

# Zero-dimensional Principal Extensions

Tomasz Downarowicz\*

Dawid Huczek\*

## Abstract

We prove that every topological dynamical system  $(X, T)$  has a zero-dimensional *principal extension*, i.e. a zero-dimensional extension  $(Y, S)$  such that for every  $S$ -invariant measure  $\nu$  on  $Y$  the conditional entropy  $h(\nu|X)$  is zero. This reduces the discussion of many entropy-related properties to the zero-dimensional case which gives access to the various useful tools of symbolic dynamics.

## 1 Introduction

All the terms used in the introduction are standard, nonetheless they are explained in the Preliminaries. Given an abstract dynamical system  $(X, T)$  in from of the action of the iterates of a continuous transformation  $T$  on a compact metric space  $X$ , it is of interest to replace it by another, more familiar system, easier to describe and handle. One of the possibilities is to lift (i.e. extend) the system to some  $(Y, S)$ , whose phase space  $Y$  is zero-dimensional. Any zero-dimensional system admits a pleasant symbolic array form in which every point is represented as an infinite array filled with discrete symbols. This representation allows one to apply many methods of symbolic dynamics, for instance to calculate the topological entropy or the entropies of the invariant measures, and, in fact, many other invariants. In case we find such a system  $(Y, S)$ , we say, that we have found a *zero-dimensional extension*. It is an elementary exercise to show that every system has a zero-dimensional extension. On the other hand, we are naturally interested in minimizing the “distance” between the original system  $(X, T)$  and its extension  $(Y, S)$ . There are at least two levels at which this minimization can be performed:

---

\*Research of both authors is supported from resources for science in years 2009-2012 as research project (grant MENII N N201 394537, Poland)

1. We may be interested in minimizing, for each invariant measure  $\mu$  on  $X$ , the increase of entropy as we lift that measure to an invariant measure  $\nu$  on  $Y$ .
2. We may want that for each  $\mu$  there exist a unique lift  $\nu$  and that  $(Y, \nu, S)$  is measure-theoretically isomorphic to  $(X, \mu, T)$ .

The increase of entropy specified in 1. is measured by the conditional dynamical entropy  $h(\nu|X)$ . In case  $\mu$  has finite dynamical entropy, this conditional entropy simply equals the difference  $h(\nu) - h(\mu)$ , but the conditional notation is universal. The best one can get in the category 1. is a *principal extension*, i.e., such that  $h(\nu|X) = 0$  for every invariant measure  $\nu$  on  $Y$ . An extension which satisfies the postulate 2. will be call here an *isomorphic extension*. Clearly, an isomorphic extension is automatically principal.

For invertible systems (i.e., such that  $T$  is a homeomorphism) with finite topological entropy and satisfying an additional assumption that it has a nonperiodic minimal factor, the existence of zero-dimensional isomorphic (hence principal) extensions has already been established as a consequence of the deep results in mean dimension theory developed by E. Lindenstrauss and B. Weiss ([L-W] and [Li]). Every such system satisfies the so-called “small boundary property”, which allows to rather easily construct its zero-dimensional isomorphic extension. If the assumption about the existence of a minimal factor is dropped, the above theory still allows to easily build a zero-dimensional principal extension: It is elementary to see that the direct product with any system of zero topological entropy is a principal extension and that the composition of two principal extensions is a principal extension. Thus, an arbitrary system of finite topological entropy is first extended to its direct product with some infinite minimal system of zero topological entropy (for example an irrational rotation of the circle, or an odometer) and since such a product already has a minimal nonperiodic factor, it can be isomorphically extended to a zero-dimensional system. The resulting extension is no longer isomorphic, but it is isomorphic to the intermediate product system.

The existence of an isomorphic zero-dimensional extension is in fact equivalent to the small boundary property, and E. Lindenstrauss provided examples showing that both assumptions (finite entropy and minimal factor) are essential. Thus, for systems with infinite entropy, even those which admit a minimal nonperiodic factor, we cannot hope to prove the existence of an isomorphic zero-dimensional extension.

We can, however, still hope to prove the existence of a principal zero-dimensional extension, but using a different method, not appealing to mean dimension or small boundary property. This is exactly what is done in this paper: we prove that every topological dynam-

ical system  $(X, T)$  has a principal zero-dimensional extension. The theorem presented herein has the twofold advantage over the same obtained via the mean dimension theory: of making no restrictions on the entropy of the dynamical system and of being achieved by more elementary methods.

Historically, the term *principal* was probably first used by F. Ledrappier in [Le]. The construction of a principal extension for systems of finite topological entropy via the mean dimension theory was heavily exploited in [B-D] in the theory of symbolic extensions. The detailed description of the passage from the small boundary property to the principal extension can be found in [D] (in earlier papers it is considered more or less obvious and left to the reader).

## 2 Preliminaries

### 2.1 Basic notions

**Dynamical systems.** Throughout this work a *dynamical system* will be a triple  $(X, T, \mathbb{S})$ , where  $X$  is a compact metric space with metric  $d$ ,  $T$  is a continuous map on  $X$  (invertible or not) and  $\mathbb{S} \in \{\mathbb{Z}, \mathbb{Z}_+ \cup \{0\}\}$  is the index set, depending on whether we consider both the negative and positive iterates of  $T$  (i.e. the action on  $X$  of  $\mathbb{Z}$ ) or only positive ones (i.e. the action on  $X$  of  $\mathbb{Z}_+ \cup \{0\}$ ) — our final result applies in both cases, but a few details of the proof differ, so we need to be able to make the distinction. For brevity, where the index set or transformation are obvious or irrelevant, we will omit them.

**Factors, extensions and conjugacies.** A dynamical system  $(X, T, \mathbb{S})$  is a *factor* of the system  $(Y, S, \mathbb{S})$  if there exists a continuous map  $\pi$  from  $Y$  onto  $X$  such that  $T \circ \pi = \pi \circ S$ . In this situation we also say that  $Y$  is an *extension* of  $X$ . If the map  $\pi$  is a homeomorphism, we say that  $X$  and  $Y$  are *conjugate*.

**Odometers.** Let  $\{p_k\}_{k \geq 1}$  be a sequence of positive integers such that  $p_{k+1}$  is a multiple of  $p_k$ . An *odometer to base  $p_k$*  is defined as the inverse limit of the sequence  $\mathbb{Z}_{p_k}$ , i.e. as the subset of the product  $\prod_{k=1}^{\infty} \mathbb{Z}_{p_k}$  (with the product topology) consisting of sequences  $\{g_k\}$  such that  $g_k = g_{k+1} \bmod p_k$ . With coordinate-wise addition  $G$  is a compact, metrizable, zero-dimensional group and with the action  $T(g_1, g_2, \dots) = (g_1 + 1, g_2 + 1, \dots)$  it becomes a zero-dimensional dynamical system. Define  $G_k$  as the set of points  $g \in G$  such that  $g_j = 0$  for  $j \leq k$ . Observe that  $G$  is a disjoint union  $G_k \dot{\cup} TG_k \dot{\cup} \dots \dot{\cup} T^{p_k-1}G_k$  and that  $G_{k+1} \subset G_k$ .  $\{G_k\}$  is a base of the topology at zero.

## 2.2 Zero-dimensional dynamical systems

A dynamical system  $(Y, S, \mathbb{S})$  is called *zero-dimensional* if  $Y$  is a zero-dimensional space, i.e. if it has a base consisting of sets which are both closed and open. A particularly important class of such systems are symbolic systems over an uncountable alphabet which we will call *array systems* and which are constructed as follows: Let  $\Lambda_k$  be a finite set with discrete topology and let  $Y_c = \prod_{k=1}^{\infty} \Lambda_k^{\mathbb{S}}$ . The points of  $Y_c$  can be thought of as arrays (infinite in at least two directions)  $\{y_{k,n}\}$  where  $y_{k,n} \in \Lambda_k$ . With the action of the horizontal shift  $S$  (i.e.  $(Sy)_{k,n} = y_{k,n+1}$ )  $Y_c$  becomes a zero-dimensional dynamical system. An array system is any closed, shift-invariant subset  $Y$  of such  $Y_c$ . Our main object of interests will be *marked* array systems, i.e. ones for which there exists a descending sequence of clopen sets  $F_k$  that are unions of cylinders associated to the symbols in  $\Lambda_k$  (which means that we can determine whether an array  $y$  belongs to  $F_k$  by viewing the symbol  $y_{k,0}$ ), and a sequence  $\{p_k\}$  such that the every point of  $Y$  visits  $F_k$  periodically with period  $p_k$  (this implies that  $p_{k+1}$  is a multiple of  $p_k$ ). This is equivalent to requiring that  $Y$  factors onto an odometer  $G$  to the base  $p_k$ . The points of such marked systems can be viewed as arrays with vertical lines inserted at regular intervals: If there is a marker at some position then there is one in the same column in every preceding row - see fig. 1.

By a *k-rectangle* we will mean a finite matrix  $C = (C_{j,n})$ ,  $j = 1, \dots, k$ ;  $n = 0, \dots, p_k - 1$  that has markers in all rows in column 0. In the array representation such as the one in Figure 1 a *k-rectangle* is a rectangle occurring in rows 1 through  $k$  of such an array, between two vertical lines. Every *k-rectangle*  $C$  can be identified with a cylinder in  $Y$  in the standard way:  $y \in C$  iff  $y_{j,n} = C_{j,n}$ ,  $j = 1, \dots, k$ ;  $n = 0, \dots, p_k - 1$ . We will denote the set of all *k-rectangles* by  $\mathcal{C}_k$ .

...	01	11	<b>00</b>	<b>01</b>	<b>00</b>	<b>01</b>	<b>10</b>	<b>11</b>	10	11	01	1...
...	12	01	<b>20</b>	<b>11</b>	<b>20</b>	<b>02</b>	<b>12</b>	<b>10</b>	12	00	20	1...
...	01	30	<b>21</b>	<b>31</b>	<b>20</b>	<b>10</b>	<b>12</b>	<b>03</b>	21	30	31	2...
...												

Figure 1: An element of a marked array system factoring onto an odometer to the base  $(2, 6, 12, \dots)$ . The boldface numbers form a 3-rectangle.

## 2.3 Invariant measures

Denote the set of all probability measures on the sigma-algebra of Borel subsets of  $X$  by  $\mathcal{M}(X)$ . It is a compact, metrizable, convex set. Let  $\mathcal{M}_T(X)$  be the set of all  $T$ -invariant probability measures on  $X$ , i.e.  $\mu$  is an element of  $\mathcal{M}_T(X)$  if and only if for any Borel set  $B$  we have  $\mu(T^{-1}(B)) = \mu(B)$ .  $\mathcal{M}_T(X)$  is a nonempty, closed and convex subset of  $\mathcal{M}(X)$ . The extreme points of  $\mathcal{M}_T(X)$  are ergodic measures, i.e. the ones for which any invariant set has measure either 0 or 1.

For any  $x \in X$  let  $\delta_x$  denote the point mass at  $x$  and let

$$M_n(\delta_x) = \frac{1}{n} (\delta_x + \delta_{Tx} + \dots + \delta_{T^{n-1}x}).$$

We will later need the following two facts, both of which are fairly obvious:

**Fact 2.1.** *For any  $n$  the set  $\mathcal{M}_T(X)$  is within the closure of the convex hull of the set  $\{M_n(\delta_x); x \in X\}$ .*

**Fact 2.2.** *Let  $U$  be an open subset of  $\mathcal{M}(X)$  containing  $\mathcal{M}_T(X)$ . There exists an  $N$  such that for any  $n > N$  and any  $x \in X$  the measure  $M_n(\delta_x)$  is in  $U$ .*

We will be particularly interested in invariant measures on marked array systems. Let  $Y$  be a marked array system and let  $d^*$  be a metric on  $\mathcal{M}(Y)$  compatible with the weak\*-topology. It is a standard fact in zero-dimensional dynamics that we can assess the proximity of measures by investigating  $k$ -rectangles alone:

**Fact 2.3.** *Let  $\mu, \nu \in \mathcal{M}_T(Y)$ . For any  $\varepsilon > 0$  there exist  $\delta > 0$  and  $k > 0$  such that if*

$$|\mu(C) - \nu(C)| < \delta$$

*for all  $k$ -rectangles  $C$ , then  $d^*(\mu, \nu) < \varepsilon$ .*

## 2.4 Entropy

We recall the basic definitions and facts of the entropy theory of dynamical systems. Let  $(X, T)$  be a dynamical system and let  $\mu \in \mathcal{M}_T(X)$ . For any finite partition  $\mathcal{A}$  of  $X$  into measurable sets we define the entropy of a partition as

$$H(\mu, \mathcal{A}) = - \sum_{A \in \mathcal{A}} \mu(A) \ln \mu(A).$$

$$H_n(\mu, \mathcal{A}) = \frac{1}{n} H(\mu, \mathcal{A}^n),$$

where  $\mathcal{A}^n = \bigvee_{j=0}^{n-1} T^{-j}(\mathcal{A})$ . The sequence  $H_n$  is known to converge to its infimum, which allows one to define

$$h(\mu, \mathcal{A}) = \lim H_n(\mu, \mathcal{A}).$$

Finally the *entropy of a measure* is given as

$$h(\mu) = \sup_{\mathcal{A}} h(\mu, \mathcal{A}),$$

where the supremum is taken over all finite partitions of  $X$ .

If  $\mathcal{A}$  and  $\mathcal{B}$  are two finite partitions of  $X$ , then we can define the conditional entropy of  $\mathcal{A}$  with respect to  $\mathcal{B}$  as

$$H(\mu, \mathcal{A}|\mathcal{B}) = \sum_{B \in \mathcal{B}} \mu(B) H(\mu_B, \mathcal{A}),$$

where  $\mu_B(A) = \mu(A|B)$ . Then we can proceed to define

$$H_n(\mu, \mathcal{A}|\mathcal{B}) = \frac{1}{n} H(\mu, \mathcal{A}^n | \mathcal{B}^n).$$

Again, the sequence  $H_n$  is known to converge to its infimum, which allows one to define

$$h(\mu, \mathcal{A}|\mathcal{B}) = \lim H_n(\mu, \mathcal{A}|\mathcal{B}).$$

Suppose now that we have a system  $(Y, S)$  that is an extension of  $(X, T)$  by a map  $\pi$ . Let  $\nu$  be an invariant measure on  $Y$ . Any partition  $\mathcal{B}$  of  $X$  can be lifted to a partition  $\pi^{-1}\mathcal{B}$  of  $Y$ . Define

$$h(\nu, \mathcal{A}|X) = \inf_{\mathcal{B}} h(\nu, \mathcal{A}|\pi^{-1}\mathcal{B}),$$

where the infimum is taken over all finite partitions of  $X$ . Finally, define

$$h(\nu|X) = \sup_{\mathcal{A}} h(\nu, \mathcal{A}|X),$$

where once again the supremum is taken over all finite partitions of  $Y$ . If the image  $\pi\nu$  of  $\nu$  by the factor map  $\pi$  has finite entropy, then it is not difficult to see that  $h(\nu|X) = h(\nu) - h(\pi\nu)$ .

We will make use of the following fact.

**Fact 2.4.** *Let  $Y$  be an array system and let  $X$  be a factor of  $Y$ . Let  $\mathcal{R}_k$  be the partition defined by cylinders of height  $k$  and length 1. Then for any measure  $\nu$  on  $Y$  we have  $h(\nu|X) = \lim_k h(\nu, \mathcal{R}_k|X)$ .*

To see that it is so, it suffices to observe two facts. Firstly, that the family  $\{\mathcal{R}_k\}$  together with its images under  $S$  generates the Borel  $\sigma$ -algebra on  $Y$ . Secondly, if  $j < k$  then  $\mathcal{R}_j \prec \mathcal{R}_k$ , and therefore  $h(\nu, \mathcal{R}_j|X) < h(\nu, \mathcal{R}_k|X)$ .

**Definition 2.5.** *Suppose a dynamical system  $(Y, S)$  is an extension of the system  $(X, T)$  via the map  $\pi$ .  $(Y, S)$  is a principal extension if  $h(\nu|X) = 0$  for every  $\nu \in \mathcal{M}_S(Y)$ .*

If  $(X, T)$  has finite topological entropy, then by the variational principle  $\pi\nu$  has finite entropy for each  $\nu$  in  $\mathcal{M}_S(Y)$  so the extension is principal if and only if  $h(\nu) = h(\pi\nu)$  for each  $\nu$ . In particular  $(Y, S)$  has the same topological entropy as  $(X, T)$ .

## 2.5 Continuity of the entropy functions

In the main proof we will consider entropy as a function of the measure, and we will need several basic facts about the continuity of this function which we state without proof. First of all:

**Fact 2.6.** *The function  $\mu \mapsto \mu(A)$  is upper semicontinuous if  $A$  is closed and lower semicontinuous if  $A$  is open.*

Since  $\mu(\text{Int}(A)) \leq \mu(A) \leq \mu(\overline{A})$  and the three are equal if the boundary of  $A$  has measure 0, we have the following:

**Fact 2.7.** *The function  $\mu \mapsto \mu(A)$  is continuous at every  $\mu$  such that  $\mu(\partial A) = 0$ .*

Using the fact that the limit defining  $h(\mu, \mathcal{A}|\mathcal{B})$  is also the infimum, we easily arrive at the following:

**Fact 2.8.** *The function  $\mu \mapsto h(\mu, \mathcal{A}|\mathcal{B})$  is upper semicontinuous at every  $\mu$  such that  $\mu(\partial A) = 0$  for every  $A \in \mathcal{A}$  and  $\mu(\pi^{-1}(\partial B)) = 0$  for every  $B \in \mathcal{B}$ .*

Finally:

**Fact 2.9.** *The function  $\mu \mapsto h(\mu, \mathcal{A}|X)$  is upper semicontinuous at every  $\mu$  such that  $\mu(\partial A) = 0$  for every  $A \in \mathcal{A}$ .*

To observe that, note that  $h(\mu, \mathcal{A}|X)$  is the infimum of any sequence  $h(\mu, \mathcal{A}|\mathcal{B}_n)$ , provided that the diameter of the largest set in  $\mathcal{B}_n$  tends to 0. Since we can construct partitions into sets of arbitrarily small diameter that all have boundaries whose measure  $\mu$  is 0, Fact 2.9 now follows.

## 3 The main result

**Theorem 3.1.** *Every topological dynamical system has a zero-dimensional principal extension.*

The proof of theorem 3.1 will occupy the remainder of this section.

### 3.1 Preliminary reshaping

Let  $(X_0, T, \mathbb{S})$  be the topological dynamical system for which we shall be constructing the zero-dimensional principal extension. Let  $G$  be an odometer to some base  $(p_k)$ . What we will actually construct is a zero-dimensional principal extension  $(Y, S)$  of the system  $X = X_0 \times G$  with the product action. Since the odometer has a unique invariant measure and its entropy is zero, for every measure  $\mu$  on  $X$  we have  $h(\mu|X_0) = 0$ ,  $Y$  will be a principal extension of  $X_0$  as well (we are using the fact that  $h(\nu|X_0) \leq h(\nu|X) + h(\pi\nu|X_0)$ ).

Let  $I$  denote the one-dimensional torus, i.e. the interval  $[0, 1]$  with the endpoints identified, and let  $\lambda$  be the Lebesgue measure on  $I$ . Any function  $f : X \rightarrow [0, 1]$  induces a partition  $\mathcal{A}_f$  of  $X \times I$  into two sets:  $\{(x, t) : 0 \leq t < f(x)\}$  and  $\{(x, t) : f(x) \leq t < 1\}$  (i.e. the sets of points below and above the graph of  $f$ ). For a family  $\mathcal{F}$  of functions we denote by  $\mathcal{A}_{\mathcal{F}}$  the partition  $\bigvee_{f \in \mathcal{F}} \mathcal{A}_f$ . Two useful observations are that  $\mathcal{A}_{\mathcal{F} \cup \mathcal{G}} = \mathcal{A}_{\mathcal{F}} \vee \mathcal{A}_{\mathcal{G}}$  and that  $\mathcal{F} \subset \mathcal{G}$  implies  $\mathcal{A}_{\mathcal{F}} \prec \mathcal{A}_{\mathcal{G}}$ .

### 3.2 Constructing the zero-dimensional principal extension

Fix a sequence of positive numbers  $\varepsilon_k$  decreasing to 0. Let  $\{\mathcal{F}_k\}$  be a sequence of families of continuous functions from  $X$  into  $[0, 1]$  such that  $\mathcal{F}_k \subset \mathcal{F}_{k-1}$ . Let  $\eta_k$  be the diameter of the largest set in  $\mathcal{A}_{\mathcal{F}_k}$  (in the product metric on  $X \times I$ ). We will require that  $\eta_k$  tend to 0 and that  $\eta_k$  be so small that  $d(x_1, x_2) < \eta_k$  in  $X$  implies that  $x_1$  and  $x_2$  belong to the same element of the family of clopen sets  $\{T^i(X_0 \times G_k)\}$ . This ensures that every cell of  $\mathcal{A}_{\mathcal{F}_k}$  is completely contained in one of the sets  $\{T^i(X_0 \times G_k) \times I\}$ .

Let  $\pi^{(1)}$  denote the projection of  $X \times I$  onto  $X$ . Consider the space of all formal arrays  $y = y_{k,n}$  ( $k > 0, n \in \mathbb{S}$ ), such that  $y_{k,n} \in \mathcal{A}_{\mathcal{F}_k}$  (we treat each finite partition  $\mathcal{A}_{\mathcal{F}_k}$  as the alphabet in row  $k$ ). For an array  $y$  define the sets  $K_{k,n}(y) = \{x \in X : d(x, \pi^{(1)}(y_{k,n})) \leq 2\eta_k\}$  ( $K_{k,n}(y)$  is the  $2\eta_k$ -neighborhood of the projection onto  $X$  of the cell of  $\mathcal{A}_{\mathcal{F}_k}$  appearing as a symbol in  $y$  at the position  $(k, n)$ ). An array  $y$  will be said to satisfy the *column condition* if for each  $n$  the sequence  $K_{k,n}(y)$  is descending (as  $k$  increases). Since the diameter of  $K_{k,n}$  tends to 0 with  $k$ , the column condition implies that the intersection  $\bigcap_{k=0}^{\infty} K_{k,n}(y)$  is a single point in  $X$  which we will denote by  $x_n(y)$ . Note that  $x_n(y)$  is within  $2\eta_k$  of each set  $\pi^{(1)}(y_{k,n})$  – a fact that will be useful later.

Let  $Y_c$  be the space of all arrays  $y$  satisfying the column condition with the additional requirement that  $x_{n+1}(y) = Tx_n(y)$ . Our choice



of  $\eta_k$  ensures that the elements of  $Y_c$  have the following two properties:

- In row  $k$  the symbols representing subsets of  $X_0 \times G_k$  occur every  $p_k$  positions.
- If  $y_{k,n}$  is a symbol representing a subset of  $X_0 \times G_k$ , then  $y_{k-1,n}$  represents a subset of  $X_0 \times G_{k-1}$ .

It is easy to see that  $Y_c$  is a marked array system (the markers in row  $k$  are the symbols representing subsets of  $X_0 \times G_k$ ) and that with the action of the horizontal shift  $S$  it forms a continuous extension of  $(X, T)$ , where the factor map is  $\pi_X(y) = x_0(y)$  (by which we mean  $x_n(y)$  for  $n = 0$ ).

We will construct  $Y$  as a subsystem of  $Y_c$  which will in some sense be the limit of an auxiliary sequence of (disjoint) mutually conjugate subsystems  $Y_k$ . We will define  $Y_k$  inductively, by constructing the maps  $\Phi_k : Y_{k-1} \rightarrow Y_k$  (these maps will in fact be block codes defined on rectangles of some order). The main goal will be to ensure that for any  $k$  and all  $j > k$  the set  $\mathcal{M}_S(Y_j)$  is contained within an open set  $\mathcal{U}_k \subset \mathcal{M}_S(Y_c)$ , where the sequence  $\{\mathcal{U}_k\}$  (which we will also define inductively) satisfies the following properties:

- U1.  $\mathcal{U}_{k+1} \subset \mathcal{U}_k$ .
- U2. For any  $k > 0$  and any measure  $\nu \in \mathcal{U}_k$  we have  $h_\nu(\mathcal{R}_k|X) < \varepsilon_k$ , where  $\mathcal{R}_k$  is the partition defined by cylinders of height  $k$  and length 1.

To begin with, let  $Y_0$  be the closure of the set of array-names of points in  $X \times I$  under the action of  $T \times \text{id}$  with respect to the partitions  $\mathcal{A}_{\mathcal{F}_k}$ . In other words,  $Y_0$  is the closure of the set of all points  $y \in Y_c$  such that for some pair  $(x, t) \in X \times I$  and for any  $k$  and  $n$  we have  $(T^n x, t) \in y_{k,n}$ . By a standard argument,  $Y_0$  is an extension of  $X \times I$  (we will denote the corresponding map by  $\pi_0$ ) as well as of  $X$  itself and the following diagram commutes:

$$\begin{array}{ccc}
 & & Y_0 \\
 & \swarrow \pi_0 & \downarrow \pi_X \\
 X \times I & & X \\
 & \searrow \pi^{(1)} & \\
 & & 
 \end{array}$$

Let the set  $\mathcal{U}_0$  be all of  $\mathcal{M}_S(Y_c)$  (all our requirements on the properties of  $\mathcal{U}_k$  only apply to the case  $k > 0$ ).

There are two important observations to be made here: Firstly, the only points in  $X \times I$  that have multiple preimages under  $\pi_0$  are

the ones whose orbits enter the graph of a function from some  $\mathcal{F}_k$  (we are using the fact that the graphs of continuous functions are closed). The product measure  $\mu \times \lambda$  of the graph of any function is 0, therefore if  $\nu$  is a measure on  $Y_0$  that factors onto a measure  $\mu \times \lambda$  on  $X \times I$  then the set of points in  $X \times I$  with multiple preimages by  $\pi_0$  has measure 0, and thus the measure-theoretic systems  $(Y, S, \nu)$  and  $(X \times I, T \times \text{id}, \mu \times \lambda)$  are isomorphic. Secondly, any  $k$ -rectangle in  $Y_0$  is associated with a unique cell of  $\mathcal{A}_{\mathcal{F}_k}^{p_k}$ , the closure of which is the image (by  $\pi_0$ ) of this rectangle. Such cells (which are exactly the cells of  $\mathcal{A}_{\mathcal{F}_k}^{p_k}$  that are contained in  $X_0 \times G_k$ ) will be called fundamental  $k$ -cells.

We will now proceed to create the systems  $Y_k$ , requiring them to have the following properties:

Y1. For each  $k$ ,  $\mathcal{M}_S(Y_k) \subset \mathcal{U}_k$ .

Y2. For each  $k$ ,  $Y_k = \Phi_k(Y_{k-1})$  and there exists an increasing sequence  $j_k$  such that  $\Phi_k$  preserves rows from  $j_k$  onwards.

Observe that the property (Y2) ensures that the diagram

$$\begin{array}{ccccccc}
 & & Y_0 & \xleftrightarrow{\Phi_0} & Y_1 & \xleftrightarrow{\Phi_1} & Y_2 & \xleftrightarrow{\Phi_2} & \cdots \\
 & \swarrow \pi_0 & \downarrow \pi_X & & \swarrow \pi_X & & \swarrow \pi_X & & \\
 X \times I & & & & & & & & \\
 & \searrow \pi^{(1)} & \downarrow & & \downarrow & & \downarrow & & \\
 & & X & & & & & & 
 \end{array}$$

commutes and that for  $j \geq j_k$  we still have the one-to-one correspondence between the  $j$ -rectangles of  $Y_k$  and the cells of  $\mathcal{A}_{\mathcal{F}_j}^{p_j}$ , since this correspondence depends only on the contents of row  $j$ . Throughout,  $\pi_k$  will denote the factor map of  $Y_k$  onto  $X \times I$  defined by composing the factorization  $\pi_0$  of  $Y_0$  with the conjugacy between  $Y_0$  and  $Y_k$ .

Now, suppose we have defined the system  $Y_{k-1}$  and the set  $\mathcal{U}_{k-1}$ . Our task is to create the set  $\mathcal{U}_k$  satisfying the requirements (U1)-(U2) and a system  $Y_k$  such that  $\mathcal{M}_S(Y_k) \subset \mathcal{U}_k$ . Let  $\mathcal{M}_{k-1}$  be the set of all measures on  $Y_{k-1}$  that factor by  $\pi_{k-1}$  onto measures of the form  $\mu \times \lambda$  on  $X \times I$ . As stated above, if  $\nu \in \mathcal{M}_{k-1}$  and  $\nu$  factors onto  $\mu \times \lambda$ , then  $(Y, \nu)$  and  $(X \times I, \mu \times \lambda)$  are measure-theoretically isomorphic. It follows that for any  $\nu$  in  $\mathcal{M}_{k-1}$  we have

$$h(\nu|X) = h(\mu \times \lambda|X) = 0,$$

so in particular  $h_\nu(\mathcal{R}_k|X) = 0$ . As we have noted earlier,  $h_\nu(\mathcal{R}_k|X)$  is upper semicontinuous at  $\nu$ , provided  $\nu(\partial R) = 0$  for every  $R \in$

$\mathcal{R}_k$ . This is the case for any  $\nu$  since cylinders have no boundaries at all. Therefore every measure in  $\mathcal{M}_{k-1}$  has a neighborhood where  $h_\nu(\mathcal{R}_k|X) < \varepsilon_k$ .  $\mathcal{M}_{k-1}$  is compact and convex, so by choosing a finite number of such neighborhoods covering  $\mathcal{M}_{k-1}$  we can simply assume that there exists some convex neighborhood  $\mathcal{V}_k$  of  $\mathcal{M}_{k-1}$  such that for any measure  $\nu \in \mathcal{V}_k$  we have  $h_\nu(\mathcal{R}_k|X) < \varepsilon_k$ .

It is clear that the set  $\mathcal{U}_k = \mathcal{U}_{k-1} \cap \mathcal{V}_k$  has the properties (U1)-(U2). We must now construct the system  $Y_k$  whose invariant measures are all in  $\mathcal{U}_k$ , also making sure to satisfy the requirement (Y2). In other words, we must ensure that every measure on  $Y_k$  must be close (in the space  $\mathcal{M}_S(Y_c)$ ) to some measure on  $Y_{k-1}$  that factors onto a product measure on  $X \times I$ . To this end we will employ the following lemma:

**Lemma 3.2.** *For any measure  $\mu$  on  $X$  and any neighborhood  $U$  of  $\mu \times \lambda$  in  $\mathcal{M}(X \times I)$  there exists a neighborhood  $U_\mu$  of  $\mu$  in  $\mathcal{M}(X)$ , an irrational rotation  $R_\mu$  of the one-dimensional torus and a number  $N_\mu$  such that for any  $(x, t) \in X \times I$  and any  $n > N$  the condition  $M_n(\delta_x) \in U_\mu$  implies that  $M_n(\delta_{(x,t)}) \in U$ , where the averaging in the product is with respect to the map  $T \times R_\mu$ .*

**Proof:** First note that for any measure  $\mu$  in  $\mathcal{M}_T(X)$  there exists an irrational rotation  $R_\mu$  disjoint from  $\mu$  (i.e. the only  $(T \times R_\mu)$ -invariant measure on  $X \times I$  with marginals  $\mu$  and  $\lambda$  is  $\mu \times \lambda$ ). Indeed, if  $\mu$  is an ergodic measure and  $e^{\alpha 2\pi i}$  is not its eigenvalue, then the rotation of the circle by  $\alpha$  is disjoint from  $\mu$ . Since an ergodic measure has at most countably many eigenvalues, for any ergodic  $\mu$  there exist at most countably many rotations that are not disjoint from  $\mu$ . If  $\mu$  is not ergodic, denote its ergodic decomposition by  $\xi$  ( $\xi$  is a measure on the set  $\mathcal{M}_T^e(X)$  of ergodic measures on  $X$ ) and consider the product  $\mathcal{M}_T^e(X) \times I$  with the measure  $\xi \times \lambda$ . The set  $\{(\nu, \alpha) : e^{\alpha 2\pi i} \text{ is an eigenvalue of } \nu\}$  is a measurable subset of the product and has measure 0 (because all its vertical sections are countable), so almost every horizontal section of this set has measure 0. Therefore there exists an  $\alpha$  such that the measures for which  $e^{\alpha 2\pi i}$  is an eigenvalue have zero mass in the ergodic decomposition of  $\mu$ . Setting  $R_\mu$  to be the rotation by  $\alpha$  we obtain a rotation disjoint from  $\mu$ .

Suppose the statement of the lemma is not true. Then there exists a sequence of measures  $M_n(\delta_{(x_n, t_n)})$  such that  $M_n(\delta_{x_n})$  converge to  $\mu$  yet the  $M_n(\delta_{(x_n, t_n)})$  all lie outside  $U$  (remember that the averaging in  $X \times I$  is with respect to  $T \times R_\mu$ ). Choose the limit  $\nu$  of some subsequence of  $M_n(\delta_{(x_n, t_n)})$ . It is a  $T \times R_\mu$ -invariant measure which is outside  $U$  and whose marginals are  $\mu$  (being the limit of  $M_n(\delta_{x_n})$ ) and

$\lambda$  (being the only  $R_\mu$ -invariant measure on  $I$ ). But the only  $T \times R_\mu$ -invariant measure with marginals  $\mu$  and  $\lambda$  is  $\mu \times \lambda$ , which is in  $U$  – a contradiction. ■

As the open set  $\mathcal{U}_k$  contains the closed convex set  $\mathcal{M}_{k-1}$ , there exists some number  $\varepsilon$  such that the convex  $\varepsilon$ -neighborhood of  $\mathcal{M}_{k-1}$  is contained in  $\mathcal{U}_k$ . From Fact 2.3 we obtain numbers  $\delta$  and  $j$  such that if two measures differ by no more than  $2\delta$  on all  $j$ -rectangles, then the distance between them is less than  $\varepsilon$ . We can assume that  $j > j_{k-1}$  (increasing  $j$  if necessary). For any measure  $\mu$  on  $X$  the partition  $\mathcal{A}_{\mathcal{F}_j}^{p_j}$  has boundaries of measure  $\mu \times \lambda$  equal to 0. Therefore there exists a neighborhood  $U$  of  $\mu \times \lambda$  in  $\mathcal{M}(X \times I)$  such that if  $\nu \in U$  then  $|\nu(A) - \mu \times \lambda(A)| < \delta$  for every  $A \in \mathcal{A}_{\mathcal{F}_j}^{p_j}$ . Applying lemma 3.2 to  $U$ , we obtain for every  $\mu \in \mathcal{M}_T(X)$  an open set  $U_\mu$  around  $\mu$  in  $\mathcal{M}(X)$ . Out of these we select a finite family  $\mathcal{W}$  of measures such that the union of  $U_\mu$  for  $\mu \in \mathcal{W}$  covers  $\mathcal{M}_T(X)$  in  $\mathcal{M}(X)$ . The union of this cover is an open set in  $\mathcal{M}(X)$ . There exists a  $j_k > j$  for which  $p_{j_k}$  is so large that every measure of the form  $M_{p_{j_k}}(\delta_x)$  is in  $U_\mu$  for some  $\mu \in \mathcal{W}$  (we are using Fact 2.2). We can also assume that  $p_{j_k}$  is larger than the number  $N_\mu$  of Lemma 3.2 for all  $\mu \in \mathcal{W}$ .

Now, let  $C$  be any  $j_k$ -rectangle in  $Y_{k-1}$ . Choose a pair  $(x_C, t_C)$  from the interior of the fundamental  $j_k$ -cell whose closure is  $\pi_{k-1}(C)$  (the cell is guaranteed to have nonempty interior, because the partition is generated by graphs of continuous functions) and a measure  $\mu_C \in \mathcal{W}$  such that  $M_{p_{j_k}}(\delta_{x_C})$  belongs to  $U_{\mu_C}$ . Recall that the fundamental  $j$ -cells partition the set  $(X_0 \times G_j) \times I$ , which is invariant under the map  $(T \times R_{\mu_C})^{p_j}$ . Thus  $(x_C, t_C)$  has a name under the action of  $(T \times R_{\mu_C})^{p_j}$  on  $(X_0 \times G_j) \times I$  with respect to the partition into the fundamental  $j$ -cells. Take the initial block of length  $q = \frac{p_{j_k}}{p_j}$  of this name. It is an ordered list of  $q$  fundamental  $j$ -cells, each associated with a unique  $j$ -rectangle in  $Y_{k-1}$ , so we have a sequence of  $j$ -rectangles  $D_1, \dots, D_q$ . Observe that the number of times at which a  $j$ -rectangle  $D$  occurs in this list equals the number of times at which  $(x_C, t_C)$  visits  $\pi_{k-1}(D)$  under the action of  $(T \times R_{\mu_C})^{p_j}$ . This number equals  $p_{j_k} M_{p_{j_k}}(\delta_{(x_C, t)}) (\pi_{k-1}(D))$ . By Lemma 3.2,

$$\left| M_{p_{j_k}}(\delta_{(x_C, t)}) (\pi_{k-1}(D)) - (\mu_C \times \lambda)(\pi_{k-1}(D)) \right| < \delta,$$

therefore the number of occurrences of  $D$  equals  $p_{j_k}(\mu_C \times \lambda)(\pi_{k-1}(D)) \pm p_{j_k} \delta$ , a fact that we will use later.

We now define the image  $C$  under  $\Phi_k$  as follows: In rows 1 through  $j$  it has the ordered list of  $j$ -rectangles  $D_1, \dots, D_q$  described above and in rows  $j + 1$  through  $j_k$  it retains the contents of  $C$ . Observe that  $\Phi_k(C)$  satisfies the column condition (the column condition concerns

pairs of symbols in one column, so it makes sense for rectangles as well as arrays). To verify this, we only need to check that every set corresponding to a symbol in row  $j + 1$  of  $\Phi_k(C)$  is contained within  $\eta_j$  of the corresponding set above it in row  $j$ ; any other pair of symbols (one above another) appears already in  $Y_{k-1}$ , so it satisfies the column condition by the inductive assumption. For each  $n \leq p_{j_k}$ , the symbols  $C_{j,n}$  and  $C_{j+1,n}$  treated as sets in the product space  $X \times I$  contain the image of  $(x_C, t_C)$  under the composition of  $n$  transformations, each being a product of  $T$  with some rotation or identity. Thus, their projections onto  $X$  both contain  $T^n(x_C)$ . Since the intersection of these projections is nonempty and the projection  $\pi^{(1)}(C_{j+1,n})$  has diameter smaller than  $\eta_{j+1}$ , its  $2\eta_{j+1}$ -neighborhood must be entirely contained in the  $2\eta_j$ -neighborhood of  $\pi^{(1)}(C_{j,n})$ .

The idea of the construction of the code  $\Phi_k$  is shown on the figure below. The bases of the three large rectangles are the sets  $X_0 \times G_j, T(X_0 \times G_j), \dots, T^{p_j-1}(X_0 \times G_j)$ . For simplicity, we imagine the transformation  $T$  as the rigid translation between these sets, except on the last one, which is mapped somehow to  $X_0 \times G_j$ . The large rectangles are Cartesian products with  $I$ . The family  $\mathcal{F}_j$  consists, in this example, of the characteristic functions of  $T^i(X_0 \times G_j)$  ( $i = 0, 1, 2$ ) and of two more functions. The partition  $\mathcal{A}_{\mathcal{F}_j}$  of  $X \times I$  is labeled  $\{0, 1, \dots, 9\}$ . The resulting fundamental  $j$ -cells are labeled 047, 048,  $\dots$ , 359 (these are our  $j$ -rectangles). The fundamental  $j_k$ -cell in  $X \times I$  corresponding to the selected  $j_k$ -rectangle  $C$  is shown in gray (with pieces of the enclosing functions from  $\mathcal{F}_{j_k}^{p_{j_k}}$ ) and the point  $(x_C, t_C)$  is inside. The  $j$ th row of  $C$  is obtained by reading the labels of the fundamental  $j$ -cells along the trajectory of  $(x_C, t_C)$  for  $q$  iterates of  $(T \times \text{id})^{p_j}$  (the black dots). In this example it starts with |057|257|258|258|057| $\dots$ . The code  $\Phi_k$  changes this row (and the ones above) by following  $(x_C, t_C)$  under the action of  $(T \times R_{\mu_c})^{p_j}$  (the gray dots). In this example the  $j$ th row of  $\Phi_k(C)$  begins with |057|357|359|258|157| $\dots$ . Notice that the projection of the  $n$ th symbol in both names contains the point  $T^n(x_C)$ .

For any point  $y \in Y_{k-1}$  define its image,  $\Phi_k(y)$ , by replacing its every  $j_k$ -rectangle  $C$  with  $\Phi_k(C)$ . The column condition and the fact that  $\Phi_k$  preserves rows from  $j$  onwards ensure that  $\Phi_k(y)$  is still in  $Y_c$  and that the condition (Y2) is satisfied.

We will now verify the condition (Y1) – that all invariant measures on  $Y_k$  are in  $\mathcal{U}_k$ . Let  $\nu$  be such a measure. By Fact 2.1, for any  $n$  the closure of the convex hull of the measures  $\{M_n(\delta_y); y \in Y_k\}$  contains  $\nu$ . In other words there exist points  $y_1, \dots, y_N$  and coefficients

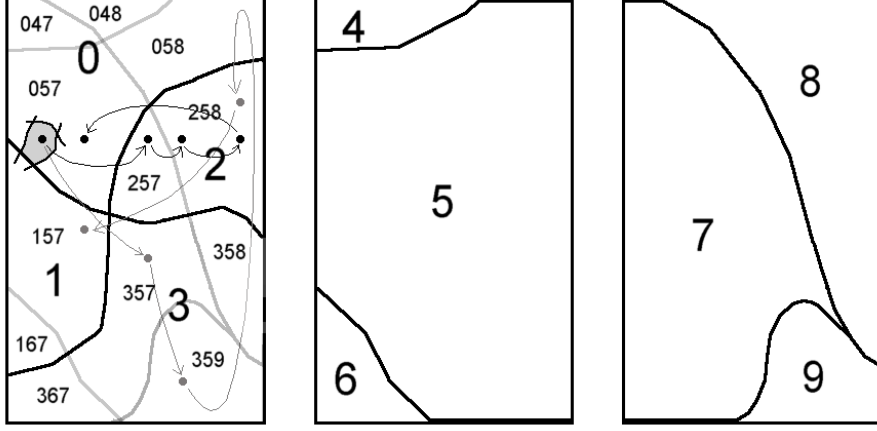


Figure 2: The code  $\Phi$

$\alpha_1, \dots, \alpha_N \in (0, 1)$  such that

$$\left| \nu(D) - \sum_{i=1}^N \alpha_i M_n(\delta_{y_i})(D) \right| < \delta \quad (3.1)$$

for all  $j$ -rectangles  $D$ .

For any  $y$  and any set  $D$  the measures  $M_n(\delta_y)(D)$  and  $M_n(\delta_{Ty})(D)$  differ by at most  $\frac{1}{n}$  (thus they are arbitrarily close for large  $n$ ) and this proximity is preserved by convex combinations. Therefore, if we choose  $n$  large enough, we can simply assume (replacing  $y_i$  with its image by at most  $p_{j_k}$  applications of  $S$ ) that  $\pi_k(y_i) \in (X_0 \times G_{j_k})$  for all  $i$ , i.e. that each  $y_i$  has a  $j_k$ -rectangle starting at coordinate 0. Furthermore, we shall pick an  $n$  that is a multiple of  $p_{j_k}$ .

For a given  $i \in \{1, \dots, N\}$  and for any  $j$ -rectangle  $D$  the number  $M_n(\delta_{y_i})(D)$  is  $\frac{1}{n}$  times the number of occurrences of  $D$  in the block  $B$  consisting of the first  $n$  coordinates of  $y_i$  in rows 1 through  $j$ . This block  $B$  forms the first  $j$  rows of a concatenation of  $j_k$ -rectangles  $\Phi_k(C_1), \dots, \Phi_k(C_Q)$ . We know that the number of occurrences of  $D$  in any  $\Phi_k(C)$  is  $p_{j_k}(\mu_C \times \lambda)(\pi_{k-1}(D)) \pm p_{j_k} \delta$ . Therefore the total number of occurrences of  $D$  in  $B$  is

$$\begin{aligned} N_B(D) &= \sum_{q=1}^Q (p_{j_k}(\mu_{C_q} \times \lambda)(\pi_{k-1}(D)) \pm p_{j_k} \delta) = \\ &= p_{j_k} \left( \sum_{q=1}^Q (\mu_{C_q} \times \lambda)(\pi_{k-1}(D)) \right) \pm Q p_{j_k} \delta. \end{aligned}$$

Therefore, if we set  $\mu_i = \frac{1}{Q} \sum_{q=1}^Q \mu_q$ , we can write

$$N_B(D) = Qp_{j_k}(\mu_i \times \lambda)(\pi_{k-1}(D)) \pm Qp_{j_k} \delta.$$

But  $Qp_{j_k} = n$ , since it is simply the length of  $B$ . So, the above statement is equivalent to

$$|N_B(D) - n(\mu_i \times \lambda)(\pi_{k-1}(D))| < n\delta$$

Dividing this by  $n$  we see that

$$|M_n(\delta(y_i))(D) - (\mu_i \times \lambda)(\pi_{k-1}(D))| < \delta,$$

for all  $i$  and all  $D$ . Taking convex combinations we obtain the estimate

$$\left| \left( \sum_{i=1}^N \alpha_i M_n(\delta y_i) \right) (D) - \left( \left( \sum_{i=1}^N \alpha_i \mu_i \right) \times \lambda \right) (\pi_{k-1}(D)) \right| < \delta,$$

Let  $\mu = \sum_{i=1}^N \alpha_i \mu_i$  and let  $\nu'$  be the (unique) measure in  $\mathcal{M}_{k-1}$  that factors onto  $\mu \times \lambda$  on  $X \times I$ . As remarked above, since  $j > j_{k-1}$ , any point  $y \in Y_{k-1}$  belongs to the  $j$ -rectangle  $D$  if and only if  $\pi_{k-1}(y)$  belongs to the set  $\pi_{k-1}(D)$  which is the closure of a fundamental  $j$ -cell. But this means that  $\nu(D) = (\mu \times \lambda)(\pi_{k-1}(D))$ , and thus

$$\nu'(D) = (\mu \times \lambda)(\pi_{k-1}(D)) = \left( \left( \sum_{i=1}^N \alpha_i \mu_i \right) \times \lambda \right) (\pi_{k-1}(D)).$$

Therefore for any  $j$ -rectangle  $D$

$$\left| \left( \sum_{i=1}^N \alpha_i M_n(\delta y_i) \right) (D) - \nu'(D) \right| < \delta,$$

Combining this with equation 3.1, we conclude that

$$|\nu(D) - \nu'(D)| < 2\delta$$

for all  $j$ -rectangles. By our choice of  $j$  and  $\delta$ ,  $d^*(\nu, \nu') < \varepsilon$ , which means that  $\nu \in \mathcal{U}_k$ , as requested in (Y1).

Now, having obtained the sequence of the  $Y_k$ 's, we define

$$Y = \bigcap_{m=1}^{\infty} \overline{\bigcup_{k=m}^{\infty} Y_k}.$$

In other words,  $Y$  is the set of all points  $y$  such that  $y = \lim_k y_k$ ,  $y_k \in Y_k$ . It is easy to see that this is a closed subsystem of  $Y_c$  and

an extension of  $X$ . The important observation is that any invariant measure  $\nu$  on  $Y$  is in every  $\mathcal{U}_k$ . This follows from the same argument as the one used to show that the invariant measures on  $Y_k$  are in  $\mathcal{U}_k$ , since this argument depended only on the properties of  $j_k$ -rectangles and  $Y$  contains no  $j_k$ -rectangles that did not occur in  $Y_k$ .

To show that  $Y$  is a principal extension of  $X$  we need to show that the conditional entropy of  $Y$  with respect to  $X$  is 0 for every measure  $\nu \in \mathcal{M}_s(Y)$ . For any  $k > 0$  and for any  $k' > k$  we have  $h_\nu(\mathcal{R}_k|X) \leq h_\nu(\mathcal{R}_{k'}|X)$ , since  $\mathcal{R}_{k'} \succ \mathcal{R}_k$ . On the other hand, since  $\nu$  is in the set  $\mathcal{U}_{k'}$ , we know that  $h_\nu(\mathcal{R}_{k'}|X) < \varepsilon_{k'}$ . It follows that for any  $k' > k$   $h_\nu(\mathcal{R}_k|X) < \varepsilon_{k'}$ , and thus  $h_\nu(\mathcal{R}_k|X) = 0$ . Thus we conclude that  $h_\nu(Y|X) = 0$ .

## References

- [B-D] M. Boyle and T. Downarowicz, *The entropy theory of symbolic extensions*, *Inventiones Math.* **156** (2004), pp 119-161.
- [D] T. Downarowicz, *Entropy Structure*, *Journal d'Analyse* **96** (2005), pp. 57-116
- [Le] F. Ledrappier *A variational principle for the topological conditional entropy*, *Springer Lec. Notes in Math.* **729**, Springer-Verlag (1979), pp 78-88
- [Li] E. Lindenstrauss, *Mean dimension, small entropy factors and an imbedding theorem*, *Publ. Math. I.H.E.S.* **89** (1999), pp 227-262
- [L-W] E. Lindenstrauss and B. Weiss, *Mean Topological Dimension*, *Israel J. Math* **115** (2000), pp 1-24