



HAL
open science

Maximal Entropy Measures for Piecewise Affine Surface Homeomorphisms

Jerome Buzzi

► **To cite this version:**

Jerome Buzzi. Maximal Entropy Measures for Piecewise Affine Surface Homeomorphisms. 2007.
hal-00171744v1

HAL Id: hal-00171744

<https://hal.science/hal-00171744v1>

Preprint submitted on 13 Sep 2007 (v1), last revised 19 Feb 2009 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MAXIMAL ENTROPY MEASURES FOR PIECEWISE AFFINE SURFACE HOMEOMORPHISMS

JÉRÔME BUZZI

ABSTRACT. We study the dynamics of piecewise affine surface homeomorphisms from the point of view of their entropy. Under the assumption of positive topological entropy, we establish the existence of finitely many ergodic and invariant probability measures maximizing entropy and prove a multiplicative lower bound for the number of periodic points. This is intended as a step towards the understanding of surface diffeomorphisms. We proceed by building a jump transformation, using not first returns but carefully selected “good” returns to dispense with Markov partitions. We control these good returns through some entropy and ergodic arguments.

CONTENTS

1. Introduction	2
1.1. Definitions and Statements	2
1.2. Outline of the Proof	4
1.3. Some Comments	5
2. Pointwise Estimates	5
2.1. Symbolic Dynamics	5
2.2. Invariant Manifolds and Lyapunov Exponents	7
2.3. Semi-Uniform Estimates	9
3. Construction of the Markov Structure	10
3.1. Markov Rectangles	11
3.2. Hyperbolic strips	14
3.3. Admissible Strips and Good Returns	17
4. Analysis of Large Return Times	18
4.1. Geometric Analysis	18
4.2. Combinatorial Estimates	19
4.3. Large Average Return Times and Entropy	21
5. Proof of the Main Results	26
5.1. Maximum Entropy Measures	26
5.2. Number of Periodic Points	28
Appendix A. Bounds on Metric Entropy	29
Appendix B. Tower Lifts	30
Appendix C. Examples	32
References	34

Date: 3/14/2007.

Key words and phrases. Dynamical systems; ergodic theory; topological entropy; variational principle; measures with maximum entropy; Markov structure; Markov tower; symbolic dynamics; coding.

1. INTRODUCTION

Robust entropy conditions can often be shown to imply a non-uniform but global hyperbolic structure with strong consequences for periodic points and probability measures with maximum entropy. This has been shown in the case of the *entropy-expanding property* [4, 8, 9] for C^∞ smooth maps in all dimensions, including plane maps of the type $(x, y) \mapsto (1.9 - x^2 + \epsilon y, 1.8 - y^2 + \epsilon x)$ for small ϵ and *all C^∞ interval maps with nonzero topological entropy*. In particular, such maps $T : M \rightarrow M$ have a finite number of ergodic invariant probability measures with maximum entropy (**maximum measures**, for short) and satisfy a **multiplicative lower bound** on the number of their periodic points:

$$\liminf_{n \rightarrow \infty, p|n} e^{-nh_{\text{top}}(T)} \#\{x \in M : T^n x = x\} > 0.$$

Remark that these results were first proved by Hofbauer [13, 14] for piecewise monotone maps on the interval. Also, the previously quoted papers contain slightly stronger results (e.g., a complete classification up to "entropy-conjugacy").

The techniques used in the above mentioned papers do not apply to diffeomorphisms (a diffeomorphism is never entropy-expanding). However, many properties of interval maps generalize to surface diffeomorphisms so the following is generally expected:

Conjecture 1. *Consider a $C^{1+\epsilon}$, $\epsilon > 0$, smooth diffeomorphism of a compact surface.*

*The collection of ergodic and invariant probability measures with maximum entropy is **countable** (possibly finite or empty) and the periodic points satisfy a multiplicative lower bound if there exists at least one measure with maximum entropy.*

Conjecture 2. *Consider a C^∞ smooth diffeomorphism of a compact surface.*

*The collection of ergodic and invariant probability measures with maximum entropy is **finite** and the periodic points satisfy a multiplicative lower bound.*

Recall that, by a result of S. Newhouse [20], all C^∞ smooth maps of compact manifolds have at least one measure of maximum entropy. Also a classical theorem of A. Katok [15] states that, if T is a $C^{1+\epsilon}$, $\epsilon > 0$, diffeomorphism of a compact surface M , the number of periodic points satisfies a **logarithmic lower bound**:

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \#\log\{x \in M : T^n x = x\} \geq h_{\text{top}}(T).$$

This paper presents the proof of the analogue of Conjecture 2 in the easier case of piecewise affine homeomorphisms.

1.1. Definitions and Statements. Let M be a compact two-dimensional manifold possibly with boundary, **affine** in the following sense. There exists a distinguished atlas with charts:

- identifying the neighborhood of any point of M with an open subset of $\{(x, y) \in \mathbb{R}^2 : x \geq 0 \ \& \ y \geq 0\}$;
- inducing affine changes of coordinates.

These charts are called the *affine charts*. The phenomena we are interested are independent of the global topology, so we could in fact restrict ourselves to the special cases $M = \mathbb{T}^2$ or $M = [0, 1]^2$.

A map continuous $T : M \rightarrow M$ is said to be **piecewise affine** if there exists a finite partition P of M such that for every $A \in P$, A and $T(A)$ are contained in affine charts which maps them to polygons of \mathbb{R}^2 with non-empty interiors and $T : A \rightarrow T(A)$ is affine w.r.t. these affine charts. It is convenient to replace the partition P by the collection \tilde{P} of the interiors of the elements of P . Such a partition \tilde{P} (a partition up to the boundaries of its elements) is called an **admissible partition** with respect to T . We drop the tilde in the sequel.

Let us recall some facts about entropy for the convenience of the reader (we refer to [26, 10] for further information). The entropy of a non-necessarily invariant subset $K \subset M$ is a measure of the “number of orbits” starting from K . Recall that the ϵ, n -ball at $x \in M$ is: $\{y \in M : \forall k = 0, 1, \dots, n-1 : d(T^k y, T^k x) < \epsilon\}$. The entropy of K is, according to Bowen [2]:

$$h(T, K) := \lim_{\epsilon \rightarrow 0} h(T, K, \epsilon) \text{ with } h(T, K, \epsilon) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log r(\epsilon, n, K)$$

where $r(\epsilon, n, K)$ is the minimum number of ϵ, n -balls with union containing K . The **topological entropy** is $h_{\text{top}}(T) := h(T, M)$.

The entropy of an ergodic and invariant probability measure μ can be defined similarly, according to [15]:

$$h(T, \mu) := \lim_{\epsilon \rightarrow 0} h(T, \mu, \epsilon) \text{ with } h(T, \mu, \epsilon) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log r(\epsilon, n, \mu)$$

where $r(\epsilon, n, \mu)$ is the minimum number of ϵ, n -balls whose union have a μ -measure at least λ , for an arbitrary constant $\lambda \in (0, 1)$.

The **variational principle** states that for $T : M \rightarrow M$ (in fact for any continuous self-map of a compact metric space):

$$(1.1) \quad h_{\text{top}}(T) = \sup_{\mu} h(T, \mu)$$

where μ ranges over the T -invariant and ergodic probability measures. A by-product of our first investigations, is the following combinatorial expression for this entropy.

The first step of the analysis follows from observations of S. Newhouse:

Proposition 1.1. *Let T be a piecewise affine homeomorphism of a compact surface. The topological entropy of T is given by:*

$$(1.2) \quad h_{\text{top}}(T) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log \#\{[A_0 \dots A_{n-1}] : A_i \in P\}.$$

It was also obtained by D. SANDS and Y. ISHII by different methods.

The variational principle (1.1) brings to the fore the ergodic and invariant probability measures μ such that $h(T, \mu) = \sup_{\nu} h(T, \nu)$. We call them **maximum measures**.

Corollary 1.2. (of the proof) *A piecewise affine homeomorphism of a compact surface has at least one maximum measure.*

Our main result is:

Theorem 1. *Let $T : M \rightarrow M$ be a piecewise affine homeomorphism of a compact affine surface. Assume that $h_{\text{top}}(T) > 0$. Then there are finitely many ergodic, invariant probability measures maximizing the entropy (or maximum measures).*

We also obtain as a by-product:

Proposition 1.3. *Let $T : M \rightarrow M$ be a piecewise affine homeomorphism of a compact affine surface. We have a multiplicative lower bound on the number of periodic points: there exists an integer $p \geq 1$ such that:*

$$(1.3) \quad \liminf_{n \rightarrow \infty, p|n} \#\{x \in M : T^n x = x\} e^{-nh_{\text{top}}(T)} > 0.$$

1.2. Outline of the Proof. We are able to analyze **large entropy measures**, i.e. invariant and ergodic probability measures with entropy close enough to the supremum. We rely on **semi-uniform estimates**, that is uniform estimates holding on subsets of lower bounded measure w.r.t. any large entropy measure and discard **entropy-negligible subsets**, i.e., subsets which have zero measure w.r.t. any large entropy measure.

We begin by the pointwise estimates of Section 2 dealing with local or semi-local properties of individual orbits. We introduce the *symbolic dynamics* showing that it has the same entropies as T . We establish that the *(local) stable $W^s(x)$ and unstable $W^u(x)$ manifolds* of points $x \in M$, i.e., the sets of points with the same past or future P -itineraries, are line segments. We prove semi-uniform lower bounds for their lengths and angles. A corollary is that the large entropy measures live off the boundary of the partition.

Section 3 is devoted to building a Markov structure representing the large entropy dynamics. We first build arrays of *Markov rectangles* which contain a significant proportion of the dynamics and are very small w.r.t. the sizes of local manifolds so they have "small holes", that is, only a small proportion of points that lie in one of them have local manifolds that don't cross the rectangle. Our techniques require replacing T by some high power T^L at this point.

We then define hyperbolic strips following the geometric picture of Markov rectangles usual in uniformly hyperbolic dynamics. We provide tools to build many such strips around typical orbits of large entropy measures. These hyperbolic strips are Markov in the sense that they can be freely concatenated as soon as they end and begin in the same rectangle.

We conclude Section 3 by the construction of the Markov structure by selecting among these hyperbolic strips a set of *admissible* ones to get a (weak) uniqueness property in the decomposition of an orbit into those (this weakness will require a more detailed ergodic analysis in Section 5). We obtain a notion of *good return times*.

The core of our analysis is in Section 4. We first relate such return times to geometric and combinatorial properties involving the visits to the Markov rectangles and the complement set of their holes. We deduce from this that invariant measures w.r.t. which the average return time is large must have small entropy.

Finally, Section 5 proves the main results by using the above Markov structure. The proof of Theorem 1 analyzes of the jump transformation (or more exactly, the tower) defined by the good return times to lift large entropy measures of T to a countable state Markov shift. We show that this lift is a finite extension and in particular preserves maximum measures so we can apply results of Gurevic [11] to conclude the proof of the Theorem.

Showing Proposition 1.3 uses a classical estimate of Vere-Jones [25] on the number of periodic sequences in the previous Markov shift together with a combinatorial argument to bring it back to T .

The three appendices (A) recall some well-known upper-bounds on measure-theoretic entropy, (B) prove a lifting theorem for the tower defined by a return time and (C) give some examples of piecewise affine maps.

1.3. Some Comments. In the setting of this paper, the distortion of diffeomorphism is replaced by more easily understood singularities. Beyond the analysis of a simple and natural class of dynamics, we illustrate an alternative approach to Markov partitions: we ask less of geometry (tolerating "holes" in our rectangles) by using more ergodic theory and entropy estimates to accommodate the resulting non-uniqueness of representation. More precisely, the "holes" are avoided by using generalized induction (that is, one waits for a good return, not necessarily the first return to the rectangle –see Appendix B). This directly yields a Markov graph with a simpler structure than those obtained by Hofbauer for piecewise monotonic maps or in [6]. However the relation between T and the Markov dynamics is more indirect.

Strengthening this link and proving that the Markov shifts obtained are sufficiently well-behaved (strongly positive recurrent, see [12]) would yield further results like a classification w.r.t. large entropy measures by the topological entropy [3] or a meromorphic extension of Artin-Mazur zeta function defined by periodic points [9].

It would be natural to apply the techniques of this paper to more general dynamics. A first direction is that of more general chaotic piecewise affine maps. Most questions are still open despite some partial results (see, e.g., [7, 24, 18]) and we should stress that new problems immediately appear. From the point of view of entropy alone:

- there exist piecewise affine continuous *maps* on surfaces and piecewise affine homeomorphisms in *dimension 3* for which the right hand side of (1.2) is strictly larger than the entropy (see Examples 1, 3 in the Appendix C);
- The example 4 in Appendix C is a piecewise affine *discontinuous map* on a surface with no maximum measure (one can give a continuous, piecewise quadratic version of it, see example 5). However, I don't know examples of continuous piecewise affine maps without maximum measures.

For diffeomorphisms, the main difficulty with our approach is to find a link between short stable/unstable manifolds and small entropy, e.g., one would need to relate small Lyapunov charts to entropy bounds for smooth diffeomorphism.

2. POINTWISE ESTIMATES

2.1. Symbolic Dynamics. An admissible partition P for the map T defines a symbolic dynamics which, in addition to its intrinsic interest, will be a key tool until we are able to show that ∂P has zero measure w.r.t. any $\mu \in \mathbb{P}_{\text{erg}}^0(T)$, where $\mathbb{P}_{\text{erg}}^h(T)$ denotes the set of ergodic, invariant probability measures of T with $h(T, \mu) > h$.

Definition 2.1. $x \in M$ is **nice** if for every $n \in \mathbb{Z}$, $T^n x$ belongs to an element A_n of the admissible partition P . The sequence $A \in P^{\mathbb{Z}}$ thus defined is the **P -itinerary** of x .

The symbolic dynamics of T, P is:

$$\Sigma := \overline{\{A \in P^{\mathbb{Z}} : \exists x \in M \forall n \in \mathbb{Z} T^n x \in A_n\}}$$

endowed with the shift map: $\sigma(A) = (A_{n+1})_{n \in \mathbb{Z}}$.

Observe that Σ being a subshift, it admits at least one maximum entropy measure by standard results (see, e.g., [26]). Hence, a “close enough” relation between the invariant measures of Σ and T will imply existence of an maximum measure also for T . By the variational principle, we shall also get that T and Σ have the same topological entropy. The well-known formula for the topological entropy of the subshift Σ :

$$h(\Sigma) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \#\{[A_0 \dots A_{n-1}] := \{x \in \Sigma : x|_0^n = A_0^n\} : A \in P^n\}$$

where $x|_0^n := x_0 \dots x_{n-1}$ will then prove the Misiurewicz-Slenk formula for T , finishing the proof of Proposition 1.1.

Neither of these two systems being an extension of the other in general, it is convenient to introduce the following common extension:

$$\Sigma \times M := \overline{\{(A, x) \in P^{\mathbb{Z}} \times M : \forall n \in \mathbb{Z} T^n x \in \text{int} A_n\}}$$
 with $\hat{T}(A, x) = (\sigma A, Tx)$.

The close relation between the measures of T and Σ alluded to above is:

Lemma 2.2. *Both maps $\pi_1 : \Sigma \times M \rightarrow \Sigma$ and $\pi_2 : \Sigma \times M \rightarrow M$ are entropy preserving: for every invariant probability measure μ on $\Sigma \times M$, $h(\sigma, \pi_1 \mu) = h(T, \pi_2 \mu) = h(\hat{T}, \mu)$. Moreover, π_1 and π_2 induce onto maps between the sets of (ergodic) invariant probability measures.*

In particular, the topological entropies of the three systems must be equal by the variational principle.

The proof of the above Lemma rests on two geometric/combinatorial properties. The first is the following observation by S. Newhouse, very specific of our setting (it is false in higher dimensions or without the invertibility assumption, see the Appendix):

Lemma 2.3. *The multiplicity entropy [5]:*

$$h_{\text{mult}}(T) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log \max_{x \in M} \text{mult}(P^n, x) \text{ with } \text{mult}(Q, x) := \#\{A \in Q : x \in \overline{A}\}$$

is zero for any piecewise affine homeomorphism of a surface.

The second is a property of linear maps:

Lemma 2.4. *Let $d \geq 1$. For each $n \geq 0$, let $T_n : \mathbb{R}^d \rightarrow \mathbb{R}^d$ be a linear map. Then*

$$\lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \max\{\#S : \forall 0 \leq k < n \text{ diam}(T_{k-1} \dots T_1 T_0 S) \leq 1 \text{ and} \\ \forall x \neq y \in S \exists 0 \leq k < n \|T_{k-1} \dots T_1 T_0(x - y)\| > \epsilon\} = 0.$$

We leave the easy proofs of Lemmas 2.3 and 2.4 to the reader.

PROOF OF LEMMA 2.2: Lemma 2.3, resp. Lemma 2.4, implies that for all $x \in M$, resp. Σ , for $i = 2$, resp. $i = 1$,

$$h(\hat{T}, \pi_i^{-1}\{x\}) = 0.$$

Now, $\pi_1 : \Sigma \times M \rightarrow \Sigma$ and $\pi_2 : \Sigma \times M \rightarrow M$ are both compact topological extensions. Hence, one can apply Bowen's result [2]:

$$h(\hat{T}, \hat{\mu}) = h(\sigma, \pi_1 \mu) = h(T, \pi_2 \mu)$$

for all invariant probability measures $\hat{\mu}$ of $\Sigma \times M$. \square

2.2. Invariant Manifolds and Lyapunov Exponents. The partition P being given, we have the following:

Definition 2.5. *The stable manifold at $A \in \Sigma$ is the following set (convex in the affine charts):*

$$W^s(A) := \bigcap_{n \geq 1} \overline{T^{-n}A_n}$$

The **unstable manifolds** $W^u(A)$ and $W^u(x)$ are defined by replacing $n \geq 1$ by $n \leq -1$ in the above equation.

The forward (+) and backward (-) upper **Lyapunov exponents** along $A \in \Sigma$ are:

$$\lambda_{\pm}^u(A) := \limsup_{n \rightarrow \pm\infty} \frac{1}{n} \log \|(T_A^n)'^{\pm 1}\| \quad \text{and} \quad \lambda_{\pm}^s(A) := \limsup_{n \rightarrow \pm\infty} \frac{1}{n} \log \|(T_A^n)'^{\mp 1}\|^{-1}$$

where T_A^n is the affine composition $(T|_{A_{n-1}}) \circ \dots \circ (T|_{A_0})$ (if $n \geq 0$) or $[(T|_{A_{-1}}) \circ \dots \circ (T|_{A_n})]^{-1}$ (if $n < 0$).

If $x \in M$ is nice then it defines a unique itinerary A and one writes $W^s(x)$ for $W^s(A)$, $\lambda_{\pm}^u(x)$ for $\lambda_{\pm}^u(A)$ and so on.

The first goal of this section is the following ‘‘non-singularity’’ result:

Proposition 2.6. *Let $\mu \in \mathbb{P}_{\text{erg}}^0(T)$. The following holds:*

- $\mu(\partial P) = 0$ (in particular, μ -a.e. $x \in M$ is nice);
- $\lambda_{-}^s(x) = \lambda_{+}^s(x) \leq -h(T, \mu) < 0 < h(T, \mu) \leq \lambda_{+}^u(x) = \lambda_{-}^u(x)$ for μ -a.e. $x \in M$ (so we can omit the index \pm);
- $W^s(x)$ and $W^u(x)$ are line segments containing x in their relative interiors $\text{int } W^s(x)$ and $\text{int } W^u(x)$ for μ -a.e. $x \in M$.

To prove Proposition 2.6, let $\mu \in \mathbb{P}_{\text{erg}}^0(T)$. As we have not yet proved that a.e. $x \in M$ is nice, we have to work in the extension $\Sigma \times M$ to be able to speak of itineraries, invariant manifolds and so on. By compactness, there exists an invariant and ergodic probability measure $\hat{\mu}$ of $\hat{T} : \Sigma \times M \leftrightarrow \Sigma \times M$ such that $\pi_2 \hat{\mu} = \mu$. We have $h(\hat{T}, \hat{\mu}) > 0$ by Lemma 2.2.

We first consider the invariant manifolds.

Claim 2.7. $\hat{\mu}$ -a.e. $(A, x) \in \Sigma \times M$, (i) $W^u(A)$ is a line segment; (ii) x is not an endpoint of this segment.

PROOF OF THE CLAIM: To begin with, observe that $W^u(\sigma A) \subset T(W^u(A))$ so that $\dim(W^u(\sigma A)) \leq \dim(W^u(A))$. As $\hat{\mu}$ is invariant and ergodic, $\dim(W^u(A))$ must be $\hat{\mu}$ -a.e. equal to a constant d 0, 1 or 2. Claim (i) above is that $d = 1$.

Let \hat{P} be the natural partition of $\Sigma \times M$ (coming from the canonical partition of Σ). As in the proof of Lemma 2.2, $h(\hat{T}, \hat{\mu}) = h(\hat{T}, \hat{\mu}, \hat{P}) = H_{\hat{\mu}}(\hat{P}|\hat{P}^-)$ where $\hat{P}^- := \bigvee_{n \geq 1} T^n \hat{P}$. Observe that $A \mapsto W^u(A)$ is \hat{P}^- -measurable.

We exclude the cases $d = 0, 2$ by contradiction. Assume first $d = 0$, i.e., $W^u(A)$ is a single point $x \in M$ for $\hat{\mu}$ -a.e. $(A, x) \in \Sigma \times M$. This implies that:

$$h(\hat{T}, \hat{\mu}) = H_{\hat{\mu}}(\hat{P}|\hat{P}^-) = \lim_{n \rightarrow \infty} \frac{1}{n} H_{\hat{\mu}}(\hat{P}^n|\hat{P}^-) \leq h_{\text{mult}}(T, P) = 0,$$

a contradiction, excluding the case $d = 0$.

Now assume $d = 2$, so $\hat{\mu}$ -a.e. $\overline{W^u(A)} = \overline{\text{int}(W^u(A))}$. By definition, two distinct unstable manifolds have disjoint interiors. Therefore, there exists a countable collection $W^u(A^1), W^u(A^2), \dots$ such that for $\hat{\mu}$ -a.e. $A \in \Sigma$, $W^u(A) = W^u(A^i)$ for some $i \geq 1$. In particular, $W^u(A) = W^u(A^{i_0})$ on a set of positive measure for some i_0 . By Poincaré recurrence, there exists an integer $n > 0$ such that $T^n(W^u(A^{i_0})) = W^u(A^{i_0})$. This implies that $\pi_1 \hat{\mu}$ is periodic, hence $0 = h(\sigma, \pi_1 \hat{\mu}) = h(\hat{T}, \hat{\mu})$. This contradiction proves (i).

We turn to (ii). If $x \in \partial W^u(A)$, then $T(x) \in \partial W^u(\sigma(A))$. Thus if (ii) is false, then $x \in \partial W^u(A)$ μ -a.e. But this implies that, for any $\epsilon > 0$, any large n ,

$$nh(\hat{T}, \hat{\mu}) = H_{\hat{\mu}}(\hat{P}^n|\hat{P}^-) \leq \log 2 + \log \max_{x \in M} \#\{A \in P^n : \bar{A} \ni x\} \leq \log 2 + (h_{\text{mult}}(T, P) + \epsilon)n.$$

As $h_{\text{mult}}(T, P) = 0$, it would follow that $h(\hat{T}, \hat{\mu}) = 0$, a contradiction. \square

We now turn to the exponents. First observe that $\lambda_+^u(A) = \lambda_-^u(A)$ $\lambda_+^s(A) = \lambda_-^s(A)$ for $\hat{\mu}$ -a.e. $A \in \Sigma$ is part of the classical Oseledec's Theorem (see, e.g., [16]).

Claim 2.8. *For $\pi_1 \hat{\mu}$ -a.e. $A \in \Sigma$, the Lyapunov exponents satisfy: $\lambda^s(A) < 0 < \lambda^u(A)$.*

Remark 2.9. *The above result will be a consequence of Ruelle-Margulis inequality [16] once we shall have proved that $\mu(\partial P) = 0$.*

PROOF: We establish the existence of a positive Lyapunov exponent μ -a.e. The existence of a negative exponent will follow by considering T^{-1} . Consider the family of norms $\|\cdot\|'_A, A \in \Sigma$ defined from an arbitrary one $\|\cdot\|_A, A \in \Sigma$ by:

$$\|v\|'_A := \|v\|_A / \|W^u(A)\|_A \quad \text{for } v \in W^u(A)$$

$\|\cdot\|_A$ being the length w.r.t. to $\|\cdot\|_A$ (this is meaningful because of the affine structure of M). As $T(W^u(A)) \supset W^u(\sigma A)$, we have that $\|T'|E^u(A)\|'_A \geq 1$ (where $E^u(A)$ is the unstable direction at A –the *invariant* family of directions defined by $W^u(A)$ for μ -a.e. $A \in \Sigma$. $T(W^u(A)) = W^u(\sigma A)$ μ -a.e. would imply $h(\hat{T}, \hat{\mu}) = H_{\hat{\mu}}(\hat{P}|\hat{P}^-) = 0$. Hence, $\|T'|E^u(A)\|'_A > 1$ on a set of positive measure and:

$$\lambda^u(A) = \int \log \|T'|E^u(B)\|'_B d\hat{\mu}(B) > 0$$

for μ -a.e. $A \in \Sigma$. \square

PROOF OF PROPOSITION 2.6: Let $\mu \in \mathbb{P}_{\text{erg}}^0(T)$. Let $\hat{\mu}$ be a lift of μ to $\Sigma \times M$. By Lemma 2.2, $h(\hat{T}, \hat{\mu}) = h(T, \mu) > 0$.

Claims 2.7 and 2.8 prove all the claims of the Proposition except $\mu(\partial P) = 0$.

Now, $W^u(A)$ and $W^s(A)$ are line segments a.e. by Claim 2.7. Their directions carry distinct Lyapunov exponents by Claim 2.8, hence they must make a.e. a non-zero angle. If $x \in \partial P$, then x would be the end point of at least one of these line segments, a contradiction. Hence $\mu(\partial P) = 0$. \square

That $\mu(\partial P) = 0$ for all ergodic invariant probability measures with nonzero entropy has the following immediate but important consequence:

Corollary 2.10. *The partially defined map $\pi : \Sigma' \rightarrow M$ with*

$$\{\pi(x)\} := \bigcap_{n \geq 0} \overline{T^n[A_{-k} \dots A_k]}$$

with Σ' the subset of Σ where the above intersection is indeed reduced to a single point, defines an entropy-preserving bijection between the sets of ergodic, invariant probability measures of T and of Σ with nonzero entropy.

2.3. Semi-Uniform Estimates. We obtain now more quantitative estimates, which we call **semi-uniform** in the sense that they are uniform on a set of uniformly lower-bounded mass for all large entropy measures. To state these results, we need the following ‘‘distortion’’ bound. By compactness of M and invertibility of T ,

$$d(T) := \sup \left\{ \log \frac{\|T'(x).u\|}{\|T'(x).v\|} : x \in M, u, v \in \mathbb{R}^2 \setminus \{0\}, \|u\| = \|v\| \right\} < \infty.$$

Proposition 2.11. *For any $\mu_0 < \frac{h_{\text{top}}(T)}{d(T)}$, there exist $h_0 < h_{\text{top}}(T)$, $\theta_0 > 0$ and $\ell_0 > 0$ such that for any $\mu \in \mathbb{P}_{\text{erg}}^{h_0}(T)$, the following properties occur jointly on a set of measure at least μ_0 :*

$$(2.1) \quad \rho(x) := \min_{\sigma=s,u} d(x, \partial W^\sigma(x)) \geq \ell_0$$

$$(2.2) \quad \alpha(x) := \angle(W^s(x), W^u(x)) > \theta_0$$

Here $\angle(W^s(x), W^u(x))$ is the angle between the two lines defined by $W^s(x)$ and $W^u(x)$. We declare $\alpha(x) = \rho(x) = 0$, if $W^s(x)$ or $W^u(x)$ fail to be line segments.

Remark 2.12. *We in fact obtain a measure $\mu_0 < 1$ arbitrarily close to 1 satisfying (2.1). However this is not the case wrt (2.2). Indeed, one can easily build a smooth surface diffeomorphism with nonzero entropy such that for some $\mu_0 > 0$ and $h_0 > 0$, there are invariant probability measures with entropy at least h_0 such that the stable and unstable directions make an arbitrarily small angle on a set of measure at least μ_0 . We do not know if these measures can be taken to have entropy arbitrarily close to the topological entropy or if piecewise affine examples exist.*

We first prove the lower bound on angles by comparing the distortion with the entropy.

Claim 2.13. *For any $0 < h_1 < h_{\text{top}}(T)$, there exists $\theta_1 > 0$ such that the set where $\alpha(x) > \theta_1$ has measure at least $h_1/d(T)$ for all measures $\mu \in \mathbb{P}_{\text{erg}}^{h_1}$.*

Ruelle-Margulis inequality applied to (T, μ) and (T^{-1}, μ) (which is valid as T' is uniformly continuous on each element of P and $\mu(\partial P) = 0$) yields:

$$(2.3) \quad h_{\text{top}}(T) \leq \frac{\lambda^u(\mu) - \lambda^s(\mu)}{2} = \frac{1}{2} \int_M \log \frac{\|T'(x)|E^u(x)\|}{\|T'(x)|E^s(x)\|} d\mu(x).$$

By continuity there exists $\theta_1 > 0$ such that, for all $u, v \in \mathbb{R}^2 \setminus \{0\}$ with $\angle(u, v) \leq \theta_1$,

$$\forall x \in M \setminus \partial P \quad \log \frac{\|T'(x).u\|}{\|T'(x).v\|} \leq h_1.$$

Therefore, setting $m := \mu(\{x \in M : \alpha(x) > \theta_1\})$:

$$2h(T, \mu) \leq m \cdot d(T) + (1 - m) \cdot h_1$$

so that, assuming $h(T, \mu) > h_1$:

$$m > \frac{2h(T, \mu) - h_1}{d(T) - h_1} \geq \frac{2h(T, \mu) - h_1}{d(T)} \geq \frac{h_1}{d(T)}.$$

This proves Claim 2.13.

Claim 2.14. *For any $\mu_3 < 1$, there exists $\ell_0 > 0$ such that*

$$(2.4) \quad \forall \mu \in \mathbb{P}_{\text{erg}}^{h_3}(T) \quad \mu(\{x \in M : d(x, \partial W^u(x)) > \ell_0\}) > \mu_3$$

for $h_3 = h_{\text{top}}(T)(1 - (1 - \mu_3)/2)$.

To prove (2.4) let $\epsilon = (1 - \mu_3)h_{\text{top}}(T)/2 > 0$, Λ be a Lipschitz constant for T , $n \geq \log 2/\epsilon$ be a large integer and $r = r(\epsilon, n) > 0$ be such that

$$\max_{x \in M} \#\{A \in P^n : B(x, r) \cap A \neq \emptyset\} \leq \frac{1}{2}e^{(h_{\text{mult}}(T, P) + \epsilon)n} \quad \text{and} \quad \#P^n \leq e^{(h_{\text{top}}(T) + \epsilon)n}.$$

Let $\mu \in \mathbb{P}_{\text{erg}}^{h_3}(T)$, $X_0 := \{x \in M : d(x, \partial W^u(x)) \leq \Lambda^{-n}r\}$ and denote by $\mu|_{X_0}$ the normalized restriction of μ to X_0 . Using standard facts about entropy (see Appendix A) we get

$$\begin{aligned} nh(T, \mu) &= H_\mu(P^n|P^-) \leq H_\mu(P^n \vee \{X_0, M \setminus X_0\}|P^-) \leq H_\mu(\{X_0, M \setminus X_0\}) \\ &\quad + \mu(X_0)H_{\mu|_{X_0}}(P^n|P^-) + (1 - \mu(X_0))H_{\mu|_{M \setminus X_0}}(P^n|P^-) \\ &\leq \log 2 + \mu(X_0) \sup_{x \in X_0} \log \#\{A \in P^n : A \cap X_0 \cap W^u(x) \neq \emptyset\} \\ &\quad + (1 - \mu(X_0)) \log \#P^n \\ &\leq \log 2 + \mu(X_0)(h_{\text{mult}}(T, P) + \epsilon)n + (1 - \mu(X_0))(h_{\text{top}}(T) + \epsilon)n. \end{aligned}$$

Hence

$$\begin{aligned} h(T, \mu) &\leq (1 - \mu(X_0))h_{\text{top}}(T) + \mu(X_0)h_{\text{mult}}(T, P) + \epsilon + \frac{1}{n} \log 2 = \\ &\quad h_{\text{top}}(T) + 2\epsilon - \mu(X_0)(h_{\text{top}}(T) - h_{\text{mult}}(T, P)). \end{aligned}$$

implying that:

$$\mu(X_0) \leq \frac{h_{\text{top}}(T) - h(T, \mu) - \epsilon}{h_{\text{top}}(T) - h_{\text{mult}}(T, P)} \leq \frac{h_{\text{top}}(T) - h(T, \mu) + \epsilon}{h_{\text{top}}(T)} < 1 - \mu_3$$

using $h_{\text{mult}}(T, P) = 0$ and $h(T, \mu) > h_3$. The claim is proved.

PROOF OF PROPOSITION 2.11: Claim 2.13 gives $\theta_0 > 0$ such that (2.2) holds on a set of measure at least $h_{\text{top}}(T)/2d(T)$ wrt all measures in $\mathbb{P}_{\text{erg}}^{h_{\text{top}}(T)/2}(T)$. Claim 2.14 applied to T and T^{-1} with $\mu_3 = 1 - h_{\text{top}}(T)/8d(T)$, shows that for

$$h_0 = h_{\text{top}}(T) \left(1 - \frac{1}{16} \frac{h_{\text{top}}(T)}{8d(T)}\right) \geq h_{\text{top}}(T)/2,$$

(2.2) and (2.1) hold jointly on a set of measure at least $h_{\text{top}}(T)/4d(T)$ w.r.t. all measures in $\mathbb{P}_{\text{erg}}^{h_3}(T)$. \square

3. CONSTRUCTION OF THE MARKOV STRUCTURE

The estimates of the previous section will allow us to build a collection of (non-uniform) ‘‘Markov rectangles’’ which will ‘‘control enough’’ of the dynamics to analyze all measures of large entropy.

3.1. Markov Rectangles.

Definition 3.1. A (Markov) rectangle¹ is a closed topological disk R contained in an affine chart and bounded by four line segments, alternatively included in stable and unstable manifolds, making respectively the unstable boundary, $\partial^u R = \partial_1^u R \cup \partial_2^u R$, and the stable one, $\partial^s R = \partial_1^s R \cup \partial_2^s R$. See Fig. 2.

A **Markov array** is a finite collection of Markov rectangles with disjoint interiors.

Not every passage of an orbit inside a rectangle is useful. We need the following properties.

Definition 3.2. A point x is **controlled** by a rectangle R if x is nice, belongs to R and if $W^s(x)$ and $W^u(x)$ each intersects ∂R in two points. $x \in R$ is **10-controlled** if moreover $\rho(x) > 10 \text{diam} R$. $x \in R$ is **s -controlled** if x is nice, $x \in R$ and $W^s(x)$ intersects ∂R in two points.

The set of controlled, 10-controlled, s -controlled points is denoted by $\kappa(R)$, $\kappa_{10}(R)$, $\kappa_s(R)$.

A point is controlled by a Markov array \mathcal{R} if it is controlled by one of the rectangles of the array. We define $\kappa(\mathcal{R})$, $\kappa_{10}(\mathcal{R})$, $\kappa_s(\mathcal{R})$ in the obvious way.

Using the previous lower bounds on the lengths and angles of invariant manifolds we shall first prove:

Lemma 3.3. There exist numbers $h_0 < h_{\text{top}}(T)$ and $\mu_0 > 0$ and a Markov array \mathcal{R} such that for all $\mu \in \mathbb{P}_{\text{erg}}^{h_0}(T)$,

$$\mu(\kappa_{10}(\mathcal{R})) > \mu_0.$$

Our analysis requires the following slightly stronger statement (i.e., we only tolerate "small holes"):

Lemma 3.4. There is $\mu_0 > 0$ such that for any $\epsilon_0 > 0$, there exist a number $h_0 < h_{\text{top}}(T)$ and a Markov array \mathcal{R} such that for any $\mu \in \mathbb{P}_{\text{erg}}^{h_0}(T)$,

- $\mu(\kappa_{10}(\mathcal{R})) > \mu_0$;
- $\mu(\mathcal{R} \setminus \kappa_{10}(\mathcal{R})) < \epsilon_0 \mu_0$.

This will be obtained by subdividing the rectangles in the Markov array from Lemma 3.3 into sub-rectangles much smaller than most stable/unstable manifolds.

The final twist is that as we replace the partition P by the convex partition $P^{\mathcal{R}}$ generated by P and the Markov array \mathcal{R} (see Fig. 1), some invariant manifolds may shrink, say $W_{\mathcal{R}}^u(x) := \bigcap_{n \geq 1} \overline{T^n P^{\mathcal{R}}(T^{-n}x)} \subsetneq W^u(x)$, diminishing the set of controlled points. Indeed, $W_{\mathcal{R}}^u(f(x)) \subsetneq W^u(f(x))$ when $W^u(x)$ crosses the boundary of a rectangle from \mathcal{R} . We shall see however that if these intersections are sufficiently separated in time, then $W_{\mathcal{R}}^u(x) = W^u(x)$ for most points $x \in \mathcal{R}$ w.r.t. large entropy measures. To guarantee that large separation, we use the following:

Definition 3.5. If \mathcal{R} is an array of Markov rectangles contained in an element of P and L is a positive integer, the **(\mathcal{R}, L) -extension** of (M, T, P, \mathcal{R}) is $(M_+, T_+, P_+, \mathcal{R}_+)$, defined in the following way:

- $M_+ = M \times \{0, \dots, L-1\}$;

¹Obviously these are quadrilaterals, but "Markov rectangle" is the usual terminology in the uniform hyperbolic theory. Our apologies to true geometers.

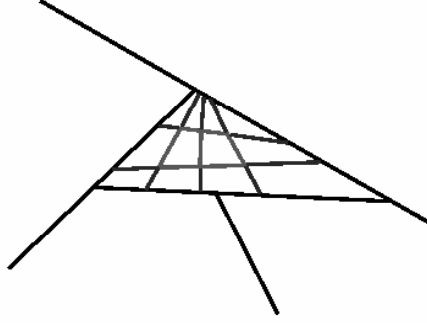


FIGURE 1. The convex partition \tilde{P} refining both P (in black) and \mathcal{R} (in red).

- $T_+(x, k) = (Tx, k + 1 \pmod L)$;
- P_+ is the finite partition of M_+ which coincides with a copy of P on each $M \times \{k\}$ for $k \neq 0$ and coincides on $M \times \{0\}$ with a copy of $P^{\mathcal{R}}$;
- $\mathcal{R}_+ = \{R \times \{0\} : R \in \mathcal{R}\}$.

Observe that the dependence on L of the above extension is trivial. The conclusion of this section is:

Proposition 3.6. *Let (M, T, P) be a piecewise affine surface homeomorphism with nonzero entropy. There exist $\mu_0 > 0$, $h_0 < h_{\text{top}}(T)$ such that for any $\epsilon_0 > 0$ and any $L_0 < \infty$, there is a Markov array \mathcal{R} and a positive integer L_0 with the following properties. Fix any $L_+ \geq L_0$ and let $(M_+, T_+, P_+, \mathcal{R}_+)$ be the (\mathcal{R}, L_+) -extension of (M, T, P, \mathcal{R}) .*

For each $\mu \in \mathbb{P}_{\text{erg}}^{h_0}(T)$, there exists an ergodic invariant probability measure μ_+ of T_+ with $\pi(\mu_+) = \mu$ (where $\pi(x, k) = x$) such that, w.r.t. the invariant manifolds defined by P_+ and the Markov array \mathcal{R}_+ :

- (i) $L_+ \cdot \mu_+(\kappa_{10}(\mathcal{R}_+)) > \mu_0$;
- (ii) $L_+ \cdot \mu_+(\mathcal{R}_+ \setminus \kappa_{10}(\mathcal{R}_+)) < \epsilon_0 \cdot \mu_0$

It is enough to prove Theorem 1 for any such extension (it is compact, so all invariant probability measures lift and it is finite-to-one so it respects entropy). This is also the case w.r.t. Proposition 1.3.

We now prove Lemmas 3.3-3.4 and Proposition 3.6. We begin by the following:

Lemma 3.7. *Given $\ell_0 > 0$ and $0 < \theta_0 < 2\pi$, there exists a finite collection of rectangles $R^{(1)}, \dots, R^{(Q)}$ such that:*

- (1) $\text{diam}(R^{(i)}) < \ell_0/10$;
- (2) any $x \in M$ with $\rho(x) > \ell_0$ and $\angle(W^u(x), W^s(x)) > \theta_0$ belongs to at least one $R^{(i)}$.

This easily implies Lemma 3.3 using Proposition 2.11 and observing that the finite collection of rectangles above can be subdivided by boundary lines like those of Fact 3.9 below so that their interiors become disjoint, defining the required Markov array \mathcal{R} .

PROOF OF LEMMA 3.7: Let

$$K_* := \{x \in M : \rho(x) > \ell_0 \text{ and } \angle(W^u(x), W^s(x)) > \theta_0\}.$$

Let $\{K_j\}_{j=1}^Q$ be a finite partition of K_* whose elements have diameter less than $\theta_0\ell_0/100$ and lie within an affine chart of M . We fix j .

Recall that the collection of closed subsets \mathcal{K} of the compact metric space M is a compact space w.r.t. the Hausdorff metric:

$$d(A, B) = \inf\{\epsilon > 0 : A \subset B(B, \epsilon) \text{ and } B \subset B(A, \epsilon)\}$$

The easy proofs of the following two facts are left to the reader.

Fact 3.8. *Let $A^n \in \Sigma(T, P)$ converge to A_+ . By taking a subsequence, $W^s(A^n)$ must also converge in the Hausdorff metric, say to $H \subset M$. Then $H \subset W^s(A_+)$.*

Fact 3.9. *Assume that K_j is as above. Then there exist two points $x_1, x_2 \in \overline{K_j}$, two non-trivial line segments L_1, L_2 and two sequences $A^1, A^2 \in \Sigma(T, P)$ with the following properties.*

- L_i is contained in the boundary of $W^s(A^i)$ as a subset of M ;
- in some affine chart, K_j lies between the two lines supporting L_1 and L_2 .

We call L_1, L_2 a pair of stable boundary lines of K_j .

Consider two one-dimensional stable manifolds $W^s(A)$ and $W^s(B)$ which intersect in a single point p . p must be the endpoint of at least one of them: otherwise, if $A_n \neq B_n$, $p \in \partial A_n \cap \partial B_n$ and both $W^s(A)$ and $W^s(B)$ must be parallel to the same segment of $\partial A_n \cap \partial B_n$, giving a large intersection. Hence, for arbitrary $x, y \in K_j$, $W^u(x)$ and $W^u(y)$ must have disjoint relative interiors or be parallel. Thus $1 - \cos \angle(W^u(x), W^u(y)) \leq (\text{diam} K_j)^2 / 2\ell_0^2$ so

$$\angle(W^u(x), W^u(y)) \leq \theta_0/50.$$

As $\angle(W^u(z), W^s(z)) > \theta_0$ for all $z \in K$, we get:

$$\angle(W^u(x), W^s(y)) > \theta_0/2.$$

Consider a pair of "stable boundary lines", resp. "unstable boundary lines", given by Fact 3.9 applied to (T, K_j) , resp. applied to (T^{-1}, K_j) . Let $R^{(j)}$ be the rectangle bounded by these four line segments. $R^{(j)}$ is contained in the intersection of two strips with almost parallel sides of width $\leq \text{diam} K_j$ and making an angle at least $\theta_0/2$. Hence

$$\text{diam}(R^{(j)}) < 5\text{diam}(K_j)/\theta_0 < \ell_0/20.$$

On the other hand, $R^{(j)} \supset K_j$, hence $\bigcup_j R^{(j)} \supset K_*$. □

PROOF OF LEMMA 3.4: Apply Lemma 3.3 to get \mathcal{R} , $\mu_0 > 0$ and $h_0 < h_{\text{top}}(T)$. Recall that $\rho(x)$ is the distance between x and the endpoints of its invariant manifolds (or 0 if one of those is not a line segment).

By Claim 2.14 applied with $\mu_+ := 1 - \epsilon_0\mu_0/4$ to T and T^{-1} , there exist $h_1 < h_{\text{top}}(T)$ and $\ell_1 = \ell_1(\epsilon_0\mu_0)$ such that

$$\forall \mu \in \mathbb{P}_{\text{erg}}^{h_1}(T) \quad \mu(\{x \in M : \rho(x) < \ell_1\}) < \epsilon_0\mu_0.$$

Let us cut each big rectangle R from \mathcal{R} into sub-rectangles R' with diameter at most $\ell_1/10$, obtaining a new Markov array \mathcal{R}' . Using Fact 3.9 again, we can do it by finitely many stable and unstable manifolds (or line segments bounding those) going across the rectangles in \mathcal{R} . Observe that $\kappa_{10}(\mathcal{R}') \supset \kappa_{10}(\mathcal{R})$ and that the

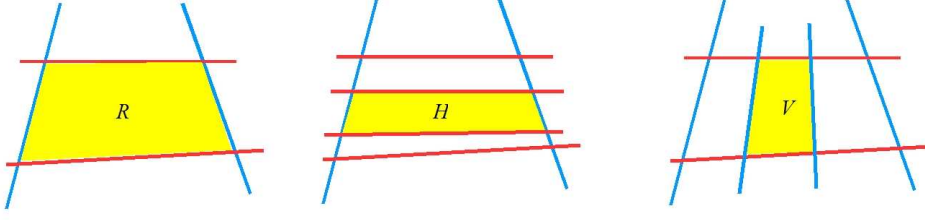


FIGURE 2. From left to right: a rectangle R , a s -rectangle H and a u -rectangle V . The (approximately) horizontal, red, resp. vertical, blue line segments are segments of stable, resp. unstable manifolds.

points in $\mathcal{R}' \setminus \kappa_{10}(\mathcal{R}')$ which have line segments as invariant manifolds, are ℓ_1 -close to an endpoint of their stable/unstable manifold. Hence

$$\mu(\kappa_{10}(\mathcal{R}')) \geq \mu_0 \text{ and } \mu(\mathcal{R}' \setminus \kappa_{10}(\mathcal{R}')) \leq \epsilon_0 \mu_0.$$

for all $\mu \in \mathbb{P}_{\text{erg}}^{h_1}(T)$. \square

PROOF OF PROPOSITION 3.6: We apply Lemma 3.4 with $\epsilon_0/2$ obtaining $\mu_0 > 0$ (independent of ϵ_0), $h_0 < h_{\text{top}}(T)$ and a Markov array \mathcal{R} . Let $P^{\mathcal{R}}$ be the convex partition previously defined. We go to the (\mathcal{R}, L_+) -extension (M_+, T_+, P_+) of (M, T, P) for some large integer L_+ to be specified. As we observed, there always exists an ergodic, T_+ -invariant measure μ_+ extending μ . Maybe after replacing it by its image under $(x, i) \mapsto (x, i + j \bmod L_+)$ for some j , we get an ergodic extension μ_+ such that

$$L_+ \cdot \mu_+(\kappa_{10}(\mathcal{R}) \times \{0\}) \geq \mu(\kappa_{10}(\mathcal{R})) \geq \mu_0.$$

As the extension is finite-to-one, μ_+ has the same entropy as μ .

Let $x \in M$. If the unstable manifold for T_+ , $W_+^u(x, 0) := \bigcap_{n \geq 1} T^n P_+(T_+^{-n}(x, 0))$ is strictly shorter than $W^u(x) \times \{0\}$, then it is bounded by $(y, 0)$ with y an intersection point of $W^u(T^{-kL_+}x)$ for some $k \geq 1$, with one of the new boundary segments, I , of $\partial P^{\mathcal{R}}$. Hence $P^{kL_+}(T^{-L_+}x)$ is determined by the past of $T^{-kL_+}x$ and I picked among finitely many choices, say m (depending only on P and \mathcal{R} but not on L_+).

A standard counting argument shows that if this happened on a subset of $M \times \{0\}$ with μ_+ -measure at least $\frac{1}{2}\epsilon_0\mu_0 \cdot L_+^{-1}$, then

$$h(T, \mu) = h(T_+, \mu_+) \leq (1 - \frac{1}{2}\epsilon_0\mu_0)h_{\text{top}}(T) + \epsilon(L_+)$$

where $\epsilon(L) \rightarrow 0$ as $L \rightarrow \infty$. This is strictly less than $h_{\text{top}}(T)$ if L_+ is large enough (which we ensure by taking L_0 large). So it is excluded for large entropy measures: the (\mathcal{R}, L) -periodic extension $(M_+, T_+, P_+, \mathcal{R}_+)$ has the required properties for all large integers L_+ . \square

3.2. Hyperbolic strips. (M, T, P) is some piecewise affine surface homeomorphism with nonzero entropy and \mathcal{R} is some Markov array with $\mathcal{R} \subset P$ (eventually (M, T, P, \mathcal{R}) will be the previously built periodic extension $(M_+, T_+, P_+, \mathcal{R}_+)$). We shall use the following picture adapted from uniformly hyperbolic dynamics to define finite itineraries that can be freely concatenated.

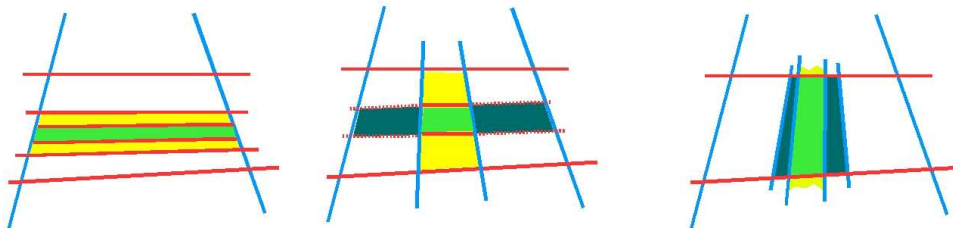


FIGURE 3. Proof of Lemma 3.12. From left to right: (a) $P_0^n(x) \subset P_0^n(x) \subset R$; (b) the u -rectangle $T^m(P_0^n(x))$ crossing the s -rectangle $P_m^n(x)$ (both inside R'); (c) $T^n P_0^n(x) \subset T^m P_m^n(x) \subset R''$ ($R, R', R'' \in \mathcal{R}$).

Definition 3.10. A quadrilateral Q u -crosses a rectangle $R \in \mathcal{R}$ if $Q \subset R$ and its boundary is the union of two subsegments of the stable boundary of R (the stable boundary of Q) with two line segments joining their endpoints (the unstable boundary). s -crossing is defined similarly.

A u -rectangle is a quadrilateral whose unstable boundary is made of two segments of unstable manifolds and which u -crosses some rectangle $R \in \mathcal{R}$. A s -rectangle is defined similarly (see Figure 2).

For $n \geq 1$, a **hyperbolic n -strip** (or just n -strip) is a s -rectangle S such that $\text{int } T^k(S)$ is included in some element of \mathcal{P} for each $k = 0, \dots, n-1$ and $T^n(S)$ is a u -rectangle. A **hyperbolic strip** is a n -strip for some $n \geq 1$.

We write $P_a^b(x)$ for $\bigcap_{k=a}^b \overline{T^{-(k-a)} P(T^k x)}$ (we assume implicitly that x is nice as this fails only on an entropy-negligible set – recall Proposition 2.6). The following is immediate.

Facts 3.11. 1. A hyperbolic n -strip is necessarily of the form $P_0^n(x)$ for some $x \in R$.

2. Two hyperbolic strips are either nested or have disjoint interiors.

We now give some tools to build hyperbolic strips.

Lemma 3.12. For $0 < m < n$, if $P_0^m(x)$ and $P_m^n(x)$ are both hyperbolic strips, then so is $P_0^n(x)$.

This is easy to show using Fig. 3. Sufficiently long invariant manifolds allow the construction of hyperbolic strips from scratch:

Lemma 3.13. Let $x \in \kappa_{10}(\mathcal{R})$ and $n \geq 1$ such that $T^n x \in \kappa_{10}(\mathcal{R})$. Then $P_0^n(x)$ is a hyperbolic strip.

Observe that the weaker condition $x \in \kappa(\mathcal{R}) \cap T^{-n}\kappa(\mathcal{R})$ does not imply that $P_0^n(x)$ is a hyperbolic strip.

PROOF: See Fig. 4. Let R, R' be the elements of \mathcal{R} containing x and $T^n x$. Consider the quadrilateral L generated by $W^s(x)$ and $T^{-n}W^u(T^n x)$. By convexity, it is contained in $P_0^n(x)$. As $T^n x \in \kappa_{10}(x)$, $\partial^s R'$ intersects the segments bounding $T^n L$ at most at a tenth of their length (counting from their intersection a, b, c, d with $W^s(T^n x)$). Consider $(abcd)$ the quadrilateral defined by these intersections. It u -crosses R' . $T^{-n}(abcd)$ s -crosses R because the preimages of a, b, c, d are each so

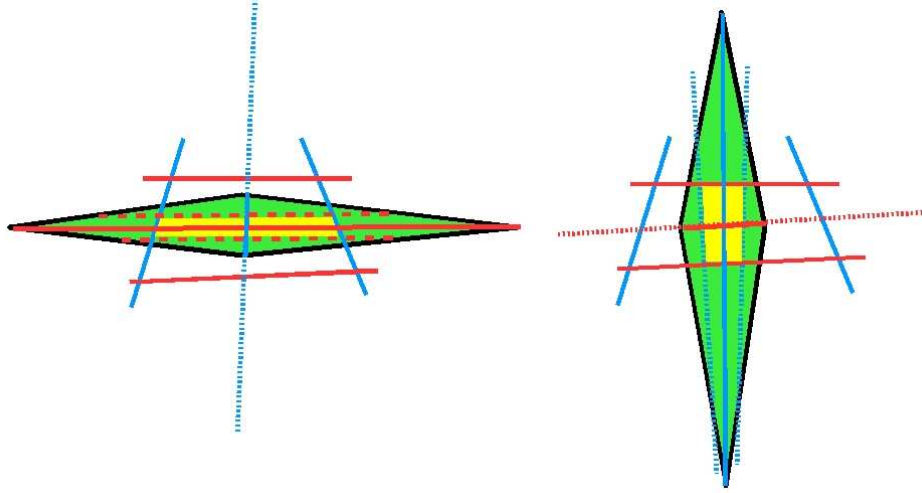


FIGURE 4. Construction of the hyperbolic strip in the proof of Lemma 3.13 (left: time 0 around R , right: time n around R' ; approximately (vertical) horizontal lines are segments of (un)stable manifolds; dashed ones are (pre)images of regular ones. The colored “diamond” is L . The rectangle inscribed in L is the hyperbolic strip.

close to the endpoints of $W^s(x)$ that they are outside of R . Therefore $R \cap T^{-n}Q$ is the desired hyperbolic strip and it must be $P_0^n(x)$. \square

Corollary 3.14. *Let $n \geq 1$ be such that $P_0^n(x)$ is a hyperbolic strip and $T^n x \in \kappa_{10}(R)$. If $m > n$ satisfies $T^m x \in \kappa_{10}(R)$, then $P_0^m(x)$ is also a hyperbolic strip.*

PROOF: By Lemma 3.13, $P_n^m(x)$ is a hyperbolic strip. Apply Lemma 3.12 to conclude. \square

We need the following technical fact.

Lemma 3.15. *Let μ be an atomless invariant probability measure. For a.e. $x \in \kappa_s(\mathcal{R})$ and all $n \geq 1$, the intersection of $W^s(x)$ with $\partial^u R$ lies in the interior of segments of $\partial P^n(x)$. In particular, $\partial^u R \cap \partial P^n(x)$ is the union of two non-trivial segments.*

PROOF: We proceed by contradiction assuming that the above fails: on a subset of $\kappa_s(R_i)$ with positive measure at least one of these intersection points coincides with a vertex z of the polygon $\partial P^n(x)$ (so $W^s(z) = W^s(x)$). Reducing this subset, we assume the vertex z to be a fixed one, say z_+ .

By Poincaré recurrence, there must exist infinitely many $n \geq 0$ such that $T^n x \in W^s(z_+)$. Considering two such integers $n_1 < n_2$, we get that $T^{n_2-n_1}(W^s(z_+)) \subset W^s(z_+)$. This implies that all points of $W^s(z_+)$ converge to a periodic orbit. Thus, the ergodic decomposition of μ has an atom, a contradiction. \square

We show that if $x \in \kappa_s(R)$, subsequent visits to $\kappa_{10}(R)$ either give a hyperbolic strip or a shadowing property which will lead to an entropy bound.

Lemma 3.16. *Let $x \in \kappa_s(R)$ and $0 \leq m < n$ be such that $T^m x, T^n x \in \kappa_{10}(R)$. Excluding a set of zero measure of points x , if $P_0^n(x)$ is not a hyperbolic strip then $P_0^m(x)$ determines $P_0^n(x)$ up to a choice of multiplicity 4.*

PROOF: $W^s(x)$ crosses R . Hence Lemma 3.15 implies that $\partial P_0^m(x) \cap \partial^u R_i$ is the union of two unstable, non-trivial segments: $[a, b], [c, d]$. Let $[a', b'], [c', d']$ be their images by $T^m|_{P_0^m(x)}$. Let Q' be the quadrilateral generated by them. By convexity $Q' \subset T^m(P_0^m(x))$.

$H := P_m^n(x)$ is a hyperbolic strip by Lemma 3.13. Q' and H intersect. If $\text{int } H \cap \{a', b', c', d'\} = \emptyset$, then Q' would go across H , and $P_0^m(x)$ would be a hyperbolic strip, contrary to assumption.

Thus, at least one of the four vertices a, b, c, d determined by $P_0^m(x)$ is contained in $\text{int } H$, this point determines H and therefore $P_0^n(x)$ as claimed. \square

3.3. Admissible Strips and Good Returns. In this section, \mathcal{R} is some Markov array with $\mathcal{R} \subset P$. Hyperbolic strips defined above have no uniqueness property: a point $x \in \kappa_s(\mathcal{R})$ sits into an infinite sequence of nested hyperbolic strips. This motivates the following notion.

Definition 3.17. *For $n \geq 1$, the **admissible n -strips** are defined by induction on n . A 1-strip is always admissible. For $n > 1$, an admissible n -strip S is a n -strip such that for all $1 \leq m < n$ such that S is included in an admissible m -strip, $T^m(S)$ meets no hyperbolic strip. An **admissible strip** is an admissible n -strip for some $n \geq 1$. For $x \in M$, an **admissible time** is such an integer n .*

Definition 3.18. *For a point $x \in M$, the **(good) return time** is $\tau = \tau(x)$, the minimum integer $\tau \geq 1$ such that both following conditions hold:*

- x belongs to an admissible τ -strip;
- $T^\tau(x) \in \kappa_s(\mathcal{R})$ belongs to an admissible strip.

If there is no such τ , then we set $\tau(x) = \infty$.

Remark 3.19. *Note that, at this point, we break the symmetry between the future and the past.*

We shall use repeatedly the following obvious observation:

Fact 3.20. *If n is the smallest integer such that $P_0^n(x)$ is a hyperbolic strip (equivalently: $P_0^n(x)$ is a n -strip which is not contained in a k -strip for any $k < n$; $P_0^n(x)$ does not meet a k -strip for any $k < n$; $P_0^n(x)$ is a hyperbolic strip which is maximum w.r.t. inclusion) then $P_0^n(x)$ is an admissible n -strip.*

Remark 3.21. *One could consider the following changes in the definition of admissibility:*

1) replacing “ $T^m(S)$ meets no hyperbolic strip” by “ $T^m(S)$ meets no admissible strip” would not change the notion. Indeed, suppose that $T^m(S)$ meets a hyperbolic strip H . Let $k \geq 1$ be the smallest integer such that H is contained in a k -strip, say H_k . The minimality of k implies that H_k is admissible and $H_k \supset H$ so that $T^m(S)$ meets H_k .

ii) replacing “ S is included in an admissible m -strip” by “ S is included in a hyperbolic m -strip” would exclude some admissible strips and so would cause a problem in the proof of the (key) Claim 4.2 (for the proof that $k = n_i$ in the notations there).

The following uniqueness property is the motivation for the definition of admissibility. Denote the one-sided symbolic dynamics by: $\Sigma_+(T, P) = \{A_0 A_1 A_2 \cdots \in P^{\mathbb{N}} : A \in \Sigma(T, P)\}$.

Lemma 3.22. *A positive itinerary $A \in \Sigma_+(T, P)$ can be decomposed in at most one way as an infinite concatenation of admissible strips.*

PROOF: Consider two distinct decompositions of A into admissible strips, that is, $n_0 = 0 < n_1 < n_2 < \dots$ and $m_0 = 0 < m_1 < m_2 < \dots$, such that $A_{n_i} \dots A_{n_{i+1}}$ and $A_{m_i} \dots A_{m_{i+1}}$ are admissible strips for all $i \geq 0$. By deleting the identical initial segments, we can assume that the decompositions differ from the beginning, say $n_1 < m_1$. It follows that the admissible m_1 -strip $H := \overline{A_0 \dots A_{m_1}}$ is contained into the n_1 -admissible strip $\overline{A_0 \dots A_{n_1}}$. Thus $T^{n_1}(H)$ meets $\overline{A_{n_1} \dots A_{n_2}}$ which is another admissible strip, contradicting admissibility. \square

4. ANALYSIS OF LARGE RETURN TIMES

In this section (M, T, P) is a piecewise affine homeomorphism with positive topological entropy. In the first two subsections, it is endowed with a Markov array $\mathcal{R} \subset P$. Eventually, (M, T, P, \mathcal{R}) will be some periodic extension of the original map as built in the previous section.

We prove that, if the average return time of some point x is very large then $P^{\tau(x)}(x)$ is already determined by some fraction of itself related to its visits to $\kappa_{10}(\mathcal{R})$ and $\mathcal{R} \setminus \kappa_{10}(\mathcal{R})$. We then deduce from this and the ergodic theorem a bound on the entropy of the measure.

4.1. Geometric Analysis. We analyze geometrically the implications of a large return time.

Proposition 4.1. *Let $x \in \kappa_s(\mathcal{R})$ and $N \geq 1$ be such that no $0 < k < N$ is a return time for x . Let:*

- $0 < N_0 < N$ be the smallest integer such that $T^{N_0}x \in \kappa_{10}(\mathcal{R})$ and $P_0^{N_0}(x)$ is a hyperbolic strip (we set $N_0 := N$ if there is no such integer);
- $0 \leq N_1 < N$ be the smallest integer such that $T^{N_1}x \in \kappa_{10}(\mathcal{R})$ (we set $N_1 := N_0 = N$ if there is no such integer);
- $0 \leq N_2 < N_0$ be the largest integer such that $T^{N_2}x \in \kappa_{10}(\mathcal{R})$ (we set $N_2 := N_1 = N_0$ if there is no such integer).
- n_1, \dots, n_r ($r \geq 0$) be the admissible times, that is the successive integers in:

$$\{0 \leq k < N : P_0^k(x) \text{ is an admissible strip}\}$$

with the convention $n_{r+1} = N$;

- $m_{i1}, \dots, m_{is(i)}$ ($s(i) \geq 0$) be the hyperbolic times, that is, for each i , the successive integers:

$$\{n_i < m < n_{i+1} : P_0^m(x) \text{ is a } m\text{-strip and } T^m x \in \kappa_s(\mathcal{R})\}.$$

with the convention $m_{is(i)} := m_{i1} := n_{i+1}$ if $s(i) = 0$.

Then $P^N(x)$ is determined, up to a choice of multiplicity $4 \cdot 2^r$, by:

- (1) the integers N_1, N_2, r and $n_i, m_{i1}, m_{is(i)}$ for $1 \leq i \leq r$;
- (2) $P(T^k x)$ for $k \in \llbracket 0, N_1 \rrbracket \cup \llbracket N_2, n_1 \rrbracket$;
- (3) $P(T^k x)$ for $k \in \bigcup_{i=1}^r \llbracket n_i, m_{i1} \rrbracket \cup \llbracket m_{is(i)}, n_{i+1} \rrbracket \setminus \llbracket 0, n_1 \rrbracket$.

($\llbracket a, b \rrbracket$ denotes the integer interval with a included and b excluded, etc.).

PROOF: We assume that $N_1 < N$, as otherwise there is nothing to show. In particular, $T^{N_1}x \in \kappa_{10}(\mathcal{R})$.

We claim that $P_0^{N_1}(x)$ determines $P_0^{N_2}(x)$ up to a choice of multiplicity 4. If $N_1 \geq N_2$, this assertion is empty so we can assume $N_1 < N_2 \leq N_0$. $T^{N_1}(x), T^{N_2}(x) \in \kappa_{10}(\mathcal{R})$ but $P_0^{N_2}(x)$ is not hyperbolic. Lemma 3.16 therefore implies that $P_0^{N_1}(x)$ determines $P_0^{N_2}(x)$ up to a choice of multiplicity 4. The Proposition now follows from the claim below. \square

Claim 4.2. *Except for an entropy-negligible subset of points $x \in M$ w.r.t. all large entropy measures the following holds. Given some $1 \leq i \leq r$, $Q := P_{n_i}^{m_{i1}}(x)$ and the integers n_i, m_{i1}, m_{is} , there are only two possibilities for $P_{n_i}^{m_{is}}(x)$ (s denotes $s(i)$).*

PROOF: Let $R, R' \in \mathcal{R}$ be the rectangles containing $T^{n_i}x, T^{m_{i1}}x$ and let ℓ be the line segment through $T^{n_i}x$, directed by $W^s(T^{n_i}x)$ and bounded by ∂R . We first show that $\ell \not\subset Q$.

By assumption $T^{n_i}x \notin \kappa_s(\mathcal{R})$, i.e., $W^s(T^{n_i}x)$ does not u -cross it: $\ell \not\subset W^s(T^{n_i}x)$. There exists a minimum $k > n_i$ such that $T^{k-n_i}\ell$ is not contained in the closure of an element of P . If one had $k > m_{i1}$, then $T^{m_{i1}-n_i}\ell \subset W^s(T^{m_{i1}}x)$ (recall that $W^s(T^{m_{i1}}x)$ crosses $R' \in P$ as $T^{m_{i1}}x \in \kappa(R')$) so that for all $k \geq m_{i1}$, $T^{k-n_i}\ell \subset T^{k-m_{i1}}W^s(T^{m_{i1}}x)$ would be inside elements of P , implying $\ell \subset W^s(T^{n_i}x)$, a contradiction. Thus, $k \leq m_{i1}$ and $\ell \not\subset Q$ as claimed.

We can assume that $T^{n_i}x$ is in the interior of Q , so $W^s(T^{n_i}x)$ cuts Q into two connected components Q_+, Q_- . If both contained stable manifolds crossing R , ℓ would be in the interior of Q by convexity. Thus, Q determines a gap in $\kappa_s(R)$ in the unstable direction. This gap must be bounded by two stable manifolds $W^s(B_+)$ ("above") and $W^s(B_-)$ ("below") going across R (recall Fact 3.8). Also at least one of $W^s(B_{\pm})$ (say $W^s(B_+)$) is not contained in Q so the interior of Q does not meet $W^s(B_+)$. Thus Q determines $W^s(B_+)$: it is the "lowermost" stable manifold "above" Q which crosses R . Q also determines $W^s(B^-)$: it is the "uppermost" stable manifold "below" $W^s(B^+)$ which crosses R .

By definition, $\mathcal{S} := P_0^{m_{is}}(x)$ is hyperbolic. Also $m_{is} \in \llbracket n_i, n_{i+1} \rrbracket$ is not admissible hence there exists an admissible time $0 < k < m_{is} < n_{i+1}$ such that \mathcal{S} is included in an admissible k -strip and $T^k(\mathcal{S})$ meets an admissible strip. As $P_0^{n_i}(x)$ is admissible, $k \geq n_i$. But n_{i+1} is the smallest admissible time after n_i . Therefore $k = n_i$.

Thus $T^{n_i}(\mathcal{S})$ meets an admissible strip which by definition is bounded "below" and "above" by pieces of stable manifolds crossing R , hence is either "above" $W^s(B_+)$ or "below" $W^s(B_-)$. This implies that $\text{int } P_0^{m_{is}}(x) \cap W^s(B_\epsilon) \neq \emptyset$ with $\epsilon = +$ or $-$ and Q determines $P_{n_i}^{m_{is}}(x)$, up to a binary choice. \square

4.2. Combinatorial Estimates. We extract from Proposition 4.1 the following complexity bound.

Proposition 4.3. *Let $\epsilon_* > 0$ and let $C_* = C_*(\epsilon_*) < \infty$ be such that*

$$\forall n \geq 0 \quad \#(P^{\mathcal{R}})^n \leq C_* e^{(h_{\text{top}}(T) + \epsilon_*)n}$$

Let $L, N \geq 1, M, R, S \geq 0$ be some integers. We consider the set $\mathcal{I} = \mathcal{I}(N, M, R, S)$ (we omit the obvious dependence on T, P, \mathcal{R}) of cylinders $P^N(x)$ for $x = (x, 0) \in \kappa_s(\mathcal{R})$ such that, in the notations of Proposition 4.1 applied to (M, T, P, \mathcal{R}) :

- $r = R$ and $\#\{m_{ij} > N_0 : 1 \leq i \leq r, 1 \leq j \leq s(i)\} = S$;
- $N_2 - N_1 = M$.
- $\min\{m_{i(j+1)} - m_{ij} : i = 1, \dots, r \text{ and } j = 0, \dots, s(i)-1\} \geq L$ (by convention $m_{i0} = n_i$ and $m_{is(i)+1} = n_{i+1}$).

Let $\rho > R/N$. Then:

$\log \#\mathcal{I} \leq (h_{\text{top}}(T) + \epsilon_*)(N - L(S - R) - M + S) + K_*(\rho, N)N + (\rho + 3/N)N \log C_*$
 where K_* is a universal² function satisfying $K_*(\rho, N) \downarrow K_*(\rho)$ when $N \rightarrow \infty$ and $K_*(\rho) \downarrow 0$ when $\rho \rightarrow 0$.

The proof of the above will use:

Lemma 4.4. *In the notation of Proposition 4.3,*

- (1) n_1 is the smallest integer such that $P_0^{n_1}x$ is hyperbolic and $n_1 \leq N_0$;
- (2) $\{n_i : 1 \leq i \leq r\} \subset \{0 \leq k < N : T^k x \in \mathcal{R} \setminus \kappa(\mathcal{R})\}$;
- (3) $\{N_0 \leq k < N : T^k x \in \kappa_{10}(\mathcal{R})\} \subset \{m_{ij} : 1 \leq i \leq r \text{ and } 1 \leq j \leq s(i)\}$.

PROOF: By definition n_1 is the smallest integer such that $P_0^{n_1}(x)$ is an admissible strip so (1) is just Fact 3.20.

$P_0^{n_i}(x)$ being an admissible strip, $T^{n_i}x \in \mathcal{R}$. $n_i < N$ so $T^{n_i}x \notin \kappa(\mathcal{R})$, proving (2).

The m_{ij} are the times $m \in \llbracket n_1, N \rrbracket$ (or, equivalently, $m \in \llbracket 0, N \rrbracket$ by property (1)) such that $T^m x \in \kappa(\mathcal{R})$ and $P_0^m(x)$ is a hyperbolic, but not admissible strip. As $P_0^{N_0}(x)$ is a hyperbolic strip and $T^{N_0}x \in \kappa_{10}(\mathcal{R})$, Corollary 3.14 gives that $N_0 \leq k < N$ and $T^k x \in \kappa_{10}(\mathcal{R})$ implies that $P_0^k(x)$ is a hyperbolic strip. This strip cannot be admissible as $T^k x \in \kappa_s(\mathcal{R})$ and $k < N$, hence such k is some m_{ij} , proving (3). \square

PROOF OF PROPOSITION 4.3: According to Proposition 4.1, given N, M, R and S , to determine an element of $\mathcal{I}(N, M, R, S)$ we need to specify:

- (1) the integers $N_1, N_2, n_1, \dots, n_R$ and $m_{i1}, m_{is(i)}$ for $i = 1, \dots, R$;
- (2) the itineraries $P^{N_1}(x), P_{N_2}^{n_1-1}(x)$ (if $n_1 > N_2$) and $P_{n_r+1}^{N-1}(x)$;
- (3) $P_{n_i}^{m_{i1}}(x), P_{m_{is(i)}}^{n_{i+1}}(x)$ for each $i = 1, \dots, r$;
- (4) a choice among $4 \cdot 2^R$.

Observe that $N_1 \leq N_2 \leq N_0$.

Using property (3) of Lemma 4.4 and the fact that $m_{ij} > n_1 \geq N_0$, it follows that:

$$\begin{aligned} \# \left(\bigcup_{i=1}^r \llbracket m_{i1}, m_{is(i)} \rrbracket \setminus \llbracket N_1, N_2 \rrbracket \right) &\geq \# \left(\bigcup_{i=1}^r \llbracket m_{i1}, m_{is(i)} \rrbracket \setminus \llbracket 0, N_0 \rrbracket \right) \\ &\geq \# \bigcup_{\substack{i=1, \dots, r \\ 1 \leq j < s(i) \\ m_{ij} > N_0}} \llbracket m_{ij}, m_{ij} + L \rrbracket \geq (S - R)(L - 1) \geq (S - R)L - S \end{aligned}$$

recalling the definitions of L and S . Hence, the number of choices for each of those items is bounded by:

- (1) $\binom{N}{2} \binom{N}{R}^3$ where $\binom{b}{a} = a!/b!(a-b)!$ is the binomial coefficient;

² K_* does not depend on any of the data $T : M \rightarrow M, N, M, R, S, L, C_*, \epsilon_*$.

- (2-3) $C_*^{R+2} \exp((h_{\text{top}}(T) + \epsilon_*)(N - (S - R)L - M + S))$ as we have to specify the itineraries for $R+2$ time intervals, all disjoint from $\llbracket N_1, N_2 \llbracket \bigcup_{i=1}^r m_{i1}, m_{is(i)} \llbracket$;
- (4) $4 \cdot 2^R$.

Recalling that³ $\binom{n}{\alpha n} \sim \frac{1}{\sqrt{2\pi\alpha(1-\alpha)}} n^{-1/2} e^{H(\alpha)n}$ as $n \rightarrow \infty$, i.e., $\log \binom{n}{\alpha n} \leq H(\alpha)n + C(\alpha)$ and that $k \in \llbracket 0, (n-1)/2 \llbracket \mapsto \binom{n}{k}$ is increasing, the stated bound follows with:

$$K_*(\rho, N) = 3H(\rho) + \rho \log 2 + 3N^{-1} \log N + N^{-1} \log 4C(\rho).$$

□

4.3. Large Average Return Times and Entropy. We are going to apply the previous estimates linking long return times either to visits to the holes ($\mathcal{R} \setminus \kappa(\mathcal{R})$) or to low entropy. We will show that for a suitable choice of the parameters of our constructions, large entropy measure have finite average return time.

Recall the good return time $\tau : \kappa_s(\mathcal{R}) \rightarrow \bar{\mathbb{N}}$ (possibly infinite) of Definition 3.18. We define $\tau_n(x)$ inductively by $\tau_1(x) = \tau(x)$ and $\tau_{n+1}(x) = \tau(T^{\tau_n(x)}(x))$ ($\tau_{n+1}(x) = \infty$ if $\tau_n(x) = \infty$).

We denote the essential supremum of a function f over a subset X with respect to a measure μ by:

$$\mu\text{-sup}_{x \in X} f(x) := \inf_{X' = X[\mu]} \sup_{x \in X'} f(x)$$

where X' ranges over the measurable subsets of X such that $\mu(X \setminus X') = 0$ (X and f are assumed to be measurable here). Our key estimate is the following:

Proposition 4.5. *There exist $h_2 < h_{\text{top}}(T)$ and $L_2 < \infty$ with the following property. Consider the Markov array \mathcal{R} defined by Proposition 3.6. For any integer $L_+ \geq L_0$ ($M_+, T_+, P_+, \mathcal{R}_+$) be the L_+ -periodic extension of Definition 3.5. Then, for each $\mu \in \mathbb{P}_{\text{erg}}^{h_2}(T_+)$,*

$$\tau_*(\mu) = \mu\text{-sup}_{x \in \mathcal{R}_+} \tau_*(x) < \infty \text{ where } \tau_*(x) := \limsup_{n \rightarrow \infty} \frac{1}{n} \tau_n(x).$$

Remark 4.6. *The simple consequence that $\tau_*(x) < \infty$ for a.e. $x \in \kappa(\mathcal{R})$ implies that almost every point in $\kappa(\mathcal{R})$ has a good return.*

One can make the following choice of L_2

$$(4.1) \quad L_2 > 2 \frac{\log 2}{h_{\text{top}}(T)} \mu_0^{-1}.$$

where μ_0 has been defined in the proof of Proposition 3.3.

The proof below does not provide a semi-uniform bound on τ_ as our estimates below depend on the speed of convergence of some ergodic averages (see Remark 4.7 below).*

PROOF OF PROPOSITION 4.5: The first step of the proof fixes a Markov array \mathcal{R} and a periodic extension of T and finds a candidate upper bound for $\tau_*(\mu)$. The second step defines a language (a collection of words of increasing lengths) with small entropy. The final step shows that large average return times imply that this language is enough to describe the measure. A large average can therefore happen only for low entropy measures.

Step 1: The Markov Array

³ $f(t) \sim g(t)$ iff $\lim_{t \rightarrow \infty} f(t)/g(t) = 1$ and $H(t) = -t \log t - (1-t) \log(1-t)$.

Recall that μ_0 is defined by Lemma 3.3 which also gives a Markov array \mathcal{R}_0 . We restrict ourselves to ergodic, T -invariant probability measures μ such that

$$\mu(\kappa_{10}(\mathcal{R}_0)) > \mu_0.$$

Let $\epsilon_0 > 0$ be small enough so that, K_* being the function defined in Proposition 4.3:

$$(4.2) \quad K_*(\epsilon_0\mu_0) < \frac{\mu_0}{100}h_{\text{top}}(T).$$

Let L_2 be so large that:

$$(4.3) \quad K_*(\epsilon_0\mu_0, L_2) < \frac{\mu_0}{100}h_{\text{top}}(T).$$

Using Proposition 3.6, one finds a refinement of the first Markov array \mathcal{R}_0 into the Markov array $\mathcal{R} = \mathcal{R}(T, \mu_0, \epsilon_0)$ and an integer $L_0 = L_0(T, \mu_0, \epsilon_0)$ (which we can assume to be larger than L_2) such that the following holds. For each (\mathcal{R}, L_+) -extension $(M_+, T_+, P_+, \mathcal{R}_+)$ with $L_+ \geq L_0$, there exists an ergodic lift μ_+ of μ such that

$$L_+\mu_+(\kappa_{10}(\mathcal{R}_+)) > \mu_0 \text{ and } L_+\mu_+(\mathcal{R}_+ \setminus \kappa(\mathcal{R}_+)) < \epsilon_0\mu_0.$$

Let $C_0 = C_0(T, \epsilon_0, \mu_0) < \infty$ be such that, $P^{\mathcal{R}}$ denoting the convex partition of M generated by P and \mathcal{R} :

$$\forall n \geq 1 \quad \#\{[A_1 \dots A_n] : A_i \in P^{\mathcal{R}}\} \leq C_0 \exp(h_{\text{top}}(T)(1 + \mu_0/100)n).$$

This is possible by Proposition 1.1 applied to T with partition $P^{\mathcal{R}}$. Possibly after increasing L_+ , we may assume that

$$\log C_0/L_+ < (\mu_0/100)h_{\text{top}}(T).$$

We omit the sharp subscript in the sequel so that M, T, P, μ, μ_0 will denote in fact $M_+, T_+, P_+, \mu_+, \mu_{0+}$. To refer to the original μ or μ_0 , we shall write μ_b or μ_{0b} . It will be a convenient exception to continue to write $P^{\mathcal{R}}$ for $P_b^{\mathcal{R}_b}$.

Let $K_1 \subset M$ and $L_1 < \infty$ be such that

$$(4.4) \quad \mu(M \setminus K_1) < \epsilon_0\mu_0^2/(10^6 \log \#P^{\mathcal{R}})$$

and, for all $x \in K_1$:

$$(4.5) \quad \forall n \geq L_0 := \frac{\epsilon_0\mu_0}{1000}L_1 \quad \left| \frac{1}{n} \#\{0 \leq k < n : T^k x \in \kappa_{10}(\mathcal{R})\} - \mu(\kappa_{10}(\mathcal{R})) \right| < \left(\frac{\mu_0}{1000} \right)^2.$$

$$\forall n \geq L_1 \quad \frac{1}{n} \#\{0 \leq k < n : T^k x \in \mathcal{R} \setminus \kappa(\mathcal{R})\} < \epsilon_0\mu_0.$$

Remark 4.7. *The above L_1 is the only estimate in the proof of this proposition which does not seem semi-uniform.*

We take:

$$(4.6) \quad \tau_{\max} := \frac{10^6 \log \#P^{\mathcal{R}}}{\mu_0^2} L_1.$$

To prove that $\tau_*(\mu) \leq \tau_{\max}$, we assume by contradiction that

$$(4.7) \quad M_1 := \{x \in \kappa(\mathcal{R}) : \limsup_{n \rightarrow \infty} \frac{1}{n} \tau_n(x) > \tau_{\max}\} \text{ has positive } \mu\text{-measure.}$$

Step 2: *Low entropy language*

For each integer $\ell \geq 1$ we define a set $C(\ell)$ of $P^{\mathcal{R}}$ -words of length ℓ as

$$C(\ell) := \bigcup_{\substack{\ell_1 + \dots + \ell_k = \ell \\ k \leq \ell/\tau_M}} C(\ell_1, \dots, \ell_k)$$

Here $C(\ell_1, \dots, \ell_k)$ is the set of all concatenations $\gamma_1 \dots \gamma_k$ where each γ_m ($1 \leq m \leq k$) is a word (i.e., a finite sequence) of length $|\gamma_m| = \ell_m$ satisfying:

- *type 1 requirement:* γ_m is an itinerary from $\mathcal{I}(\ell_m, M, R, S)$ (in the notation of Proposition 4.3) with $\ell_m \geq \frac{1000}{\mu_0} L_1$ (see below) and M, R, S integers satisfying:

$$(4.8) \quad L_+(S - R) + M - S \geq \frac{98}{100} \mu_0 \ell_m \text{ and } R \leq \frac{1}{100} \mu_0 \ell_m;$$

- *type 2 requirement:* the sum of the lengths of these segments is less than $\frac{\mu_0}{1000 \log \#P} \ell$.

Observe that

$$\#C(\ell) \leq \ell \binom{\ell}{\lfloor \ell/\tau_M \rfloor} \max_{\substack{\ell_1 + \dots + \ell_k = \ell \\ k \leq \ell/\tau_M}} \#C(\ell_1, \dots, \ell_k)$$

By Proposition 4.3, the logarithm of $\#\mathcal{I}(\ell_m, M, R, S)$, is bounded by

$$(4.9) \quad \begin{aligned} h_{\text{top}}(T) \left(1 + \frac{L_+ \mu_0}{100}\right) (\ell_k - (L_+(S - R) + M - S)) + K(\epsilon_0 \mu_0, \ell_k) \ell_k + (2R + 3) \log C_0 \\ \leq h_{\text{top}}(T) \left(1 + \frac{L_+ \mu_0}{100}\right) \left(1 - \frac{98}{100} L_+ \mu_0\right) \ell_k + \frac{2\mu_0}{100} h_{\text{top}}(T) \ell_k + 3 \log C_0 \\ \leq h_{\text{top}}(T) \left(1 - \frac{L_+ \mu_0}{100} (98 - 1.1 - 0.1)\right) + 3 \log C_0 \\ \leq h_{\text{top}}(T) (1 - 0.95 L_+ \mu_0) + 3 \log C_0 \end{aligned}$$

using the assumptions on $C(\ell)$, eq. (4.3) and that $S \ll h_{\text{top}}(T) \mu_0$ for L_+ large enough. Hence,

$$(4.10) \quad \#C(\ell) \leq \text{const} e^{h_2 \ell}$$

with $h_2 := h_{\text{top}}(T) (1 - 0.90 L_+ \mu_0)$.

Step 3: Consequence of Large Return Times

We are going to show that, for all $x \in \kappa(\mathcal{R})$,

$$(4.11) \quad \tau_n(x) > \tau_{\max} \cdot n \implies P^{\tau_n(x)}(x) \in C(\tau_n(x)).$$

Observe that this will imply that

$$h(T, \mu) \leq h_2$$

using Proposition A.2 with (4.10) and:

- $M_0 := \{x \in M : \{n \geq 0 : T^{-n}x \in M_1\} \text{ is infinite}\}$ (recall eq. (4.7));
- $a_i(x) := \min\{j \geq i : T^{-j}x \in M_1\}$ for all $i \geq 1$;
- $b_i(x) := \tau_n(T^{-a_i(x)}x)$ with n a positive integer such that $\tau_{n_i}(x) \geq \max(a_i(x), \tau_{\max} \cdot n_i)$,

concluding the proof of Proposition 4.5.

Let $x \in \kappa(\mathcal{R})$. For all large enough integers $n \geq 1$ (4.4) and the ergodic theorem give:

$$\frac{1}{\tau_n(x)} \#\{0 \leq k < \tau_n(x) : T^k x \notin K_1\} < \epsilon_0 \mu_0^2 / (10^6 \cdot \log \#P^{\mathcal{R}}).$$

Let $N := \tau_n(x)$, and for $k = 0, \dots, n-1$, let I_k be the integer interval $[[\tau_k(x), \tau_{k+1}(x)[[$ and $\ell_k := \#I_k$.

Let $B_1 \subset [[0, n[[$ be the set of those integers $0 \leq k < n$ such that

$$\#\{m \in I_k : T^m x \notin K_1\} \geq \frac{\epsilon_0 \mu_0}{1000} \ell_k.$$

The union of those segments I_k occupies only a small proportion of $[[0, N[[$:

$$\sum_{k \in B_1} \ell_k \leq \frac{1000}{\epsilon_0 \mu_0} \times \frac{\epsilon_0 \mu_0^2}{10^6 \log \#P^{\mathcal{R}}} N \leq \frac{\mu_0}{1000 \log \#P^{\mathcal{R}}} \cdot N.$$

Let $B_2 \subset [[0, n[[$ be the set of k 's such that $\ell_k \leq 1000 \mu_0^{-1} L_1$. They also occupy a small proportion:

$$\sum_{k \in B_2} \ell_k \leq 1000 \mu_0^{-1} L_1 n \leq \frac{1000 L_1}{\mu_0} \frac{N}{\tau_{\max}} \leq \frac{\mu_0}{1000 \log \#P^{\mathcal{R}}} \cdot N,$$

by the choice of τ_{\max} .

Therefore the segments I_k for $k \in B_1 \cup B_2$ satisfy the type 2 requirement in the definition of $C(\ell)$. It is enough to prove that the I_k 's for $k \in B_+ := [[0, n[[\setminus (B_1 \cup B_2)$ satisfy the type 1 requirement.

For such segments I_k , $p_1 := \min\{p \geq 0 : T^{p+\tau_k(x)} x \in K_1\}$ satisfies: $p_1 \leq \frac{\epsilon_0 \mu_0}{1000} \ell_k$ by the definition of B_1 . By the definition of B_2 , $\ell_k \gg L_1 = \frac{1000}{\epsilon_0 \mu_0} L_0$, hence (cf. N_1 from Proposition 4.3):

$$(4.12) \quad N_1 := \min\{j \geq 0 : T^{\tau_k(x)+j} x \in \kappa_{10}(\mathcal{R})\} \leq p_1 + L_0 \leq \frac{\epsilon_0 \mu_0}{500} \ell_k.$$

Also, by the choice of K_1 and using $\ell_k - p_1 > L_1$:

$$\#\{j \in [[\tau_k(x) + p_1, \tau_{k+1}(x)[[: T^j x \in R \setminus \kappa(\mathcal{R})\} < \epsilon_0 \mu_0 (\ell_k - p_1).$$

Hence, using point (2) of Lemma 4.4:

$$(4.13) \quad R \leq r' := \#\{j \in I_k : T^j x \in R \setminus \kappa(\mathcal{R})\} < 2\epsilon_0 \mu_0 \ell_k.$$

Note that this implies $R \leq \mu_0 \ell_k / 100$, one half of the type 1 requirement.

Similarly to (4.13):

$$|\#\{j \in I_k : T^j x \in \kappa_{10}(\mathcal{R})\} - \mu(\kappa_{10}(\mathcal{R})) \ell_k| < \frac{\mu_0}{500} \ell_k.$$

If, as in Proposition 4.3:

$$N_0 := \min\{j \geq 0 : T^{\tau_k(x)+j}(x) \in \kappa_{10}(\mathcal{R}) \text{ and } P^j(T^{\tau_k(x)} x) \text{ is hyperbolic}\}$$

(observe that N_0 might be large) and $S := \#\{m_{ij} > N_0 : i, j\}$ then, using point (3) of Lemma 4.4,

$$S \geq s' := \#\{j \in [[\tau_k(x) + N_0, \tau_{k+1}(x)[[: T^j x \in \kappa_{10}(\mathcal{R})\}.$$

Also, by the definition of M in Proposition 4.3, $M = [N_0]_{\kappa} - N_1$ where:

$$[t]_{\kappa} := \max\{n \in [[0, t[[: T^n x \in \kappa_{10}(R)\}$$

(we define $[t]_\kappa := t$ if there is no such integer). Hence,

$$(4.14) \quad L_+S + M = L_+S + [N_0]_\kappa - N_1 \geq L_+s' + [N_0]_\kappa - \frac{\mu_0}{1000}\ell_k$$

where, as

— **First case:** $N_0 < p_1 + L_1$. We use the trivial bound $M \geq 0$ and (4.5) to get:

$$s' \geq \#\{j \in \llbracket \tau_k(x) + p_1 + L_1, \tau_{k+1}(x) \rrbracket : T^j x \in \kappa_{10}(R)\} \geq \frac{499}{500}\mu_0(\ell_k - p_1 - L_1) \geq \frac{498}{500}\mu_0\ell_k$$

Hence,

$$(4.15) \quad L_+S + M \geq \frac{498}{500}\mu_{0b}\ell_k \geq \frac{99}{100}\mu_{0b}\ell_k.$$

— **Second case:** $N_0 \geq p_1 + L_1$. By the definition of $p_1 \leq (\epsilon_0\mu_0/1000)\ell_k$ and L_1 :

$$\begin{aligned} s' &\geq \mu(\kappa_{10}(R)) \left((1 - 10^{-3})(\ell_k - p_1) - (1 + 10^{-3})(N_0 - p_1) \right) \\ &\geq \frac{999}{1000}\mu(\kappa_{10}(\mathcal{R}))\ell_k - (1 + 10^{-3})\mu(\kappa_{10}\mathcal{R})N_0 \end{aligned}$$

Hence, using $2\mu_b(\kappa_{10}(\mathcal{R})) < 1$:

$$\begin{aligned} L_+S + M &\geq \frac{999}{1000}\mu_b(\kappa_{10}(\mathcal{R}))\ell_k - (1 + 10^{-3})\mu_b(\kappa_{10}\mathcal{R})N_0 + [N_0]_\kappa - N_1 \\ &\geq \frac{999}{1000}\mu_b(\kappa_{10}(\mathcal{R}))\ell_k - (N_0 - [N_0]_\kappa) - N_1. \end{aligned}$$

In light of (4.12), to prove that eq. (4.15) also holds in this second case it is enough to show:

Claim 4.8. For any $0 \leq t \leq \ell_k$, $t - [t]_\kappa \leq \frac{\mu_0}{250}\ell_k$.

PROOF OF CLAIM 4.8: We distinguish two cases. First assume that $[t]_\kappa < p_1 + L_0$. Then t , the first visit to $\kappa_{10}(R)$ after $[t]_\kappa$, is bounded by the first visit after $p_1 + L_0$, i.e.,

$$t - [t]_\kappa \leq t \leq p_1 + L_0 \leq \frac{\mu_0}{500}\ell_k$$

proving the claim in this case. Second we assume that $[t]_\kappa > p_1 + L_0$. Then:

$$\begin{aligned} (4.16) \quad &(1 - \mu_0/1000)\mu(\kappa_{10}(\mathcal{R}))(t - p_1) \\ &\leq \#\{j \in \llbracket \tau_k(x) + p_1, \tau_k(x) + t \rrbracket : T^j x \in \kappa_{10}(\mathcal{R})\} \\ &= \#\{j \in \llbracket \tau_k(x) + p_1, \tau_k(x) + [t]_\kappa \rrbracket : T^j x \in \kappa_{10}(\mathcal{R})\} \\ &\leq (1 + \mu_0/1000)\mu(\kappa_{10}(\mathcal{R}))([t]_\kappa + 1 - p_1) \end{aligned}$$

So that $t - p_1 \leq (1 + 3\mu_0/1000)([t]_\kappa + 1 - p_1)$. Hence,

$$t - [t]_\kappa \leq \frac{3}{1000}\mu_0[t]_\kappa + 2 - \frac{3\mu_0}{1000}p_1 \leq \frac{4}{1000}\ell_k,$$

proving the claim. \square

Eq. (4.15) together with (4.13) implies:

$$(4.17) \quad L_+(S - R) + M - S = L_+S + M - L_+R - S \geq \frac{98}{100}\mu_0L_+\ell_k$$

This gives the other half of the type 1 requirement on $P_{\tau_k(x)}^{\tau_{k+1}(x)}(x)$ for all $k \in \llbracket 0, n \rrbracket \setminus (B_1 \cup B_k)$. Hence $P^{\tau_n(x)}(x)$ belongs to $C(\tau_n(x))$. This concludes the proof of eq. (4.11) and of Proposition 4.5. \square

It also follows from the above proof that:

Corollary 4.9. *τ has eventually bounded gaps in the sense of Appendix B w.r.t. any large entropy measure.*

PROOF: Let μ be a large entropy measure. We fix τ_{\max} as in (4.6) and we proceed by contradiction assuming that for any large enough $t < \infty$ (larger than τ_{\max}), there is a set of positive μ -measure such that, for all of whose points x , there exist sequences of integers $n_k \in \mathbb{Z}$ and $m_k \in \mathbb{N}^*$ such that:

$$\tau_{m_k}(T_k^n x) > t \cdot m_k \text{ and } \sup_k d(0, \llbracket n_k, n_k + \tau_{m_k}(T_k^n x) \rrbracket) < \infty.$$

(the case with improper orbits is easy and left to the reader). It is now enough to apply eq. (4.10-4.11) and Proposition A.2 as in the above proposition to get the contradiction. \square

5. PROOF OF THE MAIN RESULTS

We finally prove the main theorem by building a Markov system from the arbitrary concatenations of admissible strips and relating it to the dynamics of the piecewise affine homeomorphism.

5.1. Maximum Entropy Measures. We prove Theorem 1 about the finite number of maximum measures.

Step 1: Tower

Fix a Markov array \mathcal{R} as in Proposition 3.6, defining a (\mathcal{R}, L_+) -periodic extension $(M_+, T_+, P_+, \mathcal{R}_+)$ of (M, T, P, \mathcal{R}) .

Observe that it is enough to prove the results for the symbolic dynamics $\Sigma := \Sigma(T_+, P_+)$ of that periodic extension, using that the periodic extension is a finite extension and Corollary 2.10. We now build an invertible tower $\hat{\Sigma}$ over Σ by defining a return time $\tau : \Sigma_{\kappa} \rightarrow \mathbb{N}^*$ for some $\Sigma_{\kappa} \subset \Sigma$ as follows.

Definition 5.1. *An extended admissible P_+ -word is a word $w_0 \dots w_n$ over P_+ such that $[w_0 \dots w_n]$ is an admissible strip. $w_0 \dots w_{n-1}$ is the associated admissible P_+ -word.*

For $A \in \Sigma$, define

$$t_1(A) = \sup\{n \geq 1 : A_0 \dots A_n \text{ is an extended admissible word}\} \in \mathbb{N}^* \cup \{-\infty, \infty\}$$

and $t_{n+1}(A) = t_1(\sigma^{t_n(A)}(A))$ (or ∞ if $t_n(A)$ is not an integer). Let

$$\Sigma_{\tau} := \{A \in \Sigma : \forall n \geq 1 \ t_n(A) \in \mathbb{N}^*\}.$$

Claim 5.2. *Σ_{τ} coincides with the set of P_+ -itineraries of the points $x \in \kappa_s(\mathcal{R}_+)$ up to a set negligible w.r.t. any large entropy measure.*

PROOF OF CLAIM: We tacitly exclude entropy-negligible subsets of points $x \in M_+$ (or Σ using Lemma 2.10). We prove first that for any $x \in \kappa_s(\mathcal{R}_+)$, the itinerary A of x belongs to $\Sigma_\tau(\mathcal{R}_+)$.

x has a finite good return time $m := \tau(x)$ by Proposition 4.5. In particular, $A_0 \dots A_m$ is admissible and $T_+^{\tau(x)}x \in \kappa_s(\mathcal{R}_+)$. By induction, A splits into a concatenation of admissible words $S^0 S^1 S^2 \dots$ (with identification of the first and last symbols of successive S^k s). The concatenations $S^0 S^1 \dots S^k$ are hyperbolic strips for all $k \geq 0$.

Now, if $H := P_{+0}^n(x)$ is hyperbolic with $n > m$, then H is contained in S^0 (an m -admissible strip) and $T^m H$ meets the hyperbolic strip S^1 . Hence H cannot be admissible, proving that $\tau(x) = t_1(A) < \infty$. $t_n(A) < \infty$ for all $n \geq 1$ follows from invariance. Hence, $A \in \Sigma_\tau$, proving the first inclusion.

Conversely, let $A \in \Sigma_\tau$ and denote by x the point with itinerary A . $[A_0 \dots A_m]$ is an admissible strip for $m = t_1(A)$ so $x \in A_0$ with $A_0 \in \mathcal{R}$. If $k = t_n(A)$, then $[A_0 \dots A_k]$ is a concatenation of admissible strips, hence a k -strip. k being arbitrarily large, it follows that $W^s(x)$, which contains the intersection of the previous strips, must cross A_0 , proving $x \in \kappa_s(\mathcal{R})$. \square

Observe that for all $A \in \Sigma_\kappa$ such that $t_1(A) < \infty$, $\sigma^{t_1(A)}(A) \in \Sigma_\kappa$. Hence t_1 is a return time and defines an invertible tower $\hat{\Sigma}$ in the sense of Appendix B. Moreover, Corollary 4.9 shows that any large entropy measure μ on $\Sigma(T_+, P_+)$ has eventually bounded gaps in the sense of Definition B.2. By Proposition B.3, any such measure can be lifted to $\hat{\Sigma}$ and the lift is a finite extension (in particular it has the same entropy as μ).

It follows also that $h(\hat{\Sigma}) := \sup_\mu h(\sigma|hS, \mu) = h_{\text{top}}(\Sigma)$ so that maximum measures of Σ lift to maximum measures of $\hat{\Sigma}$. Indeed, the above shows that $h(\hat{\Sigma}) \geq h_{\text{top}}(\Sigma)$. But point (3) of Proposition B.3 implies the converse inequality.

To prove the theorem it is therefore enough to show that $\hat{\Sigma}$ has finitely many ergodic measures of maximum entropy.

Step 2: Markov Structure

The Main Theorem will follow from $\hat{\Sigma}$ being a *Markov shift* with a finiteness property. Recall the following definition (see [12] for background):

Definition 5.3. A **Markov shift** is the left shift σ acting on:

$$\Sigma(\mathcal{G}) := \{x \in V^{\mathbb{Z}} : \forall n \in \mathbb{Z} x_n \rightarrow x_{n+1} \text{ in } \mathcal{G}\}$$

where \mathcal{G} is an oriented graph with a countable set of vertices V . Such a graph is **(strongly) irreducible** if for every $(A, B) \in V^2$, there is a path from A to B on \mathcal{G} . An **(strongly) irreducible component** is a subgraph with this property maximum w.r.t. inclusion.

To define the relevant Markov structure, we say that, given two admissible words w, w' , w' is a *follower* of w if, s being the first symbol of w' , the concatenation ws is an extended admissible word.

Claim 5.4. $\hat{\Sigma}$ is topologically conjugate to the Markov shift defined by the oriented graph \mathcal{G} with vertices (w, i) where w is any admissible word and $0 \leq i < |w|$ ($|\cdot|$ is the length of the word) and arrows:

$$(w, i) \rightarrow (w, i+1) \text{ if } i+1 < |w|, (w, |w|-1) \rightarrow (w', 0) \text{ if } w' \text{ is a follower of } w.$$

PROOF OF CLAIM: Define $p : \Sigma(\mathcal{G}) \rightarrow \hat{\Sigma}$ by $p(\alpha) = (A, \omega)$ such that for each $n \in \mathbb{Z}$, letting $(w_0 \dots w_{\ell-1}, k) := \alpha_n$, $A_n = w_k$ and $\omega_n = 1$ if and only if $k = 0$. Observe that the sequence A thus obtained is a concatenation of admissible words so $A \in \Sigma$. Also, whenever m and n are two successive integers with $\omega_m = \omega_n = 1$, $A_n A_{n+1} \dots A_m$ is an admissible word. Finally $\omega_m = 1$ for infinitely many positive and negative integers m . $\sigma^m A \in \Sigma_\kappa$ for all such m , so that $(A, \omega) \in \hat{\Sigma}$. Thus p is well-defined and is clearly a topological conjugacy. \square

Step 3: Conclusion

PROOF OF THEOREM 1: To conclude, observe that \mathcal{G} has finitely many irreducible components, corresponding to the finitely many vertices of the type $(w_0, 0)$: at most one for each rectangle in the Markov array \mathcal{R} . Hence, by a classical result of Gurevič [11], $\Sigma(\mathcal{G})$ with $h(\Sigma(\mathcal{G})) < \infty$ has finitely many maximum measures, proving the Main Theorem. \square

5.2. Number of Periodic Points. We prove Proposition 1.3 about the number of periodic points. We use the construction of the proof of Theorem 1. To prove the lower bound (1.3). It is enough to prove it for T_+ , as T_+ is an extension of T . We assume by contradiction that the number of points fixed under T_+^n is such that for any integer $p \geq 1$, there is a sequence $n_k \rightarrow \infty$ of multiples of p such that

$$(5.1) \quad \lim_{n_k \rightarrow \infty} \frac{N_{T_+}(n_k)}{e^{n_k h_{\text{top}}(T)}} = 0$$

We replace $\hat{\Sigma}$ by one of its irreducible component with maximum entropy. In the following we denote $(M_+, T_+, P_+, \mathcal{R}_+)$ by (M, T, P, \mathcal{R})

The starting point is the following estimate for $\hat{\Sigma}$. By Gurevič [11], the existence of a maximum measure for T and hence for $\hat{\Sigma}$ implies that \mathcal{G} is positive recurrent with parameter $e^{h_{\text{top}}(T)}$. By Vere-Jones [25], this implies that the number $N_{\mathcal{G}}(n)$ of loops of length n based at a given vertex satisfies, for some positive integer p :

$$(5.2) \quad \lim_{n \rightarrow \infty, p|n} N_{\mathcal{G}}(n) e^{-h_{\text{top}}(T)n} = p$$

(these computations go back to Vere-Jones [25]). We choose that p for the selection of the n_k s in eq. (5.1).

Each n -periodic sequence A in the symbolic dynamics Σ is associated to an n -periodic set

$$\bar{\pi}(A) := \bigcap_{n \geq 0} \overline{T_+^n[A_{-n} \dots A_n]}$$

$\bar{\pi}(A)$ contains at least one point fixed by T_+^n which we denote by $\pi(A)$. It remains to show that π does not identify too much points.

Consider π from the set $\Sigma(n)$ of n -periodic sequences to the set $M_+(n)$ of n -periodic T_+ -orbits. Our assumption (5.1) implies that, for some sequence $m_n \rightarrow \infty$ (for $n = n_k$) this map is at least m_n -to-1 on a subset $\Sigma'(n)$ of $\Sigma(n)$ with cardinality at least $e^{nh_{\text{top}}(T)}/3$. We use the following observation:

Lemma 5.5. *Let $A^1, \dots, A^m \in \Sigma$ be such that $\pi(A^1) = \dots = \pi(A^m) =: x$. If the finite words $A_0^i \dots A_{n-1}^i$, $i = 1, \dots, m$, are pairwise distinct, then:*

- (1) *either $T^k x$ is a vertex of P_+ for some $0 \leq k < n$;*

- (2) or there exist $r \geq (m-1)/2$ distinct integers $0 \leq n_1 < \dots < n_r < n$ such that for all $1 \leq k < \ell < n$, if v_k is the direction of the open edge containing $T^{n_k}x$, then the image by $T^{n_\ell - n_k}$ of v_k is transverse to v_ℓ .

PROOF: Let the finite words $A_0^i \dots A_{n-1}^i$, $i = 1, \dots, m$ be like in the above statement. We show that the failure of (1) implies (2).

Each word $A_0^i \dots A_{n-1}^i$ defines an element of P_+^n containing x in its closure so $\text{mult}(x, P_+^n) \geq m$. Observe that

- (1) $\text{mult}(x, P_+^{k+1}) = \text{mult}(x, P_+^k)$ if $T^k x$ is in the interior of an element of P_+ or if for all $A \in P_+^k$,

$$P_+^k T^k(A \cap B(x, \epsilon)) \subset B \text{ for some } B \in P_+ \text{ and } \epsilon > 0.$$

- (2) $\text{mult}(x, P_+^{k+1}) \leq \text{mult}(x, P_+^k) + 2$ if $T^k x$ is not a vertex of P_+ .

(2) uses that T_+ is a piecewise affine surface homeomorphism so the preimage of an edge may locally divide in at most two subsets at most two of elements of P_+^k touching x .

This implies that $\text{mult}(x, P_+^n) \leq 1 + 2\#\{0 \leq k < n : T^k x \text{ is on an edge of } P_+\}$. The Lemma follows. \square

PROOF OF PROPOSITION 1.3: Only $\text{const} \cdot n$ points of T_+ can satisfy assertion (1) in the above Lemma. As $h_{\text{mult}}(T_+, P_+) = 0$, their preimages in $\Sigma'(n)$ are in subexponential number. The remainder of $\Sigma'(n)$ corresponds to points x whose orbit stays off the vertices of P_+ and which admit distinct sequences $A^i \in \Sigma'(n)$, $i = 1, \dots, m_n$ with $\pi(A^i) = x$.

Given such an x , fix $0 < n_1 < \dots < n_r$ as in the Lemma. Pick j such that $0 \leq n_{j+1} - n_j \leq n/(m_n/2 - 1)$. Thus $T^{n_j} x$ is a vertex of $P_+^{n_{j+1} - n_j + 1}$. The number of such vertices, for given n_j , is bounded by $\text{const} \cdot \#P_+^{n/(m_n/2 - 1) + 1}$. Even taking into account the freedom in n_j , the number of such x s is bounded by: $e^{(2/m_n)(h_{\text{top}}(T) + \epsilon)n}$ for all large n . Thus, for these large $n = n_k$,

$$\#\Sigma(n) \leq 3\#\Sigma'(n) \leq 3e^{(2/m_n)(h_{\text{top}}(T) + \epsilon)n} + 3e^{\epsilon n}.$$

As $m_n \rightarrow \infty$, this contradicts the previous Vere-Jones estimate (5.2), proving the lower bound (1.3) and Proposition 1.3. \square

APPENDIX A. BOUNDS ON METRIC ENTROPY

It is a standard fact that the entropy of a measure can be computed as the average of the entropies given the past. More precisely, we have the following statement:

Lemma A.1. *Let μ be an invariant probability measure for some bimeasurable bijection $T : X \rightarrow X$. Let P be a finite, measurable partition. Then:*

$$(A.1) \quad h(T, \mu, P) = \int_X H_\mu(P|P^-)(x) \mu(dx)$$

where $H_\mu(P|P^-)(x) = -\sum_{A \in P} \mu_x(A) \log \mu_x(A)$ with $\mu_x(A) := E(\chi_A | P^-)(x)$ and P^- is the **past partition** generated by $T^n P$, $n \geq 1$.

In particular, if $N(n, x, P) = \#\{P_0^{n-1}(y) : y \in X \text{ and } P_{-\infty}^{-1}(y) = P_{-\infty}^{-1}(x)\}$ where $P_a^b(x) := (A_n)_{a \leq n \leq b, n \in \mathbb{Z}}$ with $T^n x \in A_n$, then:

$$(A.2) \quad h(T, \mu, P) \leq \int_X \frac{1}{n} \log N(n, x, P) \mu(dx)$$

PROOF: For eq. (A.1), see, e.g., [21, Ex. 4(b) p.243] for entropy as an average of conditional *information* and [21, Rem. 2.2 p. 236] for the reduction to the above, weaker statement about conditional *entropy*.

Eq. (A.2) now follows from the standard bound: for every $x \in X$, $H_\mu(A|B)(x) \leq \log \#\{a \in A : a \cap b \neq \emptyset\}$ where b is the element of B containing x . \square

In combination with Rudolph's backward Vitali Lemma [22, Theorem 3.9 p.33], it yields:

Proposition A.2. *Let μ be an ergodic, σ -invariant probability measure on $\mathcal{A}^{\mathbb{Z}}$ with finite alphabet \mathcal{A} . Assume that there exist a measurable family of subsets $W(A^-, \ell) \subset P^\ell$ (for $A^- \in \mathcal{A}^{\mathbb{Z}^-}$, $\ell \geq 1$) with cardinality bounded by $Ce^{H\ell}$ and a subset $\Sigma_0 \subset \mathcal{A}^{\mathbb{Z}}$ of positive measure such that, for all $A \in \Sigma_0$, there are sequences of integers $a_i = a_i(A)$, $b_i = b_i(A)$, $i \geq 1$, (depending measurably on A) satisfying:*

- (1) $\lim_{i \rightarrow \infty} b_i - a_i = \infty$;
- (2) $\sup_i d(0, \llbracket a_i, b_i \rrbracket) < \infty$;
- (3) $A_{a_i} A_{a_i+1} \dots A_{b_i-1} \in W(\dots A_{a_i-2} A_{a_i-1}, b_i - a_i)$

Then

$$h(\sigma, \mu) \leq H.$$

PROOF: To apply Rudolph's Backward Vitali Lemma, we need

$$(A.3) \quad a_i(A) \leq 0 \leq b_i(A)$$

for all large enough i , for all $A \in \Sigma_0$. By passing to subsequences, depending on A , we can assume the existence of the (possibly infinite) limits $\lim_{i \rightarrow \infty} a_i(A)$, $\lim_{i \rightarrow \infty} b_i(A)$ for all $A \in \Sigma_0$. Assume for instance that $\lim_i a_i(A) = -\infty$ and $\lim_{i \rightarrow \infty} b_i(A) < 0$ for a.e. $A \in \Sigma_0$, the other cases being similar or trivial. By assumption, $\inf_i b_i(A) > -\infty$ for all $A \in \Sigma_0$. Restricting Σ_0 we can assume that this infimum is some fixed number $b \in \mathbb{Z}$. Replacing Σ_0 by $\sigma^{\min(b,0)} \Sigma_0$ ensures eq. (A.3).

Rudolph Lemma implies that for any $\epsilon > 0$, for μ -a.e. A , for all large enough integers n , one can find a disjoint cover of a fraction at least $1 - \epsilon$ of $\llbracket 0, n \rrbracket$ by at most ϵn intervals $\llbracket a_i, b_i \rrbracket$ such that: $A_{a_i} \dots A_{b_i} \in W(\dots A_{a_i-2} A_{a_i-1}, b_i - a_i)$ Applying eq. (A.2) with:

$$N(A, n) \leq \binom{n}{2\epsilon n} e^{Hn} \times \#\mathcal{A}^{\epsilon n}.$$

gives that $h(\sigma, \mu) \leq H + 3\epsilon \log \epsilon + \epsilon \log \#\mathcal{A}$. We conclude by letting $\epsilon \rightarrow 0^+$. \square

APPENDIX B. TOWER LIFTS

We study towers from a point of view closely related to that of Zweimuller [27]. Let T be an ergodic invertible transformation of a probability space (X, μ) and let B be a measurable subset of X . A *return time* is a function $\tau : B \rightarrow \bar{\mathbb{N}}^* := \{1, 2, \dots, \infty\}$ which is measurable and such that $T^{\tau(x)}(x) \in B$ for all $x \in B$ with $\tau(x) < \infty$ (but τ is not necessarily the first return time).

We are interested in lifting T -invariant measures to the following **invertible tower**:

$$(B.1) \quad \hat{X} := \{(x, \omega) \in X \times \{0, 1\}^{\mathbb{Z}} : \omega_n = 1 \implies T^n x \in B \text{ and}$$

$$\tau(T^n x) = \min\{k \geq 1 : \omega_{n+k} = 1\}\} \setminus \hat{X}_*$$

with $\hat{T}(x, \omega) = (Tx, \sigma(x))$ and \hat{X}_* is the set of (x, ω) with only finitely many 1s either in the future or in the past of ω . (\hat{X}, \hat{T}) is an extension of a subset of (X, T) through $\hat{\pi} : \hat{X} \rightarrow X$ defined by $\hat{\pi}(x, \omega) = x$.

Remark B.1. The jump transformation $T^\tau : \{x \in B : \tau(x) < \infty\} \rightarrow B$ is defined by $T^\tau(x) := T^{\tau(x)}(x)$. It is closely related to \hat{T} . Indeed, T^τ is the first return map of \hat{T} on $[1] := \{(x, \omega) \in \hat{X} : \omega_0 = 1\}$ so any \hat{T} -invariant probability measure gives by restriction and normalization a T^τ -invariant measure (see [27]).

Such lifting requires that τ be "not too large" (see [27] where the classical integrability condition is studied). Our condition is in terms of the following 'iterates' of τ : the functions $\tau_m : B \rightarrow \bar{\mathbb{N}}^*$, $m \geq 1$, are defined, as before, by: $\tau_1 := \tau$ and $\tau_{m+1}(x) := \tau(T^{\tau_m(x)}(x))$ if $\tau_m(x) < \infty$, $\tau_{m+1}(x) := \infty$ otherwise.

Definition B.2. $x \in X$ has an **improper orbit** if

(B.2) $n(x) := \{n \in \mathbb{N} : T^{-n}x \in B \text{ and } \forall m \geq 1 \tau_m(T^{-n}x) < \infty\}$ is finite.

$x \in X$ has **t -gaps** for some $0 < t < \infty$ if x has an improper orbit or if there exist two integer sequences $n_k \in \mathbb{Z}$ and $m_k \in \mathbb{N}^*$ such that:

$$\tau_{m_k}(T^{n_k}x) \geq \max(t \cdot m_k, k) \text{ and } \max_{k \geq 1} d(0, [n_k, n_k + \tau_{m_k}(T^{n_k}x)]) < \infty.$$

A measure has **eventually bounded gaps**, if for some $t < \infty$, the set of points in X with t -gaps has zero measure.

Note that $\tau(T^n x) = \infty$, for a single n , implies that x has t -gaps for any $t < \infty$.

Proposition B.3. Let $T : X \rightarrow X$ be a self-map with a return time $\tau : B \rightarrow \bar{\mathbb{N}}^*$ as above. Then:

- every T -invariant ergodic probability measure μ with eventually bounded gaps can be lifted to a \hat{T} -invariant ergodic probability measure on \hat{X} ;
- any \hat{T} -invariant, ergodic probability measure $\hat{\mu}$ is a finite extension of the T -invariant measure $\hat{\pi}(\hat{\mu})$.

PROOF OF PROPOSITION B.3: We first prove the existence of a lift for μ like above. We follow the strategy of [27] and [17] (which was inspired by constructions of Hofbauer) and define the following *non-invertible tower* to get the right topology:

$$\tilde{X} := \{(x, k, \tau) \in X \times \mathbb{N} \times \mathbb{N} : \exists y \in B \tau(y) = \tau, k < \tau \text{ and } x = T^k y\}$$

$$\tilde{T}(x, k, \tau) := (T(x), k + 1, \tau) \text{ if } k + 1 < \tau, (T(x), 0, \tau(T(x))) \text{ otherwise.}$$

We write $\tilde{X}_K := \{(x, k, \tau) \in \tilde{X} : k = K\}$, $\tilde{X}_{>K} := \bigcup_{k > K} \tilde{X}_k$ and define $\tilde{\pi}(x, k, \tau) = x$. Observe that $\tilde{\pi} \circ \tilde{T} = T \circ \tilde{\pi}$ and that (\tilde{X}, \tilde{T}) is a natural extension of (\tilde{X}, \tilde{T}) through $(x, \omega) \mapsto (x, k, \ell)$ with $k \geq 0$ minimum such that $\omega_{-k} = 1$ and $\ell = \tau(T^{-k}x)$. Hence it is enough to lift μ to \tilde{X} .

Fix $t < \infty$ such that the set of points with t -gaps has zero μ -measure. Let $\tilde{\mu}_0$ be the probability measure defined by

$$\tilde{\mu}_0(\{(x, 0, \tau(x)) : x \in A\}) = \mu(A) \text{ for all Borel } A$$

(sets disjoint from the above ones have zero $\tilde{\mu}_0$ -measure). We have $\tilde{\pi}(\tilde{\mu}_0) = \mu$ but, except in trivial cases, $\tilde{T}\tilde{\mu}_0 \neq \tilde{\mu}_0$ so we consider:

$$\tilde{\mu}_n := \frac{1}{n} \sum_{k=0}^{n-1} \tilde{T}^k \tilde{\mu}_0$$

and wish to take some accumulation point $\tilde{\mu}$. We identify $\tilde{\mu}_n$ with its density with respect to $\tilde{\mu}_\infty$ the σ -finite measure defined by

$$\tilde{\mu}_\infty(\{(x, k, \tau(T^{-k}x)) : x \in A\}) = \mu(A \cap \{\tau > k\})$$

for all Borel $A \subset B$. As $\tilde{\pi}(\tilde{\mu}_n) = \mu$, we must have:

$$\frac{d\tilde{\mu}_n}{d\tilde{\mu}_\infty} \leq 1.$$

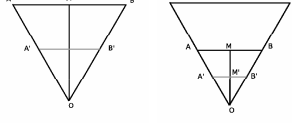


FIGURE 5. Geometry of (left) a continuous piecewise affine map with $h_{\text{mult}}(T, P) > h_{\text{top}}(T) = 0$; (right) a discontinuous piecewise affine map with no maximum measure.

Using the Banach-Alaoglu theorem, i.e., the weak star compactness of the unit ball of $L^\infty(\tilde{\mu}_\infty)$ as the dual of $L^1(\tilde{\mu}_\infty)$, we get a measure $\tilde{\mu}$ on \tilde{X} with $d\tilde{\mu}/d\tilde{\mu}_\infty \leq 1$ such that, for some subsequence $n_k \rightarrow \infty$,

$$(B.3) \quad \forall f \in L^1(\tilde{\mu}_\infty) \quad \lim_{k \rightarrow \infty} \int f d\tilde{\mu}_{n_k} = \int f d\tilde{\mu}.$$

Observe that $\tilde{\mu}$ is \tilde{T} -invariant: indeed, eq. (B.3) together with the T -invariance of μ implies that $d\tilde{\mu} \circ \tilde{T}^{-1} / d\tilde{\mu} \leq 1$ whereas $\tilde{\mu} \circ \tilde{T}^{-1}(\tilde{X}) = \tilde{\mu}(\tilde{X})$ so the previous inequality must be an equality $\tilde{\mu}$ -almost everywhere.

Using this invariance and the ergodicity of μ it is easily seen that $\tilde{\mu} = \alpha\mu$ for some $0 \leq \alpha \leq 1$. It remains to prove that $\tilde{\mu} \neq 0$ so that it can be renormalized into the announced lift of μ . Assume by contradiction that $\tilde{\mu} = 0$. Hence, for any $L < \infty$:

$$\int 1_{\tilde{X}_{\leq L}} d\tilde{\mu}_{n_k} = \int \frac{1}{n_k} \#\{0 \leq k < n_k : \tilde{T}^k(x, 0, \tau(x)) \in \tilde{X}_{\leq L}\} d\mu \rightarrow 0$$

So, possibly for a further subsequence:

$$(B.4) \quad \frac{1}{n_k} \#\{0 \leq k < n_k : \tilde{T}^k(x, 0, \tau(x)) \in \tilde{X}_{\leq L}\} \rightarrow 0 \text{ } \mu\text{-a.e.}$$

Now,

$$\#\{0 \leq k < t : \tilde{T}^k(x, 0, \tau(x)) \in \tilde{X}_0\} < n \implies t \leq \tau_n(x)$$

Hence eq. (B.4) implies that x (in fact any of its preimages in the natural extension) has t -gaps for all $t > 0$, contradicting the assumption on μ .

We show now that any \hat{T} -invariant, ergodic probability measure $\hat{\mu}$ is a finite extension of $\mu := \hat{\pi}(\hat{\mu})$. By definition of \hat{X} , $\hat{\mu}([1]) > 0$ where $[1] = \{(x, \omega) \in \hat{X} : \omega_0 = 1\}$. Let K be an integer such that for all x in a set of positive measure, one can find K distinct sequences $\omega^1, \dots, \omega^K \in \omega(x)$, each (x, ω^i) belonging to the full measure set of $\hat{\mu}$ -generic points. Hence,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \#\{0 \leq k < n : \omega_{-k}^j = 1\} = \hat{\mu}([1]).$$

If $K \cdot \hat{\mu}([1]) > 1$, there exist $j \neq j'$ and arbitrarily large integers $n_k \rightarrow \infty$ such that $\omega_{-n_k}^j = \omega_{-n_k}^{j'}$. But this implies $\omega^j = \omega^{j'}$ by Lemma 3.22. The contradiction proves $K \leq \hat{\mu}([1])^{-1} < \infty$: $\hat{\mu}$ is a finite extension of μ . \square

APPENDIX C. EXAMPLES

Positive Multiplicity Entropy

Example 1 (see [5]). *There exists a continuous, piecewise affine surface map (M, T, P) with $h_{\text{mult}}(T, P) > 0$ and $h_{\text{top}}(T) = 0$.*

Consider some triangle ABO in \mathbb{R}^2 with non-empty interior and let M, A', B' be the middle points of $[AB], [AO], [BO]$ (see Fig. 5, left). Let T be affine in each of the triangles $\tau_0 := AMO$ and $\tau_1 := BMO$ with $T(O) = O$, $T(A) = T(B) = A'$, $T(M) = B'$ so

that $T : ABO \rightarrow ABO$ is conjugate to $(\theta, r) \mapsto (1 - 2|\theta|, r/2)$ on $(-1, 1) \times (0, 1)$. Take $P = \{\tau_0, \tau_1\}$ as the admissible partition.

$h_{\text{mult}}(T, P) = \log 2$ (because all words on $\{\tau_0, \tau_1\}$ appear in the symbolic dynamics). On the other hand, the only invariant probability measure is the Dirac supported by O hence $h_{\text{top}}(T) = 0$ as claimed.

Example 2 (see Kruglikov and Rypdal [18]). *There exists a piecewise affine homeomorphism (M, T, P) with $\dim M = 3$ and $h_{\text{mult}}(T, P) = h_{\text{top}}(T) > 0$.*

Let $([0, 1]^2, T_2, P_2)$ be a piecewise affine homeomorphism with nonzero topological entropy. Consider the pyramid $M := \widehat{[0, 1]^2}$ where \widehat{A} denotes the convex set generated by $(0, 0, 0)$ and $A \times \{1\}$. Define $T : M \rightarrow M$ as the piecewise affine map with partition $P := \{\widehat{A} : A \in P_2\}$ such that: $T(0, 0, 0) = (0, 0, 0)$, $T(x, y, 1) = (T_2(x, y), 1)$ for each vertex (x, y) of P_2 . Observe that $h_{\text{top}}(T) = h_{\text{top}}(T_2)$ and that T has an obvious measure of maximum entropy carried by the invariant set $[0, 1]^2 \times \{1\}$. Finally, considering T^n around $(0, 0, 0)$ it is easy to see that $h_{\text{mult}}(T, P) = h_{\text{top}}(T_2) = h_{\text{top}}(T)$.

Example 3. *There exists a piecewise affine homeomorphism (M, S, P) with $\dim M = 3$, $h_{\text{mult}}(S, P) > 0$ and $h_{\text{top}}(S) = 0$.*

Define S from the previous example T by $S(x, y, z) := T(x, y, z)/2$ on the pyramid M so that 0 is a sink. To make S onto, add a symmetric pyramid M^- whose summit is a source.

No Maximal Measure

Example 4. *There exists a piecewise affine surface (M, T, P) discontinuous map T such that (i) the set of invariant probability measures is non-empty and compact in the usual weak star topology; (ii) there is no maximum measure.*

Remark. (i) excludes trivial examples like $T : [0, 1] \rightarrow [0, 1]$ with $T(x) = 1/4 + x/2$ for $x > 1/2$ and $T(x) = x + 1/2$ for $x \leq 1/2$ which has no invariant probability measure.

Let T be a piecewise affine map defined on the triangle $M := XYO$ with $O = (0, 0)$, $X = (-2, 2)$ and $Y = (2, 2)$. Let $A = (-1, 1)$, $B = (1, 1)$ and $M = (0, 1)$, and $A' = A/2$, $B' = B/2$ and $M' = M/2$ (see Fig. 5, right). We require:

- (1) $T|_{\overline{XYBA}}$ is the identity;
- (2) $T : AMO \rightarrow A'B'O$ is affine with $A \mapsto A'$, $M \mapsto B'$, $O \mapsto O$;
- (3) $T : MBO \rightarrow YXO$ is affine with $M \mapsto Y$, $B \mapsto X$, $O \mapsto O$;
- (4) $T(O) = O$.

It is easy to see that $\sup_{\mu} h(T, \mu) = \log 2$: for any $h < \log 2$, one can find an invariant measure on $\{0, 1\}^{\mathbb{N}}$ such that $\mu([1^K]) = 0$ for some $K < \infty$ with entropy at least h . It is then easy to construct an isomorphic T -invariant measure (with support included in $y < 1$), proving that $\sup_{\mu} h(T, \mu) \geq \log 2$. The equality follows from $h_{\text{top}}(T) = \log 2$.

On the other hand, assume that μ is an invariant and ergodic probability measure with $h(T, \mu) = 2$. μ must be supported on $y < 1$. Hence the map π that sends a point of \mathbb{R}^2 to the ray from the origin that contains it, maps (T, μ) to $(f, \pi_+ \mu)$ where $f : \theta \mapsto 1 - 2|\theta|$ on $[0, 1]$. The fibers of π are contained in intervals on which T has no distortion, hence they have zero entropy and π is entropy-preserving. This implies that $\pi_+ \mu$ is the $(1/2, 1/2)$ -Bernoulli measure. Using, say the Central Limit Theorem, we get that, for μ -a.e. $(x, y) \in ABO$, there exists a positive integer n such that

$$\#\{0 \leq k < n : f^k(\pi(x, y)) < 1/2\} \geq \frac{|\log y|}{2 \log 2} - \frac{n}{2}$$

so that $T^m(x) \in XYBA$ for some $m \leq n$. This shows that there is no maximum measure.

Let us show that the set of invariant probability measure $\mathbb{P}(T)$ is compact. Let μ_n be some sequence of T -invariant probability measures converging to some probability

measure μ_0 on the compact set M . μ_0 can fail to be T -invariant only if there is some $\epsilon_0 > 0$ such that, for any neighborhood U of $[AB] \cup [OM]$, $U' := U \setminus (\overline{XYBA} \cup \{O\})$ satisfies: $\mu_n(U') > \epsilon_0$ for all large n . But one can find U such that $T^k(U') \cap U' = \emptyset$ for all $0 < k < 2/\epsilon_0$, a contradiction.

Example 5. *There exists a continuous, piecewise quadratic surface map T such that for any invariant probability measure μ :*

$$h(T, \mu) < \sup_{\nu} h(T, \nu).$$

On the rectangle $[1, 2] \times [-1, 1]$, consider $T(x, y) := (x, T_x(y))$ with:

$$T_x(y) = \begin{cases} \frac{x(2-x)}{2} - x|y| & \text{if } |y| < 2-x \\ -\frac{x(2-x)}{2} & \text{otherwise} \end{cases}$$

For each $1 \leq x < 2$, $[-1, 1]$ is mapped into the T_x -forward invariant segment $[-\frac{x(2-x)}{2}, \frac{x(2-x)}{2}]$ on which T_x has constant slope x . Hence $h_{\text{top}}(T_x) = \log x$ for $x \neq 2$. Clearly, $T(2, y) = (2, 0)$ so $h_{\text{top}}(T_2) = 0$.

REFERENCES

- [1] Billingsley, Patrick; *Ergodic theory and information*. Reprint of the 1965 original. Robert E. Krieger Publishing Co., Huntington, N.Y., 1978.
- [2] Bowen, Rufus Entropy for group endomorphisms and homogeneous spaces. *Trans. Amer. Math. Soc.* 153 1971 401–414.
- [3] M. Boyle, J. Buzzi, R. Gomez, *Almost isomorphism of countable state Markov shifts*, *J. Reine Angew. Math.* 592 (2006), 23–47.
- [4] J. Buzzi, *Intrinsic ergodicity of smooth interval maps*, *Israel J. Math.* 100 (1997), 125–161.
- [5] J. Buzzi, *Intrinsic ergodicity of Affine Maps in $[0, 1]^d$* , *Monatsh. Math.* 124 (1997), no. 2, 97–118.
- [6] J. Buzzi, *Markov extensions for multi-dimensional dynamical systems*, *Israel J. Math.* 112 (1999), 357–380.
- [7] J. Buzzi, *Piecewise isometries have zero topological entropy*, *Ergodic Theory Dynam. Systems* 21 (2001), no. 5, 1371–1377.
- [8] J. Buzzi, *On entropy-expanding maps*, preprint CMLS 2000.
- [9] J. Buzzi, *Puzzles of quasi-finite type, zeta functions and symbolic dynamics for multi-dimensional maps*, [arxiv:math.DS/0610911](https://arxiv.org/abs/math/0610911)
- [10] A. Greven, G. Keller, G. Warnecke (eds), *Entropy*, Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2003.
- [11] B. Gurevic, *Shift entropy and Markov measures in the space of paths of a countable graph*, (Russian) *Dokl. Akad. Nauk SSSR* 192 1970 963–965; English transl. *Soviet Math. Dokl.* 11 (1970), 744–747
- [12] B. Gurevic, S. Savchenko, *Thermodynamic formalism for symbolic Markov chains with a countable number of states*, (Russian) *Uspekhi Mat. Nauk* 53 (1998), no. 2(320), 3–106; translation in *Russian Math. Surveys* 53 (1998), no. 2, 245–344
- [13] F. Hofbauer, *On intrinsic ergodicity of piecewise monotonic transformations with positive entropy*, *Israel J. Math.* 34 (1979), no. 3, 213–237 (1980); *Israel J. Math.* 38 (1981), no. 1-2, 107–115.
- [14] F. Hofbauer, *Periodic points for piecewise monotonic transformations*, *Ergodic Theory Dynam. Systems* 5 (1985), no. 2, 237–256.
- [15] A. Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, *Inst. Hautes Etudes Sci. Publ. Math.* No. 51 (1980), 137–173.
- [16] A. Katok, B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, With a supplementary chapter by Katok and Leonardo Mendoza. *Encyclopedia of Mathematics and its Applications*, 54. Cambridge University Press, Cambridge, 1995.
- [17] G. Keller, *Keller, Gerhard(D-ERL) Lifting measures to Markov extensions*, *Monatsh. Math.* 108 (1989), no. 2-3, 183–200.
- [18] B. Kruglikov, M. Rypdal, *Entropy via multiplicity*, *Discrete Contin. Dyn. Syst.* 16 (2006), no. 2, 395–410.

- [19] M. Misiurewicz, *Diffeomorphism without any measure with maximal entropy*, Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys. 21 (1973), 903–910
- [20] S. Newhouse, *Continuity properties of entropy*, Ann. of Math. (2) 129 (1989), no. 2, 215–235.
- [21] K. Petersen, *Ergodic theory*, Cambridge University Press, 1983.
- [22] D. Rudolph, *Fundamentals of measurable dynamics*, Clarendon Press, 1990.
- [23] S. Ruelle, *Mixing C^r maps of the interval without maximal measure*, Israel J. Math. 127 (2002), 253–277.
- [24] M. Tsujii, *Absolutely continuous invariant measures for expanding piecewise linear maps*, Invent. Math. 143 (2001), no. 2, 349–373.
- [25] D. Vere-Jones, *Ergodic properties of nonnegative matrices*, Pacific J. Math. 22 1967 361–386.
- [26] P. Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics, 79. Springer-Verlag, New York-Berlin, 1982.
- [27] R. Zweimuller, *Invariant measures for general(ized) induced transformations*, Proc. Amer. Math. Soc. 133 (2005), no. 8, 2283–2295.

JEROME BUZZI, CENTRE DE MATHÉMATIQUES LAURENT SCHWARTZ, ECOLE POLYTECHNIQUE,
91128 PALAISEAU CEDEX, FRANCE

E-mail address: buzzi@math.polytechnique.fr

URL: www.jeromebuzzi.com