ON LI–YORKE PAIRS

François Blanchard, Eli Glasner,
Sergiĭ Kolyada and Alejandro Maass

Abstract. The Li–Yorke definition of chaos proved its value for interval maps. In this paper it is considered in the setting of general topological dynamics. We adopt two opposite points of view.

On the one hand sufficient conditions for Li–Yorke chaos in a topological dynamical system are given. We solve a long–standing open question by proving that positive entropy implies Li–Yorke chaos.

On the other hand properties of dynamical systems without Li–Yorke pairs are investigated; in addition to having entropy 0, they are minimal when transitive, and the property is stable under factor maps, arbitrary products and inverse limits. Finally it is proved that minimal systems without Li–Yorke pairs are disjoint from scattering systems.

0. Introduction

The term 'chaos' in connection with a map was introduced by Li and Yorke [27], although without a formal definition. Today there are various definitions of what it means for a map to be chaotic, some of them working reasonably only in particular phase spaces; most of the existing ones were reviewed in [23]. Although one could say that 'as many authors, as many definitions of chaos', most of them are based on the idea of unpredictability of the behavior of trajectories when the position of the point is given with an error (instability of trajectories or sensitive dependence on initial conditions are terms usually used to describe this phenomenon). The present article mainly deals with one of these definitions, namely chaos in the sense of Li–Yorke. It is one in a series of papers by various authors on that subject ([22], [31], [13] and two forthcoming papers, [20], [21]). Our main purpose is to incorporate this notion of chaos into the general frame of Topological Dynamics; formerly it had been studied mainly in the setting of interval maps.

Let (X, T) be a topological dynamical system, with X a compact metric space with metric ϱ, and T a surjective continuous map from X to itself.

Key words and phrases. Entropy, chaos, independent sets, Ellis semigroups, almost distal systems, minimal dynamical systems.

2000 *Mathematics Subject Classification.* 37B05, 37B40, 54H20.

The definition of Li–Yorke chaos is based on ideas in [27]. A pair of points $\{x, y\} \subseteq X$ is said to be a *Li–Yorke pair* (with modulus δ) if one has simultaneously

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

A set $S \subseteq X$ is called *scrambled* if any pair of distinct points $\{x, y\} \subseteq S$ is a Li–Yorke pair. Finally, a system (X, T) is called *chaotic in the sense of Li and Yorke* if X contains an uncountable scrambled set. The first motivation for studying this notion comes from the theory of interval transformations. For such maps the existence of a Li–Yorke pair implies the existence of an uncountable scrambled set [25].

What are the connections between Li–Yorke's and other definitions of chaos? What topological properties of a system imply the existence of Li–Yorke pairs or a scrambled set? These are the questions that arise immediately in one's mind when thinking about Li–Yorke chaos, and we answer some here. But one cannot achieve a real understanding of Li–Yorke chaos without knowing what it means for a system to have the opposite property, that is, having no Li–Yorke pairs. This question is also addressed in this article.

In Section 1, after introducing general definitions and background, we survey some existing definitions of chaos and define new ones, explaining what is known about them and about their relations.

In Section 2 we answer positively the long–standing question whether positive topological entropy implies Li–Yorke chaos (Corollary 2.4). It is worth mentioning that Li–Yorke pairs are not the only ones that are necessarily found in positive–entropy systems: there must also exist asymptotic pairs, that is, pairs (x, y) such that $d(T^n x, T^n y) \to 0$ as $n \to \infty$ [8].

It is easy to show that a scattering system has Li–Yorke pairs (Remark 2.12) and other chaotic features. The scattering property, defined below, is based on the complexity of covers and is essentially weaker than positive entropy. By the time we finished writing this article Huang and Ye had obtained results in the same spirit, viz., that (1) the scattering property and (2) the existence in a transitive non–minimal system of, at least, one periodic orbit, imply Li–Yorke chaos [21].

In Section 3 we explore systems without Li–Yorke pairs. Surely distal systems have this property but, as we shall see, the class of such systems is much wider. However, it turns out that systems without Li–Yorke pairs share many of the basic properties of distal systems. We have already seen that they also have zero entropy; such systems are minimal when transitive (Theorem 3.10), and any factor of a transitive system without Li–Yorke pairs has no Li–Yorke pairs (Theorem 3.9). Finally, in Subsection 3.1 we define the 'adherence semigroup' of a dynamical system and show that systems without Li–Yorke pairs are characterized by the requirement that their adherence semigroup is minimal (Theorem 3.8), in analogy with distal systems which are characterized by the property that their enveloping semigroup is minimal. For these reasons we call systems without Li–Yorke pairs *almost distal*. One important difference with distality is that almost distal systems are not necessarily invertible, as is shown by examples in Subsection 3.3.

Disjointness was introduced in [14]; two disjoint dynamical systems have no common factors, but the property is much stronger. It was proved in [7], as a generalization of a result of [14], that minimal distal systems are disjoint from all scattering systems. Here we show (Theorem 3.12) that almost distal minimal systems are also disjoint from any scattering system.

E. Akin studies a related class of dynamical systems which he calls *semi–distal*. These are the systems in which every *recurrent* proximal pair $\{x, y\} \subset X$ is a diagonal pair (i.e., $x = y$). Among other results he shows that semi–distal is equivalent to the property that every idempotent in the adherence semigroup is minimal [1].

Acknowledgements. This paper was written down while Huang and Ye were elaborating theirs ([21]); it benefited by many remarks of theirs, and also by the contents of the various versions of their article which they sent us. We want to thank E. Akin and J. Auslander for many important remarks and suggestions. This paper was partly written during the first and third authors' visit to the Departamento de Ingeniería Matemática of the Universidad de Chile and the second, third and fourth authors' visit to the Institut de Mathématiques de Luminy (CNRS). They wish to acknowledge the support and kind hospitality offered by both institutions.

The first author was partially supported by the ECOS–CONICYT agreement and by the INTAS OPEN 97 grant 1843, the third author was partially supported by the INTAS OPEN 97 grant 1843 and by the Cátedra Presidencial fellowship, the fourth author was partially supported by the ECOS–CONICYT agreement, by the Cátedra Presidencial fellowship and Fondecyt grant 1010447.

1. Preliminaries

1.1. Definitions and notation.

By a topological dynamical system we mean a pair (X, T) where $X := (X, \varrho)$ is a compact metric space for the distance ϱ and T is a surjective continuous map from X to itself. A *factor map* $\varphi\colon (X, T) \to (X', T')$ is a continuous onto map from X to X' such that $T' \circ \varphi = \varphi \circ T$; in this situation (X, T) is said to be an *extension* of (X', T'). A measure–theoretical dynamical system is a probability space $(X, \mathcal{X}, \mu)$ together with a measurable transformation $T : X \to X$ such that $T\mu = \mu$. A topological dynamical system (X, T) together with a T-invariant probability measure μ defined on the σ-algebra of Borel sets $\mathcal{X}$, determines a measure-theoretical dynamical system $(X, \mathcal{X}, \mu, T)$. In this context, a factor map $\varphi : (X, \mathcal{X}, \mu, T) \to (X', \mathcal{X}', \mu', T')$ is a measurable map from X to X' such that $T' \circ \varphi = \varphi \circ T$ and $\varphi\mu = \mu'$.

Suppose one is given an infinite sequence of extensions $\varphi_n\colon (X_n, T_n) \to (X_{n-1}, T_{n-1})$;

the *inverse limit* is the dynamical system (X, T) where

$$X = \{(x_0, x_1, \dots, x_n, \dots) \in \prod_{i=0}^{\infty} X_i \mid x_{n-1} = \varphi_n(x_n),\ n > 0\}$$

and $T(x_0, x_1, \dots, x_n, \dots) = (T_0(x_0), T_1(x_1), \dots, T_n(x_n), \dots)$; X is endowed with the product topology. The *natural extension* of a non–invertible dynamical system (X_0, T_0) is the inverse limit of infinitely many embedded copies (X_n, T_n) of (X_0, T_0), with $\varphi_n = T_0$, that is, $\widetilde{X} = \{\widetilde{x} = (x_0, x_1, ...) \in X^{\mathbb{N}} \ :\ T_0(x_i) = x_{i-1}, i \geq 1\}$ and $\widetilde{T}(\widetilde{x}) = (T_0(x_0), x_0, x_1, ...)$. Note that the canonical map $\tilde{\pi} : (\widetilde{X}, \widetilde{T}) \to (X_0, T_0)$, the projection on the 0^{th} coordinate, is a factor map and that the system $(\widetilde{X}, \widetilde{T})$ is invertible with $T^{-1}(\widetilde{x}) = (x_1, ...)$.

Recall that $\{x, y\} \subset X$ is a Li–Yorke pair if simultaneously

$$\liminf_{n\to\infty} \varrho(T^n x, T^n y) = 0 \quad \text{and} \quad \limsup_{n\to\infty} \varrho(T^n x, T^n y) > 0.$$

A pair $\{x, y\} \subseteq X$ such that $\liminf_{n\to\infty} \varrho(T^n x, T^n y) > 0$ is said to be *distal*, and one for which $\lim_{n\to\infty} \varrho(T^n x, T^n y) = 0$ is said to be *asymptotic*. A pair $\{x, y\}$ is *proximal* if there exists a sequence n_i with $\lim_{i\to\infty} \varrho(T^{n_i} x, T^{n_i} y) = 0$. Thus a pair $\{x, y\}$ is a Li–Yorke pair iff it is proximal but not asymptotic. The sets of proximal and asymptotic pairs of (X, T) are denoted by $\mathbf{P}(X, T)$ and $\mathbf{As}(X, T)$, respectively, or simply $\mathbf{P}$ and $\mathbf{As}$ whenever there is no ambiguity. The sets of distal pairs, Li–Yorke pairs and asymptotic pairs partition X^2. It is easy to see that the image of a proximal (asymptotic) pair under a factor map is proximal (asymptotic). A *distal* dynamical system is one in which every non–diagonal pair is distal.

A subset M of X is *minimal* if it is closed, nonempty, invariant (i.e., $T(M) \subseteq M$) and contains no proper subset with these three properties. A nonempty closed set $M \subseteq X$ is minimal if and only if the orbit of every point of M is dense in M. A point is called *minimal* or *almost periodic* if it belongs to a minimal set, and a dynamical system (X, T), or the map T, is called minimal if the set X is minimal. A point $x \in X$ is called *recurrent* if for some sequence $n_i \nearrow \infty$, $\lim T^{n_i} x = x$; clearly every minimal point is recurrent.

A *transitive* system (X, T) is one for which with every pair of nonempty open sets U and V in X there is a positive integer n such that $T^n(U) \cap V \neq \emptyset$; or equivalently if there is a point $x \in X$ with dense orbit in X. Any point with dense orbit is called a *transitive point*. In a transitive system the set of transitive points is a dense G_δ subset of X and we denote it by X_{trans}. If there are no isolated points in a transitive system (X, T) then the set of points that are not transitive is either empty or dense (equivalently: if the set of transitive points has nonempty interior then the system is minimal). For more details about transitivity see [23].

A system (X, T) is said to be *topologically weakly mixing* if its cartesian square $(X \times X, T \times T)$ is transitive. A subset E of X is called *independent* for (X, T) if for any finite collection $x_1, x_2, \ldots, x_n$ of distinct points of E the point $(x_1, x_2, \ldots, x_n)$ is transitive in the product system $(X,T)^n$ $((X \times ... \times X, T \times ... \times T)$, n times). In every weakly mixing system there exists an uncountable independent set ([22], using the main theorem in [28]).

Consider a topological dynamical system (X,T). For any finite cover $\mathcal{C}$ of the compact space X, let $\mathcal{N}(\mathcal{C})$ be the minimal cardinality of a sub-cover of $\mathcal{C}$. The *topological complexity function* of the finite cover $\mathcal{C}$ of (X,T) is the non-decreasing function

$$c_T(\mathcal{C}, n) = \mathcal{N}(\mathcal{C}^n),$$

where $\mathcal{C}^n = \mathcal{C} \bigvee T^{-1}\mathcal{C} \bigvee ... \bigvee T^{-(n-1)}\mathcal{C}$. The exponential growth rate of this function is the topological entropy of $\mathcal{C}$, and the *topological entropy* of (X,T), $h(X,T)$, is the supremum of the topological entropies of finite open covers (see [10]).

Two topological dynamical systems (X,T) and (Y,S) are called *disjoint* if there exists no proper closed $T \times S$–invariant subset of the cartesian product $X \times Y$ with projections X and Y on the two coordinates; such a subset is called a *joining*. Recall that when two systems are disjoint, at least one of them is minimal.

The following notions were introduced in [7]. A dynamical system (X,T) is called *scattering* if any finite cover $\mathcal{C}$ by non–dense open sets has unbounded complexity, i.e., $c_T(\mathcal{C}, n) \to \infty$, *2–scattering* if the same condition holds but for 2–set covers only. A system is scattering if and only if its cartesian product with any minimal system is transitive [7]. Weak mixing implies scattering; scattering is strictly weaker than weak mixing [3] but the two properties are equivalent in the minimal case.

In the measure–theoretical framework, given a T–invariant measure μ, the entropy of T with respect to μ is defined as $h_\mu(X,T) = \sup h_\mu(\alpha, T)$ where the supremum ranges over the set of all finite measurable partitions α of X, $h_\mu(\alpha, T) = \lim\limits_{n\to\infty} \dfrac{1}{n} H_\mu(\alpha^{n-1})$. Topological entropy and measure–theoretic entropy are related by the *Variational Principle:* $h(X,T) = \sup\limits_\mu h_\mu(X,T)$ [10].

1.2. About chaos.

One family of definitions of chaos is based on the instability of trajectories. Given $\varepsilon > 0$, the map T is called *Lyapunov ε–unstable at $x \in X$* if for every neighborhood U of x, there is $y \in U$ and $n \geq 0$ with $\varrho(T^n x, T^n y) > \varepsilon$; T is called *unstable at x* (or the point x itself is called unstable) if there is $\varepsilon > 0$ such that T is Lyapunov ε-unstable at x.

If a dynamical system is pointwise unstable sometimes there exists no $\varepsilon > 0$ such that all points are ε–unstable. But if in addition the system is transitive then pointwise instability (in fact the instability of a point with dense orbit) implies uniform pointwise instability, i.e., the existence of such a universal $\varepsilon > 0$; this last property is also called *sensitive dependence on initial conditions* (sensitivity for short). It is easy to see that weak mixing implies sensitivity.

Auslander and Yorke [5] call chaotic a system that is both sensitive and transitive. Any non–invertible transitive system, or transitive non–minimal system with a dense set of minimal points, is chaotic in the sense of Auslander and Yorke [2]. Devaney's definition of chaos [11] is Auslander and Yorke's with the extra assumption of the existence of a dense set of periodic points (see, for example, [19]).

Now consider definitions of chaos based on Li–Yorke pairs. Li–Yorke chaos was defined in the Introduction as the existence of an uncountable scrambled set. Recently it was proved that Auslander-Yorke chaotic systems with at least one periodic point are Li–Yorke chaotic [21]. A dynamical system (X, T) is said to be *generically chaotic* if the set of Li–Yorke pairs contains a dense G_δ subset of $X \times X$. Any transitive generically chaotic system is Li–Yorke chaotic [21].

It is interesting to compare Li–Yorke and Auslander–Yorke chaoticity; neither implies the other. Sturmian systems (see Subsection 3.3) are transitive and sensitive to initial conditions, i.e., Auslander–Yorke chaotic, while having no Li–Yorke pairs. On the other hand, there are scattering systems that are not sensitive to initial conditions [3]; these systems are transitive and, by [21], Li–Yorke chaotic.

Weak mixing is a rather strong chaotic property: it implies generic (hence Li–Yorke), and Auslander–Yorke chaos (Proposition 2.12). None of the converse implications is true. Sturmian systems are Auslander–Yorke chaotic without being weakly mixing. The following example shows that weak mixing is not implied by generic chaos either, even assuming transitivity. Let (X, T) be the full shift on 2 letters and (Y, S) be an irrational rotation of the unit circle; let (Z, R) be the factor of $(X \times Y, T \times S)$ obtained by collapsing all elements of the form (x_0, y) where x_0 is the fixed point on 0. One can check that (Z, R) is transitive but not weakly mixing and has a dense set of Li–Yorke pairs with modulus $\delta > 0$, which is G_δ; on the other hand since it contains a fixed point it is Li–Yorke chaotic by [21].

Positive topological entropy and scattering can be considered as a third family of chaotic properties, based on the complexity of covers. Scattering, actually even 2–scattering, implies transitivity [7], Li–Yorke chaos and generic chaos [21].

If (X, T) is transitive but not sensitive, T has a property called uniform rigidity ([19], [2]). It implies that T is a homeomorphism and that it has zero entropy. Therefore any transitive system with positive entropy has sensitive dependence on initial conditions; it is thus Auslander–Yorke chaotic.

Recall that a minimal scattering system is weakly mixing, therefore chaotic in many senses. There are however many examples of weakly mixing systems with zero entropy (for example, any uniquely ergodic system which is, measure theoretically, weak mixing and has simple spectrum) and even minimal and uniformly rigid ones [18]. Positive entropy is another very strong notion of chaoticity. For interval maps all the existing definitions of chaos coincide more or less.

Scattering as well as Auslander–Yorke chaos have a kind of uniformity that positive entropy and Li–Yorke chaos do not possess. The last two may be considered to indicate the existence of 'a certain amount of chaos' in a dynamical system, while the first two describe 'everywhere chaotic behavior'.

Remark 1.1. Other definitions based on Li–Yorke pairs but having some common features with sensitivity may be worth considering. For instance let us say that a dynamical system (X,T) has *chaotic dependence on initial conditions* if for any $x \in X$ there is $\delta > 0$ such that for every neighborhood U of x there is $y \in U$ such that the pair (x,y) is Li–Yorke with modulus δ; when (X,T) is transitive then δ is the same for all x. A dynamical system (X,T) is said to have *weakly chaotic dependence on initial conditions* if for any $x \in X$ and every neighborhood U of x there are $y,z \in U$ such that the pair $(y,z) \in X^2$ is Li–Yorke (in other words, Li–Yorke pairs are dense near the diagonal). A dynamical system (X,T) is said to be *(weakly) ST chaotic* if it is transitive and depends (weakly) chaotically on initial conditions.

ST chaos implies Auslander–Yorke chaos. Positive topological entropy and transitivity do not imply ST chaos: there are point distal maps (which cannot be ST chaotic) with positive topological entropy. E. Akin brought to our attention the fact that a transitive system depends weakly chaotically on initial conditions if and only if it is not almost distal [1].

2. Existence of Li–Yorke pairs and Li–Yorke chaos

The main result of this section is that positive entropy implies Li–Yorke chaos. We will need some facts from the general theory of measure entropy pairs as developed in [6] and [16]. Given a topological dynamical system (X,T) endowed with a T–invariant and ergodic probability measure μ, we consider the set $E_\mu(X,T) \subset X \times X$ of entropy pairs for μ. If $h_\mu(X,T) > 0$ then this is a non-empty set. If $\varphi : (X', \mathcal{X}', \mu', T') \to (X, \mathcal{X}, \mu, T)$ is a topological factor map and (x,y) is an entropy pair for μ, there exists $(x',y') \in \varphi^{-1}(x,y)$ that is an entropy pair for μ', [6]. When T is invertible the closure $\overline{E}_\mu(X,T)$ of the set of entropy pairs for μ is characterized in [16] as follows. Call λ the independent product of μ with itself over the Pinsker factor $\Pi = (Y, \mathcal{Y}, \nu, S)$ of the (measure theoretical) system $(X, \mathcal{X}, \mu, T)$; then $\overline{E}_\mu(X,T)$ is the topological support of λ. Since μ is ergodic, so is λ (this follows from the fact that the extension $(X, \mathcal{X}, \mu, T) \to \Pi$ is a weakly mixing extension) and consequently $T \times T$ acts transitively on $\overline{E}_\mu(X,T)$.

In what follows, given a set X, we denote by Δ_X the diagonal of the cartesian product $X \times X$. Let X be a complete metric space. Call $K \subseteq X$ a *Mycielski set* if it has the form $K = \bigcup_{j=1}^\infty C_j$ with C_j a Cantor set for every j. For the reader's convenience we restate here a version of Mycielski's theorem ([28], Theorem 1) which we shall use.

Mycielski's Theorem. *Let X be a complete metric space with no isolated points. Suppose that for every natural number $n \in \mathbb{N}$, R_n is a meager subset of X^{r_n}, and let G_j, $j = 1,2,\dots$ be a sequence of non-empty open subsets of X. Then there exists Cantor subsets $C_j \subset G_j$ such that for every $n \in \mathbb{N}$ the Mycielski set $K = \bigcup_{j=1}^\infty C_j$ has the property that for every $x_1, x_2, \dots, x_{r_n}$ distinct elements of K, $(x_1, x_2, \dots, x_{r_n}) \notin R_n$.*

Theorem 2.1. *Let (X,T) be a topological dynamical system and assume that for some T–ergodic probability measure μ the corresponding measure–preserving sys-*

tem $(X, \mathcal{X}, \mu, T)$ is not measure distal. Denote $Z = supp(\mu)$; then there exists a closed invariant set $W \subseteq Z \times Z$ such that the subsystem $(W, T \times T)$ is topologically transitive and for every open set $U \subseteq X$ with $U \cap Z \neq \emptyset$ there exists a Mycielski set $K \subseteq U$ which satisfies $(K \times K) \setminus \Delta_Z \subseteq W_{trans}$. Finally, every such set K is a scrambled set of the system (X, T).

Proof. By the Furstenberg–Zimmer theorem [15], [32], [33], the ergodic system $(X, \mathcal{X}, \mu, T)$ admits a unique maximal measure distal factor $\pi : (X, \mathcal{X}, \mu, T) \to (Y, \mathcal{Y}, \nu, S)$ with the factor map π being a weakly mixing extension. Of course π is a measure–theoretical factor map and in general cannot be assumed to be a continuous map onto a topological factor. However, if we let $\mu = \int_Y \mu_y \, d\nu(y)$ be the disintegration of μ over ν and form the relative product measure

$$\lambda := \mu \times_\nu \mu = \int_Y \mu_y \times \mu_y \, d\nu(y),$$

then the measure λ is a well defined regular Borel measure on the space $X \times X$. Moreover, to say that the extension π is a weakly mixing extension is just to say that the measure–theoretical system $(X \times X, \mathcal{X} \otimes \mathcal{X}, \lambda, T \times T)$ is ergodic. Set $W = \mathrm{supp}(\lambda) \subset Z \times Z \subset X \times X$.

Applying the ergodic theorem to the system $(W, \mathcal{W}, \lambda, T \times T)$ we conclude that the subset $W_\lambda \subseteq W$ of λ–generic points in W is of λ measure 1. Since $W = \mathrm{supp}(\lambda)$ we get that the system $(W, T \times T)$ is topologically transitive and that the set W_λ is dense in W_{trans}. Now

$$1 = \lambda(W_\lambda) = \int_Y \left(\int \mathbf{1}_{W_\lambda} \, d\mu_y \times \mu_y \right) \, d\nu(y)$$

and therefore for a subset $Y_\lambda \subseteq Y$ of ν measure 1, one has $\mu_y \times \mu_y(W_\lambda) = 1$ for $y \in Y_\lambda$. Let, for any point $y \in Y_\lambda$, $S_y = \mathrm{supp}(\mu_y) \subseteq X$. Then

$$W_\lambda \cap (S_y \times S_y) \subseteq W_{\mathrm{trans}} \cap (S_y \times S_y) := L.$$

Since clearly $\mu_y \times \mu_y(W_\lambda \cap (S_y \times S_y)) = 1$, it follows that $W_\lambda \cap (S_y \times S_y)$, and a fortiori also $W_{\mathrm{trans}} \cap (S_y \times S_y)$, is a dense subset of $S_y \times S_y$. Since W_{trans} is a G_δ subset of W, it follows that L is a dense G_δ subset of $S_y \times S_y$.

Next observe that the maximality of the distal factor $(Y, \mathcal{Y}, \nu, S)$ implies that for ν–a.e. $y \in Y$ the measure μ_y is non–atomic. In fact, if this is not the case then the ergodicity of ν implies that there exists a positive integer k such that for ν–almost every y the measure μ_y is purely atomic with exactly k points in its support. This in turn implies that the extension $\pi : (X, \mathcal{X}, \mu, T) \to (Y, \mathcal{Y}, \nu, S)$ is an isometric extension and, being also a weakly mixing one, is necessarily trivial (i.e., $k = 1$ and π is an isomorphism). Since we assume that $(X, \mathcal{X}, \mu, T)$ is a non–distal system this proves our assertion. Let $Y_0 \subseteq Y_\lambda$ be the subset, of full measure, consisting of points $y \in Y_\lambda$ for which μ_y is non–atomic.

Applying Mycielski's theorem to the compact perfect space S_y — for any $y \in Y_0$ — we obtain the existence of a Mycielski set $K \subset S_y$ (which can be chosen to be dense in S_y) such that

$$K \times K \setminus \Delta_Z \subseteq L \subseteq W_{\text{trans}}.$$

Moreover, since clearly $\bigcup\{S_y : y \in Y_0\}$ is dense in $Z = \operatorname{supp}(\mu) \subseteq X$, we see that for every open subset $U \subseteq X$ with $U \cap Z \neq \emptyset$, there exists such a K inside U.

The last assertion of the theorem follows from the following observation. Since $\Delta_Z \subseteq W$, it follows that every point (x, x') of W_{trans} is, on the one hand in the proximal relation $\mathbf{P}(X,T)$, and on the other a recurrent point of $(X \times X, T \times T)$ which is not a diagonal point, hence in particular is not an asymptotic point of $(X \times X, T \times T)$. Thus every $K \subset X$ such that $K \times K \setminus \Delta_Z \subseteq W_{\text{trans}}$ is a scrambled set. ■

Remark 2.2. Since for a weakly mixing extension $\pi : (X, \mathcal{X}, \mu, T) \to (Y, \mathcal{Y}, \nu, S)$ and any positive integer $k \geq 2$ the relative product measure

$$\lambda^{(k)} := \int_Y \mu_y \times \mu_y \times \cdots \times \mu_y \, d\nu(y),$$

on $X^k = X \times X \times \cdots \times X$ is ergodic, the same proof will yield dense Mycielski sets $K \subseteq S_y$ with $(x_1, x_2, \ldots, x_n) \in W^{(k)}_{\text{trans}}$ for every $x_1, x_2, \ldots, x_n$ distinct points of K, where no w $W^{(k)}_{\text{trans}}$ is the set of transitive points of the topological system $(W^{(k)}, T \times T \times \cdots \times T)$ with $W^{(k)} = \operatorname{supp}(\lambda^{(k)})$.

Theorem 2.3. *Let (X,T) be a topological dynamical system with $h_{top}(X,T) > 0$. Let μ be a T–ergodic probability measure with $h_\mu(X,T) > 0$ and set $Z = supp(\mu)$. Then there exists a topologically transitive subsystem $(W, T \times T)$ with $W \subseteq Z \times Z$, and such that for every open $U \subseteq X$ with $U \cap Z \neq \emptyset$ there exists a Mycielski set $K \subseteq U$ with $K \times K \setminus \Delta_Z \subseteq W_{trans}$. Moreover*

1. *$K \times K \setminus \Delta_Z \subseteq E_\mu(X,T)$, where $E_\mu(X,T)$ is the set of μ-entropy pairs, and*
2. *K is a scrambled subset of (X,T).*

Proof. By the Variational Principle there exists a T–ergodic probability measure μ with $h_\mu(X,T) > 0$. Fix such a μ and let $Z = \operatorname{supp}(\mu)$. Let $\pi : (X, \mathcal{X}, \mu, T) \to (Y, \mathcal{Y}, \nu, S)$ be the unique maximal zero–entropy measure–theoretical factor of the system $(X, \mathcal{X}, \mu, T)$, i.e., $(Y, \mathcal{Y}, \nu, S)$ is the Pinsker factor of $(X, \mathcal{X}, \mu, T)$. Since $h_\mu(X,T) > 0$ the map π cannot be trivial. Since an isometric extension of a zero–entropy system has still entropy zero we conclude that the extension π is a weakly mixing extension and, using the same notation as in the proof of Theorem 2.1, that ν-a.e. μ_y is non–atomic. We can now apply the proof of Theorem 2.1, with μ as the non–distal measure, to conclude that Z and $W = \operatorname{supp}(\lambda)$, with $\lambda = \mu \times_\nu \mu$, satisfy the conclusions of that theorem.

Since by [16] the set $W \setminus \Delta_Z$ coincides with the set of μ-entropy pairs in X this proves part 1. Part 2 follows as in Theorem 2.1. ■

Recall that a dynamical system (X, T) is called *Li–Yorke chaotic* if X contains an uncountable scrambled set.

Corollary 2.4. *Let (X, T) be a topological dynamical system.*

1. *If (X, T) admits a T-invariant ergodic measure μ with respect to which the measure preserving system $(X, \mathcal{X}, \mu, T)$ is not measure distal then (X, T) is Li–Yorke chaotic.*
2. *If (X, T) has positive topological entropy then it is Li–Yorke chaotic.*

Proof. Follows directly from Theorems 2.1 and 2.3. ∎

Remark 2.5. 1. For (X, T) as in Theorem 2.3, let $\pi_P : (X, \mathcal{X}, \mu, T) \to (Y_P, \mathcal{Y}_P, \nu_P, T_P)$ be the unique maximal zero–entropy measure–theoretical factor of the system $(X, \mathcal{X}, \mu, T)$, that is, $(Y_P, \mathcal{Y}_P, \nu_P, T_P)$ is the Pinsker factor of $(X, \mathcal{X}, \mu, T)$. Since $h_\mu(X, T) > 0$ the map π_P cannot be trivial. Since a measure distal system has zero entropy, it follows that the Furstenberg–Zimmer maximal distal factor $\pi_D : (X, \mathcal{X}, \mu, T) \to (Y_D, \mathcal{Y}_D, \nu_D, T_D)$ is also a factor of $(Y_P, \mathcal{Y}_P, \nu_P, T_P)$ and we conclude that the system $(X, \mathcal{X}, \mu, T)$ is not measure distal. Thus the first assertion of Theorem 2.3 follows directly from Theorem 2.1. However for the proof of part 1 of Theorem 2.3 it is more convenient to work with the Pinsker rather than the Furstenberg–Zimmer factor of $(X, \mathcal{X}, \mu, T)$.
2. In all the assertions above we could have used Kuratowski's theorem [26], rather than Mycielski's. In fact this theorem implies that, say, in the space $S = S_y$, the collection of Cantor sets $K \subseteq S$ such that $K \times K \setminus \Delta_Z \subseteq W_{\text{trans}}$ forms a dense G_δ subset of the compact metric space 2^S of all closed non–empty subsets of S (equipped with the Hausdorff metric). From this we can deduce that if $\mathcal{U} = \{U_1, U_2, \dots, U_k\}$ is a finite open cover of S then there exists a Cantor set $K \subseteq S$ with $K \times K \setminus \Delta_Z \subseteq W_{\text{trans}}$ and such that $K \cap U_i \neq \emptyset$ for every i. Thus given $\varepsilon > 0$ we get, by choosing $\mathcal{U}$ with $\operatorname{diam}(\mathcal{U}) < \varepsilon$, a Cantor set $K \subseteq S$ with $K \times K \setminus \Delta_Z \subseteq W_{\text{trans}}$, which in addition is ε-dense in S. Of course such a K cannot be dense in S.

A theorem of Kuratowski and Ulam (see for example [29], page 56) says that if $R \subseteq X \times X$ is residual then for a residual subset $A \subseteq X$, $x \in A$ implies that $R(x)$ is a residual subset of X. Conversely if R satisfies this condition and R has the Baire property (every Borel or even analytic subset has the Baire property, see e.g. [26]) then it is easy to see that R is residual. In fact if we let $L = X \times X \setminus R$ then, being a Baire set, L can be expressed as $U \Delta M$ with U open and M meager in $X \times X$. Now by the Kuratowski–Ulam theorem $M(x)$ is a first category subset of X for x in a residual subset B of X, and if we choose non–empty open sets $V, W \subseteq X$ with $V \times W \subseteq U$ then for $x \in A \cap B \cap V$ we have:

$$\emptyset = R(x) \cap L(x) \supseteq R(x) \cap (W \setminus M(x)).$$

Since for $x \in A$ $R(x)$ is residual in X, this is impossible and we conclude that $U = \emptyset$, that $L = M$ is meager and finally that R is indeed residual.

Huang and Ye prove in [21] the following lemma (Lemma 3.1. there) by a transfiniteness argument:

Lemma 2.6. *Assume that X is a compact metric space without isolated points. If R is a symmetric relation with the property that there is a dense G_δ subset A of X such that for each $x \in A$, $R(x)$ contains a dense G_δ subset, then there is a dense subset B of X with uncountably many points such that $B \times B \setminus \Delta_X \subseteq R$.*

In view of the above remark one can use Mycielski's theorem, to obtain the following version.

Lemma 2.7. *Assume that X is a compact metric space without isolated points. If R is a symmetric residual subset of $X \times X$, then there is a Mycielski set $K \subseteq X$ which is dense in X and such that $K \times K \setminus \Delta_X \subseteq R$.*

One can then apply this modified lemma to obtain "Cantor versions" of some of the assertions of [21], for example the following one (corresponding to Corollary 4.1 of [21]).

Theorem 2.8. *Assume that (X,T) is a transitive non–periodic dynamical system containing a periodic point. Then there exists a Mycielski scrambled set for T.*

Iwanik proves the existence of Mycielski independent sets for topologically weakly mixing systems [22] (independent sets are defined in Subsection 1.1). The following theorem is a generalization of his result and has essentially the same proof.

Theorem 2.9. *Let (X,T) be a topological dynamical system and $\pi : (X,T) \to (Y,S)$ be an open surjective homomorphism such that the closed $T \times T$-invariant subset*

$$W = R_\pi = \{(x,x') \in X \times X : \pi(x) = \pi(x')\}$$

(i) is not equal to the diagonal Δ_X (i.e., π is not one–to–one), (ii) has no isolated points and (iii) is topologically transitive. Then there exists a residual subset $Y_0 \subseteq Y$ such that for every $y \in Y_0$ the set $X_y = \pi^{-1}(y)$ contains a dense Mycielski set K with $K \times K \setminus \Delta_X \subseteq W_{trans}$. In particular every such set is an independant set and (X,T) is Li–Yorke chaotic.

Remark 2.10. Since a structure theorem for minimal systems, analogous to the Furstenberg–Zimmer structure theorem exists (with a maximal PI factor replacing the maximal measure distal factor, see for example [17]), Theorem 2.1 above suggests that the following statement may hold: every minimal system (X,T) which is not a PI system admits a Mycielski set $K \subseteq X$ with $K \times K \setminus \Delta_X \subseteq W_{\mathrm{trans}}$ for some $W \subseteq X \times X$ satisfying properties (i)-(iii) of Theorem 2.9, and in particular is Li–Yorke chaotic. However a closer look at the topological structure theorem shows that one can obtain an extension $\pi : (\tilde{X},T) \to (Y,S)$, as in Theorem 2.9, only for a proximal extension $\tilde{X}$ of X. Thus such a theorem, if true, does not follow immediately from the structure theorem.

Proposition 2.11. *A weakly mixing topological dynamical system (X,T) is generically, Li–Yorke, Auslander–Yorke and weakly ST chaotic.*

Proof.. Since the set of transitive points is a dense G_δ set in X^2, the system is generically chaotic. The existence of an uncountable independent set for any weakly mixing system is proved in [22]; an independent set is obviously scrambled, therefore (X,T) is also chaotic in the sense of Li and Yorke and weakly ST chaotic. Auslander–Yorke chaos results from the fact that any weakly mixing system is sensitive. ■

Remark 2.12. Recall that a minimal scattering system is weakly mixing and has all the properties stated in the last proposition. Recently Huang and Ye proved that 2–scattering implies Li–Yorke chaos and the density of Li–Yorke pairs [21]. Here is an easy way to deduce the existence of Li–Yorke pairs when (X,T) is a non–trivial scattering system: let Y be a non–empty minimal subset of X. By [7] the product $(X\times Y, T\times T)$ is transitive. It contains $Y\times Y$, therefore intersecting the diagonal; since (X,T) is non–trivial, X is not reduced to a fixed point and $X\times Y$ is not contained in Δ_X. Let (x,y) be a transitive point in $X\times Y$: its orbit intersects any neighborhood of $X\times Y\cap\Delta_X$, hence $\liminf_{n\to\infty}\varrho(T^nx,T^ny)=0$; it also intersects any neighborhood of a non–diagonal pair, so that $\limsup_{n\to\infty}\varrho(T^nx,T^ny)>0$. Thus (x,y) is a Li–Yorke pair.

3. Almost distal systems

In this section we consider systems without Li–Yorke pairs, that is, dynamical systems for which every pair of points is either asymptotic or distal; we call them *almost distal*. Most of their properties can be obtained in two ways. One consists in applying the classical theory of Ellis semigroups to invertible almost distal systems, and then use natural extensions in order to prove the same properties for non–invertible systems; the proofs this method leads to are simple and nice in some instances, extremely awkward in others. The second method relies on a new variant of the Ellis semigroup theory, adapted to the non–invertible case. It is this variant that we sketch now.

3.1. The adherence semigroup.

The theory of the *Ellis* or *enveloping* semigroup of homeomorphisms is developed at length in [4], [9]. Suppose (X,T) is a surjective dynamical system (not necessarily invertible). Consider the set X^X of self-maps of X, endowed with the topology of pointwise convergence; it is compact Hausdorff. Define $\mathcal{A}(X,T)\in X^X$ to be the set of all pointwise limits of subnets of $(T^n), n\in\mathbb{N}$. When the meaning is clear we just write $\mathcal{A}$. Alternatively $\mathcal{A}$ can be defined as the intersection of the sets $\mathcal{A}_m=\overline{\{T^n, n\geq m\}}$, where the closure is taken in the topology of pointwise convergence.

It is essential to remark that a pair $(x,y)\in X\times X$ is proximal if and only if $px=py$ for some $p\in\mathcal{A}$, and asymptotic if and only if $px=py$ for any $p\in\mathcal{A}$.

There are two differences with the Ellis semigroup: T is not assumed to be a homeomorphism; and we do not want T and its powers to belong to $\mathcal{A}$ unless they actually belong to the orbit adherence of T. The motivations will be obvious after reading this section; essentially, we want T to act surjectively on the compact set $\mathcal{A}$. In the following proofs it is important to note that in general T need not belong to $\mathcal{A}$. Otherwise we follow [4] closely.

Proposition 3.1. *$\mathcal{A}(X,T)$ is a compact semigroup, on which the right multiplication is continuous.*

Proof. The continuity of the right multiplication follows immediately from the fact that if $p_n \to p$ pointwise, then $p_n(q(x)) \to p(q(x))$ for any $q \in X^X$. The set $\mathcal{A}$ is compact because it is closed in the compact topological space X^X. It is a semigroup: for any $n \in \mathbb{N}$, any $p \in \mathcal{A}$, the map $T^n \circ p$ belongs to $\mathcal{A}$ by continuity of T and the fact that $p \circ T^n = T^n \circ p$. For any $p, q \in \mathcal{A}$ take a net $\{\xi_n\}$ converging to p: then by continuity of the right multiplication $\{\xi_n\} \circ q$ tends to $p \circ q$ while belonging to $\mathcal{A}$, which means that $p \circ q \in \mathcal{A}$. ∎

Proposition 3.2. *(1) The formula $T(p) = p \circ T$ defines a continuous surjective transformation of $\mathcal{A}(X,T)$.*
(2) Let $\varphi\colon (X,T) \to (Y,S)$ be a factor map. Then there is a unique semigroup morphism Φ: $(\mathcal{A}(X,T),T) \to (\mathcal{A}(Y,S),S)$ such that $\Phi(p)(\varphi(x)) = \varphi(p(x))$ for $x \in X$ and $p \in \mathcal{A}(X,T)$; moreover $\Phi \circ T = S \circ \Phi$.

Proof. (1) Continuity results from Proposition 3.1; we prove surjectivity. Let $\{T^{n_k}\} \to q$; using compactness take a converging subnet of $\{T^{n_k-1}\}$ with limit p, then $T(p) = q$.
(2) The proof is similar to that of [4], Chapter 3, Theorem 7. ∎

Proposition 3.3. *A minimal left ideal of $\mathcal{A}(X,T)$ is a closed subsemigroup of $\mathcal{A}(X,T)$. Any closed subsemigroup of $\mathcal{A}(X,T)$ contains an idempotent element.*

Proof. The proof is the same as that of Corollaries 7 and 8 of [4], Chapter 6. ∎

Proposition 3.4. *The set $I \subseteq \mathcal{A}(X,T)$ is a minimal left ideal if and only if (I,T) is minimal as a dynamical system. In particular a minimal left ideal is closed.*

Proof. (1) Suppose first that I is a minimal left ideal of $\mathcal{A}(X,T)$. We prove first that $T(I) = I$. It is not hard to check that $T(I)$ is also a minimal left ideal. Let $p \in I$, $q \in \mathcal{A}$; since by Proposition 3.2(1) T acts surjectively on $\mathcal{A}$ there is $q' \in \mathcal{A}$ such that $q \circ p = T(q' \circ p)$, which shows that $I \cap T(I)$ is not empty: thus by minimality of I and $T(I)$ one concludes that $I = T(I)$. One must also show that I is closed. For any $p \in I$, $\mathcal{A} \circ p$ is a subideal of I so it is equal to I. Closure follows from the facts that the right multiplication is continuous and $\mathcal{A}$ is closed. Finally recall that $\mathcal{A} \circ p = \{(\lim\{\xi_n\}) \circ p \mid \{\xi_n\} \text{ converging net}\} = I$: this means that the orbit closure of any $p \in I$ is equal to I.
(2) Suppose $I \subseteq \mathcal{A}$ is such that (I,T) is minimal. Then I is closed and $T(I) \subset I$. For $p \in I$, $q \in \mathcal{A}$, one has $q \circ p = (\lim\{\xi_n\}) \circ p = \lim(\{\xi_n\} \circ p)$, which belongs to I by minimality of (I,T); thus I is a left ideal. If it is not minimal it contains a

minimal subideal J, and by the proof above (J,T) is a minimal system, which is a contradiction. ■

Proposition 3.5. *Let I be a minimal left ideal of $\mathcal{A}(X,T)$, and let $x \in X$. Then $I(x)$ is a minimal subset of X.*

Proof. Put $I(x) = \{p(x) \mid p \in I\}$. This set is the homomorphic image of (I,T) under the valuation map $\varphi_x : I \to X$, which commutes with T. Since (I,T) is minimal so is $(I(x),T)$. ■

Proposition 3.6. *There is a unique minimal left ideal in $\mathcal{A}(X,T)$ if and only if the proximality relation is an equivalence relation.*

Proof. Assume that there is a unique minimal left ideal $I \subseteq \mathcal{A}$ and consider $(x,y) \in \mathbf{P}$, $(y,z) \in \mathbf{P}$. Then there are minimal left ideals I' and I'' such that for $p \in I', q \in I''$ one has $p(x) = p(y)$ and $q(y) = q(z)$. Since there is a unique minimal left ideal, $I' = I'' = I$ and for any $p \in I$, $p(x) = p(y) = p(z)$. Thus (x,z) is proximal.

Conversely, assume that the proximal relation is an equivalence relation. Then for any two idempotents u, v in $\mathcal{A}$ and $x \in X$ the points $u(x)$ and $v(x)$, each proximal to x, are proximal. Now consider two minimal left ideals I and I' and using Proposition 3.3 choose an idempotent $u \in I$. Since $\{p \circ u \ : \ p \in I'\}$ is an ideal contained in I it is equal to I. It follows that there is $v' \in I'$ such that $v' \circ u = u$. The set $\{p \in I' \ : \ p \circ u = u\}$ is thus a non–empty closed subsemigroup, which by Proposition 3.3 contains an idempotent v. Then $(u(x), v(x)) \in I'((u(x), v(x)))$, which means by Proposition 3.5 and proximality of the pair that its orbit closure defines a minimal system intersecting the diagonal. We conclude that $u(x) = v(x)$ for arbitrary x, and $I \cap I' = I = I'$. ■

Proposition 3.7. *For any $x \in X$ there is an almost periodic point $y \in X$ such that $(x,y) \in \mathbf{P}(X,T)$.*

Proof. Let I be a minimal left ideal of $\mathcal{A}$ and $u \in I$ be an idempotent. Then $u(x) \in I(x)$, which is a minimal set under the action of T by Proposition 3.5, so $x' = u(x)$ is almost periodic. On the other hand since $u(x) = u(x')$ the pair (x,x') is proximal. ■

Theorem 3.8. *A system (X,T) is almost distal if and only if $(\mathcal{A}(X,T),T)$ is minimal.*

Proof. Assume that (X,T) is almost distal. Let I be a minimal ideal of $\mathcal{A}$ and u an idempotent in I (Proposition 3.3). Then $(x, u(x))$ is proximal, thus asymptotic for any $x \in X$, that is, for $q \in \mathcal{A}$, $q(x) = q(u(x)) = q \circ u(x)$, so $q = q \circ u$ belongs to I and $I = \mathcal{A}$. Then Proposition 3.4 says that $(\mathcal{A},T)$ is a minimal system.

Conversely, if $(\mathcal{A},T)$ is minimal then $\mathcal{A}$ is a minimal ideal by Proposition 3.4. Suppose (x,y) is a Li–Yorke pair. Since it is a proximal pair there is a minimal left ideal $I \subset \mathcal{A}$ such that $q(x) = q(y)$ for all q in I. Since x and y are not asymptotic, there is an $r \in \mathcal{A}$ such that $r(x) \neq r(y)$. Thus $r \notin I$, and the adherence semigroup is not minimal. ■

This is the first point of similarity that we observe between distal and almost distal systems: recall that (X, T) is distal if and only if the Ellis semigroup $\mathcal{E}(X, T)$ is minimal, or equivalently if and only if $X \times X$ is a union of minimal sets. Correspondingly (X, T) is almost distal if and only if $\mathcal{A}$ is minimal, or equivalently if and only if all ω–limit sets in $X \times X$ are minimal. Other similarities appear below.

3.2. Properties of almost distal systems.

Theorem 3.9. *Let (X, T) be almost distal and let $\pi : (X, T) \to (Y, S)$ be a factor map. Then (Y, S) is almost distal.*

Proof. By Theorem 3.8 $(\mathcal{A}(X, T), T)$ is minimal. By Proposition 3.2(2) the factor map π extends to a morphism from $(\mathcal{A}(X, T), T)$ onto $(\mathcal{A}(Y, S), S)$. It follows that $(\mathcal{A}(Y, S), S)$ is minimal, which, applying the converse part of Theorem 3.8, shows (Y, S) to be almost distal too. ■

It is worth remarking that any subsystem of an almost distal system is almost distal too.

Theorem 3.10. *Any transitive almost distal system is minimal.*

Proof. Let (X, T) be a transitive system and $x \in X$ be a transitive point. By Proposition 3.7 x is proximal to some almost periodic point y, i.e., the ω-limit set of y is a minimal set M of X. Because there are no Li–Yorke pairs, x is asymptotic to y. Therefore the ω-limit set of x belongs to M. Since x is a transitive point, M must be equal to X. ■

Proposition 3.11. *Let $\{(X_i, T_i) : i \in I\}$ be a family of almost distal systems, then the product $(X, T) = \prod_{i \in I}(X_i, T_i)$ is almost distal. In particular an inverse limit of almost distal systems is almost distal, and natural extensions preserve almost distality.*

Proof. Note that a system (X, T) is almost distal if and only if for all $x, y \in X$ and $p \in \mathcal{A}$, if $px = py$ then also $qx = qy$ for every $q \in \mathcal{A}$.

Now with obvious notation given $x, y \in X$, $p \in \mathcal{A}(X, T)$ with $px = py$ one has $px_i = py_i$ for any $i \in I$; then by almost distality $qx_i = qy_i$ for any $q \in \mathcal{A}(X_i, T_i)$, $i \in I$ and finally $qx = qy$ for $q \in \mathcal{A}(X, T)$.

To finish remark that inverse limits and natural extensions are subsystems of product systems. ■

Now we address the question of disjointness between minimal almost distal systems and scattering systems.

Theorem 3.12. *Any transitive almost distal system is disjoint from all scattering systems (in particular from all weakly mixing systems).*

Proof. By Theorem 3.10 a transitive almost distal system (X, T) is minimal, which implies that for any scattering system (Y, S) the Cartesian product $(X \times Y, T \times S)$ is transitive [7].

Let $(x,y) \in (X \times Y)_{\text{trans}}$, and let $J \subset X \times Y$ be any joining. Set

$$\mathcal{S} = \{p \in \mathcal{A}(X \times Y, T \times S) : \, p(x,y) \in J \text{ and } p(y) = y\}.$$

Because (x,y) is transitive and the projection of J on the second coordinate is equal to Y, $\mathcal{S}$ is not empty. It is a closed semigroup: first for p, $q \in \mathcal{S}$ one has $q \circ p(y) = q(y) = y$, and $p_i(y) = y$ converges to $p(y) = y$ if the net (p_i) converges to p; on the other hand since J is closed invariant $p(x,y) \in J$ and $q(x,y) \in J$ imply $q \circ p(x,y) \in J$; if (p_i) converges to p and $p_i(x,y) \in J$ then $p(x,y) \in J$ because J is closed.

It follows that there exists an idempotent v in $\mathcal{S}$. As v belongs to $\mathcal{S}$ one has $v(x,y) = (v(x),y) \in J$ and since v is an idempotent it sends both x and $v(x)$ to $v(x)$, so $(x, v(x)) \in \mathbf{P}(X,T)$ and by our assumption $(x,v(x)) \in \mathbf{As}(X,T)$. Now since x and $v(x)$ are asymptotic the pair $(v(x),y) \in J$ is transitive just like (x,y), so $J = X \times Y$. This means that the two systems are disjoint. ■

Our definitions, so far, worked for $\mathbb{Z}^+$-systems (X,T); even when the map T is invertible we only acted on it with elements of $\{T^n : n \in \mathbb{Z}^+\}$. For the next observation let us consider $\mathbb{Z}$-systems. Again we call the $\mathbb{Z}$-system (X,T) *almost distal* when every proximal pair $(x,y) \in X$ is doubly asymptotic, that is, $\lim_{|n|\to\infty} d(T^n x, T^n y) = 0$. We conclude this subsection by the remark that transitive almost distal $\mathbb{Z}$-systems, which are minimal by Theorem 3.10, have a Proximal Isometric (PI) structure as defined in [12].

Let (X,T) be a minimal dynamical system. (X,T) is said to be *strictly proximal isometric* or *strictly PI* if it can be obtained from the trivial system by a (countable) transfinite succession of proximal and isometric extensions, that is, there is a projective system of minimal dynamical systems $\{(X_\lambda, T_\lambda)\}_{\lambda \leq \theta}$, for some (countable) ordinal θ, where $X_\theta = X$, X_0 is the trivial system, $(X_\lambda, T_\lambda) = \lim_{\zeta<\lambda}(X_\zeta, T_\zeta)$ for a limit ordinal $\lambda \leq \theta$ and such that for $\lambda < \theta$ the projection map $\pi_{\lambda,\lambda+1} : X_{\lambda+1} \to X_\lambda$ defines either a proximal or an isometric extension. (X,T) is said to be *proximal isometric* or *PI* if it has a strictly PI proximal extension.

Proposition 3.13. *Any transitive almost distal $\mathbb{Z}$-system (X,T) is a PI system.*

Proof. (X,T) is minimal by Theorem 3.10. Therefore, since the proximal relation of (X,T) coincides with its asymptotic relation, which is an equivalence relation, we conclude by [12], Proposition 8.8, that the system is PI. ■

Remark 3.14. It is very likely that Proposition 3.13 holds for $\mathbb{Z}^+$-systems as well. However, in order to check this one needs to find whether the theory required for the proof of [12], Proposition 8.8, goes through for $\mathbb{Z}^+$-systems rather than $\mathbb{Z}$-systems. This we did not do. In the next section we shall see that, e.g., the (invertible) Morse minimal set (M,σ) is an example of an almost distal $\mathbb{Z}^+$-system which is not an almost distal $\mathbb{Z}$-system. It is easy to see that the proximal relation $\mathbf{P}_{\mathbb{Z}^+}(M,\sigma) = \mathbf{As}_{\mathbb{Z}^+}(M,\sigma)$ is an equivalence relation but is not closed. (M,σ) is thus an example of a minimal almost distal $\mathbb{Z}^+$-system which is not an asymptotic extension of a distal system. One important question that we leave open in the present paper is whether or not every minimal almost distal $\mathbb{Z}$-system is an asymptotic extension of a distal system.

3.3. Examples.

Here we briefly present some examples of almost distal but non–distal minimal systems. The first example illustrates the fact that an asymptotic lift of a distal system is almost distal.

Example 3.15. Sturmian systems provide the simplest example. Let Y be the 1–torus and S_α be the rotation by the irrational number $\alpha \in Y$. Code the orbits of (Y, S_α) according to the closed cover $\mathcal{C} = \{[0, 1-\alpha], [1-\alpha, 1]\}$ of Y. Call X the closed set of all bi–infinite $\mathcal{C}$–names and let σ be the shift. All points of Y have only one $\mathcal{C}$–name in X except 0 and all points of its orbit, which have two each; on the other hand there is a continuous map $\varphi \colon X \to Y$ sending an infinite name to the point of Y it was constructed from, and $S_\alpha \circ \varphi = \varphi \circ \sigma$.

It is not hard to check that if $x \neq y \in Y$, then the pair (x, y) is distal unless $\varphi(x) = \varphi(y)$; in the last case (x, y) is asymptotic under the actions of σ and σ^{-1}: (X, σ) is a finite–to–one asymptotic extension of the isometry (Y, S_α).

On the other hand, not all almost distal systems are obtained as asymptotic extensions of distal systems.

Example 3.16. For the Morse minimal subshift (M, σ) with $M \subset \{0, 1\}^{\mathbb{Z}}$, viewed as a $\mathbb{Z}^+$–system, the asymptotic and proximal relations coincide and form an equivalence (but not closed) invariant relation. Thus there are no Li–Yorke pairs and the system is almost distal. But two points forming a distal pair are mapped to the same point of the maximal distal factor (the dyadic adding machine). Viewed as a $\mathbb{Z}$ action the Morse system has no non–diagonal asymptotic points; some pairs are asymptotic under σ and distal under σ^{-1}, or the other way round.

Remark that Sturmian and Morse subshifts, when considered only on one–sided sequences, are non–invertible almost distal systems.

Example 3.17. Let T be defined on $Z = Y \times Y$ by $T(x, y) = (S_\alpha(x), S_x(y))$; the dynamical system (Z, T) is distal without being conjugate to a compact group rotation. One can apply the same construction as above by coding the first coordinate, thus obtaining a finite–to–one asymptotic extension of (Z, T).

Example 3.18. A wide class of finite–to–one asymptotic extensions of rotations or derived transformations is briefly described in [9], Section 37. It deserves to be studied in detail; this would provide matter for a complete article.

Example 3.19. Wide classes of almost one–to–one, not finite–to–one, asymptotic extensions of rotations on the 2-torus which are minimal point distal homeomorphisms (respectively minimal point distal noninvertible maps) on that torus are described in [30] (respectively [24]).

References

[1] *E. Akin*, Personal communication.

[2] *E. Akin, J. Auslander and K. Berg*, When is a transitive map chaotic?, Convergence in ergodic theory and probability (Columbus, OH, 1993), vol. 5, Ohio State Univ. Math. Res. Inst. Publ.,, de Gruyter, Berlin, 1996, pp. 25–40.

[3] *E. Akin and S. Glasner*, Residual properties and almost equicontinuity, to appear in J. d'Anal. Math.

[4] *J. Auslander*, Minimal flows and their extensions, vol. 153, North-Holland Mathematics Studies, North-Holland, Amsterdam, 1988.

[5] *J. Auslander and J. Yorke*, Interval maps, factors of maps and chaos, Tohoku Math. J. **32** (1980), 177–188.

[6] *F. Blanchard, E. Glasner and B. Host*, A variation on the variational principle and applications to entropy *pairs*, Ergodic Th. & Dynam. Sys. **17** (1997), 29–43.

[7] *F. Blanchard, B. Host and A. Maass*, Topological complexity, Ergodic Th. & Dynam. Sys. **20** (2000), 641-662.

[8] *F. Blanchard, B. Host and S. Ruette*, Asymptotic pairs in positive–entropy systems, Preprint (2000).

[9] *I.U. Bronstein*, Extensions of minimal transformation groups, (translated from the Russian), Martinus Nijhoff Publishers, The Hague, 1979.

[10] *M. Denker, C. Grillenberger and K. Sigmund*, Ergodic Theory on Compact Spaces, Lecture Notes in Math., vol. 527, Springer–Verlag, Berlin, 1976.

[11] *R. Devaney*, Chaotic dynamical systems, 2d edn., Addison-Wesley, New York, 1989.

[12] *R. Ellis, S. Glasner and L. Shapiro*, Proximal–Isometric flows, Advances Math. **17** (1975), 213–260.

[13] *X. Fu, Y. Fu, J. Duan and R.S. MacKay*, Chaotic properties of subshifts generated by a non-periodic recurrent orbit, Int. J. Bif. Chaos **10** (2000), 1067–1073.

[14] *H. Furstenberg*, Disjointness in ergodic theory, minimal sets and a problem in diophantine approximation, Math. Systems Th. **1** (1967), 1–55.

[15] *H. Furstenberg*, The structure of distal flows , Amer. J. of Math. **85** (1963), 477–515.

[16] *E. Glasner*, A simple characterization of the set of μ-entropy pairs and applications , Israel J. Math. **102** (1997), 13–27.

[17] *E. Glasner*, Proximal flows, Lecture Notes in Mathematics, vol. 517, Springer-Verlag, 1976.

[18] *S. Glasner and D. Maon*, Rigidity in topological dynamics, Ergodic Th. & Dynam. Sys. **9** (1989), 309–320.

[19] *E. Glasner and B. Weiss*, Sensitive dependence on initial conditions, Nonlinearity **6** (1993), 1067–1075.

[20] *W. Huang and XD. Ye*, Homeomorphisms with the whole compacta being scrambled sets, Preprint (1999), to appear in Ergodic Th. & Dynam. Sys.

[21] *W. Huang and XD. Ye*, Devaney's chaos or 2–scattering implies Li–Yorke chaos, Preprint (1999).

[22] *A. Iwanik*, Independent sets of transitive points, Dynamical Systems and Ergodic Theory, vol. 23, Banach Center Publications, 1989, pp. 277–282.

[23] *S. Kolyada and Ľ. Snoha*, Some aspects of topological transitivity – a survey, Grazer Math. Ber., Bericht **Nr. 334** (1997), 3–35.

[24] *S. Kolyada, Ľ. Snoha and S. Trofimchuk*, Noninvertible minimal maps, Preprint (1999), to appear in Fundamenta Mathematicae.

[25] *M. Kuchta and J. Smítal*, Two point scrambled set implies chaos, European Conference on Iteration Theory (ECIT 87), World Sci. Publishing, Singapore, 1989, pp. 427–430.

[26] *K. Kuratowski*, Applications of the Baire–category method to the problem of independent sets, Fundamenta Mathematicae **81** (1973), 65–72.

[27] *T.Y. Li and J.A. Yorke*, Period three implies chaos, Amer. Math. Monthly **82** (1975), 985–992.
[28] *J. Mycielski*, Independent sets in topological algebras, Fundamenta Mathematicae **55** (1964), 139-147.
[29] *J.C. Oxtoby*, Measure and category, Springer-Verlag, 1980.
[30] *M. Rees*, A point distal transformation of the torus, Israel J. Math. **32** (1979), 201–208.
[31] *J. Xiong and Z. Yang*, Chaos caused by a topologically mixing map, Advanced Series in Dynamical Systems, vol. 9, World Scientific, 1990, pp. 550–572.
[32] *R.J. Zimmer*, Extensions of ergodic group actions, Illinois J. Math. **20** (1976), 373-409.
[33] *R.J. Zimmer*, Extensions of ergodic actions and generalized discrete spectrum , Illinois J. Math. **20** (1976), 555-588.

Institut de Mathématiques de Luminy, UPR 9016 du CNRS, case 907, 163 avenue de Luminy, 13288 Marseille Cedex 09, France
E-mail address: blanchar@iml.univ-mrs.fr

Mathematics Department, Tel-Aviv University, Ramat-Aviv, Tel-Aviv 69978, Israel
E-mail address: glasner@math.tau.ac.il

Institute of Mathematics, NASU, Tereshchenkivs'ka 3, 01601 Kiev, Ukraine
E-mail address: skolyada@imath.kiev.ua

Departamento de Ingeniería Matemática and Centro de Modelamiento Matemático, UMR 2071 UCHILE–CNRS, Universidad de Chile, Casilla 170–3, correo 3, Santiago, Chile
E-mail address: amaass@dim.uchile.cl