Zero temperature and selection of maximizing measure

R. Leplaideur
Universit de Bretagne Occidentale

October 30, 2014

Contents

Chapter 1

Symbolic dynamics

1.1 The configuration's space

1.1.1 Topological properties

We consider the space $\Sigma = \{1, 2, ..., k\}^{\mathbb{N}}$ that is that its elements are sequences $x = (x_0, x_1, x_2, x_3...)$, where $x_i \in \{1, 2, ..., k\}, i \in \mathbb{N}$. An element in Σ will also be called an *infinite word* over the alphabet $\{1, \dots, k\}$, and x_i will be called a *digit*.

The distance between two points $x = x_0, x_1, \dots$ and $y = y_0, y_1, \dots$ is given by

$$d(x, y) = \frac{1}{2^{\min\{n,\ x_n \neq y_n\}}}.$$

We sometimes represent this distance graphically as follows:

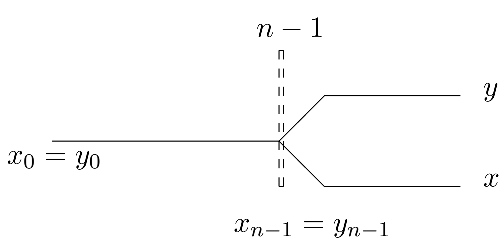

Figure 1.1: The sequence x and y coincide for digits 0 up to $n-1$ and then split.

e.g. With $k = 4$, $d(1, 2, 1, 3..), (1, 1, 1, 2, ..)) = \frac{1}{2}$.

A finite string of symbols $x_0 \dots x_{n-1}$ is also called a *word*, of length n. For a word w, its length is $|w|$. A *cylinder* (of length n) is denoted by $[x_0 \dots x_{n-1}]$. It is the set of points y such that $y_i = x_i$ for $i = 0, \dots n-1$. It will also be denoted as a *n-cylinder*.

Note that cylinders of length $n-1$ form a partition of Σ and $C_n(x)$ will denotes the unique element of this partition which contains x. That is $C_n(x) = [x_0, \dots, x_n]$. The cylinder $C_n(x)$ also coincides with the ball $B(x, \frac{1}{2^n})$. The set of cylinders of length n will be denoted by $\mathcal{C}_n(\Sigma)$.

If $\omega = \omega_0, \dots, \omega_{n-1}$ is a finite word of length n and $\omega' = \omega'_0, \dots$ is a word (of any length possibly infinite), then $\omega\omega'$ is the word

$$\omega_0 \dots \omega_{n-1}\omega'_0 \dots$$

It is called the *concatenation* of ω and ω'.

The set (Σ, d) is a compact metric space. Compactness also follows from the fact that Σ is a product of compact space. Then, note that the topology induced by the distance d coincides with the product topology. The cylinders are clopen sets and generates the topology.

General subshift of finite type

For our purpose, we need to define more general subshift of finite type. A good reference for symbolic dynamics is [83].

Definition 1. *A* transition matrix *is a $d \times d$ matrix with entries in $\{0, 1\}$.*

If $T = T_{ij}$ is a $d \times d$ transition matrix, the subshift of finite type Σ_T associated to T is the set of sequences $x = x_0x_1x_2 \dots x_n \dots$ such that for every j,

$$T_{x_jx_{j+1}} = 1.$$

Equivalently, it is the set of sequences such that the forbidden transitions $T_{ij} = 0$ never appear. If T is given we denote by $\mathcal{A}$ the set of admissible finite words for T, that is the set of words

$$\omega = \omega_0 \dots \omega_{n-1},$$

such that for every i, $T_{\omega_i\omega_{i+1}} = 1$.

Example. The full shift $\Sigma = \{1, \dots d\}^{\mathbb{N}}$ is the subshift of finite type associated to the $d \times d$ matrix with all its entries equal to 1.

The best way to understand what is the subshift of finite associated to a matrix T is to consider paths: for a given i, the set of j such that $T_{ij} = 1$ is the set of letters authorized to follow i. Then, Σ_T is the set of infinite words we can write respecting theses rules, or equivalently, the set of infinite paths that we can do.

Example. If $T = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$ has the associated graph:

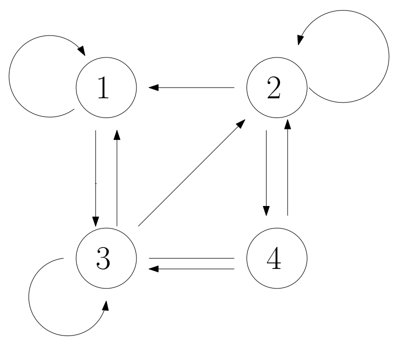

Figure 1.2:

As there are more 1's than 0's in T it is simpler to describe Σ_T as the set of infinite words with letters 1, 2, 3 and 4 such that 12, 14, 23, 41 and 44 never appear.

Definition 2. *Let T be a $d \times d$ transition matrix, let Σ_T the associated subshift of finite type. Two digits i and j are said to be* associated *if there exists a path from i to j and from j to i. We set $i \sim j$.*

Equivalently, $i \sim j$ means that there exists a word in Σ_T of the form

$$i \ldots j \ldots i \ldots j.$$

Claim 1. *The relation $\sim$ is symmetric and transitive.*

Proof of the claim. Symmetry is by definition. Transitivity follows from concatenations. If $i \sim j$ and $j \sim k$ we can form a work

$$\underbrace{i \ldots j}_{i \sim j} \underbrace{\ldots k \ldots j}_{j \sim k} \underbrace{\ldots i}_{j \sim i}.$$

□

It may be that some digit is associated to no other digit.

Definition & Proposition 3. *A digit i is said to be essential if $i \sim i$ holds. The relation $\sim$ is a relation of equivalence on the set of essential digits.*

Exercise 1
Find examples of transition matrices such that one digit is associated only to itself. Find examples of transition matrices such that one digit is associated to no digit.

Definition 4. *The equivalence classes of $\sim$ are called the irreducible component of Σ_T.*

1.1.2 Dynamics

Definition 5. *The shift $\sigma : \Sigma \to \Sigma$, is defined by*

$$\sigma(x_0, x_1, x_2, x_3, ...) = (x_1, x_2, x_3, ...).$$

The shift expands distance by a factor 2. If x and y belong to the same 1-cylinder, then

$$d(\sigma(x), \sigma(y)) = 2\, d(x, y).$$

On the other hand, if $d(x, y) = 1$, then $d(\sigma(x), \sigma(y))$ may have any value between in $0, 1, \frac{1}{2}, \frac{1}{4}, \ldots,$. The shift is Lipschitz and thus continuous.

The shift may be seen as doing infinite path on some graph with states and arrows. arrows from one state to another one indicate the allowed or forbidden transitions. An infinite path is a point in Σ or in Σ_T if we consider a subshift of finite type with transition matrix T. Then, the dynamics is just considering the path starting after the next step.

Definition 6. *Given $x \in \Sigma$ the set $\{\sigma^n(x), n \geq 0\}$ is called the orbit of x. It is denoted by $\mathcal{O}(x)$.*

The main goal in Dynamical systems is to describe orbits and their behaviors. Let us first show some simple behaviors.

Definition 7. *A point $x \in \Sigma$ is said to be periodic if there exits $k > 0$ such that $\sigma^k(x) = x$. In that case the period of x is the smaller positive integer k with this property.*

A periodic point of period 1 is called a fixed point.

Examples
$111\ldots$ is a fixed point.
In $\{1,2,3,4\}^{\mathbb{N}}$, $x := (1,3,2,3,1,3,2,3,1,3,2,3,1,3,2,3,...)$ has period 4 and

$$\{\,(1,3,2,3,1,3,2,3,1,...),(3,2,3,1,3,2,3,1,3,2,...),$$

$$(2,3,1,3,2,3,1,3,2...),(2,3,1,3,2,3,1,3,...)\,\}$$

is the orbit of x.

A n-periodic point is entirely determined by its first n-digits; actually it is the infinite concatenation of these first digits:

$$x = x_0\ldots x_{n-1}\; x_0\ldots x_{n-1}\; x_0\ldots x_{n-1}\ldots$$

Definition 8. *Given a point x in Σ, a point $y \in \Sigma$, such that, $\sigma(y) = x$ is called a* preimage of x.

*A point y such that $\sigma^n(y) = x$ is called a n-*preimage of x. *The set of*

$$\{y \,|\, \textit{there exists an } n \textit{ such that } \sigma^n(y) = x\,\}$$

is called the preimage set of x.

In $\Sigma = \{1,\ldots k\}^{\mathbb{N}}$, each point x has exactly k-preimages. They are obtained by the concatenation process ix, with $i = 1,\ldots,k$. The set of 1-preimage is $\sigma^{-1}(\{x\})$. The set of n-preimages is $(\sigma^n)^{-1}(\{x\})$ which is simply denoted by $\sigma^{-n}(\{x\})$.

Definition 9. *A Borel set A is said to be σ-invariant if it satisfies one of the following equivalent properties:*

1. *For any $x \in A$, $\sigma(x)$ belongs to A.*
2. *$\sigma^{-1}(A) \supset A$.*

e.g. If x is periodic, $\mathcal{O}(x)$ is σ-invariant.

Exercise 2
Show equivalence of both properties in Definition 9.

Exercise 3
If x is periodic, do we have $\sigma^{-1}(\mathcal{O}(x)) = \mathcal{O}(x)$?

Back to more general subshift of finite type

We have defined above a more general subshift of finite type. For such a subshift, we have also defined the irreducible components. Actually,we will give a better description of these components with respects to the dynamics.

Definition 10 (and proposition[1]). *A σ-invariant compact set $\mathbb{K}$ is said to be transitive if it satisfies one of the equivalent two following properties:*

(i) For every pair of open sets of $\mathbb{K}$, U and V, there exists $n > 0$ such that $\sigma^{-n}(U) \cap V \neq \emptyset$.

(ii) There exists a dense orbit.

Exercise 4
Show that if $\mathbb{K}$ is transitive, then the set of points in $\mathbb{K}$ with dense orbit is a G_δ-dense set.

We claim that irreducible components are also transitive components. Indeed, any open set contains a cylinder and it is thus sufficient to prove (i) with cylinders. Now, considering two cylinders of the form $[x_0 \dots x_k]$ and $[y_0 \dots y_n]$, the relation defining irreducible components shows that there exists a connection

$$x_0 \dots x_k z_1 \dots z_m y_0 \dots y_n,$$

remaining in $\mathbb{K}$.

1.1.3 Measures

We denote by $\mathcal{B}$ the Borel σ-algebra over Σ, that is, the one generated by the open sets.

We will only consider probabilities μ on Σ over this sigma-algebra. Due to the fact that cylinders are open sets and generates the topology, they also generates the σ-algebra $\mathcal{B}$. Therefore, all the values $\mu(C_n)$, where C_n runs over the cylinders of length n and n runs over $\mathbb{N}$, determine μ.

We remind the relation between Borel measures and continuous functions:

Theorem 1 (Riesz). *The set of Borel signed measures is the dual of the set $\mathcal{C}^0(\Sigma)$.*

Corollary 11. *The set of probabilities is compact and convex for the weak*-topology.*

[1]We do not prove the proposition part.

We remind that $\mu_n \xrightarrow{w^*} \mu$ means that for every continuous function $f : \Sigma \to \mathbb{R}$,

$$\int f\, d\mu_n \to_{n\to\infty} \int f\, d\mu.$$

We point out that any indicator function of a cylinder, $1\!\!1_{C_n}$, is continuous. We recall that the support of a (probability) measure μ is the set of point x such that

$$\forall\, \varepsilon,\ \ \mu(B(x, \varepsilon)) > 0.$$

In our case this is equivalent to $\mu(C_n(x)) > 0$ for every n. The support is denoted by $\mathrm{supp}\,(\mu)$.

Exercise 5
Show that $\mathrm{supp}\,(\mu)$ is compact.

Definition 12. *We say that a probability μ is* invariant for the shift σ *if for any $A \in \mathcal{B}$,*

$$\mu(\sigma^{-1}(A)) = \mu(A).$$

We will also say that μ is σ-invariant, or simply invariant as σ is the unique dynamics we shall consider.

To consider invariant measure means the following thing: if we see the action of $\mathbb{N}$ as a temporal action on the system, the systems is closed, in the sense that along the time, there is neither creation nor disappearance of mass in the system.

Exercise 6
Show that if μ is invariant, $\mathrm{supp}\,\mu$ is invariant. Is it still the case if μ is not invariant ?

1.2 Invariant measures

In this section we present some particular invariant measures in our settings and also give some more general results.

1.2.1 Examples of invariant measures

Periodic measures

If x is a point in Σ, δ_x is the *Dirac* measure at x, that is

$$\delta_x(A) = \begin{cases} 1 \text{ if } x \in A, \\ 0 \text{ otherwise.} \end{cases}$$

Then, if x is n-periodic,

$$\mu := \frac{1}{n} \sum_{j=0}^{n-1} \delta_{\sigma^j(x)}$$

is σ-invariant.

The Bernoulli product measure

Let consider $\Sigma = \{1,2\}^{\mathbb{N}}$, pick two positive numbers p and q such that $p + q = 1$. Consider the measure $\mathbb{P}$ on $\{1,2\}$ defined by

$$\mathbb{P}(\{1\}) = p, \ \mathbb{P}(\{2\}) = q.$$

Then consider the measure $\mu := \otimes\mathbb{P}$ on the product space Σ. We remind that it is defined by

$$\mu([x_0 \dots x_{n-1}]) = p^{\#\text{ of 1's in the word}} q^{\#\text{ of 2's in the word}}.$$

We claim it is an invariant measure. Indeed, for any cylinder $[x_0 \dots x_{n-1}]$,

$$\sigma^{-1}([x_0 \dots x_n]) = [1x_0 \dots x_n] \sqcup [2x_0 \dots x_n].$$

Then, $\mu([1x_0 \dots x_n]) = p\mu([1x_0 \dots x_n])$ and $\mu([2x_0 \dots x_n]) = q\mu([1x_0 \dots x_n])$.

This example corresponds to the model of tossing a coin (head identified with 1 and tail identified with 2) in an independent way a certain number of times. We are assuming that each time we toss the coin the probability of head is p and the probability of tail is q.

Therefore, $\mu([211])$ describes the probability of getting tail in the first time and head in the two subsequent times we toss the coin, when we toss the coin three times.

Remark 1. *The previous example shows that there are uncountably many σ-invariant probabilities on $\{1,2\}^{\mathbb{N}}$.* ■

Markov chain

Again, we consider the case $\Sigma = \{1,2\}^{\mathbb{N}}$. Pick p and q two positive numbers in $]0,1[$, and set

$$P = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix} = \begin{pmatrix} P(1,1) & P(1,2) \\ P(2,1) & P(2,2) \end{pmatrix}.$$

The first writing of P shows that 1 is an eigenvalue. If we solve the equation

$$(x,y).P = (x,y),$$

we find a one-dimensional eigenspace (directed by a *left eigenvector*) with $y = \dfrac{1-p}{1-q}x$. Therefore, there exists a unique left eigenvector (π_1, π_2) such that

$$(\pi_1, \pi_2).P = (\pi_1, \pi_2) \text{ and } \pi_1 + \pi_2 = 1.$$

Note that π_1 and π_2 are both positive.

The measure μ is then defined by

$$\mu([x_0 \dots x_n]) = \pi_{x_0} P(x_0, x_1) P(x_1, x_2) \dots P(x_{n-2}, x_{n-1}).$$

A simple way to see the measure μ is the following: a word $\omega = \omega_0 \dots \omega_{n-1}$ has to be seen as a path of length n, starting at state $\omega_0 \in \{1,2\}$ and finishing at state ω_{n-1}. The measure $\mu([\omega])$ is the probability of this space among all the paths of length n. This probability is then given by the initial probability of being in state ω_0 (given by π_{ω_0}) and then probabilities of transitions from the state ω_j to ω_{j+1}, these events being independent.

A probability of this form is called the *Markov measure obtained from the line stochastic matrix P and the initial probability π.*

Exercise 7
Show that the Bernoulli measure is also a Markov measure.

More generally we have

Definition 13. *A $d \times d$ matrix P such that all entries are non-negative and the sum of the elements in each line is equal to* 1 *is called a* **line stochastic matrix**.

One can show that for a line stochastic matrix there exist only one vector $\pi = (\pi_1, \pi_2, ..., \pi_d)$, such that all $\pi_j > 0$, $j \in \{1, 2, .., d\}$, $\sum_{j=1}^{d} \pi_j = 1$, and

$$\pi \ = \ \pi \ P.$$

It is called the left invariant probability vector.

Definition 14. *Given a $d \times d$ line stochastic matrix P, and its left invariant probability vector $\pi = (\varphi_1, \varphi_2, ..., \varphi_d)$, we define, μ on $\Sigma = \{1, 2, \dots, d\}$, in the following way: for any cylinder $[x_0, x_1 \cdots x_k]$*

$$\mu([x_0, x_1 \cdots x_k]) \; = \pi_{x_0} \, P(x_0, x_1) \, P(x_1, x_2) \, P(x_2, x_3) \, ... \, P(x_{k-1}, x_k).$$

This measure is invariant for the shift and it is called the Markov measure associated to P and π.

1.2.2 General results on invariant measures

The definition of invariant measure involves Borel sets. The next result gives another characterization of invariant measures (see [117]):

Proposition 15. *The measure μ is invariant if and only if for any continuous function f*

$$\int f(x) \, d\mu(x) = \int f(\sigma(x)) \, d\mu(x).$$

We denote by $\mathcal{M}_\sigma$ the set of invariant probabilities on Σ. Proposition 15 shows it is a closed subset of probabilities for the weak*-topology, hence it is compact and convex.

Definition 16. *An extremal measure in $\mathcal{M}_\sigma$ is said to be ergodic.*

This definition is however not useful and clearly not easy to check. The next proposition gives other criteria for a measure to be ergodic.

Proposition 17. *A probability μ is ergodic if and only if it satisfies one of the following properties:*

1. *Every invariant Borel set has full measure or zero measure.*
2. *For $f : \Sigma \to \mathbb{R}$ continuous, $f = f \circ \sigma$ μ-a.e. implies f is constant.*

Exercise 8
Show that a Markov measure is ergodic.

We can now state the main theorem in Ergodic Theory:

Theorem 2. (Birkhoff Ergodic Theorem) *Let μ be σ-invariant and ergodic. Then, for every continuous function $f : \Sigma \to \mathbb{R}$ there exist a Borel set K, such that $\mu(K) = 1$, and for every x in K*

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} f(\sigma^{k-1}(x)) = \int f \, d\mu.$$

The Birkhoff Theorem says that, under the assumption of ergodicity, a time average is equal to a spatial average. Here is an example of application:

"Average cost of car ownership rises to \$8,946 per year."

What does the term average mean ? One can imagine that we count how many one fixed person spends every year for his car, and then do the average cost. This is a time-average. The main problem of this average is to know if it representative of the cost of anybody.

On the contrary, one can pick some region, then count how many people spend in 1 year for their car, and take the average amount. This is a spatial average. The main problem is to know if it represents how much each person is going to spend along the years (at beginning the car is new, and then get older !).

The ergodic assumption means that the repartition of old and new cars in the space is "well" distributed and/or that the chosen person in the first way to compute the average cost is "typical". Then, the Birkhoff theorem says that both averages are equal.

Notation. We set $S_n(f)(x) := f(x) + \ldots + f \circ \sigma^{n-1}(x)$.

We finish this section with some application of ergodicity:

Proposition 18. *Let μ be an invariant ergodic probability. Let x be a "generic " points in $\operatorname{supp}\mu$. Then, x returns infinitely many times as closed as wanted to itself.*

Proof. Pick $\varepsilon > 0$ and consider the ball $B(x, \varepsilon)$. It is a clopen set, hence $\mathbb{1}_{B(x,\varepsilon)}$ is continuous. The point x is generic for μ, then $\lim_{n\to\infty} \frac{1}{n} S_n(\mathbb{1}_{B(x,\varepsilon)})(x) = \mu(B(x, \varepsilon))$ and this last term is positive. Therefore, there are infinitely many n such that $\mathbb{1}_{B(x,\varepsilon)}(\sigma^n(x)) = 1$. □

1.3 Symbolic and other dynamics

The Symbolic dynamics is very useful to study other dynamics.

Definition 19. *Two dynamical systems (X, T) and (Y, f) are said to be conjugated if there exists a bijection $\Theta : X \to Y$ such that the following graph commutes :*

$$\begin{array}{ccc} X & \xrightarrow{T} & X \\ \Theta\downarrow & \circlearrowleft & \downarrow \\ Y & \xrightarrow[f]{} & Y \end{array}$$

If the map Θ is onto, we say that the systems are semi-conjugated.

If X and Y are topological space we usually require Θ to be a homeomorphism or a continuous function (for semi-conjugacy).

Example. The systems $x \mapsto 2x$ in $[0, 1]$ is semi conjugated to the full shift $\{0, 1\}^{\mathbb{N}}$. More precisely, the map Θ which associates to a sequence $(x_n) \in \{0, 1\}^{\mathbb{N}}$ the point $\overline{x} := \sum_{n=0}^{+\infty} x_n 2^{-(n+1)}$ is continuous, and one-to-one except for dyadic points (which is a countable set of points). Note that the ω-limit set of the dyadic points is reduced to the two fixed-points 0 and 1.

Theorem 3. *If (X, T) is an irreducible mixing Axiom-A diffeomoprhism then it is semi-conjugated to a subshift of finie-type (Σ_A, σ). The semi-conjugacy is Hölder continuous and one-to-one except on some F_σ invariant subset, where it is finite-to-one.*

Proof. See [26] □

1.4 Ergodic optimization and temperature zero

The set $\mathcal{M}_\sigma$ of invariant measures is usually big. It is thus natural to find a way to singularize some measures.

One first way is to consider maximizing measures:

Definition 20. *Let $A : \Sigma \to \mathbb{R}$ be a continuous function. An invariant measure μ is said to be A-maximizing if*

$$\int A\, d\mu = \max \left\{ \int A\, d\nu,\ \nu \in \mathcal{M}_\sigma \right\} =: m(A).$$

Note that this maximum is well defined because A is continuous and $\mathcal{M}_\sigma$ is compact for the weak*-topology. This yields that $\mu \mapsto \int A\, d\mu$ is a continuous map on $\mathcal{M}_\sigma$, and then is bounded and reaches its bounds.

This will be studied in Chapter **??**. There, it will be explained some tools, and we emphasize that *generically* for the $\mathcal{C}^0$topology there exists a unique maximizing measure.

Nevertheless it is extremely easy to construct non-generic examples: take two disjoint periodic orbits, say $\mathcal{O}(x)$ and $\mathcal{O}(y)$, set $K = \mathcal{O}(x) \sqcup \mathcal{O}(y)$, and then pick

$A(\omega) = -d(\omega, K)$. It has two invariant ergodic maximizing measures, the periodic measure associated to $\mathcal{O}(x)$ and the periodic measure associated to $\mathcal{O}(y)$.

We also emphasize that any convex combination of these two measures is also invariant and A-maximizing.

Another way to singularis some measure is the Thermodynamic formalism. This will be the topic of Chapter 2.

Without entering too much into the theory, for a real parameter β we shall associate to $\beta.A$ a functional $\mathcal{P}(\beta)$, and for each β some measure μ_β called *equilibrium state for the potential* $\beta.A$. In statistical mechanics, β represents the inverse of the temperature. Then, $\beta \to +\infty$ means that the temperature goes to 0.

The relations with ergodic optimization are the following:

1. $\beta \mapsto \mathcal{P}(\beta)$ is convex and admits an asymptote for $\beta \to +\infty$. The slope is given by $m(A) = \max \displaystyle\int A\, d\mu$.
2. Any accumulation point for μ_β as $\beta \to +\infty$ is A-maximizing. Then the question is to know if there is *convergence* and if yes, how does μ_β select the limit ?

These are the mains points in Chapters 3 for general results and 5 for a specific example.

Chapter 2

Thermodynamic Formalism and Equilibrium states

2.1 Motivation-main definitions

2.1.1 Motivations and definition

There are several motivations for the thermodynamic formalism of a dynamical system. We present here the simplest to be understood, which is perhaps not the more relevant.

A dynamical systems admits, in general, various ergodic measures. One natural problem would be to find a way to singularize some among others.

Given $A : \Sigma \to \mathbb{R}$, the thermodynamic formalism aims to singularize measures via the maximizing principal :

$$\mathcal{P}(A) := \sup_{\mu \in \mathcal{M}_\sigma} \left\{ h_\mu + \int A \, d\mu \right\}. \tag{2.1}$$

The quantity h_μ is the Kolmogorov entropy for the measure μ. It is a non-negative real number, bounded by $\log d$ if $\Sigma = \{1, \ldots d\}^{\mathbb{N}}$. Roughy speaking it measures the chaos seen by the measure μ.

Definition 21. *Any measure which realizes the maximum in* (2.1) *is called an* equilibrium state *for A. The function A is called the* potential *and $\mathcal{P}(A)$ is* the pressure *of the potential.*

For a given σ-invariant measure μ, the quantity $h_\mu + \displaystyle\int A \, d\mu$ is called the *free-energy* of the measure (with respect to the potential A).

If we consider some fixed $A : \Sigma \to \mathbb{R}$ and β a real parameter, we write $\mathcal{P}(\beta)$ instead of $\mathcal{P}(\beta.A)$. It is an easy exercise to check that $\beta \mapsto \mathcal{P}(\beta)$ is convex, thus continuous. In the rest of course, we will focus on this function, called the *pressure function*. In statistical mechanics, β is the inverse of the temperature.

2.1.2 Entropy and existence of equilibrium states

We refer the reader to [26][114] [8] [76] [88] for the results we use from Thermodynamic Formalism. We present here some results, the ones which are the most important for our purpose.

We emphasize that the results we present here are stated for Σ but they holds for any general irreducible subshift of finite type.

Entropy

For general results on entropy see also [10] and [100]. The complete description of entropy is somehow complicated and not relevant for the purpose of this course.

A simple definition for the general case could be the following:

Theorem 4 (and definition)**.** *Let μ be a σ-invariant ergodic probability. Then, for μ-a.e. $x = x_0x_1x_2\ldots$,*

$$\lim_{n\to\infty} -\frac{1}{n}\log\mu([x_0x_1\ldots x_n])$$

exists and is independent of x. It is equal to h_μ.

If μ_0 and μ_1 are both invariant and ergodic probabilities, for every $\alpha \in [0,1]$ set $\mu_\alpha := \alpha.\mu_1 + (1-\alpha).\mu_0$. Then,

$$h_{\mu_\alpha} = \alpha h_{\mu_1} + (1-\alpha)h_{\mu_0}.$$

The definition for ergodic measure means that we can consider

$$\mu([x_0x_1\ldots x_{n-1}]) \approx e^{-nh_\mu},$$

for a.e. x. Roughly speaking entropy explains that very long cylinder have all there same measure.

Examples

• Let us consider $\Sigma = \{1,2\}^{\mathbb{N}}$ and μ the Bernoulli measure given by the line stochastic matrix $P = \begin{pmatrix} p & q \\ p & q \end{pmatrix}$. We have seen before that

$$\mu([x_0\ldots x_{n-1}]) = p^{\#\text{ of 1's in the word}}q^{\#\text{ of 2's in the word}}.$$

We also have seen before that $\frac{1}{n}\#$ of 1's in the word $\to_{n\to\infty} \mu([1]) = p$, and $\frac{1}{n}\#$ of 2's in the word $\to_{n\to\infty} \mu([2]) = q$. Therefore,

$$h_\mu = -p\log p - q\log q.$$

• More generally, if μ is the Markov measure associated to a line stochastic matrix P, and, $\pi = (\pi_1, \pi_2, , , , \pi_d)$ is the stationary vector, then,

$$h(\mu) = -\sum_{i,j=1}^{n} \pi_i \; P(i,j) \; \log P(i,j).$$

Exercise 9
Check that the limit in Theorem 4 exists if μ is Markov.

Proposition 22. *The (metric) entropy is upper-semi continuous : If μ_n converges to μ for the weak*-topology, then*

$$h_\mu \geqslant \limsup_{n\to\infty} h_{\mu_n}.$$

Existence of equilibrium states

As an immediate consequence of Proposition 22 we get:

Theorem 5. *If A is continuous, then there exists at least one equilibrium state for A.*

We remind that $\beta \mapsto \mathcal{P}(\beta)$ is convex. It is thus Lebesgue a.e. differentiable (actually everywhere except on some possibly empty countable set). Then we get:

Lemma 23. *Let β be such that $\mathcal{P}$ is differentiable at β. Then, for every equilibrium state μ_β for $\beta.A$,*

$$\mathcal{P}'(\beta) = \int A\, d\mu_\beta.$$

Proof. Let β' be any real number. Let μ_β be any equilibrium state for $\beta.A$. The definition of being equilibrium state yields

$$h_{\mu_\beta} + \beta.\int A\, d\mu_\beta = \mathcal{P}(\beta),$$

and the definition of the pressure yields for β', $\mathcal{P}(\beta') \geqslant h_{\mu_\beta} + \beta'.\int A\,d\mu_\beta$. This means that the graph of $\mathcal{P}$ is above the line $\beta' \mapsto\geqslant h_{\mu_\beta} + \beta'.\int A\,d\mu_\beta$ and touches it for $\beta' = \beta$. This is an equivalent definition for the tangent to a convex graph. □

The function $\beta \mapsto \mathcal{P}(\beta)$ is convex, thus the slope of the graphe increases.

Proposition 24. *The graph of $\mathcal{P}(\beta)$ admits an asymptote if β goes to $+\infty$. The slope is given by $m(A) = \sup \int A\,d\mu$.*

The proof will be done later.

Theorem 6 (see [41])**.** *Assume that A is continuous. Any accumulation point for μ_β (and for the weak* topology) is a A-maximizing measure.*

The spirit of the proof is the following. If μ_β maximizes $h_\nu + \beta.\int A\,d\nu$, it also maximizes $\frac{1}{\beta}h_\nu + \int A\,d\nu$. If $\beta \to +\infty$ it maximizes $\int A\,d\nu$, since h_ν is non-negative and bounded from above.

Exercise 10
Prove Theorem 6.

2.2 Uniqueness of Equilibrium State

Uniqueness does not always hold. Nevertheless the key result is the following:

Theorem 7. *If $A : \Sigma \to \mathbb{R}$ is Hölder continuous, then there is a unique equilibrium state for A. Moreover it is a* Gibbs measure *and $\beta \to \mathcal{P}(\beta)$ is analytic.*

We want to emphasize this result. If existence of equilibrium state is done via a general result of maximization of a semi-continuous function, uniqueness is obtained for Hölder continuous potential via a completely different way: the key tool is an operator acting on continuous and Hölder continuous functions. Then, the pressure and the equilibrium state are related to the spectral properties of this operator.

2.2.1 The Transfer operator

In this section we consider a fixed α-Hölder potential $A : \Sigma \to \mathbb{R}$.

We recall that $A : \Sigma \to \mathbb{R}$ is said to be **α-Hölder**, $0 < \alpha < 1$, if there exists $C > 0$, such that, for all x, y we have $|\, A(x) - A(y)\,| \leq C\, d(x,y)^\alpha$.

For a fixed value α, we denote by $\mathcal{H}_\alpha$ the set of α-Hölder functions $A : \Sigma \to \mathbb{R}$. $\mathcal{H}_\alpha$ is a vector space.

For a fixed α, the norm we consider in the set $\mathcal{H}_\alpha$ of α-Hölder potentials A is

$$||A||_\alpha = \sup_{x \neq y} \frac{|A(x) - A(y)|}{d(x,y)^\alpha} + \sup_{x \in \Sigma} |A(x)|.$$

For a fixed α, the vector space $\mathcal{H}_\alpha$ is complete with the above norm.

Definition 25. *We denote by $\mathcal{L}_A : \mathcal{C}^0(\Sigma) \to \mathcal{C}^0(\Sigma)$ the Transfer operator corresponding to the potential A, which is given in the following way: for a given ϕ we will get another function $\mathcal{L}_A(\phi) = \varphi$, such that,*

$$\varphi(x) = \sum_{a,\ ax_0 \in \mathcal{A}} e^{A(ax)}\, \phi(ax).$$

In another form

$$\varphi(x) = \varphi(x_0 x_1 \cdots) = \sum_{a, ax_0 \in \mathcal{A}} e^{A(ax_0x_1x_2,\ldots)}\, \phi(ax_0x_1x_2\ldots).$$

The transfer operator is also called the Ruelle-Perron-Frobenius operator. It had been introduced by Ruelle and extends in some sense the matrices with positive entries. We remind that for such matrices, the Perron-Frobenius theorem gives information on the spectrum.

It is immediate to check that $\mathcal{L}_A$ acts on continuous functions. It also acts on α-Hölder functions if A is α-Hölder.

Consequently the dual operator acts on measures:

$$\begin{aligned} \mathcal{L}_A^* : \mu \ &\mapsto \ \nu \\ \int \psi\, d\nu &:= \int \mathcal{L}_A(\psi)\, d\mu. \end{aligned}$$

Theorem 8 (see [?]). *Let λ_A be the spectral radius of $\mathcal{L}_A$. Then, λ_A is an eigenvalue for $\mathcal{L}_A^*$: there exists a probability measure ν_A such that*

$$\mathcal{L}_A^*(\nu_A) = \lambda_A \nu_A.$$

This probability is called the eigenmeasure and/or the conformal measure.

Then, the main ingredient to prove uniqueness of the equilibrium state is

Theorem 9. *The operator $\mathcal{L}_A$ is quasi-compact on $\mathcal{H}_\alpha$: λ_A is simple isolated and the unique eigenvalue with maximal radius. The rest of the spectrum is contained in a disk $\mathbb{D}(0, \rho\lambda_A)$ with $0 < \rho < 1$.*

From Theorem 9 we get a unique H_A, up to the normalization $\displaystyle\int H_A\, d\nu_A = 1$, such that

$$\mathcal{L}_A(H_A) = \lambda_A H_A.$$

Exercise 11
Show that the measure defined by $\mu_A = H_A\nu_A$ is σ-invariant.

This measure is actually a *Gibbs* measure: there exists $C_A > 0$ such that for every $x = x_0x_1\ldots$ and for every n,

$$e^{-C_A} \leqslant \frac{\mu_A([x_0\ldots x_{n-1}])}{e^{S_n(A)(x)-n\log\lambda_A}} \leqslant e^{C_A}. \tag{2.2}$$

These two inequalities yields that the free energy for μ_A is $\log\lambda_A$. The left-side inequality yields that for any other ergodic measure ν,

$$h_\nu + \int A\, d\nu < \log\lambda_A.$$

In particular we get $\mathcal{P}(A) = \lambda_A$ and μ_A is the unique equilibrium state for A.

The same work can be done for $\beta.A$ instead of A. Now, the spectral gap obtained in Theorem 9 and general results for perturbations of spectrum of operators yield that $\beta \mapsto \mathcal{P}(\beta)$ is locally analytic. One argument of convexity shows that it is globally analytic.

2.2.2 Conformal measure

If X is a compact metric space and $f : X \to X$ is a Borel map, and ν is a measure, the dynamical system is said to be *non-singular* if $f_*\nu$ is absolutely continuous with respect to ν, that is

$$\nu(B) = 0 \Longrightarrow \nu(f^{-1}(B)) = 0.$$

Example. if f is $\mathcal{C}^1$ on some manifold, the system is non-singular for the Lebesgue measure.

If the system (X, f, ν) is non-singular, it is normal to try to find some invariant measure μ absolutely continuous with respect to the measure ν. This means that we have

$$d\mu = h.d\nu.$$

Invariance means

$$\begin{aligned}\int \phi \circ f \, d\mu &= \int \phi \, d\mu \\ \int \phi \circ f \, h \, d\nu &= \int \phi \, h \, d\nu.\end{aligned}$$

Therefore, it is natural to consider the operator $\mathcal{T} : \phi \mapsto \phi \circ T$ and its dual. The Radon-Nikodim derivative is actually a fixed point for the dual operator. It turns out that the dual operator exactly is the Transfer operator.

This is well defined as soon as we get some non-singular measure ν. For the general case, we need first to find this *conformal* measure. This is why we define the transfer operator as acting on continuous functions, to get that its dual acts on measures.

Then, using notations of Theorem 8 we get that ν_A is a conformal measure with Jacobian $e^{A-\mathcal{P}(A)}$: for any Borel set B where σ

$$\nu_A(B) = \int_{\sigma(B)} e^{A \circ \sigma^{-1}(x) - \mathcal{P}(A)} \, d\nu_A, \tag{2.3}$$

where σ^{-1} is the inverse branch from $\sigma(B)$ to B.

2.2.3 Some more results

The theory described above can be generalized to other cases. In the special case of Σ_T (irreducible shift of finite type) we can actually prove that λ_A is the unique dominating eigenvalue. Moreover we get for every ψ Hölder continuous,

$$\mathcal{L}_A^n(\psi) = e^{n\mathcal{P}(A)} \int \psi \, d\nu_\beta . \phi + e^{n(\mathcal{P}(A)-\varepsilon)} \psi_n, \tag{2.4}$$

where ε is a positive real number (depending on A), ψ_n is continuous and $||\psi_n||_\infty \leqslant C.||\psi||_\infty$, for every n and C is a constant (depending on A). From this one get

$$\mathcal{P}(A) = \lim_{n\to\infty} \frac{1}{n} \log \mathcal{L}_A^n(\mathbb{1}),$$

which yields $\dfrac{d\mathcal{P}}{d\beta}(\beta) = \int A \, d\mu_\beta$.

We finish this subsection by some important remark. We have seen the double-inequality (2.2)

$$e^{-C_A} \leqslant \frac{\mu_A([x_0 \dots x_{n-1}])}{e^{S_n(A)(x)-n\log\lambda_A}} \leqslant e^{C_A}.$$

We emphasize here that C_A is proportional to $||A||_\infty$. Therefore, replacing A by $\beta.A$ and letting $\beta \to +\infty$ implies $C_{\beta.A} \to +\infty$. More generally, doing $\beta \to +\infty$ will explodes all the constants.

Nevertheless we have:

Proposition 26. *Assume that A is α-Hölder. There exists a universal constant C such that $\frac{1}{\beta}\log H_{\beta.A}$ is α-Hölder with norm bounded by $C.||A||_\alpha$.*

Consequently, "at the scale " $\frac{1}{\beta}\log$ we recover bounded quantities.

2.2.4 A complete and exact computation for one example

Let us now assume that A depends on two coordinates, that is

$$A(x_0x_1x_2\dots) = A(x_0, x_1).$$

We denote by $A(i,j)$ the value of A in the cylinder $[ij]$, $i, j \in \{1, 2, .., d\}$. In this case, the Transfer operator takes a simple form:

$$\mathcal{L}_A(\phi)(x_0x_1x_2\dots) = \sum_{a\in\{1,2,\dots,d\}} e^{A(ax_0)}\,\phi(ax_0x_1x_2\dots).$$

Let M be the matrix will all positive entries given by $M_{i,j} = e^{A(i,j)}$.

Lemma 27. *The spectral radius of $\mathcal{L}_A$ is also the spectral radius of M.*

Proof. Assume that ϕ is a function depending only on one coordinate, *i.e.,*

$$\phi(x_0x_1x_2\dots) = \phi(x_0).$$

Then, set by abuse of notation ϕ the vector $(\phi(1), \phi(2), ..., \phi(d))$. Then, for every j

$$\mathcal{L}_A(\phi)(j) = \sum_{j=1}^{s} M_{ij}.\phi(i),$$

which can be written as $\mathcal{L}_A(\phi) = M^*.\phi$. This yields that the spectral radius λ_M of M is lower or equal to λ_A.

We remind that the spectral radius is given by

$$\lambda_A := \limsup_{n\to\infty} \frac{1}{n} \log |||\mathcal{L}_A^n|||, \text{ and } |||\mathcal{L}_A^n||| = \sup_{||\psi||=1} ||\mathcal{L}^n(\psi)||_\infty.$$

The operator $\mathcal{L}_A$ is positive and this shows that for every n, $|||\mathcal{L}_A^n||| = ||\mathcal{L}_A^n(\mathbb{1})||_\infty$. Now, $\mathbb{1}$ depends only on 1 coordinate, which then $\mathcal{L}_A^n(\mathbb{1}) = M^n(\mathbb{1})$. This yields $\lambda_A \leqslant \lambda_M$. □

Theorem 10. (Perron-Frobenius) *Let $B = (b_{ij})$ be a $d \times d$ matrix with positive entries. Then, the spectral radius of B, say λ, is a simple dominated eigenvalue. The associated eigenspace is generated by some "positive" vector $u = (u_1, \cdots, u_d)$ with $u_i > 0$.*

Proof. Consider in $\mathbb{R}^d$ the simplex $\mathcal{S} = \{u = (u_1, \cdots, u_d),\ u_i > 0,\ u_1 + \ldots u_d = 1\}$. As $b_{ij} > 0$, then $B.\mathcal{S}$ is a compact sub-simplex in the interior $\overset{\circ}{\mathcal{S}}$ of $\mathcal{S}$.

The distance between $B.\mathcal{S}$ and $\partial\mathcal{S}$ is thus positive.

We equip $B.\mathcal{S}$ with the Finsler metric : if two points say $\xi \neq \xi'$ are in $B.\mathcal{S}$, let consider the line $(\xi\xi')$. It intersects $\partial\mathcal{S}$ in two points, say ζ and ζ'. Assume that these 4 points are in the order

$$\zeta, \xi, \xi', \zeta'.$$

Then we set $d(\xi, \xi') = \log \dfrac{|\zeta - \xi'||\xi - \zeta'|}{|\zeta - \xi||\zeta' - \xi'|}$.

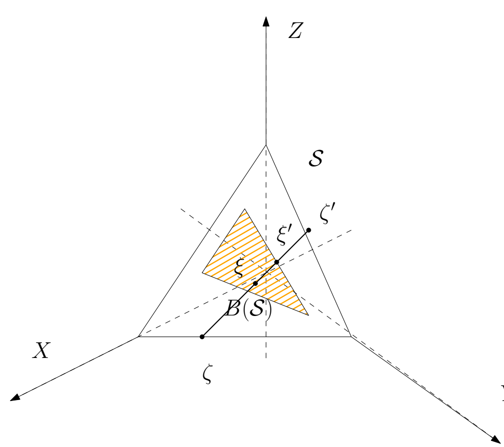

Figure 2.1:

It turns out that $d(.,.)$ is a metric and that B is a contraction for $d(.,.)$ on $B.\mathcal{S}$. Therefore, there exists a unique $u = (u_1, \ldots, u_d) \in B\mathcal{S} \subset \mathcal{S}$ and a positive real number $\bar{\lambda}$ such that

$$B.u = \bar{\lambda}.u.$$

The same argument shows that u is the unique eigenvector (up to a multiplicative factor) in $(\mathbb{R}_+)^d$.

It remains to show that $\bar{\lambda} = \lambda$ and that λ is simple.

□

By Theorem 10 there exists an eigenvector say $\phi_A = (\phi_A(1), \dots, \phi_A(d))$ with positive entries for the matrix M^*. The associated eigenvalue is also the spectral radius of $\mathcal{L}_A$ (due to Lemma 27).

Let us define the 2×2 matrix $P_A = P_A(i, j)$ with

$$P_A(i, j) = \frac{e^{A(i,j)}\phi_A(j)}{\lambda_A \phi_A(i)}.$$

Note that P_A is a line stochastic matrix. Following Subsec. , there exists an invariant Markov measure given by a vector $\pi = (\pi_1, \dots, \pi_d)$. We can indeed recover this result with the Perron-Frobenius theorem.

Actually, Theorem 10 applied to the adjoint matrix M yields an eigenvector vector with positive entries $u = (u_1, \dots, u_d)$. Set $\nu_i = \phi_A(i).u_i$ and $\nu = (\nu_1, \dots, \nu_d)$. We may also assume that $\nu_1 + \dots + \nu_d = 1$. Then

$$\nu.P_A = \nu.$$

The associated invariant measure μ_A is defined by

$$\mu_A([x_0 \dots x_{n-1}]) = \nu_{x_0} P_A(x_0, x_1) \dots P_A(x_{n-2}, x_{n-1}).$$

The exact computation yields $\mu_A([x_0 \dots x_{n-1}]) = u_{x_0} e^{S_n(A)(x) - n \log \lambda_A}.\phi_A(x_{n-1})$. Since u and ϕ_A have positive entries, this shows that μ_A is a Gibbs measure.

Things can be summarized as follows:

Let $M = (M_{ij})$ be the matrix with entries $e^{A(i,j)}$. Let $\mathbf{r} = (r_1, \dots r_d)$ be the right-eigenvector associated to λ with normalization $\sum r_i = 1$. Let $\mathbf{l} = (l_1, \dots, l_d)$ be the left-eigenvector for λ with renormalization $\sum l_i r_i$=1. Then, $\mathbf{r}$ is the eigenmeasure ν_A and $\mathbf{l}$ is the density H_A.
The Gibbs measure of the cylinder $[i_0 \dots i_{n-1}]$ is
$\mu_A = ([i_0 \dots i_{n-1}]) = l_{i_0} e^{S_n(A)(x) - n \log \lambda_A} r_{i_{n-1}}$

Chapter 3

Maximizing measures, Ground states and temperature zero

3.1 Selection at temperature zero

3.1.1 The main questions

We remind Definition 20 of a A-maximizing measure: it is a σ-invariant probability μ such that

$$\int A\,d\mu = \max_{\nu\in\mathcal{M}_\sigma}\int A\,d\nu.$$

Existence of maximizing probabilities follows from the compactness of $\mathcal{M}_\sigma$.

Consequently to this definition, the first questions we are interested in are related to maximizing measures. We can for instance address:

1. For a given potential A, how big is the set of maximizing measures ?
2. How can we construct/get maximizing measures ?
3. For a given maximizing measure, how its supports does look like ?

We already mentioned above the relation between maximizing measures and equilibrium states (see Theorem 6): assume A is Hölder continuous. Any accumulation point for μ_β is a A-maximizing measure.

This motivates a new definition:

Definition 28. *Let A be Hölder continuous. A σ-invariant probability measure μ is called a* ground state *(for A) if it is an accumulation point for μ_β as β goes to $+\infty$.*

Remark 2. *The Hölder continuity of A is only required to get uniqueness for the equilibrium state for $\beta.A$.* ■

Clearly a ground state is a A-maximizing measure but *a priori* a maximizing measure is not necessarily a ground state. The natural question we can address are then:

1. Is there convergence for μ_β as β goes to $+\infty$?
2. If there is convergence, how does the family of measures μ_β select the limit ?

Concerning the first question, there is one known example of non-convergence (see [38]). Let us now detail the last question. Obviously, if there is a unique maximizing measure, there is convergence because there is a unique possible accumulation point. Let us thus assume that there are at least two different maximizing measures $\mu_{max,1}$ and $\mu_{max,2}$. Just by linearity, any convex combination of both measure $\mu_t = t\mu_{max,1} + (1-t)\mu_{\text{max},2}$, $t \in [0,1]$, is also a A-maximizing measure. The question of the selection is then to determine why the family (or even a subfamily) will converge to some specific limit, if there are several possible choices.

The notion of *ground states* comes from Statistical Mechanics. Perhaps the most famous example is the supraconductivity phenomenon: consider a one-dimensional lattice with spins which can be up (say +1) or down (say -1). If the temperature decreases, then, spins change and at very low temperature they are all up or down. This modify the magnetic properties of the material. In statistical Mechanics β is the inverse of the temperature, thus, doing $\beta \to +\infty$ means to reach the zero temperature. The goal is thus to furnish mathematical tools to understand why materials have a strong tendency to be highly ordered at low temperature. They reach a crystal or quasi crystal configuration.

A dual viewpoint for maximizing measure is "generic" maximizing orbits. If μ is a maximizing measure, then for μ almost every x,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} A \circ \sigma^j(x) = m(A).$$

It is then normal to study the problem with this viewpoint:

1. How can we detect that an orbit is A-maximizing ?
2. How related to maximizing measures is the set of maximizing orbits ?
3. How are A-maximizing orbits ?

We will see that this “orbital” viewpoint leads to solve a cohomological inequality:

$$A \geqslant m(A) + V \circ T - V. \tag{3.1}$$

Definition 29. *A coboundary is a function of the form $\psi \circ \sigma - \psi$.*

Exercise 12
Show that a coboundary has zero integral for any invariant measure.

3.1.2 Uniqueness and lock-up on periodic orbits

We mentioned above the notion of crystals or quasi-crystals. In our settings crystals means periodic orbit. We will not deal with the question of quasi-crystal. Consequently, an important point we want to emphasize here is the role of the periodic orbits. Nevertheless, the regularity of the potential is of prime importance in that study.

Theorem 11 (see [24])**.** *Generically for the $\mathcal{C}^0$-norm the potential A has a unique maximizing measure. This measure is not-supported by a periodic orbit.*

We shall give the proof of this theorem inspired from [24]. We emphasize that this proofs can be extended (concerning the uniqueness of the maximizing measure) to any separable space. Another proof for Hölder continuous potentials can be found in [41].

Proof. The set $\mathcal{C}^0(\Sigma)$ is separable. Let (ψ_n) be a dense sequence in $\mathcal{C}^0(\Sigma)$. Two different measures, say μ and ν must give different values for some ψ_n. This means

$$\begin{aligned}
\{A,\ \#\mathcal{M}_{max} > 1\} &= \left\{A,\ \exists n\ \exists \mu,\nu \in \mathcal{M}_{max}\ \int \psi_n\,d\nu \neq \int \psi_n\,d\mu\right\}\\
&= \bigcup_n \left\{A,\ \exists \mu,\nu \in \mathcal{M}_{max}\ \int \psi_n\,d\nu \neq \int \psi_n\,d\mu\right\}\\
&= \bigcup_n\bigcup_m \left\{A,\ \exists \mu,\nu \in \mathcal{M}_{max}\ \left|\int \psi_n\,d\nu - \int \psi_n\,d\mu\right| \geqslant \frac{1}{m}\right\}.
\end{aligned}$$

Set $F_{n,m} := \left\{A,\ \exists \mu,\nu \in \mathcal{M}_{max}\ \left|\int \psi_n\,d\nu - \int \psi_n\,d\mu\right| \geqslant \frac{1}{m}\right\}$. We claim that theses sets are closed. For this we need a lemma:

Lemma 30. *Let (A_k) be a sequence of continuous potentials converging to A. Let μ_k be any maximizing measure for A_k and μ be an accumulation point for μ_k.*

Then, $\lim_{k\to+\infty} m(A_k) = m(A)$ and μ is a A-maximizing measure.

Proof of Lemma 30. For any $\varepsilon > 0$, and for k sufficiently big,

$$A - \varepsilon \leqslant A_k \leqslant A + \varepsilon.$$

This shows $m(A) - \varepsilon \leqslant m(A_k) \leqslant m(A) + \varepsilon$. Furthermore, we have $m(A_k) = \int A_k\, d\mu_k$, $m(A) = \int A\, d\mu$ and (up to a subsequence) $\lim_{k\to+\infty} \int A_k\, d\mu_k = \int A\, d\mu$ because μ_k converges to μ for the weak* topology and A_k goes to A for the strong topology. □

We can thus show that the set $F_{n,m}$ is closed in $\mathcal{C}^0(\Sigma)$. Indeed, considering a sequence A_k converging to A (for the strong topology), we get two sequences (μ_k) and (ν_k) of A_k-maximizing measures such that

$$\left| \int \psi_n\, d\mu_k - \int \psi_n\, d\nu_k \right| \geqslant \frac{1}{m}.$$

We pick a subsequence such that μ_k and ν_k converge for this subsequence. Lemma 30 shows that the two limits, say μ and ν are A-maximizing and they satisfy

$$\left| \int \psi_n\, d\mu - \int \psi_n\, d\nu \right| \geqslant \frac{1}{m}.$$

To complete the proof concerning generic uniqueness, we need to prove that the sets $F_{n,m}$ have empty interior.

For such A, the function $\varepsilon \mapsto m(A + \varepsilon.\psi_n)$ is convex but not differentiable at $\varepsilon = 0$. Assume μ and ν are A-maximizing and

$$\int \psi_n\, d\mu \geqslant \int \psi_n\, d\nu + \frac{1}{m},$$

then the right derivative if bigger than $\int \psi_n\, d\mu$ and the left derivative is lower than $\int \psi_n\, d\nu$.

Now, a convex function is derivable Lebesgue everywhere (actually everywhere except on a countable set), which proves that there are infinitely many ε accumulating on 0 such that $A + \varepsilon.\psi_n$ cannot be in $F_{n,m}$.

Let us now prove that generically, the unique maximizing measure is not supported on a periodic orbit.

Let us consider some periodic orbit $\mathcal{O}$ and $\mu_{\mathcal{O}}$ the associated invariant measure. If A is such that $\mu_{\mathcal{O}}$ is not A-maximizing, then for every A_ε closed, $\mu_{\mathcal{O}}$ is still not

A_ε-maximizing (see the proof of Lemma 30). This proves that the set of A such that $\mu_{\mathcal{O}}$ is A-maximizing is a closed set in $\mathcal{C}^0$.

To prove it has empty interior, let consider A such that $\mu_{\mathcal{O}}$ is A-maximizing, and some measure μ closed to $\mu_{\mathcal{O}}$ for the weak* topology. It can be chosen such that $\displaystyle\int A\,d\mu \geqslant m(A) - \varepsilon$ with $\varepsilon > 0$ as small as wanted.

There exists a set K_ε such that $\mu_{\mathcal{O}}(K_\varepsilon) < \varepsilon$ and $\mu(K_\varepsilon) > 1 - \varepsilon$, and then we can find a continuous function $0 \leqslant \psi_\varepsilon \leqslant 1$, null on the periodic orbit $\mathcal{O}$ and such that $\displaystyle\int \psi_\varepsilon\,d\mu > 1 - 2\varepsilon$.

Then,

$$\int A + 2\varepsilon\psi_\varepsilon\,d\mu_{\mathcal{O}} = m(A) < m(A) - \varepsilon + 2\varepsilon - 4\varepsilon^2 \leqslant \int A + 2\varepsilon.\psi_\varepsilon\,d\mu_\varepsilon.$$

This shows that $\mu_{\mathcal{O}}$ is not $A + 2\varepsilon\psi_\varepsilon$-maximizing, thus the set of potential such that $\mu_{\mathcal{O}}$ is maximizing has empty interior.

As there are only countably many periodic orbits, this proves that generically, a periodic orbit is not maximizing. □

Now we mention a recent result which was conjectured for many years:

Theorem 12 (see [39]). *Generically for the Lipschitz norm, the potential A has a unique maximizing measure and it is supported by a periodic orbit*

We have just seen that "generically" the maximizing measure is unique. This solve the problem of convergence. Nevettheless, it is extremely simple to get example with non-uniqueness for the maximizing measure. Let $\mathbb{K}$ be any σ-invariant compact set such that it contains the support of at least two different invariant measure (in other words $\mathbb{K}$ is not uniquely ergodic). Then set

$$A := -d(.,\mathbb{K}).$$

Then A is Lipschitz and any measure with support in $\mathbb{K}$ is A-maximizing. It is so simple to obstruct examples where uniqueness fails that generic uniqueness cannot be seen as sufficient to consider the problem as solved.

Moreover, the problem of selection is so fascinating, that it is just for itself interesting.

3.1.3 First selection: entropy criterion

In this section we consider $A : \Sigma \to \mathbb{R}$ Lipschitz continuous. Note that in that specific case, results also hold for Hölder continuous potentials. We denote by $\mathcal{M}_{max}$ the set of maximizing measures.

Definition 31. *The Mather set of A is the union of the support of all the A-maximizing measures.*

Theorem 13. *Any Ground state has maximal entropy among the set of maximizing measures. In other word, any accumulation point μ_∞ for μ_β satisfies*

$$h_{\mu_\infty} = \max\left\{h_\nu, \nu \in \mathcal{M}_{max}\right\}.$$

Proof. First, note that $\mathcal{M}_{max}$ is closed, thus compact in $\mathcal{M}_\sigma$. The entropy is upper-semi-continuous, then, there exists measures in $\mathcal{M}_{max}$ with maximal entropy.

Let μ_∞ be such a measure, and set $h_{max} = h_{\mu_\infty}$. Then

$$h_{\max} + \beta.m(A) = h_{\mu_\infty} + \beta.\int A\,d\mu_\infty \leqslant \mathcal{P}(\beta). \tag{3.2}$$

We remind that $\beta \mapsto \mathcal{P}(\beta)$ admits an asymptote as $\beta \to +\infty$, which means that there exists some h such that

$$\mathcal{P}(\beta) = h + \beta.m(A) + o(\beta),$$

with $\lim\limits_{\beta\to+\infty} o(\beta) = 0$. Then, Inequality (3.2) shows that $h_{max} \leqslant h$. Now consider μ_∞ any accumulation point for μ_β. Theorem 6 says that it is A-maximizing. On the other hand

$$h + \beta.m(A) + o(\beta) = \mathcal{P}(\beta) = h_{\mu_\beta} + \beta.\int A\,d\mu_\beta \leqslant h_{\mu_\beta} + \beta.m(A)$$

yields $h_{max} \geqslant h_{\mu_\infty} \geqslant \limsup\limits_{\beta\to+\infty} h_{\mu_\beta} \geqslant h \geqslant h_{max}$. □

Remark 3. *In Statistical Mechanics, h_{max} is called the residual entropy: it is the entropy of the system at temperature zero, when it reach its ground state.* ■

Consequently, if the Mather set admits a unique measure of maximal entropy, μ_β converges to this measure as $\beta \to +\infty$.

We have seen in Proposition 24 that the pressure function $\beta \mapsto \mathcal{P}(\beta)$ admits an asymptote as $\beta \to +\infty$. Actually this asymptote is given by

$$h_{max} + \beta.m(A).$$

Theorem 13 justifies the study of the set of maximizing orbits. This is the goal of next section.

3.2 The Aubry set and the subcohomological inequality

3.2.1 Calibrated subactions

Consider $A : \Sigma \to \mathbb{R}$ Lipschitz and $\beta > 0$. The transfer operator yields for every x,

$$e^{\mathcal{P}(\beta)} H_\beta(x) = \sum_{i,\ ix_0 \in \mathcal{A}} e^{\beta.A(ix)} H_\beta(ix), \tag{3.3}$$

where all the subscribed $\beta.A$ have been replaced by β for simplicity.

We have seen that if β goes to $+\infty$ there is control on the constant, but this control exists at the $\frac{1}{\beta}$ log-scale.

Exercise 13
Show that $\lim_{\beta \to +\infty} \frac{1}{\beta} \log(e^{\beta.a} + e^{\beta.b}) = \max(a, b)$.

Actually, Proposition 26 shows that the $\frac{1}{\beta} \log H_\beta$ form an equicontinuous family. We can consider accumulation point, say V, it is Lipschitz continuous. For simplicity we keep writing $\beta \to +\infty$ even if we actually consider subsequences. Then, taking $\frac{1}{\beta}$ log of (3.3), doing $\beta \to +\infty$, Proposition 24 implies

$$m(A) + V(x) = \max_i \{V(ix) + A(ix)\}. \tag{3.4}$$

Definition 32. *A continuous function $u : \Sigma \to \mathbb{R}$ is called a* calibrated subaction *for $A : \Sigma \to \mathbb{R}$, if for any $x \in \Sigma$, we have*

$$u(x) = \max_{\sigma(y)=x} [A(y) + u(y) - m(A)]. \tag{3.5}$$

Consequently any accumulation point for $\frac{1}{\beta} \log H_\beta$ is a calibrated subaction. Now, Equality (3.4) yields for every i and every x such that ix_0 is admissible,

$$A(ix) \leqslant m(A) + V(x) - V(ix),$$

which can be rewritten as $A(y) = m(A) + V \circ \sigma(y) - V(y) + g(y)$, where g is a non-positive function. By Proposition 26 it is also Lipschitz continuous.

Theorem 14. *There exists an invariant set $\mathcal{A}$ called the* Aubry set *such that*

1. *$\mathcal{A}$ contains the Mather set, or equivalently any A-maximizing measure has its support in $\mathcal{A}$.*

2. *Restricted to $\mathcal{A}$, $A - m(A)$ is equal to a Lipschitz coboundary.*

This theorem answers to the question of detection of maximizing orbits. The terminology is borrowed from the Aubry-Mather Theory (see [106] [40] [48] [104] [56]).

Definition 33. *Given A, the Mañé potential is:*

$$S_A(x,y) :=$$

$$\lim_{\epsilon\to 0}\Big[\sup\Big\{\sum_{i=0}^{n-1}\big[A(\sigma^i(z)) - m_A\big]\ \Big|\ n\in\mathbb{N},\ \sigma^n(z)=y,\ d(z,x)<\epsilon\Big\}\Big].$$

Proof of Theorem 14 . We set $\mathcal{A} := \{x \in \Sigma \,|\, S_A(x,x) = 0\}$. Let check that $\mathcal{A}$ satisfies the required conditions.

First of all, let z be in Σ. We can write Then we get

$$S_n(A - m(A))(z) = S_n(g)(z) + V\circ\sigma^n(z) - V(z). \tag{3.6}$$

Now, pick y and consider z such that $\sigma^n(z) = y$. This yields

$$S_n(A - m(A))(z) = S_n(g)(z) + V(y) - V(z) \leqslant g(z) + V(y) - V(z) \leqslant V(y) - V(z). \tag{3.7}$$

Therefore, for every x and y, continuity of V shows

$$S_A(x,y) \leqslant g(x) + V(y) - V(x) \leqslant V(y) - V(x). \tag{3.8}$$

Now, we show that if μ is a A-maximizing ergodic measure then $\operatorname{supp}\mu \subset \mathcal{A}$. This will, in particular, imply that $\mathcal{A}$ is not empty. Consider x generic for μ (and also in the support of μ). V is Lipschitz, thus bounded and x is generic for μ, then Equality (3.6) yields $\lim_{n\to\infty}\frac{1}{n}S_n(A - m(A))(x) = 0$. Consequently g is a non-positive function satisfying $\int g\,d\mu = 0$. Since μ is ergodic and g is continuous, we have $g_{|\operatorname{supp}\mu} \equiv 0$. This implies

$$A(x) = m(A) + V\circ\sigma(x) - V(x),$$

for every $x \in \operatorname{supp}\mu$.

Now, x returns infinitely many times as closed as wanted to itself (see Proposition 18). This yields that for a given $\varepsilon = \frac{1}{2^N}$, there are infinitely many n_i such that

$$x = x_0x_1\ldots x_{n_i-1}x_0x_1\ldots x_{N-1}\ldots.$$

Equivalently, this means that the word $x_0 \ldots x_{N-1}$ appears infinitely many times into $x = x_0x_1 \ldots$.

This yields that for such n_i, the word

$$z = x_0x_1 \ldots x_{n_i}x,$$

coincides with x for at least $n_i + N$ digits. Lipschitz regularity for A and g yields

$$|S_{n_i}(A)(z) - S_{n_i}(A)(x)| \leqslant C.\frac{1}{2^N}, \text{ and } |S_{n_i}(g)(z) - S_{n_i}(g)(x)| \leqslant C.\frac{1}{2^N}.$$

Remind that $g \circ \sigma^k(x) = 0$ for every k because x beings to $\operatorname{supp} \mu$, and that V is also Lipschitz continuous. Therefore we get

$$|S_{n_i}(A - m(A))(z)| \leqslant C\frac{1}{2^N}, \sigma^{n_i}(z) = x \text{ and } d(z,x) \leqslant \frac{1}{2^{n_i+N}}.$$

This yields that $S_A(x,x) \geqslant 0$ and Inequality (3.8) yields $S_A(x,x) \leqslant 0$. Therefore, x belongs to $\mathcal{A}$.

Now, we prove that $A - m(A)$ restricted to $\mathcal{A}$ is a coboundary. Let x be in $\mathcal{A}$. Note that Inequality (3.8) yields

$$0 \leqslant S_A(x,x) \leqslant g(x) \leqslant 0,$$

which show that g is equal to 0 on $\mathcal{A}$. Therefore $A - m(A)$ is a coboundary on $\mathcal{A}$.

The last point to check is that $\mathcal{A}$ is σ-invariant. Let x be in $\mathcal{A}$. Let $\varepsilon_0 > 0$ be fixed and suppose z is such that $\sigma^n(z) = x$ and $d(x,z) < \varepsilon$. Then,

$$\begin{aligned}
S_n(A)(z) &= A(z) + A \circ \sigma(z) + \ldots + A \circ s^{n-1}(z) \\
&= A(\sigma(z)) + \ldots + A \circ \sigma^{n-2}(\sigma(z)) + A(z) \\
&= A(\sigma(z)) + \ldots + A \circ \sigma^{n-2}(\sigma(z)) + A(x) + A(z) - A(x) \\
&= S_n(A)(\sigma(z)) + A(z) - A(x).
\end{aligned}$$

Note that $d(\sigma(z), \sigma(x)) = 2d(z,x) < 2\varepsilon$ and $|A(z) - A(x)| \leqslant C.d(z,x) < C\varepsilon$. Taking the supremum over all the possible n for fixed ε, and then letting $\varepsilon \to 0$ shows

$$S_A(x,x) = S_A(\sigma(x), \sigma(x)).$$

□

Proposition 34. *The Aubry set is compact.*

Proof. The definition of S_A yields

$$S_A(x,x) = \lim_{\varepsilon \to 0} \sup_n \{S_n(g)(z),\ \sigma^n(z) = x\ d(x,z) < \varepsilon\}.$$

The function g is non-positive, thus $S_n(g)(z)$ is always non-positive. To realize $S_A(x,x) = 0$, it can be done if and only if for every n, there exists z, such that $\sigma^n(z) = x$ and $S_n(g)(z) = 0$. Lipschitz continuity of g shows that this condition is closed: if x does not satisfies this condition, there exists n such that for every z, such that $\sigma^n(z) = x$, $S_n(g)(z) < 0$. This is obviously true for every x' sufficiently closed to x. □

3.2.2 Description of the Aubry set. The case of locally constant potential

Here we consider a potential depending on 2 coordinates : $A(x_0x_1x_2\ldots) = A(x_0, x_1)$. We set $A(i,j)$ for $A(x)$ with $x = ij\ldots$.

Note that in that case H_β is constant one each 1-cylinder $[i]$: remind that in that case $\mathcal{L}_\beta$ is just a matrix and H_β is its left dominating eigenvector. This also holds for $\dfrac{1}{\beta}\log H_\beta$ and then for any accumulation point V.

If x is in Σ and z is such that $\sigma^n(z) = x$ and $d(x,z) < 1$, then, $V(x) = V(z)$. Hence,

$$S_n(A - m(A))(z) = S_n(g)(z) \leqslant 0.$$

If y is in $\mathcal{A}$, $g_{|\mathcal{A}} \equiv 0$, then $S_n(A - m(A))(y) = V \circ \sigma^n(y) - V(y)$. If $\sigma^n(y)$ and y are in the same 1-cylinder, then $S_n(A - m(A))(y) = 0$. Furthermore, if z is the periodic orbit given by the concatenation of $y_0y_1\ldots y_{n-1}$, then $S_n(A - m(A))(z) = S_n(A - m(A))(y)$ because all the transitions $z_i \to z_{i+1}$ are the same than for y (for $i \leqslant n-1$).

This shows that $m(A)$ is reached by periodic orbits.

Definition 35. *A periodic orbit obtained as the concatenation of $z_0\ldots z_{n-1}$ is said to be* simple *if all the digits z_i are different.*

Example. $123123123\ldots$ is a simple periodic orbit of length 3. $121412141214\ldots$ is not a simple periodic orbit.

One simple periodic orbit of length n furnishes n *bricks*, that are the words producing the n points of the orbit :

Example. The bricks of $123123123\ldots$ are 123, 231 and 312.

Then, the Aubry set $\mathcal{A}$ is constructed as follows :

1. List all the simple periodic orbits. This is a finite set.

2. Pick the ones such that their Birkhoff means are maximal. This maximal vague is $m(A)$. Such a simple periodic orbit is also said to be maximizing.
3. Consider the associated bricks for all these simple maximizing periodic orbits.
4. The set $\mathcal{A}$ is the subshift of finite type constructed from these bricks:
   (a) Two bricks can be combined if they have a common digit. On one of the simple loop one glues the other simple loop. A brick $x_0x_1\dots x_n$ and $x_ny_1\dots y_k$ produce the new periodic orbit $x_0x_1\dots x_ny_1y_2\dots y_kx_nx_0x_1\dots x_ny_1y_2\dots y_kx_n\dots$

   Example. 123 and 345 produce the new orbit 123453123453....

   (b) $\mathcal{A}$ is the closure of the set of all the periodic orbits obtained with this process.

We can also define $\mathcal{A}$ from its transition matrix. Set $T_{ij}=0$ for every j if i does not appear in any of the bricks. If i appears in a brick, set $T_{ij}=1$ if ij appears in a brick (for $j\neq i$). If i is also a brick (that means that the fixed point $iiiiii\dots$ is a maximizing orbit) set also $T_{ii}=1$. Set $T_{ij}=0$ otherwise. Then $\mathcal{A}=\Sigma_T$.

Example. If the bricks are (up to permutations) abc, cde, fgh, gi and fj, the transition matrix restricted to these letters is

$$T=\begin{pmatrix} 0&1&0&0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0&0&0\\ 1&0&0&1&0&0&0&0&0&0\\ 0&0&0&0&1&0&0&0&0&0\\ 0&0&1&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&1&0&0&1\\ 0&0&0&0&0&0&0&1&1&0\\ 0&0&0&0&0&1&0&0&0&0\\ 0&0&0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0&0&0 \end{pmatrix}.$$

Figure 3.1: The graph for T

The set $\mathcal{A}$ is a subhift of finite type. It can thus be decomposed in irreducible components, say $\mathcal{A}_1\dots,\mathcal{A}_r$. Each component admits a unique measure of maximal entropy, say $\mu_1,\dots,\mu_r$. Let h_i be the associated entropies. We assume that the order has been chosen such that

$$h_1\geqslant h_2\geqslant\dots\geqslant h_r.$$

In that case, the topological entropy for $\mathcal{A}$ is h_1. More precisely, assume that j_0 is such that

$$h_1 = h_2 = \ldots = h_{j_0} > h_{j_0+1} \geqslant h_{j_0+2} \ldots .$$

Then $\mathcal{A}$ admits exactly j_0 ergodic measures of maximal entropy h_1. Any ground state is a convex combination of these j_0 ergodic measures.

In that special case, it is proved that there is only one ground state:

Theorem 15 (see [29, 79, 37])**.** *If A depends only on two coordinates, then μ_β converges as $\beta \to +\infty$.*

Example. In the previous example, $\mathcal{A}$ has two irreducible components. The first one has entropy $\frac{1}{3}\log 2 \sim 0.23$. The second one has entropy ~ 0.398. In that case the ground state is the unique measure of maximal entropy, which has support in the second irreducible component.

3.2.3 Some consequences of Theorem 15

A dual viewpoint for the selection problem is the following. For every β μ_β has full support. In some sense the measure μ_β can be seen as a set of points (the set of Generic points see [105]). This set of point is dense.

However, as $\beta \to +\infty$ this set remains dense but accumulates itself on the Aubry set $\mathcal{A}$. More precisely, it is going to accumulate on the irreducible components which have positive weight for some ground state.

The selection problem is thus to determine what are the components of $\mathcal{A}$ where the set accumulates itself.

It may happen that an irreducible component of $\mathcal{A}$ has maximal entropy but has no weight at temperature zero. In [79], the author introduced the notion of isolation rate between the irreducible components and showed that only the most isolated component have weight at temperature zero.

One natural question is to study the speed of convergence for μ_β. More precisely, if C is a cylinder such that $C \cap \mathcal{A} = \emptyset$, then $\lim\limits_{\beta\to+\infty} \mu_\beta(C) = 0$ (because any possible accumulation point gives 0-measure to C). It is thus natural to study the possible limit:

$$\lim_{\beta\to+\infty} \frac{1}{\beta}\log(\mu_\beta(C)).$$

Chapter 4

The Peierl's barrier and the Large deviation principle

4.1 Irreducible components of the Aubry set

4.1.1 Definition of the Peierl's barrier

We have seen above that if A depends on two coordinates the Aubry set $\mathcal{A}$ is a subshift of finite type. It thus has well-defined irreducible components, each one being the support of a unique measure of maximal entropy. For more general case, there are no reasons why $\mathcal{A}$ should be a subshift of finite type. Actually it can be any invariant subset as it was shown above : pick any compact set $\mathcal{A}$ and consider $A := -d(.,\mathcal{A})$.

In that condition it is far from obvious to define the irreducible components of $\mathcal{A}$ and to determine the measures of maximal entropy.

Definition 36. *The Peierl's barrier between x and y is defined by*

$$h(x,y) := \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\left\{S_n(A-m(A))(z),\ \sigma^n(z)=y\ d(x,z)<\varepsilon\right\}.$$

We remind that we got $A = m(A) + V\circ\sigma - V + g$, where V is a calibrated subaction (obtained via a converging subsequence for $\dfrac{1}{\beta}\log H_\beta$) and g is a non-positive Lipschitz function. Replacing this expression of A into the definition of the Peierl's barrier we get

$$h(x,y) = \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\left\{S_n(g)(z),\ \sigma^n(z)=y\ d(x,z)<\varepsilon\right\}.$$

This shows that to compute $h(x,y)$, we morally have to find a sequence of pre-images for y which converges as fast as possible to x.

We shall see later that of prime importance equality $h(x,y)+h(y,x)=h(x,x)$ is.

Theorem 16. *For any x, the Peierl's barrier $y\mapsto h(x,y)$ is a Lipschitz calibrated subaction. Moreover $h(x,x)=0$ if and only if x belongs to $\mathcal{A}$.*

Proof. Pick x and y. Consider z such that $\sigma^n(z)=y$, $d(x,z)<\varepsilon$. Note that in Σ, z is just the concatenation $z_0\ldots z_{n-1}y$. For y' close to y (namely $y_0=y'_0$), we consider $z':=z_0\ldots z_{n-1}y'$; the Lipschitz regularity for g yields

$$|S_n(g)(z)-S_n(z')|\leqslant C.d(y,y'),$$

for some constant C. If we consider a sequence of z realizing the lim sup and then do $\varepsilon\to 0$, we get

$$h(x,y)\leqslant h(x,y')+C.d(y,y').$$

The same argument shows $h(x,y')\leqslant h(x,y)+C.d(y,y')$ and then $y\mapsto h(x,y)$ is Lipschitz continuous.

Let us show it is a calibrated subaction. For this consider y, n and ε such that $\sigma^n(z)=y$ and $d(x,z)<\varepsilon$. Then, $d(x,z)<\varepsilon$ and $s^{n+1}(z)=\sigma(y)$. Moreover

$$S_{n+1}(A-m(A))(z)=S_n(A-m(A))(z)+A(y)-m(A). \tag{4.1}$$

Considering a sequence of z realizing the lim sup for $h(x,y)$, taking the limit along the subsequence and then doing $\varepsilon\to 0$ we get

$$h(x,y)+A(y)-m(A)\leqslant h(x,\sigma(y)).$$

This shows that $y\mapsto h(x,y)$ is a subaction. It remains to show that for a fixed value for $y'=\sigma(y)$, the equality is achieved for one preimages of y'. This follows from taking the z's and the n's in Equality (4.1) which realize the lim sup for the left hand-side of the equality. Then we get

$$h(x,y')\leqslant h(x,y)+A(y)-m(A),$$

with $\sigma(y)=y'$. As the reverse inequality holds, the global equality holds.

both definition of the Mañé potential and the Peierl's barrier are very similar, except that in one we consider the supremum and in the other we consider a lim sup. This immediatey shows that

$$h(x,y)\leqslant S_A(x,y),$$

and then $h(x,x)=0$ yields $S_A(x,x)=0$ (due to Inequality 3.8), thus x belongs to $\mathcal{A}$. Let us prove the converse. Set x in $\mathcal{A}$. Consider $\rho>0$ small and ε_0 such that for every $\varepsilon<\varepsilon_0$,

$$\sup_n\{S_n(A-m(A))(z),\ \sigma^n(z)=x\ d(z,x)<\varepsilon\}\geqslant-\rho. \tag{4.2}$$

In the following, we assume for simplicity that x is not periodic, but the proof can be easily extended to that case.

Inequality(4.2) holds for every ε. We thus construct a subsequence (n_k) by induction. We pick any ε and consider n_0 realizing the supremum up to $-\rho$. Then we get

$$S_{n_0}(A-m(A))(z_0) \geqslant -2\rho \text{ with } \sigma^{n_0}(z_0)=x \text{ and } d(x,z_0)<\varepsilon.$$

Now we use Inequality(4.2) but with[1] $\varepsilon_1 < \min\{d(x,z), \sigma^n(z)=y\ n\leqslant n_0\}$.

We get $n_1>n_0$ and z_1 such that

$$S_{n_1}(A-m(A))(z_1) \geqslant -2\rho \text{ with } \sigma^{n_1}(z_1)=x \text{ and } d(x,z_1)<\varepsilon_1.$$

We then proceed by induction with $\varepsilon_{k+1} < \min\{d(x,z), \sigma^n(z)=y\ n\leqslant n_k\}$. Then we have

$$\begin{array}{rcl} -2\rho & \leqslant & S_{n_k}(A-m(A))(z_k), \text{ with } \sigma^{n_k}(z_k)=y \text{ and } d(x,z_k)<\varepsilon \\ & \text{then} & \\ -2\rho & \leqslant & \limsup_{n\to\infty}\{S_n(A-m(A))(z),\ \sigma^n(z)=x\ d(x,z)<\varepsilon\}. \end{array}$$

This holds for every $0<\varepsilon<\varepsilon_0$ and then $h(x,x)\geqslant -2\rho$. Then we do $\rho\to 0$, and $h(x,x)\geqslant 0$. the reverse equality is always true, so $h(x,x)=0$. □

4.1.2 Definition of the irreducible components of the Aubry set

Here, we show that the Peierl's barrier allow to define "irreducible" components of the Aubry set.

Lemma 37. *For any x, y and z*

$$h(x,y) \geqslant h(x,z)+h(z,y). \tag{4.3}$$

Proof. Let $\varepsilon>0$ be fixed. Consider a preimage y' of y close to z and a preimage z' of z close to x. For small ε, z' satisfies $z'=z'_0\ldots z'_{n-1}z$ and then $y'':=z'_0\ldots z'_{n-1}y'$ is also a preimage of y. The cocycle relation yields, if $\sigma^m(y')=y$,

$$S_{n+m}(A-m(A))(y'') = S_n(A-m(A))(y'') + S_m(A-m(A))(y').$$

The Lipschitz regularity shows that $S_n(A-m(A))(y'')$ differs from $S_n(A-m(A))(z')$ of a term $\pm C.\varepsilon$. If we assume that the n's and the $m's$ are chosen to realize the respective lim sup, the term form the left-hand side is lower than the lim sup. Then the Lemma is proved. □

[1]Here we use that this distance is positive.

Lemma 38. *For any x in $\mathcal{A}$, $h(x,\sigma(x)) + h(\sigma(x), x) = 0$.*

Proof. Lemma 37 and Theorem 16 show that

$$h(x,\sigma(x)) + h(\sigma(x), x) \leqslant h(x,x) = 0.$$

It remains to prove it is non-negative. The Peierl's barrier is a subaction thus

$$h(x,\sigma(x)) \geqslant A - m(A) + h(x,x) = A - m(A).$$

Now, consider y a preimage of x ε-close to $\sigma(x)$. Say $y = y_0 \dots y_{n-1}x$. Then, $x' := x_0 y_0 \dots y_{n-1}x$ is a preimage of x ε-close[2] to x. In particular $A(x') = A(x) \pm C.\varepsilon$. We emphasize that is is equivalent to have $x' \to x$ or $y \to \sigma(x)$.

We assume that these quantities are chosen such that $S_{n+1}(A-m(A))(x')$ converges to the $\limsup$ if n goes to $+\infty$. Now,

$$S_{n+1}(A - m(A))(x') = A(x') - m(A) + S_n(A - m(A))(y).$$

Doing $n \to +\infty$, the right-hand side term is lower than $h(\sigma(x), x) + A(x) - m(A)$ and the left-hand side term goes to $h(x,x) = 0$. This yields

$$0 \leqslant h(\sigma(x), x) + A(x) - m(A) \leqslant h(\sigma(x), x) + h(x, \sigma(x)).$$

□

Lemma 39. *Let x y and z be in $\mathcal{A}$. If $h(x,y)+h(y,x) = 0$ and $h(y,z)+h(z,y) = 0$, then $h(x,z) + h(z,x) = 0$.*

Proof. Inequality (4.3) shows

$$h(x,z) + h(z,x) \geqslant h(x,y) + h(y,z) + h(z,y) + h(y,x) = 0 + 0 = 0.$$

It also yields $h(x,z) + h(z,x) \leqslant h(x,x) = 0$. □

Lemma 39 proves that $h(x,y) + h(y,x) = 0$ is a transitive relation. Since it is obviously symmetric and reflexive, then it is an equivalence relation on $\mathcal{A}$.

Definition 40. *The equivalence classes for the relation*

$$h(x,y) + h(y,x) = 0,$$

are called irreducible components of $\mathcal{A}$.

Note that x and $\sigma(x)$ belong to the same class, which shows that the classes are invariant. The continuity for the Peierl's barrier has been proved with respect to the second variable for a fixed first variable. It is thus not clear that an irreducible component is closed.

[2] Actually $\frac{\varepsilon}{2}$-close to x.

4.1.3 The locally constant case

If the potential is locally constant, we have seen in Subsection 3.2.2 that the Aubry set $\mathcal{A}$ is a subshift of finite type, for which the notion of irreducible component has already been defined (see 4). We have to check that the two notions coincide.

We have seen that the irreducible components of a subshift of finite type are exactly the transitive components. We shall use this description to show that irreducible components of the Aubry (in the sense of the Peierl's barrier) set are the irreducible components (with respect to subshifts).

Lemma 41. *Assume A depend only on two coordinates. Let x and y be in Σ. Assume that $d(z,x) \leqslant \frac{1}{4}$, $\sigma^n(z) = y$ and there exists $1 \leqslant k < n$ such that $d(\sigma^k(z), x) \leqslant \frac{1}{4}$. Then,*

$$S_n(A - m_A)(z) \leqslant S_{n-k}(A - m_A)(\sigma^k(z)).$$

Proof. We consider a calibrated subaction V. We have seen it only depends on one coordinate. We remind that for every ξ,

$$A(\xi) = m_A + V \circ \sigma(\xi) - V(\xi) + g(\xi) \tag{4.4}$$

holds, where g is a non-positive function (depending only on 2 coordinates). Now we have

$$S_k(A - m_A)(z) = S_k(g)(z) + V(\sigma^k(z)) - V(z) = S_k(g)(z) \leqslant 0.$$

□

From Lemma 41 we claim that for every x and y,

$$h(x,y) = S_A(x,y) = \max\left\{ S_n(A - m_A)(z)\, \sigma^n(z) = y\ d(z,x) \leqslant \frac{1}{4} \right\}. \tag{4.5}$$

Lemma 42. *Let consider some y in Σ. The map $x \mapsto h(x,y)$ is continuous*

Proof. Consider a sequence (x_n) converging to x. Assume that all these x_n and x coincide for at least 2 digits. Then,

$$d(z, x_n) \leqslant \frac{1}{4} \iff d(z,x) \leqslant \frac{1}{4}.$$

For y and n, consider any z realizing the maximum into the definition of $S_A(x_n, y)$. It also realizes the maximum for $S_A(x,y)$ and more generally for every $S_A(x_k, y)$. □

Lemma 42 shows that any irreducible components of the Aubry in the sense of definition 40 is closed. It is also invariant. We thus have to show it is transitive.

Let us consider some components and pick two open sets U and V (for the components). We still assume that Equality (4.4) holds. Consider $x \in U \cap \mathcal{A}$ and $y \in V \cap \mathcal{A}$. By definition

$$h(x, y) + h(y, x) = 0.$$

By Equality (4.5), $h(x, y)$ is realized by some $S_n(A - m_A)(z)$ and $h(y, x)$ is realized by some $S_m(A - m_A)(z')$. Moreover we can also assume that z belongs to U and z' belongs to V.

Indeed, if it is not the case, we can always follow preimages of x in the component $\mathcal{A}$ which are exactly on the set $g^{-1}(\{0\})$.

From the two pieces of orbits $z, \sigma(z), \ldots, \sigma^n(z)$ and $z', \sigma(z'), \ldots, \sigma^m(z')$ we can construct a periodic orbit. Denote by ξ the point of this periodic orbit in U. Then we claim that

$$h(y, \xi) = h(y, x) = S_m(A - m_A)(z') \text{ and } h(\xi, y) = h(x, y) = S_n(A - m_A)(z).$$

This shows that ξ belongs to $\mathcal{A}$ and to the same component than y. Therefore $\sigma^{-m}(U) \cap V \neq \emptyset$ and the component is transitive.

4.2 On the road to solve the subcohomological inequality

4.2.1 Peierl's barrier and calibrated subactions

The Peierl's barrier satisfies an important property. It shows that a calibrated subaction is entirely determined by its values on the Aubry set.

Theorem 17 (see [50]Th. 10)**.** *] Any calibrated subaction u satisfies for any y*

$$u(y) = \sup_{\mathbf{x} \in \mathcal{A}} [h(\mathbf{x}, y) + u(\mathbf{x})], \tag{4.6}$$

Proof. We show that any calibrated subaction u is entirely determined by its values on the Aubry set $\mathcal{A}$.

Let us thus consider some calibrated subaction u. Let y be in Σ. Let y_{-1} be any preimage of y such that

$$u(y) = A(y_{-1}) - m(A) + u(y_{-1}).$$

More generally we consider a sequence y_{-n} such that $\sigma(y_{-n}) = y_{-n+1}$ and

$$u(y_{-n+1}) = A(y_{-n}) - m(A) + u(y_{-n}).$$

We claim that any accumulation point for (y_{-n}) belongs to $\mathcal{A}$. Indeed, let us consider some converging sequence y_{-n_k}, converging to x.

For $k' > k$ we get $\sigma^{n_{k'}-n_k}(y_{-n_{k'}}) = y_{-n_k}$. Therefore we have

$$S_{n_{k'}-n_k}(A - m(A))(y_{-n_{k'}}) = u(y_{-n_k}) - u(y_{-n_{k'}}),$$

and $y_{n_k} \to x$ yields

$$S_A(x, x) \geqslant 0.$$

since $S_A(x, x) \leqslant 0$ always holds (see Inequality (3.8)), this shows that the limit point x belongs to $\mathcal{A}$.

Now we have $S_{n_k}(A - m(A))(y_{n_k}) = u(y) - u(y_{n_k})$, which yields

$$h(x, y) \geqslant u(y) - u(x).$$

In particular $u(y) \leqslant \sup_{x' \in \mathcal{A}}\{h(x', y) + u(x')\}$ holds.

Actually the reasoning we have just done allow to get a finer result. Consider z in Σ. the we can write

$$A(z) = m(A) + u \circ \sigma(z) - u(z) + g(z),$$

where g is a non-positive Lipschitz function. Now, consider z_n such that $\sigma^n(z) = y$. We get

$$S_n(A - m(A))(z_n) = u(y) - u(z) + S_n(g)(z) \leqslant u(y) - u(z).$$

This shows that

$$h(x', y) \leqslant u(y) - u(x')$$

always holds (consider any x' and take a subsequence of z_n converging to it). Then Equality 4.6 holds. □

Moreover, the irreducible component get a special importance :

Theorem 18. *If x and z are in the same irreducible component of $\mathcal{A}$, then for any y,*

$$h(x, y) + u(x) = h(z, y) + u(z).$$

Proof. We remind tow inequalities and one equality:

$$\begin{aligned} h(x, y) &\geqslant h(x, z) + h(z, y), \\ u(x) \geqslant h(z, x) + u(z), & \\ h(x, z) + h(z, x) &= 0. \end{aligned}$$

Then,

$$\begin{aligned} u(x)+h(x,y) &\geqslant u(x)+h(x,z)+h(z,y) \\ &\geqslant u(x)-h(z,x)+h(z,y)+u(z)-u(z)=u(x)-u(z)-h(z,x)+h(z,y)+u(z) \\ &\geqslant h(z,y)+u(z). \end{aligned}$$

Exchanging the roles of x and z shows that the reverse inequality also holds. □

4.2.2 Selection of calibrated subactions

We remind that for a given $\beta>0$, the equilibrium state μ_β is also a Gibbs measure obtained by the product of the *eigenfunction* H_β and the *eigenprobability*[3] ν_β. We also remind that any accumulation point for the family $\dfrac{1}{\beta}\log H_\beta$ is a calibrated subaction.

It is thus natural to study the *selection* of the calibrated subaction:

1. What are the accumulation points for $\dfrac{1}{\beta}\log H_\beta$ as $\beta\to+\infty$?
2. Is there convergence for $\dfrac{1}{\beta}\log H_\beta$ as $\beta\to+\infty$?

Uniqueness of the maximizing measure gives a partial answer to these questions:

Theorem 19. *Assume that there is a unique A-maximizing measure, then all the calibrated subactions are equal up to an additive constant.*

Proof. In that case $\mathcal{A}$ is uniquely ergodic and has thus a single irreducible component. If x_0 is any point of $\mathcal{A}$, Theorems 18 and 17 show that any calibrated subaction is entirely determined by its value on x_0. □

We point out that even in that simple case, the convergence for $\dfrac{1}{\beta}\log H_\beta$ is not clear.

In that direction we mention one of the results in [81]. For simplicity we state it using the setting of [12].

Theorem 20 (see [81])**.** *Assume that* $\Sigma=\{1,2,3\}^{\mathbb{N}}$ *and* A *satisfies*

$$A(x):=\begin{cases} -d(x,1^\infty) \ \textit{if}\ x=1\ldots, \\ -3d(x,2^\infty) \ \textit{if}\ x=2\ldots, \\ -\alpha<0 \ \textit{if}\ x=3\ldots. \end{cases}$$

[3] which is also the conformal measure.

Then, $\nu_\beta \to \delta_{1^\infty}$ *as* $\beta \to +\infty$ *and* $\frac{1}{\beta}\log H_\beta$ *converges.*

This theorem shows that flatness is a selection criterion: the Aubry set in that case is reduced to $\{1^\infty\}\cup\{2^\infty\}$ and the two unique ergodic maximizing measures are the dirac measures δ_{1^∞} and δ_{2^∞}. The potential is "more flat" in 1^∞ than in 2^∞. Then the Theorem says that the locus where the potential is flatter wins the eigenmeasure. In that case it is sufficient to determine all the calibrated subactions.

More generally if the Aubry set $\mathcal{A}$ is not a subshift of finite type, the problem concerning selection is that

1. there is no theory for measure of maximal entropy for general subshifts.
2. there is no necessary existence or uniqueness of conformal measure (one of the key point in Theorem 20 to select calibrated subactions).

We shall also see that the problem of selection of subaction is related to the multiplicity of an eigenvector in the Max-Plus formalism.

4.3 Large deviation for the Gibbs measure

In the study of Large Deviations when temperature goes to zero one is interested in the limits of the following form:

$$\lim_{\beta\to\infty}\frac{1}{\beta}\log\mu_\beta(C), \tag{4.7}$$

where C is a for a fixed cylinder on Σ.

In principle, the limit may not exist. We remind that general references in Large deviation are [45] [46]).

Definition 43. *We say there exists a Large Deviation Principle for the one parameter family* μ_β, $\beta>0$, *if there exists a non-negative function* I, *where* $I:\Sigma\to\mathbb{R}\cup\{\infty\}$ *(which can have value equal to infinity in some points), which it is lower semi-continuous and for any cylinder set* $C\subset\Sigma$,

$$\lim_{\beta\to+\infty}\frac{1}{\beta}\log\mu_{\beta\,A}(C) = -\inf_{x\in C} I(x).$$

In the affirmative case an important point is to be able to identify such function I.

Theorem 21 (see [9, 99]). *Assume that A admits a unique maximizing measure. Let V be a calibrated subaction. Then, for any cylinder $[i_0 i_1 ... i_n]$, we have*

$$\lim_{\beta \to \infty} \frac{1}{\beta} \log \mu_\beta([i_0 i_1 ... i_n]) = - \inf_{x \in [i_0 i_1 ... i_n]} \{I(x)\},$$

where

$$I(x) = \sum_{n=0}^{\infty} [\, V \circ \sigma - V - (\, A - m(A))\,]\, \sigma^n\,(x).$$

In the case the potential A depends on two coordinates and the maximizing probability is unique we get

$$I(x) = I(x_0, x_1 ..., x_k, ..) = \sum_{j=0}^{\infty} [\, V(x_{j+1}) - V(x_j) - A(x_j, x_{j+1}) + m(A)\,].$$

Chapter 5

One explicit example with all the computations

Preliminaries: Max-Plus algebra We equip $\mathbb{R} \cup \{-\infty\}$ with two new binary operators :

1. $\oplus$ is defined my $a \oplus b = \max(a, b)$, with the rule $-\infty < a$, for any $a \in \mathbb{R}$.
2. $\otimes$ is defined by $a \otimes b = a + b$ with the rule $a \otimes (-\infty) = -\infty$.

We claim that $(\mathbb{R} \cup \{-\infty\}, \oplus, \otimes)$ is a commutative algebra. The two neutral elements are respectively $-\infty$ and 0.

Lemma 44. *Let a, b, c and d be real numbers. Let v_1 and v_2 and t be such that*

$$t \otimes \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Then, $t = \max(a, d, \frac{b+c}{2})$.

Proof. We get the two equations

$$v_1 + t = \max(v_1 + a, v_2 + b) \quad v_2 + t = \max(v_1 + c, v_2 + d).$$

This yields $t \geqslant a$ and $t \geqslant d$. From $v_1 + t \geqslant v_2 + b$ and $v_2 + t \geqslant v_1 + c$ we get $2t \geqslant c + b$. This proves $t \geqslant \max(a, d, \frac{b+c}{2})$.

Assume that $t > \max(a, d)$, then we must have $v_1 + t = v_2 + b$ and $v_2 + t = v_1 + c$, which finishes the proof. □

The example. We consider the following particular case: $\Sigma := \{1,2,3\}^{\mathbb{N}}$ and A is a non-positive potential depending only on two coordinates. For each pair (i,j) we set $A(i,j) = -\varepsilon_{ij}$. We assume that

$$\varepsilon_{11} = \varepsilon_{22} = 0,$$

and for every other pair $\varepsilon_{ij} > 0$.

Note that there are only two ergodic A-maximizing measures, namely, δ_{1^∞} and δ_{2^∞}, which are the Dirac measures at $1^\infty := 111\ldots$ and at $2^\infty := 222\ldots$. The Aubry set is exactly the union of these two fixed points and each one is an irreducible component.

We remind that for each β, the unique equilibrium state is given by

$$\mu_\beta = H_\beta \nu_\beta,$$

where H_β and ν_β are the eigenvectors for the transfer operator $\mathcal{L}_\beta$. Its dominating single eigenvalue (*i.e.,* its spectral radius) is $e^{\mathcal{P}(\beta)}$.

We have seen in Lemma 23 that $\mathcal{P}'(\beta) = \displaystyle\int A\, d\mu_\beta$. This quantity is negative because A is non-positive and negative on a set of positive measure (remind that for any β μ_β has full support). We have seen (see the comments after Remark 3) that the asymptote of the pressure is of the form

$$h_{max} + \beta.m(A),$$

where h_{max} is the residual entropy and is the entropy of the Aubry set. In our case, we have $h_{max} = 0$ and $m(A) = 0$. Then, $\displaystyle\lim_{\beta\to+\infty} \mathcal{P}(\beta) = 0$. The first role of the max-plus algebra seems to determine how the pressure goes to 0 as β goes to $+\infty$.

Proposition 45. *There exists a positive sub exponential function g and a positive real number ρ such that $\mathcal{P}(\beta) = g(\beta)e^{-\rho.\beta}$.*

Proof. First we consider any accumulation point $-\rho$ for $\dfrac{1}{\beta}\log \mathcal{P}(\beta)$. We shall prove that this $-\rho$ is actually unique, namely, it does not depend on the chosen subsequence.

First, we consider a subsequence which realizes $-\rho$ in $\displaystyle\lim_{\beta\to+\infty} \frac{1}{\beta}\log(e^{\mathcal{P}(\beta)} - 1)$. By Proposition 26, the family $\left\{\dfrac{1}{\beta}\log H_\beta\right\}$ is relatively compact, and we can always extract extract another subsequence such that $\dfrac{1}{\beta}\log H_\beta$ converges. We denote by V this limit. V is a calibrated subaction.

For simplicity we shall write $\beta \to +\infty$ even if we consider a restricted subsequence.

Moreover, H_β and V only depend on one coordinate, and we shall write $H_\beta(i)$ or $V(i)$ for $H_\beta(x)$ and $V(x)$ with $x = i\ldots$.

The eigenfunction H_β is an eigenvector for $\mathcal{L}_\beta$ and this yields

$$\begin{aligned} e^{\mathcal{P}(\beta)}H_\beta(1) &= e^{\beta.A(1,1)}H_\beta(1) + e^{\beta.A(2,1)}H_\beta(2) + e^{\beta.A(3,1)}H_\beta(3) \\ e^{\mathcal{P}(\beta)}H_\beta(2) &= e^{\beta.A(1,2)}H_\beta(1) + e^{\beta.A(2,2)}H_\beta(2) + e^{\beta.A(3,2)}H_\beta(3). \end{aligned}$$

Replacing with the values for A we get the two following equations:

$$(e^{\mathcal{P}(\beta)} - 1)H_\beta(1) = e^{-\beta\varepsilon_{21}}H_\beta(2) + e^{-\beta\varepsilon_{31}}H_\beta(3), \tag{5.1a}$$

$$(e^{\mathcal{P}(\beta)} - 1)H_\beta(2) = e^{-\beta\varepsilon_{12}}H_\beta(1) + e^{-\beta\varepsilon_{32}}H_\beta(3). \tag{5.1b}$$

Now, $\dfrac{\mathcal{P}(\beta)}{\beta} \to_{\beta\to+\infty} 0$ yields $\displaystyle\lim_{\beta\to+\infty} \frac{e^{\mathcal{P}(\beta)} - 1}{\mathcal{P}(\beta)} = 1$ and, finally, $\displaystyle\lim_{\beta\to+\infty} \frac{1}{\beta}\log(e^{\mathcal{P}(\beta)} - 1) = -\rho$.

Taking $\dfrac{1}{\beta}\log$ and doing $\beta \to +\infty$ in (5.1a) and (5.1b) we get

$$-\rho + V(1) = \max(-\varepsilon_{21}V(2), -\varepsilon_{31}V(3)), \tag{5.2a}$$

$$-\rho + V(2) = \max(-\varepsilon_{12}V(1), -\varepsilon_{32}V(3)), \tag{5.2b}$$

which can be written under the form

$$-\rho \otimes \begin{pmatrix} V(1) \\ V(2) \end{pmatrix} = \begin{pmatrix} -\infty & -\varepsilon_{21} & -\varepsilon_{31} \\ -\varepsilon_{12} & -\infty & -\varepsilon_{32} \end{pmatrix} \otimes \begin{pmatrix} V(1) \\ V(2) \\ V(3) \end{pmatrix}. \tag{5.3}$$

Now, we use Theorem 17 to get an expression of $V(3)$ in terms of $V(1)$ and $V(2)$. Indeed, we have

$$V(3) = \max(V(1) + h(1^\infty, 3), V(2) + h(2^\infty, 3)),$$

where h is the Peierl's barrier and 3 means any point starting with 3. Copying the work done to get Equality (4.5), we claim that

$$h(1^\infty, 3) = -\varepsilon_{13} \text{ and } h(2^\infty, 3) = -\varepsilon_{23}.$$

This yields,

$$\begin{pmatrix} V(1) \\ V(2) \\ V(3) \end{pmatrix} = \begin{pmatrix} 0 & -\infty \\ -\infty & 0 \\ -\varepsilon_{13} & -\varepsilon_{23} \end{pmatrix} \otimes \begin{pmatrix} V(1) \\ V(2) \end{pmatrix}. \tag{5.4}$$

Merging (5.3) and (5.4) we finally get

$$-\rho\otimes\begin{pmatrix}V(1)\\V(2)\end{pmatrix}=\begin{pmatrix}-\varepsilon_{13}-\varepsilon_{31} & -\varepsilon_{21}\oplus(-\varepsilon_{31}-\varepsilon_{23})\\ -\varepsilon_{12}\oplus(-\varepsilon_{32}-\varepsilon_{13}) & -\varepsilon_{23}-\varepsilon_{32}\end{pmatrix}\otimes\begin{pmatrix}V(1)\\V(2)\end{pmatrix}. \tag{5.5}$$

Then, Lemma 44 shows that $\frac{1}{\beta}\log\mathcal{P}(\beta)$ admits a unique accumulation point as $\beta\to+\infty$, thus converges.

Setting $g(\beta):=\mathcal{P}(\beta).e^{\rho\beta}$ we get the proof of the proposition. □

From Lemma 44 we also get

$$-\rho=\max\begin{cases}A(1,3)+A(3,1)=-\varepsilon_{13}-\varepsilon_{31},\\[2ex] A(3,2)+A(2,3)=-\varepsilon_{32}-\varepsilon_{23},\\[2ex] \dfrac{A(2,1)+A(1,2)}{2}=-\dfrac{\varepsilon_{12}+\varepsilon_{21}}{2},\\[2ex] \dfrac{A(2,1)+A(1,3)+A(3,2)}{2}=-\dfrac{\varepsilon_{21}+\varepsilon_{13}+\varepsilon_{32}}{2},\\[2ex] \dfrac{A(1,2)+A(2,3)+A(3,1)}{2}=-\dfrac{\varepsilon_{12}+\varepsilon_{23}+\varepsilon_{31}}{2},\\[2ex] \dfrac{A(2,3)+A(3,1)+A(1,3)+A(3,2)}{2}.\end{cases} \tag{5.6}$$

We emphasize that the last quantity is actually the mean value of the two first ones, and then $-\rho\geqslant\frac{A(2,3)+A(3,1)+A(1,3)+A(3,2)}{2}$ always holds, as soon as, $-\rho\geqslant A(1,3)+A(3,1)$ and $-\rho\geqslant A(3,2)+A(2,3)$ hold.

We can now finish the proof of the convergence of μ_β. We recall that any accumulation point must be of the form

$$\alpha.\delta_{1^\infty}+(1-\alpha)\delta_{2^\infty},$$

with $\alpha\in[0,1]$. It is thus sufficient to show that $\dfrac{\mu_\beta([1])}{\mu_\beta([2])}$ converges as $\beta\to+\infty$ to prove the convergence of μ_β. However, we emphasize that this is very particular to our case (two ergodic A-maximizing measures). This reasoning may not work for a more general case. Nevertheless, one of the by-product results of our proof is that for the general case, it seems possible to determine the convergence we get here, in a similar (but more complex) way. The complexity is an issue which is due, essentially, to the large amount of possible combinatorics.

First, we get some equations for the measure ν_β. We remind that this measure is $\beta.A$-conformal. Then, equality 2.3 yields,

$$\begin{aligned}
\nu_\beta([1]) &= \nu_\beta\left(\bigsqcup_{n=1}^{\infty}[1^n2]\sqcup[1^n3]\right)\\
&= \sum_{n=1}^{+\infty}\nu_\beta([1^n2])+\nu_\beta([1^n3])\\
&= \sum_{n=1}^{+\infty}e^{\beta.S_n(A)(1^n2)-n\mathcal{P}(\beta)}\nu_\beta([2])+\sum_{n=1}^{+\infty}e^{\beta.S_n(A)(1^n3)-n\mathcal{P}(\beta)}\nu_\beta([3])\\
&= \frac{e^{-\beta\varepsilon_{12}-\mathcal{P}(\beta)}}{1-e^{-\mathcal{P}(\beta)}}\nu_\beta([2])+\frac{e^{-\beta\varepsilon_{13}-\mathcal{P}(\beta)}}{1-e^{-\mathcal{P}(\beta)}}\nu_\beta([3]).
\end{aligned}$$

We remind that $\nu_\beta([1])+\nu_\beta([2])+\nu_\beta([3])=1$, then, we get a linear equation between $\nu_\beta([1])$ and $\nu_\beta([2])$. Doing the same work with the cylinder $[2]=\bigsqcup_{n=1}^{+\infty}[2^n1]\sqcup[2^n3]$ we get the following system:

$$\begin{cases}(e^{\mathcal{P}(\beta)}-1+e^{-\beta.\varepsilon_{13}})\nu_\beta([1])+(e^{-\beta.\varepsilon_{13}}-e^{-\beta.\varepsilon_{12}})\nu_\beta([2])=e^{-\beta.\varepsilon_{13}},\\ (e^{-\beta.\varepsilon_{23}}-e^{-\beta.\varepsilon_{21}})\nu_\beta([1])+(e^{\mathcal{P}(\beta)}-1+e^{-\beta.\varepsilon_{23}})\nu_\beta([2])=e^{-\beta.\varepsilon_{23}}.\end{cases} \tag{5.7}$$

The determinant of the system is

$$\begin{aligned}\Delta(\beta):=(e^{\mathcal{P}(\beta)}-1)^2+(e^{\mathcal{P}(\beta)}-1)(e^{-\beta.\varepsilon_{13}}+e^{-\beta.\varepsilon_{23}})+\\ e^{-\beta(\varepsilon_{12}+\varepsilon_{23})}+e^{-\beta.(\varepsilon_{21}+\varepsilon_{13})}-e^{\beta.(\varepsilon_{12}+\varepsilon_{21})},\end{aligned}$$

and, we get

$$\frac{\nu_\beta([1])}{\nu_\beta([2])}=\frac{(e^{\mathcal{P}(\beta)}-1)e^{-\beta.\varepsilon_{13}}+e^{-\beta.(\varepsilon_{12}+\varepsilon_{23})}}{(e^{\mathcal{P}(\beta)}-1)e^{-\beta.\varepsilon_{23}}+e^{-\beta.(\varepsilon_{21}+\varepsilon_{13})}}. \tag{5.8}$$

On the other hand, Equations (5.2a) and (5.2a) yield the following formula:

$$\frac{H_\beta(1)}{H_\beta(2)}=\frac{(e^{\mathcal{P}(\beta)}-1)e^{-\beta.\varepsilon_{31}}+e^{-\beta.(\varepsilon_{21}+\varepsilon_{32})}}{(e^{\mathcal{P}(\beta)}-1)e^{-\beta.\varepsilon_{32}}+e^{-\beta.(\varepsilon_{12}+\varepsilon_{31})}}. \tag{5.9}$$

Therefore, from Equations (5.8) and (5.9) we get

$$\begin{aligned}
\frac{\mu_\beta([1])}{\mu_\beta([2])} &= \frac{H_\beta(1)}{H_\beta(2)}\frac{\nu_\beta([1])}{\nu_\beta([2])}\\
&= \frac{(e^{\mathcal{P}(\beta)}-1)e^{-\beta.\varepsilon_{13}}+e^{-\beta.(\varepsilon_{12}+\varepsilon_{23})}}{(e^{\mathcal{P}(\beta)}-1)e^{-\beta.\varepsilon_{23}}+e^{-\beta.(\varepsilon_{21}+\varepsilon_{13})}}\frac{(e^{\mathcal{P}(\beta)}-1)e^{-\beta.\varepsilon_{31}}+e^{-\beta.(\varepsilon_{21}+\varepsilon_{32})}}{(e^{\mathcal{P}(\beta)}-1)e^{-\beta.\varepsilon_{32}}+e^{-\beta.(\varepsilon_{12}+\varepsilon_{31})}}\\
&= \frac{\left((e^{\mathcal{P}(\beta)}-1)e^{\beta.(\varepsilon_{12}+\varepsilon_{21}-\varepsilon_{13})}+e^{\beta.(\varepsilon_{21}-\varepsilon_{23})}\right)}{\left((e^{\mathcal{P}(\beta)}-1)e^{\beta.(\varepsilon_{12}+\varepsilon_{21}-\varepsilon_{23})}+e^{\beta.(\varepsilon_{12}-\varepsilon_{13})}\right)}\times\\
&\qquad \frac{\left((e^{\mathcal{P}(\beta)}-1)e^{\beta.(\varepsilon_{12}+\varepsilon_{21}-\varepsilon_{31})}+e^{\beta.(\varepsilon_{12}-\varepsilon_{32})}\right)}{\left((e^{\mathcal{P}(\beta)}-1)e^{\beta.(\varepsilon_{12}+\varepsilon_{21}-\varepsilon_{32})}+e^{\beta.(\varepsilon_{21}-\varepsilon_{31})}\right)}.
\end{aligned} \tag{5.10}$$

Convergence will follow from the next proposition.

Proposition 46. *The function g admits a limit as β goes to $+\infty$.*

Proof. We remind that ν_β is the eigenmeasure for the dual transfer operator. This yields:

$$\begin{aligned}
e^{\mathcal{P}(\beta)} &= \int \mathcal{L}_\beta(\mathbb{1})\,d\nu_\beta\\
&= (1+e^{-\beta.\varepsilon_{21}}+e^{-\beta.\varepsilon_{31}})\nu_\beta([1])+(1+e^{-\beta.\varepsilon_{12}}+e^{-\beta.\varepsilon_{32}})\nu_\beta([2])\\
&\qquad +(e^{-\beta.\varepsilon_{13}}+e^{-\beta.\varepsilon_{23}}+e^{-\beta.\varepsilon_{33}})\nu_\beta([3]).
\end{aligned} \tag{5.11}$$

Let us set

$$\begin{cases}
X := e^{\mathcal{P}(\beta)},\\
A := e^{-\beta.\varepsilon_{13}},\\
A' := e^{-\beta.\varepsilon_{23}},\\
B := e^{-\beta.(\varepsilon_{12}+\varepsilon_{23})},\\
B' := e^{-\beta.(\varepsilon_{21}+\varepsilon_{13})},\\
C := e^{-\beta.(\varepsilon_{12}+\varepsilon_{21})},\\
a := e^{-\beta.\varepsilon_{21}}+e^{-\beta.\varepsilon_{31}},\\
b := e^{-\beta.\varepsilon_{12}}+e^{-\beta.\varepsilon_{32}},\\
c := e^{-\beta.\varepsilon_{13}}+e^{-\beta.\varepsilon_{23}}+e^{-\beta.\varepsilon_{33}}.
\end{cases}$$

From the system (5.7), we get exact values for $\nu_\beta([1])$, $\nu_\beta([2])$ and $\nu_\beta([3])$. Replacing these values in (5.11), this yields

$$X = \frac{(1+a)(A(X-1)+B)+(1+b)(A'(X-1)+B')+c((X-1)^2-C)}{(X-1)^2+(A+A')(X-1)+B+B'-C}. \tag{5.12}$$

This can also be written

$$X-1=\frac{a(A(X-1)+B)+b(A'(X-1)+B')+(c-1)((X-1)^2-C)}{(X-1)^2+(A+A')(X-1)+B+B'-C},$$

and this yields

$$\begin{aligned}(X-1)^3+(A+A'+1-c)(X-1)^2\\ +(B+B'-C-a.A-b.A')(X-1)+C(c-1)-a.B-b.B'=0.\end{aligned} \tag{5.13}$$

Remember that all the terms $A, A', B, \dots$ go exponentially fast to 0 as $\beta \to +\infty$, $X-1$ behaves like $g(\beta)e^{-\rho.\beta}$, and, moreover g is sub-exponential. We can thus use Taylor development to replace $X-1$ by $g(\beta)e^{-\rho.\beta}$, and, keep in each summand of (5.13) the largest term; larger here means that we are comparing it to other terms, if there is a sum, but also to terms of the other summands.

• Terms in $(X-1)^2$ and $(X-1)^3$. Note that A, A' and c go exponentially fast to 0 as $\beta \to +\infty$, which shows that $(A+A'+1-c)(X-1)^2$ has as fordominating term $g^2(\beta)e^{-2\rho.\beta}$, whereas $(X-1)^3$ has as dominating term $g^3(\beta)e^{-3\rho.\beta}$, which is exponentially smaller than $g^2(\beta)e^{-2\rho.\beta}$.

• Term in $(X-1)$. Note that the term B is cancelled by part of $b.A'$ and similarly B' is cancelled by part of $a.A$. The term in $X-1$ is actually equal to

$$\begin{aligned}e^{-\beta.(\varepsilon_{12}+\varepsilon_{23})}+e^{-\beta.(\varepsilon_{21}+\varepsilon_{13})}-e^{-\beta(\varepsilon_{12}+\varepsilon_{21})}\\ -e^{-\beta.(\varepsilon_{21}+\varepsilon_{13})}-e^{-\beta(\varepsilon_{13}+\varepsilon_{31})}-e^{-\beta.(\varepsilon_{12}+\varepsilon_{23})}-e^{-\beta(\varepsilon_{23}+\varepsilon_{32})}\\ =-e^{-\beta(\varepsilon_{12}+\varepsilon_{21})}-e^{-\beta(\varepsilon_{13}+\varepsilon_{31})}-e^{-\beta(\varepsilon_{23}+\varepsilon_{32})}.\end{aligned}$$

Remind that this term will be multiplied by $g(\beta)e^{-\rho.\beta}$ and compared to $g^2(\beta)e^{-2\rho.\beta}$. The fact that $\rho \leqslant \dfrac{\varepsilon_{12}+\varepsilon_{21}}{2}$ shows that the term $-e^{-\beta(\varepsilon_{12}+\varepsilon_{21})}(X-1)$ is exponentially smaller than $g^2(\beta)e^{-2\rho.\beta}$.

• Term without $(X-1)$. With a change of sign it is equal to

$$\begin{aligned}e^{-\beta.(\varepsilon_{12}+\varepsilon_{21})}\oplus e^{-\beta(\varepsilon_{12}+\varepsilon_{23}+\varepsilon_{21})}\oplus e^{-\beta(\varepsilon_{12}+\varepsilon_{23}+\varepsilon_{31})}\\ \oplus e^{-\beta(\varepsilon_{21}+\varepsilon_{12}+\varepsilon_{13})}\oplus e^{-\beta(\varepsilon_{21}+\varepsilon_{13}+\varepsilon_{32})}.\end{aligned}$$

Note that the second and the fourth terms are exponentially smaller than the first one. Therefore, the remaining term is

$$e^{-\beta.(\varepsilon_{12}+\varepsilon_{21})}\oplus e^{-\beta(\varepsilon_{12}+\varepsilon_{23}+\varepsilon_{31})}\oplus e^{-\beta(\varepsilon_{21}+\varepsilon_{13}+\varepsilon_{32})}.$$

Remark 4. *We emphasize that comparing all these terms allows to recover recover 5.6: comparing the terms in* 2ρ *and the terms in* ρ*, leads to compare* $-\rho$ *with* $-(\varepsilon_{13}+\varepsilon_{31})\oplus(-\varepsilon_{23}-\varepsilon_{32})$*, and comparing the term with* 2ρ*, with the terms without* ρ*, leads to compare* -2ρ *with* $-(\varepsilon_{12}+\varepsilon_{21})\oplus(-\varepsilon_{12}-\varepsilon_{23}-\varepsilon_{31})\oplus(-\varepsilon_{21}-\varepsilon_{13}-\varepsilon_{32})$*.* ■

Now, considering the dominating terms at exponential scale in (5.13) yields an equality of the form

$$\widetilde{a}g^2(\beta)-\widetilde{b}g(\beta)-\widetilde{c}=\text{ term exponentially small,} \tag{5.14}$$

where $\widetilde{a}$ is either 0 or 1, $\widetilde{b}\in\{0,1,2\}$ and $\widetilde{c}\in\{0,1,2,3\}$ and not all the coefficients $\widetilde{a}$, $\widetilde{b}$ and $\widetilde{c}$ are zero[1]. In the left-hand side of (5.14) we get terms of order $-2\rho\beta$ (at exponential scale) and in the right-hand side we have terms with higher order. Therefore at least two of $\widetilde{a},\widetilde{b}$ or $\widetilde{c}$ must be non-zero. Moreover, considering any accumulation point G for $g(\beta)$, as β goes to $+\infty$ and remembering that g is positive show that $\widetilde{G}+\widetilde{c}=0$ is impossible, *i.e.,* $\widetilde{a}=1$ necessarily holds. Hence we have

$$G^2-\widetilde{b}G-\widetilde{c}=0. \tag{5.15}$$

Such equation admits all its roots in $\mathbb{R}$, and, at least one of them is non-negative. But, the key point here is that the roots form a finite set, and this set contains the set of accumulation points for $g(\beta)$ as $\beta\to+\infty$. On the other hand, g is a continuous function, thus the set of accumulation points for g is an interval. This shows that it is reduced to a single point, and then $g(\beta)$ converges as $\beta\to+\infty$. □

We remind that $\mathcal{P}(\beta)$ goes to 0 as β goes to $+\infty$, and then $e^{\mathcal{P}(\beta)}-1$ behaves like $g(\beta)e^{-\beta.\rho}$. We replace this in Equation (5.10). The final expression is thus:

$$\begin{aligned}\frac{\mu_\beta([1])}{\mu_\beta([2])} &= \frac{\left(g(\beta)e^{\beta.(\varepsilon_{12}+\varepsilon_{21}-\varepsilon_{13}-\rho)}+e^{\beta.(\varepsilon_{21}-\varepsilon_{23})}\right)}{\left(g(\beta)e^{\beta.(\varepsilon_{12}+\varepsilon_{21}-\varepsilon_{23}-\rho)}+e^{\beta.(\varepsilon_{12}-\varepsilon_{13})}\right)}\\ &\qquad\times\frac{\left(g(\beta)e^{\beta.(\varepsilon_{12}+\varepsilon_{21}-\varepsilon_{31}-\rho)}+e^{\beta.(\varepsilon_{12}-\varepsilon_{32})}\right)}{\left(g(\beta)e^{\beta.(\varepsilon_{12}+\varepsilon_{21}-\varepsilon_{32}-\rho)}+e^{\beta.(\varepsilon_{21}-\varepsilon_{31})}\right)}.\end{aligned} \tag{5.16}$$

We know by Proposition 46 that $g(\beta)$ converges to a nonnegative limit, and then $\dfrac{\mu_\beta([1])}{\mu_\beta([2])}$ behaves like $(g(\beta)+1)^r(g(\beta))^s e^{\beta.t}$ for r and s in $\{-2,-1,0,1,2\}$ and $t\in\mathbb{R}$. We also remind that $g(\beta)$ is subexponential. Therefore $\dfrac{\mu_\beta([1])}{\mu_\beta([2])}$ admits a limit in $\mathbb{R}_+\cup\{+\infty\}$ as $\beta\to+\infty$. If the limit is 0, μ_β goes to δ_{2^∞}, if the limit is $+\infty$, μ_β goes to δ_{1^∞}, if it is equal to $\alpha\in]0,+\infty[$, μ_β goes to $\dfrac{\alpha}{\alpha+1}\delta_1+\dfrac{1}{\alpha+1}\delta_{2^\infty}$.

[1] because we exactly consider the dominating exponential scale.

Bibliography

[1] S. Addas-Zanata and F. Tal, Support of maximizing measures for typical C^0 dynamics on compact manifolds. Discrete Contin. Dyn. Syst. 26 (2010), no. 3

[2] S. Addas-Zanata and F. Tal, Maximizing measures for endomorphisms of the circle. Nonlinearity 21 (2008), no. 10, 2347-2359.

[3] S. Addas-Zanata and F. Tal, On maximizing measures of homeomorphisms on compact manifolds. Fund. Math. 200 (2008), no. 2, 145-159

[4] V. Anagnostopoulou, and O. Jenkinson, Which beta-shifts have a largest invariant measure? J. Lond. Math. Soc. (2) 79 (2009), no. 2, 445-464

[5] V. Anagnostopoulou, K. Daz-Ordazand, O. Jenkinson and C. Richard, Entrance time functions for flat spot maps. Nonlinearity 23 (2010), no. 6, 147-1494

[6] V. Anagnostopoulou, K. Daz-Ordazand, O. Jenkinson and C. Richard, Anagnostopoulou, V.; Daz-Ordaz, K.; Jenkinson, O.; Richard, C. The flat spot standard family: variation of the entrance time median. Dyn. Syst. 27 (2012), no. 1, 29-43.

[7] V. Anagnostopoulou, K. Daz-Ordazand, O. Jenkinson and C. Richard, Sturmian maximizing measures for the piecewise-linear cosine family. Bull. Braz. Math. Soc. (N.S.) 43 (2012), no. 2, 285-302.

[8] V. Baladi, Positive Transfer Operators and decay of correlations, World Scientific

[9] A. T. Baraviera, A. O. Lopes and P. Thieullen, A large deviation principle for equilibrium states of Hölder potencials: the zero temperature case, *Stochastics and Dynamics* **6** (2006), 77-96.

[10] A. T. Baraviera, L. Cioletti, A. O. Lopes, J. Mohr and R. R. Souza, On the general one-dimensional XY model: positive and zero temperature, selection and non-selection. *Rev.Math. Phys.* 23 (2011),no. 10, 1063-1113, 82Bxx.

[11] R. Bamn, J. Kiwi, J. Rivera-Letelier and R. Urza, On the topology of solenoidal attractors of the cylinder. Ann. Inst. H. Poincar Anal. Non Linaire 23 (2006), no. 2, 209-236

[12] A. Baraviera, R. Leplaideur and A. O. Lopes, Selection of measures for a potential with two maxima at the zero temperature limit, *SIAM Journal on Applied Dynamical Systems*, Vol. 11, n 1, 243-260 (2012)

[13] A. T. Baraviera, A. O. Lopes and J. K. Mengue, On the selection of subaction and measure for a subclass of potentials defined by P. Walters. *Ergodic Theory and Dynamical Systems*, Available on CJO2012.

[14] R. Bartle, The elements of integration, John Wiley (1966)

[15] R. Bissacot and E. Garibaldi, Weak KAM methods and ergodic optimal problems for countable Markov shifts, *Bull. Braz. Math. Soc.* 41, N 3, 321-338, 210.

[16] R. Bissacot and R. Freire Jr., On the existence of maximizing measures for irreducible countable Markov shifts: a dynamical proof, to appear in Erg. Theo. and Dyn. Systems

[17] J. Bochi and A. Navas, A geometric path from zero Lyapunov exponents to rotation cocycles, Arxiv (2011)

[18] T. Bousch Le poisson n'a pas d'artes, Ann. Inst. H. Poincar Probab. Statist. 36 (2000), no. 4, 489-508.

[19] T. Bousch, Le poisson n'a pas d'arêtes, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, **36** (2000), 489-508.

[20] T. Bousch, Un lemme de Ma bilatral, C. R. Math. Acad. Sci. Paris 335 (2002), no. 6, 533-536

[21] T. Bousch and O. Jenkinson, Cohomology classes of dynamically non-negative C^k functions. Invent. Math. 148 (2002), no. 1, 207-217.

[22] T. Bousch and J. Mairesse Asymptotic height optimization for topical IFS, Tetris heaps, and the finiteness conjecture. J. Amer. Math. Soc. 15 (2002), no. 1, 77-111

[23] X. Bressaud e A. Quas, Rate of approximation of minimizing measures, *Nonlinearity* **20** no. 4, (2007), 845-853.

[24] T. Bousch, La condition de Walters. *Ann. Sci. ENS*, 34, (2001)

[25] T. Bousch, Le lemme de Ma-Conze-Guivarc'h pour les systmes amphidynamiques rectifiables, Ann. Fac. Sci. Toulouse Math. (6) 20 (2011), no. 1, 1-14

[26] R. Bowen, Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms, Lecture Notes in Math., vol. 470. Springer, revised edition, 2008. With a preface by David Ruelle, Edited by Jean-René Chazottes.

[27] F. Branco, Sub-actions and maximizing measures for one-dimensional transformations with a critical point. Discrete Contin. Dyn. Syst. 17 (2007), no. 2, 271-280.

[28] S. Branton, Sub-actions for Young towers. Discrete Contin. Dyn. Syst. 22 (2008), no. 3, 541-556.

[29] J. Brémont, Gibbs measures at temperature zero. *Nonlinearity*, 16(2): 419–426, 2003.

[30] J. Brémont, Finite flowers and maximizing measures for generic Lipschitz functions on the circle. Nonlinearity 19 (2006), no. 4, 813-828.

[31] J. Brémont, Entropy and maximizing measures of generic continuous functions. C. R. Math. Acad. Sci. Paris 346 (2008), no. 3-4, 199-201.

[32] J. Brémont and S. Seuret, The singularity spectrum of the fish's boundary. Ergodic Theory Dynam. Systems 28 (2008), no. 1, 49-66.

[33] J. Brmont, Dynamics of injective quasi-contractions. Ergodic Theory Dynam. Systems 26 (2006), no. 1, 1944.

[34] X. Bressaud and A. Quas, Rate of approximation of minimizing measures. Nonlinearity 20 (2007), no. 4, 845-853.

[35] D. Collier and I. D. Morris, Approximating the maximum ergodic average via periodic orbits. Ergodic Theory Dynam. Systems 28 (2008), no. 4, 1081-1090

[36] W. Chou and R. J. Duffin, An additive eigenvalue problem of physics related to linear programming, *Advances in Applied Mathematics* **8**, 486-498, 1987.

[37] J.R. Chazottes, J.M. Gambaudo and E. Ulgade, Zero-temperature limit of one dimensional Gibbs states via renormalization: the case of locally constant potentials, *Erg. Theo. and Dyn. Sys.* 31 (2011), no. 4, 1109-1161

[38] J. R. Chazottes and M. Hochman, On the zero-temperature limit of Gibbs states, *Commun. Math. Phys.* vol 297, N1, 2010.

[39] Contreras, G. 2013, arXiv:1307.0559

[40] G. Contreras and R. Iturriaga. *Global minimizers of autonomous Lagrangians*, 22° Colóquio Brasileiro de Matemática, IMPA, 1999.

[41] G. Contreras, A. O. Lopes and Ph. Thieullen. Lyapunov minimizing measures for expanding maps of the circle, Ergodic Theory and Dynamical Systems, Vol 21, 1379-1409, 2001.

[42] G. Contreras, A. Lopes and E. Oliveira, Ergodic Transport Theory, periodic maximizing probabilities and the twist condition, preprint Arxiv.

[43] J. P. Conze and Y. Guivarc'h, Croissance des sommes ergodiques et principe variationnel, manuscript circa (1993).

[44] A. Davie, M. Urbanski annd A. Zdunik, Maximizing measures of metrizable non-compact spaces. Proc. Edinb. Math. Soc. (2) 50 (2007), no. 1, 123-151

[45] A. Dembo and O. Zeitouni, Large Deviations Techniques and Applications, Springer Verlag, 1998.

[46] R. Ellis, Entropy, Large Deviations, and Statistical Mechanics, Springer Verlag, 2005

[47] R. Exel and A. Lopes, C^*-Algebras, approximately proper equivalence relations and Thermodynamic Formalism, Erg Theo and Dyn Sys, 1051-1082, Vol 24, t (2004).

[48] A. Fathi, Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens, *Comptes Rendus de l'Académie des Sciences, Série I, Mathématique* Vol 324 1043-1046, 1997.

[49] P. Fernandez, Medida e Integração, (1976) IMPA

[50] E. Garibaldi and A. O. Lopes, On Aubry-Mather theory for symbolic Dynamics, Ergodic Theory and Dynamical Systems, Vol 28 , Issue 3, 791-815 (2008)

[51] E. Garibaldi and A. Lopes, Functions for relative maximization, Dynamical Systems, v. 22, 511-528, 2007

[52] E. Garibaldi and A. O. Lopes, The effective potential and transshipment in thermodynamic formalism at temperature zero, to appear in Stoch. and Dyn. (2013).

[53] E. Garibaldi, A. Lopes and P. Thieullen, On calibrated and separating sub-actions, Bull. of the Bras. Math. Soc. Vol 40, 577-602, (4)(2009)

[54] E. Garibaldi and Ph. Thieullen, Minimizing orbits in the discrete Aubry-Mather model, Nonlinearity, 24 (2011), no. 2, 563-611

[55] E. Garibaldi and Ph. Thieullen, Description of some ground states by Puiseux technics, *Journ. of Statis. Phys*, 146, no. 1, 125-180, (2012)

[56] D. A. Gomes, Viscosity solution method and the discrete Aubry-Mather problem, *Discrete and Continuous Dynamical Systems, Series A* **13** (2005), 103-116.

[57] D. A. Gomes, A. O. Lopes and J. Mohr, The Mather measure and a large deviation principle for the entropy penalized method. Commun. Contemp. Math. 13 (2011), no.2, 235-268.

[58] A. Grant, Finding Optimal Orbits of Chaotic Systems, Phd thesis, Univ. of Maryland (2005)

[59] B. R. Hunt Maximum local Lyapunov dimension bounds the box dimension of chaotic attractors, Nonlinearity 9 (1996), no. 4, 845-852.

[60] B. R. Hunt and Y. Guocheng, Optimal orbits of hyperbolic systems. *Nonlinearity* **12**, (1999), 1207-1224.

[61] B. Hunt B and E. Ott, Optimal periodic orbits of chaotic systems occur at low period, Phys. Rev. E 54 32837 (1996)

[62] E. Harriss and O. Jenkinson, Flattening functions on flowers. Ergodic Theory Dynam. Systems 27 (2007), no. 6, 1865-1886

[63] Godofredo Iommi, Ergodic Optimization for Renewal Type Shifts, Monatshefte fr Mathematik, Volume 150, Number 2 (2007), 91-95,

[64] O. Jenkinson. Ergodic optimization, Discrete and Continuous Dynamical Systems, Series A, V. 15, 197-224, 2006.

[65] O. Jenkinson and J. Steel, Majorization of invariant measures for orientation-reversing maps. Ergodic Theory Dynam. Systems 30 (2010), no. 5, 1471-1483.

[66] O. Jenkinson and I. D. Morris, Lyapunov optimizing measures for C 1 expanding maps of the circle. Ergodic Theory Dynam. Systems 28 (2008), no. 6, 1849-1860

[67] O. Jenkinson, A partial order on $\times 2$ -invariant measures. Math. Res. Lett. 15 (2008), no. 5, 893-900.

[68] O. Jenkinson, Optimization and majorization of invariant measures. Electron. Res. Announc. Amer. Math. Soc. 13 (2007), 112

[69] O. Jenkinson, Every ergodic measure is uniquely maximizing. Discrete Contin. Dyn. Syst. 16 (2006), no. 2, 383-392.

[70] O. Jenkinson, Maximum hitting frequency and fastest mean return time. Nonlinearity 18 (2005), no. 5, 2305-2321

[71] O. Jenkinson, Directional entropy of rotation sets. C. R. Acad. Sci. Paris Sr. I Math. 332 (2001), no. 10, 921-926.

[72] O. Jenkinson, Geometric barycentres of invariant measures for circle maps. Ergodic Theory Dynam. Systems 21 (2001), no. 2, 511532.

[73] O. Jenkinson, R. D. Mauldin and M. Urbanski, Zero Temperature Limits of Gibbs-Equilibrium States for Countable Alphabet Subshifts of Finite Type, Journ. of Statis. Physics, Volume 119, Numbers 3-4 (2005)

[74] T. Jordan and M. Rams, Multifractal analysis of weak Gibbs measures for non-uniformly expanding C 1 maps. Ergodic Theory Dynam. Systems 31 (2011), no. 1, 143164.

[75] S. Karlin and H. Taylor, A First Course in Stochastic Processes, Academic Press

[76] G. Keller, Gibbs States in Ergodic Theory, Cambrige Press, 1998.

[77] T. Kempton, Zero Temperature Limits of Gibbs Equilibrium States for Countable Markov Shifts, *J. Stat Phys* 143, 795-806, 2011.

[78] O. Lanford, Entropy and Equilibrium States in Classical Statistical Mechanics. Statistical mechanics and mathematical problems. Battelle Rencontres, Seattle, Wash., 1971. Lecture Notes in Physics, 20. Springer-Verlag, Berlin-New York, 1973.

[79] R. Leplaideur, A dynamical proof for convergence of Gibbs measures at temperature zero, Nonlinearity, 18, N 6, 2847-2880, 2005.

[80] R. Leplaideur, Totally dissipative measures for the shift and conformal s -finite measures for the stable holonomies. Bull. Braz. Math. Soc. 41 (2010), no. 1, 136.

[81] R. Leplaideur, Flatness is a criterion for selection of maximizing measures. J. Stat. Phys. 147 (2012), no. 4, 728-757.

[82] R. Leplaideur, Thermodynamic formalism for a family of non-uniformly hyperbolic horseshoes and the unstable Jacobian. Ergodic Theory Dynam. Systems 31 no. 2, 423-447, (2011).

[83] D. Lind and B. Marcus. *An introduction to Symbolic Dynamics and Coding.* Cambridge University Press, 1995.

[84] A. O. Lopes, Entropy and Large Deviation, *NonLinearity*, Vol. 3, N. 2, 527-546, 1990.

[85] A. O. Lopes, Entropy, Pressure and Large Deviations, Cellular Automata, Dynamical Systems and Neural Networks, E. Goles e S. Martinez (eds.), Kluwer, Massachusets, pp. 79-146, 1994.

[86] A. O. Lopes, J. Mohr, R. Souza and Ph. Thieullen, Negative entropy, zero temperature and stationary Markov chains on the interval, *Bulletin of the Brazilian Mathematical Society* **40**, 1-52, 2009.

[87] A. O. Lopes and E. R. Oliveira, On the thin boundary of the fat attractor, preprint (2012)

[88] A. O. Lopes. Thermodynamic Formalism, Maximizing Probabilities and Large Deviations. Work in progress. Lecture Notes - Dynamique en Cornouaille (2012)

[89] A. O. Lopes and J. Mengue, Zeta measures and Thermodynamic Formalism for temperature zero, *Bulletin of the Brazilian Mathematical Society* 41 (3) pp 449-480 (2010)

[90] A. Lopes and J. Mengue, Duality Theorems in Ergodic Transport, Journal of Statistical Physics. Vol 149, issue 5, pp 921-942 (2012)

[91] A. O. Lopes, E. R. Oliveira and Ph. Thieullen, The dual potential, the involution kernel and transport in ergodic optimization, *preprint* Arxiv.

[92] A. Lopes and P. Thieullen, Eigenfunctions of the Laplacian and associated Ruelle operator, Nonlinearity, Volume 21, Number 10, pp-2239-2254 (2008)

[93] A. Lopes, V. Rosas and R. Ruggiero Cohomology and Subcohomology for expansive geodesic flows, Discrete and Continous Dynamical Systems" Volume: 17, N. 2, pp 403-422 (2007).

[94] A. Lopes and P. Thieullen, Subactions for Anosov Flows, Erg Theo and Dyn Syst vol 25 Issue 2, pp605-628 (2005)

[95] A. Lopes and P. Thieullen, Mather measures and the Bowen-Series transformation, Annal Inst Henry Poincare - Anal non-linear, 663-682, vol 5, (2006)

[96] A.O. Lopes and E. Oliveira, Entropy and variational principles for holonomic probabilities of IFS, *Disc. and Cont. Dyn. Systems* series A, vol 23, N 3, 937-955, 2009.

[97] A. O. Lopes, E. R. Oliveira and Ph. Thieullen, The dual potential, the involution kernel and transport in ergodic optimization, *preprint*, 2008.

[98] A.O. Lopes, An analogy of charge distribution on Julia sets with the Brownian motion , *J. Math. Phys.* 30 9, 2120-2124, 1989.

[99] A. Lopes and J. Mengue, Selection of measure and a Large Deviation Principle for the general one-dimensional XY model, preprint, UFRGS, 2011.

[100] A.Lopes, J. K. Mengue, J. Mohr and R. R. Souza Entropy and Variational Principle for one-dimensional Lattice Systems with a general a-priori measure: positive and zero temperature, preprint (2012)

[101] A. Lopes, E. R. Oliveira and D. Smania, Ergodic Transport Theory and Piecewise Analytic Subactions for Analytic Dynamics, Bull. of the Braz. Math Soc. Vol 43 (3) 467-512 (2012)

[102] A. Lopes and S. Lopes, Uma introdução aos Processos Estocásticos para estudantes de Matemática, a ser publicado

[103] E. L. Lima, Curso de Anlise , Vol II, IMPA (1981)

[104] R. Mañé. Generic properties and problems of minimizing measures of Lagrangian systems, *Nonlinearity*, Vol 9, 273-310, 1996.

[105] H. McCluskey and A. Manning. Hausdorf dimension for horseshoes. *Ergod th. & Dynam. Sys.*, (3):251–260, 1983.

[106] J. Mather, Action minimizing invariant measures for positive definite Lagrangian Systems, *Math. Z.*, 207 (2), pp 169-207, 1991

[107] I. D. Morris, Entropy for Zero-Temperature Limits of Gibbs-Equilibrium States for Countable-Alphabet Subshifts of Finite Type, Journ. of Statis. Physics, Volume 126, Number 2 (2007), 315-324,

[108] I. D. Morris, A rapidly-converging lower bound for the joint spectral radius via multiplicative ergodic theory. Adv. Math. 225 (2010), no. 6, 3425-3445.

[109] I. D. Morris, Ergodic optimization for generic continuous functions. Discrete Contin. Dyn. Syst. 27 (2010), no. 1, 383-388

[110] I. D. Morris, The Ma-Conze-Guivarc'h lemma for intermittent maps of the circle. Ergodic Theory Dynam. Systems 29 (2009), no. 5, 16031611

[111] I. D. Morris, Maximizing measures of generic Hlder functions have zero entropy. Nonlinearity 21 (2008), no. 5, 993-1000

[112] I. D. Morris, A sufficient condition for the subordination principle in ergodic optimization. Bull. Lond. Math. Soc. 39 (2007), no. 2, 214-220

[113] I. D. Morris, Entropy for zero-temperature limits of Gibbs-equilibrium states for countable-alphabet subshifts of finite type. J. Stat. Phys. 126 (2007), no. 2, 315-324

[114] W. Parry and M. Pollicott. *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Astérisque Vol 187-188 1990

[115] V. Pit, Invariant relations for the Bowen-Series transform, Conform. Geom. Dyn. 16 (2012), 103-123.

[116] M. Pollicott and M. Yuri, Dynamical systems and Ergodic Theory, Cambrige Press, 1998

[117] M. Pollicott and R. Sharp, Livsic theorems, maximizing measures and the stable norm. Dyn. Syst. 19 (2004), no. 1, 75-88.

[118] A. Quas and J. Siefken, Ergodic optimization of supercontinuous functions on shift spaces, Ergodic Theory and Dynamical Systems / Volume 32 / Issue 06 2012, 2071-2082

[119] Y. Sinai, Probabilty Theory, an introductory course,Springer Veralag

[120] O. Sarig, Thermodynamic formalism for countable Markov shifts, *Ergodic Theory and Dynamical Systems* 19, 1565-1593, 1999

[121] O. Sarig, Lecture notes on thermodynamic formalism for topological Markov shifts, *Penn State*, 2009.

[122] E. A. da Silva, R. R. da Silva and R. R. Souza The Analyticity of a Generalized Ruelle's Operator, to appear Bull. Braz. Math. Soc.

[123] R. R. Souza, Sub-actions for weakly hyperbolic one-dimensional systems, *Dynamical Systems* 18 **(2)**, 165-179, 2003.

[124] F. Spitzer. A Variational characterization of finite Markov chains. *The Annals of Mathematical Statistics.* (43): N.1 303-307, 1972.

[125] J. Steel, Concave unimodal maps have no majorisation relations between their ergodic measures. Proc. Amer. Math. Soc. 139 (2011), no. 7, 2553-2558

[126] P. Walters, Ergodic Theory, Springer Verlag