Dynamical systems II: Introductory Lectures (Master)

Tien-Cuong Dinh

January 01, 2011
available at http://www.math.jussieu.fr/∼dinh

Preface

This text is written for the students in the Master program at the University of Paris 6 and is the continuation of the lectures given by Patrice Le Calvez (Systèmes dynamiques: cours fondamental I, 2010-2011). A large part is translated from another text by Patrice Le Calvez (Systèmes dynamiques: cours fondamental II, 2009-2010).

Contents

Chapter 1

Topological entropy

1.1 Open covers, entropy and generators

Let X be a topological space. *An open cover* of X if a family $\mathscr{U} = (U_i)_{i\in I}$ of open subsets U_i of X whose union is equal to X, that is, we have $\cup_{i\in I} U_i = X$. Denote by $N(\mathscr{U})$ the minimal number of elements in $\mathscr{U}$ needed in order to cover X.

An open cover $\mathscr{V} = (V_j)_{j\in J}$ is said to be *finer* than $\mathscr{U}$ (or is *a refinement* of $\mathscr{U}$) if any element V_j of $\mathscr{V}$ is contained in some element U_i of $\mathscr{U}$. We will write in this case $\mathscr{U} \preceq \mathscr{V}$ and $\mathscr{V} \succeq \mathscr{U}$. We write also $\mathscr{U} \sim \mathscr{V}$ when both relations $\mathscr{U} \preceq \mathscr{V}$ and $\mathscr{V} \preceq \mathscr{U}$ are satisfied.

Example 1.1.1. Consider the case where X is the circle $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$. Let $\mathscr{U}$ (resp. $\mathscr{V}$ or $\mathscr{W}$) be the set of the images in X of the intervals $]a, b[$ with $a, b \in \mathbb{Q}$ (resp. $a, b \in \mathbb{Q}$ with $b - a = 2$ or $b - a = 1$). We have

$$\mathscr{U} \sim \mathscr{V} \preceq \mathscr{W}, \quad N(\mathscr{U}) = N(\mathscr{V}) = 1 \quad \text{and} \quad N(\mathscr{W}) = 2.$$

For arbitrary open covers $\mathscr{U}$ and $\mathscr{V}$, the open sets $U_i \cap V_j$ define a new open cover that we denote by $\mathscr{U} \vee \mathscr{V}$. We have $\mathscr{U} \preceq \mathscr{U} \vee \mathscr{V}$ and $\mathscr{V} \preceq \mathscr{U} \vee \mathscr{V}$. In fact, any open cover $\mathscr{W}$ such that $\mathscr{U} \preceq \mathscr{W}$ and $\mathscr{V} \preceq \mathscr{W}$ is finer than $\mathscr{U} \vee \mathscr{V}$.

The following properties are elementary. Their proofs are left to the reader.

Proposition 1.1.2. *Let $\mathscr{U}, \mathscr{V}, \mathscr{U}^1, \ldots, \mathscr{U}^n$ be open covers of X.*

1. *$N(\mathscr{U}) = 1$ if and only if X is an element of $\mathscr{U}$;*
2. *If $\mathscr{U} \preceq \mathscr{V}$ then $N(\mathscr{U}) \leq N(\mathscr{V})$. If $\mathscr{U} \sim \mathscr{V}$ then $N(\mathscr{U}) = N(\mathscr{V})$;*
3. *$N(\bigvee_{k=1}^n \mathscr{U}^k) \leq \prod_{k=1}^n N(\mathscr{U}^k)$;*
4. *If $\pi : Y \to X$ is a continuous map, then $N(\pi^{-1}(\mathscr{U})) \leq N(\mathscr{U})$ with equality when π is surjective.*

Let $T : X \to X$ be a continuous map. From now on assume for simplicity that X is compact. We have the following result.

Proposition 1.1.3. *Let $\mathscr{U}$ be an open cover of X. Then the sequence*

$$u_n = N\Big(\bigvee_{i=0}^{n-1} T^{-i}(\mathscr{U})\Big)$$

is sub-multiplicative, that is, $u_{n+m} \leq u_n u_m$ for $n, m \geq 1$. In particular, $\frac{1}{n}\log u_n$ converge to $\inf_{n\geq 1} \frac{1}{n}\log u_n$ which is finite when X is compact.

Proof. Using the last two assertions of the previous proposition, we obtain

$$\begin{aligned}
u_{n+m} &= N\Big(\bigvee_{i=0}^{n+m-1} T^{-i}(\mathscr{U})\Big) \\
&= N\Big(\Big(\bigvee_{i=0}^{n-1} T^{-i}(\mathscr{U})\Big) \vee T^{-n}\Big(\bigvee_{i=0}^{m-1} T^{-i}(\mathscr{U})\Big)\Big) \\
&\leq N\Big(\bigvee_{i=0}^{n-1} T^{-i}(\mathscr{U})\Big) N\Big(T^{-n}\Big(\bigvee_{i=0}^{m-1} T^{-i}(\mathscr{U})\Big)\Big) \\
&\leq u_n u_m.
\end{aligned}$$

This gives the first assertion. We also deduce the second assertion using Exercise 1.6.1 below.

When X is compact, u_n is finite for every n. Therefore, $\inf_{n\geq 1} \frac{1}{n}\log u_n$ is finite. □

Definition 1.1.4. The limit of $\frac{1}{n}\log u_n$ is denoted by $h(T, \mathscr{U})$. It is called *the entropy* of T relatively to $\mathscr{U}$.

Here are some basic properties of the entropy.

Proposition 1.1.5. *With the above notations, we have*

1. *If $\mathscr{U} \preceq \mathscr{V}$ then $h(T, \mathscr{U}) \leq h(T, \mathscr{V})$;*
2. *$h(T, \mathscr{U}) = h(T, \bigvee_{i=0}^{m} T^{-i}(\mathscr{U}))$ for every $m \geq 0$;*
3. *$h(T^m, \bigvee_{i=0}^{m-1} T^{-i}(\mathscr{U})) = mh(T, \mathscr{U})$ for every $m \geq 1$;*
4. *If T is a homeomorphism, then*

$$h(T^{-1}, \mathscr{U}) = h(T, \mathscr{U}) = h(T, T(\mathscr{U})) = h(T, T^{-1}(\mathscr{U})).$$

Proof. 1) This property is a direct consequence of the definition of entropy.

2) Define $\mathscr{V} = \bigvee_{i=0}^{m} T^{-i}(\mathscr{U})$. It follows from 1) that

$$h(T, \mathscr{U}) \leq h(T, \mathscr{V}).$$

It remains to check the converse inequality. We have for any $n \geq 0$

$$\bigvee_{i=0}^{n} T^{-i}(\mathscr{V}) \preceq \bigvee_{i=0}^{n+m} T^{-i}(\mathscr{U}).$$

Therefore, from the definition of entropy, we obtain

$$h(T, \mathscr{V}) \leq h(T, \mathscr{U}).$$

3) From the definition of entropy, we have

$$mh(T, \mathscr{U}) = h(T^m, \bigvee_{i=0}^{m-1} \mathscr{U}).$$

Then, the assertion 2) implies the result.

4) The last two equalities are direct consequence of the fact that T is a homeomorphism. The hypothesis implies also that

$$\bigvee_{i=0}^{m} T^{-i}(\mathscr{U}) = T^{-m} \bigvee_{i=0}^{m} T^{i}(\mathscr{U}).$$

Therefore,

$$N\Big(\bigvee_{i=0}^{m} T^{-i}(\mathscr{U})\Big) = N\Big(T^{-m} \bigvee_{i=0}^{m} T^{i}(\mathscr{U})\Big) = N\Big(\bigvee_{i=0}^{m} T^{i}(\mathscr{U})\Big).$$

This implies the first equality in 4). □

Definition 1.1.6. We call *topological entropy* of T the following non-negative number

$$h(T) = \sup \big\{h(T, \mathscr{U}), \quad \mathscr{U} \quad \text{open cover of } X\big\}.$$

Note that $h(T)$ may be infinite. We give some basic properties of the topological entropy.

Proposition 1.1.7. *With the above notations, we have*

1. $h(T^n) = nh(T)$ *for* $n \geq 0$. *In particular,* $h(\mathrm{id}) = 0$*;*
2. *If* T *is a homeomorphism then* $h(T^n) = |n|h(T)$ *for* $n \in \mathbb{Z}$*;*
3. *If* S *is a factor of* T*, then* $h(S) \leq h(T)$*. In particular, if* S *is conjugate to* T *then* $h(S) = h(T)$*;*

4. If $Y \subset X$ is an invariant closed set, then $h(T_{|Y}) \leq h(T)$.

Proof. 1) This is a consequence of part 3) in Proposition 1.1.5.

2) This is a consequence of the above property and of part 4) in Proposition 1.1.5.

3) Assume that $S : Y \to Y$ is semi-conjugate to T via the continuous surjective map $\pi : X \to Y$. If $\mathscr{U}$ is an open cover of Y, $\mathscr{V} = \pi^{-1}(\mathscr{U})$ is an open cover of X. The identity $\pi \circ T^n = S^n \circ \pi$ implies that

$$\pi^{-1}\Big(\bigvee_{i=0}^{m} S^{-i}(\mathscr{U})\Big) = \bigvee_{i=0}^{m} T^{-i}(\mathscr{V}).$$

We then obtain

$$N\Big(\bigvee_{i=0}^{m} S^{-i}(\mathscr{U})\Big) = N\Big(\bigvee_{i=0}^{m} T^{-i}(\mathscr{V})\Big).$$

Finally, this implies

$$h(S, \mathscr{U}) \leq h(T, \mathscr{V}).$$

The result follows.

4) If $\mathscr{U} = (U_i)_{i \in I}$ is an open cover of Y, there are open sets $\widehat{U}_i$ of X such that $\widehat{U}_i \cap Y = U_i$. Consider the cover $\widehat{\mathscr{U}}$ of X formed by these open sets and the open set $\widehat{U} = X \setminus Y$. Since Y is invariant, we have $T^{-n}(\widehat{U}) \subset \widehat{U}$ for $n \geq 0$. It follows that

$$N\Big(\bigvee_{i=0}^{n} T_{|Y}^{-i}(\mathscr{U})\Big) \leq N\Big(\bigvee_{i=0}^{n} T^{-i}(\widehat{\mathscr{U}})\Big).$$

This give the result. □

We now introduce some notions which allow us to simplify computations on entropy.

Definition 1.1.8. A family $(\mathscr{U}^\alpha)$ of open covers is said to be *generating* if for every open cover $\mathscr{U}$ there is an α such that $\mathscr{U} \preceq \mathscr{U}^\alpha$.

This notion is independent of T. The following proposition is straightforward, see Proposition 1.1.5.

Proposition 1.1.9. *If $(\mathscr{U}^\alpha)_{\alpha \in A}$ is a generating family of open covers, then*

$$h(T) = \sup_{\alpha \in A} h(T, \mathscr{U}^\alpha).$$

Definition 1.1.10. We say that $\mathscr{U}$ is a *generator* if the sequence

$$\Big(\bigvee_{i=0}^{n-1} T^{-i}(\mathscr{U})\Big)_{n \geq 1}$$

is a generating family.

Note that this notion depends on T. The following useful property is straightforward, see Proposition 1.1.5.

Proposition 1.1.11. *Let $\mathscr{U}$ be a generator of (X,T). Then*

$$h(T) = h(T, \mathscr{U}).$$

In particular, if such a cover exists and if X is compact then $h(T)$ is finite.

We will close this section by giving some examples.

Example 1.1.12. Consider the Bernouilli one-sided shift

$$\sigma : A^{\mathbb{N}} \to A^{\mathbb{N}}$$

$$(x_n)_{n\geq 0} \mapsto (x_{n+1})_{n\geq 0}$$

where A is a finite alphabet of cardinal $p \geq 2$. Consider also the cover $\mathscr{U} = (U_a)_{a\in A}$ where U_a is the cylinder $U_a = \{x \in A^{\mathbb{N}},\ x_0 = a\}$. This cover is a generator. Indeed, the cover $\bigvee_{i=0}^{n-1} \sigma^{-i}(\mathscr{U})$ is the family of the cylinder blocks based on the first n coordinates. It contains p^n cylinder blocks and its diameter tends to 0 when $n \to \infty$. It follows that

$$h(T) = h(T, \mathscr{U}) = \lim_{n\to\infty} \frac{1}{n} \log \Big(N\Big(\bigvee_{i=0}^{n-1} \sigma^{-i}(\mathscr{U}) \Big)\Big) = \lim_{n\to\infty} \frac{1}{n} \log p^n = \log p.$$

Example 1.1.13. Consider now the Bernouilli two-sided shift

$$\sigma : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$$

$$(x_n)_{n\in\mathbb{Z}} \mapsto (x_{n+1})_{n\in\mathbb{Z}}$$

where A is a finite alphabet of cardinal $p \geq 2$. Consider also the cover $\mathscr{U} = (U_a)_{a\in A}$ where U_a is the cylinder $U_a = \{x \in A^{\mathbb{Z}},\ x_0 = a\}$. This cover is not a generator. However, the sequence

$$\Big(\bigvee_{i=-n}^{n} \sigma^{-i}(\mathscr{U}) \Big)_{n\geq 0}$$

is a generating family. It follows from Proposition 1.1.5 that

$$h(\sigma) = \lim_{n\to\infty} h\Big(\sigma, \bigvee_{i=-n}^{n} \sigma^{-i}(\mathscr{U})\Big) = \lim_{n\to\infty} h\Big(\sigma, \bigvee_{i=0}^{2n} \sigma^{-i}(\mathscr{U})\Big) = h(\sigma, \mathscr{U}) = \log p,$$

where the last identity is obtained as in the previous example.

Example 1.1.14. Consider the map T on the circle $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ defined by $x \mapsto px$ where p is an integer such that $|p| \geq 2$. We will prove that $h(T) = \log|p|$. Since $T^2(x) = p^2(x)$ and $h(T^2) = 2h(T)$, we can replace T with T^2 in order to assume that p is positive.

Since T is a factor of the above one-sided shift σ which is of entropy $\log p$, we have that $h(T) \leq \log p$. Fix a constant $\delta > 0$ small enough and consider a cover $\mathscr{U}$ whose elements are intervals of length $< \delta$. We only need to show that $h(T, \mathscr{U}) \geq \log p$. Indeed, such covers form a generating family.

If I is an element of $\mathscr{U}$, then $T^{-1}(I)$ is a disjoint union of p intervals of length less than δ/p which are separated by intervals of length larger than $(1-\delta)/p > \delta$. It follows that the intersection of $T^{-1}(I)$ with any interval J of length less than δ is connected. So, this intersection is an interval of length less than δ. We deduce that all the elements of $\mathscr{U} \vee T^{-1}(\mathscr{U})$ are intervals of length less than δ/p.

Arguing by induction, we see that the elements of the cover

$$\bigvee_{i=0}^{n-1} T^{-i}(\mathscr{U})$$

are intervals of length less than δp^{-n+1}. Therefore,

$$N\Big(\bigvee_{i=0}^{n-1} T^{-i}(\mathscr{U})\Big) \geq \delta^{-1} p^{n-1}.$$

It is now easy to deduce that $h(T, \mathscr{U}) \geq \log p$.

Example 1.1.15. Let T be a homeomorphism of $\mathbb{T}$. We will show that $h(T) = 0$. Observe that if $\mathscr{U}, \mathscr{V}$ are two covers by intervals of length less than $1/2$, so is $\mathscr{U} \vee \mathscr{V}$. Consider sub-covers $\mathscr{U}' \subset \mathscr{U}$ and $\mathscr{V}' \subset \mathscr{V}$ of minimal cardinal. Let I and J be respectively elements of $\mathscr{U}'$ and $\mathscr{V}'$. If $I \cap J$ is not empty, its left extremal point is either the left extremal point of I or the one of J. It follows that

$$\#\mathscr{U}' \vee \mathscr{V}' \leq \#\mathscr{U}' + \#\mathscr{V}'.$$

Therefore, we have

$$N(\mathscr{U} \vee \mathscr{V}) \leq N(\mathscr{U}) + N(\mathscr{V}).$$

Fix a constant $0 < \delta < 1/2$ small enough such that

$$\operatorname{dist}(T^{-1}(x), T^{-1}(y)) < 1/2 \quad \text{when} \quad \operatorname{dist}(x, y) < \delta.$$

Fix also a cover $\mathscr{U}$ by intervals of length $< \delta$. It is enough to show that $h(T, \mathscr{U}) = 0$. Indeed, the family of these covers $\mathscr{U}$ is generating. Observe that $\mathscr{U}$ and $T^{-1}(\mathscr{U})$ are covers by intervals of length less than $1/2$. Therefore, $\mathscr{U} \vee T^{-1}(\mathscr{U})$ is a cover by intervals of length less than δ. We deduce from the above discussion that

$$N(\mathscr{U} \vee T^{-1}(\mathscr{U})) \leq N(\mathscr{U}) + N(T^{-1}(\mathscr{U})) = 2N(\mathscr{U}).$$

By induction, we obtain

$$N\Big(\bigvee_{i=0}^{n-1} T^{-i}(\mathscr{U})\Big) \leq nN(\mathscr{U}).$$

Hence, $h(T,\mathscr{U})=0$.

1.2 Bowen-Dinaburg's definition

In this section, assume for simplicity that X is a compact metric space. We will give an equivalent definition of the topological entropy due to Bowen and Dinaburg.

Let d denote the distance on X. Define a sequence of new distances d_n by

$$d_n(x,y) = \max_{0\leq i\leq n-1} d(T^i(x),T^i(y)) \quad \text{for} \quad n\geq 1.$$

It is not difficult to check that d_n is a distance. Denote by $B_n(x,\epsilon)$ the ball of center x and of radius ϵ with respect to d_n. We often call it *Bowen ball.* Even when X is a manifold, this ball is not necessarily connected.

Example 1.2.1. Let T be the map on $\mathbb{T}$ defined by $x\mapsto 4x$. The ball $B_2(0,1/4)$ is not connected. It contains 3 connected components.

Definition 1.2.2. A subset S of X is said to be (n,ϵ)*-separated* if $d_n(x,y)\geq\epsilon$ for all distinct points x,y in S. A subset R of X is (n,ϵ)*-spanning* if $d_n(x,R)<\epsilon$ for every $x\in X$.

Denote by $s(n,\epsilon)$ the maximal cardinal of (n,ϵ)-separated sets and $r(n,\epsilon)$ the minimal cardinal of (n,ϵ)-spanning sets. Denote by $\mathscr{U}^\epsilon$ the family of all the balls $B(x,\epsilon)$ of radius ϵ in X with respect to the distance d. For simplicity, let

$$N(n,\epsilon) = N\Big(\bigvee_{i=0}^{n-1} T^{-i}(\mathscr{U}^\epsilon)\Big).$$

These three quantities satisfy the following properties.

Proposition 1.2.3. *We have*

$$N(n,\epsilon)\leq r(n,\epsilon)\leq s(n,\epsilon)\leq N(n,\frac{\epsilon}{2}).$$

Proof. Consider an (n,ϵ)-spanning set R whose cardinal is minimal, i.e. equal to $r(n,\epsilon)$. So, the (n,ϵ)-balls with center in R with respect to the distance d is an open cover. Since any (n,ϵ)-ball is an element of

$$\bigvee_{i=0}^{n-1} T^{-i}(\mathscr{U}^\epsilon)$$

we deduce that

$$N(n,\epsilon) \leq r(n,\epsilon).$$

Consider now an (n,ϵ)-separated set S of maximal cardinal $s(n,\epsilon)$. Since the cardinal is maximal, the family of (n,ϵ)-balls with center in S is a cover. Indeed, otherwise, we can add to S any point outside these balls in order to get a larger (n,ϵ)-separated set. As a direct consequence, we obtain $r(n,\epsilon) \leq s(n,\epsilon)$.

Finally, since S is (n,ϵ)-separated, each element of

$$\bigvee_{i=0}^{n-1} T^{-i}(\mathscr{U}^{\epsilon/2})$$

contains at most one point in S. The last equality in the proposition follows. □

We deduce that the topological entropy can be computed using the following formulas.

Corollary 1.2.4. *We have*

$$\begin{aligned} h(T) &= \lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\log N(n,\epsilon) \\ &= \lim_{\epsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log r(n,\epsilon) \\ &= \lim_{\epsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log s(n,\epsilon) \\ &= \lim_{\epsilon\to 0}\liminf_{n\to\infty}\frac{1}{n}\log r(n,\epsilon) \\ &= \lim_{\epsilon\to 0}\liminf_{n\to\infty}\frac{1}{n}\log s(n,\epsilon), \end{aligned}$$

where $\lim_{\epsilon\to 0}$ *can be replaced with* $\sup_{\epsilon>0}$.

Proof. All the considered quantities increase when ϵ decreases to 0. So, we can replace $\lim_{\epsilon\to 0}$ with $\sup_{\epsilon>0}$.

Observe that the family of covers $(\mathscr{U}^{\epsilon})_{\epsilon>0}$ is generating. So, the first equality is a consequence of Propositions 1.1.5 and 1.1.9. The other equalities are then deduced from Proposition 1.2.3. □

The above corollary implies that $h(T)$ does not depend on the choice of the distance d with a fixed topology on X. This is why it is called topological entropy.

Corollary 1.2.5. *If T is 1-Lipschitz, then $h(T)=0$. If T is A-Lipschitz and X is a subset of a smooth manifold of dimension d, then $h(T)\leq d\log^+ A$ where $\log^+ A = \max(\log A, 0)$.*

Proof. Assume that T is 1-Lipschitz. If S is an (n,ϵ)-separated set then S is ϵ-separated, that is, $\mathrm{dist}\,(x,y)\geq\epsilon$ for all distinct points $x,y\in S$. Therefore, its cardinal is bounded by a constant independent of n. We then deduce that $h(T)=0$.

Assume now that T is A-Lipschitz and X is a subset of a smooth manifold of dimension d. We only have to consider the case $A>1$. If S is (n,ϵ)-separated then S is $A^{-n}\epsilon$-separated. It follows that $\#S\leq c(A^{-n}\epsilon)^{-d}$ for some constant $c>0$. It is easy to deduce from Corollary 1.2.4 that $h(T)\leq d\log A$. □

1.3 Variational principle

Assume always that X is compact. Recall that the family $\mathscr{M}_T$ of invariant probability measures is a non-empty convex set which is compact with respect to the weak topology on measures. For each measure μ in $\mathscr{M}_T$ denote by $h_\mu(T)$ its entropy, see [7] for definition. The main result in this section is the so-called *the variational principle* due to Goodwyn-Goodman.

Theorem 1.3.1. *Let $T:X\to X$ be a continuous map on a compact metric space X. Then*

$$h(T)=\sup_{\mu\in\mathscr{M}_T}h_\mu(T).$$

The proof of this important result that we will give here is due to Misiurewicz. We first recall some basic lemmas of measure theory.

Lemma 1.3.2. *Any Borel probability measure μ on a metric space is regular. That is, if A is a Borel set and $\epsilon>0$ is a constant, there are a closed set F and an open set U such that $F\subset A\subset U$ and $\mu(U\setminus F)<\epsilon$.*

Proof. We only have to check that the family $\mathscr{C}$ of Borel sets satisfying the above property is a σ-algebra containing all the closed sets. Observe that if A is in $\mathscr{C}$ then its complement is also in $\mathscr{C}$. Moreover, X is an element of $\mathscr{C}$. In order to show that $\mathscr{C}$ is a σ-algebra, we only have to check that if A_n is a sequence in $\mathscr{C}$ then $\cup A_n$ is also an element of $\mathscr{C}$.

Choose closed sets F_n and open sets U_n such that

$$F_n\subset A_n\subset U_n\quad\text{and}\quad\mu(U_n\setminus F_n)<2^{-n-1}\epsilon.$$

We have

$$\cup F_n\subset\cup A_n\subset\cup U_n$$

and

$$\mu(\cup U_n\setminus\cup F_n)\leq\sum\mu(U_n\setminus F_n)<\sum 2^{-n-1}\epsilon=\epsilon.$$

Define

$$U=\cup U_n\quad\text{and}\quad F=\cup_{n\leq N}F_n$$

for some integer N large enough. Since N is large, we have $\mu(U \setminus F) < \epsilon$. It is clear that U is open and F is closed. So, $\cup A_n$ is an element of $\mathscr{C}$ and $\mathscr{C}$ is a σ-algebra.

Now, we show that any closed set A belongs to $\mathscr{C}$. Choose $F = A$ and

$$U_m = \{x \in X, \quad \mathrm{dist}\,(x, A) < 1/m\}$$

for $m \geq 1$. Clearly, U_m is open and $F \subset A \subset U_m$. Since $A = \cap U_m$, we have

$$\lim_{m\to\infty} \mu(U_m \setminus A) = \mu(\cap U_m \setminus A) = 0.$$

Therefore, if $U = U_m$ with m large enough, we have $\mu(U \setminus F) < \epsilon$. Hence, A is an element of $\mathscr{C}$. □

Lemma 1.3.3. *Let X be a metric space and (μ_n) a sequence of Borel probability measures which converges weakly to a measure μ. Then for any Borel set A such that $\mu(\partial A) = 0$, we have*

$$\lim_{m\to\infty} \mu_n(A) = \mu(A).$$

Proof. Define continuous functions f_k with $k \geq 1$ by

$$f_k(x) = \max(1 - k\mathrm{dist}\,(x, A), 0).$$

This sequence decreases to the characteristic function $\chi_{\overline{A}}$ of $\overline{A}$. Since μ_n converges to μ, we have

$$\limsup_{n\to\infty} \mu_n(A) \leq \limsup_{n\to\infty} \int f_k d\mu_n = \lim_{n\to\infty} \int f_k d\mu_n = \int f_k d\mu.$$

It follows that

$$\limsup_{n\to\infty} \mu_n(A) \leq \inf_{k\geq 1} \int f_k d\mu = \lim_{k\to\infty} \int f_k d\mu = \mu(\overline{A}) = \mu(A).$$

The last identity is a consequence of the property that $\mu(\partial A) = 0$. The obtained inequality together with the similar one for the complement of A implies the result. □

Lemma 1.3.4. *Let μ be a Borel probability measure on a compact metric space X. Then, for every $\epsilon > 0$, there is a finite partition $\mathscr{P} = (P_i)_{i\in I}$ of Borel sets such that for every $i \in I$, we have* $\mathrm{diam}(P_i) < \epsilon$ *and* $\mu(\partial P_i) = 0$.

Proof. Observe that for each $x \in X$, there are only a finite or countable number of $r > 0$ such that $\mu(\partial B(x, r)) > 0$. Since X is compact, there is a finite cover of X by balls $(B_i)_{1\leq i\leq m}$ of diameter less than ϵ such that $\mu(\partial B_i) = 0$. Now, it is enough to define

$$P_i := B_i \setminus \cup_{1\leq j<i} B_j.$$

□

Theorem 1.3.1 is a consequence of the following two propositions.

Proposition 1.3.5. *If μ is an invariant probability measure, then*

$$h_\mu(T) \leq h(T).$$

Proof. First, observe that it is enough to prove the following inequality

$$h_\mu(T) \leq 1 + \log 2 + h(T).$$

Indeed, the same inequality applied to T^m gives

$$mh_\mu(T) \leq 1 + \log 2 + mh(T).$$

After dividing by m, we obtain the result by taking the limit when m tends to infinity.

Choose a measurable partition $\mathscr{P} = (P_i)_{1\leq i\leq r}$. Fix a constant $\epsilon > 0$ small enough. By Lemma 1.3.2, we can find closed sets $Q_i \subset P_i$ such that $\mu(P_i \backslash Q_i) < \epsilon$. Define also

$$Q_0 = X \setminus \bigcup_{1\leq i\leq r} Q_i \quad \text{and} \quad P_0 = \varnothing.$$

Consider now two partitions $\mathscr{Q} = (Q_i)_{0\leq i\leq r}$ and $\mathscr{P}' = (P_i)_{0\leq i\leq r}$. It is not difficult to see that $\mu(P_i \Delta Q_i) \leq r\epsilon$ for $0 \leq i \leq r$ (one can distinguish the case $i = 0$ and the case $i \neq 0$). Therefore, since ϵ is small, we obtain

$$H(\mathscr{P}|\mathscr{Q}) = H(\mathscr{P}'|\mathscr{Q}) \leq 1.$$

It follows that

$$h_\mu(T, \mathscr{P}) \leq h_\mu(T, \mathscr{Q}) + H(\mathscr{P}|\mathscr{Q}) \leq h_\mu(T, \mathscr{Q}) + 1,$$

see [7].

For $1 \leq i \leq r$, define

$$U_i = Q_0 \cup Q_i = X \setminus \bigcup_{\substack{1\leq j\leq r\\ j\neq i}} Q_j.$$

Consider now the open covers $\mathscr{U} = (U_i)_{1\leq i\leq r}$ and $\bigvee_{k=1}^{n-1} T^{-k}(\mathscr{U})$. Each element of the last cover has the form

$$\bigcap_{0\leq k<n} T^{-k}(U_{i_k}) = \bigcap_{0\leq k<n} T^{-k}(Q_0 \cup Q_{i_k}) = \bigcap_{0\leq k<n} \Big(T^{-k}(Q_0) \cup T^{-k}(Q_{i_k})\Big).$$

So, it is a union of at most 2^n elements (possibly empty) of the partition $\bigvee_{0\leq k<n} T^{-k}(\mathscr{Q})$.

We also see that each non-empty element in the last partition appears at most 2^n times in the above description of elements of $\bigvee_{k=1}^{n-1} T^{-k}(\mathscr{U})$. If M is the number of non-empty elements in $\bigvee_{0\leq k<n} T^{-k}(\mathscr{Q})$, we have

$$H_\mu\Big(\bigvee_{0\leq k<n} T^{-k}(\mathscr{Q})\Big) \leq \log M \leq \log\Big(2^n N\Big(\bigvee_{0\leq k<n} T^{-k}(\mathscr{U})\Big)\Big),$$

see [7]. It follows that

$$h_\mu(T,\mathscr{Q}) \leq \log 2 + h(T,\mathscr{U}) \leq \log 2 + h(T),$$

and hence

$$h_\mu(T,\mathscr{P}) \leq 1 + \log 2 + h(T).$$

This implies the result. □

The proof of the following proposition contains a useful construction of invariant measures.

Proposition 1.3.6. *For any $\epsilon > 0$, there is an invariant probability measure μ such that*

$$h_\mu(T) \geq \limsup_{n\to\infty} \frac{1}{n} \log s(n,\epsilon).$$

Proof. For each $n \geq 1$ choose an (n,ϵ)-separated set S_n of maximal cardinal $s(n,\epsilon)$. Consider the sequences of probability measures

$$\nu_n = \frac{1}{s(n,\epsilon)} \sum_{x\in S_n} \delta_x$$

and

$$\mu_n := \frac{1}{n}\sum_{i=0}^{n-1} T_*^i(\nu_n).$$

Choose a strictly increasing sequence (n_k) such that

$$\lim_{k\to\infty} \frac{1}{n_k} \log s(n_k,\epsilon) = \limsup_{n\to\infty} \frac{1}{n} \log s(n,\epsilon)$$

and such that μ_{n_k} converges to a probability measure μ.

The measure μ in invariant. Indeed, we have

$$T_*(\mu) - \mu = \lim_{k\to\infty} \frac{1}{n_k}(T^{n_k})_*(\nu_{n_k}) - \frac{1}{n_k}\nu_{n_k}.$$

The last measure is of mass at most equal to $2/n_k$. So, it converges to 0 as k tends to infinity.

Now, consider a partition $\mathscr{P} = (P_i)_{i\in I}$ of diameter less than ϵ such that $\mu(\partial P_i) = 0$, see Lemma 1.3.4. We will show that

$$h_\mu(T, \mathscr{P}) \geq \limsup_{n\to\infty} \frac{1}{n} \log s(n, \epsilon).$$

This implies the lemma since $h_\mu(T) \geq h_\mu(T, \mathscr{P})$.

Consider the partitions $\mathscr{P}^n = \bigvee_{0\leq i<n} T^{-i}(\mathscr{P})$. Since S_n is (n, ϵ)-separated, each element of $\mathscr{P}^n$ contains at most one element of S_n. Therefore,

$$H_{\nu_n}(\mathscr{P}^n) = \log s(n, \epsilon).$$

Claim. For any partition $\mathscr{Q} = (Q_j)_{j\in J}$ and for any $q \leq n$, we have

$$qH_{\nu_n}(\mathscr{Q}^n) \leq nH_{\mu_n}(\mathscr{Q}^q) + 2q^2 \log \#J.$$

Here, $\mathscr{Q}^n$ denotes the partition $\bigvee_{0\leq i<n} T^{-i}(\mathscr{Q})$.

We first assume the claim and complete the proof of the proposition. Applying the claim to $\mathscr{P}^{n_k}$ and dividing the obtained inequality to qn_k give

$$\frac{1}{n_k} \log s(n_k, \epsilon) = \frac{1}{n_k} H_{\nu_{n_k}}(\mathscr{P}^{n_k}) \leq \frac{1}{q} H_{\mu_{n_k}}(\mathscr{P}^q) + \frac{2q}{n_k} \log \#I.$$

Observe that since μ is invariant, the boundary of the elements of $\mathscr{P}^q$ have zero μ measure. So, by Lemma 1.3.3, the last estimate gives

$$\limsup_{n\to\infty} \frac{1}{n} \log s(n, \epsilon) \leq \frac{1}{q} H_\mu(\mathscr{P}^q).$$

When q tends to infinity, we obtain

$$\limsup_{n\to\infty} \frac{1}{n} \log s(n, \epsilon) \leq h_\mu(T, \mathscr{P}) \leq h_\mu(T)$$

which is the desired inequality.

It remains to prove the claim. Fix $q \geq 1$. For $r < q$ define the integer j_r by

$$n - 1 - q < r + qj_r - 1 \leq n - 1.$$

We have

$$\mathscr{Q}^n = \Big(\bigvee_{j=0}^{j_r-1} T^{-jq-r}(\mathscr{Q}^q)\Big) \vee \Big(\bigvee_{i=0}^{r-1} T^{-i}(\mathscr{Q})\Big) \vee \Big(\bigvee_{i=qj_r+r}^{n-1} T^{-i}(\mathscr{Q})\Big).$$

Hence

$$\begin{aligned} H_{\nu_n}(\mathscr{Q}^n) &\leq \sum_{j=0}^{j_r-1} H_{\nu_n}(T^{-jq-r}(\mathscr{Q}^q)) + \sum_{i=0}^{r-1} H_{\nu_n}(T^{-i}(\mathscr{Q})) + \sum_{i=qj_r+r}^{n-1} H_{\nu_n}(T^{-i}(\mathscr{Q})) \\ &\leq \sum_{j=0}^{j_r-1} H_{\nu_n}(T^{-jq-r}(\mathscr{Q}^q)) + 2q \log \#J. \end{aligned}$$

Taking the sum over r gives

$$qH_{\nu_n}(\mathscr{Q}^n) \leq \sum_{i=0}^{n-1} H_{\nu_n}(T^{-i}(\mathscr{Q}^q)) + 2q^2 \log \#J.$$

In order to obtain the claim, we use the concavity of the function $\phi(t) = -t\log t$. If $\mathscr{S} = (S_k)_{k\in K}$ is a partition, then

$$\begin{aligned} H_{\mu_n}(\mathscr{S}) &= \sum_{k\in K} \phi(\mu_n(S_k)) \\ &= \sum_{k\in K} \phi\Big(\frac{1}{n}\sum_{i=0}^{n-1}(T^i)_*(\nu_n)(S_k)\Big) \\ &\geq \sum_{k\in K} \frac{1}{n}\sum_{i=0}^{n-1} \phi\Big((T^i)_*(\nu_n)(S_k)\Big) \\ &= \frac{1}{n}\sum_{i=0}^{n-1} H_{(T^i)_*(\nu_n)}(\mathscr{S}) \\ &= \frac{1}{n}\sum_{i=0}^{n-1} H_{\nu_n}(T^{-i}(\mathscr{S})). \end{aligned}$$

It is enough to apply the above property to $\mathscr{S} = \mathscr{Q}^q$. □

Definition 1.3.7. A measure μ in $\mathscr{M}_T$ is called *measure of maximal entropy* if it satisfies $h_\mu(T) = h(T)$.

An important problem is to study the existence and uniqueness of such a measure. It is possible that there is no measure of or that there are infinitely many. We will give below a partial answer of the above question.

Definition 1.3.8. The map T is said to be *expansive* if there is a constant $\epsilon_0 > 0$ such that for all distinct points x, y in X we have

$$\operatorname{dist}(T^n(x), T^n(y)) \geq \epsilon_0$$

for some $n \geq 0$.

Proposition 1.3.9. *Assume that T is expansive. Then its topological entropy is finite and T admits an invariant probability measure of maximal entropy.*

Proof. Let $\epsilon_0 > 0$ be the constant in Definition 1.3.8. Let $\mathscr{U}^{\epsilon_0/2}$ be the family of all balls of radius $\epsilon_0/2$. Since T is expansive and X is compact, one can check that $\mathscr{U}^{\epsilon_0/2}$ is a generator. Indeed, define

$$t_n(x) = \sup\{\operatorname{dist}(y, x), \text{ for } y \in B_n(x, \epsilon_0/2)\}.$$

This function is continuous and decreases to 0 when n tends to infinity. It follows that t_n converges uniformly to 0. We deduce that for any $\epsilon > 0$, if n is large enough, $\bigvee_{i=0}^{n-1} T^{-i}(\mathscr{U}^{\epsilon_0/2})$ is finer than $\mathscr{U}^\epsilon$. As a consequence, $\mathscr{U}^{\epsilon_0/2}$ is a generator.

Hence, by Proposition 1.2.3

$$h(T) = \lim_{n\to\infty} \frac{1}{n} \log N\Big(\bigvee_{i=0}^{n-1} T^{-i}(\mathscr{U}^{\epsilon_0/2})\Big) \leq \limsup_{n\to\infty} \frac{1}{n} \log s(n, \epsilon_0/2).$$

By Proposition 1.3.6, there is a measure μ with entropy larger or equal to the last limit. It should be a measure of maximal entropy. We have moreover that

$$N(\bigvee_{i=0}^{n-1} \mathscr{U}^{\epsilon_0/2}) \leq N(\mathscr{U}^{\epsilon_0/2})^n.$$

It follows that $h(T)$ is bounded by $N(\mathscr{U}^{\epsilon_0/2})$ which is finite. □

Example 1.3.10. Let T be an affine map on the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ given by $x \mapsto px + \alpha$ where p is an integer such that $|p| \geq 2$ and $\alpha \in \mathbb{R}$. Then T is expansive. One can show that the normalized Lebesgue measure on $\mathbb{T}$ has maximal entropy $\log |p|$.

Example 1.3.11. Consider the Bernouilli one-sided shift $\sigma : A^{\mathbb{N}} \to A^{\mathbb{N}}$ where A is an alphabet of cardinal $p \geq 2$. This map is expansive and we can show that the equilibrium measure has maximal entropy $\log p$. Recall that here the equilibrium measure μ is defined by

$$\mu(\{(x_n)_{n\geq 0},\ x_i = a_i \text{ for } i \leq m\}) = p^{-m-1}$$

for all $m \geq 0$ and $a_i \in A$.

The following deep theorem is due to Newhouse. The proof uses a result due to Yomdin.

Theorem 1.3.12. *Assume that T is a smooth map on a compact manifold. Then the function $\mu \mapsto h_\mu(T)$ is upper semi-continuous on $\mathscr{M}_T$. In particular, T admits an invariant probability measure with maximal entropy.*

Recall that the above function is upper semi-continuous if for any constant c the set $\{\mu \in \mathscr{M}_T,\ h_\mu(T) \geq c\}$ is closed in $\mathscr{M}_T$. The second assertion in Theorem 1.3.12 is a consequence of the first one because $\mathscr{M}_T$ is compact and $h(T)$ is finite.

1.4 Entropy and action on homology

In this section, assume that X is an oriented compact manifold. The map T induces a linear map T_* on the homology group

$$T_* : H_p(X, \mathbb{R}) \to H_p(X, \mathbb{R}) \quad \text{with } 0 \leq p \leq d = \dim X.$$

In fact, consider a continuous map $\pi : Z \to X$ defined on an oriented compact manifold Z of dimension p. It defines a homology class c in $H_p(X, \mathbb{R})$. Then,

$T_*(c)$ is defined by $T \circ \pi$. It is not difficult to see that the construction induces a well-defined map T_* on $H_p(X, \mathbb{R})$. We also have $(T^n)_* = (T_*)^n$.

We recall here some properties of linear operators on finite dimensional vector spaces. Let E be a real or complex vector space and $L : E \to E$ be a linear map. If $\| \ \|$ is a norm on E, the *spectral radius* of L is defined by the formula

$$\rho(L) = \lim_{n\to\infty} \|L^n\|^{1/n}.$$

Since the dimension of E is finite, all the norms on E are equivalent. Therefore, the above limit does not depend on the choice of the norm. When E is complex, using a suitable basis of E, we obtain that L is associated with a Jordan matrix M. We easily see that the above limit exists and is equal to the maximal modulus of the eigenvalues of M. If E is real, we can complexify the map L and reduce the problem to the previous case.

Denote by $\rho_p(T)$ the spectral radius of the map $T_* : H_p(X, \mathbb{R}) \to H_p(X, \mathbb{R})$. There are two special cases to consider first. For $p = 0$, since the image of a point is a point, we deduce that $T_* = \mathrm{id}$ on $H_0(X, \mathbb{R})$ and hence $\rho_0(T) = 1$. When $p = d$, since the dimension of $H_d(X, \mathbb{R})$ is equal to 1, T_* is just a multiplication by a constant. This constant is moreover an integer because T_* preserves $H_d(X, \mathbb{Z}) \simeq \mathbb{Z}$. It is called *the degree* of T which is positive when T preserves the orientation of X. We have $\deg(T \circ S) = \deg(T) \deg(S)$ and $\deg(T) = \pm 1$ if T is an homeomorphism. More generally, we have $\rho(T^n) = \rho(T)^n$ for $n \geq 0$.

We will discuss the following conjecture due to Shub which is still open.

Conjecture 1.4.1. *For $\mathscr{C}^1$ maps*

$$h(T) \geq \max_p \log \rho_p(T).$$

Note that in general ρ_p is easier to compute or to estimate than $h(T)$. We have the following result due to Misiurewicz and Przytycki which is false in general when T is only continuous.

Theorem 1.4.2. *For $\mathscr{C}^1$ maps*

$$h(T) \geq \log \deg(T).$$

The following result was proved by Manning.

Theorem 1.4.3. *We have for continuous maps*

$$h(T) \geq \log \rho_1(T).$$

Proof. Fix a Riemannian metric on X. Fix also a constant $\delta_0 > 0$ such that every ball of radius $4\delta_0$ is convex with respect to the geodesics. Choose $0 < \delta < \delta_0$ such that every ball of radius δ is sent by T into a ball of radius δ_0. Choose a smooth closed curve u in X such that its class c in $H_1(X,\mathbb{R})$ satisfies

$$\lim_{n\to\infty}\|(T^n)_*(c)\|^{1/n}=\rho_1(T).$$

We will show that $T^n(u)$ is homotopic to a curve u_n of length $\lesssim r(n+1,\delta)$. This will give the result because

$$\log\rho_1(T)\leq\limsup_{n\to\infty}\frac{1}{n}\log\mathrm{length}(u_n)\leq\limsup_{n\to\infty}\frac{1}{n}\log r(n+1,\delta)\leq h(T),$$

where the last inequality is a consequence of Corollary 1.2.4. We also used here the property that the norm of the class of u_n is bounded by a constant times the length of u_n. This can be seen as a consequence of the Poincaré duality.

Claim. Let u be a connected curve of length l. Then we can find

$$m\leq(l\delta^{-1}+1)r(n+1,\delta)$$

points $A_0,\ldots,A_{m-1}$ and m curves v_k of length $\leq\delta$ such that

1. v_k joins A_k and A_{k+1} and u is the sum of v_k;

2. A_k and A_{k+1} belong to a same Bowen $(n+1,\delta)$-ball.

Assuming the claim, we complete the proof of the theorem. Define v_k^i the shortest geodesic joining $T^i(A_i)$ and $T^i(A_{i+1})$. Their lengths are $\leq 2\delta_0$. We see that v_k^0 is homotopic to v_k. The choice of δ also implies that v_k^1 is homotopic to $T(v_k)$. By induction, we obtain that v_k^n is homotopic to $T^n(v_k)$ and its length is $\leq 2\delta_0$. So, it is enough to define u_n as the sum of the v_k^n. It is homotopic to $T^n(u)$ and its length is bounded by a constant times $r(n+1,\delta)$. This completes the proof.

It remains to prove the claim. Since we can divide u into curves of length at most equal to δ, it is enough to consider the case $l=\delta$ and $m\leq r(n+1,\delta)$. Recall that we can cover X by $r=r(n+1,\delta)$ Bowen $(n+1,\delta)$-balls $B_1,\ldots,B_r$. So, u is covered by these balls. Suppose that u joins two points A and A'. Define $A_0=A$. We can assume that B_1 contains A_0. If A' belongs to B_0 the claim is true for $m=1$. Otherwise, denote by u_1' the largest arc of u joining a point A_1' to A' which contains no point of B_1. Therefore, u_1' is covered by $B_2,\ldots,B_r$. If A_1 is a point in $u\setminus u_1'$ close enough to A_1 and u_1 is the arc joining A_1 and A', then u_1 is also covered by $B_2,\ldots,B_r$ and A_1 is a point in B_1. The construction of A_1 is finished. We only have to repeat it in order to construct $A_2,\ldots,A_m$. $\square$

Finally, we have the following deep result due to Yomdin.

Theorem 1.4.4. *Assume that T is of class $\mathscr{C}^\infty$. Let $Z \subset X$ be a manifold which is smooth up to the boundary. Then*

$$h(T) \geq \limsup_{n\to\infty} \frac{1}{n} \log \mathrm{volume}(T^n(Z)),$$

where the points of $T^n(Z)$ are counted with multiplicity. Here, the volume means the p-dimensional Hausdorff measure.

We deduce the following result.

Corollary 1.4.5. *Conjecture 1.4.1 is true for smooth (i.e. $\mathscr{C}^\infty$) maps.*

Proof. Fix a norm on $H_p(X, \mathbb{R})$. Choose a class c in this space which is represented by a subvariety Z of dimension p such that $\|T^n_*(c)\| \geq \rho_p(T)^n$. The class $T^n_*(c)$ is represented by $T^n(Z)$. So, using Poincaré's duality, we can choose a smooth p-form φ on X such that

$$\int_{T^n(Z)} \varphi \geq \rho_p(T)^n.$$

The left-hand side member of the last inequality is smaller than a constant times the volume of $T^n(Z)$. The constant is independent of n. Therefore, Yomdin's theorem implies that $h(T) \geq \rho_p(T)$. This gives the corollary. $\square$

1.5 Some other results and notions

We always assume that X is a metric compact space. Let μ be an invariant probability measure. Following Brin-Katok, we can define for $x \in X$ the following quantities

$$h^+_\mu(x) := \sup_{\epsilon>0} \limsup_{n\to\infty} -\frac{1}{n} \log \mu(B_n(x,\epsilon))$$

and

$$h^-_\mu(x) := \sup_{\epsilon>0} \liminf_{n\to\infty} -\frac{1}{n} \log \mu(B_n(x,\epsilon)).$$

We call them *the upper and lower local entropies* of T at x. They give the asymptotic measure of the Bowen ball $B_n(x,\epsilon)$ as n tends to infinity.

The following result is due to Brin-Katok.

Theorem 1.5.1. *We have $h^+_\mu = h^-_\mu$ μ-almost everywhere. Denote by h_μ this function. Then we have*

$$h_\mu(T) = \int h_\mu(x) d\mu(x).$$

Moreover, h_μ is an invariant function, i.e. $h_\mu \circ T = h_\mu$ μ-almost everywhere, and if μ is ergodic then $h_\mu = h_\mu(T)$ μ-almost everywhere.

Note also that the function $\mu \mapsto h_\mu(T)$ is affine on the convex compact set of invariant probability measures.

1.6 Exercises

We will give in this section some exercises. We always assume that $T : X \to X$ is a continuous map on a compact space.

Exercise 1.6.1. *Let* (v_n) *be a sub-additive sequence of positive numbers. Show that* $\frac{1}{n}v_n$ *converge to* $\inf_{n\geq 1}\frac{1}{n}v_n$.

Exercise 1.6.2. *Show that the family of all finite covers is generating. Assume that* X *is a metric space. For* $n \geq 1$*, let* $\mathscr{U}^n$ *be a cover by open sets of diameter* $\leq 1/n$*. Show that* $(\mathscr{U}^n)$ *is a generating family.*

Exercise 1.6.3. *Assume that* X *is metric. If* $\mathscr{U}$ *is an open cover of* X*, define the diameter of* $\mathscr{U}$ *by*

$$\mathrm{diam}(\mathscr{U}) = \sup_{U\in\mathscr{U}} \mathrm{diam}(U).$$

Show that $(\mathscr{U}^n)_{n\geq 1}$ *is generating if and only if* $\inf_{n\geq 1}\mathrm{diam}\mathscr{U}^n = 0$*. Assume moreover that* $\mathscr{U}^n \preceq \mathscr{U}^{n+1}$ *for every* n*. Deduce that* $(\mathscr{U}^n)$ *is generating if and only if* $\mathrm{diam}(\mathscr{U}^n) \to 0$ *as* $n \to \infty$*. Deduce that in this case, we have*

$$h(T) = \lim_{n\to\infty} h(T, \mathscr{U}^n).$$

Exercise 1.6.4. *If* T *is a homeomorphism of* $[0, 1]$*, show that* $h(T) = 0$.

Exercise 1.6.5. *Compute the entropy of the map* $T(z) = z^2$ *on the Riemann sphere.*

Exercise 1.6.6. *Let* T *and* S *be two continuous maps on compact metric spaces. Define* $R = (T, S)$*. Show that*

$$h(R) = h(T) + h(S).$$

Exercise 1.6.7. *Let* T *be the multiplication by* 2 *on a torus. Compute* $h(T)$.

Exercise 1.6.8. *Find a homeomorphism* T *such that* $h(T)$ *is infinite.*

Exercise 1.6.9. *Prove the last assertion of Theorem 1.5.1. Show that the set* $\{h^+_\mu > h_\mu(T)\}$ *has zero* μ*-measure.*

Exercise 1.6.10. *Assume that* X *is compact and* T *admits an invariant probability measure with maximal entropy. Show that* T *admits such a measure which is moreover ergodic.*

Exercise 1.6.11. *Let* $T : X \to X$ *be an expansive continuous map on a compact space.*

1. *Show that* $h(T)$ *is finite.*
2. *Show that the set* $\mathrm{Per}_n(T)$ *of periodic points of period* $n \geq 1$ *is finite.*

3. *Show that*
$$h(T) \geq \limsup_{n\to\infty} \frac{1}{n} \log \#\mathrm{Per}_n(T).$$

4. *If T is an homeomorphism, show that X is finite and $h(T) = 0$. Hint: choose ϵ small enough and N such that $t_N < \epsilon$. Show that if B is a Bowen $(N, \epsilon_0/2)$-ball then $T^{-n}(B)$ is contained in a Bowen $(N+k, \epsilon_0/2)$-ball. Consider maximal connected families of Bowen $(N, \epsilon_0/2)$-balls and show that each of them contains only one point.*

Exercise 1.6.12. *Show that T is expansive if and only if it admits an open cover which is a generator.*

Exercise 1.6.13. *Let $T : X \to X$ be a continuous map on a compact space. Assume there is a finite family of closed sets F_i such that $X = \cup F_i$ and $T(F_i) \subset F_i$. Show that*
$$h(T) = \sup_i h(T_{|F_i}).$$
Is the result true when X is metric and the family F_i is countable ?

Exercise 1.6.14. *Let α be in $\mathbb{T}^1$. Compute the topological entropy of*
$$F : \mathbb{T}^2 \to \mathbb{T}^2$$
$$(x, y) \mapsto (x + y, y + \alpha).$$

Chapter 2

Some basic dynamical systems

In this chapter, we will study some fundamental dynamical systems. We will consider in particular the questions of the topological entropy and invariant measures of maximal entropy.

2.1 Subshifts of finite type

We consider here the case of the two-sided shift $\sigma : X \to X$ where

$$X = \{1, \ldots, p\}^{\mathbb{Z}}.$$

The case of one-sided shift can be treated using the same approach. Consider a square matrix $A = (A_{i,j})_{1\leq i,j\leq p}$ whose entries are equal to 0 or 1. The set

$$X_A = \{I = (i_k)_{k\in\mathbb{Z}} \in X,\ A_{i_k,i_{k+1}} = 1 \text{ for every } k \in \mathbb{Z}\}$$

is closed in X and is invariant by σ. We will study the restriction σ_A of the map σ to X_A.

Assume that no line and no column of A vanishes identically. Otherwise, the study can be reduced to the case with a smaller p. This conditions is equivalent to the property that for every $i \in \{1, \ldots, p\}$ and $k_0 \in \mathbb{Z}$, the set X_A contains an element $(i_k)_{k\in\mathbb{Z}}$ such that $i_{k_0} = i$.

We call *word of length* n any sequence $w = (i_k)_{0\leq k\leq n}$ with elements in $\{1, \ldots, p\}$ and we says that w joins i_0 and i_n. Such a word is *admissible* if it can be extended to a word in X_A. This property is equivalent to the property that $A_{i_k,i_{k+1}} = 1$ for $k = 0, \ldots, n-1$. The cylinders which generate the topology of X_A are *admissible cylinders* which are defined by

$$C_w^{k_0} = \{I' = (i'_k)_{k\in\mathbb{Z}} \in X_A,\ i'_{k+k_0} = i_k \text{ for every } k = 0, \ldots, n\},$$

where w is an admissible word and $k_0 \in \mathbb{Z}$. The basis of this cylinder is $\{k_0, \ldots, n + k_0\}$.

We now compute the topological entropy of σ_A. We have the following lemma.

Lemma 2.1.1. *For every i, j in $\{1, \ldots, p\}$ the number of admissible words of length $n \geq 1$ joining i and j is equal to the coefficient $A^n_{i,j}$ of index (i, j) of the matrix A^n.*

Proof. A direct computation gives

$$A^n_{i,j} = \sum_{1 \leq i_k \leq p} A_{i,i_1} A_{i_1,i_2} \ldots A_{i_{n-1},j}.$$

Observe that $A_{i,i_1} A_{i_1,i_2} \ldots A_{i_{n-1},j}$ is equal to 1 when $(i, i_1, \ldots, i_{n-1}, j)$ is admissible and is equal to 0 otherwise. The lemma follows. □

Proposition 2.1.2. *The map σ_A is positively transitive if and only if for all i, j in $\{1, \ldots, p\}$, there is an $n \geq 1$ such that $A^n_{i,j} > 0$. In this case, we say that A is irreducible.*

Proof. Consider the cylinders

$$C^0_i = \{I \in X_A,\ i_0 = i\}.$$

If σ_A is positively transitive, then for all i, j in $\{1, \ldots, p\}$, there is $n \geq 1$ such that

$$C^0_i \cap \sigma_A^{-n}(C^0_j) \neq \varnothing.$$

Therefore, there is an admissible word of length n joining i and j. The previous lemma implies that $A^n_{i,j} > 0$.

Assume now that for all i, j there is $n \geq 1$ such that $A^n_{i,j} > 0$. Consider two cylinders $C^{k_0}_w$ and $C^{k'_0}_{w'}$. We want to show that there is $k \geq 0$ such that $C^{k_0}_w \cap \sigma_A^{-k}(C^{k'_0}_{w'}) \neq \varnothing$. Replacing the above cylinders with smaller ones, we can assume that they have the same basis $\{-N, \ldots, N\}$. Observe that there is an admissible word w'' of length $n \geq 1$ joining the last coordinate of w to the first coordinate of w'. Therefore, the word $ww''w'$ is admissible. Hence, $C^{k_0}_w \cap \sigma_A^{-2N-n}(C^{k'_0}_{w'}) \neq \varnothing$. This completes the proof. □

Theorem 2.1.3. *The entropy of σ_A is equal to $\log \rho(A)$ where $\rho(A)$ is the spectral radius of A.*

Proof. Consider the open cover $\mathscr{U}$ of X_A by cylinders $(C^0_i)_{1 \leq i \leq p}$. As in the case of two-sided shift, we see that the familly

$$\Big(\bigvee_{i=-n+1}^{n-1} \sigma_A^{-i}(\mathscr{U}) \Big)_{n \geq 0}$$

is generating. It follows that

$$h(\sigma_A) = \lim_{n\to\infty} h\Big(\sigma_A, \bigvee_{i=-n+1}^{n-1} \sigma_A^{-i}(\mathscr{U})\Big) = \lim_{n\to\infty} h\Big(\sigma_A, \bigvee_{i=0}^{2n-2} \sigma_A^{-i}(\mathscr{U})\Big) = h(\sigma_A, \mathscr{U}).$$

Observe now that the cover $\bigvee_{i=0}^{n-1} \sigma_A^{-i}(\mathscr{U})$ is equal to the cover by admissible cylinders C_w^0 with basis $\{0, \ldots, n-1\}$. These cylinders are mutually disjoint. The cardinal of this cover is equal to

$$N\Big(\bigvee_{i=0}^{n-1} \sigma_A^{-i}(\mathscr{U})\Big) = \sum_{i,j} A_{i,j}^{n-1}.$$

Consider on the space of matrices the norm which is equal to the sum of the absolute values of the entries. We then obtain from the above discussion

$$h(\sigma_A) = \lim_{n\to\infty} \frac{1}{n} \log \|A^{n-1}\| = \log \rho(A).$$

This completes the proof. □

2.2 Markov and Parry measures

In tis section, we will construct for σ_A an invariant probability measure of maximal entropy. We then show that it is the unique probability measure of maximal entropy. We will need the following theorem of Perron-Frobenius.

Theorem 2.2.1. *Let A be a square matrix of rank p with positive entries. Assume that there is $N \geq 1$ such that the entries of A^N are strictly positive. Then*

1. *The matrix A admits an eigenvector with strictly positive coefficients. Moreover, any eigenvector of A with positive coefficients is colinear to v.*

2. *The eigenvalue λ associated with v is simple, strictly positive and the other eigenvalues satisfy $|\lambda'| < \lambda$.*

Proof. Define

$$C = \{(x_1, \ldots, x_p) \in \mathbb{R}^p,\ x_i \geq 0\}.$$

Fix also an affine hyperplane H such that $H \cap C$ is a non-empty convex set which does not contain 0. Since the entries of A are positive, A sends C to C. So, it induces a map $A_H : H \to H$ which sends $H \cap C$ into $H \cap C$. Fixed points of A_H correspond to eigenvectors of A in C. Since A^N has only strictly positive entries, any vector in $C \setminus \{0\}$ is sent by A^N to the interior of C. Therefore, A_H^N sends $H \cap C$ to the interior of $H \cap C$.

We introduce a distance on the interior of $H \cap C$. Recall that if a, b, c, d are four points in an affine line, their cross-ratio is given by

$$[a, b, c, d] = \frac{(d-a)(c-b)}{(c-a)(d-b)}.$$

This quantity does not depend of the choice of the coordinate on the line. It is also invariant by homography.

If x, y are two points in the interior of $H \cap C$, the line joining x, y intersects the boundary of C at two points denoted by a on the side of x and b on the side of y. We have $[a, b, x, y] > 1$. Define

$$d(x, y) = \log[a, b, x, y] = \log \frac{(y-a)(x-b)}{(x-a)(y-b)}.$$

Define also $d(x, y) = 0$ when $x = y$. This function is continuous and symmetric on x, y. We can show that it defines a distance, that is, the triangle inequality is satisfied, but we don't need this property here.

We show that

$$d(A_H(x), A_H(y)) \leq d(x, y).$$

Define $x' = A_H(x)$, $y' = A_H(y)$ and a', b' in the boundary of $H \cap C$ constructed as above but for x', y'. Since A has positive entries, $A_H(a)$ is between a', x' and $A_H(b)$ is between y', b'. Therefore,

$$[a, b, x, y] = [A_H(a), A_H(b), x', y'] \leq [a', b', x', y'].$$

Hence,

$$d(A_H(x), A_H(y)) \leq d(x, y)$$

with equality only when $x = y$ or $A_H(a) = a'$, $A_H(b) = b'$. Since the entries of A^N are strictly positive, we also have for $x \neq y$

$$d(A_H^N(x), A_H^N(y)) < d(x, y).$$

The above properties imply that A_H admits a unique fixed point in $H \cap C$ that we denote by v. The coefficients of v are strictly positive and $A_H^n(w) \to v$ for every $w \in H \cap C$. This proves the first part of the theorem.

Let λ denote the eigenvalue of A associated with v. Clearly, λ is strictly positive. We prove now the second part of the theorem. We apply the first part to the transposed A^t of the matrix A. Let f denote the unique eigenvector of A^t with positive coefficients and such that $\sum f_i v_i = 1$. Let λ^* be the associated eigenvalue. Consider f as a linear form on $\mathbb{R}^p$. We have $f \circ A = \lambda^* f$ and $f(v) = 1$. Applying the first identity to v yields $\lambda^* = \lambda$. The hyperplane $\ker(f)$ is invariant by A and its intersection with C is reduced to $\{0\}$. Therefore, we only have to check that the spectral radius of A restricted to $\ker(f)$ is strictly smaller than λ.

The first part of the proof can be applied to $H = v + \ker(f)$. The map A_H is then given by

$$A_H(v + w) = v + \frac{1}{\lambda} A(w).$$

Fix a constant $\epsilon > 0$ small enough and an integer n large enough. Then

$$U_\epsilon = \{v + w,\ w \in \ker(f),\ \|w\| \leq \epsilon\}$$

is contained in $H \cap C$. Moreover, $A_H^n(U_\epsilon) \subset U_{\epsilon/2}$ as we have seen in the first part. Hence, if $\|w\| \leq \epsilon$, we have $\|A^n(w)\| \leq \frac{\epsilon}{2}\lambda^n$. It follows that the spectral radius of A restricted to $\ker(f)$ is smaller or equal to $2^{-1/n}\lambda$. This completes the proof. □

Remarks 2.2.2. (a) For i, j in $\{1, \dots, p\}$, we have

$$\lim_{n\to\infty} \frac{1}{\lambda^n} A^n_{i,j} = v_i f_j.$$

Indeed, we have

$$\lim_{n\to\infty} \frac{1}{\lambda^n} A^n = v f^t$$

since both sides give linear maps which fix v and vanish on $\ker(f)$.

(b) If the entries of A are 0 or 1, then $\lambda > 1$. Indeed,

$$p\lambda^N > \mathrm{trace}(A^N) \geq p.$$

(c) If A is a *stochastic matrix*, i.e. $\sum_i A_{i,j} = 1$ for any j, then $\lambda = 1$. Indeed, A and A_H are equal when H is given by

$$H = \{(x_1, \dots, x_p) \in \mathbb{R}^p,\ x_1 + \cdots + x_p = 1\}.$$

The vector v is then fixed by A. We have $f = (1, \dots, 1)$ and $\lim_{n\to\infty} A^n_{i,j} = v_i$.

We will consider some measures invariant by the shift σ which are called *Markov measures*. Let $M = (M_{i,j})$ be a stochastic matrix of rank p (we do not assume here that M^N has strictly positive coefficients for some N). We have seen that the hyperplane

$$H = \{(x_1, \dots, x_p) \in \mathbb{R}^p,\ x_1 + \cdots + x_p = 1\}$$

is invariant by M and the convex compact set $H \cap C$ is invariant by $M_{|H}$. We see in the proof of the above theorem that there is at least one vector v in $H \cap C$ which is invariant by M (the uniqueness requires an extra condition, e.g. the entries of M^N are strictly positive for some N). Observe also that the vector $(1, \dots, 1)$ is fixed by M^t.

Proposition 2.2.3. *There is a unique Borel probability measure μ such that for every $w = (i_0, \dots, i_m) \in \{1, \dots, p\}^{m+1}$ and every $k_0 \in \mathbb{Z}$, we have*

$$\mu(C_w^{k_0}) = \Big(\prod_{k=0}^{m-1} M_{i_k, i_{k+1}}\Big) v_{i_m}.$$

Moreover, this measure is invariant by σ.

Proof. If $w = (i_0, \ldots, i_m)$ and $i \in \{1, \ldots, p\}$, write

$$iw = (i, i_0, \ldots, i_m) \quad \text{and} \quad wi = (i_0, \ldots, i_m, i).$$

Let μ denote the above function. Since v belongs to H, we have

$$\sum_{i=1}^{p} \mu(C_i^{k_0}) = \sum_{i=1}^{p} v_i = 1.$$

Moreover, for any w as above and $k_0 \in \mathbb{Z}$, we have

$$\mu(C_w^{k_0}) = \sum_{i=1}^{p} \mu(C_{iw}^{k_0-1}) = \sum_{i=1}^{p} \mu(C_{wi}^{k_0}).$$

This is a consequence of the properties that M is stochastic and that v is an eigenvector associated to the eigenvalue 1.

The Carathéodory extension theorem says that there is a unique measure, still denoted by μ, satisfying the above relations.

It is clear that μ is invariant. The coefficient $M_{i,j}$ is the probability of being in C_i^0 at the times $k-1$ when we are in C_j^0 at the times k. □

Proposition 2.2.4. *If all the entries of M^N are strictly positive for some integer $N \geq 1$, then μ is mixing.*

Proof. We only have to show that

$$\lim_{n\to\infty} \mu(C_w^{k_0} \cap \sigma^{-n}(C_{w'}^{k_0'})) = \mu(C_w^{k_0})\mu(C_{w'}^{k_0'})$$

for all cylinders $C_w^{k_0}$ and $C_{w'}^{k_0'}$. Write $w = (i_0, \ldots, i_m)$ and $w' = (i_0', \ldots, i_{m'}')$. If n is large enough, the intersection

$$C_w^{k_0} \cap \sigma^{-n}(C_{w'}^{k_0'})$$

is a cylinder. Hence, we have

$$\mu(C_w^{k_0} \cap \sigma^{-n}(C_{w'}^{k_0'})) = \Big(\prod_{k=0}^{m-1} M_{i_k,i_{k+1}}\Big) M_{i_m,i_0'}^{n-m-k_0+k_0'} \Big(\prod_{k=0}^{m'-1} M_{i_k',i_{k+1}'}\Big) v_{i_{m'}'}.$$

The last expression converges to $\mu(C_w^{k_0})\mu(C_{w'}^{k_0'})$ since by Remarks 2.2.2, we have

$$\lim_{n\to\infty} M_{i_m,i_0'}^n = v_{i_m}.$$

The result follows. □

The following proposition gives us the entropy of Markov measures.

Proposition 2.2.5. *With the above notations, we have*

$$h_\mu(\sigma) = -\sum_{i,j} M_{i,j}\log(M_{i,j})v_j.$$

Proof. Since the partition $\mathscr{P} = (C_i^0)_{1\le i\le p}$ is a generator, we have

$$h_\mu(\sigma) = h_\mu(\sigma,\mathscr{P}) = \lim_{n\to\infty} H_\mu\Big(\mathscr{P}|\bigvee_{i=1}^{n}\sigma^{-i}(\mathscr{P})\Big),$$

see [7]. We also have

$$\begin{aligned} H_\mu\Big(\mathscr{P}|\bigvee_{i=1}^{n}\sigma^{-i}(\mathscr{P})\Big) &= -\sum_{i_0,\ldots,i_n}\mu(C^0_{i_0,\ldots,i_n})\log\Big(\frac{\mu(C^0_{i_0,\ldots,i_n})}{\mu(C^1_{i_1,\ldots,i_n})}\Big)\\ &= -\sum_{i_0,\ldots,i_n}\mu(C^0_{i_0,\ldots,i_n})\log M_{i_0,i_1}\\ &= -\sum_{i_0,i_1}\mu(C^0_{i_0,i_1})\log M_{i_0,i_1}\\ &= -\sum_{i,j}\mu(C^0_{i,j})\log M_{i,j}. \end{aligned}$$

The result follows. □

Consider now a square matrix A of rank p whose entries are 0 or 1 such that if $A_{i,j} = 0$ then $M_{i,j} = 0$, i.e. $M \le cA$ for a constant c large enough. In this case, the support of μ is contained in X_A since non-admissible cylinders have zero measure. So, μ is invariant by σ_A.

Assume that there is an integer $N \ge 1$ such that all the entries of A^N are strictly positive. We say that A is *irreducible and aperiodic*. Recall that A admits an eigenvalue $\lambda > 1$ which is the only eigenvalue whose modulus is equal to the spectral radius of A. There is a particular Markov measure, called *Parry measure*, which is defined as follows.

Let v, A^t and f be as above. Consider the matrix M such that

$$M_{i,j} = \frac{A_{i,j}f_i}{\lambda f_j}\cdot$$

This matrix is stochastic since

$$\sum_{i=1}^{p} A_{i,j}f_i = \lambda f_j.$$

The associated Markov measure μ_A is called in this case *the Parry measure*. It is supported by X_A. We easily check that the positive eigenvector of M is $(f_1v_1,\ldots,f_pv_p)$. If all the entries of A are equal to 1, we obtain here the equilibrium measure of σ.

The following result says that Parry measure maximizes the entropy.

Proposition 2.2.6. *We have*

$$h_{\mu_A}(\sigma_A) = h(\sigma_A) = \log \rho(A).$$

Proof. It suffices to apply Proposition 2.2.5. We have

$$\begin{aligned} h_\mu(\sigma_A) &= -\sum_{i,j} \frac{A_{i,j} f_i}{\lambda f_j} \log\Big(\frac{A_{i,j} f_i}{\lambda f_j}\Big) v_j f_j \\ &= \sum_{i,j} (-\log(A_{i,j} f_i) + \log f_j + \log \lambda) \frac{A_{i,j} f_i v_j}{\lambda}. \end{aligned}$$

Moreover, since $A_{i,j} = 0, 1$, we have $(\log A_{i,j}) A_{i,j} = 0$ and then

$$\sum_{i,j} -\log(A_{i,j} f_i) \frac{A_{i,j} f_i v_j}{\lambda} = -\sum_{i,j} (\log f_i) \frac{A_{i,j} f_i v_j}{\lambda} = -\sum_i (\log f_i) f_i v_i$$

$$\sum_{i,j} \log f_j \frac{A_{i,j} f_i v_j}{\lambda} = \sum_j (\log f_j) f_j v_j$$

and

$$\sum_{i,j} \log \lambda \frac{A_{i,j} f_i v_j}{\lambda} = \log \lambda \sum_i f_i v_i = \log \lambda.$$

The proposition follows. □

Finally, we have the following theorem.

Theorem 2.2.7. *Parry measure μ_A is the unique invariant probability measure of σ_A which is of maximal entropy.*

Proof. Assume that there is another measure μ of maximal entropy. Since μ_A is ergodic, μ is not absolutely continuous with respect to μ_A. We deduce that there is a Borel set B such that $\mu(B) > 0$ and $\mu_A(B) = 0$. These two measures are regular. So, there is a sequence B_l such that each B_l is a finite union of cylinders and such that

$$\lim_{l\to\infty} \mu(B_l) = \mu(B) > 0 \quad \text{and} \quad \lim_{l\to\infty} \mu_A(B_l) = \mu_A(B) = 0.$$

Consider the generator $\mathscr{P} = (C_i^0)_{1\le i\le p}$. We have

$$\begin{aligned} h_\mu(\sigma_A) &= h_\mu(\sigma_A, \mathscr{P}) = \lim_{n\to\infty} \frac{1}{2n-1} H_\mu\Big(\bigvee_{i=0}^{2n-2} \sigma_A^{-i}(\mathscr{P})\Big) \\ &= \lim_{n\to\infty} \frac{1}{2n-1} H_\mu\Big(\bigvee_{i=-n+1}^{n-1} \sigma_A^{-i}(\mathscr{P})\Big). \end{aligned}$$

Observe that for every admissible cylinder $C_w^{k_0}$ associated with a word w of length m, we have

$$\mu_A(C_w^{k_0}) = \Big(\prod_{k=0}^{m-1} A_{i_k,i_{k+1}}\Big) f_{i_0} v_{i_m} \lambda^{-m} \geq f_{i_0} v_{i_m} \lambda^{-m} \geq C\lambda^{-m},$$

where $C = \inf_{i,j} f_i v_j > 0$. Now, we will use the concavity of the function $\phi(x) = -x\log x$. We have

$$-\sum_{i=1}^{p} x_i \log x_i \leq a \log\Big(\frac{p}{a}\Big)$$

for positive numbers x_i such that

$$\sum_{1\leq i\leq p} x_i = a \leq 1.$$

Fix an integer l. The choice of B_l implies that if n is large enough, B_l is a union of elements of

$$\mathscr{P}_n = \bigvee_{i=-n+1}^{n-1} \sigma_A^{-i}(\mathscr{P}).$$

Hence,

$$\begin{aligned}
H_\mu(\mathscr{P}_n) &= \sum_{P\in\mathscr{P}_n, P\subset B_l} \phi(\mu(P)) + \sum_{P\in\mathscr{P}_n, P\not\subset B_l} \phi(\mu(P)) \\
&\leq \mu(B_l)\log\Big(\frac{\#\{P\in\mathscr{P}_n, P\subset B_l\}}{\mu(B_l)}\Big) \\
&\quad +(1-\mu(B_l))\log\Big(\frac{\#\{P\in\mathscr{P}_n, P\not\subset B_l\}}{(1-\mu(B_l))}\Big) \\
&\leq \mu(B_l)\log\Big(\frac{\lambda^{2n-1}\mu_A(B_l)}{C\mu(B_l)}\Big) + (1-\mu(B_l))\log\Big(\frac{\lambda^{2n-1}(1-\mu_A(B_l))}{C(1-\mu(B_l))}\Big) \\
&= \mu(B_l)\log\Big(\frac{\mu_A(B_l)}{\mu(B_l)}\Big) + (1-\mu(B_l))\log\Big(\frac{1-\mu_A(B_l)}{1-\mu(B_l)}\Big) \\
&\quad +(2n-1)\log\lambda - \log C.
\end{aligned}$$

So, if l is large enough, the last quantity is strictly smaller than $(2n-1)\log\lambda$. This is a contradiction since

$$\log\lambda = h_\mu(\sigma_A) = \inf_{n\geq 1}\frac{1}{2n-1} H_\mu\Big(\bigvee_{i=0}^{2n-2}\sigma_A^{-i}(\mathscr{P})\Big) = \inf_{n\geq 1}\frac{1}{2n-1} H_\mu\Big(\bigvee_{i=-n+1}^{n-1}\sigma_A^{-i}(\mathscr{P})\Big).$$

This completes the proof. □

In the rest of this section, we will study the link between periodic orbits and the maximal entropy measure. We have the following lemma.

Lemma 2.2.8. *For any $n \geq 1$, we have*

$$\#\mathrm{Fix}(\sigma_A^n) = \mathrm{trace}(A^n).$$

Proof. Fixed points of σ_A^n correspond to admissible words whose first and last coordinates are equal. So, by Lemma 2.1.1, we have

$$\#\mathrm{Fix}(\sigma_A^n) = \sum_{1 \leq i \leq p} A_{i,i}^n = \mathrm{trace}(A^n).$$

□

Proposition 2.2.9. *If A is irreducible and aperiodic, then*

$$\lim_{n\to\infty} \frac{1}{n} \log \#\mathrm{Fix}(\sigma_A^n) = h(\sigma_A) = \log \rho(A).$$

Proof. Let λ_i denote the eigenvalues of A. Then

$$\#\mathrm{Fix}(\sigma_A^n) = \mathrm{trace}(A^n) = \sum \lambda_i^n = \lambda^n + o(\lambda^n).$$

Hence,

$$\lim_{n\to\infty} \frac{\#\mathrm{Fix}(\sigma_A^n)}{\lambda^n} = 1.$$

This gives the results. □

Finally, we have the following theorem. Define for $n \geq 1$

$$\mu_n = \frac{1}{\#\mathrm{Fix}(\sigma_A^n)} \sum_{I \in \mathrm{Fix}(\sigma_A^n)} \delta_I$$

where δ_I denotes the Dirac mass at I.

Theorem 2.2.10. *Assume that A is irreducible and aperiodic. Then we have*

$$\lim_{n\to\infty} \mu_n = \mu_A.$$

Proof. We only have to check that for any cylinder $C_w^{k_0}$ that

$$\lim_{n\to\infty} \mu_n(C_w^{k_0}) = \mu_A(C_w^{k_0}).$$

Write $P_n = \mathrm{Fix}(\sigma_A^n)$ and $w = (w_0, \ldots, w_m)$. We have

$$\begin{aligned} \mu_n(C_w^{k_0}) &= \frac{\#P_n \cap C_w^{k_0}}{\mathrm{trace}(A^n)} \\ &= \frac{1}{\mathrm{trace}(A^n)} \Big(\prod_{k=0}^{m-1} A_{i_k, i_{k+1}} \Big) A_{i_m, i_0}^{n-m}. \end{aligned}$$

We have seen that

$$\lim_{n\to\infty} \frac{1}{\lambda^n}\mathrm{trace}(A^n) = 1$$

and

$$\lim_{n\to\infty} \frac{1}{\lambda^n} A^n_{i_m,i_0} = v_{i_m} f_{i_0}.$$

Recall also that $(v_1 f_1, \dots, v_p f_p)$ is a positive eigenvector of M associated to the eigenvalue 1. Therefore,

$$\begin{aligned} \lim_{n\to\infty} \mu_n(C^{k_0}_w) &= \Big(\prod_{k=0}^{m-1} A_{i_k,i_{k+1}}\Big)\frac{v_{i_m} f_{i_0}}{\lambda^m} \\ &= \Big(\prod_{k=0}^{m-1} M_{i_k,i_{k+1}}\Big) v_{i_m} f_{i_m} = \mu_A(C^{k_0}_w). \end{aligned}$$

The theorem follows. □

2.3 Hyperbolic automorphisms of tori

We first consider the case of linear maps. In what follows, E is a Banach space and $\| \ \|_0$ is a fixed norm on E. We will use in particular the case where E is a finite dimensional vector space.

Definition 2.3.1. Let $u : E \to E$ be a linear endomorphism. We say that u is *hyperbolic* if there is a decomposition $E = E^s \oplus E^u$ into closed subspaces, a norm $\| \ \|$ equivalent to $\| \ \|_0$ and a number $\lambda \in]0, 1[$ such that

1. $u_{|E^s}$ is an endomorphism of E^s, i.e. $u(E^s) \subset E^s$;
2. $u_{|E^u}$ is an automorphism of E^u, in particular $u(E^u) = E^u$;
3. $\|u_{E^s}\| \le \lambda$ and $\|u^{-1}_{E^u}\| \le \lambda$;
4. for every $x \in E$, we have $\|x\| = \max(\|x^s\|, \|x^u\|)$, where $x = x^s + x^u$ is the decomposition of x in $E^s \oplus E^u$.

We say that $E = E^s \oplus E^u$ is a *hyperbolic decomposition*, that $\| \ \|$ is an adapted norm and that u is λ-*hyperbolic*.

The proof of the following proposition is left to the reader.

Proposition 2.3.2. *Let u be a hyperbolic endomorphism as above. Then E^s is the set of points x such that* $\lim_{n\to\infty} u^n(x) = 0$ *and E^u is the set of points x such that there is a sequence $x_n \to 0$ such that $u^n(x_n) = x$. If u is an automorphism, then u^{-1} is also hyperbolic and E^u is the set of points x such that* $\lim_{n\to\infty} u^{-n}(x) = 0$.

The proposition shows that the hyperbolic decomposition is unique.

Remark 2.3.3. When the dimension of E is finite, u is hyperbolic if and only if no eigenvalue of u has modulus 1. The eigenvalues of u restricted to E^s have modulus strictly smaller than 1 and the eigenvalues of u restricted to E^u have modulus strictly larger than 1.

We now study hyperbolic automorphisms on tori. Let A be a hyperbolic automorphism on $\mathbb{R}^r$ which is given by a hyperbolic matrix with integer entries and with determinant ± 1. So, A induces an automorphism on $\mathbb{T}^r$ that we denote by $\widehat{A}$. Let $\mathbb{R}^r = E^s \oplus E^u$ be the hyperbolic decomposition of A. We have the following proposition.

Proposition 2.3.4. *For any $\widehat{x} \in \mathbb{T}^r$, the sets*

$$W^s(\widehat{x}) = \{\widehat{y} \in \mathbb{T}^r,\ \lim_{k\to\infty} \operatorname{dist}(\widehat{A}^k(\widehat{y}), \widehat{A}^k(\widehat{x})) = 0\}$$

and

$$W^u(\widehat{x}) = \{\widehat{y} \in \mathbb{T}^r,\ \lim_{k\to-\infty} \operatorname{dist}(\widehat{A}^k(\widehat{y}), \widehat{A}^k(\widehat{x})) = 0\}$$

are projections in $\mathbb{T}^r$ of the affine spaces $x + E^s$ and $x + E^u$ respectively, where x is a point in $\pi^{-1}(x)$ and $\pi : \mathbb{R}^r \to \mathbb{T}^r$ is the canonical projection.

Proof. We prove the result for $W^s(\widehat{x})$. The case of $W^u(\widehat{x})$ is treated in the same way. First, for $y \in x + E^s$, we have

$$\lim_{k\to\infty} \|A^k(y) - A^k(x)\| = \lim_{k\to\infty} \|A^k(y - x)\| = 0.$$

It follows that

$$\lim_{k\to\infty} \operatorname{dist}(\widehat{A}^k(\widehat{y}), \widehat{A}^k(\widehat{x})) = 0$$

since $\operatorname{dist}(\widehat{y}, \widehat{x}) \le \operatorname{dist}(x, y)$. Hence, $\pi(x + E^s) \subset W^s(\widehat{x})$.

Let $\epsilon > 0$ be a constant small enough such that the inverse image of any ball B of radius ϵ in $\mathbb{T}^r$ by π is a disjoint union of balls which are sent bijectively to B. Fix also a constant $\delta > 0$ such that $\|A(x) - A(y)\| < \epsilon$ when $\|x - y\| < \delta$. Consider a point $\widehat{y}$ in $W^s(\widehat{x})$. For n large enough, we have $\operatorname{dist}(\widehat{A}^n(\widehat{y}), \widehat{A}^n(\widehat{x})) < \delta$. The choice of δ, ϵ implies that there is a unique $k_n \in \mathbb{Z}^r$ such that

$$\|A^n(y + k_n) - A^n(x)\| = \operatorname{dist}(\widehat{A}^n(\widehat{y}), \widehat{A}^n(\widehat{x})) < \delta.$$

We deduce that

$$\operatorname{dist}(\widehat{A}^{n+1}(\widehat{y}), \widehat{A}^{n+1}(\widehat{x})) < \epsilon.$$

The uniqueness of k_n implies that $k_{n+1} = k_n$. By Proposition 2.3.2, $y + k_n$ belongs to $x + E^s$ and hence, $\widehat{y} \in \pi(x + E^s)$. This completes the proof. $\square$

Example 2.3.5. The following example is often called *Arnold's cat*. It is the simplest example of a hyperbolic automorphism on a torus. Consider the map A on $\mathbb{R}^2$ given by the matrix

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Its eigenvalues are $\frac{3+\sqrt{5}}{2}$ and $\frac{3-\sqrt{5}}{2}$. The line E^s is given by the equation $x_2 = \frac{-\sqrt{5}-1}{2}x_1$ and the line E^u is given by $x_2 = \frac{\sqrt{5}-1}{2}x_1$. The slopes are irrational. The images of $x + E^s$ and $x + E^u$ are dense in $\mathbb{T}^2$.

We have the following proposition for general hyperbolic automorphisms on tori.

Proposition 2.3.6. *For all $\widehat{x}$ and $\widehat{y}$, the sets $W^s(\widehat{x})$ and $W^u(\widehat{y})$ are dense in $\mathbb{T}^r$. Moreover, they intersect transversally in a dense countable set.*

Proof. Clearly, E^s intersects E^u transversally in one point. So, $W^s(\widehat{x})$ and $W^u(\widehat{y})$ intersect transversally. In fact, the last intersection is finite or countable. Since the directions of $W^s(\widehat{x})$ and $W^u(\widehat{y})$ are constant, if they are dense, their intersection is also dense. So, it is enough to check that $W^s(\widehat{x})$ is dense. The case of $W^u(\widehat{y})$ can be treated in the same way.

Using a translation, we can assume that $\widehat{x} = 0$. Since periodic points are dense, it is enough to show that they belong to $\overline{W}^s(0)$. We will use the property that $W^s(0)$ is invariant. Let $\widehat{a}$ be a periodic point of period q. Choose a point $\widehat{b}$ in $W^s(0) \cap W^u(\widehat{a})$. Such a point exists by Proposition 2.3.4. Now, observe that $\lim_{n\to\infty} \widehat{A}^{-nq}(\widehat{b}) = \widehat{a}$. It follows that $\widehat{a} \in \overline{W}^s(0)$ since this set is invariant and contains the points $A^{-nq}(\widehat{b})$. □

Corollary 2.3.7. *The automorphism $\widehat{A}$ is positively transitive.*

Proof. Let U and V be two open sets of $\mathbb{T}^r$. We want to show that $U \cap \widehat{A}^{-n}(V) \neq \varnothing$ for some $n \geq 0$. Replace $\widehat{A}$ by an iterate, we can assume that there is a fixed point $\widehat{a}$ in V. Now, it is enough to observe that U contains a point $\widehat{b}$ which belongs to $W^s(\widehat{a})$. Its orbit converges to $\widehat{a}$. This gives the result. □

We will now study the perturbations of the above maps. First recall Picard's fixed point theorem.

Theorem 2.3.8. *Let $T : E \to E$ be a λ-hyperbolic endomorphism as above. Define $\epsilon_0 = 1 - \lambda$. If $\varphi : E \to E$ is a Lipschitz map with Lipschitz constant $\epsilon < \epsilon_0$. Then $f = T + \varphi$ admits a unique fixed point $x \in E$. Moreover, we have*

$$\|x\| < \frac{\|\varphi(0)\|}{\epsilon_0 - \epsilon}.$$

We will need the following version with parameters.

Theorem 2.3.9. *Let X, Y be two metric spaces such that Y is complete. Let $\Phi : X \times Y \to Y$ be a continuous application. Suppose there is a constant $0 < \lambda < 1$ such that for any $x \in X$ and $y, y' \in Y$ we have*

$$\mathrm{dist}\,(\Phi(x,y), \Phi(x,y')) \leq \lambda \mathrm{dist}\,(y,y').$$

Then for every $x \in X$ there is a unique solution $y = \theta(x)$ of the equation $\Phi(x,y) = y$. Moreover, the map $\theta : X \to Y$ is continuous.

Definition 2.3.10. Let $F : X \to X$ be a continuous map on a metric space X. Let $\alpha > 0$ be a constant. We say that a sequence $(x_i)_{i\in\mathbb{Z}}$ is *an α-pseudo-orbit* of F if for any $i \in \mathbb{Z}$

$$\mathrm{dist}\,(F(x_i), x_{i+1}) < \alpha.$$

We have the following fundamental property.

Proposition 2.3.11. *Let $T : E \to E$ be a λ-hyperbolic endomorphism as above. Define $\epsilon_0 = 1 - \lambda$. Let $\varphi : E \to E$ be a Lipschitz map with Lipschitz constant $\epsilon < \epsilon_0$. Define $f = T + \varphi$. Then for any α-pseudo-orbit $(x_i)_{i\in\mathbb{Z}}$ of f there is a unique orbit $(y_i)_{i\in\mathbb{Z}}$ such that*

$$\sup_{i\in\mathbb{Z}} \|y_i - x_i\| < \infty.$$

Moreover, we have

$$\sup_{i\in\mathbb{Z}} \|y_i - x_i\| \leq \frac{\alpha}{\epsilon_0 - \epsilon}.$$

Proof. Let $\mathscr{E}$ denote the space of bounded sequences $z = (z_i)_{i\in\mathbb{Z}}$ in E, endowed with the norm

$$\|z\| = \sup_{i\in\mathbb{Z}} \|z_i\|.$$

This is a Banach space which can be decomposed as $\mathscr{E} = \mathscr{E}^s \oplus \mathscr{E}^u$, where $\mathscr{E}^s$ and $\mathscr{E}^u$ are the spaces of bounded sequences in E^s and E^u respectively. Define

$$\mathscr{T} : \mathscr{E} \to \mathscr{E}$$

$$(z_i)_{i\in\mathbb{Z}} \mapsto (T(z_{i-1}))_{i\in\mathbb{Z}}.$$

Observe that this map is λ-hyperbolic with respect to the above decomposition. Define also

$$\mathscr{F} : \mathscr{E} \to \mathscr{E}$$

$$(z_i)_{i\in\mathbb{Z}} \mapsto (\varphi(x_{i-1} + z_{i-1}) + T(x_{i-1}) - x_i)_{i\in\mathbb{Z}}.$$

This function is well-defined because the sequence

$$(\varphi(x_{i-1} + z_{i-1}) + T(x_{i-1}) - x_i)_{i\in\mathbb{Z}} = (\varphi(x_{i-1} + z_{i-1}) - \varphi(x_{i-1}) + f(x_{i-1}) - x_i)_{i\in\mathbb{Z}}$$

is bounded and is ϵ-Lipschitz. We deduce that $\mathscr{T} + \mathscr{F}$ has a unique fixed point $z = (z_i)_{i\in\mathbb{Z}}$ and that

$$\|z\| \leq \frac{\|\mathscr{F}(0)\|}{\epsilon_0 - \epsilon} \leq \frac{\alpha}{\epsilon_0 - \epsilon},$$

where the last inequality is a consequence of the property that $(x_i)_{i\in\mathbb{Z}}$ is an α-pseudo-orbit.

Define $y_i = z_i + x_i$. We obtain from the equation

$$z_i = \varphi(x_{i-1} + z_{i-1}) + T(x_{i-1}) - x_i + T(z_{i-1})$$

that $y = (y_i)_{i\in\mathbb{Z}}$ is an orbit of f satisfying

$$\|y - x\| \leq \frac{\alpha}{\epsilon_0 - \epsilon}.$$

The uniqueness of z implies also that y is the unique orbit satisfying the condition in the proposition. □

We have the following proposition on the perturbations of hyperbolic automorphisms.

Proposition 2.3.12. *Let $T : E \to E$ be a λ-hyperbolic automorphism on a Banach space as above. Define $\epsilon_1 = \min(\|T^{-1}\|^{-1}, 1 - \lambda)$. Let $\varphi : E \to E$ be a bounded map which is ϵ-Lipschitz with $\epsilon < \epsilon_1$. Then there is a unique homeomorphism $h : E \to E$ such that $h - \mathrm{id}$ is bounded and $h \circ T \circ h^{-1} = T + \varphi$.*

This result is a consequence of the following three lemmas (we only have to take $\psi = 0$ and $h = \mathrm{id} + \beta$).

Lemma 2.3.13. *The map $f = T + \varphi$ is a homeomorphism.*

Proof. Observe that $f(x) = y$ if and only if $T(x) + \varphi(x) = y$ and this property is equivalent to

$$x = T^{-1}(y) - T^{-1} \circ \varphi(x)$$

because T^{-1} is linear. The function

$$\Phi(x, y) = T^{-1}(y) - T^{-1} \circ \varphi(x)$$

is continuous and for any y the map $x \mapsto \Phi(x, y)$ is Lipschitz with Lipschitz constant $\|T^{-1}\|\epsilon$. By hypothesis, this constant is smaller than 1. So, we can apply the version with parameters of the fixed point theorem. It says that there is a continuous function $\theta : E \to E$ such that $f(x) = y$ if and only if $x = \theta(y)$. The lemma follows. □

Lemma 2.3.14. *Let φ and ψ be two bounded maps which are ϵ-Lipschitz with $\epsilon < \epsilon_1$. Then there is a unique bounded continuous map $\beta : E \to E$ such that*

$$(T + \varphi) \circ (\mathrm{id} + \beta) = (\mathrm{id} + \beta) \circ (T + \psi).$$

Proof. We know from the last lemma that $g = T + \psi$ is a homeomorphism. So, the equation

$$(T + \varphi) \circ (\mathrm{id} + \beta) = (\mathrm{id} + \beta) \circ (T + \psi)$$

can be re-written as

$$(T + \varphi) \circ (\mathrm{id} + \beta) \circ g^{-1} = \mathrm{id} + \beta$$

which is also equivalent to

$$T \circ \beta \circ g^{-1} + \varphi \circ (\mathrm{id} + \beta) \circ g^{-1} + T \circ g^{-1} - \mathrm{id} = \beta.$$

We used here that T is linear.

Define $\mathscr{E} = \mathscr{C}(E, E)$ the Banach space of continuous bounded maps $\beta : E \to E$ endowed with the norm

$$\|\beta\| = \sup_{x \in E} \|\beta(x)\|.$$

Define two maps $\mathscr{T} : \mathscr{E} \to \mathscr{E}$ and $\mathscr{F} : \mathscr{E} \to \mathscr{E}$ by

$$\mathscr{T}(\beta) = T \circ \beta \circ g^{-1}$$

and

$$\mathscr{F}(\beta) = \varphi \circ (\mathrm{id} + \beta) \circ g^{-1} + T \circ g^{-1} - \mathrm{id}.$$

Because $\mathrm{id} = g \circ g^{-1}$ and $g = T + \psi$, then

$$\mathscr{F}(\beta) = \varphi \circ (\mathrm{id} + \beta) \circ g^{-1} - \psi \circ g^{-1}.$$

Observe that $\mathscr{T}$ is linear and λ-hyperbolic with respect to the decomposition $\mathscr{E} = \mathscr{E}^s \oplus \mathscr{E}^u$, where $\mathscr{E}^s = \mathscr{C}(E, E^s)$ and $\mathscr{E}^u = \mathscr{C}(E, E^u)$. The map $\mathscr{F}$ is ϵ-Lipschitz. Since $\epsilon < \epsilon_1 < 1 - \lambda$, $\mathscr{T} + \mathscr{F}$ admits a unique fixed point. The lemma follows. $\square$

Lemma 2.3.15. *Under the hypothesis of the last lemma, the map $h = \mathrm{id} + \beta$ is a homeomorphism.*

Proof. Recall that $f = T + \varphi$ and $g = T + \psi$. The last lemma implies the existence of β, β' such that

$$f \circ (\mathrm{id} + \beta) = (\mathrm{id} + \beta) \circ g$$

and

$$(\mathrm{id} + \beta') \circ f = g \circ (\mathrm{id} + \beta').$$

So, we have

$$f \circ (\mathrm{id} + \beta) \circ (\mathrm{id} + \beta') = (\mathrm{id} + \beta) \circ (\mathrm{id} + \beta') \circ f$$

and

$$g \circ (\mathrm{id} + \beta') \circ (\mathrm{id} + \beta) = (\mathrm{id} + \beta') \circ (\mathrm{id} + \beta) \circ g.$$

Since the map

$$(\mathrm{id} + \beta) \circ (\mathrm{id} + \beta') - \mathrm{id} = \beta' + \beta \circ (\mathrm{id} + \beta')$$

is bounded and continuous, the last lemma applied to (f, f) gives

$$(\mathrm{id} + \beta) \circ (\mathrm{id} + \beta') = \mathrm{id}.$$

We obtain in the same way that

$$(\mathrm{id} + \beta') \circ (\mathrm{id} + \beta) = \mathrm{id}.$$

It follows that $\mathrm{id} + \beta$ is a homeomorphism of E and its inverse map is equal to $\mathrm{id} + \beta'$. $\square$

We will now apply the above results to the case of homeomorphisms and diffeomorphisms of tori. Recall that any continuous map $F : \mathbb{T}^r \to \mathbb{T}^r$ is homotopic to a unique linear map $\widehat{F_*} : \mathbb{T}^r \to \mathbb{T}^r$, see [7]. Denote also by F_* the lift of $\widehat{F_*}$ to $\mathbb{R}^r$. It is a linear map from $\mathbb{R}^r$ to $\mathbb{R}^r$.

Theorem 2.3.16. *Let $F : \mathbb{T}^r \to \mathbb{T}^r$ be a homeomorphism such that F_* is a hyperbolic automorphism of $\mathbb{R}^r$. Then $\widehat{F_*}$ is a factor of F. More precisely, there is a unique continuous map $H : \mathbb{T}^r \to \mathbb{T}^r$ homotopic to* id *such that $H \circ F = \widehat{F_*} \circ H$.*

Proof. Consider the hyperbolic decomposition $\mathbb{R}^r = E^s + E^u$ associated with F_*. Fix also a lift f of F to $\mathbb{R}^r$. We want to show the existence of a continuous map $\beta : \mathbb{R}^r \to \mathbb{R}^r$ which is invariant under the translations by vectors with integer coordinates (i.e. is $\mathbb{Z}^r$-periodic) and such that

$$(\mathrm{id} + \beta) \circ f = F_* \circ (\mathrm{id} + \beta).$$

This equation can be written as

$$\beta = F_* \circ f^{-1} + F_* \circ \beta \circ f^{-1} - \mathrm{id}.$$

Denote by $\mathscr{E}$ the Banach space of continuous maps $\beta : \mathbb{R}^r \to \mathbb{R}^r$ which are $\mathbb{Z}^r$-periodic. Denote also by $\mathscr{E}^s$ and $\mathscr{E}^u$ the spaces of continuous maps $\beta : \mathbb{R}^r \to E^s$ and $\beta : \mathbb{R}^r \to E^u$ respectively, which are $\mathbb{Z}^r$-periodic.

Observe that

$$\mathscr{T} : \mathscr{E} \to \mathscr{E}$$
$$\beta \mapsto F_* \circ \beta \circ f^{-1}$$

is linear and hyperbolic with the hyperbolic decomposition $\mathscr{E} = \mathscr{E}^s \oplus \mathscr{E}^u$. Since the map

$$\mathscr{F} : \mathscr{E} \to \mathscr{E}$$
$$\beta \mapsto F_* \circ f^{-1} - \mathrm{id}$$

is constant, i.e. independent of β, we deduce from the fixed point theorem the existence of a unique $\beta \in \mathscr{E}$ such that

$$\beta = F_* \circ f^{-1} + F_* \circ \beta \circ f^{-1} - \mathrm{id}.$$

We have

$$(\mathrm{id} + \beta) \circ f = F_* \circ (\mathrm{id} + \beta).$$

The application $h = \mathrm{id} + \beta$ is a lift of a continuous map $H : \mathbb{T}^r \to \mathbb{T}^r$ such that $H \circ F = \widehat{F}_* \circ H$. It remains to show that H is surjective. This is the consequence of the fact that $\deg(H) \neq 0$. $\square$

Remark 2.3.17. If F and $\widehat{F}_*$ are as above, then $h(F) \geq h(\widehat{F}_*)$. We can prove that the topological entropy of $\widehat{F}_*$ is equal to $\max \rho_p(\widehat{F}_*)$. So, Shub conjecture holds for homeomorphisms on tori.

Let $\widehat{A} : \mathbb{T}^r \to \mathbb{T}^r$ be a hyperbolic linear automorphism on $\mathbb{T}^r$ induced by a linear map $A : \mathbb{R}^r \to \mathbb{R}^r$. We will show that it is $\mathscr{C}^1$-*structurally stable*. That is, if $F : \mathbb{T}^r \to \mathbb{T}^r$ is a map close enough to $\widehat{A}$ for the $\mathscr{C}^1$-topology, then F is conjugate to $\widehat{A}$. In other words, there is $\epsilon > 0$ such that F is conjugate to $\widehat{A}$ when

$$\sup_{a \in \mathbb{T}^r} \mathrm{dist}\,(F(a), \widehat{A}(a)) \leq \epsilon$$

$$\sup_{a \in \mathbb{T}^r} \mathrm{dist}\,(F^{-1}(a), \widehat{A}^{-1}(a)) \leq \epsilon$$

$$\sup_{a \in \mathbb{T}^r} \|DF(a) - \widehat{A}\| \leq \epsilon$$

and

$$\sup_{a \in \mathbb{T}^r} \|DF^{-1}(a) - \widehat{A}^{-1}\| \leq \epsilon.$$

Here is the precise statement.

Theorem 2.3.18. *Let $A : E \to E$ be a λ-hyperbolic linear automorphism and $E = \mathbb{R}^r = E^s \oplus E^u$ be the corresponding decomposition as above. Let $F : \mathbb{T}^r \to \mathbb{T}^r$ be a $\mathscr{C}^1$-map such that*

$$\sup_{a \in \mathbb{T}^r} \mathrm{dist}\,(F(a), \widehat{A}(a)) \leq \epsilon_2 = \frac{1}{4} \min_{k \in \mathbb{Z}^r - 0} \|k\|$$

and

$$\sup_{a \in \mathbb{T}^r} \|DF(a) - A\| \leq \epsilon < \epsilon_1 = \min(\|A^{-1}\|^{-1}, 1 - \lambda).$$

Then F is a diffeomorphism of class $\mathscr{C}^1$ and there is a unique homeomorphism $H : \mathbb{T}^r \to \mathbb{T}^r$, homotopic to id, *and such that $H \circ F = \widehat{A} \circ H$. In particular, we have $h(F) = h(\widehat{A})$.*

We have seen that the first inequality in the theorem implies that $F_* = A$. Fix also a lift f of F to $\mathbb{R}^r$. We can write $f = A + \psi$ where ψ is a map of class $\mathscr{C}^1$ which is $\mathbb{Z}^r$-periodic. We first show that f and F are diffeomorphisms. For this purpose, we need the following differentiable version of the fixed point theorem with parameters.

Lemma 2.3.19. *Let E and F be two normed vector spaces. Assume that F is a Banach space. Let $\Phi : E \times F \to F$ be a map of class $\mathscr{C}^p$, $p \geq 1$, such that*

$$\sup_{x,y} \|D_2\Phi(x, y)\| = \lambda < 1.$$

Then, for every $x \in E$, there is a unique solution $y = \theta(x)$ of the equation $\Phi(x, y) = y$. Moreover, the map θ is of class $\mathscr{C}^p$ and we have

$$D\theta(x) = (\mathrm{id} - D_2\Phi(x, \theta(x)))^{-1} \circ D_1\Phi(x, \theta(x)).$$

Lemma 2.3.20. *The map $f = A + \psi$ is a diffeomorphism of class $\mathscr{C}^1$.*

Proof. Observe that $f(x) = y$ if and only if

$$A(x) + \psi(x) = y.$$

This equation is also equivalent to

$$x = A^{-1}(y) - A^{-1} \circ \psi(x).$$

The map

$$\Phi(x, y) = A^{-1}(y) - A^{-1} \circ \psi(x)$$

is of class $\mathscr{C}^1$ and the map $x \mapsto \Phi(x, y)$ is $\|A^{-1}\|\epsilon$-Lipschitz for each y fixed since

$$\sup_{x,y} \|D_2\Phi(x, y)\| \leq \|A^{-1}\|\epsilon < 1.$$

According to last lemma, we obtain a function $\theta : E \to E$ of class $\mathscr{C}^1$ such that $f(x) = y$ if and only if $x = \theta(y)$. The lemma follows. □

It is left to the reader to show that F is also a diffeomorphism.

End of the proof of Theorem 2.3.18. By Proposition 2.3.12, there is a homeomorphism $h : \mathbb{R}^r \to \mathbb{R}^r$ such that $h \circ f = A \circ h$ and such that $h - \mathrm{id}$ is bounded. Theorem 2.3.16 implies the existence of a continuous map $h' : \mathbb{R}^r \to \mathbb{R}^r$ such that $h' \circ f = A \circ h'$ and such that $h' - \mathrm{id}$ is $\mathbb{Z}^r$-periodic. We then deduce from the uniqueness that $h = h'$ and this is a lift of a map $H : \mathbb{T}^r \to \mathbb{T}^r$ which is homotopic to id. Write $h = \mathrm{id} + k$ and $h^{-1} = \mathrm{id} + l$. Since k is $\mathbb{Z}^r$-periodic, using that $h^{-1} \circ h = \mathrm{id}$, it is not difficult to see that l is also $\mathbb{Z}^r$-periodic. We then deduce that H is a homeomorphism. This map H satisfies the theorem. □

Here is a local version of the last theorem. It is due to Hartman-Grobman.

Theorem 2.3.21. *Let $f : U \to \mathbb{R}^r$ be a $\mathscr{C}^1$ map defined on an open set U of $\mathbb{R}^r$. Assume there is a point $x_0 \in U$ such that $f(x_0) = x_0$ and that the differential $T = Df(x_0)$ at x_0 is hyperbolic. Then there is a neighbourhood $V \subset U$ of x_0, a neighbourhood W of 0 and a homeomorphism $h : V \to W$ with $h(x_0) = 0$ such that for every $x \in V$, the point $f(x)$ belongs to V if and only if $T(h(x))$ belongs to W and in this case we have $f(x) = h^{-1}(T(h(x)))$.*

It is enough to apply Proposition 2.3.12 to an extension of f to $E = \mathbb{R}^r$. More precisely, we have the following easy lemma. We can assume that $x_0 = 0$.

Lemma 2.3.22. *For any $\epsilon > 0$, we can construct a map $\widetilde{f} : \mathbb{R}^r \to \mathbb{R}^r$ which is equal to f in a neighbourhood of 0 and such that $\widetilde{f} - T$ is bounded and ϵ-Lipschitz.*

Note that if f is of class $\mathscr{C}^p$, we can construct $\widetilde{f}$ of class $\mathscr{C}^p$.

2.4 Stable manifolds and hyperbolic sets

We have the following result.

Proposition 2.4.1. *Let $T : E \to E$ be a λ-hyperbolic linear endomorphism on $E = E^s \oplus E^u$ as above. Let $\varphi : E \to E$ be a bounded Lipschitz map with Lipschitz constant $\epsilon < \epsilon_0 = 1 - \lambda$ such that $\varphi(0) = 0$. Define $f = T + \varphi$. Then the set*

$$W^s(0) = \{x \in E, \quad (f^n(x))_{n \geq 0} \text{ is bounded}\}$$

is the graph of a map $\psi : E^s \to E^u$ which is $(\lambda + \epsilon)$-Lipschitz. Moreover, f is $(\lambda + \epsilon)$-Lipschitz on this graph and

$$W^s(0) = \{x \in E, \quad \lim_{n \to \infty} (f^n(x))_{n \geq 0} = 0\}.$$

Note that if φ is of class $\mathscr{C}^p$, $p \geq 1$, such that $D\varphi(0) = 0$, then ψ is also of class $\mathscr{C}^p$ and $D\psi(0) = 0$.

We can deduce from the above result the following theorem.

Theorem 2.4.2. *Let $f : U \to \mathbb{R}^r$ be a map of class $\mathscr{C}^p$, $p \geq 1$, defined in a neighbourhood of a fixed point x_0. Assume that $Df(x_0)$ is λ-hyperbolic, $0 < \lambda < 1$. Let $\mathbb{R}^r = E^s \oplus E^u$ be the corresponding hyperbolic decomposition and $\| \ \|$ the adapted norm. Let λ' be a constant such that $\lambda < \lambda' < 1$. Then there is $\delta > 0$ such that the set*

$$W^s_\delta(x_0) = \{x \in U, \ \|f^n(x) - x_0\| \leq \delta \text{ for all } n \geq 0\}$$

is the graph of a $\mathscr{C}^p$ map

$$\psi : x_0 + (E^s \cap \overline{B}(0, \delta)) \to x_0 + (E^u \cap \overline{B}(0, \delta))$$

such that

$$\psi(x_0) = x_0 \quad \textit{and} \quad D\psi(x_0) = 0.$$

Moreover, we have for every $x_0 \in W_\delta^s(x_0)$ *and every* $n \geq 0$

$$\|f^n(x) - x_0\| \leq \lambda'^n \|x - x_0\|.$$

Definition 2.4.3. The set $W_\delta^s(x_0)$ is called *the local stable manifold* of f at x_0. When all the eigenvalues of $Df(x_0)$ have modulus strictly smaller than 1, this manifold is equal to $\overline{B}(x_0, \delta)$. In this case, we say that x_0 is *an attractive fixed point*. When $Df(x_0)$ is invertible, the stable manifold $W_\delta^u(x_0)$ of f^{-1} is called *the local unstable manifold* of f at x_0. The sum of the dimension of stable manifold and the dimension of the unstable manifold is equal to r. When all the eigenvalues of $Df(x_0)$ have modulus strictly larger than 1, the local unstable manifold is equal to $\overline{B}(x_0, \delta)$ and we say that x_0 is *a repelling fixed point*. When the stable and unstable manifolds have positive dimension, x_0 is called *a saddle fixed point*.

We also have the following result on global stable manifolds.

Theorem 2.4.4. *Assume that* $f : M \to M$ *is a diffeomorphism of class* $\mathscr{C}^p$, $p \geq 1$, *on a manifold* M. *Let* x_0 *be a fixed point such that* $T_{x_0}f$ *is hyperbolic and let* $T_{x_0}M = E^s \oplus E^u$ *be the corresponding decomposition. Then, the set*

$$W^s(x_0) = \{x \in M,\ \lim_{n\to\infty} f^n(x) = x_0\}$$

is the image of an injective immersion $\theta^s : E^s \to M$ *of class* $\mathscr{C}^p$ *with* $\theta^s(0) = x_0$ *such that* $T_0\theta^s$ *is the inclusion map* $i^s : E^s \to T_{x_0}M$. *The set*

$$W^u(x_0) = \{x \in M,\ \lim_{n\to\infty} f^{-n}(x) = x_0\}$$

is the image of an injective immersion $\theta^u : E^u \to M$ *of class* $\mathscr{C}^p$ *with* $\theta^u(0) = x_0$ *such that* $T_0\theta^u$ *is the inclusion map* $i^u : E^u \to T_{x_0}M$.

Definition 2.4.5. The manifolds $W^s(x_0)$ and $W^u(x_0)$ are respectively *the (global) stable and unstable manifolds* of f at x_0. When x_0 is an attractive fixed point, $W^s(x_0)$ is an open subset of M. It is called *the attractive basin* of x_0. When x_0 is a repelling fixed point, $W^s(x_0)$ is also an open subset of M. It is called *the repelling basin* of x_0. When x_0 is a periodic point of period p such that $T_{x_0}f^p$ is hyperbolic, the stable and unstable manifolds are defined by

$$W^s(x_0) = \{x \in M,\ \lim_{n\to\infty} \operatorname{dist}(f^n(x), f^n(x_0)) = 0\}$$

and

$$W^u(x_0) = \{x \in M,\ \lim_{n\to\infty} \operatorname{dist}(f^{-n}(x), f^{-n}(x_0)) = 0\}.$$

Consider now a more general setting. Let f be a diffeomorphism of class $\mathscr{C}^1$ on a smooth Riemannian manifold M. Let X be an invariant compact subset of M. We say that X is *hyperbolic* if at every point $x \in X$ there is a decomposition of the tangent space at x: $E(x) = E^s(x) \oplus E^u(x)$, a number $0 < \mu < 1$ and a constant $C > 0$ such that

1. For every $v \in E^s(x)$ and every $n \geq 0$, we have $\|Tf^n(x).v\| \leq C\mu^n\|v\|$.
2. For every $v \in E^u(x)$ and every $n \geq 0$, we have $\|Tf^{-n}(x).v\| \leq C\mu^n\|v\|$.

It is not difficult to see that this decomposition is unique and the maps $x \mapsto E^s(x)$ and $x \mapsto E^u(x)$ are continuous. In particular, the dimensions of stable and unstable spaces are locally constant and the angle between $E^s(x)$ and $E^u(x)$ is uniformly bounded from below by a positive constant. We have the following theorem.

Theorem 2.4.6. *Assume that the dimension of $E^s(x)$ and $E^u(x)$ are equal to r_s and r_u respectively. Let $\delta > 0$ be a constant, small enough. Define for $x \in X$*

$$W^s_{loc}(x) = \{y \in M,\ \mathrm{dist}\,(f^n(y), f^n(x)) \leq \delta \textit{ for } n \geq 0\}$$

and

$$W^u_{loc}(x) = \{y \in M,\ \mathrm{dist}\,(f^{-n}(y), f^{-n}(x)) \leq \delta \textit{ for } n \geq 0\}.$$

1. *Each set $W^s_{loc}(x)$ (resp. $W^u_{loc}(x)$) is the image of the unit ball in $\mathbb{R}^{r_s}$ (resp. $\mathbb{R}^{r_u}$) of a $\mathscr{C}^1$-embedding map which is equal to x at 0 and depends continuously on x for the $\mathscr{C}^1$ topology.*
2. *For every $\mu' \in]\mu, 1[$, there is a constant $C > 0$ such that for $n \geq 0$ we have $\mathrm{dist}\,(f^n(y), f^n(x)) \leq C\mu'^n$ when $y \in W^s_{loc}(x)$ and $\mathrm{dist}\,(f^{-n}(y), f^{-n}(x)) \leq C\mu'^n$ when $y \in W^u_{loc}(x)$.*
3. *For each $x \in X$, the sets*

$$W^s(x) = \{y \in M,\ \lim_{n\to\infty} \mathrm{dist}\,(f^n(y), f^n(x)) = 0\}$$

and

$$W^u(x) = \{y \in M,\ \lim_{n\to\infty} \mathrm{dist}\,(f^{-n}(y), f^{-n}(x)) = 0\}$$

are the images of $\mathbb{R}^{r_s}$ and $\mathbb{R}^{r_u}$ by $\mathscr{C}^1$ embedding in M. We also have

$$W^s(x) = \bigcup_{n\geq 0} f^{-n}(W^s_{loc}(f^n(x)) \quad \textit{and} \quad W^u(x) = \bigcup_{n\geq 0} f^n(W^u_{loc}(f^n(x)).$$

2.5 Markov partitions and examples

In this section, we will study some dynamical systems which are factors of subshifts of finite type. This is often applied to the restriction of a dynamical system to the most interesting part like the set of non-wandering points. We first describe the general idea.

Let $T : X \to X$ be a topological dynamical system. Decompose X into a finite union $X = \cup_{i\in I} R_i$ of closed subsets R_i such that the interiors of R_i are mutually disjoint. Consider a sequence $i = (i_n)_{n\in\mathbb{N}} \in I^{\mathbb{N}}$. It is *the code* of the orbit of a point a if $T^n(a) \in R_{i_n}$ for every $n \geq 0$. The above decomposition of X is *good* if for any i the set $H(i) := \cap_{n\geq 0} T^{-n}(R_{i_n})$ contains at most one point. Assume that this is the case. Let Z denote the set of i such that $H(i)$ is not empty. Then Z is invariant by σ. It is closed when X is compact. Let $h : Z \to X$ be the map $i \mapsto H(i)$. We see that T is semi-conjugate with $\sigma_{|Z}$ via the map h.

We can deduce dynamical properties of T from properties of $\sigma_{|Z} : Z \to Z$. In several cases, Z is the set

$$X_A = \{i \in \{1, \ldots, p\}^{\mathbb{N}},\ A_{i_k,i_{k+1}} = 1 \text{ for } k \in \mathbb{N}\}$$

which is associated to a square matrix A with entries 0 or 1. In this case, T is a factor of a subshift of finite type and we will say that the above decomposition of X is a *Markov partition or decomposition.* Note that this is not exactly a partition in the usual sense. Note also that we can define Markov partition for invertible dynamical systems using $I^{\mathbb{Z}}$.

We now consider some examples. The first and simplest example is the following. Let $T : \mathbb{T} \to \mathbb{T}$ be the map induced by $x \mapsto px$ with p integer and $|p| \geq 2$. Consider the partition

$$\mathbb{T} = \bigcup_{i=1}^{p} \Big(\Big[\frac{i}{p}, \frac{i+1}{p}\Big] + \mathbb{Z}\Big).$$

This is a Markov partition. In this case, we have $Z = \{1, \ldots, p\}^{\mathbb{N}}$ and T is semi-conjugated to σ.

Consider now *the logistic family* of maps $(f_\lambda)_{\lambda>0}$ defined by

$$f_\lambda : \mathbb{R} \to \mathbb{R}$$

$$x \mapsto \lambda x(1 - x).$$

Observe that each point admits at most two preimages, 0 is a fixed point and 1 is sent to this fixed point. The interval $[0, 1]$ is invariant when $\lambda \leq 4$. When $\lambda \geq 1$, if $x < 0$ or $x > 1$, its orbit is strictly decreasing and converges to $-\infty$. In this case, the set of points with bounded orbits is

$$X_\lambda = \bigcap_{n\geq 0} f_\lambda^{-n}([0, 1]).$$

This is a compact subset of $[0,1]$ which is equal to $[0,1]$ if $\lambda \in [1,4]$ and is disconnected when $\lambda > 4$. If $x \notin X_\lambda$, the orbit of x converges to $-\infty$.

The most interesting case is the case where $\lambda < 4$ which was studied by many people. There are here different types of dynamics. When $\lambda = 4$, the restriction of f_λ to $X_\lambda = [0,1]$ is a factor of the Bernouilli shift on $\{0,1\}^{\mathbb{N}}$. This can be shown using the Markov partition

$$[0,1] = [0,1/2] \cup [1/2,1].$$

When λ is strictly larger than 4, the situation is even simpler. Using the same partition, one can prove that f_λ restricted to X_λ is conjugated to the above shift. We will prove this property in the case where λ is strictly larger than $2+\sqrt{5}$. This condition implies that $|f'_\lambda(x)| > 1$ on X_λ. Under the condition $\lambda > 4$, we need another tool: the Schwarzian derivative.

Proposition 2.5.1. *Assume that $\lambda > 2+\sqrt{5}$. For any $i = (i_n)_{n\geq 0}$ there is a unique point $x = h(i) \in X_\lambda$ such that for any $n \geq 0$ we have*

$$0 \leq f_\lambda^n(x) \leq 1/2 \text{ if } i_n = 0$$

and

$$1/2 \leq f_\lambda^n(x) \leq 1 \text{ if } i_n = 1.$$

Moreover, the map h is a homeomorphism from $\{0,1\}^{\mathbb{N}}$ to X_λ which conjugates f_λ to the shift σ.

Proof. The two solutions of the equation $f(x) = 1$ are

$$\frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{1}{\lambda}}.$$

We deduce that

$$f^{-1}([0,1]) = \Delta_0 \cup \Delta_1$$

where

$$\Delta_0 = \Big[0, \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{\lambda}}\Big] \quad \text{and} \quad \Delta_1 = \Big[\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{\lambda}}, 1\Big].$$

Observe that the minimum of $|f'|$ on $\Delta_0 \cup \Delta_1$ is given by

$$m = f'\Big(\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{\lambda}}\Big) = -f'\Big(\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{\lambda}}\Big).$$

It follows that

$$m = 2\lambda\sqrt{\frac{1}{4} - \frac{1}{\lambda}} = \sqrt{\lambda^2 - 4\lambda} > \sqrt{(2+\sqrt{5})^2 - 4(2+\sqrt{5})} = 1.$$

Fix an $i = (i_n)_{n\geq 0}$. We deduce that for $N \geq 0$ the set

$$\bigcap_{n=0}^{N} f_\lambda^{-n}(\Delta_{i_n})$$

is an interval of length $\leq m^{-N}$. This implies that

$$\bigcap_{n\geq 0} f_\lambda^{-n}(\Delta_{i_n})$$

is reduced to a point $h(i)$. So, h defines a bijection from $\{0,1\}^{\mathbb{N}}$ to X_λ which is a Cantor set. Therefore, f_λ restricted to X_λ is conjugated to σ. $\square$

The last example we consider in this section is the horseshoe. This example is due to Smale. Let $R = [a,b] \times [c,d]$ be a rectangle of $\mathbb{R}^2$ and $f : R \to f(R)$ a diffeomorphism such that

(1) $f(R) \cap R$ is the union of two vertical rectangles $R_0 = [a_0, b_0] \times [c,d]$ and $R_1 = [a_1, b_1] \times [c,d]$ where $a < a_0 < b_0 < a_1 < b_1 < b$.

(2) $f^{-1}(R) \cap R$ is the union of two horizontal rectangles $f^{-1}R_0 = [a,b] \times [c_0, d_0]$ and $f^{-1}(R_1) = [a,b] \times [c_1, d_1]$ where $c < c_0 < d_0 < c_1 < d_1 < d$.

(3) There is $\lambda < 1 < \mu$ such that f restricted to $f^{-1}(R_0)$ is an affine map whose linear part is $(x,y) \mapsto (\lambda x, \mu y)$ and f restricted to $f^{-1}(R_1)$ is an affine map whose linear part is $(-\lambda x, -\mu y)$.

We can extend this map to a global homeomorphism on the Riemann sphere $S = \mathbb{R}^2 \cup \{\infty\}$. The extended map is still denoted by f. Denote by D_0 the lower haft-disc of diameter $[a,b] \times \{c\}$ and D_1 the upper haft-disc of diameter $[a,b] \times \{d\}$. Let D denote the topological disc $D_0 \cup D_1 \cup R$ and D' the closure of its complement. We can construct f with the following properties:

(4) $f(D_0) \subset \mathrm{int}(D_0)$ and $\cap_{n\geq 0} f^n(D_0)$ is reduced to an attractive fixed point that we denote by z_0.

(5) $f(D) \subset \mathrm{int}(D)$ and $\cap_{n\geq 0} f^{-n}(D')$ is equal to ∞ which is a repelling fixed point.

Observe that we have $f(D_1) \subset D_0$. The set of points $z \in D$ whose orbits do not converge to z_0, is the set of points whose orbits are in R. So, it is equal to $\cap_{n\geq 0} f^{-n}(R)$. The set of points which are not in the basin of ∞ is $\cap_{n\geq 0} f^n(R)$.

For every $n \geq 1$, the set $\cap_{0\leq k\leq n} f^{-k}(R)$ is the union of 2^n horizontal rectangles of width $\mu^{-n}(d-c)$ and then $\cap_{k\geq 0} f^{-k}(R)$ is a Cantor family of horizontal segments. In the same way, for $n \geq 1$, the set $\cap_{0\leq k\leq n} f^k(R)$ is the union of 2^n

vertical rectangles of length $\lambda^n(b-a)$ and $\cap_{k\geq 0} f^k(R)$ is a Cantor family of vertical segments. We conclude that the maximal invariant set in R, i.e. $X = \cap_{k\in\mathbb{Z}} f^k(R)$, is a Cantor set. The restriction of f to X is conjugated to the two-sided shift σ on $\{0,1\}^{\mathbb{Z}}$ via the homeomorphism

$$h : (i_k)_{k\in\mathbb{Z}} \mapsto \cap_{k\in\mathbb{Z}} f^{-k}(R_{i_k}).$$

The decomposition $\{R_0 \cap X, R_1 \cap X\}$ is a Markov partition for f restricted to X.

As consequences of the above description of f, we obtain the following properties.

(a) The restriction of f to X is topologically transitive.

(b) Periodic points are dense in X and also in $\Omega = X \cup \{z_0\} \cup \{\infty\}$.

(c) For every point $z = (x, y) \in X$, the segment $W^s_{loc}(z) = [a, b] \times \{y\}$ is the set of points z' such that

$$\lim_{k\to\infty} \operatorname{dist}(f^k(z'), f^k(z)) = 0$$

and the segment $W^u_{loc}(z) = \{x\} \times [a, b]$ is the set of points z' such that

$$\lim_{k\to-\infty} \operatorname{dist}(f^k(z'), f^k(z)) = 0.$$

(d) There are 4 fixed points : an attractive point z_0, a repelling point ∞ and two other points which are saddles.

(e) There are $2 + 2^n$ fixed points for f^n : z_0, ∞ and 2^n other points which are saddles. If z is such a saddle point, we have

$$W^s(z) = \cup_{k\geq 0} f^{-k}(W^s_{loc}(z)) \quad \text{and} \quad W^n(z) = \cup_{k\geq 0} f^k(W^u_{loc}(z)).$$

(f) For all saddle periodic points z, z', the intersection of the manifolds $W^s(z)$ and $W^u(z')$ is an infinite set.

(g) The topological entropy of f is equal to $\log 2$. There is a unique invariant probability measure of maximal entropy which is the limit of the measures equidistributed on the periodic points.

(h) The sets X and Ω are hyperbolic.

2.6 Exercises

Exercise 2.6.1. *Show that if a linear map on $\mathbb{T}^r$ is expansive then it is hyperbolic.*

Exercise 2.6.2. 1. *Let A be a square matrix of rank 2 with integer entries and determinant 1. Show that the following conditions are equivalent*

a) $|\mathrm{trace}(A)| > 2$*;*

b) A is hyperbolic;

c) $\widehat{A}$ is mixing for the Haar measure.

2. *Let A be a square matrix of rank 2 with integer entries and determinant -1. Show that the following conditions are equivalent*

a) $|\mathrm{trace}(A)| \neq 0$*;*

b) A is hyperbolic;

c) $\widehat{A}$ is mixing for the Haar measure.

3. *Find a matrix A of rank 4 with integer entries and determinant 1 such that A is not hyperbolic and $\widehat{A}$ is mixing.*

Exercise 2.6.3. *Using a dynamical argument, show that the eigenvalues of a square matrix of rank 2 and determinant 1 are always irrational.*

Exercise 2.6.4. *Let $F : \mathbb{T}^r \to \mathbb{T}^r$ be a homeomorphism such that F_* is hyperbolic. We give another proof of the existence of a unique continuous map $H : \mathbb{T}^r \to \mathbb{T}^r$ homotopic to* id *and such that $H \circ F = \widehat{F_*} \circ H$.*

1. *Let f be a lift of F to $\mathbb{R}^r$. Show that there is an $M > 0$ such that for $x \in \mathbb{R}^r$ there exists an unique $y = h(x) \in \mathbb{R}^r$ such that the sequence $(f^k(x) - F_*^k(y))_{y\in\mathbb{Z}}$ is bounded and $\|f^k(x) - F_*^k(y)\| \leq M$ for $k \in \mathbb{Z}$.*

2. *Show that $h \circ f = F_* \circ h$ and that $h - \mathrm{id}$ is $\mathbb{Z}^r$-periodic.*

3. *Conclude.*

Exercise 2.6.5. *(closing lemma) Let $\widehat{A}$ be a hyperbolic automorphism of $\mathbb{T}^r$. Show that for any $\epsilon > 0$ there is $\delta > 0$ such that any periodic δ-pseudo-orbit $(x_k)_{k\in\mathbb{Z}}$ of period q can be ϵ-approximated by a periodic orbit of the same period.*

Exercise 2.6.6. *Let p be an integer with $|p| > 1$. Show that the Lebesgue measure on $\mathbb{T}^1$ is the only invariant measure of maximal entropy for $x \mapsto px$.*

Exercise 2.6.7. *Let $T : X \to X$ be an expansive map where X is compact. Show that any decomposition $X = \cup R_i$ is good in the sense of Section 2.5 when the diameters of R_i are small enough.*

Exercise 2.6.8. *For* $\alpha \in]0,1[$, *define* $g_\alpha : [0,1] \to [0,1]$ *by*

$$g_\alpha(x) = \alpha x + \gamma \quad \textit{if } x \in [0, \gamma]$$

and

$$g_\alpha(x) = \beta(1-x) \quad \textit{if } x \in [\gamma, 1],$$

where

$$\beta = 1 + \frac{1}{\alpha} \quad \textit{and} \quad \gamma = \frac{1}{1+\alpha}\cdot$$

1) Draw the graph of g_α *and show that* $\{[0,\gamma],[\gamma,1]\}$ *is a Markov partition.*

2) Deduce that g_α *is semi-conjugated to a subshift of finite type via a map* $h_\alpha : \{0,1\}^{\mathbb{N}} \to [0,1]$.

3) Compute the topological entropy of g.

Exercise 2.6.9. *1) Show that the map*

$$h : x \mapsto \sin^2\left(\frac{\pi x}{2}\right)$$

conjugates $f : [0,1] \to [0,1]$ *and* $g : [0,1] \to [0,1]$ *where*

$$f(x) = 2x \textit{ if } x \in [0, 1/2]$$

$$f(x) = 1 - 2x \textit{ if } x \in [1/2, 1]$$

and

$$g(x) = 4x(1-x).$$

2) Construct a Markov partition for g.

3) Let $\varphi : [0,1] \to \mathbb{R}$ *be a continuous function. Show that there is* $c \in \mathbb{R}$ *such that for almost every* x *with respect to the Lebesgue measure we have*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \varphi(g^i(x)) = c.$$

4) Compute this number c *and prove that*

$$\lim_{n\to\infty} \frac{1}{\#\mathrm{Fix}(g^n)} \sum_{x \in \mathrm{Fix}(g^n)} \varphi(x) = c.$$

Chapter 3

Dynamics of complex polynomials

See the lectures on complex dynamical systems by Dinh and Sibony available at http://www.math.jussieu.fr/∼dinh.

Bibliography

[1] Beardon A., *Iteration of rational functions. Complex analytic dynamical systems*, Graduate Texts in Mathematics, **132**, Springer-Verlag, New York, 1991.

[2] Berteloot F., Mayer V., *Rudiments de dynamique holomorphe*, Paris; EDP Sciences, Les Ulis, 2001.

[3] Carleson L., Gamelin T.W., *Complex dynamics*, Universitext: Tracts in Mathematics, Springer-Verlag, New York, 1993.

[4] Dinh T.-C., Sibony N., Dynamics of holomorphic maps, Introductory Lectures (Master), available at http://www.math.jussieu.fr/∼dinh

[5] Dinh T.-C., Sibony N., Dynamics in several complex variables: endomorphisms of projective spaces and polynomial-like mappings. Holomorphic dynamical systems, 165-294, Lecture Notes in Math., **1998**, Springer, Berlin, 2010.

[6] Katok A., Hasselblatt B., *Introduction to the modern theory of dynamical systems*, Encyclopedia of Mathematics and its Applications, **54**, Cambridge University Press, Cambridge, 1995.

[7] Le Calvez P., Systèmes dynamiques : cours fondamental I, Master 2, 2010-2011.

[8] Le Calvez P., Systèmes dynamiques : cours fondamental II, Master 2, 2009-2010.

[9] Sibony N., Dynamique des applications rationnelles de $\mathbb{P}^k$, *Panoramas et Synthèses*, **8** (1999), 97-185.

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

T.-C. Dinh, UPMC Univ Paris 06, UMR 7586, Institut de Mathématiques de Jussieu, F-75005 Paris, France. `dinh@math.jussieu.fr`, `http://www.math.jussieu.fr/∼dinh`