ASYMPTOTIC PAIRS IN POSITIVE-ENTROPY SYSTEMS.

F. BLANCHARD, B. HOST, AND S. RUETTE

Abstract. We show that in a topological dynamical system (X, T) of positive entropy there exist proper (positively) asymptotic pairs, that is, pairs (x, y) such that $x \neq y$ and $\lim_{n\to+\infty} d(T^n x, T^n y) = 0$. More precisely we consider a T-ergodic measure μ of positive entropy and prove that the set of points that belong to a proper asymptotic pair is of measure 1. When T is invertible, the stable classes (i.e., the equivalence classes for the asymptotic equivalence) are not stable under T^{-1}: for μ-almost every x there are uncountably many y that are asymptotic with x and such that (x, y) is a Li-Yorke pair with respect to T^{-1}. We also show that asymptotic pairs are dense in the set of topological entropy pairs.

1. Introduction

In this article a *topological dynamical system* is a compact metric space X endowed with a homeomorphism $T : X \to X$, except in subsection 3.3 where we drop the assumption that T is invertible; the distance on X is denoted by d.

Classically in Topological Dynamics one considers the asymptotic behaviour of pairs of points. In this article, even when the systems considered are invertible, the definitions of asymptoticity, proximality and Li-Yorke pairs that we use are those fitted to an $\mathbb{N}$-action. A pair $(x, y) \in X \times X$ is said to be *proximal* if $\liminf_{n\to+\infty} d(T^n x, T^n y) = 0$, and (x, y) is called an *asymptotic pair* if $\lim_{n\to+\infty} d(T^n x, T^n y) = 0$; the set of asymptotic pairs is denoted by $\mathbf{A}$. An asymptotic pair (x, y) with $x \neq y$ is said to be *proper*. Asymptoticity is an equivalence relation; the equivalence class of a point is called its *stable class*. We call a proximal pair that is not asymptotic a *Li-Yorke pair*: in 1975 Li and Yorke introduced such pairs in a tentative definition of chaos [16].

It is proven in [3] that positive entropy implies the existence of a topologically 'big' set of Li-Yorke pairs. Here we prove by ergodic methods that in any topological dynamical system with positive topological entropy there is a measure-theoretically 'rather big' set of proper asymptotic pairs; this is obvious for a symbolic system but not in general. The set of asymptotic pairs of any topological dynamical system has been shown to be first category in [14]: it is a small set, but not too small according to the present result. We also show that a 'rather big' set of T-asymptotic pairs are Li-Yorke under the action of T^{-1}.

In [13] Huang and Ye construct a completely scrambled system, that is to say, a dynamical system (X, T) such that all proper pairs in $(X \times X)$ are Li-Yorke. They ask whether such a system may have positive entropy. That it may not is a direct consequence of our Proposition 1. This statement formally generalizes a previous result of Weiss [22], showing that any system (X, T) such that $(X \times X, T \times T)$ is recurrent has entropy 0; recurrence of $(X \times X, T \times T)$ means that any pair (x, y),

Ergod. Th. & Dynam. Syst., **22**, 671-686, 2002.

$x \neq y$, comes back arbitrarily close to itself under the action of powers of T, which implies that it cannot be asymptotic.

Then we study the behaviour of T-asymptotic pairs under T^{-1}. Anosov diffeomorphisms on a manifold have stable and unstable foliations; points belonging to the same stable foliation are asymptotic under T and tend to diverge under T^{-1}, while pairs belonging to the unstable foliation behave the opposite way. Our results show that any positive-entropy system retains a faint flavour of this situation: there is a universal $\delta > 0$ such that outside a 'small' set the stable class of x is non-empty and contains an uncountable set of points y such that $\limsup_{n\to+\infty} d(T^{-n}x, T^{-n}y) \geq \delta$.

We also obtain a result about entropy pairs [4]: the set of asymptotic pairs $\mathbf{A}$ is dense in the set of entropy pairs $E(X,T)$. The proof relies on two facts: that the union of the sets of μ-entropy pairs for all ergodic measures μ is dense in the set of topological entropy pairs [2], and the characterization of the set $E_\mu(X,T)$ of μ-entropy pairs as the support of some measure on $X \times X$ [12].

The article is organized as follows. Section 2 contains some background in Ergodic Theory, in particular the old but not very familiar definition of an *excellent partition.* In Section 3 using an ad-hoc excellent partition we show that every system of positive entropy admits 'many' asymptotic pairs, and that this is also true for non-invertible systems. In the next section, after recalling the definition of the relative independent square of a measure, we use this notion to show that asymptotic pairs are dense in the set of entropy pairs. In Section 5, we show that a system of positive entropy has 'many' pairs that are asymptotic for T and Li-Yorke for T^{-1}. In the last section we show that the sets constructed above are uncountable.

Some results are stated several times in increasingly strong form; Propositions 5 and 6 are strongest. We chose this organization in order to avoid a long preliminary section containing all the required background. Most tools are introduced just before the statements that require them for their proofs.

A final remark about the methods. It is not very satisfactory to prove a purely topological result – the existence of many asymptotic pairs in any positive-entropy topological dynamical system – in a purely ergodic way. Proving it topologically is a good challenge. On the other hand Ergodic Theory is a powerful tool; it is not the first time that it demonstrates its strength in a neighbouring field. Here it also permits to prove results that are probabilistic in nature.

We are grateful to X.D. Ye for providing the initial motivation, to W. Huang and him for several valuable observations and to S. Kolyada for various interesting remarks. The referee made significant remarks and corrected many English mistakes.

2. Background

Here are some classical definitions and results from Ergodic Theory, and some technical Lemmas that will be needed in the sequel.

A measure-theoretic dynamical system $(X, \mathcal{A}, T, \mu)$ is a Lebesgue probability space $(X, \mathcal{A}, \mu)$ endowed with a measurable transformation $T\colon X \to X$ which preserves μ. In this article unless stated otherwise T is assumed to be one-to-one and bi-measurable. The σ-algebra $\mathcal{A}$ is assumed to be complete for μ. All measures are assumed to be probability measures; since quasi-invariant measures are not considered in this article, an ergodic measure is always assumed to be invariant.

2.1. **Partitions.** All partitions of X are assumed to consist of atoms belonging to the σ-algebra $\mathcal{A}$. Given a partition $\mathcal{P}$ of X and $x \in X$, denote by $\mathcal{P}(x)$ the atom of $\mathcal{P}$ containing x.

If $(\mathcal{P}_i; i \in I)$ is a countable family of finite partitions, the partition $\mathcal{P} = \bigvee_{i\in I} \mathcal{P}_i$ is called a *measurable partition* [17]. The sets $A \in \mathcal{A}$ which are union of atoms of $\mathcal{P}$ form a sub-σ-algebra of $\mathcal{A}$ denoted by $\sigma(\mathcal{P})$ or $\mathcal{P}$ if there is no ambiguity. Every sub-σ-algebra of $\mathcal{A}$ coincides with a σ-algebra constructed in this way outside a set of measure 0.

A sub-σ-algebra $\mathcal{F}$ of $\mathcal{A}$ which is T-invariant, that is, $T^{-1}\mathcal{F} = \mathcal{F}$, is called a *factor*. Equivalently, a factor is given by a measure-theoretical system $(Y, \mathcal{B}, S, \nu)$ and a measurable map $\varphi\colon X \to Y$ such that $\varphi \circ T = S \circ \varphi$; the corresponding T-invariant sub-σ-algebra of $\mathcal{A}$ is $\varphi^{-1}\mathcal{B}$.

Given a measurable partition $\mathcal{P}$, put $\mathcal{P}^- = \bigvee_{n=1}^{\infty} T^{-n}\mathcal{P}$ and $\mathcal{P}^T = \bigvee_{n=-\infty}^{+\infty} T^{-n}\mathcal{P}$. Define in the same way $\mathcal{F}^-$ and $\mathcal{F}^T$ if $\mathcal{F}$ is a sub-σ-algebra of $\mathcal{A}$. The measurable partition $\mathcal{P}$ (resp. the sub-σ-algebra $\mathcal{F}$) is called *generating* if $\sigma(\mathcal{P}^T)$ (resp. $\mathcal{F}^T$) is equal to $\mathcal{A}$.

2.2. **Entropy.** For the definition of the conditional entropy $H(\mathcal{P} \mid \mathcal{F})$ of a finite measurable partition $\mathcal{P}$ with respect to the sub-σ-algebra $\mathcal{F}$, of the entropy $h_\mu(\mathcal{P}, T) = H(\mathcal{P} \mid \mathcal{P}^-)$ of a partition $\mathcal{P}$ with respect to T and of the entropy $h_\mu(X, T)$, refer to [17], [18], [21].

The *Pinsker factor* Π_μ of $(X, \mathcal{A}, T, \mu)$ is the maximal factor with entropy 0; a finite partition $\mathcal{P}$ is measurable with respect to Π_μ if and only if $h_\mu(\mathcal{P}, T) = 0$.

We do not give the proofs of the next two results; they can be found in [17].

Lemma 1. *If $\mathcal{F}$ is a generating sub-σ-algebra then $\Pi_\mu \subset \mathcal{F}^-$.*

Pinsker Formula. *For any finite partitions $\mathcal{P}$ and $\mathcal{Q}$ one has*

$$H(\mathcal{Q} \vee \mathcal{P} \mid \mathcal{Q}^- \vee \mathcal{P}^-) - H(\mathcal{P} \mid \mathcal{P}^-) = H(\mathcal{Q} \mid \mathcal{Q}^- \vee \mathcal{P}^T) \ . \tag{1}$$

The next technical Lemma compares the entropy of a partition with the conditional entropy of this partition with respect to the past of another.

Lemma 2. *Let $(X, \mathcal{A}, T, \mu)$ be a measure-theoretic dynamical system, and let $\mathcal{P}_1 \prec \mathcal{P}_2 \prec \cdots \prec \mathcal{P}_k$ be finite partitions. Then*

$$H(\mathcal{P}_1 \mid \mathcal{P}_1^-) - H(\mathcal{P}_1 \mid \mathcal{P}_2^-) = H(\mathcal{P}_2 \mid \mathcal{P}_1 \vee \mathcal{P}_2^-) - H(\mathcal{P}_2 \mid \mathcal{P}_1^T \vee \mathcal{P}_2^-) \tag{2}$$

and

$$H(\mathcal{P}_1 \mid \mathcal{P}_1^-) - H(\mathcal{P}_1 \mid \mathcal{P}_k^-) \le \sum_{i=1}^{k-1} \Big(H(\mathcal{P}_i \mid \mathcal{P}_i^-) - H(\mathcal{P}_i \mid \mathcal{P}_{i+1}^-) \Big) \ . \tag{3}$$

Proof. Obviously $\mathcal{P}_k = \mathcal{P}_1 \vee \cdots \vee \mathcal{P}_k$. A repeated use of the Pinsker Formula (1) yields

$$H(\mathcal{P}_k \mid \mathcal{P}_k^-) = H(\mathcal{P}_1 \mid \mathcal{P}_1^-) + H(\mathcal{P}_2 \mid \mathcal{P}_2^- \vee \mathcal{P}_1^T) + \cdots + H(\mathcal{P}_k \mid \mathcal{P}_k^- \vee \mathcal{P}_{k-1}^T) \ ;$$

also, using the elementary formula for conditional entropy of partitions inductively one gets

$$H(\mathcal{P}_k \mid \mathcal{P}_k^-) = H(\mathcal{P}_1 \mid \mathcal{P}_k^-) + H(\mathcal{P}_2 \mid \mathcal{P}_k^- \vee \mathcal{P}_1) + \cdots + H(\mathcal{P}_k \mid \mathcal{P}_k^- \vee \mathcal{P}_{k-1}) \ .$$

Combining these two equalities one obtains

$$H(\mathcal{P}_1 \mid \mathcal{P}_1^-) - H(\mathcal{P}_1 \mid \mathcal{P}_k^-) = \sum_{i=2}^{k} \big(H(\mathcal{P}_i \mid \mathcal{P}_k^- \vee \mathcal{P}_{i-1}) - H(\mathcal{P}_i \mid \mathcal{P}_i^- \vee \mathcal{P}_{i-1}^T) \big) \ .$$

For $k = 2$ this is (2).

For $k > 2$, remark that $\mathcal{P}_i^- \prec \mathcal{P}_k^-$ for $i \leq k$ so that $H(\mathcal{P}_i \mid \mathcal{P}_k^- \vee \mathcal{P}_{i-1}) \leq H(\mathcal{P}_i \mid \mathcal{P}_i^- \vee \mathcal{P}_{i-1})$, hence

$$H(\mathcal{P}_1 \mid \mathcal{P}_1^-) - H(\mathcal{P}_1 \mid \mathcal{P}_k^-) \leq \sum_{i=2}^{k} \big(H(\mathcal{P}_i \mid \mathcal{P}_i^- \vee \mathcal{P}_{i-1}) - H(\mathcal{P}_i \mid \mathcal{P}_i^- \vee \mathcal{P}_{i-1}^T) \big) \ .$$

Applying (2) (with $\mathcal{P}_{i-1}$ and $\mathcal{P}_i$ in place of $\mathcal{P}_1$ and $\mathcal{P}_2$) to each term in the sum, the inequality above becomes

$$H(\mathcal{P}_1 \mid \mathcal{P}_1^-) - H(\mathcal{P}_1 \mid \mathcal{P}_k^-) \leq \sum_{i=2}^{k} \big(H(\mathcal{P}_{i-1} \mid \mathcal{P}_{i-1}^-) - H(\mathcal{P}_{i-1} \mid \mathcal{P}_i^-) \big)$$

which is (3) up to a change of index. □

2.3. **Excellent partitions.** For any measure-theoretic dynamical system $(X, \mathcal{A}, T, \mu)$ there exists a generating measurable partition with the property that $\displaystyle\bigcap_{k=1}^{\infty} T^{-k}\mathcal{P}^- = \Pi_\mu$. In the finite-entropy case any finite generating partition has this property. The existence of such a partition in the general case was proven by Rohlin and Sinaĭ and permitted to show that the class of K-systems and the class of completely positive entropy systems coincide [19]; they gave a construction from which the one in Subsection 3.1 is derived. The name "excellent" was coined by one of the present authors in a later article.

Definition 1. *Let $(X, \mathcal{A}, T, \mu)$ be a measure-theoretic dynamical system. A measurable partition $\mathcal{P}$ is said to be* excellent *if it is generating and there is an increasing sequence of finite measurable partitions $(\mathcal{P}_n)_{n \geq 1}$ such that $\mathcal{P}_n \to \mathcal{P}$ and $H(\mathcal{P}_n \mid \mathcal{P}_n^-) - H(\mathcal{P}_n \mid \mathcal{P}^-) \to 0$ as $n \to \infty$.*

Lemma 3. [17] *If $\mathcal{P}$ is an excellent partition, then $\displaystyle\bigcap_{k=1}^{\infty} T^{-k}\mathcal{P}^- = \Pi_\mu$.*

Proof. Let $\mathcal{Q}$ be a finite partition, measurable with respect to $\bigcap_{k=1}^{\infty} T^{-k}\mathcal{P}^-$, and let the partitions $\mathcal{P}_n$ be as in Definition 1. Applying the Pinsker formula twice one obtains

$$\begin{aligned} H(\mathcal{Q} \mid \mathcal{Q}^-) &= H(\mathcal{P}_n \vee \mathcal{Q} \mid \mathcal{P}_n^- \vee \mathcal{Q}^-) - H(\mathcal{P}_n \mid \mathcal{P}_n^- \vee \mathcal{Q}^T) \\ &= H(\mathcal{P}_n \mid \mathcal{P}_n^-) + H(\mathcal{Q} \mid \mathcal{P}_n^T \vee \mathcal{Q}^-) - H(\mathcal{P}_n \mid \mathcal{P}_n^- \vee \mathcal{Q}^T) \ . \end{aligned}$$

When n goes to infinity $H(\mathcal{Q} \mid \mathcal{P}_n^T \vee \mathcal{Q}^-)$ goes to 0, since $\mathcal{P}_n^T$ tends to $\mathcal{P}^T = \mathcal{A}$; on the other hand we assumed that $T^n\mathcal{Q}$ is measurable with respect to $\mathcal{P}^-$ for $n \in \mathbb{Z}$, so $\mathcal{P}_n^- \vee \mathcal{Q}^T$ is contained in $\mathcal{P}^-$ and

$$0 \leq H(\mathcal{P}_n \mid \mathcal{P}_n^-) - H(\mathcal{P}_n \mid \mathcal{P}_n^- \vee \mathcal{Q}^T) \leq H(\mathcal{P}_n \mid \mathcal{P}_n^-) - H(\mathcal{P}_n \mid \mathcal{P}^-) \ .$$

By our assumption the majoration tends to 0. Thus $H(\mathcal{Q} \mid \mathcal{Q}^-) = 0$, which means that $\mathcal{Q}$ is coarser than the Pinsker σ-algebra. As this is true for any finite partition $\mathcal{Q}$ measurable with respect to $\bigcap_{k=1}^{\infty} T^{-k}\mathcal{P}^-$, one has $\bigcap_{k=1}^{\infty} T^{-k}\mathcal{P}^- \subset \Pi_\mu$.

The reverse inclusion is due to the fact that $\mathcal{P}$ is generating, so that $\Pi_\mu \subset \bigcap_{k=1}^{\infty} T^{-k}\mathcal{P}^-$ by Lemma 1. □

3. EXISTENCE OF ASYMPTOTIC PAIRS

Let (X,T) be a topological dynamical system, and let $\mathcal{B}$ be the Borel σ-algebra of X. Given two topological dynamical systems (X,T) and (Y,S) a continuous onto map $\pi : (X,T) \to (Y,S)$ such that $\pi \circ T = S \circ \pi$ is called a topological factor map.

The definitions of proximal, asymptotic and Li-Yorke pairs are given at the very beginning of the Introduction. Recall that $\mathbf{A}$ is the set of all asymptotic pairs in $X \times X$. See [21] for the definition of topological entropy, and for the

Variational Principle. *The topological entropy $h(X,T)$ of the system (X,T) is equal to the supremum of the entropies $h_\mu(X,\mathcal{B},T,\mu)$ where μ ranges over the set of ergodic T-invariant measures.*

3.1. Construction of an excellent partition.

The next Lemma establishes a connection between asymptotic pairs and entropy. It is our main tool. It is based on the construction of excellent partitions in [17].

Lemma 4. *Let μ be an ergodic measure on X.*

ı) The system $(X,\mathcal{B},T,\mu)$ admits an excellent partition $\mathcal{P}$, such that any pair of points belonging to the same atom of $\mathcal{P}^-$ is asymptotic.

ıı) Moreover, if $h_\mu(X,T) > 0$ then the σ-algebras $\mathcal{P}^-$ and $\mathcal{B}$ do not coincide up to sets of μ-measure 0.

Proof. ı) Let $(\mathcal{Q}_n)_{n\geq 1}$ be an increasing sequence of finite partitions such that the maximal diameter δ_n of an element of $\mathcal{Q}_n$ goes to 0 as $n \to \infty$, and $(\epsilon_n)_{n\geq 1}$ be a sequence of positive numbers such that $\sum_{i=1}^{\infty} \epsilon_n < \infty$.

We construct inductively an increasing sequence $(k_n)_{n\geq 1}$ of non-negative integers such that, if

$$\mathcal{P}_i = T^{-k_1}\mathcal{Q}_1 \vee T^{-k_2}\mathcal{Q}_2 \cdots \vee T^{-k_i}\mathcal{Q}_i$$

for $i \geq 1$ one has

$$H(\mathcal{P}_i \mid \mathcal{P}_i^-) - H(\mathcal{P}_i \mid \mathcal{P}_{i+1}^-) < \epsilon_i \ . \tag{4}$$

Put $k_1 = 0$ and $\mathcal{P}_1 = \mathcal{Q}_1$. Take $n \geq 2$, and suppose that the sequence is already defined up to k_{n-1} and the bound (4) holds for $1 \leq i \leq n-2$.

By Lemma 2 (2) one has for $k \geq 0$

$$\begin{aligned} D_k &\stackrel{\text{def}}{=} H(\mathcal{P}_{n-1} \mid \mathcal{P}_{n-1}^-) - H(\mathcal{P}_{n-1} \mid \mathcal{P}_{n-1}^- \vee T^{-k}\mathcal{Q}_n^-) \\ &= H(T^{-k}\mathcal{Q}_n \mid \mathcal{P}_{n-1} \vee \mathcal{P}_{n-1}^- \vee T^{-k}\mathcal{Q}_n^-) - H(T^{-k}\mathcal{Q}_n \mid \mathcal{P}_{n-1}^T \vee T^{-k}\mathcal{Q}_n^-) \ . \end{aligned}$$

By T-invariance of μ the second equality above becomes

$$D_k = H(\mathcal{Q}_n \mid T^{k+1}\mathcal{P}_{n-1}^- \vee \mathcal{Q}_n^-) - H(\mathcal{Q}_n \mid \mathcal{P}_{n-1}^T \vee \mathcal{Q}_n^-) \ ;$$

when k goes to infinity the conditioning σ-algebra in the first term tends to the conditioning σ-algebra in the second term, and the difference D_k tends to 0.

Fix k_n so that $D_{k_n} < \epsilon_n$, which, putting $\mathcal{P}_n = \mathcal{P}_{n-1} \vee T^{-k_n}\mathcal{Q}_n$, is Property (4) at rank $i = n-1$. Setting $\mathcal{P} = \bigvee_{n\in\mathbb{N}} \mathcal{P}_n$ completes our construction.

It remains to check that $\mathcal{P}$ is excellent.

By construction, $\mathcal{P}^T$ is finer than $\bigvee_{n\geq 1} \mathcal{Q}_n$, and this partition spans $\mathcal{B}$ because of our hypotheses on $(\mathcal{Q}_n)$. Thus $\mathcal{P}$ is generating.

The sequence $(\mathcal{P}_n)$ increases to $\mathcal{P}$; moreover

$$H(\mathcal{P}_n \mid \mathcal{P}_n^-) - H(\mathcal{P}_n \mid \mathcal{P}^-) = \lim_{k\to\infty} \big(H(\mathcal{P}_n \mid \mathcal{P}_n^-) - H(\mathcal{P}_n \mid \mathcal{P}_{n+k}^-)\big) \ ,$$

and by Lemma 2 (3) one gets

$$H(\mathcal{P}_n \mid \mathcal{P}_n^-) - H(\mathcal{P}_n \mid \mathcal{P}^-) \leq \sum_{i=n}^{\infty} \big(H(\mathcal{P}_i \mid \mathcal{P}_i^-) - H(\mathcal{P}_i \mid \mathcal{P}_{i+1}^-)\big) < \sum_{i=n}^{\infty} \epsilon_i \ , \tag{5}$$

a quantity which vanishes as $n \to \infty$: the second condition for excellence of $\mathcal{P}$ holds.

Let x, y belong to the same atom of $\mathcal{P}^-$. For each $i \geq 1$, $T^i x$ and $T^i y$ belong to the same atom of $\mathcal{P}$, thus $T^{i+k_n}x$ and $T^{i+k_n}y$ belong to the same atom of $\mathcal{Q}_n$ for all $n \geq 1$. For all $k > k_n$ the points $T^k x$ and $T^k y$ belong to the same atom of $\mathcal{Q}_n$, and $d(T^k x, T^k y) \leq \delta_n$, thus x, y are asymptotic.

ii) Assume that $\mathcal{P}^- = \mathcal{B}$, then by (5) one obtains $H(\mathcal{P}_n|\mathcal{P}_n^-) \to 0$. In addition

$$H(\mathcal{P}_n|\mathcal{P}_n^-) = h_\mu(\mathcal{P}_n, T) \geq h_\mu(\mathcal{Q}_n, T) \to h_\mu(X, T),$$

so $h_\mu(X, T) = 0$. This completes the proof. □

3.2. **The invertible case.** For $x \in X$ denote by $\mathbf{A}(x)$ the set of points of X that are asymptotic to x.

Proposition 1. *Let (X, T) be an invertible topological dynamical system with positive topological entropy. Then (X, T) has proper asymptotic pairs.*

More precisely, the set of points belonging to a proper asymptotic pair has measure 1 for any ergodic measure on X with positive entropy.

Proof. Let μ be an ergodic measure on X with $h_\mu(X, T) > 0$; the existence of μ follows from the Variational Principle. Let $\mathcal{P}$ be the excellent partition for $(X, \mathcal{B}, T, \mu)$ constructed in Lemma 4.

Let J be the set of points of X which belong to a proper asymptotic pair. J is measurable and invariant under T.

By ergodicity $\mu(J) = 0$ or 1; assume that $\mu(J) = 0$. Then $\mathbf{A}(x) = \{x\}$ for almost every x, thus $\mathcal{P}^-(x) = \{x\}$ by construction. Then $\mathcal{B} = \sigma(\mathcal{P}^-)$ up to sets of measure 0; by Lemma 4 *ii*) this contradicts $h_\mu(X, T) > 0$. □

Remark 1. Although $\mathbf{A}$ is Borel, the set J of points that belong to a proper asymptotic pair may not be Borel. Nevertheless J is measurable (modulo null sets) for all Borel measures.

Remark 2. As $h(X, T) = h(X, T^{-1})$, there are also proper asymptotic pairs for T^{-1}. It will be shown later that the stable classes of x under T and T^{-1} do not coincide.

Proposition 2. *Let $\pi : (X, T) \to (Y, S)$ be a topological factor map, collapsing all proper asymptotic pairs of X. Then $h(Y, S) = 0$.*

Proof. The difficulty here comes from the fact that the system Y can have proper asymptotic pairs [23].

Denote the Borel σ-algebra of Y by $\mathcal{B}_Y$. Let ν be an ergodic measure on Y: ν has a preimage μ under π, which is T-ergodic [8]. Let $\mathcal{P}$ be the excellent partition of $(X, \mathcal{B}, T, \mu)$ constructed in Lemma 4. When two points belong to the same atom of $\mathcal{P}^-$ they are asymptotic: they are collapsed by π and belong to the same atom of $\pi^{-1}\mathcal{B}_Y$. This means that the σ-algebra $\pi^{-1}(\mathcal{B}_Y)$ is contained in the σ-algebra $\mathcal{P}^-$. As $\pi^{-1}(\mathcal{B}_Y)$ is invariant by T, it follows from Lemma 3 that it is contained in Π_μ. Thus, for any finite partition $\mathcal{Q}$ of Y, the partition $\pi^{-1}(\mathcal{Q})$ of X is Π_μ-measurable and

$$h_\nu(\mathcal{Q}, S) = h_\mu(\pi^{-1}(\mathcal{Q}), T) = 0 \ .$$

Therefore $h_\nu(Y, S) = 0$; the conclusion follows from the Variational Principle. □

3.3. **The non-invertible case.** Let (X, T) be a non-invertible topological system: X is a compact metric space for the distance d, and $T : X \to X$ is continuous and onto but not one-to-one.

X evidently admits proper asymptotic pairs, namely any pair (x, y) with $x \neq y$ and $T^n x = T^n y$ for some $n > 0$; when (X, T) is a subshift all asymptotic pairs are of this kind. It is nevertheless not obvious, and interesting to know, that the almost-everywhere result of Proposition 1 holds in the non-invertible case too.

Proposition 3. *Let (X, T) be a non-invertible topological dynamical system. The set of points belonging to a proper asymptotic pair has measure* 1 *for any ergodic measure of positive entropy.*

Proof. Recall the definition of the natural extension $(\tilde{X}, \tilde{T})$ of (X, T): denote by $\tilde{x} = (x_n; n \in \mathbb{Z})$ a point of $X^{\mathbb{Z}}$, and by $\tilde{X}$ the closed subset of $X^{\mathbb{Z}}$ consisting of points $\tilde{x}$ such that $x_{n+1} = T x_n$ for all n. $\tilde{X}$ is invariant by the shift $\tilde{T}$, which is a homeomorphism of $\tilde{X}$. Moreover, the map $\pi : \tilde{x} \mapsto x_0$ is onto by compactness and satisfies $T \circ \pi = \pi \circ \tilde{T}$.

The topology of $\tilde{X}$ is defined by the distance

$$\tilde{d}(\tilde{x}, \tilde{y}) = \sum_{n \in \mathbb{Z}} 2^{-|n|} d(x_n, y_n) \ .$$

Thus a pair $(\tilde{x}, \tilde{y})$ is asymptotic in $\tilde{X}$ if and only if the pair (x_0, y_0) is asymptotic in X.

Let J be the subset of X consisting of all points belonging to a proper asymptotic pair, and let $\tilde{J}$ have the same definition in $\tilde{X}$. Since T is onto, $T^{-1}J \subset J$. Let $z \in \pi(\tilde{J})$. Choose $\tilde{x} \in \tilde{J}$ with $x_0 = z$, then there exists $\tilde{y} \neq \tilde{x}$ such that $(\tilde{x}, \tilde{y})$ is asymptotic. There exists $k \geq 0$ such that $y_{-k} \neq x_{-k}$, thus (x_{-k}, y_{-k}) is a proper asymptotic pair in X, and $x_{-k} \in J$. It follows from $z = x_0 = T^k x_{-k}$ that $z \in T^k J$. Finally $\pi(\tilde{J}) \subset \bigcup_{k \geq 0} T^k J$.

Let μ be an ergodic measure on X with $h_\mu(X, T) > 0$. It lifts to an ergodic measure $\tilde{\mu}$ on $\tilde{X}$, with $h_{\tilde{\mu}}(\tilde{X}, \tilde{T}) > 0$. By Proposition 1, $\tilde{\mu}(\tilde{J}) = 1$, thus $\mu(\pi(\tilde{J})) = 1$, which by the inclusion above implies that $\mu(T^k J) > 0$ for some k.

For every k, $T^{-k}(T^k J) \subset J$: if $T^k x \in T^k J$ there exist $y \in J$ and z such that $T^k x = T^k y$ and (y, z) is a proper asymptotic pair. Then either (x, z) or (x, y) is a proper asymptotic pair depending on whether $x = y$ or not, and $x \in J$. By the inclusion above it follows that $\mu(J) > 0$, and since μ is ergodic $\mu(J) = 1$. □

4. Relatively independent squares

4.1. **Background.** Let (X, T) be a topological dynamical system, $\mathcal{B}$ be its Borel σ-algebra, and μ be an ergodic measure.

For the definition and classical properties of conditional expectations used in this section see [6], [7], [1]. We shall use the

Martingale Theorem. *Let* $(\mathcal{G}_n)_{n\geq 1}$ *be a decreasing sequence of sub-*σ*-algebras of* $\mathcal{B}$ *and let* $\mathcal{G} = \bigcap_{n\geq 1} \mathcal{G}_n$*. For every* $f \in L^2(\mu)$*,* $\mathbb{E}(f \mid \mathcal{G}_n) \to \mathbb{E}(f \mid \mathcal{G})$ *in* $L^2(\mu)$ *and almost everywhere.*

The definition of the relatively independent (or conditional) product of two systems can be found in [20].

Definition 2. *Let* $\mathcal{G}$ *be a sub-*σ*-algebra of* $\mathcal{B}$*. The* conditional square $\mu \underset{\mathcal{G}}{\times} \mu$ *of* μ *relatively to* $\mathcal{G}$ *is the measure on* $(X \times X, \mathcal{B} \otimes \mathcal{B})$ *determined by*

$$\forall A, B \in \mathcal{B},\ \mu \underset{\mathcal{G}}{\times} \mu(A \times B) = \int \mathbb{E}(\mathbb{1}_A \mid \mathcal{G})(x)\, \mathbb{E}(\mathbb{1}_B \mid \mathcal{G})(x)\, d\mu(x)\ .$$

$\mu \underset{\mathcal{G}}{\times} \mu$ is a probability measure, and its two projections on X are equal to μ.

By standard arguments for every pair of bounded Borel functions f, g on X one has

$$\int f(x)g(y)\, d(\mu \underset{\mathcal{G}}{\times} \mu)(x,y) = \int \mathbb{E}(f \mid \mathcal{G})(x)\, \mathbb{E}(g \mid \mathcal{G})(x)\, d\mu(x)\ .$$

The following lemma states the properties of conditional squares that will be used.

Lemma 5. *Let* $\mathcal{G}$ *be a sub-*σ*-algebra of* $\mathcal{B}$*.*

- $\imath$) $\mu \underset{\mathcal{G}}{\times} \mu$ *is concentrated on the diagonal* Δ *of* $X \times X$ *if and only if the* σ*-algebras* $\mathcal{G}$ *and* $\mathcal{B}$ *are equal up to null sets.*
- $\imath\imath$) *If the* σ*-algebra* $\mathcal{G}$ *is invariant by* T*, then the measure* $\mu \underset{\mathcal{G}}{\times} \mu$ *is invariant by* $T \times T$*.*
- $\imath\imath\imath$) *Let* f *be a bounded* $\mathcal{G}$*-measurable function on* X*. Then* $f(x) = f(y)$ *for* $\mu \underset{\mathcal{G}}{\times} \mu$*-almost all* $(x, y) \in X \times X$*.*
- $\imath\nu$) *Let* $(\mathcal{G}_n)_{n\geq 1}$ *be a decreasing sequence of* σ*-algebras with* $\bigcap_{n\geq 1} \mathcal{G}_n = \mathcal{G}$*. Then for all* $A, B \in \mathcal{B}$ *one has*

$$\mu \underset{\mathcal{G}}{\times} \mu(A \times B) = \lim_{n\to\infty} \mu \underset{\mathcal{G}_n}{\times} \mu(A \times B)\ ,$$

and the sequence $(\mu \underset{\mathcal{G}_n}{\times} \mu; n \geq 1)$ *converges weakly to* $\mu \underset{\mathcal{G}}{\times} \mu$*.*

Proof. $\imath$) If $\mathcal{G} = \mathcal{B}$, then for all $A, B \in \mathcal{B}$ we have $\mathbb{E}(\mathbb{1}_A \mid \mathcal{G}) = \mathbb{1}_A$ and $\mathbb{E}(\mathbb{1}_B \mid \mathcal{G}) = \mathbb{1}_B$ μ-a.e.; then by definition $\mu \underset{\mathcal{G}}{\times} \mu(A \times B) = \mu(A \cap B)$; the measure $\mu \underset{\mathcal{G}}{\times} \mu$ is the image of μ under the map $x \mapsto (x, x)$, thus it is concentrated on Δ.

If $\mu \underset{\mathcal{G}}{\times} \mu$ is concentrated on Δ, for all $A \in \mathcal{B}$ one has $\mu \underset{\mathcal{G}}{\times} \mu(A \times (X \setminus A)) = 0$, that is,

$$\int \mathbb{E}(\mathbb{1}_A \mid \mathcal{G})(x)\, \mathbb{E}(\mathbb{1}_{X\setminus A} \mid \mathcal{G})(x)\, d\mu(x) = 0\ ,$$

thus the product of the two conditional expectations is equal to 0 a.e.. As the sum of these two functions is equal to 1, each of them is equal to 0 or 1 a.e.. It follows that $\mathbb{E}(\mathbb{1}_A \mid \mathcal{G}) = \mathbb{1}_A$ a.e., and A is measurable with respect to $\mathcal{G}$. The σ-algebras $\mathcal{G}$ and $\mathcal{B}$ are equal up to null sets.

$\imath\imath$) Obvious.

$\imath\imath\imath$) By definition

$$\int f(x)\overline{f(y)}\, d(\mu \underset{\mathcal{G}}{\times} \mu)(x,y) = \int |f(x)|^2\, d\mu(x) \text{ because } f \text{ is } \mathcal{G}\text{-measurable thus}$$

$$\int |f(x) - f(y)|^2\, d(\mu \underset{\mathcal{G}}{\times} \mu)(x,y) = 0\ .$$

$\imath\nu$) When f and g are bounded measurable functions on X, by the Martingale Theorem

$$\begin{aligned}\int f(x)g(y)\, d(\mu \underset{\mathcal{G}_n}{\times} \mu)(x,y) &= \int \mathbb{E}(f \mid \mathcal{G}_n)(x)\, \mathbb{E}(g \mid \mathcal{G}_n)(x)\, d\mu(x) \\ &\to \int \mathbb{E}(f \mid \mathcal{G})(x)\, \mathbb{E}(g \mid \mathcal{G})(x)\, d\mu(x) \\ &= \int f(x)g(y)\, d(\mu \underset{\mathcal{G}}{\times} \mu)(x,y)\ . \end{aligned} \tag{6}$$

For $f = \mathbb{1}_A$ and $g = \mathbb{1}_B$ this is the first part of $\imath\nu$). The family of continuous functions F on $X \times X$ such that

$$\int F(x,y)\, d(\mu \underset{\mathcal{G}_n}{\times} \mu)(x,y) \to \int F(x,y)\, d(\mu \underset{\mathcal{G}}{\times} \mu)(x,y)$$

is a closed subspace of $\mathcal{C}(X \times X)$. By equation (6) it contains all functions $f(x)g(y)$ where f and g belong to $\mathcal{C}(X)$ and their linear combinations. By density it is equal to $\mathcal{C}(X \times X)$, which completes the proof. □

We consider now the case where $\mathcal{G}$ is associated to a measurable partition, also denoted by $\mathcal{G}$.

Lemma 6. *Let $\mathcal{G}$ be a measurable partition. Then the set*

$$\Delta_{\mathcal{G}} = \big\{(x,y) \in X \times X; y \in \mathcal{G}(x)\big\}$$

belongs to $\mathcal{B} \otimes \mathcal{B}$, and $\mu \underset{\mathcal{G}}{\times} \mu$ is concentrated on this set.

Proof. Let $(\mathcal{G}_n)_{n\geq 1}$ be an increasing sequence of finite partitions with $\bigvee_{n\geq 1} \mathcal{G}_n = \mathcal{G}$. Whenever A, B are two distinct atoms of $\mathcal{G}_n$ it follows immediately from the definition that $\mu \underset{\mathcal{G}_n}{\times} \mu(A \times B) = 0$. By Lemma 5 $\imath\nu$), $\mu \underset{\mathcal{G}}{\times} \mu(A \times B) = 0$. Thus, for all n the measure $\mu \underset{\mathcal{G}}{\times} \mu$ is concentrated on $\Delta_{\mathcal{G}_n}$. But the intersection of these sets is $\Delta_{\mathcal{G}}$, and the result follows. □

4.2. **The 'construction $\mathcal{C}$'.** In the sequel we use several times the following construction, referred to as *the construction* $\mathcal{C}$, with the same notation.

Let (X,T) be a topological dynamical system, $\mathcal{B}$ be its Borel σ-algebra and $\mathbf{A}$ be the set of asymptotic pairs; it is a Borel subset of $X \times X$, invariant under $T \times T$.

Let μ be an invariant ergodic measure. Using Lemma 4, choose an excellent partition $\mathcal{P}$, such that any pair of points belonging to the same atom of $\mathcal{P}^-$ is asymptotic, and put $\mathcal{F} = \sigma(\mathcal{P}^-)$. By Lemma 4 again if $h_\mu(X,T) > 0$, $\mathcal{F}$ is not equal to $\mathcal{B}$ up to μ-null sets. In the notation of Lemma 6

$$\Delta_{\mathcal{F}} \subset \mathbf{A}\ .$$

For every $n \geq 0$ put

$$\mathcal{F}_n = T^{-n}\mathcal{F} \ \text{ and } \ \nu_n = \mu \underset{\mathcal{F}_n}{\times} \mu\ ;$$

one has

$$\Delta_{\mathcal{F}_n} = (T \times T)^{-n} \Delta_{\mathcal{F}} \subset \mathbf{A} \text{ and } \nu_n = (T \times T)^{-n} \nu_0 \ ;$$

thus ν_n is concentrated on $\mathbf{A}$. Moreover, the sequence of sets $(\Delta_{\mathcal{F}_n})_{n \geq 0}$ is increasing; the sequence $(\mathcal{F}_n)_{n \geq 0}$ of σ-algebras is decreasing and its intersection is equal to Π_μ up to sets of μ-measure 0 by Lemma 3.

Define

$$\lambda = \mu \underset{\Pi_\mu}{\times} \mu \ .$$

From Lemma 5 $\imath\nu$) one gets

Corollary 1. *$\imath$) For every $A, B \in \mathcal{B}$, $\nu_n(A \times B) \to \lambda(A \times B)$ as $n \to \infty$, and the sequence $(\nu_n)_{n \geq 0}$ of measures on $X \times X$ converges weakly to λ.*

$\imath\imath$) For every closed subset F of $X \times X$ with $(T \times T)F \supset F$ one has $\lambda(F) \geq \nu_0(F)$.

$\imath\imath\imath$) For every open subset U of $X \times X$ with $(T \times T)U \subset U$ one has $\lambda(U) \leq \nu_0(U)$.

Proof. $\imath$) Immediate from Lemma 5 $\imath\nu$).
$\imath\imath$) Since F is closed and $\nu_n \to \lambda$ weakly one has

$$\lambda(F) \geq \limsup_{n \to \infty} \nu_n(F) \ .$$

But the sequence $\nu_n(F) = \nu_0\big((T \times T)^n F\big)$ is increasing and the result follows.
$\imath\imath\imath$) Immediate from $\imath\imath$). □

The next result shows that a 2-set partition of positive entropy separates some asymptotic pair. Significantly, it does the same for some entropy pair [4].

Corollary 2. *Let $\mathcal{Q} = (A_1, A_2)$ be a Borel partition with $h_\mu(\mathcal{Q}, T) > 0$. Then there exists an asymptotic pair (x_1, x_2) with $x_1 \in A_1$ and $x_2 \in A_2$.*

Proof. If the result is false, then $(A_1 \times A_2) \cap \mathbf{A} = \emptyset$, and $\nu_n(A_1 \times A_2) = 0$ for all n, thus by Corollary 1 $\imath$)

$$0 = \lambda(A_1 \times A_2) = \int \mathbb{E}(\mathbb{1}_{A_1} \mid \Pi_\mu)(x) \, \mathbb{E}(\mathbb{1}_{A_2} \mid \Pi_\mu)(x) \, d\mu(x) \ .$$

As the two conditional expectations in the integral are non-negative and have sum equal to 1, each of them is equal to 0 or 1 a.e., which means that the sets A_1 and A_2 belong to the σ-algebra Π_μ; thus $h_\mu(Q, T) = 0$, which contradicts the assumption. □

4.3. **Application to entropy pairs.** The definition of entropy pairs of a topological system (X, T) is given in [5]. The set $E(X, T)$ of entropy pairs is a $T \times T$ invariant subset of $X \times X$, and $E(X, T) \cup \Delta$ is closed. The system (X, T) has entropy pairs if and only if its entropy is positive.

The reader should be reminded of the definition of entropy pairs for an invariant measure μ [4]. Let $x, y \in X$ with $x \neq y$. A partition $\mathcal{Q} = (A, B)$ is said to *separate* x and y if x belongs to the interior of A and y to the interior of B. (x, y) is said to be an *entropy pair for* μ if for any partition $\mathcal{Q}$ separating x and y one has $h_\mu(\mathcal{Q}, T) > 0$. Call $E_\mu(X, T)$ the set of entropy pairs for μ. This set is non-empty if and only if $h_\mu(X, T) > 0$.

It is shown in [2] that $E(X, T) = \overline{\bigcup_\mu E_\mu(X, T)}$, where the union is taken over the family of ergodic measures.

Moreover, Glasner shows in [12] that for any ergodic measure μ, $E_\mu(X,T)$ is the set of non-diagonal points in the topological support of $\mu \underset{\Pi_\mu}{\times} \mu$ (this result also follows easily from the definition of $\mu \underset{\mathcal{F}}{\times} \mu$ and Lemma 5).

Proposition 4. *The closure $\overline{\mathbf{A}}$ of $\mathbf{A}$ in $X \times X$ contains the set $E(X,T)$ of entropy pairs.*

Proof. Let μ be an ergodic measure on X. In the notation of the 'construction $\mathcal{C}$', for every n, the measure ν_n is concentrated on the closed set $\overline{\mathbf{A}}$, and so is the weak limit λ of the sequence (ν_n). By Glasner's result $E_\mu(X,T) \subset \overline{\mathbf{A}}$. As this is true for any ergodic μ, the result of [2] quoted above gives the conclusion. □

Corollary 3. *If (X,T) admits an invariant measure μ of full support such that $(X,\mathcal{B},T,\mu)$ is a K-system, then asymptotic pairs are dense in $X \times X$.*

Proof. For such a measure μ the Pinsker σ-algebra Π_μ is trivial, it follows that $\lambda = \mu \times \mu$, its support is $X \times X$, and $E_\mu(X,T) \cup \Delta = X \times X = \overline{\mathbf{A}}$. □

In this case $E(X,T) \cup \Delta = X \times X$, as shown in [11] by different means.

5. Li-Yorke pairs and instability in negative times

Lemma 7. *Let $(X,\mathcal{B},T,\mu)$ be an ergodic system, and $\lambda = \mu \underset{\Pi_\mu}{\times} \mu$. Then $(X \times X, \mathcal{B} \otimes \mathcal{B}, T \times T, \lambda)$ is ergodic.*

Proof. Assume that λ is not ergodic. According to Theorems 7.5 and 8.2 in [9] there exists a non-trivial isometric extension of (X,Π_μ,T,μ) (in the measure-theoretic sense) which is a factor of $(X,\mathcal{B},T,\mu)$. An ergodic isometric extension is a factor of an ergodic group extension (Theorem 8.2 in [9]), thus an ergodic isometric extension of a 0-entropy system also has entropy 0, and this contradicts the characterization of Π_μ as the largest factor of X with entropy 0. □

In the next Proposition ν_0 is defined as in the 'construction $\mathcal{C}$' above.

Proposition 5. *Let (X,T) be a topological system, μ an ergodic measure of positive entropy and*

$$\delta = \sup\big\{d(x,y); (x,y) \in E_\mu(X,T)\big\} > 0 \ .$$

For ν_0-almost every pair $(x,y) \in X \times X$ one has

$$\lim_{n\to+\infty} d(T^n x, T^n y) = 0 \ ; \tag{7}$$

$$\liminf_{n\to+\infty} d(T^{-n}x, T^{-n}y) = 0 \text{ and } \limsup_{n\to+\infty} d(T^{-n}x, T^{-n}y) \geq \delta \ ; \tag{8}$$

in particular (x,y) is a Li-Yorke pair for T^{-1}.

Proof. Let U be an open set in $X \times X$, with $\lambda(U) > 0$. For every $M \geq 0$ we write

$$U_M = \bigcup_{m \geq M} (T \times T)^m U \ .$$

U_M is open, and $(T \times T)U_M = U_{M+1} \subset U_M$. Moreover, $\lambda(U_M) \geq \lambda(U) > 0$. By ergodicity of λ, $\lambda(U_M) = 1$. By Corollary 1 *iii*), $\nu_0(U_M) \geq \lambda(U_M) = 1$. Let

$$V = \bigcap_{M \geq 0} U_M = \bigcap_{M \geq 0} \bigcup_{m \geq M} (T \times T)^m U \ .$$

V is invariant by $T \times T$, and $\nu_0(V) = 1$.

For every integer $r > 1$, we can cover $\mathrm{Supp}(\lambda)$ by a finite number of open balls of radius $1/r$, each of them intersecting $\mathrm{Supp}(\lambda)$. Taking the union of all these families we obtain a sequence $(U_k)_{k\geq 1}$ of open sets, with $U_k \cap \mathrm{Supp}(\lambda) \neq \emptyset$ for all k; each point of $\mathrm{Supp}(\lambda)$ belongs to U_k for infinitely many values of k; the diameter of U_k tends to 0 as $k \to \infty$. To each k we associate a set V_k as above, and write $G = \bigcap_{k\geq 1} V_k$. We have $\nu_0(G) = 1$.

Let (x, y) be a point in G. For each k, $(T \times T)^{-n}(x, y) \in U_k$ for infinitely many values of n, thus the negative orbit of (x, y) is dense in $\mathrm{Supp}(\lambda)$.

By Glasner's result [12], $\mathrm{Supp}(\lambda) = E_\mu(X, T) \cup S(\mu)$, where $S(\mu) = \{(x, x); x \in \mathrm{Supp}(\mu)\}$. Thus we can choose a pair (x_0, y_0) in $E_\mu(X, T)$ with $d(x_0, y_0) = \delta$ and another pair (z_0, z_0) in $S(\mu)$. It follows that for all $(x, y) \in G$ both (x_0, y_0) and (z_0, z_0) are in the closure of the negative orbit of (x, y), thus $\limsup_{n\to+\infty} d(T^{-n}x, T^{-n}y) \geq \delta$ and $\liminf_{n\to+\infty} d(T^{-n}x, T^{-n}y) = 0$. Finally, every pair $(x, y) \in G$ satisfies Eq. (8)

Recall that ν_0 is concentrated on $\Delta_{\mathcal{F}}$, that is, $\nu_0(\Delta_{\mathcal{F}}) = 1$, and that every pair in $\Delta_{\mathcal{F}}$ is positively asymptotic. Thus $\nu_0(\Delta_{\mathcal{F}} \cap G) = 1$, and every pair in this set satisfies Eq. (7). □

Remark 3. Assume that μ is a weakly mixing invariant measure on X, different from a Dirac measure. Then the same argument as in the proof of Proposition 5 shows that there exists a G_δ-set G of $X \times X$, invariant under $T \times T$, dense in $\mathrm{Supp}(\mu) \times \mathrm{Supp}(\mu)$, with $\mu \times \mu(G) = 1$ and such that every pair $(x, y) \in G$ is Li-Yorke. More precisely, there exists $\delta > 0$ such that for every $(x, y) \in G$

$$\liminf_{n\to+\infty} d(T^n x, T^n y) = 0 \text{ and } \limsup_{n\to+\infty} d(T^n x, T^n y) \geq \delta \ .$$

Here no assumption of positive entropy is needed. This is related to Iwanik's result on independent sets in topologically weakly mixing systems [15].

6. There are uncountably many asymptotic pairs

Up to now most of the results were existence results: we have shown that a system of positive entropy has asymptotic pairs, and even pairs which are asymptotic for positive times and Li-Yorke for negative times. It is interesting to know how large a stable class is, and in particular whether it can be countable. We prove that the answer is negative for a.e. class. We need more probabilistic tools.

6.1. **Conditional measures.**

Here X is a compact metric space, endowed with its Borel σ-algebra $\mathcal{B}$. Let $\mathcal{M}(X)$ be the set of probability measures on X, endowed with the topology of weak convergence. It is a compact metrizable space. A proof of the next result can be found in [10].

Lemma 8. *Let μ be a probability measure on X, and $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{B}$. There exists a map $x \mapsto \mu_x$ from X to $\mathcal{M}(X)$, measurable with respect to $\mathcal{F}$, and such that for every bounded function f on X*

$$\mathbb{E}(f \mid \mathcal{F})(x) = \int f(y)\, d\mu_x(y) \quad \text{for } \mu\text{-a.e. } x. \tag{9}$$

This map is called a regular version *of the conditional probability.*

We continue to use the notation of this Lemma.

By definition of the conditional square (see Sec. (4.1)) the equality

$$\mu \underset{\mathcal{F}}{\times} \mu(K) = \int \mu_x \otimes \mu_x(K)\, d\mu(x) \tag{10}$$

holds whenever $K = A \times B$ where A, B are Borel sets in X. By standard arguments, it holds for every Borel subset K of $X \times X$. Thus for every bounded Borel function f on $X \times X$ one has

$$\int f(x,y)\, d(\mu \underset{\mathcal{F}}{\times} \mu)(x,y) = \int \Big(\int f(x,y)\, d\mu_x(y) \Big)\, d\mu(x)\ . \tag{11}$$

We establish now a condition for the measure μ_x to be atomless μ-almost-everywhere. It is easy to check that the function $(x,y) \mapsto \mu_x(\{y\})$ is Borel, thus the set $\{x \in X; \mu_x \text{ is atomless}\}$ is measurable.

Lemma 9. *Let Δ be the diagonal of $X \times X$. Then $\mu \underset{\mathcal{F}}{\times} \mu(\Delta) = 0$ if and only if μ_x is atomless for μ-almost all $x \in X$.*

Proof. We write $\nu = \mu \underset{\mathcal{F}}{\times} \mu$.

By Fubini's Theorem and Eq. (10),

$$\nu(\Delta) = \int \mu_x \otimes \mu_x(\Delta)\, d\mu(x) = \int \mu_x(\{x\})\, d\mu(x)\ .$$

The 'if' part of the Lemma is now immediate. Now assume that $\nu(\Delta) = 0$. One has

$$\mu_x(\{x\}) = 0 \text{ for } \mu\text{-almost all } x\ . \tag{12}$$

As the map $x \mapsto \mu_x$ is $\mathcal{F}$-measurable, it follows from Lemma 5 *iii*) that $\mu_x = \mu_y$ for ν-almost all (x,y), thus

$$\mu_x(\{x\}) = \mu_y(\{x\}) \text{ for } \nu\text{-almost all } (x,y)\ . \tag{13}$$

As the first projection of ν on X is μ, it follows from Eqs. (12) and (13) that

$$\mu_y(\{x\}) = 0 \text{ for } \nu\text{-almost all } (x,y)\ . \tag{14}$$

Using Eq. (11) with $f(x,y) = \mu_y(\{x\})$ one gets

$$0 = \int \mu_y(\{x\})\, d\nu(y) = \int \Big(\int \mu_y(\{x\})\, d\mu_y(x) \Big)\, d\mu(y)\ .$$

Hence $\int \mu_y(\{x\})\, d\mu_y(x) = 0$ for μ-almost all y.

But for all y the measure μ_y is larger than its discrete part $\tau_y = \sum_z \mu_y(\{z\})\delta_z$, where δ_z is the Dirac mass at z, and for μ-almost all y we have

$$0 = \int \mu_y(\{x\})\, d\tau_y(x) = \sum_z (\mu_y(\{z\}))^2$$

thus $\mu_y(\{z\}) = 0$ for all z and μ_y is atomless. □

6.2. **Application to asymptotic pairs.** The next result is a topological counterpart of Proposition 5.

Proposition 6. *Assume that $h(X,T) > 0$. There exist $\delta > 0$, an uncountable subset F of X, and for every $x \in F$ an uncountable subset F_x of X such that for every $y \in F_x$ the relations* (7) *and* (8) *hold, that is,*

$$\lim_{n\to+\infty} d(T^n x, T^n y) = 0 \ ;$$

$$\liminf_{n\to+\infty} d(T^{-n}x, T^{-n}y) = 0 \text{ and } \limsup_{n\to+\infty} d(T^{-n}x, T^{-n}y) \geq \delta \ .$$

Here one can choose any δ such that $0 < \delta < \sup\{d(x,y); (x,y) \in E(X,T)\}$.

Proof. Let $D = \sup\{d(x,y); (x,y) \in E(X,T)\}$. By compactness there exists $(x',y') \in E(X,T)$ with $d(x',y') = D$. Thus for $0 < \delta < D$ there exist an ergodic measure μ and a μ-entropy pair (x_0,y_0) close to (x',y') such that $d(x_0,y_0) \geq \delta$. Recall that in the notations of the 'construction $\mathcal{C}$' $\nu_0 = \mu \underset{\mathcal{F}}{\times} \mu$; denote by μ_x a regular version of the conditional probability given $\mathcal{F}$, as in Lemma 8.

Lemma 10. *With the assumptions of Proposition 6 for μ-almost every x the measure μ_x is atomless.*

Proof. (of the Lemma) Assume that the conclusion does not hold. By Lemma 9, $\nu_0(\Delta) > 0$. As Δ is invariant under $T \times T$, by Corollary 1 *ii*) $\lambda(\Delta) \geq \nu_0(\Delta) > 0$. By Lemma 7, λ is ergodic for $T \times T$, thus $\lambda(\Delta) = 1$.

By Lemma 5 *i*) it means that $\Pi_\mu = \mathcal{B}$ up to μ-null sets, thus $h_\mu(X,T) = 0$. This is impossible because there exists an entropy pair (x_0,y_0) for μ. □

We continue the proof of Proposition 6. By Proposition 5, the relations (7) and (8) hold for ν_0-almost every $(x,y) \in X \times X$.

For $x \in X$, let F_x be the set of all points $y \in X$ such that these relations hold for (x,y). Since $\nu_0 = \mu \underset{\mathcal{F}}{\times} \mu$ one has $\int \mu_x(F_x)\, d\mu(x) = 1$, thus $\mu_x(F_x) = 1$ for μ-almost all x. Let

$$F = \{x \in X; \mu_x(F_x) = 1\} \cap \{x \in X; \mu_x \text{ is atomless}\} \ .$$

Then $\mu(F) = 1$. The measure μ is ergodic and of positive entropy, thus atomless. Hence the set F is uncountable. For $x \in F$, $\mu_x(F_x) = 1$ and μ_x is atomless, therefore F_x is an uncountable set. □

References

[1] Billingsley, B.: Probability and Measure. John Wiley & Sons, New York, 1995.

[2] Blanchard, F., Glasner, E., Host, B.: *A variation on the variational principle and applications to entropy pairs*, Ergodic Th. Dyn. Sys. **17** (1997), no. 1, 29-43.

[3] Blanchard, F., Glasner, E., Kolyada, S., Maass, A.: *A note on Li-Yorke pairs*, preprint (2000).

[4] Blanchard, F., Host, B., Maass, A., Martínez, S., Rudolph, D.J.: *Entropy pairs for a measure*, Ergodic Th. Dynam. Sys. **15** (1995), no. 4, 621-632.

[5] Blanchard, F.: *A disjointness theorem involving topological entropy*, Bull. Soc. Math. France **121** (1993), no. 4, 465-478.

[6] Doob, J. L.: Stochastic Processes. John Wiley & Sons, New York, 1990.

[7] Doob, J. L.: Measure Theory. Springer-Verlag, Berlin-Heidelberg-New York, 1994.

[8] Denker, M., Grillenberger, C., Sigmund, K.: *Ergodic Theory on Compact Spaces*, Lecture Notes in Mathematics 527, Springer-Verlag, Berlin-Heidelberg-New York, 1976.

[9] Furstenberg, H.: *Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204-256.

[10] Furstenberg, H.: Recurrence in Ergodic Theory and Combinatorial number theory. Princeton University Press, Princeton, 1981.
[11] Glasner, E., Weiss, B.: *Strictly ergodic, uniform positive entropy models*, Bull. Soc. Math. France **122** (1994), 399-412.
[12] Glasner, E.: *A simple characterization of the set of μ-entropy pairs and applications*, Israel J. Math. **102** (1997), 13-27.
[13] Huang, W., Ye, X.D.: *Homeomorphisms with the whole compacta being scrambled sets*, Ergod. Th. Dynam. Sys., to appear.
[14] Huang, W., Ye, X.D.: *Devaney's chaos or 2-scattering imply Li-Yorke's chaos*, preprint (1999).
[15] Iwanik, A.: Independent sets of transitive points, *Dynamical systems and ergodic theory (Warsaw, 1986)*, 277-282 Banach Center Publ. **23**, PWN, Warsaw, 1989.
[16] Li, T.Y., Yorke, J.A.: *Period three implies chaos*, Amer. Math. Monthly **82** (1975), no. 10, 985-992
[17] Parry, W.: Entropy and generators in Ergodic Theory. Benjamin, New York, 1969.
[18] Parry, W.: Topics in Ergodic Theory. Cambridge Tracts in Math., Cambridge University Press, Cambridge, 1981.
[19] Rohlin, V.A., Sinaĭ, Ya.G.: *Construction and properties of invariant measurable partitions*, Sov. Math. **2** (1961), no. 6, 1611-1614.
[20] Rudolph, D.J.,: Fundamentals of measurable dynamics, Ergodic theory on Lebesgue spaces. The Clarendon Press Oxford University Press, New York, 1990.
[21] Walters, P.: An introduction to Ergodic Theory. Springer, Berlin-Heidelberg-New York, 1982.
[22] Weiss, B.: *Multiple recurrence and doubly minimal systems*, Contemporary Mathematics, Vol 215, 1998, 189-196.
[23] Ye, XD.: personal communication.

François Blanchard, Institut de Mathématiques de Luminy - CNRS UPR 9016 - 163, avenue de Luminy, case 907 - 13288 Marseille cedex 9 - France
E-mail address: blanchar@iml.univ-mrs.fr

Bernard Host, Université de Marne la Vallée - Cité Descartes, 5, boulevard Descartes - Champs sur Marne, 77454 Marne-la-Vallée Cedex 2 - France
E-mail address: host@math.univ-mlv.fr

Sylvie Ruette, Institut de Mathématiques de Luminy - CNRS UPR 9016 - 163, avenue de Luminy, case 907 - 13288 Marseille cedex 9 - France
E-mail address: ruette@iml.univ-mrs.fr