SCIENCE CHINA
Mathematics

Special Issue on Differentiable Dynamical Systems
• ARTICLES •

September 2020 Vol. 63 No. 9: 1757–1776
https://doi.org/10.1007/s11425-019-1607-5

Regionally proximal relation of order d along arithmetic progressions and nilsystems

Dedicated to Professor Shantao Liao

Eli Glasner[1], Wen Huang[2], Song Shao[2] & Xiangdong Ye[2,*]

[1] *School of Mathematics, Tel Aviv University, Tel Aviv* 6997801*, Israel;*
[2] *CAS Wu Wen-Tsun Key Laboratory of Mathematics and School of Mathematics, University of Science and Technology of China, Hefei* 230026*, China*

Email: eli.glasner@gmail.com, wenh@mail.ustc.edu.cn, songshao@ustc.edu.cn, yexd@ustc.edu.cn

Received May 22, 2019; accepted September 26, 2019; published online August 4, 2020

Abstract The regionally proximal relation of order d along arithmetic progressions, namely $\mathbf{AP}^{[d]}$ for $d \in \mathbb{N}$, is introduced and investigated. It turns out that if (X, T) is a topological dynamical system with $\mathbf{AP}^{[d]} = \Delta$, then each ergodic measure of (X, T) is isomorphic to a d-step pro-nilsystem, and thus (X, T) has zero entropy. Moreover, it is shown that if (X, T) is a strictly ergodic distal system with the property that the maximal topological and measurable d-step pro-nilsystems are isomorphic, then $\mathbf{AP}^{[d]} = \mathbf{RP}^{[d]}$ for each $d \in \mathbb{N}$. It follows that for a minimal ∞-pro-nilsystem, $\mathbf{AP}^{[d]} = \mathbf{RP}^{[d]}$ for each $d \in \mathbb{N}$. An example which is a strictly ergodic distal system with discrete spectrum whose maximal equicontinuous factor is not isomorphic to the Kronecker factor is constructed.

Keywords regionally proximal relation, pro-nilsystem, discrete spectrum, equicontinuous factor

MSC(2010) 37B05, 37A99

Citation: Glasner E, Huang W, Shao S, et al. Regionally proximal relation of order d along arithmetic progressions and nilsystems. Sci China Math, 2020, 63: 1757–1776, https://doi.org/10.1007/s11425-019-1607-5

1 Introduction

This paper is dedicated to the counterpart of the study of multiple ergodic averages in ergodic theory in the setting of topological dynamics. The regionally proximal relation of order d along arithmetic progressions, namely $\mathbf{AP}^{[d]}$ for $d \in \mathbb{N}$, is introduced and investigated.

In some sense an equicontinuous system is the simplest system in topological dynamics. In the study of topological dynamics, one of the first problems was to characterize the equicontinuous structure relation $S_{eq}(X)$ of a system (X, T), i.e., to find the smallest closed invariant equivalence relation $R(X)$ on (X, T) such that $(X/R(X), T)$ is equicontinuous. A natural candidate for $R(X)$ is the so-called regionally proximal relation $\mathbf{RP}(X)$ introduced by Ellis and Gottschalk [10]. By the definition, $\mathbf{RP}(X)$ is closed, invariant, and reflexive, but not necessarily transitive. The problem was then to find conditions under which $\mathbf{RP}(X)$ is an equivalence relation. It turns out to be a difficult problem. Starting with Veech [34],

* Corresponding author

researchers, including MacMahon [31], Ellis and Keynes [11] and Bronstein [5], came up with various sufficient conditions for $\mathbf{RP}(X)$ to be an equivalence relation. Note that in our case where $T: X \to X$ is homeomorphism and (X, T) is minimal, $\mathbf{RP}(X)$ is always an equivalence relation. Using the relative version of equicontinuity, Furstenberg [13] gave the structure theorem of a minimal distal system, which had a very important influence both in topological dynamics and ergodic theory.

The connection between ergodic theory and additive combinatorics was built in the 1970s with Furstenberg's beautiful proof of Szemerédi's theorem via ergodic theory [14]. For a measurable system $(X, \mathcal{X}, \mu, T)$, Furstenberg asked about the convergence (both in the sense of $L^2(\mu)$ and almost surely) of the multiple ergodic averages

$$\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^n x)\cdots f_d(T^{dn}x), \tag{1.1}$$

where $f_1, \ldots, f_d \in L^\infty(X, \mu)$. After nearly 30 years' efforts of many researchers, this problem for the case of L^2-convergence was finally solved in [21, 35]. In their proofs the notion of characteristic factors, introduced by Furstenberg and Weiss [16], plays a great role. Loosely speaking, to understand the multiple ergodic averages $\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^n x)\cdots f_d(T^{dn}x)$, one can replace each function f_i by its conditional expectation with respect to some d-step pro-nilsystem (the 1-step pro-nilsystem is the Kroneker factor). Thus one can reduce the problem to the study of the same average in a nilsystem. In [21], some very useful tools, such as dynamical parallelepipeds, ergodic uniformity seminorms, structure theory involving pro-nilsystems for ergodic systems, were introduced and obtained (for the details we refer to the recent book by Host and Kra [22]).

In the topological setting, Host et al. [23] obtained a topological structure theorem involving pro-nilsystems for all minimal distal systems, which can be viewed as an analog of the purely ergodic structure theory of [21] and the refinement of the Furstenberg's structure theorem for minimal distal systems. In [23], a certain generalization of the regionally proximal relation, namely $\mathbf{RP}^{[d]}$ (the regionally proximal relation of order d), is introduced and used to produce the maximal pro-nilfactors. Precisely, in [23] it is shown that if a system is minimal and distal then $\mathbf{RP}^{[d]}$ is an equivalence relation and $(X/\mathbf{RP}^{[d]}, T)$ is the maximal d-step pro-nilfactor of the system. The maximal pro-nilfactor of order d, namely $(X/\mathbf{RP}^{[d]}, T)$ can be seen as the characteristic factor of the minimal system (X, T). In [33], Shao and Ye showed that all these results in fact hold for arbitrarily minimal systems of abelian group actions. In a recent paper by Glasner et al. [18], the same question is considered for a general group G, and similar results are proved. Applications of the above structure theorems can be found in [24, 26].

Earlier the counterpart of characteristic factors in topological dynamics was studied by Glasner [17] from a different point of view, where the characteristic factors for the action $T \times T^2 \times \cdots \times T^n$ are considered. To be precise, let (X, T) be a topological system and $d \in \mathbb{N}$. Let $\sigma_d = T \times T^2 \times \cdots \times T^d$. (Y, T) is said to be a *topological characteristic factor of order* d if there exists a dense G_δ set Ω of X such that for each $x \in \Omega$ the orbit closure $L = \overline{\mathcal{O}}(x^d, \sigma_d)$ is $\pi \times \cdots \times \pi$ (d times) saturated, where $x^d = (x, \ldots, x)$ (d times) and $\pi: X \to Y$ is the corresponding factor map, i.e., $(x_1, x_2, \ldots, x_d) \in L$ if and only if $(x'_1, x'_2, \ldots, x'_d) \in L$ whenever for all $1 \leqslant i \leqslant d$, $\pi(x_i) = \pi(x'_i)$. In [17], it is shown that if (X, T) is a distal minimal system, then its largest class d distal factor (in the Furstenberg's tower of a minimal distal system) is a topological characteristic factor of order d; if (X, T) is a weakly mixing system, then the trivial system is its topological characteristic factor.

It is a long open question whether for a minimal distal system in Glasner's theorem in [17] one can replace the largest class d distal factor by the maximal pro-nilfactor of order d. Indeed, this is the case where we consider characteristic factors along cubes of minimal systems. In [6], the topological characteristic factors along cubes of minimal systems are studied. It is shown that up to proximal extensions the pro-nilfactors are the topological characteristic factors along cubes of minimal systems. In particular, for a distal minimal system, the maximal $(d-1)$-step pro-nilfactor is the topological cubic characteristic factor of order d [6].

In this paper, we try to give another way to study the counterpart of characteristic factors in topological dynamics. Note that for a minimal system, the maximal pro-nilfactor of order d is obtained by the regionally proximal relation of order d, i.e., $\mathbf{RP}^{[d]}$. Here we propose a direct approach, i.e., we consider the regionally proximal relation of order d along arithmetic progressions, namely $\mathbf{AP}^{[d]}$ for $d \in \mathbb{N}$.

It turns out that if (X,T) is a topological dynamical system with $\mathbf{AP}^{[d]} = \Delta$, then each ergodic measure of (X,T) is isomorphic to a d-step pro-nilsystem, and thus (X,T) has zero entropy. We also show that if (X,T) is a strictly ergodic distal system with the property that the maximal topological and measurable d-step pro-nilsystems are isomorphic, then $\mathbf{AP}^{[d]} = \mathbf{RP}^{[d]}$ for each $d \in \mathbb{N}$. It then follows that for a minimal ∞-pro-nilsystem, $\mathbf{AP}^{[d]} = \mathbf{RP}^{[d]}$ for each $d \in \mathbb{N}$. We construct an example (X,T) which is a uniquely ergodic minimal distal system with discrete spectrum whose maximal equicontinuous factor is not isomorphic to the Kronecker factor.

To finish the introduction we make the following conjecture.

Conjecture 1.1. Let (X,T) be a minimal distal system. Then $\mathbf{AP}^{[d]} = \mathbf{RP}^{[d]}$ for any $d \in \mathbb{N}$.

Unfortunately, we cannot achieve this currently.

2 Preliminaries

2.1 Topological dynamical systems

A *transformation* of a compact metric space X is a homeomorphism of X to itself. A *topological dynamical system* (t.d.s.) or just a *system*, is a pair (X,T), where X is a compact metric space and $T : X \to X$ is a transformation. We use $\rho(\cdot,\cdot)$ to denote a compatible metric in X. In the sequel, and if there is no room for confusion, *in any t.d.s. we will always use T to indicate the transformation.*

A system (X,T) is *transitive* if there exists $x \in X$ whose orbit $\mathcal{O}(x,T) = \{T^n x : n \in \mathbb{Z}\}$ is dense in X and such a point is called a *transitive point*. The system is *minimal* if the orbit of every point is dense in X. This is equivalent to saying that X and the empty set are the only closed invariant subsets of X.

Let (X,T) be a system and let $\mathcal{B}(X)$ be the Borel σ-algebra. Let $\mathcal{M}(X)$ be the set of Borel probability measures in X. A measure $\mu \in \mathcal{M}(X)$ is T*-invariant* if for every Borel set B of X, $\mu(T^{-1}B) = \mu(B)$. Denote by $\mathcal{M}(X,T)$ the set of invariant probability measures. A measure $\mu \in \mathcal{M}(X,T)$ is *ergodic* if for any Borel set B of X satisfying $\mu(T^{-1}B\Delta B) = 0$ we have $\mu(B) = 0$ or $\mu(B) = 1$. Denote by $\mathcal{M}^e(X,T)$ the set of ergodic measures. The system (X,T) is *uniquely ergodic* if $\mathcal{M}(X,T)$ consists of only one element, and it is *strictly ergodic* if in addition it is minimal.

A *homomorphism* between the t.d.s. (X,T) and (Y,T) is a continuous onto map $\pi : X \to Y$ which intertwines the actions; one says that (Y,T) is a *factor* of (X,T) and that (X,T) is an *extension* of (Y,T). One also refers to π as a *factor map* or an *extension* and uses the notation $\pi : (X,T) \to (Y,T)$. The systems are said to be *conjugate* if π is a bijection. An extension π is determined by the corresponding closed invariant equivalence relation

$$R_\pi = \{(x_1,x_2) : \pi(x_1) = \pi(x_2)\} = (\pi \times \pi)^{-1}\Delta_Y \subset X \times X,$$

where Δ_Y is the diagonal on Y.

2.2 Distality and proximality

Let (X,T) be a t.d.s. A pair $(x,y) \in X \times X$ is a *proximal* pair if

$$\inf_{n\in\mathbb{Z}} \rho(T^n x, T^n y) = 0$$

and is a *distal* pair if it is not proximal. Denote by $\mathbf{P}(X,T)$ or $\mathbf{P}_X$ the set of proximal pairs of (X,T). The t.d.s. (X,T) is *distal* if (x,y) is a distal pair whenever $x,y \in X$ are distinct.

An extension $\pi : (X,T) \to (Y,T)$ is *proximal* if $R_\pi \subset \mathbf{P}(X,T)$ and is *distal* if $R_\pi \cap \mathbf{P}(X,T) = \Delta_X$. Observe that when Y is trivial (reduced to one point) the map π is distal if and only if (X,T) is distal.

2.3 Independence

The notion of *independence* was first introduced and studied in [29, Definition 2.1]. It corresponds to a modification of the notion of the *interpolator* studied in [19, 28] and was discussed in depth in [25].

Definition 2.1. Let (X,T) be a t.d.s. Given a tuple $\mathcal{A}=(A_1,\dots,A_k)$ of subsets of X we say that a subset $F\subset\mathbb{Z}$ is an *independence set* for $\mathcal{A}$ if for any nonempty finite subset $J\subset F$ and any $s=(s(j):j\in J)\in\{1,\dots,k\}^J$ we have

$$\bigcap_{j\in J}T^{-j}A_{s(j)}\neq\emptyset.$$

We denote the collection of all independence sets for $\mathcal{A}$ by $\mathrm{Ind}(A_1,\dots,A_k)$ or $\mathrm{Ind}\mathcal{A}$.

Definition 2.2. Let (X,T) be a t.d.s. A pair $(x_1,x_2)\in X\times X$ is called an Ind_{ap}*-pair* (ap for arithmetic progression) if for every pair of neighborhoods U_1, U_2 of x_1 and x_2 respectively, and every $d\in\mathbb{N}$ there is some $n\in\mathbb{N}$ such that for each $(t_1,\dots,t_d)\in\{1,2\}^d$,

$$T^{-n}U_{t_1}\cap T^{-2n}U_{t_2}\cap\cdots\cap T^{-nd}U_{t_d}\neq\emptyset.$$

Denote by $\mathrm{Ind}_{ap}(X,T)$ or $\mathrm{Ind}_{ap}(X)$ the set of all Ind_{ap}-pairs of (X,T).

2.4 Dynamical parallelepipeds

Let X be a set, and let $d\geqslant 1$ be an integer. We view elements in $\{0,1\}^d$ as a sequence $\epsilon=\epsilon_1\cdots\epsilon_d$ of 0's and 1's. We denote X^{2^d} by $X^{[d]}$. A point $\boldsymbol{x}\in X^{[d]}$ can be written as $\boldsymbol{x}=(x_\epsilon:\epsilon\in\{0,1\}^d)$.

Definition 2.3. Let (X,T) be a topological dynamical system and let $d\geqslant 1$ be an integer. We define $\mathbf{Q}^{[d]}(X)$ to be the closure in $X^{[d]}$ of elements of the form

$$(T^{\boldsymbol{n}\cdot\epsilon}x=T^{n_1\epsilon_1+\cdots+n_d\epsilon_d}x:\epsilon=\epsilon_1\cdots\epsilon_d\in\{0,1\}^d),$$

where $x\in X$ and $\boldsymbol{n}=(n_1,\dots,n_d)\in\mathbb{Z}^d$. When there is no ambiguity, we write $\mathbf{Q}^{[d]}$ instead of $\mathbf{Q}^{[d]}(X)$. An element of $\mathbf{Q}^{[d]}(X)$ is called a (dynamical) *parallelepiped of dimension* d.

As examples, $\mathbf{Q}^{[2]}$ is the closure in $X^{[2]}=X^4$ of the set

$$\{(x,T^mx,T^nx,T^{n+m}x):x\in X,m,n\in\mathbb{Z}\}$$

and $\mathbf{Q}^{[3]}$ is the closure in $X^{[3]}=X^8$ of the set

$$\{(x,T^mx,T^nx,T^{m+n}x,T^px,T^{m+p}x,T^{n+p}x,T^{m+n+p}x):x\in X,m,n,p\in\mathbb{Z}\}.$$

Let (X,T) be a system and $d\geqslant 1$ be an integer. The *diagonal transformation* of $X^{[d]}$ is the map $T^{[d]}:X^{[d]}\to X^{[d]}$ defined by $(T^{[d]}\boldsymbol{x})_\epsilon=Tx_\epsilon$ for every $\boldsymbol{x}\in X^{[d]}$ and every $\epsilon\in\{0,1\}^d$.

Definition 2.4. *Face transformations* are defined inductively as follows: Let $T^{[0]}=T$, $T^{[1]}_1=\mathrm{id}\times T$. If $\{T^{[d-1]}_j\}_{j=1}^{d-1}$ is defined already, then the set

$$\begin{aligned}T^{[d]}_j&=T^{[d-1]}_j\times T^{[d-1]}_j,\quad j\in\{1,2,\dots,d-1\},\\ T^{[d]}_d&=\mathrm{id}^{[d-1]}\times T^{[d-1]}.\end{aligned}\tag{2.1}$$

The *face group* of dimension d is the group $\mathcal{F}^{[d]}(X)$ of transformations of $X^{[d]}$ spanned by the face transformations. The *parallelepiped group* of dimension d is the group $\mathcal{G}^{[d]}(X)$ spanned by the diagonal transformation $T^{[d]}$ and the face transformations $\mathcal{F}^{[d]}(X)$. We often write $\mathcal{F}^{[d]}$ and $\mathcal{G}^{[d]}$ instead of $\mathcal{F}^{[d]}(X)$ and $\mathcal{G}^{[d]}(X)$, respectively. For $\mathcal{G}^{[d]}$ and $\mathcal{F}^{[d]}$, we use similar notations to that used for $X^{[d]}$: namely, an element of either of these groups is written as $S=(S_\epsilon:\epsilon\in\{0,1\}^d)$. In particular, $\mathcal{F}^{[d]}=\{S\in\mathcal{G}^{[d]}:S_\emptyset=\mathrm{id}\}$.

For convenience, we denote the orbit closure of $\boldsymbol{x}\in X^{[d]}$ under $\mathcal{F}^{[d]}$ by $\overline{\mathcal{F}^{[d]}}(\boldsymbol{x})$, instead of $\overline{\mathcal{O}(\boldsymbol{x},\mathcal{F}^{[d]})}$. It is easy to verify that $\mathbf{Q}^{[d]}$ is the closure in $X^{[d]}$ of

$$\{Sx^{[d]}:S\in\mathcal{F}^{[d]},x\in X\}.$$

If x is a transitive point of X, then $\mathbf{Q}^{[d]}$ is the orbit closure of $x^{[d]}$ under the group $\mathcal{G}^{[d]}$.

2.5 Nilmanifolds and nilsystems

Let G be a group. For $g, h \in G$ and $A, B \subset G$, we write $[g, h] = ghg^{-1}h^{-1}$ for the commutator of g and h and $[A, B]$ for the subgroup spanned by $\{[a, b] : a \in A, b \in B\}$. The commutator subgroups G_j, $j \geqslant 1$, are defined inductively by setting $G_1 = G$ and $G_{j+1} = [G_j, G]$. Let $d \geqslant 1$ be an integer. We say that G is *d-step nilpotent* if G_{d+1} is the trivial subgroup.

Let G be a d-step nilpotent Lie group and Γ be a discrete cocompact subgroup of G. The compact manifold $X = G/\Gamma$ is called a *d-step nilmanifold*. The group G acts on X by left translations and we write this action as $(g, x) \mapsto gx$. The Haar measure μ of X is the unique probability measure on X invariant under this action. Fix $\tau \in G$ and let T be the transformation $x \mapsto \tau x$ of X. Then (X, μ, T) is called a *d-step nilsystem*. In the topological setting we omit the measure and just say that (X, T) is a d-step nilsystem. For more details on nilsystems, refer to [22].

We need to use inverse limits of nilsystems, so we recall the definition of a sequential inverse limit of systems. If $(X_i, T_i)_{i\in\mathbb{N}}$ are systems with $\mathrm{diam}(X_i) \leqslant 1$ and $\pi_i : X_{i+1} \to X_i$ are factor maps, the *inverse limit* of these systems is defined to be the compact subset of $\prod_{i\in\mathbb{N}} X_i$ given by

$$\{(x_i)_{i\in\mathbb{N}} : \pi_i(x_{i+1}) = x_i\},$$

and we denote it by

$$\varprojlim (X_i, T_i)_{i\in\mathbb{N}}.$$

It is a compact metric space endowed with the distance $\rho((x_i)_{i\in\mathbb{N}}, (y_i)_{i\in\mathbb{N}}) = \sum_{i\in\mathbb{N}} 1/2^i \rho_i(x_i, y_i)$, where ρ_i is the metric in X_i. We note that the maps T_i induce naturally a transformation T on the inverse limit.

The following structure theorem characterizes inverse limits of nilsystems using dynamical parallelepipeds.

Theorem 2.5 (See [23, Theorem 1.2]). *Assume that (X, T) is a transitive topological dynamical system and let $d \geqslant 2$ be an integer. The following properties are equivalent*:

(1) *If $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{Q}^{[d]}$ have $2^d - 1$ coordinates in common, then $\boldsymbol{x} = \boldsymbol{y}$.*

(2) *If $x, y \in X$ are such that $(x, y, \ldots, y) \in \mathbf{Q}^{[d]}$, then $x = y$.*

(3) *X is an inverse limit of $(d-1)$-step minimal nilsystems.*

A transitive system satisfying one of the equivalent properties above is called a *$(d-1)$-step pro-nilsystem* or *system of order* $(d-1)$.

2.6 Regionally proximal relation of order d

Definition 2.6. Let (X, T) be a system and let $d \in \mathbb{N}$. The points $x, y \in X$ are said to be *regionally proximal of order d* if for any $\delta > 0$, there exist $x', y' \in X$ and a vector $\boldsymbol{n} = (n_1, \ldots, n_d) \in \mathbb{Z}^d$ such that $\rho(x, x') < \delta, \rho(y, y') < \delta$, and

$$\rho(T^{\boldsymbol{n}\cdot\epsilon}x', T^{\boldsymbol{n}\cdot\epsilon}y') < \delta \quad \text{for every nonempty } \ \epsilon \subset [d].$$

In other words, there exists $S \in \mathcal{F}^{[d]}$ such that $\rho(S_\epsilon x', S_\epsilon y') < \delta$ for every $\epsilon \neq \emptyset$. The set of regionally proximal pairs of order d is denoted by $\mathbf{RP}^{[d]}$ (or by $\mathbf{RP}^{[d]}(X, T)$ in case of ambiguity), and is called *the regionally proximal relation of order d*.

It is easy to see that $\mathbf{RP}^{[d]}$ is a closed and invariant relation. Observe that

$$\mathbf{P}(X, T) \subseteq \cdots \subseteq \mathbf{RP}^{[d+1]} \subseteq \mathbf{RP}^{[d]} \subseteq \cdots \subseteq \mathbf{RP}^{[2]} \subseteq \mathbf{RP}^{[1]} = \mathbf{RP}(X, T).$$

The following theorems proved in [23] (for minimal distal systems) and in [33] (for general minimal systems) tell us conditions under which (x, y) belongs to $\mathbf{RP}^{[d]}$ and the relation between $\mathbf{RP}^{[d]}$ and d-step pro-nilsystems, which are defined in Theorem 2.5.

Theorem 2.7. *Let (X, T) be a minimal system and let $d \in \mathbb{N}$. Then*

(1) *$(x, y) \in \mathbf{RP}^{[d]}$ if and only if $(x, y, \ldots, y) \in \mathbf{Q}^{[d+1]}$ if and only if $(x, y, \ldots, y) \in \overline{\mathcal{F}^{[d+1]}}(x^{[d+1]})$.*

(2) *$\mathbf{RP}^{[d]}$ is an equivalence relation.*

(3) *(X, T) is a d-step pro-nilsystem if and only if $\mathbf{RP}^{[d]} = \Delta_X$.*

2.7 ∞-step pro-nilsystems

The regionally proximal relation of order d allows to construct the maximal d-step pro-nilfactor of a system. In [23], it was shown that for a minimal distal system (X,T) the quotient of X under $\mathbf{RP}^{[d]}(X,T)$ is the maximal d-step pro-nilfactor of X. In general one has the following theorem.

Theorem 2.8 (See [33]). *Let $\pi:(X,T)\to(Y,T)$ be a factor map between minimal systems and let $d\in\mathbb{N}$. Then,*

(1) $\pi\times\pi(\mathbf{RP}^{[d]}(X,T))=\mathbf{RP}^{[d]}(Y,T)$.

(2) (Y,T) *is a d-step pro-nilsystem if and only if* $\mathbf{RP}^{[d]}(X,T)\subset R_\pi$.

In particular, $X_d=X/\mathbf{RP}^{[d]}(X,T)$, the quotient of (X,T) under $\mathbf{RP}^{[d]}(X,T)$, is the maximal d-step pro-nilfactor of X.

It follows that for any minimal system (X,T),

$$\mathbf{RP}^{[\infty]}=\bigcap_{d\geqslant 1}\mathbf{RP}^{[d]}$$

is a closed invariant equivalence relation (we write $\mathbf{RP}^{[\infty]}(X,T)$ in case of ambiguity). Now we formulate the definition of ∞-step pro-nilsystems.

Definition 2.9 (See [8]). A minimal system (X,T) is an ∞-*step pro-nilsystem* or *a system of order* ∞, if the equivalence relation $\mathbf{RP}^{[\infty]}$ is trivial, i.e., coincides with the diagonal.

Remark 2.10. Similar to Theorem 2.8, one can show that the quotient of a minimal system (X,T) under $\mathbf{RP}^{[\infty]}$ is the maximal ∞-step pro-nilfactor of (X,T).

3 Regionally proximal relation of order d along arithmetic progressions

Now we introduce the notion of the regionally proximal relation of order d along arithmetic progressions.

3.1 Definition of $\mathbf{AP}^{[d]}(X,T)$

3.1.1 *Definition*

In [23] $\mathbf{RP}^{[d]}$ was introduced based on d-dimensional parallelepipeds. Now we define a relation based on Furstenberg's original average.

Definition 3.1. Let (X,T) be a t.d.s. and $d\in\mathbb{N}$. We say that $(x,y)\in X\times X$ is a *regionally proximal pair of order d along arithmetic progressions* if for each $\delta>0$ there exist $x',y'\in X$ and $n\in\mathbb{Z}$ such that $\rho(x,x')<\delta,\rho(y,y')<\delta$ and

$$\rho(T^{in}(x'),T^{in}(y'))<\delta\quad\text{for each}\quad 1\leqslant i\leqslant d.$$

The set of all such pairs is denoted by $\mathbf{AP}^{[d]}(X)$ and is called the *regionally proximal relation of order d along arithmetic progressions.*

For a relation B on X let $\mathbf{R}(B)$ be the smallest closed invariant equivalence generated by B.

Remark 3.2. (1) When $d=1$, $\mathbf{AP}^{[1]}(X)$ is nothing but the regionally proximal relation $\mathbf{RP}(X)$.

(2) Note that for $\boldsymbol{n}=(n,n,\dots,n)\in\mathbb{Z}^d$, one has

$$\{\boldsymbol{n}\cdot\epsilon:\epsilon=(\epsilon_1,\epsilon_2,\dots,\epsilon_d)\in\{0,1\}^d\setminus\{\emptyset\}\}=\{n,2n,\dots,dn\}.$$

It follows easily that $\mathbf{AP}^{[d]}(X)\subset\mathbf{RP}^{[d]}(X)$ for each $d\in\mathbb{N}$, and hence $\mathbf{R}(\mathbf{AP}^{[d]}(X))\subset\mathbf{RP}^{[d]}(X)$.

Lemma 3.3. *Let $k\in\mathbb{N}$. Then* $\mathbf{AP}^{[d]}(X,T)=\mathbf{AP}^{[d]}(X,T^k)$.

Proof. First, we note that $\mathbf{AP}^{[d]}(X,T)\supset\mathbf{AP}^{[d]}(X,T^k)$. Now let $(x,y)\in\mathbf{AP}^{[d]}(X,T)$. Then for each $\delta>0$ there exist $x',y'\in X$ and $n\in\mathbb{Z}$ such that $\rho(x,x')<\delta$, $\rho(y,y')<\delta$ and $\rho(T^{in}(x'),T^{in}(y'))<\delta_1$ for each $1\leqslant i\leqslant d$, where $\delta_1<\delta$ is such that $\rho(z_1,z_2)<\delta_1$ implies $\rho(T^jz_1,T^jz_2)<\delta$ for each $1\leqslant j\leqslant dk$. Then we know that $(x,y)\in\mathbf{AP}^{[d]}(X,T^k)$. □

3.1.2 *Comparing* $\mathbf{AP}^{[d]}$ *with* $\mathbf{RP}^{[d]}$

In order to show that $\mathbf{RP}^{[d]}$ is an equivalence relation in [33] (see also [23]) one proves that $(x,y)\in\mathbf{RP}^{[d]}$ if and only if for each neighborhood U of y there is $\boldsymbol{n}=(n_1,\ldots,n_{d+1})\in\mathbb{Z}^{d+1}$ such that $T^{\boldsymbol{n}\cdot\epsilon}(x)\in U$ for each $\epsilon\neq\emptyset$. Since $\mathbf{AP}^{[1]}=\mathbf{RP}^{[1]}$, it is natural to ask if for $d=1$, $(x,y)\in\mathbf{AP}^{[1]}$, if and only if for each neighborhood U of y there is $n\in\mathbb{N}$ such that $T^nx,T^{2n}x\in U$. Unfortunately this is not the case as the following example shows.

Consider $T:\mathbb{T}^2\to\mathbb{T}^2,(x,y)\mapsto(x+\alpha,x+y)$, where α is irrational. Then $T^n(x,y)=(x+n\alpha, y+nx+a(n)\alpha)$ with $a(n)=\frac{1}{2}n(n-1)$. It is easy to see that

$$\mathbf{RP}^{[1]}=\{((x,y_1),(x,y_2)):x,y_1,y_2\in\mathbb{T}\}.$$

Let $y\in\mathbb{T}=[0,1)$ and $y\neq0,\frac{1}{3},\frac{2}{3}$. We claim it is not true that for each neighborhood U of $(0,y)$ there is $n\in\mathbb{N}$ such that $T^n(0,0),T^{2n}(0,0)\in U$. Assume that this is the case, i.e., for each $\epsilon>0$ there is $n\in\mathbb{N}$ such that

$$-\epsilon<n\alpha \pmod 1<\epsilon,\quad -\epsilon<2n\alpha \pmod 1<\epsilon,$$

and

$$-\epsilon<y-a(n)\alpha \pmod 1<\epsilon,\quad -\epsilon<y-a(2n)\alpha \pmod 1<\epsilon.$$

A simple calculation shows that $3y=0 \pmod 1$, which leads to a contradiction. Thus, we do not have the property for $\mathbf{AP}^{[1]}$ as for $\mathbf{RP}^{[1]}$.

3.1.3 *A question*

It is easy to check that $\mathbf{AP}^{[d]}(X)$ is a closed $T\times T$ invariant relation. We do not know if it is an equivalence relation, i.e.,

Question 1. *Is it true that for a minimal t.d.s.* $\mathbf{AP}^{[d]}(X)$ *is an equivalence relation? If not, is this true when* (X,T) *is also distal?*

3.2 Systems with $\mathbf{AP}^{[d]}(X)=X\times X$

In this subsection, we show that in some cases we have $\mathbf{AP}^{[d]}(X)=X\times X$. Glasner [17] studied the diagonal action $\sigma_d=T\times T^2\times\cdots\times T^d$ and showed the following theorem.

Theorem 3.4 (See [17])**.** *Let* (X,T) *be a minimal weakly mixing t.d.s. Then for each* $d\in\mathbb{N}$ *there is a dense* G_δ *subset* K_d *of* X *such that for each* $x\in K_d$*, the orbit of* $(x,\ldots,x)$ *under* σ_d *is dense in* X^d*.*

Using this result we have the following proposition.

Proposition 3.5. *Let* (X,T) *be a minimal t.d.s. Then the following statements are equivalent*:

(1) *Each pair is an* Ind_{ap}*-pair, i.e.,* $\mathrm{Ind}_{ap}(X)=X\times X$.

(2) (X,T) *is weakly mixing.*

(3) $\mathbf{AP}^{[d]}(X)=X\times X$ *for some* $d\geqslant2$.

Proof. (1) $\Rightarrow$ (2). Assume that U_0 and U_1 are two non-empty open subsets of X. Then there is $n\in\mathbb{N}$ such that

$$T^{-n}U_0\cap T^{-2n}U_0\cap T^{-3n}U_0\cap T^{-4n}U_1\neq\emptyset,$$

which implies that $N(U_0,U_0)\cap N(U_0,U_1)\neq\emptyset$, and hence (X,T) is weakly mixing.

(2) $\Rightarrow$ (1). Let $d\geqslant1$, $A_d=\{0,1\}^d=\{S_1,\ldots,S_{2^d}\}$ and $s=S_1S_2\cdots S_{2^d}\in\{0,1\}^{d2^d}=\{s_1,\ldots,s_{d2^d}\}$. Assume that U_0 and U_1 are two non-empty open subsets of X. By Theorem 3.4 there are $x\in X$ and $n\in\mathbb{N}$ such that

$$(T^nx,T^{2n}x,\ldots,T^{(d2^d)n}x)\in\prod_{i=1}^{d2^d}U_{s_i}.$$

This implies that for each $(t_1,\dots,t_d)\in A_d$,

$$T^{-n}U_{t_1}\cap T^{-2n}U_{t_2}\cap\cdots\cap T^{-nd}U_{t_d}\neq\emptyset,$$

i.e., each pair is an Ind_{ap}-pair.

(1) ⇒ (3) is obvious. To show (3) ⇒ (1) we observe that $\mathbf{RP}(X)=X\times X$ which implies weak mixing by the well-known results (see, for example, [1]). □

Remark 3.6. (1) Let (X,T) be a weakly mixing t.d.s. If in addition (X,T) is TE[1)], then

$$(X,T)\times(X,T^2)\times\cdots\times(X,T^d)$$

is weakly mixing (and TE) (see [27, Corollary 4.2]). Without the assumption of TE, this is not true in general.

(2) Let (X,T) be a t.d.s. If there is a dense G_δ set X_0 such that for each $x\in X_0$, $(x,x,\dots,x)$ has a dense orbit under σ_d in X^d, then using the method of the previous theorem we get that $\mathbf{AP}^{[d]}(X)=X\times X$ for each $d\geqslant 1$.

To show the following property we need a lemma from [15].

Lemma 3.7 (See [15, Theorem 1.24])**.** *Let $\mathbb{N}=N_1\cup N_2\cup\cdots\cup N_d$ for some $d\in\mathbb{N}$. Then there is $1\leqslant i\leqslant d$ such that N_i is piece-wise syndetic.*

Theorem 3.8. *Let $\pi:(X,T)\to(Y,T)$ be a proximal extension between two t.d.s. Then*

$$\mathbf{AP}^{[d]}(X)\supset R_\pi=\{(x,y)\in X^2:\pi(x)=\pi(y)\}$$

for any $d\in\mathbb{N}$. In particular, if (X,T) is proximal, then $\mathbf{AP}^{[d]}(X)=X\times X$ for any $d\in\mathbb{N}$.

Proof. Assume to the contrary that there are $d\in\mathbb{N}$ and a pair $(x_1,x_2)\in R_\pi$ but $(x_1,x_2)\notin\mathbf{AP}^{[d]}$. This implies that there is $\epsilon_0>0$ such that if $0<\epsilon\leqslant\epsilon_0$ then for each $m\in\mathbb{N}$ there is $1\leqslant i\leqslant d$ satisfying

$$\rho(T^{im}x_1,T^{im}x_2)\geqslant\epsilon.$$

Let

$$E_i=\{m\in\mathbb{N}:\rho(T^{im}x_1,T^{im}x_2)\geqslant\epsilon\},\quad 1\leqslant i\leqslant d.$$

Then $\mathbb{N}=E_1\cup\cdots\cup E_d$, and then by Lemma 3.7 there is $1\leqslant i\leqslant d$ such that E_i is piece-wise syndetic. This implies that $E_1\supset iE_i$ is piecewise syndetic. Thus the orbit closure of (x_1,x_2) under $T\times T$ contains a minimal point which is not on the diagonal, which leads to a contradiction, since π is proximal. □

To finish the section we ask the following question.

Question 2. *Let (X,T) be a minimal system and assume that (x,y) is proximal. Is it true that $(x,y)\in\mathbf{AP}^{[d]}$ for each $d\in\mathbb{N}$?*

3.3 Factors and extensions

The following property follows directly by the definition.

Proposition 3.9. *Let $\pi:(X,T)\to(Y,T)$ be a factor map between two systems. Then*

$$\pi\times\pi(\mathbf{AP}^{[d]}(X))\subset\mathbf{AP}^{[d]}(Y).$$

Generally we do not have $\pi\times\pi(\mathbf{AP}^{[d]}(X))=\mathbf{AP}^{[d]}(Y)$. For example, let (X_1,T_1) and (X_2,T_2) be two non-trivial proximal t.d.s. with $X_1\cap X_2=\emptyset$ and $X=X_1\cup X_2$. Assume that $T:X\to X$ is such that $T(x)=T_i(x)$ if $x\in X_i$. Then (X,T) has two minimal points. It is clear that $\mathbf{AP}^{[d]}(X,T)=X_1\times X_1\cup X_2\times X_2\neq X\times X$. Let (Y,S) be the t.d.s. obtained by collapsing the two minimal points. Then $\mathbf{AP}^{[d]}(Y)=Y\times Y$ since (Y,S) is proximal. Choose $(y_1,y_2)\in Y\times Y$ such that y_i is not minimal and $y_i\in X_i$. It is clear that $(y_1,y_2)\notin\mathbf{AP}^{[d]}(X)$.

1) TE means topologically ergodic, i.e., for all non-empty open sets $U,V\subseteq X$, $N(U,V)$ is syndetic.

Question 3. *Let $\pi : (X, T) \to (Y, T)$ be a factor map between two minimal systems. Is it true that $\pi \times \pi(\mathbf{AP}^{[d]}(X)) = \mathbf{AP}^{[d]}(Y)$?*

Proposition 3.10. *Let (X, T) be the inverse limit of (X_i, T_i) with bonding maps $\pi_i : X_{i+1} \to X_i$. If $x = (x_1, x_2, \ldots)$ and $y = (y_1, y_2, \ldots) \in X$ are such that $(x_i, y_i) \in \mathbf{AP}^{[d]}(X_i)$, then $(x, y) \in \mathbf{AP}^{[d]}(X)$.*

Proof. Let $\epsilon > 0$, and let $U \times V$ be a neighborhood of (x, y). Then there are $i \in \mathbb{N}$ and a neighborhood $U_i \times V_i$ of (x_i, y_i) such that $\pi_i^{-1}(U_i) \subset U$ and $\pi_i^{-1}(V_i) \subset V$. Since $(x_i, y_i) \in \mathbf{AP}^{[d]}(X_i)$ there are $x_i' \in U_i$ and $y_i' \in V_i$ such that there is $n > 0$ large enough with $\rho(T_i^{jn} x_i', T_i^{jn} y_i') < \epsilon'$ $(1 \leqslant j \leqslant d,\ \epsilon' > 0)$ and $\rho(\pi_{i-1,k} T_i^{jn} x_i', \pi_{i-1,k} T_i^{jn} y_i') < \frac{\epsilon}{2}$, where $1 \leqslant j \leqslant d$, $1 \leqslant k \leqslant i - 2$. Note that $\pi_{i-1,k} : X_i \to X_k$ is defined by $\pi_{i-1,k} = \pi_k \cdots \pi_{i-1}$. Now choose i with $\sum_{m=i+1}^{\infty} \frac{\mathrm{diam}(X_i)}{2^m(1+\mathrm{diam}(X_i))} < \frac{\epsilon}{2}$.

Put $x' = (\cdots x_i' \cdots) \in \pi_i^{-1}(U_i)$ and $y' = (\cdots y_i' \cdots) \in \pi_i^{-1}(V_i)$. Then

$$T^{jn}(x') = (\cdots T_i^{jn}(x_i') \cdots) \quad \text{and} \quad T^{jn}(y') = (\cdots T_i^{jn}(y_i') \cdots).$$

Thus we have $\rho(T^{jn} x', T^{jn} y') < \epsilon$. □

4 Systems with $\mathbf{AP}^{[d]} = \Delta$

In this section we discuss the structure of a t.d.s. with $\mathbf{AP}^{[d]} = \Delta$, and we show that each ergodic measure of (X, T) is isomorphic to a system of order d, and in particular (X, T) has zero entropy.

4.1 Metric description

4.1.1 *Some known results*

Let $d \in \mathbb{N}$. A factor $(Z, \mathcal{Z}, \nu, T)$ of X is *characteristic* for averages

$$\frac{1}{N} \sum_{n=0}^{N-1} f_1(T^n x) \cdots f_d(T^{dn} x) \tag{4.1}$$

if the limiting behavior of (4.1) only depends on the conditional expectation of f_i with respect to Z,

$$\left\| \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} (T^n f_1 T^{2n} f_2 \cdots T^{dn} f_d - T^n \mathrm{E}(f_1 \mid \mathcal{Z}) T^{2n} \mathrm{E}(f_2 \mid \mathcal{Z}) \cdots T^{dn} \mathrm{E}(f_d \mid \mathcal{Z})) \right\|_{L^2} = 0$$

for any $f_1, \ldots, f_d \in L^\infty(X, \mathcal{X}, \mu)$. The system Z is a universal characteristic factor if it is a characteristic factor of X, and a factor of any other characteristic factor of X. The universal characteristic factor of (4.1) always exists [21, 35], and is denoted by $(Z_{d-1}, \mathcal{Z}_{d-1}, \mu_{d-1}, T)$.

Theorem 4.1 (See [21]). *Let $(X, \mathcal{X}, \mu, T)$ be an ergodic system and $d \in \mathbb{N}$. Then the system $(Z_{d-1}, \mathcal{Z}_{d-1}, \mu_{d-1}, T)$ is a (measure theoretic) inverse limit of $(d-1)$-step nilsystems. $(Z_{d-1}, \mathcal{Z}_{d-1}, \mu_{d-1}, T)$ is called a system of order $d - 1$.*

We also need the following classic result by Furstenberg [14].

Theorem 4.2 (See [14]). *Let $(X, \mathcal{X}, \mu, T)$ be a measure preserving transformation and let $A \in \mathcal{X}$ be a set with positive measure. Then for every integer $k \geqslant 1$,*

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(A \cap T^{-n} A \cap T^{-2n} A \cap \cdots \cap T^{-kn} A) > 0.$$

Remark 4.3. In fact by [21, Theorem 1.1], one can replace lim inf in Theorem 4.2 by lim, i.e.,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(A \cap T^{-n} A \cap T^{-2n} A \cap \cdots \cap T^{-kn} A) > 0.$$

4.1.2 *Consequences*

Let $(X,\mathcal{B},\mu,T)$ be an ergodic measure preserving transformation and $(Z_k,\mathcal{Z}_k,\mu_k,T)$ be the k-step nil-factor of $(X,\mathcal{B},\mu,T)$. Let $\mu=\int_{Z_k}\mu_z d\mu_k(z)$ be the disintegration of μ over μ_k. Pairs in the support of the measure

$$\lambda_k=\int_{Z_k}\mu_z\times\mu_z d\mu_k(z)$$

are called L_k^μ-pairs, where $L_k^\mu=\mathrm{Supp}(\lambda_k)$.

Now we may obtain the following theorem related to L_k^μ.

Theorem 4.4. *Let (X,T) be a t.d.s. and μ an ergodic Borel measure on X. Let $k\geqslant 1$ be an integer. Then $L_k^\mu\subset\mathbf{AP}^{[k]}(X)$. Moreover,*

$$\bigcap_{k=1}^{\infty}L_k^\mu\subset\mathrm{Ind}_{ap}(X).$$

Proof. Let $(x_0,x_1)\in L_k^\mu$. Then for any neighborhood $U_0\times U_1$ of (x_0,x_1),

$$\lambda_k(U_0\times U_1)=\int_{Z_k}\mathrm{E}(1_{U_0}\mid\mathcal{Z}_k)\mathrm{E}(1_{U_1}\mid\mathcal{Z}_k)d\mu_k>0.$$

By Theorem 4.1, we have

$$\begin{aligned}
&\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(U_0\cap T^{-n}U_1\cap T^{-2n}U_1\cap\cdots\cap T^{-(k+1)n}U_1)\\
&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int_X 1_{U_0}(x)1_{U_1}(T^nx)1_{U_1}(T^{2n}x)\cdots 1_{U_1}(T^{(k+1)n}x)d\mu(x)\\
&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\int_{Z_k}\mathrm{E}(1_{U_0}\mid\mathcal{Z}_k)(z)\mathrm{E}(1_{U_1}\mid\mathcal{Z}_k)(T^nz)\cdots\mathrm{E}(1_{U_1}\mid\mathcal{Z}_k)(T^{(k+1)n}z)d\mu_k(z)\\
&\geqslant\liminf_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}a^{k+2}\int_{Z_k}1_{A_a}(z)1_{A_a}(T^nz)\cdots 1_{A_a}(T^{(k+1)n}z)d\mu_k(z)\\
&=\liminf_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}a^{k+2}\mu(A_a\cap T^{-n}A_a\cap T^{-2n}A_a\cap\cdots\cap T^{-(k+1)n}A_a),
\end{aligned}$$

where $a>0$ and $A_a=\{z\in Z_k:\mathrm{E}(1_{U_0}\mid\mathcal{Z}_k)(z)>a$ and $\mathrm{E}(1_{U_1}\mid\mathcal{Z}_k)(z)>a\}$.

As $\mathrm{E}(1_{U_0}\mid\mathcal{Z}_k)\leqslant 1$ and $\mathrm{E}(1_{U_1}|\mathcal{Z}_k)\leqslant 1$, one has

$$\begin{aligned}
0<b&:=\int_{Z_k}\mathrm{E}(1_{U_0}\mid\mathcal{Z}_k)\mathrm{E}(1_{U_1}\mid\mathcal{Z}_k)d\mu_k\\
&=\int_{A_a}\mathrm{E}(1_{U_0}\mid\mathcal{Z}_k)\mathrm{E}(1_{U_1}\mid\mathcal{Z}_k)d\mu_k+\int_{Z_k\setminus A_a}\mathrm{E}(1_{U_0}\mid\mathcal{Z}_k)\mathrm{E}(1_{U_1}\mid\mathcal{Z}_k)d\mu_k\\
&\leqslant\mu_k(A_a)+a\mu_k(Z_k\setminus A_a).
\end{aligned}$$

Hence there exists $a>0$ such that $\mu_k(A_a)=b-a\mu_k(Z_k\setminus A_a)>0$. So

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(U_0\cap T^{-n}U_1\cap T^{-2n}U_1\cap\cdots\cap T^{-(k+1)n}U_1)>0$$

following Theorem 4.2. In particular, there is some $x\in X$ such that $x\in U_0$ and $T^{jn}x\in U_1$ for $j=1,2,\ldots,k+1$.

Given $\epsilon>0$, let $U_0\times U_1$ be a neighborhood of (x_0,x_1) with diameters of U_0 and U_1 less than ϵ. By the above discussion, letting $x_0'=x$ and $x_1'=T^n(x)$, we get that $T^{jn}x_0',T^{jn}x_1'\in U_1$ for $j=1,2,\ldots,k$. This implies that $(x_0,x_1)\in\mathbf{AP}^{[k]}(X)$.

Now assume that

$$(x_0, x_1) \in \bigcap_{k=1}^{\infty} L_k^{\mu}$$

and $U_0 \times U_1$ is a neighborhood of (x_0, x_1). Let $(i_0, i_1, \ldots, i_m) \in \{0,1\}^{m+1}$. In a similar discussion to the above, we get that there is $n > 0$ such that

$$U_{i_0} \cap T^{-n}U_{i_1} \cap \cdots \cap T^{-mn}U_{i_m} \neq \emptyset.$$

Since m is arbitrary, for a fixed $k \in \mathbb{N}$ by choosing suitable $(i_0, i_1, \ldots, i_m)$ we get that $(x_0, x_1) \in \mathrm{Ind}_{ap}(X)$. To see this, we show for every pair of neighborhoods U_0, U_1 of x_0 and x_1 respectively, and every $d \in \mathbb{N}$ there is some $n \in \mathbb{N}$ such that for each $(t_1, \ldots, t_d) \in \{0,1\}^d$,

$$T^{-n}U_{t_1} \cap T^{-2n}U_{t_2} \cap \cdots \cap T^{-nd}U_{t_d} \neq \emptyset.$$

Let $\{0,1\}^d = \{t^{(k)} = (t_1^{(k)}, \ldots, t_d^{(k)}) : k = 1, 2, \ldots, 2^d\}$. Let $m = d \cdot 2^d$ and

$$(i_1, \ldots, i_m) = (t^{(1)}, t^{(2)}, \ldots, t^{(2^d)}) \in \{0,1\}^{d \cdot 2^d}.$$

There is $n > 0$ such that

$$T^{-n}U_{i_1} \cap \cdots \cap T^{-mn}U_{i_m} \neq \emptyset.$$

Set $V_{t^{(k)}} = T^{-n}U_{t_1^{(k)}} \cap T^{-2n}U_{t_2^{(k)}} \cap \cdots \cap T^{-nd}U_{t_d^{(k)}}$. Then one has

$$T^{-n}U_{i_1} \cap \cdots \cap T^{-mn}U_{i_m} = V_{t^{(1)}} \cap T^{-n}V_{t^{(2)}} \cap T^{-2n}V_{t^{(2)}} \cap \cdots \cap T^{-(2^d-1)n}V_{t^{(2^d)}}.$$

It follows that $V_{t^{(k)}} \neq \emptyset$ for all $k \in \{1, 2, \ldots, 2^d\}$. Since $\{0,1\}^d = \{t^{(k)} : k = 1, 2, \ldots, 2^d\}$, we have that for each $(t_1, \ldots, t_d) \in \{0,1\}^d$,

$$T^{-n}U_{t_1} \cap T^{-2n}U_{t_2} \cap \cdots \cap T^{-nd}U_{t_d} \neq \emptyset.$$

The proof is completed. □

A direct application of the above theorem is as follows.

Theorem 4.5. *Let (X, T) be a t.d.s. with $\mathbf{AP}^{[d]}(X) = \Delta$ for some integer $d \geqslant 1$. Then for each ergodic Borel measure μ, (X, μ, T) is measure theoretically isomorphic to a d-step pro-nilsystem.*

Likewise, if $\mathbf{AP}^{[\infty]}(X) = \Delta$, then for each ergodic Borel measure μ, (X, μ, T) is measure theoretically isomorphic to an ∞-step pro-nilsystem.

4.2 Topological description

Let (X, T) be a t.d.s. A pair $(x, y) \in X \times X$ is said to be *asymptotic* when $\lim_{n\to+\infty} d(T^n x, T^n y) = 0$. The set of asymptotic pairs of (X, T) is denoted by $\mathrm{Asym}(X, T)$.

The notion of the *entropy pair* was introduced by Blanchard [2,3]. Let (X, T) be a t.d.s. and x and x' be two distinct points of X. Call $(x, x') \in X^2$ an *entropy pair* of (X, T) if for every open cover $\{U, V\}$ of X with $x \in \mathrm{int}(U^c), x' \in \mathrm{int}(V^c)$ we have that the entropy $h_{\mathrm{top}}(\{U, V\}) > 0$. Let $E(X, T)$ be the set of entropy pairs.

It is proved[2)] that Ind_{ap} has the lifting property. Since $\mathbf{P} \subset \mathbf{RP}^{[d]}$, we know that $X/\mathbf{RP}^{[d]}$ is distal, and hence has zero entropy. Here, we have the following proposition.

Proposition 4.6. *Let (X, T) be a t.d.s. Then*

(1) $\mathrm{Asym}(X, T) \subset \mathbf{AP}^{[d]}(X)$. *Consequently, $X/\mathbf{R}(\mathbf{AP}^{[d]}(X))$ has zero topological entropy.*

(2) $E(X, T) \subset \mathrm{Ind}_{ap}(X, T) \subset \mathbf{AP}^{[d]}(X, T)$. *This also implies that $X/\mathbf{R}(\mathbf{AP}^{[d]}(X))$ has zero topological entropy.*

(3) *If $\mu \in M^e(X, T)$ is not measure theoretically isomorphic to an ∞-step nilsystem, then* $\mathrm{Ind}_{ap} \neq \Delta_X$, *where an ∞-step nilsystem means that it is an inverse limit of minimal nilsystems. This also implies that $X/\mathbf{R}(\mathbf{AP}^{[d]}(X))$ has zero topological entropy.*

2) Huang W, Li H, Ye X. Localization and dynamical Ramsey property. Preprint

Proof. (1) It is easy to see that $\mathrm{Asym}(X,T) \subset \mathbf{AP}^{[d]}(X)$. It follows that $\overline{\mathrm{Asym}(X,T)} \subset \mathbf{AP}^{[d]}(X)$. By the result of [4] we know that $X/\mathbf{R}(\mathbf{AP}^{[d]}(X))$ has zero topological entropy.

(2) It was shown in [28] that if $(x_1,x_2) \in E(X,T)$ then each neighborhood $U_1 \times U_2$ of (x_1,x_2) has an independence set of positive density. By the famous Szemerédi's theorem, each positive density set contains arbitrarily long arithmetic progressions, which implies that $E(X,T) \subset \mathrm{Ind}_{ap}(X,T) \subset \mathbf{AP}^{[d]}(X,T)$. Thus, one gets that $X/\mathbf{R}(\mathbf{AP}^{[d]}(X))$ has zero topological entropy.

(3) The proof is similar to that of [8, Theorem 6.4]. If $\mathrm{Ind}_{ap}(X) = \Delta_X$, then by Theorem 4.4 we have $\bigcap_{k\in\mathbb{N}} L_k^{\mu} \subset \mathrm{Ind}_{ap}(X) = \Delta_X$. It is easy to verify that $\bigcap_{k\in\mathbb{N}} L_k^{\mu} = L_\infty^{\mu}$ where L_∞^{μ} is the support of

$$\lambda_\infty = \int_{Z_\infty} \mu_z \times \mu_z d\mu_\infty(z)$$

with (Z_∞,μ_∞) the inverse limit of (Z_k,μ_k). Thus for each ergodic measure μ, (X,μ,T) is measure theoretically isomorphic to an ∞-step nilsystem.

If $\mu \in M^e(X,T)$ is not measure theoretically isomorphic to an ∞-step nilsystem, then $\mathrm{Ind}_{ap} \neq \Delta_X$. Since $\mathrm{Ind}_{ap}(X,T)$ has the lifting property and $\mathrm{Ind}_{ap}(X,T) \subset \mathbf{AP}^{[d]}(X,T)$, we conclude that $X/\mathbf{R}(\mathbf{AP}^{[d]}(X))$ has zero topological entropy. □

5 For a d-step pro-nilsystem, $\mathrm{AP}^{[i]} = \mathrm{RP}^{[i]}$, $1 \leqslant i \leqslant d$

In this section we show that for a d-step pro-nilsystem,

$$\mathbf{AP}^{[i]} = \mathbf{RP}^{[i]}, \quad 1 \leqslant i \leqslant d.$$

Hence at least in this case, $\mathbf{AP}^{[i]}$ is an equivalence relation.

Let X be a compact metric space and let $\mathcal{M}(X)$ be the collection of regular Borel probability measures on X provided with the weak star topology. Then $\mathcal{M}(X)$ is a compact metric space in which X is embedded by the mapping $x \mapsto \delta_x$, where δ_x is the Dirac measure at x. If $\phi : X \to Y$ is a continuous map between compact metric spaces, then ϕ induces a continuous map $\phi^* : \mathcal{M}(X) \to \mathcal{M}(Y)$ which extends ϕ, where $(\phi^*\mu)(A) = \mu(\phi^{-1}A)$ for all Borel sets $A \subseteq Y$.

Definition 5.1. An extension $\pi : (X,T) \to (Y,T)$ of t.d.s. is said to have a *relatively invariant measure* (RIM for short) if there exists a continuous homomorphism $\lambda : Y \to \mathcal{M}(X)$ of t.d.s. such that $\pi^* \circ \lambda : Y \to \mathcal{M}(Y)$ is just the (Dirac) embedding.

In other words, π is an RIM extension if and only if for every $y \in Y$ there is a $\lambda_y \in \mathcal{M}(X)$ with $\mathrm{supp}\lambda_y \subseteq \pi^{-1}(y)$ and the map $y \mapsto \lambda_y : Y \to \mathcal{M}(X)$ is a homomorphism of t.d.s; this map λ is called a *section* for π. Note that $\pi : X \to \{\star\}$ has an RIM if and only if X has an invariant measure if and only if $\mathcal{M}(X)$ has a fixed point, where $\{\star\}$ stands for the trivial system.

Definition 5.2. An extension $\phi : (Z,T) \to (Y,T)$ is called a *group extension* with group G if the following conditions are fulfilled:

(1) G is a compact Hausdorff topological group, acting continuously on Z from the right as a group of automorphisms of the system Z; this means that there is a continuous mapping $(x,g) \mapsto xg : Z \times G \to Z$ such that

(a) (right action) $\forall\, x \in Z, \forall\, g_1,g_2 \in G$, $x(g_1g_2) = (xg_1)g_2, xe_G = x$;

(b) (automorphisms) $\forall\, g \in G, \forall\, x \in Z$, $T(xg) = (Tx)g$.

(2) The fibers of ϕ are precisely the G-orbits in Z, i.e., for all $x \in Z$, $\phi^{-1}(\phi(x)) = xG$.

A basic theorem about equicontinuous extension is the following result.

Theorem 5.3 (See [9])**.** *Let $\pi : X \to Y$ be an extension of minimal systems. Then π is equicontinuous if and only if it is a factor of a group extension, i.e., we have the following commutative diagram with ϕ*

a group extension:

$$\begin{array}{ccc} X & \xleftarrow{\ \psi\ } & Z. \\ {\scriptstyle\pi}\big\downarrow & \swarrow{\scriptstyle\phi} & \\ Y & & \end{array}$$

Glasner [20, Proposition 3.8] showed that every distal extension has an RIM; for our purpose we need a little more.

Proposition 5.4. *Let (X,T) be a minimal system and let $\pi:(X,T)\to(Y,T)$ be a distal extension. Then π has an RIM with a section λ such that* $\mathrm{Supp}\lambda_y=\pi^{-1}(y)$ *for all $y\in Y$.*

Proof. One can find the proof of the first part of the statements in [20] or [7, Chapter V, (6.5)]. Since we need to show the second part of the statements, we give the whole proof of the results for completeness.

Let $\pi:(X,T)\to(Y,T)$ be a factor between two minimal systems. Then by Furstenberg structure theorem for distal extensions π is distal if and only if there exist a countable ordinal ζ and a directed family of factors $(X_\theta,T),\theta\leqslant\zeta$ such that

(1) $X_0=Y$ and $X_\zeta=X$;

(2) for $\theta<\zeta$ the extension $\pi_\theta:X_{\theta+1}\to X_\theta$ is equicontinuous; and

(3) for a limit ordinal $\xi\leqslant\zeta$,

$$X_\xi=\varprojlim_{\theta<\xi}X_\theta.$$

For convenience if a section satisfies $\mathrm{Supp}\lambda_y=\pi^{-1}(y)$ for all $y\in Y$ then we say it is a *section with full support*. Hence to prove the result, we need to show (I) each equicontinuous extension has a section with full support; (II) a (transfinite) composition of RIM with full support section has an RIM with full support section.

(I) *Each equicontinuous extension has a section λ and* $\mathrm{Supp}\lambda_y=\pi^{-1}(y)$ *for all $y\in Y$.*

Let $\pi:X\to Y$ be an equicontinuous extension of minimal systems. Now by Theorem 5.3 we have the following diagram with ϕ a group extension:

$$\begin{array}{ccc} X & \xleftarrow{\ \psi\ } & Z. \\ {\scriptstyle\pi}\big\downarrow & \swarrow{\scriptstyle\phi} & \\ Y & & \end{array}$$

Now assume that ϕ satisfies all conditions in Definition 5.2. For $x\in Z$ and $f\in C(Z)$, let

$$\hat\lambda_{\phi(x)}(f)=\int_K f(xg)d\mu(g),$$

where μ is the Haar measure on the group G. Then $Y\to\mathcal{M}(Z),y\mapsto\hat\lambda_y$ is a section for ϕ. Since for all $x\in Z$, $\phi^{-1}(\phi(x))=xG$, by the definition of $\hat\lambda$, we have

$$\mathrm{Supp}\hat\lambda_{\phi(x)}=\mathrm{Supp}\mu=xG=\phi^{-1}(\phi(x))$$

for all $x\in Z$.

Now let $\lambda=\psi_*\circ\hat\lambda:Y\to\mathcal{M}(X)$, where $\psi_*:\mathcal{M}(Z)\to\mathcal{M}(X)$ is the map induced by $\psi:Z\to X$. Then λ is a section for π, and

$$\mathrm{Supp}\lambda_y=\psi(\mathrm{Supp}\hat\lambda_y)=\psi(\phi^{-1}(y))=\pi^{-1}(y)$$

for all $y\in Y$. This ends the proof for equicontinuous extensions.

(II) *A (transfinite) composition of RIM with full support section has an RIM with full support section.*

To prove the statement, it suffices to show two cases. One is the composition of two extensions with the properties, and the other is the inverse limit of extensions.

First, let $\pi_1 : X \to Y$ and $\pi_2 : Y \to Z$ be open factor extensions between minimal systems and let π_1 and π_2 have RIMs with full support sections. Let $\lambda^1 : Y \to \mathcal{M}(X)$ and $\lambda^2 : Z \to \mathcal{M}(Y)$ be two sections. Define $\eta : Z \to \mathcal{M}(X)$ such that for each $z \in Z$,

$$\eta_z(f) = \int_Y \bigg(\int_X f d\lambda^1_y(x) \bigg) d\lambda^2_z(y),$$

for each $f \in C(X)$. To check that η is a section we need to show that η is continuous and $(\pi_2 \circ \pi_1)^*(\eta_z) = \delta_z$. The continuity of η follows from that of λ^i, $i = 1, 2$. Now we check that $(\pi_2 \circ \pi_1)^*(\eta_z) = \delta_z$. In fact,

$$\begin{aligned}(\pi_2 \circ \pi_1)^*(\eta_z)(B) &= \eta_z(\pi_1^{-1} \circ \pi_2^{-1}(B)) \\ &= \int_Y \bigg(\int_X 1_{\pi_2^{-1}(B)} d\delta_y \bigg) d\lambda^2_z(y) = \int_Y 1_B d\delta_z = \delta_z(B),\end{aligned}$$

for each $B \in \mathcal{B}(Z)$, since λ^i, $i = 1, 2$ is a section.

Finally, we show $\mathrm{Supp}(\eta_z) = (\pi_2 \circ \pi_1)^{-1}(z)$ for each $z \in Z$. Fix $z \in Z$ and assume that $x \in (\pi_2 \circ \pi_1)^{-1}(z)$ and U is an open neighborhood of x. Then

$$\eta_z(U) = \int_Y \bigg(\int_X 1_U d\lambda^1_y(x) \bigg) d\lambda^2_z(y) = \int_Y \lambda^1_y(U) d\lambda^2_z(y) > 0,$$

since (1) $\pi_1(U)$ is open in Y, and $\pi_2 \circ \pi_1(U)$ is open in Z, (2) for $y \in \pi_1(U)$, $\lambda^1_y(U) > 0$ and $\lambda^2_z(\pi_1(U)) > 0$.

Next, we discuss the inverse limit. Assume that X is an inverse limit of X_n. Let $\pi_n : X \to X_n$ and $\pi_{n,m} : X_n \to X_m$ if $n \geqslant m$ (we set $\pi_{n,n} = \mathrm{id}$). For any $x \in X_1$ and $f \in C(X)$ define

$$\eta_x(f) = \lim_{n\to\infty} (\eta_n)_x(f_n),$$

if f is a limit of $f_n \circ \pi_n$ with $f_n \in C(X_n)$. Here, $(\eta_n)_x \in \mathcal{M}(X_n)$ is defined by induction using the previous argument.

It is easy to check that η_x is well defined. Moreover, if $f = f_n \pi_n$ for some $n \in \mathbb{N}$ then $\eta_x(f) = (\eta_{n+i})_x(f_n \pi_{n+i,n}) = (\eta_n)_x(f_n)$ for $i \geqslant 0$.

Then we check that $\eta : X_1 \to \mathcal{M}(X)$ is a section. To show the continuity of η, assume that $y_n \to y$. We show $\eta_{y_n} \to \eta_y$, i.e., for each $f \in C(X), \eta_{y_n}(f) \to \eta_y(f)$. This follows from the facts that when f is close to $f_k \pi_k$ in $C(X)$, $\eta_z(f)$ is close to $\eta_z(f_k \pi_k) = (\eta_k)_z(f_k)$ for each $z \in X_1$ uniformly; and $\eta_k : X_1 \to \mathcal{M}(X_k)$ is continuous.

We are left to show that $(\pi_1)^* \eta_{x_1}(B) = \delta_{x_1}(B)$ for each $B \in \mathcal{B}(X_1)$. In fact,

$$(\pi_1)^* \eta_{x_1}(B) = \eta_{x_1}(\pi_1^{-1}(B)) = \lim_{n\to\infty} (\eta_n)_{x_1}(\pi_{n,1}^{-1}(B)) = \delta_{x_1}(B).$$

To show η is full, we note that $\{\pi_n^{-1}(U_n) : U_n \text{ is open in } X_n,\ n \in \mathbb{N}\}$ is a base for the topology of X. Fix $x_1 \in X_1$ and let $x \in \pi_1^{-1}(x_1)$ and U be an open neighborhood of x. Then there is $n \in \mathbb{N}$ such that $U \supset \pi_n^{-1}(U_n)$ and $x_1 \in \pi_{n,1}(U_n)$, where U_n is an open set in X_n. Then $\eta_{x_1}(U) \geqslant (\eta_n)_{x_1}(U_n) > 0$. The proof is completed. □

Now we have the following result.

Proposition 5.5. *Let (X, T) be a strictly ergodic system with a unique invariant measure μ and let $\pi : (X, T) \to (Y, T)$ be a distal extension. Let $\mu = \int_Y \mu_y d\nu(y)$ be the disintegration of μ over ν, where $v = \pi_*(\mu)$. Then there is $Y_0 \subset Y$ with a full measure such that for each $y \in Y_0$,* $\mathrm{Supp}(\mu_y) = \pi^{-1}(y)$.

Proof. Since π is distal, it has an RIM by Proposition 5.4. Let $\lambda : Y \to M(X)$ be a section for π. Then $\widetilde{\mu} = \int \lambda_y \, d\, \nu(y)$ is an invariant measure of (X, T). By unique ergodicity, $\widetilde{\mu} = \mu$. Since the disintegration is unique, there is $Y_0 \subset Y$ with full measure such that for each $y \in Y_0$, $\mu_y = \lambda_y$. Thus the result follows from Proposition 5.4. □

The following lemmas come from [8].

Lemma 5.6. *Let (X,T) be a minimal system of order n. Then the maximal measurable and topological factors of order d coincide, where $d \leqslant n$.*

Lemma 5.7. *Let (X,T) be a minimal ∞-step pro-nilsystem. Then (X,T) is an inverse limit of minimal d_i-step nilsystems.*

Recall that for $d \in \mathbb{N}$, $X_d = X/\mathbf{RP}^{[d]}$.

Lemma 5.8. *Let (X,T) be a minimal system. If $X_n = X_{n+1}$ then $X_k = X_n$ for any $k \geqslant n$.*

Now it is time to give the main result of this section.

Theorem 5.9. *Let (X,T) be a unique ergodic minimal distal system such that for each $d \geqslant 1$, Z_d is isomorphic to X_d. Then for $d \geqslant 1$, $\mathbf{AP}^{[d]} = \mathbf{RP}^{[d]}$.*

Consequently, for a minimal ∞-nilsystem, we have for $d \geqslant 1$, $\mathbf{AP}^{[d]} = \mathbf{RP}^{[d]}$.

Proof. We use induction. It is clear that for $d = 1$, $\mathbf{AP}^{[1]} = \mathbf{RP}^{[1]}$. Assume that $\mathbf{AP}^{[d]} = \mathbf{RP}^{[d]}$ for $1 \leqslant d \leqslant n$. Let μ be the unique ergodic measure on (X,T). Let $\pi : X \to X_{n+1} = X/\mathbf{RP}^{[n+1]}$ be the factor map and $\nu = \pi(\mu)$. By the assumption, π can be viewed as the factor map from X to Z_{n+1}.

Let $\mu = \int_{X_{n+1}} \mu_z d\nu(z)$ be the disintegration of μ over ν and

$$\lambda = \int_{X_{n+1}} \mu_z \times \mu_z d\nu(z).$$

By Theorem 4.4, $\mathrm{Supp}(\lambda) \subset \mathbf{AP}^{[n+1]}$.

We are going to show that $\mathrm{Supp}(\lambda) = R_\pi$. First we note that $\lambda(R_\pi) = 1$, so $\mathrm{Supp}(\lambda) \subset R_\pi$. By Proposition 5.5 there is a measurable set $Y_0 \subset X_{n+1}$ with a full measure such that for any $y \in Y_0$, $\mathrm{Supp}(\mu_y) = \pi^{-1}(y)$. Let $W = \mathrm{Supp}(\lambda)$. Since

$$\lambda(W) = \int_{Y_0} \mu_y \times \mu_y(W) d\nu(y) = 1,$$

we have that for a.e. $y \in Y$, $\mu_y \times \mu_y(W) = 1$. This implies that

$$\mathrm{Supp}(\mu_y) \times \mathrm{Supp}(\mu_y) \subset W, \quad \text{a.e.} \quad y \in Y.$$

Thus by the distality of π, we have $\mathrm{Supp}(\lambda) = R_\pi$. Thus, $R_\pi = \mathrm{Supp}(\lambda) \subset \mathbf{AP}^{[n+1]}$. Since $\mathbf{AP}^{[n+1]} \subset \mathbf{RP}^{[n+1]}$, we conclude that $\mathbf{AP}^{[n+1]} = \mathbf{RP}^{[n+1]}$. This ends the proof of the first statement of the theorem.

When (X,T) is a minimal ∞-step pro-nilsystem, (X,T) is uniquely ergodic (see [8]). The result follows from what we just proved, Lemmas 5.6–5.8 and an inverse limit argument. □

6 An example

In general it is not difficult to find a system whose maximal measurable and topological factors of order d do not coincide, where $d \leqslant n$. In fact Lehrer [30] showed the following result: every ergodic system has a uniquely ergodic and topologically mixing model. Pick any non-periodic ergodic system with discrete spectrum, and by Lehrer's result let (X,T) be its uniquely ergodic and topologically mixing model. Since (X,T) is topologically mixing, its maximal equicontinuous factor Z_1 is trivial.

By Lemma 5.6, for a minimal system of order n, the maximal measurable and topological factors of order d coincide, where $d \leqslant n$. It is natural to ask that for a distal minimal system, if the maximal measurable and topological factors of order d coincide.

In this section we construct a strictly ergodic distal system such that Z_1 is not isomorphic to X_1, i.e., we want to give the following example.

Example 6.1. There is a uniquely ergodic minimal distal system (X,T) with discrete spectrum whose maximal equicontinuous factor is not equal to (X,T).

Proof. Let us state the general idea. Let $T_\alpha:\mathbb{T}\to\mathbb{T}$, $x\mapsto x+\alpha, x\in\mathbb{T}$, and $m_{\mathbb{T}}$ be the unique measure of the irrational rotation T_α on $\mathbb{T}$. Then $m_{\mathbb{T}}$ is the Lebesgue measure on $\mathbb{T}$. We construct $T:\mathbb{T}^2\to\mathbb{T}^2$ having the form of $T(x,y)=(x+\alpha,y+u(x))$ such that $(\mathbb{T}^2,T)$ is minimal distal and uniquely ergodic with the unique measure $\mu=m_{\mathbb{T}^2}=m_{\mathbb{T}}\times m_{\mathbb{T}}$, where $u:\mathbb{T}\to\mathbb{T}$ is continuous. At the same time $(\mathbb{T},T_\alpha)$ is the maximal equicontinuous factor of $(\mathbb{T}^2,T)$.

Step 1. The construction of u.

Let us construct u using some results of [12]. Choose an irrational α and a subsequence $\{n_k\}$ of integers with $n_k\neq 0, n_{-k}=-n_k$ such that

$$h(\theta)=\sum_{k\neq 0}\frac{1}{|k|}(\mathrm{e}^{2\pi\mathrm{i}n_k\alpha}-1)\mathrm{e}^{2\pi\mathrm{i}n_k\theta}$$

and $g(\mathrm{e}^{2\pi\mathrm{i}\theta})=\mathrm{e}^{2\pi\mathrm{i}\lambda h(\theta)}$ (where $\lambda\in\mathbb{R}$ will be determined later) are C^∞-functions of $[0,1)$ and $\mathbb{T}$ respectively. It is clear that

$$h(\theta)=H(\theta+\alpha)-H(\theta),\quad \text{where } H(\theta)=\sum_{k\neq 0}\frac{1}{|k|}\mathrm{e}^{2\pi\mathrm{i}n_k\theta}.$$

Thus, $H(\cdot)\in L^2(0,1)$ is a measurable function. However, $H(\cdot)$ cannot correspond to a continuous function since $\sum_{k\neq 0}\frac{1}{|k|}=\infty$ and here the series is not Cesero summable at $\theta=0$ (see [36]). Therefore, for some λ, $\mathrm{e}^{2\pi\mathrm{i}\lambda H(\theta)}$ cannot be a continuous function either.

Considering $R(\mathrm{e}^{2\pi\mathrm{i}\theta})=\mathrm{e}^{2\pi\mathrm{i}\lambda H(\theta)}$, we get $R(\mathrm{e}^{2\pi\mathrm{i}\alpha}s)/R(s)=g(s)$ with $R:\mathbb{T}\to\mathbb{T}$ measurable but not continuous.

Put $u(x)=\lambda h(x)+\beta$, where α and β are irrational such that $T_{\alpha,\beta}:\mathbb{T}^2\to\mathbb{T}^2$, $(x,y)\mapsto(x+\alpha,y+\beta)$ is minimal on $\mathbb{T}^2$ and thus uniquely ergodic.

Step 2. The system $(\mathbb{T}^2,T)$ with $T(x,y)=(x+\alpha,y+u(x))$ is strictly ergodic, and $(\mathbb{T}^2,T,\mu)$ is isomorphic to $(\mathbb{T}^2,T_{\alpha,\beta},m_{\mathbb{T}^2})$.

It is clear that $m_{\mathbb{T}^2}$ is an invariant measure for T. Define $\phi:\mathbb{T}^2\to\mathbb{T}^2$, $(x,y)\mapsto(x,y-\lambda H(x))$. It is clear that ϕ is measurable and $m_{\mathbb{T}^2}$ is an invariant measure for ϕ. Moreover, we have the following commuting diagram:

$$\begin{array}{ccc}
(\mathbb{T}^2,m_{\mathbb{T}^2}) & \xrightarrow{T} & (\mathbb{T}^2,m_{\mathbb{T}^2}) \\
\phi\downarrow & & \downarrow\phi \\
(\mathbb{T}^2,m_{\mathbb{T}^2}) & \xrightarrow[T_{\alpha,\beta}]{} & (\mathbb{T}^2,m_{\mathbb{T}^2}),
\end{array}$$

since $\phi\circ T(x,y)=(x+\alpha,y+\beta-\lambda H(x))=T_{\alpha,\beta}\circ\phi(x,y)$. By the fact that ϕ is a measurable isomorphism it follows that $(\mathbb{T}^2,m_{\mathbb{T}^2},T)$ is ergodic (as $(\mathbb{T}^2,m_{\mathbb{T}^2},T_{\alpha,\beta})$ is ergodic). This implies that $(\mathbb{T}^2,m_{\mathbb{T}^2},T)$ is uniquely ergodic (see [15, Proposition 3.10]). Since $\mathrm{Supp}(m_{\mathbb{T}^2})=\mathbb{T}^2$, it follows that $(\mathbb{T}^2,T)$ is minimal.

Step 3. $(\mathbb{T},T_\alpha)$ is the maximal equicontinuous factor of $(\mathbb{T}^2,T)$.

To show this fact we need some preparation. Let $\pi:(\mathbb{T}^2,T)\to(\mathbb{T},T_\alpha)$ be the projection to the first coordinate and $\rho=\int_0^1u(x)dx$. We show that u is an unbounded motion, i.e., there exists $x'\in\mathbb{T}$ such that

$$\sup_{n\geqslant 1}|u(x')+u(x'+\alpha)+\cdots+u(x'+(n-1)\alpha)-n\rho|=+\infty.$$

This is equivalent to say the following lemma.

Lemma 6.2. *There exists $x'\in\mathbb{T}$ such that*

$$\sup_{n\geqslant 1}|\lambda h(x')+\lambda h(x'+\alpha)+\cdots+\lambda h(x'+(n-1)\alpha)-n\rho^*|=+\infty, \tag{6.1}$$

where $\rho^*=\int_0^1\lambda h(x)dx=0$.

The proof of Lemma 6.2 will be given at the end of the proof. By Lemma 6.2 there exists $x' \in \mathbb{T}$ such that $\sup_{n\geqslant 1} |\sum_{j=0}^{n-1} u(x'+j\alpha)-n\rho| = +\infty$. Without loss of generality, we assume that $\sup_{n\geqslant 1}\{\sum_{j=0}^{n-1} u(x' + j\alpha) - n\rho\} = +\infty$.

We need another well-known lemma (see, for example, [32, Lemma 4.1]). Note that the degree of u is zero.

Lemma 6.3. *There exist $x_1, x_2 \in \mathbb{T}$ such that*

$$\sup_{n\geqslant 1}\{u(x_1) + u(x_1+\alpha) + \cdots + u(x_1 + (n-1)\alpha) - n\rho\} \leqslant 2$$

and

$$\inf_{n\geqslant 1}\{u(x_2) + u(x_2+\alpha) + \cdots + u(x_2 + (n-1)\alpha) - n\rho\} \geqslant -2.$$

Now we are ready to show that $(\mathbb{T}, T_\alpha)$ is the maximal equicontinuous factor of $(\mathbb{T}^2, T)$. Since $\mathbf{RP}^{[1]}(\mathbb{T}^2, T)$ is $T \times T$-invariant and closed it remains to prove that

$$\mathbf{RP}^{[1]}(\mathbb{T}^2, T) \supset \{(x_1, y_1), (x_1, y_2) : y_1, y_2 \in \mathbb{T}\}.$$

To do this consider

$$\mathbb{T}_{\infty,+} = \left\{x \in \mathbb{T} : \sup_{n\geqslant 1}\{u(x) + u(x+\alpha) + \cdots + u(x + (n-1)\alpha) - n\rho\} = +\infty\right\}.$$

It is G_δ- and T_α-invariant. As $x' \in \mathbb{T}_{\infty,+}$ we know that $x' + i\alpha \in \mathbb{T}_{\infty,+}, i \in \mathbb{N}$.

Fix $y_1, y_2 \in \mathbb{T}$. For $\epsilon > 0$ let

$$U_1 = (x_1 - \epsilon, x_1 + \epsilon) \times (y_1 - \epsilon, y_1 + \epsilon) \quad \text{and} \quad U_2 = (x_1 - \epsilon, x_1 + \epsilon) \times (y_2 - \epsilon, y_2 + \epsilon).$$

Choose $i_* \in \mathbb{N}$ such that $x_1^* = x' + i_*\alpha \pmod{\mathbb{Z}} \in (x_1 - \epsilon, x_1 + \epsilon)$. Since $x_1^* \in \mathbb{T}_{\infty,+}$, there exists $m \in \mathbb{N}$ such that

$$u(x_1^*) + u(x_1^* + \alpha) + \cdots + u(x_1^* + (m-1)\alpha) - m\rho \geqslant 3.$$

Now consider $Q : (x_1 - \epsilon, x_1 + \epsilon) \to \mathbb{R}$, $x \mapsto \sum_{j=0}^{m-1} u(x + j\alpha) - m\rho + y_1$. Then by Lemma 6.3,

$$\begin{aligned} Q(x_1 - \epsilon, x_1 + \epsilon) &\supset \left[y_1 + \sum_{j=0}^{m-1} u(x_1 + j\alpha) - m\rho, y_1 + \sum_{j=0}^{m-1} u(x_1^* + j\alpha) - m\rho\right] \\ &\supset [y_1 + 2, y_1 + 3]. \end{aligned}$$

Thus, there is $x^* \in (x_1 - \epsilon, x_1 + \epsilon)$ such that $Q(x^*) = y_2 + \sum_{j=0}^{m-1} u(x_1 + j\alpha) - m\rho \pmod{\mathbb{Z}}$. Now we have $(x^*, y_1) \in U_1$ and $(x_1, y_2) \in U_2$. Moreover,

$$\begin{aligned} T^m(x^*, y_1) &= \left(x^* + m\alpha, y_1 + \sum_{j=0}^{m-1} u(x^* + j\alpha)\right) \\ &= (x^* + m\alpha, y_1 + Q(x^*) + m\rho) = \left(x^* + m\alpha, y_2 + \sum_{j=0}^{m-1} u(x_1 + j\alpha)\right), \\ T^m(x_1, y_2) &= \left(x_1 + m\alpha, y_2 + \sum_{j=0}^{m-1} u(x_1 + j\alpha)\right). \end{aligned}$$

This implies that

$$\|T^m(x^*, y_1) - T^m(x_1, y_2)\| \leqslant \|x^* - x_1\| < \epsilon,$$

i.e., we have proved that $((x_1, y_1), (x_1, y_2)) \in \mathbf{RP}^{[1]}(\mathbb{T}^2, T)$. It follows that

$$\mathbf{RP}^{[1]}(\mathbb{T}^2, T) = \{((x, y_1), (x, y_2)) : x, y_1, y_2 \in \mathbb{T}\} = R_\pi,$$

since π is distal. $\square$

To show Lemma 6.2 we need the following lemma.

Lemma 6.4 (See [12, Theorem 3.1]). *Let (Ω_0, T_0) be a strictly ergodic system and μ_0 its unique ergodic measure. Let $\Omega = \Omega_0 \times \mathbb{T}$ and let $T : \Omega \to \Omega$ be defined by $T(w_0, s) = (T_0(w_0), g(w_0)s)$, where $g : \Omega_0 \to \mathbb{T}$ is a continuous function. Then if the equation $g^k(w_0) = R(T_0(w_0))/R(w_0)$ has a solution $R : \Omega_0 \to \mathbb{T}$ which is measurable but not equal almost everywhere to a continuous function, then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f \circ T^n(w)$$

cannot exist for all continuous functions f and all $w \in \Omega$.

Proof of Lemma 6.2. Let

$$\mathbb{T}_\infty = \Big\{x \in \mathbb{T} : \sup_{n\geqslant 1} |\lambda h(x) + \lambda h(x+\alpha) + \cdots + \lambda h(x+(n-1)\alpha) - n\rho^*| = +\infty\Big\}. \tag{6.2}$$

It is clear that $\mathbb{T}_\infty$ is a G_δ- and T_α-invariant subset, and thus if it is not empty then it is a dense G_δ subset of $\mathbb{T}$.

Assume to the contrary that $\mathbb{T}_\infty = \emptyset$. We claim that there exists $M \in \mathbb{N}$ such that

$$|\lambda h(x) + \lambda h(x+\alpha) + \cdots + \lambda h(x+(n-1)\alpha) - n\rho^*| \leqslant M$$

for any $n \geqslant 1$ and $x \in \mathbb{T}$. If the claim does not hold, then there exist $x_k \in \mathbb{T}$ and $n_k \to +\infty$ such that

$$|\lambda h(x_k) + \lambda h(x_k+\alpha) + \cdots + \lambda h(x_k+(n_k-1)\alpha) - n_k\rho^*| > k. \tag{6.3}$$

Consider

$$\mathbb{T}_l = \{x \in \mathbb{T} : \exists\, n \in \mathbb{N} \text{ s.t. } |\lambda h(x) + \lambda h(x+\alpha) + \cdots + \lambda h(x+(n-1)\alpha) - n\rho^*| > l\},$$

$l \in \mathbb{N}$. It is clear that $\mathbb{T}_l$ is open and $\mathbb{T}_\infty = \bigcap_{l\in\mathbb{N}} \mathbb{T}_l$. Now we show that $\mathbb{T}_l$ is a dense open subset for any $l \in \mathbb{N}$.

Fix $l \in \mathbb{N}$. For any non-empty open subset V of $\mathbb{T}$, there exists $r = r(V) \in \mathbb{N}$ such that $\bigcup_{i=0}^r T_\alpha^{-i} V = \mathbb{T}$. Choose $k > l + r|\rho^*| + r\max_{x\in\mathbb{T}} |\lambda h(x)|$. By (6.3) we have

$$|\lambda h(x_k+i\alpha) + \lambda h(x_k+i\alpha+\alpha) + \cdots + \lambda h(x_k+i\alpha+(n_k-i-1)\alpha) - (n_k-i)\rho^*| > l$$

for $i = 0, 1, \ldots, r$, i.e., $x_k + i\alpha \in \mathbb{T}_l$. Since $\bigcup_{i=0}^r T_\alpha^{-i} V = \mathbb{T}$, there exists $0 \leqslant i \leqslant r$ with $x_k + i\alpha \in V \cap \mathbb{T}$. This implies that $\mathbb{T}_l$ is dense, and hence $\mathbb{T}_\infty$ is dense, which leads to a contradiction. This proves the claim.

Consider now $S : \mathbb{T}\times\mathbb{R} \to \mathbb{T}\times\mathbb{R}$, $(x, y) \mapsto (x+\alpha, y+\lambda h(x) - \rho^*)$. Since

$$S^n(x, y) = (x + n\alpha, y + \lambda h(x) + \cdots + \lambda h(x+(n-1)\alpha) - n\rho^*)$$

for any $n \geqslant 0$, we have $\{S^n(0,0) : n \geqslant 0\} \in \mathbb{T}\times[-M, M]$. Thus, $E = \overline{\{S^n(0,0) : n \geqslant 0\}} \subset \mathbb{T}\times[-M, M]$ is an S-invariant compact subset. This deduces that there is a minimal subset $F \subset E$.

As S is distal, (F, S) is a minimal distal system and $p : \mathbb{T}\times\mathbb{R} \to \mathbb{T}$, $(x, y) \mapsto x$ is a factor map from (F, S) to $(\mathbb{T}, T_\alpha)$.

Let $I(x) = \{y \in \mathbb{R} : (x, y) \in F\}$ for any $x \in \mathbb{T}$. Fix $x \in \mathbb{T}$. We claim that $|I(x)| = 1$. In fact let $y_1^* = \max I(x)$ and $y_2^* = \min I(x)$. Then $y_2^* \leqslant y_1^*$. As (F, S) is minimal, there are $\{n_k\}$ such that $S^{n_k}(x, y_2^*) \to (x, y_1^*)$. This implies

$$y_2^* + \lambda h(x) + \cdots + \lambda h(x+(n_k-1)\alpha) - n_k\rho^* \to y_1^*$$

and we assume that

$$y_1^* + \lambda h(x) + \cdots + \lambda h(x+(n_k-1)\alpha) - n_k\rho^* \to y_3^* \in I(x).$$

Thus, $y_1^* \leqslant y_3^*$ and hence $y_1^* = y_3^*$. This implies $y_1^* = y_2^*$, i.e., $|I(x)| = 1$. This ends the proof of the claim.

By what we just proved we know that there exists $\tilde{g} : \mathbb{T} \to \mathbb{R}$ continuous such that

$$\{(x, g(x)) : x \in \mathbb{T}\} = F.$$

Note that the continuity of g follows from the fact that the projection $p : E \to \mathbb{T}$ is one to one.

Since $SF = F$ we get that $(x + \alpha, \tilde{g}(x + \alpha)) = (x + \alpha, \tilde{g}(x) + \lambda h(x) - \rho^*)$ for any $x \in \mathbb{T}$. As $\rho^* = 0$ we know that $\lambda h(x) = \tilde{g}(x + \alpha) - \tilde{g}(x)$, $\forall\, x \in \mathbb{T}$.

Now consider $U : \mathbb{T}^2 \to \mathbb{T}^2$, $(w_1, w_2) \mapsto (w_1 \mathrm{e}^{2\pi \mathrm{i}\alpha}, g(w_1)w_2)$, where $g(\mathrm{e}^{2\pi\mathrm{i}\theta}) = \mathrm{e}^{2\pi\mathrm{i}\lambda h(\theta)}$. Since $R(\mathrm{e}^{2\pi\mathrm{i}\alpha}s)/R(s) = g(s)$ with $R : \mathbb{T} \to \mathbb{T}$ is measurable but not continuous, by Lemma 6.4 there exist $f \in C(\mathbb{T}^2)$ and $(w_1, w_2) \in \mathbb{T}^2$ such that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f \circ U^n(w_1, w_2) \quad \text{does not exist.}$$

On the other hand since $\lambda h(x) = \tilde{g}(x + \alpha) - \tilde{g}(x), \forall\, x \in \mathbb{T}$, by writing $w_1 = \mathrm{e}^{2\pi\mathrm{i}x_1}$ and $w_2 = \mathrm{e}^{2\pi\mathrm{i}y_1}$ we have

$$\begin{aligned} \frac{1}{N} \sum_{n=0}^{N-1} f \circ U^n(w_1, w_2) &= \frac{1}{N} \sum_{n=0}^{N-1} f\big(\mathrm{e}^{2\pi\mathrm{i}(x_1+n\alpha)}, \mathrm{e}^{2\pi\mathrm{i}(y_1+\tilde{g}(x_1+n\alpha)-\tilde{g}(x_1))}\big) \\ &= \frac{1}{N} \sum_{n=0}^{N-1} \tilde{H}(n\alpha) = \frac{1}{N} \sum_{n=0}^{N-1} \tilde{H}(T_\alpha^n(0)) \to \int_0^1 \tilde{H}(t)dt, \end{aligned}$$

by the unique ergodicity of $(\mathbb{T}, T_\alpha)$, where $\tilde{H}(t) = f\big(\mathrm{e}^{2\pi\mathrm{i}(x_1+t)}, \mathrm{e}^{2\pi\mathrm{i}(y_1+\tilde{g}(x_1+t)-\tilde{g}(x_1))}\big)$ is a periodic continuous function of period 1 for t. It is a contradiction. Thus, we have proved that $\mathbb{T}_\infty \neq \emptyset$, and this ends the proof. □

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 11431012, 11971455, 11571335 and 11371339).

References

1 Auslander J. Minimal Flows and Their Extensions. North-Holland Mathematics Studies, vol. 153. Amsterdam: North-Holland, 1988

2 Blanchard F. Fully positive topological entropy and topological mixing: Symbolic dynamics and its applications. Contemp Math, 1992, 135: 95–105

3 Blanchard F. A disjointness theorem involving topological entropy. Bull Soc Math France, 1993, 121: 465–478

4 Blanchard F, Host B, Ruette S. Asymptotic pairs in positive-entropy systems. Ergodic Theory Dynam Systems, 2002, 22: 671–686

5 Bronstein I. Extensions of Minimal Transformation Groups. Netherlands: Springer, 1979

6 Cai F, Shao S. Topological characteristic factors along cubes of minimal systems. Discrete Contin Dyn Syst, 2019, 39: 5301–5317

7 de Vries J. Elements of Topological Dynamics. Dordrecht: Kluwer Academic Publishers, 1993

8 Dong P, Donoso S, Maass S, et al. Infinite-step nilsystems, independence and complexity. Ergodic Theory Dynam Systems, 2013, 33: 118–143

9 Ellis R, Glasner S, Shapiro L. Proximal-isometric (PJ) flows. Adv Math, 1975, 17: 213–260

10 Ellis R, Gottschalk W. Homomorphisms of transformation groups. Trans Amer Math Soc, 1960, 94: 258–271

11 Ellis R, Keynes H. A characterization of the equicontinuous structure relation. Trans Amer Math Soc, 1971, 161: 171–183

12 Furstenberg H. Strict ergodicity and transformation of the torus. Amer J Math, 1961, 83: 573–601

13 Furstenberg H. The structure of distal flows. Amer J Math, 1963, 85: 477–515

14 Furstenberg H. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. J Anal Math, 1977, 31: 204–256

15 Furstenberg H. Recurrence in Ergodic Theory and Combinatorial Number Theory. Princeton: Princeton University Press, 1981

16 Furstenberg H, Weiss B. A mean ergodic theorem for $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(Tn^2 x)$. In: Convergence in Ergodic Theory and Probability. Ohio State University Mathematical Research Institute Publications, vol. 5. Berlin: de Gruyter, 1996, 193–227
17 Glasner E. Topological ergodic decompositions and applications to products of powers of a minimal transformation. J Anal Math, 1994, 64: 241–262
18 Glasner E, Gutman Y, Ye X. Higher order regionally proximal equivalence relations for general minimal group actions. Adv Math, 2018, 333: 1004–1041
19 Glasner E, Weiss B. Quasi-factors of zero-entropy systems. J Amer Math Soc, 1995, 8: 665–686
20 Glasner S. Relatively invariant measures. Pacific J Math, 1975, 58: 393–410
21 Host B, Kra B. Nonconventional ergodic averages and nilmanifolds. Ann of Math (2), 2005, 161: 397–488
22 Host B, Kra B. Nilpotent Structures in Ergodic Theory. Mathematical Surveys and Monographs, vol. 236. Providence: Amer Math Soc, 2018
23 Host B, Kra B, Maass A. Nilsequences and a structure theorem for topological dynamical systems. Adv Math, 2010, 224: 103–129
24 Host B, Kra B, Maass A. Variations on topological recurrence. Monatsh Math, 2016, 179: 57–89
25 Huang W, Li H, Ye X. Family-independence for topological and measurable dynamics. Trans Amer Math Soc, 2012, 364: 5209–5242
26 Huang W, Shao S, Ye X. Nil Bohr-sets and almost automorphy of higher order. Mem Amer Math Soc, 2016, 241: 1143
27 Huang W, Ye X. An explicit scattering, non-weakly mixing example and weak disjointness. Nonlinearity, 2002, 15: 849–862
28 Huang W, Ye X. A local variational relation and applications. Israel J Math, 2006, 151: 237–280
29 Kerr D, Li H. Independence in topological and C^*-dynamics. Math Ann, 2007, 338: 869–926
30 Lehrer E. Topological mixing and uniquely ergodic systems. Israel J Math, 1987, 57: 239–255
31 McMahon D. Relativized weak disjointness and relatively invariant measures. Trans Amer Math Soc, 1978, 236: 225–237
32 Qiao Y X. Topological complexity, minimality and systems of order two on torus. Sci China Math, 2016, 59: 503–514
33 Shao S, Ye X. Regionally proximal relation of order d is an equivalence one for minimal systems and a combinatorial consequence. Adv Math, 2012, 231: 1786–1817
34 Veech W. The equicontinuous structure relation for minimal Abelian transformation groups. Amer J Math, 1968, 90: 723–732
35 Ziegler T. Universal characteristic factors and Furstenberg averages. J Amer Math Soc, 2007, 20: 53–97
36 Zygmund A. Trigonormetric Series, 2nd ed. Cambridge: Cambridge University Press, 1959