A}, \mu)$ be a probability space and $\mathscr{Z}$ a sub-σ-algebra. Let A, B, U and V be random variables such that*

$$(A, U) \perp\!\!\!\perp_{\mathscr{Z}} (B, V).$$

Then we have

$$\begin{aligned}\mathcal{L}(A, B \,|\, U, V, \mathscr{Z}) &= \mathcal{L}(A \,|\, U, \mathscr{Z}) \otimes \mathcal{L}(B \,|\, V, \mathscr{Z}) \\ &= \mathcal{L}(A \,|\, U, V, \mathscr{Z}) \otimes \mathcal{L}(B \,|\, U, V, \mathscr{Z})\end{aligned}$$

Proof. It follows from the fact that if A', B', U' and V' are respectively A, B, U and V-measurable random variables:

$$\begin{aligned}\mathbb{E}[A' \cdot B' \cdot U' \cdot V' \,|\, \mathscr{Z}] &= \mathbb{E}[A' \cdot U' \,|\, \mathscr{Z}] \cdot \mathbb{E}[B' \cdot V' \,|\, \mathscr{Z}] \\ &= \mathbb{E}[\mathbb{E}[A' \,|\, U, \mathscr{Z}] \cdot U' \,|\, \mathscr{Z}] \cdot \mathbb{E}[\mathbb{E}[B' \,|\, V, \mathscr{Z}] \cdot V' \,|\, \mathscr{Z}] \\ &= \mathbb{E}[\mathbb{E}[A' \,|\, U, \mathscr{Z}] \cdot U' \cdot \mathbb{E}[B' \,|\, V, \mathscr{Z}] \cdot V' \,|\, \mathscr{Z}].\end{aligned}$$

□

Proposition 3.3.6. *Let ξ be the coordinate process on $\mathbf{Y}$. For every $n \geq 1$, ξ is relatively very weak Bernoulli over $\tau^n\xi$.*

Proof. Set $\eta := \tau^n\xi$. Relative very weak Bernoullicity was defined in Definition **??**. We recall some notation: take $\lambda \in \mathscr{P}(\{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}})$ to be the law of (η, ξ), and for $I, J \subset \mathbb{Z}$ and $a, b \in \{0,1\}^{\mathbb{Z}}$, $\lambda_\ell(\cdot \,|\, a_I, b_J)$ is the conditional law of $\xi_{[0,\ell[}$ given that $\eta_I = a_I$ and $\xi_J = b_J$.

Let $\varepsilon > 0$. We need to show that there exists $\ell \geq 1$ such that for every $m \geq 1$ and for $k \geq 1$ large enough, we have

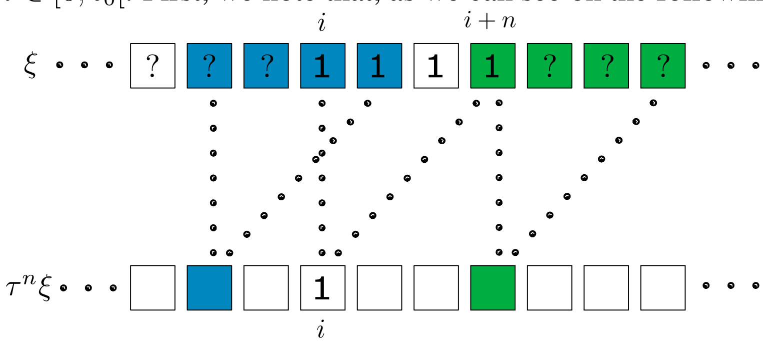
$$\int \bar{d}_\ell \left(\lambda_\ell(\cdot \,|\, a_{[-k,k]}, b_{[-m,0]}), \lambda_\ell(\cdot \,|\, a_{[-k,k]})\right) d\lambda(a,b) \leq \varepsilon. \qquad (3.26)$$

Let $m \geq 1$. We start by noting that there must be some «1» that appears in η: indeed, the law of large numbers tells us that there exists $\ell_0 \geq 1$ such that

$$\mu(\underbrace{\{\exists i \in [0, \ell_0[\,;\ \eta_i = 1\}}_{:=A}) \geq 1 - \varepsilon. \tag{3.27}$$

We then set $\ell := \lceil \frac{1}{\varepsilon} \rceil \ell_0$. Next, we take $k \geq \ell_0$ so that $\eta_{[-k,k]}$ determines entirely A.

We fix $i \in [0, \ell_0[$. First, we note that, as we can see on the following image



if $\eta_i = 1$, then $(\xi_{]-\infty,i[}, \eta_{]-\infty,i[})$ is $\xi_{]-\infty,i[}$-measurable and $(\xi_{]i,\infty[}, \eta_{]i,\infty[})$ is $\xi_{]i+n,\infty[}$-measurable. So, since the variables $\{\xi_j\}_{j\in\mathbb{Z}}$ are independent, given $\{\eta_i = 1\}$ the variables $(\xi_{]-\infty,i[}, \eta_{]-\infty,i[})$ and $(\xi_{]i,\infty[}, \eta_{]i,\infty[})$ are independent. Finally, using Lemma **??**, for $a \in A^{\mathbb{Z}}$ such that $a_i = 1$, we get:

$$\begin{aligned}\mathcal{L}(\xi_{]-\infty,i[}, \xi_{]i,\infty[} \,|\, \eta_{[-k,k]} = a_{[-k,k]}) &= \mathcal{L}(\xi_{]-\infty,i[} \,|\, \eta_{[-k,i[} = a_{[-k,i[}, \eta_i = 1) \\ &\quad \otimes \mathcal{L}(\xi_{]i,\infty[} \,|\, \eta_{]i,k]} = a_{]i,k]}, \eta_i = 1) \\ &= \mathcal{L}(\xi_{]-\infty,i[} \,|\, \eta_{[-k,k]} = a_{[-k,k]}) \otimes \mathcal{L}(\xi_{]i,\infty[} \,|\, \eta_{]i,k]} = a_{[-k,k]}).\end{aligned}$$

Therefore, if $\eta_{[-k,k]}$ is chosen so that there exists $i \in [0, \ell_0[$ such that $\eta_i = 1$, we see that $\xi_{[-m,0[}$ and $\xi_{[\ell_0,\ell[}$ are independent given $\eta_{[-k,k]}$.

We are now ready to prove (**??**). For any $b \in \{0,1\}^{\mathbb{Z}}$ and any $a \in \{0,1\}^{\mathbb{Z}}$ such that there exists $i \in [0, \ell_0[$ such that $a_i = 1$, the fact that $\xi_{[-m,0]}$ and $\xi_{[\ell_0,\ell[}$ are relatively independent given $\{\eta_{[-k,k]} = a_{[-k,k]}\}$ implies that the measures $\lambda_\ell(\cdot \,|\, a_{[-k,k]}, b_{[-m,0]})$ and $\lambda_\ell(\cdot \,|\, a_{[-k,k]})$ have the same marginal on the coordinates of $[\ell_0, \ell[$. So the relative product of those measures over $\xi_{[\ell_0,\ell[}$ is a coupling under which the copies of $\xi_{[\ell_0,\ell[}$ coincide. It follows that

$$\bar{d}_\ell \left(\lambda_\ell(\cdot \,|\, a_{[-k,k]}, b_{[-m,0]}), \lambda_\ell(\cdot \,|\, a_{[-k,k]})\right) \leq \ell_0/\ell \leq \varepsilon. \tag{3.28}$$

By combining (**??**) and (**??**), we can conclude that

$$\int \bar{d}_\ell \left(\lambda_\ell(\cdot \mid a_{[-k,k]}, b_{[-m,0]}), \lambda_\ell(\cdot \mid a_{[-k,k]})\right) d\lambda(a,b) \leq 2\varepsilon.$$

□

Proof of Theorem **??**. First of all, (**??**) follows directly from Proposition **??**. Next, from Proposition **??**, it follows that $\mathscr{F}_0$ is relatively very weak Bernoulli over $\mathscr{F}_n$, so $\mathscr{F}_{n+1}$ is relatively very weak Bernoulli over $\mathscr{F}_n$ (by part (iii) of Lemma **??**), so $\mathscr{F}_{n+1}$ is relatively Bernoulli over $\mathscr{F}_n$ (by part (i) of Lemma **??**). □

3.3.2 A cellular automaton on Ornstein's K-process

Here, we consider the non-Bernoulli K-system introduced by Ornstein in [**?**]. A more detailed presentation of this system is given in [**?**, Part III], but we give a sketch of the construction for completeness. It is a process defined on the alphabet $\{0, e, f, s\}$. We set $h(r)$, $s(r)$ and $f(r)$ to be integers depending on $r \in \mathbb{N}$ used in the construction of the process. For $r \geq 1$, an r-block is a random sequence of length $h(r)$ on the alphabet $\{0, e, f, s\}$, whose law we define inductively.

To get a 1-block, take $k_1 \in [\![1, f(1) - 1]\!]$ chosen uniformly at random, and consider a sequence that starts with a string of k_1 «f», followed by a string of $h(0)$ «0», and ends with a string of $f(1) - k_1$ «e»:

$$\underbrace{f \cdots\cdots f}_{k_1}\underbrace{0 \cdots\cdots 0}_{h(0)}\underbrace{e \cdots\cdots e}_{f(1)-k_1}$$

This construction implies that $h(1) = h(0) + f(1)$.

To get an r-block, take $k_r \in [\![1, f(r) - 1]\!]$ chosen uniformly at random, and 2^r i.i.d. random variables $(\xi_i^{(r-1)})_{i \in [\![1,2^r]\!]}$ such that each $\xi_i^{(r-1)}$ is an $(r-1)$-block. The r-block is then built as follows:

$$\underbrace{f \cdots f}_{k_r}\underbrace{s \cdots s}_{s(r)}\ \xi_1^{(r-1)}\ \underbrace{s \cdots s}_{2s(r)}\ \xi_2^{(r-1)}\ \cdots\cdots\ \underbrace{s \cdots s}_{2^r s(r)}\ \xi_{2^r}^{(r-1)}\ \underbrace{s \cdots s}_{(2^r+1)s(r)}\underbrace{e \cdots e}_{f(r)-k_r}$$

So an r-block starts with a string of k_r «f», and ends with a string of $f(r) - k_r$ «e». In between, we put all the $(r-1)$-blocks separated by strings of «s» so that each $\xi_i^{(r-1)}$ is placed in between two strings of «s» of respective lengths $is(r)$ and $(i+1)s(r)$. In particular, $h(r)$ is entirely determined by $h(r-1)$, $f(r)$ and $s(r)$.

Ornstein's K-system is then built by constructing an increasing sequence of towers $(\mathcal{T}_r)_{r\geq 1}$ such that $X := \bigcup_{r\geq 1} \mathcal{T}_r$. A tower $\mathcal{T}_r$ is given by its base F_r for which the sets $\{T^i F_r\}_{i\in[0,h(r)[}$ are disjoint and

$$\mathcal{T}_r := \bigsqcup_{i=0}^{h(r)-1} T^i F_r.$$

Through a cutting and stacking method, Ornstein builds in [?] the towers $(\mathcal{T}_r)_{r\geq 1}$ along with a process ξ so that the law of $\xi_{[0,h(r)[}$ given F_r is the law of an r-block. In other words, this means that the columns of the form

$$C_{\mathbf{a}} := \bigsqcup_{i=0}^{h(r)-1} T^i(F_r \cap \{\xi_{[0,h(r)[} = \mathbf{a}\}), \text{ for } \mathbf{a} \in \{0, e, f, s\}^{h(r)},$$

partition $\mathcal{T}_r$ according to the law of an r-block. Denote $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ the resulting dynamical system. A proper choice of $h(r)$, $s(r)$ and $f(r)$ assures that this construction gives a finite measure. Then ξ is a factor map onto the system

$$\mathbf{Y} := (\{0, e, f, s\}^{\mathbb{Z}}, \mathscr{B}, \nu, S),$$

where ν is the law of ξ.

Since ξ is a process on the alphabet $\{0, e, f, s\}$, the local function (**??**) becomes:

$$\begin{array}{rccc} \tau_0: & \{0,e,f,s\}^2 & \longrightarrow & \{0,e,f,s\} \\ & (\alpha_1,\alpha_2) & \mapsto & \begin{cases} \alpha_1 & \text{si } \alpha_1 = \alpha_2 \\ 0 & \text{otherwise.} \end{cases} \end{array}$$

From now on, τ denotes the corresponding cellular automaton. Similarly to what we did in Section **??**, we prove

Theorem 3.3.7. *On the system $\mathbf{Y} := (\{0, e, f, s\}^{\mathbb{Z}}, \mathscr{B}, \nu, S)$, the filtration given by $\mathscr{F} := (\sigma(\tau^{|n|}))_{n\leq 0}$ is a weak Pinsker filtration. That is, for every $n \leq -1$, $\mathscr{F}_{n+1}$ is relatively Bernoulli over $\mathscr{F}_n$ and we have*

$$h_\nu(\mathscr{F}_n) \underset{n\to-\infty}{\longrightarrow} 0. \tag{3.29}$$

The overall structure of the proof will resemble Section **??**, but the details are adapted to the specific structure of Ornstein's process. First, the convergence to 0 of the entropy follows from Proposition **??**. We could also adapt the proof of Proposition **??** to get that convergence, but it does not give a better rate of convergence than Proposition **??**, so we do not give any details.

Proposition 3.3.8. *If ξ is the process defined above, then for every $n \geq 1$, ξ is relatively very weak Bernoulli over $\tau^n\xi$.*

Proof. We set $\eta := \tau^n\xi$. Let $\varepsilon > 0$. Once again, we need to show that there exists $\ell \geq 1$ such that for every $m \geq 1$ and for $k \geq 1$ large enough, we have

$$\int \bar{d}_\ell \left(\lambda_\ell(\cdot \,|\, a_{[-k,k]}, b_{[-m,0]}), \lambda_\ell(\cdot \,|\, a_{[-k,k]})\right) d\lambda(a,b) \leq \varepsilon,$$

where λ is the law of (η, ξ) and, for $I, J \subset \mathbb{Z}$, $\lambda_\ell(\cdot \,|\, a_I, b_J)$ is the conditional law of $\xi_{[0,\ell[}$ given that η_I equals a_I and that ξ_J equals b_J.

Let $m \geq 1$. We choose r so that $s(r+1) \geq n+1$. By construction of ξ, we know that for any r-block in ξ, there exists $i \in [1, 2^{r+1}]$ such that the said r-block will come after a string of $i \cdot s(r+1)$ «s» and be followed by a string of $(i+1) \cdot s(r+1)$ «s». Therefore, by knowing the positions of all the strings of «s» longer that $s(r+1)$, we know the position of every r-block.

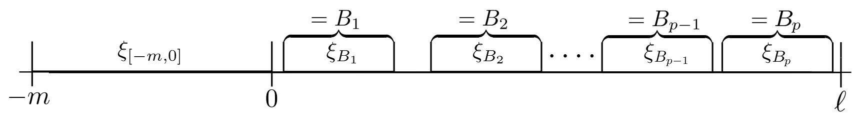
However, since we chose to have $s(r+1) \geq n+1$, we can say that, for $k \in \mathbb{Z}$, we have $\xi_{[k,k+s(r+1)[} = (s, ..., s)$ if and only if $\eta_{[k,k+s(r+1)-n[} = (s, ..., s)$. This means that the positions of the r-blocks contained on a segment $[k_1, k_2]$ are $\eta_{[k_1-N,k_2+N]}$-measurable, for N large enough (for example $N = (2^{r+1}+1)s(r+1)$).

By choosing r large enough, we can also assume that $\mu(\mathcal{T}_r) \geq 1 - \varepsilon/2$. Using Birkhoff's ergodic theorem, for ℓ large enough, the set

$$A := \left\{ x \in X;\ \frac{1}{\ell}\sum_{j=0}^{\ell-1} \mathbb{1}_{\mathcal{T}_r}(T^j(x)) > 1 - \varepsilon \right\},$$

satisfies $\mu(A) > 1 - \varepsilon$.

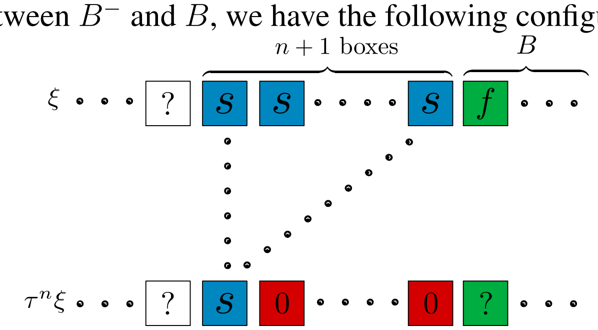
In other words, for $x \in A$, the number of elements in the sequence $\xi_{[0,\ell[}(x)$ that are part of an r-block is greater than $(1-\varepsilon)\ell$. However, among the intervals on which those r-blocks are supported, two of them may not be included in $[0, \ell[$, and can intersect $\mathbb{Z} \backslash [0, \ell[$. But, if $h(r)/\ell \leq \varepsilon/2$, there are at most $\varepsilon\ell$ elements in those two intervals. To sum up, we get that the number of elements in the sequence $\xi_{[0,\ell[}(x)$ that are part of an r-block, and for which the position of that r-block is contained on the segment $[0, \ell[$, is greater than $(1-2\varepsilon)\ell$. Then, we choose $k \geq 1$ so that the positions of the r-blocks contained in $[-m, \ell[$ are $\eta_{[-k,k]}$-measurable (in particular, A is $\eta_{[-k,k]}$-measurable). So we have the following configuration for $\xi_{[-m,\ell[}$:



where the B_i are the positions of the r-blocks supported on $[0, \ell[$, and we have shown that $\# \bigsqcup_{i=1}^{p} B_i \geq (1 - 2\varepsilon)\ell$.

Denote by $I_\ell := \{B_i\}_{1 \leq i \leq p}$ the random variable that gives the positions of the r-blocks on the segment $[0, \ell[$. By construction of ξ, we know that, given I_ℓ, for any r-block B, the variables ξ_B and ξ_{B^c} are independent. Moreover, we know that any r-block is between two strings of at least $n + 1$ «s». Therefore, we see that if I_ℓ is fixed, for any r-block B, η_B is ξ_B-measurable and η_{B^c} is ξ_{B^c}-measurable.

Let us give details on the proof of that last claim: we write B^c as the union of B^- and B^+, the infinite intervals that come before and after B respectively. Given the structure of our automaton, it is always true that η_{B^+} is ξ_{B^+}-measurable. At the boundary between B^- and B, we have the following configuration:



Indeed, in the construction of the blocks, we see that ξ must put an «f» in the first box of B. Therefore, we must have «0» in the red boxes. So, the values that η takes on the $n + 1$ boxes preceding B are determined. For the rest of the boxes of B^-, it comes from the structure of τ that the values of η are determined by ξ_{B^-} since we are at a distance $\geq n + 1$ from B. So we have shown that η_{B^-} is (ξ_{B^-})-measurable. A similar reasoning at the boundary between B and B^+ shows that η_B is ξ_B-measurable. And since it is always true that η_{B^+} is ξ_{B^+}-measurable, we have proven that η_B is ξ_B-measurable and η_{B^c} is ξ_{B^c}-measurable.

But, we also know from the structure of ξ that, given I_ℓ, ξ_B and ξ_{B^c} are independent. The previous paragraph enables us to use Lemma **??** to extend that to: given $I_\ell \vee \eta_{[-k,k]}$, ξ_B and ξ_{B^c} are independent. Finally, since I_ℓ is $\eta_{[-k,k]}$-measurable, this yields that ξ_B and ξ_{B^c} are relatively independent given $\eta_{[-k,k]}$.

This independence tells us that, for every sequences a and b, $\lambda_\ell(\cdot \,|\, a_{[-k,k]}, b_{[-m,0]})$ and $\lambda_\ell(\cdot \,|\, a_{[-k,k]})$ have the same marginals on the coordinates of the r-blocks B contained in $[0, \ell[$. Moreover, if a is chosen so that $\{\eta_{[-k,k]} = a_{[-k,k]}\}$ is a subset of A, we know that the positions of the r-blocks cover at least $(1 - 2\varepsilon)\ell$ elements in $[0, \ell[$. Then, by considering the relative product of $\lambda_\ell(\cdot \,|\, a_{[-k,k]}, b_{[-m,0]})$ and $\lambda_\ell(\cdot \,|\, a_{[-k,k]})$ over $\{\xi_{B_i}\}_{1 \leq i \leq p}$, we get:

$$\bar{d}_\ell \left(\lambda_\ell(\cdot \,|\, a_{[-k,k]}, b_{[-m,0]}), \lambda_\ell(\cdot \,|\, a_{[-k,k]}) \right) \leq 2\varepsilon.$$

Finally, since $\mu(A) \geq 1 - \varepsilon$, this yields

$$\int \bar{d}_\ell \left(\lambda_\ell(\cdot \,|\, a_{[-k,k]}, b_{[-m,0]}), \lambda_\ell(\cdot \,|\, a_{[-k,k]}) \right) d\nu(a,b) \leq 3\varepsilon.$$

□

Remark 3.3.9. We see that the proofs of Theorem **??** and Theorem **??** are very similar. In both cases, we have a process ξ, whose conditional law given $\tau^n \xi$ is made of random blocks separated by deterministic blocks, and the random blocks are filled independently from each other. The main difference that prevents Ornstein's K-process from being Bernoulli is that the position of r-blocks is determined by the long sequences of «s», and this creates correlations over long distances. But once we condition by $\tau^n \xi$, those sequences of «s» are entirely determined. Therefore we are left with filing independently all the r-blocks, and the past has no longer a significant influence on the future.

In that sense, when we look at the relative structure of Ornstein's K-process over τ^n, the non-Bernoulli aspects disappear. However, when we look at the asymptotic properties of the weak Pinsker filtration obtained by applying $\{\tau^n\}_{n \geq 1}$, whether we start with a Bernoulli process or with a non-Bernoulli K-process, we get different results. Therefore, getting a better understanding of the classification of the various properties of weak Pinsker filtrations could help to develop a new classification of non-Bernoulli K-systems.

Abstract

In this thesis, we explore the possible structures of measure preserving dynamical systems of the form $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ and their factor σ-algebras $\mathscr{B} \subset \mathscr{A}$.

The first two chapters investigate various ways in which a factor σ-algebra $\mathscr{B}$ can sit in a dynamical system $\mathbf{X} := (X, \mathscr{A}, \mu, T)$, i.e. we study some possible structures of the *extension* $\mathscr{A} \to \mathscr{B}$. In the first chapter, we consider the concepts of *super-innovations* and *standardness* of extensions, which are inspired from the theory of filtrations. An important focus of our work is the introduction of the notion of *confined extensions*, which first interested us because they have no super-innovation. We give several examples and study additional properties of confined extensions, including several lifting results. Then, we show our main result: the existence of non-standard extensions. Finally, this result finds an application to the study of dynamical filtrations, i.e. filtrations of the form $(\mathscr{F}_n)_{n \leq 0}$ such that each $\mathscr{F}_n$ is a factor σ-algebra. We show that there exist *non-standard I-cosy dynamical filtrations*.

The second chapter furthers the study of confined extensions by finding a new kind of such extensions, in the setup of Poisson suspensions: we take an infinite σ-finite measure-preserving dynamical system (X, μ, T) and a compact extension $(X \times G, \mu \otimes m_G, T_\varphi)$, then we consider the corresponding Poisson extension $((X \times G)^*, (\mu \otimes m_G)^*, (T_\varphi)_*) \longrightarrow (X^*, \mu^*, T_*)$. We give conditions under which that extension is confined and build an example which fits those conditions.

Lastly, the third chapter focuses on a family of dynamical filtrations: *weak Pinsker filtrations*. The existence of those filtrations on any ergodic system comes from a recent result by Austin [**?**], and they present themselves as a potential tool to describe positive entropy systems. We explore the links between the asymptotic structure of weak Pinsker filtrations and the properties of the underlying dynamical system. Naturally, we also ask whether, on a given system, the structure of weak Pinsker filtrations is unique up to isomorphism. We give a partial answer, in the case where the underlying system is Bernoulli. We conclude our work by giving two explicit examples of weak Pinsker filtrations.

Keywords: Confined extensions, compact extensions, joinings, Poisson suspensions, entropy, Bernoulli systems, K-systems, dynamical filtrations, weak Pinsker filtrations

Résumé

Dans cette thèse, nous explorons les structures possibles des systèmes dynamiques de la forme $\mathbf{X} := (X, \mathscr{A}, \mu, T)$ et leurs tribus facteur $\mathscr{B} \subset \mathscr{A}$.

Les deux premiers chapitres étudient les différentes façons dont une tribu facteur $\mathscr{B}$ peut s'inclure dans un système dynamique $\mathbf{X} := (X, \mathscr{A}, \mu, T)$, c'est-à-dire que nous étudions certaines structures possibles de l'*extension* $\mathscr{A} \to \mathscr{B}$. Dans le premier chapitre, nous considérons les concepts de *super-innovations* et de *standardité* des extensions, inspirés de la théorie des filtrations. Un point important est l'introduction de la notion d'*extensions confinées*, qui nous intéressent parce qu'elles n'ont pas de super-innovation. Nous donnons plusieurs exemples et étudions des propriétés supplémentaires de ces extensions, y compris des résultats de relèvement. Ensuite, nous montrons notre résultat principal : l'existence d'extensions non-standard. Enfin, ce résultat trouve une application dans l'étude des filtrations dynamiques, qui sont les filtrations de la forme $(\mathscr{F}_n)_{n \leq 0}$ telles que chaque $\mathscr{F}_n$ est une tribu facteur. Nous montrons qu'il existe des *filtrations dynamiques I-confortables non standard*.

Le deuxième chapitre approfondit l'étude des extensions confinées en trouvant un nouveau type de telles extensions, dans le cadre des suspensions de Poisson : nous prenons un système dynamique (X, μ, T) en mesure σ-finie infinie et une extension compacte $(X \times G, \mu \otimes m_G, T_\varphi)$, puis nous considérons l'extension de Poisson correspondante $((X \times G)^*, (\mu \otimes m_G)^*, (T_\varphi)_*) \longrightarrow (X^*, \mu^*, T_*)$. Nous donnons des conditions sous lesquelles cette extension est confinée et construisons un exemple qui correspond à ces conditions.

Enfin, le troisième chapitre se concentre sur une famille de filtrations dynamiques : les *filtrations de Pinsker faible*. L'existence de ces filtrations sur tout système ergodique provient d'un résultat récent d'Austin [**?**], et elles se présentent comme un outil potentiel pour décrire les systèmes à entropie positive. Nous explorons les liens entre la structure asymptotique des filtrations de Pinsker faible et les propriétés du système dynamique sous-jacent. Naturellement, nous demandons aussi si, sur un système donné, la structure des filtrations de Pinsker faible est unique à isomorphisme près. Nous donnons une réponse partielle, dans le cas où le système sous-jacent est un schéma de Bernoulli. Nous concluons notre travail en donnant deux exemples explicites de filtrations de Pinsker faible.

Mots clés: Extensions confinées, extensions compactes, couplages, suspensions de Poisson, entropie, schémas de Bernoulli, K-systèmes, filtrations dynamiques, filtrations de Pinsker faible