A CONVEX ANALYSIS APPROACH TO ENTROPY FUNCTIONS, VARIATIONAL PRINCIPLES AND EQUILIBRIUM STATES

ANDRZEJ BIŚ, MARIA CARVALHO, MIGUEL MENDES, AND PAULO VARANDAS

ABSTRACT. Using methods from Convex Analysis, for each generalized pressure function we define an upper semi-continuous affine entropy-like map, establish an abstract variational principle for both countably and finitely additive probability measures and prove that equilibrium states always exist. We show that this conceptual approach imparts a new insight on dynamical systems without a measure with maximal entropy, may be used to detect second-order phase transitions, prompts the study of finitely additive ground states for non-uniformly hyperbolic transformations and grants the existence of finitely additive Lyapunov equilibrium states for singular value potentials generated by linear cocycles over continuous self-maps.

1. INTRODUCTION

Classical thermodynamic formalism. The modern theory of Dynamical Systems has its origins in the end of the nineteenth century with the pioneering work of Poincaré, who aimed at a complete description of the solutions of the differential equations modeling the three body problem in Celestial Mechanics. It is in the course of this investigation that Poincaré encounters the intriguing phenomenon of what was later named homoclinic tangencies. Several decades after this first surfacing of such complex dynamical behavior a common opinion had grown among researchers that geometric methods were insufficient to fully describe the asymptotic behavior of dynamical systems in general. In the meantime, some of the ideas from Statistical Mechanics (e.g. Boltzman ergodic hypothesis), dated from the previous century, were spreading around, and the main ergodic results of that period, namely von Neumann's and Birkhoff's ergodic theorems, were published. Yet, Kolmogorov's groundbreaking proposal made in the late 1950's of bringing both Probability and Entropy Theories into the realm of Dynamical Systems faced intrinsic difficulties, inasmuch as the latter domain of study had not yet been established as an independent domain of research.

The theory of uniformly hyperbolic dynamical systems, essentially born in the sixties and building over the existence of invariant foliations with exponential contracting and expanding behavior, brought forward both the models and the axiomatic framework which granted some understanding of the homoclinic behavior and support to the realization of Kolmogorov's insight. Meanwhile, the paradigmatic example of Smale's horseshoe made clear that subshifts of finite type could codify the hyperbolic dynamical systems, and that such coding was possible due to the existence of finite partitions with a Markov property. What's more, the finite Markov partitions for hyperbolic dynamical systems (both diffeomorphisms and flows) had a further

benefit: they conveyed the thermodynamic formalism from Statistical Mechanics to Dynamical Systems, through keystone contributions by Sinai, Ruelle, Bowen, Ratner and Walters [18, 20, 76, 80, 85, 91], among others.

The thermodynamic formalism for dynamical systems aims to prove the existence of invariant probability measures which maximize the topological pressure, besides reporting on their statistical properties. Such measures, called equilibrium states, are often Gibbs, and include as specific examples both probability measures absolutely continuous with respect to Lebesgue and measures of maximal entropy. For uniformly hyperbolic diffeomorphisms and flows, when restricted to a basic piece of the non-wandering set, equilibrium states exist and are unique for every Hölder continuous potential (cf. [19, 80, 85]). In the case of diffeomorphisms, the basic known strategy to prove this remarkable fact is to (semi-)conjugate the dynamics to a subshift of finite type, *via* a finite Markov partition. Finer properties, including the analiticity of the pressure map and the relation between pressure, periodic orbits and dynamical zeta functions, were later addressed by Parry and Pollicott [67].

Non-uniform hyperbolicity and phase transitions. The study of dynamical systems with weaker forms of hyperbolicity, such as partial hyperbolicity and dominated splittings, is still under development (cf. [17]). In particular, an extension of the thermodynamic formalism beyond the scope of uniform hyperbolicity has been facing several intricacies. A very fruitful strategy to overcome some of them is the construction of induced and tower dynamics with hyperbolic behavior (cf. [70, 82]). These induced maps can be well described by countable shifts, for which the thermodynamic formalism is more or less settled [83], but introduces two new hindrances. On the one hand, not all invariant probability measures may be lifted to the induced tower dynamics (see [22] and references therein); on the other hand, equilibrium states for the tower dynamics may induce probability measures that are σ-finite on the phase space but this transfer process depends on the integrability of the return time function. While in the case of C^∞-surface diffeomorphisms with positive entropy there are finitely many measures of maximal entropy (cf. [27]), the previous two issues are among the reasons why the theory of thermodynamic formalism for non uniformly hyperbolic dynamical systems remains incomplete even for partially hyperbolic systems.

It is in this context that phase transitions (characterized either by the non-analiticity of the pressure map or by a discontinuity on the number of its equilibrium states) have been thoroughly studied. Phase transitions are reasonably well understood in the one-dimensional context (see [34, 51, 72, 71] and references therein) due to the recent tools to detect and characterize the sources of non-hyperbolic behavior. Beyond hyperbolicity even the existence of equilibrium states is far from being established, though the non-uniform hyperbolicity may be enough in special contexts, as happens in the case of Sinai billiard maps (cf. [5]). Even so, Newhouse described in [64] sufficient conditions for the entropy map to be upper semi-continuous, which in particular ensure that C^∞ maps of compact manifolds have equilibrium states for every continuous potential. In addition, a codification mechanism by symbolic extensions, and its existence for smooth dynamics, were also explored by Downarowicz and Newhouse in [38]. Nonetheless, one can ask under what general conditions do equilibrium states exist, seeing that there are C^r-surface diffeomorphisms, $1 \leqslant r < +\infty$, having no measure with maximal entropy (cf. [26]). One of the main goals in this paper is to prove that they always exist if one drops the requirement that they are countably additive.

The use of Convex Analysis. One of the fundamental mathematical tools used in Thermodynamics, Statistical Physics and Stochastic Finance is convexity. For instance, regarding the theory of lattice gases, Israel observes in [52] that the pressure exists in the thermodynamic limit as a convex function in the space of interactions; and uses the pioneer work of Bishop and Phelps on Convex Analysis [14] to construct equilibrium states, to prove that they satisfy Dobrushin-Lanford-Ruelle condition and to describe phase transitions. The aforesaid paper [14] is also a cornerstone for several recent applications in Economics, namely on equilibrium theory, risk measures and stochastic finance (cf. [47] and references therein), where abstract variational principles appear associated to convex risk measures and penalty functions.

An important application of the Convex Analysis methods in the thermodynamic formalism within statistical mechanics is due to Israel and Phelps [53] who, inspired by mathematical models in the classical theory of lattice gases [80], investigated the differentiability, the tangent functionals and the variational entropy (defined by means of the Legendre-Fenchel transform) in the case of generalized pressure functions acting on the space of affine real-valued continuous functions whose domain is a compact convex set. Apart from the natural convexity and continuity assumptions, and the fact that the pressure function acts on a set of affine maps, such pressure maps were assumed to also satisfy a strong positivity condition (cf. [53, p.136]), a hypothesis which seldom occurs in the thermodynamic formalism context. Encouraged by the works of Israel and Phelps [53] and Lopes et al [59], recent formulations of variational principles using Convex Analysis have appeared in [33, 32, 48], extending [53] to the dynamical context of shifts on Borel standard spaces and dynamical systems whose transfer operators have a spectral gap. Since the usual Kolmogorov-Sinai metric entropy is not adequate to the context of shifts on Borel standard spaces, these authors define a notion of entropy (which they also call variational entropy) for measures that are eigenvectors of a normalized transfer operator's dual, and prove, as a variational principle, that it coincides with the Legendre transform of the spectral radius of a suitable operator. Similar variational principles have been obtained in the context of weighted shift spaces acting on continuous or L^1-functions (see [2, 7]). In both situations, it may happen that the pressure function, defined as the logarithm of the spectral radius of the corresponding transfer operator, even if well adapted to the transfer operator, is hardly related to the classical notion of topological pressure for dynamical systems.

In what follows, we will use Banach function spaces that contain non-continuous observable maps, which appear naturally in the context of piecewise smooth dynamical systems (see Subsection 7.5). We will focus on the space $B_m(X)$ of bounded measurable functions on a compact metric space, in which case the Convex Analysis dualities produce an entropy function on the space of finitely additive measures. In the seventies, Salomon Bochner wrote that, contrary to popular mathematical opinions, finitely additive measures were more interesting and perhaps more important than countable additive ones (see [60]). Surely, finitely additive measures arise in Analysis, Statistics, Measure Theory and Dynamical Systems (see [12, 33] and references therein), and are closely related to the Banach-Tarski paradox (cf. [90]) and the characterization of amenable groups (cf. [68]). We will show in this paper that they can also be used to detect second-order phase transitions in the classical thermodynamic formalism (see Theorem 1 and Subsection 7.6).

Our main contributions. We address the following problem: to find an abstract variational principle, valid for real-valued convex, monotone and translation invariant functions defined on a suitable Banach space of potentials, which is powerful enough to be applied to either the

classical topological pressure for a continuous self-map or to the more recently defined notions of pressure for semigroup actions. Inspired by some recent applications of convex analysis to risk measures in stochastic finance [47] and by differentiability results on the classical topological pressure function [92], we obtain a variational principle under great generality, which conveys an upper semi-continuous entropy-like map acting either on the space of finitely additive normalized set functions or on the space probability measures (cf. Theorem 1). Furthermore, we prove that these entropy-like functions, which are Fenchel-Legendre transforms of abstract pressure functions, are affine. This is a crucial property for both the study of differentiability properties of the pressure functions (see Theorem 3 and Corollary 4) and the characterization of the entropy-like function determined by the classical topological pressure, when restricted to the set of invariant probability measures, as the upper semi-continuous concave envelope of the measure-theoretic entropy (cf. Subsection 6.1.1 for the precise definitions and Theorem 5).

In the special case of entropy functions defined over the space of probability measures, our contributions not only recover the classical variational principle for the topological pressure of continuous potentials but also improve the thermodynamic formalism of dynamics which do not admit equilibrium states, and allow us to address the more general family of bounded potentials (see Theorem 5 for the precise statement). The latter turned out to be of special interest in order to consider the pressure functions linked to non-additive sequences of continuous potentials as singular value potentials associated to linear cocycles. Indeed, in this context we show that the pressure function of such non-additive sequences coincides with the pressure function of a single bounded observable, for which our results guarantee the existence of equilibrium states (see Theorem 8.6 and Subsection 8.2 for more details).

The need to deal with finitely additive measures (respectively, probability measures) is a consequence of the duality theorem for bounded measurable functions (respectively, continuous functions with compact support) employed in the proof of Theorem 1. For finitely additive set functions there are known versions of Poincaré recurrence theorem [9], Birkhoff ergodic theorem [75], Central limit theorem [55, 74], a strong law of large numbers [30, 31] and a Radon-Nikodym theorem [61], besides information on the accumulation points of sequences of finitely additive set functions [65]. Other useful properties of these set functions may be found in [73, 40, 93, 36, 24] and in the comprehensive book [86].

Finally, we observe that the abstract variational principle in Theorem 1 does not summon any dynamics, which makes it applicable in many contexts. In [13], using several suitable notions of pressure, we have applied it to semigroup actions which are finitely generated by continuous maps, a setting which comprises free semigroups, countable sofic groups and groups endowed with a reference measure.

2. Main results

2.1. **Entropy functions for abstract pressure functions.** Let (X, d) be a metric space and let $\mathfrak{B}$ stand for the σ-algebra of Borel subsets of X. Denote by $\mathbf{B}$ a Banach space over the field $\mathbb{R}$ equal to either

$$
\begin{aligned}
B_m(X) &= \big\{\varphi : X \to \mathbb{R} \mid \varphi \text{ is Borel measurable and bounded}\big\} &\quad (2.1) \\
\text{or } C_b(X) &= \big\{\varphi \in B_m(X) \mid \varphi \text{ is continuous}\big\} \\
\text{or else } C_c(X) &= \big\{\varphi \in C_b(X) \mid \varphi \text{ has compact support}\big\}
\end{aligned}
$$

endowed with the norm $\|\varphi\|_\infty = \sup_{x \in X} |\varphi(x)|$. In what follows, $\mathcal{P}_a(X)$ will stand for the set of $\mathfrak{B}$-measurable, regular and normalized finitely additive set functions on X, which we will simply call *finitely additive probability measures*, with the total variation norm (cf. [41, IV.2.15]). Recall that the total variation norm in $\mathcal{P}_a(X)$ is given by

$$\|\mu - \nu\| = \sup \left\{ \left| \int \psi \, d\mu - \int \psi \, d\nu \right| : \ \psi \in \mathbf{B} \text{ and } \|\psi\|_\infty \leqslant 1 \right\}.$$

As the set of regular, finitely additive measures endowed with the total variation norm is a Banach space isometrically isomorphic to the topological dual of $C_b(X)$ (cf. [1, Theorem 14.9]), the set $\mathcal{P}_a(X)$ is compact in the weak* topology as a consequence of the classical Banach-Alaoglu theorem. For future use, we denote by $\mathcal{P}(X) \subset \mathcal{P}_a(X)$ the set of Borel σ-additive probability measures on X endowed with the weak* topology, and by $C(X)$ the space of real valued continuous maps whose domain is X. Inspired by [47], we introduce the following axiomatic definition of pressure function.

Definition 2.2. A function $\Gamma : \mathbf{B} \to \mathbb{R}$ is called a *pressure function* if it satisfies the following conditions:

(C$_1$) *Monotonicity*: $\varphi \leqslant \psi \ \Rightarrow \ \Gamma(\varphi) \leqslant \Gamma(\psi) \quad \forall \varphi, \psi \in \mathbf{B}$.

(C$_2$) *Translation invariance*: $\Gamma(\varphi + c) = \Gamma(\varphi) + c \quad \forall \varphi \in \mathbf{B} \quad \forall c \in \mathbb{R}$.

(C$_3$) *Convexity*: $\Gamma(t \, \varphi + (1 - t) \, \psi) \leqslant t \, \Gamma(\varphi) + (1 - t) \, \Gamma(\psi) \quad \forall \varphi, \psi \in \mathbf{B} \quad \forall t \in [0, 1]$.

We note that, in comparison to endomorphisms, pressure functions and the abstract framework of [53], here we do not require the pressure function to preserve co-boundary type functions. This is crucial in [13], where we deal with very abstract pressure functions associated to group and semigroup actions.

It is not hard to check that properties (C$_1$) and (C$_2$) imply that any pressure function is Lipschitz continuous, meaning that $|\Gamma(\varphi) - \Gamma(\psi)| \leqslant \|\varphi - \psi\|_\infty$ for every $\varphi, \psi \in \mathbf{B}$.

Our first result, inspired by [47], establishes a general variational principle and the existence of finitely additive equilibrium states for any pressure function.

Theorem 1. *Let $\Gamma : \mathbf{B} \to \mathbb{R}$ be a pressure function. Then*

$$\Gamma(\varphi) = \max_{\mu \in \mathcal{P}_a(X)} \left\{ \mathfrak{h}(\mu) + \int \varphi \, d\mu \right\} \qquad \forall \varphi \in \mathbf{B} \tag{2.3}$$

where

$$\mathfrak{h}(\mu) = \inf_{\varphi \in \mathcal{A}_\Gamma} \left\{ \int \varphi \, d\mu \right\} \qquad and \qquad \mathcal{A}_\Gamma = \left\{ \varphi \in \mathbf{B} : \Gamma(-\varphi) \leqslant 0 \right\}. \tag{2.4}$$

Moreover, $\mathfrak{h}(\mu)$ is affine and upper semi-continuous; and if $\alpha : \mathcal{P}_a(X) \to \mathbb{R} \cup \{-\infty, +\infty\}$ is another function taking the role of $\mathfrak{h}$ in (2.3), then $\alpha \leqslant \mathfrak{h}$. In addition, one has

$$\mathfrak{h}(\mu) = \inf_{\varphi \in \mathbf{B}} \left\{ \Gamma(\varphi) - \int \varphi \, d\mu \right\} \qquad \forall \mu \in \mathcal{P}_a(X).$$

If X is locally compact and $\mathbf{B} = C_c(X)$ then the maximum in (2.3) is attained in $\mathcal{P}(X)$.

Some comments are in order. Theorem 1 can be understood as a variant of the Fenchel-Moreau Theorem in Convex Analysis (see Theorem VI.5.3 in [43]). We stress that we do not request compactness of the space X, just local compactness when considering $\mathbf{B} = C_c(X)$ due to the use of the Riesz-Markov Theorem, and that the supremum in (2.3) can be computed

taking only extremal measures (cf. (5.3)). We also highlight the fact that (2.4) asserts that the function $\mathfrak{h}$ is computed by averaging potentials whose additive inverses have non-positive pressure. Observe, in addition, that $\mathfrak{h}(\mu) \geqslant 0$ if and only if $\Gamma(\varphi) \geqslant \int \varphi \, d\mu \quad \forall \mu \in \mathcal{P}_a(X)$, a condition which extends a characterization of invariant probability measures for maps with upper semi-continuous Kolmogorov-Sinai entropy (cf. [91, Theorems 9.11 & 9.12]).

Remark 2.5. The entropy function $\mathfrak{h} = \mathfrak{h}_{\Gamma, \mathbf{B}}$ depends on both the pressure function Γ and its domain. For instance, if X is a compact metric space and $\Gamma : B_m(X) \to \mathbb{R}$ is a pressure function then $\Gamma \big|_{C(X)} : C(X) \to \mathbb{R}$ is a pressure function as well, though with a different domain; and, according to (2.4), the corresponding functions $\mathfrak{h}$ are computed respectively by

$$\mathfrak{h}_{B_m(X)}(\mu) = \inf \left\{ \int \varphi \, d\mu \colon \varphi \in B_m(X) \text{ and } \Gamma(-\varphi) \leqslant 0 \right\} \quad \forall \mu \in \mathcal{P}_a(X)$$

and

$$\mathfrak{h}_{C(X)}(\mu) = \inf \left\{ \int \varphi \, d\mu \colon \varphi \in C(X) \text{ and } \Gamma(-\varphi) \leqslant 0 \right\} \qquad \forall \mu \in \mathcal{P}(X).$$

In particular, as $\mathfrak{h}_{B_m(X)}$ is computed on the larger space of finitely additive measures, one has that $\mathfrak{h}_{B_m(X)}(\mu) \leqslant \mathfrak{h}_{C(X)}(\mu)$ for every $\mu \in \mathcal{P}(X)$.

Remark 2.6. It is straightforward to conclude that μ_0 attains the minimum in (2.4) (that is, $\mathfrak{h}(\mu_0) = \int \varphi_0 \, d\mu_0$ for some $\varphi_0 \in \mathcal{A}_\Gamma$) if and only if $\Gamma(-\varphi_0) = 0$ and μ_0 is an equilibrium state of $-\varphi_0$. For instance, the minimum is attained when μ_0 is a measure of maximal entropy (in which case $\varphi_0 = h_{\text{top}}(f)$); or when μ_0 is the Sinai-Ruelle-Bowen measure of a C^2 Axiom A attractor (in which case $\varphi_0 = -\log \text{Jac}(Df_{|E^u})$; cf. [19]).

2.2. **Uniqueness of the equilibrium states.** The variational principle stated in Theorem 1 ensures that there always exist normalized finitely additive measures for which the right-hand side of (2.3) attains the supremum; that is, the set $\mathcal{E}_\varphi(\Gamma) = \left\{ \mu \in \mathcal{P}_a(X) \colon \Gamma(\varphi) = \mathfrak{h}(\mu) + \int \varphi \, d\mu \right\}$ is non-empty. This raises the subtle question of whether they are unique. To address this issue we consider functionals tangent to the pressure in our abstract framework.

Given a pressure function $\Gamma : \mathbf{B} \to \mathbb{R}$ and a potential $\varphi \in \mathbf{B}$, we say $\mu \in \mathcal{P}_a(X)$ is a *tangent functional to Γ at φ* (also known as *sub-differential*) if

$$\Gamma(\varphi + \psi) - \Gamma(\varphi) \geqslant \int \psi \, d\mu \qquad \forall \, \psi \in \mathbf{B}. \tag{2.7}$$

As in [92], the continuity of Γ and the Hahn-Banach Theorem guarantee that the space $\mathcal{T}_\varphi(\Gamma)$ of tangent functionals to Γ at φ is non-empty for every φ; and it is easily seen to be a convex and weak* compact set. The use of the Hahn-Banach theorem makes that argument for the existence of such measures is not constructive. The next result states that the space of tangent functionals to Γ at $\varphi \in \mathbf{B}$ coincides with the space of finitely additive probability measures attaining the maximum on (2.3), and that typical potentials have a unique equilibrium state. More precisely:

Theorem 2. *Let $\Gamma : \mathbf{B} \to \mathbb{R}$ be a pressure function. Then*

$$\mathcal{E}_\varphi(\Gamma) = \mathcal{T}_\varphi(\Gamma) \qquad \forall \, \varphi \in \mathbf{B}.$$

Moreover, if $\mathbf{B} = C_b(X)$ or $\mathbf{B} = C_c(X)$, then there exists a residual subset $\mathfrak{R} \subset \mathbf{B}$ such that $\# \mathcal{E}_\varphi(\Gamma) = 1$ for every $\varphi \in \mathfrak{R}$.

The previous result extends the work of Walters [92] within the context of the classical topological pressure, for which the uniqueness of equilibrium states is tied in with the differentiability of the pressure.

A pressure function $\Gamma : \mathbf{B} \to \mathbb{R}$ is said to be *locally affine at* $\varphi \in \mathbf{B}$ if there exist a neighborhood $\mathcal{V}$ of 0 in $\mathbf{B}$ and a unique $\mu_\varphi \in \mathcal{P}_a(X)$ such that

$$\Gamma(\varphi + \psi) - \Gamma(\varphi) = \int \psi \, d\mu_\varphi \qquad \forall \, \psi \in \mathcal{V}. \tag{2.8}$$

In particular, $\mathcal{T}_\varphi(\Gamma) = \{\mu_\varphi\}$. If Γ is locally affine at all elements of $\mathbf{B}$, then μ_φ does not depend on φ, and Γ is said to be *affine*. We say that the pressure function $\Gamma : \mathbf{B} \to \mathbb{R}$ is *Fréchet differentiable* at $\varphi \in \mathbf{B}$ if there exists a unique $\mu_\varphi \in \mathcal{P}_a(X)$ such that

$$\lim_{\psi \to 0} \frac{1}{\|\psi\|_\infty} \left| \Gamma(\varphi + \psi) - \Gamma(\varphi) - \int \psi \, d\mu_\varphi \right| = 0. \tag{2.9}$$

It is known (cf. [92, Theorem 6]) that the local affine property is equivalent to Fréchet differentiability when one considers the classical topological pressure function. We will show that the same statement for general pressure functions is also true.

Theorem 3. *Let $\Gamma : \mathbf{B} \to \mathbb{R}$ be a pressure function. The following assertions are equivalent:*

(a) Γ *is locally affine at* φ.
(b) *There exists a unique tangent functional in* $\mathcal{T}_\varphi(\Gamma)$ *and*

$$\lim_{\psi \to 0} \sup \left\{ \|\mu - \mu_\varphi\| : \ \mu \in \mathcal{T}_{\varphi+\psi}(\Gamma) \right\} = 0.$$

(c) Γ *is Fréchet differentiable at* φ.

Therefore, the following statements are mutually equivalent as well: $(\bar{a})$ Γ is affine, $(\bar{b})$ $\bigcup_{\varphi \in \mathbf{B}} \mathcal{T}_\varphi(\Gamma)$ is a singleton, and $(\bar{c})$ Γ is everywhere Fréchet differentiable.

As being affine is a rigid condition, the previous theorem also conveys the information that a pressure function Γ is rarely everywhere Fréchet differentiable. Consequently, typical pressure functions either exhibit more than one tangent functional at some element of $\mathbf{B}$, or these do not vary continuously in the operator norm (see [37] for examples). The previous discussion prompts us to consider the weaker notion of Gateaux differentiability. A pressure function $\Gamma : \mathbf{B} \to \mathbb{R}$ is said to be *Gateaux differentiable at* $\varphi \in \mathbf{B}$ if, for every $\psi \in \mathbf{B}$ the directional pressure map $t \in \mathbb{R} \mapsto \Gamma(\varphi + t\psi)$ is differentiable, that is, the limit

$$d\,\Gamma(\varphi)(\psi) \ := \ \lim_{t \to 0} \frac{1}{t} \left[\Gamma(\varphi + t\psi) - \Gamma(\varphi) \right]$$

exists and is finite for every $\psi \in \mathbf{B}$. Concerning real valued convex functions on Banach spaces, Walters proved in [92, Corollary 2] a criterion for Gateaux differentiability. In our setting, the corresponding statement reads as follows.

Corollary 4. *Let $\Gamma : \mathbf{B} \to \mathbb{R}$ be a pressure function. Then Γ is Gateaux differentiable at φ if and only if there exists a unique tangent functional in $\mathcal{T}_\varphi(\Gamma)$.*

2.3. **Existence of equilibrium states.** The duality results in Theorem 1 can also be used to obtain a new insight on the thermodynamic formalism of a continuous self-map on a compact metric space. Indeed, in this context we provide the following variational principle regarding the classical topological pressure (we refer the reader to Subsection 7.2 for the definitions).

Theorem 5. *Let* $f : X \to X$ *be a continuous transformation of a compact metric space* X *with* $h_{\text{top}}(f) < +\infty$. *The upper semi-continuous entropy map* $\mathfrak{h}_f : \mathcal{P}(X) \to \mathbb{R}$ *given by*

$$\mathfrak{h}_\mu(f) = \inf_{\varphi \in C(X)} \left\{ P_{\text{top}}(f, \varphi) - \int \varphi \, d\mu \right\} \qquad \forall \, \mu \in \mathcal{P}(X)$$

satisfies:

(a) $0 \leqslant h_\mu(f) \leqslant \mathfrak{h}_\mu(f) \quad \forall \, \mu \in \mathcal{P}_f(X)$.

(b) *For every continuous potential* $\varphi : X \to \mathbb{R}$,

$$P_{\text{top}}(f, \varphi) = \max_{\mu \in \mathcal{P}(X)} \left\{ \mathfrak{h}_\mu(f) + \int \varphi \, d\mu \right\} = \max_{\mu \in \mathcal{P}_f(X)} \left\{ \mathfrak{h}_\mu(f) + \int \varphi \, d\mu \right\} \tag{2.10}$$

where $\mathcal{P}_f(X)$ *denotes the space of* f*-invariant Borel probability measures on* X *endowed with the weak* topology.*

(c) *Every measure* $\mu \in \mathcal{P}(X)$ *which attains the maximum* (2.10) *is* f*-invariant.*

(d) *The restriction of* $\mathfrak{h}_f$ *to* $\mathcal{P}_f(X)$ *is the upper semi-continuous concave envelope of Kolmogorov-Sinai metric entropy.*

The elements of the set

$$\mathscr{D} = \left\{ \mu \in \mathcal{P}_f(X) \colon 0 \leqslant h_\mu(f) < \mathfrak{h}_\mu(f) \right\} \tag{2.11}$$

are the f-invariant probability measures where the entropy function $\mu \in \mathcal{P}_f(X) \mapsto h_\mu(f)$ fails to be upper semi-continuous. The classical variational principle together with Theorem 5 (b) imply that

$$\sup_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu(f) + \int \varphi \, d\mu \right\} = \sup_{\mu \in \mathcal{P}_f(X)} \left\{ \mathfrak{h}_\mu(f) + \int \varphi \, d\mu \right\}$$

even if $\mathscr{D} \neq \emptyset$. This shows that the measures in $\mathscr{D}$ (which play a role for instance if the entropy spectrum $\{h_\mu(f) \colon \mu \in \mathcal{P}_f(X)\}$ is not an interval) do not affect the supremum. Still, such measures have a say when we look for equilibrium states (see an example in Subsection 7.4). More details on the relation between the star-entropy and the $\mathfrak{h}$-entropy are given in Proposition 7.4. Finally, it is worthwhile mentioning that one can obtain a similar statement to the one of Theorem 5 using finitely additive measures instead of probability measures (see Remark 7.11).

Remark 2.12. When $h_{\text{top}}(f) < +\infty$ and, additionally, the entropy map $\mu \in \mathcal{P}_f(X) \mapsto h_\mu(f)$ is upper semi-continuous (as happens, for instance, when f is expansive), Theorem 5 and [91, Theorem 9.12] imply that $\mathfrak{h}_\mu(f) = h_\mu(f)$ for every $\mu \in \mathcal{P}_f(X)$. Thus, under these assumptions, the map $\mathfrak{h}_f : \mathcal{P}(X) \to \mathbb{R}$ is an extension of the metric entropy (with negative values at $\mathcal{P}(X) \setminus \mathcal{P}_f(X)$).

The previous results pave the way to the description of multifractal analysis for Birkhoff averages, large deviations or ergodic optimization in both hyperbolic and non-hyperbolic contexts. Actually, while upper semi-continuity of entropy and uniqueness of equilibrium states are useful ingredients to provide a full description of the entropy map, the dimension of the level sets associated to Birkhoff averages and the maximizing probability measures, these properties

may fail beyond the realm of uniform hyperbolicity. We will give a simple illustration through an application in ergodic optimization (see Subsection 6.2 for definitions and the concept of zero temperature limits). Taking into account Theorem 5 and [54, Theorem 4.1] (replacing the usual entropy function by the upper semi-continuous map $\mu \mapsto \mathfrak{h}_\mu(f)$) we obtain the following information.

Corollary 6. *Let $f : X \to X$ be a continuous map on a compact metric space X such that $h_{\mathrm{top}}(f) < +\infty$ and let $\varphi \colon X \to \mathbb{R}$ be a continuous observable. Then*

$$\frac{1}{t} P_{\mathrm{top}}(f, t\varphi) = \max_{\nu \in \mathcal{P}_f(X)} \left\{ \frac{1}{t} \mathfrak{h}_\nu(f) + \int \varphi \, d\nu \right\} \quad \underset{t \to +\infty}{\longrightarrow} \quad \max_{\nu \in \mathcal{P}_f(X)} \int \varphi \, d\nu.$$

Moreover, if for each $t > 0$ one chooses an f-invariant probability measure $\nu_t \in \mathcal{P}_f(X)$ satisfying $P_{\mathrm{top}}(f, t\varphi) = \mathfrak{h}_{\nu_t}(f) + \int t\,\varphi \, d\nu_t$ then any weak* accumulation point $\nu_\infty \in \mathcal{P}_f(X)$ of $(\nu_t)_{t>0}$ as $t \to +\infty$ satisfies:

(a) $\int \varphi \, d\nu_\infty = \max\limits_{\mu \in \mathcal{P}_f(X)} \int \varphi \, d\mu$

(b) $\mathfrak{h}_{\nu_\infty}(f) = \max\limits_{\nu \in \mathcal{M}_f(X, \varphi)} \mathfrak{h}_\nu(f) \geqslant \max\limits_{\nu \in \mathcal{M}_f(X, \varphi)} h_\nu(f)$

(c) $\lim\limits_{t \to +\infty} \mathfrak{h}_{\nu_t}(f) = \mathfrak{h}_{\nu_\infty}(f)$

where $\mathcal{M}_f(X, \varphi)$ stands for the set of f-invariant φ-maximizing probability measures obtained through zero-temperature limits.

2.4. **Non-additive thermodynamic formalism.** The previous Convex Analysis conclusions also apply to the thermodynamic formalism and the ergodic optimization for sub-additive sequences of continuous potentials, which include as special cases the singular value potentials associated to linear cocycles, a topic that goes beyond the scope of hyperbolic dynamical systems (see Section 8 for definitions and a discussion on previous results). This is so in spite of the fact that the Legendre-Fenchel duality is not available in this context, since the space of sub-additive sequences of continuous functions is not a vector space. As far as we know, this approach brings forth novelty even in the case when f is the one-sided full shift.

In what follows, given a cocycle $A \in C(X, GL(\ell, \mathbb{R}))$ over a continuous map $f : X \to X$, we denote by $P(f, A, \Phi_{\vec{\alpha}})$ the topological pressure of its singular value potential $\Phi_{\vec{\alpha}}$ (see Definition 8.1, equation (8.9) and Subsection 8.2.2 for the precise definitions).

Theorem 7. *Let f be a continuous self-map of a compact metric space (X, d). There is an upper semi-continuous map $\mathfrak{h}_f : \mathcal{P}_a(X) \to \mathbb{R}$ such that, for every cocycle $A \in C(X, GL(\ell, \mathbb{R}))$, every vector $\vec{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \in \mathbb{R}^\ell$ with $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_\ell$ and the corresponding non-additive sequence $\Phi_{\vec{\alpha}}$ of singular value potentials, there exists a map $\psi_{\Phi_{\vec{\alpha}}} \in B_m(X)$ such that*

$$P(f, A, \Phi_{\vec{\alpha}}) = \max_{\mu \in \mathcal{P}_a(X)} \left\{ \mathfrak{h}_\mu(f) + \int \psi_{\Phi_{\vec{\alpha}}} \, d\mu \right\}$$

and

$$\forall \, \mu \in \mathcal{P}_f(X) \qquad \psi_{\Phi_{\vec{\alpha}}}(x) = \sum_{i=1}^{k} \alpha_i \cdot \lambda_i(A, x) \qquad \text{at } \mu\text{-almost every } x \in X.$$

Moreover, the set of finitely additive equilibrium states is non-empty for every linear cocycle in $C(X, GL(\ell, \mathbb{R}))$ and the zero temperature limits of finitely additive equilibrium states have the largest value of $\mathfrak{h}_f$ amongst the Lyapunov optimizing measures.

3. A variational principle: Proof of Theorem 1

We begin this section by recalling some duality results on the Banach spaces we will consider. If X is locally compact, it follows from the Riesz-Markov representation Theorem that the dual of $C(X)$ can be identified with the collection of all finite signed measures on $(X, \mathfrak{B})$ equipped with the weak* topology, and that the subset of its positive normalized elements corresponds to the space $\mathcal{P}(X)$ (cf. [57, pp. 253]). This is clearly the case if X is a compact metric space, where $C(X) = C_c(X) = C_b(X)$. In the non-compact setting, the dual of $C_b(X)$ is identified with the space of regular Borel and finitely additive measures endowed with the total variation norm (cf. Theorem 14.9 in [1]). An extension of the Riesz-Markov Theorem informs that the dual of $B_m(X)$ is represented by the space of Borel finitely additive measures with the topology induced by the total variation norm (see [46, 50]), whose subset of positive normalized elements corresponds to $\mathcal{P}_a(X)$.

For the sake of completeness and rigor, we include a proof of the first part of Theorem 1 along the lines of that provided by Föllmer and Schied (see [47]). Let (X, d) be a metric space, $\mathbf{B} = B_m(X)$ and $\Gamma : \mathbf{B} \to \mathbb{R}$ be a pressure function. Define

$$\mathcal{A}_\Gamma = \big\{ \varphi \in \mathbf{B} : \Gamma(-\varphi) \leqslant 0 \big\} \qquad \text{and} \qquad \mathfrak{h}(\mu) := \inf_{\varphi \in \mathcal{A}_\Gamma} \left\{ \int \varphi \, d\mu \right\}.$$

We start showing that, for every $\varphi \in \mathbf{B}$, one has

$$\Gamma(\varphi) \;\geqslant\; \sup_{\mu \in \mathcal{P}_a(X)} \left\{ \mathfrak{h}(\mu) + \int \varphi \, d\mu \right\}. \tag{3.1}$$

Due to translation invariance it is clear that $\Gamma(\varphi - \Gamma(\varphi)) = 0$. Thus $\widetilde{\varphi} := \Gamma(\varphi) - \varphi$ belongs to $\mathcal{A}_\Gamma$. Therefore, for every $\mu \in \mathcal{P}_a(X)$

$$\mathfrak{h}(\mu) \;\leqslant\; \int \widetilde{\varphi} \, d\mu = \Gamma(\varphi) - \int \varphi \, d\mu$$

which implies (3.1). Conversely, given $\varphi \in \mathbf{B}$ we need to find $\mu_\varphi \in \mathcal{P}_a(X)$ such that

$$\Gamma(\varphi) \;\leqslant\; \mathfrak{h}(\mu_\varphi) + \int \varphi \, d\mu_\varphi. \tag{3.2}$$

Yet, it is enough to do it for φ such that $\Gamma(\varphi) = 0$ since the general case follows from taking $\varphi - \Gamma(\varphi)$ and the translation invariance property of Γ. So, consider $\varphi \in \mathbf{B}$ such that $\Gamma(\varphi) = 0$. Thus the observable $-\varphi$ does not belong to the set

$$\mathcal{B}_\Gamma \;=\; \{\psi \in \mathbf{B} : \Gamma(-\psi) < 0\}$$

which is convex and open due to the convexity and continuity of Γ. Therefore, by the geometric version of the Hahn-Banach Theorem there is a continuous, not identically zero, linear functional $L : \mathbf{B} \to \mathbb{R}$ which separates the sets $\{-\varphi\}$ and $\mathcal{B}_\Gamma$ in the sense that

$$L(-\varphi) \;\leqslant\; \inf_{\psi \in \mathcal{B}_\Gamma} L(\psi). \tag{3.3}$$

By linearity of L, this is equivalent to saying that

$$L(\varphi) + \inf_{\psi \in \mathcal{B}_\Gamma} L(\psi) \;\geqslant\; 0. \tag{3.4}$$

Lemma 3.5. L *is positive and* $L(1) > 0$, *where* 1 *stands for the constant function equal to one.*

Proof. Consider $\psi \in \mathbf{B}$ such that $\psi \geqslant 0$. Firstly, let us see that for every $\lambda > 0$ we have $\lambda\psi + c \in \mathcal{B}_\Gamma$, where $c = \Gamma(0) + 1$. Indeed, by translation invariance and monotonicity of Γ one has

$$\Gamma(-\lambda\psi - c) = \Gamma(-\lambda\psi) - c \leqslant \Gamma(0) - c < 0.$$

Due to (3.3), this in turn implies that

$$L(-\varphi) \leqslant L(\lambda\psi + c) = \lambda L(\psi) + cL(1) \qquad \forall\, \lambda > 0.$$

Thus, if $L(\psi) < 0$ then $L(-\varphi) = -\infty$, leading to a contradiction with the fact that L is a bounded functional. This proves that L is a positive functional. In particular, $L(1) \geqslant 0$.

Let us now prove that $L(1) \neq 0$. As L is linear and not identically zero, we may take $\psi_0 \in \mathbf{B}$ such that $L(\psi_0) > 0$ and $\|\psi_0\|_\infty < 1$. Write $\psi_0 = \psi_0^+ - \psi_0^-$, where $\psi_0^+ = \max\{\psi_0, 0\}$ and $\psi_0^- = \max\{-\psi_0, 0\}$. Then, as L is positive, we have

$$L(\psi_0^+) = L(\psi_0) + L(\psi_0^-) \geqslant L(\psi_0) > 0 \qquad \text{and} \qquad L(1 - \psi_0^+) \geqslant 0$$

since $1 - \psi_0^+ \geqslant 0$. Using again both the linearity and the monotonicity of L we finally conclude that

$$L(1) = L(1 - \psi_0^+) + L(\psi_0^+) \geqslant L(\psi_0^+) > 0.$$

$\square$

The previous lemma indicates that the continuous linear operator $\frac{L}{L(1)}$ is positive and normalized. Therefore, according to an extension of the Riesz-Markov Representation Theorem [41], there is a finitely additive probability measure $\mu_\varphi \in \mathcal{P}_a(X)$ (which belongs to $\mathcal{P}(X)$ if $\mathbf{B} = C_c(X)$ and X is locally compact) such that

$$\int \psi\, d\mu_\varphi = \frac{L(\psi)}{L(1)} \qquad \forall\, \psi \in \mathbf{B}.$$

We are left to show that μ_φ satisfies (3.2). Observe that for every $\psi \in \mathcal{A}_\Gamma$ and every $\varepsilon > 0$ we have that $\psi + \varepsilon \in \mathcal{B}_\Gamma$. In other words, $\mathcal{A}_{\Gamma,\varepsilon} := \{\psi + \varepsilon : \psi \in \mathcal{A}_\Gamma\} \subset \mathcal{B}_\Gamma$. Hence

$$\mathfrak{h}(\mu_\varphi) = \inf_{\psi \in \mathcal{A}_\Gamma} \int \psi\, d\mu_\varphi \leqslant \inf_{\psi \in \mathcal{B}_\Gamma} \int \psi\, d\mu_\varphi \leqslant \inf_{\psi \in \mathcal{A}_{\Gamma,\varepsilon}} \int \psi\, d\mu_\varphi = \inf_{\psi \in \mathcal{A}_\Gamma} \int \psi\, d\mu_\varphi + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary we conclude that $\mathfrak{h}(\mu_\varphi) = \inf_{\psi \in \mathcal{B}_\Gamma} \int \psi\, d\mu_\varphi$. Consequently,

$$\mathfrak{h}(\mu_\varphi) + \int \varphi\, d\mu_\varphi = \left(\inf_{\psi \in \mathcal{B}_\Gamma} \int \psi\, d\mu_\varphi\right) + \frac{L(\varphi)}{L(1)} = \left(\inf_{\psi \in \mathcal{B}_\Gamma} \frac{L(\psi)}{L(1)}\right) + \frac{L(\varphi)}{L(1)}$$

$$= \frac{1}{L(1)} \left(L(\varphi) + \inf_{\psi \in \mathcal{B}_\Gamma} L(\psi)\right) \geqslant 0 = \Gamma(\varphi)$$

where we have used relation (3.4) in the last step. This completes the proof of (3.2).

Endowing $\mathcal{P}_a(X)$ with the total variation distance, the function $\mathfrak{h}$ is upper semi-continuous since it is defined as the infimum of the family of continuous functions $\left(\mu \in \mathcal{P}_a(X) \mapsto \int \varphi\, d\mu\right)_{\varphi \in \mathcal{A}_\Gamma}$ (cf. [4, 18]).

We proceed by showing the maximality (hence uniqueness) of the function $\mathfrak{h}$ among those which satisfy (2.3). Let α be such a map. Then,

$$\Gamma(-\psi) \geqslant \alpha(\mu) + \int -\psi\, d\mu \qquad \forall\, \psi \in \mathbf{B} \quad \forall\, \mu \in \mathcal{P}_a(X)$$

or, equivalently,

$$\alpha(\mu) \;\leqslant\; \Gamma(-\psi) + \int \psi \, d\mu \qquad \forall\, \psi \in \mathbf{B} \quad \forall\, \mu \in \mathcal{P}_a(X)$$

which implies that

$$\alpha(\mu) \;\leqslant\; \inf_{\psi \in \mathbf{B}} \left\{ \Gamma(-\psi) + \int \psi \, d\mu \right\} \qquad \forall\, \mu \in \mathcal{P}_a(X). \tag{3.6}$$

Moreover, as $\mathcal{A}_\Gamma \subset \mathbf{B}$ and $\Gamma(-\psi) \leqslant 0$ for every $\psi \in \mathcal{A}_\Gamma$ we conclude that

$$
\begin{aligned}
\alpha(\mu) \;&\leqslant\; \inf_{\psi \in \mathbf{B}} \left\{ \Gamma(-\psi) + \int \psi \, d\mu \right\} \;\leqslant\; \inf_{\psi \in \mathcal{A}_\Gamma} \left\{ \Gamma(-\psi) + \int \psi \, d\mu \right\} \\
&\leqslant\; \inf_{\psi \in \mathcal{A}_\Gamma} \left\{ \int \psi \, d\mu \right\} = \mathfrak{h}(\mu) \qquad \forall\, \mu \in \mathcal{P}_a(X).
\end{aligned}
$$

The previous reasoning using $\alpha = \mathfrak{h}$ allows us to conclude that

$$\mathfrak{h}(\mu) \;=\; \inf_{\psi \in \mathbf{B}} \left\{ \Gamma(-\psi) + \int \psi \, d\mu \right\} \qquad \forall\, \mu \in \mathcal{P}_a(X). \tag{3.7}$$

As $\mathbf{B}$ is a vector space the equality (3.7) can be rewritten as

$$\mathfrak{h}(\mu) \;=\; \inf_{\psi \in \mathbf{B}} \left\{ \Gamma(-\psi) - \int -\psi \, d\mu \right\} = \inf_{\psi \in \mathbf{B}} \left\{ \Gamma(\psi) - \int \psi \, d\mu \right\}.$$

We are left to prove that $\mathfrak{h}$ is affine. The convex set $\mathcal{P}_a(X)$ is compact in the weak* topology (cf. [41, Theorem 2, V.4.2]) and so, by the Krein-Milman Theorem, it is the closed convex hull of its extreme points. Moreover, we can use the the Choquet Representation Theorem (cf. [3, Theorem 6.6] or [91, p.153]) to express each member of $\mathcal{P}_a(X)$ in terms of the extreme elements of $\mathcal{P}_a(X)$. More precisely, if $E_a(X)$ denotes the set of extreme points of $\mathcal{P}_a(X)$ and μ belongs to $\mathcal{P}_a(X)$ then there is a unique measure $\mathbb{P}_\mu$ on the Borel subsets of the compact metrizable space $\mathcal{P}_a(X)$ such that $\mathbb{P}_\mu(E_a(X)) = 1$ and

$$\int_X \psi(x) \, d\mu(x) = \int_{E_a(X)} \left(\int_X \psi(x) \, dm(x) \right) d\mathbb{P}_\mu(m) \qquad \forall\, \psi \in B_m(X). \tag{3.8}$$

Hence every $\mu \in \mathcal{P}_a(X)$ is a generalized convex combination of extreme finitely additive probability measures. We write $\mu = \int_{E_a(X)} m \, d\mathbb{P}_\mu(m)$ and call this equality the decomposition in extremes of the finitely additive probability measure μ.

Lemma 3.9. *Given $\mu \in \mathcal{P}_a(X)$ whose decomposition in extremes is $\mu = \int_{E_a(X)} m \, d\mathbb{P}_\mu(m)$, then*

$$\mathfrak{h}(\mu) = \int \mathfrak{h}(m) \, d\mathbb{P}_\mu(m).$$

In particular, the function $\mathfrak{h}$ is affine.

Proof. Recall that $\mathfrak{h}(\mu) = \inf_{\psi \in \mathcal{A}_\Gamma} \int \psi \, d\mu$ and the map $\psi \in B_m(X) \mapsto \int \psi \, d\mu$ is continuous. By (3.8) there exists a probability measure $\mathbb{P}_\mu$ giving full weight to the space $E_a(X)$ of extreme measures of $\mathcal{P}_a(X)$ and satisfying $\int \psi \, d\mu = \int_{E_a(X)} \left(\int \psi \, dm \right) d\mathbb{P}_\mu(m)$. In particular,

$$\mathfrak{h}(\mu) = \inf_{\psi \in \mathcal{A}_\Gamma} \int_{E_a(X)} \left(\int \psi(x) \, dm(x) \right) d\mathbb{P}_\mu(m).$$

Moreover, as $\mathbb{P}_\mu$ is countably additive, the Monotone Convergence Theorem (cf. [77, Theorem IV.15, Vol. I]) for nets of continuous maps when applied to the net $(\int \psi \, d\mu)_{\psi \in \mathcal{A}_\Gamma}$ implies that

$$
\begin{aligned}
\mathfrak{h}(\mu) &= \inf_{\psi \in \mathcal{A}_\Gamma} \int_{E_a(X)} \left(\int \psi(x) \, dm(x) \right) d\mathbb{P}_\mu(m) \\
&= \int_{E_a(X)} \left(\inf_{\psi \in \mathcal{A}_\Gamma} \int \psi(x) \, d\, m(x) \right) d\mathbb{P}_\mu(m) = \int_{E_a(X)} \mathfrak{h}(m) \, d\mathbb{P}_\mu(m).
\end{aligned}
$$

Finally, by Choquet Theorem, given $\mu_1, \mu_2 \in \mathcal{P}_a(X)$ there exist unique probability measures $\mathbb{P}_{\mu_1}$, $\mathbb{P}_{\mu_2}$ giving full weight to $E_a(X)$ and such that $\mu_i = \int_{E_a(X)} m \, d\mathbb{P}_{\mu_i}(m)$ for $i = 1, 2$. In particular, for each $0 < \alpha < 1$, one has

$$
\alpha\mu_1 + (1-\alpha)\mu_2 = \int_{E_a(X)} m \, d \, [\alpha \, d\mathbb{P}_{\mu_1} + (1-\alpha) \, d\mathbb{P}_{\mu_2}](m)
$$

and, using the first part of the lemma, one gets

$$
\mathfrak{h}(\alpha\mu_1 + (1-\alpha)\mu_2) = \int_{E_a(X)} \mathfrak{h}(m) d[\alpha\mathbb{P}_{\mu_1} + (1-\alpha)\mathbb{P}_{\mu_2}](m) = \alpha\mathfrak{h}(\mu_1) + (1-\alpha)\mathfrak{h}(\mu_2).
$$

This ends the proof of Theorem 1. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

4. Tangent functionals: Proof of Theorem 2

The argument we will present follows closely the one in [91, Theorems 9.14 and 9.15]. Consider $\varphi \in \mathbf{B}$ and assume that $\mu \in \mathcal{E}_\varphi(\Gamma)$. Then, by Theorem 1,

$$
\Gamma(\varphi + \psi) - \Gamma(\varphi) \geqslant \mathfrak{h}(\mu) + \int (\varphi + \psi) \, d\mu - \mathfrak{h}(\mu) - \int \varphi \, d\mu = \int \psi \, d\mu \qquad \forall \psi \in \mathbf{B}.
$$

This shows that $\mathcal{E}_\varphi(\Gamma) \subseteq \mathcal{T}_\varphi(\Gamma)$. To establish the converse inclusion, fix $\mu \in \mathcal{T}_\varphi(\Gamma)$ and note that

$$
\Gamma(\varphi + \psi) - \Gamma(\varphi) \quad \geqslant \quad \int \psi \, d\mu \qquad \forall \psi \in \mathbf{B}
$$

$$
\Updownarrow
$$

$$
\Gamma(\varphi + \psi) - \int (\varphi + \psi) \, d\mu \quad \geqslant \quad \Gamma(\varphi) - \int \varphi \, d\mu \qquad \forall \psi \in \mathbf{B}.
$$

This equivalence together with the variational principle in Theorem 1, the fact that $\mathbf{B}$ is a vector space and equation (3.7) imply that

$$
\mathfrak{h}(\mu) = \inf_{\psi \in \mathbf{B}} \left\{ \Gamma(\varphi + \psi) - \int (\varphi + \psi) \, d\mu \right\} \geqslant \Gamma(\varphi) - \int \varphi \, d\mu.
$$

Since the reverse inequality

$$
\mathfrak{h}(\mu) \leqslant \Gamma(\varphi) - \int \varphi \, d\mu
$$

is an immediate consequence of (2.3), we conclude that $\mathcal{T}_\varphi(\Gamma) \subseteq \mathcal{E}_\varphi(\Gamma)$. The second claim in the statement of Theorem 2 is a consequence of [62] (see also [69, page 12]), which ensures that the convex function Γ, acting on the separable Banach space $\mathbf{B} = C_b(X)$ or $\mathbf{B} = C_c(X)$, admits a unique tangent functional for every φ in a residual subset of $\mathbf{B}$.

5. Fréchet differentiability: Proof of Theorem 3

In this section we will show the characterization of Fréchet differentiability of the pressure functional in terms of tangent functionals.

$(a) \Rightarrow (b)$. As Γ is locally affine at φ, there exist (a unique) $\mu_\varphi \in \mathcal{P}_a(X)$ (respectively, a signed probability measure if X is locally compact and $\mathbf{B} = C_c(X)$) such that, for every $\psi_1, \psi_2 \in \mathbf{B}$ whose norms are small enough, one has

$$\Gamma(\varphi + \psi_1) - \Gamma(\varphi + \psi_2) = \int (\psi_1 - \psi_2)\, \mu_\varphi.$$

This implies that $\mathcal{T}_{\varphi + \psi}(\Gamma) = \{\mu_\varphi\}$ for every $\psi \in \mathbf{B}$ with small enough norm. Thus,

$$\lim_{\psi \to 0} \sup \left\{ \|\mu - \mu_\varphi\| : \ \mu \in \mathcal{T}_{\varphi + \psi}(\Gamma) \right\} = \lim_{\psi \to 0} \{0\} = 0.$$

$(b) \Rightarrow (c)$. Assume that there exists a unique tangent functional $\mu_\varphi \in \mathcal{T}_\varphi(\Gamma)$ and that

$$\lim_{\psi \to 0} \sup \left\{ \|\mu - \mu_\varphi\| : \ \mu \in \mathcal{T}_{\varphi + \psi}(\Gamma) \right\} = 0. \tag{5.1}$$

As $\mu_\varphi \in \mathcal{T}_\varphi(\Gamma)$, one has

$$\Gamma(\varphi + \psi) - \Gamma(\varphi) \geqslant \int \psi \, d\mu_\varphi \qquad \forall\, \psi \in \mathbf{B}.$$

Moreover, the uniqueness of the tangent functional μ_φ at φ and Theorem 2 imply that

$$\Gamma(\varphi) = \mathfrak{h}(\mu_\varphi) + \int \varphi \, d\mu_\varphi \qquad \text{and} \qquad \Gamma(\varphi) > \mathfrak{h}(\mu) + \int \varphi \, d\mu \quad \forall\, \mu \neq \mu_\varphi.$$

So, given $\psi \in \mathbf{B}$ and $\mu \in \mathcal{T}_{\varphi + \psi}(\Gamma) = \mathcal{E}_{\varphi + \psi}(\Gamma)$, one gets

$$0 \leqslant \Gamma(\varphi + \psi) - \Gamma(\varphi) - \int \psi \, d\mu_\varphi = \mathfrak{h}(\mu) + \int (\varphi + \psi)\, d\mu - \Gamma(\varphi) - \int \psi \, d\mu_\varphi$$

$$\leqslant \mathfrak{h}(\mu) + \int (\varphi + \psi)\, d\mu - \mathfrak{h}(\mu) - \int \varphi \, d\mu - \int \psi \, d\mu_\varphi$$

$$= \int \psi \, d\mu - \int \psi \, d\mu_\varphi \ \leqslant \ \|\psi\|_\infty \, \|\mu - \mu_\varphi\|.$$

Therefore, by assumption (5.1), one has

$$\lim_{\psi \to 0} \frac{1}{\|\psi\|_\infty} \left| \Gamma(\varphi + \psi) - \Gamma(\varphi) - \int \psi \, d\mu_\varphi \right| = 0.$$

$(c) \Rightarrow (a)$. The map $\mathfrak{h}$ is affine, but the pressure function Γ is not associated to an underlying dynamics. Therefore, part of the argument to prove Theorem 6 of [92] (the analogue of Theorem 3 for the topological pressure) has to be adapted to the general setting we are dealing with. Assume that Γ is Fréchet differentiable at $\varphi \in \mathbf{B}$. Then, as Γ is convex, there is a unique tangent functional to Γ at φ (cf. [78, Chapt. IV, §44]), say $\mathcal{T}_\varphi(\Gamma) = \{\mu_\varphi\}$, and μ_φ satisfies

$$\lim_{\psi \to 0} \frac{1}{\|\psi\|_\infty} \left| \Gamma(\varphi + \psi) - \Gamma(\varphi) - \int \psi \, d\mu_\varphi \right| = 0.$$

By (2.7), one also has

$$\Gamma(\varphi + \psi) - \Gamma(\varphi) \geqslant \int \psi \, d\mu_\varphi \qquad \forall \psi \in \mathbf{B}.$$

We are left to prove the reverse inequality for ψ inside a neighborhood of 0 in $\mathbf{B}$.

Lemma 5.2. *Let $(\mu_n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{P}_a(X)$ such that $\lim_{n \to +\infty} \mathfrak{h}(\mu_n) + \int \varphi \, d\mu_n = \Gamma(\varphi)$. Then $\lim_{n \to +\infty} \|\mu_n - \mu_\varphi\| = 0$.*

Proof. Given $\varepsilon > 0$, take $\delta > 0$ so that

$$\psi \in \mathbf{B}, \ \|\psi\|_\infty \leqslant \delta \qquad \Rightarrow \qquad 0 \leqslant \Gamma(\varphi + \psi) - \Gamma(\varphi) - \int \psi \, d\mu_\varphi \leqslant \varepsilon \, \|\psi\|_\infty.$$

For each $n \in \mathbb{N}$ consider $\varepsilon_n = \Gamma(\varphi) - \mathfrak{h}(\mu_n) - \int \varphi \, d\mu_n$, and let $N \in \mathbb{N}$ be such that for every $n \geqslant N$ one has $0 \leqslant \varepsilon_n < \varepsilon\delta$. Therefore, if $\|\psi\|_\infty \leqslant \delta$ and $n \geqslant N$,

$$\int \psi \, d\mu_n - \int \psi \, d\mu_\varphi = \Gamma(\varphi) + \int \psi \, d\mu_n - \Gamma(\varphi) - \int \psi \, d\mu_\varphi$$

$$= \mathfrak{h}(\mu_n) + \int \varphi \, d\mu_n + \varepsilon_n + \int \psi \, d\mu_n - \Gamma(\varphi) - \int \psi \, d\mu_\varphi$$

$$= \mathfrak{h}(\mu_n) + \int (\varphi + \psi) \, d\mu_n + \varepsilon_n - \Gamma(\varphi) - \int \psi \, d\mu_\varphi$$

$$\leqslant \Gamma(\varphi + \psi) + \varepsilon_n - \Gamma(\varphi) - \int \psi \, d\mu_\varphi$$

$$= \Gamma(\varphi + \psi) - \Gamma(\varphi) - \int \psi \, d\mu_\varphi + \varepsilon_n < 2\delta\varepsilon.$$

Since these estimates are also valid for $-\psi$, one has $\left| \int \psi \, d\mu_n - \int \psi \, d\mu_\varphi \right| < 2\delta\varepsilon$ for every $\|\psi\|_\infty \leqslant \delta$ and $n \geqslant N$. Thus, if $n \geqslant N$,

$$\|\mu_n - \mu_\varphi\| = \sup \left\{ \left| \int \psi \, d\mu_n - \int \psi \, d\mu_\varphi \right| : \|\psi\|_\infty \leqslant 1 \right\}$$

$$= \frac{1}{\delta} \sup \left\{ \left| \int \psi \, d\mu_n - \int \psi \, d\mu_\varphi \right| : \|\psi\|_\infty \leqslant \delta \right\} < \frac{1}{\delta} \, 2\delta\varepsilon = 2\varepsilon.$$

$\square$

Recall that, by Lemma 3.9, one has $\mathfrak{h}(\mu) = \int \mathfrak{h}(m) \, d\mathbb{P}_\mu(m)$ for every $\mu \in \mathcal{P}_a(X)$ whose decomposition in extreme points is $\mu = \int_{E_a(X)} m \, d\mathbb{P}_\mu(m)$. Thus,

$$\Gamma(\varphi) = \sup \left\{ \mathfrak{h}(\mu) + \int \varphi \, d\mu \ \middle| \ \mu \in E_a(X) \right\}. \tag{5.3}$$

Lemma 5.4. *The tangent functional μ_φ is an extreme point of $\mathcal{P}_a(X)$ and*

$$\Gamma(\varphi) > \sup \left\{ \mathfrak{h}(\mu) + \int \varphi \, d\mu \ \middle| \ \mu \in E_a(X) \ \text{and} \ \mu \neq \mu_\varphi \right\}.$$

Proof. Using (5.3), one can choose a sequence $(\mu_n)_{n \in \mathbb{N}}$ of extreme finitely additive probability measures with

$$\lim_{n \to +\infty} \left(\mathfrak{h}(\mu_n) + \int \varphi \, d\mu_n \right) = \Gamma(\varphi).$$

Then, by Lemma 5.2, one has $\lim_{n \to +\infty} \|\mu_n - \mu_\varphi\| = 0$. Since distinct extreme points in $\mathcal{P}_a(X)$ have norm distance equal to 2 (cf. [3]), there is $N \in \mathbb{N}$ such that $\mu_n = \mu_N$ for every $n \geqslant N$. Therefore, $\mu_N = \mu_\varphi$, so μ_φ is an extreme point of $\mathcal{P}_a(X)$. In addition, observe that, since $\mathcal{T}_\varphi(\Gamma) = \mathcal{E}_\varphi(\Gamma) = \{\mu_\varphi\}$, the previous argument also shows that one cannot have

$$\Gamma(\varphi) = \sup \left\{ \mathfrak{h}(\mu) + \int \varphi \, d\mu \ \mid \ \mu \in E_a(X) \ \text{and} \ \mu \neq \mu_\varphi \right\}.$$

Thus, $\Gamma(\varphi) > \sup \left\{ \mathfrak{h}(\mu) + \int \varphi \, d\mu \ \mid \ \mu \in E_a(X) \ \text{and} \ \mu \neq \mu_\varphi \right\}$.   □

Lemma 5.5. *There is a neighborhood $\mathcal{U}$ of μ_φ in the total variation norm such that*

$$\mathfrak{h}(\mu_\varphi) > \sup \left\{ \mathfrak{h}(\mu) \ \mid \ \mu \in \mathcal{U}, \ \mu \in E_a(X) \ \text{and} \ \mu \neq \mu_\varphi \right\}.$$

Proof. Let $a = \Gamma(\varphi) - \sup \left\{ \mathfrak{h}(\mu) + \int \varphi \, d\mu \ \mid \ \mu \in E_a(X) \ \text{and} \ \mu \neq \mu_\varphi \right\}$ and $\mathcal{U}$ the neighborhood of μ_φ given by

$$\mathcal{U} = \left\{ \mu \in \mathcal{P}_a(X) : \ \left| \int \varphi \, d\mu - \int \varphi \, d\mu_\varphi \right| < a/2 \right\}.$$

If $\mu \in \mathcal{U}$ is an extreme of $\mathcal{P}_a(X)$ and $\mu \neq \mu_\varphi$, then

$$\begin{aligned}
\mathfrak{h}(\mu) \ &\leqslant \ \mathfrak{h}(\mu) + \int \varphi \, d\mu - \int \varphi \, d\mu_\varphi + \frac{a}{2} \ \leqslant \ \Gamma(\varphi) - a - \int \varphi \, d\mu_\varphi + \frac{a}{2} \\
&= \ \Gamma(\varphi) - \int \varphi \, d\mu_\varphi - \frac{a}{2} = \mathfrak{h}(\mu_\varphi) - \frac{a}{2}
\end{aligned}$$

where the last equality is due to the fact that $\mathcal{T}_\varphi(\Gamma) = \mathcal{E}_\varphi(\Gamma) = \{\mu_\varphi\}$.   □

We are finally ready to show that Γ is locally affine at φ. Let a be as in the proof of Lemma 5.5. Then, for every $\psi \in B_m(X)$ satisfying $\|\psi - \varphi\|_\infty < a/2$, one has

$$\sup \left\{ \mathfrak{h}(\mu) + \int \psi \, d\mu \ \mid \ \mu \in E_a(X) \ \text{and} \ \mu \neq \mu_\varphi \right\}$$
$$\leqslant \Gamma(\varphi) - a + \|\psi - \varphi\|_\infty \leqslant \Gamma(\psi) - a + 2 \|\psi - \varphi\|_\infty \ < \ \Gamma(\psi).$$

Thus, by (5.3), all such maps ψ which are $a/2$-close to φ have μ_φ as unique extreme Γ-equilibrium state in $\mathcal{P}_a(X)$. In particular,

$$\|\psi - \varphi\|_\infty < a/2 \qquad \Rightarrow \qquad \Gamma(\psi) = \mathfrak{h}(\mu_\varphi) + \int \psi \, d\mu_\varphi.$$

So Γ is locally affine at φ. This ends the proof of the first part of Theorem 3.

Regarding the second list of equivalent assertions stated in Theorem 3, firstly assume that Γ is everywhere Fréchet differentiable. The previous equivalent conditions imply that Γ is locally affine at every $\varphi \in \mathbf{B}$, so, by the connectedness of the vector space $\mathbf{B}$, we conclude that Γ is affine. Conversely, affine functions are clearly Fréchet differentiable. Hence items $(\bar{a})$ and $(\bar{b})$ are equivalent. Assume now that Γ is affine. Then there exists $\mu \in \mathcal{P}_a(X)$ such that for every $\varphi \in B$ one has

$$\Gamma(\varphi + \psi) - \Gamma(\varphi) = \int \psi \, d\mu \qquad \forall \, \psi \in B.$$

Thus μ is a tangent functional to every $\varphi \in \mathbf{B}$. As any element in $\mathcal{P}_a(X)$ is determined by its integrals over $\mathbf{B}$, the previous equality implies that μ is the unique tangent functional at every

$\varphi \in B$. So, $\bigcup_{\varphi \in \mathbf{B}} \mathcal{T}_\varphi(\Gamma) = \{\mu\}$, and $(\bar{a})$ implies $(\bar{c})$. Finally, condition $(\bar{c})$ implies $(\bar{a})$ due to the corresponding local property $(c) \Rightarrow (a)$. This completes the proof of Theorem 3.

6. Pressure, entropy and zero temperature limits

We start by defining the topological and free energy for a continuous transformation of a compact metric space, along with a recollection of some of their properties.

6.1. Entropy and pressure.
Let $f : X \to X$ be a continuous transformation of a compact metric space (X, d). Given $n \in \mathbb{N}$, define the dynamical distance $d_n \colon X \times X \to [0, +\infty)$ by

$$d_n(x, z) = \max \left\{ d(x, z), d(f(x), f(z)), \ldots, d(f^n(x), f^n(z)) \right\}$$

which generates the same topology as d. For every $x \in X$, $n \in \mathbb{N}$ and $\varepsilon > 0$, we denote by $B_n^f(x, \varepsilon)$ the open ball centered at x with radius ε with respect to the metric d_n. Having fixed $n \in \mathbb{N}$ and $\varepsilon > 0$, we say that a set $E \subset X$ is (n, ε)–separated by f if $d_n(x, z) \geqslant \varepsilon \qquad \forall x \neq z \in E$. Denote by $s_n(f, \varepsilon)$ the maximal cardinality of all (n, ε)–separated subsets of X by f. Due to the compactness of X, the number $s_n(f, \varepsilon)$ is finite. The *topological entropy* of f is defined by

$$h_{\text{top}}(f) = \lim_{\varepsilon \to 0^+} \limsup_{n \to +\infty} \frac{1}{n} \log s_n(f, \varepsilon).$$

More generally, given a continuous map $\varphi : X \to \mathbb{R}$ (also called a potential), the *topological pressure* of f and φ is defined by

$$P_{\text{top}}(f, \varphi) = \lim_{\varepsilon \to 0^+} \limsup_{n \to +\infty} \frac{1}{n} \log P_n(f, \varphi, \varepsilon) \tag{6.1}$$

where, for every $n \in \mathbb{N}$,

$$P_n(f, \varphi, \varepsilon) = \sup_E \left\{ \sum_{x \in E} e^{S_n^f \varphi(x)} : \ E \subset X \text{ is } (n, \varepsilon)\text{-separated} \right\} \tag{6.2}$$

and $S_n^f \varphi(x) = \varphi(x) + \varphi(f(x)) + \cdots + \varphi(f^n(x))$. This way, one assigns to each point $x \in X$ the weight $e^{S_n^f \varphi(x)}$ determined by the potential φ along the block of the first n iterates of f at x. In particular, $P_{\text{top}}(f, 0) = h_{\text{top}}(f)$.

As X is compact, $C(X)$ is a subspace of $B_m(X)$ where $\|.\|_\infty$ is the norm of the uniform convergence. The pressure map

$$P_{\text{top}}(f, .) : \ C(X) \quad \to \quad \mathbb{R} \cup \{+\infty\}$$

satisfies, for every $\varphi, \psi \in C(X)$ and constant $c \in \mathbb{R}$, the following properties [91]:

(1) $h_{\text{top}}(f) + \min \varphi \ \leqslant \ P_{\text{top}}(f, \varphi) \ \leqslant \ h_{\text{top}}(f) + \max \varphi$.

(2) $P_{\text{top}}(f, .)$ is either finite valued or constantly $+\infty$.

(3) If $P_{\text{top}}(f, .) < +\infty$, then $P_{\text{top}}(f, .)$ is convex.

(4) $P_{\text{top}}(f, \varphi + c) = P_{\text{top}}(f, \varphi) + c$.

(5) $\varphi \leqslant \psi \quad \Rightarrow \quad P_{\text{top}}(f, \varphi) \leqslant P_{\text{top}}(f, \psi)$.

(6) $P_{\text{top}}(f, \varphi + \psi \circ f - \psi) = P_{\text{top}}(f, \varphi)$.

(7) If $P_{\text{top}}(f, .) < +\infty$, then $|P_{\text{top}}(f, \varphi) - P_{\text{top}}(f, \psi)| \leqslant \|\varphi - \psi\|_\infty$.

We observe that (6.1) and (6.2) can be used to define a pressure function on the space $B_m(X)$ of bounded potentials which, by some abuse of notation, we still denote by $P_{\text{top}}(f, .) : B_m(X) \to \mathbb{R} \cup \{+\infty\}$. To avoid any confusion, in order to distinguish the pressure functions we will mention their domains.

Denote by $\mathcal{P}_f(X)$ the space of f-invariant Borel probability measures on X endowed with the weak* topology. Given $\mu \in \mathcal{P}_f(X)$ and a continuous potential $\varphi : X \to \mathbb{R}$, the *free energy* of f, μ and φ is given by

$$P_\mu(f, \varphi) = h_\mu(f) + \int \varphi \, d\mu$$

where $h_\mu(f)$ is the *metric entropy* of f with respect to μ (definition and properties may be read in [91, Chapter 4]). A measure $\mu \in \mathcal{P}_f(X)$ is called an *equilibrium state* for the potential φ if $P_\mu(f, \varphi) = \sup_{\nu \in \mathcal{P}_f(X)} \left\{ h_\nu(f) + \int \varphi \, d\nu \right\}$.

6.1.1. **Star-entropy**. The lack of upper semi-continuity of the entropy map has led some authors to regularize the notion of metric entropy. For instance, given $\mu \in \mathcal{P}_f(X)$, the concept of *star-entropy* was introduced in [64] and later explored by Viana and Yang in [89], being defined by

$$h_\mu^*(f) = \sup \left\{ \limsup_{n \to +\infty} h_{\mu_n}(f) \mid (\mu_n)_{n \in \mathbb{N}} \text{ is a sequence in } \mathcal{P}_f(X) \text{ with } \lim_{n \to +\infty} \mu_n = \mu \right\}.$$

It is known that, for every $\mu \in \mathcal{P}_f(X)$, one has

$$h_\mu(f) \ \leqslant \ h_\mu^*(f) \ \leqslant \ h_\mu(f) + h_{loc}(f) \tag{6.3}$$

where $h_{loc}(f)$ stands for the local entropy of f, defined by

$$h_{loc}(f) = \lim_{\varepsilon \to 0^+} \lim_{\delta \to 0^+} \limsup_{n \to +\infty} \frac{1}{n} \sup_{x \in X} \log s_n(f, \delta, B_n^f(x, \varepsilon))$$

and $s_n(f, \delta, B_n^f(x, \varepsilon))$ denotes the maximal cardinality of an (n, δ)-separated subset of $B_n^f(x, \varepsilon)$. The first inequality of (6.3) is a straightforward consequence of the definition of h^*, while the second was proved in [64] (see also [25, Appendix B]), from which we conclude that $h_{loc}(f)$ bounds the defect in upper semi-continuity of the map $\mu \in \mathcal{P}_f(X) \mapsto h_\mu(f)$. The star-entropy function $h^* : \mathcal{P}_f(X) \to \mathbb{R}$ is related to the entropy structures of Boyle and Downarowicz (cf. [21]) since h^* is precisely the upper semi-continuous envelope of the metric entropy. More precisely, when the topological entropy of f is finite the map h^* is bounded from above by $T = $ the topological entropy of f and one has (see [38, p. 466–467])

$$h^* = \inf \left\{ T : \mathcal{P}_f(X) \to \mathbb{R} \mid T \text{ is continuous and } T(\mu) \geqslant h_\mu(f) \ \forall \mu \in \mathcal{P}_f(X) \right\}. \tag{6.4}$$

The advantage of considering the star-entropy is that the function $\mu \in \mathcal{P}_f(X) \mapsto h_\mu^*(f)$ is upper semi-continuous when $\mathcal{P}_f(X)$ is endowed with the weak*-topology. In particular, defining the star-pressure by

$$P_{\text{top}}^*(f, \varphi) = \sup_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu^*(f) + \int \varphi \, d\mu \right\}$$

one guarantees that there always exists an f-invariant probability measure which attains the supremum. Yet, the behavior of the star-entropy function may differ substantially from the usual metric entropy map (we refer the reader to [81] for examples of smooth interval maps without

an invariant probability measure with maximal entropy), and so the corresponding maximum values and maximal measures might fail to describe standard physical quantities.

Denote the topological entropy of f by $h_{\text{top}}(f)$ and assume that it is finite. Let h^{**} be the upper semi-continuous concave envelope of the metric entropy, defined in $\mathcal{P}_f(X)$ by

$$h^{**} = \inf \left\{ T : \mathcal{P}_f(X) \to \mathbb{R} \mid T \text{ is continuous, affine and } T(\mu) \geqslant h_\mu(f) \ \forall \mu \in \mathcal{P}_f(X) \right\}. \quad (6.5)$$

Thus

$$h_\mu(f) \ \leqslant \ h_\mu^*(f) \ \leqslant \ h_\mu^{**}(f) \ \leqslant \ h_{\text{top}}(f) \quad \forall \mu \in \mathcal{P}_f(X).$$

6.2. **Ergodic optimization.** Given $\varphi \in C(X)$, the map $\mu \in \mathcal{P}_f(X) \mapsto \int \varphi \, d\mu$ is continuous and defined on a compact metric space. Hence it has a maximum, which is realized by f-invariant probability measures. These are referred to as φ-*maximizing probability measures*, and may be not unique. It is therefore useful to find φ-maximizing probability measures which are the most chaotic, meaning those which carry larger metric entropy. It is known that, if the metric entropy map is upper semi-continuous, then the φ-maximizing probability measures obtained through zero-temperature limits have this property (cf. [54, Theorem 4.1]). More precisely, on the one hand, if $h_{\text{top}}(f) < +\infty$, then

$$\frac{1}{t} P_{\text{top}}(f, t\varphi) = \sup_{\nu \in \mathcal{P}_f(X)} \left\{ \frac{1}{t} h_\nu(f) + \int \varphi \, d\nu \right\}$$

$$\underset{t \to +\infty}{\longrightarrow} \ \sup_{\nu \in \mathcal{P}_f(X)} \int \varphi \, d\nu = \max_{\nu \in \mathcal{P}_f(X)} \int \varphi \, d\nu.$$

On the other hand, if for each $t > 0$ large enough there exists a unique equilibrium state μ_t for f with respect to $t\varphi$, then any weak* accumulation probability measure $\mu \in \mathcal{P}_f(X)$ of $(\mu_t)_{t>0}$ as t goes to $+\infty$ is a φ-maximizing probability measure.

Two difficulties to carry on this approach for dynamical systems which may not be expansive is that the entropy function may not be upper semi-continuous, and so equilibrium states likely fail to exist at sufficiently small temperatures. We refer the reader to [87] for a class of non-uniformly expanding maps for which the entropy function is not upper semi-continuous, and for which one can only ensure that $t\varphi$ has an equilibrium state for small values of the parameter t. The results in Theorem 5 guarantee a way to bypass these issues using the variational $\mathfrak{h}$-entropy, which still satisfies the variational principle with respect to the classical pressure function.

7. Variational principles: Proof of Theorem 5

In the mid seventies the thermodynamic formalism was brought from statistical mechanics to dynamical systems by the pioneering work of Sinai, Ruelle and Bowen [19], which established a powerful correspondence between one-dimensional lattices and uniformly hyperbolic dynamics and conveyed several notions from one setting to the other. The success of this approach ultimately relies on a variational principle for the topological pressure, along with the construction of equilibrium states as the class of pressure maximizing invariant probability measures. In this section we first show that Theorem 1 extends the classical thermodynamic formalism for continuous maps on compact metric spaces, and then we complete the proof of Theorem 5.

7.1. Classical variational principle. Given a continuous transformation $f : X \to X$ acting on a compact metric space (X, d), the variational principle (cf. [91, §9.3]) states that, given a continuous potential $\varphi : X \to \mathbb{R}$,

$$P_{\text{top}}(f, \varphi) = \sup_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu(f) + \int_X \varphi \, d\mu \right\}.$$

Moreover, the previous least upper bound coincides with the supremum evaluated on the set of ergodic probability measures (cf. [91, Corollary 9.10.1]). A probability measure which attains the maximal value is called an equilibrium state for f and the potential φ. For instance, an equilibrium state for f and the potential $\varphi \equiv 0$ is a measure with maximal entropy.

In the event that $h_{\text{top}}(f) < +\infty$ and the entropy function $\mu \in \mathcal{P}_f(X) \mapsto h_\mu(f)$ is upper semi-continuous, the metric entropy satisfies (see [91, Theorem 9.12])

$$h_\mu(f) = \inf_{\varphi \in C(X)} \left\{ P_{\text{top}}(f, \varphi) - \int \varphi d\mu \right\}$$
$$= \sup \left\{ \limsup_{n \to +\infty} h_{\mu_n} \mid (\mu_n)_{n \in \mathbb{N}} \text{ is a sequence in } \mathcal{P}_f(X) \text{ with } \lim_{n \to +\infty} \mu_n = \mu \right\}.$$

In this case, an easy computation using [91, Theorem 9.12] also yields

$$h_\mu(f) = \inf_{\varphi \in \Lambda} \int \varphi \, d\mu$$

where $\Lambda = \{P_{\text{top}}(f, \varphi) - \varphi \colon \varphi \in C(X)\}$.

7.2. New variational principle. Keeping the classical notion of topological pressure and summoning Theorem 1, we replace the entropy map (metric or star) acting on the space $\mathcal{P}_f(X)$ by a more general real valued function $\mathfrak{h}_f$ whose domain is the space $\mathcal{P}(X)$ of the Borel probability measures on X. More precisely, assume that $h_{\text{top}}(f) < +\infty$; then $P_{\text{top}}(f, .) : C(X) \to \mathbb{R}$ is a pressure function (cf. Section 6) to which we may apply Theorem 1. This way, we conclude that the map $\mathfrak{h}_f : \mathcal{P}(X) \to \mathbb{R}$ given by

$$\mathfrak{h}_\mu(f) = \inf_{\varphi \in \mathcal{A}_{P_{\text{top}}}} \int \varphi \, d\mu$$

where $\mathcal{A}_{P_{\text{top}}} = \{\varphi \in C(X) : P_{\text{top}}(f, -\varphi) \leqslant 0\}$ is upper semi-continuous, satisfies

$$\mathfrak{h}_\mu(f) = \inf_{\varphi \in C(X)} \left\{ P_{\text{top}}(f, \varphi) - \int \varphi \, d\mu \right\} \qquad \forall \, \mu \in \mathcal{P}(X) \tag{7.1}$$

and

$$P_{\text{top}}(f, \varphi) = \max_{\mu \in \mathcal{P}(X)} \left\{ \mathfrak{h}_\mu(f) + \int \varphi \, d\mu \right\} \qquad \forall \, \varphi \in C(X). \tag{7.2}$$

It is immediate from (7.1) that $\mathfrak{h}_\mu(f) \leqslant h_{\text{top}}(f)$ for every $\mu \in \mathcal{P}(X)$. Moreover, using the aforementioned strategy, it is clear that, given $\varphi \in C(X)$, there exists $\mu_\varphi \in \mathcal{P}(X)$ such that $P_{\text{top}}(f, \varphi) = \mathfrak{h}_{\mu_\varphi}(f) + \int \varphi \, d\mu_\varphi$. In the special case of $\varphi \equiv 0$ one gets both the equality

$$h_{\text{top}}(f) = \max_{\mu \in \mathcal{P}(X)} \mathfrak{h}_\mu(f)$$

and $\mu_0 \in \mathcal{P}(X)$ where $\mathfrak{h}_f$ attains its maximum value $h_{\text{top}}(f)$. This ends the proof of Theorem 5.

Corollary 7.3. *Given $\varphi \in C(X)$, every $\mu_\varphi \in \mathcal{P}(X)$ attaining the maximum at (7.2) is f-invariant.*

Proof. Recall that $\mu \in \mathcal{P}(X)$ is said to be f-invariant if $\int (\psi \circ f) \, d\mu = \int \psi \, d\mu$ for every ψ in $C(X)$. Fix $\mu_\varphi \in \mathcal{P}(X)$ such that $P_{\text{top}}(f, \varphi) = \mathfrak{h}_{\mu_\varphi}(f) + \int \varphi \, d\mu_\varphi$, and consider $\psi \in C(X)$. By the variational relation (7.2) applied to both $\varphi + \psi \circ f - \psi$ and $\varphi + \psi - \psi \circ f$ we may take $\mu_1, \mu_2 \in \mathcal{P}(X)$ such that

$$P_{\text{top}}(f, \varphi + \psi \circ f - \psi) = \mathfrak{h}_{\mu_1}(f) + \int \varphi \, d\mu_1 + \int (\psi \circ f) \, d\mu_1 - \int \psi \, d\mu_1$$

and

$$P_{\text{top}}(f, \varphi + \psi - \psi \circ f) = \mathfrak{h}_{\mu_2}(f) + \int \varphi \, d\mu_2 + \int \psi \, d\mu_2 - \int (\psi \circ f) \, d\mu_2.$$

Using the equalities

$$P_{\text{top}}(f, \varphi + \psi \circ f - \psi) = P_{\text{top}}(f, \varphi) = P_{\text{top}}(f, \varphi + \psi - \psi \circ f)$$

(cf. Subsection 6) together with (7.2), we conclude that

$$\begin{aligned} \mathfrak{h}_{\mu_\varphi}(f) + \int \varphi \, d\mu_\varphi &= \mathfrak{h}_{\mu_1}(f) + \int \varphi \, d\mu_1 + \int (\psi \circ f) \, d\mu_1 - \int \psi \, d\mu_1 \\ &\geqslant \mathfrak{h}_{\mu_\varphi}(f) + \int \varphi \, d\mu_\varphi + \int (\psi \circ f) \, d\mu_\varphi - \int \psi \, d\mu_\varphi \end{aligned}$$

so $\int (\psi \circ f) \, d\mu_\varphi - \int \psi \, d\mu_\varphi \leqslant 0$. In a similar way, we deduce that

$$\begin{aligned} \mathfrak{h}_{\mu_\varphi}(f) + \int \varphi \, d\mu_\varphi &= \mathfrak{h}_{\mu_2}(f) + \int \varphi \, d\mu_2 + \int \psi \, d\mu_2 - \int (\psi \circ f) \, d\mu_2 \\ &\geqslant \mathfrak{h}_{\mu_\varphi}(f) + \int \varphi \, d\mu_\varphi + \int \psi \, d\mu_\varphi - \int (\psi \circ f) \, d\mu_\varphi \end{aligned}$$

hence $\int \psi \, d\mu_\varphi - \int (\psi \circ f) \, d\mu_\varphi \leqslant 0$. □

A straightforward consequence one draws from the classical variational principle and the fact that the metric entropy is always non-negative is that the pressure function determines $\mathcal{P}_f(X)$, in the sense that, if $h_{\text{top}}(f) < +\infty$, then (cf. [91, Theorem 9.11])

$$\mu \in \mathcal{P}_f(X) \quad \Leftrightarrow \quad \int \varphi \, d\mu \leqslant P_{\text{top}}(f, \varphi) \quad \forall \varphi \in C(X).$$

Therefore, by (7.1), if $h_{\text{top}}(f) < +\infty$ and μ belongs to $\mathcal{P}(X)$ then

$$\mu \in \mathcal{P}_f(X) \quad \Leftrightarrow \quad \mathfrak{h}_\mu(f) \geqslant 0.$$

7.3. **Linking $\mathfrak{h}$ and h^{**}.**
Here we prove the second part of Theorem 5, relating the h^{**}-entropy (defined by (6.5)) with the $\mathfrak{h}$-entropy function as a consequence of the following result.

Proposition 7.4. *Let $f : X \to X$ be a continuous transformation acting on a compact metric space X with $h_{\text{top}}(f) < +\infty$. Then:*

(a) $P_{\text{top}}(f, \varphi) = \max_{\mu \in \mathcal{P}_f(X)} \left\{ h_f^*(\mu) + \int \varphi \, d\mu \right\} \quad \forall \varphi \in C(X).$

(b) $h_\mu(f) \leqslant h_\mu^*(f) \leqslant h_\mu^{**}(f) = \mathfrak{h}_\mu(f) \quad \forall \mu \in \mathcal{P}_f(X).$

The remainder of this subsection is devoted to the proof of this proposition. We start relating $\mathfrak{h}_\mu(f)$, $h_\mu(f)$ and $h_\mu^*(f)$ when μ belongs to $\mathcal{P}_f(X)$. As previously mentioned, for every $\mu \in \mathcal{P}_f(X)$ one has $h_\mu(f) \leqslant h_\mu^*(f)$. Moreover, by the classical variational principle, for each $\mu \in \mathcal{P}_f(X)$ we get

$$h_\mu(f) \;\leqslant\; P_{\text{top}}(f, \varphi) - \int \varphi \, d\mu \qquad \forall \, \varphi \in C(X).$$

So

$$\mathfrak{h}_\mu(f) = \inf_{\varphi \in C(X)} \left\{ P_{\text{top}}(f, \varphi) - \int \varphi \, d\mu \right\} \;\geqslant\; h_\mu(f). \qquad (7.5)$$

If $h_\mu^*(f) > \mathfrak{h}_\mu(f)$ for some $\mu \in \mathcal{P}_f(X)$, then there exists $\nu \in \mathcal{P}_f(X)$ satisfying $h_\nu(f) > \mathfrak{h}_\nu(f)$, contradicting (7.5). This proves that $h_\mu(f) \leqslant h_\mu^*(f) \leqslant \mathfrak{h}_\mu(f)$ for every $\mu \in \mathcal{P}_f(X)$. These inequalities, together with Theorem 1, yield

$$\sup_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu(f) + \int \varphi \, d\mu \right\} \;\leqslant\; \max_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu^*(f) + \int \varphi \, d\mu \right\}$$

$$\leqslant\; \max_{\mu \in \mathcal{P}_f(X)} \left\{ \mathfrak{h}_\mu(f) + \int \varphi \, d\mu \right\}.$$

On the other hand, as $\mathcal{P}_f(X) \subset \mathcal{P}(X)$, one has for every $\varphi \in C(X)$

$$\max_{\mu \in \mathcal{P}_f(X)} \left\{ \mathfrak{h}_\mu(f) + \int \varphi \, d\mu \right\} \;\leqslant\; \max_{\mu \in \mathcal{P}(X)} \left\{ \mathfrak{h}_\mu(f) + \int \varphi \, d\mu \right\}.$$

Therefore, from both the classical and the new variational principle (7.2) we deduce that

$$\max_{\mu \in \mathcal{P}_f(X)} \left\{ \mathfrak{h}_\mu(f) + \int \varphi \, d\mu \right\} \;\leqslant\; \max_{\mu \in \mathcal{P}(X)} \left\{ \mathfrak{h}_\mu(f) + \int \varphi \, d\mu \right\}$$

$$=\; P_{\text{top}}(f, \varphi) \;=\; \sup_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu(f) + \int \varphi \, d\mu \right\}$$

$$\leqslant\; \max_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu^*(f) + \int \varphi \, d\mu \right\}$$

$$\leqslant\; \max_{\mu \in \mathcal{P}_f(X)} \left\{ \mathfrak{h}_\mu(f) + \int \varphi \, d\mu \right\}.$$

and, consequently,

$$P_{\text{top}}(f, \varphi) \;=\; \max_{\mu \in \mathcal{P}_f(X)} \left\{ \mathfrak{h}_\mu(f) + \int \varphi \, d\mu \right\} \;=\; \max_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu^*(f) + \int \varphi \, d\mu \right\}. \qquad (7.6)$$

Notice that the supremum is attained since $\mathfrak{h}_f$ is upper semi-continuous on $\mathcal{P}_f(X)$, so the map $\mu \in \mathcal{P}_f(X) \mapsto \mathfrak{h}_\mu(f) + \int \varphi \, d\mu$ is upper semi-continuous as well; similar comment regarding h^*. This proves item (a) of the proposition.

Remark 7.7. The previous argument also works if we replace h^* by h^{**}. Thus

$$P_{\text{top}}(f, \varphi) \;=\; \max_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu^{**}(f) + \int \varphi \, d\mu \right\}. \qquad (7.8)$$

We proceed by showing that $h^{**} = \mathfrak{h}$ in $\mathcal{P}_f(X)$. According to [39, A.2.1], the map h^{**} is smaller than any other upper semi-continuous concave map that also upper bounds the metric entropy. Therefore, one has $h^{**} \leqslant \mathfrak{h}$, since $\mathfrak{h}$ is upper semi-continuous, affine (hence concave) and satisfies (7.5). So, regarding item (b) of the proposition, we are left to show that $h^{**} \geqslant \mathfrak{h}$ in $\mathcal{P}_f(X)$.

To prove this inequality, we just need to adapt the argument in [91, Theorem 9.12] and this way conclude that, for every $\mu \in \mathcal{P}_f(X)$,

$$h_\mu^{**}(f) \;\geqslant\; \inf_{\varphi \in C(X)} \left\{ P_{\text{top}}(f, \varphi) - \int \varphi \, d\mu \right\}$$

which, due to (7.1), yields $h_\mu^{**}(f) \geqslant \mathfrak{h}_\mu(f)$.

Take $\mu_0 \in \mathcal{P}_f(X)$ and, as $h_{\text{top}}(f) < +\infty$, consider a real number $b > h_{\mu_0}^{**}(f)$. Let C be the set

$$C = \left\{ (\mu, t) \in \mathcal{P}_f(X) \times \mathbb{R} \colon \; 0 \leqslant t \leqslant h_\mu^{**}(f) \right\}.$$

Since h^{**} is concave, C is a convex set: if (μ, t) and (ν, s) belong to C and $0 \leqslant p \leqslant 1$, then $p\mu + (1-p)\nu \in \mathcal{P}_f(X)$ and

$$0 \leqslant pt + (1-p)s \;\leqslant\; ph_\mu^{**}(f) + (1-p)h_\nu^{**}(f) \;\leqslant\; h_{p\mu+(1-p)\nu}^{**}(f).$$

Moreover, if we consider C as a subset of the dual $C(X)^*$ with the weak* topology, then (μ_0, b) does not belong to the closure $\overline{C}$ of C. Indeed, if there was a sequence $\big((\mu_n, t_n)\big)_n \in C$ converging to (μ_0, b), then, since h^{**} is upper semi-continuous, one would get

$$b = \lim_{n \to +\infty} t_n \;\leqslant\; \limsup_{n \to +\infty} h_{\mu_n}^{**}(f) \;\leqslant\; h_{\mu_0}^{**}(f)$$

contradicting the choice of b. Therefore, there is a continuous functional $F : C(X)^* \times \mathbb{R} \to \mathbb{R}$ separating the disjoint convex sets $\overline{C}$ and $\{(\mu_0, b)\}$, that is, such that

$$F((\mu, t)) \;<\; F((\mu_0, b)) \qquad \forall\, (\mu, t) \in \overline{C}. \tag{7.9}$$

Since we are using the weak* topology on $C(X)^*$, there are $\psi \in C(X)$ and $\gamma \in \mathbb{R}$ such that $F((\mu, t)) = \int \psi \, d\mu + t\gamma$ for every $(\mu, t) \in C(X)^* \times \mathbb{R}$. Thus (7.9) becomes

$$\int \psi \, d\mu + t\gamma \;<\; \int \psi \, d\mu_0 + b\gamma \qquad \forall\, (\mu, t) \in \overline{C}.$$

In particular,

$$\int \psi \, d\mu + h_\mu^{**}(f)\gamma \;<\; \int \psi \, d\mu_0 + b\gamma \qquad \forall\, \mu \in \mathcal{P}_f(X)$$

so, when $\mu = \mu_0$, one obtains

$$\int \psi \, d\mu_0 + h_{\mu_0}^{**}(f)\gamma \;<\; \int \psi \, d\mu_0 + b\gamma.$$

Therefore $\gamma > 0$, and so

$$\int \frac{\psi}{\gamma} \, d\mu + h_\mu^{**}(f) \;<\; \int \frac{\psi}{\gamma} \, d\mu_0 + b \qquad \forall\, \mu \in \mathcal{P}_f(X).$$

Consequently, by the variational principle (7.8),

$$P_{\text{top}}\big(f, \frac{\psi}{\gamma}\big) \;\leqslant\; \int \frac{\psi}{\gamma} \, d\mu_0 + b.$$

Hence,

$$b \geqslant P_{\text{top}}\left(f, \frac{\psi}{\gamma}\right) - \int \frac{\psi}{\gamma} \, d\mu_0 \geqslant \inf_{\varphi \in C(X)} \left\{ P_{\text{top}}(f, \varphi) - \int \varphi \, d\mu_0 \right\}$$

which implies that

$$h_{\mu_0}^{**}(f) \geqslant \inf_{\varphi \in C(X)} \left\{ P_{\text{top}}(f, \varphi) - \int \varphi \, d\mu_0 \right\}.$$

The proof of the proposition is complete.

Remark 7.10. If $\mu_0 \in \mathcal{P}_f(X)$ is an equilibrium state for f and φ with respect to the classical variational principle, that is,

$$P_{\text{top}}(f, \varphi) = h_f(\mu_0) + \int \varphi \, d\mu_0$$

then μ_0 attains the maximum at (7.2) as well. Indeed,

$$P_{\text{top}}(f, \varphi) \geqslant \mathfrak{h}_{\mu_0}(f) + \int \varphi \, d\mu_0 \geqslant h_f(\mu_0) + \int \varphi \, d\mu_0 = P_{\text{top}}(f, \varphi).$$

The converse is false, though, as the example in Subsection 7.4 illustrates.

Remark 7.11. When the metric entropy map is not upper semi-continuous, the map $\mathfrak{h}_f = h^{**}$ is a strict upper bound for the metric entropy. Thus, Theorem 5 motivates the search for an optimal upper semi-continuous upper bound. Clearly, Theorem 1 applied to the pressure function $P_{\text{top}}(f, .) \colon B_m(X) \to \mathbb{R} \cup \{+\infty\}$ provides in general a better bound than $\mathfrak{h}_f$, since it guarantees that

$$P_{\text{top}}(f, \varphi) = \max_{\mu \in \mathcal{P}_a(X)} \left\{ \mathfrak{h}_\infty(\mu) + \int \varphi \, d\mu \right\} \qquad \forall \, \varphi \in B_m(X)$$

where the upper semi-continuous map $\mathfrak{h}_\infty$ is defined by

$$\mathfrak{h}_\infty(\mu) = \inf_{\varphi \in B_m(X)} \left\{ P_{\text{top}}(f, \varphi) - \int \varphi \, d\mu \right\} \qquad \forall \, \mu \in \mathcal{P}_a(X)$$

and so $h_\mu(f) \leqslant \mathfrak{h}_\infty(\mu) \leqslant \mathfrak{h}_\mu(f)$ for every $\mu \in \mathcal{P}(X)$.

7.4. **An example without a measure with maximal entropy.** Given $\varphi \in C(X)$, denote by $\mathcal{P}_\varphi(f, X) \subset \mathcal{P}_f(X)$ the space of (classical) equilibrium states for f and φ. Both $\mathcal{P}_\varphi(f, X)$ and $\mathcal{T}_\varphi(P_{\text{top}})$ are convex sets, but whereas $\mathcal{T}_\varphi(P_{\text{top}})$ is always non-empty and compact for the weak* topology, this is sometimes not true for $\mathcal{P}_\varphi(f, X)$, as we will now check. In general, one has $\mathcal{P}_\varphi(f, X) \subset \mathcal{T}_\varphi(P_{\text{top}})$, with equality if and only if the metric entropy map is upper semi-continuous at every element of $\mathcal{T}_\varphi(P_{\text{top}})$ (cf. [91]). As stated in Theorem 2, it is the set $\mathcal{E}_\varphi(\Gamma)$ of the f-invariant probability measures which maximize the map $\Gamma(\varphi) = \max_{\mu \in \mathcal{P}_f(X)} \left\{ \mathfrak{h}_\mu(f) + \int \varphi \, d\mu \right\}$ that fills in the gap between $\mathcal{P}_\varphi(f, X)$ and $\mathcal{T}_\varphi(P_{\text{top}})$.

Let us briefly recall the example given on [91, p. 193] of a homeomorphism without a measure with maximal entropy. We start describing the β-shift. Let $\beta > 1$ be given and take the expansion of 1 in powers of β^{-1}, that is, $1 = \sum_{n=1}^{+\infty} a_n \beta^{-n}$ where

$$a_1 = [\beta] \qquad \text{and} \qquad a_n = \left[\beta^n - \sum_{j=1}^{n-1} a_j \beta^{n-j}\right] \qquad \forall \, n \geqslant 2.$$

Then $0 \leqslant a_n \leqslant k-1$ for all $n \in \mathbb{N}$, where $k = [\beta]+1$. So we can consider $a = (a_n)_{n \in \mathbb{N}}$ as a point in the space $\Sigma_k^+ = \prod_{i=1}^{+\infty} \{0, 1, \cdots, k-1\}$, within which we define the lexicographical ordering, that is, $(x_n)_{n \in \mathbb{N}} < (y_n)_{n \in \mathbb{N}}$ if $x_j < y_j$ for the smallest j with $x_j \neq y_j$. Let $\sigma_+ : \Sigma_k^+ \to \Sigma_k^+$ be the one-sided shift transformation. Note that $\sigma_+^n(a) \leqslant a$ for every $n \in \mathbb{N}_0$. Define

$$Y_\beta = \Big\{ x = (x_n)_{n \in \mathbb{N}} \in \Sigma_k^+ : \quad \sigma_+^n(x) \leqslant a \quad \forall\, n \in \mathbb{N}_0 \Big\}.$$

This is a non-empty closed subset of Σ_k^+, and one has $\sigma_+(Y_\beta) = Y_\beta$ and $h_{\mathrm{top}}(\sigma_{+|Y_\beta}) = \log \beta$. Besides, if $\Sigma_k = \prod_{i=-\infty}^{+\infty} \{0, 1, \cdots, k-1\}$ and

$$X_\beta = \Big\{ x = (x_n)_{n \in \mathbb{Z}} \in \Sigma_k : \quad (x_i, x_{i+1}, \cdots) \in Y_\beta \quad \forall\, i \in \mathbb{Z} \Big\}$$

then X_β is closed in Σ_k, invariant under the two-sided shift σ and $h_{\mathrm{top}}(\sigma_{|X_\beta}) = \log \beta$ as well.

Now choose an increasing sequence $(\beta_n)_{n \in \mathbb{N}}$ such that $1 < \beta_n < 2$ and $\lim_{n \to +\infty} \beta_n = 2$. Let $f_n : X_{\beta_n} \to X_{\beta_n}$ denote the two-sided β_n-shift and consider on Σ_k a metric d_n inducing the product topology and satisfying $d_n(x, y) \leqslant 1$ for every $x, y \in \Sigma_k$. Define a new space X as the disjoint union of all the spaces X_{β_n} together with a compactification point x_∞, and put on X the metric

$$\rho(x, y) = \begin{cases} \frac{1}{n^2}\, d_n(x, y), & \text{if } x, y \in X_{\beta_n} \\[2mm] \sum_{j=n}^{p} \frac{1}{j^2}, & \text{if } x \in X_{\beta_n},\ y \in X_{\beta_p} \text{ and } n < p \\[2mm] \sum_{j=n}^{+\infty} \frac{1}{j^2}, & \text{if } x = x_\infty \text{ and } y \in X_{\beta_n}. \end{cases}$$

Then (X, ρ) is a compact metric space and the sequence of subsets $\big(X_{\beta_n}\big)_{n \in \mathbb{N}}$ converges to x_∞, that is, the sequence

$$n \in \mathbb{N} \quad \mapsto \quad \tau_n = \inf \big\{ \rho(z, x_\infty) : \ z \in X_{\beta_n} \big\}$$

converges to 0. Moreover, the map $f : X \to X$ defined as $f_{|X_{\beta_n}} = f_n$ and $f(x_\infty) = x_\infty$ is a homeomorphism of (X, ρ); and the Borel f-invariant probability measures are given by

$$\sum_{n=1}^{+\infty} p_n\, \mu_n + \Big(1 - \sum_{n=1}^{+\infty} p_n\Big) \delta_{x_\infty}$$

where $\mu_n \in \mathcal{P}_{f_n}(X_{\beta_n})$ for every $n \in \mathbb{N}$, and the numbers p_n are non-negative and satisfy $\sum_{n=1}^{+\infty} p_n \leqslant 1$. Hence the ergodic elements of $\mathcal{P}_f(X)$ are either ergodic measures in $\mathcal{P}_{f_n}(X_{\beta_n})$ for some n or δ_{x_∞}. Therefore, if $\mathcal{E}_f(X)$ stands for the subset of ergodic measures in $\mathcal{P}_f(X)$, then

$$\begin{aligned} h_{\mathrm{top}}(f) &= \sup \big\{ h_\mu(f) : \mu \in \mathcal{E}_f(X) \big\} = \sup_{n \in \mathbb{N}} \sup \big\{ h_{\mu_n}(f_n) : \mu_n \in \mathcal{E}_{f_n}(X_{\beta_n}) \big\} \\ &= \sup_{n \in \mathbb{N}} h_{\mathrm{top}}(f_n) = \lim_{n \to +\infty} \log \beta_n = \log 2. \end{aligned}$$

Now, if f had a maximal entropy measure, then there should exist an ergodic maximal entropy measure μ. Thus μ would belong to $\mathcal{E}_{f_n}(X_{\beta_n})$ for some n, and so $h_\mu(f) = \log \beta_n < \log 2$.

Let us look instead for a maximizing probability measure of h^*.

Lemma 7.12. *Let $\varepsilon > 0$ be given and, for each $n \in \mathbb{N}$, consider $\mu_n \in \mathcal{P}_{f_n}(X_{\beta_n})$ such that $h_{\mathrm{top}}(f_n) = h_\mu^*(f_n)$. Then any accumulation point of $(\mu_n)_{n \in \mathbb{N}}$ in the weak* topology is δ_{x_∞}.*

Proof. Take $\psi \in C(X)$. Our aim is to show that $\lim_{n \to +\infty} \int \psi \, d\mu_n = \psi(x_\infty)$. As ψ is continuous on the compact X, the subsets $(X_{\beta_n})_{n \in \mathbb{N}}$ are pairwise disjoint and converge to x_∞ with respect to the metric ρ, then the sequence of continuous maps $(\psi_n = \psi_{|X_{\beta_n}})_{n \in \mathbb{N}}$ converges uniformly to $\psi(x_\infty)$. Therefore,

$$\left| \int \psi \, d\mu_n - \psi(x_\infty) \right| = \left| \int (\psi_n - \psi(x_\infty)) \, d\mu_n \right| \leqslant \|\psi_n - \psi(x_\infty)\|_\infty \overset{n \to +\infty}{\longrightarrow} 0.$$

$\square$

From Lemma 7.12 and the upper semi-continuity of h^* we also conclude that

$$h^*_{\delta_{x_\infty}}(f) \geqslant \limsup_{n \to +\infty} h^*_{\mu_n}(f) = \limsup_{n \to +\infty} h^*_{\mu_n}(f_n) = \log 2.$$

Since by definition $h^*_{\delta_{x_\infty}}(f) \leqslant \mathfrak{h}_{\delta_{x_\infty}}(f) \leqslant h_{\text{top}}(f) = \log 2$, the measure δ_{x_∞} maximizes both h^* and $\mathfrak{h}$. On the contrary, $h_{\delta_{x_\infty}}(f) = 0$. In particular, as $h_{\delta_{x_\infty}}(f) < \mathfrak{h}_{\delta_{x_\infty}}(f)$, the measure δ_{x_∞} belongs to $\mathscr{D}$ (cf. (2.11)).

7.5. **Pressure derived from Ruelle-Perron-Frobenius transfer operators.** Some of the statistical properties of equilibrium states are often proved using transfer operators, and the topological pressure arises as the logarithm of the spectral radius of such an operator (cf. [56]). However, the spectral theory of these operators is more powerful when the transfer operator preserves the spaces of Hölder continuous or bounded variation potentials. In what follows, we recall some of these concepts and show that Theorem 1 also imparts a new insight in the thermodynamic formalism of piecewise continuous maps.

Let $f : X \to X$ be a piecewise continuous map on a metric space (X, d) and assume that $\kappa := \sup_{x \in X} \#f^{-1}(x) < +\infty$. Then, given a potential $\varphi \in B_m(X)$, the *Ruelle-Perron-Frobenius transfer operator with weight* φ is (well) defined by

$$
\begin{aligned}
\mathcal{L}_\varphi : \quad & B_m(X) && \to && B_m(X) \\
& \psi && \mapsto && \mathcal{L}_\varphi(\psi) : \; x \in X \mapsto \sum_{f(y) = x} e^{\varphi(y)} \, \psi(y).
\end{aligned}
\tag{7.13}
$$

Denote by $r(\mathcal{L}_\varphi)$ the spectral radius of $\mathcal{L}_\varphi$ which, according to Gelfand's formula (cf. [42]), may be computed by $r(\mathcal{L}_\varphi) = \lim_{n \to +\infty} \sqrt[n]{\|\mathcal{L}_\varphi^n\|}$.

Lemma 7.14. *The function $P : B_m(X) \to \mathbb{R}$ given by $P(\varphi) = \log r(\mathcal{L}_\varphi)$ is a pressure function.*

Proof. Fix $\varphi \in B_m(X)$. Since the space $B_m(X)$ is endowed with the supremum norm and $\mathcal{L}_\varphi$ is a positive operator, for every $n \in \mathbb{N}$ one has

$$\|\mathcal{L}_\varphi^n\| = \sup_{\|\psi\|_\infty = 1} \|\mathcal{L}_\varphi^n(\psi)\|_\infty = \|\mathcal{L}_\varphi^n(\mathbf{1})\|_\infty.$$

So $r(\mathcal{L}_\varphi) = \lim_{n \to +\infty} \sqrt[n]{\|\mathcal{L}_\varphi^n(\mathbf{1})\|_\infty}$, which is bounded by $\kappa \, e^{\|\varphi\|_\infty}$. Now, given $a \in [0, 1]$, $\varphi_1, \varphi_2 \in B_m(X)$ and $n \in \mathbb{N}$, we write

$$\mathcal{L}_{a\varphi_1 + (1-a)\varphi_2}^n(\mathbf{1})(x) = \sum_{f^n(y) = x} e^{S_n(a\varphi_1 + (1-a)\varphi_2)(y)} = \sum_{f^n(y) = x} \left(e^{S_n \varphi_1(y)}\right)^a \left(e^{S_n \varphi_2(y)}\right)^{1-a}$$

and apply Hölder's inequality to get

$$\|\mathcal{L}_{a\varphi_1 + (1-a)\varphi_2}^n(\mathbf{1})\|_\infty \leqslant \|\mathcal{L}_{\varphi_1}^n(\mathbf{1})\|_\infty^a \|\mathcal{L}_{\varphi_2}^n(\mathbf{1})\|_\infty^{1-a}.$$

Taking logarithm, dividing by n and letting n go to $+\infty$, we obtain

$$\log\left(r(\mathcal{L}_{a\,\varphi_1+(1-a)\,\varphi_2})\right) \leqslant a\,\log\left(r(\mathcal{L}_{\varphi_1})\right) + (1-a)\,\log\left(r(\mathcal{L}_{\varphi_2})\right)$$

thereby showing the convexity of the function P. The monotonicity follows from the positivity of the operator $\mathcal{L}_\varphi$ and the proof of the translation invariance is immediate. $\square$

Consequently, Theorem 1 yields the following variational principle.

Corollary 7.15. *Let $f : X \to X$ be a piecewise continuous map on a metric space X such that $\kappa := \sup_{x \in X} \#f^{-1}(x) < +\infty$. Given $\varphi \in B = B_m(X)$, there exists an upper semi-continuous map $\mathfrak{h}_B \colon \mathcal{P}_a(X) \to \mathbb{R}$ such that*

$$\log r(\mathcal{L}_\varphi) = \max_{\mu \in \mathcal{P}_a(X)} \left\{ \mathfrak{h}_B(\mu) + \int \varphi\,d\mu \right\}$$

In particular, there is $\mu_\varphi \in \mathcal{P}_a(X)$ satisfying Rohlin's-like formula (cf. [79])

$$\mathfrak{h}_B(\mu_\varphi) = \int \log\left(r(\mathcal{L}_\varphi)\,e^{-\varphi}\right) d\mu_\varphi.$$

We illustrate this result with the following class of examples.

Example 7.16. Consider a C^1-piecewise expanding map $f : X \to X$ whose domain is the union of a finite number of subintervals $X_i = [a_i, b_i)$ or $X_i = (a_i, b_i]$, where $a_i < b_i$, within which f is continuous. Let $\varphi = -\log|f'|$, which we assume to be piecewise continuous and bounded, though it may not exhibit any further regularity. The corresponding transfer operator is given by

$$\psi \in B_m(X) \quad \mapsto \quad \mathcal{L}_\varphi(\psi)(x) = \sum_{f(y)=x} \frac{1}{|f'(y)|}\,\psi(y)$$

and Corollary 7.15 ensures that there exists $\mu_\varphi \in \mathcal{P}_a(X)$ such that

$$\mathfrak{h}_B(\mu_\varphi) + \int \log|f'|\,d\mu_\varphi = \log r(\mathcal{L}_\varphi).$$

For instance, the Lorenz maps satisfy the previous assumptions with $X = [-1, 0) \cup (0, 1]$.

We observe that, while dealing with a transfer operator acting on a suitable Banach space $\mathcal{X}$ and exhibiting a spectral gap, Giulietti et al [48, Theorem F] showed a variational principle similar to the one in Corollary 7.15 with an entropy-like function $\mathbf{h}_\mathcal{X}$ computed by

$$\mathbf{h}_\mathcal{X}(\mu) = \inf_{\phi \in \mathcal{X}} \left\{ \log \lambda_\mathcal{X}(\phi) - \int \phi\,d\mu \right\}$$

for every f-invariant probability measure μ, where $\lambda_\mathcal{X}(\phi)$ denotes the spectral radius of the transfer operator $\mathcal{L}_\phi : \mathcal{X} \to \mathcal{X}$. In general, since $\mathcal{X} \subsetneq B_m(X)$ one has $h_\mu(f) \leqslant \mathfrak{h}_B(\mu) \leqslant \mathbf{h}_\mathcal{X}(\mu)$ for every $\mu \in \mathcal{P}_f(X)$. In the special case that f is a Ruelle expanding map on a compact metric space X and $\mathcal{X} = C^\alpha(X)$, $\alpha > 0$, the spectral radius of the operator $\mathcal{L}_\phi$ acting on both spaces $C^\alpha(X)$ and $C(X)$ coincide and the three notions of entropy (with $B = C(X)$) are the same.

7.6. **Finitely additive equilibrium states and second order phase transition.** Consider the Manneville-Pomeau family of maps $f_\alpha : [0,1] \to [0,1]$, $\alpha > 0$, given by

$$f_\alpha(x) = \begin{cases} x\left(1 + 2^\alpha x^\alpha\right) & \text{if } x \in [0, \tfrac{1}{2}[\\ 2x - 1 & \text{if } x \in [\tfrac{1}{2}, 1]. \end{cases} \tag{7.17}$$

It is known that the metric entropy map of each f_α is upper semi-continuous, and that this family exhibits phase transitions with respect to the potentials $\varphi_{\alpha,t} = -t \log |f'_\alpha|$, parameterized by $t \in \mathbb{R}$. For instance:

(MP$_1$) If $\alpha > 0$ and $t \in \,]-\infty, 1[$, there exists a unique equilibrium state $\mu_{\alpha,t}$ for f_α and $\varphi_{\alpha,t}$.

(MP$_2$) If $\alpha > 0$, the map $t \in [1, +\infty[\,\mapsto P_{\text{top}}(-t \log |f'_\alpha|)$ is equal to zero and the Dirac measure δ_0 is an equilibrium state with respect to $\varphi_{\alpha,t}$.

(MP$_3$) If $0 < \alpha < 1$, there exist two equilibrium states with respect to $\varphi_{\alpha,1}$, namely the Dirac δ_0 and an f_α-invariant probability measure $\mu_{\alpha,1}$ which is absolutely continuous with respect to the Lebesgue measure; moreover, the map $t \in \mathbb{R} \mapsto P_{\text{top}}(-t \log |f'_\alpha|)$ is not C^1.

(MP$_4$) If $\alpha \geqslant 1$, there exists a unique equilibrium state for $\varphi_{\alpha,t}$ for any $t \in \mathbb{R}$; moreover, the map $t \in \mathbb{R} \mapsto P_{\text{top}}(-t \log |f'_\alpha|)$ is C^1, but not C^2, and there is an f_α-invariant, σ-finite and infinite measure which is absolutely continuous with respect to the Lebesgue measure.

We refer the reader to [23, 58, 88] for an ample discussion on phase transitions of the Manneville-Pomeau family.

It is worth noticing that the C^1-smoothness of the pressure is compatible with the presence of second-order phase transitions. Yet, in the special case of the Manneville-Pomeau maps and $\alpha \geqslant 1$, the second order phase transitions for the potential $\varphi_{\alpha,t} \in C^\alpha([0,1]) \subset C^0([0,1])$ can be detected as first order phase transitions when one considers the pressure function defined on the space of bounded observables or, equivalently, when one replaces the space of invariant probability measures by the larger set of invariant finitely additive measures. More precisely, let $P(f_\alpha, \cdot) : B_m([0,1]) \to \mathbb{R}$ be the pressure function defined by $P(f_\alpha, \varphi) = \log r(\mathcal{L}_\varphi)$ (recall Lemma 7.14 and the definition of the transfer operator $\mathcal{L}_\varphi$ in (7.13)). Then

$$P(f_\alpha, \varphi) \;=\; \max_{\mu \in \mathcal{P}_a([0,1])} \left\{ \mathfrak{h}(\mu) + \int \varphi \, d\mu \right\} \tag{7.18}$$

where

$$\mathfrak{h}(\mu) = \inf_{\varphi \in B_m([0,1])} \left\{ \log r(\mathcal{L}_\varphi) - \int \varphi \, d\mu \right\}$$

for every finitely additive Borel probability measure μ. As $P(f_\alpha, \log |f'_\alpha|) = \log r(\mathcal{L}_{-\log |f'_\alpha|}) = 0$ and

$$\|\mathcal{L}_\varphi^n\|_\infty \geqslant \|\mathcal{L}_\varphi^n(\mathbf{1})\|_\infty \geqslant (\mathcal{L}_\varphi^n(\mathbf{1}))(0) \geqslant e^{n\varphi(0)}$$

for every $\varphi \in B_m([0,1])$, one concludes that $\mathfrak{h}(\delta_0) \geqslant 0$. Moreover, for every f_α-invariant probability measure μ,

$$\mathfrak{h}(\mu) = \inf_{\varphi \in B_m([0,1])} \left\{ \log r(\mathcal{L}_\varphi) - \int \varphi \, d\mu \right\} \leqslant \inf_{\varphi \in C^0([0,1])} \left\{ \log r(\mathcal{L}_\varphi) - \int \varphi \, d\mu \right\}$$

$$= \inf_{\varphi \in C^0([0,1])} \left\{ P(f_\alpha, \varphi) - \int \varphi \, d\mu \right\} \;=\; h_\mu(f_\alpha)$$

where the last equality is due to the fact that the metric entropy map of f_α is upper semi-continuous, while the last but one is a consequence of the conjugation between f_α and the doubling map, for which such an equality holds (see e.g. [29]). Therefore, $\mathfrak{h}(\delta_0) \leqslant h_\mu(\delta_0) = 0$; hence $\mathfrak{h}(\delta_0) = 0$. In the particular case of $\varphi = -\log|f_\alpha'|$ we get

$$P(f_\alpha, -\log|f_\alpha'|) \; = \; 0 \; = \; \max_{\mu \in \mathcal{P}_a([0,1])} \left\{ \mathfrak{h}(\mu) - \int \log|f_\alpha'| \, d\mu \right\}$$

and the Dirac measure δ_0 attains the maximum.

A second equilibrium state appears while one looks for absolutely continuous finitely additive invariant measures. Consider the Lebesgue measure on $[0,1]$ (which we abbreviate into Leb) and the Banach space $L^\infty(\text{Leb}, [0,1])$ of equivalent classes of the essentially bounded real-valued Lebesgue measurable maps $\varphi : [0,1] \to \mathbb{R}$, under the relation $\varphi \backsim \psi$ if and only if $\varphi = \psi$ at Leb almost everywhere. Endow $L^\infty(\text{Leb}, [0,1])$ with the essential supremum norm. The Hewitt-Yosida representation theorem (cf. [49]) informs that the dual of $L^\infty(\text{Leb}, [0,1])$ is the space of bounded finitely additive measures m on the Borel subsets of $[0,1]$ which are absolutely continuous with respect to Leb (in the sense that $\text{Leb}(A) = 0$ implies $m(A) = 0$), with the norm of total variation. Denote by $\mathcal{P}_{a,\text{Leb}}([0,1])$ the set of normalized elements of $\left(L^\infty(\text{Leb}, [0,1])\right)^*$ and by $\mathcal{P}_{a,\text{Leb}}(f_\alpha) \subset \mathcal{P}_{a,\text{Leb}}([0,1])$ the subset of those elements which are f_α-invariant. According to [33], the space $\mathcal{P}_{a,\text{Leb}}(f_\alpha)$ is non-empty.

Given $\mu \in \left(L^\infty(\text{Leb}, [0,1])\right)^*$, consider the upper semi-continuous entropy function

$$H(\mu) = \begin{cases} \inf_{\varphi \in L^\infty(\text{Leb},[0,1])} \; \left\{ \log r(\widehat{\mathcal{L}_\varphi}) - \int \varphi \, d\mu \right\} & \text{if } \mu \text{ is } f_\alpha\text{-invariant} \\ -\infty & \text{otherwise} \end{cases} \tag{7.19}$$

where $r(\widehat{\mathcal{L}_\varphi})$ is the spectral radius of a transfer operator $\widehat{\mathcal{L}_\varphi}$ obtained as the extension to $L^\infty(\text{Leb}, [0,1])$ of the operator $\mathcal{L}_\varphi$ acting on $B_m([0,1])$ (for more details we refer the reader to [6]). The previous definition allows us to build a pressure function on $L^\infty(\text{Leb}, [0,1])$. Given $\varphi \in L^\infty(\text{Leb}, [0,1])$, define

$$\widehat{P}(f_\alpha, \varphi) \; = \; \sup_{\mu \in \mathcal{P}_{a,\text{Leb}}(f_\alpha)} \left\{ H(\mu) + \int \varphi \, d\mu \right\}. \tag{7.20}$$

By the Legendre-Fenchel duality,

$$\widehat{P}(f_\alpha, \varphi) \; = \; \log r(\widehat{\mathcal{L}_\varphi}).$$

Moreover, $\widehat{P}$ satisfies the following variational principle (cf. [6, Theorem 1])

$$\widehat{P}(f_\alpha, \varphi) \; = \; \max_{\mu \in \mathcal{P}_{a,\text{Leb}}(f_\alpha)} \left\{ H(\mu) + \int \varphi \, d\mu \right\}. \tag{7.21}$$

In addition, as the spectral radius can be evaluated using the constant map $\mathbf{1}$, which belongs to both $B_m([0,1])$ and $L^\infty(\text{Leb}, [0,1])$, one has

$$\log r(\widehat{\mathcal{L}_{-\log|f_\alpha'|}}) \; = \; \log r(\mathcal{L}_{-\log|f_\alpha'|}) \; = \; 0.$$

Consequently,

$$0 \; = \; P(f_\alpha, -\log|f_\alpha'|) \; = \; \widehat{P}(f_\alpha, -\log|f_\alpha'|) \; = \; \max_{\mu \in \mathcal{P}_{a,Leb}(f_\alpha)} \left\{ H(\mu) - \int \log|f_\alpha'| d\mu \right\}.$$

Let m_α be a finitely additive probability measure absolutely continuous with respect to Leb where the previous maximum is attained, that is,

$$H(m_\alpha) \;=\; \int \log |f'_\alpha| \, dm_\alpha.$$

Then

$$
\begin{aligned}
0 \;&=\; \widehat{P}(f_\alpha, -\log|f'_\alpha|) \;=\; H(m_\alpha) - \int \log|f'_\alpha| \, dm_\alpha \\
&=\; \inf_{\varphi \in L^\infty(Leb,[0,1])} \left\{ \log r(\widehat{\mathcal{L}_\varphi}) - \int \varphi \, dm_\alpha \right\} - \int \log|f'_\alpha| \, dm_\alpha \\
&\leqslant\; \inf_{\varphi \in B_m([0,1])} \left\{ \log r(\mathcal{L}_\varphi) - \int \varphi \, dm_\alpha \right\} - \int \log|f'_\alpha| \, dm_\alpha \\
&=\; \mathfrak{h}(m_\alpha) - \int \log|f'_\alpha| \, dm_\alpha \;\leqslant\; P(f_\alpha, -\log|f'_\alpha|) \;=\; 0.
\end{aligned}
$$

Thus, summoning the previous information on m_α and δ_0, and Corollary 4, one deduces that:

Corollary 7.22. *If $(f_\alpha)_{\alpha \geqslant 1}$ is the Manneville-Pomeau family* (7.17), *then $\# \mathcal{T}_{\varphi_{\alpha,1}}(P_{\text{top}}) \geqslant 2$. In particular, for every $\alpha \geqslant 1$ the map $P_{\text{top}} : C^0([0,1]) \to \mathbb{R}$ is Gateaux differentiable, though its extension to $B_m([0,1])$ is not Gateaux differentiable at $\varphi_{\alpha,1}$.*

8. Non-additive sequences of continuous potentials: Proof of Theorem 7

In this section, we fix a continuous endomorphism of a compact metric space X and, instead of the classic pressure function with respect to a given potential $\varphi : X \to \mathbb{R}$ and the sequence of sums $(S_n\varphi)_{n \in \mathbb{N}}$, we consider non-additive sequences of continuous potentials. Although these objects arise naturally in the study of Lyapunov exponents and dimension theory, the non-additive thermodynamic formalism is still barely understood. We refer the reader to [10, 11, 45, 35] for a thorough discussion on these topics.

Let $f : X \to X$ be a continuous map on a compact metric space (X, d). We say that a sequence $\Phi = (\varphi_n)_{n \in \mathbb{N}} \in C(X)^{\mathbb{N}}$ of continuous potentials is

(1) *sub-additive* if $\varphi_{m+n} \leqslant \varphi_m + \varphi_n \circ f^m \quad \forall \, m, \, n \in \mathbb{N}$;

(2) *almost additive* if there exists a uniform constant $C > 0$ such that

$$\varphi_m + \varphi_n \circ f^m - C \;\leqslant\; \varphi_{m+n} \;\leqslant\; \varphi_m + \varphi_n \circ f^m + C \qquad \forall \, m, \, n \in \mathbb{N};$$

(3) *asymptotically additive* if for any $\varepsilon > 0$ there exists $\varphi_\varepsilon \in C(X)$ such that

$$\limsup_{n \to +\infty} \frac{1}{n} \left\| \varphi_n - \sum_{j=0}^{n-1} \varphi_\varepsilon \circ f^j \right\|_\infty < \varepsilon.$$

It is known that every almost additive sequence is asymptotically additive, and that for every asymptotically additive sequence $\Phi = (\varphi_n)_{n \in \mathbb{N}} \in C(X)^{\mathbb{N}}$ there exists $\varphi \in C(X)$ such that

$$\limsup_{n \to +\infty} \frac{1}{n} \left\| \varphi_n - \sum_{j=0}^{n-1} \varphi \circ f^j \right\|_\infty = 0$$

(cf. [35, 44]). Therefore, both the variational principle and the existence of finitely additive equilibrium states established in Theorem 5 admit an immediate generalization to this context

(the modifications necessary to deal with sequences in other Banach spaces are left as an easy exercise to the interested reader.) Henceforth, we will aim at the more general context of sub-additive sequences of continuous potentials.

Definition 8.1. Given a sequence $\Phi = (\varphi_n)_{n \in \mathbb{N}} \subset C(X)$ of continuous potentials, the *non-additive topological pressure* is defined by

$$P(f, \Phi) = \lim_{\varepsilon \to 0^+} \limsup_{n \to +\infty} \frac{1}{n} \log \left(\sup_E \sum_{x \in E} e^{\varphi_n(x)} \right) \tag{8.2}$$

where the supremum is taken over the (n, ε)-separated subsets E of X.

This definition coincides with the usual notion of topological pressure $P_{\text{top}}(f, \varphi)$ when there is $\varphi \in C(X)$ such that $\varphi_n = \sum_{j=0}^{n-1} \varphi \circ f^j$ for every $n \in \mathbb{N}$. It is known (cf. [11]) that every almost additive sequence of continuous potentials which have bounded distortion admits a unique equilibrium state, which is a Gibbs measure. More recently, it was proved in [35] that any almost additive or asymptotically additive sequence of continuous potentials have the same pressure of an additive sequence associated to a continuous potential. However, it is not known whether this potential inherits the distortion properties of the original almost additive sequence. Moreover, for sub-additive sequences of continuous potentials no general construction of equilibrium states is known, though for these sequences Cao, Feng and Huang [28] established the following general variational principle.

Theorem 8.3. [28] *Given a continuous self-map $f : X \to X$ of a compact metric space X, if $\Phi = (\varphi_n)_{n \in \mathbb{N}}$ is a sub-additive sequence of continuous potentials such that $P(f, \Phi) > -\infty$, then*

$$P(f, \Phi) = \sup_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu(f) + \mathcal{F}_*(\Phi, \mu) \right\} \tag{8.4}$$

where, for every f-invariant probability measure μ,

$$\mathcal{F}_*(\Phi, \mu) := \lim_{n \to +\infty} \frac{1}{n} \int \varphi_n \, d\mu.$$

We note that if $\Phi = (\varphi_n)_{n \in \mathbb{N}}$ is a sub-additive sequence of continuous potentials then $\mathcal{F}_*(\Phi, \mu)$ is well defined for every $\mu \in \mathcal{P}_f(X)$. In fact, the sequence of real numbers $(a_n)_{n \in \mathbb{R}}$ given by $a_n = \int \varphi_n \, d\mu$ is sub-additive, hence $\lim_{n \to +\infty} \frac{1}{n} a_n = \inf_{n \in \mathbb{N}} \frac{1}{n} a_n$ by Fekete's Lemma.

8.1. **An alternative variational principle for sub-additive sequences.** One might expect a counterpart of Theorem 1 for a more general context of Banach spaces of sequences of functions. This faces non-trivial difficulties, though. Contrary to what happens within the simpler case of almost additive sequences, albeit providing a convex cone in the space of sequences of potentials, sub-additivity is not preserved under multiplication by negative numbers. This is a major obstruction since the entropy in Theorem 1 is defined using observable maps φ such that $-\varphi$ has non-positive pressure, which makes the Banach space generated by sub-additive sequences not suitable to this approach. To overcome this difficulty we will combine two pressure functions to which we apply Theorem 1, thus reducing the previous problem, concerning sub-additive sequences of functions, to the construction of equilibrium states for a single bounded potential (see Lemma 8.5).

In order to apply Kingman's Sub-additive Ergodic Theorem, we need to narrow our analysis to the set

$$\mathcal{S}_b = \left\{ \Phi = (\varphi_n)_{n \in \mathbb{N}} \in C(X)^{\mathbb{N}} \colon \Phi \text{ is sub-additive and } \inf_{x \in X} \left[\inf_{n \in \mathbb{N}} \frac{1}{n} \varphi_n(x) \right] > -\infty \right\}.$$

This set comprises relevant families of sequences of continuous potentials arising within the theory of linear cocycles, as we will detail on Subsection 8.2. We also observe that $\Phi \in \mathcal{S}_b$ if and only if it is sub-additive and $\mathcal{F}_*(\Phi, \mu) > -\infty$ for every f-invariant probability measure μ (cf. [84, pp. 336–337]). Moreover:

Lemma 8.5. *Given $\Phi = (\varphi_n)_{n \in \mathbb{N}} \in \mathcal{S}_b$, the map ψ_Φ defined by*

$$x \in X \quad \mapsto \quad \psi_\Phi(x) = \inf_{n \in \mathbb{N}} \frac{1}{n} \varphi_n(x)$$

is measurable, upper semi-continuous, belongs to $B_m(X)$ and satisfies

$$\int \psi_\Phi \, d\mu = \mathcal{F}_*(\Phi, \mu) \qquad \forall \, \mu \in \mathcal{P}_f(X).$$

Proof. For every $\Phi = (\varphi_n)_{n \in \mathbb{N}} \in \mathcal{S}_b$, the corresponding map ψ_Φ is measurable and upper semi-continuous, hence upper bounded on the compact X. Since Φ belongs to $\mathcal{S}_b$, the map ψ_Φ is also lower bounded. Moreover, by the Kingman's Sub-additive Ergodic Theorem (see [91]), the maps $\inf_{n \in \mathbb{N}} \frac{1}{n} \varphi_n$ and $\liminf_{n \to +\infty} \frac{1}{n} \varphi_n$ coincide in a set with full measure and, for every $\mu \in \mathcal{P}_f(X)$ one has $\int \liminf_{n \to +\infty} \frac{1}{n} \varphi_n \, d\mu = \mathcal{F}_*(\Phi, \mu)$. So, $\int \psi_\Phi \, d\mu = \mathcal{F}_*(\Phi, \mu)$ as well. □

In the remaining of this subsection we restrict to $\mathcal{S}_b$ and consider $P(f, .)$ in order to improve the variational relation (8.4). More precisely, we will show the following counterpart of Theorem 5 in this context, using $B_m(X)$ instead of $C(X)$.

Theorem 8.6. *Let f be a continuous self-map of a compact metric space (X, d) whose topological entropy is finite. Then there exists an affine and upper semi-continuous entropy map $\mathfrak{h}_1 : \mathcal{P}_a(X) \to \mathbb{R}$ such that*

$$P(f, \, \Phi) = \max_{\mu \in \mathcal{P}_a(X)} \left\{ \mathfrak{h}_1(\mu) + \int \psi_\Phi \, d\mu \right\} \qquad \forall \, \Phi \in \mathcal{S}_b.$$

Proof. An effortless computation shows that the map $\Gamma_1 : B_m(X) \to \mathbb{R}$ defined by

$$\psi \in B_m(X) \quad \mapsto \quad \Gamma_1(\psi) = \sup_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu(f) + \int \psi \, d\mu \right\}$$

is a pressure function. Therefore, we may apply Theorem 1 and conclude that there exists an affine and upper semi-continuous map $\mathfrak{h}_1 : \mathcal{P}_a(X) \to \mathbb{R}$ such that, for every $\mu \in \mathcal{P}_a(X)$,

$$\mathfrak{h}_1(\mu) = \inf_{\psi \in \mathcal{A}_{\Gamma_1}} \int \psi \, d\mu$$

and, for each $\psi \in B_m(X)$,

$$\Gamma_1(\psi) = \max_{\mu \in \mathcal{P}_a(X)} \left\{ \mathfrak{h}_1(\mu) + \int \psi \, d\mu \right\}.$$

Besides, $P(f, \Phi) \geqslant \inf_{\mu \in \mathcal{P}_f(X)} \mathcal{F}_*(\Phi, \mu) > -\infty$ for every $(\varphi_n)_{n \in \mathbb{N}} \in \mathcal{S}_b$ (cf. [28]). In addition, from Lemma 8.5 and Theorem 8.3 one deduces that, for every $\Phi = (\varphi_n)_{n \in \mathbb{N}} \in \mathcal{S}_b$,

$$P(f, \Phi) = \Gamma_1(\psi_\Phi).$$

That is,

$$\sup_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu(f) + \mathcal{F}_*(\Phi, \mu) \right\} = \sup_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu(f) + \int \psi_\Phi d\mu \right\}$$

$$= \max_{\mu \in \mathcal{P}_a(X)} \left\{ \mathfrak{h}_1(\mu) + \int \psi_\Phi d\mu \right\}.$$

$\square$

8.2. **Application to linear cocycles and Lyapunov equilibrium states.**

Non-uniform hyperbolicity is defined in terms of *Lyapunov exponents*: a diffeomorphism is non-uniformly hyperbolic if it has no zero Lyapunov exponents and there exists at least one positive and one negative exponent. These numbers measure the exponential asymptotic rates of contraction or expansion along fixed directions, and became a fundamental tool to characterize chaotic dynamics. Linear cocycles turn to be a powerful mean to attest the abundance of non-uniformly hyperbolic behavior, as it allows to detach the underlying dynamics from the action it induces on a vector space. Here we are mainly interested in the existence of Lyapunov equilibrium states for linear cocycles. Some recent contributions on this topic comprise [8, 15, 44, 45].

8.2.1. **Lyapunov exponents.**

We start recalling some preliminary notions. Let f be a continuous map on a compact metric space (X, d). Given an integer $\ell \geqslant 1$, a field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and a measurable matrix-valued map $A : X \to GL(\ell, \mathbb{K})$, the *linear cocycle* generated by A and driven by f is the map

$$\begin{aligned} F_A : X \times \mathbb{K}^\ell &\to X \times \mathbb{K}^\ell \\ (x, v) &\mapsto (f(x), A(x)v). \end{aligned}$$

Its iterates are $F_A^n(x, v) = (f^n(x), A^n(x)v)$, where

$$A^n(x) = A(f^{n-1}(x)) \cdots A(f(x))A(x)$$

for every $n \in \mathbb{N}$, $A^0(x) = x$ and, if f is invertible,

$$A^n(x) = A(f^n(x))^{-1} \cdots A(f^{-1}(x))^{-1}$$

when $n < 0$. We shall also refer to the cocycle as a pair (f, A). A natural example of linear cocycle is given by the *derivative cocycle* associated to a diffeomorphism $f \in \mathrm{Diff}^1(X)$ on a compact Riemannian manifold X, in which case the cocycle is generated by $A(x) = Df(x)$ for each $x \in X$.

Oseledets' Theorem asserts that, under mild conditions, the Lyapunov exponents of the cocycle (f, A) are well defined. More precisely, given an f-invariant probability measure μ, if $\log \|A^\pm\| \in L^1(\mu)$ then for μ-almost every $x \in X$ there exist an integer $k(x) \geqslant 1$, a splitting $\mathbb{K}^\ell = E_x^{1,A} \oplus \cdots \oplus E_x^{k(x),A}$ and real numbers (called *Lyapunov exponents*)

$$\lambda_1(A, \mu, x) > \cdots > \lambda_{k(x)}(A, \mu, x)$$

such that, for every $v \in E_x^{i,A} \setminus \{0\}$ and $1 \leqslant i \leqslant k(x)$,

$$A(x)\left(E_x^{i,A}\right) = E_{f(x)}^{i,A} \qquad \text{and} \qquad \lambda_i(A, \mu, x) = \lim_{n \to +\infty} \frac{1}{n} \log \|A^n(x)v\|.$$

If, in addition, μ is ergodic, then $k(x)$, the Lyapunov exponents $\lambda_i(A, \mu, x)$ and the dimensions of the subspaces $E_x^{i,A}$ are μ-almost everywhere constant, in which case one simplifies the notation by writing $\lambda_i(A, \mu)$.

Remark 8.7. *If X is a compact Riemannian manifold, f is a C^1-diffeomorphism on X and μ is an f-invariant and ergodic probability measure then the Lyapunov exponents of μ are defined as the corresponding Lyapunov exponents for the derivative cocycle $A = Df$.*

8.2.2. **Singular value sub-additive potentials**. In what follows, $\wedge^k L$ stands for the kth exterior power of the linear map L. Assume that the linear cocycle $A : X \to GL(\ell, \mathbb{K})$ is continuous. Then the Lyapunov exponents can be computed using exterior powers and a family of sub-additive sequences of potentials. More precisely, if μ is an f-invariant and ergodic probability measure and one takes for each $k \in \mathbb{N}$ the sub-additive sequence $\Phi_k = (\varphi_{k,n})_{n \in \mathbb{N}}$ of the continuous functions

$$x \in X \quad \mapsto \quad \varphi_{k,n}(x) = \log \| \wedge^k A^n(x) \|$$

then

$$\lim_{n \to +\infty} \frac{1}{n} \varphi_{k,n}(x) = \sum_{i=1}^{k} \lambda_i(A, \mu) \qquad \text{at } \mu\text{-almost every } x \in X.$$

Motivated by applications in dimension theory and aiming to apply their results to Falconer's singular value function and affine iterated function systems with invertible affinities, Bochi and Morris [15] studied the following continuous parameterized family of sub-additive sequences of potentials. Given $\vec{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \in \mathbb{R}^\ell$ with $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_\ell$, consider the sequence $\Phi_{\vec{\alpha}} = (\varphi_{\vec{\alpha},n})_{n \in \mathbb{N}}$ defined by

$$\varphi_{\vec{\alpha},n}(x) = \log \Big(\prod_{i=1}^{\ell} \mathfrak{s}_i(A^n(x))^{\alpha_i} \Big)$$

where $\mathfrak{s}_i(L)$ denotes the ith singular value of the linear map L. Then it is shown in [15] that if μ is an f-invariant and ergodic probability one has

$$\lim_{n \to +\infty} \frac{1}{n} \varphi_{\vec{\alpha},n}(x) = \sum_{i=1}^{k} \alpha_i \cdot \lambda_i(A, \mu) \qquad \text{at } \mu\text{-almost every } x \in X. \tag{8.8}$$

Now, the variational principle for the previous family of sub-additive sequences, established by Theorem 8.3, says that, if $h_{\text{top}}(f) < +\infty$ and $P(f, A, \Phi_{\vec{\alpha}})$ is the pressure function defined by (8.2) when $\Phi = \Phi_{\vec{\alpha}}$, then

$$P(f, A, \Phi_{\vec{\alpha}}) = \sup_{\mu \in \mathcal{P}_f(X)} \Big\{ h_\mu(f) + \int \sum_{i=1}^{k} \alpha_i \cdot \lambda_i(A, \mu, x) \, d\mu \Big\}. \tag{8.9}$$

Invariant measures attaining the previous equality, so called *Lyapunov equilibrium states*, may be hard to find. In the special context of cocycles over a full shift, the metric entropy function is upper semi-continuous and so these equilibrium states always exist. Moreover, within totally disconnected spaces, it has been shown under great generality that, for typical one-step cocycles and Hölder continuous fiber-bunched cocycles, the previous families of sequences of potentials satisfy a quasi-additivity property, and so they have unique Lyapunov equilibrium states (cf. [45, 66]). More recently, Bochi and Morris (cf. [15, Theorem 1]) proved that, in this setting,

there are finitely many Lyapunov equilibrium states, these have full support, and are unique for potentials at large temperatures.

It is known that the existence of Lyapunov equilibrium states for the family $(\beta \Phi_{\vec{\alpha}})_{\beta > 0}$ carries information on Lyapunov optimizing measures. For example, given Lyapunov equilibrium states μ_β with respect to $\beta \Phi_{\vec{\alpha}}$, $\beta > 0$, any weak* accumulation point of $(\mu_\beta)_{\beta > 0}$ is an ergodic optimizing measure for the potential $\Phi_{\vec{\alpha}}$. Additionally, in the case of 2×2 one-step dominated cocycles, Bochi and Rams [16] proved that Lyapunov optimizing measures always exist and, under an additional strong domination condition, these have zero topological entropy. More information about the behavior of zero temperature limits may be read in [63].

8.3. **Proof of Theorem 7.** Let f be a continuous map on a compact metric space (X, d). For each cocycle $A \in C(X, GL(\ell, \mathbb{R}))$ and vector $\vec{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_\ell) \in \mathbb{R}^\ell$ with $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_\ell$ consider the corresponding non-additive sequence $\Phi_{\vec{\alpha}}$ of singular value potentials. The map

$$\psi_{\Phi_{\vec{\alpha}}}(x) := \inf_{n \in \mathbb{N}} \frac{1}{n} \varphi_{\vec{\alpha}, n}(x), \quad x \in X$$

defined by Lemma 8.5 is a measurable, upper semi-continuous and bounded function on X. Combining equations (8.8) and (8.9) one gets

$$P(f, A, \Phi_{\vec{\alpha}}) = \sup_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu(f) + \int \psi_{\Phi_{\vec{\alpha}}} \, d\mu \right\}.$$

Consider now the function $\mathbf{P}(f, \cdot) \colon B_m(X) \to \mathbb{R}$ defined by

$$\mathbf{P}(f, \psi) = \sup_{\mu \in \mathcal{P}_f(X)} \left\{ h_\mu(f) + \int \psi \, d\mu \right\}.$$

It is immediate to check that this is a pressure function. Thus, applying Theorem 1, one obtains an affine and upper semi-continuous entropy entropy function

$$\mathfrak{h}^{\vec{\alpha}} \colon \mathcal{P}_a(X) \to \mathbb{R} \cup \{-\infty, +\infty\}$$

satisfying

$$P(f, A, \Phi_{\vec{\alpha}}) = \mathbf{P}(f, \psi_{\Phi_{\vec{\alpha}}}) = \max_{\mu \in \mathcal{P}_a(X)} \left\{ \mathfrak{h}^{\vec{\alpha}}_\mu(f) + \int \psi_{\Phi_{\vec{\alpha}}} \, d\mu \right\}. \tag{8.10}$$

The set of finitely additive probability measures attaining the maximum is f-invariant (recall the proof of Corollary 7.3). Moreover, by Theorem 1 there is an upper semi-continuous map $\mathfrak{h}_f \colon \mathcal{P}_a(X) \to \mathbb{R}$ such that, for every non-additive sequence $\Phi_{\vec{\alpha}}$ of singular value potentials, one has

$$P(f, A, \Phi_{\vec{\alpha}}) = \max_{\mu \in \mathcal{P}_a(X)} \left\{ \mathfrak{h}_\mu(f) + \int \psi_{\Phi_{\vec{\alpha}}} \, d\mu \right\}.$$

We are left to show that the set of finitely additive equilibrium states is non-empty for every linear cocycle in $C(X, GL(\ell, \mathbb{R}))$, and that the zero temperature limits of finitely additive equilibrium states have the largest value of $\mathfrak{h}^{\vec{\alpha}}_f$ amongst the Lyapunov optimizing measures. This is a simple consequence of equality (8.10) and the fact that, as

$$\sup_{\mu \in \mathcal{P}_a(X)} \mathfrak{h}^{\vec{\alpha}}_\mu(f) \leqslant h_{\text{top}}(f) < +\infty$$

then one has

$$\frac{1}{t} \mathbf{P}(f, t\Phi_{\vec{\alpha}}) = \max_{\mu \in \mathcal{P}_a(X)} \left\{ \frac{1}{t} \mathfrak{h}^{\vec{\alpha}}_\mu(f) + \int \psi_{\Phi_{\vec{\alpha}}} \, d\mu \right\} \xrightarrow[t \to +\infty]{} \max_{\mu \in \mathcal{P}_a(X)} \int \psi_{\Phi_{\vec{\alpha}}} \, d\mu$$

and

$$\frac{1}{t} P(f, A, t\Phi_{\vec{\alpha}}) = \sup_{\mu \in \mathcal{P}_f(X)} \left\{ \frac{1}{t} h_\mu(f) + \int \psi_{\Phi_{\vec{\alpha}}} \, d\mu \right\}$$

$$\xrightarrow[t \to +\infty]{} \sup_{\mu \in \mathcal{P}_f(X)} \int \psi_{\Phi_{\vec{\alpha}}} \, d\mu = \sup_{\mu \in \mathcal{P}_f(X)} \int \sum_{i=1}^{k} \alpha_i \cdot \lambda_i(A, \mu, x) \, d\mu.$$

Therefore, there exists an f-invariant finitely additive probability measure $\mu_{\Phi_{\vec{\alpha}}} \in \mathcal{P}_a(X)$ such that

$$\int \psi_{\Phi_{\vec{\alpha}}} \, d\mu_{\Phi_{\vec{\alpha}}} = \sup_{\mu \in \mathcal{P}_f(X)} \int \sum_{i=1}^{k} \alpha_i \cdot \lambda_i(A, \mu, x) \, d\mu.$$

References

[1] C. D. Aliprantis and K. C. Border. Infinite Dimensional Analysis. Springer-Verlag Berlin, 3rd edition, 2006. 2.1, 3

[2] A. B. Antonevich, V.I. Bakhtin and A. V. Lebedev. *On t-entropy and variational principle for the spectral radii of transfer and weighted shift operators.* Ergodic Theory Dynam. Systems 31 (2011) 995–1045. 1

[3] L. Asimov and A. Ellis. Convexity Theory and its Applications in Functional Analysis. London Math. Soc. Monographs 16, Academic Press, New York, 1980. 3, 5

[4] R. Baire. Leçons sur les Fonctions Discontinues, professées au Collège de France. Gauthier-Villars, 1905. 3

[5] V. Baladi and M. Demers. *On the measure of maximal entropy for finite horizon Sinai Billiard maps.* J. Amer. Math. Soc. 33:2 (2020) 381–449. 1

[6] V. I. Bakhtin. *On t-entropy and variational principle for the spectral radius of weighted shift operators.* Ergodic Theory Dynam. Systems 30 (2010) 1331–1342. 7.6, 7.6

[7] V. I. Bakhtin and A. V. Lebedev. *Entropy statistic theorem and variational principle for t-entropy are equivalent.* J. Math. Anal. Appl. 474 (2019) 59–71. 1

[8] B. Bárány, A. Käenmäki and I. Morris. *Domination, almost additivity and thermodynamical formalism for planar matrix cocycles.* Israel J. Math. 239 (2020) 173–214. 8.2

[9] E. Barone and K.P.S. Bhaskara Rao. *Poincaré recurrence theorem for finitely additive measures.* Rend. Mat. 7:1:4 (1981) 521–526. 1

[10] L. Barreira. *A non-additive thermodynamic formalism and applications to dimension theory of hyperbolic dynamical systems.* Ergodic Theory Dynam. Systems 16 (1996) 871–927. 8

[11] L. Barreira. *Non-additive thermodynamic formalism: equilibrium and Gibbs measures.* Discrete Contin. Dyn. Syst. 16 (2006) 279–305. 8, 8

[12] K. P. S. Bhaskara and M. Bhaskara. Theory of charges: a study of finitely additive measures. Academic Press, London, 1983. 1

[13] A. Biś, M. Carvalho, M. Mendes and P. Varandas. *Entropy functions for semigroup actions.* ArXiv:2009.07212 1, 2.1

[14] E. Bishop and R. Phelps. *The support functionals of a convex set.* Proc. Sympos. Pure Math. Vol. VII, Amer. Math. Soc., Providence R.I., 1963, pp. 27–35. 1

[15] J. Bochi and I. Morris. *Equilibrium states of generalised singular value potentials and applications to affine iterated function systems.* Geom. Funct. Analysis 28 (2018) 995–1028. 8.2, 8.2.2, 8.2.2

[16] J. Bochi and M. Rams. *The entropy of Lyapunov-optimizing measures of some matrix cocycles.* J. Modern Dynam. 10 (2016) 255–286. 8.2.2

[17] C. Bonatti, L. Díaz and M. Viana. Dynamics Beyond Uniform Hyperbolicity. Encyclopaedia of Mathematical Sciences 102, Springer-Verlag Berlin Heidelberg, 2004. 1

[18] N. Bourbaki. General Topology. Chapters 5-10. Elements of Mathematics, Springer-Verlag Berlin, 1998. 1, 3

[19] R. Bowen. Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms. Lectures Notes in Math. 470, Springer-Verlag Berlin, 1975. 1, 2.6, 7

[20] R. Bowen. *Symbolic dynamics for hyperbolic flows.* Amer. J. Math. 95 (1973) 429–460. 1

[21] M. Boyle and T. Downarowicz. *The entropy theory of symbolic extensions.* Invent. Math. 156 (2004) 119–161. 6.1.1

[22] H. Bruin and M. Todd. *Equilibrium states for potentials with $\sup \phi - \inf \phi < h_{top}(f)$.* Comm. Math. Phys. 283 (2008) 579–611. 1

[23] H. Bruin, D. Terhesiu and M. Todd. *The pressure function for infinite equilibrium states.* Israel J. Math. 232 (2019) 775–826. 7.6

[24] A. Bufetov. *Finitely-additive measures on the asymptotic foliations of a Markov compactum.* Moscow Math. Journal 14:2 (2014) 205–224 1

[25] J. Buzzi. *Intrinsic ergodicity of smooth interval maps.* Israel J. Math. 100 (1997) 125–161. 6.1.1

[26] J. Buzzi. C^r *surface diffeomorphisms with no maximal entropy measure.* Ergodic Theory and Dynam. Systems 34:6 (2014) 1770–1793. 1

[27] J. Buzzi, S. Crovisier and O. Sarig. *Measures of maximal entropy for surface diffeomorphisms.* Ann. of Math. 195:2 (2022) 421–508. 1

[28] Y. Cao, D. Feng and W. Huang. *The thermodynamic formalism for sub-additive potentials.* Discrete Contin. Dyn. Syst. 20 (2008) 639–657. 8, 8.3, 8.1

[29] M. Carvalho, F. B. Rodrigues and P. Varandas. *Semigroup actions of expanding maps.* J. Stat. Phys. 166 (2017) 114–136. 7.6

[30] R. Chen. *A finitely additive version of Kolmogorov's law of iterated logarithm.* Israel J. Math. 23 (1976) 209–220. 1

[31] R. Chen. *Some finitely additive versions of the strong law of large numbers.* Israel J. Math. 24 (1976) 244–259. 1

[32] L. Cioletti, E. Silva and M. Stadlbauer. *Thermodynamic formalism for topological Markov chains on standard Borel spaces.* Discrete Contin. Dyn. Syst. 39:11 (2019) 6277–6298. 1

[33] L. Cioletti, A. van Enter and R. Ruviaro. *The double transpose of the Ruelle operator.* ArXiv:1710.03841v2 1, 7.6

[34] D. Coronel and J. Rivera-Letelier. *Sensitive dependence of geometric Gibbs states at positive temperature.* Comm. Math. Phys. 368 (2019) 383–425. 1

[35] N. Cuneo. *Additive, almost additive and asymptotically additive potential sequences are equivalent.* Comm. Math. Phys. 377 (2020) 2579–2595. 8, 8, 8

[36] M. Cuth, O. Kalenda and P. Kaplicky. *Finitely additive measures and complementability of Lipschitz-free spaces.* Israel J. Math. 230:1 (2019) 409–442. 1

[37] H. Daniëls and A. van Enter. *Differentiability properties of the pressure in lattice systems.* Comm. Math. Phys. 71 (1980) 65–76. 2.2

[38] T. Downarowicz and S. Newhouse. *Symbolic extensions and smooth dynamical systems.* Invent. Math. 160 (2005) 453–499. 1, 6.1.1

[39] T. Downarowicz. Entropy in Dynamical Systems. New Mathematical Monographs 18, Cambridge University Press, Cambridge, 2011. 7.3

[40] L. Dubins. *On Lebesgue-like extensions of finitely additive measures.* Ann. Probab. 2 (1974) 456–463. 1

[41] N. Dunford and J. Schwartz. Linear operators I John Wiley & Sons Inc., 1958. 2.1, 3, 3

[42] N. Dunford and J. Schwartz. Linear operators II. Interscience Publishers Inc., 1963. 7.5

[43] R. Ellis. Entropy, Large Deviations and Statistical Mechanics. Classics in Mathematics. Springer-Verlag Berlin, 2006. 2.1

[44] D.-J. Feng and W. Huang. *Lyapunov spectrum of asymptotically sub-additive potentials.* Comm. Math. Phys. 297 (2010) 1–43. 8, 8.2

[45] D.-J. Feng and A. Käenmäki. *Equilibrium states of the pressure function for products of matrices.* Discrete Cont. Dyn. Sys. 30 (2011) 699–708. 8, 8.2, 8.2.2

[46] G. Fichtenholz and L. Kantorovich. *Sur les opérations linéaires dans l'espace des fonctions bornées.* Studia Math. 5 (1934) 69–98. 3

[47] H. Föllmer and A. Schied. Stochastic Finance, an Introduction in Discrete Time. de Gruyter, Studies in Mathematics 27, Berlin, 2002. 1, 1, 2.1, 2.1, 3

[48] P. Giulietti, B. Kloeckner, A. O. Lopes and D. Marcon. *The calculus of thermodynamical formalism.* J. Eur. Math. Soc. 20 (2018) 2357–2412 1, 7.5

[49] E. Hewitt and K. Yosida. *Finitely additive measures.* Trans. Amer. Math. Soc. 72 (1952) 46–66. 7.6

[50] T. Hildebrandt. *On bounded functional operations.* Trans. Amer. Math. Soc. 36:4 (1934) 868–875. 3

[51] G. Iommi and M. Todd. *Transience in dynamical systems.* Ergodic Theory Dynam. Systems 33:5 (2013) 1450–1476. 1

[52] R. Israel. *Convexity in the Theory of Lattice Gases.* Princeton Series in Physics, Princeton University Press, Princeton N.J., 1979. 1

[53] R. Israel and R. Phelps. *Some convexity questions arising in statistical mechanics.* Math. Scand. 54 (1984) 133–156. 1, 2.1

[54] O. Jenkinson. *Ergodic optimization in dynamical systems.* Ergodic Theory and Dynam. Systems 39:10 (2019) 2593–2618. 2.3, 6.2

[55] R. Karandikar. *A general principle for limit theorems in finitely additive probability.* Trans. Amer. Math. Soc. 273:2 (1982) 541–550. 1

[56] G. Keller. *Equilibrium States in Ergodic Theory.* London Mathematical Society Student Texts, Cambridge University Press, Cambridge, 1998. 7.5

[57] J. Kingman and S. Taylor. Introduction to Measure and Probability. Cambridge University Press, Cambridge, 1966. 3

[58] A. O. Lopes. *The zeta function, nondifferentiability of pressure, and the critical exponent of transition.* Adv. Math. 101 (1993) 133–165. 7.6

[59] A. O. Lopes, J. Mengue, J. Mohr and R. Souza. *Entropy and variational principle for one-dimensional lattice systems with a general a priori probability: positive and zero temperature.* Ergodic Theory and Dynam. Systems 35:6 (2015) 1925–1961. 1

[60] D. Maharan. *Finitely additive measures on the integers.* Sankhya: The Indian Journal of Statistics 38 (1976) 44–59. 1

[61] H. B. Maynard. *A Radon-Nikodym theorem for finitely additive bounded measures.* Pacific J. Math. 83:2 (1979) 401–413. 1

[62] S. Mazur. *Über konvexe Mengen in linearen normierten Räumen.* Studia Math. 4 (1933) 70–84. 4

[63] R. Mohammadpour. *Zero temperature limits of equilibrium states for subadditive potentials and approximation of the maximal Lyapunov exponent.* Topol. Methods Nonlinear Anal. 55:2 (2020) 697–710. 8.2.2

[64] S. Newhouse. *Continuity properties of the entropy.* Ann. of Math. 129 (1989) 215–237. 1, 6.1.1, 6.1.1

[65] O. Nikodym. *A theorem on infinite sequences of finitely additive real valued measures.* Rend. Sem. Mat. Padova 24 (1955) 265–286. 1

[66] K. Park. *Quasi-multiplicativity of typical cocycles.* Comm. Math. Phys. 376:3 (2020) 1957–2004. 8.2.2

[67] W. Parry and M. Pollicott. *Zeta functions and the periodic orbit structure of hyperbolic dynamics.* Astérisque 187–188 (1990). 1

[68] A. Paterson. Amenability. Mathematical Surveys and Monographs, 29, American Mathematical Society, Providence R.I., 1988. 1

[69] R. Phelps. Convex Functions, Monotone Operators and Differentiability. Lecture Notes in Math. 1364, Springer-Verlag New York, 2nd edition, 1993. 4

[70] V. Pinheiro. *Expanding measures.* Ann. Inst. H. Poincaré Anal. Non Linéaire 28 (2011) 889–939. 1

[71] F. Przytycki, J. Rivera-Letelier and S. Smirnov. *Equivalence and topological invariance of conditions for non-uniform hyperbolicity in the iteration of rational maps.* Invent. Math. 151:1 (2003) 29–63. 1

[72] F. Przytycki and J. Rivera-Letelier. *Geometric pressure for multimodal maps of the interval.* Mem. Amer. Math. Soc. 259 (2019) 1246, v+81 pp. 1

[73] R. Purves and W. Sudderth. *Some finitely additive probability.* Ann. Probab. 4:2 (1976) 259–276. 1

[74] S. Ramakrishnan. *Central limit theorems in a finitely additive setting.* Illinois J. Math. 28:1 (1984) 139–161. 1

[75] S. Ramakrishnan. *A finitely additive generalization of Birkhoff's ergodic theorem.* Proc. Amer. Math. Soc. 96:2 (1986) 299–305. 1

[76] M. Ratner. *Markov partitions for Anosov flows on n-dimensional manifolds.* Israel J. Math. 15 (1973) 92–114. 1

[77] M. Reed and B. Simon. Methods of Modern Mathematical Physics: Functional Analysis. Academic Press, London New York, 1981. 3

[78] A. Roberts and D. Varberg. Convex Functions. Academic Press, London New York, 1973. 5

[79] V. A. Rohlin. *Exact endomorphisms of a Lebesgue space.* Izv. Akad. Nauk SSSR Ser. Mat. 25 (1961) 499–530; English translation: Amer. Math. Soc. Transl. 39:2 (1964) 1–36. 7.15

[80] D. Ruelle. Thermodynamic Formalism. The Mathematical Structures of Classical Equilibrium Statistical Mechanics. Addison-Wesley, New York, 1978. 1, 1

[81] S. Ruette. *Mixing C^r maps of the interval without maximal measure.* Israel J. Math. 127 (2002) 253–277. 6.1.1

[82] O. Sarig. *Symbolic dynamics for surface diffeomorphisms with positive entropy.* J. Amer. Math. Soc. 26:2 (2013) 341–426. 1

[83] O. Sarig. *Thermodynamic formalism for countable Markov shifts. Hyperbolic dynamics, fluctuations and large deviations* Proc. Sympos. Pure Math. 89, Amer. Math. Soc., Providence R.I., 2015, 81–117. 1

[84] S. Schreiber. *On growth rates of sub-additive functions for semi-flows.* J. Diff. Equations 148 (1998) 334–350. 8.1

[85] Ya. Sinai. *Gibbs measures in ergodic theory.* Russian Math. Surveys 27 (1972) 21–69. 1

[86] J. Toland. The Dual of $L_\infty(X, \mathcal{L}, \lambda)$, Finitely Additive Measures and Weak Convergence. Springer Briefs in Mathematics, Springer Nature Switzerland AG, 2020. 1

[87] P. Varandas and M. Viana. *Existence, uniqueness and stability of equilibrium states for non-uniformly expanding maps.* Ann. Inst. H. Poincaré Anal. Non Linéaire 27 (2010) 555–593. 6.2

[88] R. Venegeroles. *Thermodynamic phase transitions for Pomeau-Manneville maps.* Physical Review E 86:2 (2012) 021114. 7.6

[89] M. Viana and J. Yang. Oral presentation at *International Conference on Dynamical Systems*, IMPA - Rio de Janeiro, 2015. 6.1.1

[90] S. Wagon. The Banach-Tarski paradox. Encyclopedia of Mathematics and its Applications, vol. 24, Cambridge University Press, Cambridge, 1985. 1

[91] P. Walters. An Introduction to Ergodic Theory. Springer-Verlag New York, 1975. 1, 2.1, 2.12, 3, 4, 6.1, 6.1, 7.1, 7.2, 7.3, 7.4, 8.1

[92] P. Walters. *Differentiability properties of the pressure of a continuous transformation on a compact metric space.* J. Lond. Math. Soc. 46:3 (1992) 471–481. 1, 2.2, 2.2, 2.2, 2.2, 5

[93] D. Zubov. *Finitely additive measures on the unstable leaves of Anosov diffeomorphisms.* Functional Analysis and its Applications 53 (2019) 232–236. 1

Faculty of Mathematics, Łódź University, Poland
E-mail address: andrzej.bis@wmii.uni.lodz.pl

CMUP & Faculdade de Ciências da Universidade do Porto, Portugal.
E-mail address: mpcarval@fc.up.pt

CMUP & Faculdade de Engenharia da Universidade do Porto, Portugal.
E-mail address: migmendx@fe.up.pt

CMUP & Departamento de Matemática, Universidade Federal da Bahia, Salvador, Brazil.
E-mail address: paulo.varandas@ufba.br