Séminaires & Congrès

COLLECTION SMF

ENTROPY THEORY ON THE INTERVAL

Jérôme Buzzi

ÉCOLE DE THÉORIE ERGODIQUE

Numéro 20 Y. Lacroix, P. Liardet, J.-P. Thouvenot, éds.

SOCIÉTÉ MATHÉMATIQUE DE FRANCE

Séminaires & Congrès
20, 2010, p. 39–82

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 *École thématique de théorie ergodique* at the C.I.R.M., Marseille, in April 2006.

***Résumé* (Theorie de l'entropie sur un intervalle).** — Nous présentons un survol de la théorie de l'entropie des applications de l'intervalle. Nous nous plaçons du point de vue de la théorie ergodique et considérons notamment les mesures d'entropie maximale et les points périodiques. Les outils principaux sont (i) une forme adaptée du diagramme de Markov introduit par Hofbauer, (ii) la propriété de pistage et ses conséquences (borne sur l'entropie, propriété de rang 1 faible), (iii) les sous-décalages markoviens à espace d'états dénombrable fortement positivement récurrents. Les preuves ne sont données que pour une sélection de résultats. Cet article est basé sur les conférences prononcées à l'occasion de l'*École thématique de théorie ergodique* qui s'est tenue au C.I.R.M. en avril 2006.

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

2000 Mathematics Subject Classification. — 37Axx, 37Bxx.
Key words and phrases. — Topological and combinatorial dynamics, ergodic theory, symbolic dynamics, entropy, variational principle, interval maps, piecewise monotone maps, horseshoes, measures maximizing entropy, periodic orbits, Artin-Mazur zeta function, kneading invariants, strongly positive recurrent Markov shifts, Markov diagram, shadowing, weak rank one.

J.B. wishes to thank the organizers and the participants of the *École thématique de théorie ergodique* at the C.I.R.M. in Marseille.

(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 [**56**] 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 [**68**]); (2) the Birkhoff ergodic theorem, explained in any textbook on ergodic theory (see, *e.g.*, [**86**]).

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 continuous. The above topology induces 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 periodic 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.*, [**32**]).

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

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

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 alphabets "contained" in the considered dynamical system. We refer to the many excellent texts (see, *e.g.*, [**55**, **66**, **69**, **79**, **80**, **81**]) 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:

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

[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.

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 $\psi : X \to Y$ 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 [66] for an introduction and [69] 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 [64]).

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 theorem:

***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 $n \to \infty$:*

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

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*