A minicourse on entropy theory on the interval

Jerome Buzzi

A MINICOURSE ON ENTROPY THEORY ON THE INTERVAL

by

Jérôme Buzzi

Abstract. — We give a survey of the entropy theory of interval maps as it can be analyzed using ergodic theory, especially measures of maximum entropy and periodic points. The main tools are (i) a suitable version of Hofbauer's Markov diagram, (ii) the shadowing property and the implied entropy bound and weak rank one property, (iii) strongly positively recurrent countable state Markov shifts. Proofs are given only for selected results. This article is based on the lectures given at the *Ecole thématique de théorie ergodique* at the C.I.R.M., Marseilles, in April 2006.

Contents

1. Introduction

We are going to give a very selective survey of interval dynamics, mainly (but not exclusively) those defined by maps with finitely many critical points or discontinuities (see Definition 1.1 below). We focus on "complexity" as defined through entropy as seen from an *ergodic theory* point of view. Ergodic theory will be for us both a powerful tool and a guide to the "right" questions. In particular, we shall concentrate on aspects not dealt previously in book form (like the classic treatise [**67**] or [**2**] for another, non-ergodic, point of view on the same subject) around measures of maximum or large entropy.

As the lectures given in Luminy, these notes are intended to be accessible to readers with only a basic knowledge of dynamical systems. From a technical point of view we shall only assume (1) measure theory (e.g., the very first chapters of ([**86**]); (2) the Birkhoff ergodic theorem, explained in any textbook on ergodic theory (see, e.g., [**107**]).

We shall deal mainly with the following rather well-behaved class of one-dimensional dynamics:

Definition 1.1. — *I will always denote a compact interval of $\mathbb{R}$. A self-map of I is* **piecewise monotone** *if there exists a partition of I into finitely many subintervals (the "pieces") on each of which the restriction of f is continuous and strictly monotone. Note that one can always subdivide the pieces. The natural partition is the set of interiors (relatively to $\mathbb{R}$) of the pieces in such a partition with minimum cardinality.*

We denote the set of piecewise monotone maps by $\mathrm{PMM}(I)$ and its topology is defined by $d(f_1, f_2) < \epsilon$ if f_1 and f_2 both admit natural partitions with n pieces with endpoints a_j^i, b_j^i such that $|a_j^1 - a_j^2| < \epsilon$, $|b_j^1 - b_j^2| < \epsilon$ and $|f^1(a_j^1 + (b_j^1 - a_j^1)t) - f^2(a_j^2 + (b_j^2 - a_j^2)t)| < \epsilon$ for all $t \in [0, 1]$.

Remark 1.2. — *A piecewise monotone map is* not *assumed to be continous. The above topology induce a topology on $C^0(I)$ which is neither stronger nor weaker than the usual one.*

Let us outline these notes. We shall recall in Section 2 some general facts about entropy for smooth, topological or probabilistic dynamical systems. In Section 3, after recalling the basics of the symbolic dynamics of piecewise monotone maps, we give combinatorial and geometric formulations of the entropy on the interval using in particular the special form of its symbolic dynamics defined by the kneading invariants. We discuss the continuity and monotonicity of the entropy function over maps and over invariant probability measures in Section 4.

To get to the global structure we shall use Hofbauer's Markov diagram explained in Section 5. This will leave a part of the dynamics which we analyze using "shadowing" in Section 6. The main part of the dynamics is reduced to a Markov shift with countably many states but most of the properties of the finite case (section 7). We apply these tools to the piecewise monotone maps getting a precise description of their measures of large or maximum entropy and their peirodic points, including a

complete classification from this point of view (section 8). We conclude in Section 9 by mentioning further works which either analyze more precisely the piecewise monotone maps or apply the techniques presented here to more general settings, following the idea that one-dimensional dynamics should be the gateway to (more) general results.

Remark 1.3. — *Theorems, Problems and Questions are numbered consecutively and independently throughout the paper. All other items are numbered in a common sequence within each section.*

Exercises should be rather straightforward and quick applications of techniques and ideas exposed in the text. **Problems** are more ambitious projects that (I believe) can be solved by standard techniques - but have not been done yet, to the best of my knowledge. **Questions** are problems that I don't know how to handle.

2. Generalities on Entropies

The dynamical entropies have a long story and are related to very basic notions in statistical physics, information theory and probability theory (see, e.g., [**56**]).

Putting the work of Shannon on a completely new level of abstraction, Kolmogorov and Sinai defined in 1958 the **measured entropy**[1] of any endomorphism T of a probability space (X, μ) as follows. For any finite measurable partition P, we get a P-valued process whose law is the image of μ by the map $x \mapsto (P(x), P(Tx), \dots) \in P^{\mathbb{N}}$. This process has a mean Shannon entropy:

$$h(T, \mu, P) := \lim_{n \to \infty} \frac{1}{n} H(\mu, P^n)$$

where $H(\mu, Q) = \sum_{A \in Q} -\mu(A) \log \mu(A)$ and

$$P^n := \{\langle A_0 \dots A_{n-1}\rangle := A_0 \cap T^{-1}A_1 \cap \dots T^{-n+1}A_{n-1} \neq \emptyset : A_i \in P\}$$

The elements of P^n are called the (geometric) P, n-**cylinders**.

The Kolmogorov-Sinai $h(T, \mu)$ is then the supremum of the Shannon entropies of all processes over finite alphabet "contained" in the considered dynamical system. We refer to the many excellent texts (see, e.g., [**66, 101, 102, 84, 87, 100**]) for more information and only quote a few facts here.

This supremum can look forbidding. However, Sinai showed: if $P_1, P_2, \dots$ is an increasing[2] sequence of finite measurable partitions such that $\{T^{-k}P_n : k, n \in \mathbb{N}\}$ generates the σ-algebra of measurable subsets of X, then:

$$h(T, \mu) = \sup_{n \geq 1} h(T, \mu, P_n)$$

Note the case where all P_n's are the same partition (said then to be **generating** under T).

The measured entropy is an invariant of measure-preserving conjugacy: if (X, T, μ) and (Y, S, ν) are two measure preserving maps of probability spaces and if $\psi : X \to Y$

---

[1] We prefer this nonstandard terminology to the usual, but cumbersome *measure-theoretic* and even more to the confusing *metric entropy*.
[2] That is, the elements of P_n are union of elements of P_{n+1} for all n.

is a bimeasurable bijection $\psi : X \to Y$ of probability spaces such that $\nu = \mu \circ \psi$ and $\psi \circ T = S \circ \psi$ (i.e., ψ is an isomorphism of (X, T, μ) and (Y, S, ν)) then T and S have the same entropy.

Stunningly, this invariant is complete for this notion of isomorphism among Bernoulli automorphisms according to Ornstein's theory (see [**84**] for an introduction and [**87**] for a complete treatment). Ornstein theory also (and perhaps more importantly) shows that many natural systems are measure-preserving conjugate to such a system (see [**80**]).

Let us note that the above define the entropy wrt not necessarily ergodic invariant probability measure. One shows also that $h(T, \mu)$ is an *affine function* of μ so that, if $\mu = \int \mu_x \nu(dx)$ is the ergodic decomposition of μ, then $h(T, \mu) = \int h(T, \mu_x) \nu(dx)$. We also note that $h(T^n, \mu) = |n| h(T, \mu)$ for all $n \geq 1$ (for all $n \in \mathbb{Z}$ if T is invertible).

From our point of view, the real meaning of measured entropy is given by the Shannon-McMillan-Breiman theoreom:

Theorem 1 (Shannon-McMillan-Breiman). — *Let T be a map preserving a probability measure μ on a space X. Assume that it is ergodic. Let P be a finite measurable partition of X. Denote by $P^n(x)$ the set of points y such that $f^k x$ and $f^k y$ lie in the same element of P for $0 \leq k < n$. Then, as $n \to \infty$:*

$$\frac{1}{n} \log \mu(P^n(x)) \to h(T, P, \mu) \qquad a.e. \text{ and in } L^1(\mu)$$

Corollary 2.1. — *Let T and P be a as above. For $0 < \lambda < 1$, let $r(P, n, \mu, \lambda)$ be the minimum cardinality of a collection of P, n-cylinders the union of which has μ-measure at least λ. Then:*

$$h(T, \mu, P) = \lim_{n \to \infty} \frac{1}{n} \log r(P, n, \mu, \lambda)$$

Exercise 2.2. — *Consider $X = \{0, 1\}^{\mathbb{N}}$ together with the shift σ and the product probability μ induced by $(p, 1 - p)$. Show that:*

$$h(\sigma, \mu) = -p \log p - (1 - p) \log(1 - p)$$

The following general fact is especially useful in dimension 1:

Theorem 2 (Rokhlin formula). — *Let T be an endomorphism of a probability space (X, μ). Assume that there is a generating countable measurable partition of X into pieces X_i on each of which the restriction $T|X_i$ is one-to-one and non-singular wrt μ and $-\sum_i \mu(X_i) \log \mu(X_i) < \infty$. Let $JT = d\mu \circ (T|X_i)/d\mu$ if $x \in X_i$. Then:*

$$h(T, \mu) = \int \log JF \, d\mu$$

Compare this with Pesin's formula (below).

Exercise 2.3. — *Recover the result of the previous exercise by applying Rokhlin formula to a suitable dynamical system on the interval.*

2.1. Bowen-Dinaburg formula. — R. Adler, A. Konheim, and M. McAndrew defined in 1965 a topological counterpart of the measured entropy for arbitrary continuous maps of compact spaces. Where measured entropy counts orbit segments "typical" for the measure (Corollary 2.1), topological entropy counts all orbit segments.

The original definition of topological entropy was given in terms of open covers. In the case of a compact metric space (X, d), we shall use the Bowen-Dinaburg formula which we now explain.

Definition 2.4. — *For a real number $\epsilon > 0$ and a positive integer n, the (ϵ, n)-ball at some $x \in X$ is:*

$$B(\epsilon, n, x) := \{y \in X : \forall 0 \leq k < n \ d(T^k y, T^k y) < \epsilon\}$$

The (ϵ, n)-covering number of $Y \subset X$ is:

$$r(\epsilon, n, Y) := \min\{\#C : Y \subset \bigcup_{y \in C} B(\epsilon, n, y)\}$$

The entropy of Y is then:

$$h(T, Y) := \lim_{\epsilon \to 0^+} h(T, Y, \epsilon) \ \text{with} \ h(T, Y, \epsilon) := \limsup_{n \to \infty} \frac{1}{n} \log r(\epsilon, n, Y)$$

The **topological entropy** *of T is:*

$$h_{\text{top}}(T) := h(T, X).$$

Thus, $h_{\text{top}}(T)$ counts *all* orbit segments.

Remark 2.5. — *1. $h_{\text{top}}(T)$ is an invariant of topological conjugacy.*

2. One can replace $\limsup$ by $\liminf$ in the definition of $h(T, Y, \epsilon)$. The number $h_{\text{top}}(T) = h(T, X)$ is then left unchanged.

Question 1. — *It is unknown if for $X = Y$ the $\limsup$ in $h(T, X, \epsilon)$ is in fact a limit.*

Exercise 2.6. — *Show that if T is Lipschitz with constant L on a compact manifold of dimension d then $h_{\text{top}}(T) \leq d \log^+ L$. Note: A very important refinement is Ruelle-Margulis inequality (see 3 and its descendants. It is false if T is only piecewise continuous and Lipschitz, see [16, 62].*

Exercise 2.7. — *Let $\sigma : \mathcal{A}^{\mathbb{N}} \to \mathcal{A}^{\mathbb{N}}$ be the shift map: $(\sigma(A))_n = A_{n+1}$. Let Σ be a* **subshift** *on a finite alphabet $\mathcal{A}$, i.e., Σ is a closed, σ-invariant subset of $\mathcal{A}^{\mathbb{N}}$. Show that*

$$h_{\text{top}}(\sigma|\Sigma) = \lim_{n \to \infty} \frac{1}{n} \log \#\{w \in \mathcal{A}^n : \exists x \in \Sigma \ s.t. \ x_0 \ldots x_{n-1} = w\}$$

If Σ is the **subshift of finite type** *(see [65] for background) defined by a matrix $A : \mathcal{A} \times \mathcal{A} \to \{0, 1\}$ according to $x \in \Sigma$ iff $x \in \mathcal{A}^{\mathbb{N}}$ and, for all $n \in \mathbb{N}$, $A(x_n, x_{n+1}) = 1$, then*

$$h_{\text{top}}(\sigma|\Sigma) = \log \rho(A)$$

where $\rho(A)$ is the spectral radius of A.

2.2. Katok's formula. — Let T be an endomorphism of a probability space (X, μ). Assume that X is also a compact metric space (the measurable structure being the Borel one). Katok [**51**] observed that one can compute the measured entropy by a variant of the Bowen-Dinaburg formula:

Definition 2.8. — *The (ϵ, n)-covering number of a probability μ wrt a parameter $\lambda \in (0, 1)$ is:*

$$r(\epsilon, n, \mu, \lambda) := \min \left\{ \# : \mu \left(\bigcup_{y \in C} B(\epsilon, n, y) \right) > \lambda \right\}$$

Proposition 2.9 (Katok). — *Assume that (T, μ) is ergodic. Then the measured entropy of (T, μ) is equal to:*

$$h(T, \mu) = \lim_{\epsilon \to 0^+} h(T, \mu, \epsilon) \ with \ h(T, \mu, \epsilon) = \limsup_{n \to \infty} \frac{1}{n} \log r(\epsilon, n, \mu, \lambda)$$

for any $0 < \lambda < 1$. $h(T, \mu, \epsilon)$ is independent of λ.

Remark 2.10. — *One can replace $\limsup$ by $\liminf$ in the above definition. However it is not known if this is in fact a limit*

2.3. Ruelle-Margulis Inequality. — Let $f : M \to M$ be a self-map of a compact manifold with an ergodic and invariant probability measure μ and an open subset U such that $\mu(U) = 1$ and $f|U$ is C^1.

The Lyapunov exponents of (f, μ) are numbers $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_d \geq -\infty$ such that, for a.e. $x \in M$:

$$\dim \overline{\{v \in T_x M : \lambda(x, v) = \lambda\}} = \#\{i : \lambda_i = \lambda\}.$$

Oseledets Theorem ensures that such a collection of d numbers exists if $\log^+ \|f'\|$ is μ-integrable (see, e.g., [**61**]).

Theorem 3 (Ruelle-Margulis inequality). — *If $\|f'\|$ is bounded over U, then*

$$h(f, \mu) \leq \sum_{i=1}^{d} \lambda_i^+$$

See [**53**] for refinements (i.e., how fast can one let f blow up near ∂U).

2.4. Variational Principle. — The above formulas imply immediately that $h(T, \mu) \leq h_{\text{top}}(T)$ for all measurable maps T of a compact metric space and all invariant probability measures μ. In fact much more is true:

Theorem 4 (Goodman, Dinaburg). — *For a continuous map of a metric space, the variational principle holds:*

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

The supremum can be taken either over all invariant probability measures; or over all ergodic and invariant probability measures.

The classical proof is due to Misiurewicz (see the textbooks, e.g., [**84**])

Remark 2.11. — *1. This of course "justifies" the definition of topological entropy to the extent that this quantity is shown to depend only on the measurable structure of $T : X \to X$ and not on its topological (or metric) structure.*

2. It was not at all obvious that one could find measures with "almost all the topological complexity" of the system. For instance, the opposite is true in the setting of Birkhoff theorem: e.g., for the shift on $\{0,1\}^{\mathbb{N}}$, the set of points for which $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \phi \circ \sigma^k(x)$ exists for all continuous function ϕ is of full measure but is of first Baire category (i.e., it is negligible from the topological point of view).

Exercise 2.12. — *Prove this last assertion.*

The variational principle leads to the study of invariant probability measures that achieve the topological entropy. More generally, one makes the following

Definition 2.13. — *For a Borel map $T : X \to X$, the entropy is defined as*

$$h(T) := \sup_{\mu} h(T, \mu)$$

where μ ranges over all the invariant probability measures. The **maximum measures** *are the ergodic, invariant probability measure maximizing entropy, i.e., such that $h(T, \mu) = h(T)$.*

Recall that the entropy is an affine function of the measure, hence the invariant probability measures maximizing entropy are exactly the closed convex hull of the ergodic ones, i.e., the maximum measures.

Maximum measures sometimes do not exist and sometimes are not unique. For instance, for each finite r, there are C^r maps with no maximum measures and also C^r maps with infinitely many maximum measures [**15, 90**] (see [**70**] for examples of diffeomorphisms). But we shall see that for many nice systems, such measures exist, have some uniqueness and describe important characteristics of the dynamics, e.g., the repartition of the periodic points.

2.5. Misiurewicz's local entropy. — A key aspect of measured entropy is the uniformity (or lack thereof) in the limits, when μ ranges over the invariant probability measures, $h(f, \mu) = \lim_{\epsilon \to 0} h(f, \mu, \epsilon)$. It corresponds to the existence of "complexity at arbitrarily small scales". The simplest way[3] to measure this is the following quantity introduced by Misiurewicz (under the name *topological conditional entropy*) [**71**]:

Definition 2.14 (**Misiurewicz**). — *The* **local entropy** *of a self-map f of a metric space (X, d) is:*

$$h_{\mathrm{loc}}(f) := \lim_{\epsilon \to 0} h_{\mathrm{loc}}(f, \epsilon) \ \text{with} \ h_{\mathrm{loc}}(f, \epsilon) := \lim_{\delta \to 0} \limsup_{n \to \infty} \sup_{x \in X} \frac{1}{n} \log r(\delta, n, B(\epsilon, n, x)).$$

---

[3] Much deeper results have been obtained in the study of symbolic extension entropy, see the references in Section 9.

Exercise 2.15. — *Show that, if X is compact and f is continuous, then $h_{\mathrm{loc}}(f) = \lim_{\epsilon \to 0} \tilde{h}_{\mathrm{loc}}(f, \epsilon)$ where $\tilde{h}_{\mathrm{loc}}(f, \epsilon) := \sup_{x \in X} \lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{n} \log r(\delta, n, B(\epsilon, n, x))$.*

Exercise 2.16. — *Show that if X is compact, but f not necessarily continuous, then $h_{\mathrm{loc}}(f^n) = |n| h_{\mathrm{loc}}(f)$ for all $n \in \mathbb{N}$ (or $n \in \mathbb{Z}$ if f is invertible).*

A simple but important consequence motivating the above definition is:

Proposition 2.17. — *$h_{\mathrm{loc}}(f)$ bounds the defect in upper semicontinuity of $\mu \mapsto h(f, \mu)$ over the set of invariant probability measure with the weak star topology: for any sequence $\mu_n \to \mu$,*

$$\limsup_{n \to \infty} h(f, \mu_n) \le h(f, \mu) + h_{\mathrm{loc}}(f)$$

In particular, if $h_{\mathrm{loc}}(f) = 0$, then the above map is upper semi-continuous and therefore achieves its maximum: there exists at least one maximum entropy measure.

Exercise 2.18. — *Prove this result.* Hint: *for any $0 < \epsilon' < \epsilon$, for all n large enough, $h(f, \mu_n, \epsilon) \le h(f, \mu, \epsilon')$.*

$h_{\mathrm{loc}}(f) = 0$ for C^∞ maps follows from Yomdin's theory to which we now turn.

2.6. Yomdin's theory. — Yomdin's theory analyzes the local complexity of differentiable maps. It explains why C^∞ interval maps are so much like piecewise monotone maps (see [**15**] or section 9).

Theorem 5 (following Yomdin). — *Let $f : \mathbb{R}^d \to \mathbb{R}^d$ be a C^r map with all partial derivatives up to order r bounded by 1, except for the derivatives of order 1 maybe.*

Let $\sigma : [0, 1]^k \to \mathbb{R}^d$ extend to a C^r map defined on a neighborhood of $[0, 1]^k$ with all partial derivatives up to order r bounded by 1.

Then there exists a tree $\mathcal{T}$ (i.e., a collection of finite sequences such that if $\tau_1 \dots \tau_n \in \mathcal{T}$ then $\tau_1 \dots \tau_k \in \mathcal{T}$ for all $1 \le k \le n$) and a family of maps $\phi_\tau : [0, 1]^k \to [0, 1]^k$, $\tau \in \mathcal{T}$, with the following properties. Set for any $\tau = \tau_1 \dots \tau_n \in \mathcal{T}_n$ (the subset of sequences of length n),

$$\Phi_{\tau_1 \dots \tau_n} := \phi_{\tau_1 \dots \tau_n} \circ \phi_{\tau_1 \dots \tau_{n-1}} \circ \cdots \circ \phi_{\tau_1}.$$

Then, for all $n \ge 1$,

- *$\sigma^{-1}(B(1, n, \sigma(0))) \subset \bigcup_{\tau \in \mathcal{T}_n} \Phi_\tau([0, 1]^k) \subset \sigma^{-1}(B(2, n, \sigma(0)))$*
- *for each $\tau \in \mathcal{T}_n$, ϕ_τ, Φ_τ and $f^n \circ \sigma \circ \Phi_\tau$ have all their partial derivatives up to order r bounded by 1;*
- *$\#\mathcal{T}_n \le C(r, d) \mathrm{Lip}^+(f)^{n(k/r)+1}$ where $\mathrm{Lip}^+(f)$ is the maximum of the Lipschitz constant of f and 1.*

A slightly weaker claim was prove by Yomdin [**110**]. However simple modifications (made in [**15**]) yields the above statement. We do not give a proof here, but only sketch the main ideas.

Sketch of Proof. — The first step of the proof is to observe that one can approximate $f^n \circ \sigma : [0,1]^k \to \mathbb{R}^d$ by its Taylor polynomial, after restricting it to a cube of linear size $o(\mathrm{Lip}^+(f)^{-n/r})$. There are essentially $\mathrm{Lip}^+(f)^{n(k/r)}$ such cubes in the domain $[0,1]^k$. This accounts for the corresponding factor in the above Theorem.

The second step of the proof considers the intersection with the unit cube of the graph of the polynomial approximation of $f^n \circ \sigma$. This is a **semi-algebraic subset**, i.e., a subset of $\mathbb{R}^d$ defined by finitely many equalities and inequalities involving only polynomials. It turns out that such subsets can be written as the union of the images of a *constant number* of C^r maps with all partial derivatives bounded by 1. Here by a constant number, we mean one that depends only on the order of differentiability r, dimension d and the list of the degrees of the polynomials defining the subset. This natural but deep fact was explained in [**39**] and a complete proof can be found in [**13**].

Applications. — The initial motivation of Yomdin was the following theorem:

Theorem 6 (Yomdin). — *Let f be a C^r self-map of a compact d-dimensional Riemannian manifold M. Define the volume growth of $f \in C^r(M)$, $r \geq 1$, by:*

$$v(f) := \max_{0 \leq k \leq d} \sup_S \limsup_{n \to \infty} \frac{1}{n} \log \int_S \|\Lambda^k (f^n)'(x)\| \, dvol_S$$

where S ranges over all compact subsets of k-dimensional C^r submanifolds of M, vol_S is the volume induced by the restriction to S of the Riemannian structure of M and $\|\Lambda^k (f^n)'(x)\|$ is the Jacobian corresponding to vol_S and $vol_{f^n S}$. Thus $v(f)$ is the growth rate of all volumes with multiplicity.

Then,

(2.19) $$v(f) \leq h_{\mathrm{top}}(f) + \frac{d}{r} \log^+ \mathrm{Lip}^+(f)$$

In particular, for C^∞ maps, the logarithm of the spectral radius of the action of f on the total homology is bounded by the topological entropy.

Exercise 2.20. — *Deduce the above theorem from Theorem 5.* Hint: *For any $0 \leq k \leq n$, bound $\int_B \|\Lambda^k (f^n)'(x)\| \, dvol_B$ for B contained in an (ϵ, n)-ball.*

Note: The last assertion is *Shub conjecture* in the case of C^∞ maps. The conjecture remains open for less regular functions, in particular C^1 functions (for which it was stated).

The bound (2.19) is sharp among maps, see [**110**].

Exercise 2.21. — *Prove that for a C^r map of a compact d-dimensional manifold, $h_{\mathrm{loc}}(f) \leq \frac{d}{r} \log \mathrm{Lip}^+(f)$.* Hint: *apply Theorem 5 to a suitable F representing a scaled version of f and σ being a suitable chart and consider $\bigcup_{w \in \mathcal{T}_n} \sigma \circ \Phi_\tau(Q_\epsilon)$ where Q_ϵ is an ϵ-dense subset of $[0,1]^d$.*

Even more interestingly (for higher dimensional dynamics), Yomdin's theory implies a sort of submultiplicative property among families of maps of the form $f^n \circ \sigma$ where f is the dynamics and σ is a parametrized disk (see [**19**]).

3. Computing the Entropy on the Interval

We now turn to one-dimensional dynamics. Here entropy is somewhat "explicit". This first appears in two results of Misiurewicz which say that the entropy for interval maps is essentially a combinatorial phenomenon.

3.1. Symbolic Dynamics. — We first recall the general definition of the symbolic dynamics of a piecewise map before explaining the classical description of the subshifts thus obtained in the class of piecewise monotone maps.

Definition 3.1. — *Let f be a self-map of a Baire topological space X with a collection P of disjoint open subsets with dense union. The symbolic dynamics of f wrt P is the left-shift σ acting on:*

$$\Sigma_+(f, P) := \overline{\{\iota(x) : x \in X'\}} \subset P^{\mathbb{N}}$$

where the topology on $P^{\mathbb{N}}$ is the product of the discrete topology of P and $X' := \bigcap_{k \geq 0} f^{-k} \bigcup_{A \in P} A$ and

$$(3.2) \qquad\qquad \iota(x) := A \in P^{\mathbb{N}} \text{ such that } \forall k \geq 0 \ f^k x \in A_k.$$

*$\iota(x)$ is called the **itinerary** of x.*

Exercise 3.3. — *For $A_0 \ldots A_{n-1} \in P^n$, let*

$$\langle A_0 \ldots A_n \rangle := A_0 \cap f^{-1} A_1 \cap \cdots \cap f^{-n+1} A_{n-1} \subset X.$$

Check that $\Sigma_+(f, P) = \{A \in P^{\mathbb{N}} : \forall n \geq 0 \ \langle A_0 \ldots A_n \rangle \neq \emptyset\}$.

For piecewise monotone map, the symbolic dynamics is very close to the interval dynamics:

Exercise 3.4. — *Show that for a piecewise monotone map f with its natural partition P, ι is well-defined except on a countable set. Show that ι is one-to-one except on a set negligible wrt any nonatomic invariant probability measure.* Hint: *Consider the homtervals of f, i.e., the maximum open intervals on which f^k is monotone and continuous for all $k \geq 1$.*

Give a counter-example for a smooth interval map. Can you obtain $h_{\text{top}}(\sigma | \Sigma_+(f, P)) > h_{\text{top}}(f)$ or the reverse inequality? See [**16, 62**] *for some piecewise affine examples in higher dimensions.*

3.2. Kneading theory. — Let (T, P) be a piecewise monotone map.

Endpoints of pieces in P have no itinerary in the proper sense. However, one can define left/right itineraries of any point $x \in I$:

$$\iota(x\pm) := \lim_{\epsilon \to 0^+} \iota(x \pm \epsilon).$$

If $c_0 < c_1 < \cdots < c_N$ are the endpoints of the intervals in the partition P of $I = [c_0, c_N]$, the set of kneading invariants is:

$$K(T, P) := \{\iota(\gamma) : \gamma \in \{c_0+, c_1-, c_1+, \ldots, c_N^-\}\}$$

These are obviously invariants of topological conjugacy. In fact, they determine completely the symbolic dynamics (hence, they are (essentially) complete invariants of topological conjugacy – see [**67**] for the details).

Definition 3.5. — *T defines a* **modified lexicographic order** *on $P^{\mathbb{N}}$ according to:*

$$A \preceq_T B \iff A = B \text{ or } (A_0 \ldots A_{n-1} = B_0 \ldots B_{n-1} \text{ and}$$
$$\langle A_0 \ldots A_{n-1} A_n \rangle < \langle B_0 \ldots B_{n-1} B_n \rangle)$$

where $I < J$ means that for every $x \in I$, $y \in J$, $x < y$.

This is the standard lexicographic order if T is increasing on each element of P.

Exercise 3.6. — *Show that $\iota : (I, \leq) \to (\Sigma_+, \preceq_T)$ is order preserving.*

Proposition 3.7. — *Let (T, P) be a piecewise monotone map. $A \in P^{\mathbb{N}}$ belongs to $\Sigma_+(T, P)$ iff*

$$(3.8) \qquad\qquad \forall n \geq 0 \quad \iota(e_n) \preceq_T \sigma^n A \preceq_T \iota(f_n)$$

where $e_n = c_i+$ and $f_n = c_{i+1}-$ if $(c_i, c_{i+1}) = A_n$.

Exercise 3.9. — *Prove the above.*

It is remarkable that the symbolic dynamics of piecewise monotone maps admits such a simple description, even though it is not a finite one as for subshifts of finite type (but this is impossible in general as, for instance, their topological entropy can take uncountably many values — see 3.12). We shall see that these subshifts are however very close to being of finite types and share many of their properties. The notion of *quasi-finite type* described in Section 9 put this observation to work.

Remark 3.10. — *(3.8) is satisfied in particular by the kneading invariants themselves. This can be formulated abstractly as a property of a finite alphabet $\mathcal{A}$ where each element is given a sign and two sequences in $\mathcal{A}^{\mathbb{N}}$. The abstract version of (3.8) is then necessary and sufficient to ensure that these data are realized (up to an obvious identification) as the kneading invariants of some piecewise monotone map. See [**67**].*

3.3. Misiurewicz and Szlenk's lap numbers. — This formula says that the topological entropy of a piecewise monotone map is combinatorial:

Theorem 7 (**Misiurewicz-Szlenk**). — *Let $f : [0, 1] \to [0, 1]$ be a piecewise monotone and piecewise continous map. Let P be its natural partition. Consider the partition P^n into P, n-cylinders. The topological entropy of f can be computed as:*

$$(3.11) \qquad\qquad h_{\text{top}}(f) = \lim_{n \to \infty} \frac{1}{n} \log \#P^n$$

*$\#P^n$ is called the n-**lap number** of f.*

Using the Hofbauer diagram below one should be able to solve the following:

Problem 1. — *The rate of convergence in the above formula is exponential.*

One can give a very easy ergodic proof of the Misiurewicz-Slenk formula:

Recognize in eq. (3.11) the topological entropy of the symbolic dynamics $\Sigma_+(f, P)$. The equality $h_{\text{top}}(\sigma|\Sigma_+(f, P)) = h_{\text{top}}(f)$ will follow rather easily from the variational principle. Indeed, according to this principle it is enough to see that one can identify invariant and ergodic probability measures of non-zero entropy of both systems in an entropy preserving way. Conclude by applying Exercise 3.4.

Exercise 3.12. — *Prove that if $\beta > 1$ and $T : x \mapsto \beta x \mod 1$ on $[0, 1]$, then $h_{\text{top}}(T) = \log \beta$ (you may consider first the case $\beta \in \mathbb{N}$).* Hint: *Use Ruelle's inequality.*

3.4. Misiurewicz's Horseshoes. —

Definition 3.13. — *A* **horseshoe** *is a collection of pairwise disjoint compact intervals $J_1, \ldots, J_N$ and an integer $T \geq 1$ such that $f^T(J_i)$ contains a neighborhood of $J_1 \cup \cdots \cup J_N$. The entropy of the horseshoe is $\log N/T$.*

A **piecewise monotone horseshoe** *is a horseshoe such that, in the above notation, $f^T|J_i$ is continuous and strictly monotone.*

Theorem 8 (Misiurewicz). — *Let f be (i) a continuous map or (ii) a piecewise monotone map of a compact interval. If the topological entropy of f is not zero, then it is the supremum of the entropies of its horseshoes (its piecewise monotone horseshoes if f is piecewise monotone).*

Complete proofs can be found in [**2**] and also [**93**]. These are combinatorial proofs. We give one using ergodic theory and techniques involved in the *spectral decompositon* of A.M. Blokh [**5**]. We consider only the case $f \in C^0(I)$.

Exercise 3.14. — *Prove Misiurewicz's theorem for piecewise monotone maps either by adapting the proof below or by using Hofbauer's Markov diagram (see below).*

Ergodic construction of Misiurewicz horseshoes. — We have to build a horseshoe of entropy at least H for every $H < h_{\text{top}}(f)$. Fix such a H. By the variational principle, we can find an ergodic and invariant probability measure μ such that $h(f, \mu) > H$. By Katok's entropy formula, we can find $\epsilon_0 > 0$ so that $h(f, \mu, \epsilon_0) > H$.

Step 1: *A Cycle*

Let $\text{supp}'\mu$ be the support of μ minus the countably many points x such that $\mu([x, x + t])\mu([x - t, x]) = 0$ for $t > 0$ small enough. In the rest of the construction, x will be a point of $\text{supp}'\mu$.

If I is an open interval, define $F(I)$ as the interior of $f(I)$. For $r > 0$ and $x \in [0, 1]$, define:

$$V_r(x) := \bigcup_{k \geq 0} F^k(B(x, r)) \text{ and } C_r(x) \text{ is the c.c. of } V_r(x) \text{ containing } x.$$

As $x \in \text{supp}'\mu$, there exists $k \geq 1$ such that $F^k(B(x, r)) \cap B(x, r)$ has positive measure. Hence $V_r(x)$ has finitely many, say $n(x, r)$, connected components.

Exercise 3.15. — *Check that F permutes these connected components and that $x \in V_r(x)$.*

We claim that

(3.16) $$\sup_{r>0} n(x,r) < \infty.$$

An entropy argument shows that the proportion $\xi(x,r)$ of $0 \leq k < n(x,r)$ such that $\text{diam}(F^k C_r(x)) > \epsilon_0$ satisfies: $\xi(x,r) \geq H/\log(\epsilon_0^{-1}+1)$. Indeed, $r(\epsilon_0, n, C_r(x)) \leq (\epsilon_0^{-1}+1)^{\xi(x,r)n}$ so $H < h(f, \mu, \epsilon_0) \leq \xi(x,r)\log(\epsilon_0^{-1}+1)$. But, considering the lengths of the pairwise disjoint $F^k(C_r(x))$, we have: $(1 - \xi(x,r))n(x,r)\epsilon_0 \leq 1$, proving eq. (3.16). This implies that $\bigcap_{r>0} C_r(x) =: C(x)$ is a non-trivial interval defining a cycle $V(x)$ of period $p(x)$ with $F^{p(x)}(C(x)) = C(x)$ (if $F^{p(x)}(C(x)) \subsetneq C(x)$, then taking $r > 0$ so small that $F^{p(x)}(C_r(x)) \cup B(x,r) \subsetneq C(x)$ we would get $C_r(x) \subsetneq C(x)$, a contradiction).

If y is another point of $\text{supp}'\mu$, we claim that $V(y) = V(x)$. Indeed, observe that $\mu(V(x)) = 1$ by ergodicity, so $\text{supp}\mu \subset \overline{V(x)}$. As $y \in \text{supp}'\mu$, $y \in V(x)$. Hence $B(y,r) \subset V(x)$ for $r > 0$ small enough, so $V(y) \subset V(x)$. By symmetry, $V(y) = V(x)$. Henceforth, we drop the dependence on x, simplifying $V(x), C(x), p(x)$ to C, V, p.

Step 2: *Strong Mixing*

Call a subinterval $I \subset [0,1]$ such that $\bigcup_{k\geq 0} F^k(I) \subsetneq V$ **negligible**. Introduce the following relation on $[0,1]$: $x \sim y$ iff $[x,y]$ is negligible. Observe that $x \sim y$ implies $f(x) \sim f(y)$ as $[f(x), f(y)] \subset f([x,y])$ and that the equivalence classes of $\sim$ are compact subintervals.

Exercise 3.17. — *Show that $f : [0,1] \to [0,1]$ induces a continuous self-map of $[0,1]/\sim$ and that, up to a homeomorphism, this new map is a continuous map f' of $[0,1]$ which admits an invariant and ergodic probability measure μ' satisfying $h(f', \mu', \epsilon_0') > H$ for some $\epsilon_0' > 0$.*

By this exercise one can assume that no non-trivial subinterval of V is negligible. Hence for every non-trivial subinterval $J \subset C$, $\bigcup_{k\geq 0} F^k J = V$.

Observe that $f^p(C) = C$ implies the existence of a fixed point $f^p(x_0) = x_0$ in the closure $\bar{C}$. In particular, there is an integer $n_0 > 0$ such that $f^{n_0} J \ni x_0$ by the following:

Exercise 3.18. — *Show that either one can find $x_0 \in C$ (in which case we set $y_0 := x_0$), or there exists $y_0 \in V$ such that $f(y_0) = x_0$.*

Observe that $f^{n_0} J$ cannot be negligible, as it would imply that all connected components of $\bigcup_{k=0}^{n_0-1} f^k J$ must also be negligible, but this union must have full μ-measure, a contradiction.

If $x_0 \in \partial C$, let $z_0 \in C$ such that $\mu([x_0, z_0]) > 0$. Otherwise, let $w_0 \in \partial C$ such that $\mu([x_0, w_0]) > 0$ and $z_0 \in (x_0, w_0)$ such that $\mu([x_0, z_0])\mu([z_0, w_0]) > 0$. In both cases, set $K_0 := [x_0, z_0]$.

As $f^{n_0} J$ is not negligible, there exists $n_1 > n_0$ such that $f^{n_1} J \ni z_0$, so $f^{n_1} J \supset K_0$.

Step 3: *Conclusion*

Divide V into subintervals $I_1, \ldots, I_L$, each of length $\epsilon_0/2$. By Step 2, for each $i = 1, \ldots, L$, a positive integer t_i such that $f^{t_i}(I_i) \supset K_0$. Let $t_* := \max t_i$.

By Katok's entropy formula, $\mu(K_0) > 0$ implies that, for all large n, there exist $x_1 < x_2 < \cdots < x_N$ points in K_0 which are (ϵ_0, n)-separated with $N \geq 2ne^{t_* H}e^{nH}$: for every $i = 1, \ldots, N-1$, there exists $n_i < n$ such that $|f^{n_i}x_i - f^{n_i}x_{i+1}| \geq \epsilon_0$. Hence, $f^{n_i}([x_i, x_{i+1}]) \supset I_j$ for some $j = j_i \in \{1, 2, \ldots, L\}$. Thus, $f^{n_i + t_j}([x_i, x_{i+1}]) \supset K_0$.

The integers $n_i + t_{j_i}$ belong to $[0, n + t_*]$. Hence, one can select $i_1, \ldots, i_M$ from $\{1, \ldots, N\}$ with $M = N/(n+t_*) \geq e^{(n+t_*)H}$, $j(i_1) = \cdots = j(i_M)$ and $T := n_{i_1} + t_{j(i_1)}$ such that:

- $J_\ell := I_{i_\ell} \subset\subset K_0$ for $\ell = 1, \ldots, M$;
- $J_\ell \cap J_{\ell'} = \emptyset$ if $\ell \neq \ell'$;
- $f^T(J_i) \supset K_0$

That is, f has a horseshoe with entropy at least H, proving Misiurewicz theorem.

3.5. Entropy and length. — Entropy is reflected in the growth rate of volume. The situation is especially simple on the interval [**74, 110**].

***Proposition 3.19* (Misiurewicz-Przyticky, Yomdin).** — *For $f \in C^1(I)$ define the growth of length under iterations of f to be:*

$$\ell(f) := \limsup_{n \to \infty} \frac{1}{n} \log \int_I |(f^n)'(x)| \, dx.$$

Then $\ell(f) \geq h_{\text{top}}(f)$. For $f \in \text{PMM}(I) \cap C^1(I)$ or $f \in C^\infty(I)$, this is in fact an equality:

$$(3.20) \qquad\qquad h_{\text{top}}(f) = \ell(f).$$

The above remain true for $f \in \text{PMM}(I)$, provided that $\ell(f)$ is suitably redefined using the variation of f^n (minus that occuring at its discontinuities). The proof is the same for $f \in \text{PMM}(I) \cap C^0(I)$ and slightly more delicate for the general case.

This formula has generalizations to C^∞-smooth self-maps in arbitrary dimension – see [**77**] and also [**60**]. One can compare it with the following consequence of Rokhlin formula (Theorem 2). Let f be a piecewise monotone, piecewise C^1 map of the interval. If μ is an invariant probability measure with absolutely continuous ergodic components, then:

$$h(f, \mu) = \int \limsup_{n \to \infty} (1/n) \log |(f^n)'(x)| \, d\mu$$

This has been generalized to arbitrary invariant probability measures absolutely continuous (this is Pesin's formula) or not and self-maps on manifolds of arbitrary dimension by taking into account the relevant *dimensions* (see [**109, 48**]) and considering the *Lyapunov exponents* instead of the single derivatives (see [**53, 64**]).

Exercise 3.21. — *Check (3.20) in the case where f maps each of its piece to the full interval (up to its endpoints).*

PROOF: We first prove $\ell(f) \geq h_{\text{top}}(f)$ for all C^1 maps. Let $\epsilon > 0$. For all large n, there exists $x_1 < \cdots < x_N$ a (ϵ, n)-separated subset with $N \geq n e^{n(h_{\text{top}}(f)-\epsilon)}/\epsilon$. As in the proof of Theorem 8, we can find a subset $y_1 < \cdots < y_M$ with $M \geq N/n$ and $0 \leq m \leq n$ such that $|f^k y_i - f^k y_{i+1}| \geq \epsilon$. Therefore,

$$\frac{1}{n} \log \int_I |(f^m)'(x)| \, dx \geq e^{m(h_{\text{top}}(f)-\epsilon)}$$

which implies the claimed inequality.

We prove the converse inequality. The C^∞ case follows from Yomdin's theory - Theorem 5 - we leave its details to the diligent reader. For $f \in \text{PMM}(I) \cap C^1(I)$, we observe:

$$\int_I |(f^n)'(x)| \, dx = \sum_{A \in P^n} |f^n(A)| \leq \#P^n$$

so that the Misiurewicz-Szlenk formula is enough to conclude. $\qquad\square$

Remark 3.22. — *Misiurewicz and Przytycki [74] have shown that, in any dimensions, the logarithm of the degree is a lower bound for the entropy for C^1 maps (but not C^0 maps). This is a (rather) special case of the Shub entropy conjecture mentioned before.*

4. Entropy as a function

4.1. Local entropy. — Recall that local entropy quantifies the complexity at arbitrarily small scales. The situation is not so bad for "reasonable" interval maps:

Proposition 4.1. — *For $f \in \text{PMM}(I)$ or $f \in C^\infty(I)$, $h_{\text{loc}}(f) = 0$.*

Exercise 4.2. — *If $K \subset I$ is such that $f^k|K$ is monotone for $0 \leq k < n$, then $r(\epsilon, n, K) \leq n(|I|/\epsilon + 1) + 1$.*

PROOF: We have already explained how the assertion for $f \in C^\infty(I)$ follows from Yomdin's theory. Assume $f \in \text{PMM}(I)$.

Consider first $f \in \text{PMM}(I)$. The exercise below implies that $r(\epsilon, n, J) \leq Cn\#\{A \in P^n : A \cap J \neq \emptyset\}$ where P is the natural partition of f. Now, if $K \subset B(\epsilon, n, x)$, then each $f^k K$, $0 \leq k < n$, meets at most two elements of P, hence $\#\{A \in P^n : A \cap J \neq \emptyset\} \leq 2^n$ so that $h_{\text{loc}}(f) \leq 2$. Applying the same to f^N we get, by Exercise 2.16 $h_{\text{loc}}(f) = (1/N)h_{\text{loc}}(f^N) \leq \log 2/N$. Letting $N \to \infty$, we get the claim for piecewise monotone maps. $\qquad\square$

4.2. Measured entropy. — Recall that a map $\phi : X \to \mathbb{R}$ is *lower semi-continuous* at $x \in X$, if, for $\epsilon > 0$, there exists a neighborhood of x on which ϕ is at least $\phi(x) - \epsilon$: that is, ϕ cannot "collapse". *Upper semi-continuity* of ϕ is lower semi-continuity for $-\phi$: it means that ϕ cannot "explode". In particular an upper semi-continuous function over a non-empty compact subset achieves its maximum.

The preceding result on the local entropy yields the:

Corollary 4.3. — *For any $f \in \mathrm{PMM}(I)$ or $f \in C^\infty(I)$, the measured entropy is upper semi-continuous on the compact of invariant probability measure and therefore f admits measures of maximum entropy.*

Note that such an existence proof says nothing about the multiplicity or the structure of these measures. These two points will require a rather complete analysis of the one-dimensional dynamics through Hofbauer's Markov diagram below.

Fact 4.4. — *If $f \in C^0(I)$ or $f \in \mathrm{PMM}(I)$ and $h(f, \mu_0) > 0$, then $\mu \mapsto h(f, \mu)$ is not lower semi-continuous at $\mu = \mu_0$.*

This follows from the density of periodic measures (invariant and ergodic probability measures defined by periodic orbits) among all invariant measures (which holds because of specification properties enjoyed by these maps (see [**5, 17, 47**])).

Fact 4.5. — *If $f \in C^\infty(I)$ or $f \in \mathrm{PMM}(I)$, then $\mu \mapsto h(f, \mu)$ is upper semi-continuous.*
For any $r < \infty$ there are examples of $f \in C^r(I)$ with invariant probability measures μ_0 at which the entropy fails to be upper semicontinuous.

The first fact follows from $h_{\mathrm{loc}}(f) = 0$ (see Proposition 4.1). For $r < \infty$, examples on the interval are given in [**15, 90**]. *Note:* In these examples, $h(f, \mu_0) = 0$.

Remark 4.6. — *S. Ruette [**92**] observed that this gives examples of topologically mixing, smooth interval maps with the same nonzero topological entropy which are not Borel conjugate.*

Remark 4.7. — *M. Rychlik [**94**] proved that the measured entropy of the unique absolutely continuous invariant probability measure of a family of piecewise C^2, piecewise expanding map has modulus of continuity $x \log 1/x$.*

The following should be tractable by using the perturbative results of Keller and Liverani [**57**] about the transfer operator.

Problem 2. — *Study the regularity of the measured entropy for families of piecewise expanding, piecewise monotone maps.*

4.3. Topological entropy. — We consider the dependance of the topological entropy on the map.

4.3.1. Continuity. —

Theorem 9. — *The map $f \mapsto h_{\mathrm{top}}(f)$ is lower semi-continuous both over $C^0([0,1])$ and over $\mathrm{PMM}([0,1])$.*
This map is upper semi-continuous (and therefore continuous) over $C^\infty([0,1])$ and over $\mathrm{PMM}([0,1])$ but not over $C^r([0,1])$ for any finite r.

Remark 4.8. — *In our choice of topology over $\mathrm{PMM}(I)$, two maps can be close only if they have the same number of pieces. This is not always the chosen definition but it is necessary to get the upper semi-continuity.*

It is easy to find a counter-example to the upper semi-continuity in the $C^0(I)$ topology even among piecewise monotone maps:

Exercise 4.9. — *Let $f_t(x) = (1-t)(1-2t|x|/(1-t))$ if $|x| < 1 - t$ and $f_t(x) = 0$ otherwise. Show that $h_{\mathrm{top}}(f_t) = \log(2t)$ pour $t < 1$ and $h_{\mathrm{top}}(f_1) = 0$.*
Can you find such an example among polynomials?

The basic phenomenon here is that a periodic point becomes a cycle of intervals containing a horseshoe. This in fact the only obstruction – see [**2**, Prop. 4.5.3]. We refer to [**67**, Th. 9.1] or [**2**, Cor. 4.5.5] for two rather different proofs: the original one, due to Milnor and Thurston, which involves the "kneading determinant" which is a sort of zeta function, and a combinatorial one, closer to the spirit of this paper. We indicate below the outline of the proof in the smooth case.

The proof of the above theorem will use the:

Exercise 4.10. — *Show that for every $\epsilon > 0$, $C(g^n) \subset B(C(f^n), \epsilon)$ for all g close enough to $f \in \mathrm{PMM}(I)$.*

Exercise 4.11. — *Assume that for all $\alpha > 0$, there exist $\epsilon_0 > 0$ and a continuous function $C_\alpha : (0, \infty) \to (0, \infty)$ such that, for all g in a neighborhood of f, all $\delta > 0$:*

$$(4.12) \qquad \forall n \geq 1 \ r_g(\delta, n, B_g(\epsilon_0, n, x)) \leq C_\alpha(\delta)e^{\alpha n}.$$

Show that $g \mapsto h_{\mathrm{top}}(g)$ is then upper semi-continuous at $g = f$.

PROOF: The lower semi-continuity follows from Theorem 8: recall that a horseshoe is defined by finitely many conditions like (i) $f^n|[a, b]$ continuous; (ii) $\mathrm{int} f^n([a, b]) \supset [c, d]$ and each of these is equivalent to $f^n|[a, b]$ continuous and $f^n(a') < c$ and $f^n(b') > d$ for some $a', b' \in [a, b]$. This is obviously an open condition in $C^0(I)$. It is also open in $\mathrm{PMM}(I)$ by Exercise 4.10.

The upper semi-continuity of the topological entropy in the C^∞ smooth case follows from Yomdin's theory which gives, not only that $h_{\mathrm{loc}}(f) = 0$, but the very strong uniformity stated as the assumption in (4.12). By the above exercise, this concludes the proof of the proposition.

We refer the reader to the references quoted above for the proof in the piecewise monotone case. □

Lower semi-continuity also holds for surface diffeomorphisms of class $C^{1+\epsilon}$ according to a classical result [**51**] of Katok which shows the existence of horesehoes. This is false for surface homeomorphisms as can be deduced from examples of Rees. The natural case of C^1 diffeomorphisms is still open. It fails however in higher dimensions, even in the analytic category.

Exercise 4.13. — *Find a family of maps $F_\lambda : [0, 1]^2 \to [0, 1]^2$, $\lambda \in [0, 1]$ with $F_\lambda(t, x) = (\lambda t, f_t(x))$ such that $\lambda = 1$ is a discontinuity point for $\lambda \mapsto h_{\mathrm{top}}(F_\lambda)$.*

The following question seems not to have been studied though the characterization of entropy as an eigenvalue of a transfer operator should allow the application of techniques like [**57**].

Question 2. — *Study the modulus of continuity of topological entropy in families of piecewise monotone maps.*

4.3.2. Monotonicity. — The following fact was proved by Sullivan, Milnor, Douady and Hubbard:

Theorem 10. — *Let $Q_t(x) = 4tx(x - x)$ be the quadratic family $Q_t : [0,1] \to [0,1]$. Then $t \mapsto h_{\mathrm{top}}(Q_t)$ is non-decreasing for $t \in [0,1]$.*

Remark 4.14. — *This sounds very natural. One must be careful however that the similar statements involving other families of the type $t \mapsto tQ(x)$ for a smooth map $Q : [0,1] \to [0,1]$ satisfying $Q'(x) = 0$ iff $x = 1/2$ for $x \in [0,1]$. See [**11**]. See [**76, 38, 68**] for further results.*

Most of the proofs rely on complex dynamics (e.g., Teichmuller theory or quasiconformal maps). Let us describe the strategy of the remarkable "real"[4] proof found by Tsujii [**103**].

Recall how the kneading invariants of a piecewise monotonic map define its symbolic dynamics (Proposition 3.7) and therefore its topological entropy. Observe that $Q_t(0) = Q_t(1) = 0$ and Q_t is continuous, hence there is only one non-trivial kneading invariant, $\kappa(Q_t) := \iota(Q_t(1/2))$ and that, with respect to the lexicographic order defined by any of the Q_t, the symbolic dynamics as a subset and therefore the topological entropy as a number are non-decreasing functions of $\kappa(Q_t)$.

Proposition 3.7 implies that the map that associates to such a kneading invariant the topological entropy of the corresponding unimodal map is nondecreasing. Hence monotonicity of $t \mapsto \kappa(Q_t)$ (for the above order on $\{-1, 0, 1\}^{\mathbb{N}}$ will imply the monotonicity of the entropy. Hence the above theorem is reduced to showing the following property:

$$(4.15) \qquad \text{if } n := \min\{k \geq 1 : Q_t^k(0) = 0\} \text{ then, } \frac{\partial}{\partial t}Q_t^n(0) \text{ and } \frac{\partial}{\partial x}Q_t^n(Q_t(0))$$

Tsujii proves this last property by considering the following **Ruelle transfer operator** (see [**3**] for this very important notion):

$$R(\Psi)(x) = \sum_{y \in Q^{-1}(x)} \frac{\Psi(y)}{(\partial Q_t(y)/\partial y)^2}$$

acting on the space E linearly generated by $1/(z - Q^k(0))$, $k \geq 0$ (there are exactly n such functions because of the assumption on t). A computation shows that

$$\det(\mathrm{Id}_E - R|E) = \frac{\partial}{\partial t}Q_t^n(0)/\frac{\partial}{\partial x}Q_t^n(Q_t(0))$$

Now, Tsujii proves that $R|E$ is a contraction. This proves that $\phi(z) := \det(\mathrm{Id}_E - zR|E) \neq 0$ for all $|z| < \rho$ for some $\rho > 1$. Clearly $\phi(0) = 1$ and $\phi(z)$ is real for all real z. Hence, $\phi(1) > 0$, concluding Tsujii's proof.

---

[4] With the following restriction: as remarked by Tsujii in his paper, the map R below is a local version of Thurston map on the Teichmuller space - see [**63**].

5. The Markov Diagram

The main tool for the detailed analysis of the complexity of interval maps with nonzero entropy is Hofbauer's **Markov diagram**, which reduces much of their study to that of a countable state Markov shift with very good properties. It is inspired from a classical construction in the theory of subshifts of finite type and sofic subshifts: *the canonical extension*, which builds for any sofic subshift a topological extension which is a subshift of finite type with the same entropy — see [**65**].

5.1. Abstract construction. — Let us describe this construction, or rather a variant introduced in [**15**] which allows slightly stronger statements (and, more importantly, is necessary for generalizations, see remark 5.3 below).

The Hofbauer construction applies to any subshift. Let $\Sigma_+ \subset \mathcal{A}^{\mathbb{N}}$ be a one-sided subshift (i.e., a closed, shift-invariant subset of $\mathcal{A}^{\mathbb{N}}$, not necessarily of finite type). Let $\Sigma \subset \mathcal{A}^{\mathbb{Z}}$ be its natural extension:

$$\Sigma := \{A \in \mathcal{A}^{\mathbb{Z}} : \forall p \in \mathbb{Z} \; A_p A_{p+1} \cdots \in \Sigma_+\}$$

with the action of the (invertible) shift $\sigma : A \mapsto (A_{n+1})_{n \in \mathbb{Z}}$

Definition 5.1. — *A* **word** *is a map from a finite integer interval* $\{a, a+1, \ldots, b\}$ *to* $\mathcal{A}$*, up to an integer translation of the indices. The* **follower set** *of a word* $A_{-n} \ldots A_0$ *is:*

$$\mathrm{fol}(A_{-n} \ldots A_0) = \{B_0 B_1 \cdots \in \Sigma_+ : \exists B \in \Sigma \; B_{-n} \ldots B_0 = A_{-n} \ldots A_0\}.$$

A **minimal word** *of* Σ *(or* Σ_+*) is* $A_{-n} \ldots A_0$ *such that:*

$$\mathrm{fol}(A_{-n} \ldots A_0) \subsetneq \mathrm{fol}(A_{-n+1} \ldots A_0).$$

The **minimal form** *of* $A_{-n} \ldots A_0$ *is*

$$\min(A_{-n} \ldots A_0) := A_{-k} \ldots A_0$$

where $k \leq n$ *is maximum such that* $A_{-k} \ldots A_0$ *is minimal.*

Definition 5.2. — *The* **Markov diagram** $\mathcal{D}$ *of a subshift* Σ *is the oriented graph whose vertices are the minimal words of* Σ *and whose arrows are:*

$$A_{-n} \ldots A_0 \to B_{-m} \ldots B_0 \iff B_{-m} \ldots B_0 = \min(A_{-n} \ldots A_0 B_0).$$

The corresponding Markov shift is:

$$\hat{\Sigma} := \{\alpha \in \mathcal{D}^{\mathbb{Z}} : \forall p \in \mathbb{Z} \; \alpha_p \to \alpha_{p+1} \; \text{on } \mathcal{D}\},$$

with the left shift σ *may be called the* **Hofbauer shift***. There is a natural, continuous projection:*

$$\hat{\pi} : \alpha \in \hat{\Sigma} \longmapsto A \in \Sigma$$

with A_n *the last symbol of the word* α_n *for each* $n \in \mathbb{Z}$*.*

Remark 5.3. — *The vertices of Hofbauer's original diagram were simply the follower sets, as in the classical construction for sofic subshifts* [**65**]*. It can also be used*

beyond piecewise monotonic maps [**18**] *but it leads to a less tight relationship of the dynamics. In particular, periodic points may have several lifts*[(5)].

Exercise 5.4. — *Show that a subshift has a finite Markov diagram iff it is a subshift of finite type.* Note: *Hofbauer's Markov diagram is finite if and only if the subshift is sofic.*

Exercise 5.5. — *Compute the Markov diagram of the even shift, that is, $\Sigma \subset \{0,1\}^{\mathbb{Z}}$ obtained by excluding the words $01^{2n+1}0$, $n \geq 0$.*

5.2. Eventual Markov property and partial isomorphism. — We are going to see that the Hofbauer shift $\hat{\Sigma}$ is partially isomorphic with Σ, the (invertible) original subshift.

Definition 5.6. — *Let Σ be a subshift on a finite alphabet. $A \in \Sigma$ is* **eventually Markov** *(or just Markov) at time $p \in \mathbb{Z}$ if there exists $N = N(x,p)$ such that:*

$$\forall n \geq N \, \mathrm{fol}(A_{p-n} \dots A_p) = \mathrm{fol}(A_{p-N} \dots A_p)$$

The **eventually Markov part** $\Sigma_M \subset \Sigma$ *is the set of $A \in \Sigma$ which are eventually Markov at all times $p \in \mathbb{Z}$.*

Remark 5.7. — *Σ_M is a topological version of the so-called "variable length Markov chains".*

The following was shown by Hofbauer for piecewise monotone maps and the original Markov diagram (building on work of Y. Takahashi) and then generalized to arbitrary subshifts [**15, 18**]:

Theorem 11 (Hofbauer, Buzzi). — *The natural projection from the Hofbauer shift $\hat{\Sigma}$ to the subshift Σ defined by:*

$$\hat{\pi} : (\alpha_p)_{p \in \mathbb{Z}} \longmapsto (A_p)_{p \in \mathbb{Z}} \quad \text{with } A_p \in P \text{ the last symbol of } \alpha_p$$

is well-defined and is a Borel isomorphism from $\hat{\Sigma}$ to $\Sigma \setminus \Sigma_+$.

Exercise 5.8. — *Show that if Σ is the symbolic dynamics defined by an irrational rotation T with P a partition into two disjoint intervals, then Σ_M is empty.*

Lemma 5.9. — *Let $\alpha_0 \to \dots \alpha_n$ be a finite path on $\mathcal{D}$. Let $B_{-k} \dots B_0 = \alpha_0$ and $A = \hat{\pi}(\alpha) \in \Sigma$. Then,*

$$\alpha_n = \min(B_{-k} \dots B_0 A_1 A_2 \dots A_n).$$

Exercise 5.10. — *Prove the lemma.*

[(5)]This creates only a finite number of extra periodic orbits on the interval but can be a serious issue for generalizations

PROOF OF THEOREM 11: The following measurable map is a candidate to be a partial inverse to $\hat{\pi}$:

$$\iota : A \in \Sigma_M \longmapsto \alpha \in \hat{\Sigma} \quad \text{with } \alpha_p = \lim_{n\to\infty} \min(A_{p-n} \ldots A_p)$$

The above limit is to be understood in the most trivial sense: it is equal to $A_{p-N} \ldots A_p$ if this is the value of $\min(A_{p-n} \ldots A_p)$ for all large n (i.e., use the discrete topology on the countable set of finite words). $A \in \Sigma_M$ is equivalent to the existence of this "limit" for all $p \in \mathbb{Z}$.

Step 1: $\hat{\pi} : \hat{\Sigma} \to \Sigma_M$ *is well-defined*

Let $\alpha \in \hat{\Sigma}$. Let $A := \hat{\pi}(\alpha) \in \mathcal{A}^{\mathbb{Z}}$. We first show that for each $p \in \mathbb{Z}$ and $n \geq 0$, $A_p \ldots A_{p+n}$ is a word of $\Sigma_+(T, P)$. But the lemma applied to $\alpha_p \to \alpha_{p+1} \ldots$ implies that it is a suffix of such a word, which gives the property so $A \in \Sigma$.

Let us check that $A \in \Sigma_M$. For any $n \geq 0$, $\alpha_{-n} = D^n_{-k(n)} \ldots D^n_0$ with $D^n_0 = A_{-n}$. Let $m := k(0)$. By the Lemma applied to $\alpha_{-n} \to \cdots \to \alpha_0$, we have:

$$(5.11) \qquad \min(D^n_{-k(n)} \ldots D^n_{-1} A_{-n} \ldots A_0) = D^0_{-m} \ldots D^0_0$$

But this implies that

$$\forall n \geq m \ \min(A_{-n} \ldots A_0) = D^0_{-m} \ldots D^0_0$$

so that A is Markov at time 0. The same applies to any time, hence $A \in \Sigma_M$.

Step 2: $\iota(\Sigma_M) \subset \hat{\Sigma}$

Let $A \in \Sigma_M$ and $\alpha := \iota(A) \in \mathcal{D}^{\mathbb{Z}}$. Let $p \in \mathbb{Z}$. We have, for n large enough,

$$\alpha_p = \min(A_{p-n} \ldots A_p) \text{ and } \alpha_{p+1} = \min(A_{p+1-n} \ldots A_{p+1})$$

Hence $\alpha_p = A_{p-\ell} \ldots A_p = \min(A_{p-\ell-1} A_{p-\ell} \ldots A_p)$, so:

$$\sigma([A_{p-\ell-1}]) \supset [A_{p-\ell} \ldots A_p]$$

which implies:

$$\sigma([A_{p-\ell-1}]) \supset [A_{p-\ell} \ldots A_{p+1}]$$

giving that $\min(A_{p-\ell-1} \ldots A_{p+1}) = \min(A_{p-\ell} \ldots A_{p+1})$. By induction:

$$\alpha_{p+1} = \min(A_{p-n} \ldots A_{p+1}) = \min(A_{p-\ell} \ldots A_{p+1})$$

We have shown:

$$\alpha_p = A_{p-\ell} \ldots A_p \text{ and } \alpha_{p+1} = \min(A_{p-\ell} \ldots A_{p+1})$$

and this is the definition of $\alpha_p \to \alpha_{p+1}$. Hence $\iota(\Sigma_M) \subset \hat{\Sigma}$.

Step 3: $\hat{\pi} : \hat{\Sigma} \to \Sigma_M$ *is a bijection*

By the previous steps, $\iota \circ \hat{\pi}(\alpha)$ is well-defined for $\alpha \in \hat{\Sigma}$. Eq. (5.11) implies that

$$\alpha_0 = D^0_{-m} \ldots D^0_0 = A_{-m} \ldots A_0 = \min(A_{-n} \ldots A_0) \quad \forall n \geq m.$$

Hence, $\iota \circ \hat{\pi} : \hat{\Sigma} \to \hat{\Sigma}$ is the identity.

The previous steps also show that $\hat{\pi} \circ \iota$ is well-defined over Σ_M and it is obviously the identity there. This concludes the proof of the theorem. $\qquad \square$

Remark 5.12. — *G. Keller introduced another point of view in* [**54**] *by observing that the one-sided Markov shift defined by the Markov diagram contains a (non necessarily invariant) copy of Σ so that one can lift any probability measure of Σ there and try and make it invariant by pushing it forward. It leads to other type of lifting theorems. See Remark 6.5 below for a (quick) comparison between the two approaches.*

5.3. Markov diagrams for piecewise monotone maps. — Hofbauer discovered that the Markov diagram of a piecewise monotone map is very special. Recall how the symbolic dynamics is defined by the finite set of **kneading invariants** $K(T, P)$ (see Section 3.2 above).

Theorem 12 (Hofbauer). — *Let $T : [0, 1] \to [0, 1]$ be a piecewise monotone map with natural partition P. Let $\Sigma_+(T, P)$ be its symbolic dynamics.*

All the vertices of the Markov diagram are obtained from the kneading invariants:

$$\mathcal{D} = \{A_0 \dots A_n : 0 \leq n < N(A) \text{ and } A \in K(T, P)\}$$

where $N(A) := \inf\{n \geq 0 : \min(A_0 \dots A_n) \neq A_0 \dots A_n\}$ (usually infinite) and the arrows are exactly:

  (i) $A_0 \dots A_n \to \min(A_0 \dots A_{n+1})$;
  (ii) $A_0 \dots A_{S(A,i+1)-1} \to \min(A_{S(A,i)} \dots A_{S(A,i+1)})$ *if $S(A, i+1) < \infty$;*
  (iii) $A_0 \dots A_{S(A,i+1)} \to Q$ *for $Q \in P(A, i)$ if $S(A, i+1) < \infty$.*

where $P(A, i) \subset P$ and $S : K(T, P) \times \mathbb{N} \to \{1, 2, \dots, \infty\}$ satisfies for each $i \geq 0$, $S(A, i+1) = S(A, i) + S(B, j)$ or ∞ (in the finite case $(B, j) \in K(T, P) \times \mathbb{N}$ is determined by $B_0 \dots B_{S(B,j)} = A_{S(A,i)} \dots A_{S(A,i+1)}$).

Exercise 5.13. — *For $\beta > 1$, compute the Markov diagram of the β-**transformation***: $T_\beta : [0, 1] \to [0, 1]$ *defined by $T_\beta(x) = \beta x - [\beta x] \in [0, 1)$ ([·] is the integer part).* Hint: *Introduce $\{R_1 < R_2 < \dots\} := \{r \geq 1 : \iota(1^-)_n \neq (0, \beta^{-1})\}$ and show that, up to finitely many vertices, it can be represented as a "linear graph" $0 \to 1 \to 2 \to \dots$ with additional "backwards" arrows that you will specify.*

Show that the Markov diagram of T_β can be finite only if β is algebraic.

Proposition 5.14 (Hofbauer). — *Let $T : [0, 1] \to [0, 1]$ be unimodal, that is, T is continuous and there exists $c \in (0, 1)$ such that $T|[0, c]$ and $T|[c, 1]$ are continuous and strictly monotone, normalized by $T(0) = T(1) = 0$. Assume that $c < T(c) < 1$ (otherwise the situation is easy to analyze). Let $A = \iota(c^-)$ be the unique nontrivial kneading invariant.*

The Markov diagram of T has the following structure.

Its vertices are: $I_n := \min(A_0 \dots A_n)$, for $n \geq 0$, together with a transient vertex $(c, 1)$.

Its arrows are:

  − $(c, 1) \to I_0, I_1$;
  − $I_n \to I_{n+1}$ *for all $n \geq 0$;*
  − $I_{S(n)-1} \to I_{S(Q(n+1))}$ *if $Q(n+1) < \infty$*

where $S : \mathbb{N} \to \mathbb{N}$ and $Q : \mathbb{N} \to \mathbb{N}$ satisfy $S(n+1) = S(n) + S(Q(n+1))$ if $Q(n) < \infty$, $Q(n) < n$ unless it is infinite.

Remark 5.15. — *The simplicity already observed on the level of symbolic dynamics (which gives rise to the notion of* quasi finite type *described in section 9) is again visible in the Markov Diagrams of piecewise monotone maps. They are not finite (except in very special cases) but they are "almost so". This is made precise by the notion of* strongly positive recurrence *(see section 7).*

Remark 5.16. — *F. Hofbauer translated the relatively simple admissibility condition on the kneading invariants to one on the above Q-function. This results allowed the construction of quadratic map with finely tuned properties (see, e.g., [49]).*

6. Non-Markov sequences, Shadowing and Weak Rank One

We have showed that any invertible subshift Σ contains a Markov part Σ_M which is the continuous, Borel isomorphic image of the Markov shift defined by the Markov diagram of Σ. To use this, one needs to control the complement $\Sigma \setminus \Sigma_M$. For the symbolic dynamics of piecewise monotone maps (and many other kind of subshifts, see Section 9) we use the following notion.

6.1. Non-Markov Sequences. —

Definition 6.1. — *Let $\Sigma_+ \subset \mathcal{A}^{\mathbb{N}}$ be a subshift over a finite alphabet and Σ its natural extension. Let $S \subset \Sigma_+$. $A \in \Sigma$ is* **shadowed** *by S if there exist arbitrarily large integers n such that $A_{-n} \ldots A_0$ is the beginning of a sequence in S. A measure μ on Σ is shadowed if μ-a.e. sequence is.*

Theorem 13 (**after Hofbauer**). — *If (T, P) is a piecewise monotone map, then the non-Markov part of its invertible symbolic dynamics $\Sigma(T, P)$ carries only zero entropy measure and therefore the partial isomorphism of Theorem 11 is an entropy-conjugacy.*

More precisely, except for finitely many periodic sequences, a sequence of $\Sigma(T, P)$ is shadowed at all times by $K(T, P)$ iff it is non-Markov at all times.

Moreover, $\hat{\pi}$ is a period-preserving bijection between the periodic orbits of $\hat{\Sigma}$ and their image in $\Sigma_+(T, P)$ except for finitely many periodic orbits.

PROOF: We check that all invariant and ergodic probability measures μ with $\mu(\Sigma_M) = 0$ are shadowed by $K(T, P)$, the kneading set. This extends immediately to nonergodic measures, using the ergodic decomposition and yields the result as $K(T, P)$ is finite (so that $h_{\text{top}}(\sigma, K(T, P)) = 0$). We shall finally prove the converse.

Step 1: *Almost every A Markov at 0 belongs to Σ_M*

Let X_p is the set of $A \in \Sigma$ which are Markov at time p. We claim that $X_p \subset X_{p+1}$. Indeed, for $A \in \Sigma$, $[\ldots]_+$ denoting the symbolic cylinders in Σ_+ (i.e., the set of sequences in Σ_+ starting with a given word),

$$(6.2) \quad \sigma^{n+1}[A_{p-n-1} \ldots A_0]_+ = \sigma^n[A_{p-n} \ldots A_0]_+ \iff \sigma(A_{p-n-1}) \supset [A_{p-n} \ldots A_0]_+$$

as $\sigma|[A_{-n}\ldots A_0]$ is one-to-one. But

$$\sigma(A_{p-n-1}) \supset [A_{p-n}\ldots A_p]_+$$
$$\implies \sigma(A_{p+1-(n+1)-1}) \supset [A_{(p+1)-(n+1)}\ldots A_p]_+ \supset [A_{(p+1)-(n+1)}\ldots A_{p+1}]_+$$

proving the claim. On the other hand, $\sigma^{-1}(X_p) = X_{p+1}$ so X_p and X_{p+1} both have the same measure for any invariant probability measure.

Thus $X_p \subset X_{p+1}$ implies $X_p = X_{p+1}$ modulo a set negligible wrt all invariant probability measures so that:

$$X_0 = \bigcap_{p\in\mathbb{Z}} X_p = \Sigma_M$$

modulo a negligible set wrt any invariant probability measure as claimed.

Step 2: *Geometric consequence of non-Markovianness*

By eq. (6.2), A is non-Markov at time 0 implies on the interval that, for arbitrarily large integers n:
$$T\langle A_{-n-1}\rangle \not\supset \langle A_{-n}\ldots A_0\rangle$$
The right hand side is an interval so it is **connected** implying that
$$(\partial T\langle A_{-n-1}\rangle) \cap \langle A_{-n}\ldots A_0\rangle \neq \emptyset$$

This implies that $A_{-n-1}\ldots A_0$ is the beginning of the itinerary of one endpoint of $\langle A_{-n-1}\rangle$, i.e., a sequence in $K(T,P)$. Thus A non-Markov at time 0 implies that A is shadowed at time 0 by $K(T,P)$.

Step 3: *Shadowed sequences are non-Markov*

We prove the converse implication (up to finitely many periodic sequences).

Let $A \in \Sigma(T,P)$ be a sequence which is shadowed at time 0: there are infinitely many integers $n \to \infty$ such that (*) $A_{-n}\ldots A_0$ is the beginning of $K \in K(T,P)$ (we may assume K to be fixed as $K(T,P)$ is finite). Assume also that A is Markov at time 0: on the interval we have: $TA_{-n-1} \supset \langle A_{-n}\ldots A_0\rangle$ for all large n (note that this holds whether or not P is generating).

K is the itinerary of, say, the left endpoint of $(c,c') \in P$. Assume for simplicity that $T|(c,c')$ is increasing.

Take a large integer n with both properties. Let $(d,e) := \langle A_{-n}\ldots A_0\rangle$. By the the assumptions, $TA_{-n-1} = T(c,c') \supset (d,e)$ and $Tc \in [d,e]$, so that c is the left endpoint of $\langle A_{-n-1}\ldots A_0\rangle$.

Consider $n' > n$ also satisfying the first property. Let $p = n' - n$. T^p maps $\langle A_{-n'}\ldots A_0\rangle$ to $\langle A_{-n}\ldots A_0\rangle$ as A is Markov at time 0. Hence the endpoint c, which is the left endpoint of both previous intervals, is p-periodic. More precisely, for any $\epsilon > 0$, there exists $\delta > 0$ such that $0 < t < \delta \implies f^p(c+t) \in (c,c+\epsilon)$. This implies that $(A_k)_{-n'\leq k\leq 0}$ is p-periodic.

Taking A non-Markov for all times and $n' \to \infty$, we get that A is a periodic sequence from $K(T,P)$, finishing the proof of the theorem. $\square$

Remark 6.3. — *The non-Markov part of $\Sigma(T,P)$ is non-trivial for many piecewise monotone maps. For instance, for an infinitely renormalizable map, it carries the corresponding odometer. Even for β-transformation, it can carries non-periodic measures, for instance ones isomorphic to rotations (see [26]). Note that this does not says that one cannot find a Borel isomorphism between $\Sigma(T,P)$ and $\hat{\Sigma}$, just that it is not given by the natural map $\hat{\pi}$.*

Exercise 6.4. — *Find a piecewise monotone map (T,P) with an invariant probability measures μ on the* one-sided *Markov shift defined by $\mathcal{D}$ which is not sent by $\hat{\pi}$ to an isomorphic one on $\Sigma_+(T,P)$. Hint: Use for instance that, for piecewise invertible maps T with invariant probability measure μ, the measure-theoretic Jacobian $d\mu \circ T/d\mu$ is an invariant of measured conjugacy.*

Remark 6.5. — *G. Keller has developped another approach to Hofbauer's construction by trying to keep the smoothness of the interval map in the picture. In particular he showed that, for "nice" interval maps, like unimodal maps with negative Schwartzian derivative and no attracting periodic orbits, the measures that cannot be lifted to the Markov shift have been characterized by G. Keller and H. Bruin [54, 12] as those having a zero or negative Lyapunov exponent. Assumptions on Lyapunov exponents are more general (recall Ruelle's inequality) but are also seem much harder to generalize.*

6.2. Entropy Bound. —

Proposition 6.6. — *If μ is an invariant probability measure on Σ shadowed by $S \subset \Sigma_+$, then*

$$h(\sigma,\mu) \leq h_{\text{top}}(S,\sigma)$$

Recall that $h_{\text{top}}(S,\sigma) = \limsup_{n\to\infty} \frac{1}{n} \log \#\{A_0 \ldots A_n : A \in S\}$.

PROOF: Observe that the canonical partition of $P^{\mathbb{Z}}$ (which we'll also denote by P) is generating so that, according to Corollary 2.1, it is enough to find, for some $\lambda > 0$, for arbitrarily large $n \geq 1$, a small collection of P, n-cylinders with a union of measure $\geq \lambda$.

Recall that $S \subset \Sigma_+$ and let $S_n := \{A_0 \ldots A_{n-1} : A \in S\}$. Let $\epsilon > 0$. By definition of the topological entropy, there exists $N_0 < \infty$ such that for all $n \geq N_0$,

$$\#S_n \leq e^{n(h_{\text{top}}(\sigma,S)+\epsilon)}.$$

Increase N_0 if necessary so that the binominal coefficients satisfy:

$$\binom{2n/N_0}{n} \leq e^{\epsilon n}$$

According to the main assumption the following measurable function over Σ is finite a.e.:

$$n(A) := \min\{n \geq N_0 : A_{-n+1} \ldots A_0 \in S_n\}$$

Thus there exists $M_0 < \infty$ such that

$$\mu(\{x \in \Sigma : n(x) > M_0\}) < \epsilon/\log \#P$$

By the ergodic theorem, there exists $N_1 < \infty$ and X_1 such that $\mu(X_1) > \lambda$ and for all $n \geq N_1$, for all $x \in X_1$:

$$\#\{0 \leq k < n : n(\sigma^k x) \leq M_0\} < \frac{\epsilon}{\log \#P} n$$

For $n \geq \max\left((\log \#P/\epsilon)M_0 + N_1\right)$ we cover X_1 with a small number of P, n-cylinders.

For $x \in X_1$, we are going to define $c(x) = (a, b)$ where a, b are integer sequences satisfying:

$$0 < a_r < b_r < a_{r-1} < b_{r_1} < \cdots < a_1 < b_1 \leq a_0 = n$$

and:

  (i) for $1 \leq s \leq r$, $\ell_s := b_s - a_s + 1 \geq N_0$;
 (ii) $A_{a_s} \ldots A_{b_s} \in S_{\ell_s}$;
(iii) $\sum_{s=1}^r |b_s - a_s + 1| > \left(1 - 2\frac{\epsilon}{\log \#P}\right) n$

We claim that this will prove:

$$r(P, n, X_1, \lambda) \leq \binom{2n/N_0}{n}(\#P)^{n(\epsilon/\log \#P)+M_0} \exp n(h_{\text{top}}(S)+\epsilon) \leq \exp n(h_{\text{top}}(S)+4\epsilon)$$

and therefore the entropy bound as $\epsilon > 0$ is arbitrary. We conclude the proof of the proposition by proving this claim. Define $c(x) = (a, b)$ over X_1 by setting: $a_0 = n$ and then, inductively:

  − $b_{s+1} := \max\{b < a_s : n(\sigma^b x) \leq M_0\}$;
  − $a_{s+1} := b_{s+1} - n(\sigma^{b_{s+1}} x)$

and $r := \max\{s \geq 0 : a_s \geq 0\}$. Observe that c takes at most $\binom{2n/N_0}{n}$ distinct values.

$x|\bigcup_{s=1}^r [a_s, b_s]$ can be described by specifying for each $1 \leq s \leq r$, $x|[a_s, b_s] \in S_{\ell_s}$: there are at most $\exp n(h_{\text{top}}(\sigma, S) + \epsilon)$ choices.

Let $I_* := [0, a_r[\cup\bigcup_{s=1}^r]b_s, a_{s-1}[$. For $k \in I_*$, either $k \in]a_{r+1}, b_{r+1}[$ (which has length at most M_0) or $n(T^k x) > M_0$, therefore this subset of $[0, n[$ has cardinality at most $M_0 + (\epsilon/\log \#P)n$ by the choice of M_0. Thus there are at most $(\#P)^{\#I_*} \leq (\#P)^{M_0+(\epsilon/\log \#P)n}$ distinct choices for $x|I_*$.

Multiplying these numbers of choices we get the claim.   □

6.3. Weak Rank One. — The non-Markov part of the symbolic dynamics of a piecewise monotone map is shadowed by a finite set. As the entropy of a finite set is zero, Proposition 6.6 implies that the entropy of any invariant probability measure carried by this part is zero. Can one go further?

A motivation would be to find other type of arguments to control $\Sigma \setminus \Sigma_M$, especially in higher dimensions where the kneading set is replaced by a set with positive entropy.

Here, finiteness is much stronger than having zero entropy so it is natural to ask whether one can say more about these measures:

What are the measures shadowed by finitely many sequences?

The good news is that this property can be formulated as a natural generalization of the classical, ergodic theory notion of rank one [**58**]: an automorphism T of a probability space (X, P) has **rank one** if, for every finite measurable partition P, for every $\epsilon > 0$ and $N < \infty$, there exists a finite word ω with length at least N such that; for almost every $x \in X$, for all $n \geq n_0(x, \epsilon, N)$, one can write the n, P-itinerary of x as $G^0 \omega^1 G^1 \ldots G^{r-1} \omega^r G^r$ where:

- the sum of the lengths of the gaps G^s is at most ϵn;
- $|\omega^s| = |\omega|$ and $\sum_{s=1}^{r} \#\{0 \leq i < |\omega| : \omega_i^s \neq \omega_i\} \leq \epsilon n$.

Indeed, the propery of weak rank one is obtained by loosening the "approximate copies of a long word" to "approximate copies of long prefixes of a word":

Definition 6.7. — *An automorphism T of a probability space (X, P) has **weak rank one** [**20**] if, for every finite measurable partition P, for every $\epsilon > 0$ and $N < \infty$, there exists a finite word ω such that; for almost every $x \in X$, for all $n \geq n_0(x, \epsilon, N)$, one can write the n, P-itinerary of x as $G^0 \omega^1 G^1 \ldots G^{r-1} \omega^r G^r$ where:*

- *the sum of the lengths of the gaps G^s is at most ϵn;*
- *$N \leq |\omega^s| \leq |\omega|$ and $\sum_{s=1}^{r} \#\{0 \leq i < |\omega_s| : \omega_i^s \neq \omega_i\} \leq \epsilon n$.*

Proposition 6.8. — *An automorphism T of a probability space (X, μ) is weak rank one iff, there exists a finite, measurable, generating partition P with respect to which μ is shadowed by a single sequence [**20**].*

Weak rank one is weaker than the usual generalizations of rank one:

Proposition 6.9. — *All systems which have locally finite rank [**36**] are weak rank one and the inclusion is strict. However there exists zero entropy loosely Bernoulli systems which are not weak rank one. In the $\bar{d}$-metric (see [**87, 100**]), the generic system is not weak rank one (i.e., the weak rank one systems are contained in a countable union of closed subset with empty interiors).*

Exercise 6.10. — *Recall that system is of rank $r < \infty$ if it satisfies the definition of rank one with the modification that instead of a single word ω one has r, possibly distinct, words $\omega(1), \ldots, \omega(r)$. Show that such systems are weak rank one.*

The bad news is:

Question 3. — *Find a* bona fide *ergodic property holding for all systems with weak rank one, but not all some zero entropy systems.*

6.4. Shadowing Sequences. — A **shadowing sequence** is a sequence which shadows an invariant probability measure. Not every sequence is shadowing:

Exercise 6.11. — *Show that any word appearing in a shadowing sequence must appear infinitely many times.*

Question 4. — *What can one say about the repetition times of x (i.e., $n(\ell) \geq 1$ minimum such that $x_{n(\ell)} \ldots x_{n(\ell)+\ell-1} = x_0 \ldots x_{\ell-1}$)? And about the return times of μ? Is there a relationship?*

In fact, shadowing sequences are exceptional from a measure-theoretic point of view:

Exercise 6.12. — *Show that the set of sequences in $\{0,1\}^{\mathbb{N}}$ which shadow at least one measure has zero measure wrt the $(\frac{1}{2}, \frac{1}{2})$-Bernoulli measure.* Hint: *Use the ergodic theorem to restrict the measures that could be shadowed.*

Does the same occur in "concrete" situations:

Question 5. — *Is the set of $\beta > 1$ such that the itinerary of 1^{-} is a shadowing sequence a set of zero Lebesgue measure? (It is known [26] that this set is uncountable).*

One the other hand, it is common from the topological point of view:

Exercise 6.13. — *Show that the set of sequences that shadow some measure is of second Baire category (contains a dense countable intersection of open subsets).* Hint: *Consider, for each $n \geq 1$, $G_n := \{x \in \{0,1\}^{\mathbb{N}} : x = x_0 \ldots x_{n+m-1}(x_0 \ldots x_{n-1})^p \ldots$ for some $p \geq (n+m)^2\}$. See [20].*

Some shadowing sequences x satisfy: $\mu_x := \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \delta_{\sigma^k x}$ exists in the weak star topology.

Exercise 6.14. — *Show that when the limit μ_x exists, x can shadow only that measure. Build two shadowing sequences one such that the measure exists and one such that it does not.*

Say that two measures are *co-shadowed* if there exists a sequence which shadows both of them.

Exercise 6.15. — *Show that two measures that are shadowed by two possible distinct sequences are co-shadowed iff they have the same support. Generalize this to a countable collection of measures [20].*

Question 6. — *Does there exists a sequence which shadows uncountable many distinct measures? Does there exists a "universal shadowing sequence", i.e., such that any shadowable measure with full support is shadowed by it? Do these sequences form a set of second Baire category?*

7. SPR Markov shifts

The heart of the entropy theory of interval maps presented here is the representation of their dynamics by combinatorial systems. These are countable state Markov shifts but of a special kind, called *Strongly (or Stably) Positive Recurrent* Markov shifts. They are the closest to the finite state Markov shifts: most of the classical results generalize to them as we are going to explain.

For countable Markov shift we refer to [59, 96, 97, 91] and especially to the treatise [42] which considers the SPR ones in more details.

Recall the following definitions. A (countable state) **Markov shift** is $\Sigma = \Sigma(G)$ where G is a countable oriented graph and $\Sigma := \{x \in G^{\mathbb{Z}} : \forall p \in \mathbb{Z} \ x_p \to x_{p+1} \text{ on } G\}$

endowed with the left-shift σ. Note that $\Sigma(G)$ is the vertex shift of G - see [**65**] for the edge shift). One can define an obvious one-sided Markov shift $\Sigma_+(G)$ where the left-shift is again denoted by σ.

   G or Σ are called **irreducible** if there is a path from any vertex to any other vertex in G (i.e., G is *strongly connected*) (or, equivalently, if Σ is topologically transitive).

   The **spectral decomposition** of a Markov shift is the at most countable collection of irreducible Markov shifts defined by the maximum subgraphs of its defining graph which are strongly connected components. Any ergodic invariant probability measure is carried on a single piece of the spectral decomposition.

   The **period** of an irreducible G is the highest common factor of all the lengths of its loops.

   The **outdegree** of a vertex $v \in G$ is $\#\{w \in G : v \to w\}$. The outdegree of G is the supremum of the outdegrees of all vertices.

7.1. Definition. — Let Σ be an irreducible Markov shift defined by a countable oriented graph G. The basic classification of such objects from our point of view is due to Vere-Jones [**104, 105, 106**], which remarked that one could extend the probabilistic classification of stochastic matrices (related to the underlying Markov chain) to a large class of positive matrices.

Theorem 14 **(Vere-Jones)**. — *Fix an arbitrary vertex v of G. Let f_n be the number of first returns after time n to v. Let ℓ_n be the number of returns after time n to v. Let R be the radius of convergence of $\ell(z) := \sum_{n \geq 1} \ell_n z^n$:*

$$R = \exp -h(\Sigma)$$

Assume that R is finite. Then $\sum_{n \geq 1} f_n R^{-n} \leq 1$.

(1) $\sum_{n \geq 1} f_n R^{-n} < 1$: G is transient *and A admits no left or right positive eigenvectors;*

(2) $\sum_{n \geq 1} f_n R^{-n} = 1$ and $\sum_{n \geq 1} n f_n R^{-n} = \infty$: G is null recurrent *and A has both a left and a right eigenvectors, say ℓ and r, with $\sum_i \ell_i \cdot v_i = \infty$;*

(3) $\sum_{n \geq 1} f_n R^{-n} = 1$ and $\sum_{n \geq 1} n f_n R^{-n} < \infty$: G is positive recurrent *and A has both a left and a right eigenvectors, say ℓ and r, with $\sum_i \ell_i \cdot v_i < \infty$;*

Moreover, one can distinguish among the last type, those graphs which are SPR *defined by the condition:*

$$\limsup_{n \to \infty} |f_n|^{1/n} < R.$$

In the above cases, the iterates of A have the following behavior as $n \to \infty$ for any (and then all) i, j:

(1) $\sum_{n \geq 0} A_{ij}^n R^{-n} < \infty$;

(2) $\sum_{n \geq 0} A_{ij}^n R^{-n} = \infty$ and $A_{ij}^n R^{-n} \to 0$;

(3) $A_{ij}^n R^{-n} \to d \cdot \ell_i r_j \neq 0$ where d is the period of G and r and ℓ are positive eigenvectors: $\sum_i \ell_i A_{ij} = R\ell_j$ and $\sum_j A_{ij} r_j = R r_i$ normalized by $\sum_k \ell_k \cdot f_k = 1$.

In the SPR case, we additionally have:

$$\limsup_{n \to_i nfty} \frac{1}{n} \log \left(A_{ij}^n R^{-n} - d \cdot \ell_i r_j \right) < 0$$

B.M. Gurevič [**40**] discovered the link with dynamics and that there was in fact a probability measure hidden in these positive matrices.

Theorem 15 (Gurevič). — *In the Vere-Jones theorem, $R = \exp -h(G)$, where $h(G)$ is the supremum of the measured entropies of Σ. This supremum is achieved by at most one measure (called the* maximum measure*).*
  G is positive recurrent if and only if there exists a maximum measure. In this case, this measure μ_M is Markov and satisfies:

$$\mu_M([v_1 \dots v_n]) = R^{-n} \ell_{v_1} r_{v_n} \ or \ 0 \ if \ [v_1 \dots v_n] = \emptyset.$$

Exercise 7.1. — *Show that the above data defines a probability measure μ_M which is invariant.*

Proposition 7.2. — *Such measures as μ_M are finite extension of a Bernoulli measure [**50**]. The period of the extension is called the period of the measure.*

Exercise 7.3. — *Show that a positive recurrent graph is SPR iff the maximum measure is exponentially filling: for any non-empty open $U \subset \Sigma$:*

$$\limsup_{n \to \infty} \frac{1}{n} \log \mu \left(\bigcup_{0 \le k < n} \sigma^{-k} U \right) < 0$$

The SPR graphs have other characterizations:

Proposition 7.4 (Gurevič). — *Let G be irreducible with finite entropy. G is SPR iff any subgraph G' obtained by removing one arrow satisfies: $h(G') < h(G)$.*

Exercise 7.5. — *Show that if G is irreducible with finite entropy, then G is SPR iff any subgraph G' obtained by removing one arrow is still positive recurrent.*

From our point of view the following characterization is fundamental, showing that graphs that are "simple at infinity" are good:

Theorem 16 (Gurevič-Zargaryan). — *Let G be an irreducible Markov shift with finite entropy. Define its* **entropy at infinity** *by:*

$$h_\infty(G) := \inf_{F \subset\subset G} \sup_{\epsilon > 0} \{h(\sigma, \mu) : \mu \in \mathrm{Prob}(\Sigma) \ s.t. \ \mu(\{x \in \Sigma : x_0 \in F\}) < \epsilon\}$$

where F ranges over the finite subsets of vertices of G. G is SPR iff $h_\infty(G) < h(G)$.

7.2. Artin-Masur Zeta Function. —

Definition 7.6. — *If F is a subset of the vertices of the graph G, the Artin-Masur* **zeta function at** F *is the formal power series:*

$$\zeta_F^G(z) := \exp \sum_{n \geq 1} \frac{z^n}{n} \#\{x \in \Sigma(G) : \sigma^n x = x \ and \ \{x_0, x_1, \dots\} \cap F \neq \emptyset\}$$

(provided each cardinality above is finite).

For finite $F \subset G$, $\zeta_F^G(z)$ *is a* **semi-local** *zeta function. For F reduced to a single vertex a, $\zeta_a^G(z) := \zeta_{\{a\}}^G(z)$ is called the* **local** *zeta function.*

The full zeta function $\zeta_G^G(z)$ may fail to be defined or to have the expected properties (see [**42**, Thm. 9.4 and following]).

Theorem 17 (Gurevich-Savchenko, Buzzi). — *Let G be an irreducible countable oriented graph with $h(G) < \infty$. Let d be its period, i.e., the largest common divisor of the length of all loops in G.*

(1) *Each semi-local zeta function of G defines a holomorphic function over $|z| < \exp{-h(G)}$.*

(2) *These functions have a meromorphic extension to $|z| < \exp{-h_\infty(G)}$.[**25**]*

(3) *G is SPR iff any of its local zeta functions ζ_a^G has a meromorphic extension to $|z| < R_a$ for $R_a > e^{-h(G)}$ without zeroes and whose only singularities are simple poles at $z = e^{-h(G)}e^{2i\pi k/d}$, for $k = 0, \dots, d-1$.[**42**]*

7.3. Classification. —

The following result shows that such systems are in some sense very simple (or completely chaotic depending on your point of view).

Theorem 18 (Boyle-Buzzi-Gomez). — *To each irreducible Markov shift Σ associate its entropy $h(G)$ and the list of the periods of its maximum measures (repeated according to multiplicities).*

This data is a complete invariant for entropy-conjugacy among SPR Markov shifts.

Remark 7.7. — *In fact a stronger notion of conjugacy is established in [**9**]: one finds in each graph a vertex and a Borel conjugacy between the sets of all sequences that visit infinitely often in the past and in the future these vertices, except for those that remain in two finite subgraphs and a (necessarily countable) set of periodic sequences.*

Remark 7.8. — *For subshifts of finite type this is a classical result of Adler and Marcus [**1**].*

The idea of the proof of the above is quite simple. Let us sketch it. Consider two SPR graphs. One can assume these graphs to be loop graphs with many first return loops of each large length after discarding some subshifts of finite type (which have strictly smaller entropy by the SPR property). Then one modifies these graphs by removing loops of increasing length from one or the other until they both have the same number of loops of all lengths. These removals are operated on first return loops. A first return loop ℓ is removed and first return loops $L\ell^n$, $n \geq 1$, for any first return loop $L \neq \ell$ are added back so that no other sequence has been removed

from the shift. The Markov shift obtained in this limit is Borel isomorphic, except
for the set of removed periodic sequences, to the initial ones *if*, for instance, one has
been able to remove only first return loops that were already present in these initial
graphs. The SPR assumption ensures that the number of loops to be removed is so
small that the previous condition can easily be fulfilled (see Theorem 17).

8. Application to Piecewise Monotone Maps

Let (T, P) be a piecewise monotone map with its symbolic dynamics $\Sigma_+(T, P)$ and
the natural extension of it, $\Sigma(T, P)$.

By Theorem 13, $\Sigma_{(}T, P)$, after discarding a subset carrying only zero entropy
measures and finitely many periodic sequences, is Borel isomorphic to the Markov
shift $\hat{\Sigma}$ defined by its Markov diagram $\mathcal{D}$.

We claim that $h_\infty(\mathcal{D}) = 0$ so that, for every $H > 0$, the spectral decomposition of
$\hat{\Sigma}$ contains only finitely many irreducible Markov shifts with entropy $\geq H$ and that
each of these is SPR.

This follows from the description of Theorem 12. $\mathcal{D}$ has finite outdegree (bounded
by $\#P$) and, after removing the finite subset

$$\mathcal{D}_N := \{\min(K_0 \ldots K_n) : K \in K(T, P) \text{ and } n \leq N\}$$

the remaining graph $\mathcal{D} \setminus \mathcal{D}_N$ has outdegree at most 2 and the following property: if
$\alpha_0 \to \ldots \alpha_n$ is a path on this graph with α_0 and α_n with outdegree > 1 in $\mathcal{D} \setminus \mathcal{D}_N$,
then $n \geq N$.

Exercise 8.1. — *Check that the above property holds and that it implies that*
$h_\infty(\mathcal{D}) = 0$.

8.1. Measures of Maximum or Large Entropy. — We prove in this section
the following:

Theorem 19 (**Hofbauer, Boyle-Buzzi-Gomez**). — *A piecewise monotone map
with nonzero entropy has a finite number of ergodic measure, each of which is up to
a period, a Bernoulli probability measure* [**43, 50**].

*More precisely, the topological entropy together with the list with repetitions of the
periods of the maximum measure is a complete invariant for the entropy-conjugacy of
the natural extensions of these maps* [**9**].

Remark 8.2. — *This is false if one removes the words "natural extensions of".*

Question 7. — *Is a piecewise monotone map with nonzero topological entropy al-
ways entropy-conjugate to* some *countable state Markov shift?*

The first part of the theorem follows from Gurevič result that a Markov shift has
exactly one maximum measure (ergodic, invariant probability measure with maximum
entropy) on each piece of the spectral decomposition which is positive recurrent.
Moreover these measures are Markov and this implies by [**50**] that they are Bernoulli
up to a period.

Exercise 8.3. — *Show that there is a bijection which is entropy and ergodicity preserving between a dynamical system and its natural extension.*

Hence the above implies that piecewise monotone maps with nonzero topological entropy have a finite number of maximum measures.

By exploiting the very special structure of the Markov diagram of a piecewise monotone map one can obtain quantitative bounds.

Theorem 20 (Hofbauer,Buzzi). — *A unimodal map with positive topological entropy has exactly one maximum measure* [**44**].

A piecewise monotone map $T : [0,1] \to [0,1]$ with nonzero entropy and N pieces under the normalization $T(\{0,1\}) \subset \{0,1\}$ has at most $4(N-1)$ maximum measures (only $3(N-1)$ if it is continuous) [**14**].

However the following very natural conjecture remains open:

Question 8. — *A piecewise monotone map with nonzero entropy and N pieces has at most $N-1$ maximum measures.*

Remark 8.4. — *Another approach to this problem is to use the following results: (1) for a piecewise monotone map with slope constant in absolute value, the maximum measures are exactly the ergodic absolutely continuous invariant probability measures (this is not totally obvious because of the discontinuity* [**33**]*); (2) any piecewise monotone map with nonzero entropy admits as a topological factor a piecewise monotone map of the previous type* [**69**]*; (3) the number of distinct ergodic absolutely continuous invariant probability measure of, say, an interval map with piecewise constant slope and N pieces is bounded by $N-1$. The problem is to understand what measures are collapsed under the factor map and to relate the number of maximum measures to the successive renormalizations of the map.*

8.2. Periodic points. — The above estimate $h_\infty(\mathcal{D}) = 0$ yields, using Theorem 17:

Theorem 21 (Milnor-Thurston, Hofbauer). — *Let (T,P) be a piecewise monotone map with non zero entropy. The Artin-Masur zeta function:*

$$\zeta_T(z) := \exp \sum_{n \geq 1} \frac{z^n}{n} \#\{A \in \Sigma_+(T,P) : \sigma^n A = A\}$$

is holomorphic on $|z| < e^{-h_{\mathrm{top}}(T)}$. It has a meromorphic extension to $|z| < 1$ [**69**] *(see* [**89**] *for a later, simpler proof).*

The periodic orbits are equidistributed according to the maximum measures:

$$\frac{1}{\#\{A \in \Sigma_+(T,P) : T^n A = A\}} \sum_{A \in \Sigma_+(T,P) \ s.t. \ A = T^n A} \delta_A \xrightarrow[n \in p\mathbb{N}]{n \to \infty} \frac{1}{\sum_{i=1}^r p_i} \sum_{i=1}^r p_i \mu_i$$

where $p := \mathrm{lcm}(p_1, \ldots, p_r)$, $\mu_1, \ldots, \mu_r$ are the maximum measures on $\Sigma_+(T,P)$ and p_i is the period of μ_i.

9. Further Results

9.1. Generalization to Other Dynamics. — The results presented above are well-known in the classical hyperbolic setting, e.g., for subshifts of finite type [65]. They in turn can be vastly generalized to various non-uniformly expanding systems.

To begin with, one can define a class of subshifts over finite alphabets, the **subshifts of quasi-finite type** [23], which includes both the subshifts of finite type and the symbolic dynamics of piecewise monotone maps and satisfy most of the properties explained in this paper:

Definition 9.1. — *A subshift Σ over a finite alphabet is of* **quasi-finite type** *if the minimum words of length n, $\mathcal{M}_n$, satisfy:*

$$\limsup_{n\to\infty} \frac{1}{n} \log \#\mathcal{M}_n < h_{\mathrm{top}}(\sigma|\Sigma).$$

This class also contains new dynamics, e.g., multidimensional β-transformations, i.e., self-maps of $[0,1]^d$ defined by $x \mapsto A.x \mod \mathbb{Z}^d$ where A is an affine transformation such that $\|A^N.v\| > \|v\|$ for some N and all $v \in \mathbb{R}^d \setminus \{0\}$. The entropy theory of arbitrary piecewise affine maps in higher dimension is mostly unknown [62] beyond similar maps and piecewise affine surface homeomorphisms [24].

One can further extend the approach explained here to **entropy-expanding** maps, which are multidimensional smooth maps with critical points such that the entropy of smooth submanifolds is smaller than their topological entropy. This class includes arbitrary C^∞ interval maps with nonzero entropy [15, 29] and maps like:

$$(x,y) \mapsto (1.9 - x^2 + \epsilon y, 1.8 - y^2 + \epsilon x)$$

for small $|\epsilon|$. They can be treated either geometrically [19, 22] or, under a "general position assumption", by symbolic techniques with the introduction of **puzzles of quasi-finite type** [25] which generalize to combinatorial puzzles (as defined in complex dynamics) the notions above. One obtains in this way, e.g., that they possess finitely many maximum measures with Bernoulli extensions (up to a period).

9.2. Thermodynamical Formalism. — The **thermodynamical formalism** introduced in the uniformly hyperbolic setting by Sinai and Ruelle involves the introduction of a weight in the countings defining the entropy. This generalization of entropy is called "pressure" (it is in fact the free energy per site of a one-dimensional crystal formally associated to the dynamical system) [88, 107, 55]. The measures maximizing the pressure are called the equilibrium measures.

This generalization is very important for several reasons. For instance, if f is a smooth uniformly expanding map of the circle, the equilibrium measure for the weight $|f'|$ is the absolutely continuous invariant measure. Most of the entropy theory can generalized to this setting either when in one-dimension ([4, 89] and the references in the book [3]) or when there is enough uniformity in the expansion (see, e.g., [21, 28, 27]) so that the regularity properties of the weight (bounded variation or Holder continuity) also hold in the symbolic representation.

9.3. Further Questions. — We have not been able to discuss *stability issues*, that is, the (dis)continuity of the map that associates to a piecewise monotone map its set of maximum measures (see [**85**]).

We can only refer the reader to the already well-developed *dimension theory* [**83, 48**] which in particular relates Hausdorff dimension, entropy and Lyapunov exponents or the developping *symbolic extension theory* [**6, 7, 34**] which has already obtained deep refinements of the local entropy estimates presented in section 3.

References

[1] R. Adler, B. Marcus, *Topological entropy and equivalence of dynamical systems,* Mem. Amer. Math. Soc. 20 (1979), no. 219.

[2] L. Alsedà, J. Llibre, M. Misiurewicz, *Combinatorial dynamics and entropy in dimension one. Second edition,* Advanced Series in Nonlinear Dynamics, 5. World Scientific Publishing Co., Inc., River Edge, NJ, 2000.

[3] V. Baladi, *Positive transfer operators and decay of correlations,* Advanced Series in Nonlinear Dynamics, 16. World Scientific Publishing Co., Inc., River Edge, NJ, 2000.

[4] V. Baladi, G. Keller, *Zeta functions and transfer operator for piecewise monotone maps,* Commun. Math. Phys. (1990).

[5] A. Blokh, *The "spectral" decomposition for one-dimensional maps,* Dynamics reported, 1–59, Dynam. Report. Expositions Dynam. Systems (N.S.), 4, Springer, Berlin, 1995.

[6] M. Boyle, D. Fiebig, U. Fiebig,*Residual entropy, conditional entropy and subshift covers,* Forum Math. 14 (2002), no. 5, 713–757.

[7] M. Boyle, T. Downarowicz, *The entropy theory of symbolic extensions,* Invent. Math. 156 (2004), no. 1, 119–161.

[8] R. Bowen, *Entropy for group endomorphisms and homogeneous spaces,* Trans. Amer. Math. Soc. 153 1971 401–414.

[9] M. Boyle, J. Buzzi, R. Gomez, *Almost isomorphism for countable state Markov shifts,* Journal für die reine und angewandte Mathematik, 592 (2006).

[10] K. M. Brucks, H. Bruin, *Topics from One-Dimensional Dynamics,* Series: London Mathematical Society Student Texts (No. 62), Cambridge University Press.

[11] H. Bruin, *Non-monotonicity of entropy of interval maps,* Phys. Lett. A 202 (1995) 359-362.

[12] H. Bruin, G. Keller, *Equilibrium states for S-unimodal maps,* Ergod. th. dynam. syst. **18** (1998), 765-789.

[13] D. Burguet, *A proof of the Yomdin-Gromov Algebraic Lemma,* Preprint CMLS, 2005. http://arxiv.org/abs/math.DS/0509277.

[14] J. Buzzi, *Number of equilibrium states of piecewise monotone maps of the interval,* Proc. Amer. Math. Soc. 123 (1995), no. 9, 2901–2907. Erratum: Proc. Amer. Math. Soc. 125 (1997), no. 10, 3131.

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

[16] J. Buzzi, *Intrinsic ergodicity of affine maps in $[0,1]^d$,* Monatsh. Math. 124 (1997), no. 2, 97–118.

[17] J. Buzzi, *Specification on the interval,* (French summary) Trans. Amer. Math. Soc. 349 (1997), no. 7, 2737–2754.

[18] J. Buzzi, *Markov extensions for multi-dimensional dynamical systems,* Israel J. Math. 112 (1999), 357–380.

[19] J. Buzzi, *Ergodicité intrinsèque de produits fibrés d'applications chaotiques unidimensionelles,* Bull. Soc. Math. France 126 (1998), no. 1, 51–77.

[20] J. Buzzi, *Sur une généralisation de la notion de rang un définie par le pistage,* preprint de l'Institut de Mathématiques de Luminy (1997), `http://www.jeromebuzzi.com`.

[21] J. Buzzi, *Thermodynamical formalism for piecewise invertible maps: absolutely continuous invariant measures as equilibrium states,* in: Smooth ergodic theory and its applications (Seattle, WA, 1999), 749–783, Proc. Sympos. Pure Math., 69, Amer. Math. Soc., Providence, RI, 2001.

[22] J. Buzzi, *On entropy-expanding maps*, 2000, `http://www.jeromebuzzi.com`.

[23] J. Buzzi, *Subshifts of quasi-finite type,* Invent. Math. 159 (2005), no. 2, 369–406.

[24] J. Buzzi, *Piecewise Affine Surface Homemorphisms,* preprint 2006, `http://www.jeromebuzzi.com`.

[25] J. Buzzi, *Puzzles of quasi-finite type, zeta functions and symbolic dynamics for multidimensional maps* preprint 2006, `http://arxiv.org/abs/math.DS/0610911`.

[26] J. Buzzi, K. Johnson, *Beta-transformations with a non-Markov part carrying a given rotation,* preprint 2001, `http://math.polytechnique.fr/cmat/buzzi`.

[27] J. Buzzi, F. Paccaut, B. Schmidt, *Conformal measures for multidimensional piecewise invertible maps,* Ergodic Theory Dynam. Systems 21 (2001), no. 4, 1035–1049.

[28] J. Buzzi, G. Keller, *Zeta functions and transfer operators for multidimensional piecewise affine and expanding maps,* Ergodic Theory Dynam. Systems 21 (2001), no. 3, 689–716.

[29] J. Buzzi, S. Ruette, *Large entropy implies existence of a maximal entropy measure for interval maps,* Discrete Contin. Dyn. Syst. 14 (2006), no. 4, 673–688.

[30] W.-C. Cheng, S. Newhouse, *Pre-image entropy,* Ergodic Theory Dynam. Systems 25 (2005), no. 4, 1091–1113.

[31] T. Cover, J. Thomas, *Elements of information theory,* A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1991.

[32] M. Denker, *Measures with maximal entropy,* Théorie ergodique (Actes Journées Ergodiques, Rennes, 1973/1974), pp. 70–112. Lecture Notes in Math., Vol. 532, Springer, Berlin, 1976.

[33] M. Denker, G. Keller, M. Urbański, *On the uniqueness of equilibrium states for piecewise monotone mappings,* Studia Math. 97 (1990), no. 1, 27–36.

[34] T. Downarowicz, S. Newhouse, *Symbolic extensions and smooth dynamical systems,* Invent. Math. 160 (2005), no. 3, 453–499.

[35] T. Downarowicz, J. Serafin, *Possible entropy functions,* Israel J. Math. 135 (2003), 221–250.

[36] S. Ferenczi, *Systems of finite rank,* Colloq. Math. 73 (1997), no. 1, 35–65.

[37] I. Cornfeld, S. Fomin, Ya. Sinaǐ, *Ergodic theory*, Grundlehren der Mathematischen Wissenschaften, 245. Springer-Verlag, New York, 1982.

[38] R. Galeeva, M. Martens, Ch. Tresser,*Inducing, slopes, and conjugacy classes,* Israel J. Math. 99 (1997), 123–147.

[39] M. Gromov, *Entropy, homology and semialgebraic geometry,* Séminaire Bourbaki, Vol. 1985/86. Astérisque No. 145-146 (1987), 5, 225–240.

[40] B. Gurevič, *Shift entropy and Markov measures in the space of paths of a countable graph,* Dokl. Akad. Nauk SSSR 192 1970 963–965.

[41] B. Gurevič, *Stably recurrent nonnegative matrices,* (Russian) Uspekhi Mat. Nauk 51 (1996), no. 3(309), 195–196; translation in Russian Math. Surveys 51 (1996), no. 3, 551–552.

[42] B. Gurevich, S. Savchenko, *Thermodynamic formalism for symbolic Markov chains with a countable number of states,* Uspekhi Mat. Nauk 53 (1998), no. 2(320), 3–106; translation in: Russian Math. Surveys 53 (1998), no. 2, 245–344

[43] F. Hofbauer, *On intrinsic ergodicity of piecewise monotone transformations with positive entropy,* Israel J. Math. 34 (1979), no. 3, 213–237; II. Israel J. Math. 38 (1981), no. 1-2, 107–115.

[44] F. Hofbauer, *Maximal measures for piecewise monotoneally increasing transformations on $[0,1]$,* Ergodic theory (Proc. Conf., Math. Forschungsinst., Oberwolfach, 1978), pp. 66–77, Lecture Notes in Math., 729, Springer, Berlin, 1979.

[45] F. Hofbauer, *Periodic points for piecewise monotone transformations,* Ergodic Theory Dynam. Systems 5 (1985), no. 2, 237–256.

[46] F. Hofbauer, *Piecewise invertible transformations,* Probab. Th. Rel. Fields (1986).

[47] F. Hofbauer, *Generic properties of invariant measures for continuous piecewise monotonic transformations,* Monatsh. Math. 106 (1988), no. 4, 301–312.

[48] F. Hofbauer, *Hausdorff dimension and pressure for piecewise monotonic maps of the interval,* J. London Math. Soc. (2) 47 (1993), no. 1, 142–156.

[49] F. Hofbauer, G. Keller, *Quadratic maps without asymptotic measure,* Comm. Math. Phys. 127 (1990), no. 2, 319–337.

[50] Sh. Ito, H. Murata, H. Totoki, *Remarks on the isomorphism theorem for weak Bernoulli transformations in the general case,* Publ. Res. Inst. Math. Sci. 7, 541–580 (1971/1972).

[51] A. Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms,* Inst. Hautes Etudes Sci. Publ. Math. No. 51 (1980), 137–173.

[52] 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.

[53] A. Katok, J.-M. Strelcyn, F. Ledrappier, F. Przytycki, *Invariant manifolds, entropy and billiards; smooth maps with singularities,* Lecture Notes in Mathematics, 1222. Springer-Verlag, Berlin, 1986.

[54] G. Keller, *Lifting measures to Markov extensions,* Monat. Math. **108** (1989), 183–200.

[55] G. Keller, *Equilibrium States in Ergodic Theory,* London Mathematical Society Student Texts 42 Cambridge University Press 1998.

[56] G. Keller, A. Greven and G. Warnecke (eds), *Entropy,* Princeton Series in Applied Mathematics Princeton University Press (2003).

[57] G. Keller, C. Liverani, *Stability of the spectrum for transfer operators,* Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 28 (1999), no. 1, 141–152.

[58] J. King, *Joining-rank and the structure of finite rank mixing transformations,* J. Analyse Math. 51 (1988), 182–227.

[59] B. Kitchens, *Symbolic dynamics. One-sided, two-sided and countable state Markov shifts,* Universitext. Springer-Verlag, Berlin, 1998.

[60] O. Kozlovski, *An integral formula for topological entropy of C^∞ maps,* Ergodic Theory Dynam. Systems **18** (1998), no. 2, 405–424.

[61] U. Krengel, *Ergodic theorems. With a supplement by Antoine Brunel,* de Gruyter Studies in Mathematics, 6. Walter de Gruyter & Co., Berlin, 1985.

[62] B. Kruglikov, M. Rypdal, *Entropy via multiplicity*, http://arxiv.org/abs/math.DS/0505019.

[63] G. Levin, *On explicit connections between dynamical and parameter spaces,* J. Anal. Math. 91 (2003), 297–327.

[64] F. Ledrappier, L.-S. Young, *The metric entropy of diffeomorphisms;* Ann. of Math. (2) 122 (1985), 509–539 (I) and 540–574 (II).

[65] D. Lind, B. Marcus, *An introduction to symbolic dynamics and coding,* Cambridge University Press, Cambridge, 1995.

[66] R. Mañé, *Ergodic theory and differentiable dynamics,* Ergebnisse der Mathematik und ihrer Grenzgebiete (3), 8. Springer-Verlag, Berlin, 1987.

[67] W. de Melo, S. van Strien, *One-dimensional dynamics,* Springer 1993.

[68] J. Milnor, Ch. Tresser, *On entropy and monotonicity for real cubic maps,* With an appendix by Adrien Douady and Pierrette Sentenac. Comm. Math. Phys. 209 (2000), no. 1, 123–178.

[69] J. Milnor, W. Thurston, *On iterated maps of the interval,* in: Dynamical Systems, Lecture Notes in Mathematics **1342**, Springer, 465–564.

[70] M. Misiurewicz, *Diffeomorphism without any measure with maximal entropy,* Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 21 (1973), 903–910.

[71] M. Misiurewicz, *Topological conditional entropy,* Studia Math. 55 (1976), no. 2, 175–200.

[72] M. Misiurewicz, *A short proof of the variational principle for a Z_+^N action on a compact space* in: International Conference on Dynamical Systems in Mathematical Physics (Rennes, 1975), pp. 147–157. Asterisque, No. 40, Soc. Math. France, Paris, 1976.

[73] M. Misiurewicz, *Horseshoes for continuous mappings of an interval,* Dynamical systems (Bressanone, 1978), pp. 125–135, Liguori, Naples, 1980.

[74] M. Misiurewicz, F. Przytycki, *Topological entropy and degree of smooth mappings,* Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 25 (1977), no. 6, 573–574.

[75] M. Misiurewicz, W. Szlenk, *Entropy of piecewise monotone mappings,* Studia Math. 67 (1980), no. 1, 45–63.

[76] M. Misiurewicz, E. Visinescu, *Kneading sequences of skew tent maps,* (Ann. Inst. H. Poincaré Probab. Statist. 27 (1991), no. 1, 125–140.

[77] S. Newhouse, *Entropy and volume,* Ergodic Theory Dynam. Systems 8* (1988), Charles Conley Memorial Issue, 283–299.

[78] S. Newhouse, *Continuity properties of entropy,* Ann. of Math. (2) 129 (1989), no. 2, 215–235. Erratum: 131 (1990), no. 2, 409–410.

[79] D. Ornstein, *Bernoulli shifts with the same entropy are isomorphic,* Advances in Math. 4 1970 337–352 (1970).

[80] D. Ornstein, B. Weiss, *On the Bernoulli nature of systems with some hyperbolic structure,* Ergodic Theory Dynam. Systems 18 (1998), no. 2, 441–456.

[81] W. Parry, *Intrinsic Markov chains,* Trans. Amer. Math. Soc. 112 (1964), 55–66.

[82] W. Parry, *Entropy and generators in ergodic theory*, W. A. Benjamin, Inc., New York-Amsterdam 1969.

[83] Ya. Pesin, *Dimension theory in dynamical systems,* Contemporary views and applications. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1997.

[84] K. Petersen, *Ergodic theory,* Cambridge Studies in Advanced Mathematics, 2. Cambridge University Press, Cambridge, 1983.

[85] P. Raith, *Stability of the maximal measure for piecewise monotonic interval maps,* Ergodic Theory Dynam. Systems 17 (1997), no. 6, 1419–1436.

[86] W. Rudin, *Real and complex analysis,* Third edition. McGraw-Hill Book Co., New York, 1987.

[87] D. Rudolph, *Fundamentals of measurable dynamics*, Clarendon Press, 1990.

[88] D. Ruelle, *Thermodynamic formalism. The mathematical structures of equilibrium statistical mechanics. Second edition,* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2004.

[89] D. Ruelle, *Dynamical zeta functions for piecewise monotone maps of the interval,* CRM Monograph Series, 4. American Mathematical Society, Providence, RI, 1994.

[90] S. Ruette, *Mixing C^r maps of the interval without maximal measure,* Israel J. Math. 127 (2002), 253–277.

[91] S. Ruette, *On the Vere-Jones classification and existence of maximal measures for countable topological Markov chains,* Pacific J. Math. 209 (2003), no. 2, 366–380.

[92] S. Ruette, *C^n interval maps not Borel conjugate to any C^∞ map,* Proc. Amer. Math. Soc. 132 (2004), no. 4, 1091–1093.

[93] S. Ruette, *Chaos for continuous interval maps - a survey of relationship between the various sorts of chaos,* (preliminary version - 123 pages) at `http://www.math.u-psud.fr/ ˜ruette/publications.html`.

[94] M. Rychlik, *Regularity of the metric entropy for expanding maps,* Trans. Amer. Math. Soc. 315 (1989), no. 2, 833–847.

[95] M. Rychlik, E. Sorets, *Regularity and other properties of absolutely continuous invariant measures for the quadratic family,* Comm. Math. Phys. 150 (1992), no. 2, 217–236

[96] O. Sarig, *Thermodynamic formalism for countable Markov shifts,* Ergodic Theory Dynam. Systems 19 (1999), no. 6, 1565–1593.

[97] O. Sarig, *Thermodynamic formalism for null recurrent potentials,* Israel J. Math. 121 (2001), 285–311.

[98] B. Schmitt, *Ergodic theory and thermodynamic of one-dimensional Markov expanding endomorphisms,* in: Dynamical systems (Temuco, 1991/1992), 93–123, Travaux en Cours, 52, Hermann, Paris, 1996.

[99] C. Shannon, W. Weaver, *The Mathematical Theory of Communication,* The University of Illinois Press, Urbana, Ill., 1949.

[100] P. Shields, *The ergodic theory of discrete sample paths,* Graduate Studies in Mathematics, 13. American Mathematical Society, Providence, RI, 1996.

[101] Ya. Sinai, *Introduction to ergodic theory,* Mathematical Notes, 18. Princeton University Press, Princeton, N.J., 1976.

[102] Ya. Sinai, *Topics in ergodic theory,* Princeton Mathematical Series, 44. Princeton University Press, Princeton, NJ, 1994.

[103] M. Tsujii, *A simple proof for monotonicity of entropy in the quadratic family,* Ergodic Theory Dynam. Systems 20 (2000), no. 3, 925–933.

[104] D. Vere-Jones, *Geometric ergodicity in denumerable Markov chains,* Quart. J. Math. (Oxford, 2) **13** (1962), 7–28.

[105] D. Vere-Jones, *Ergodic properties of nonnegative matrices I,* Pacific J. Math. 22 1967 361–386.

[106] D. Vere-Jones, *Ergodic properties of nonnegative matrices. II,* Pacific J. Math. 26 1968 601–620.

[107] P. Walters, *An introduction to ergodic theory,* Graduate Texts in Mathematics, 79. Springer-Verlag, New York-Berlin, 1982.

[108] B. Weiss, *Single orbit dynamics,* CBMS Regional Conference Series in Mathematics, 95. American Mathematical Society, Providence, RI, 2000.

[109] L.-S. Young, *Dimension, entropy and Lyapunov exponents,* Ergodic Theory Dynamical Systems 2 (1982), no. 1, 109–124.

[110] Y. Yomdin, *Volume growth and entropy,* Israel J. Math. 57 (1987), no. 3, 285–300.

JÉRÔME BUZZI, Centre de Mathématiques, Ecole polytechnique, 91128 Palaiseau Cedex, France
E-mail : buzzi@math.polytechnique.fr • *Url :* www.jeromebuzzi.com