MEAN DIMENSION, SMALL ENTROPY FACTORS AND AN EMBEDDING THEOREM

ELON LINDENSTRAUSS

Abstract. In this paper we show how the notion of mean dimension is connected in a natural way to the following two questions: what points in a dynamical system (X, T) can be distinguished by factors with arbitrarily small topological entropy, and when can a system (X, T) be embedded in $\left(([0,1]^d)^{\mathbb{Z}}, \text{shift}\right)$. Our results apply to extensions of minimal $\mathbb{Z}$-actions, and for this case we also show that there is a very satisfying dimension theory for mean dimension.

1. Introduction

In Lindenstrauss and Weiss (1999), the notion of mean dimension of a dynamical system is developed. It is a new invariant for dynamical systems, suggested by M. Gromov, that can give interesting information on a dynamical system even when the usual invariants of topological entropy and topological dimension are infinite. We denote the mean dimension of a system by $\text{mdim}(X)$. In addition to mean dimension, which behaves rather like topological dimension, in Lindenstrauss and Weiss (1999) we have also defined an analogue of Minkowski dimension $\text{mdim}_{\text{M}}(X, d)$ that is not a topological invariant but depends on the metric (one can turn it into a topological invariant by taking the infimum of this value for all metrics d compatible with the given topology), and an analogue of the definition of zero dimension in the inductive definition of dimension, the Small Boundary Property (SBP). In the general setting of amenable group actions on compact metric spaces B. Weiss and myself have been able to show the following implications:

$$\text{mdim}(X) \leq \text{mdim}_{\text{M}}(X, d) \qquad \text{for all } d\text{, and}$$
$$X \text{ has SBP} \implies \text{mdim}(X) = 0.$$

We also mention that if the topological entropy of X is finite then $\text{mdim}_{\text{M}}(X, d) = 0$ for all d, that if the dimension of X is finite then $\text{mdim}(X) = 0$, and that if X is uniquely ergodic then X has the SBP. Thus the collection of systems with $\text{mdim}(X) = 0$ is rather large and

Date: September 20, 1998.

Key words and phrases. dimension theory, entropy, factors, embeddings.

can potentially unify arguments given for systems with finite dimension, finite entropy or a unique ergodic measure.

In this paper, I use an analogue of the Rokhlin Tower Lemma in measurable dynamics and the Baire Category Theorem for the space of functions from X to a suitable K to complete the dimension theory for mean dimension, and prove that for systems where the Tower Lemma holds,

$$\exists \text{ metric } d: \ \operatorname{mdim}(X) = \operatorname{mdim}_{\mathrm{M}}(X, d) \tag{1.1}$$

$$X \text{ has SBP} \iff \operatorname{mdim}(X) = 0. \tag{1.2}$$

Unfortunately, the proof I give to the Tower Lemma works only for extensions of minimal $\mathbb{Z}$ actions. For these systems, however, equations (1.1)–(1.2) give a very satisfying dimension theory for mean dimension. Recall that a system (X, T) is minimal if X has no T invariant proper closed subsets (we will implicitly assume throughout that the minimal systems we will consider are non-trivial, i.e. infinite).

These techniques also shed light on two problems that have not been completely understood for quite some time. The first of these questions is a very natural question raised in Shub and Weiss (1991) — *When can one lower the topological entropy of a system by taking (continuous) factors?* We recall that a factor of a dynamical system (X, T) is a dynamical system (Y, S) (together with a onto map $\Phi\colon X \to Y$) such that the following diagram commutes

$$\begin{array}{ccc} X & \xrightarrow{T} & X \\ \big\downarrow{\scriptstyle \Phi} & & \big\downarrow{\scriptstyle \Phi} \\ Y & \xrightarrow{S} & Y \end{array}$$

It seems that the correct question to ask is for what pairs of points $x, y \in X$ can one find (for every $\epsilon > 0$) factor mappings Φ_ϵ into a system with topological entropy less than ϵ such that $\Phi_\epsilon(x) \neq \Phi_\epsilon(y)$. If all (nontrivial) pairs of points in X can be distinguished then, as shown in Lindenstrauss (1995), for any factor (Y, S) of (X, T) and any $\eta \in [h_{\text{top}}(Y), h_{\text{top}}(X)]$ one can find an intermediate system (Z, R) such that (Z, R) is a factor of (X, T), $h_{\text{top}}(Z) = \eta$, and such that the factor map $X \longrightarrow Y$ factors through Z. The existence of such factors is proved in Shub and Weiss (1991) for uniquely ergodic systems, and in Lindenstrauss (1995) for finite dimensional systems. In Lindenstrauss (1995) it is also shown that for some systems, no two points can be distinguished by finite entropy factors — or, in other words, these systems have no finite entropy factors.

The examples given there have positive mean dimension, and, as shown in Lindenstrauss and Weiss (1999), no factors with zero mean dimension. In this paper we show that for extensions of minimal $\mathbb{Z}$ actions every two points can be distinguished by low entropy factors iff the system has zero mean dimension. Moreover, there is a unique factor of X, the maximal zero mean dimensional factor, such that x and y can be distinguished by low (or finite) entropy factors iff they map into different points in the maximal zero mean dimensional factor.

Another question that has been considered by previous authors and on which we can, using mean dimension theory, considerably clarify the picture is *when can a dynamical system be embedded in* $([0,1]^{\mathbb{Z}}, \textit{shift})$. The initial motivation for this result is Beboutov's theorem that every real flow (X, T_t) whose fixed point set can be embedded in $\mathbb{R}$ can be embedded in the space of continuous functions on $\mathbb{R}$, with the natural action of $\mathbb{R}$ (see Kakutani (1968)). A dynamical system can have a trivial obstruction to being embeddable in $([0,1]^{\mathbb{Z}}, \mathrm{shift})$ if it has too many periodic points. But, for example, it was not clear for a long time if every minimal dynamical system is embeddable in this system. Jaworski proved that if (X, T) is finite dimensional and has no periodic points then it is embeddable in $([0,1]^{\mathbb{Z}}, \mathrm{shift})$ (Jaworski (1974); a more accessible source is Auslander (1988), Chapter 13, pp 183–194). In Lindenstrauss and Weiss (1999) it is shown that a necessary condition for (X, T) to be embeddable in $\left(([0,1]^d)^{\mathbb{Z}}, \mathrm{shift}\right)$ is that $\mathrm{mdim}(X) \leq d$ — and so there are many minimal systems that are not embeddable in $\left(([0,1]^d)^{\mathbb{Z}}, \mathrm{shift}\right)$. The different behavior for $\mathbb{R}$ actions from that of $\mathbb{Z}$ actions is not too surprising considering the fact that $[0,1]^{\mathbb{Z}}$ is a compact space, whereas the space of continuous functions from $\mathbb{R}$ to $[0,1]$ is huge.

In this paper, we give a partial converse of the necessary condition that $\mathrm{mdim}(X) < d$. We show (for extensions of minimal $\mathbb{Z}$ actions) that if $\mathrm{mdim}(X) < cd$ for some $c < 1$ then X can be embedded in $\left(([0,1]^d)^{\mathbb{Z}}, \mathrm{shift}\right)$. In particular we get two new results that do not involve at all the notion of mean dimension: any uniquely ergodic (or more precisely strictly ergodic) system and any minimal system with finite entropy can be embedded in $([0,1]^Z, \mathrm{shift})$.

These results are (hopefully) only part of some larger picture. It would be interesting to extend these results to more general $\mathbb{Z}$-actions, and to more general groups. Even extending the results to $\mathbb{Z}^2$ seems to require new ideas. But I would like to remark that the obstruction to extending these results is not purely technical. Especially troublesome seem to be the periodic points of X. Indeed, if the set of periodic points of X is not zero dimensional then X does not have the SBP, and as we have seen periodic points do indeed obstruct embedding X into $\left(([0,1]^d)^{\mathbb{Z}}, \mathrm{shift}\right)$. The problem of handling the case where there are many periodic points has been successfully handled in Lindenstrauss (1995) for the special case of finite dimensional systems, where it is shown that for theses systems every two points can be distinguished by low entropy factors regardless of the dimension of the periodic points.

Another nice question that remains open is what is the largest constant c such that $\mathrm{mdim}(X) < cd$ implies that X can be embedded in $([0,1]^{d\mathbb{Z}}, \mathrm{shift})$. The bound we get is that $c \geq 1/36$.

Overview: In the next section, §2, we review the necessary definitions and results we need from Lindenstrauss and Weiss (1999). In §3 we prove the Tower Lemma for extensions of minimal systems. In §4 we prove the existence of a metric d such that $\mathrm{mdim}_{\mathrm{M}}(X,d) = \mathrm{mdim}(X)$, using a Baire Category argument.

We prove the embedding theorem in §5. The proof of the embedding result is more difficult than the proofs of the main results of §4 and §6, and seems harder to generalize to more general setups.

The main result in §6 is that every system with $\mathrm{mdim}(X) = 0$ has the SBP. We also describe in this section the implications of the SBP on small entropy factors, and the existence of a maximal mean dimension zero factor. We conclude §6 with a surprising corollary that exhibits a sharp dichotomy between systems with $\mathrm{mdim}(X) = 0$ and systems with $\mathrm{mdim}(X) > 0$.

§2, §3 and the beginning of §4 (up to Lemma 4.4) contain ideas and definitions that are used throughout. The remainder of §4, §5 and §6 can be read independently of each other.

Acknowledgments: This work is part of the author's PhD thesis, conducted under the guidance of Prof. Benjamin Weiss of the Hebrew University. I would like to thank him for many helpful discussions and insights, and for his constant encouragement. I would also like to thank Prof. Mikhael Gromov who first suggested the notion of mean dimension, with whom I had a very helpful discussion on this and related subjects.

2. Preliminaries

We consider a compact metric space X, and an invertible map $T\colon X \mapsto X$. Like topological dimension, the mean dimension will be defined using open covers of X, and since X is compact, all covers are supposed to be finite. We will say that a cover β refines α $(\beta \succ \alpha)$, if every member of β is a subset of some member of α. We also define the order of a cover α by

$$\mathrm{ord}(\alpha) = \left(\max_{x \in X} \sum_{U \in \alpha} 1_U(x) \right) - 1$$

and define $\mathcal{D}(\alpha) = \inf_{\beta\succ\alpha} \operatorname{ord}(\beta)$. Recall that the topological dimension of a space X is $\le D$ iff every open cover α can be refined by a cover β with $\operatorname{ord}(\beta) \le D$, i.e. if and only if $\mathcal{D}(\alpha) \le D$ for all α.

We state without proof a few facts about open covers and $\mathcal{D}(\alpha)$ (for proofs see Lindenstrauss and Weiss (1999)):

Definition 2.1. *A continuous map $f\colon X \to Y$ will be called α-compatible if it is possible to find a finite open cover of Y, β, such that $f^{-1}(\beta) \succ \alpha$. We will use the notation $f \succ \alpha$ to denote that f is α compatible.*

If X is compact, to see that a continuous $f\colon X \to Y$ is α compatible it is enough to check that for every $y \in Y$, $f^{-1}(y)$ is a subset of some $U \in \alpha$.

Proposition 2.2. *If α is an open cover of X, then*

$$\mathcal{D}(\alpha) \le k$$

iff there is an α-compatible continuous function $f\colon X \to K$ where K has topological dimension k.

Proposition 2.3. *Let α, β be open covers of X, and set $\alpha \vee \beta = \{U \cap V : U \in \alpha,\ V \in \beta\}$. Then*

$$\mathcal{D}(\alpha \vee \beta) \le \mathcal{D}(\alpha) + \mathcal{D}(\beta).$$

We use the notation $\alpha_a^b = \bigvee_{i=a}^b T^{-i}\alpha$.

Definition 2.4. *If (X,T) is a dynamical system, then the mean dimension of (X,T), denoted by* $\operatorname{mdim}(X,T)$ *(or* $\operatorname{mdim}(X)$ *if T is understood), is defined by:*

$$\operatorname{mdim}(X,T) = \sup_{\alpha} \lim_{n\to\infty} \frac{\mathcal{D}(\alpha_0^{n-1})}{n}, \tag{2.1}$$

where α runs over all finite open covers of X.

By sub-additivity of $\mathcal{D}$, the limit in (2.1) exists. We mention some important basic properties of mean dimension:

1. If Y is a T-invariant subset of X then $\operatorname{mdim}(Y,T) \le \operatorname{mdim}(X,T)$. However, If (Y,S) is a factor of (X,T), $\operatorname{mdim}(Y,S)$ might well be bigger than $\operatorname{mdim}(X,T)$.
2. If (X,T) is finite dimensional then $\operatorname{mdim}(X,T) = 0$.
3. If $X = [0,1]^{\mathbb{Z}}$, and $\sigma\colon [0,1]^{\mathbb{Z}} \to [0,1]^{\mathbb{Z}}$ the shift transformation then $\operatorname{mdim}([0,1]^{\mathbb{Z}},\sigma) = 1$. More generally, $\operatorname{mdim}(([0,1]^d)^{\mathbb{Z}},\sigma) = d$. All proper factors of either of these systems have strictly positive mean dimension.

4. For any dynamical system (X,T), $\operatorname{mdim}(X,T^n) = n \operatorname{mdim}(X,T)$.
5. If (X_i,T_i), is a sequence of dynamical systems, $1 \le i < I$ with $I \le \infty$, then

$$\operatorname{mdim}(X_1 \times X_2 \times \cdots, T_1 \times T_2 \times \cdots) \le \sum_{i<I} \operatorname{mdim}(X_i,T_i). \tag{2.2}$$

Throughout this paper, we will use the notation $\mathbf{n}$ to denote the set $\{0,\dots,n-1\}$, for any $n \in \mathbb{N}$. If $n \in N$ and $a \in \mathbb{R}$ (or $\mathbb{Z}$) we take $a \bmod n$ to be the unique $0 \le r < n$ such that $a - r \in n\mathbb{Z}$. Finally, if $A, B \subset \mathbb{Z}$, we let $A + B$ denote the set $\{a + b \colon a \in A, b \in B\}$.

3. A Rokhlin-type lemma for some dynamical systems

The Rokhlin Tower Lemma in ergodic theory is the following theorem:

Theorem 3.1. *If $(X,\mathcal{B},\mu,T)$ is a measure preserving system, μ a probability measure, and if in addition the measure of the set of periodic points is zero, then for every N and $\delta > 0$ there is a B such that the sets in the collection $\{T^{-i}B\}_{i=0}^{N-1}$ are disjoint and*

$$\mu(\bigcup_{i=0}^{N-1} T^{-i}B) \ge 1 - \delta$$

An alternative way to phrase this is that there is, for every N and $\delta > 0$, a measurable function $n\colon X \to \{0,\dots,N-1\}$ such that $n(Tx) = n(x) + 1 \bmod N$, except for a set of measure at most δ. This relatively simple lemma is a very powerful tool in ergodic theory. One of the typical uses of this lemma is for constructing partitions of X with various properties by reducing this to a question of partitioning orbits $\{x, Tx, \dots, T^{N-1}x\}$ of points in X.

There is a natural notion to replace the condition that a set has measure less than δ for dynamical systems:

Definition 3.2. *Let (X,T) be a dynamical system. For a set $C \subset X$ define the* **orbitwise capacity** *of C to be*

$$\operatorname{ocap}(C) = \lim_{n\to\infty} \frac{1}{n} \sup_{x\in X} \sum_{i=0}^{n-1} 1_C(T^i x).$$

If $\operatorname{ocap}(C) = 0$ *we shall say the set C is* **uniformly small**.

We remark that as $\sup_{x\in X} \sum_{i=0}^{n-1} 1_C(T^i x)$ is sub-additive in n, the limit above exists and in fact

$$\operatorname{ocap}(C) = \inf_n \frac{1}{n} \sup_{x\in X} \sum_{i=0}^{n-1} 1_C(T^i x).$$

The definition of uniformly small sets is due to Shub and Weiss (1991). Note that $\operatorname{ocap}(A \cup B) \leq \operatorname{ocap}(A) + \operatorname{ocap}(B)$, and so in particular the union of any finite number of uniformly small sets is uniformly small. We would also like to remark that it is not hard to see that for closed sets C

$$\operatorname{ocap}(C) = \sup_{\mu \in M_{T\ \mathrm{inv}}(X)} \mu(C),$$

where $M_{T\ \mathrm{inv}}(X)$ denotes the collection of T-invariant measures on X.

Lemma 3.3. *Let (X,T) be an extension of a nontrivial (i.e. infinite) minimal system. Then for any N there is a continuous function $n\colon X \to \mathbb{R}$ such that the set $E = \{x \in X : n(Tx) \neq n(x) + 1\}$ satisfies $E \cap T^n E = \emptyset$ for all $n = 1, \ldots, N$.*

Proof. It clearly suffices to prove the lemma for the case X minimal, since if $n\colon X \to \mathbb{R}$ is as in the lemma and Y is an extension of X (with $\phi\colon Y \to X$ a factor map) then $n \circ \phi\colon Y \to \mathbb{R}$ attests the validity of the lemma also for Y.

As (X,T) is minimal we can find open sets U, $U' \subset X$ with $\bar{U} \subset U'$ and such that

$$U' \cap T^{-k} U' = \emptyset \qquad \text{for all } 0 < k \leq N. \tag{3.1}$$

In fact, we can take U' to be any ball of small enough radius. Note that as U is open, (X,T) minimal, there is some M such that

$$\bigcup_{k=0}^{M-1} T^{-k} U = X.$$

Let $w\colon X \to [0,1]$ be a continuous function such that

$$\operatorname{supp} w \subset U' \qquad \text{and} \qquad w|_{\bar{U}} \equiv 1.$$

We will use w to define a markovian random walk on X, which for every starting point x will end after a finite number of steps, as follows: At any point y we will get to during the random walk, we finish the random walk with probability $w(y)$ and move to $T^{-1}y$ with probability $1 - w(y)$.

As the orbit of every point eventually enters the set U on which $w = 1$ this random walk will indeed stop after a finite number of steps.

Let

$$n(x) = E(\#\text{of steps in the random walk starting at } x).$$

Note that if $x \notin U'$ then the random walk starting at x will always move in the next step to $T^{-1}x$, so

$$x \notin U' \qquad \Longrightarrow \qquad n(T^{-1}x) = n(x) - 1$$

Thus

$$E = \{x : n(T(x)) \neq n(x) + 1\} \subset T^{-1}U'$$

And the result follows from (3.1). □

In this form, it does not seem that this Rokhlin-type Lemma can be extended to a more general setup. The following corollary of Lemma 3.3, is what is needed to prove Theorems 4.3 and 6.2, and seems more likely to be true in more general setups.

Corollary 3.4. *Let (X, T) be an extension of a non-trivial minimal system. Then for any δ there is a continuous $n\colon X \to \mathbb{R}$ such that*

$$E = \{x : n(Tx) \neq n(x) + 1 \text{ or } n(x) \notin \mathbb{Z}\}$$

has orbit capacity less than δ.

Proof. Let $\tilde{n}(x)$ be as in Lemma 3.3 for $N > 1/5\delta$. Define

$$g_\epsilon(a) = \begin{cases} \lfloor a \rfloor & \text{if } 0 \le \{a\} < 1 - \epsilon \\ \lfloor a \rfloor + \frac{\{a\} - (1-\epsilon)}{\epsilon} & \text{if } \{a\} \ge 1 - \epsilon \end{cases}$$

g_ϵ is a continuous function $\mathbb{R} \to \mathbb{R}$. Notice that

$$g_\epsilon(a + 1 - \epsilon) = g_\epsilon(a) + 1 \qquad \text{iff } \epsilon \le \{a\} \le 1 - \epsilon, \tag{3.2}$$

and under these conditions $g_\epsilon(a) \in \mathbb{Z}$. Also notice that for any $a \in \mathbb{R}$

$$\sum_{k=0}^{N-1} 1_{[\epsilon, 1-\epsilon]}(\{a + k(1-\epsilon)\}) \ge N - 2\lceil N\epsilon \rceil. \tag{3.3}$$

Now, take $\epsilon < 1/10N$ and define

$$n(x) = g_\epsilon((1-\epsilon)\tilde{n}(x)).$$

For every $x \in X$, we know that

$$\tilde{n}(T^k x) \neq \tilde{n}(T^{k-1}x) + 1 \qquad \text{for at most one } 0 \le k < N.$$

Set k_0 to be this exceptional k if it exists. Then by (3.2) and (3.3) we see that for all $0 \le k < k_0$ except at most two such k's

$$n(T^k x) = n(T^{k-1}x) + 1.$$

and the same is true for all but at most two integers k in the range $k_0 < k < N$. Hence if

$$E = \{x : n(T(x)) \neq n(x) + 1 \text{ or } n(x) \notin \mathbb{Z}\}$$

then

$$\operatorname{ocap}(E) \leq \sup_{x \in X} \frac{\sum_{i=0}^{N-1} 1_E(T^i x)}{N} \leq \frac{5}{N} < \delta.$$

□

We remark that just as in measurable dynamics, the periodic points form an inherent obstruction to Rokhlin type results of the type presented here. In the subsequent sections, we shall call a continuous function $n\colon X \to \mathbb{R}$ such that $n(Tx) \neq n(x) + 1$ only rarely a **level function**, since they can be though of as assigning to every $x \in X$ its position in a Rokhlin-like tower.

4. Connections with the metric mean dimension

In Lindenstrauss and Weiss (1999) we presented another definition of a mean dimension that is metric dependent. We first recall the definition:

For an open cover α, define the mesh of α according to a semi-metric or metric d by

$$\operatorname{mesh}(\alpha, d) = \max_{U \in \alpha} \operatorname{diam}(U).$$

Definition 4.1. *Let X be a dynamic system, $d(\cdot,\cdot)$ a metric on X. Define*

$$d_a^b(x, y) = \max_{a \leq n \leq b} d(T^n x, T^n y).$$

Set

$$S(X, \epsilon, d) = \lim_{n \to \infty} \frac{1}{n} \inf_{\operatorname{mesh}(\alpha, d_0^{n-1}) < \epsilon} \log |\alpha|. \tag{4.1}$$

S is monotone nondecreasing as $\epsilon \to 0$, and we wish to measure just how fast it increases. We define the **metric mean dimension** *of X (for the given metric d), $\operatorname{mdim}_{\mathrm{M}}(X, d)$, as*

$$\operatorname{mdim}_{\mathrm{M}}(X, d) = \underline{\lim}_{\epsilon \to 0} \frac{S(X, \epsilon, d)}{|\log \epsilon|}. \tag{4.2}$$

Notice that

$$h_{\mathrm{top}}(X, T) = \lim_{\epsilon \to 0} S(X, \epsilon, d),$$

so essentially $\operatorname{mdim}_{\mathrm{M}}(X, d)$ measures how fast the terms that approximate the entropy $S(X, \epsilon, d)$ tend to ∞ as $\epsilon \to 0$ (and in particular, $\operatorname{mdim}_{\mathrm{M}}(X, d) = 0$ if $h_{\mathrm{top}}(X) < \infty$ for any metric d).

The limit in (4.1) exists, since if both $\operatorname{mesh}(\alpha, d_n) < \epsilon$ and $\operatorname{mesh}(\beta, d_m) < \epsilon$ then

$$\operatorname{mesh}(\alpha \vee T^{-m}\beta, d_0^{m+n}) < \epsilon \quad \text{and} \quad \log|\alpha \vee T^{-m}\beta| \leq \log|\alpha| + \log|\beta|$$

hence the sequence

$$a_n = \inf_{\operatorname{mesh}(\alpha, d_0^{n-1}) < \epsilon} \log|\alpha|$$

is sub-additive ($a_{n+m} \leq a_n + a_m$), and so

$$S(X, \epsilon, d) \equiv \lim_{n\to\infty} \frac{a_n}{n} = \inf_n \frac{a_n}{n}. \tag{4.3}$$

Now $\operatorname{mdim}_{\mathrm{M}}(X, d)$ is metric dependent. The main result regarding this mean dimension in Lindenstrauss and Weiss (1999) was the following:

Theorem 4.2 (Lindenstrauss and Weiss (1999))**.** *For any metric d on X,*

$$\operatorname{mdim}(X) \leq \operatorname{mdim}_M(X, d)$$

This result is true for all dynamical system, with the acting group being any countable discrete amenable group. Using Corollary 3.4, we can complete the picture for extensions of minimal $\mathbb{Z}$ actions, by showing that the following is true:

Theorem 4.3. *If (X, T) is an extension of a minimal system, then there is a metric d such that*

$$\operatorname{mdim}(X) = \operatorname{mdim}_M(X, d)$$

This theorem has a close analogue in the standard theory of topological dimension. It is well known that for any separable metric space Z, with metric d, the topological dimension of Z and the Hausdorff dimension according to d satisfy

$$\dim(Z) \leq \dim_H(Z, d),$$

(this is the analogue of Theorem 4.2). Furthermore, there is a metric d' on Z such that equality holds, that is

$$\dim(Z) = \dim_H(Z, d') \tag{4.4}$$

(see Hurewicz and Wallman (1941), chapter VII, and especially Theorem VII.5).

Recall how (4.4) is usually proved: For simplicity we assume Z is compact. Consider all continuous maps from Z to a compact convex subset $K \subset \mathbb{R}^M$ with nonempty interior, for M large enough (greater than $2\dim(Z) + 1$), and let $\|\cdot\|$ be some norm on $\mathbb{R}^M$. The usual choice is $K = [0, 1]^M$, but this makes no difference in the proof. Endow this space of functions, $C(Z,K)$, with the uniform convergence topology, i.e. the topology given by the metric

$$d(f,\tilde{f}) = \sup_{z\in Z} \|f(z) - \tilde{f}(z)\|.$$

One shows that for a dense G_δ set of functions $f \in C(Z,K)$, the map f is a homeomorphism of Z onto the subset $f(Z) \subset K$ (where we use on $f(Z)$ the topology induced from that of K). $f(Z)$ also inherits a metric from K — the metric given by the norm $\|\cdot\|$. The proof that there is a metric d' for which $\dim(Z) = \dim_H(Z,d')$ is completed by showing that for a dense G_δ subset of $C(Z,K)$ we have $\dim_H(f(Z),\|\cdot\|) = \dim(Z)$.

We follow the same procedure here. For our uses, it is not enough to map our dynamical system X to a metric space. Since we are considering properties of the dynamical system (X,T), our mappings must preserve the dynamics. However, this is easily attainable as follows: Let K be a compact convex subset of some Banach space (finite or infinite dimensional as needed). We will consider mappings from (X,T) to the dynamical system $(K^{\mathbb{Z}},\sigma)$, where the topology on $K^{\mathbb{Z}}$ is the usual product topology and σ is the two sided shift

$$\sigma\colon (\dots,k_{-1},k_0,k_1,\dots) \mapsto (\dots,k_0,k_1,k_2,\dots)$$

(each k_i is in K). To any map $f \in C(X,K)$ there corresponds a map $I_f\colon X \to K^{\mathbb{Z}}$ that respects the $\mathbb{Z}$ action on these spaces (i.e. $I_f \circ T = \sigma \circ I_f$) as follows

$$I_f\colon x \mapsto (\dots,f(T^{-1}x),f(x),f(Tx),\dots).$$

It is easy to see that conversely, all maps $F\colon X \to K^{\mathbb{Z}}$ that respect the $\mathbb{Z}$ action can be obtained in this way, but we will not need this.

Let D be some standard metric on $K^{\mathbb{Z}}$. What we would like is to prove that for a dense G_δ set of functions $f \in C(X,K)$ the map I_f is an embedding of (X,T) into $(K^{\mathbb{Z}},\sigma)$, and that for a dense G_δ set of $f \in C(X,K)$ we have

$$\operatorname{mdim}_{\mathrm{M}}(I_f(X),D) = \operatorname{mdim}(X).$$

While it is true that for systems with finite mean dimension, if $\dim(K)$ is large enough, then for a dense G_δ set of functions $f \in C(X,K)$ the map I_f is an embedding, the proof of this result is somewhat more elaborate than the proof that for some metric d the metric mean dimension is equal to the mean dimension, and so we defer it to the next section. Instead we will work with infinite dimensional K such as the Hilbert cube, for which the result that for a dense G_δ set of $f \in C(X,K)$, the associated map I_f is an embedding is a triviality, since it is well known that for a dense G_δ set of $f \in C(X,K)$, f itself is an embedding (see for example Hurewicz and Wallman (1941), Theorem V.4). This approach has the additional advantage that we will only need to use the weaker Rokhlin type result, Corollary 3.4, while for the embedding theorem we will need the full force of Lemma 3.3.

We begin with some notations. Let K be compact and convex, inside some Banach space. For any collection $\mathcal{F}$ of vectors from K, $\operatorname{co}(\mathcal{F})$ denotes their convex hull. For $w \in K^{\mathbb{Z}}$ we will use the standard notation w_a^b (or the equivalent $w|_a^b$) to designate the $b-a+1$ coordinates of w at places $a \ldots b$. This notation will also be used for functions from X to $K^{\mathbb{Z}}$. Recall also the related notation

$$\alpha_a^b = \bigvee_{i=a}^{b} T^{-i}\alpha$$

for open covers α of X introduced in Section 2. If $\|\cdot\|$ is a norm on the space containing K, we define a seminorm on $K^{\mathbb{Z}}$ as follows

$$\|x\|_a^b = \max_{a \le n \le b} \|x_n\|.$$

We use these notation also for finite products K^N, where we index the coordinates of K^N starting from 0. Finally, as the standard metric on $K^{\mathbb{Z}}$ we will use the metric $D(\cdot,\cdot)$ defined by

$$D(x,y) = \sum_{i \in Z} 2^{-|i|}\|x_i - y_i\|.$$

The following lemma is useful in that it allows us to estimate $\operatorname{mdim}(I_f(X), D)$ working with seminorm $\|x\|' := \|x\|_0^0$ on $K^{\mathbb{Z}}$, instead of the somewhat awkward metric D.

Lemma 4.4. *Let (Y,σ) be a dynamical system embedded in $(K^{\mathbb{Z}},\sigma)$, K as above, and $\|\cdot\|$ be a norm on the Banach space containing K. Then $S(I_f(X),\epsilon,D) \le S(I_f(X),\epsilon/10,\|\cdot\|')$*

Proof:

For any x, $y \in K^{\mathbb{Z}}$, and integer n, if

$$\|x-y\|_{-\log_2(\operatorname{diam}(K)/\epsilon)}^{n+\log_2(\operatorname{diam}(K)/\epsilon)} < \epsilon$$

then $D_0^{n-1}(x,y) < 10\epsilon$. Hence for any open cover α

$$\operatorname{mesh}(\alpha, D_0^{n-1}) \le 10 \operatorname{mesh}(T^{\log_2(\operatorname{diam}(K)/\epsilon)}\alpha, \|\cdot\|_0^{n+2\log_2(\operatorname{diam}(K)/\epsilon)}).$$

Thus for any ϵ

$$
\begin{aligned}
S(I_f(X),\epsilon,D) &= \lim_{n\to\infty} \inf\left\{\frac{\log|\alpha|}{n} : \alpha \text{ covers } I_f(X) \text{ and } \operatorname{mesh}(\alpha, D_0^{n-1}) < \epsilon\right\}\\
&\le \lim_{n\to\infty} \inf\left\{\frac{\log|\alpha|}{n} : \alpha \text{ covers } I_f(X) \text{ and } \operatorname{mesh}(\alpha, \|\cdot\|_0^{n+2\log_2(\operatorname{diam}(K)/\epsilon)}) < \epsilon/10\right\}\\
&= S(I_f(X),\epsilon/10,\|\cdot\|').
\end{aligned}
$$

□

Lemma 4.5. *Let K be compact and convex, $\mathcal{F} \subset K^N$ finite. Then there is an open cover γ of $\operatorname{co}(\mathcal{F})$ with*

$$
\operatorname{mesh}(\gamma, \|\cdot\|_0^{N-1}) \le \epsilon \qquad \textit{and}
$$
$$
|\gamma| \le C\epsilon^{-|\mathcal{F}|+1},
$$

where C depends only on $\mathcal{F}$.

Proof. Suppose $\mathcal{F} = \{f(0),\dots,f(r)\}$. Cover $\operatorname{co}(\mathcal{F})$ by the $\lceil 4r\operatorname{diam}(K)/\epsilon + 1\rceil^r$ sets

$$
\left\{f(0) + \sum_{i=1}^{r} a(i)(f(i)-f(0)) :\ \forall i \ge 1\ a(i) \in \left(\frac{n(i)\epsilon}{2r\operatorname{diam}(K)}, \frac{(n(i)+2)\epsilon}{2r\operatorname{diam}(K)}\right)\right\}
$$

where each $n(i) \in \mathbb{Z}$ is in the range

$$
-\lceil 2r\operatorname{diam}(K)/\epsilon\rceil - 1 \le n(i) < \lceil 2r\operatorname{diam}(K)/\epsilon\rceil.
$$

An easy calculation shows that these sets have $\|\cdot\|_0^{N-1}$ diameter $\le \epsilon$ □

Lemma 4.6. *Let β be a cover of X with $\operatorname{ord}(\beta) < \Delta$. Suppose we are given for every $U \in \beta$ a point $p_U \in U$ and $v_U \in K^M$. Then is possible to find a continuous function $F\colon X \to K^M$ with the following properties:*

1. *$F(p_U) = v_u$ for all $U \in \beta$,*
2. *For all $x \in X$, $F(x) \in \operatorname{co}(F(x_U) :\ x \in U \in \beta)$.*

In particular, $F(X)$ is contained in a finite union of Δ dimensional polytopes.

Proof:

Let $\{\phi_U(x)\}_{U\in\beta}$ a partition of unity subordinate to β — that is, a collection of continuous functions $X \to [0,1]$ such that

$$
\sum_{U\in\beta} \phi_U(x) = 1 \qquad \text{for all } x \in X
$$

and $\mathrm{supp}(\phi_U) \subset U$, and we can further assume that $\phi_U(p_U) = 1$ for all $U \in \beta$. We now define F to be

$$F(x) = \sum_{U \in \beta} \phi_U(x) v_U.$$

$F(x)$ clearly satisfies conditions of the Lemma. $\square$

Theorem 4.7. *Let K be compact and convex, D the standard metric on $K^{\mathbb{Z}}$. Then for a dense G_δ set of $f \in C(X; K)$,*

$$\mathrm{mdim}_{\mathrm{M}}\left(I_f(X), D\right) \leq \mathrm{mdim}(X).$$

Proof:

Let $A(\epsilon, \eta) \subset C(X; K)$ be the set

$$A(\epsilon, \eta) = \{f : \exists \epsilon_0 < \epsilon \text{ s.t. } S(I_f(X), \epsilon_0, \|\cdot\|') < |\log(\epsilon_0)|(\mathrm{mdim}(X) + \eta)\}.$$

Using Lemma 4.4, we see that

$$\{f : \mathrm{mdim}_{\mathrm{M}}\left(I_f(X), D\right) \leq \mathrm{mdim}(X)\} = \bigcap_{r,n=1}^{\infty} A(\tfrac{1}{n}, \tfrac{1}{r})$$

so it suffice to prove that $A(\epsilon, \eta)$ are open and dense.

That $A(\epsilon, \eta)$ is open is very easy — assume that

$$S(I_f(X), \epsilon_0, \|\cdot\|') < |\log(\epsilon_0)|(\mathrm{mdim}(X) + \eta).$$

Then there is an open cover α of $I_f(X)$ and a $k \in \mathbb{N}$ such that

$$|\alpha| \leq {\epsilon_0}^{-(\mathrm{mdim}(X)+\eta)k}, \qquad \text{and} \tag{4.5}$$

$$\mathrm{mesh}(\alpha, \|\cdot\|_0^{k-1}) < \epsilon_0. \tag{4.6}$$

By extending the sets in α (which are open in the relative topology on $I_f(X)$) to open sets in $K^{\mathbb{Z}}$, we can find a collection $\tilde{\alpha}$ of open sets in $K^{\mathbb{Z}}$ that satisfies the above two conditions, and covers $I_f(X)$. If f' is sufficiently close to f then $\tilde{\alpha}$ will also cover $I_{f'}(X)$. Using (4.3) we immediately deduce that

$$S(I_{f'}(X), \epsilon_0, \|\cdot\|') \leq \frac{\log|\alpha|}{k} \leq |\log(\epsilon_0)|(\mathrm{mdim}(X) + \eta)$$

hence $f' \in A(\epsilon, \eta)$.

It remains to be seen that $A(\epsilon, \eta)$ is dense. Let $\tilde{f}$ be any function in $C(X, K)$, and take any $\varepsilon > 0$. Take α to be an open cover of X such that

$$\forall U \in \alpha \qquad \mathrm{diam}(\tilde{f}(U)) < \varepsilon.$$

Take M big enough so that

$$\mathcal{D}(\alpha_0^{M-1}) < (\mathrm{mdim}(X) + \eta/4)M, \tag{4.7}$$

Let $\beta \succ \alpha_0^{M-1}$ be such that $\mathrm{ord}(\beta) = \mathcal{D}(\alpha_0^{M-1})$, and use Lemma 4.6 with (arbitrary) $p_U \in U$ for every $U \in \beta$ and

$$v_U = (\tilde{f}(p_U), \tilde{f}(Tp_U), \dots, \tilde{f}(T^{M-1}p_U))$$

to find an $F\colon X \to K^M$ such that

$$F(p_U) = v_U \qquad \text{for every } U \in \beta, \tag{4.8}$$

$$F(x) \in \mathrm{co}(F(x_U):\ x \in U \in \beta) \qquad \forall x \in X. \tag{4.9}$$

According to Corollary 3.4 there is a level function $n\colon X \to \mathbb{R}$ such that the set

$$E = \{x:\ n(Tx) \neq n(x) + 1 \text{ or } n(x) \notin \mathbb{Z}\}$$

satisfies

$$\mathrm{ocap}(E) < \frac{\eta}{(8M + 100)(1 + 2\,\mathrm{ord}(\beta))}.$$

Set $\underline{n}(x) = \lfloor n(x) \rfloor \bmod M$, $\bar{n}(x) = \lceil n(x) \rceil \bmod M$ and $n'(x) = \{n(x)\}$. We can now define $f\colon X \to K$ by

$$f(x) = (1 - n'(x))F\big(T^{-\underline{n}(x)}x\big)\big|_{\underline{n}(x)} + n'(x)F\big(T^{-\bar{n}(x)}x\big)\big|_{\bar{n}(x)}. \tag{4.10}$$

At every x such that $n(x) \notin \mathbb{Z}$ the function f is continuous, for in this case in a neighborhood of x the functions $\underline{n}(x)$ and $\bar{n}(x)$ are constant and $n'(x)$ is continuous. At x with $n(x) \in \mathbb{Z}$, if x' is sufficiently close to x, then either $n(x) - \epsilon' < n(x') < n(x)$, in which case $\bar{n}(x') = n(x)$ and $n'(x) > 1 - \epsilon'$, or $n(x) \leq n(x') < n(x) + \epsilon'$, hence $\underline{n}(x') = n(x)$ and $n'(x') < \epsilon'$. In both cases, we can estimate $\|f(x) - f(x')\|$ (for x with $n(x) \in \mathbb{Z}$) as follows:

$$\begin{aligned}\|f(x') - f(x)\| &\leq \epsilon \operatorname{diam} K + \big\|F(T^{n(x)}x')|_{n(x)} - F(T^{n(x)}x)|_{n(x)}\big\| \\ &\leq \epsilon(\operatorname{diam} K + 1)\end{aligned}$$

(the second inequality holds for all x' in some sufficiently small neighborhood of x.) Thus f is continuous. Notice that if $\underline{n}(x) = 0$ and $x \notin \bigcup_{i=0}^{M} T^{-i}E$ (hence in particular $n'(x) = 0$) then

$$I_f(x)|_0^{M-1} = F(x).$$

Claim 1: $\sup_{x\in X}\|f(x)-\tilde f(x)|<\varepsilon$**.**

Indeed, by equation (4.10) and (4.9),

$$f(x)\in \operatorname{co}\left(\left\{F(p_U)\big|_{\underline{n}(x)}: T^{-\underline{n}(x)}x\in U\right\}\cup\left\{F(p_V)\big|_{\bar n(x)}: T^{-\bar n(x)}x\in V\right\}\right) \tag{4.11}$$

But each of the elements of K on the right hand side of (4.11) is within ε of $\tilde f(x)$. Indeed, for any $0\le n<M$, if $T^{-n}x\in U$,

$$\left\|F(p_U)\big|_n-\tilde f(x)\right\|\le \left\|F(p_U)\big|_n - v_U\big|_n\right\| + \left\|v_U\big|_n-\tilde f(x)\right\|. \tag{4.12}$$

The first term of the right hand side of (4.12) is 0 by equation (4.8). Recall that by definition $v_U\big|_n=\tilde f(T^n p_U)$. Now $T^n(p_U)$ and x are in both in $T^n(U)$, and hence

$$\left\|v_U\big|_n-\tilde f(x)\right\|\le \operatorname{diam}(\tilde f(T^nU)).$$

As $U\in\beta\succ\alpha_0^{M-1}$, there is a $U_n\in\alpha$ such that $U\subset T^{-n}(U_n)$. This shows that

$$\operatorname{diam}(\tilde f(T^nU))\le \max_{U'\in\alpha}\operatorname{diam}(\tilde f(U'))<\varepsilon$$

Thus the left hand side of (4.12) is $<\varepsilon$.

Claim 2: $f\in A(\epsilon,\eta)$ **for all** ϵ

It clearly suffices to show that if ϵ is small enough, $I_f(X)$ can be covered by an open cover γ with

$$|\gamma|\le \epsilon^{N(\operatorname{mdim}(X)+\eta)}$$
$$\operatorname{mesh}(\gamma,\|\cdot\|_0^{N-1})<\epsilon.$$

Naturally, we can assume $\eta\ll\operatorname{mdim}(X)$

By (4.7) and (4.9), $F(X)$ is contained in the union of a finite number of $\operatorname{ord}(\beta)\le(\operatorname{mdim}(X)+\eta/4)M$ dimensional polytopes in K^M. By Lemma 4.5, for any $\epsilon>0$, there is an open cover $\gamma'(\epsilon)$ of $F(X)$ with

$$|\gamma'(\epsilon)|\le C'\epsilon^{-(\operatorname{mdim}(X)+\eta/4)M}$$
$$\operatorname{mesh}(\gamma'(\epsilon),\|\cdot\|_0^{M-1})<\epsilon.$$

Using this we can bound the number of sets with diameter at most ϵ needed to cover $f(X)\subset K$. Indeed,

$$f(X)\subset\bigcup_{i=0}^{M-1}\left\{(1-\lambda)F(x)|_i+\lambda F(y)|_{i+1\bmod M}:\ x,y\in X,\ \lambda\in[0,1)\right\},$$

so $f(X)$ is a subset of a finite union of $2\operatorname{ord}(\beta)+1$ dimensional polytopes in K. Again using Lemma 4.5, we see that there is a cover $\gamma''(\epsilon)$ of $f(X)$ with

$$|\gamma''(\epsilon)| \leq C'' \epsilon^{-2\operatorname{ord}(\beta)-1}$$
$$\operatorname{mesh}(\gamma''(\epsilon), \|\cdot\|) < \epsilon.$$

By slight abuse of notation we can consider $\gamma'(\epsilon)$ and $\gamma''(\epsilon)$ as collections of open sets in $K^{\mathbb{Z}}$ (instead of open sets in K^M or K respectively), by replacing every $U \in \gamma'(\epsilon)$ with

$$\{y \in K^{\mathbb{Z}} : \ y|_0^{M-1} \in U\}, \qquad \text{etc.}$$

Take $N > 10(2\operatorname{ord}(\beta)+1)\eta^{-1}M$ such that for every $x \in X$

$$\frac{1}{N}\sum_{i=0}^{N-1} 1_E(T^i x) \leq \frac{\eta}{(8M+50)(2\operatorname{ord}(\beta)+1)}.$$

Consider now a point $I_f(x) \in I_f(X)$. Set

$$J = \{j : \ 0 \leq j < N-1 \text{ and } n(T^j x) = 0 \text{ and } n(T^k x) = k-j \text{ for all } j < k < j+M\}.$$

Then

$$|\mathbf{N} \setminus (J+\mathbf{M})| \leq |\{0 \leq k < N : \ T^k x \in \bigcup_{i=-M}^{M} T^i E\}| + M \leq \frac{\eta N}{2(2\operatorname{ord}(\beta)+1)}$$

(we recall that $\mathbf{N} = \{0, \dots, N-1\}$.) Furthermore, for any $j \in J$,

$$I_f(x)\big|_j^{j+M-1} = F(T^j x). \tag{4.13}$$

Using equation (4.13), we see that there is an element of

$$\gamma(J,\epsilon) := \bigvee_{j \in J} \sigma^{-j}\gamma'(\epsilon) \vee \bigvee_{j \in \mathbf{N}\setminus(J+\mathbf{M})} \sigma^{-j}\gamma''(\epsilon)$$

that contains $I_f(x)$. Let

$$\mathcal{J} = \left\{ \begin{array}{c} J \subset \{0,\dots,N-M\} : \ J \cap (J+i) = \emptyset \text{ for all } 0 < i < M, \text{ and} \\ |\mathbf{N} \setminus (J+\mathbf{M})| \leq \dfrac{\eta N}{2(2\operatorname{ord}(\beta)+1)} \end{array} \right\}$$

We now take

$$\gamma(\epsilon) = \bigcup_{J \in \mathcal{J}} \gamma(J,\epsilon)$$

and claim that $\gamma(\epsilon)$ is the sought after cover of $I_f(X)$.

That $\gamma(\epsilon)$ is indeed a cover is clear, since we showed that for an arbitrary $x \in X$, there is a $J \in \mathcal{J}$ and a $U \in \gamma(J,\epsilon)$ such that $I_f(x) \in U$. Also, by construction, any set of $\gamma(J,\epsilon)$ (and hence any set of $\gamma(\epsilon)$) has $\|\cdot\|_0^{N-1}$ diameter at most ϵ. It remains to bound $|\gamma(\epsilon)|$.

$$\begin{aligned}|\gamma(J,\epsilon)| &= |\gamma'(\epsilon)|^{|J|} \times |\gamma''(\epsilon)|^{|\mathbf{N}\setminus(J+\mathbf{M})|}\\ &\le C\epsilon^{-|J|(\mathrm{mdim}(X)+\eta/4)M-|\mathbf{N}\setminus(J+\mathbf{M})|(2\,\mathrm{ord}(\beta)+1)}\\ &\le C\epsilon^{-(N+M)(\mathrm{mdim}(X)+\eta/4)-N\eta/2} \le C\epsilon^{-N(\mathrm{mdim}(X)+7\eta/8)},\end{aligned}$$

and so

$$|\gamma(\epsilon)| \le C|\mathcal{J}|\epsilon^{-N(\mathrm{mdim}(X)+7\eta/8)}.$$

As long as ϵ is very small,

$$|\gamma(\epsilon)| \le \epsilon^{-(\mathrm{mdim}(X)+\eta)M}.$$

□

Proof of Theorem 4.3:

Let $K = [0,1]^{\mathbb{N}}$ (the Hilbert cube). As X is separable and metric a dense G_δ set of functions $f\colon X \to K$ are an embedding, hence in particular for a dense G_δ set of functions $f\colon X \to K$, $I_f\colon X \to K^{\mathbb{Z}}$ is an embedding (see Hurewicz and Wallman (1941), Theorem V.4).

From Theorem 4.7 we know that for a dense G_δ set of functions $f\colon X \to K$,

$$\mathrm{mdim}_{\mathrm{M}}\left(I_f(X), D\right) \le \mathrm{mdim}(X).$$

If we take f to be an embedding, then D on $I_f(X)$ gives rise to a metric $\tilde{D}$ on X defined by

$$\tilde{D}(x,y) = D(I_f(x), I_f(y)),$$

and hence

$$\mathrm{mdim}_{\mathrm{M}}\,(X,\tilde{D}) \le \mathrm{mdim}(X). \tag{4.14}$$

On the other hand, we already know by Theorem 4.2 that $\mathrm{mdim}(X) \le \mathrm{mdim}_{\mathrm{M}}\,(X,\tilde{D})$. Thus equality holds in equation (4.14), and the theorem is proved. □

5. An embedding theorem

The main result of this section is that if $\mathrm{mdim}(X) < CD$ then (X,T) can be embedded in $\big(([0,1]^D)^{\mathbb{Z}}, \sigma\big)$. Like Theorem 4.3, this theorem has an analogue in dimension theory — the theorem that every space M of topological dimension d can be embedded in $[0,1]^{2d+1}$. This dimension theoretic result is proved by showing (using the Baire Category Theorem) that embeddings are a dense G_δ subset of the space $C(M;[0,1]^{2d+1})$ of all continuous functions from M to $[0,1]^{2d+1}$ with the uniform convergence topology.

As in the previous section, we will work with the maps $I_f \colon X \to K^{\mathbb{Z}}$ where K is compact and convex, and what we shall prove is that if $\dim K$ is larger than some constant C times the mean dimension of X, then for a dense G_δ set of functions $f \in C(X,K)$, the map I_f is an embedding.

Theorem 5.1. *Let (X,T) be an extension of a minimal system, K a convex set with non-empty interior. If $\operatorname{mdim} X < \dim K/36$, then (X,T) can be embedded in $K^{\mathbb{Z}}$. Indeed, in this case for a dense G_δ set of functions $f \in C(X;K)$, the map I_f is an embedding.*

We begin by providing the abstract framework for applying the Baire Category Theorem. This part is nearly identical to the corresponding part in the proof of the well known dimension-theoretic embedding theorem.

Notice that if I_f is an embedding, the inverse image of every point in $I_f(X)$ is, of course, a single point. Hence I_f will be α compatible for every open cover α. Conversely, if $\alpha(i)$ is any sequence of covers such that the mesh of $\alpha(i)$ tends to zero as $i \to \infty$, and if I_f is $\alpha(i)$ compatible for all i, then clearly I_f is an embedding. Indeed, in this case for any $x \in I(f)(X)$ its inverse image $I(f)^{-1}(x)$ must be a subset of some $U(i) \in \alpha(i)$. As $\operatorname{diam} U(i) \leq \operatorname{mesh}(\alpha(i)) \to 0$ we conclude that $I(f)^{-1}(x)$ is a single point, hence $I(f)$ is an embedding. For any open α, let

$$\mathcal{F}_\alpha = \{f \in C(X;K) : \ I_f \succ \alpha\}.$$

Lemma 5.2. *Let X and Y be compact metric spaces, and α and open cover of X. Then $\{f \in C(X,Y) : \ f \succ \alpha\}$ is open in $C(X,Y)$.*

Proof:

Assume $f \succ \alpha$. Let β be an open cover of Y such that $f^{-1}(\beta) \succ \alpha$. For any open $V \subset Y$ we define

$$V_{-\epsilon} = \{y : \ d(y, Y \setminus V) > \epsilon\}.$$

For small enough ϵ, the collection of open sets

$$\beta(\epsilon) = \{V_{-\epsilon} : \ V \in \beta\}$$

covers Y.

Now assume $d(f(x), f'(x)) < \epsilon$ for all x. We show that $f' \succ \alpha$. Indeed, if $f'(x) \in V_{-\epsilon}$ then $f(x) \in V$, so

$$f'^{-1}(\beta(\epsilon)) \succ f^{-1}(\beta) \succ \alpha.$$

□

Lemma 5.3. *Let (X,T) be a dynamical system, $\alpha(i)$ be some sequence of open covers of X with $\mathrm{mesh}(\alpha(i)) \to 0$. Then the set of all $f \in C(X, K)$ such that I_f is an embedding is equal to $\bigcap_{i=1}^{\infty} \mathcal{F}_{\alpha(i)}$, and every $\mathcal{F}_{\alpha(i)}$ is open in $C(X; K)$ (according to the uniform convergence topology).*

Proof:

We have already seen in the discussion preceding Lemma 5.2 that

$$\{f \in C(X; K),\ I_f \text{ is an embedding}\} = \bigcap_{i=1}^{\infty} \mathcal{F}_{\alpha(i)}.$$

It remains to show that for any finite open cover α of X, $\mathcal{F}_\alpha$ is open. The set of $F \in C(X, K^{\mathbb{Z}})$ such that $F \succ \alpha$ is open, and the map

$$I \colon f \mapsto I_f$$

is a continuous map from $C(X, K)$ to $C(X, K^{\mathbb{Z}})$, and so the set of f's such that $I_f \succ \alpha$ is also open as the inverse image of an open set. □

We now begin to prove the harder part of showing that $\mathcal{F}_\alpha$ is dense. If α is an open cover of X, we set $\delta(\alpha) = \delta(\alpha, d)$ to be the Lebesgue constant of the cover α (with respect to the metric d which will often be implicit). Recall that the Lebesgue constant of a cover is the largest $\delta > 0$ such that for every $x \in X$ there is a $U \in \alpha$ such that $d(x, X \setminus U) \geq \delta$.

Lemma 5.4. *Let α be an open cover of a compact metric space Y, and β a cover with mesh finer than $\delta(\alpha)$. Let F be a continuous function from Y to some space Z such that for any x and y in Y such that $F(x) = F(y)$ there is an element of β containing both. Then $F \succ \alpha$*

Proof. Since Y is compact, we only need to prove $F^{-1}(x)$ is contained in some $U \in \alpha$. We know that there is some $U' \in \alpha$ that contains a ball of radius $\delta(\alpha)$ around x. Every $y \in Y$ with $F(x) = F(y)$ is in some $V \in \beta$ with $x \in V$. Since $\mathrm{mesh}(\beta) < \delta(\alpha)$, $V \subset U'$, and hence $y \in U'$. □

Lemma 5.5. *Let $n > m$ and r be integers, M an $n \times m$ matrix with entries in $\{1, \ldots, r\}$ such that no value appears twice in a row or in a column. Then for almost all $t_1, \ldots\ t_r \in \mathbb{R}$,* *the columns of*

$$A(t_1, \ldots, t_r) := \left(t_{M_{i,j}}\right)_{i,j}$$

are linearly independent.

Proof. First, notice that it is enough to prove for the case $m = n$, for we can simply ignore the last $n - m$ rows of A. Thus, we need to show that for almost all $t_1, \ldots, t_r$,

$$\det(A(t_1, \ldots, t_r)) \neq 0$$

which will follow if we prove that the polynomial $\det(A(t_1, \ldots, t_r))$ is non-zero. We now use induction on n. Let $a = M_{1,1}$, and assume that a appears exactly s times in M. To simplify notations, we shall assume $a = 1$. We can write

$$\det(A(t_1, \ldots, t_r)) = f_0(t_2, \ldots, t_r) + t_1 f_1(t_2, \ldots, t_r) + \cdots + t_1^s f_s(t_2, \ldots, t_r).$$

Notice that f_s is, up to sign, the determinant of the minor of A that remains after throwing away all columns and rows in which t_a appears, or 1 if no rows are left. In the former case, the minor thus formed is a smaller matrix that also satisfies the assumptions of the Lemma and so by induction $f_s(t_2, \ldots, t_r) \not\equiv 0$ (and hence $\det(A(t_1, \ldots, t_r)) \not\equiv 0$), in the latter $f_s = 1$ and again $\det(A(t_1, \ldots, t_r)) \not\equiv 0$. $\square$

Recall that we use the notation $v|_r^s$ also for vectors in K^N for finite N, with the convention that the coordinates are numbered from 0 to $N - 1$.

Lemma 5.6. *Let Δ be an integer, let β be a cover of X with $\operatorname{ord}(\beta) < \Delta \dim(K)$, and let $\epsilon > 0$. Suppose we are given for every $U \in \beta$ a point $p_U \in U$ and $v_U \in K^N$. Then is possible to find a continuous function $F\colon X \to K^N$ with the following properties:*

1. $\left\|(F(p_U) - v_U)\big|_k\right\| < \epsilon$ *for all* $k = 0, \ldots, N - 1$,
2. *For all* $x \in X$, $F(x) \in \operatorname{co}(F(p_U) : x \in U \in \beta)$.
3. *If for some* $0 \leq \ell, j < N - 4\Delta$, *and* $\lambda, \lambda' \in (0, 1]$,

$$\lambda F(x)|_\ell^{\ell+4\Delta-1} + (1-\lambda) F(y)|_{\ell+1}^{\ell+4\Delta} = \lambda' F(x')|_j^{j+4\Delta-1} + (1-\lambda') F(y')|_{j+1}^{j+4\Delta}$$

then there is a $U \in \beta$ such that both x and $x' \in U$.

Proof:

Let $\{\phi_U(x)\}_{U \in \beta}$ a partition of unity subordinate to β — that is, a collection of continuous functions $X \to [0, 1]$ such that

$$\sum_{U \in \beta} \phi_U(x) = 1 \qquad \text{for all } x \in X$$

and $\operatorname{supp}(\phi_U) \subset U$, and we can further assume that $\phi_U(p_U) = 1$ for all $U \in \beta$. We chose for every $U \in \beta$ a value $F_U \in K^N$ such that

$$\left\| (F_U - v_U)\big|_k \right\| < \epsilon \qquad \text{for } k = 0, \ldots, N-1,$$

and define F to be

$$F(x) = \sum_{U \in \beta} \phi_U(x) F_U.$$

$F(x)$ clearly satisfies the first two conditions of the Lemma. We claim that for almost every choice of values F_U, this function also satisfies the third condition.

First notice that for any $S_i, S_i' \subset \beta$ with $|S_i|, |S_i'| \leq \Delta \dim(K)$, the following two collections of vectors in $K^{4\Delta}$ obey the conditions of Lemma 5.5, where $0 \leq \ell < M - 4\Delta$ and $\ell + 1 \leq k < M - 4\Delta$:

1st collection: All the vectors of the form $F_{U_i}\big|_{\ell+i}^{\ell+4\Delta-1+i}$ where i is 0, 1, and if $\ell < M - 4\Delta - 1$ also $i = 2$, and $U_i \in S_i$.

2nd collection: All the vectors of the form $F_{U_i}\big|_{\ell+i}^{\ell+4\Delta-1+i}$ or $F_{V_i}\big|_{k+i}^{k+4\Delta-1+i}$, where i is 0 or 1, the set $U_i \in S_i$, and $V_i \in S_i'$.

Indeed, these collections contain at most $4\Delta \dim(K)$ vectors of $K^{4\Delta}$ which is $4\Delta \dim(K)$ dimensional. Also, for every $U \in \alpha$ and $0 \leq k \leq N-1$ the variable $F_U\big|_k$ appears at most once in every row and column. Thus, from Lemma 5.5, for almost every choice of F_U's, both of the above collections of vectors are linearly independent for every choice of S_i, S_i'. We can assume that the F_U we have chosen satisfy this property.

Assume that for $0 \leq \ell \leq j < N - 4\Delta$, $\lambda, \lambda' \in (0,1]$,

$$\lambda F(x)|_{\ell}^{\ell+4\Delta-1} + (1-\lambda) F(y)|_{\ell+1}^{\ell+4\Delta} = \lambda' F(x')|_{j}^{j+4\Delta-1} + (1-\lambda') F(y')|_{j+1}^{j+4\Delta}. \tag{5.1}$$

If $j > \ell + 1$ set

$$S_0 = \{U \in \beta : \ x \in U\}, \qquad S_1 = \{U \in \beta : \ y \in U\}$$
$$S_0' = \{U \in \beta : \ x' \in U\}, \qquad S_1' = \{U \in \beta : \ y' \in U\}.$$

We now use property 2 in the statement of the Lemma which we have already proved to deduce from (5.1) that the 2nd collection above is not linearly independent — a contradiction. Similarly, we use the fact that the 1st collection is linearly independent to deduce that $j = \ell + 1$ is impossible.

There remains the case $j = \ell$. In this case, the linear independence of the vectors in the 1st collection shows that for any $U \ni x$

$$\phi_U(x) = \phi_U(x')$$

for otherwise we again get a non trivial linear relation. Thus if we take some U such that $\phi_U(x) \neq 0$, both x and x' are in this U. □

Lemma 5.7. *Let M be an even integer, and $n\colon X \to \mathbb{R}$ a function such that*

$$\{x: \ n(Tx) \neq n(x) + 1\} \cap \{x: \ n(T^{k+1}x) \neq n(T^k x) + 1\} = \emptyset$$

for all $1 \leq k \leq 5M$. Then for any x_1 and $x_2 \in X$, there is an $1 \leq r \leq 4M$ such that for $r \leq s \leq r + M/2 - 1$ and $i = 1$, 2

$$n(T^s x_i) \bmod M = (n(T^r x_i) \bmod M) + s - r \tag{5.2}$$

Remark: Notice that the mod operation in the right hand side of (5.2) is performed *before* adding $s - r$. Hence, in particular, this implies that $n(T^r x_i) \bmod M \leq M/2 + 1$

Proof:

By the condition on N there is (for $i = 1$, 2) at most one j_i such that

$$n(T^{j_i+1}x_i) \neq n(T^{j_i}x_i) + 1.$$

We extend this to the case that j_i is undefined by setting $j_i = -1$,

At least one of the intervals

$$\begin{aligned} &\{0, \dots, \min(j_1, j_2) - 1\}, \\ &\{\min(j_1, j_2) + 1, \dots, \max(j_1, j_2) - 1\}, \qquad \text{or} \\ &\{\max(j_1, j_2) + 1, \dots, 9M/2 - 1\} \end{aligned}$$

is of size $3M/2 - 1$. We denote this interval as $\{a, \dots, b\}$. There are now two cases:

- If there is no $j \in \{b - M/2 + 1, \dots, b\}$ such that $n(T^j x_i) \bmod M \in [M - 1, M)$ for either $i = 1$ or 2, then we can take $r = b - M/2 + 1$.
- If there is $j \in \{b - M/2 + 1, \dots, b\}$ such that (for example) $n(T^j x_i) \bmod M \in [M - 1, M)$ then either $r = j - M + 1$ or $r = j - M/2 + 1$ will work.

□

The following lemma finishes the proof of Theorem 5.1.

Lemma 5.8. *If* $\operatorname{mdim}(X) < \dim(K)/36$*, then for any open cover* α *of* X *the set* $\mathcal{F}_\alpha$ *is dense in* $C(X,K)$*.*

Proof:

Assume $\tilde{f} \in C(X,K)$ and $\epsilon > 0$ are given. Let β be an open cover of X fine enough so that $\operatorname{diam}(\tilde{f}(U)) < \epsilon/2$ for every $U \in \beta$ and such that $\operatorname{mesh}(\beta) < \delta(\alpha)$. Let $\epsilon' > 0$ be chosen so that $36(1+\epsilon')\operatorname{mdim}(X) < \dim(K)$.

Choose N so large such that

$$\mathcal{D}(\beta_0^{N-1}) < N \operatorname{mdim}(X)(1+\epsilon')$$

and

$$N\left(1 - 36(1+\epsilon')\frac{\operatorname{mdim}(X)}{\dim(K)}\right) \gg 1.$$

We take $\beta' \succ \beta_0^{N-1}$ to be an open cover of X with $\operatorname{ord}(\beta') = \mathcal{D}(\beta_0^{N-1})$. Let M be a positive even integer roughly proportional to N which we will fix later, $\Delta = \lceil \operatorname{ord}(\beta')/\dim(K) \rceil$. For every $U \in \beta'$ fix some $p_U \in U$, and set

$$v_U = (\tilde{f}(p_U), \tilde{f}(Tp_U), \dots, \tilde{f}(T^{N-1}p_U))$$

We now construct a function $F\colon X \to K^M$ using Lemma 5.6, for the cover β' and the parameters Δ and $\epsilon/2$.

From this function F we now construct an $f\colon X \to \mathbb{R}$ as follows. Use Lemma 3.3 to find an $n\colon X \to \mathbb{R}$ such that

$$\{x:\ n(Tx) \neq n(x)+1\} \cap \{x:\ n(T^{k+1}x) \neq n(T^k x)+1\} = \emptyset$$

for all $1 \le k \le 100N$. Set $\underline{n}(x) = \lfloor n(x) \rfloor \bmod M$, $\bar{n}(x) = \lceil n(x) \rceil \bmod M$, and $n'(x) = \{n(x)\}$. Now define f by

$$f(x) = (1-n'(x))F(T^{-\underline{n}(x)}x)\big|_{\underline{n}(x)} + n'(x)F(T^{-\bar{n}(x)}x)\big|_{\bar{n}(x)}. \tag{5.3}$$

It is rather straightforward to see that f is continuous and $\sup_{x\in X}\|f(x) - \tilde{f}(x)\| < \epsilon$, and as we have given a detailed proof of the analogous claim in the proof of Theorem 4.7, we omit the details.

It remains to be shown that $f \in \mathcal{F}_\alpha$.

Suppose that $I_f(x) = I_f(x')$. We will show that if M has been defined properly this implies that there is an $U \in \beta$ such that both x and $x' \in U$, and so using Lemma 5.4 $I_f \succ \alpha$, which establishes the claim.

Let C be an integer that will also be determined later. By Lemma 5.7 there are r and s satisfying $C \leq r < s \leq C + 9M/2$ such that $s - r > M/2 - 10$ and for any $r \leq j \leq s$

$$n(T^j x) \bmod M = (n(T^r x) \bmod M) + j - r$$
$$n(T^j x') \bmod M = (n(T^r x') \bmod M) + j - r,$$

so if we take $\lambda = n'(T^r x)$, $a = \underline{n}(T^r x)$ and $b = a + s - r - 1$

$$I_f(x)\Big|_r^s = (1-\lambda)F(T^{r-a}x)\Big|_a^{b-1} + \lambda F(T^{r-a-1}x)\Big|_{a+1}^b.$$

and a similar equation holds for x', with parameters b', a' and λ'. As long as

$$b - a \geq 4\Delta, \tag{5.4}$$

we can use property 3 in Lemma 5.6 to find a $U \in \beta' \succ \beta_0^{N-1}$ so that both $x, x' \in T^{-r+a}U$. We know that $0 \leq a < M/2 + 10$, and $C \leq r \leq C + 4M + 10$ hence $C - M/2 - 10 \leq r - a \leq C + 4M + 20$. Now, since $U \in \beta' \succ \beta_0^{N-1}$ there is some $U' \in \beta_0^{N-1}$ such that

$$T^{-r+a}U \subset T^{-r+a}U' \in T^{-r+a}\beta_0^{N-1} = \beta_{r-a}^{N-1+r-a} \succ \beta_{C+4M+20}^{N+C-M/2-10}. \tag{5.5}$$

If we chose C and M so that

$$C + 4M + 20 \leq 0 \leq N + C - M/2 - 10 \tag{5.6}$$

then (5.5) implies that there will be a $V \in \beta$ containing both x and x'.

It only remains to find adequate M and C. Since $b - a \geq M/2 - 10$ and

$$\Delta = \lceil \mathrm{ord}(\beta')/\dim(K) \rceil \leq N\frac{\mathrm{mdim}(X)}{\dim(K)}(1+\epsilon') + 1,$$

the inequality (5.4) will be satisfied if

$$M/2 - 20 \geq 4N\frac{\mathrm{mdim}(X)}{\dim(K)}(1+\epsilon') \tag{5.7}$$

and we can find C that satisfies (5.6) if $N \geq 9M/2 + 30$. Use (5.7) (with equality sign instead of $\geq$) to define M. Then in order to satisfy the second inequality it suffices that

$$N \geq 36N\frac{\mathrm{mdim}(X)}{\dim(K)}(1+\epsilon') + 1000$$

and this is satisfied when $36(1+\epsilon')\frac{\mathrm{mdim}(X)}{\dim(K)} < 1$ and N is big enough. $\square$

6. Some theorems regarding systems with $\mathrm{mdim}(X) = 0$

In this section, we consider zero mean dimensional extensions of minimal dynamical systems. We recall that in Lindenstrauss and Weiss (1999) we have seen that the collection of systems (X, T) with zero mean dimension is rather rich, and contains all dynamical systems with finite topological entropy, all dynamical systems for which X has finite topological dimension and all uniquely ergodic systems.

The basic tool in investigating these systems is the notion of uniformly small sets, which we defined in §3.

Definition 6.1. *A dynamical system (X, T) has the* **small-boundary property** *(SBP) if every point $x \in X$ and every open $U \ni x$ there is a neighborhood $V \subset U$ of x with uniformly small boundary.*

Our first structure theorem on systems with $\mathrm{mdim}(X) = 0$ is the following:

Theorem 6.2. *If (X, T) is an extension of a minimal system with $\mathrm{mdim}(X) = 0$ then (X, T) has the SBP.*

We note that as shown in Lindenstrauss and Weiss (1999), §5 the converse is also true — any dynamical system with the SBP must have zero mean dimension.

Just as in the previous proofs, we shall use the Baire Category Theorem to prove this result. Again, we can find an analogy to the proof of a standard result in dimension theory. Recall that there are two standard definitions of topological dimension: one using covers, the so-called Lebesgue cover dimension, and an inductive definition. Our proof is similar to the harder direction in the proof that the Lebesgue cover dimension is the same as the inductive dimension (Hurewicz and Wallman (1941), Theorem V.5).

Indeed, this is not so surprising, since our definition of mean dimension is based on the definition of the Lebesgue cover dimension, and the SBP is similar to the inductive definition of zero dimension — a space has zero topological dimension if for every point $x \in X$ and every open $U \ni x$ there is a neighborhood $V \subset U$ of x with empty boundary.

Our basic strategy is to consider for $f \in C(X, [0, 1])$ the image $I_f(X) \subset [0, 1]^{\mathbb{Z}}$ which we know is isomorphic to (X, T) for a dense G_δ subset of functions in $C(X, [0, 1])$. A natural countable basis for the topology of $[0, 1]^{\mathbb{Z}}$ are the cylinder sets

$$\tilde{C}^n(p_{-n}, \ldots, p_n; q_{-n}, \ldots, q_n) = \{x \in [0, 1]^{\mathbb{Z}} : \ \forall -n \leq i \leq n \ x_i \in (p_i, q_i)\},$$

for all $n \in \mathbb{N}$ and rational $p_{-n}, \ldots, p_n, q_{-n}, \ldots, q_n$. The intersection of these sets with $I(f)(X)$ form a basis $C^n(p_{-n}, \ldots, p_n; q_{-n}, \ldots, q_n)$ for the topology of $I(f)(X)$. We show that for a dense G_δ set of f's, the boundary (in $I_f(X)$) of $C^n(p_{-n}, \ldots, p_n; q_{-n}, \ldots, q_n)$ is uniformly small (as subsets of the dynamical system $(I_f(X), \sigma)$). If, in addition, I_f is an embedding of X in $K^{\mathbb{Z}}$, then this gives us a basis for the topology of X with uniformly small boundaries, hence X has the SBP. To show that $\partial C^n(p_{-n}, \ldots, p_n; q_{-n}, \ldots q_n)$ is small for all n and p_i, $q_i \in \mathbb{Q}$ it is clearly suffice to show that for any one $t \in [0, 1]$ (hence also for all $q \in \mathbb{Q}$), for a dense G_δ set of f's

$$I_f(X) \cap \{x \in [0,1]^{\mathbb{Z}} : \ x_0 = t\} \cong \{x \in X : \ f(x) = t\} \tag{6.1}$$

is uniformly small. Thus the main part of the proof is to show that indeed, the set in (6.1) is uniformly small for a generic f.

We will again use the notation $\mathbf{N}$ for the set $\{0, \ldots, N-1\}$.

Lemma 6.3. *Let $E \subset X$ be closed, $\epsilon > 0$ arbitrary. Then there is an open $U \supset E$ such that*

$$\operatorname{ocap}(U) \leq \operatorname{ocap}(E) + \epsilon$$

Proof:

Let N be large enough so that for all $x \in X$

$$\frac{\sum_{i=0}^{N-1} 1_E(T^i x)}{N} < \operatorname{ocap}(E) + \epsilon.$$

What this means is that the intersection of every $N' = \lfloor N(\operatorname{ocap}(E) + \epsilon) \rfloor + 1$ sets from the collection

$$\{T^{-k}E : \ 0 \leq k < N\}$$

is empty. Let

$$\rho(x) = \min_{\substack{I \subset \mathbf{N} \\ |I| \geq N'}} \sum_{i \in I} d(T^i x, E).$$

$\rho(x)$ is strictly positive and continuous, and so there is some $\delta > 0$ such that $\rho(x) > \delta$ for all X. Set

$$U = \{x \colon d_0^{N-1}(x, E) < \frac{\delta}{N'}\}.$$

If there was an $I \subset \mathbf{N}$ with $|I| \geq N'$ such that $\bigcup_{i \in I} T^{-i}U \neq \emptyset$, we would take $x \in \bigcup_{i \in I} T^{-i}U$, and then for all $i \in I$ we would have $d(T^i x, E) < \delta/N'$. Summing over $i \in I$ we would get

$\rho(x) < \delta$, a contradiction. Thus

$$\sup_{x \in X} \frac{\sum_{i=0}^{N-1} 1_U(T^i x)}{N} < N(\mathrm{ocap}(E) + \epsilon)$$

and since $\sup_{x \in X} \sum_{i=0}^{n-1} 1_U(T^i x)$ is sub-additive in n

$$\mathrm{ocap}(U) = \inf \sup_{x \in X} \frac{\sum_{i=0}^{n-1} 1_U(T^i x)}{n} < \mathrm{ocap}(E) + \epsilon.$$

□

Lemma 6.4. *Let (X, T) be any dynamical system, $t \in [0,1]$. The set of functions $f \colon X \to [0,1]$ such that the set $\{x \in X : f(x) = t\}$ is uniformly small is a G_δ subset of $C(X, [0,1])$.*

Proof: First note that

$$\Big\{ f : \{f(x) = t\} \text{ is small} \Big\} = \bigcap_{n=1}^{\infty} \big\{ f : \mathrm{ocap}(\{x : \ f(x) = t\}) < 1/n \big\} \tag{6.2}$$

The sets $\big\{ f : \mathrm{ocap}(\{x : \ f(x) = t\}) < 1/n \big\}$ are open. Indeed, suppose for some $f \in C(X, [0,1])$ the set

$$E := \{x : \ f(x) = t\}$$

has $\mathrm{ocap}(E) < 1/n$. Use Lemma 6.3 to find a $U \supset E$ that is open with $\mathrm{ocap}(U) < 1/n$. Let

$$\delta = \min_{x \in X \setminus U} |f(x) - t| > 0.$$

For any $\tilde{f}$ with $\sup_x |f(x) - \tilde{f}(x)| < \delta$, we see that

$$\{x : \ \tilde{f}(x) = t\} \subset U$$

hence $\mathrm{ocap}(\{x : \ \tilde{f}(x) = t\}) < 1/n$. □

We recall that if $v \in [0,1]^N$, we will use to the notation $v|_k$ to designate the k'th coordinate of v, where the coordinates of v are numbered between 0 and $N-1$.

The following lemma is quite similar in many respects to Lemma 5.6, and is used in a similar way.

Lemma 6.5. *Let β be a cover of X with $\mathrm{ord}(\beta) < \Delta$. Suppose we are given a $t \in [0,1]$ and, for every $U \in \beta$, a point $p_U \in U$ and $v_U \in [0,1]^M$. Then is possible to find a continuous function $F \colon X \to [0,1]^M$ with the following properties:*

1. $\left\| (F(p_U) - v_U)\big|_k \right\| < \epsilon$ *for all* $k = 0, \ldots, M-1$,
2. *For all* $x \in X$, $\ F(x) \in \mathrm{co}(F(x_U) : \ x \in U \in \beta)$.
3. *For every* $x \in X$, *no more than* Δ *of the coordinates of* $F(x)$ *are equal to* t.

Proof:

Let $\{\phi_U(x)\}_{U\in\beta}$ a partition of unity subordinate to β — we recall that this means that each $\phi_U \in C(X,[0,1])$, that

$$\sum_{U\in\beta} \phi_U(x) = 1 \qquad \text{for all } x \in X$$

and $\operatorname{supp}(\phi_U) \subset U$, and we again assume that $\phi_U(p_U) = 1$ for all $U \in \beta$. We choose for every $U \in \beta$ a value $F_U \in K^M$ such that

$$\left\| (F_U - v_U)\big|_k \right\| < \epsilon \qquad \text{for } k = 0, \ldots, M-1,$$

and define F to be

$$F(x) = \sum_{U\in\beta} \phi_U(x) F_U.$$

$F(x)$ clearly satisfies the first two conditions of the Lemma.

We now show that for almost every choice of values F_U, this function also satisfies the third condition. Each point in $F(X)$ is contained in a at most Δ-dimensional affine subspace of $\mathbb{R}^M$ spanned by at most $\Delta+1$ of the vectors F_U. Generically each one of these subspaces will not intersect any of the $M-\Delta-1$ dimensional subspaces

$$\{v \in \mathbb{R}^M : \forall i \in I \ v_i = t\} \qquad \text{where } I \subset \{0,\ldots,M-1\} \text{ with } |I| \geq \Delta+1.$$

□

Lemma 6.6. *If (X,T) is an extension of a minimal system with* $\operatorname{mdim}(X) = 0$, *then for any $\epsilon > 0$ the sets*

$$\{f : \operatorname{ocap}(\{x : \ f(x) = t\}) < \epsilon\}$$

are dense in $C(X,[0,1])$

Proof:

Let $\tilde{f} \in C(X;[0,1])$ and $\varepsilon > 0$ be arbitrary. We show there is an f such that the orbit capacity of $\{x : \ f(x) = t\}$ is less than ϵ, and such that f is within ε of $\tilde{f}$. Take α be a cover such that

$$\operatorname{diam} \tilde{f}(U) < \varepsilon/2 \qquad \text{for all } U \in \alpha.$$

Let M be large enough so that $\mathcal{D}(\alpha_0^{M-1}) < \epsilon M/4$, and let $\beta \succ \alpha_0^{M-1}$ such that

$$\operatorname{ord}(\beta) = \mathcal{D}(\alpha_0^{M-1}).$$

Pick for every U, a point $p_U \in U$ and set

$$v_u = (\tilde{f}(p_U), \tilde{f}(Tp_U), \dots, \tilde{f}(T^{M-1}p_U)).$$

Let $F\colon X \to [0,1]^M$ be a function as in Lemma 6.5 for β, $\varepsilon/2$, $\Delta = M\epsilon/4$, p_U, v_U and t.

Find a level function $n\colon X \to \mathbb{R}$ (using Corollary 3.4) so that the set E defined by

$$E = \{x : n(x) \notin \mathbb{Z} \text{ or } n(Tx) \neq n(x) + 1\}$$

satisfies $\operatorname{ocap}(E) < \epsilon/4M$, and let $\underline{n}(x) = \lfloor n(x) \rfloor \bmod M$, $\bar{n}(x) = \lceil n(x) \rceil \bmod M$, and $n'(x) = \{n(x)\}$. Define $f \in C(X, [0,1])$ by

$$f(x) = (1 - n'(x))F(T^{-\underline{n}(x)}x)|_{\underline{n}(x)} + n'(x)F(T^{-\bar{n}(x)}x)|_{\bar{n}(x)}$$

Now, as in Theorem 4.7 and Lemma 5.8, it is easy to see that properties 1–2 (in the statement of Lemma 6.5) of F imply that

$$\sup_{x\in X} |f(x) - \tilde{f}(x)| < \varepsilon.$$

It remains to verify that $\operatorname{ocap}(\{x : \ f(x) = t\}) < \epsilon$. Take $N \geq 100M/\epsilon$ so that for every $x \in X$

$$\frac{1}{N}\sum_{i=0}^{N-1} 1_E(T^i x) < \frac{\epsilon}{4M}.$$

We show that for all $x \in X$,

$$\frac{1}{N}\sum_{i=0}^{N-1} 1_{f(T^i x) = t} < \epsilon, \tag{6.3}$$

proving the lemma. Let

$$J = \left\{ \begin{array}{l} 0 \leq j \leq N - M : \ n(T^j x) \bmod M = 0 \text{ and} \\ \qquad\qquad\quad n(T^{j+k}x) = n(T^j x) + k \text{ for } k = 1, \dots, \text{M-1} \end{array} \right\}$$

Two things are clear: first, for every j and $j' \in J$ the distance $|j - j'| > M$. Second, from the condition on the function $n(\cdot)$,

$$|J + \mathbf{M}| > N - M - \epsilon N/4 > (1 - \epsilon/2)N.$$

Now for every $j \in J$, by the third property of F from Lemma 6.5, t can appear at most $\Delta = \epsilon M/4$ times in the finite sequence

$$f(T^j x), f(T^{j+1}x), \dots, f(T^{j+M-1}x).$$

Hence the number of times t can appear in

$$f(x), f(Tx), \dots, f(T^{N-1}x)$$

is at most

$$|J|(\epsilon M/4) + |\mathbf{N} \setminus (J+\mathbf{M})| \leq \frac{N\epsilon M}{4M} + \frac{\epsilon N}{2} < \epsilon N$$

which is exactly what is required in equation (6.3). $\square$

The proof of Theorem 6.2 is now straight forward:

Proof of Theorem 6.2:

Let (X,T) be an extension of a minimal system with $\mathrm{mdim}(X)=0$. By Theorem 5.1, for a dense G_δ set of $f \in C(X,[0,1])$ the map I_f is an embedding. By Lemmas 6.4 and 6.6, for every $t \in \mathbb{Q} \cap (0,1)$, for a dense G_δ set of $f \in C(X,[0,1])$ the subset $\{x_0 = t\} \cap I_f(X)$ is small. Hence there is an f such that I_f is an embedding, and such that for every $t \in \mathbb{Q} \cap (0,1)$, and every i, the set $\{x_i = t\} \cap I_f(X)$ is small. As finite intersections of the sets

$$\{x \in I_f(X) : t < x_i < s\}$$

for $i \in \mathbb{Z}$, and t, $s \in \mathbb{Q}$, form a basis for the topology of $I_f(X) \cong X$ we are done. $\square$

We now present applications of this result. In Lindenstrauss (1995), section 4, the SBP property is used to construct small entropy factors. The argument, at least in the case we are most interested in where there are no periodic points, is also given (somewhat implicitly) in Shub and Weiss (1991). We can summarize the result we need from Lindenstrauss (1995), section 4 in the following theorem:

Theorem 6.7 (Shub and Weiss (1991), Lindenstrauss (1995))**.** *If (X,T) has the SBP, then for any $a \neq b \in X$ and $\epsilon > 0$ there is a factor map ϕ such that $h_{\mathrm{top}}(\phi(X)) < \epsilon$ and $\phi(a) \neq \phi(b)$.*

The proof of this result is not too hard. One important observation is the following lemma that is needed to relate our definition of the SBP to the discussion in Lindenstrauss (1995), section 4.

Lemma 6.8. *If (X,T) has the SBP, then for any open sets U, $U' \subset X$ with $\bar{U} \subset U'$, there is an open $U \subset V \subset U'$ with* $\mathrm{ocap}(\partial V) = 0$.

Proof: Using the SBP, for any $x \in \partial U$ there is a open set $V_x \subset U'$ with ∂V_x uniformly small (i.e. $\mathrm{ocap}(\partial V_x) = 0$). A finite number of these, say $V_{x_1}, \dots, V_{X_N}$ suffice to cover ∂U. Set

$$V = \bigcup_{i=1}^{N} V_{x_i} \cup U.$$

Then $\partial V \subset \bigcup_{i=1}^N \partial V_{x_i}$, and so as the finite union of uniformly small sets is easily seen to be uniformly small. $\square$

Rough sketch of Proof of Theorem 6.7:

We only treat the simpler case that a and b are non-periodic. In this case, using Lemma 6.8, and Lemma 6.3, one can construct a countable collection of sets

$$\mathcal{C} = \{U_1, U_2, \dots\}$$

such that

- $a \in U_1, \;\; b \notin \bar{U}_1$,
- for every $U \in \mathcal{C}$ and every open $V \supset \partial U$, there is $V' \in \mathcal{C}$ with $\partial U \subset V' \subset V$,
- $\operatorname{ocap} U_i \to 0$ arbitrarily fast.

We define an equivalence relation $\sim$ on X as follows:

$$x \sim y \qquad \iff \qquad \forall n \in \mathbb{Z} \text{ and } U \in \mathcal{C},\; 1_{T^{-n}U}(x) = 1_{T^{-n}U}(y).$$

One needs to verify that $\sim$ is a closed equivalence relation — that is

$$\{(x,y): \; x \sim y\}$$

is a closed subset of $X \times X$. In this case, $X/{\sim}$ can be given in a natural way a nice topology, such that the map $x \mapsto x/{\sim}$ is continuous (we note that in Lindenstrauss (1995), p. 248 the definition of the quotient topology is faulty. However, this is used nowhere in that paper). Notice also that $a \not\sim b$. Since in addition $\sim$ is T-invariant, i.e. $Tx \sim Ty$ iff $x \sim y$, it is possible to define a continuous $\tilde{T}\colon X/{\sim} \to X/{\sim}$ so that $(X/\sim, \tilde{T})$ is a factor of (X,T).

Finally, one needs to estimate $h_{\mathrm{top}}(X/\sim)$. For any $U \in \mathcal{C}$, Set $\mathcal{S}_U \subset \{0,1\}^{\mathbb{Z}}$ by

$$\mathcal{S}_U = \overline{\{(\dots, 1_U(T^{-1}x), 1_U(x), 1_U(Tx), \dots): \; x \in X\}}$$

(where the closure is taking according to the usual product topology on $\{0,1\}^{\mathbb{Z}}$), and take σ to be the shift operation on $\{0,1\}^{\mathbb{Z}}$. One bounds $h_{\mathrm{top}}(X/\sim, \tilde{T})$ by showing

$$h_{\mathrm{top}}(X/\sim, \tilde{T}) \le \sum_i h_{\mathrm{top}}(\mathcal{S}_{U_i}, \sigma).$$

The sum on the right hand side can be made arbitrarily small if $\operatorname{ocap}(U_i)$ are very small and tend very rapidly to zero. $\square$

Corollary 6.9. *If (X,T) is an extension of a minimal system with $\operatorname{mdim}(X) = 0$ then for any two distinct points $a, b \in X$ and $\epsilon > 0$ there is a factor map ϕ such that $h_{\mathrm{top}}(\phi(X)) < \epsilon$ and $\phi(a) \neq \phi(b)$.*

This Corollary has a rather strong converse, which is our next aim. Before stating it, we first prove some auxiliary results.

Definition 6.10. *Let (X_i, T_i) for $i = 1, 2, \dots$ be a sequence of dynamical systems, and assume that for every $i > j \geq 1$ we have a factor map $\pi(i,j)\colon X_i \to X_j$, with $\pi(j,k) \circ \pi(i,j) = \pi(i,k)$ for every $i > j > k$. The inverse limit $\varprojlim(X_i, T_i)$ is a dynamical system (X, T) defined as follows:*

$$X = \{(x_1, x_2, \dots,) : \ x_i \in X_i \text{ and } \pi(i,j)x_i = x_j\},$$

with the topology inherited from $X_1 \times X_2 \times \dots$ and $T\colon X \to X$ is simply

$$T\colon (x_1, x_2, \dots) \mapsto (T_1(x_1), T_2(x_2), \dots)$$

Notice that if all (X_i, T_i) are factors of some $(\hat{X}, \hat{T})$ with factor maps $\phi(i)\colon \hat{X} \to X_i$ such that for $i > j$

$$\phi(i,j) \circ \phi(i) = \phi(j)$$

then $\varprojlim(X_i, T_i)$ is also a factor of $(\hat{X}, \hat{T})$.

Proposition 6.11. *If (X_i, T_i) are dynamical systems with $\operatorname{mdim}(X_i) = 0$ for all i, and $\pi(i,j)$ are as above, then*

$$\operatorname{mdim}(\varprojlim X_i) = 0.$$

Proof:

Let

$$(X, T) = \varprojlim(X_i, T_i),$$

and take α be a finite open cover of $X \subset X_1 \times X_2 \times \dots$. Then α has a refinement α' of the form

$$\alpha' = \bigvee_{i=1}^{n} \pi_i^{-1}(\alpha(i)),$$

where $\alpha(i)$ is an open cover of X_i and $n \in \mathbb{N}$. Then for large enough N,

$$\begin{aligned}
\frac{\mathcal{D}(\alpha|_0^{N-1})}{N} &\leq \frac{\mathcal{D}(\bigvee_{i=1}^{n} \alpha'(i)|_0^{N-1})}{N} \\
&\leq \sum_{i=1}^{n} \frac{\mathcal{D}(\alpha(i)|_0^{N-1})}{N} \\
&\leq \sum_{i=1}^{n} \operatorname{mdim}(X_i) + \epsilon = \epsilon.
\end{aligned}$$

□

Proposition 6.12. *Let (X,T) be a dynamical system. Then X has a* **universal zero mean dimensional factor***, i.e. a factor (Y,S) with factor map $\phi\colon X \to Y$ such that* $\mathrm{mdim}(Y) = 0$*, and such that for any factor map $\psi\colon X \to Z$ with* $\mathrm{mdim}(Z) = 0$ *there is a factor map $\psi'\colon Y \to Z$ such that $\psi = \psi' \circ \phi$.*

Proof:

For any factor map $\psi_Z\colon X \to Z$ with $\mathrm{mdim}(Z) = 0$ denote

$$N(\psi_Z) = \{(x,y) : \psi_Z(x) \neq \psi_Z(y)\} \subset X \times X.$$

As $N(\psi_Z)$ is open, we can find a sequence Z_i such that

$$\bigcup_{\mathrm{mdim}(Z)=0} N(\psi_Z) = \bigcup_{i=1}^{\infty} N(\psi_{Z_i}).$$

Take

$$\phi(i) = \psi_{Z_1} \times \psi_{Z_2} \times \dots \psi_{Z_i}$$
$$Y_i = \phi(i)(X)$$

and for $i > j$ $\pi(i,j)\colon Y_i \to Y_j$ the projection on the first j coordinates. Now take

$$Y = \varprojlim(Y_i)$$

which is also a factor of (X,T). By equation (2.2) we cited from Lindenstrauss and Weiss (1999), $\mathrm{mdim}(Y_i) = 0$ for all i, and so $\mathrm{mdim}(Y) = 0$. It is easy to see that Y has the required universality property. □

The universality of the above factor implies the following result, which can be regarded as a strong converse to Corollary 6.9.

Theorem 6.13. *Let (X,T) be any dynamical system, x, $y \in X$. If there is a finite entropy factor map of X that distinguishes between x and y then the images of x and y in the universal zero mean dimensional factor of X are distinct.*

Remark: For minimal systems, the universal zero mean dimensional factor will also be minimal, and so in this case any two point of X that project to distinct points in the universal zero mean dimensional factor can be distinguished by factors with *arbitrarily small entropy.*

We conclude this section by two interesting observation.

Proposition 6.14. *If (X,T) is an extension of a minimal system, then X is the inverse limit of systems with finite entropy iff* $\mathrm{mdim}(X) = 0$

Proof:

Suppose $\mathrm{mdim}(X) = 0$. Since we know that finite entropy factor maps of X separate points, we can take the Z_i in the proof of Proposition 6.12 to be finite entropy factors of (X, T), and then the resulting factor Y will be the inverse limit of finite entropy systems. However, since $\mathrm{mdim}(X) = 0$, from the universality property of Y, the factor transformation $\mathrm{id}\colon X \to X$ factors through Y, so $Y \cong X$.

The converse follows from Lemma 6.11. □

Recall the definition of $S(X, \epsilon, d)$ used to define $\mathrm{mdim}_{\mathrm{M}}(X, d)$ in Section 4,

$$S(X, \epsilon, d) = \lim_{n\to\infty} \frac{1}{n} \inf_{\mathrm{mesh}(\alpha, d_0^{n-1})<\epsilon} \log |\alpha|.$$

If $\mathrm{mdim}(X) > 0$, then for all metrics d on X, we know that $S(X, \epsilon, d)$ is eventually bigger than $(\mathrm{mdim}(X) - o(1))|\log \epsilon|$. One might ask whether lower rates of increase are possible. For extensions of minimal systems, there is a clear dichotomy:

Corollary 6.15. *Suppose (X, T) is an extension of a minimal system. If $\mathrm{mdim}(X) = 0$, then for any monotone function $\phi\colon (0, 1) \to \mathbb{R}^+$ with $\phi(\epsilon) \to \infty$ as $\epsilon \to 0$, there is a metric d such that*

$$\lim_{\epsilon\to 0} \frac{S(X, \epsilon, d)}{\phi(\epsilon)} = 0$$

If $\mathrm{mdim}(X) > 0$, then for all metrics d

$$\underline{\lim}_{\epsilon\to 0} \frac{S(X, \epsilon, d)}{|\log \epsilon|} \geq \mathrm{mdim}(X)$$

Proof:

We only need to prove the first part, as the second one is a restatement of the fact that $\mathrm{mdim}_{\mathrm{M}}(X, d) \geq \mathrm{mdim}(X)$ for all metrics d.

If $\mathrm{mdim}(X) = 0$, then we know that $X = \varprojlim_{i\to\infty} X_i$ with $h_{\mathrm{top}}(X_i) < \infty$ for every i. By definition of inverse limits, we have factor maps $\pi(i)\colon X \to X_i$ and, for $i > j$, $\pi(i, j)\colon X_i \to X_j$. Let $\tilde{d}_{(i)}$ be a metric on X_i. $\tilde{d}_{(i)}(\pi(i)x, \pi(i)y)$ is a semi-metric on X, which we will denote by $d_{(i)}$. We can assume that $d_{(i)}(x, y) \leq 1$ for every x, y and i. Set $a_1 = 1$, and chose inductively, for $k > 1$, $a_k < a_{k-1}/2$ such that

$$\phi(4a_k) > k h_{\mathrm{top}}(X_k) \tag{6.4}$$

We now take D to be the metric

$$D(x, y) = \sum_{k=1}^{\infty} a_k d_{(k)}(x, y).$$

on X. We will denote by $D_{(k)}$ the semi-metric

$$D_{(k)}(x,y)=\sum_{i=1}^{k} a_i d_{(i)}(x,y).$$

This semi-metric $D_{(k)}$ can be identified with the metric $\tilde{D}_{(k)}$ on X_k defined by

$$\tilde{D}_{(k)}(x,y)=\sum_{i=1}^{k} a_i \tilde{d}_{(i)}(\pi(k,i)x,\pi(k,i)y).$$

Let $\epsilon>0$ be given. We claim that

$$S(X,\epsilon,d)\le h_{\text{top}}(X_k)$$

where k is the smallest integer such that

$$\sum_{i=k+1}^{\infty} a_i<\epsilon/2. \tag{6.5}$$

Indeed, any open cover α of X with $\text{mesh}(\alpha, D_{(k)}\big|_0^{n-1})<\epsilon/2$, satisfies $\text{mesh}(\alpha, D\big|_0^{n-1})<\epsilon$. Thus

$$S(X,\epsilon,d)\le S(X_k,\epsilon/2,\tilde{D}_{(k)})\le h_{\text{top}}(X_k).$$

Since k is the smallest integer such that inequality (6.5) holds,

$$2a_k\ge\sum_{i=k}^{\infty} a_i\ge\epsilon/2$$

hence $\phi(4a_k)\le\phi(\epsilon)$. Using (6.4), we conclude that $kh_{\text{top}}(X_k)\le\phi(\epsilon)$, or

$$\frac{S(X,\epsilon,d)}{\phi(\epsilon)}\le\frac{h_{\text{top}}(X_k)}{\phi(\epsilon)}\le\frac{1}{k}\longrightarrow 0 \qquad \text{as } \epsilon\to 0.$$

□

References

[1] J. Auslander, *Minimal flows and their extensions*, Amsterdam, North-Holland (1988).

[2] W. Hurewicz and H. Wallman, *Dimension Theory*, Princeton University Press (1941).

[3] A. Jaworski, University of Maryland Ph.D. thesis, (1974).

[4] S. Kakutani, *A proof of Bebutov's theorem*, J. of Differential Eq. **4** (1968), pp 194–201.

[5] E. Lindenstrauss, *Lowering Topological Entropy*, J. d'Analyse Math. **67** (1995), pp 231–267.

[6] E. Lindenstrauss and B. Weiss, *On Mean Dimension*, to appear in Israel J. of Math.

[7] M. Shub and B. Weiss, *Can one always lower topological entropy?*, Ergod. Th. & Dynam. Sys.**11** (1991), pp 535–546.

Institute of Mathematics, The Hebrew University, Jerusalem 91904, Israel

E-mail address: elon@math.huji.ac.il