ISRAEL JOURNAL OF MATHEMATICS 112 (1999), 357–380

MARKOV EXTENSIONS FOR MULTI-DIMENSIONAL DYNAMICAL SYSTEMS

BY

JÉRÔME BUZZI*

Laboratoire de Topologie, Université de Bourgogne / CNRS
B.P. 47870, 21078 Dijon Cedex, France
e-mail: buzzi@iml.univ-mrs.fr

ABSTRACT

By a result of F. Hofbauer [11], piecewise monotonic maps of the interval can be identified with topological Markov chains with respect to measures with large entropy. We generalize this to arbitrary piecewise invertible dynamical systems under the following assumption: the total entropy of the system should be greater than the topological entropy of the boundary of some reasonable partition separating almost all orbits. We get a sufficient condition for these maps to have a finite number of invariant and ergodic probability measures with maximal entropy. We illustrate our results by quoting an application to a class of multi-dimensional, non-linear, non-expansive smooth dynamical systems.

1. Introduction

The motivation of this paper is the study of the "complexity" of chaotic multi-dimensional dynamical systems. By a chaotic dynamical system, we mean one that has positive topological entropy: the number of distinguishable orbits of given length, given some arbitrarily small precision, grows exponentially with the length (see below p. 361). One would like to find models for large classes of such dynamical systems, i.e., prove isomorphism "in the sense of entropy" with simple systems.

In this paper we prove that an arbitrary piecewise invertible dynamical system is isomorphic "in the sense of entropy" with a countable topological Markov chain as soon as one can find some invertibility partition (see below p. 359) such that:

(1) the topological entropy of the boundary of the partition is smaller than the total topological entropy;

* Part of this work was done at Université Paris-Sud, dép. de mathématiques, Orsay.
Received June 25, 1997

(2) the partition separates orbits outside some set which is negligible "in the sense of entropy".

We also prove a finiteness result.

We conclude by quoting a non-trivial application to a class of multi-dimensional dynamical systems. We show that fibered perturbations of products of chaotic smooth interval maps have a finite number of invariant and ergodic probability measures with maximal entropy.

ACKNOWLEDGEMENT: I would like to express my gratitude to Ph. Thieullen who introduced me to this subject and provided me with many invaluable ideas, not speaking of his constant encouragement.

1.1 SOME DEFINITIONS. Before stating our Main Theorem, we have to give some definitions.

ISOMORPHISM IN THE SENSE OF ENTROPY. Through the variational principle (see, e.g., [8]), the topological entropy $h_{\mathrm{top}}(f)$ of (X,f) appears often as the supremum $h(f)$ of the entropies $h_\mu(f)$ of the (infinitely-many) invariant and ergodic probability measures $\mu \in \mathrm{M}_{\mathrm{erg}}(X,f)$ carried by a dynamical system. Hence it is natural to focus on the invariant and ergodic probability measures with entropy close to the supremum.

For technical reasons, we shall consider the natural extensions instead of the original maps (see Remark 2.1(c)). Recall that the **natural extension** $(\mathcal{X},\mathcal{F})$ of an endomorphism (X,f) of a measurable space can be defined as the left-shift $\mathcal{F}$ acting on the set $\mathcal{X}$ of complete orbits:

$$\mathcal{X} = \{x \in X^{\mathbb{Z}}\colon \forall p \in \mathbb{Z} \quad x_{p+1} = f(x_p)\} \quad \text{and} \quad \mathcal{F}\colon (x_p)_{p\in\mathbb{Z}} \mapsto (x_{p+1})_{p\in\mathbb{Z}}.$$

Loosely speaking, $(\mathcal{X},\mathcal{F})$ and (X,f) are essentially the same from the point of view of complexity. For instance, their invariant probability measures are in a one-to-one relationship in a way which preserves ergodicity and entropy.

These remarks motivate the following definition of isomorphism. Let (X,f) and (Y,g) be endomorphisms of measurable spaces. Assume that they have positive **total entropies**:

$$h(f) \stackrel{\text{def}}{=} \sup\{h_\mu(f)\colon \mu \in \mathrm{M}_{\mathrm{erg}}(X,f)\} > 0 \quad \text{and} \quad h(g) > 0.$$

Consider their natural extensions $(\mathcal{X},\mathcal{F})$ and $(\mathcal{Y},\mathcal{G})$.

Definition 1.1: We say that (X, f) and (Y, g) are **isomorphic in the sense of entropy**, or simply h-**isomorphic**, if there exist invariant subsets $\mathcal{X}' \subset \mathcal{X}$, $\mathcal{Y}' \subset \mathcal{Y}$ of the natural extensions, such that:

(1) the restricted systems $(\mathcal{X}', \mathcal{F})$ and $(\mathcal{Y}', \mathcal{G})$ are conjugated by some bi-measurable invertible map;
(2) the complements of these sets are h-**negligible**, i.e., there exists $0 \leq H < h(f)$ such that
$$\mu \in \mathrm{M_{erg}}(\mathcal{X}, \mathcal{F}) \quad \text{and} \quad h_\mu(\mathcal{F}) > H \implies \mu(\mathcal{X} \setminus \mathcal{X}') = 0,$$
and likewise for $\mathcal{Y} \setminus \mathcal{Y}'$.

One could also further concentrate on the measures (if they exist) which realize the supremum of the entropy: that is the original point of view of **intrinsic ergodicity** introduced by B. Weiss [20].

TOPOLOGICAL MARKOV CHAINS. The simplest class of combinatorial dynamical systems which contains systems having entropy with any prescribed non-negative value is the class of topological Markov chains, provided that we allow the underlying graphs to be countably infinite. We recall that a **topological Markov chain** is the set $\Sigma(G)$ of bi-infinite paths on some countable (maybe finite) oriented graph G together with the left-shift σ:
$$\Sigma(G) = \{g \in G^{\mathbb{Z}} \colon \forall p \in \mathbb{Z} \quad g_p \overset{G}{\to} g_{p+1}\} \quad \text{and} \quad \sigma\colon (g_p)_{p\in\mathbb{Z}} \mapsto (g_{p+1})_{p\in\mathbb{Z}}.$$

Of course, not every chaotic smooth dynamical system is h-isomorphic to a topological Markov chain (consider the product of an irrational rotation with anything of positive topological entropy).

We recall that $\Sigma(G)$ splits into its **irreducible sub-chains**, $\Sigma(H)$, where H ranges over the sub-graphs maximal with respect to inclusion having the following property: for every $(g, h) \in H^2$ there exists a path from g to h. Each irreducible sub-chain is a maximal topologically transitive subset of $\Sigma(G)$. B. M. Gurevič [9, 10] proved that each irreducible sub-chain carries at most one invariant probability measure with maximal entropy, so the number of irreducible sub-chains with large entropy bounds the number of such measures.

PIECEWISE INVERTIBLE DYNAMICAL SYSTEMS. We study (X, P, f) such that:

(1) the space X is metric;
(2) the "partition" P is a countable collection of pairwise disjoint subsets of X, the union of which
$$Y = \bigcup_{A\in P} A$$

is open and dense in X. Each $A \in P$ is non-empty, open and relatively compact, it is also connected and locally connected;

(3) the map $f\colon Y \to X$ is such that for every $A \in P$, there is a homeomorphism $f_A\colon U \to V$ with open sets $U \supset \overline{A}$, $V \supset \overline{f(A)}$, such that its restriction to A coincides with that of f.

Recall that the iterated partition is

$$P^n = \{[A_0 \cdots A_{n-1}] \overset{\text{def}}{=} A_0 \cap f^{-1}A_1 \cap \cdots \cap f^{-n+1}A_{n-1} \neq \emptyset\colon A_0, \ldots, A_{n-1} \in P\},$$

the set of **n-cylinders** for $n = 1, 2, \ldots$. We write $P^n(x)$ for the element of P^n containing some point $x \in X$, if that element exists.

We shall need to measure the quality of the coding through P:

Definition 1.2: We write $P_n(x)$ for the connected component of the cylinder $P^n(x)$ containing x. We say that P **h-separates** if for all $x \in X$ outside some h-negligible set, we have

$$\lim_{n\to\infty} \operatorname{diam}(P_n(x)) = 0.$$

The **multiplicity entropy** of (X, P, f) is

$$h_{\text{mult}}(P, f) = \limsup_{n\to\infty} \frac{1}{n} \log \operatorname{mult}(P^n) \text{ with } \operatorname{mult}(Q) = \max_{x\in X} \#\{A \in Q\colon x \in \overline{A}\}.$$

(X, P, f) is not supposed to satisfy the Markov property, i.e., we do not assume that the sets $f(A)$ for $A \in P$ are unions of elements of P. Such a property would imply that the symbolic dynamics of (X, f) would be the topological Markov chain defined by P together with the arrows $A \to B \iff f(A) \supset B$, bringing us quite close to our goal.

Now one can force this property by going to an extension. Following G. Keller's version [14] of F. Hofbauer's construction [11], set

$$P_1 = P \quad \text{and} \quad P_{n+1} = \{f(A) \cap B \neq \emptyset\colon A \in P_n \text{ and } B \in P\}$$

for $n \geq 1$. The **Markov extension** is defined as the disjoint union of the elements of $\bigcup_{n\geq 1} P_n$ together with the obvious dynamics (see below for the connected case).

Clearly, the Markov extension satisfies the Markov property with respect to the obvious partition. Moreover, this construction is always possible. However, one can expect it to give trivial results in many cases. For instance, the Markov extension could have no invariant probability measures. F. Hofbauer [11] has shown that for piecewise monotonic maps with positive entropy: the map, the Markov extension and the topological Markov chain defined using the Markov property are all h-isomorphic (with a constant $H = 0$).

To get something similar in spirit for non-linear multi-dimensional systems, it turns out that we have to modify this construction by taking the *connected components* of the intersections $f(A) \cap B$ above, instead of the intersections themselves. This will ensure that "bad measures are related to the border of the partition" (in a sense made precise in section 2) and we shall get an h-isomorphism (with a constant H given by the entropy of the boundary of P).

Definition 1.3: The **connected Markov diagram** of (X, P, f) is $\mathcal{D} = \bigcup_{n \geq 1} \mathcal{D}_n$ with

$$\mathcal{D}_1 = P,$$
$$\mathcal{D}_{n+1} = \{C\colon C \neq \emptyset \text{ is a connected component of } f(A) \cap B \text{ for } A \in \mathcal{D}_n \text{ and } B \in P\}.$$

Moreover, $\mathcal{D}$ has a natural graph structure:

$$U \to V \iff V \text{ is a connected component of } f(U) \cap B \text{ for some } B \in P.$$

$\mathcal{D}$ thus defines a topological Markov chain $\Sigma(\mathcal{D})$.

Remark that $\mathcal{D}$ is countable. Indeed, its elements are open (each element of P is open and locally connected) so that any element $U \in \mathcal{D}$ can have only countably-many successors included in any $B \in P$, which has compact closure.

Having taken the connected components instead of the cylinders, it is no longer possible to work with the symbolic dynamics. One could replace it with a version of the "Yoccoz puzzle". But it turns out to be easier to work in a Markov extension.

Definition 1.4: The **connected Markov extension** of (X, P, f) is $(\hat{X}, \hat{P}, \hat{f})$ defined as

$$\hat{X} = \{(x, D) \in X \times \mathcal{D}\colon x \in D\},$$
$$\hat{P} = \{\pi_{\mathcal{D}}^{-1}(D)\colon D \in \mathcal{D}\} \qquad \text{with } \pi_{\mathcal{D}}\colon (x, D) \mapsto D,$$
$$\hat{f} = (x, D) \mapsto (f(x), E),$$

where E is the successor of D in $\mathcal{D}$ containing $f(x)$, if it exists.

Remark that $\hat{f}$ is defined only on $\hat{Y} \stackrel{\text{def}}{=} \{(x,D) \in \hat{X}: x \in Y \cap f^{-1}(Y)\}$.

We have the natural map $\hat{\pi}$: $(x,D) \mapsto x$ from $\hat{X}$ to X. $\hat{\pi}$ is in general infinite-to-one. On the domain $\hat{Y}$ of $\hat{f}$ we have

$$\hat{\pi} \circ \hat{f} = f \circ \hat{\pi}.$$

ENTROPIES. We recall the definition of the **Bowen topological entropy** [1]. Given a positive integer n, let $d_n(x,y)$ be the distance $\max_{0 \le k < n} d(f^k x, f^k y)$. An (ϵ, n)-ball is a ball with radius $\epsilon > 0$ with respect to d_n. An (ϵ, n)-covering is a covering by (ϵ, n)-balls and the (ϵ, n)-covering number of $S \subset X$ is the minimal cardinality of an (ϵ, n)-covering of S and is denoted by $r(\epsilon, n, S)$. If f is defined only on a subset of X (as here), then we adapt these definitions by requiring an (ϵ, n)-cover of S to cover only the set of points in S which define orbits with a length at least n.

The topological entropy of a (non-necessarily invariant) subset S of X is

$$h_{\text{top}}(S, f) = \lim_{\epsilon \to 0+} \limsup_{n \to \infty} \frac{1}{n} \log r(\epsilon, n, S).$$

Our entropy condition will be about the **boundary** of P. Set

$$\partial P = \bigcup_{A \in P} \partial A.$$

We have to assume that *the boundary of our system is small from the point of view of the entropy.*

To define the dynamics on the boundary we embed (X, f) in a continuous extension, namely the action of the left-shift σ on $\widetilde{X \times P}$:

$$\overline{\{(x, A) \in Y^{\mathbb{N}} \times P^{\mathbb{N}}: \forall n \in \mathbb{N} \quad x_{n+1} = f(x_n) \text{ and } x_n \in A_n\}} \subset X^{\mathbb{N}} \times P^{\mathbb{N}}$$

(P being endowed with the discrete topology). The **border entropy** is defined as

$$h_B(P, f) = h_{\text{top}}(\{(x, A) \in \widetilde{X \times P}: x_0 \in \partial P\}, \sigma),$$

$\widetilde{X \times P}$ being endowed with any metric of the form

$$d((x, A), (y, B)) = d(x_0, y_0) + \sum_{n \ge 0} 2^{-n} d_P(A_n, B_n)$$

with d_P a bounded metric on P.

Remark 1.5: If (X, P, f) is a piecewise invertible dynamical system such that X is compact and f extends continuously to the whole of X, then

$$h_B(P, f) \leq h_{\text{top}}(\partial P, f) + h_{\text{mult}}(P, f)$$

provided, if P is infinite, that, identifying P with $\{1, 2, \dots\}$, we choose the distance $d_P(n, m) = |1/n - 1/m|$.

We defer the proof to the Appendix, at the end of the paper.

1.2 MAIN RESULTS. We state and discuss our results.

MAIN THEOREM: *Let (X, P, f) be a piecewise invertible dynamical system. Assume that*

(H1) *P h-separates,*

(H2) *$h_B(P, f) < h(f)$.*

Then (X, f), $(\hat{X}, \hat{f})$ and $\Sigma(\mathcal{D})$ are all h-isomorphic. Moreover,

$$\limsup_{n\to\infty} h\Big(\Sigma\Big(\mathcal{D} \setminus \bigcup_{k\leq n} \mathcal{D}_k\Big)\Big) < h(f). \tag{1.1}$$

The point of the inequality (1.1) is the following:

COROLLARY 1.6: *Let (X, P, f) satisfies hypothesis* (H1)–(H2) *of the Main Theorem. Assume in addition that*

(H3) *for each $n \geq 1$, $\mathcal{D}_n$ is finite.*

Then for some constant $H < h(f)$, there are finitely many irreducible sub-chains having total entropy greater than H.

Using B.M. Gurevič's result, we get:

COROLLARY 1.7: *Let (X, P, f) satisfy hypotheses* (H1)–(H3). *Then there can only be finitely many invariant and ergodic probability measures with maximal entropy.*

Remark 1.8: We shall prove slightly stronger statements (Theorems A, B and C). In particular, for applications to perturbations as in Theorem 1.9 below, it turns out to be interesting to have an estimate, uniform with respect to f, on the rate of convergence of the lim sup in the inequality (1.1).

To illustrate our results we give the following application:

THEOREM 1.9 [6]: *Let $f_1, \dots, f_d$: $[0,1] \to [0,1]$ be C^∞ interval maps with positive topological entropy and* **non-degenerate** *critical orbits: for each $i = 1, \dots, d$, for all points c with $f_i'(c) = 0$, we have*

$$f_i''(c) \neq 0 \quad \text{and} \quad f_i'(f_i^k(c)) \neq 0 \quad (\forall k \geq 1).$$

Then every small enough C^∞-perturbation F: $[0,1]^d \to [0,1]^d$ of the direct product $F_0 = f_1 \times \cdots \times f_d$, which is fibered, that is, of the form

$$F(x_1, \dots, x_d) = (\bar{f}_1(x_1), \bar{f}_2(x_1, x_2), \dots, \bar{f}_d(x_1, \dots, x_d))$$
$$\text{for some maps } \bar{f}_i\colon [0,1]^i \to [0,1],$$

is isomorphic in the sense of entropy with a topological Markov chain which has a finite number of irreducible sub-chains.

COROLLARY: *Every such perturbation F: $[0,1]^d \to [0,1]^d$ has a finite, non-empty set of invariant and ergodic probability measures with maximal entropy.*

We remark that we get for these multi-dimensional systems the same results as F. Hofbauer [11] on the interval.

We stress that the maps in this example are non-linear and non-expanding.

It is enough to check that hypotheses (H1)–(H3) are satisfied for these maps. The proof will appear elsewhere, as it uses quite different techniques from the ones of this paper, notably entropy estimates by approximation of differentiable submanifolds by semi-algebraic sets and related results of Y. Yomdin [21].

We think that our technique can be applied to many other interesting classes of multi-dimensional dynamical systems. For instance, we can prove that piecewise expanding maps in arbitrary dimensions have generically a non-zero and finite number of ergodic absolutely continuous invariant measures [7].

COMMENTS. The starting point of our work was the result by F. Hofbauer [11] about piecewise monotonic interval map with positive entropy. Remark that in this case our Main Theorem gives exactly F. Hofbauer's result: taking P to be the partition into monotonicity intervals we see that the boundary of the partition is finite (hence has zero entropy) and that the cylinders are connected, so that the connected diagram and the usual diagram are the same.

There had been previous attempts to use Markov diagrams in higher dimension, by F. Hofbauer [12] and then by G. Keller [15]. Indeed, we are indebted to G. Keller's paper, although our approach is quite different: for instance we get a *pointwise isomorphism*, a result that was thought specific of the interval. We also have benefited from S. Newhouse's presentation [16] of F. Hofbauer's result.

In [2, 3, 4], we showed how to deal with an infinite boundary but we had to assume that the cylinders were connected— a very restrictive geometric condition which seemed to prevent the study of *non-linear multi-dimensional* mappings. The removal of this assumption is the most important point in this paper. Another point, less important and more technical, is that working with a Markov extension, more obviously related to the dynamics than $\Sigma(\mathcal{D})$, allows several simplifications. The only price to be paid is that we have to discard a h-negligible subset not only in (X, f) but also in $(\hat{X}, \hat{f})$.

Remark 1.10: As was pointed to us by G. Keller, it can be interesting (for example, to have "margins" for distortion estimates) to redefine the Markov diagram as $\bigcup_{n\geq 0} \mathcal{D}'_n$ with $\mathcal{D}'_0 = \{X\}$ and $\mathcal{D}'_{n+1}$ the set of connected components of $f(A \cap Z)$ for $A \in \mathcal{D}'_n$ and $Z \in P$. It makes only trivial changes from our point of view. More precisely, Theorem A remains true for this variant of the construction. Theorem B fails for obvious reasons: for instance, if $f(A) = X$ for all $A \in P$ then this variant of the Markov diagram has only one point. But one can recover Theorems B and C very easily, by putting labels on the Markov diagram: label $U \to V$ with $B \in P$ if V is a connected component of $f(U \cap B)$ and define the topological Markov chain over the alphabet $\mathcal{D}' \times P$ in the obvious way.

1.3 Outline of the proof. We start by establishing the h-isomorphism of (X, f) with its connected Markov extension $(\hat{X}, \hat{f})$. We proceed by trying to find an inverse to $\hat{\pi}$. This inverse will only be partially defined and will not be onto either. Thus we get a measurable isomorphism between subsets (Proposition 2.2). To see that the complements of these subsets carry only measures with small entropy, we prove that every such measure is "shadowed" by the boundary of the partition (Propositions 2.6 and 2.7). An abstract ergodic result (Theorem 2.4) shows then that the entropy of such a measure is bounded by the topological entropy of the shadowing set, finishing the proof of h-isomorphism of (X, f) with $(\hat{X}, \hat{f})$.

We continue by checking the isomorphism of the Markov extension $(\hat{X}, \hat{f})$ with the topological Markov chain $\Sigma(\mathcal{D})$. We have a natural map: $\hat{x} \mapsto (\pi_{\mathcal{D}}(\hat{f}^n \hat{x}))_{n \in \mathbb{N}}$. From the condition that P h-separates, one deduces that the sets (in $\hat{X}$ as well as in $\Sigma(\mathcal{D})$) where this map fails to be one-to-one are h-negligible. To check that this map is h-almost onto, one remarks that the paths belonging to $\Sigma(\mathcal{D})$ which are not in the image correspond, in some sense, to itineraries of orbits going through ∂P. Now if a set is not negligible for some measure, then its topological entropy bounds the entropy of the measure and this finishes the proof of the isomorphim.

Finally, inequality (1.1) (the smallness of the entropy at infinity) follows from the fact that paths on $\mathcal{D} \setminus \bigcup_{k \leq n} \mathcal{D}_k$ are shadowed in some sense by the boundary of P if n is very large. This gives the needed bound on the entropy of the corresponding measures.

2. Isomorphism with the Markov extension

We prove in this section the following:

THEOREM A: *Let (X, P, f) by a piecewise invertible dynamical system. Set $\Delta P = \bigcup_{A \in P} f_A(\partial A)$. Assume that:*
(H1) P h-separates,
(H2′) $h_{\text{top}}(\Delta P, f) < h(f)$.
Then (X, f) and $(\hat{X}, \hat{f})$ are h-isomorphic.

Remark 2.1: (a) Theorem A is slightly stronger than the corresponding part of the Main Theorem as (H2′) only counts real orbits and does not take into account $h_{\text{mult}}(P, f)$, i.e., the separation by P of close-by points.

(b) The assumption that P h-separates makes things a little simpler, but one can remove it provided one replaces the entropy condition (H2′) by: $h_{\text{top}}(\Delta P, f) + h_{\text{loc}}(P, f) < h(f)$ with $h_{\text{loc}}(P, f)$ defined as the growth rate with respect to n, when $\epsilon \to 0+$, of the maximum number of ϵ, n-orbits within any connected component of an arbitrary n-cylinder.

(c) We are going to check that $\hat{\pi}$ is a pointwise isomorphism between almost all of the *natural extensions*. Going to the natural extensions is really necessary. Let us give an example. Consider β-transformations, i.e., $x \mapsto \beta x \bmod 1$ on $[0, 1[$ with parameter $\beta > 1$. Fix $1 < \beta < (1 + \sqrt{5})/2$ such that the orbit of 1 is not eventually periodic. $P = \{]0, \beta^{-1}[,]\beta^{-1}, 1[\}$. The element of the Markov diagram containing $\hat{f}(\hat{x})$ determines the element of P containing $\hat{\pi}(\hat{x})$. Indeed, every element of the diagram has exactly one predecessor, except for $]0, \beta^{-1}[$. But, as $\beta \cdot |]\beta^{-1}, 1[| < \beta^{-1}$, all the predecessors of that element are subsets of $]0, \beta^{-1}[$. Hence, if $\hat{\pi}$ is an isomorphism between $(\hat{X}, \hat{f}, \hat{\mu})$ and (X, f, μ) so that $\hat{\pi}^{-1}(P)$ is a generating partition, we get that

$$\bigvee_{n \geq 1} \hat{f}^{-n} \hat{\pi}^{-1}(P) \supset \bigvee_{n \geq 1} \hat{f}^{-n} \hat{P} \supset \hat{\pi}^{-1}(P).$$

But this implies that $\hat{\mu}$ and therefore μ have zero entropy. ∎

2.1 Existence of an isomorphism between subsets. Remark that we have an obvious injection:

$$i\colon x \in Y = \bigcup_{A \in P} A \longmapsto (x, P(x)) \in \hat{X}.$$

We have the following easy formula:

$$\hat{f}^n(x, D) = (f^n(x), f^n(C_x(D \cap P_{n+1}(x)))), \tag{2.1}$$

where $C_x(\cdot)$ denotes the connected component containing x.

In particular, $\hat{f}^n(i(x)) = (f^n(x), f^n(P_{n+1}(x)))$.

Let $(\mathcal{X}, \mathcal{F})$ resp. $(\hat{\mathcal{X}}, \hat{\mathcal{F}})$ be the natural extensions of (X, f) resp. $(\hat{X}, \hat{f})$; $\hat{\pi}\colon \hat{X} \to X$ extends to a unique map $\hat{\pi}\colon \hat{\mathcal{X}} \to \mathcal{X}$.

Proposition 2.2: *Consider the following invariant subsets of the natural extensions:*

$$\hat{\mathcal{X}}' = \{\hat{x} \in \hat{\mathcal{X}}\colon \forall p \in \mathbb{Z} \exists n \geq 0 \ \hat{x}_p = \hat{f}^n(i(\hat{\pi}(\hat{x}_{p-n})))\},$$
$$\mathcal{X}' = \{x \in \mathcal{X}\colon \forall p \in \mathbb{Z} \text{ the sequence } (f^n(P_{n+1}(x_{p-n})))_{n \in \mathbb{N}} \text{ stabilizes}\}$$

(by "stabilization" we mean that the sequence is constant for large n).

Then the restriction $\hat{\pi}\colon (\hat{\mathcal{X}}', \hat{\mathcal{F}}) \to (\mathcal{X}', \mathcal{F})$ is a well-defined measurable isomorphism.

Proof: We first show that this restriction of $\hat{\pi}$ is well-defined, i.e., that $\hat{\pi}(\hat{\mathcal{X}}') \subset \mathcal{X}'$. Let $\hat{x} \in \hat{\mathcal{X}}'$. Set $(x_p, D_p) = \hat{x}_p$ for $p \in \mathbb{Z}$. We prove that $x = (x_p)_{p \in \mathbb{Z}} \overset{\text{def}}{=} \hat{\pi}(\hat{x}) \in \mathcal{X}'$.

As $\hat{x} \in \hat{\mathcal{X}}'$, there exists an integer $n \geq 0$ such that $\hat{x}_0 = \hat{f}^n(i(x_{-n}))$, i.e.

$$D_0 = f^n(P_{n+1}(x_{-n})). \tag{2.2}$$

Let $m \geq n$. As $\hat{x} \in \hat{\mathcal{X}}$, $\hat{x}_0 = \hat{f}^m(\hat{x}_{-m})$ so that

$$D_0 = f^m(C_{x_{-m}}(D_{-m} \cap P_{m+1}(x_{-m})))$$

according to (2.1). Hence $f^m(P_{m+1}(x_{-m})) \supset D_0$.

On the other hand, $f^{m-n}(P_{m+1}(x_{-m})) \subset P_{n+1}(x_{-n})$: this, with (2.2), gives the reverse inclusion so that we have the equality $f^m(P_{m+1}(x_{-m})) = D_0$ for all $m \geq n$.

This is the required stabilizing property for $p = 0$. The case for general $p \in \mathbb{Z}$ is similar and therefore $\hat{\pi}(\hat{x}) \in \mathcal{X}'$, as was to be proved.

We note that we have also proved that, if $\hat{x}_p = \hat{f}^n(i(x_{p-n}))$ for some n, then $\hat{x}_p = \hat{f}^m(i(x_{p-m}))$ for all $m \geq n$.

Now, we build the inverse j of $\hat{\pi}$: $\hat{\mathcal{X}}' \to \mathcal{X}'$. Let $x \in \mathcal{X}'$. Define $\hat{x} = j(x)$ by setting, for all $p \in \mathbb{Z}$,

$$\hat{x}_p = \hat{f}^n(i(x_{p-n})) = (x_p, f^n(P_{n+1}(x_{p-n}))) \quad \text{for large } n.$$

Each $\hat{x}_p$ is well-defined by the stabilization property.

Taking n large enough we have $\hat{x}_p = \hat{f}^n(i(x_{p-n}))$ and $\hat{x}_{p+1} = \hat{f}^{n+1}(i(x_{p-n}))$ so that $\hat{x}_{p+1} = \hat{f}(\hat{x}_p)$. Thus $\hat{x} \in \hat{\mathcal{X}}$. Now it is clear that $\hat{x} \in \hat{\mathcal{X}}'$.

We check that j is indeed the inverse of $\hat{\pi}$. Clearly, $\hat{\pi} \circ j = \mathrm{Id}$. But the definition of $\hat{\mathcal{X}}'$ makes it clear that $\hat{\pi}|\hat{\mathcal{X}}'$ is one-to-one. Thus $j = \hat{\pi}^{-1}$ as claimed. ∎

2.2 Shadowing of a measure. We shall prove that bad measures, i.e., measures such that the good set $\mathcal{X}'$ or $\hat{\mathcal{X}}'$ has not total measure, are "related" to the border of the partition. The relationship will be the following.

In this subsection we only assume that (X, f) is a measurable self-map of some standard space with natural extension $(\mathcal{X}, \mathcal{F})$.

Definition 2.3: We say that a measure $\mu \in \mathrm{M}(\mathcal{X}, \mathcal{F})$ is **shadowed** by a subset S of X if the following holds: for every $\epsilon > 0$, for μ-a.e. $x \in \mathcal{X}$, one can find arbitrarily large integers m, n and a point $s \in S$ such that:
(2.3)

$$d(x_{-n+k}, f^k(s)) < \epsilon \quad \text{for at least } (1-\epsilon)(m+n+1) \text{ of the } k\text{'s in } [0, m+n].$$

We say that $[-n, m]$ is a ϵ-**shadowing interval** for x with respect to S.

The shadowing property gives a bound on the metric entropy in terms of Bowen's topological entropy of the subset which shadows:

Theorem 2.4 ([2]): *Let (X, f) be some measurable endomorphism with X some standard space. Let μ be an invariant probability measure on the natural extension of (X, f) which is shadowed by some subset S of X. Then*

$$h(\mathcal{F}, \mu) \leq h_{\mathrm{top}}(S, f).$$

We refer to [2, p. 59–61] for the proof. We note that it is very similar to the proof of Theorem C below.

The shadowing property will be proved first under the following "symbolic" form:

Given a measurable partition P of X, we say that μ is **P-shadowed** by S if shadowing holds but with (2.3) replaced by the condition:

$$x_{-n} \text{ and } s \text{ are in the same element of } P_{n+m+1} \tag{2.4}$$

(the number ϵ does not play any role here).

We have the following easy fact:

LEMMA 2.5: *Assume that P separates μ-almost every orbit. Then P-shadowing implies shadowing (for μ).*

The definition of shadowing (in particular the ϵ) was chosen to make this true and even obvious.

2.3 Bad measures on $(\mathcal{X}, \mathcal{F})$.

PROPOSITION 2.6: *Let μ be an invariant and ergodic probability measure of $(\mathcal{X}, \mathcal{F})$ such that $\mu(\mathcal{X} \setminus \mathcal{X}') > 0$. Then μ is P-shadowed by $\Delta P = \bigcup_{A \in P} f_A(\partial A)$.*

Proof: Let μ be as above. As $\mathcal{X}'$ is invariant and μ is ergodic, $\mu(\mathcal{X} \setminus \mathcal{X}') = 1$. Write $\mathcal{X}' = \bigcap_{p \in \mathbb{Z}} \mathcal{X}_p$ with

$$\mathcal{X}_p = \{x \in \mathcal{X}\colon \text{the sequence } (f^n(P_{n+1}(x_{p-n})))_{n \in \mathbb{N}} \text{ stabilizes}\}.$$

We relate the property "$x \in \mathcal{X}_p$" to the boundary of P. To start with, $x \notin \mathcal{X}_p$ if and only if there exist infinitely-many integers $n \geq 0$ such that

$$f^{n+1}(P_{n+2}(x_{p-n-1})) \subsetneq f^n(P_{n+1}(x_{p-n})).$$

This strict inclusion occurs if and only if $f(P(x_{p-n-1}))$ does not cover $P_{n+1}(x_{p-n})$. These two sets always meet and $P_{n+1}(x_{p-n})$ is a connected set by definition. Hence, the strict inclusion occurs if and only if

$$\partial f(P(x_{p-n-1})) \cap P_{n+1}(x_{p-n}) \neq \emptyset. \tag{2.5}$$

Obviously, (2.5) implies $\partial f(P(x_{(p-1)-(n-1)-1})) \cap P_{(n-1)+1}(x_{(p-1)-(n-1)}) \neq \emptyset$. Therefore, $\mathcal{X} \setminus \mathcal{X}_p \subset \mathcal{X} \setminus \mathcal{X}_{p-1}$. But, by definition, $\mathcal{F}^{-1}\mathcal{X}_p = \mathcal{X}_{p+1}$, so that $\mu(\mathcal{X}_p) = \mu(\mathcal{X}_{p-1})$. Hence, $\mathcal{X}_p = \mathcal{X}_{p-1}$ modulo μ. In particular

$$\mathcal{X} \setminus \mathcal{X}' = \bigcap_{p \in \mathbb{Z}} \mathcal{X} \setminus \mathcal{X}_p \mod \mu.$$

Now given any constant L, one can find a shadowing interval $[-n, m]$ with $n, m \geq L$. Indeed, just take $m = L$, and the existence of n as required is then a consequence of $x \notin \mathcal{X}_m$, using (2.5) and

$$\partial f(A) = f_A(\partial A) \subset \Delta P \qquad \forall A \in P.$$

Hence μ is P-shadowed by ΔP. ∎

By Lemma 2.5 and the assumption that P h-separates, this proposition implies that any bad measure on $(\mathcal{X}, \mathcal{F})$ is shadowed by ΔP or has small entropy. Hence, applying Theorem 2.4 and using the hypothesis $h_{\mathrm{top}}(\Delta P, f) < h(f)$, we see that the entropy of any bad measure on $(\mathcal{X}, \mathcal{F})$ is small.

Thus, the bad set $\mathcal{X} \setminus \mathcal{X}'$ is therefore h-negligible.

2.4 Bad measures on $(\hat{\mathcal{X}}, \hat{\mathcal{F}})$. To finish the proof of Theorem A, we have to see that $\hat{\mathcal{X}} \setminus \hat{\mathcal{X}}'$ is h-negligible.

We remark that the result of the previous subsection implies (using the isomorphism $\hat{\pi}$: $\hat{\mathcal{X}}' \to \mathcal{X}'$) that $h(\hat{f}) \geq h(f) > h_{\mathrm{top}}(\Delta P, f)$. Hence, to get an h-isomorphism, it will be enough to prove that bad measures of $(\hat{\mathcal{X}}, \hat{\mathcal{F}})$ have entropy less than $h_{\mathrm{top}}(\Delta P, f)$.

Proposition 2.7: *Let $\hat{\mu}$ be an invariant and ergodic probability measure of $(\hat{\mathcal{X}}, \hat{\mathcal{F}})$. Assume that $\hat{\mu}(\hat{\mathcal{X}} \setminus \hat{\mathcal{X}}') > 0$. Then $\hat{\pi}_* \hat{\mu}$ is P-shadowed by $\bigcup_{k=0}^{K} f^k(\Delta P)$ for some finite K.*

Remark that $h_{\mathrm{top}}(\bigcup_{k=0}^{K} f^k(\Delta P), f) = h_{\mathrm{top}}(\Delta P, f)$.

Proof: Let $\hat{\mu}$ be as above. By ergodicity, $\hat{\mu}(\hat{\mathcal{X}} \setminus \hat{\mathcal{X}}') = 1$. Write $\hat{\mathcal{X}}' = \bigcap_{p \in \mathbb{Z}} \hat{\mathcal{X}}_p$ with $\hat{\mathcal{X}}_p = \{\hat{x} \in \hat{\mathcal{X}} \colon \exists n \geq 0 \, \hat{x}_p = \hat{f}^n(i(\hat{\pi}(\hat{x}_{p-n})))\}$.

It is obvious that $\hat{\mathcal{X}}_p \subset \hat{\mathcal{X}}_{p+1}$. Now $\hat{\mathcal{F}}^{-1} \hat{\mathcal{X}}_p = \hat{\mathcal{X}}_{p+1}$ so that $\hat{\mathcal{X}}_p$ and $\hat{\mathcal{X}}_{p+1}$ have the same measure. Hence $\hat{\mathcal{X}}_p = \hat{\mathcal{X}}_{p+1}$ modulo $\hat{\mu}$: $\hat{\mu}$-almost all points must belong to $\hat{\mathcal{X}} \setminus \hat{\mathcal{X}}_p$ for each $p \in \mathbb{Z}$.

Using formula (2.1) for $\hat{f}^n$, we see that this means that, for $\hat{\mu}$-almost all $\hat{x}$, for each $p \in \mathbb{Z}$,

$$P_{n+1}(x_{p-n}) \not\subset D_{p-n} \qquad \forall n \geq 0.$$

Hence the connected set $P_{n+1}(x_{p-n})$ must meet the boundary of D_{p-n} in X. This boundary is contained in the union $\partial P \cup \Delta P \cup f(\Delta P) \cup \cdots \cup f^{\ell-1}(\Delta P)$ with ℓ the **level** of D_{p-n}: the level of an element D of $\mathcal{D}$ is the smallest integer k such that $D \in \mathcal{D}_k$, i.e., such that there exists $A \in P_k$ with $f^{k-1}(A) = D$.

Take K so large that the set of points with level at most K has positive measure. Then it is possible to find for almost every $\hat{x}$, for all $p \in \mathbb{Z}$, an arbitrarily large integer n such that the level of D_{p-n} is bounded by K; thus

$$P_{n+1}(x_{p-n}) \text{ meets } \partial P \cup f^k(\Delta P) \cup \cdots \cup f^{K-1}(\Delta P).$$

Applying f to both sides reduces n by one on the left-hand side and changes the right-hand side to $\Delta P \cup \cdots \cup f^K(\Delta P)$. ∎

The previous proposition gives a bound on the entropy of $\hat{\pi}_*\hat{\mu}$ when $\hat{\mu}$ is a bad measure on $(\hat{X}, \hat{f})$. We need a bound on the entropy of $\hat{\mu}$ itself. It is provided by the following result:

PROPOSITION 2.8: *Let $\hat{\pi}$: $(\hat{X}, \hat{f}, \hat{\mu}) \to (X, f, \mu)$ be an extension of endomorphisms of Lebesgue probability spaces. Assume that $\hat{\pi}$ is countable-to-one. Then*

$$h(\hat{f}, \hat{\mu}) = h(f, \hat{\pi}_*\hat{\mu}).$$

We remark that, in contrast to the case of an extension of automorphisms (see [17, 18]), $(\hat{X}, \hat{f}, \hat{\mu})$ is not necessarily a finite extension of (X, f, μ).

Proof: We may assume that $\hat{X} = X \times \mathbb{N}$ (see, for instance, [18, p. 40]). We may also assume that $\hat{\mu}$ is ergodic. Set $\mu = \hat{\pi}_*\hat{\mu}$. It is obvious that $h(\hat{f}, \hat{\mu}) \geq h(f, \mu)$. We prove the reverse inequality. Let $\epsilon > 0$.

$\hat{\mu}$ being ergodic, we have that (see [19, p. 72])

$$h(\hat{f}, \hat{\mu}) = \sup_{\hat{Q}} \limsup_{n \to \infty} \frac{1}{n} \log r(\hat{Q}, n, \hat{\mu}), \tag{2.6}$$

where $\hat{Q}$ ranges over a family of finite measurable partitions of $\hat{X}$ which is stable by finite joining and the union of which generates the measurable subsets, and $r(\hat{Q}, n, \hat{\mu})$ is the minimal number of elements of $\hat{Q}^n$ necessary to have a measure at least c, where c is any constant with $0 < c < 1$.

We shall consider the finite measurable partitions $\hat{Q}$ of the form $\hat{\pi}^{-1}(Q) \vee \hat{T}$ with Q some finite partition of X and $\hat{T}$ the partition of $X \times \mathbb{N}$ generated by the subsets $X \times \{i\}$ for $i = 0, 1, \ldots, r-1$ for some $r < \infty$. We fix such a partition $\hat{Q}$. Set $\beta > 0$ so small that

$$C_n^{2\beta n} \leq e^{\epsilon n} \quad \text{and} \quad \beta < \epsilon / 2 \log \#\hat{Q}.$$

Pick some $d_0 \in \mathbb{N}$ such that $\hat{\mu}([d_0]) > 0$. Here we have used the notation

$$[d_0 \cdots d_n] \stackrel{\text{def}}{=} \{\hat{x} \in \hat{X}: \pi_2(\hat{f}^k(\hat{x})) = d_k \text{ for } 0 \leq k \leq n\}$$

(a cylinder on $\hat{X}$).

Choose a non-decreasing sequence of finite subsets $G_n \subset \mathbb{N}$, $n \geq 1$, with the property that:

$$\hat{\mu}(S_n) := \hat{\mu}\left(\bigcup_{d_1,\ldots,d_n \in G_n} [d_0 \cdots d_n]\right) > \hat{\mu}([d_0])(1 - 1/4 - 1/2 - \cdots - 1/2^{n+1})$$
$$> \hat{\mu}([d_0])/2.$$

Set $S_* = \bigcap_{n\geq 1} S_n$. Clearly $\hat{\mu}(S_*) > 0$. By the ergodicity of $\hat{\mu}$, if M is large enough, then $\bigcup_{m=0}^{M} \hat{f}^{-m}(S_*)$ has measure at least $1-\beta$. Now pick $N \geq \log \#\hat{Q} \cdot \epsilon^{-1} M$ so large that $C_n^{3n/N} \leq e^{\epsilon n}$.

Let R be the finite and measurable partition of X generated by Q and by the sets $\hat{\pi}(X \times \{i\} \cap \hat{f}^{-1}(X \times \{j\}))$ for $i, j \in G_N \cup \{d_0\}$. Remark that $\hat{\pi}$ restricted to $X \times \{i\}$, for any $i \in \mathbb{N}$, is one-to-one. Hence, for $\hat{x} \in S_N$, the element of R^N which contains x completely determines the element of $\hat{Q}^N$ containing $\hat{x}$. Now

$$h(f,\mu) \geq \limsup_{n\to\infty} \frac{1}{n} \log r(R, n, \mu).$$

Hence, if n is very large, there exists $A \subset X$ such that $\mu(A) > 3/4$ and A is the union of at most $e^{n(h(f,\mu)+\epsilon)}$ elements of R^n.

By Birkhoff's ergodic theorem, there exist $\hat{A} \subset \hat{X}$ such that $\hat{\mu}(\hat{A}) > 3/4$ and an integer K such that, for $n \geq K$ and $\hat{x} \in \hat{A}$

$$\frac{1}{n}\#\left\{0 \leq k \leq n-1 \colon \hat{f}^k(\hat{x}) \in \bigcup_{m=0}^{M} \hat{f}^{-m} S_N\right\} > 1 - 2\beta.$$

Set $\hat{B} = \hat{A} \cap \hat{\pi}^{-1}(A)$. $\hat{\mu}(\hat{B}) > 1/2$. Now take $\hat{x} \in \hat{B}$. To specify the element of $\hat{Q}^n$ which contains $\hat{x}$, it is enough to know the following:

(1) the element of R^n containing $\hat{\pi}(\hat{x})$;
(2) the position of disjoint segments of times $[n_i - m_i, n_i + N]$ included in $[0, n-1]$ such that $\hat{f}^{n_i}(\hat{x}) \in S_N$ and $0 \leq m_i \leq M$;
(3) the element of $\hat{Q}$ containing $\hat{f}^k(\hat{x})$ for k not included in such a segment or contained in a sub-segment $[n_i - m_i, n_i - 1]$.

Bounding from above the number of possibilities for each of these three choices, we get that the number of elements of $\hat{Q}^n$ necessary to cover $\hat{B}$ is at most

$$e^{n(h(f,\mu)+\epsilon)} \cdot C_n^{3n/N} \cdot \#\hat{Q}^{(2\beta+M/N)n+(M+N)} \leq \text{const} \cdot e^{n(h(f,\mu)+4\epsilon)}.$$

Thus, $h(\hat{f},\hat{\mu},\hat{Q}) \leq h(f,\mu)+4\epsilon$. But $\epsilon>0$ is arbitrarily small and $\hat{Q}$ ranges in a generating, stable-by-joining family of partitions. ∎

Thus the bad set $\hat{\mathcal{X}} \setminus \hat{\mathcal{X}}'$ is h-negligible.

Theorem A is proved.

3. Isomorphism with the Markov chain

Let $(\Sigma_+(\mathcal{D}),\sigma)$ be the one-sided topological Markov chain defined by the Markov diagram $\mathcal{D}$:

$$\Sigma_+(\mathcal{D}) = \{D \in \mathcal{D}^{\mathbb{N}}\colon \forall p \in \mathbb{N}\ D_p \to D_{p+1} \text{ on } \mathcal{D}\}$$

and

$$\sigma\colon (D_p)_{p\in\mathbb{N}} \longmapsto (D_{p+1})_{p\in\mathbb{N}}.$$

Let $\pi_{\mathcal{D}}\colon \hat{x} \in \hat{X} \mapsto (\pi_{\mathcal{D}}(\hat{f}^n\hat{x}))_{n\in\mathbb{N}} \in \Sigma_+(\mathcal{D})$ be the coding map, defined on the set of $\hat{x} \in \hat{X}$ with complete positive orbits. It extends to a map between the natural extensions. Recall that

$$\widetilde{X \times P} = \overline{\{(x,A) \in Y^{\mathbb{N}} \times P^{\mathbb{N}}\colon \forall n \in \mathbb{N} \quad x_n \in A_n \text{ and } x_{n+1} = f(x_n)\}} \subset X^{\mathbb{N}} \times P^{\mathbb{N}},$$

this set being endowed with the left-shift σ.

THEOREM B: *Let (X,P,f) be a piecewise invertible dynamical system. Assume that:*

(H1) *P h-separates,*

(H2″) *the extended boundary: $\partial\widetilde{X\times P} \stackrel{\text{def}}{=} \{(x,A) \in \widetilde{X\times P}\colon x_0 \in \partial P\}$ is h-negligible.*

Then the coding map $\pi_{\mathcal{D}}$ defines a h-isomorphism between $(\hat{X},\hat{f})$ and $\Sigma_+(\mathcal{D})$.

Using Katok's entropy formula [13] one sees that, given some invariant and ergodic probability measure μ of $(\widetilde{X\times P},\sigma)$ such that $\mu(\partial\widetilde{X\times P})>0$, one has $h(\mu) \leq h_{\mathrm{top}}(\partial\widetilde{X\times P},\sigma)$. But this last quantity is the border entropy $h_B(P,f)$.

Hence, condition (H2) of the Main Theorem implies condition (H2″) above: Theorem B implies the corresponding part of the Main Theorem.

In fact, we shall build an isomorphism between $(\hat{X},\hat{f})$ and $\Sigma_+(D)$ modulo h-negligible sets (this is slightly stronger than an isomorphism between the natural extensions, which is the definition of h-isomorphism).

It is convenient to consider the left-shift σ acting on $\widetilde{X\times\mathcal{D}}$ defined as

$$\overline{\{(x,D) \in Y^{\mathbb{N}} \times \Sigma_+(\mathcal{D})\colon \forall n \in \mathbb{N} \quad x_{n+1} = f(x_n) \text{ and } x_n \in D_n\}} \subset X^{\mathbb{N}} \times \mathcal{D}^{\mathbb{N}}.$$

Remark that

$$\hat{X} \subset \widetilde{X \times \mathcal{D}}$$

modulo the injection $\hat{x} \mapsto (f^n(x), \hat{\pi}_{\mathcal{D}}(\hat{f}^n \hat{x})))_{n \in \mathbb{N}}$, defined for all $\hat{x}$ which have a complete $\hat{f}$-orbit.

Conversely, it is easy to see that

$$\widetilde{X \times \mathcal{D}} \setminus \hat{X} \subset \bigcup_{k \geq 0} \sigma^{-k}(\partial \widetilde{X \times \mathcal{D}})$$

with $\partial \widetilde{X \times \mathcal{D}}$ the set of points $(x, D) \in \widetilde{X \times \mathcal{D}}$ such that $x_0 \in \partial P$. We prove that $\widetilde{X \times \mathcal{D}} \setminus \hat{X}$ is h-negligible. Let μ be an invariant and ergodic probability measure of $(\widetilde{X \times \mathcal{D}}, \sigma)$ such that $\mu(\widetilde{X \times \mathcal{D}} \setminus \hat{X}) > 0$. We have to prove that μ has small entropy, i.e., entropy smaller than a constant $H < h(f)$.

Obviously, $\mu(\partial \widetilde{X \times \mathcal{D}}) > 0$. By Proposition 2.8, μ projects on $\widetilde{X \times P}$, to a measure with the same entropy, as this projection is countable-to-one. Now $\partial \widetilde{X \times \mathcal{D}}$ projects to $\partial \widetilde{X \times P}$, hence this last set has positive measure for the projected measure. By (H2''), the projected measure has small entropy.

Thus we have proved that $(\hat{X}, \hat{f})$ and $(\widetilde{X \times \mathcal{D}}, \sigma)$ are h-isomorphic.

We now turn to the h-isomorphism between $(\widetilde{X \times \mathcal{D}}, \sigma)$ and the topological Markov chain $(\Sigma_+(\mathcal{D}), \sigma)$.

We have the projection

$$(x, D) \in \widetilde{X \times \mathcal{D}} \longmapsto D \in \Sigma_+(\mathcal{D}).$$

It is onto: given $D \in \Sigma_+(\mathcal{D})$, $n \geq 0$, $D_0 \cap f^{-1}(D_1) \cap \cdots \cap f^{-n}(D_n)$ is a non-empty open set, hence it meets $\bigcap_{k \geq 0} f^{-k}(Y)$, the G_δ-dense set of points in X which have a complete orbit. Pick some point x_n in the intersection. Let $(x^n, D^n) \in Y^{\mathbb{N}} \times \mathcal{D}^{\mathbb{N}}$ be defined by $x_k^n = f^k(x_n)$ for $k \geq 0$, $D_0^n = D_0$ and, for each $k \geq 0$, D_{k+1}^n is the successor of D_k^n containing $f^k(x^n)$. Using the relative compactness of the elements of P, (x^n, D^n) converges to (x^*, D) for some $x^* \in X^{\mathbb{N}}$, maybe after passing to a subsequence. Thus $(x^*, D) \in \widetilde{X \times \mathcal{D}}$ and D belongs to the projection of $\widetilde{X \times \mathcal{D}}$.

This proves that the projection is onto. The problem is the defect in injectivity.

We remark that $h_{\text{top}}(\Delta P, f) \leq h_B(P, f)$; therefore, by Theorem A, (X, f) and $(\hat{X}, \hat{f})$ are h-isomorphic. Hence, by what we have just shown, $(\widetilde{X \times \mathcal{D}}, \sigma)$ and (X, f) are h-isomorphic. In particular, the condition (H1) (P h-separates) implies that the non-injectivity set of $\pi_{\mathcal{D}}$, i.e., the set of $(x, D) \in \widetilde{X \times \mathcal{D}}$ such that $\bigcap_{n \geq 0} f^{-n} D_n \supsetneq \{x\}$, is h-negligible within $\widetilde{X \times \mathcal{D}}$.

It remains to see that the image by π_D of the non-injectivity set is h-negligible within $\Sigma_+(\mathcal{D})$. It is enough to see that if some invariant and ergodic probability measure on $\Sigma_+(\mathcal{D})$ gives a positive measure to this set, then it can be lifted to a measure on $\widetilde{X \times \mathcal{D}}$ which gives a positive measure to the non-injectivity set. Therefore the lifted measure (and thus the original measure on $\Sigma_+(\mathcal{D})$) will have small entropy. It is enough to apply the following:

FACT: *Let μ be some Borelian invariant probability measure on the shift $(\mathcal{A}^{\mathbb{N}}, \sigma)$ over some countable set $\mathcal{A}$. Let F be some metric space, and fix, for each $A \in \mathcal{A}$, some continuous map g_A: $F_A \to F$ defined on some compact subset F_A of F. Define the measurable endomorphism $\sigma \times g$ of $\mathcal{A}^{\mathbb{N}} \times F$ by*

$$(\sigma \times g)(a, x) = (\sigma(a), g_{a_0}(x)).$$

Assume that there exists $(a^*, x^*) \in \mathcal{A}^{\mathbb{N}} \times F$ such that a^* is μ-generic and that, for all $n \geq 0$, $(\sigma \times g)^n(a^*, x^*)$ is defined, i.e., if $(b, y) = (\sigma \times g)^n(a^*, x^*)$ then $y \in F_{b_0}$.

Then, there exists at least one $(\sigma \times g)$ invariant probability measure ν on $\mathcal{A}^{\mathbb{N}} \times F$ projecting to μ.

The proof of this fact is standard: take ν to be an accumulation point, in the vague topology, of the sequence of measures:

$$\nu_n = \frac{1}{n} \sum_{k=0}^{n-1} \delta_{(\sigma \times g)^k (a^*, x^*)}.$$

Let $\epsilon > 0$. If the finite subset $\mathcal{A}_*$ of $\mathcal{A}$ is large enough then, setting $[\mathcal{A}_*] = \{a \in \mathcal{A}^{\mathbb{N}}\colon a_0 \in \mathcal{A}_*\}$, $\mu([\mathcal{A}_*] \cap \sigma^{-1}[\mathcal{A}_*]) > 1 - \epsilon$ and therefore $\nu_n([\mathcal{A}_*] \cap \sigma^{-1}[\mathcal{A}_*] \times F) > 1 - \epsilon$, for all large n. It follows that, for K_* the compact union of the images $g_A(F_A)$, $A \in \mathcal{A}_*$,

$$\nu([\mathcal{A}_*] \times K_*) \geq \limsup_{n \to \infty} \nu_n([\mathcal{A}_*] \times K_*) \geq 1 - \epsilon.$$

Letting ϵ decrease to zero, we see that ν is a probability measure.

This also implies that $\nu(h \circ (\sigma \times g) \cdot \chi_{[\mathcal{A}_*] \times K_*}) \leq \nu(h) + \|h\|_\infty \epsilon$ for an arbitrary continuous function h: $\mathcal{A}^{\mathbb{N}} \times F \to \mathbb{R}$ with compact support, giving the $(\sigma \times g)$-invariance of ν.

This proves the h-isomorphism of $\widetilde{X \times \mathcal{D}}$ with $\Sigma_+(\mathcal{D})$.

In conclusion, we have proved the h-isomorphism of $(\hat{X}, \hat{f})$ with the topological Markov chain $\Sigma_+(\mathcal{D})$ defined by the Markov diagram: this is Theorem B.

4. Small entropy at infinity

We prove the following uniform version of the statement in the Main Theorem:

THEOREM C: *Let $(X,P,f), f \in \mathcal{F}$, be a family of piecewise invertible dynamical systems. Define*

$$H_\Delta \stackrel{\text{def}}{=} \limsup_{n\to\infty} \frac{1}{n} \log \max_{f\in\mathcal{F}} \#\{A \in P^n \colon A \cap \Delta P = \emptyset\},$$

and let H_P be the supremum, for all $f \in \mathcal{F}$, of the entropies of the invariant and ergodic probability measures μ of (X,f) such that P does not separate μ-almost all orbits.

Then, for every $H > \max(H_\Delta, H_P)$, there exists $N < \infty$ such that, for all $f \in \mathcal{F}$,

$$h\Big(\Sigma\Big(\mathcal{D} \setminus \bigcup_{\ell=1}^{N} \mathcal{D}_\ell\Big)\Big) \leq H.$$

The proof of the corresponding statement in the Main Theorem is similar to the proof below and is therefore omitted.

Remark 4.1: We have the following estimate for H_Δ. If X is compact and f extends to a continuous map on the whole of X: for every $\epsilon > 0$, there exists $\delta > 0$ and $C < \infty$ such that

$$\#\{A \in P^n \colon A \cap \Delta P = \emptyset\} \leq C \cdot r(\delta, n, \partial P) \mathrm{mult}(P^n) e^{\epsilon n}.$$

We defer the proof to the Appendix. ∎

The reason for this smallness of the entropy is the following. Recall that the **level** of $D \in \mathcal{D}$ is the smallest ℓ such that $D \in \mathcal{D}_\ell$. One has the following: an orbit in $\hat{X}$ which remains at very high levels in the Markov diagram projects on X to an orbit made up of very long pieces of orbits starting in the boundary of P. Indeed, we have the following variant of P-shadowing:

PROPOSITION 4.2: *Let N be a positive integer. Let $\hat{x} \in \hat{X}$ be such that*

$$\pi_{\mathcal{D}}(\hat{f}^k \hat{x}) \notin \bigcup_{\ell=1}^{N} \mathcal{D}_\ell \quad \forall 0 \leq k < n.$$

Then there exist integers $n_0 < n_1 < \cdots < n_s = n$ such that

$$n_{i+1} - n_i > N \quad \text{and} \quad P_{n_{i+1}-n_i+1}(f^{n_i}(x)) \cap \Delta P \neq \emptyset \qquad \text{for all } i = 0, 1, \ldots, s-1,$$

i.e., each $[n_i, n_{i+1}]$ is a P-shadowing interval for $x = \hat{\pi}(\hat{x}) \in X$ with respect to ΔP.

Here n_0 is non-positive: $f^{n_0}(\hat{x})$ is to be understood as some pre-image of x by $f^{|n_0|}$; $|n_0|$ is bounded above by the level of $\hat{x}$. The other n_i's are positive.

Proof: Let r be the level of $\hat{x}$. By definition, $\pi_{\mathcal{D}}(\hat{x}) \in \mathcal{D}$ is the image by f^{r-1} of some set $A \in P_r$, i.e., A is a connected component of some r-cylinder. Define $f^{-r}(x)$ to be the unique pre-image of x by f^r in this set. Define accordingly $f^{-r+i}(x) = f^i(f^{-r}(x))$ for $0 \le i \le r$.

Using $\hat{\pi}_{\mathcal{D}}(\hat{x}) = f^r(P_{r+1}(f^{-r}(x)))$ and the formula (2.1) for $\hat{f}^n$ we get that

$$f^{n+r}(P_{n+r+1}(f^{-r}(x))) = \pi_{\mathcal{D}}(\hat{f}^n(\hat{x})).$$

Now take $0 \le k \le n+r$ minimal such that $F_k := f^k(P_{k+1}(f^{n-k}(x))) = \hat{\pi}_{\mathcal{D}}(\hat{f}^n(\hat{x}))$. By the previous claim such a k exists. As $\pi_{\mathcal{D}}(\hat{f}^n(\hat{x})) \in \mathcal{D}_{k+1}$, $k \ge N \ge 1$. Notice that the sets F_k are non-increasing with respect to k. As k is minimal we have that $F_k \subsetneq F_{k-1}$. But F_k is the image by f^{k-1} of some connected component of $f(A) \cap P_k(f^{n-k+1}(x))$ for some $A \in P$. Therefore, $f(A)$ meets but does not contain $P_k(f^{n-k+1}(x))$. This last set being connected, we get that

$$\partial f(A) = f_A(\partial A) \text{ meets } P_k(f^{n-k+1}(x)).$$

Hence

$$P_k(f^{n-k+1}(x)) \cap \Delta P \neq \emptyset.$$

Setting $n_s = n$ and $n_{s-1} = n-k+1$, an easy induction on n completes the proof. ∎

We turn to the proof of Theorem C.

We are given $H > \max(H_\Delta, H_P)$. Let $\epsilon \stackrel{\text{def}}{=} (H - H_\Delta)/4$. Pick N so large that

$$\#\{A \in P^n\colon A \cap \Delta P \neq \emptyset\} \le \exp(n(H - 3\epsilon)) \qquad \forall n \ge N \quad \forall f \in \mathcal{F}$$

and that $C_n^{2n/N} \le e^{\epsilon n}$ for $n \ge 0$. Fix now $f \in \mathcal{F}$.

It is enough to bound the entropy of invariant and ergodic probability measures $\hat{\mu}$ carried by $\Sigma(\mathcal{D} \setminus \bigcup_{k \le N} \mathcal{D}_k)$, the natural extension of the corresponding one-sided topological Markov chain. We recall that $h(\hat{f}, \hat{\mu}, \hat{\pi}^{-1}(P)) = h(\mathcal{F}, \mu, P)$ by Proposition 2.8 (here $\mu = \hat{\pi}_* \hat{\mu}$ and P is identified with its counterpart in the natural extension). We may assume that P separates μ-almost all orbits, as otherwise $h(\mu) \le H_P < H$. This implies that: $h(\hat{f}, \hat{\mu}, \hat{\pi}^{-1}(P)) = h(\hat{f}, \hat{\mu}) = h(\mathcal{F}, \mu, P) = h(\mathcal{F}, \mu)$.

Proposition 4.2 is easily seen to imply that, for μ-almost every $x \in \mathcal{X}$, there exists $n(x) \geq N$ such that $[-n(x), 0]$ is a P-shadowing interval for x with respect to ΔP; $n(\cdot)$ can be defined in a measurable way. Let N_0 be so large that $n(\cdot) \leq N_0$ on a set of measure greater than $1 - \epsilon/\log \#P$.

By the ergodic theorem, there exist $E \subset \mathcal{X}$ with $\mu(E) > 1/2$ and an integer $L \geq \epsilon^{-1} \log \#P \cdot N_0$ such that, for all $x \in E$ and $n \geq L$,

$$\frac{1}{n}\#\{0 \leq k < n\colon n(\mathcal{F}^k x) \leq N_0\} > 1 - \epsilon/\log \#P.$$

Thus, we can divide $[0, n-1]$ into P-shadowing intervals with length between N and N_0 leaving out a fraction of $[0, n-1]$ at most $\epsilon/\log \#P + N_0/L \leq 2\epsilon/\log \#P$.

Hence to specify the element of P^n containing x_k for $0 \leq k < n$, it is enough to give:

(1) the position of the shadowing intervals;
(2) for each one of these, the element of P^l corresponding to it (l being the length of the shadowing interval);
(3) the element of P containing x_k for the times k outside these shadowing intervals.

Thus we get at most

$$C_n^{2n/L} \cdot e^{n(H-3\epsilon)} \cdot \#P^{2\epsilon/\log \#P \cdot n} \leq e^{nH}$$

possibilities. Hence, using the entropy formula (2.6),

$$h(\hat{f}, \hat{\mu}) = h(\mathcal{F}, \mu, P) \leq H,$$

proving Theorem C. ∎

Appendix

Proof of Remark 1.5: We claimed that, for X compact and f continuous on X,

$$h_B(P, f) \leq h_{\text{top}}(\partial P, f) + h_{\text{mult}}(P, f).$$

Let $\epsilon > 0$. By replacing (X, P, f) with (X, P^N, f^N) for N large enough we may assume that $\log \text{mult}(P) \leq h_{\text{mult}}(P, f) + \epsilon$ (the elements of P^N are not necessarily connected but it does not matter here). We endow P^N with the distance $d(A_1, \ldots, A_N; B_1, \ldots, B_N) = \max_k d_P(A_k, B_k)$, where d_P is the distance on P.

Recall that we have identified the metric space (P, d_P) with $\{1, 1/2, 1/3, \dots\}$. For simplicity, we now pretend $N = 1$.

Writing $P = \{P_1, P_2, P_3, \dots\}$, let Q be the finite partition

$$\left\{ P_1, P_2, \dots, P_{[2\epsilon^{-1}]+1}, \bigcup_{k>[2\epsilon^{-1}]+1} P_k \right\}.$$

Observe that the corresponding partition of P has a d_P-diameter less than $\epsilon/2$.

As X is compact and Q is finite, there exists $0 < r < \epsilon/2$ such that every ball with radius r meets at most $\mathrm{mult}(P)$ elements of Q.

Let $n \geq 1$. Let C_n be an (r, n)-cover of ∂P with $C_n \leq \mathrm{const} \cdot e^{n(h_{\mathrm{top}}(\partial P, f)+\epsilon)}$. Consider $A = [A_0 \cdots A_{n-1}] \in Q^n$ such that $\overline{A}$ meets ∂P. Fix some $x \in \overline{A} \cap \partial P$. Pick $y \in C_n$ such that $d_n(x, y) < r$. A_k meets $B(f^k(y), r)$ for each $k = 0, 1, \dots, n-1$. Therefore, when x ranges in the (r, n)-ball around a given y, there is at most $\mathrm{mult}(P)$ choices for each $A_k \in Q$. But each $A_0 \cdots A_{n-1}$ determines an $(\epsilon/2, n)$-ball in P^n. Hence, we have found an (ϵ, n)-cover for $\widetilde{\partial X \times P}$, so that the claim is proved as $\epsilon > 0$ was arbitrary. ∎

As for Remark 4.1, that is, for P finite and f continuous,

$$\#\{A \in P^n \colon A \cap \Delta P = \emptyset\} \leq C \cdot r(\delta, n, \partial P) \mathrm{mult}(P^n) e^{\epsilon n};$$

it follows from a proof entirely similar to the one just given. ∎

References

[1] R. Bowen, *Entropy for group endomorphisms and homogeneous spaces,* Transactions of the American Mathematical Society **153** (1971), 401–414.

[2] J. Buzzi, *Entropies et représentation markovienne des applications régulières de l'intervalle,* Thèse, Université Paris-Sud, Orsay, 1995.

[3] J. Buzzi, *Intrinsic ergodicity of smooth interval maps,* Israel Journal of Mathematics **100** (1997), 125–161.

[4] J. Buzzi, *Intrinsic ergodicity of affine maps on* $[0,1]^d$, Monatshefte für Mathematik **124** (1997), 97–118.

[5] J. Buzzi, *Entropy, volume growth and Lyapunov exponents,* submitted.

[6] J. Buzzi, *Ergodicité intrinsèque de produits fibrés d'applications unidimensionnelles chaotiques,* Bulletin de la Société Mathématique de France **126** (1998), 51–77.

[7] J. Buzzi, *Absolutely continuous invariant measures for generic piecewise affine expanding maps*, International Journal of Bifurcation and Chaos, to appear.

[8] M. Denker, C. Grillenberg and K. Sigmund, *Ergodic theory on compact spaces*, Lecture Notes in Mathematics **527**, Springer-Verlag, Berlin, 1976.

[9] B. M. Gurevič, *Topological entropy of enumerable Markov chains*, Soviet Mathematics Doklady **10** (1969), 911–915.

[10] B. M. Gurevič, *Shift entropy and Markov measures in the path space of a denumerable graph*, Soviet Mathematics Doklady **11** (1970), 744–747.

[11] F. Hofbauer, *On intrinsic ergodicity of piecewise monotonic transformations with positive entropy, I*, Israel Journal of Mathematics **34** (1979), 213–237; *II*, **38** (1981), 107–115.

[12] F. Hofbauer, *Piecewise invertible dynamical systems*, Probability Theory and Related Fields **72** (1986), 359–386.

[13] A. Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Publications Mathématiques de l'Institut des Hautes Études Scientifiques **51** (1980), 137–173.

[14] G. Keller, *Markov extensions, zeta functions, and Fredholm theory for piecewise invertible dynamical systems*, Transactions of the American Mathematical Society **314** (1989), 433–497.

[15] G. Keller, *Lifting measures to Markov extensions*, Monatshefte für Mathematik **108** (1989), 183–200.

[16] S. E. Newhouse, *On some results of Hofbauer on maps on the interval*, Proceedings of the conference *Dynamical systems and related topics* held in Nagoya in 1990 and published as volume 9 of the *Advanced Series in Dynamical Systems*, World Science Publishing, River Edge, NJ, 1991, pp. 407–421.

[17] D. Newton and W. Parry, *On a factor automorphism of a normal dynamical system*, Annals of Mathematical Statistics **37** (1966), 1528–1533.

[18] W. Parry, *Entropy and Generators in Ergodic Theory*, W. A. Benjamin, New York, 1969.

[19] D. J. Rudolph, *Fundamentals of Measurable Dynamics*, Clarendon Press, Oxford, 1990.

[20] B. Weiss, *Intrinsically ergodic systems*, Bulletin of the American Mathematical Society **76** (1970), 1266–1269.

[21] Y. Yomdin, *Volume growth and entropy*, Israel Journal of Mathematics **57** (1987), 285–300.