



Nilsequences and a structure theorem for topological dynamical systems [☆]

Bernard Host ^a, Bryna Kra ^{b,*}, Alejandro Maass ^{c,d}

^a *Laboratoire d'analyse et de mathématiques appliquées, Université Paris-Est Marne la Vallée & CNRS UMR 8050, 5 Bd. Descartes, Champs sur Marne, 77454 Marne la Vallée Cedex 2, France*

^b *Department of Mathematics, Northwestern University, 2033 Sheridan Road Evanston, IL 60208-2730, USA*

^c *Departamento de Ingeniería Matemática, Universidad de Chile, Casilla 170/3 correo 3, Santiago, Chile*

^d *Centro de Modelamiento Matemático UMI 2071 UCHILE-CNRS, Casilla 170/3 correo 3, Santiago, Chile*

Received 15 May 2009; accepted 17 November 2009

Available online 27 November 2009

Communicated by Gil Kalai

Abstract

We characterize inverse limits of nilsystems in topological dynamics, via a structure theorem for topological dynamical systems that is an analog of the structure theorem for measure preserving systems. We provide two applications of the structure. The first is to nilsequences, which have played an important role in recent developments in ergodic theory and additive combinatorics; we give a characterization that detects if a given sequence is a nilsequence by only testing properties locally, meaning on finite intervals. The second application is the construction of the maximal nilfactor of any order in a distal minimal topological dynamical system. We show that this factor can be defined via a certain generalization of the regionally proximal relation that is used to produce the maximal equicontinuous factor and corresponds to the case of order 1.

© 2009 Elsevier Inc. All rights reserved.

Keywords: Nilsystems; Distal systems; Nilsequences; Regionally proximal relation

[☆] The first author was partially supported by the Institut Universitaire de France, the second author by NSF grant 0555250, and the third author by the Millennium Nucleus Information and Randomness P04-069F, CMM-Fondap-Basal fund. This work was begun during the visit of the authors to MSRI and we thank the institute for its hospitality.

* Corresponding author.

E-mail addresses: bernard.host@univ-mlv.fr (B. Host), kra@math.northwestern.edu (B. Kra), amaass@dim.uchile.cl (A. Maass).

1. Introduction

1.1. Nilsequences

The connection between ergodic theory and additive combinatorics started in the 1970's, with Furstenberg's beautiful proof of Szemerédi's Theorem via ergodic theory. Furstenberg's proof paved the way for new combinatorial results via ergodic methods, as well as leading to numerous developments within ergodic theory. More recently, the interaction between the fields has taken a new dimension, with ergodic objects being imported into the finite combinatorial setting. Some objects at the center of this interchange are nilsequences and the nilsystems on which they are defined. They enter, for example, in ergodic theory into convergence of multiple ergodic averages [10] and into the theory of multicorrelations [4]. In number theory, they arise in finding patterns in the primes (see [8] and the companion articles [7] and [9]). In combinatorics, they are used to find intricate patterns in subsets of integers with positive upper density [5].

Nilsequences are defined by evaluating a function along the orbit of a point in the homogeneous space of a nilpotent Lie group. In a variety of situations, nilsequences have been used to test for a lack of uniformity of a function. Yet, the local properties of nilsequences are not well understood. It is difficult to detect if a given sequence is a nilsequence, particularly if one only knows *local* information about the sequence, meaning properties that can only be tested on finite intervals.

We recall the definition of a nilsequence. A *basic d -step nilsequence* is a sequence of the form $(f(T^n x): n \in \mathbb{Z})$, where (X, T) is a d -step nilsystem, $f: X \rightarrow \mathbb{C}$ is a continuous function, and $x \in X$. A *d -step nilsequence* is a uniform limit of basic d -step nilsequences. (See Section 2.3 for the definition of a nilsystem.) We give a characterization of nilsequences of all orders that can be tested locally, generalizing the work in [14] that gives such an analysis for 2-step nilsequences.

We look at finite portions, the “windows”, of a sequence and we are interested in finding a copy of the same finite window up to some given precision. To make this clear, we introduce some notation. For a sequence $\mathbf{a} = (a_n: n \in \mathbb{Z})$, integers k, j, L , and a real $\delta > 0$, if each entry in the window $[k - L, k + L]$ is equal to the corresponding entry in the window $[j - L, j + L]$ up to an error of δ , then we write

$$\mathbf{a}_{[k-L, k+L]} =_{\delta} \mathbf{a}_{[j-L, j+L]}. \quad (1)$$

The characterization of almost periodic sequences (which are exactly 1-step nilsequences) by compactness can be formulated as follows:

Proposition. *The bounded sequence $\mathbf{a} = (a_n: n \in \mathbb{Z})$ of complex numbers is almost periodic if and only if for all $\varepsilon > 0$, there exist an integer $L \geq 1$ and a real $\delta > 0$ such that for any $k, n_1, n_2 \in \mathbb{Z}$ whenever $\mathbf{a}_{[k-L, k+L]} =_{\delta} \mathbf{a}_{[k+n_1-L, k+n_1+L]}$ and $\mathbf{a}_{[k-L, k+L]} =_{\delta} \mathbf{a}_{[k+n_2-L, k+n_2+L]}$ then $|a_k - a_{k+n_1+n_2}| < \varepsilon$.*

We give a similar characterization for a $(d - 1)$ -step nilsequence \mathbf{a} : if in every interval of a given length the translates of the sequence \mathbf{a} along finite sums (i.e. cubes) of any sequence $\mathbf{n} = (n_1, \dots, n_d)$ are δ -close to the original sequence *except* possibly at the sum $n_1 + \dots + n_d$, then we also have control over the distance between \mathbf{a} and the translate by $n_1 + \dots + n_d$.

The general case is:

Theorem 1.1. Let $\mathbf{a} = (a_n: n \in \mathbb{Z})$ be a bounded sequence of complex numbers and let $d \geq 2$ be an integer. The sequence \mathbf{a} is a $(d - 1)$ -step nilsequence if and only if for every $\varepsilon > 0$ there exist an integer $L \geq 1$ and real $\delta > 0$ such that for any $(n_1, \dots, n_d) \in \mathbb{Z}^d$ and $k \in \mathbb{Z}$, whenever

$$\mathbf{a}_{[k+\epsilon_1 n_1+\dots+\epsilon_d n_d-L, k+\epsilon_1 n_1+\dots+\epsilon_d n_d+L]} =_{\delta} \mathbf{a}_{[k-L, k+L]}$$

for all choices of $\epsilon_1, \dots, \epsilon_d \in \{0, 1\}$ other than $\epsilon_1 = \dots = \epsilon_d = 1$, then we have $|a_{k+n_1+\dots+n_d} - a_k| < \varepsilon$.

In fact, we can replace the approximation in (1) in both the hypothesis and conclusion by any other approximation that defines pointwise convergence and have the analogous result.

1.2. A structure theorem for topological dynamical systems

We prove a structure theorem for topological dynamical systems that gives a characterization of inverse limits of nilsystems. Theorem 1.1 follows from this structure theorem, exactly as it does in the case for $d = 2$ in [14], where the proof of this implication can be found. The structure theorem for topological dynamical systems can be viewed as an analog of the purely ergodic structure theorem of [10]. We introduce the following structure:

Definition 1.1. Let (X, T) be a topological dynamical system and let $d \geq 1$ be an integer. We define $\mathbf{Q}^{[d]}(X)$ to be the closure in X^{2^d} of elements of the form

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

where $\mathbf{n} = (n_1, \dots, n_d) \in \mathbb{Z}^d$, $x \in X$, and we denote a point of X^{2^d} by $(x_\epsilon: \epsilon \in \{0, 1\}^d)$. 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^4 of the set

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

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

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

In each of these, the indices m, n and m, n, p can be taken in \mathbb{N} rather than \mathbb{Z} , giving rise to the same object. This is obvious if T is invertible, but can also be proved without the assumption of invertibility. Thus, throughout the article, we assume that all maps are invertible.

We use these parallelepipeds structures to characterize nilsystems:

Theorem 1.2. Assume that (X, T) is a transitive topological dynamical system and let $d \geq 2$ be an integer. The following properties are equivalent:

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

- (2) If $x, y \in X$ are such that $(x, y, \dots, y) \in \mathbf{Q}^{[d]}(X)$, then $x = y$.
 (3) X is an inverse limit of $(d - 1)$ -step minimal nilsystems.

(For definitions of all the objects, see Section 3.) We note that the use of both d and $d - 1$ is necessary throughout the article, and this leads us to use whichever is notationally more convenient at various times in the proofs.

The first property clearly implies the second, since $(y, y, \dots, y) \in \mathbf{Q}^{[d]}(X)$ for all $y \in X$. The second property implies that the system is distal (see Section 3). The second property plus the assumption of distality implies the first property (see Section 4), which together give that the first two properties are equivalent.

Systems satisfying these properties play a key role in the article and so we define:

Definition 1.2. A transitive system satisfying either of the first two equivalent properties of Theorem 1.2 is called a *system of order $d - 1$* .

The implication (3) \Rightarrow (1) in Theorem 1.2 follows from results in [13] and is reviewed here in Proposition 4.6. The implication (1) \Rightarrow (3) is proved in Section 6, using completely different methods from that used in [14] for $d = 3$, and proceeds by introducing an invariant measure on X .

1.3. The regionally proximal relation and generalizations

We give a second application of Theorem 1.2 in topological dynamics. The study of maximal equicontinuous factors is classical (see, for example [1]). The maximal equicontinuous factor is the topological analog of the Kronecker factor in ergodic theory and recovers the continuous eigenvalues of a system. There are several ways to construct this factor, but the standard method is as a quotient of the *regionally proximal relation*. The first step in generalizing this relation was carried out in [14], where the concept of a double regionally proximal relation is introduced and is used in the distal case to define the maximal 2-step nilfactor. In this article we generalize this relation for higher levels and for $d \geq 1$ we define the *regionally proximal relation of order d* , referring to it as $\mathbf{RP}^{[d]}$. While these generalizations were motivated by the study of abstract parallelepipeds in additive combinatorics [11], they require new techniques. Although we defer the definition of the regionally proximal relation of order d until Section 3, we summarize its uses.

Proposition 1.1. Assume that (X, T) is a transitive topological dynamical system and that $d \geq 1$ is an integer. If the regionally proximal relation of order d on X is trivial, then the system is distal.

In a distal system, we show that $\mathbf{RP}^{[d]}$ is an equivalence relation and that it defines the *maximal d -step topological nilfactor* of the system.

Theorem 1.3. Assume that (X, T) is a distal minimal system and that $d \geq 1$ is an integer. Then the regionally proximal relation of order d on X is a closed invariant equivalence relation and the quotient of X under this relation is its maximal d -step nilfactor.

The maximal d -step (topological) nilfactor is the topological analog of the ergodic theoretic factor \mathcal{Z}_d constructed in [10]. These ergodic factors are characterized by inverse limits of d -step nilsystems. In this direction, we prove in the distal case that $\mathbf{RP}^{[d]}$ is trivial if and only if the system itself is an inverse limit of d -step nilsystems.

To prove Theorem 1.3 we show in Proposition 4.5 that the quotient of X under $\mathbf{RP}^{[d]}$ is its maximal factor of order d . From Theorem 1.2, we deduce that the notions of a system of order d and an inverse limit of d -step nilsystems are equivalent, giving us the conclusion.

We conjecture that the hypothesis of distality in Theorem 1.3 is superfluous, but were unable to prove this.

1.4. Guide to the paper

The article is divided into two somewhat distinct parts. In the first part (Sections 3 and 4), we develop the topological theory of parallelepipeds and the associated theory of generalized regionally proximal relations. With the topological methods developed in these sections, we are able to prove all but the implication “(1) \Rightarrow (3)” of Theorem 1.2. In Section 3, we state the properties of parallelepiped structures and the relation with generalized regionally proximal pairs and show how the conditions of Theorem 1.2 imply that the system is distal. In Section 4, we prove that in the distal case, the main structural properties of parallelepipeds (the “property of closing parallelepipeds”) allows us to show that first two conditions in Theorem 1.2 are equivalent and to show that regionally proximal relation of order d gives rise to the maximal factor of order d . The proof of the remaining implication is carried out in Section 6 and relies heavily on ergodic theoretic notions of Section 5. However, the interaction of the topological and measure theoretic structures plays a key role in the analysis, and it is only via measure theoretic methods that we are finally able to obtain the general topological results.

2. Background

2.1. Topological dynamical systems

A *transformation* of a compact metric space X is a homeomorphism of X to itself. A *topological dynamical system*, referred to more succinctly as just a *system*, is a pair (X, T) , where X is a compact metric space and $T : X \rightarrow X$ is a transformation. We use $d_X(\cdot, \cdot)$ to denote the metric in X and when there is no ambiguity, we write $d(\cdot, \cdot)$. We also make use of a more general definition of a topological system. That is, instead of just a single transformation T , we consider commuting homeomorphisms T_1, \dots, T_k of X or a countable abelian group of transformations. We summarize some basic definitions and properties of systems in the classical setting of one transformation. Extensions to the general case are straightforward.

A *factor* of a system (X, T) is another system (Y, S) such that there exists a continuous and onto map $p : X \rightarrow Y$ satisfying $S \circ p = p \circ T$. The map p is called a *factor map*. If p is bijective, the two systems are (*topologically*) *conjugate*. In a slight abuse of notation, when there is no ambiguity, we denote all transformations (including ones in possibly distinct systems) by T .

A system (X, T) is *transitive* if there exists some point $x \in X$ whose orbit $\{T^n x : n \in \mathbb{Z}\}$ is dense in X and we call such a point a *transitive point*. The system is *minimal* if the orbit of any point is dense in X . This property is equivalent to saying that X and the empty set are the only closed invariant sets in X .

2.2. Distal systems

The system (X, T) is *distal* if for any pair of distinct points $x, y \in X$,

$$\inf_{n \in \mathbb{Z}} d(T^n x, T^n y) > 0. \quad (2)$$

In an arbitrary system, pairs satisfying property (2) are called *distal pairs*. The points x and y are *proximal* if $\liminf_{n \rightarrow \infty} d(T^n x, T^n y) = 0$.

The following proposition summarizes some basic properties of distal systems:

Proposition 2.1. (See Auslander [1, Chapters 5 and 7].)

- (1) *The Cartesian product of a finite family of distal systems is a distal system.*
- (2) *If (X, T) is a distal system and Y is a closed and invariant subset of X , then (Y, T) is a distal system.*
- (3) *A transitive distal system is minimal.*
- (4) *A factor of a distal system is distal.*
- (5) *Let $p: X \rightarrow Y$ be a factor map between the distal systems (X, T) and (Y, T) . If (Y, T) is minimal, then p is an open map.*

Up to the obvious changes in notation, this proposition holds for systems with a countable abelian group of transformations acting on the space X .

For later use, we note the following lemma on distal systems:

Lemma 2.1. *Let (X, T) and (Y, T) be two minimal systems and assume that (Y, T) is distal. If X_1 is a nonempty invariant subset of X and $\Phi: X_1 \rightarrow Y$ is a continuous map on X_1 with the induced topology and commuting with the transformations T , then Φ has a continuous extension to X .*

Proof. Let $\Gamma \subset X \times Y$ be the graph of Φ :

$$\Gamma = \{(x, \Phi(x)): x \in X_1\}.$$

Let $\bar{\Gamma}$ be the closure of Γ in $X \times Y$. We claim that $\bar{\Gamma}$ is the graph of some map $\Phi': X \rightarrow Y$.

The projection of $\bar{\Gamma}$ on X is a closed invariant subset of X containing X_1 , and by minimality this projection is equal to X . Assume that $x \in X$ and $y, y' \in Y$ are such that (x, y) and (x, y') belong to $\bar{\Gamma}$. Let $x_1 \in X_1$ and choose a sequence $(n_i)_{i \in \mathbb{N}}$ of integers such that $T^{n_i} x \rightarrow x_1$ and such that the sequences $(T^{n_i} y)_{i \in \mathbb{N}}$ and $(T^{n_i} y')_{i \in \mathbb{N}}$ converge in Y , to the points z and z' , respectively, as $i \rightarrow \infty$. Then (x_1, z) and (x_1, z') belong to $\bar{\Gamma} \cap (X_1 \times Y)$.

On the other hand, since Φ is continuous on X_1 , we have that $\bar{\Gamma} \cap (X_1 \times Y) = \Gamma$ and thus $z = \Phi(x_1) = z'$. Since (Y, T) is distal, we conclude that $y = y'$ and we have that $\bar{\Gamma}$ is the graph of a map $\Phi': X \rightarrow Y$.

The restriction of Φ' to X_1 is equal to Φ and because its graph is closed, Φ' is continuous. Finally, since X_1 is invariant and nonempty, it is dense in X . By minimality and density, we conclude that $\Phi' \circ T = T \circ \Phi'$. \square

2.3. Nilsystems and nilsequences

Definition 2.1. Let $d \geq 1$ be an integer and assume that G is a d -step nilpotent Lie group and that $\Gamma \subset G$ is a discrete, cocompact subgroup of G . The compact manifold $X = G/\Gamma$ is a d -step nilmanifold and G acts naturally on X by left translations: $x \mapsto \tau.x$ for $\tau \in G$.

If T is left multiplication on X by some fixed element of G , then (X, T) is called a d -step nilsystem.

A d -step nilsystem is an example of a distal system. In particular if the nilsystem is transitive, then it is minimal. Also, the closed orbit of a point in a d -step nilsystem is topologically conjugate to a d -step nilsystem. See [2,17], and [15] for proofs and general references on nilsystems.

We also make use of inverse limits of nilsystems and so we recall the definition of an inverse limit of systems (restricting ourselves to the case of sequential inverse limits). If $(X_i, T_i)_{i \in \mathbb{N}}$ are systems and $\pi_i : X_{i+1} \rightarrow X_i$ are factor maps, the *inverse limit* of the systems is defined to be the compact subset of $\prod_{i \in \mathbb{N}} X_i$ given by

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

It is a compact metric space endowed with the distance

$$d(x, y) = \sum_{i \in \mathbb{N}} 1/2^i d_i(x_i, y_i).$$

We note that the maps T_i induce a transformation T on the inverse limit.

Many properties of the systems (X_i, T_i) also pass to the inverse limit, including minimality, distality, and unique ergodicity.

We return to the definition of a nilsequence:

Definition 2.2. If $(X = G/\Gamma, T)$ is a d -step nilsystem, where T is given by multiplication by the element $\tau \in G$, $f : X \rightarrow \mathbb{C}$ is a continuous function, and $x \in X$, the sequence $(f(\tau^n.x) : n \in \mathbb{Z})$ is a *basic d -step nilsequence*. A uniform limit of basic d -step nilsequences is a *nilsequence*.

Equivalently, a d -step nilsequence is given by $(f(T^n x) : n \in \mathbb{Z})$, where (X, T) is an inverse limit of d -step nilsystems, $f : X \rightarrow \mathbb{C}$ is a continuous function and $x \in X$.

The two statements in the definition are shown to be equivalent in Lemma 14 in [14]. Moreover, in the definition of a d -step nilsequence, we can assume that the system is minimal. Namely, considering the closed orbit of x_0 , this is a transitive and so minimal system.

The 1-step nilsystems are translations on compact abelian Lie groups and 1-step nilsequences are exactly almost periodic sequences (see [17]). Examples of 2-step nilsequences and a detailed study of them are given in [12].

3. Dynamical parallelepipeds: first properties

3.1. Notation

Let X be a set, let $d \geq 1$ an integer, and write $[d] = \{1, 2, \dots, d\}$. We view $\{0, 1\}^d$ in one of two ways, either as a sequence $\epsilon = \epsilon_1, \dots, \epsilon_d$ of 0's and 1's written without commas or parentheses;

or as a subset of $[d]$. A subset ϵ corresponds to the sequence $(\epsilon_1, \dots, \epsilon_d) \in \{0, 1\}^d$ such that $i \in \epsilon$ if and only if $\epsilon_i = 1$ for $i \in [d]$.

If $\mathbf{n} = (n_1, \dots, n_d) \in \mathbb{Z}^d$ and $\epsilon \subset [d]$, we define

$$\mathbf{n} \cdot \epsilon = \sum_{i=1}^d n_i \epsilon_i = \sum_{i \in \epsilon} n_i.$$

We denote X^{2^d} by $X^{[d]}$. A point $\mathbf{x} \in X^{[d]}$ can be written in one of two equivalent ways, depending on the context:

$$\mathbf{x} = (x_\epsilon : \epsilon \in \{0, 1\}^d) = (x_\epsilon : \epsilon \subset [d]).$$

For $x \in X$, we write $x^{[d]} = (x, x, \dots, x) \in X^{[d]}$. The diagonal of $X^{[d]}$ is $\Delta^{[d]} = \{x^{[d]} : x \in X\}$.

A point $\mathbf{x} \in X^{[d]}$ can be decomposed as $\mathbf{x} = (\mathbf{x}', \mathbf{x}'')$ with $\mathbf{x}', \mathbf{x}'' \in X^{[d-1]}$, where $\mathbf{x}' = (x_{\epsilon 0} : \epsilon \in \{0, 1\}^{d-1})$ and $\mathbf{x}'' = (x_{\epsilon 1} : \epsilon \in \{0, 1\}^{d-1})$. We can also isolate the first coordinate, writing $X_*^{[d]} = X^{2^{d-1}}$ and then writing a point $\mathbf{x} \in X^{[d]}$ as $\mathbf{x} = (x, \mathbf{x}_*)$, where $\mathbf{x}_* = (x_\epsilon : \epsilon \neq \emptyset) \in X_*^{[d]}$.

The *faces* of dimension r of a point in $\mathbf{x} \in X^{[d]}$ are defined as follows. Let $J \subset [d]$ with $|J| = d - r$ and $\xi \in \{0, 1\}^{d-r}$. The elements $(x_\epsilon : \epsilon \in \{0, 1\}^d, \epsilon_J = \xi)$ of $X^{[r]}$ are called *faces* of dimension r of \mathbf{x} , where $\epsilon_J = (\epsilon_i : i \in J)$. Thus any face of dimension r defines a natural projection from $X^{[d]}$ to $X^{[r]}$, and we call this the *projection along this face*.

Identifying $\{0, 1\}^d$ with the set of vertices of the Euclidean unit cube, a Euclidean isometry of the unit cube permutes the vertices of the cube and thus the coordinates of a point $x \in X^{[d]}$. These permutations are the *Euclidean permutations* of $X^{[d]}$. Examples of Euclidean permutations are permutations of digits, meaning a permutation of $\{0, 1\}^d$ induced by a permutation of $[d]$, and symmetries, such as replacing ϵ_i by $1 - \epsilon_i$ for some i . For $d = 2$, an example of a digit permutation is the map $(00, 01, 10, 11) \mapsto (00, 10, 01, 11)$ and an example of a symmetry is the map $(00, 01, 10, 11) \mapsto (01, 00, 11, 10)$.

3.2. Dynamical parallelepipeds

We recall that $\mathbf{Q}^{[d]}$ is the closure in X^{2^d} of elements of the form

$$(T^{n_1 \epsilon_1 + \dots + n_d \epsilon_d} x : \epsilon \in \{0, 1\}^d),$$

where $\mathbf{n} = (n_1, \dots, n_d) \in \mathbb{Z}^d$ and $x \in X$ (Definition 1.1). It follows immediately from the definition that $\mathbf{Q}^{[d]}$ contains the diagonal.

Some other basic structural properties of $\mathbf{Q}^{[d]}$ are:

- (1) Any face of dimension r of any $\mathbf{x} \in \mathbf{Q}^{[d]}$ belongs to $\mathbf{Q}^{[r]}$. (This condition is trivial for $d = 2$.)
- (2) $\mathbf{Q}^{[d]}$ is invariant under the Euclidean permutations of $X^{[d]}$.
- (3) If $\mathbf{x} \in \mathbf{Q}^{[d]}$, then $(\mathbf{x}, \mathbf{x}) \in \mathbf{Q}^{[d+1]}$.

Lemma 3.1. *Let $d \geq 1$ be an integer, (X, T) and (Y, T) be systems, and $\pi : X \rightarrow Y$ be a factor map. Then $\mathbf{Q}^{[d]}(Y)$ is the image of $\mathbf{Q}^{[d]}(X)$ under the map $\pi^{[d]} := \pi \times \dots \times \pi$ (2^d times).*

We can rephrase the definition of $\mathbf{Q}^{[d]}$ using some groups of transformations on $X^{[d]}$. We define:

Definition 3.1. Let (X, T) be a system and $d \geq 1$ be an integer. The *diagonal transformation* of $X^{[d]}$ is the map given by $(T^{[d]} \mathbf{x})_\epsilon = Tx_\epsilon$ for every $\mathbf{x} \in X^{[d]}$ and every $\epsilon \in [d]$.

For $j \in [d]$, the *face transformation* $T_j^{[d]} : X^{[d]} \rightarrow X^{[d]}$ is defined for every $\mathbf{x} \in X^{[d]}$ and $\epsilon \in [d]$ by:

$$T_j^{[d]} \mathbf{x} = \begin{cases} (T_j^{[d]} \mathbf{x})_\epsilon = Tx_\epsilon & \text{if } j \in \epsilon, \\ (T_j^{[d]} \mathbf{x})_\epsilon = x_\epsilon & \text{if } j \notin \epsilon. \end{cases}$$

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 and the face transformations. 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, $\mathcal{F}^{[d]} = \{S \in \mathcal{G}^{[d]} : S_\emptyset = \mathbf{Id}\}$.

We note that the group $\mathcal{G}^{[d]}$ satisfies the three properties (3.2)–(3.2) above, with $\mathbf{Q}^{[d]}$ replaced by $\mathcal{G}^{[d]}$. Moreover, for $S \in \mathcal{F}^{[d]}$, we have that $(S, S) \in \mathcal{F}^{[d+1]}$. As well, $\mathcal{F}^{[d]}$ is invariant under digit permutations.

The following lemma follows directly from the definitions:

Lemma 3.2. Let (X, T) be a system and let $d \geq 1$ be an integer. Then $\mathbf{Q}^{[d]}$ 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 closed orbit of $x^{[d]}$ under the group $\mathcal{G}^{[d]}$.

3.3. Definition of the regionally proximal relations

In this section, we discuss the relation $\mathbf{RP}^{[d]}$ and its relation to $\mathbf{Q}^{[d+1]}$.

Definition 3.2. Let (X, T) be a system and let $d \geq 1$ be an integer. 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 $\mathbf{n} = (n_1, \dots, n_d) \in \mathbb{Z}^d$ such that $d(x, x') < \delta$, $d(y, y') < \delta$, and

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

(In other words, there exists $S \in \mathcal{F}^{[d]}$ such that $d(S_\epsilon \cdot x', S_\epsilon \cdot y') < \delta$ for every $\epsilon \neq \emptyset$.) We call this the *regionally proximal relation of order d* and denote the set of regionally proximal points by $\mathbf{RP}^{[d]}$ (or by $\mathbf{RP}^{[d]}(X)$ in case of ambiguity).

Since $\mathbf{RP}^{[d+1]}$ is finer than $\mathbf{RP}^{[d]}$, we have defined a nested sequence of closed and invariant relations.

Lemma 3.3. Assume that (X, T) is a transitive system and that $d \geq 1$ is an integer. Then $(x, y) \in \mathbf{RP}^{[d]}$ if and only if there exists $\mathbf{a}_* \in X_*^{[d]}$ such that

$$(x, \mathbf{a}_*, y, \mathbf{a}_*) \in \mathbf{Q}^{[d+1]}.$$

Proof. Assume that $(x, y) \in \mathbf{RP}^{[d]}$. Let $\delta > 0$ and let x', y' and S be as in the definition of regionally proximal points. As transitive points are dense in X , there exists a transitive point z with $d(z, x') < \delta$ and, for every $\epsilon \neq \emptyset$, $d(S_\epsilon \cdot z, S_\epsilon \cdot x') < \delta$. There exists an integer k such that $d(T^k z, y') < \delta$ and that, for every $\epsilon \neq \emptyset$, $d(S_\epsilon \cdot T^k z, S_\epsilon \cdot y') < \delta$. We have that $d(z, x) < 2\delta$, $d(T^k z, y) < 2\delta$ and $d(S_\epsilon \cdot T^k z, S_\epsilon \cdot z) < 3\delta$.

Define $\mathbf{z} \in X^{[d+1]}$ by $z_{\epsilon 0} = S_\epsilon \cdot z$ and $z_{\epsilon 1} = S_\epsilon \cdot T^k z$ for $\epsilon \in \{0, 1\}^d$. Then $\mathbf{z} = (S, S)(T_{d+1}^{[d+1]})^k z^{[d+1]}$ and thus this point belongs to $\mathbf{Q}^{[d+1]}$. We have that $d(z_{\emptyset}, x) < 2\delta$, $d(z_{0\dots 01}, y) < 2\delta$ and $d(z_{\epsilon 0}, z_{\epsilon 1}) < 3\delta$ for every $\epsilon \in \{0, 1\}^d$ different from \emptyset . Letting $\delta \rightarrow 0$ and passing to a subsequence, we have a point of $\mathbf{Q}^{[d+1]}$ of the announced form.

Conversely, if $(x, \mathbf{a}_*, y, \mathbf{a}_*) \in \mathbf{Q}^{[d+1]}$ with $\mathbf{a}_* \in X_*^{[d]}$, then for every $\delta > 0$, there exist $z \in X$, $\mathbf{n} \in \mathbb{Z}^d$, and $p \in \mathbb{Z}$ such that $d(z, x) < \delta$, $d(T^p z, y) < \delta$, and $d(T^{\mathbf{n}\cdot\epsilon} z, a_\epsilon) < \delta$ and $d(T^{\mathbf{n}\cdot\epsilon+p} z, a_\epsilon) < \delta$ for every nonempty $\epsilon \subset [d]$. Thus $(x, y) \in \mathbf{RP}^{[d]}$. \square

Corollary 3.1. Assume that (X, T) is a transitive system and that $d \geq 1$ is an integer. The relation $\mathbf{RP}^{[d]}(X)$ is a closed, symmetric relation that is invariant under T .

If $\phi : X \rightarrow Y$ is a factor map and if $(x, y) \in \mathbf{RP}^{[d]}(X)$, then $(\phi(x), \phi(y)) \in \mathbf{RP}^{[d]}(Y)$.

Proof. This follows immediately from the definition and Lemma 3.3. \square

If the first property of Theorem 1.2 holds, then the relation $\mathbf{RP}^{[d]}$ is trivial: if $(x, \mathbf{a}_*, y, \mathbf{a}_*) \in \mathbf{Q}^{[d+1]}$, then $(x, \mathbf{a}_*) \in \mathbf{Q}^{[d]}$ and so $(x, \mathbf{a}_*, x, \mathbf{a}_*) \in \mathbf{Q}^{[d+1]}$. By the first property of Theorem 1.2, $x = y$.

3.4. Reduction to the distal case

We show that systems verifying the conditions of Theorem 1.2 are distal.

Proposition 3.1. Assume (X, T) is a transitive system and that $d \geq 1$ is an integer. If x and y are proximal and the closed orbit of y is a minimal set, then $(x, y, y, \dots, y) \in \mathbf{Q}^{[d]}$.

Proof. First we claim that for every $\eta > 0$, there exists $n \in \mathbb{N}$ such that $d(T^n x, y) < \eta$ and $d(T^n y, y) < \eta$. Since x and y are proximal, there exists a sequence $(m_i : i \geq 1)$ and a point $z \in X$ such that $T^{m_i} x \rightarrow z$ and $T^{m_i} y \rightarrow z$. We have that z belongs to the closed orbit of y , which is minimal, and so y belongs to the closed orbit of z . Thus there exists p such that $d(T^p z, y) < \eta/2$. By continuity of T^p , for i sufficiently large we have that $d(T^{m_i+p} x, y) < \eta$ and $d(T^{m_i+p} y, y) < \eta$. Setting $n = m_i + p$ for some sufficiently large i , we have n that satisfies the claim.

Fix $\delta > 0$. Applying the claim for $\eta = \delta/d$, we find some n_1 such that $d(T^{n_1} x, y) < \delta/d$ and $d(T^{n_1} y, y) < \delta/d$.

Taking η with $0 < \eta < \delta/d$ such that $d(T^{n_1} u, T^{n_1} v) \leq \delta/d$ when $d(u, v) \leq \eta$, and then taking n_2 associated to this η , from the claim we have that: $d(T^{n_1 \epsilon_1 + n_2 \epsilon_2} x, y) < 2\delta/d$ and $d(T^{n_1 \epsilon_1 + n_2 \epsilon_2} y, y) < 2\delta/d$ for all $\epsilon_1, \epsilon_2 \in \{0, 1\}^2$ other than $\epsilon_1 = \epsilon_2 = 0$.

Thus by induction, there is a sequence of integers n_1, \dots, n_d such that $d(T^{n^\epsilon}x, y) < \delta$ for all $\emptyset \neq \epsilon \subset [d]$. Taking $\delta \rightarrow 0$, we have the statement of the proposition. \square

Corollary 3.2. *Assume that (X, T) is a transitive system. If the second property of Theorem 1.2 holds, then X is distal.*

Proof. We first show that any point in X is minimal, i.e. its closed orbit is minimal, and so the system is minimal. Every $x \in X$ is proximal to some minimal point y (see [1]). By the previous proposition and the hypothesis, $x = y$ and so x is a minimal point. Applying the proposition to any pair of proximal points, the statement follows. \square

4. Parallelepipeds in distal systems

4.1. Minimal distal systems and parallelepiped structures

Lemma 4.1. *Let (X, T) be a minimal distal system and let $d \geq 1$ be an integer. Then $(\mathbf{Q}^{[d]}, \mathcal{G}^{[d]})$ is a minimal distal system.*

Proof. Since (X, T) is distal, so is the system $(X^{[d]}, \mathcal{G}^{[d]})$. Since $\mathbf{Q}^{[d]}$ is a closed and invariant subset of $X^{[d]}$ under the face transformations, the system $(\mathbf{Q}^{[d]}, \mathcal{G}^{[d]})$ is also distal. By the second part of Lemma 3.2, the system is transitive and thus is minimal. \square

Using the Ellis semigroup, Eli Glasner [6] showed us a proof that this lemma holds without the assumption of distality.

Although we do not make use of the following proposition in the sequel, we include it as it is an analog of the geometric property of parallelepipeds in a vector space:

Proposition 4.1. *Let (X, T) be a minimal distal system and let $d \geq 1$ be an integer. The relation \sim_{d-1} defined on $\mathbf{Q}^{[d-1]}$ by*

$$\mathbf{x} \sim_{d-1} \mathbf{x}' \text{ if and only if the element } (\mathbf{x}, \mathbf{x}') \in X^{[d]} \text{ belongs to } \mathbf{Q}^{[d]}$$

is an equivalence relation.

Proof. By Property (2) of Section 3.2, we have that the relation is symmetric and by Property (3), it is reflexive. We are left with showing that the relation is transitive. Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{Q}^{[d-1]}$ and assume that $(\mathbf{u}, \mathbf{v}) \in \mathbf{Q}^{[d]}$ and $(\mathbf{v}, \mathbf{w}) \in \mathbf{Q}^{[d]}$.

Choose $z \in X$. By Lemma 4.1, the system $(\mathbf{Q}^{[d]}, \mathcal{G}^{[d]})$ is minimal and so it is the closed orbit of $z^{[d]}$ under the group $\mathcal{G}^{[d]}$. There exists a sequence $(S_i : i \geq 1)$ such that $S_i(\mathbf{u}, \mathbf{v}) \rightarrow z^{[d]} = (z^{[d-1]}, z^{[d-1]})$ as $i \rightarrow \infty$. Writing $S_i = (S'_i, S''_i)$ with $S'_i, S''_i \in \mathcal{G}^{[d-1]}$, we have that $S'_i \mathbf{u} \rightarrow z^{[d-1]}$ and $S''_i \mathbf{v} \rightarrow z^{[d-1]}$.

Passing to a subsequence if needed, we can assume that $S''_i \mathbf{w}$ converges to some point $\hat{\mathbf{z}} \in X^{[d-1]}$ as $i \rightarrow \infty$. We have that

$$(S''_i, S'_i)(\mathbf{v}, \mathbf{w}) \rightarrow (z^{[d-1]}, \hat{\mathbf{z}}) \in X^{[d]}.$$

But for each $i \in \mathbb{N}$, $(S_i'', S_i'') \in \mathcal{G}^{[d]}$ and thus $(z^{[d-1]}, \hat{\mathbf{z}})$ belongs to the closed orbit of (\mathbf{v}, \mathbf{w}) under $\mathcal{G}^{[d]}$ and so $(z^{[d-1]}, \hat{\mathbf{z}}) \in \mathbf{Q}^{[d]}$.

On the other hand, $S_i(\mathbf{u}, \mathbf{w}) = (S_i'\mathbf{u}, S_i''\mathbf{w})$ converges to $(z^{[d-1]}, \hat{\mathbf{z}})$ and this point belongs to the closed orbit of (\mathbf{u}, \mathbf{w}) under $\mathcal{G}^{[d]}$. By distality this orbit is minimal and so it follows that (\mathbf{u}, \mathbf{w}) also belongs to the orbit closure of $(z^{[d-1]}, \hat{\mathbf{z}})$. In particular, $(\mathbf{u}, \mathbf{w}) \in \mathbf{Q}^{[d]}$ and the relation \sim_{d-1} is transitive. \square

Corollary 4.1. *Let (X, T) be a minimal distal system and let $d \geq 1$ be an integer. If $\mathbf{x}, \mathbf{y} \in \mathbf{Q}^{[d+1]}$ and $x_\epsilon = y_\epsilon$ for all $\epsilon \neq \emptyset$, then $(x_\emptyset, y_\emptyset) \in \mathbf{RP}^{[d]}$.*

Proof. We write $\mathbf{x} = (x_\emptyset, \mathbf{a}_*, \mathbf{z})$ with $\mathbf{a}_* \in X_*^{[d]}$ and $\mathbf{z} \in \mathbf{Q}^{[d]}$. By hypothesis, $\mathbf{y} = (y_\emptyset, \mathbf{a}_*, \mathbf{z})$ and by transitivity of relation \sim_{d+1} , we have that $(x_\emptyset, \mathbf{a}_*, y_\emptyset, \mathbf{a}_*) \in \mathbf{Q}^{[d+1]}$. We conclude via Lemma 3.3. \square

4.2. Completing parallelepipeds

Notation. For $x \in X$ and $d \geq 1$, write

$$\mathbf{Q}^{[d]}(x) = \{\mathbf{y} \in \mathbf{Q}^{[d]}: y_\emptyset = x\}.$$

In this section, we show:

Proposition 4.2. *For $x \in X$ and $d \geq 1$, $\mathbf{Q}^{[d]}(x)$ is the closed orbit of $x^{[d]}$ under the action of the group $\mathcal{F}^{[d]}$.*

Proposition 4.2 follows from the more general Proposition 4.3 below.

In this section (and only in this section), we make use of yet another notation for the points of $X^{[d]}$:

Notation. For $\epsilon \subset [d]$, define

$$\sigma_d(\epsilon) = \sum_{k=1}^d \epsilon_k 2^{k-1}.$$

For $0 \leq j < 2^d$, set

$$E(d, j) = \{\epsilon \subset [d]: \sigma_d(\epsilon) \leq j\}.$$

For $x \in X$ and $d \geq 1$, let $\mathbf{K}^{[d]}(x)$ denote the closed orbit of $x^{[d]}$ under $\mathcal{F}^{[d]}$.

We remark that $\mathbf{K}^{[d]}(x)$ is minimal under the action of $\mathcal{F}^{[d]}$. Moreover, if $d \geq 2$ and $\mathbf{y} \in \mathbf{K}^{[d-1]}(x)$, then $(\mathbf{y}, \mathbf{y}) \in \mathbf{K}^{[d]}(x)$. As well, $\mathbf{K}^{[d]}(x)$ is invariant under digit permutations.

Proposition 4.3. *Assume that $d \geq 1$ is an integer and let $0 \leq j < 2^d$. Assume that $\mathbf{x} \in X^{[d]}$ satisfies the hypothesis*

$H(d, j)$: for every r and every face F of dimension r of $\{0, 1\}^d$ included in $E(d, j)$, the projection of \mathbf{x} along F belongs to $\mathbf{Q}^{[r]}$.

Then there exists $\mathbf{w} \in \mathbf{K}^{[d]}(x_\emptyset)$ such that $w_\epsilon = x_\epsilon$ for every $\epsilon \in E(d, j)$.

First we remark that this proposition implies Proposition 4.2. Indeed, if $\mathbf{x} \in \mathbf{Q}^{[d]}(x)$ then $x_\emptyset = x$. Moreover, \mathbf{x} satisfies the hypothesis $H(d, 2^d - 1)$ and thus agrees with a point of $\mathbf{K}^{[d]}(x)$ on $E(d, 2^d - 1)$, which is the set of all $\epsilon \subset [d]$.

Proof. For $d = 1$, the result is obvious since $\mathbf{K}^{[1]}(x_\emptyset) = \{x_\emptyset\} \times X$. For $d > 1$ and $j = 0$, there is nothing to prove.

We proceed by induction: take $d > 1$ and $j > 0$ and assume that the result holds for $d - 1$ and all values of j and for d and $j' < j$.

Assume that $\mathbf{x} \in X^{[d]}$ satisfies the hypothesis $H(d, j)$ and write $x = x_\emptyset$.

4.2.1. We first make a reduction. We assume that the result holds under the additional hypothesis

(*) \mathbf{x} is of the form $x_\epsilon = x_\emptyset$ for $\epsilon \in E(d, j - 1)$

and we show that it holds in the general case.

Assume that \mathbf{x} satisfies $H(d, j)$. By the induction hypothesis, there exists $\mathbf{v} \in \mathbf{K}^{[d]}(x)$ such that $v_\epsilon = x_\epsilon$ for all $\epsilon \in E(d, j - 1)$. By minimality, the point $x^{[d]}$ lies in the closed $\mathcal{F}^{[d]}$ -orbit of \mathbf{v} , meaning that there exists a sequence $(S_\ell: \ell \geq 1)$ in $\mathcal{F}^{[d]}$ such that $S_\ell \mathbf{v} \rightarrow x^{[d]}$. Passing to a subsequence, we can assume that $S_\ell \mathbf{x} \rightarrow \mathbf{x}'$. We have that $x'_\epsilon = x$ for all $\epsilon \in E(d, j - 1)$ and \mathbf{x}' satisfies property (*).

Property $H(d, j)$ is invariant under the action of $\mathcal{F}^{[d]}$ and under passage to limits. Thus since \mathbf{x}' lies in the closed $\mathcal{F}^{[d]}$ -orbit of \mathbf{x} , \mathbf{x}' satisfies $H(d, j)$. Using the result of the proposition with the additional assumption of (*), we have that there exists $\mathbf{v}' \in \mathbf{K}^{[d]}(x)$ such that $v'_\epsilon = x'_\epsilon$ for $\epsilon \in E(d, j)$.

Since the system is distal and \mathbf{x}' belongs to the closed $\mathcal{F}^{[d]}$ -orbit of \mathbf{x} , we also have that \mathbf{x} belongs to the closed $\mathcal{F}^{[d]}$ -orbit of \mathbf{x}' . There exists a sequence $(S'_\ell: \ell \geq 1)$ such that $S'_\ell \mathbf{x}' \rightarrow \mathbf{x}$. Passing to a subsequence, we have that $S'_\ell \mathbf{v}' \rightarrow \mathbf{u}$. Thus $\mathbf{u} \in \mathbf{K}^{[d]}(x)$ and $u_\epsilon = x_\epsilon$ for $\epsilon \in E(d, j)$.

4.2.2. We now assume \mathbf{x} satisfies $H(d, j)$ and (*) and assume that $j \neq 2^d - 1$. Again, we write $x = x_\emptyset$.

Let $\eta \in \{0, 1\}^d$ be defined by $\sigma_d(\eta) = j$. By hypothesis, there exists some k with $1 \leq k \leq d$ such that $\eta_k = 0$. Choose k to be the largest k with this property.

Define the map $\Phi: \{0, 1\}^{d-1} \rightarrow \{0, 1\}^d$ by

$$\Phi(\epsilon) = \epsilon_1 \dots \epsilon_{k-1} 0 \epsilon_k \dots \epsilon_{d-1}.$$

Setting

$$\theta = \eta_1 \dots \eta_{k-1} 1 \dots 1 \in \{0, 1\}^{d-1},$$

we have that $\Phi(\theta) = \eta$.

Set $i = \sigma_{d-1}(\theta)$. It is easy to check that for $\alpha \in \{0, 1\}^{d-1}$,

$$\sigma_{d-1}(\alpha) < i \quad \text{if and only if} \quad \sigma_d(\Phi(\alpha)) < \sigma_d(\Phi(\theta)) = j. \tag{3}$$

In particular, $\Phi(E(d - 1, i)) \subset E(d, j)$.

Define $\mathbf{u} \in X^{[d-1]}$ to be the projection of \mathbf{x} on $X^{[d-1]}$ along the face defined by $\epsilon_k = 0$. In other words,

$$u_\epsilon = x_{\Phi(\epsilon)}, \quad \epsilon \in \{0, 1\}^{d-1}.$$

Moreover, if F is a face of $\{0, 1\}^{d-1}$, then $\Phi(F)$ is a face of $\{0, 1\}^d$. Since \mathbf{x} satisfies $H(d, j)$, we have that \mathbf{u} satisfies $H(d - 1, i)$.

We have that $u_\theta = x$ and by the induction hypothesis, there exists $\mathbf{v} \in \mathbf{K}^{[d-1]}(x)$ with $v_\epsilon = u_\epsilon$ for all $\epsilon \in E(d - 1, i)$.

Define the map $\Psi : \{0, 1\}^d \rightarrow \{0, 1\}^{d-1}$ by

$$\Psi(\epsilon) = \epsilon_1 \dots \epsilon_{k-1} \epsilon_{k+1} \dots \epsilon_d.$$

By definition, $\Psi \circ \Phi$ is the identity and $\Psi(\eta) = \theta$. On the other hand, $\Phi \circ \Psi(\epsilon) = \epsilon_1 \dots \epsilon_{k-1} 0 \epsilon_{k+1} \dots \epsilon_d$. In particular,

$$\sigma_d(\Phi \circ \Psi(\epsilon)) \leq \sigma_d(\epsilon) \quad \text{for every } \epsilon \in \{0, 1\}^d.$$

Define $\mathbf{w} \in X^{[d]}$ by $w_\epsilon = v_{\Psi(\epsilon)}$ for $\epsilon \in \{0, 1\}^d$. In other words, \mathbf{w} is obtained by duplicating \mathbf{v} on two opposite faces. We check that $\mathbf{w} \in \mathbf{K}^{[d]}(x)$.

To see this, let \mathbf{v}' be obtained from \mathbf{v} by the digit permutation that exchanges the digits $k - 1$ and $d - 1$. Then $\mathbf{v}' \in \mathbf{K}^{[d-1]}(x)$ and so $(\mathbf{v}', \mathbf{v}') \in \mathbf{K}^{[d]}(x)$. We obtain \mathbf{w} from the point $(\mathbf{v}', \mathbf{v}')$ by the digit permutation that exchanges the digits k and d .

We claim that $\Psi(E(d, j - 1)) \subset E(d - 1, i - 1)$. To show this, we take $\epsilon \in E(d, j - 1)$ and distinguish two cases. First assume there exists some m with $k + 1 \leq m \leq d$ with $\epsilon_m = 0$. Then one of the $d - k$ last coordinates of $\Psi(\epsilon) = 0$ and by definition of θ , $\sigma_{d-1}(\Psi(\epsilon)) < \sigma_{d-1}(\theta) = i$.

Now assume that there is no such m . Because $\sigma_d(\epsilon) < \sigma_d(\eta)$ and $\eta_k = 0$, we have that $\epsilon_k = 0$. Then $\Phi(\Psi(\epsilon)) = \epsilon$. Thus

$$\sigma_d(\Phi(\Psi(\epsilon))) = \sigma_d(\epsilon) < j$$

and applying (3) with $\alpha = \Psi(\epsilon)$, we have that $\sigma_{d-1}(\Psi(\epsilon)) < i$. This proves the claim.

We check that \mathbf{w} satisfies the conclusion of the proposition. First for w_η , we have that $w_\eta = v_{\Psi(\eta)} = v_\theta = u_\theta$ since $\theta \in E(d - 1, i)$, and $u_\theta = x_{\Phi(\theta)} = x_\eta$. Thus $w_\eta = x_\eta$. Next, if $\epsilon \in E(d, j - 1)$, then $x_\epsilon = x$. On the other hand, $w_\epsilon = v_{\Psi(\epsilon)} = u_{\Psi(\epsilon)}$, where the last equality holds because by the claim we have $\Psi(\epsilon) \in E(d - 1, i - 1)$. But $u_{\Psi(\epsilon)} = x_{\Phi \circ \Psi(\epsilon)} = x$, because $\sigma_d(\Phi \circ \Psi(\epsilon)) \leq \sigma_d(\epsilon) \leq j - 1$. This \mathbf{w} is as announced.

4.2.3. We are left with considering the case that $j = 2^d - 1$. The hypothesis $H(d, 2^d - 1)$ means that $\mathbf{x} = (x, x, \dots, x, y) \in \mathbf{Q}^{[d]}$ and we have to show that this lies in $\mathbf{K}^{[d]}(x)$.

We start with a general property. Writing a point $\mathbf{x} \in X^{[d]}$ as $\mathbf{x} = (\mathbf{x}', \mathbf{x}'')$, define the projection $\phi: \mathbf{K}^{[d]}(x) \rightarrow \mathbf{Q}^{[d-1]}$ by $\phi(\mathbf{x}) = \mathbf{x}''$. The range of ϕ is invariant under the group $\mathcal{G}^{[d-1]}$ and thus by Lemma 4.1, it is equal to $\mathbf{Q}^{[d-1]}$. By distality, the map ϕ is open.

Assume $(x, x, \dots, x, y) \in \mathbf{Q}^{[d]}$. Write $\mathbf{v} = (x, \dots, x, y) \in X^{[d-1]}$. Let $\delta > 0$. Since $(x^{[d-1]}, x^{[d-1]}) \in \mathbf{K}^{[d]}(x)$, by the openness of ϕ , there exists δ' with $0 < \delta' < \delta$ such that if $\mathbf{u} \in \mathbf{Q}^{[d-1]}$ is δ' -close to $x^{[d-1]}$, there exists \mathbf{z} that is δ -close to $x^{[d-1]}$ and $(\mathbf{z}, \mathbf{u}) \in \mathbf{K}^{[d]}(x)$.

Since $(x^{[d-1]}, \mathbf{v}) \in \mathbf{Q}^{[d]}$, there exists $\mathbf{u} \in \mathbf{Q}^{[d-1]}$ and $n \in \mathbb{Z}$ such that \mathbf{u} is at most distance δ' from $x^{[d-1]}$ and $(T^{[d-1]})^n \mathbf{u}$ is at most distance δ from \mathbf{v} . Taking \mathbf{z} as above, we have that $(\mathbf{z}, (T^{[d-1]})^n \mathbf{u}) \in \mathbf{K}^{[d]}(x)$ and is δ -close to $(x^{[d-1]}, \mathbf{v})$.

Letting δ go to 0, we have that $(x^{[d-1]}, \mathbf{v}) \in \mathbf{K}^{[d]}(x)$. \square

The next result follows directly from Proposition 4.3 and the definition of $\mathbf{Q}^{[d]}$. It shows that $\mathbf{Q}^{[d]}$ verifies properties that are generalizations of the 2- and 3-dimensional parallelepiped structures as defined in [14]. In particular, $\mathbf{Q}^{[d]}$ satisfies the “property of closing parallelepipeds”. This plays a key role in our study of the first condition in Theorem 1.2.

Proposition 4.4. *Let (X, T) be a minimal distal system and let $d \geq 1$ be an integer. Assume that $x_\epsilon, \epsilon \in [d]$ with $\epsilon \neq [d]$, are points in X such that the face $(x_\epsilon: j \notin \epsilon)$ belongs to $\mathbf{Q}^{[d-1]}$ for each $j \in [d]$. Then there exists $x_{[d]} \in X$ such that $(x_\epsilon: \epsilon \in [d]) \in \mathbf{Q}^{[d]}$.*

Although we have given the last coordinate in the statement of this proposition a particular role, using Euclidean permutations the analogous statement holds for any other fixed coordinate, provided that the corresponding faces lie in $\mathbf{Q}^{[d-1]}$.

4.3. Strong form of the regionally proximal relation

Corollary 4.2. *Let (X, T) be a minimal distal system and let $d \geq 1$ be an integer. Let $x, y \in X$ and $\mathbf{b}_* \in X_*^{[d+1]}$ with $(x, \mathbf{b}_*) \in \mathbf{Q}^{[d+1]}$. Then $(y, \mathbf{b}_*) \in \mathbf{Q}^{[d+1]}$ if and only if $(y, x, x, \dots, x) \in \mathbf{Q}^{[d+1]}$.*

Proof. We write $\mathbf{u} = (x, \mathbf{b}_*)$, $\mathbf{v} = (y, \mathbf{b}_*)$, and $\mathbf{y} = (y, x, x, \dots, x) \in X^{[d+1]}$. By Proposition 4.3, we have that \mathbf{u} belongs to $\mathbf{K}^{[d+1]}(x)$ and, by minimality, there exists a sequence $(S_n: n \geq 1)$ in $\mathcal{F}^{[d+1]}$ such that $S_n \mathbf{u} \rightarrow x^{[d+1]}$. Then $S_n \mathbf{v} \rightarrow \mathbf{y}$ and \mathbf{y} belongs to the closed orbit of \mathbf{v} under $\mathcal{F}^{[d+1]}$. By distality, this last property implies that \mathbf{v} belongs to the closed orbit of \mathbf{y} . Since $\mathbf{Q}^{[d+1]}$ is closed and invariant under $\mathcal{F}^{[d+1]}$, we have that $\mathbf{y} \in \mathbf{Q}^{[d+1]}$ if and only if $\mathbf{v} \in \mathbf{Q}^{[d+1]}$. \square

Corollary 4.3. *Let (X, T) be a minimal distal system and let $d \geq 1$ be an integer. Let $x, y \in X$. Then $(x, y) \in \mathbf{RP}^{[d]}$ if and only if $(y, x, x, \dots, x) \in \mathbf{Q}^{[d+1]} = \mathbf{K}^{[d+1]}(y)$.*

Proof. For $\mathbf{a}^* \in X_*^{[d]}$, apply the preceding corollary with $\mathbf{b}_* = (\mathbf{a}_*, x, \mathbf{a}_*)$ and use Lemma 3.3. \square

The combination of the previous corollaries allows to prove that each coordinate in a parallelepiped of $\mathbf{Q}^{[d]}$ can be replaced by another point that is regionally proximal of order d with it and the resulting point is still a parallelepiped.

We finish with a comment about the regionally proximal relation of order d . In [16], McMahon (see also Auslander [1, Corollary 10, Chapter 9]) proves that in the definition of the regionally proximal relation, the point x' (see Definition 3.2 with $d = 1$) can be taken to be x . The same result can be stated for the regionally proximal relation of order d in the distal case. In fact, a similar argument to the one used to prove Lemma 3.3 allows us to show that: $(x, y, \dots, y) \in \mathbf{Q}^{[d+1]} = \mathbf{K}^{[d+1]}(x)$ if and only if for any $\delta > 0$ there exist $y' \in X$ and a vector $\mathbf{n} = (n_1, \dots, n_d) \in \mathbb{Z}^d$ such that for any nonempty $\epsilon \subset [d]$

$$d(y, y') < \delta, \quad d(T^{\mathbf{n}-\epsilon}x, y) < \delta, \quad \text{and} \quad d(T^{\mathbf{n}-\epsilon}y', y) < \delta.$$

4.4. Summarizing

4.4.1. We show that the second property in Theorem 1.2 implies the first one. Assume that the transitive system (X, T) satisfies the second property. By Corollary 3.2, the system is distal.

If $\mathbf{x}, \mathbf{y} \in \mathbf{Q}^{[d+1]}$ agree on all coordinates other than the coordinate indexed by \emptyset , then $\mathbf{x} = \mathbf{y}$ by Corollary 4.2. By permutation of coordinates we deduce that the first property of Theorem 1.2 is satisfied.

The first two properties of this theorem are thus equivalent. From the above discussion, Proposition 1.1 follows: these properties mean that the relation $\mathbf{RP}^{[d]}$ is trivial.

4.4.2.

Proposition 4.5. *Let (X, T) be a minimal distal system and let $d \geq 1$ be an integer. Then the relation $\mathbf{RP}^{[d]}$ is a closed invariant equivalence relation on X .*

The quotient of X under this equivalence relation is the maximal factor of order d of X .

The second statement means that this quotient is a system of order d and that every system of order d which is a factor of X is a factor of this quotient.

Proof. In order to prove the first statement, we are left with showing that the relation is transitive. Assume that (x, y) and $(y, z) \in \mathbf{RP}^{[d]}$. By Corollary 4.3 applied to the pair (x, y) , $(y, x, x, \dots, x) \in \mathbf{Q}^{[d+1]}$. By Corollary 4.2 applied to the pair (y, z) , $(z, x, x, \dots, x) \in \mathbf{Q}^{[d+1]}$ and by Corollary 4.3 again, $(x, z) \in \mathbf{RP}^{[d]}$.

We show now the second part of the proposition. Let Y be the quotient of X under the equivalence relation $\mathbf{RP}^{[d]}$ and let ϕ denote the factor map. Let $(a, b) \in \mathbf{RP}^{[d]}(Y)$. Then $(a, b, b, \dots, b) \in \mathbf{Q}^{[d+1]}(Y)$. By Lemma 3.1, there exists $\mathbf{x} \in \mathbf{Q}^{[d+1]}(X)$ satisfying $\phi^{[d+1]}(\mathbf{x}) = (a, b, b, \dots, b)$.

Write $x_\emptyset = x$ and $x_{00\dots 01} = y$. For every $\epsilon \neq \emptyset$, $\phi(x_\epsilon) = b = \phi(y)$. Thus $(x_\epsilon, y) \in \mathbf{RP}^{[d]}(X)$. Using Corollary 4.3 and Corollary 4.2, we can replace x_ϵ by y in \mathbf{x} and obtain an element of $\mathbf{Q}^{[d+1]}(X)$. Doing this for all $\epsilon \neq \emptyset$, we have that $(x, y, y, \dots, y) \in \mathbf{Q}^{[d+1]}(X)$. By Corollary 4.3, this means that $(x, y) \in \mathbf{RP}^{[d]}(X)$. Thus that $\phi(x) = \phi(y)$ and so $a = b$.

Let W be a system of order d and let $\psi : X \rightarrow W$ be a factor map. Take Y and ϕ as above and let $x, y \in X$. If $\phi(x) = \phi(y)$, then $(x, y) \in \mathbf{RP}^{[d]}(X)$. Thus by Corollary 3.1, $(\psi(x), \psi(y)) \in \mathbf{RP}^{[d]}(W)$ and thus $\psi(x) = \psi(y)$. \square

4.4.3. In order to complete the proofs of Theorems 1.2 and 1.3, we are left with showing that the notions of a system of order d and an inverse limit of d -step minimal nilsystems are equivalent.

In one direction, a result from Appendix B of [13], translated into our current vocabulary, states that a d -step minimal nilsystem is a system of order d . This property easily passes to inverse limits, and so we have:

Proposition 4.6. *Let (X, T) be an inverse limit of minimal $(d - 1)$ -step nilsystems and let $d \geq 2$ be an integer. Then (X, T) is a system of order $d - 1$.*

We are left with showing the converse, which is:

Theorem 4.1. *Assume that (X, T) is a transitive system of order $d - 1$. Then it is an inverse limit of $(d - 1)$ -step minimal nilsystems.*

We recall that the hypothesis of this theorem means that if $\mathbf{x}, \mathbf{y} \in \mathbf{Q}^{[d]}$ have $2^d - 1$ coordinates in common, then $\mathbf{x} = \mathbf{y}$. In particular, this implies that the system is distal and minimal.

The proof of this theorem is carried out in the next two sections.

5. Ergodic preliminaries

The result of Theorem 4.1 is established in the next section using invariant measures on X . In this section, we summarize the background material and give some preliminary results.

5.1. Inverse limits of nilsystems

A *measure preserving system* is defined to be a quadruple (X, \mathcal{B}, μ, T) , where (X, \mathcal{B}, μ) is a probability space and $T : X \rightarrow X$ is a measure preserving transformation. In general, we omit the σ -algebra \mathcal{B} from the notation and write (X, μ, T) .

Throughout, we make use both of the vocabulary of topological dynamics and of ergodic theory, leading to possible confusion. In general, it is clear from the context whether we are referring to a measure preserving system or a topological system, and so we just refer to either as a *system*. Topological factor maps were already defined. We recall that an *ergodic theoretic factor map* between the measure preserving systems (X, μ, T) and (X', μ', T') is a measurable map $\pi : X \rightarrow X'$ (defined almost everywhere), mapping the measure μ to μ' and commuting with the transformations (almost everywhere). If the map π is invertible (almost everywhere), we say that the two systems are *isomorphic*.

Inverse limits of nilsystems in the topological sense were discussed in Section 2.3. We make this notion precise in the measure theoretic sense, in this case also we consider only sequential inverse limits. A d -step nilsystem (X, T) , endowed with its Haar measure μ , is ergodic if and only if (X, T) is a minimal topological system; in this case, μ is its unique invariant measure. Therefore, every inverse limit (in the topological sense) of d -step minimal nilsystems is uniquely ergodic.

Now, let $(X, \mu, T) = \varprojlim (X_j, \mu_j, T_j)$ be an inverse limit in the ergodic theoretic sense of a sequence of d -step ergodic nilsystems. Recall that each nilsystem (X_j, T_j) is endowed with its Borel σ -algebra and μ_j its Haar measure. This means that for every $j \in \mathbb{N}$, there exist ergodic theoretic factor maps $\pi_j : (X_{j+1}, \mu_{j+1}, T_{j+1}) \rightarrow (X_j, \mu_j, T_j)$ and $p_j : (X, \mu, T) \rightarrow (X_j, \mu_j, T_j)$ satisfying $\pi_j \circ p_{j+1} = p_j$ for every j such that the Borel σ -algebra \mathcal{B} of X is spanned by the σ -algebras $p_j^{-1}(\mathcal{B}_j)$, where \mathcal{B}_j denotes the Borel σ -algebra of X_j .

Every ergodic theoretic factor map between ergodic nilsystems is equal almost everywhere to a topological factor map. A short proof of this fact is given in Appendix A (Theorem A.1). Therefore, the factor maps π_j in the definition of an inverse limit (in the ergodic sense) can be assumed to be topological factor maps. It follows that (X, μ, T) can be identified with the topological inverse limit.

This allows us, in the sequel, to not distinguish between the notions of topological and ergodic theoretic inverse limits of d -step ergodic nilsystems.

5.2. Ergodic uniformity seminorms and nilsystems

Let (X, μ, T) be an ergodic system. For points in $X^{[d]}$ and transformations of these spaces we use the same notation as in the topological setting. In Section 3 of [10], a measure $\mu^{[d]}$ on $X^{[d]}$ and a seminorm $\| \cdot \|_d$ on $L^\infty(\mu)$ are constructed.

We recall the properties of these objects:

Proposition 5.1. *Assume (X, μ, T) is an ergodic system and that $d \geq 1$ is an integer.*

- (1) *The measure $\mu^{[d]}$ is invariant and ergodic under the action of the group $\mathcal{G}^{[d]}$.*
- (2) *Each one-dimensional marginal of $\mu^{[d]}$ is equal to μ and each of its two-dimensional marginals (meaning the image under the map $\mathbf{x} \mapsto (x_\epsilon, x_\theta)$ for $\epsilon \neq \theta \subset [d]$) is equal to $\mu \times \mu$.*
- (3) *If $p : (X, \mu, T) \rightarrow (Y, \nu, T)$ is a factor map then, $\nu^{[d]}$ is the image of $\mu^{[d]}$ under the map $p^{[d]} : X^{[d]} \rightarrow Y^{[d]}$.*

For every $f \in L^\infty(\mu)$, the d -th seminorm $\| \| f \|_d$ of f is defined by

$$\| \| f \|_d^{2^d} = \int \prod_{\epsilon \subset [d]} f(x_\epsilon) d\mu^{[d]}(\mathbf{x}). \tag{4}$$

We have that:

Lemma 5.1. *Assume that (X, μ, T) is an ergodic system and let $d \geq 1$ be an integer.*

- (1) *For every $f \in L^\infty(\mu)$, $|\int f d\mu| \leq \| \| f \|_d$.*
- (2) *If $p : (X, \mu, T) \rightarrow (Y, \nu, T)$ is a factor map, then $\| \| f \|_d = \| \| f \circ p \|_d$ for every function $f \in L^\infty(\nu)$.*

We summarize some of the main results of [10]:

Theorem 5.1. *Assume that (X, μ, T) is an ergodic system and that $d \geq 1$ is an integer. The following properties are equivalent:*

- (1) *(X, μ, T) is measure theoretically isomorphic to an inverse limit of $(d - 1)$ -step ergodic nilsystems.*
- (2) *$\| \cdot \|_d$ is a norm on $L^\infty(\mu)$ (equivalently $\| \| f \|_d = 0$ implies that $f = 0$).*
- (3) *There exists a measurable map $J : X_*^{[d]} \rightarrow X$ such that $x_\emptyset = J(x_\epsilon : \emptyset \neq \epsilon \subset [d])$ for $\mu^{[d]}$ -almost every $\mathbf{x} \in X^{[d]}$.*

Using these properties, it follows that any measure theoretic factor of an inverse limit of $(d - 1)$ -step nilsystems is isomorphic in the ergodic theoretic sense to an inverse limit of $(d - 1)$ -step nilsystems.

Theorem 5.2. (See [10, Theorem 1.2].) Assume that (X, μ, T) is an ergodic system, $d \geq 1$ is an integer, and $f_\epsilon \in L^\infty(\mu)$ for $\emptyset \neq \epsilon \subset [d]$. The averages

$$\frac{1}{N^d} \sum_{0 \leq n_1, \dots, n_d < N} \prod_{\substack{\epsilon \subset [d] \\ \epsilon \neq \emptyset}} f_\epsilon(T^{\mathbf{n} \cdot \epsilon} x) \tag{5}$$

converge in $L^2(\mu)$ as $N \rightarrow +\infty$.

Letting F denote the limit of these averages, we have that for every $g \in L^\infty(\mu)$,

$$\int g(x)F(x) d\mu(x) = \int g(x_\emptyset) \prod_{\substack{\epsilon \subset [d] \\ \epsilon \neq \emptyset}} f_\epsilon(x_\epsilon) d\mu^{[d]}(\mathbf{x}). \tag{6}$$

Lemma 5.2. Let (X, μ, T) , d , f_ϵ , $\emptyset \neq \epsilon \subset [d]$, and F be as in Theorem 5.2. Then

$$\|F\|_{L^\infty(\mu)} \leq \prod_{\substack{\epsilon \subset [d] \\ \epsilon \neq \emptyset}} \|f_\epsilon\|_{L^{2^d-1}(\mu)}.$$

Proof. Let $g \in L^\infty(\mu)$ and choose a function h with $h^{2^d-1} = g$. By (6) and the Hölder Inequality,

$$\left| \int gF d\mu \right| \leq \left(\prod_{\substack{\epsilon \subset [d] \\ \epsilon \neq \emptyset}} \int |h(x_\emptyset)f_\epsilon(x_\epsilon)|^{2^d-1} d\mu^{[d]}(\mathbf{x}) \right)^{1/2^d-1}.$$

Since each two-dimensional marginal of $\mu^{[d]}$ is equal to $\mu \times \mu$, this can be rewritten as

$$\left(\prod_{\substack{\epsilon \subset [d] \\ \epsilon \neq \emptyset}} \int |h(x)f_\epsilon(y)|^{2^d-1} d\mu(x) d\mu(y) \right)^{1/2^d-1} = \|g\|_{L^1(\mu)} \prod_{\substack{\epsilon \subset [d] \\ \epsilon \neq \emptyset}} \|f_\epsilon\|_{L^{2^d-1}(\mu)}$$

and the result follows. \square

5.3. Dual functions

Here again, (X, μ, T) is an ergodic system. Following the notation and terminology of [13], for every $f \in L^\infty(\mu)$, the limit function

$$\lim_{N \rightarrow +\infty} \frac{1}{N^d} \sum_{n_1, \dots, n_d=0}^{N-1} \prod_{\emptyset \neq \epsilon \subset [d]} f(T^{\mathbf{n} \cdot \epsilon} x) \tag{7}$$

is called the *dual function of order d of f* and is written $\mathcal{D}_d f$. It is worth noting that $\mathcal{D}_d f$ is only defined as an element of $L^2(\mu)$, and thus is defined almost everywhere.

By (6) and (4), for every $f \in L^\infty(\mu)$ we have that

$$\int f \mathcal{D}_d f \, d\mu = \|f\|_d^{2^d}. \tag{8}$$

It follows from Lemma 5.2 that:

Lemma 5.3. *If (X, μ, T) is an ergodic system and $d \geq 1$ is an integer, then for every $f \in L^\infty(\mu)$:*

$$\|\mathcal{D}_d f\|_{L^\infty(\mu)} \leq \|f\|_{L^{2^d-1}(\mu)}^{2^d-1}.$$

Moreover, the map \mathcal{D}_d extends to a continuous map from $L^{2^d-1}(\mu)$ to $L^\infty(\mu)$.

We remark that if $0 \leq f \leq g$, then $0 \leq \mathcal{D}_d f \leq \mathcal{D}_d g$.

Lemma 5.4. *If (X, μ, T) is an ergodic system and $d \geq 1$ is an integer, then for every $A \subset X$ we have $\mathcal{D}_d \mathbf{1}_A(x) > 0$ for μ -almost every $x \in A$.*

Proof. Let $B = \{x \in A: \mathcal{D}_d \mathbf{1}_A(x) = 0\}$. By part (5.1) of Lemma 5.1 and (8), since $\mathcal{D}_d \mathbf{1}_B \leq \mathcal{D}_d \mathbf{1}_A$ we have that

$$\mu(B)^{2^d} \leq \|\mathbf{1}_B\|_d^{2^d} = \int_B \mathcal{D}_d \mathbf{1}_B(x) \, d\mu(x) \leq \int_B \mathcal{D}_d \mathbf{1}_A(x) \, d\mu(x) = 0.$$

Thus $\mu(B) = 0$. \square

Using the definition (7) of the dual function, we immediately deduce:

Lemma 5.5. *Let $p: (X, \mu, T) \rightarrow (X', \mu', T')$ be a measure theoretic factor map. For every $f \in L^\infty(\mu')$ we have $(\mathcal{D}_d f) \circ p = \mathcal{D}_d(f \circ p)$.*

Using Theorem 5.2, we deduce that:

Lemma 5.6. *Let (X, T) be a minimal topological dynamical system and μ be an invariant ergodic measure on X . Then the measure $\mu^{[d]}$ is concentrated on the subset $\mathbf{Q}^{[d]}$ of $X^{[d]}$.*

Proof. By Theorem 5.2, the measure $\mu^{[d]}$ is a weak limit of averages of Dirac masses at points of the form $(T^{\mathbf{n}-\epsilon} x: \epsilon \in [d])$ for $\mathbf{n} \in \mathbb{Z}^d$ and $x \in X$. Since all of these points belong to $\mathbf{Q}^{[d]}$, the measure $\mu^{[d]}$ is concentrated on this set. \square

Lemma 5.7. *Let (X, T) be a minimal system of order $d - 1$ and let μ be an invariant ergodic measure on X . Let d_X denote a distance on X defining the topology of this space and for every $x \in X$ and $r > 0$, let $B(x, r)$ denote the ball centered at x of radius r with respect to the distance d_X . Then for every $\eta > 0$, there exists $\delta > 0$ such that for every $x \in X$, $\mathcal{D}_d \mathbf{1}_{B(x, \delta)} = 0$ μ -almost everywhere on the complement of $B(x, \eta)$.*

Proof. By definition of a system of order $d - 1$, the last coordinate of an element of $\mathbf{Q}^{[d]}$ is a function of the other ones. Using the symmetries of $\mathbf{Q}^{[d]}$, we have that the same property holds with the first coordinate substituted for the last one. Therefore, writing $\mathbf{Q}_*^{[d]}$ for $\mathbf{Q}^{[d]}$ without the first coordinate, there exists a map $J : \mathbf{Q}_*^{[d]} \rightarrow X$ such that for every $\mathbf{x} \in \mathbf{Q}^{[d]}$,

$$x_\emptyset = J(x_\epsilon : \epsilon \in [d], \epsilon \neq \emptyset).$$

The graph of this map is the closed subset $\mathbf{Q}^{[d]}$ of $\mathbf{Q}_*^{[d]} \times X$ and thus is continuous.

Fix $\eta > 0$. Since J is uniformly continuous and satisfies $J(x, \dots, x) = x$ for every x , there exists $\delta > 0$ such that for every $x \in X$, the set

$$(X \setminus B(x, \eta)) \times B(x, \delta) \times \dots \times B(x, \delta)$$

has empty intersection with $\mathbf{Q}^{[d]}$. Thus by Lemma 5.6 it has zero $\mu^{[d]}$ -measure. By Theorem 5.2 and the definition of $\mathcal{D}_d \mathbf{1}_B$, we have that

$$\int \mathbf{1}_{X \setminus B(x, \eta)} \mathcal{D}_d \mathbf{1}_{B(x, \delta)} d\mu = 0. \quad \square$$

5.4. Systems with continuous dual functions

It is convenient to give a name to the following, although we only make use of it within proofs:

Definition 5.1. Let (X, T) be a minimal system and let μ an ergodic invariant measure on X . We say that (X, T, μ) has property $\mathcal{P}(d)$ if whenever $f_\epsilon, \emptyset \neq \epsilon \in [d]$, are continuous functions on X , the averages (5) converge everywhere and uniformly.

If this property holds, then in particular, for every continuous function f on X , the averages (7) converge everywhere and uniformly for every continuous function f on X . The limit of these averages coincides almost everywhere with the function $\mathcal{D}_d f$ defined above and so we also denote it by $\mathcal{D}_d f$.

Proposition 5.2. Let (X, T) be an inverse limit of minimal $(d - 1)$ -step nilsystems and let μ be the invariant measure of this system. Then (X, μ, T) has property $\mathcal{P}(d)$.

Proof. Assume first that (X, μ, T) is a $(d - 1)$ -step ergodic nilsystem. In [13] (Corollary 5.2), the convergence of the averages (5) is shown to hold everywhere and this convergence is uniform when the functions $f_\epsilon, \emptyset \neq \epsilon \in [d]$, are continuous.

Assume now that (X, μ, T) is as in the statement. Every continuous function on X can be approximated uniformly by a continuous function arising from one of the nilsystems which are factors of X . By density, the result also holds in this case. \square

We now establish some properties of systems with property $\mathcal{P}(d)$. We write $\mathcal{C}(X)$ for the algebra of continuous functions on X . We always assume that $\mathcal{C}(X)$ is endowed with the norm of uniform convergence.

By Lemma 5.3 and density:

Lemma 5.8. Assume that the ergodic system (X, μ, T) has property $\mathcal{P}(d)$.

- For every $f \in L^{2^d-1}(\mu)$, the function $\mathcal{D}_d f$ is equal μ -almost everywhere to a continuous function on X , which we also denote by $\mathcal{D}_d f$, called the dual function of f .
- The map $f \mapsto \mathcal{D}_d f$ is continuous from $L^{2^d-1}(\mu)$ to $\mathcal{C}(X)$.

Lemma 5.9. Let (X, μ, T) be a system with property $\mathcal{P}(d)$, (Y, T) be a minimal system, $p: X \rightarrow Y$ a topological factor map, and ν be the image of μ under p . Then (Y, T, ν) has property $\mathcal{P}(d)$.

Proof. Let $f_\epsilon, \emptyset \neq \epsilon \subset [d]$, be continuous functions on Y . Then the averages

$$\frac{1}{N^d} \sum_{0 \leq n_1, \dots, n_d < N} \prod_{\substack{\epsilon \subset [d] \\ \epsilon \neq \emptyset}} f_\epsilon(T^{n_\epsilon} p(x))$$

converge uniformly on X and thus the averages (5) converge uniformly on Y . \square

6. Using a measure

In this section, we prove Theorem 4.1 which completes the proof of Theorem 1.2: any transitive system (X, T) of order $d - 1$ is an inverse limit of $(d - 1)$ -step minimal nilsystems. By Corollary 3.2, (X, T) is distal and thus is minimal. The method we use is completely different from that used in [14] for $d = 3$, and proceeds by introducing an invariant measure on X .

We start by reducing the proof of Theorem 4.1 to the following:

Proposition 6.1. Let (X, T) be a minimal system of order $d - 1$, μ be an invariant ergodic measure on X , and let (Y, T) be an inverse limit of minimal $(d - 1)$ -step nilsystems with Haar measure ν . Let $\Psi: (Y, \nu, T) \rightarrow (X, \mu, T)$ be a measure theoretic isomorphism. Then Ψ coincides ν -almost everywhere with a topological isomorphism.

Proof of Theorem 4.1 (assuming Proposition 6.1). By Lemma 5.6, the measure $\mu^{[d]}$ is concentrated on the subset $\mathbf{Q}^{[d]}$ of $X^{[d]}$. Since (X, T) is a system of order $d - 1$, there exists a continuous map $J: \mathbf{Q}_*^{[d]} \rightarrow X$ such that

$$x_\emptyset = J(x_\epsilon: \epsilon \subset [d], \epsilon \neq \emptyset) \quad \text{for every } \mathbf{x} \in \mathbf{Q}^{[d]}$$

and so this property holds $\mu^{[d]}$ -almost everywhere. (Again, $\mathbf{Q}_*^{[d]}$ denotes $\mathbf{Q}^{[d]}$ without the first coordinate.) By Theorem 5.1, (X, μ, T) is isomorphic in the ergodic theoretic sense to an inverse limit (Y, ν, T) of $(d - 1)$ -step ergodic nilsystems. By Proposition 6.1, (X, T) and (Y, S) are isomorphic in the topological sense and we are finished. \square

6.1. Proof of Proposition 6.1

To prove Proposition 6.1, we start with a lemma:

Lemma 6.1. *Let (Y, ν, T) be a system with Property $\mathcal{P}(d)$, (X, T) be a minimal system of order $d - 1$, μ be an invariant probability measure on X , and $\Psi : (Y, \nu, T) \rightarrow (X, \mu, T)$ be a measure theoretic factor map. Then Ψ agrees ν -almost everywhere with some topological factor map.*

Proof. We can assume that there exists a Borel invariant subset Y_0 of full measure and that Ψ is a Borel map from Y_0 to X , mapping the measure ν to the measure μ and such that $\Psi(Tx) = T\Psi(x)$ for every $x \in Y_0$.

We claim that:

Claim 6.1. For every open subset U of X , there exists an open subset \tilde{U} of Y equal to $\Psi^{-1}(U)$ up to a ν -negligible set.

To see this, if (X, μ) is a probability space and $A, B \subset X$, write $A \subset_{\mu} B$ if $\mu(A \setminus B) = 0$. The notations $A \supset_{\mu} B$ and $A =_{\mu} B$ are defined similarly.

Assume that $U \neq \emptyset$, as otherwise the claim holds trivially. Let $x \in U$. By Lemma 5.7 there exists an open subset W_x containing x and included in U such that the set

$$U_x := \{x \in X : \mathcal{D}_d \mathbf{1}_{W_x} > 0\}$$

satisfies

$$U_x \subset_{\mu} U. \tag{9}$$

Define

$$\tilde{U}_x = \{y \in Y : \mathcal{D}_d(\mathbf{1}_{W_x} \circ \Psi)(y) > 0\}.$$

By Lemmas 5.8 and 5.5, \tilde{U}_x is an open subset of Y and

$$\tilde{U}_x =_{\nu} \Psi^{-1}(U_x).$$

We have that U is the union of the open sets W_x for $x \in U$. Since U is σ -compact, there exists a countable subset Γ of U such that the union $\bigcup_{x \in \Gamma} U_x$ is equal to U . Define

$$\tilde{U} = \bigcup_{x \in \Gamma} \tilde{U}_x.$$

Then

$$\tilde{U} =_{\nu} \Psi^{-1}\left(\bigcup_{x \in \Gamma} U_x\right).$$

By (9), $\tilde{U} \subset_{\nu} \Psi^{-1}(U)$. By Lemma 5.4, for every $x \in \Gamma$ we have that $U_x \supset_{\mu} W_x$. Thus $\tilde{U}_x \supset_{\nu} \Psi^{-1}(W_x)$ and

$$\tilde{U} \supseteq \bigcup_{x \in \Gamma} \Psi^{-1}(W_x) = \Psi^{-1}\left(\bigcup_{x \in \Gamma} W_x\right) = \Psi^{-1}(U).$$

This completes the proof of the claim.

Claim 6.2. There exists an invariant subset Y_1 of full measure such that the restriction of Ψ to Y_1 (endowed with the induced topology) is continuous.

To prove this claim, we let $(U_j: j \geq 1)$ be a countable basis for the topology of X .

For every $j \geq 1$, by Claim 6.1 there exists an open subset \tilde{U}_j of Y such that the symmetric difference

$$Z_j := \tilde{U}_j \Delta \Psi^{-1}(U_j)$$

has zero ν -measure. Define

$$Y_1 = Y_0 \setminus \bigcup_{n \in \mathbb{Z}} \bigcup_{j \geq 1} T^n Z_j.$$

(Recall that Y_0 is the invariant subset of Y where the map Φ is defined.)

For every $j \geq 1$, $\Psi^{-1}(U_j) \cap Y_1 = \tilde{U}_j \cap Y_1$. Every nonempty open subset U of X is the union of some of the sets U_j , and if \tilde{U} is the union of the corresponding sets \tilde{U}_j we have that $\Psi^{-1}(U) \cap Y_1 = \tilde{U} \cap Y_1$. This proves the claim.

We combine these results to complete the proof of Lemma 6.1. Since (Y, T) is minimal, the measure ν has full support in Y and the subset Y_1 given by Claim 6.2 is dense in Y . Since (X, T) is distal, the result now follows from Lemma 2.1. \square

Using this, we return to the proposition:

Proof of Proposition 6.1. There exist a Borel invariant subset Y_0 of Y of full measure, a Borel invariant subset X_0 of full measure, and a Borel bijection $\Psi : Y_0 \rightarrow X_0$ with Borel inverse, mapping ν to μ and commuting with the transformations.

Recall that (X, T) is a system of order $d - 1$ and that (Y, ν, S) satisfies property $\mathcal{P}(d)$. By Lemma 6.1, there exist a subset Y_1 of Y_0 of full measure and a topological factor map $\Phi : Y \rightarrow X$ that coincides with Ψ on Y_1 .

By Lemma 5.9, (X, μ, T) has property $\mathcal{P}(d)$. Recall that (Y, T) is a system of order $d - 1$. Using Lemma 6.1 again, there exist a subset X_1 of X_0 of full measure and a topological factor map $\Theta : X \rightarrow Y$ that coincides with Ψ^{-1} on X_1 .

The subset $Y_1 \cap \Psi^{-1}(X_1)$ has full measure in Y and for y in this set, we have $\Theta \circ \Phi(y) = y$. Since the measure ν has full support, this equality holds everywhere and $\Theta \circ \Phi = \text{Id}_Y$. By the same argument, $\Phi \circ \Theta = \text{Id}_X$ and we are done. \square

Acknowledgment

We thank Eli Glasner for helpful conversations during the preparation of this article.

Appendix A. Rigidity properties of inverse limits of nilsystems

In this section, we assume that $d > 1$ is an integer and establish some “rigidity” properties of inverse limits of $(d - 1)$ -step nilsystems, meaning some continuity properties.

A property of nilsystems of this type (Theorem A.1) was used in Section 5.1 in the discussion on the definition of inverse limits, and so the reader may be concerned about a possible vicious circle in the argument. The way to avoid this is to first carry out the results in this section for nilsystems, and not inverse limits of nilsystems. This suffices to establish the property needed in Section 5.1. Then it is easy to check that the same proofs extend to the general case.

Throughout the remainder of this section, we assume that (X, T) is an inverse limit of minimal $(d - 1)$ -step nilsystems and that μ is the invariant measure of this system. We recall that (X, T) is a system of order $d - 1$ and has property $\mathcal{P}(d)$ of continuity of dual functions (Proposition 5.2). We first give a slight improvement of Lemma 5.7, maintaining the same notation:

Lemma A.1. *For every $x \in X$ and every neighborhood U of X , there exists a neighborhood V of x such that if f is a continuous function on X whose support lies in V , then the support of the function $\mathcal{D}_d f$ is contained in U .*

Proof. Pick $\eta > 0$ such that the ball $B(x, 2\eta)$ is contained in U . Let δ be as in Lemma 5.7 and let $V = B(x, \delta)$.

Assume that $f \in \mathcal{C}(X)$ has support contained in V and assume that $|f| \leq 1$. We have that $|\mathcal{D}_d f| \leq \mathcal{D}_d |f| \leq \mathcal{D}_d \mathbf{1}_{B(x, \delta)}$. By the choice of δ , $\mathcal{D}_d f$ is equal to zero almost everywhere on the complement of $B(x, \eta)$.

Since the function $\mathcal{D}_d f$ is continuous and since the measure μ has full support in X , $\mathcal{D}_d f$ vanishes everywhere outside the closed ball $\bar{B}(x, \eta)$, which is included in U . \square

Lemma A.2. *If f is a nonnegative continuous function on X , then $\mathcal{D}_d f(x) > 0$ for every $x \in X$ such that $f(x) > 0$.*

Proof. It follows immediately from property $\mathcal{P}(d)$ that for every $x \in X$, there exists a probability measure $\mu_x^{[d]}$ on $X_*^{[d]}$ such that

$$\frac{1}{N^d} \sum_{0 \leq n_1, \dots, n_d < N} \prod_{\substack{\epsilon \subset [d] \\ \epsilon \neq \emptyset}} f_\epsilon(T^{n_\epsilon} x) \rightarrow \int \prod_{\substack{\epsilon \subset [d] \\ \epsilon \neq \emptyset}} f_\epsilon(y_\epsilon) d\mu_x^{[d]}(\mathbf{y}_*)$$

as $N \rightarrow +\infty$ for any continuous functions $f_\epsilon, \emptyset \neq \epsilon \subset [d]$, on X .

By construction, the measure $\delta_x \times \mu_x^{[d]}$ is concentrated on the closed orbit $\mathbf{K}^{[d]}(x)$ of the point $x^{[d]} \in X^{[d]}$ under the group of face transformations $\mathcal{F}^{[d]}$, and is invariant under these transformations. Since (X, T) is distal, the action of these transformations on $\mathbf{K}^{[d]}(x)$ is minimal and thus the topological support of the measure $\delta_x \times \mu_x^{[d]}$ is equal to $\mathbf{K}^{[d]}(x)$. Therefore, the point $x_*^{[d]} \in X_*^{[d]}$ belongs to the topological support of the measure $\mu_x^{[d]}$.

If f is a nonnegative continuous function on X with $f(x) > 0$ then,

$$\mathcal{D}_d f(x) = \int \prod_{\substack{\epsilon \subset [d] \\ \epsilon \neq \emptyset}} f(y_\epsilon) d\mu_x^{[d]}(\mathbf{y}_*) > 0,$$

because the function in the integral is positive at the point $x_*^{[d]}$ which belongs to the support of the measure $\mu_x^{[d]}$. \square

Lemma A.3. *The algebra of functions spanned by $\{\mathcal{D}_d f: f \in \mathcal{C}(X)\}$ is dense in $\mathcal{C}(X)$ under the uniform norm.*

Proof. By Lemmas A.1 and A.2, for distinct $x, y \in X$, there exists a continuous function f on X with $\mathcal{D}_d f(x) \neq \mathcal{D}_d f(y)$. Recall that $\mathcal{D}_d f$ is a continuous function on X . Noting that $\mathcal{D}_d 1 = 1$, the statement follows from the Stone–Weierstrass Theorem. \square

Theorem A.1. *Let $p: (X, \mu, T) \rightarrow (X', \mu', T')$ be a measure theoretic factor map between inverse limits of $(d - 1)$ -step ergodic nilsystems. Then the factor map $p: X \rightarrow X'$ is equal almost everywhere to a topological factor map.*

Proof. Let \mathcal{A} be a countable subset of $\mathcal{C}(X')$ that is dense under the uniform norm. By Lemmas 5.8 and A.3, $\{\mathcal{D}_d f: f \in \mathcal{A}\}$ is included in $\mathcal{C}(X')$ and is dense in this algebra.

By Lemma 5.5, for every $f \in \mathcal{A}$ we have that $\mathcal{D}_d f \circ p = \mathcal{D}_d(f \circ p)$ almost everywhere. By Lemma 5.8, $\mathcal{D}_d(f \circ p)$ is μ -almost everywhere equal to a continuous function on X . Therefore, there exists $X_0 \subset X$ of full measure such that for every $f \in \mathcal{A}$, the function $(\mathcal{D}_d f) \circ p$ coincides on X_0 with a continuous function on X . The same property holds for every function belonging to the algebra spanned by \mathcal{A} . Since X_0 is dense in X , by density the same property holds for every continuous function on X .

This defines a homomorphism of algebras $\kappa: \mathcal{C}(X') \rightarrow \mathcal{C}(X)$ with $\kappa f(x) = f(p(x))$ for every $x \in X_0$ and every $f \in \mathcal{C}(X')$, and κ commutes with the transformations T and T' . Thus there exists a continuous map $p': X \rightarrow X'$ such that $\kappa f = f \circ p'$ for all $f \in \mathcal{C}(X')$. \square

Theorem A.2. *Let (X, T, μ) be an ergodic inverse limit of $(d - 1)$ -step nilsystems, G be a Polish group, and $(g, x) \mapsto g \cdot x$ be a Borel action of G on X by measure preserving transformations commuting with T . There exists a continuous action $(g, x) \mapsto g * x$ of G on X , commuting with T , such that for every $g \in G$, $g * x = g \cdot x$ for μ -almost every $x \in X$.*

By hypothesis, the map $(g, x) \mapsto g \cdot x$ is Borel from $G \times X$ to X . The action of G on X we want must be such that the map $(g, x) \mapsto g * x$ is continuous from $G \times X$ to X .

Proof. By Theorem A.1, for every $g \in G$ there exists a continuous map $x \mapsto g * x$, commuting with T and preserving the measure μ , such that $g * x = g \cdot x$ for μ -almost every $x \in X$. For $g, h \in G$, we have that for μ -almost every $x \in X$, $g * (h \cdot x) = gh * x$. By density, the same equality holds everywhere. Therefore, the map $(g, x) \mapsto g * x$ is an action of G on X . We are left with showing that this map is jointly continuous.

Let $f \in \mathcal{C}(X)$. For $g \in G$, write $f_g(x) = f(g * x)$. For each $g \in G$, the function f_g is continuous and the map $x \mapsto g * x$ commutes with T . By Proposition 5.2, $\mathcal{D}_d f_g(x) = \mathcal{D}_d f(g * x)$ for every $x \in X$.

For each $g \in G$, the functions f_g and $x \mapsto g \cdot x$ are equal almost everywhere and represent the same element of $L^{2^d-1}(\mu)$. Since the action $(g, x) \mapsto g \cdot x$ of G on X is Borel and measure preserving, by [3] we have that the map $g \mapsto f_g$ is continuous from G to $L^{2^d-1}(\mu)$. By Lemma 5.3, the map $g \mapsto \mathcal{D}_d f_g$ is continuous from G to $\mathcal{C}(X)$, meaning that the function $(g, x) \mapsto \mathcal{D}_d f_g(x) = \mathcal{D}_d f(g * x)$ is continuous on $G \times X$.

By density (Lemma A.3), for every function $h \in \mathcal{C}(X)$, the function $(g, x) \mapsto h(g * x)$ is continuous on $G \times X$. We deduce that the map $(g, x) \mapsto g * x$ is continuous from $G \times X$ to X . \square

References

- [1] J. Auslander, *Minimal Flows and Their Extensions*, North-Holland Math. Stud., vol. 153, North-Holland Publishing Co., Amsterdam, 1988.
- [2] L. Auslander, L. Green, F. Hahn, *Flows on Homogeneous Spaces*, Ann. of Math. Stud., vol. 53, Princeton Univ. Press, 1963.
- [3] H. Becker, A.S. Kechris, *The Descriptive Theory of Polish Group Actions*, London Math. Soc. Ser., vol. 232, Cambridge Univ. Press, 1996.
- [4] V. Bergelson, B. Host, B. Kra, Multiple recurrence and nilsequences, *Invent. Math.* 160 (2005) 261–303, with an appendix by I.Z. Ruzsa.
- [5] N. Frantzikinakis, M. Wierdl, A Hardy field extension of Szemerédi’s theorem, *Adv. Math.* 222 (2009) 1–43.
- [6] E. Glasner, Personal communication.
- [7] B. Green, T. Tao, An inverse theorem for the Gowers $U^3(G)$ norm, *Proc. Edinb. Math. Soc.* 51 (2008) 73–153.
- [8] B. Green, T. Tao, Linear equations in primes, *Ann. of Math.*, in press.
- [9] B. Green, T. Tao, The Möbius function is strongly orthogonal to nilsequences, preprint.
- [10] B. Host, B. Kra, Nonconventional averages and nilmanifolds, *Ann. of Math.* 161 (2005) 398–488.
- [11] B. Host, B. Kra, Parallelepipeds, nilpotent groups, and Gowers norms, *Bull. Soc. Math. France* 136 (2008) 405–437.
- [12] B. Host, B. Kra, Analysis of two step nilsequences, *Ann. Inst. Fourier* 58 (2008) 1407–1453.
- [13] B. Host, B. Kra, Uniformity norms on ℓ^∞ and applications, *J. Anal. Math.* 108 (2009) 219–276.
- [14] B. Host, A. Maass, Nilsystèmes d’ordre deux et parallélépipèdes, *Bull. Soc. Math. France* 135 (2007) 367–405.
- [15] A. Leibman, Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold, *Ergodic Theory Dynam. Systems* 25 (1) (2005) 201–213.
- [16] D. McMahon, Relativized weak disjointness and relatively invariant measures, *Trans. Amer. Math. Soc.* 236 (1978) 225–237.
- [17] W. Parry, Dynamical systems on nilmanifolds, *Bull. Lond. Math. Soc.* 2 (1970) 37–40.