

# Ergodic Optimization and Ground States: a brief Introduction

Artur O. Lopes

Dedicated to the memory of Gonzalo Contreras

May 14, 2026

## Abstract

Our goal in this short note is to briefly and succinctly describe some basic concepts and properties of Ergodic Optimization for readers unfamiliar with the subject. We avoid technical issues in order to provide a global overview of this topic. We will not attempt to cover all of the many contributions of various authors, who have greatly enriched the theory with invaluable results. The author has made a personal selection of the topics to be addressed, keeping in mind two main objectives: to motivate the reasons for studying the subject, and to describe schematically and pictorially its relationship with relevant concepts and properties of Statistical Mechanics, which is one of the sources of inspiration for the theory. We will not present new results or detailed proofs. Some examples will be provided. We describe some procedures that may help in obtaining explicit solutions. We present some references that by no means aim to exhaust the bibliography on the subject, where possible, minimizing the number of references.

## 1 Introduction

Consider the set  $\Omega = \{1, 2, \dots, d\}^{\mathbb{N}}$  and the shift  $\sigma$  acting on  $\Omega$ :

$$\sigma(x) = \sigma(x_1, x_2, \dots, x_n, \dots) = (x_2, x_3, \dots, x_n, \dots).$$

The metric on  $\Omega$  is:

$$d(x, y) = \left(\frac{1}{2}\right)^{\min\{n: x_n \neq y_n\}}, \text{ where } x = (x_1, x_2, \dots) \text{ and } y = (y_1, y_2, \dots).$$

A Borel probability  $\mu$  is  $\sigma$ -invariant if for any continuous function  $f : \Omega \rightarrow \mathbb{R}$  we have

$$\int (f \circ \sigma) d\mu = \int f d\mu. \tag{1}$$

Equivalently,  $\mu$  is  $\sigma$ -invariant if for any Borel set  $E \subset \Omega$ , we have that

$$\mu(\sigma^{-1}(E)) = \mu(E). \quad (2)$$

Denote by  $\mathcal{M}_\sigma$  the set of  $\sigma$ -invariant probabilities over the Borel sigma-algebra.

When considering the iteration  $\sigma^n(x)$ , the  $n$  does not represent time, but rather the translation in the lattice  $\mathbb{N}$ . One can think of  $\{1, 2, \dots, d\}$  as the set of possible spins located at each site  $k$  of the lattice  $\mathbb{N}$ .

From the point of view of Statistical Mechanics,  $\mu \in \mathcal{M}_\sigma$  means that  $\mu$  is in thermodynamic equilibrium.

A cylinder set of size  $r$  is a set of the form  $\overline{a_1, a_2, \dots, a_r} \subset \Omega$ , given by

$$\overline{a_1, a_2, \dots, a_r} = \{(a_1, a_2, \dots, a_r, x_{r+1}, x_{r+2}, \dots) \mid \text{where } x_j \in \{1, 2, \dots, d\}, j \geq (r+1)\}.$$

The set of all cylinders of all sizes  $r \in \mathbb{N}$  generates the Borel sigma-algebra on  $\Omega$ .

**Definition 1.** *The Kolmogorov entropy of a  $\sigma$ -invariant probability  $\mu$  is given by*

$$h(\mu) = - \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{(a_1, a_2, \dots, a_m) \in \{1, 2, \dots, d\}^m} \mu(\overline{a_1, a_2, \dots, a_m}) \log(\mu(\overline{a_1, a_2, \dots, a_m})). \quad (3)$$

For general properties of entropy see [44] or [43].

In Statistical Mechanics, a function  $H : \Omega \rightarrow \mathbb{R}$  describes the interaction among the sequences  $(x_1, x_2, \dots, x_n, \dots)$  and is usually called a Hamiltonian.

The relation between problems in the symbolic space  $\{1, 2, \dots, d\}^{\mathbb{Z}}$  and those in  $\{1, 2, \dots, d\}^{\mathbb{N}}$  can be properly described via Proposition 1.2 in [39], or via Appendix section 9 in [36]. Ergodic results for probabilities acting on the space  $\{1, 2, \dots, d\}^{\mathbb{Z}}$  can be deduced from the use of the Ruelle operator (a powerful tool) acting on functions defined on the space  $\Omega = \{1, 2, \dots, d\}^{\mathbb{N}}$ . This is not an issue here.

## 2 Thermodynamic Formalism

General references for Thermodynamic Formalism are [39], [33], [36], [20] and [44].

Given a Hölder (sometimes merely continuous) function

$$A : \Omega = \{1, 2, \dots, d\}^{\mathbb{N}} \rightarrow \mathbb{R},$$

in this section we are interested in the study of the set of probabilities  $\mu$  maximizing the Topological Pressure:

$$P(A) := \sup_{\rho \in \mathcal{M}_\sigma} \left( \int A d\rho + h(\rho) \right), \quad (4)$$

here  $h(\rho)$  is the Kolmogorov entropy of  $\rho$ .

One is interested in probabilities  $\mu_A$  which maximize  $P(A)$ , that is, those satisfying

$$P(A) = h(\mu_A) + \int A d\mu_A.$$

Any such probability  $\mu_A$  is called an equilibrium state (or measure) for  $A$ .

The study of (4) for the case of functions  $A$  that are merely continuous (see [44]) presents some conceptual differences from the setting where  $A$  is of Hölder class (see [39]). Here we are interested in the latter case.

In the above setting, a function  $A$  is usually called a potential. In accordance with Statistical Mechanics,  $A$  should represent  $-H = A$ , where  $H$  is the Hamiltonian. The minus sign is related to the fact that for a probability  $\mu$  in equilibrium (under the action of the Hamiltonian  $H$ ), states of high energy are less probable.

By equilibrium under the action of the Hamiltonian  $H$  we mean the one associated with

$$P(-H) := \sup_{\rho \in \mathcal{M}_\sigma} \left( - \int H d\rho + h(\rho) \right). \quad (5)$$

Given  $H$ , if there exists more than one equilibrium state for (6), we say that the Hamiltonian  $H$  exhibits a phase transition.

In Statistical Mechanics, if  $H : \Omega \rightarrow \mathbb{R}$  is the Hamiltonian, then  $\int H d\mu$  is the mean energy of the state  $\mu \in \mathcal{M}(\sigma)$ .

**Second Law of Thermodynamics:** There is a tendency in nature for equilibrium to be reached in states of higher entropy. This fact can be formalized as follows: for a system under the action of Hamiltonian  $H$ , given a real value  $E$  (representing energy), consider the set

$$\mathfrak{E}_E = \{ \mu \in \mathcal{M}_\sigma \mid \int H d\mu = E \}.$$

Under the above conditions, the probability at equilibrium is the one maximizing  $h(\mu)$  among elements  $\mu \in \mathfrak{E}_E$ .

If the function  $A$  is of Hölder class, the  $\sigma$ -invariant probability  $\mu_A$  is unique; it has positive entropy, is ergodic, mixing, and positive on cylinder

sets (see [39]). The Ruelle Theorem is a helpful tool for obtaining its ergodic properties (see [39], [36] and [1]). Explicit examples of equilibrium probabilities for some nontrivial potentials are presented in [12].

When  $A = 0$ , we obtain the equilibrium probability  $\mu_0$ , which is the maximal entropy measure: the independent probability with weights  $\frac{1}{d}$ , and  $P(0) = \log d$ .

A large value of the entropy of a  $\sigma$ -invariant probability  $\mu$  corresponds to a large randomness of the samples; more uncertainty. Our dynamical entropy  $h(\mu)$ , more than concave, is affine (see [44]).

In agreement with Statistical Mechanics — in the one-dimensional lattice  $\mathbb{N}$  — it is natural to introduce a parameter  $\beta > 0$ , where  $\beta = \frac{1}{T}$ , and  $T > 0$  represents temperature; and to consider  $P(\beta A)$ .

Via the Legendre transform one can relate the maximizing solutions of the Second Law problem (described above) with the maximization of pressure as in (6); the so-called MaxEnt method (see Appendix 9.2 in [36] and [30]).

In this case, in consonance with the Second Law of Thermodynamics, if  $H$  is a Hamiltonian  $H : \Omega \rightarrow \mathbb{R}$  (which can represent energy), the probability  $\mu_{-\beta H} \in \mathcal{M}_\sigma$  is the equilibrium probability for the system under the influence of  $H$ , at temperature  $T = \frac{1}{\beta}$ . Then, in our notation,  $-\beta H = \beta A$ .

More precisely, in accordance with Statistical Mechanics, we will be interested in equilibrium probabilities  $\mu_{-\frac{1}{T} H}$  for

$$P\left(-\frac{1}{T} H\right) := \sup_{\rho \in \mathcal{M}_\sigma} \left( -\frac{1}{T} \int H d\rho + h(\rho) \right). \quad (6)$$

Consider the spin lattice  $\{+, -\}^{\mathbb{N}}$ . The spin up is associated with  $+$  and spin down with  $-$ .

A measurement could result in  $(+, -, -, +, +, -, -, \dots)$ , which is a random sample. Consider a physical system at positive temperature  $T$  subject to a Hamiltonian  $H : \{+, -\}^{\mathbb{N}} \rightarrow \mathbb{R}$ . In this case one is interested, for instance, in the probability  $\mu_{-\frac{1}{T} H}$  of the cylinder set  $\overline{+, -, ++} \subset \{+, -\}^{\mathbb{N}}$ , under the equilibrium regime for  $H$  and temperature  $T > 0$ ; this probability  $\mu_{-\frac{1}{T} H}$  provides statistics of samples. If  $\mu_{-\frac{1}{T} H}(\overline{+, -, ++})$  is very small, this means that the arrangement  $+, -, ++$  is quite unlikely to be observed in the first 4 sites of the lattice  $\mathbb{N}$ .

The probability  $\delta_{+,+,+,+, \dots} = \delta_{+\infty}$  represents a type of magnetization (all spins up) on the lattice  $\mathbb{N}$ . There is no randomness in this case. The entropy of the  $\sigma$ -invariant probability  $\delta_{+\infty}$  is zero.

$T = 0$  means absolute zero (around  $-273$  degrees Celsius). In real-world experiments, the temperature of the system under consideration can be lowered to approach zero temperature.

It is known from experiments in Physics that when decreasing the temperature  $T \rightarrow 0$  of a metal, the randomness of the corresponding equilibrium state  $\mu_{-\frac{1}{T}H}$  also decreases. In most cases, when  $T \rightarrow 0$ , there is a tendency toward decreasing entropy and the metal tends to exhibit magnetic properties. How can this be described in mathematical terms?

For low temperatures, the spins of atoms at each site become aligned with each other, and this can be properly described as something like  $\delta_{+,+,+,+, \dots} = \delta_{+\infty}$  (a “*non-random*” probability), or alternatively  $\delta_{-,-,-,-, \dots} = \delta_{-\infty}$ .

Another possibility of magnetization could be  $\frac{1}{2}(\delta_{+,+,+,+, \dots} + \delta_{-,-,-,-, \dots})$ ; or, more generally, a  $\sigma$ -invariant probability supported on a periodic orbit (for the shift  $\sigma$ ).

### 3 Ergodic Optimization

General references on the topic are [4], [22] and [36].

A continuous function  $A : \Omega \rightarrow \mathbb{R}$  can be seen as a cost, or profit, and  $\int A d\mu$ , for  $\mu \in \mathcal{M}_\sigma$ , represents the (dynamical)  $\mu$ -mean value of the cost. One could consider minimization or maximization of  $\int A d\mu$  as well; simply replace  $A$  by  $-A$ . Given  $A : \Omega \rightarrow \mathbb{R}$ , maximizing  $\mu \rightarrow \int A d\mu$  is the main problem in Ergodic Optimization.

An extensive bibliography on Ergodic Optimization appears in [27]. The study of the case of potentials  $A$  that are merely continuous (see [26], [28] and [6]) presents some conceptual differences from the setting where  $A$  is of Hölder class (see [4], [22] and [9]). Here we are interested in the latter case.

Given a continuous function  $A : \Omega \rightarrow \mathbb{R}$ , consider the problem:

$$\alpha(A) := \sup_{\rho \in \mathcal{M}_\sigma} \left\{ \int A(x) d\rho(x) \right\}.$$

$\alpha(A)$  is called the maximizing value of  $A$ .

In Ergodic Optimization one is interested in probabilities  $\mu^A$  which maximize  $\alpha(A)$ , that is, those satisfying

$$\alpha(A) = \int A d\mu^A$$

Any such probability  $\mu^A$  is called a maximizing probability for  $A$ . If the potential  $A$  is of Hölder class, the maximizing probability  $\mu^A$  need not be unique (see Section 6 in [36]).

Several explicit examples illustrating the theory are presented in [17].

If  $\mu^A$  maximizes the function  $A$ , it also maximizes  $A$  plus any constant.

A natural question is: given a Hölder potential  $A$ , when is  $\mu^A$  unique? Using the Hölder norm, uniqueness holds for a dense set of Hölder potentials — adapting results from convex analysis (see [9] for a proof).

For a fixed Hölder potential  $A$ , consider  $\beta \rightarrow \infty$ , the Pressure  $P(\beta A)$  (which is convex and analytic in  $\beta$ , as proved in [39]), and the associated family of equilibrium probabilities  $\mu_{\beta A}$ ,  $\beta > 0$ . One can show that any weak limit  $\mu$  of a subsequence  $\lim_{\beta_n \rightarrow \infty} \mu_{\beta_n A} = \mu$  will be a maximizing probability for  $A$  (see [9], [4] or [36]). ; more results on the topic in Section 4.

$\beta_n \rightarrow \infty$  means  $T_n = \frac{1}{\beta_n} \rightarrow 0$ . Ergodic Optimization is the natural setting in which to analyze equilibrium probabilities at zero temperature.

Given a  $k$ -periodic orbit  $\{x, \sigma(x), \sigma^2(x), \dots, \sigma^{k-1}(x)\}$ , where  $\sigma^k(x) = x$ , we say that  $\rho = \sum_{j=0}^{k-1} \frac{1}{k} \delta_{\sigma^j(x)}$  is the  $\sigma$ -invariant probability associated with such an orbit; it has zero entropy (see [44]).

The set of  $\sigma$ -invariant probabilities associated with periodic orbits (of all periods) is weakly dense in the set of  $\sigma$ -invariant probabilities (see [41]).

Given a continuous potential  $A : \Omega \rightarrow \mathbb{R}$  and  $\epsilon > 0$ , there exists a probability  $\rho$  supported on a periodic orbit such that  $\int A d\rho > \alpha(A) - \epsilon$ .

**Fourth Law of Thermodynamics:** (Nernst law) The entropy vanishes when there is only one equilibrium state at absolute zero temperature.

In the same spirit as this postulate, we emphasize the fact that the entropy of an invariant probability supported on a unique periodic orbit is zero.

Another source of inspiration for Ergodic Optimization is the Aubry–Mather Theory (see [15] and [14]).

**Theorem 2.** *The analogue of the Mañé Conjecture in Aubry–Mather Theory for the ergodic setting: the property that the maximizing measure  $\mu^A$  is supported on a unique periodic orbit is dense in the set of Hölder potentials  $A$ .*

This quite important result was proved by Gonzalo Contreras in 2016 in [13] (partial results in [9]); it is somehow in consonance with Nernst law.

Related work appears in [6], [28], [38] and [40].

A continuous function  $u : \Omega \rightarrow \mathbb{R}$  is called a subaction for  $A$  if

$$u(\sigma(x)) \geq u(x) + A(x) - \alpha(A), \quad \text{for all } x \in \Omega. \quad (7)$$

A continuous function  $u : \Omega \rightarrow \mathbb{R}$  is called a calibrated subaction for  $A$  if

$$u(y) = \max_{\sigma(x)=y} [u(x) + A(x) - \alpha(A)], \quad \text{for all } y \in \Omega. \quad (8)$$

The above equation corresponds to the discrete-time version of the Lax–Oleinik semigroup equation used to investigate the Hamilton–Jacobi equation in Aubry–Mather theory (see [23]).

**Theorem 3.** (see [9]) *Given a Hölder function  $A : \Omega \rightarrow \mathbb{R}$ , there exists a Hölder function  $u$  which is a calibrated subaction for  $A$ .*

In order to find an explicit subaction  $u$ , it is necessary to first guess the value  $\alpha(A)$ ; this is not a straightforward question in general. We will address this issue later.

If we add a constant  $c$  to a calibrated subaction  $u$  for the potential  $A : \Omega \rightarrow \mathbb{R}$ , we obtain another one for  $A$ . Given  $A$ , we say that the calibrated subaction is unique when it is unique up to the addition of a constant. When  $\mu_A$  is unique there is only one calibrated subaction. There exist Hölder potentials  $A$  for which the calibrated subaction is not unique; in this case, one can use the Mañé potential to find different subactions (see [23] and [3]).

If  $u : \Omega \rightarrow \mathbb{R}$  is a subaction for  $A : \Omega \rightarrow \mathbb{R}$ , it follows that

$$R_u(x) = R(x) := u(\sigma(x)) - u(x) - A(x) + \alpha(A) \geq 0.$$

Given  $u$ , we will be interested in the points  $x \in \Omega$  such that  $R(x) = 0$ .

We call the contact locus of a subaction  $u : \Omega \rightarrow \mathbb{R}$  the set

$$\mathbb{M}_A(u) := \{x \in \Omega \mid A(x) - u \circ \sigma(x) + u(x) = \alpha(A)\} = \{x \in \Omega \mid R(x) = 0\}.$$

Determining the set of points  $x$  where  $R(x) = 0$  is of great importance: it helps to localize the support of maximizing measures.

**Theorem 4.** (see [9] and [4]) — *Suppose  $\mu$  is maximizing for  $A$  and  $u$  is a subaction. Then,*

$$(u \circ \sigma)(x) - u(x) = A - \alpha(A), \quad (9)$$

for all  $x$  in the support of  $\mu$ ; that is,  $R(x) = 0$ .

Moreover, if a probability  $\mu \in \mathcal{M}_\sigma$  is such that (9) holds for all points  $x$  in its support, then  $\mu$  is maximizing for  $A$ .

**Claim:** The supports of maximizing probabilities are contained in  $\mathbb{M}_A(u)$  for any subaction  $u$ .

*Proof.* It is known that  $R(x) \geq 0$  for all  $x$ . Then, for  $\mu$  maximizing  $A$ ,

$$0 \leq \int R(x)d\mu(x) = \int (u(\sigma(x)) - u(x) - A(x) + \alpha(A))d\mu = \int (-A(x) + \alpha(A))d\mu = 0.$$

Then  $R(x) = 0$  for  $\mu$ -a.e. point  $x$ .  $\square$

Note, however, that even when  $\mu^A$  is unique, given a calibrated subaction  $u$  for  $A$ , the set  $\mathbb{M}_A(u)$  may strictly contain the support of  $\mu^A$ .

If we interpret  $x \rightarrow u(\sigma(x)) - u(x) = \mathfrak{d}(u)(x)$  as the discrete time derivative of the function  $u$ , the equality

$$\mathfrak{d}(u)(x) = u(\sigma(x)) - u(x) = A(x) - \alpha(A), \quad (10)$$

for  $x$  in the support of  $\mu^A$ , can be rephrased as:  $u$  is the **primitive** of  $A - \alpha(A)$  on the support of the maximizing probability  $\mu^A$ .

Using the 1/2-algorithm described in [16] and [17], it is possible in some cases to guess the explicit expression of a subaction for a given potential  $A$ ; for instance, when  $A$  is the indicator function defined on a cylinder set in  $\{0, 1\}^{\mathbb{N}}$ . The expression for the implementation of the iterative procedure on a desktop is in [16] and the Code in Python is available in [18]; the performance of the 1/2-iteration method is analyzed in [16].

You do not need to know  $\alpha(A)$  to apply the 1/2-method. Using your desktop you derive explicit results which can be checked with pen and paper.

The plan is to consider the associated binary expansion on the interval  $[0, 1]$ , and from the pictures obtained from a certain iterative process to guess the solution: we will use the iterative 1/2-method. The idea is first to identify elements  $w \in \Omega = \{0, 1\}^{\mathbb{N}}$  with elements  $x \in [0, 1]$ . We will elaborate on this.

Suppose the symbolic space is  $\{0, 1\}^{\mathbb{N}}$ . One can visualize, via the binary expansion, elements  $x = (x_1, x_2, x_3, \dots, x_n, \dots) \in \{0, 1\}^{\mathbb{N}}$  as

$$z = x_1 2^{-1} + x_2 2^{-2} + x_3 2^{-3} + \dots + x_n 2^{-n} + \dots \in [0, 1]. \quad (11)$$

Figure 1 illustrate this point of view; it helps to see geometric pictures.

We will explain how to use the procedure of guessing the subaction  $u$  of a potential  $A$  through examples. Take  $A = I_{\overline{01}}$ : the 1/2-iteration method produces the picture described in Figure 2 (it is not necessary many iterations). From this, a natural candidate is to set  $u = -I_{\overline{0}} + 1/2 I_{\overline{1}}$ . We substitute this function  $u$  into the subaction equation (7) and verify that it is indeed satisfied. By inspection we get that the value  $\alpha(A) = 1/2$ ,  $R = 1/2 I_{\overline{00}} + 1/2 I_{\overline{11}}$  (which is zero on the set  $\{(01)^\infty, (10)^\infty\}$ ); it follows from Theorem 4 that  $\frac{1}{2}\delta_{(01)^\infty} + \frac{1}{2}\delta_{(10)^\infty}$  is a maximizing probability for such  $A$ . Thus, we obtain

a rigorous mathematical result - it can be verified by hand - using the hint produced by the above-mentioned method obtained via computer. A similar procedure for finding  $u$  for the potential  $A = I_{\overline{011111}}$  is described in Figure 3 (see also Example 37 in the Section 6.2 in [36]). We use the software Mathematica to generate the pictures.

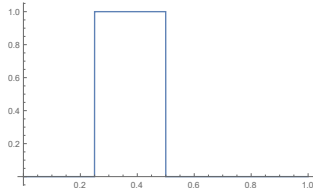


Figure 1: The graph of  $x \rightarrow A(x)$  “defined” for  $x \in [0, 1]$  corresponding to the potential  $A = I_{\overline{01}}$ .

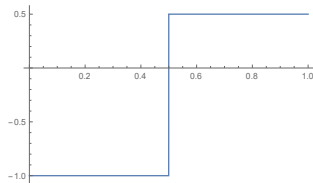


Figure 2: The graph of the subaction  $x \rightarrow u(x)$  “defined” for  $x \in [0, 1]$  obtained from the 1/2-iteration method when  $A = I_{\overline{01}}$ . From the computer-generated picture we guess that the subaction of  $A$  is  $u = -I_{\overline{0}} + 1/2 I_{\overline{1}}$ . One can verify rigorously (by hand) that such  $u$  is indeed a solution to equation (7). In this case  $\frac{1}{2}\delta_{(01)^\infty} + \frac{1}{2}\delta_{(10)^\infty}$  is the maximizing probability and  $\alpha(A) = 1/2$ . Not many iterations were required to obtain this picture.

A point  $x \in \Omega$  is said to be *non-wandering with respect to*  $A : \Omega \rightarrow \mathbb{R}$ , if for every  $\epsilon > 0$ , there exists an integer  $k \geq 1$  and a point  $y \in \Omega$  such that

$$d(x, y) < \epsilon, \quad d(x, \sigma^k(y)) < \epsilon \quad \text{and} \quad \left| \sum_{j=0}^{k-1} (A - \alpha(A)) \circ \sigma^j(y) \right| < \epsilon.$$

For practical purposes we can assume  $\alpha(A) = 0$ . In this case the above means that pseudo-orbits beginning at  $x$  can have an arbitrarily small cumulative cost along the sequence.

Up to resolution  $\epsilon$ ,  $\{y, \sigma(y), \dots, \sigma^{k-1}(y)\}$  looks like a periodic orbit.

We denote by  $\Omega(A)$  the *set of non-wandering points* with respect to the observable  $A \in C^0(\Omega)$ . When the observable is Hölder,  $\Omega(A)$  contains the support of all maximizing probability measures for  $A$  (see [9], [23] or [22]).

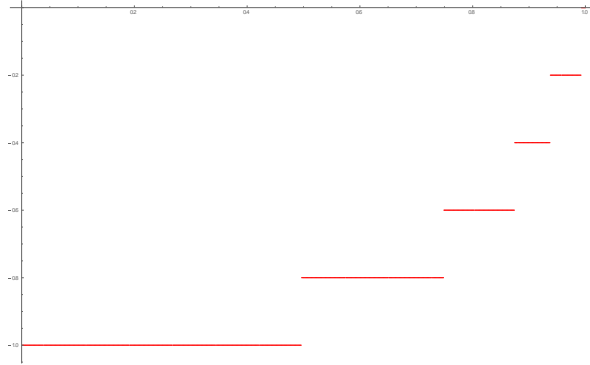


Figure 3: The graph of  $x \rightarrow u(x)$  on  $[0, 1]$  obtained from the 1/2-iteration method when  $A = I_{\overline{01111}}$ . From the above picture one can guess the expression for a subaction  $u$  in the case of this potential  $A$ . One can verify that such  $u$  is a solution to the subaction equation (7), where  $\alpha(A) = 0.2$ .

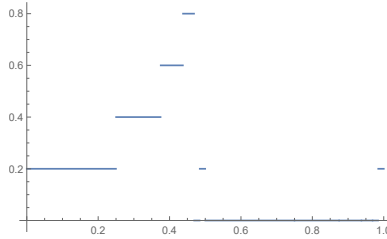


Figure 4: The graph on  $[0, 1]$  corresponding to the function  $x \rightarrow R(x)$  for the potential  $I_{\overline{01111}}$  (using the subaction  $u$  obtained from Figure 3). It is known that the function  $R$  vanishes on the support of the maximizing probability. From the computer-generated function  $R$ , it is possible to see that the maximizing probability is  $\mu^A = \frac{1}{5} \sum_{j=0}^4 \delta_{\sigma^j((01111)^\infty)}$ . Note that the function  $R$  can vanish outside the support of  $\mu^A$ . Also note that  $\mu^A(\bar{1}) > \mu^A(\bar{0})$ .

There is no mention of subactions in the above definition of  $\Omega(A)$ .

We are interested in finding  $u$  so that  $\Omega(A) = \mathbb{M}_A(u)$ .

A subaction  $u \in C^0(\Omega)$  is said to be separating (with respect to  $A$ ) if it satisfies  $\mathbb{M}_A(u) = \Omega(A)$ ; such  $u$  provides the smallest possible  $\mathbb{M}_A(u)$ .

The next two results were proved in [21]:

**Theorem 5.** *Given an  $\alpha$ -Hölder potential  $A : \Omega \rightarrow \mathbb{R}$ , there exists an  $\alpha$ -Hölder separating subaction for  $A$ . Furthermore, in the  $\alpha$ -Hölder topology, the subset of  $\alpha$ -Hölder separating subactions is generic among all  $\alpha$ -Hölder subactions.*

**Theorem 6.** *The subset of Lipschitz continuous potentials  $A$  such that for some subaction  $u$ , the set  $\mathbb{M}_A(u) = \Omega(A)$  coincides with the support of the unique (ergodic) maximizing probability (with support in a periodic orbit) is dense in the Lipschitz topology.*

The following result is a consequence of the continuously varying support property (see [9]).

**Theorem 7.** *In the case where the unique maximizing measure for  $A$  is supported on a single periodic orbit, there exists a neighborhood  $U$  of  $A$  in the Lipschitz topology such that, for all  $C$  in  $U$ , the maximizing probability for  $C$  coincides with the maximizing probability for  $A$ .*

## 4 Ground States

Consider a variable parameter  $\beta \rightarrow \infty$ , and the associated family of equilibrium probabilities  $\mu_{\beta A}$ ,  $\beta = \frac{1}{T} > 0$ .

Any  $\nu \in \mathcal{M}_\sigma$  such that, for some sequence  $\beta_n$ , we have the weak convergence

$$\lim_{\beta_n \rightarrow \infty} \mu_{\beta_n A} = \nu,$$

is called a ground state for  $A$ . One can show that such  $\nu$  is maximizing for  $A$  (see [4], [22], [31], [11] or [36]).

When, for a given  $A$  and  $\nu$ , we have  $\lim_{\beta \rightarrow \infty} \mu_{\beta A} = \nu$ , we say that there exists selection of a maximizing probability at zero temperature. The measure  $\nu$  is the selected probability. If the maximizing probability  $\mu^A$  is unique, then  $\nu = \mu^A$ ; a unique ground state.

Suppose there exist two distinct maximizing probabilities  $\mu_1, \mu_2$ ; they could both be ground states for  $A$  (distinct examples appear in [3] and [32]).

Different subsequences  $\lim_{\beta_n \rightarrow \infty} \mu_{\beta_n A}$  could possibly converge to each of the two maximizing probabilities. Indeed, this can happen, as mentioned in [8].

In an independent work, when  $\Omega = \{0, 1\}^{\mathbb{N}}$ , the authors of [7] consider a certain subclass of potentials where the two maximizing probabilities for  $A$  are  $\mu_1 = \delta_{0\infty}$  and  $\mu_2 = \delta_{1\infty}$ ; two magnetic states at zero temperature. In some of the examples,  $A$  is of Hölder class.

**Theorem 8.** *By specifying certain explicit values of  $A$  on  $\overline{00^n 1}$ ,  $\overline{01^n 0}$ ,  $\overline{11^n 0}$  and  $\overline{10^n 1}$ ,  $n \in \mathbb{N}$ , one can produce examples where*

a)  $\lim_{\beta \rightarrow +\infty} \mu_{\beta A} = \delta_{1^\infty}$ ,

or,

b) for an explicit value  $c$ , one has  $\lim_{\beta \rightarrow +\infty} \mu_{\beta A} = \frac{1}{1+c^2} \delta_{0^\infty} + \frac{c^2}{1+c^2} \delta_{1^\infty}$ ,

or,

c) There exists a sequence  $\beta_n \rightarrow \infty$  such that

$\lim_{k \rightarrow \infty} \mu_{\beta_{2k} A} = \delta_{0^\infty}$  and  $\lim_{k \rightarrow \infty} \mu_{\beta_{2k+1} A} = \delta_{1^\infty}$ .

In another direction, in [2] (see also [37] and [32]) for a more general setting) the authors prove the following result:

**Theorem 9.** *When, for a given  $A$ , we have  $\lim_{\beta \rightarrow \infty} \mu_{\beta A} = \mu^A$ , using properties of an associated subaction, one can establish the existence of a Large Deviation Principle (LDP) when temperature goes to zero: there exists a lower semicontinuous deviation function  $I : \Omega \rightarrow [0, \infty]$  such that for any cylinder  $C$ ,*

$$\lim_{\beta \rightarrow +\infty} \frac{1}{\beta} \log \mu_{\beta A}(C) = - \inf_{x \in C} I(x) := Q(C) \leq 0.$$

When  $\lim_{\beta \rightarrow \infty} \mu_{\beta A}(C) = \mu^A(C) = 0$ ,  $Q(C) < 0$ , we obtain the exponential rate of convergence as  $\beta \rightarrow \infty$ :

$$0 \not\prec \mu_{\beta A}(C) \sim e^{\beta Q(C)}$$

One can show that  $I(x) = 0$  for  $x$  in the support of  $\mu^A$ . Indeed,  $I$  is given by

$$I(x) = \sum_{n \geq 0} (u \circ \sigma^n - u - (A - \alpha(A))) \circ \sigma^n(x) = \sum_{n \geq 0} R(\sigma^n(x)), \quad (12)$$

where  $u(x)$  is any calibrated subaction for  $A$ .

Expression (12) further highlights the importance of the role of subactions.

## 5 Maximization under constraints

For a Hölder (or continuous)  $A$ , the problem of finding maximizing probabilities for  $A$  with constraints was considered in [24]. It is a kind of generalization of the concept of rotation number to the setting of symbolic dynamics (see also [29] and [20]).

Let  $\varphi : \Omega \rightarrow \mathbb{R}^n$  be a continuous map with coordinate continuous functions  $\varphi_1, \dots, \varphi_n : \Omega \rightarrow \mathbb{R}$ ; that is,  $\varphi = (\varphi_1, \dots, \varphi_n)$ . We can define an induced map  $\varphi_* : \mathcal{M}_\sigma \rightarrow \mathbb{R}^n$  given by

$$\varphi_*(\mu) = \left( \int \varphi_1 d\mu, \dots, \int \varphi_n d\mu \right) \in \mathbb{R}^n.$$

$\varphi_*$  is a continuous and affine map.

We call  $\varphi_*(\mu)$  the rotation vector of the measure  $\mu \in \mathcal{M}_\sigma$ .

The image  $\varphi_*(\mathcal{M}_\sigma) \subset \mathbb{R}^n$  is a convex compact set. We call  $\varphi_*(\mathcal{M}_\sigma)$  the rotation set of  $\varphi$ .

For  $h \in \varphi_*(\mathcal{M}_\sigma)$ , the fiber  $\varphi_*^{-1}(h)$  is called the rotation class of  $h$ .

$$\{\mu \in \mathcal{M}_\sigma \mid \varphi_*(\mu) = h\} = \varphi_*^{-1}(h) \subset \mathcal{M}_\sigma$$

is a convex compact set.

For  $A \in C^0(\Omega)$  and  $\varphi$ , we define the so-called beta function  $\beta_{A,\varphi} : \varphi_*(\mathcal{M}_\sigma) \rightarrow \mathbb{R}$ , such that for  $h \in \mathbb{R}^n \in \varphi_*(\mathcal{M}_\sigma)$ :

$$\beta_{A,\varphi}(h) = \sup_{\mu \in \mathcal{M}_\sigma} \left\{ \int A d\mu \mid \varphi_*(\mu) = h \right\}.$$

We call the function  $\varphi$  the constraint and the function  $A$  the potential to be maximized.

Given  $A$ ,  $\varphi$  and  $h \in \mathbb{R}^n$ , we are interested in probability measures  $\mu$  belonging to the rotation class of  $h$  that maximize the integral  $\int A d\mu$ .

Consider the set

$$\mathcal{M}_{A,\varphi}(h) = \left\{ \mu \mid \varphi_*(\mu) = h \text{ and } \int A d\mu = \beta_{A,\varphi}(h) \right\} \neq \emptyset.$$

If  $\mu \in \mathcal{M}_{A,\varphi}(h)$ , we say that  $\mu$  is an  $(A, \varphi, h)$ -maximizing probability.

The following results were proved in [24].

**Theorem 10.** *Given  $\varphi$ , consider a fixed  $h \in \varphi_*(\mathcal{M}_\sigma)$ . There exists a residual subset  $\mathcal{G} = \mathcal{G}(h) \subset C^0(\Omega)$  such that, for each potential  $A \in \mathcal{G}$ , the set  $\mathcal{M}_{A,\varphi}(h)$  contains a unique probability measure.*

**Theorem 11.** *Let  $\varphi \in C^0(\Omega, \mathbb{R}^n)$  be a locally constant function. Every point of the interior of the rotation set  $\varphi_*(\mathcal{M}_\sigma)$  is the rotation vector of an ergodic probability measure.*

An example of a locally constant function is  $\varphi = (I_{0\overline{1}}, 3I_{\overline{1}0}, -4.3I_{\overline{1}\overline{1}})$ .

**Theorem 12.** *Suppose  $\varphi \in C^0(\Omega, \mathbb{Q})$  is a locally constant function, and  $A$  is a Hölder potential. Consider a rational vector  $r \in \text{int}(\varphi_*(\mathcal{M}_\sigma))$ . Then,*

$$\beta_{A,\varphi}(r) = \sup \left\{ \int A d\mu : \mu \in \varphi_*^{-1}(r), \mu \text{ a periodic probability measure} \right\}.$$

## 6 Changing the dynamics

One can consider a different dynamics and similar results are true.

Consider the transformation  $T(x) = 2x \pmod{1}$  acting on  $[0, 1]$ , and the set  $\mathcal{M}_T$  of  $T$ -invariant probabilities. Given a Hölder potential  $A : [0, 1] \rightarrow \mathbb{R}$ , we say that  $\mu^A \in \mathcal{M}_T$  is maximizing for  $A$ , if

$$\alpha(A) := \sup \left\{ \int A d\rho \mid \rho \in \mathcal{M}_T \right\} = \int A d\mu^A.$$

In a similar way as before we can say that a continuous function  $u : [0, 1] \rightarrow \mathbb{R}$  is a subaction for  $A$  if

$$u(T(x)) - u(x) - A(x) + \alpha(A) \geq 0. \quad (13)$$

By definition we set  $R$  as

$$R(x) := u(T(x)) - u(x) - A(x) + \alpha(A) \geq 0, \quad (14)$$

One can show that for any point  $x$  in the support of a maximizing measure for  $A$ , we have  $R(x) = 0$ . In the same way as before, we can take advantage of the 1/2 iterative procedure (using your desktop) to get explicit results; which can be checked with pen and paper. In this case, the method is more direct than the case of the symbolic space; you do not need to use the binary expansion (11) (see [16]).

We show in Figure 5 the graph of a potential  $A : [0, 1] \rightarrow \mathbb{R}$  (which is linear by part), the graph of the calibrated subaction  $u$  (we get from the iterative procedure), and the graph of  $R$ . It follows from a simple inspection analysis of the function  $R$  that the maximizing probability is  $\mu^A = \frac{1}{2}\delta_{1/3} + \frac{1}{2}\delta_{2/3}$ .

In Figure 6 we exhibit one more application of the use of the 1/2 iterative procedure taken from [17] (see also [10]).

In order to test its efficiency, in [17] the method was applied to the potential  $A(x) = -|x - K|$ , where  $K$  is the Cantor set in  $[0, 1]$ . In this case, a large number of iterations are needed to get a reasonable approximation of  $u$  (producing  $x \rightarrow R(x)$ , such that  $R(x)$  is very close to zero for  $x \in K$ ).

Some related references on Ergodic Optimization for one-dimensional maps are [19], [25] and [42].

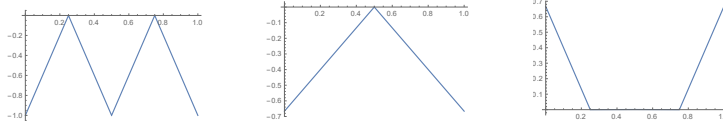


Figure 5: From left to right: the graph of the potential  $A$ , the graph of the calibrated subaction  $u$  and the graph of  $R$ . The dynamics is given by the map  $T(x) = 2x \pmod{1}$  and the maximizing probability is  $\mu^A = \frac{1}{2}\delta_{1/3} + \frac{1}{2}\delta_{2/3}$ .

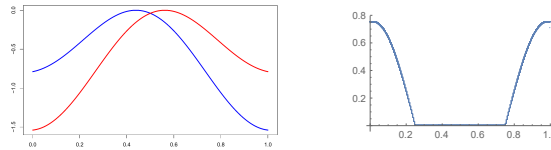


Figure 6: Case  $A(x) = \sin^2(2\pi x)$  - From the  $1/2$  iterative procedure: on the left side, we show the approximated subaction  $u$  which is given by the supremum of the two functions in red and in blue. The graph of  $R$  using the approximation of the subaction  $u$  is shown in the right-hand picture. The orbit of period 2 given by  $\frac{1}{2}\delta_{1/3} + \frac{1}{2}\delta_{2/3}$  is inside the set  $R = 0$ , and therefore is a maximizing probability  $\mu^A$  for  $A$ . The expression of the two functions on the left side is given explicitly by a convergent series (see Section 5 in [17]), and the use of the method shows a very good approximation of the analytical expression.

## References

- [1] V. Baladi, Positive Transfer Operators and Decay of Correlations, World Scientific (2000)
- [2] A. Baraviera, A. O. Lopes and Ph. Thieullen, A Large Deviation Principle for Gibbs states of Hölder potentials: the zero temperature case. Stoch. and Dyn. (6), 77-96, (2006).
- [3] A. Baraviera, R. Leplaideur and A. O. Lopes, Selection of ground states in the zero temperature limit for a one-parameter family of potentials, SIAM Journal on Appl. Dyn. Systems, Vol. 11, no. 1, 243-260 (2012)
- [4] A. Baraviera, R. Leplaideur and A. O. Lopes, Ergodic Optimization, Zero Temperature Limits and the Max-Plus Algebra, XXIX Colóquio Brasileiro de Matemática (2013)

- [5] J. Bochi, Ergodic optimization of Birkhoff averages and Lyapunov exponents. In Proc. Internat. Congr. Math. (Rio de Janeiro 2018), Volume III, World Sci. Publ., Singapore, pp. 1843-1866 (2018).
- [6] J. Bochi and Y. Zhang, Ergodic optimization of prevalent super-continuous functions. *Int. Math. Res. Not. IMRN*, 19, 5988-6017 (2016).
- [7] R. Bissacot, E. Garibaldi and Ph. Thieullen, Zero-temperature phase diagram for double-well type potentials in the summable variation class, *ETDS* 38, no. 3, 863-885 (2018)
- [8] J.-R. Chazottes, J.-M. Gambaudo and E. Ugalde, Zero-temperature limit of one-dimensional Gibbs states via renormalization: the case of locally constant potentials, *Erg. Theo. and Dyn. Sys.* 1109-1161(2010).
- [9] G. Contreras, A. O. Lopes and Ph. Thieullen. Lyapunov minimizing measures for expanding maps of the circle, *ETDS*, 21,1379-1409 (2001).
- [10] J. P. Conze and Y. Guivarc'h. Croissance des sommes ergodiques et principe variationnel. manuscript circa (1993).
- [11] D. Coronel and J. Rivera-Letelier, Sensitive dependence of Gibbs measures at low temperatures, *Jou. St. Phys*, 160, Issue 6, 1658-1683 (2015)
- [12] L. Cioletti, M. Denker, A. O. Lopes and M. Stadlbauer, Spectral Properties of the Ruelle Operator for Product Type Potentials on Shift Spaces, *Jour. of the London Math. Soc.*, Volume 95, Issue 2, 684-704 (2017)
- [13] G. Contreras, Ground states are generically a periodic orbit, *Invent. Math.*, no. 2, 383-412 (2016)
- [14] G. Contreras and R. Iturriaga. Global minimizers of autonomous Lagrangians, 22° Colóquio Brasileiro de Matemática, IMPA (1999).
- [15] A. Fathi, Weak KAM Theorem in Lagrangian Dynamics, Lecture Notes, Pisa (2005)
- [16] H. H. Ferreira, A. O. Lopes and E. R. Oliveira, An iterative process for approximating subactions, "Modeling, Dynamics, Optimization IV", Editors: Alberto Pinto and David Zilberman, 187-212 (2021)
- [17] H. H. Ferreira, A. O. Lopes and E. R. Oliveira, Explicit examples in Ergodic Optimization, *S. P. Jou. of Math. Scien*, Vol. 14, 443-489 (2020)

- [18] H. H. Ferreira, Python Code used to generate the figures for the explicit examples, written in Jupyter notebooks.  
[https://github.com/hermes-hf/Explicit\\_examples\\_ergodic](https://github.com/hermes-hf/Explicit_examples_ergodic)
- [19] R. Gao, W. Shen and R. Zhang, Typicality of periodic optimization over an expanding circle map, arXiv (2025)
- [20] P. Giulietti, B. Kloeckner, A. O. Lopes and D. Marcon, The calculus of thermodynamical formalism, Journ. of the European Math Society, Vol 20, Issue 10, pages 2357-2412 (2018)
- [21] E. Garibaldi, A. O. Lopes and P. Thiellien, On calibrated and separating sub-actions, Bull. of the Bras. Math. Soc. Vol. 40, 577-602 (2009).
- [22] E. Garibaldi, Ergodic Optimization in the Expanding Case (2017)
- [23] E. Garibaldi and A. O. Lopes, On Aubry–Mather theory for symbolic dynamics, ETDS, Vol. 28, Issue 3, 791-815 (2008)
- [24] E. Garibaldi and A. O. Lopes, Functions for relative maximization, Dynamical Systems, v. 22, 511-528 (2007)
- [25] W. Huang, L. Xu and D. Yang, Lyapunov Optimizing Measures and Periodic Measures for  $C^2$  Expanding Maps, Acta Mathematica Sinica, Volume 41, pages 2259-2274 (2025)
- [26] O. Jenkinson, Ergodic optimization, Discrete and Continuous Dynamical Systems, Series A, V. 15, 197-224 (2006)
- [27] O. Jenkinson, Ergodic optimization in dynamical systems, Ergodic Theory and Dynamical Systems, pp. 2593–2618 (2019)
- [28] O. Jenkinson, X. Li and Y. Zhang, Typical Uniqueness in Ergodic Optimization, arXiv (2025)
- [29] T. Kucherenko and C. Wolf, Geometry and entropy of generalized rotation sets, Israel Journal of Mathematics 199, 791-829 (2014)
- [30] S. Lalley, Distribution of Periodic Orbits of Symbolic and Axiom A Flows, Adv. in Appl. Math, 8, 154-193 (1987)
- [31] R. Leplaideur, A dynamical proof for the convergence of Gibbs measures at temperature zero, Nonlinearity, 18(6):2847–2880 (2005)

- [32] R. Leplaideur and J. K. Mengue, On the selection of subaction and measure for perturbed potentials, *ETDS*, Vol. 46, 3, 660-694 (2026)
- [33] A. O. Lopes, J. K. Mengue, J. Mohr and R. R. Souza, Entropy and Variational Principle for one-dimensional Lattice Systems with a general a priori probability: positive and zero temperature, *Erg. Theory and Dyn. Systems*, 35 (6), 1925-1961 (2015)
- [34] A. O. Lopes, E. R. Oliveira and Ph. Thiullen, The dual potential, the involution kernel and transport in ergodic optimization, *Dynamics, Games and Science*, Springer Verlag, 331-398 (2015)
- [35] A. O. Lopes, E. R. Oliveira, W. de S. Pedra and V. Vargas, Grand-canonical Thermodynamic Formalism via IFS: volume, temperature, gas pressure and grand-canonical topological pressure, *Sao Paulo Journal of Math. Scien.*, Volume 19, article number 30 - pp 1-34 (2025)
- [36] A. O. Lopes, Thermodynamic Formalism, Maximizing Probabilities and Large Deviations — Notes UFRGS — online homepage
- [37] J. K. Mengue, Large Deviations for Equilibrium Measures and Selection of Subaction. *Bull. of the Braz. Math. Soc.*, v. 49, p. 17-42 (2018)
- [38] I. Morris, Prevalent uniqueness in ergodic optimisation, *Proc. Amer. Math. Soc.* 149, no. 4, 1631-1639 (2021)
- [39] W. Parry and M. Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics, *Astérisque* Vol. 187-188 (1990).
- [40] A. Quas and J. Siefken, Ergodic optimization of super-continuous functions on shift spaces. *ETDS*, 32, 6, 2071-2082 (2012)
- [41] K. Sigmund, On dynamical systems with the specification property. *Trans. Amer. Math. Soc.* 190, 285-299 (1974)
- [42] Rafael R. Souza, Sub-actions for weakly hyperbolic one-dimensional systems, *Dynamical Systems*, Volume 18, Issue 2, 165-179 (2003)
- [43] M. Viana and K. Oliveira, *Foundations of Ergodic Theory*, Cambridge University Press (2016)
- [44] P. Walters, *Introduction to Ergodic Theory*, Springer Verlag (1982)