PhD

3rd CYCLE

FCUP
FCTUC
2025

Contributions to the theory of metric mean dimension

Gustavo Sperotto Pessil

Contributions to the theory of metric mean dimension

Gustavo Sperotto Pessil
UC|UP Joint PhD Program in Mathematics
Faculty of Sciences of the University of Porto and Faculty of Science and Technology of the University of Coimbra
2025

Contributions to the theory of metric mean dimension

Gustavo Sperotto Pessil

Thesis carried out as part of the UC|UP Joint PhD Program in Mathematics

Department of Mathematics of the University of Porto

2025

Supervisors
Paulo Varandas, University of Aveiro
Maria Carvalho, University of Porto

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

1 2 9 0
UNIVERSIDADE D
COIMBRA

Contributions to the theory of metric mean dimension

Gustavo Sperotto Pessil

UC|UP Joint PhD Program in Mathematics

Programa Interuniversitário de Doutoramento em Matemática

PhD Thesis | Tese de Doutoramento

September 2025

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my advisors, Paulo Varandas and Maria Carvalho, for welcoming me in Porto and for their unwavering support throughout this journey. I greatly appreciate their trust and enthusiasm towards my ideas, which has made me grow as a more confident researcher.

I thank my friends in the department for the stimulating discussions, the joys of football matches and the mutual support that kept me motivated. Special thanks to Etubi, João and Alessio for the incredible conference trips we shared.

To my family, João, Carla and Laura, thank you for always believing in me and for making me feel that I could venture into the world, knowing I had a safe place back home.

To Julia, my beloved partner and greatest supporter, thank you for always being there for me, for collecting my broken pieces after a long day and for reminding me there's life beyond the PhD.

To all my friends from Porto Alto, who came into my life from different corners of the world, thank you for the countless amazing moments we've had together. Special thanks to Ivan and Lucas, who became my brothers of heart.

I am grateful to Fundação para a Ciência e a Tecnologia (FCT) for funding my scholarship (UI/BD/152212/2021) end to Centro de Matemática da Universidade do Porto (CMUP) for funding my conference trips.

Abstract

Metric mean dimension is a geometric invariant of dynamical systems with infinite topological entropy. In recent years, it has received increasing attention due to its intrinsic connections with the embbeding problem, coding and information theory, geometric analysis and thermodynamic formalism.

Firstly, we provide two notions of measure-theoretic metric mean dimension - which also depend on the metric structure of the space - both satisfying a classical variational principle, and study their properties. We also obtain a variational principle in terms of local metric mean dimension, and make use of our results to obtain closed formulas for the upper metric mean dimension with potential of two classes of systems.

Secondly, we investigate the metric mean dimension of subshifts of compact type. We prove that the metric mean dimensions of a continuous map and its inverse limit coincide, generalizing Bowen's entropy formula. Building upon this result, we extend the notion of metric mean dimension to discontinuous maps in terms of suitable subshifts. As an application, we show that the metric mean dimension of the Gauss map and that of induced maps of the Manneville-Pomeau family is equal to the box dimension of the corresponding set of discontinuity points, which also coincides with a critical parameter of the pressure operator associated to the geometric potential.

Finally, we relate the metric mean dimension with the fractal structure of the phase space and the Hölder regularity of the map. Afterwards we improve our general estimates in a family of interval maps by computing the metric mean dimension in a way similar to the Misiurewicz formula for the entropy, which in particular shows that our bounds are sharp. As an application, we determine the metric mean dimension of the classical Weierstrass functions. Of independent interest, we develop a dynamical analogue of Minkowski-Bouligand dimension for maps acting on Ahlfors regular spaces.

Resumo

A *dimensão métrica média* é um invariante geométrico de sistemas dinâmicos com entropia topológica infinita. Nos últimos anos, este conceito tem recebido atenção crescente devido às suas conexões intrínsecas com o problema de imersão, a teoria de codificação e de informação, a análise geométrica e o formalismo termodinâmico.

Primeiramente, apresentamos duas noções de dimensão métrica média em termos da teoria da medida - as quais também dependem da estrutura métrica do espaço - ambas satisfazendo um princípio variacional clássico, e estudamos suas propriedades. Além disso, obtemos um princípio variacional em termos da dimensão métrica média local e utilizamos nossos resultados para obter fórmulas fechadas para a dimensão métrica média superior com potencial em duas classes de sistemas.

Em segundo lugar, investigamos a dimensão métrica média de subshifts de tipo compacto. Provamos que as dimensões métricas médias de um mapa contínuo e de seu limite inverso coincidem, generalizando a fórmula da entropia de Bowen. Com base nesse resultado, estendemos a noção de dimensão métrica média para mapas descontínuos, utilizando subshifts apropriados. Como aplicação, mostramos que a dimensão métrica média do mapa de Gauss e dos mapas induzidos pela família de Manneville-Pomeau é igual à dimensão box do conjunto de pontos de descontinuidade, que também coincide com um parâmetro crítico do operador pressão associado ao potencial geométrico.

Por fim, relacionamos a dimensão métrica média com a estrutura fractal do espaço de fase e com a regularidade Hölder do mapa. Posteriormente, aprimoramos nossas estimativas gerais em uma família de mapas de intervalo, calculando a dimensão métrica médiade maneira análoga à fórmula de Misiurewicz para a entropia, o que, em particular, mostra que nossas estimativas são otimais. Aplicando nossos resultados, determinamos a dimensão métrica média das funções de Weierstrass. Como resultado de interesse independente, desenvolvemos um análogo dinâmico da dimensão de Minkowski-Bouligand para mapas que atuam em espaços Ahlfors regulares.

Table of contents

List of figures

Chapter 1

Introduction

The theory of mean dimension has its origins in the work of Gromov [27], where spaces of holomorphic maps arising in geometric analysis are studied from a dynamical point of view, under a natural group action. This approach is grounded in a new invariant of infinite dimensional systems called *(topological) mean dimension*, which is a dynamical analogue of the Lebesgue covering dimension. Gromov's program has been followed by many authors and, for instance, a fomula for the mean dimension of the system of Brody curves was successfully obtained by Tsukamoto [55]. Mean dimension was further studied by Lindenstrauss and Weiss in [43], where it is used to give a negative answer to Auslander's problem of whether every minimal dynamical system can be embedded in $([0,1]^{\mathbb{Z}}, \sigma)$. They show that a necessary condition is for the mean dimension of the system to be at most 1 and then construct an example attaining a bigger value. Conversely, Lindenstrauss proved in [40] that any minimal system with mean dimension less than $N/36$ can be embedded in $\big(([0,1]^N)^{\mathbb{Z}}, \sigma\big)$ and asked what is the optimal constant $c > 0$ so that the embedding result holds whenever the mean dimension is less than cN. This question was finally settled by Gutman and Tsukamoto in [30], where they prove that the optimal value is $c = 1/2$.

Besides the applications to the embedding problem, Lindestrauss and Weiss [43] developed another dynamical invariant called *metric mean dimension*, which is the central subject of this thesis. It is a dynamical analogue of the Minkowski or box-counting dimension and depends on the *metric structure* of the underlying space, not only on the topology. Moreover, just like the Lebesgue covering dimension is always at most the box dimension, they showed that the mean dimension is less or equal than the metric mean dimension, for any metric that generates the topology. This makes the metric mean dimension very useful in obtaining (possibly tight) upper bounds on mean dimension, as it is much easier to compute due to its local nature (cf. [58]). The metric mean dimension vanishes if the topological entropy is finite (consequently, the mean dimension is also zero, but from now on we focus only on metric mean dimension). If, otherwise, the entropy is infinite, it measures the speed at which the entropy at a given scale diverges as this scale goes to zero, just like the box dimension quantifies the growth of the ε-covering numbers of the space as $\varepsilon \to 0$. The choice of the metric has impact precisely on the speed of this convergence. Moreover, the metric mean dimension is always bounded from above by the box dimension of the phase space.

Ergodic aspects of the theory first appeared in the pioneering work of Lindenstrauss and Tsukamoto [41], where they provided a variational principle linking dynamical systems (metric mean dimension) and information theory (rate distortion function). This was followed by several authors, who replaced the rate distortion function by other measure theoretic objects more familiar to ergodic theorists, such as Kolmogorov-Sinai entropy (Gutman-Śpiewak [29]) and Brin-Katok entropy (Shi [53]), just to mention a few. The general structure of such variational principles (see more details in Chapter 3) is

$$\overline{\text{mdim}}_M(X,d,T) = \limsup_{\varepsilon} \sup_{\mu} \frac{h(\mu,\varepsilon)}{\log(1/\varepsilon)}, \tag{1.1}$$

where $h(\mu,\varepsilon)$ is an entropy-like function of μ at scale ε and the maximization is over ergodic, or invariant, probability measures. A recurrent question in the literature asks whether, or under which conditions, we may exchange the order of $\limsup_\varepsilon$ and $\sup_\mu$. An affirmative answer would yield the existence of a *single* probability measure that captures almost full complexity of the system over, *infinitely many*, small scales. It is known that this order can be exchanged under the marker property (cf. [42, 57]). However, for general systems the order exchange may fail, as illustrated in Example 3.6.3 and Remark 3.6.4.

Meanwhile, Tsukamoto introduced in [57] the concept of *upper metric mean dimension with potential*, and proved a double variational principle similar to the one derived earlier in [42] for the (topological) mean dimension. Tsukamoto's definition of upper metric mean dimension with potential is inspired by the topological pressure, though the potential varies with the scale ε. One is led to ask whether there is a measure-theoretic upper metric mean dimension $H\colon \mathscr{P}_T(X) \to \mathbb{R}$ satisfying a classical variational principle (like the one established in [60, Theorem 9.10]), that is,

$$\overline{\text{mdim}}_M(X,d,T,\varphi) = \sup_{\mu \in \mathscr{P}_T(X)} \Big\{ H(\mu) + \int \varphi \, d\mu \Big\}. \tag{1.2}$$

As happens with the topological pressure, there may exist several maps H satisfying such a relation, and the selection of a particular one depends on the preferred mechanism to generate them, or on their applications. We stress that in (1.1) the maximization on the space of probability measures is done for a fixed scale, which only afterwards is made to decrease to zero. Therefore, it does not comply with formula (1.2). The main goal in Chapter 3 is to provide two definitions of H - for general continuous maps acting on a compact metric space - satisfying (1.2), and study their properties. Our first approach is an application of *convex analysis* and provides an abstract variational definition. The second, which we refer to as the *double envelope* approach, describes the complexity of a measure in terms of finite scale entropies, but one must take into account how the measure can be approximated in the simplex of invariant measures. In the examples we consider, the two notions coincide but, in general, we only have an inequality between them.

A natural class of systems to which the mean dimension theory can be applied is the one of subshifts on compact alphabets. Firstly, the embedding theorem of Gutman-Tsukamoto guarantees that subshifts of the Euclidean cube contain a wide variety of dynamical systems. Secondly, since alphabets of infinite cardinality can give rise to many subshifts with infinite entropy, we need other invariants to quantify their complexity. Tsukamoto [56] provided formulas for the mean dimension of full shifts with compact finite dimensional (with respect to the Lebesgue covering dimension) alphabets. For the metric mean dimension of full shifts, endowed with a natural metric, it is easy to prove that it coincides with the *box dimension* of the alphabet (see Example 2.4.1), resembling the fact that topological entropy of a full shift expresses the *cardinality* of the alphabet. Building on the bridge between mean dimension theory and information theory developed by Lindenstrauss and Tsukamoto [41], Gutman and Śpiewak [28] studied almost lossless analog compression. The signals to be compressed are modeled, in the language of ergodic theory, by invariant measures of $(\mathbb{R}^{\mathbb{Z}}, \sigma)$. They searched for uniform estimates on compression rates of signals that belong to a given compactly supported set of trajectories, that is, estimates valid for all invariant measures supported on a given subshift of $([0,1]^{\mathbb{Z}}, \sigma)$, and the resulting lower bound is given in terms of the metric mean dimension of that subshift.

In Chapter 4 we study a class of subshifts introduced by Friedland in [25], which we refer to as *subshifts of compact type*. They are a natural generalization of subshifts of finite type to compact alphabets, where the transition set is now an arbitrary subset $\Gamma \subset X \times X$. We prove that the unilateral and bilateral subshifts associated to a closed transition set Γ have the same metric mean dimension and, as a consequence, we show that taking the inverse limit of a continuous map does not change its metric mean dimension, generalizing a result of Bowen [8]. This allows us to extend the notion of metric mean dimension to a discontinuous map T as the metric mean dimension of the subshift of compact type with transition set $\Gamma = \mathrm{graph}(T)$. As an application of this concept, we consider a class of interval maps with infinitely many full branches, like the Gauss map or an induced map of the Manneville-Pomeau family, and show that the metric mean dimension coincides with the box dimension of the set of critical points. Again, this result is parallel to the known fact that the entropy of a map composed by finitely many, say c, full branches is $\log c$. In our setting, the role of counting branches (which represents the number of critical points of the map) is played by the computation of the box dimension of the (now infinite) set of critical points.

Another application of subshifts of compact type is that they provide a natural framework for studying free semigroup actions, and this was actually the main goal of Friedland [25] when working with this class of systems. For a finite collection of endomorphisms, we consider a transition set given by the union of their graphs and its induced subshift. Friedland proposed that the entropy of the action should be given by the entropy of the corresponding subshift, since the same holds in the case of a single map. We define the metric mean dimension of the action in a similar manner and prove that it is related through a variational principle to a more recent definition, proposed by Carvalho, Rodrigues and Varandas [14]. Their approach is inspired by Bufetov (see [9]), where one selects randomly which map will be used to evolve time on each step; and our variational principle states that by optimizing this probabilistic choice, one attains Friedland's definition, which in general is an upper bound. We use this variational principle to construct two maps with zero dynamical complexity (more precisely, both of their second iterates are constant maps) such that their action has positive metric mean dimension and, in particular, infinite entropy (see Subsection 4.3.4).

A classical result of Yano [62] states that on a compact manifold of dimension bigger than one, a C^0-generic homeomorphism has infinite entropy. This was generalized by Carvalho, Rodrigues and Varandas [13], who refined Yano's construction to prove that the metric mean dimension of a C^0-generic homeomorphism is maximal, that is, it is equal to the dimension of the manifold. This was further extended to C^0-generic non-invertible maps by Acevedo [1] and to C^0-generic conservative homeomorphisms by Lacerda and Romaña [39]. These facts indicate that the information provided by the metric mean dimension of a finite dimensional dynamical system is likely to be eclipsed by the overly flexible topological category. This is due to the *metric* nature of the metric mean dimension, which is more restrictive than pure topology. It seems that the C^0 class of maps is not the best to be investigated with this invariant. Indeed, the metric mean dimension should be more related to analytical regularity conditions on the map and other metric properties of the system, such as the fractal structure of the underlying space. On the other hand, any Lipschitz map on a finite dimensional space has finite entropy and, therefore, zero metric mean dimension. Consequently, the utilization of metric mean dimension within finite dimensional systems should be the most fruitful when aimed at maps with intermediate regularity, that is, between C^0 and C^1. The dynamical theory of such maps, which we refer to as *rough dynamics*, is far less developed than its neighbours' and it is expected that the mean dimension theory should provide its main invariants. This is suggested, for instance, by the works of de Faria, Hazard and Tresser [17] and Hazard [32], where they prove genericity of infinite entropy on Hölder and Sobolev spaces.

In Chapter 5 we provide a sharp upper bound on the metric mean dimension in terms of the Hölder regularity of the map and the fractal structure of the space. It was clear from the beginning that the metric mean dimension should be related to the box dimension. Perhaps the main feature of this part of the text is the appearance of a recent notion of dimension called *Assouad spectrum*, introduced by Fraser and Yu [23], which serves as a one parameter interpolation between the upper box dimension and the Assouad dimension. This dimension, at a given parameter $\alpha \in (0,1)$, essentially quantifies how many ε-balls are necessary to cover an ε^{α}-ball. Therefore, it is naturally associated to dynamical covers generated by an α-Hölder map.

In fractal geometry, a longstanding and active research direction is to try to understand all sorts of fractal dimensions of highly oscillating functions in the interval. The most famous example is the Hausdorff dimension of the graph of the continuous nowhere differentiable Weierstrass functions, established by Shen [52]. A theory of rough dynamics should include a dynamical counterpart of this program, which is naturally framed in terms of mean dimension theory.

1.1 Structure of the thesis

In Chapter 2 we introduce some basic terminology and include a glossary with the main definitions and some preliminary information.

In Chapter 3 we provide two notions of measure-theoretic metric mean dimension, both satisfying a classical variational principle, and study their properties. We also obtain a variational principle in terms of local metric mean dimension, and make use of our results to deduce closed formulas for the upper metric mean dimension with potential of two classes of systems (see Subsections 2.4 and 3.6).

In Chapter 4 we investigate the metric mean dimension of subshifts of compact type. We prove that the metric mean dimensions of a continuous map and its inverse limit coincide, generalizing Bowen's entropy formula. Building upon this result, we extend the notion of metric mean dimen- sion to discontinuous maps in terms of suitable subshifts. As an application, we show that the metric mean dimension of the Gauss map and that of induced maps of the Manneville-Pomeau family is equal to the box dimension of the corresponding set of discontinuity points, which also coincides with a critical parameter of the pressure operator associated to the geometric potential.

In Chapter 5 we relate the metric mean dimension with the fractal structure of the phase space and the Hölder regularity of the map. Afterwards we improve our general estimates in a family of interval maps by computing the metric mean dimension in a way similar to the Misiurewicz formula for the entropy, which in particular shows that our bounds are sharp. Of independent interest, we develop a dynamical analogue of Minkowski-Bouligand dimension for maps acting on Ahlfors regular spaces.

Chapter 2

Notation and preliminaries

Throughout this thesis, let (X,d) be a compact metric space and $T\colon X\to X$ be a continuous map. Denote by $C^0(X)$ the space of real valued continuous maps whose domain is X, endowed with the uniform norm; by $\mathfrak{B}$ the σ-algebra of Borel sets of X; by $\mathscr{P}(X)$ the set of Borel probability measures on X with the weak*-topology; by $\mathscr{P}_T(X)$ its subset of T-invariant measures; by $\mathscr{E}_T(X)$ the set of its ergodic elements; and, given $\mu\in\mathscr{P}(X)$, let $\mathrm{supp}(\mu)$ stand for its support.

We denote the open ball of center x and radius ε by $B_\varepsilon(x,d)$, and omit d if the metric is clear. For each positive integer $n\in\mathbb{N}$ and each continuous potential $\varphi\colon X\to\mathbb{R}$, we denote the Bowen metric

$$d_n(x,y)=\max_{0\le j\le n-1} d(T^j(x),T^j(y)) \qquad \forall x,y\in X$$

which is equivalent to d and the Birkhoff sum $S_n\varphi=\sum_{j=0}^{n-1}\varphi\circ T^j$.

For a vector space V, we say that $f\colon V\to\mathbb{R}\cup\{+\infty\}$ is *convex* if the set $C=\{x\in V\colon f(x)<+\infty\}$, which we refer to as the *support of* f, is convex and $f(\lambda x+(1-\lambda)y)\le\lambda f(x)+(1-\lambda)f(y)$ for all $x,y\in C$ and $\lambda\in(0,1)$. We say that $f\colon V\to\mathbb{R}\cup\{-\infty\}$ is *concave* if $-f$ is convex.

2.1 Metric mean dimension with potential

Given a subset $K\subset X$, a continuous potential $\varphi\in C^0(X)$ and $\varepsilon>0$, consider the following infimum

$$S(K,d,\varphi,\varepsilon)=\inf\left\{\sum_{i=1}^{\ell}(1/\varepsilon)^{\sup_{U_i}\varphi}\colon\ \{U_i\}_{1\le i\le\ell}\text{ finite open cover of }K,\ \mathrm{diam}(U_i,d)<\varepsilon\right\} \qquad (2.1)$$

and the limit

$$P(K,d,T,\varphi,\varepsilon)=\lim_{n\to+\infty}\frac{1}{n}\log S(K,d_n,S_n\varphi,\varepsilon),$$

which exists since the sequence $\big(\log S(K,d_n,S_n\varphi,\varepsilon)\big)_n$ is sub-additive in the variable n. In the case $\varphi=0$, we drop it from the notation and write $h_\varepsilon(K,d,T)=P(K,d,T,0,\varepsilon)$. Recall that the *topological entropy of* (K,T) is given by $h_{top}(K,T)=\lim_{\varepsilon\to0^+}h_\varepsilon(K,d,T)$. The upper/lower metric mean dimension with potential, as defined in [57], extends the concept of metric mean dimension introduced in [43] (corresponding to the particular case of $\varphi=0$) as follows:

Definition 2.1.1 *The upper/lower metric mean dimension with potential of* (K,d,T,φ) *is given by*

$$\overline{\mathrm{mdim}}_M(K,d,T,\varphi) = \limsup_{\varepsilon\to 0^+} \frac{P(K,d,T,\varphi,\varepsilon)}{\log(1/\varepsilon)}$$
$$\underline{\mathrm{mdim}}_M(K,d,T,\varphi) = \liminf_{\varepsilon\to 0^+} \frac{P(K,d,T,\varphi,\varepsilon)}{\log(1/\varepsilon)}.$$

If the limit exists, we refer to it as the metric mean dimension with potential of (K,d,T,φ).

Again, if $\varphi = 0$, we drop it from the notation. A subset $E \subset K$ is said to be ε-separated with respect to the metric d if $d(x,y) \geq \varepsilon$ for every $x,y \in E$; it is ε-spanning with respect to the metric d if for every $x \in K$ there exists some $y \in E$ such that $d(x,y) \leq \varepsilon$. The notion of upper metric mean dimension with potential can be equivalently defined if one replaces $S(K,d,\varphi,\varepsilon)$ by either

$$S_1(K,d,\varphi,\varepsilon) = \sup\Big\{\sum_{x\in E}(1/\varepsilon)^{\varphi(x)} \colon\ E\subset K \text{ is } \varepsilon\text{-separated}\Big\}$$

or

$$S_2(K,d,\varphi,\varepsilon) = \inf\Big\{\sum_{x\in E}(1/\varepsilon)^{\varphi(x)} \colon\ E\subset K \text{ is } \varepsilon\text{-spanning}\Big\}.$$

Since S, S_1 and S_2 coincide, up to multiplying the scale by a constant, we will abuse the notation and use them indistinctively as follows: for $i = 1,2$,

$$P(K,d,T,\varphi,\varepsilon) = \lim_{n\to+\infty} \frac{1}{n}\log S(K,d_n,S_n\varphi,\varepsilon) = \limsup_{n\to+\infty} \frac{1}{n}\log S_i(K,d_n,S_n\varphi,\varepsilon). \tag{2.2}$$

The *upper/lower box dimension* of (X,d) is defined as

$$\overline{\dim}_B(X,d) = \limsup_{\varepsilon\to 0^+} \frac{\log S(X,d,0,\varepsilon)}{\log(1/\varepsilon)}$$
$$\underline{\dim}_B(X,d) = \liminf_{\varepsilon\to 0^+} \frac{\log S(X,d,0,\varepsilon)}{\log(1/\varepsilon)}.$$

Again, S can be replaced by S_1 or S_2. If the upper and lower box dimensions of X coincide, their value is the box dimension of X.

2.2 Katok entropy

Given an ergodic probability measure $\mu \in \mathscr{E}_T(X)$, $n \geq 1$, $\varepsilon > 0$ and $\delta \in (0,1)$, denote

$$N_\mu(X,d,T,n,\varepsilon,\delta) = \inf_B \big\{S(B,d_n,0,\varepsilon) \colon\ \mu(B) > 1-\delta\big\},$$

where the infimum is taken over all measurable subsets $B \subset X$ with measure μ bigger than $1-\delta$. In fact, the same value is attained if, instead of being just measurable, we assume that B ranges over finite unions of $(n,\varepsilon)-$dynamical balls whose union has measure bigger than $1-\delta$. We will always consider such B's. The $\delta-$*Katok entropy of* μ *at scale* ε, defined in [34], is given by

$$h^K_\mu(X,d,T,\varepsilon,\delta) = \limsup_{n\to+\infty} \frac{1}{n}\log N_\mu(X,d,T,n,\varepsilon,\delta).$$

If the dynamical system (X,d,T) is clear, we will omit it from the notation. The previous notion can be extended to non-ergodic invariant probability measures via integration: given $\mu \in \mathscr{P}_T(X)$, define

$$h^K_\mu(\varepsilon,\delta) = \int_{\mathscr{E}_T(X)} h^K_m(\varepsilon,\delta)\, d\mathbb{P}_\mu(m) \tag{2.3}$$

where $\mu = \int_{\mathscr{E}_T(X)} m\, d\mathbb{P}_\mu(m)$ is the ergodic decomposition of μ. We note that, by definition, the map h^K_m is measurable and integrable; this way, the function

$$\mu \in \mathscr{P}_T(X) \quad \mapsto \quad h^K_\mu(\varepsilon,\delta)$$

is also affine.

The following variational principle links Katok entropy and metric mean dimension (see [53] and [16]).

Theorem 2.2.1 *Let (X,d) be a compact metric space and $T\colon X\to X$ be a continuous map. Then, for every $\delta\in(0,1)$,*

$$\begin{aligned}\overline{\mathrm{mdim}}_M(X,d,T) &= \limsup_{\varepsilon\to 0^+} \frac{\sup_{\mu\in\mathscr{E}_T(X)} h^K_\mu(\varepsilon,\delta)}{\log(1/\varepsilon)}\\ \underline{\mathrm{mdim}}_M(X,d,T) &= \liminf_{\varepsilon\to 0^+} \frac{\sup_{\mu\in\mathscr{E}_T(X)} h^K_\mu(\varepsilon,\delta)}{\log(1/\varepsilon)}.\end{aligned}$$

For further use, we also include the following easy consequence of the previous variational principle.

Corollary 2.2.2 *Let (X,d) be a compact metric space and $T\colon X\to X$ be a continuous map. Then,*

$$\begin{aligned}\overline{\mathrm{mdim}}_M(X,d,T) &= \overline{\mathrm{mdim}}_M\Big(\bigcap_{k=0}^{\infty} T^k(X),d,T\Big)\\ \underline{\mathrm{mdim}}_M(X,d,T) &= \underline{\mathrm{mdim}}_M\Big(\bigcap_{k=0}^{\infty} T^k(X),d,T\Big).\end{aligned}$$

2.3 Basic properties of the metric mean dimension

It the following properties are immediate consequences of the definition of metric mean dimension.

Proposition 2.3.1 *Let (X_1,d_1) and (X_2,d_2) be compact metric spaces and $T_1\colon X_1\to X_1$ and $T_2\colon X_2\to X_2$ be continuous maps.*

(a) *If there exists a surjective Lipschitz map $\pi\colon X_2 \to X_1$ such that $T_1 \circ \pi = \pi \circ T_2$, then*

$$\begin{aligned}\overline{\mathrm{mdim}}_M(X_1,d_1,T_1) &\leq \overline{\mathrm{mdim}}_M(X_2,d_2,T_2)\\ \underline{\mathrm{mdim}}_M(X_1,d_1,T_1) &\leq \underline{\mathrm{mdim}}_M(X_2,d_2,T_2).\end{aligned}$$

(b) *Given a positive integer k,*

$$\begin{aligned}\overline{\mathrm{mdim}}_M(X_1,d_1,T_1^k) &\leq k\,\overline{\mathrm{mdim}}_M(X_1,d_1,T_1)\\ \underline{\mathrm{mdim}}_M(X_1,d_1,T_1^k) &\leq k\,\underline{\mathrm{mdim}}_M(X_1,d_1,T_1),\end{aligned}$$

and the inequality can be strict (see Remark 2.4.1).

(c) *For every $K \subset X_1$,*

$$\begin{aligned}\overline{\mathrm{mdim}}_M(\overline{K},d_1,T_1) &= \overline{\mathrm{mdim}}_M(K,d_1,T_1)\\ \underline{\mathrm{mdim}}_M(\overline{K},d_1,T_1) &= \underline{\mathrm{mdim}}_M(K,d_1,T_1).\end{aligned}$$

(d) *For every $K \subset X_1$ such that, $T_1(K) \subset K$,*

$$\begin{aligned}\overline{\mathrm{mdim}}_M(K,d_1,T_1) &\leq \overline{\dim}_B(K,d_1)\\ \underline{\mathrm{mdim}}_M(K,d_1,T_1) &\leq \underline{\dim}_B(K,d_1).\end{aligned}$$

(e) *For every $\varphi,\, \psi \in C^0(X_1)$,*

$$\begin{aligned}\overline{\mathrm{mdim}}_M(X_1,d_1,T_1^k,\varphi) &= \overline{\mathrm{mdim}}_M(X_1,d_1,T_1,\varphi+\psi\circ T_1-\psi)\\ \underline{\mathrm{mdim}}_M(X_1,d_1,T_1^k,\varphi) &= \underline{\mathrm{mdim}}_M(X_1,d_1,T_1,\varphi+\psi\circ T_1-\psi).\end{aligned}$$

2.4 Toy models

In this section we present the two main, though not the only, classes of examples in the text. They will show up in most chapters to illustrate our results and demonstrate their sharpness.

2.4.1 Shifts on compact alphabets

Let $\mathbb{K} = \mathbb{N}$ or $\mathbb{Z}$. Fix some $\rho > 1$ and endow the product space $X^{\mathbb{K}}$ with the metric

$$d^{\mathbb{K}}(x,y) = \sup_{i\in\mathbb{K}} \frac{d(x_i,y_i)}{\rho^{|i-1|}}. \tag{2.4}$$

We will always consider $\rho = 2$ unless stated otherwise, in which case we denote it by $d_\rho^{\mathbb{K}}$ or simply by d_ρ, provided there is no risk of ambiguity. A subset $Y \subset X^{\mathbb{K}}$ is called a *subshift* if it is closed and invariant by the shift action

$$\sigma_{\mathbb{K}}\big((x_i)_{i\in\mathbb{K}}\big) = (x_{i+1})_{i\in\mathbb{K}}.$$

The *full shift* $\sigma_{\mathbb{K}}\colon X^{\mathbb{K}} \to X^{\mathbb{K}}$ is the simplest example where the metric mean dimension can be computed:

$$\begin{aligned} \overline{\mathrm{mdim}}_M(X^{\mathbb{K}}, d^{\mathbb{K}}, \sigma_{\mathbb{K}}) &= \overline{\dim}_B(X,d) \\ \underline{\mathrm{mdim}}_M(X^{\mathbb{K}}, d^{\mathbb{K}}, \sigma_{\mathbb{K}}) &= \underline{\dim}_B(X,d). \end{aligned} \tag{2.5}$$

Proof: We will prove the unilateral case with $\rho = 2$. For each $\varepsilon > 0$, let $\ell = \ell(\varepsilon)$ be a positive integer such that $2^{-\ell}\operatorname{diam} X < \varepsilon$ and $\mathscr{U} = \{U_1, ..., U_N\}$ be a minimal open cover of X satisfying $\operatorname{diam}(\mathscr{U}, d) < \varepsilon$. Then the open cover

$$\alpha_{n,\varepsilon} = \{U_{j_1} \times ... \times U_{j_{n+\ell}} \times X \times X \cdots : j_i = 1, ..., N\}$$

of $X^{\mathbb{N}}$ satisfies $\operatorname{diam}(\alpha_{n,\varepsilon}, d_n^{\mathbb{N}}) \le \varepsilon$. Hence $h_\varepsilon(X^{\mathbb{N}}, d^{\mathbb{N}}, \sigma_{\mathbb{N}}) \le \log N$, and we obtain

$$\begin{aligned} \overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d_2, \sigma_{\mathbb{N}}) &\le \overline{\dim}_B(X,d) \\ \underline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d_2, \sigma_{\mathbb{N}}) &\le \underline{\dim}_B(X,d). \end{aligned}$$

For the converse inequality, Let $E \subset X$ be a maximal ε-separated subset. Then

$$\{(x_i)_{i\in\mathbb{N}} \in X^{\mathbb{N}} \colon x_1, ..., x_n \in E\}$$

is an ε-separated subset of $X^{\mathbb{N}}$ with respect to the metric $d_n^{\mathbb{N}}$. Hence $S_1(X^{\mathbb{N}}, d_n^{\mathbb{N}}, \varepsilon) \ge n \log \#E$, and we obtain

$$\begin{aligned} \overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d_2, \sigma_{\mathbb{N}}) &\ge \overline{\dim}_B(X,d) \\ \underline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d_2, \sigma_{\mathbb{N}}) &\ge \underline{\dim}_B(X,d). \end{aligned}$$

□

This simple example reveals the dependence of the metric mean dimension on the metric structure of the phase space. For instance, consider the symbolic space $X = \{0,1\}^{\mathbb{N}}$ with the product topology. This topology is generated by both distances

$$d^1(x,y) = \sum_{n\in\mathbb{N}} \frac{|x_n - y_n|}{2^n} \quad \text{and} \quad d^2(x,y) = \sum_{n\in\mathbb{N}} \frac{|x_n - y_n|}{3^n}.$$

Endowing the product space $X^{\mathbb{N}}$ with the metrics d_2^j, for $j = 1, 2$, given by

$$d_2^j(a,b) = \sup_{n\in\mathbb{N}} \frac{d^j(a_n, b_n)}{2^n}$$

which induce the same topology in $X^{\mathbb{N}}$, we get

$$\begin{aligned} \overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d_2^1, \sigma^{\mathbb{N}}) &= \dim_B(X, d_1) = 1 \\ \overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d_2^2, \sigma^{\mathbb{N}}) &= \dim_B(X, d_2) = \log 2/\log 3. \end{aligned} \tag{2.6}$$

If one fixes the topology and considers the space of generating distances (with a proper metric structure), the metric mean dimension can be seen as a function of this parameter, which in general is not continuous anywhere (cf. [47]).

Under a dimensional homogeneity assumption on the alphabet[1], we will extend the formula (2.5) to all continuous potentials $\varphi \in C^0(X^{\mathbb{N}})$ (see Theorem C):

$$\overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d^{\mathbb{N}}, \sigma_{\mathbb{N}}, \varphi) = \dim_B(X,d) + \max_{\mu \in \mathscr{E}_{\sigma_{\mathbb{N}}}(X^{\mathbb{N}})} \int \varphi \, d\mu \qquad \forall \varphi \in C^0(X).$$

2.4.2 Accumulating horseshoes

In this subsection we consider the interval $[0,1]$ with the Euclidean distance. We will present a $2-$parameter family of interval maps firstly introduced by Kolyada and Snoha (cf. [38]).

Fix a sequence $a = (a_k)_{k\geq 0}$ of real numbers and a sequence $b = (b_k)_{k\geq 1}$ of positive odd integers satisfying the following conditions:

$$\begin{aligned} &\text{(C1)} \quad 0 = a_0 < a_1 < \ldots < a_k \to 1. \\ &\text{(C2)} \quad (a_k - a_{k-1})_k \text{ decreases to zero.} \\ &\text{(C3)} \quad (b_k)_k \text{ is strictly increasing.} \end{aligned} \tag{2.7}$$

For each $k \in \mathbb{N}$, denote by $f_k \colon [0,1] \to [0,1]$ the unique continuous piecewise affine map with b_k equidistributed full branches such that f_k is increasing in the first one - and since b_k is odd, also in the last. Let $J_k = [a_{k-1}, a_k]$ and $T_k \colon J_k \to [0,1]$ be the unique increasing affine map from J_k onto $[0,1]$. Define

$$\begin{array}{rccc} T_{a,b}\colon & [0,1] & \to & [0,1] \\ & x \in J_k & \mapsto & T_k^{-1} \circ f_k \circ T_k \\ & x = 1 & \mapsto & 1 \end{array} \tag{2.8}$$

Each map is determined by a sequence of horseshoes accumulating at the extreme point 1 (see Figure 2.1). The sequence $(a_k)_k$ defines the intervals $(J_k)_k$ where the horseshoes lie, and each b_k represents the number of branches in the $k-$th horseshoe. In particular, the topological entropy of $T_{a,b}$ restricted to J_k is $\log b_k$. To fix notation we write

$$\varepsilon_k = \frac{|J_k|}{b_k}, \tag{2.9}$$

which represents the length of the injectivity domains in J_k.

Given $k \in \mathbb{N}$, the set J_k can be partitioned into b_k intervals $J_k(1), \cdots, J_k(b_k)$ with the same length in such a way that

$$T(J_k(i)) = J_k \qquad \forall\, i \in \{1, \cdots, b_k\}.$$

[1]We say that (X,d) is *box-dimensionally homogeneous* if $\dim_B(B_\varepsilon(x)) = \dim_B(X)$, for every $x \in X$ and $\varepsilon > 0$.

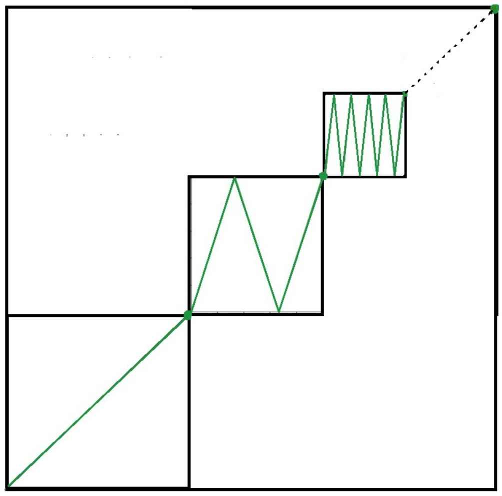

Fig. 2.1 Graph of the map $T_{a,b}$.

Similarly, $J_k(i)$ can be partitioned into b_k intervals $J_k(i,1), \cdots, J_k(i,b_k)$ with the same length so that

$$T^2(J_k(i,j)) = J_k \qquad \forall\, i,j \in \{1,\cdots,b_k\}.$$

Inductively, for every $n \in \mathbb{N}$ and $(i_1,\cdots,i_n)$ such that $i_j \in \{1,\cdots,b_k\}$, we can split $J_k(i_1,\cdots,i_n)$ into b_k intervals

$$J_k(i_1,\cdots,i_n,1),\ J_k(i_1,\cdots,i_n,2),\ \cdots,\ J_k(i_1,\cdots,i_n,b_k)$$

with the same length and satisfying

$$T^{n+1}J_k(i_1,\cdots,i_n,i) = J_k \qquad \forall\, i \in \{1,\cdots,b_k\}.$$

We may rename these subsets so that the order of the intervals $J_k(i_1,\cdots,i_n,i)$ partitioning $J_k(i_1,\cdots,i_n)$ is increasing in i if i_1 is odd, and decreasing in i if i_1 is even. This choice fits the fact that f_k is increasing in $J_k(i_1)$ if i_1 is odd, and decreasing otherwise. This way,

$$T^s_{a,b}\big(J_k(i_1,\cdots,i_n)\big) \;=\; J_k(i_{s+1},i_{s+2},\cdots,i_n) \qquad \forall\, s \in \{0,\cdots,n-1\}. \tag{2.10}$$

Clearly, $T_{a,b}$ has infinite topological entropy for every a and b. The metric mean dimension has been computed for specific choices of the parameters (cf. [59]) and we will establish a general expression, resembling Misiurewicz formula for the entropy, which now takes into account the scale at which the dynamics of each horseshoe can be detected. Namely, we will prove in Theorem I that

$$\begin{aligned} \overline{\mathrm{mdim}}_M([0,1],|\cdot|,T_{a,b}) &= \limsup_{k\to+\infty} \frac{\log b_k}{\log(1/\varepsilon_k)} \\ \underline{\mathrm{mdim}}_M([0,1],|\cdot|,T_{a,b}) &= \liminf_{k\to+\infty} \frac{\log b_k}{\log(1/\varepsilon_k)} \end{aligned}$$

and that the supremum of the Hölder exponents of $T_{a,b}$ is given by $1 - \overline{\mathrm{mdim}}_M([0,1],|\cdot|,T_{a,b})$.

Furthermore, we will extend the previous formula to all continuous potentials $\varphi \in C^0([0,1])$ in terms of ergodic optimization (see (3.34) in Example 3.6.3), namely

$$\overline{\mathrm{mdim}}_M([0,1],|\cdot|,T_{a,b},\varphi) = \max_{\mu \in \mathscr{E}_{T_{a,b}}\big([0,1]\big)} \int \big(\mathscr{D}+\varphi\big)\,d\mu \qquad \forall\, \varphi \in C^0([0,1]),$$

where

$$\mathscr{D}(x) = \begin{cases} \limsup_k \frac{\log b_k}{\log(1/\varepsilon_k)} & \text{if } x = 1 \\ 0 & \text{otherwise.} \end{cases}$$

In fact, it is a consequence of our results that $\mu \mapsto \int \mathscr{D}\,d\mu$ is the *unique* map $H\colon \mathscr{E}_{T_{a,b}}\big([0,1]\big) \to \mathbb{R}$ satisfying

$$\overline{\mathrm{mdim}}_M([0,1],|\cdot|,T_{a,b},\varphi) = \max_{\mu \in \mathscr{E}_{T_{a,b}}\big([0,1]\big)} \Big\{H(\mu) + \int \varphi\,d\mu\Big\} \qquad \forall\, \varphi \in C^0([0,1]).$$

Remark 2.4.1 *Under a suitable choice of parameters, say $a_k = \sum_{n=1}^{k} 6/\pi n^2$ and $b_k = 3^{k-1}$, the map $T_{a,b}$ reveals that the inequality in Proposition 2.3.1 (b) can be strict: by Proposition 2.3.1 (d), we have*

$$\mathrm{mdim}_M([0,1],|\cdot|,T^2_{a,b}) \;\leq\; 1 \;<\; 2 \;=\; 2\,\mathrm{mdim}_M([0,1],|\cdot|,T_{a,b}).$$

Chapter 3

Variational principles

The results in this chapter are based in the article

[11] M. Carvalho, G. Pessil and P. Varandas: *A convex analysis approach to the metric mean dimension: limits of scaled pressures and variational principles.* Advances in Mathematics 436 (2024) 109407.

3.1 Preamble

Over the last decade, a major development has been obtained in studying ergodic aspects of the mean dimension theory. Lindenstrauss and Tsukamoto [41] established, under mild conditions, a variational principle between the metric mean dimension and the L^p rate-distortion function $\mathscr{R}_{\mu,p}$ of each $\mu \in \mathscr{P}_T(X)$. More precisely, they showed that, for every $p \in \mathbb{N}$,

$$\overline{\mathrm{mdim}}_M(X,d,T) = \limsup_{\varepsilon \to 0^+} \sup_{\mu \in \mathscr{P}_T(X)} \frac{\mathscr{R}_{\mu,p}(\varepsilon)}{\log(1/\varepsilon)}. \tag{3.1}$$

Actually, Gutman and Śpiewak showed in [29] that it suffices to take the previous supremum over $\mathscr{E}_T(X)$, and obtained a new variational principle linking the upper metric mean dimension to the metric entropy h_μ, namely

$$\overline{\mathrm{mdim}}_M(X,d,T) = \limsup_{\varepsilon \to 0^+} \frac{1}{\log(1/\varepsilon)} \sup_{\mu \in \mathscr{E}_T(X)} \inf_{\mathrm{diam}(P) \le \varepsilon} h_\mu(P) \tag{3.2}$$

where the infimum is taken over the Borel partitions P of X with diameter at most ε and $h_\mu(P)$ stands for the entropy of P. Problem 3 in [29] asked whether the metric mean dimension could be expressed in terms of Brin-Katok local entropy h^{BK}. An affirmative answer to this problem was given by Shi in [53], where it is proved that

$$\overline{\mathrm{mdim}}_M(X,d,T) = \limsup_{\varepsilon \to 0^+} \frac{1}{\log(1/\varepsilon)} \sup_{\mu \in \mathscr{E}_T(X)} h^{BK}_\mu(\varepsilon). \tag{3.3}$$

These equations raise the question of whether we can define a meaningful notion of measure-theoretic metric mean dimension, in the sense that it should satisfy a variational principle. Furthermore, we would like $H\colon \mathscr{P}_T(X) \to \mathbb{R}$ to satisfy a classical variational principle in the more general thermodynamic formalism setting, that is,

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) = \sup_{\mu \in \mathscr{P}_T(X)} \Big\{ H(\mu) + \int \varphi \, d\mu \Big\}. \tag{3.4}$$

The first natural attempt is to drop $\sup_\mu$ in one of the equations (3.1)-(3.3) to define $\tilde{H}(\mu)$ and ask whether maximizing $\tilde{H}$ over invariant measures suffices to approximate the metric mean dimension, that is, whether we may exchange the order of $\limsup_\varepsilon$ and $\sup_\mu$ in (3.1)-(3.3). The validity of the order exchange is equivalent to the existence of a probability measure that captures almost full complexity of the system over, infinitely many, small scales. In Example 3.6.3, we consider a class of systems such that each ergodic measure is trapped in a finite (topological) entropy region and, therefore, each of such measures detect arbitrarily low complexity, in terms of metric mean dimension, at small scales, due to the normalization term. Hence, in general the order exchange fails and we need another way to define the complexity captured by a measure.

3.1.1 Convex analysis approach

Our starting point is an application of general results from Convex Analysis. In [6], the authors established an abstract variational principle for the so called *pressure functions* acting on a Banach space of potentials on a compact metric space, for which equilibrium states always exist. The main result of [6] ensures the existence of a conjugate of the pressure function, which acts on the set of Borel finitely additive probability measures if the space of potentials is made up from bounded measurable maps, and on the space of Borel probability measures if the space of potentials is $C^0(X)$. The first aim of our work is to show that the upper metric mean dimension with potential is a pressure function, so [6, Theorem 1] provides a suitable definition of a measure-theoretic upper metric mean dimension, defined on $\mathscr{P}(X)$ whenever the potentials belong to $C^0(X)$. This is the content Theorem A.

Theorem A *Let (X,d) be a compact metric space and $T\colon X \to X$ be a continuous map such that $\overline{\mathrm{mdim}}_M(X,d,T) < \infty$. Then, for every $\varphi \in C^0(X)$,*

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) = \max_{\mu \in \mathscr{P}_T(X)} \Big\{ M(\mu) + \int \varphi \, d\mu \Big\}, \tag{3.5}$$

where

$$M(\mu) = \inf_{\varphi \in \mathscr{C}} \int \varphi \, d\mu \quad \text{and} \quad \mathscr{C} = \big\{ \varphi \in C^0(X) \colon \overline{\mathrm{mdim}}_M(X,d,T,-\varphi) \le 0 \big\} \tag{3.6}$$

for every $\mu \in \mathscr{P}(X)$. The map M is concave, supported in $\mathscr{P}_T(X)$, upper semicontinuous, bounded from above by $\overline{\mathrm{mdim}}_M(X,d,T)$ and satisfies

$$M(\mu) = \inf_{\varphi \in C^0(X)} \Big\{ \overline{\mathrm{mdim}}_M(X,d,T,\varphi) - \int \varphi \, d\mu \Big\} \quad \forall \mu \in \mathscr{P}(X).$$

In addition, if for a a subset of probability measures $\mathscr{A} \subseteq \mathscr{P}(X)$ and a function $\tau\colon \mathscr{A} \to \mathbb{R} \cup \{-\infty\}$ we have

$$\sup_{\mu \in \mathscr{A}} \left\{ \tau(\mu) + \int \varphi \, d\mu \right\} \; \leq \; \overline{\mathrm{mdim}}_M(X,d,T,\varphi) \quad \forall \varphi \in C^0(X),$$

then

$$\tau(\mu) \; \leq \; M(\mu) \quad \forall \mu \in \mathscr{A}. \tag{3.7}$$

It is straightforward to conclude that μ attains the infimum in (3.6) (that is, $M(\mu) = \int \varphi_0 \, d\mu$ for some $\varphi_0 \in \mathscr{C}$) if and only if $\overline{\mathrm{mdim}}_M(X,d,T,-\varphi_0) = 0$ and μ is an equilibrium state of $-\varphi_0$. Indeed, if $\overline{\mathrm{mdim}}_M(X,d,T,-\varphi_0) = 0$ and μ is an equilibrium state of $-\varphi_0$, then

$$0 \, = \, \overline{\mathrm{mdim}}_M(X,d,T,-\varphi_0) \, = \, M(\mu) + \int -\varphi_0 \, d\mu$$

so $M(\mu) = \int \varphi_0 \, d\mu$. Conversely, if $M(\mu) = \int \varphi_0 \, d\mu$ for some $\varphi_0 \in \mathscr{C}$, then

$$0 \geq \overline{\mathrm{mdim}}_M(X,d,T,-\varphi_0) \geq M(\mu) - \int \varphi_0 \, = \, 0$$

so $\overline{\mathrm{mdim}}_M(X,d,T,-\varphi_0) \, = \, 0$ and μ is an equilibrium state of $-\varphi_0$.

A similar approach was proposed in a recent preprint [61] by Yang, Chen and Zhou, where the authors define a notion of measure-theoretic upper metric mean dimension. Our strategy is different from, though akin to, the one we read in [61], whose authors define the measure-theoretic upper metric mean dimension for a continuous map $T\colon X \to X$ of a compact metric space (X,d) by

$$\mu \in \mathscr{P}(X) \quad \mapsto \quad F(\mu) \, = \, \inf_{\varphi \in \tilde{\mathscr{C}}} \int \varphi \, d\mu$$

where

$$\tilde{\mathscr{C}} \, = \, \{\varphi \in C^0(X)\colon \; \overline{\mathrm{mdim}}_M(X,d,T,-\varphi) \, = \, 0\}.$$

Moreover, the authors showed that the map F satisfies the variational principle

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) \, = \, \sup_{\mu \in \mathscr{P}(X)} \left\{ F(\mu) + \int \varphi \, d\mu \right\} \qquad \forall \varphi \in C^0(X).$$

From the latter equality, we deduce by applying (3.7) to $\mathscr{A} = \mathscr{P}(X)$ that

$$F(\mu) \leq M(\mu) \qquad \forall \mu \in \mathscr{P}(X).$$

Yet, by definition, $\tilde{\mathscr{C}} \subseteq \mathscr{C}$; hence

$$M(\mu) \leq F(\mu) \qquad \forall \mu \in \mathscr{P}(X).$$

Thus, $M = F$, so we may add that F depends on the choice of the metric d (cf. Example 3.6.1), takes the value $-\infty$ for non-invariant measures (cf. Lemma 3.2.10) and always attains the maximum at some invariant measure. Therefore, Theorem A recovers the main contribution of [61].

3.1.2 Double envelope approach

Although Theorem A provides a map $M\colon \mathscr{P}(X) \to \mathbb{R}$ satisfying a variational principle, the variational nature of the definition of M prevents us from easily computing it in specific examples or drawing dynamical information from it. We will introduce another suitable notion of measure-theoretic upper metric mean dimension, built on Katok's description of the metric entropy, which also satisfies a variational principle like (3.4).

For finite entropy systems, the existence of equilibrium states is not always guaranteed. The lack of upper semicontinuity of the entropy map has led authors to regularize the notion of metric entropy, leading to the so called star entropy h^*, defined as the upper semicontinuous envelope of entropy: for every $\mu \in \mathscr{P}_T(X)$,

$$h^*(\mu) = \sup\left\{\limsup_k h(\mu_k) : (\mu_k)_k \subset \mathscr{P}_T(X) \text{ such that } \mu_k \to \mu\right\}.$$

Since metric mean dimension is highly dependent on the metric structure of the phase space, the role of the scales must be taken into account. In what follows, we consider a two variable semicontinuous envelope - hence *double* in the title - namely, depending on measures and scales *simultaneously*. An intrinsic aspect of this approach is the selection of an entropy-like function of a measure μ at a fixed scale ε, say $h(\mu,\varepsilon)$. For a given measure $\mu \in \mathscr{P}_T(X)$, we pair a scale ε with an approximating measure μ_ε and maximize the upper limits of $\frac{h(\mu_\varepsilon,\varepsilon)}{\log(1/\varepsilon)}$ over all sequences $(\varepsilon,\mu_\varepsilon) \to (0,\mu)$. We proceed by using Katok entropy, though our approach allows a generalization which will be established in Subsection 3.3.3.

Given $\mu \in \mathscr{P}_T(X)$ and $\delta \in (0,1)$, we define

$$\begin{array}{rccc} H^K_\delta\colon & \mathscr{P}_T(X) & \to & \mathbb{R} \\ & \mu & \mapsto & \sup_{(\mu_\varepsilon)_\varepsilon \in \mathscr{M}(\mu)} \limsup_{\varepsilon\to 0^+} \frac{h^K_{\mu_\varepsilon}(\varepsilon,\delta)}{\log(1/\varepsilon)} \end{array} \tag{3.8}$$

where $\mathscr{M}(\mu)$ stands for the space of sequences of probability measures in $\mathscr{P}_T(X)$ which converge to μ in the weak*-topology, and $h^K_\nu(\varepsilon,\delta)$ is the δ-Katok entropy of ν at scale ε.

Theorem B *Let (X,d) be a compact metric space and $T\colon X \to X$ be a continuous map such that $\overline{\mathrm{mdim}}_M(X,d,T) < +\infty$. Then, for every $\delta \in (0,1)$, the map H^K_δ is upper semicontinuous, bi-Lipschitz invariant and satisfies, for every $\varphi \in C^0(X)$,*

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) = \max_{\mu\in\mathscr{P}_T(X)} \left\{H^K_\delta(\mu) + \int \varphi\, d\mu\right\}. \tag{3.9}$$

In particular, $H^K_\delta(\mu) \le M(\mu)$ for every $\mu \in \mathscr{P}_T(X)$.

The more acessible definition of H^K_δ allows us to obtain explicit formulas for the metric mean dimension with potential of some dynamical systems. We illustrate this feature for full shifts on alphabets satisfying a (rather weak) box dimension homogeneity condition. Given a compact metric space (X,d) we recall that the space $X^{\mathbb{N}}$ is endowed with the metric

$$d_2(x,y) = \sup_{n\in\mathbb{N}} \frac{d(x_n,y_n)}{2^{n-1}}, \tag{3.10}$$

where $x = (x_n)_{n\in\mathbb{N}}$ and $y = (y_n)_{n\in\mathbb{N}}$. For each $\varphi \in C^0(X^{\mathbb{N}})$ the following result provides an exact formula for the upper metric mean dimension $\overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d_2, \sigma, \varphi)$, yielding a reformulation of the variational principles (3.5) and (3.9) in this setting. The obtained formula splits the upper metric mean dimension in two separate objects: the box dimension of the alphabet and an ergodic optimization term on the shift space.

Theorem C *Let (X,d) be a compact metric space such that $\dim_B U = \dim_B X$ for every nonempty open set $U \subset X$. Then,*

$$\overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d_2, \sigma, \varphi) = \dim_B X + \max_{\mu \in \mathscr{E}_\sigma(X^{\mathbb{N}})} \int \varphi \, d\mu \qquad \forall\, \varphi \in C^0(X^{\mathbb{N}}). \tag{3.11}$$

For instance, for every positive integer D,

$$\overline{\mathrm{mdim}}_M(([0,1]^D)^{\mathbb{N}}, d_\rho, \sigma, \varphi) = D + \max_{\mu \in \mathscr{E}_\sigma(([0,1]^D)^{\mathbb{N}})} \int \varphi \, d\mu \qquad \forall\, \varphi \in C^0(([0,1]^D)^{\mathbb{N}}).$$

For more information when $D = 1$, see Example 3.6.2. To the author's knowledge, this is the first explicit calculation of the upper metric mean dimension with potential of a dynamical system.

3.1.3 Local metric mean dimension

Inspired by the concepts of entropy point and local entropy introduced in [63], we define similarly the *local metric mean dimension function* by

$$\begin{aligned} \mathscr{D} \colon X &\to \mathbb{R} \\ x &\mapsto \inf\Big\{\overline{\mathrm{mdim}}_M(U,d,T) \colon\ U \text{ is an open neighborhood of } x\Big\}. \end{aligned} \tag{3.12}$$

The map $\mathscr{D}$ is upper semi-continuous and for any ergodic measure $\mu \in \mathscr{E}_T(X)$, it is constant μ-almost everywhere. There are two main motivations for introducing this concept. Firstly, it satisfies the following variational principle.

Theorem D *Let (X,d) be a compact metric space and $T\colon X \to X$ be a continuous map such that $\overline{\mathrm{mdim}}_M(X,d,T) < +\infty$. Then*

$$\overline{\mathrm{mdim}}_M(X,d,T) = \max_{\mu \in \mathscr{P}_T(X)} \int \mathscr{D} \, d\mu = \max_{\mu \in \mathscr{E}_T(X)} \int \mathscr{D} \, d\mu. \tag{3.13}$$

In addition, a measure $\mu \in \mathscr{P}_T(X)$ attains the previous maximum if and only if

$$\mathscr{D}|_{\mathrm{supp}(\mu)} \equiv \overline{\mathrm{mdim}}_M(X,d,T).$$

In order to explain our second motivation, we start by rephrasing Corollary 16 of [44] in our setting, which may be seen as a refinement of (3.7).

Proposition 3.1.1 *Let (X,d) be a compact metric space and $T: X \to X$ be a continuous map such that $\overline{\mathrm{mdim}}_M(X,d,T) < +\infty$. If $\tau\colon \mathscr{P}(X) \to \mathbb{R} \cup \{-\infty\}$ is concave, upper semicontinuous and satisfies*

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) = \max_{\mu \in \mathscr{P}(X)} \left\{ \tau(\mu) + \int \varphi \, d\mu \right\} \qquad \forall \varphi \in C^0(X), \tag{3.14}$$

then $\tau(\mu) = M(\mu)$ for every $\mu \in \mathscr{P}(X)$.

Extending the map H^K_δ to all non-invariant probability measures by the value $-\infty$, we obtain an upper semicontinuous map, bounded from above, and satisfying (3.14). If for a given system we are able to show, additionally, that H^K_δ is concave, then we can conclude that

$$H^K_\delta(\mu) = M(\mu) \qquad \forall \mu \in \mathscr{P}(X),$$

that is, the two seemingly independent approaches coincide. It turns out that in the computation of some examples (see Example 3.6.3), the obtained formula is exactly

$$H^K_\delta(\mu) = \int \mathscr{D} \, d\mu \qquad \forall \mu \in \mathscr{P}_T(X).$$

3.2 Pressure functions

In this section we collect some information from [6] concerning pressure functions. Let (X,d) be a locally compact metric space and $\mathbf{B}(X)$ be a Banach space over $\mathbb{R}$ equal to either

$$\begin{aligned} B_d(X) &= \{\phi\colon X \to \mathbb{R} : \phi \text{ is measurable and bounded}\} \\ \text{or } C^0(X) &= \{\phi \in B_d(X) : \phi \text{ is continuous}\} \\ \text{or else } C_c(X) &= \{\phi \in C^0(X) : \phi \text{ has compact support}\} \end{aligned}$$

endowed with the norm $\|\varphi\|_\infty = \sup_{x\in X} |\varphi(x)|$. In what follows, $\mathscr{P}_a(X)$ will stand for the set of normalized finitely additive set functions on the Borel σ-algebra of X, which we will simply call *finitely additive probability measures*, with the total variation norm, defined by

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

We recall the notion of pressure function used in [6].

Definition 3.2.1 *We say that a map $\Upsilon\colon \boldsymbol{B}(X) \to \mathbb{R}$ is a pressure function if it satisfies the following conditions:*

- *Monotonicity:* $\varphi \le \psi \implies \Upsilon(\varphi) \le \Upsilon(\psi) \quad \forall \varphi, \psi \in \boldsymbol{B}(X)$
- *Translation invariance:* $\Upsilon(\varphi + c) = \Upsilon(\varphi) + c \quad \forall \varphi \in \boldsymbol{B}(X), \forall c \in \mathbb{R}$
- *Convexity:* $\Upsilon(t\varphi + (1-t)\psi) \le t\Upsilon(\varphi) + (1-t)\Upsilon(\psi) \quad \forall \varphi, \psi \in \boldsymbol{B}(X), \forall t \in (0,1).$

For instance, the topological pressure of a continuous self-map $T\colon X \to X$ of a compact metric space X, whose topological entropy is finite, is a pressure function (cf. [60, Theorem 9.7]). More generally, if the weighted topological entropy is finite, the weighted topological pressure function defined in [21] is a pressure function.

We start by observing that a monotone and translation invariant map $\Upsilon\colon \mathbf{B}(X) \to \mathbb{R} \cup \{\pm\infty\}$ is finite valued or constantly ∞. Indeed, given φ, $\psi \in \mathbf{B}(X)$, then

$$\Upsilon(\varphi) - \|\varphi-\psi\|_\infty = \Upsilon(\varphi - \|\varphi-\psi\|_\infty) \le \Upsilon(\varphi) \le \Upsilon(\psi + \|\varphi-\psi\|_\infty) = \Upsilon(\psi) + \|\varphi-\psi\|_\infty.$$

Theorem 3.2.2 (Theorem 1 of [6]; see also [7]) *Let (X,d) be a locally compact metric space and $\Upsilon\colon \boldsymbol{B}(X) \to \mathbb{R}$ be a pressure function. Then*

$$\Upsilon(\varphi) = \max_{\mu \in \mathscr{P}_a(X)} \left\{ \mathfrak{h}(\mu) + \int \varphi \, d\mu \right\} \qquad \forall\, \varphi \in \boldsymbol{B}(X) \tag{3.15}$$

where the map $\mathfrak{h} = \mathfrak{h}_{\Upsilon,\boldsymbol{B}(X)}\colon \mathscr{P}_a(X) \to \mathbb{R} \cup \{-\infty,+\infty\}$ satisfies

$$\mathfrak{h}(\mu) = \inf_{\varphi \in \mathscr{A}_\Upsilon} \int \varphi \, d\mu \qquad \textit{and} \qquad \mathscr{A}_\Upsilon = \{\varphi \in \boldsymbol{B}(X) \,:\, \Upsilon(-\varphi) \le 0\}.$$

The map $\mathfrak{h}$ is concave, upper semicontinuous, bounded from above by $\Upsilon(0)$ and, if $\tau\colon \mathscr{P}_a(X) \to \mathbb{R} \cup \{-\infty,+\infty\}$ takes the role of $\mathfrak{h}$ in (3.15)*, then $\tau \le \mathfrak{h}$. Moreover, one has*

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

In case X is compact and $\boldsymbol{B}(X) = C^0(X)$, then the maximum is attained at $\mathscr{P}(X)$.

The variational principle (3.15) ensures that there always exist normalized finitely additive measures which attain the supremum; that is, the set

$$\mathscr{E}_\varphi(\Upsilon) = \left\{ \mu \in \mathscr{P}_a(X)\colon\ \Upsilon(\varphi) = \mathfrak{h}(\mu) + \int \varphi \, d\mu \right\}$$

is non empty.

Definition 3.2.3 *Consider a pressure function $\Upsilon\colon \boldsymbol{B}(X) \to \mathbb{R}$ and a potential $\varphi \in \boldsymbol{B}(X)$. We say that $\mu \in \mathscr{P}_a(X)$ is a tangent functional to Υ at φ if*

$$\Upsilon(\varphi + \psi) - \Upsilon(\varphi) \ge \int \psi \, d\mu \qquad \forall\, \psi \in \boldsymbol{B}(X).$$

We denote by $\mathscr{T}_\varphi(\Upsilon)$ the set of tangent functionals to Υ at φ.

The next result states that, for every potential φ, the set $\mathscr{T}_\varphi(\Upsilon)$ coincides with the space of finitely additive equilibrium states of φ and established a sufficient condition for the uniqueness of such finitely additive equilibrium states.

Theorem 3.2.4 (Theorem 2 of [6]) *Let (X,d) be a locally compact metric space and $\Upsilon\colon \boldsymbol{B}(X) \to \mathbb{R}$ be a pressure function. Then*

$$\mathscr{E}_\varphi(\Upsilon) = \mathscr{T}_\varphi(\Upsilon) \qquad \forall\, \varphi \in \boldsymbol{B}(X).$$

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

It is known that the uniqueness of equilibrium states for the topological pressure associated to a dynamical system is tied in with the differentiability of the pressure. In the wider setting of pressure functions one has the following generalization.

Definition 3.2.5 *A pressure function $\Upsilon\colon \boldsymbol{B}(X) \to \mathbb{R}$ is locally affine at $\varphi \in \boldsymbol{B}(X)$ if there exists a neighborhood $\mathscr{V}$ of $0 \in \boldsymbol{B}(X)$ and a unique $\mu_\varphi \in \mathscr{P}_a(X)$ such that*

$$\Upsilon(\varphi+\psi) - \Upsilon(\varphi) = \int \psi \, d\mu_\varphi \qquad \forall\, \psi \in \mathscr{V}.$$

Recall that $\Upsilon\colon \mathbf{B}(X) \to \mathbb{R}$ is *Fréchet differentiable at* $\varphi \in \boldsymbol{B}(X)$ if there exists a unique $\mu_\varphi \in \mathscr{P}_a(X)$ such that

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

Theorem 3.2.6 (Theorem 3 of [6]) *Let Let (X,d) be a locally compact metric space and $\Upsilon\colon \boldsymbol{B}(X) \to \mathbb{R}$ be a pressure function. The following are equivalent:*

(a) Υ *is locally affine at* φ.

(b) There exists a unique tangent functional $\mu_\varphi \in \mathscr{T}_\varphi(\Upsilon)$ *and*

$$\lim_{\psi \to 0} \sup \left\{ \|\mu - \mu_\varphi\| \colon \, \mu \in \mathscr{T}_{\varphi+\psi}(\Upsilon) \right\} = 0$$

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

In particular, the following are also equivalent:

(d) Υ *is affine.*

(e) $\bigcup_{\varphi \in \boldsymbol{B}(X)} \mathscr{T}_\varphi(\Upsilon)$ *is a singleton.*

(f) Υ *is everywhere Fréchet differentiable.*

As being affine is a rigid condition that often does not hold, one may consider the weaker notion of Gateaux differentiability. A pressure function $\Upsilon\colon \mathbf{B}(X) \to \mathbb{R}$ is *Gateaux differentiable at* $\varphi \in \boldsymbol{B}(X)$

if the directional pressure map $t \in \mathbb{R} \mapsto \Upsilon(\varphi + t\psi)$ is differentiable for every $\psi \in \mathbf{B}(X)$: that is, given $\psi \in \mathbf{B}(X)$, the limit

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

exists and is finite.

Theorem 3.2.7 (Corollary 4 of [6]) *A pressure function $\Upsilon\colon \boldsymbol{B}(X) \to \mathbb{R}$ is Gateaux differentiable at φ if and only if there exists a unique tangent functional in $\mathscr{T}_\varphi(\Upsilon)$.*

3.2.1 Variational principle *via* convex analysis

In this section we will prove Theorem A. It will be essentially an application of the previous results. Fix a compact metric space (X,d) and a continuous map $T\colon X \to X$ satisfying $\overline{\mathrm{mdim}}_M(X,d,T) < +\infty$.

Proposition 3.2.8 *The upper metric mean dimension map*

$$\varphi \in C^0(X) \quad\longmapsto\quad \overline{\mathrm{mdim}}_M(X,d,T,\varphi) = \limsup_{\varepsilon\to 0^+} \frac{P(X,d,T,\varphi,\varepsilon)}{\log(1/\varepsilon)}$$

is a pressure function.

Proof: The monotonicity is clear. For the convexity, let $\eta = t\,\varphi + (1-t)\,\psi$ for some $\varphi,\psi \in C^0(X)$ and $t \in (0,1)$. Given a finite set $E \subset X$, a positive integer $n \geq 1$ and $\varepsilon > 0$, by the Hölder Inequality we have

$$\sum_{x\in E} (1/\varepsilon)^{S_n\eta(x)} \leq \Big(\sum_{x\in E} (1/\varepsilon)^{S_n\varphi(x)}\Big)^t \Big(\sum_{x\in E} (1/\varepsilon)^{S_n\psi(x)}\Big)^{1-t}.$$

Taking the supremum over (n,ε)- separated sets $E \subset X$, we get

$$P(X,d,T,\eta,\varepsilon) \leq t\,P(X,d,T,\varphi,\varepsilon) + (1-t)\,P(X,d,T,\psi,\varepsilon).$$

To conclude, divide by $\log(1/\varepsilon)$ and make $\varepsilon \to 0^+$.

Regarding the translation invariance condition, given $\varphi \in C^0(X)$ and $c \in \mathbb{R}$, one has for every $\varepsilon > 0$

$$S(X,d,\varphi+c,\varepsilon) = \inf_{(U_i)_{1\leq i\leq n}} \left\{\sum_{i=1}^n (1/\varepsilon)^{\sup_{U_i}(\varphi+c)}\right\} = (1/\varepsilon)^c \inf_{(U_i)_{1\leq i\leq n}} \left\{\sum_{i=1}^n (1/\varepsilon)^{\sup_{U_i}\varphi}\right\}$$

and therefore,

$$\begin{aligned}
P(X,d,T,\varphi+c,\varepsilon) &= \lim_{n\to+\infty} \frac{1}{n} \log S(X,d_n,S_n(\varphi+c),\varepsilon)\\
&= \lim_{n\to+\infty} \frac{1}{n} \log S(X,d_n,S_n\varphi,\varepsilon) + \big(\log 1/\varepsilon\big)c\\
&= P(X,d,T,\varphi,\varepsilon) + \big(\log 1/\varepsilon\big)c.
\end{aligned}$$

Again conclude dividing by $\log(1/\varepsilon)$ and making $\varepsilon \to 0^+$. $\square$

Remark 3.2.9 *The main reason we are working with* upper *metric mean dimension is that it is* convex. *In the literature, many results concerning the upper metric mean dimension have their analogous counterparts phrased in terms of lower metric mean dimension. This is not the case here, since convexity is essential in our approach.*

Now that we have established that the upper metric mean dimension with potential is a pressure function, we can apply Theorem 3.2.2, thus obtaining $M = \mathfrak{h}$. Now, we proceed by showing that M is supported on invariant probability measures. First, we verify that every equilibrium state must be invariant.

Lemma 3.2.10 *Every equilibrium state μ_φ of $\varphi \in C^0(X)$ with respect to the variational principle* (3.5) *is invariant under T.*

Proof: Given $\varphi \in C^0(X)$, let $\mu_\varphi \in \mathscr{P}(X)$ be such that

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) = M(\mu_\varphi) + \int \varphi \, d\mu_\varphi.$$

We will show that $\int (\psi \circ T)\, d\mu_\varphi = \int \psi \, d\mu_\varphi$ for every $\psi \in C^0(X)$. Fix $\psi \in C^0(X)$ and let μ_1 and μ_2 be equilibrium states associated to the potentials $\varphi + \psi \circ T - \psi$ and $\varphi - \psi \circ T + \psi$, respectively. Then

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi + \psi \circ T - \psi) = M(\mu_1) + \int \varphi \, d\mu_1 + \int (\psi \circ T)\, d\mu_1 - \int \psi \, d\mu_1$$

and

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi - \psi \circ T + \psi) = M(\mu_2) + \int \varphi \, d\mu_2 - \int (\psi \circ T)\, d\mu_2 + \int \psi \, d\mu_2.$$

Moreover, by Proposition 2.3.1 (e), we know that

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) = \overline{\mathrm{mdim}}_M(X,d,T,\varphi + \psi \circ T - \psi) = \overline{\mathrm{mdim}}_M(X,d,T,\varphi - \psi \circ T + \psi).$$

Therefore,

$$\begin{aligned} M(\mu_\varphi) + \int \varphi \, d\mu_\varphi &= M(\mu_1) + \int \varphi \, d\mu_1 + \int (\psi \circ T)\, d\mu_1 - \int \psi \, d\mu_1 \\ &\geq M(\mu_\varphi) + \int \varphi \, d\mu_\varphi + \int (\psi \circ T)\, d\mu_\varphi - \int \psi \, d\mu_\varphi \end{aligned}$$

which yields $\int (\psi \circ T)\, d\mu_\varphi \leq \int \psi \, d\mu_\varphi$. Analogously,

$$\begin{aligned} M(\mu_\varphi) + \int \varphi \, d\mu_\varphi &= M(\mu_2) + \int \varphi d\mu_2 - \int (\psi \circ T)\, d\mu_2 + \int \psi \, d\mu_2 \\ &\geq M(\mu_\varphi) + \int \varphi \, d\mu_\varphi - \int (\psi \circ T)\, d\mu_\varphi + \int \psi \, d\mu_\varphi \end{aligned}$$

so $\int (\psi \circ T)\, d\mu_\varphi \geq \int \psi \, d\mu_\varphi$. The proof is complete. Hence, the maximum in the variational principle provided by Theorem A can be computed on the set of T-invariant probability measures. $\square$

The previous lemma implies that, if we consider the map $\tau\colon \mathscr{P}(X) \to \mathbb{R}\cup\{-\infty\}$ given by

$$\tau(\mu) = \begin{cases} M(\mu) & \text{if } \mu \in \mathscr{P}_T(X) \\ -\infty & \text{otherwise} \end{cases},$$

we obtain a concave, upper semicontinuous map satisfying

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) = \max_{\mu\in\mathscr{P}(X)} \left\{\tau(\mu) + \int \varphi\, d\mu\right\} \qquad \forall\, \varphi \in C^0(X).$$

By Proposition 3.1.1, $\tau(\mu) = M(\mu)$ for every $\mu \in \mathscr{P}(X)$ and we get the following

Proposition 3.2.11 *For every non-invariant probability measure $\mu \in \mathscr{P}(X)\setminus\mathscr{P}_T(X)$, we have*

$$M(\mu) = -\infty.$$

The final piece of Theorem A is the maximality (3.7).

Lemma 3.2.12 *Let $\mathscr{A} \subseteq \mathscr{P}(X)$. If a function $\tau\colon \mathscr{A} \to \mathbb{R}\cup\{-\infty\}$ satisfies*

$$\sup_{\mu\in\mathscr{A}} \left\{\tau(\mu) + \int \varphi\, d\mu\right\} \le \overline{\mathrm{mdim}}_M(X,d,T,\varphi) \quad \forall\, \varphi \in C^0(X),$$

then

$$\tau(\mu) \le M(\mu) \quad \forall\, \mu \in \mathscr{A}.$$

Proof: We have

$$\tau(\mu) - \int \varphi\, d\mu \le \overline{\mathrm{mdim}}_M(X,d,T,-\varphi) \qquad \forall\, \mu \in \mathscr{A},\ \forall\, \varphi \in C^0(X),$$

which, since $C^0(X)$ is a vector space, implies that for every $\mu \in \mathscr{A}$,

$$\begin{aligned}\tau(\mu) &\le \inf_{\varphi\in C^0(X)} \left\{\overline{\mathrm{mdim}}_M(X,d,T,-\varphi) + \int \varphi\, d\mu\right\} \\ &= \inf_{\varphi\in C^0(X)} \left\{\overline{\mathrm{mdim}}_M(X,d,T,\varphi) - \int \varphi\, d\mu\right\} \\ &= M(\mu).\end{aligned}$$

□

3.3 Scaled formalism

Let (X,d) be a compact metric space and $T\colon X\to X$ be a continuous map such that $\overline{\mathrm{mdim}}_M(X,d,T) < +\infty$. In this section we will study the map

$$\begin{array}{rccc} H^K_\delta\colon & \mathscr{P}_T(X) & \to & \mathbb{R} \\ & \mu & \mapsto & \sup_{(\mu_\varepsilon)_\varepsilon\in\mathscr{M}(\mu)} \limsup_{\varepsilon\to 0^+} \frac{h^K_{\mu_\varepsilon}(\varepsilon,\delta)}{\log(1/\varepsilon)} \end{array}, \tag{3.16}$$

where $\mathscr{M}(\mu)$ stands for the space of sequences of probability measures in $\mathscr{P}_T(X)$ which converge to μ in the weak*-topology, and $h^K_\nu(\varepsilon,\delta)$ denotes the δ-Katok entropy of $\nu \in \mathscr{P}_T(X)$ at scale ε, as defined in Section 2.2.

Remark 3.3.1 A priori, *the map H^K_δ may depend on δ, though in several examples its value is independent of this parameter (cf. Section 3.6). Moreover, given $\mu \in \mathscr{P}_T(X)$,*

$$\lim_{\delta\to 0^+} H^K_\delta(\mu) = \sup_{\delta\in(0,1)} H^K_\delta(\mu) \tag{3.17}$$

since the map $\delta \in (0,1) \mapsto H^K_\delta(\mu)$ is non-increasing.

To introduce a potential in the previous concept, given a continuous map $\varphi\colon X \to \mathbb{R}$ and $\mu \in \mathscr{E}_T(X)$, take

$$N_\mu(\varphi,\varepsilon,\delta,n) = \inf\{S(A,d_n,S_n\varphi,\varepsilon)\colon\ \mu(A) > 1-\delta\}$$

and

$$P^K_\mu(\varphi,\varepsilon,\delta) = \limsup_{n\to+\infty} \frac{1}{n}\log N_\mu(\varphi,\varepsilon,\delta,n).$$

Similarly, we now extend the previous notion to $\mu \in \mathscr{P}_T(X)$ using its ergodic decomposition, that is,

$$P^K_\mu(\varphi,\varepsilon,\delta) = \int_{\mathscr{E}_T(X)} P^K_m(\varphi,\varepsilon,\delta)\,d\mathbb{P}_\mu(m)$$

if $\mu = \int_{\mathscr{E}_T(X)} m\,d\mathbb{P}_\mu(m) \in \mathscr{P}_T(X)$.

Remark 3.3.2 *Given $\delta \in (0,1)$ and $\mu \in \mathscr{E}_T(X)$, one has $N_\mu(\varphi,\varepsilon,\delta,n) \le S(X,d_n,S_n\varphi,\varepsilon)$ for every $n \in \mathbb{N}$ and $\varepsilon > 0$; so $P^K_\mu(\varphi,\varepsilon,\delta)$ is bounded from above by $P(X,d,T,\varphi,\varepsilon)$.*

The following connection between h^K_μ and P^K_μ is a straightforward adaptation of Proposition 2.2 in [15].

Lemma 3.3.3 *Let (X,d) be a compact metric space, $T\colon X\to X$ be a continuous map, $\varphi\colon X\to\mathbb{R}$ be a continuous potential and $\mu\in\mathscr{E}_T(X)$ be an ergodic probability measure. Then, given $\tau > 0$ and $1 > \delta_1 > \delta_2 > \delta_3 > 0$, one has for every $0 < \varepsilon < \tau$*

$$\frac{h^K_\mu(\varepsilon,\delta_1)}{\log(1/\varepsilon)} + \int \varphi\,d\mu - \tau \le \frac{P^K_\mu(\varphi,\varepsilon,\delta_2)}{\log(1/\varepsilon)} \le \frac{h^K_\mu(\varepsilon,\delta_3)}{\log(1/\varepsilon)} + \int \varphi\,d\mu + \tau. \tag{3.18}$$

This lemma has a main consequence: having fixed $\delta \in (0,1)$, if we define

$$\begin{array}{rccc} P^K_{\varphi,\delta}\colon & \mathscr{P}_T(X) & \to & \mathbb{R} \\ & \mu & \mapsto & \sup_{(\mu_\varepsilon)_\varepsilon\in\mathscr{M}(\mu)} \limsup_{\varepsilon\to 0^+} \frac{P^K_{\mu_\varepsilon}(\varphi,\varepsilon,\delta)}{\log(1/\varepsilon)} \end{array} \tag{3.19}$$

then, taking lim sup in (3.18), one gets

$$H^K_{\delta_1}(\mu) + \int \varphi\,d\mu \le P^K_{\varphi,\delta_2}(\mu) \le H^K_{\delta_3}(\mu) + \int \varphi\,d\mu \qquad \forall \mu \in \mathscr{P}_T(X).$$

Therefore, noticing that (see Remark 3.3.2)

$$P^K_{\varphi,\delta} \leq \overline{\mathrm{mdim}}_M(X,d,T,\varphi) \qquad \forall\,\delta \in (0,1) \quad \forall\,\varphi \in C^0(X)$$

we conclude that:

Lemma 3.3.4 *For every $\delta \in (0,1)$ and $\mu \in \mathscr{P}_T(X)$, one has*

$$H^K_\delta(\mu) \leq \inf_{\varphi\in C^0(X)} \left\{\overline{\mathrm{mdim}}_M(X,d,T,\varphi) - \int \varphi\, d\mu\right\} = M(\mu).$$

3.3.1 Basic properties

For future use, let us state some properties of the map H^K_δ.

Lemma 3.3.5 *For every $\delta \in (0,1)$, the map H^K_δ is upper semi-continuous.*

Proof: Consider $\mu \in \mathscr{P}_T(X)$ and a sequence $(\mu_n)_{n\in\mathbb{N}}$ in $\mathscr{P}_T(X)$ converging, in the weak*-topology, to μ. Given $n \in \mathbb{N}$, for every $\eta > 0$ there is a sequence $(\mu^{(n)}_{\varepsilon_k})_{k\in\mathbb{N}}$ in $\mathscr{M}(\mu_n)$ such that

$$H^K_\delta(\mu_n) - \eta < \limsup_{k\to+\infty} \frac{h^K_{\mu_{\varepsilon_k}}(\varepsilon_k,\delta)}{\log(1/\varepsilon_k)} \leq H^K_\delta(\mu_n).$$

Moreover, the sequence $(\mu^{(n)}_{\varepsilon_n})_{n\in\mathbb{N}}$ belongs to $\mathscr{M}(\mu)$. Thus,

$$H^K_\delta(\mu_n) < \limsup_{n\to+\infty} \frac{h^K_{\mu_{\varepsilon_n}}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} + \eta \leq H^K_\delta(\mu) + \eta.$$

So, for every $\eta > 0$,

$$\limsup_{n\to+\infty} H^K_\delta(\mu_n) \leq H^K_\delta(\mu) + \eta$$

which implies that

$$\limsup_{n\to+\infty} H^K_\delta(\mu_n) \leq H^K_\delta(\mu).$$

□

Lemma 3.3.6 *Let $T\colon X \to X$ and $T'\colon X' \to X'$ be continuous maps on compact metric spaces (X,d) and (X',d'), and $\Phi\colon X \to X'$ be a bi-Lipschitz conjugacy between them. Then,*

$$H^K_\delta(X,d,T,\mu) = H^K_\delta(X',d',T',\Phi_*(\mu)) \qquad \forall\,\mu \in \mathscr{P}_T(X).$$

Proof: We start by recalling a few simple facts about the push-forward map $\mu \mapsto \Phi_*\mu = \mu \circ \Phi^{-1}$:
(a) A sequence $(\mu_\varepsilon)_\varepsilon$ in $\mathscr{P}_T(X)$ converges to $\mu \in \mathscr{P}_T(X)$ if and only if $\Phi_*\mu_\varepsilon$ in $\mathscr{P}_{T'}(X')$ converges to $\Phi_*\mu \in \mathscr{P}_{T'}(X')$.
(b) $m \in \mathscr{E}_T(X)$ if and only if $\Phi_* m \in \mathscr{E}_{T'}(X')$.

(c) $\nu = \int m \, d\mathbb{P}_\nu(m)$ is the ergodic decomposition of $\nu \in \mathscr{P}_T(X)$ if and only if $\Phi_*\nu = \int \Phi_*(m) \, d\mathbb{P}_\nu(m)$ is the ergodic decomposition of $\Phi_*\nu \in \mathscr{P}_{T'}(X')$.

Take now constants $C_1, C_2 > 0$ such that

$$C_1 \, d(x,x') \le d'(\Phi(x),\Phi(x')) \le C_2 \, d(x,y) \qquad \forall x, x' \in X.$$

Then, by considering separated subsets, we get for any $A \subset X$

$$S(A,d_n,0,\varepsilon) \le S(\Phi(A),d'_n,0,C_1\varepsilon) \le S(A,d_n,0,C_3\varepsilon)$$

where $C_3 = C_1/C_2$. Thus, for every $m \in \mathscr{E}_T(X)$,

$$\begin{aligned} N_m(\varepsilon,\delta,n) &= \inf\{S(A,d_n,0,\varepsilon) \,:\, m(A) > 1-\delta\} \\ &\le \inf\{S(\Phi(A),d'_n,0,C_1\varepsilon) \,:\, \Phi_*m(\Phi(A)) > 1-\delta\} \\ &= N_{\Phi_*m}(C_1\varepsilon,\delta,n) \\ &\le \inf\{S(A,d_n,0,C_3\varepsilon) \,:\, m(A) > 1-\delta\} \\ &= N_m(C_3\varepsilon,\delta,n). \end{aligned}$$

Hence,

$$h^K_m(\varepsilon,\delta) \le h^K_{\Phi_*m}(C_1\varepsilon,\delta) \le h^K_m(C_3\varepsilon,\delta).$$

Consequently, for every $\nu \in \mathscr{P}_T(X)$ with ergodic decomposition given by $\nu = \int m \, d\mathbb{P}_\nu(m)$ we have

$$\begin{aligned} h^K_\nu(\varepsilon,\delta) &= \int h^K_m(\varepsilon,\delta) \, d\mathbb{P}_\nu(m) \\ &\le \int h^K_{\Phi_*m}(C_1\varepsilon,\delta) \, d\mathbb{P}_\nu(m) \\ &= h^K_{\Phi_*\nu}(C_1\varepsilon,\delta) \\ &\le \int h^K_m(C_3\varepsilon,\delta) \, d\mathbb{P}_\nu(m) \\ &= h^K_\nu(C_3\varepsilon,\delta). \end{aligned}$$

Thus,

$$h^K_\nu(\varepsilon,\delta) \le h^K_{\Phi_*\nu}(C_1\varepsilon,\delta) \le h^K_\nu(C_3\varepsilon,\delta).$$

We now proceed to evaluate H^K_δ. Given $\mu \in \mathscr{P}_T(X)$, take sequences $\mu_n \to \mu$ and $\varepsilon_n \to 0^+$ such that

$$H^K_\delta(\mu) = \lim_{n\to+\infty} \frac{h^K_{\mu_n}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)}.$$

Then,

$$H^K_\delta(\mu) \le \limsup_{n\to+\infty} \frac{h^K_{\Phi_*\mu_n}(C_1\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} \le H^K_\delta(\Phi_*\mu).$$

For the converse inequality, take sequences $\mu_n \to \mu$ and $\varepsilon_n \to 0^+$ such that

$$H_\delta^K(\Phi_*\mu) = \lim_{n\to+\infty} \frac{h^K_{\Phi_*\mu_n}(C_1\varepsilon_n,\delta)}{\log(1/C_1\varepsilon_n)}.$$

Then,

$$H_\delta^K(\Phi_*\mu) \le \limsup_{n\to+\infty} \frac{h^K_{\mu_n}(C_3\varepsilon_n,\delta)}{\log(1/C_1\varepsilon_n)} \le H_\delta^K(\mu).$$

□

It is equally straightforward to show the following power rule for the map H_δ^K.

Lemma 3.3.7 *Let (X,d,T) be as above and $p\in\mathbb{N}$. Then*

$$H_\delta^K(X,d,T^p,\mu) \le pH_\delta^K(X,d,T,\mu) \qquad \forall \mu\in\mathscr{P}_T(X).$$

3.3.2 Limits of scaled equilibrium states

Let's now prove the variational principle in Theorem B.

Fix $\delta\in(0,1)$. The upper bound follows immediately from Lemma 3.3.4 and Theorem A:

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) = \max_{\mu\in\mathscr{P}_T(X)} \Big\{M(\mu)+\int\varphi\,d\mu\Big\} \ge \max_{\mu\in\mathscr{P}_T(X)} \Big\{H_\delta^K(\mu)+\int\varphi\,d\mu\Big\}. \tag{3.20}$$

We are left to show the other inequality. For that, we proceed by constructing $\mu_0\in\mathscr{P}_T(X)$ such that

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) \le H_\delta^K(\mu_0)+\int\varphi\,d\mu_0.$$

It was proved in [16] that, for every $\delta\in(0,1)$ and every non-negative $\varphi\in C^0(X)$,

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) = \limsup_{\varepsilon\to0^+}\sup_{\mu\in\mathscr{E}_T(X)} \frac{P_\mu^K(\varphi,\varepsilon,\delta)}{\log(1/\varepsilon)}. \tag{3.21}$$

We note that the translation invariance property on both sides of the previous equality allows us to drop the condition $\varphi\ge0$. Thus, given $\varphi\in C^0(X)$, applying Lemma 3.3.3 and (3.21) we obtain, for every $1>\delta_1>\delta_3>0$,

$$\limsup_{\varepsilon\to0^+}\sup_{\mu\in\mathscr{E}_T(X)} \frac{h_\mu^K(\varepsilon,\delta_1)}{\log(1/\varepsilon)}+\int\varphi\,d\mu \le \overline{\mathrm{mdim}}_M(X,d,T,\varphi) \le \limsup_{\varepsilon\to0^+}\sup_{\mu\in\mathscr{E}_T(X)} \frac{h_\mu^K(\varepsilon,\delta_3)}{\log(1/\varepsilon)}+\int\varphi\,d\mu.$$

These inequalities imply that, for every $\delta\in(0,1)$,

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) = \limsup_{\varepsilon\to0^+}\sup_{\mu\in\mathscr{E}_T(X)} \frac{h_\mu^K(\varepsilon,\delta)}{\log(1/\varepsilon)}+\int\varphi\,d\mu.$$

Therefore, we may find a sequence $(\varepsilon_n)_n$ of positive real numbers converging to 0 and such that

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) = \lim_{n\to+\infty} \sup_{\mu\in\mathscr{E}_T(X)} \frac{h^K_\mu(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} + \int \varphi\, d\mu.$$

Thus, given $\eta > 0$, there is $p \in \mathbb{N}$ such that, for every $n \geq p$,

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) - \eta < \sup_{\mu\in\mathscr{E}_T(X)} \frac{h^K_\mu(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} + \int \varphi\, d\mu$$

and so, for every $n \geq p$ there exists $\mu_{\varepsilon_n} \in \mathscr{E}_T(X)$ satisfying

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) - \eta < \frac{h^K_{\mu_{\varepsilon_n}}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} + \int \varphi\, d\mu_{\varepsilon_n}.$$

Taking a subsequence if necessary, we may assume that the sequence $\left(\mu_{\varepsilon_n}\right)_{n\geq p}$ converges in the weak*-topology to $\mu_0 \in \mathscr{P}_T(X)$. Then, by (3.16), for every $\eta > 0$,

$$H^K_\delta(\mu_0) \geq \limsup_{n\to+\infty} \frac{h^K_{\mu_{\varepsilon_n}}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} \geq \overline{\mathrm{mdim}}_M(X,d,T,\varphi) - \int \varphi\, d\mu_0 - \eta. \tag{3.22}$$

Hence,

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) \leq H^K_\delta(\mu_0) + \int \varphi\, d\mu_0.$$

The proof of Theorem B is complete.

Remark 3.3.8 *From* (3.20) *and* (3.22) *we conclude that, up to a subsequence, the equilibrium state* μ_0 *satisfies*

$$H^K_\delta(\mu_0) = \lim_{n\to+\infty} \frac{h^K_{\mu_{\varepsilon_n}}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)}.$$

We finish this subsection observing that if one defines H^K_δ without considering measures varying with the scale, say

$$\tilde{H}_\delta\colon\ \mu \quad\mapsto\quad \limsup_{\varepsilon\to 0^+} \frac{h^K_\mu(\varepsilon,\delta)}{\log(1/\varepsilon)},$$

then the equality (3.9) may not be valid: see Example 3.6.3 and Remark 3.6.4. This is related to the existence of probability measures capturing the dynamical complexity at all scales, which is a too strong requirement (see Section VIII in [41]). However, in some cases those probabilities *do* exist, as happens in Example 3.6.2.

3.3.3 Other entropies

The reasoning to prove Theorem B may be extended to maps

$$F\colon\ \mathscr{P}_T(X) \times (0,1) \to \mathbb{R}$$

satisfying

$$\overline{\text{mdim}}_M(X,d,T,\varphi) = \limsup_{\varepsilon\to 0^+} \sup_{\mu \in \mathscr{P}_T(X)} \left\{ \frac{F(\mu,\varepsilon)}{\log(1/\varepsilon)} + \int \varphi\, d\mu \right\}. \tag{3.23}$$

Actually, if we define

$$\begin{array}{rccc} F^*: & \mathscr{P}_T(X) & \to & \mathbb{R} \\ & \mu & \mapsto & \sup_{(\mu_\varepsilon)_\varepsilon \in \mathscr{M}(\mu)} \limsup_{\varepsilon\to 0^+} \frac{F(\mu_\varepsilon,\varepsilon)}{\log(1/\varepsilon)} \end{array}$$

where $\mathscr{M}(\mu)$ is the space of sequences of probability measures in $\mathscr{P}_T(X)$ which converge to μ in the weak*-topology, then the previous argument also shows that

$$\overline{\text{mdim}}_M(X,d,T,\varphi) = \max_{\mu \in \mathscr{P}_T(X)} \left\{ F^*(\mu) + \int \varphi\, d\mu \right\}.$$

For instance, F may be any of the examples referred to in (3.1), (3.2) and (3.3).

Remark 3.3.9 *Example 3.6.3 provides a negative answer to the question of whether we can exchange the order of* $\limsup_\varepsilon$ *and* $\sup_\mu$ *in* (3.23)*. However, Theorem B shows that the regularization F^* of F is sufficient to surpass the obstructions to such an exchange of order.*

3.4 Ergodic optimization for shifts

In this section we prove Theorem C. The estimates will be done for unilateral sequences, though they are also valid for bilateral ones. Let (X,d) be a compact metric space such that $\dim_B U = \dim_B X$ for every nonempty open set $U \subset X$. By Theorem A we have

$$\overline{\text{mdim}}_M(X^{\mathbb{N}},d_\rho,\sigma,\varphi) = \max_{\mu \in \mathscr{P}_\sigma(X^{\mathbb{N}})} \left\{ M(\mu) + \int \varphi\, d\mu \right\} \qquad \forall\, \varphi \in C^0(X^{\mathbb{N}}).$$

Since $\overline{\text{mdim}}_M(X^{\mathbb{N}},d_\rho,\sigma) = \overline{\dim}_B X$ (cf. [59, Theorem 5]), Theorem A also indicates that

$$\max_{\mu \in \mathscr{P}_\sigma(X^{\mathbb{N}})} M(\mu) = \overline{\dim}_B X.$$

Thus,

$$\overline{\text{mdim}}_M(X^{\mathbb{N}},d_\rho,\sigma,\varphi) \le \overline{\dim}_B X + \max_{\mu \in \mathscr{P}_\sigma(X^{\mathbb{N}})} \int \varphi\, d\mu \qquad \forall\, \varphi \in C^0(X^{\mathbb{N}}). \tag{3.24}$$

For the converse inequality, recall from Lemma 3.3.4 that, given $\delta > 0$, we have

$$H^K_\delta(\mu) \le M(\mu) \le \overline{\text{mdim}}_M(X^{\mathbb{N}},d_\rho,\sigma) = \overline{\dim}_B X \qquad \forall \mu \in \mathscr{P}_\sigma(X^{\mathbb{N}}). \tag{3.25}$$

Theorem 3.4.1 *Let (X,d) be a compact metric space such that* $\dim_B U = \dim_B X$ *for every nonempty open set $U \subset X$. Then, for every $\delta \in (0,1)$,*

$$H^K_\delta(\mu) = \dim_B X \qquad \forall\, \mu \in \mathscr{P}_\sigma(X^{\mathbb{N}}).$$

Consequently,

$$\overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d_\rho, \sigma, \varphi) \geq \dim_B X + \max_{\mu \in \mathscr{P}_\sigma(X^{\mathbb{N}})} \int \varphi \, d\mu \qquad \forall \varphi \in C^0(X^{\mathbb{N}}) \tag{3.26}$$

and, bringing (3.24) and (3.26) together, the proof of Theorem C is complete, up to the proof of Theorem 3.4.1.

Proof of Theorem 3.4.1: To show that

$$H^K_\delta(\mu) \geq \dim_B X \qquad \forall \mu \in \mathscr{P}_\sigma(X^{\mathbb{N}})$$

we start relating the map H^K_δ with the local box dimension of the alphabet set X.

Lemma 3.4.2 *For every $\delta \in (0,1)$ and $\rho > 1$, given a fixed point $(y, y, \cdots) \in X^{\mathbb{N}}$ one has*

$$\lim_{\gamma \to 0^+} \overline{\dim}_B B_\gamma(y) \leq H^K_\delta(\delta^{\mathbb{N}}_{\{y\}})$$

where $\delta^{\mathbb{N}}_{\{y\}} \in \mathscr{P}_\sigma(X^{\mathbb{N}})$ is the product measure of the Dirac measure supported on y and $B_\gamma(y)$ stands for the open ball in X centered at y with radius γ in the metric d_ρ.

Proof: Fix $\delta \in (0,1)$ and $\rho > 1$. Given $\varepsilon > 0$ and $\gamma > 0$, let $E = E(\varepsilon) \subset B_\gamma(y)$ be a maximal ε-separated subset of $B_\gamma(y)$. Take

$$\nu^\gamma_\varepsilon = \frac{1}{\#E} \sum_{e \in E} \delta_{\{e\}} \qquad \text{and} \qquad \mu^\gamma_\varepsilon = (\nu^\gamma_\varepsilon)^{\mathbb{N}}. \tag{3.27}$$

Then, clearly any two elements in the n-cylinder

$$z, w \in E \times \cdots \times E \times X \times X \times \cdots \subset X^{\mathbb{N}}$$

which differ at some of the first n coordinates satisfy $d_n(z,w) \geq \varepsilon$. Let $L = L(n, \varepsilon, \delta)$ be the maximal positive integer such that $L(\#E)^{-n} < \delta$, where $\#E$ stands for the cardinal number of E. Then, any subset $A \subset X^{\mathbb{N}}$ satisfying

$$\mu^\gamma_\varepsilon(A) > 1 - \delta$$

must contain an (n, ε)-separated subset $\mathscr{F}$ whose cardinal number satisfies

$$\#\mathscr{F} \geq (\#E)^n - L > (1-\delta)(\#E)^n$$

since $X^{\mathbb{N}} \setminus A$ can contain at most L sub-cylinders $\{y_1\} \times \cdots \times \{y_n\} \times Y \times Y \times \cdots$. Thus,

$$h^K_{\mu^\gamma_\varepsilon}(\varepsilon, \delta) \geq \log \#E.$$

Let $(\varepsilon^\gamma_n)_n$ be a sequence of positive real numbers converging to 0 such that

- $\overline{\dim}_B B_\gamma(y) = \lim_{n \to +\infty} \frac{\log \#E(\varepsilon^\gamma_n)}{\log(1/\varepsilon^\gamma_n)}$;

- $\lim_{n\to+\infty} \mu^{\gamma}_{\varepsilon^{\gamma}_n} = \mu^{\gamma} \in \mathscr{P}(X^{\mathbb{N}})$ in the weak*-topology.

Then,

$$H^K_\delta(\mu^\gamma) \geq \limsup_{n\to+\infty} \frac{h^K_{\mu^{\gamma}_{\varepsilon^{\gamma}_n}}(\varepsilon^{\gamma}_n,\delta)}{\log(1/\varepsilon^{\gamma}_n)} \geq \limsup_{n\to+\infty} \frac{\log \#E(\varepsilon^{\gamma}_n)}{\log(1/\varepsilon^{\gamma}_n)} = \overline{\dim}_B B_\gamma(y).$$

As $\mathrm{supp}(\mu^\gamma) \subset B_\gamma(y)^{\mathbb{N}}$, we also have

$$\lim_{\gamma\to 0^+} \mu^\gamma = \delta^{\mathbb{N}}_{\{y\}}$$

so, by the upper semi-continuity of H^K_δ, we get

$$H^K_\delta(\delta^{\mathbb{N}}_{\{y\}}) \geq \limsup_{\gamma\to 0^+} H^K_\delta(\mu^\gamma) \geq \lim_{\gamma\to 0^+} \overline{\dim}_B B_\gamma(y).$$

□

The next proposition generalizes the previous information.

Proposition 3.4.3 *Given $\delta \in (0,1)$ and $\rho > 1$, for every periodic point by σ with minimal period p, say $\xi = (y_1,\cdots,y_p,y_1,\cdots,y_p,\cdots) \in X^{\mathbb{N}}$, the Dirac periodic probability measure $\mu_\xi \in \mathscr{P}_\sigma(X^{\mathbb{N}})$ supported on the orbit of ξ satisfies*

$$\frac{1}{p} \lim_{\gamma\to 0^+} \overline{\dim}_B B^{\gamma}_p(\xi) \leq H^K_\delta(\mu_\xi)$$

where $B^{\gamma}_p(y) = B_\gamma(y_1)\times\cdots\times B_\gamma(y_p) \subset Y^p$ and $B_\gamma(a)$ stands for the open ball in X centered at a with radius γ.

Proof: Endow the space X^p with the metric

$$d^{\max}\big((z_1,z_2,\cdots,z_p),(w_1,w_2,\cdots,w_p)\big) = \max\big\{d(z_1,w_1),\cdots,d(z_p,w_p)\big\}$$

and consider in $(X^p)^{\mathbb{N}}$ the metric $d^{\max}_{\rho^p}$ given by Definition 2.4 when applied to the metric space $(X^p, d^{\max})$ and the value ρ^p:

$$d^{\max}_{\rho^p}(\alpha,\beta) = \sup_{n\in\mathbb{N}\cup\{0\}} \frac{\max_{1\leq i\leq p} d(\alpha_{i+np}-\beta_{i+np})}{\rho^{pn}}.$$

Then the map

$$\begin{array}{rccc} \Phi\colon & \left(X^{\mathbb{N}}, d_\rho, \sigma^p\right) & \to & \left((X^p)^{\mathbb{N}}, d^{\max}_{\rho^p}, \sigma\right) \\ & (z_1,z_2,\cdots) & \mapsto & \big((z_1,\cdots,z_p),(z_{p+1},\cdots,z_{2p}),\cdots\big) \end{array}$$

satisfies $\Phi\circ\sigma^p=\sigma\circ\Phi$ and is bi-Lipschitz. In fact, given $z=(z_1,z_2,\cdots)$ and $w=(w_1,w_2,\cdots)$ in $X^{\mathbb{N}}$, one has

$$\begin{aligned} d^{\max}_{\rho^p}(\Phi(z),\Phi(w)) &= \sup_{n\in\mathbb{N}\cup\{0\}} \frac{\max_{1\le i\le p} d(z_{i+np}-w_{i+np})}{\rho^{pn}} \\ &\le \rho^{p-1}\sup_{n\in\mathbb{N}\cup\{0\}}\max_{1\le i\le p}\frac{|z_{i+np}-w_{i+np}|}{\rho^{i+np-1}} \\ &= \rho^{p-1}d_\rho(z,w) \end{aligned}$$

and

$$\begin{aligned} d^{\max}_{\rho^p}(\Phi(z),\Phi(w)) &= \sup_{n\in\mathbb{N}\cup\{0\}} \frac{\max_{1\le i\le p} d(z_{i+np}-w_{i+np})}{\rho^{pn}} \\ &\ge \sup_{n\in\mathbb{N}\cup\{0\}}\max_{1\le i\le p}\frac{d(z_{i+np}-w_{i+np})}{\rho^{i+pn-1}} \\ &= d_\rho(z,w). \end{aligned}$$

Now, take a periodic point $\xi=(y_1,\cdots,y_p,y_1,\cdots,y_p,\cdots)\in X^{\mathbb{N}}$ by σ, with minimal period p, and let μ_ξ be the Dirac periodic probability measure supported on the orbit of ξ. We notice that, for each $i\in\{1,\cdots,p\}$, the push-forward $\Phi_*(\delta_{\{\sigma^i(\xi)\}})$ is the Dirac measure $\delta_{\{\Phi(\sigma^i(\xi))\}}$ in $(X^p)^{\mathbb{N}}$ supported on the fixed point

$$\Phi(\sigma^i(\xi)) = \Big((y_{1+i},\cdots,y_p,y_1,\cdots,y_i),(y_{1+i},\cdots,y_p,y_1,\cdots,y_i),\cdots\Big).$$

For every $\gamma>0$ and $\varepsilon>0$, consider the probability measure $\mu^\gamma_\varepsilon(i)$ in $(X^p)^{\mathbb{N}}$ defined in (3.27), within the proof of Lemma 3.4.2, but now with respect to the fixed point $\Phi(\sigma^i(\xi))$ by σ. Then,

$$h^K_{\mu^\gamma_\varepsilon(i)}(\varepsilon,\delta) \ge \log\#E_i(\varepsilon)$$

where $E_i(\varepsilon)$ is a maximal ε-separated subset of

$$B^\gamma_i = B_\gamma(y_{i+1},\cdots,y_p,y_1,\cdots,y_i) = B_\gamma(y_{i+1})\times\cdots\times B_\gamma(y_p)\times B_\gamma(y_1)\times\cdots\times B_\gamma(y_i).$$

Therefore,

$$h^K_{\frac{1}{p}\sum_{i=0}^{p-1}\mu^\gamma_\varepsilon(i)}(\varepsilon,\delta) = \frac{1}{p}\sum_{i=0}^{p-1}h^K_{\mu^\gamma_\varepsilon(i)}(\varepsilon,\delta) \ge \frac{1}{p}\sum_{i=0}^{p-1}\log\#E_i(\varepsilon).$$

We now observe that $B^\gamma_i(\xi)$ is given by permutations in the coordinates of $B^\gamma_p(\xi)$ for every $i=1,...,p$. Hence they all have the same upper box dimension, which can be estimated using the same sequence of scales, say $(\varepsilon^\gamma_n)_n$. Assume (by taking a subsequence if necessary) that those scales are such that the sequence $\big(\mu^\gamma_{\varepsilon^\gamma_n}(i)\big)_n$ converges in the weak*-topology, say

$$\lim_{n\to+\infty}\mu^\gamma_{\varepsilon^\gamma_n}(i) = \mu^\gamma(i) \qquad \forall\, i\in\{1,\cdots,p\}.$$

Consequently,

$$H^K_\delta\Big(\frac{1}{p}\sum_{i=1}^{p}\mu^\gamma(i)\Big) \geq \limsup_{n\to+\infty}\frac{h^K_{\frac{1}{p}\sum_{i=0}^{p-1}\mu^\gamma_{\varepsilon^\gamma_n}(i)}(\varepsilon^\gamma_n,\delta)}{\log(1/\varepsilon^\gamma_n)} \geq \limsup_{n\to+\infty}\frac{1}{p}\sum_{i=1}^{p}\frac{\log\#E_i(\varepsilon^\gamma_n)}{\log(1/\varepsilon^\gamma_n)} = \overline{\dim}_B B^\gamma_p(\xi).$$

Moreover, as $\mathrm{supp}(\mu^\gamma(i)) \subset B^\gamma_i$ for every i,

$$\lim_{\gamma\to 0^+}\frac{1}{p}\sum_{i=1}^{p}\mu^\gamma(i) = \frac{1}{p}\sum_{i=1}^{p}\delta_{\{\Phi(\sigma^i(\xi))\}} = \frac{1}{p}\sum_{i=1}^{p}\Phi_*\delta_{\{\sigma^i(\xi)\}} = \Phi_*(\mu_\xi).$$

Finally, by the upper semi-continuity of H^K_δ,

$$H^K_\delta(\Phi_*(\mu_\xi)) \geq \lim_{\gamma\to 0^+}\overline{\dim}_B B^\gamma_p(\xi).$$

To complete the proof of Proposition 3.4.3, we summon Lemmas 3.3.6 and 3.3.7 to deduce that

$$\lim_{\gamma\to 0^+}\dim_B B^\gamma_p(x) \leq H^K_\delta\Big((X^p)^{\mathbb{N}}, d^{\max}_{\rho^p}, \sigma, \Phi_*\mu_\xi\Big) = H^K_\delta\Big(X^{\mathbb{N}}, d_\rho, \sigma^p, \mu_\xi\Big) \leq p H^K_\delta\Big(X^{\mathbb{N}}, d_\rho, \sigma, \mu_\xi\Big).$$

□

Let us resume the proof of Theorem 3.4.1. Under the assumption that every nonempty open set in $U \subset X$ satisfies $\dim_B U = \dim_B X$, we know that for every $\gamma > 0$,

$$\dim_B B^\gamma_p(\xi) = \dim_B\Big(B_\gamma(y_1)\times\cdots\times B_\gamma(y_p)\Big) = \sum_{i=1}^{p}\dim_B B_\gamma(y_i) = p\dim_B X.$$

Therefore, by Proposition 3.4.3, for every periodic probability measure μ_ξ supported on a periodic point $\xi = (y_1,\cdots,y_p,y_1,\cdots)$ with minimal period p, we have

$$\dim_B X = \frac{1}{p}\lim_{\gamma\to 0^+}\dim_B B^\gamma_p(\xi) \leq H^K_\delta\Big(X^{\mathbb{N}}, d_\rho, \sigma, \mu_\xi\Big).$$

In addition, by the specification property of the shift map (cf. [54, Proposition 2]), the set of periodic probability measures is dense in $\mathscr{P}_\sigma(X^{\mathbb{N}})$. Thus, from the upper semi-continuity of H^K_δ we obtain

$$H^K_\delta(\mu) \geq \dim_B X \qquad \forall\mu\in\mathscr{P}_\sigma(X^{\mathbb{N}}).$$

We end the proof of Theorem C by bringing together the last inequality and (3.25).

□

Remark 3.4.4 *We note that, under the assumptions of Theorem C, one has*

$$M(\mu) = \dim_B X \qquad \forall\mu\in\mathscr{P}_\sigma(X^{\mathbb{N}}).$$

3.4.1 Dimension homogeneity assumption

We stress that some homogeneity hypothesis on the box dimension structure of the space Y is indeed necessary. For instance, if a space of positive upper box dimension contains an isolated point (as happens with $X = \{0\} \cup \{1/n\colon\ n \in \mathbb{N}\}$ endowed with the Euclidean metric, whose box dimension is $1/2$; cf. [36, Lemma 3.1]), then Theorem 3.4.1 does not hold for the Dirac probability measure supported on that point. Consequently, Theorem C is not valid for a well chosen potential, as the following result illustrates.

Proposition 3.4.5 *Let (X,d) be a compact metric space and σ the shift map on $(X^{\mathbb{N}}, d_\rho)$.*

(a) Given $y \in X$, one has

$$H^K_\delta(\delta^{\mathbb{N}}_{\{y\}}) \;=\; M(\delta^{\mathbb{N}}_{\{y\}}) \;=\; \lim_{\gamma\to 0^+} \overline{\dim}_B\, B_\gamma(y).$$

(b) Assume that (X,d) has positive upper box dimension and contains a point y_0 such that

$$\lim_{\gamma\to 0^+} \overline{\dim}_B\, B_\gamma(y_0) \;<\; \overline{\dim}_B X.$$

Then there exists $\varphi \in C^0(X^{\mathbb{N}})$ such that

$$\overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d_\rho, \sigma, \varphi) \;<\; \overline{\dim}_B X \;+\; \max_{\mu\in\mathscr{E}_\sigma(X^{\mathbb{N}})} \int \varphi\, d\mu.$$

Proof: (a) Fix $y \in X$. For each $\gamma > 0$ consider a map $\varphi_\gamma \in C^0(X^{\mathbb{N}})$ satisfying

$$\varphi_\gamma(x_1, x_2, \cdots) = \begin{cases} 0 & \text{if } x_1 \in B_{\gamma/2}(y) \\ -\overline{\dim}_B Y - 1 & \text{if } x_1 \notin B_\gamma(y) \end{cases}$$

and always bounded from above by 0. Given $\varepsilon > 0$, let $U_1, \cdots, U_N$ be a minimal open cover of $B_\gamma(y)$ with diameter at most ε. Complete it to an open cover of the whole space X by adding a minimal ε-cover $U_{N+1}, \cdots, U_{N+P}$ of the complement of $\cup_{i=1}^n U_i$. Without loss of generality we may suppose that none of the sets $U_{N+1}, \cdots, U_{N+P}$ intersects $B_\gamma(y)$. Take $\ell = \ell(\varepsilon)$ such that $\mathrm{diam}(X,d)/\rho^{\ell-1} < \varepsilon$ and consider the open cover of $X^{\mathbb{N}}$ given by

$$\alpha = \big\{U_{j_1} \times \cdots \times U_{j_{\ell+n}} \times X \times X \times \cdots \;:\; 1 \le j_1, \cdots, j_{\ell+n} \le N+P\big\}.$$

Then the diameter of α with respect to the metric $d_{\rho,n}$ satisfies $\mathrm{diam}(\alpha, d_{\rho,n}) \le \varepsilon$ and

$$
\begin{aligned}
\sum_{U \in \alpha} (1/\varepsilon)^{\sup_U S_n\varphi_\gamma} &= \sum_{U_{j_1}\times\cdots\times U_{j_{\ell+n}}} (1/\varepsilon)^{\sum_{\nu=1}^n \sup_{U_{j_\nu}} \varphi_\gamma} \\
&= (N+P)^\ell \sum_{U_{j_1}\times\cdots\times U_{j_n}} (1/\varepsilon)^{\sum_{\nu=1}^n \sup_{U_{j_\nu}} \varphi_\gamma} \\
&\le (N+P)^\ell \sum_{U_{j_1}\times\cdots\times U_{j_n}} (1/\varepsilon)^{-(\overline{\dim}_B Y+1)\#\left\{\nu:\ j_\nu\in\{N+1,\cdots,N+P\}\right\}} \\
&= (N+P)^\ell \Big(\sum_{k=0}^n \binom{n}{k} P^k N^{n-k} (\varepsilon^{\overline{\dim}_B Y+1})^k\Big) \\
&= (N+P)^\ell \Big(N+P\varepsilon^{\overline{\dim}_B Y+1}\Big)^n.
\end{aligned}
$$

We observe that as P is smaller than the ε-covering number of X then one has $P \le \varepsilon^{-(\overline{\dim}_B X+1/2)}$ for small enough $\varepsilon>0$. Thus, for those values of ε,

$$
P(X^{\mathbb{N}}, d_\rho, \sigma, \varphi_\gamma, \varepsilon) \le \log\Big(N+P\varepsilon^{\overline{\dim}_B X+1}\Big) \le \log\Big(N+\varepsilon^{1/2}\Big)
$$

and so

$$
\overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d_\rho, \sigma, \varphi_\gamma) \le \limsup_{\varepsilon\to 0^+} \frac{\log\Big(N+\varepsilon^{1/2}\Big)}{\log(1/\varepsilon)} = \overline{\dim}_B B_\gamma(y). \tag{3.28}
$$

On the other hand,

$$
\max_{\mu\in\mathscr{E}_\sigma(X^{\mathbb{N}})} \int \varphi_\gamma\, d\mu = \int \varphi_\gamma\, d\delta_{\{y\}}^{\mathbb{N}} = 0 \qquad \forall \gamma>0. \tag{3.29}
$$

Therefore

$$
\begin{aligned}
M(\delta_{\{y\}}^{\mathbb{N}}) &= \inf_{\psi\in C^0(X^{\mathbb{N}})} \left\{\overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d, \sigma, \psi) - \int \psi\, d\delta_{\{y\}}^{\mathbb{N}}\right\} \\
&\le \lim_{\gamma\to 0^+} \overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d, \sigma, \varphi_\gamma) \\
&\le \lim_{\gamma\to 0^+} \overline{\dim}_B B_\gamma(y).
\end{aligned}
$$

To complete the proof of item (a) we summon Lemmas 3.4.2 and 3.3.4.

(b) Let $y_0 \in X$ be such that $\lim_{\gamma\to 0^+} \overline{\dim}_B B_\gamma(y_0) < \overline{\dim}_B X$ and $\gamma_0>0$ satisfying

$$
\overline{\dim}_B B_{\gamma_0}(y_0) < \overline{\dim}_B X.
$$

Consider φ_{γ_0} associated to y_0 as previously defined. Then, taking into account (3.28) and (3.29) we conclude that

$$
\overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d_\rho, \sigma, \varphi_{\gamma_0}) < \overline{\dim}_B X + \max_{\mu\in\mathscr{E}_\sigma(X^{\mathbb{N}})} \int \varphi_{\gamma_0}\, d\mu.
$$

$\square$

3.5 Metric mean dimension points

In this section we prove Theorem D. Recall that the map $\mathscr{D}\colon X \to \mathbb{R}$ is given by

$$\mathscr{D}(x) = \inf\left\{\overline{\mathrm{mdim}}_M(U,d,T) \,:\, U \text{ is an open neighborhood of } x\right\}.$$

We start by establishing some useful properties of the map $\mathscr{D}$.

Lemma 3.5.1 *The map $\mathscr{D}$ is upper semi-continuous and*

$$0 \leq \mathscr{D} \leq \overline{\mathrm{mdim}}_M(X,d,T).$$

In particular, $\mathscr{D}$ is measurable and μ-integrable with respect to every $\mu \in \mathscr{P}(X)$.

Proof: Given $x \in X$ and an open neighborhood U of x, let $(x_n)_{n\in\mathbb{N}}$ be a sequence of elements in U converging to x. Then, for sufficiently large n, one has $x_n \in U$ and

$$\mathscr{D}(x_n) \leq \overline{\mathrm{mdim}}_M(U,d,T).$$

We conclude the proof by taking $\limsup_n$ and $\inf_U$. □

Lemma 3.5.2 *For any $\mu \in \mathscr{E}_T(X)$, the map $\mathscr{D}$ is constant μ-almost everywhere.*

Proof: We begin by showing that $\mathscr{D}$ is increasing along orbits. Given an open nonempty subset U of X, let $E \subset U$ be $(n+1,\varepsilon)$-separated. Then, by the Pigeonhole Principle, there is an (n,ε)-separated subset $A \subset T(E) \subset T(U)$ with cardinality $\#A \geq \#E/S(X,d,0,\varepsilon)$. Hence,

$$S_1(T(U),d_n,0,\varepsilon) \geq S_1(U,d_{n+1},0,\varepsilon)/S(X,d,0,\varepsilon)$$

and so $\mathscr{D}(x) \leq \mathscr{D}(T(x))$ for all $x \in X$.

We observe now that if $x \in X$ is recurrent, that is, $\lim_{k\to+\infty} T^{n_k}(x) = x$ for some subsequence of the orbit of x by T, then by the upper semi-continuity of $\mathscr{D}$ one has

$$\limsup_{k\to+\infty} \mathscr{D}(T^{n_k}(x)) \leq \mathscr{D}(x).$$

As $\mathscr{D}$ is increasing along orbits, the previous inequality implies that $\mathscr{D}$ is constant along orbits of recurrent points. Thus, given $\mu \in \mathscr{E}_T(X)$, by the Poincaré Recurrence Theorem we deduce that $\mathscr{D}$ is constant μ-almost everywhere. □

Proof of Theorem C:

We will first prove the equality (3.13) in Theorem D under the assumption that $T\colon X \to X$ is a homeomorphism. We will prove a general result, taking into account the role of the potentials, of which (3.13) is an immediate consequence.

Theorem 3.5.3 *Let (X,d) be a compact metric space and $T: X \to X$ be a homeomorphism such that $\overline{\text{mdim}}_M(X,d,T) < +\infty$. Then, for every $\varphi \in C^0(X)$,*

$$\overline{\text{mdim}}_M(X,d,T,\varphi) \leq \max_{\mu \in \mathscr{P}_T(X)} \int \big(\mathscr{D}+\varphi\big)(x)\, d\mu(x) = \max_{\mu \in \mathscr{E}_T(X)} \int \big(\mathscr{D}+\varphi\big)(x)\, d\mu(x).$$

Proof: Fix $\delta \in (0,1)$ and let μ_0 be the equilibrium state constructed in the proof of Theorem B, namely the probability measure which is the weak*-limit of the sequence $\big(\mu_{\varepsilon_n}\big)_{n\in\mathbb{N}}$ in $\mathscr{E}_T(X)$ as $(\varepsilon_n)_n$ goes to 0 and satisfies

$$\overline{\text{mdim}}_M(X,d,T,\varphi) = H_\delta^K(\mu_0) + \int \varphi\, d\mu_0 = \lim_{n\to\infty} \frac{h^K_{\mu_{\varepsilon_n}}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} + \int \varphi\, d\mu_0.$$

Note that, given an open set $U \subset X$ such that $\mu(U) > 0$, for large enough n one has $\mu_{\varepsilon_n}(U) > 0$.

Lemma 3.5.4 *[63, Theorem 3.7] If $T: X \to X$ is a homeomorphism on a compact metric space (X,d), for every ergodic probability measure $\mu \in \mathscr{E}_T(X)$, every open set $U \subset X$ such that $\mu(U) > 0$, every $\varepsilon > 0$ and every $\delta \in (0,1)$ we have*

$$P(U,d,T,0,\varepsilon) \geq h^K_\mu(\varepsilon,\delta).$$

By Lemma 3.5.4 and Remark 3.3.8 we have

$$H_\delta^K(\mu_0) = \lim_{n\to+\infty} \frac{h^K_{\mu_{\varepsilon_n}(\varepsilon_n,\delta)}}{\log(1/\varepsilon_n)} \leq \overline{\text{mdim}}_M(U,d,T).$$

Thus, for any $x \in \text{supp}(\mu_0)$,

$$H_\delta^K(\mu_0) \leq \inf_{x\in U} \overline{\text{mdim}}_M(U,d,T) = \mathscr{D}(x).$$

Therefore,

$$H_\delta^K(\mu_0) \leq \inf_{x\in \text{supp}(\mu_0)} \mathscr{D}(x) \leq \int \mathscr{D}\, d\mu_0.$$

Hence,

$$\overline{\text{mdim}}_M(X,d,T,\varphi) \leq \int \big(\mathscr{D}+\varphi\big)\, d\mu_0 \leq \max_{\mu \in \mathscr{P}_T(X)} \int \big(\mathscr{D}+\varphi\big)(x)\, d\mu(x).$$

As $\mu \mapsto \int \mathscr{D} d\mu$ is affine, the previous maximum is also attained at some ergodic probability measure. This ends the proof of Theorem 3.5.3.

□

For homeomorphisms, the equality (3.13) in Theorem D follows from Theorem 3.5.3, when $\varphi \equiv 0$, and the fact that $\mathscr{D} \leq \overline{\text{mdim}}_M(X,d,T)$ (cf. Lemma 3.5.1). We will extend this result to non-invertible maps in Chapter 3 (see Proposition 4.4.1). For now, let us assume that (3.13) holds for arbitrary continuous maps.

Consider a measure $\mu \in \mathscr{P}_T(X)$ satisfying

$$\overline{\text{mdim}}_M(X,d,T) = \int \mathscr{D}\, d\mu.$$

Then, as $\mathscr{D} \leq \overline{\text{mdim}}_M(X,d,T)$, there exists a full measure set $Z \subset X$ such that $\mathscr{D}|_Z \equiv \overline{\text{mdim}}_M(X,d,T)$. Therefore, since $\text{supp}(\mu) \subset \overline{Z}$ and $\mathscr{D}$ is upper semi-continuous, we get $\mathscr{D}|_{\text{supp}(\mu)} \equiv \overline{\text{mdim}}_M(X,d,T)$. The proof of Theorem D is complete. □

Remark 3.5.5 *A strict inequality may happen in Theorem 3.5.3. For instance, if one considers the shift map on $X^{\mathbb{Z}}$, we always have*

$$\mathscr{D} \equiv \overline{dim_B} X$$

and so, if (X,d) has positive upper box dimension and contains an isolated point, then using the potential φ provided by Proposition 3.4.5 we get

$$\overline{mdim_M}(X^{\mathbb{Z}}, d_\rho, \sigma, \varphi) < \max_{\mu \in \mathscr{E}_\sigma(X^{\mathbb{Z}})} \int \big(\mathscr{D} + \varphi\big)(x)\, d\mu(x).$$

Another consequence of Theorem 3.5.3 is the following relation between $\mu \mapsto \int \mathscr{D}\, d\mu$ and $\mu \mapsto M(\mu)$ (see (3.6)).

Corollary 3.5.6 *Let (X,d) be a compact metric space and $T\colon X \to X$ be a homeomorphism such that $\overline{\text{mdim}}_M(X,d,T) < +\infty$. Then,*

$$M(\mu) \leq \int \mathscr{D}\, d\mu \qquad \forall\, \mu \in \mathscr{P}_T(X).$$

In addition, the equality in Theorem 3.5.3 holds for every $\varphi \in C^0(X)$ if and only if

$$M(\mu) = \int \mathscr{D}\, d\mu \qquad \forall\, \mu \in \mathscr{P}_T(X).$$

Proof: We recall that M is given by $M(\mu) = \inf_{\varphi \in \mathscr{C}} \int \varphi\, d\mu$, where

$$\mathscr{C} = \big\{\varphi \in C^0(X) \colon \overline{\text{mdim}}_M(X,d,T,-\varphi) \leq 0\big\}.$$

Now observe that, by Theorem 3.5.3, for every continuous potential satisfying $\psi \geq \mathscr{D}$ we have

$$\overline{\text{mdim}}_M(X,d,T,-\psi) \leq \max_{\mu \in \mathscr{P}_T(X)} \int \big(\mathscr{D} - \psi\big)\, d\mu \leq 0$$

that is, $\psi \in \mathscr{C}$. By upper semi-continuity of $\mathscr{D}$, there exists a decreasing sequence of continuous potentials φ_n converging to $\mathscr{D}$ as $n \to \infty$. Then, by the Monotone Convergence Theorem,

$$M(\mu) = \inf_{\varphi \in \mathscr{C}} \int \varphi\, d\mu \leq \lim_{n \to \infty} \int \varphi_n\, d\mu = \int \mathscr{D}\, d\mu \qquad \forall\, \mu \in \mathscr{P}_T(X).$$

To conclude, assume that the inequality in Theorem 3.5.3 is indeed an equality for any potential $\varphi \in C^0(X)$. Then, $\mu \mapsto \int \mathscr{D}\, d\mu$ satisfies the variational principle (3.5). By the maximality of $\mu \mapsto M(\mu)$ among all those maps satisfying this variational principle (provided by Theorem A), we get

$$M(\mu) \geq \int \mathscr{D}\, d\mu \qquad \forall\, \mu \in \mathscr{P}_T(X).$$

□

Remark 3.5.7 *A strict inequality may happen in Corollary 3.5.6. For instance, if one considers the shift map on $X^{\mathbb{Z}}$, where X has positive upper box dimension and contains an isolated point y_0, by Proposition 3.4.5,*

$$\int \mathscr{D}\, d\delta_{y_0}^{\mathbb{Z}} \;=\; \overline{dim}_B X \;>\; 0 \;=\; \lim_{\gamma\to 0^+} \overline{\dim}_B B_\gamma(y) \;=\; M(\delta_{\{y\}}^{\mathbb{Z}}).$$

An element $x \in X$ is said to be as a *metric mean dimension point* if $\mathscr{D}(x) > 0$; it is a *full metric mean dimension point* if $\mathscr{D}(x) = \overline{\mathrm{mdim}}_M(X,d,T)$. Denote the set of such points by $D_p(X,d,T)$ and $D_p^f(X,d,T)$, respectively. The following is an immediate consequence of Theorem D:

Corollary 3.5.8 *Under the assumptions of Theorem D, there exists an ergodic probability measure $\mu \in \mathscr{E}_T(X)$ such that* $\mathrm{supp}(\mu) \subseteq D_p^f(X,d,T)$.

3.6 Examples

The first example confirms that the measure-theoretic upper metric mean dimension defined in the statement of Theorem A depends on the metric.

Example 3.6.1 *Consider the symbolic space $X = \{0,1\}^{\mathbb{N}}$ with the product topology. Recall that we can generate the product topology on $X^{\mathbb{N}}$ using two distinct metrics d_2^1 and d_2^2 satisfying (see* (2.6)*)*

$$\begin{aligned} \overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d_2^1, \sigma^{\mathbb{N}}) &= 1 \\ \overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d_2^2, \sigma^{\mathbb{N}}) &= \log 2/\log 3. \end{aligned}$$

Let M_1 and M_2 denote the measure-theoretic upper metric mean dimension maps assigned by Theorem A to $(X^{\mathbb{N}}, d_2^1, \sigma^{\mathbb{N}})$ and $(X^{\mathbb{N}}, d_2^2, \sigma^{\mathbb{N}})$. Then

$$\max_{\mu \in \mathscr{P}(X^{\mathbb{N}})} M_1(\mu) = \frac{\log 2}{\log 3} < 1 = \max_{\mu \in \mathscr{P}(X^{\mathbb{N}})} M_2(\mu).$$

The next example illustrate that there *can* exist a measure capturing almost full complexity over all scales.

Example 3.6.2 *Denote by $|\cdot|$ the Euclidean distance in $[0,1]$ and, given any $\rho > 1$, consider the space $[0,1]^{\mathbb{N}}$ with the metric d_ρ (defined in* (2.4)*). Let $\sigma\colon [0,1]^{\mathbb{N}} \to [0,1]^{\mathbb{N}}$ be the shift map. From Theorem 3.4.1 we already know that, for every $\delta \in (0,1)$,*

$$H_\delta^K(\mu) = 1 \qquad \forall\, \mu \in \mathscr{P}_\sigma([0,1]^{\mathbb{N}}). \tag{3.30}$$

Nevertheless, we can improve the argument in the proof of Lemma 3.4.2 when computing H_δ^K at a Dirac measure supported on a fixed point, thereby showing that each approximating measure μ_ε detects separation at all *scales.*

Given $x \in [0,1]$ and $\varepsilon > 0$, denote by $I(\varepsilon)$ the amount of odd numbers between 1 and $\lceil 1/\varepsilon \rceil$ (thus $I(\varepsilon) \geq 1/2\varepsilon$) and by $B = B_\varepsilon(x) \subset [0,1]$ the open interval of radius ε centered at x. Partition B into $\lceil 1/\varepsilon \rceil$ intervals with the same length, say $B(1), \cdots, B(\lceil 1/\varepsilon \rceil)$. Iterate by partitioning $B(i_1, \cdots, i_n)$ into $B(i_1, \cdots, i_n, 1), \cdots, B(i_1, \cdots, i_n, \lceil 1/\varepsilon \rceil)$ so that the length of $B(i_1, \cdots, i_n)$ is at least ε^{n+1}. We observe that, if

$$\begin{aligned} y &\in B(i_1^1, \cdots, i_n^1) \times B(i_1^2, \cdots, i_n^2) \times \cdots \times B(i_1^k, \cdots, i_n^k) \\ z &\in B(j_1^1, \cdots, j_n^1) \times B(j_1^2, \cdots, j_n^2) \times \cdots \times B(j_1^k, \cdots, j_n^k) \end{aligned}$$

where the i's and j's are all odd, then $d_k(y,z) \geq \varepsilon^{n+1}$.

Let ν_ε be a probability measure in $[0,1]$ determined by

- $\nu_\varepsilon(B) = 1$;
- $\nu_\varepsilon(B(i)) = I(\varepsilon)^{-1}$, *for every odd i;*
- $\nu_\varepsilon(B(i_1, \cdots, i_n)) = I(\varepsilon)^{-n}$, *for every odd i's.*

Define $\mu_\varepsilon = \nu_\varepsilon^{\mathbb{N}}$, which is clearly ergodic, converges to $\delta_{\{x\}}^{\mathbb{N}}$ as $\varepsilon \to 0^+$ and satisfies

$$\mu_\varepsilon\big(B(i_1^1, \cdots, i_n^1) \times B(i_1^2, \cdots, i_n^2) \times \cdots \times B(i_1^k, , i_n^k)\big) = I(\varepsilon)^{-kn} \qquad \forall \text{ odd } i's.$$

Let $L = L(\varepsilon, k, n, \delta) \in \mathbb{N}$ be the maximal integer such that $LI(\varepsilon)^{-kn} < \delta$. Then, any set A satisfying $\mu_\varepsilon(A) > 1 - \delta$ must intersect at least $I(\varepsilon)^{kn} - L$ cylinders, with odd indices, of the partition level $B(i_1^1, \cdots, i_n^1) \times B(i_1^2, \cdots, i_n^2) \times \cdots \times B(i_1^k, \cdots, i_n^k)$, where

$$I(\varepsilon)^{kn} - L > (1-\delta) I(\varepsilon)^{kn}$$

since $[0,1]^{\mathbb{N}} \setminus A$ can contain at most L of such sets. Hence, A must have a (d_k, ε^{n+1})-separated subset with cardinality $(1-\delta) I(\varepsilon)^{kn}$. Thus,

$$h_{\mu_\varepsilon}^K(\varepsilon^{n+1}, \delta) \geq \limsup_{k \to +\infty} \frac{1}{k} \log\big((1-\delta) I(\varepsilon)^{kn}\big) = \log I(\varepsilon)^n \geq \log(1/2\varepsilon)^n.$$

Fix $\varepsilon > 0$. Given $\varepsilon' > 0$, let $n = n(\varepsilon, \varepsilon')$ be such that $\varepsilon^{n+1} \leq \varepsilon' \leq \varepsilon^n$. Then

$$\begin{aligned} H_\delta^K(\mu_\varepsilon) &\geq \limsup_{\varepsilon' \to 0^+} \frac{h_{\mu_\varepsilon}^K(\varepsilon', \delta)}{\log(1/\varepsilon')} \\ &\geq \liminf_{\varepsilon' \to 0^+} \frac{h_{\mu_\varepsilon}^K(\varepsilon', \delta)}{\log(1/\varepsilon')} \\ &\geq \liminf_{n \to +\infty} \frac{h_{\mu_\varepsilon}^K(\varepsilon^n, \delta)}{\log(1/\varepsilon^{n+1})} \\ &\geq \liminf_{n \to +\infty} \frac{n-1}{n+1} \Big(\frac{\log(1/\varepsilon) - \log 2}{\log(1/\varepsilon)} \Big) \\ &= 1 - \frac{\log 2}{\log(1/\varepsilon)}. \end{aligned} \tag{3.31}$$

Finally, the upper semi-continuity of H^K_δ yields

$$H^K_\delta(\delta^{\mathbb{N}}_{\{x\}}) \geq \limsup_{\varepsilon\to 0^+} H^K_\delta(\mu_\varepsilon) = 1.$$

We emphasize that the interesting feature of this construction, in contrast to the proof of Lemma 3.4.2, is the fact that the lower bound (3.31) of $H^K_\delta(\mu_\varepsilon)$ is done by using the measure μ_ε for all scales $\varepsilon' > 0$. More precisely, the inequality

$$\liminf_{\varepsilon'\to 0^+} \frac{h^K_{\mu_\varepsilon}(\varepsilon',\delta)}{\log(1/\varepsilon')} \geq 1 - \frac{\log 2}{\log(1/\varepsilon)}$$

indicates that the complexity at every scale $\varepsilon' > 0$ (in terms of metric mean dimension) captured by the measure μ_ε is at least $1-\frac{\log 2}{\log(1/\varepsilon)}$. This means precisely that μ_ε detects separation of points at all scales.

Now we make use of the map $\mu \mapsto H^K_\delta(\mu)$ to obtain an exact formula for the upper metric mean dimension with potential for the maps $T_{a,b}$ defined in Subsection 2.4.2. The formula is expressed in terms of ergodic optimization.

Example 3.6.3 *Let a and b be parameters satisfying conditions (C1)-(C3) in* (2.7) *and $T_{a,b}$ be the corresponding interval map defined in* (2.8). *We will make use of the formula*

$$\overline{\text{mdim}}_M([0,1],|\cdot|,T_{a,b}) \;=\; \limsup_{k\to+\infty} \frac{\log b_k}{\log(1/\varepsilon_k)},$$

where $\varepsilon_k = |J_k|/b_k$ (see Theorem I), and extend it to all continuous potentials $\varphi \in C^0([0,1])$.

Recall that we can partition each interval J_k in b_k^n subintervals $J_k(i_1,\cdots,i_n)$ of length ε_k/b_k^n satisfying

$$T^s_{a,b}\big(J_k(i_1,\cdots,i_n)\big) \;=\; J_k(i_{s+1},i_{s+2},\cdots,i_n) \qquad \forall s \in \{0,\cdots,n-1\}.$$

If $x \in J_k(i_1,\cdots,i_n)$ and $y \in J_k(j_1,\cdots,j_n)$, where $(i_1,\cdots,i_n) \neq (j_1,\cdots,j_n)$ and the i's and the j's are all odd, in at most n iterates their images lie in $J_k(i_l)$ and $J_k(j_l)$, respectively, whose distance is at least ε_k; hence $d_n(x,y) \geq \varepsilon_k$.

We are ready to define a suitable sequence of invariant probability measures $(\mu_k)_k$ which converges to $\delta_{\{1\}}$. We start by recalling that, for every $k \in \mathbb{N}$,

$$\#\{1 \leq i \leq b_k \,:\, i \text{ is odd}\} \;=\; \frac{b_k+1}{2}.$$

Let μ_k be defined by

- $\mu_k(J_k) = 1$;
- $\mu_k(J_k(i)) = \frac{2}{b_k+1}$ *for every odd* $1 \leq i \leq b_k$;
- $\mu_k(J_k(i_1,\cdots,i_n)) = (\frac{2}{b_k+1})^n$ *for every odd* $1 \leq i_1,\cdots,i_n \leq b_k$ *and every* $n \in \mathbb{N}$.

Clearly, $(\mu_k)_k$ converges to $\delta_{\{1\}}$ and each μ_k is T-invariant. In fact,

$$T^{-1}J_k(i_1,\cdots,i_n) = \bigcup_{i=1}^{b_k} J_k(i,i_1,\cdots,i_n)$$

and

$$\mu_k(T^{-1}J_k(i_1,\cdots,i_n)) = \sum_{1\le i\le b_k;\, i\ odd} \Big(\frac{2}{b_k+1}\Big)^{n+1} = \Big(\frac{2}{b_k+1}\Big)^{n} = \mu_k(J_k(i_1,\cdots,i_n)).$$

Moreover, being a Bernoulli probability measure, μ_k is ergodic.

Let $L = L(\delta,n,k) \in \mathbb{N}$ be the maximal positive integer satisfying

$$L\Big(\frac{2}{b_k+1}\Big)^{n} < \delta.$$

Then, any subset $A \subset [0,1]$ with $\mu_k(A) > 1-\delta$ intersects at least $(\frac{b_k+1}{2})^n - L$ sets $J_k(i_1,\cdots,i_n)$ with odd indices and

$$\Big(\frac{b_k+1}{2}\Big)^{n} - L > \Big(\frac{b_k+1}{2}\Big)^{n}(1-\delta)$$

since $[0,1]\setminus A$ contains at most L of such intervals. Therefore, A must contain an (ε_k,d_n)-separated subset of cardinality bigger than $\Big(\frac{b_k+1}{2}\Big)^{n}(1-\delta)$ and

$$h^K_{\mu_k}(\varepsilon_k,\delta) \;\ge\; \limsup_{n\to+\infty} \frac{1}{n}\log\Big(\Big(\frac{b_k+1}{2}\Big)^{n}(1-\delta)\Big) \;=\; \log\Big(\frac{b_k+1}{2}\Big).$$

Hence, for every $\delta\in(0,1)$,

$$H^K_\delta(\delta_{\{1\}}) \;\ge\; \limsup_{k\to+\infty} \frac{\log\big(\frac{b_k+1}{2}\big)}{\log(1/\varepsilon_k)} \;=\; \limsup_{k\to+\infty}\frac{\log b_k}{\log(1/\varepsilon_k)}.$$

Since H^K_δ is always bounded from above by the upper metric mean dimension, we conclude

$$H^K_\delta(\delta_{\{1\}}) \;=\; \limsup_{k\to+\infty}\frac{\log b_k}{\log(1/\varepsilon_k)} \;=\; \overline{\mathrm{mdim}}_M([0,1],|\cdot|,T_{a,b}).$$

Let us now use the previous information to evaluate H^K_δ for all $T_{a,b}$-invariant measures.

Claim: For every $\delta\in(0,1)$,

$$H^K_\delta(\mu) \;=\; M(\mu) \;=\; \int \mathscr{D}\,d\mu \;=\; \mu\big(\{1\}\big)\;\limsup_{k\to+\infty}\frac{\log b_k}{\log(1/\varepsilon_k)} \qquad \forall\,\mu\in\mathscr{P}_{T_{a,b}}\big([0,1]\big).$$

In particular, the map H^K_δ is affine and the unique probability measure which maximizes H^K_δ is the Dirac mass $\delta_{\{1\}}$.

Proof of the Claim: *We will start by showing that, for every $\mu \in \mathscr{P}_{T_{a,b}}\big([0,1]\big)$,*

$$H^K_\delta(\mu) \;\leq\; \mu\big(\{1\}\big)\, H^K_\delta(\delta_{\{1\}}) \;=\; \mu(\{1\})\; \limsup_{k\to+\infty} \frac{\log b_k}{\log(1/\varepsilon_k)}.$$

In particular, this implies that

$$\mu\big(\{1\}\big) = 0 \quad \Rightarrow \quad H^K_\delta(\mu) = 0. \tag{3.32}$$

Let $\mu \in \mathscr{P}_{T_{a,b}}\big([0,1]\big)$ and $(\mu_\varepsilon)_\varepsilon \in \mathscr{M}(\mu)$. Denote the ergodic decomposition of μ_ε by

$$\mu_\varepsilon = \int_{\mathscr{E}_{T_{a,b}}([0,1])} m\, d\mathbb{P}_\varepsilon(m).$$

Given $\delta \in (0,1)$, select $k = k(\delta)$ such that $\mu\big([0,a_k)\big) > 1-\mu\big(\{1\}\big)-\delta$. Then, for small enough $\varepsilon > 0$, we have $\mu_\varepsilon\big([0,a_k)\big) > 1-\mu\big(\{1\}\big)-\delta$ and consequently, $\mu_\varepsilon\big([a_k,1)\big) < \delta+\mu\big(\{1\}\big) - \mu_\varepsilon\big(\{1\}\big)$. Since for every ergodic probability measure other than $\delta_{\{1\}}$ there exists some $n \geq 1$ such that $J'_n = [a_n, a_{n+1})$ has full mass, we can write

$$\mu_\varepsilon = \sum_{n=1}^{\infty} \int_{m(J'_n)=1} m\, d\mathbb{P}_\varepsilon(m) + \mu_\varepsilon\big(\{1\}\big)\, \delta_{\{1\}}.$$

Moreover, for every positive integer $n \leq k$, if $m(J'_n) = 1$ then $m([0,a_k)) = 1$. Yet, as T restricted to $[0,a_k)$ has finite topological entropy, we know that

$$\limsup_{\varepsilon\to 0^+} \frac{1}{\log(1/\varepsilon)} \sum_{n=1}^{k} \int_{m(J'_n)=1} h^K_m(\varepsilon,\delta)\, d\mathbb{P}_\varepsilon(m) = 0.$$

Since we also have $h^K_{\delta_{\{1\}}}(\varepsilon,\delta) = 0$, in order to estimate $H^K_\delta(\mu)$, we are reduced to the ergodic components m of μ_ε such that $m\big([a_k,1)\big) = 1$. So,

$$\begin{aligned}
\limsup_{\varepsilon\to 0^+} \frac{h^K_{\mu_\varepsilon}(\varepsilon,\delta)}{\log(1/\varepsilon)} &= \limsup_{\varepsilon\to 0^+} \frac{1}{\log(1/\varepsilon)} \int_{m\left([a_k,1)\right)=1} h^K_m(\varepsilon,\delta)\, d\mathbb{P}_\varepsilon(m) \\
&\leq \limsup_{\varepsilon\to 0^+} \frac{h_\varepsilon(X,d,T)}{\log(1/\varepsilon)}\, \mu_\varepsilon\big([a_k,1)\big) \\
&\leq \limsup_{\varepsilon\to 0^+} \frac{h_\varepsilon(X,d,T)}{\log(1/\varepsilon)} \Big(\delta + \mu(\{1\}) - \mu_\varepsilon\big(\{1\}\big)\Big) \\
&\leq \limsup_{k\to+\infty} \frac{\log b_k}{\log(1/\varepsilon_k)} \Big(\delta + \mu\big(\{1\}\big)\Big).
\end{aligned} \tag{3.33}$$

Therefore,

$$H^K_\delta(\mu) \;\leq\; \limsup_{k\to+\infty} \frac{\log b_k}{\log(1/\varepsilon_k)} \Big(\mu(\{1\}) + \delta\Big) \qquad \forall\, \delta \in (0,1),$$

and so, by Remark 3.3.1,

$$H^K_\delta(\mu) \;\leq\; \sup_{0<\delta'<1} H^K_{\delta'}(\mu) \;=\; \lim_{\delta'\to 0^+} H^K_{\delta'}(\mu) \;\leq\; \limsup_{k\to+\infty} \frac{\log b_k}{\log(1/\varepsilon_k)}\; \mu(\{1\}).$$

Now we prove the reverse inequality. Let $\nu \in \mathscr{P}_{T_{a,b}}\big([0,1]\big)$ and consider the decomposition $\nu = t\delta_{\{1\}} + (1-t)\mu$, where $t = \nu(\{1\})$ and $\mu \in \mathscr{P}_{T_{a,b}}\big([0,1]\big)$ satisfies $\mu(\{1\}) = 0$. Take the sequence $(\varepsilon_k)_k$ as in (2.9) *and the sequence of probability measures $(\mu_k)_k$ in $\mathscr{M}(\delta_{\{1\}})$ used on the previous page to compute $H^K_\delta(\delta_{\{1\}})$, which satisfy*

$$\limsup_{k\to+\infty} \frac{h^K_{\mu_k}(\varepsilon_k,\delta)}{\log(1/\varepsilon_k)} \;=\; H^K_\delta(\delta_{\{1\}}) \;=\; \limsup_{k\to+\infty} \frac{\log b_k}{\log(1/\varepsilon_k)}.$$

Let $(\nu_k)_k = \big(t\mu_k + (1-t)\mu\big)_k$ in $\mathscr{M}(\nu)$. As $\eta \mapsto h^K_\eta(\varepsilon,\delta)$ is affine (cf. (2.3)*), we get*

$$\begin{aligned}
H^K_\delta(\nu) \;&\geq\; \limsup_{k\to+\infty} \frac{h^K_{\nu_k}(\varepsilon_k,\delta)}{\log(1/\varepsilon_k)} && \textit{by definition of } H^K_\delta(\nu)\\
&=\; \limsup_{k\to+\infty} \Big(t\,\frac{h^K_{\mu_k}(\varepsilon_k,\delta)}{\log(1/\varepsilon_k)} + (1-t)\frac{h^K_{\mu}(\varepsilon_k,\delta)}{\log(1/\varepsilon_k)} \Big) &&\\
&=\; \limsup_{k\to+\infty} t\,\frac{h^K_{\mu_k}(\varepsilon_k,\delta)}{\log(1/\varepsilon_k)} && \textit{by } (3.32)\\
&=\; t\, H^K_\delta(\delta_{\{1\}}) &&\\
&=\; \nu(\{1\}) \; \limsup_{k\to+\infty} \frac{\log b_k}{\log(1/\varepsilon_k)}. &&
\end{aligned}$$

Thus,

$$H^K_\delta(\mu) \;=\; \mu(\{1\}) \; \limsup_{k\to+\infty} \frac{\log b_k}{\log(1/\varepsilon_k)} \qquad \forall\, \mu \in \mathscr{P}_{T_{a,b}}\big([0,1]\big).$$

Moreover, by definition,

$$\mathscr{D}(x) = \begin{cases} \limsup_k \frac{\log b_k}{\log(1/\varepsilon_k)} & \textit{if } x = 1\\ 0 & \textit{otherwise,} \end{cases}$$

and therefore

$$H^K_\delta(\mu) \;=\; \mu(\{1\}) \; \limsup_{k\to+\infty} \frac{\log b_k}{\log(1/\varepsilon_k)} \;=\; \int \mathscr{D}\, d\mu \qquad \forall\, \mu \in \mathscr{P}_{T_{a,b}}\big([0,1]\big).$$

The discussion at the end of Subsection 3.1.3 yields $H^K_\delta(\mu) = M(\mu)$ for every $\mu \in \mathscr{P}_{T_{a,b}}\big([0,1]\big)$. In particular, this implies that the map $\mu \mapsto H^K_\delta(\mu) = M(\mu)$ is affine.

□

As a consequence of the previous claim, we obtain the following closed formula for the upper metric mean dimension with potential of $T_{a,b}$ with respect to the Euclidean distance

$$\overline{\mathrm{mdim}}_M([0,1],|\cdot|,T_{a,b},\varphi) = \max_{\mu\in\mathscr{E}_{T_{a,b}}\left([0,1]\right)} \int \big(\mathscr{D}+\varphi\big)\,d\mu \qquad \forall\, \varphi \in C^0([0,1]). \tag{3.34}$$

Remark 3.6.4 *Example 3.6.3 evinces an obstruction if one tries to prove* (3.9) *using the map*

$$\widetilde{H}_\delta \colon \ \mu \in \mathscr{P}_{T_{a,b}}\big([0,1]\big) \quad \mapsto \quad \limsup_{\varepsilon \to 0^+} \frac{h^K_\mu(\varepsilon,\delta)}{\log(1/\varepsilon)}$$

instead of H^K_δ. *Given* $\mu \in \mathscr{P}_{T_{a,b}}\big([0,1]\big)$, *applying* (3.33) *to the constant sequence* $(\mu)_\varepsilon \in \mathscr{M}(\mu)$, *we obtain*

$$\widetilde{H}_\delta(\mu) \ \le \ \limsup_{\varepsilon \to 0^+} \frac{h_\varepsilon(X,d,T)}{\log(1/\varepsilon)}\,\delta \ = \ \delta \ \limsup_{k \to +\infty} \frac{\log b_k}{\log(1/\varepsilon_k)},$$

which implies that

$$\widetilde{H}_\delta(\mu) \ = \ 0 \qquad \forall\, \mu \in \mathscr{P}_{T_{a,b}}\Big([0,1]\Big).$$

3.7 Open questions

Despite satisfying the variational principle (3.13), the definition of $\mathscr{D}$ as a quantifier of local complexity is too coarse. This is expressed in Remark 3.5.5 and what causes $\mathscr{D}$ to be so large is the fact that, at any point x, it considers non-dynamical neighbourhoods of x. A truly dynamical measurement of local complexity should take into account only trajectories following the one of x. This can be done by replacing neighbourhoods by stable sets, namely, defining

$$\mathscr{D}_1(x) \ = \ \lim_{\gamma \to 0^+} \limsup_{\varepsilon \to 0^+} \frac{1}{\log(1/\varepsilon)} \limsup_{n \to +\infty} \frac{1}{n} \log S\big(B_\gamma(x,d_n),d_n,\varepsilon\big),$$

where $B_\gamma(x,d_n) = \{y \in X \colon d(T^j(x),T^j(y)) < \gamma, \ \forall 0 \le j \le n-1\}$, or

$$\mathscr{D}_2(x) \ = \ \lim_{\gamma \to 0^+} \overline{\mathrm{mdim}}_M\big(B_\gamma(x,d_{\mathbb{N}}),d,T\big),$$

where $B_\gamma(x,d_{\mathbb{N}}) = \{y \in X \colon d(T^j(x),T^j(y)) < \gamma, \ \forall j \ge 0\}$.

These quantities coincide with $\mathscr{D}$ for the maps $T_{a,b}$ in Subsection 2.4.2. Since they properly avoid what originally caused $\mathscr{D}$ to be coarse in Remark 3.5.5, we ask

Question 1 *Do the previous results hold if we replace* $\mathscr{D}$ *by* $\mathscr{D}_1$ *or* $\mathscr{D}_2$*?*

Another natural question is whether the maps M and H^K_δ coincide.

Question 2 *Is there a dynamical system* (X,d,T) *with an invariant probability measure* $\mu \in \mathscr{P}_T(X)$ *satisfying* $H^K_\delta(\mu) < M(\mu)$*?*

Chapter 4

Discontinuous maps and free semigroup actions

The results in this chapter are based in the paper

[50] G. Pessil: *Metric mean dimension via subshifts of compact type.* Nonlinearity 38 (2025) 045018.

4.1 Preamble

Symbolic dynamics on finite alphabets are classical mathematical objects that have brought forward a great variety of dynamics and intervened in major achievements in Dynamical Systems and Ergodic Theory. For subshifts on finite alphabets the topological entropy quantifies the exponential growth rate of the number of finite words of fixed length. Furthermore, for a Markov subshift the entropy equals the logarithm of the spectral radius of the generating graph, which may be read as the spectral radius of a linear operator.

For compact alphabets this interpretation of complexity in terms of counting words is no longer feasible. Nevertheless, we may still follow the same guiding principle by using a dimensional approach. More precisely, for a given scale, one identifies all but finitely many letters prior to counting and defines the entropy at that scale; the topological entropy is thus obtained by refining the scale. However, in this setting, its value is often infinite; a natural measurement of complexity is, therefore, the rate of such divergence. For instance, the metric mean dimension of a full shift on a compact alphabet, endowed with a suitable metric, is exactly the box dimension of the alphabet (see (2.5)).

There is an asymmetry in the definition of topological entropy since it looks only at the future orbits of points, and the same happens with the metric mean dimension. When the map is invertible, it turns out that its inverse has the same entropy. On the contrary, when the map is not invertible, the definition of entropy cannot be reversed in time. Yet, a non-invertible map induces a shift homeomorphism on the corresponding inverse limit space, and Bowen showed in [8] that the entropy of such a shift homeomorphism is equal to the entropy of the original map.

Generalizing the class of subshifts of finite type, Friedland introduced in [25] the analogous concept within the compact alphabet setting, which we refer to as subshifts of compact type, and proved that the topological entropy of the unilateral and bilateral subshifts induced by a closed transition set coincide. This relation was already known for subshifts of finite type on finite alphabets, in which case the entropy is equal to the spectral radius of the transition matrix. In the particular case where the transition set is the graph of a continuous map, the associated bilateral subshift is a reformulation of the inverse limit of the map. Altogether, Friedland's results generalize Bowen's formula to arbitrary subshifts of compact type. We shall establish analogous properties for the metric mean dimension (see Theorems E and F). As a byproduct, we provide an upper bound for the metric mean dimension of an arbitrary map acting on a compact metric space in terms of spectral radii of scale-dependent matrices (see Section 4.5).

An application of our results, which is of independent interest, is that they motivate a definition of metric mean dimension for discontinuous maps on compact metric spaces in terms of suitable subshifts of compact type. To illustrate the scope of this new concept, we will show that the metric mean dimension of an interval map with infinitely many full branches coincides with the box dimension of its sets of critical points (see Theorem G). In particular, this setting comprises the Gauss map and the Young-induced maps of the Manneville-Pomeau family.

4.1.1 Subshifts of compact type

Let (X,d) be a compact metric space and $\mathbb{K}=\mathbb{N}$ or $\mathbb{Z}$. Consider a subset $\Gamma \subset X\times X$, whose role will be to to prescribe transitions, and the induced set of admissible sequences

$$\Gamma_{\mathbb{K}} = \{(x_i)_{i\in\mathbb{K}} : (x_i,x_{i+1}) \in \Gamma,\ \forall\, i\in\mathbb{K}\},$$

which is shift invariant and we always assume to be non-empty. We refer to $\overline{\Gamma_{\mathbb{K}}}$ as a *subshift of compact type* and note that $\Gamma_{\mathbb{K}}$ is already closed whenever Γ is. This model was proposed by Friedland (cf. [25]) in order to assign a notion of entropy to finitely generated free semigroup actions. We refer the reader to Section 4.3 for more information regarding such actions.

It was proved in [25, Theorem 3.1] that the topological entropy of the unilateral and bilateral subshifts induced by a given closed set coincide. Our first result shows that this also holds in the case of the metric mean dimension.

Theorem E *Let (X,d) be a compact metric space and $\Gamma\subset X\times X$ be a closed subset. Then,*

$$\begin{aligned}\overline{\mathrm{mdim}}_M(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma_{\mathbb{N}}) &= \overline{\mathrm{mdim}}_M(\Gamma_{\mathbb{Z}},d^{\mathbb{Z}},\sigma_{\mathbb{Z}})\\ \underline{\mathrm{mdim}}_M(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma_{\mathbb{N}}) &= \underline{\mathrm{mdim}}_M(\Gamma_{\mathbb{Z}},d^{\mathbb{Z}},\sigma_{\mathbb{Z}}).\end{aligned}$$

A particular instance to which we can apply Theorem E is when we consider $\mathbb{K}=\mathbb{Z}$ and $\Gamma=\mathrm{graph}(T)$, where $T\colon X\to X$ is a continuous map. In this case, the subshift

$$\Gamma_{\mathbb{Z}} = \{(x_i)_{i\in\mathbb{Z}} \colon T(x_i)=x_{i+1}\}$$

is referred to as the *inverse limit of T*. In the entropy setting, Bowen's formula is an immediate consequence of Friedland's [25, Theorem 3.1], since (X,T) and $(\Gamma_{\mathbb{N}},\sigma_{\mathbb{N}})$ are topologically conjugate. Even though the metric mean dimension is not a topological invariant (though it is a bi-Lipschitz one), we show that it is the same for these two systems. The next result summarizes this information.

Theorem F *Let (X,d) be a compact metric space and $T\colon X \to X$ be a continuous map. Then*

$$\begin{aligned} \overline{\mathrm{mdim}}_M(X,d,T) &= \overline{\mathrm{mdim}}_M(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma_{\mathbb{N}}) = \overline{\mathrm{mdim}}_M(\Gamma_{\mathbb{Z}},d^{\mathbb{Z}},\sigma_{\mathbb{Z}}) \\ \underline{\mathrm{mdim}}_M(X,d,T) &= \underline{\mathrm{mdim}}_M(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma_{\mathbb{N}}) = \underline{\mathrm{mdim}}_M(\Gamma_{\mathbb{Z}},d^{\mathbb{Z}},\sigma_{\mathbb{Z}}). \end{aligned} \tag{4.1}$$

Some comments are in order. Clearly, Theorem F hints that it is worthwhile investigating the metric mean dimension of subshifts on compact alphabets, a topic that has recently gained attention (see [28], [30], [56]), though far less known than the finite alphabet case. We remark that Theorems E and F were first proved in the author's Master Thesis [49]. A version of Theorem F for surjective maps was recently published in [10, Lemma 3.8].

4.1.2 Discontinuous maps

An interesting consequence of Theorem F is that, since it expresses the metric mean dimension of a map in terms of its induced subshift, we may regard one of the equalities in (4.1) as a *definition* of the metric mean dimension of an arbitrary (not necessarily continuous) map. The main difference from the continuous setting is that, for a discontinuous map $T\colon X \to X$, the transition set $\Gamma = \mathrm{graph}(T)$ is no longer closed and therefore the sequence space induced by it is not a subshift. However, since the metric mean dimension does not distinguish a set from its closure (see Proposition 2.3.1 (c)), it is natural to consider the subshift of compact type $\overline{\Gamma_{\mathbb{N}}}$, which always satisfies

$$\overline{\mathrm{mdim}}_M(\overline{\Gamma_{\mathbb{N}}},d^{\mathbb{N}},\sigma_{\mathbb{N}}) = \overline{\mathrm{mdim}}_M(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma_{\mathbb{N}}) \quad \text{and} \quad \underline{\mathrm{mdim}}_M(\overline{\Gamma_{\mathbb{N}}},d^{\mathbb{N}},\sigma_{\mathbb{N}}) = \underline{\mathrm{mdim}}_M(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma_{\mathbb{N}}).$$

Definition 4.1.1 *Let (X,d) be a compact metric space, $T\colon X \to X$ be a (not necessarily continuous) map and $\Gamma = \mathrm{graph}(T)$. The upper/lower metric mean dimension of (X,d,T) are given by*

$$\begin{aligned} \overline{\mathrm{mdim}}_M(X,d,T) &= \overline{\mathrm{mdim}}_M(\overline{\Gamma_{\mathbb{N}}},d^{\mathbb{N}},\sigma_{\mathbb{N}}) \\ \underline{\mathrm{mdim}}_M(X,d,T) &= \underline{\mathrm{mdim}}_M(\overline{\Gamma_{\mathbb{N}}},d^{\mathbb{N}},\sigma_{\mathbb{N}}) \end{aligned}$$

For more information about this definition, we refer the reader to Subection 4.6.2. Let us see an interesting use of this concept. It is known (cf. [35, Proposition 15.2.13]) that if $T\colon [0,1] \to [0,1]$ is a piecewise monotone map composed by finitely many, say c, full branches, then

$$h_{top}(T) = \log c. \tag{4.2}$$

Our next result establishes an analogue formula for the metric mean dimension of piecewise monotone maps with infinitely many full branches. In this context, the role of counting branches (which represents the number of critical points of the map) is played by the computation of the box dimension of the (now infinite) set of critical points.

Let $I \subset \mathbb{R}$ be a compact interval and $A \subset I$ be a compact set defined by $A = I \setminus \cup_{k\geq 1} J_k$, where $(J_k)_k$ are pairwise disjoint open intervals ordered non-increasingly in length such that $|I| = \sum_{k\geq 1} |J_k|$. We note that the latter equality implies that A has zero Lebesgue measure; and every compact subset of $\mathbb{R}$ with zero Lebesgue measure has the previous structure. The intervals $(J_k)_k$ are referred to as the *cut-out sets of* A and are deeply related to the box dimension of A (cf. [18, Propositions 3.6 and 3.7]), which can be any value in $[0,1]$.

Consider a map $T_A: I \to I$ satisfying:

$$\begin{aligned}
&\text{(H1)} \quad T_A|_{J_k} \text{ is strictly monotone, for every } k \geq 1.\\
&\text{(H2)} \quad T_A(J_k) = \overset{\circ}{I}, \text{ for every } k \geq 1\\
&\text{(H3)} \quad T_A|_{J_k} \text{ is differentiable and } |T_A'|_{J_k}| \geq \eta > 0, \text{ for some } \eta > 0 \text{ and every } k \geq 1.
\end{aligned} \tag{4.3}$$

The next result relates the metric mean dimension of T_A with the box dimension of A, both with respect to the Euclidean distance $|\cdot|$. Regarding condition (H3), we refer the reader to Subsection 4.6.1.

Theorem G *Let $A \subset I$ be a compact subset with zero Lebesgue measure and $T_A: I \to I$ be any map satisfying conditions* (H1)-(H3) *in* (4.3). *Then*

$$\begin{aligned}
\overline{\text{mdim}}_M(I, |\cdot|, T_A) &= \overline{\dim}_B(A, |\cdot|)\\
\underline{\text{mdim}}_M(I, |\cdot|, T_A) &\leq \underline{\dim}_B(A, |\cdot|).
\end{aligned}$$

Moreover, if $\dim_B(A, |\cdot|)$ *exists then*

$$\text{mdim}_M(I, |\cdot|, T_A) = \dim_B(A, |\cdot|). \tag{4.4}$$

Example 4.1.2 *Consider the* Gauss map

$$\begin{array}{cccc}
G\colon & [0,1] & \to & [0,1]\\
& 0 & \mapsto & 0\\
& x \neq 0 & \mapsto & \frac{1}{x}\,(\textit{mod } 1)
\end{array}$$

to which we apply Theorem G using $A = \{1/n : n \in \mathbb{N}\} \cup \{0\}$ and $\eta = 1$. Thus,

$$\text{mdim}_M([0,1], |\cdot|, G) = \dim_B(A, |\cdot|) = 1/2.$$

Example 4.1.3 *We may also apply Theorem G to Young induced transformations. For instance, let $f_\alpha\colon [0,1] \to [0,1]$ be the Manneville-Pomeau map with parameter $\alpha > 0$, defined as*

$$f_\alpha(x) = \begin{cases} x + 2^\alpha x^{\alpha+1} & \textit{if } 0 \leq x \leq \frac{1}{2}\\ 2x - 1 & \textit{if } \frac{1}{2} < x \leq 1 \end{cases}.$$

In this case, the Young tower induces a transformation $T_\alpha\colon [\frac{1}{2}, 1] \to [\frac{1}{2}, 1]$ (cf. [2, §3.5]) which satisfies conditions (H1)-(H3) with $\eta = 1$ and $A = \{z_n : n \in \mathbb{N}\} \cup \{0\}$, where

$$z_n \in \left[\frac{1}{(n+n_0+1)^{1/\alpha}}, \frac{1}{(n+n_0)^{1/\alpha}}\right]$$

for some $n_0 \in \mathbb{N}$ and every $n \in \mathbb{N}$. Therefore,

$$\mathrm{mdim}_M\big([1/2\,,\,1],|\cdot|,T_\alpha\big) \,=\, \dim_B(A,|\cdot|) \,=\, \frac{\alpha}{1+\alpha}.$$

Regarding the two previous examples, it was proved in [51] that they belong in the class of EMR maps (cf [33, Definition 2.3]). It is also known that the pressure function

$$P(t) \,=\, \lim_{n\to+\infty} \frac{1}{n}\log \sum_{T^n(x)=x} \left(\prod_{i=0}^{n-1} |T'(T^i(x))|^{-t}\right)$$

associated to the parameterized geometric potential possesses a unique transition point

$$s_\infty \,=\, \sup\{t \geq 0 \,:\, P(t) = +\infty\} \,=\, \inf\{t \geq 0 \,:\, P(t) < +\infty\}.$$

This value was computed explicitly in [33, Theorem 2.11], where it was shown to be equal to the box dimension of the corresponding set of discontinuities. By Theorem G we further conclude that, for these maps, s_∞ is precisely the metric mean dimension with respect to the Euclidean distance. More precisely:

Corollary 4.1.4 *Assume that* $\dim_B(A,|\cdot|)$ *exists and that* T_A *satisfies the additional condition*

$$\text{(H4)} \quad \sup_k \sup_{x,y\in J_k} \left|\frac{T'(x)}{T'(y)}\right| \,<\, +\infty.$$

Then

$$\mathrm{mdim}_M(I,|\cdot|,T_A) \,=\, \dim_B(A,|\cdot|) \,=\, s_\infty.$$

Since the transition point s_∞ does not change if we replace T by one of its iterates (cf. [33, Remark 2.13]), under the assumptions of Corollary 4.1.4 we deduce

$$\mathrm{mdim}_M(I,d,T_A) \,=\, \mathrm{mdim}_M(I,d,T_A^k), \qquad \forall\, k \geq 1.$$

4.2 Unilateral and bilateral sequences

In this section we show that the unilateral and bilateral subshifts of compact type generated by a closed transition set have the same metric mean dimension. The next arguments hold for both upper and lower metric mean dimension, so to simplify the notation we will not distinguish them.

In what follows, $\mathbb{K} = \mathbb{N}$ or $\mathbb{Z}$ and, given a closed set $\Gamma \subset X \times X$, we consider the subshift of compact type

$$\Gamma_{\mathbb{K}} \,=\, \{(x_i)_{i\in\mathbb{K}} : (x_i, x_{i+1}) \in \Gamma,\ \forall\, i \in \mathbb{K}\}$$

endowed with the distance

$$d^{\mathbb{K}}(x,y) \,=\, \sup_{i\in\mathbb{K}} \frac{d(x_i,y_i)}{2^{|i|-1}}.$$

Denote $Z = \bigcap_{k=0}^{\infty} \sigma_{\mathbb{N}}^k(\Gamma_{\mathbb{N}})$ and keep the notation $d^{\mathbb{N}}$ and $\sigma_{\mathbb{N}}$ for the distance and dynamics restricted to Z, respectively. By Corollary 2.2.2, to prove Theorem E it is enough to show that

$$\operatorname{mdim}_M(Z, d^{\mathbb{N}}, \sigma_{\mathbb{N}}) = \operatorname{mdim}_M(\Gamma_{\mathbb{Z}}, d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}).$$

Firstly, note that the projection $\pi\big((x_i)_{i\in\mathbb{Z}}\big) = (x_i)_{i\in\mathbb{N}}$ is Lipschitz, $\pi \circ \sigma_{\mathbb{Z}} = \sigma_{\mathbb{N}} \circ \pi$ and $\pi(\Gamma_{\mathbb{Z}}) = Z$. Hence, the restriction

$$\pi\colon \Gamma_{\mathbb{Z}} \longrightarrow Z$$

satisfies the conditions in Proposition 2.3.1 (a), and therefore

$$\operatorname{mdim}_M(Z, d^{\mathbb{N}}, \sigma_{\mathbb{N}}) \leq \operatorname{mdim}_M(\Gamma_{\mathbb{Z}}, d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}).$$

To show the reverse inequality we will make use of the variational principle provided by Theorem 2.2.1. Given a subshift $Y \subset X^{\mathbb{K}}$, a finite set of coordinates $i_1, \cdots, i_k \in \mathbb{K}$ and a collection $U_1, \cdots, U_k \subset X$ of open subsets, the associated cylinder is defined by

$$C(Y; i_1, \cdots, i_k; U_1, \cdots, U_k) = \big\{(x_i)_{i\in\mathbb{K}} \in Y \colon x_{i_j} \in U_j, \forall j = 1, \cdots, k\big\}. \tag{4.5}$$

These sets are generators of the topology of Y and every dynamical ball is a cylinder. Thus, for any ergodic measure $\nu \in \mathscr{E}_{\sigma_{\mathbb{K}}}(Y)$, we can compute

$$N_\nu(Y, d^{\mathbb{K}}, \sigma_{\mathbb{K}}, n, \varepsilon, \delta) = \inf_B \big\{S(B, d_n^{\mathbb{K}}, \varepsilon) \colon \nu(B) > 1 - \delta\big\}$$

by taking the infimum over finite unions of cylinders in Y whose union has measure ν bigger than $1-\delta$. Note that, for every ergodic measure $\mu \in \mathscr{E}_{\sigma_{\mathbb{Z}}}(\Gamma_{\mathbb{Z}})$, its push-forward $\pi_*\mu \in \mathscr{E}_{\sigma_{\mathbb{N}}}(Z)$ is ergodic. Moreover, every finite union of cylinders in Z with measure $\pi_*\mu$ bigger than $1-\delta$ has as a pre-image by π a finite union of cylinders with measure μ bigger than $1-\delta$. Consequently,

$$N_\mu(\Gamma_{\mathbb{Z}}, d_{\mathbb{Z}}, \sigma_{\mathbb{Z}}, n, \varepsilon, \delta) = \inf_{\substack{B\subset\Gamma_{\mathbb{Z}}\\ \mu(B)>1-\delta}} S(B, d_n^{\mathbb{Z}}, \varepsilon) \leq \inf_{\substack{A\subset Z\\ \pi_*\mu(A)>1-\delta}} S(\pi^{-1}(A), d_n^{\mathbb{Z}}, \varepsilon). \tag{4.6}$$

We proceed by estimating the right hand side in (4.6). For each $\varepsilon > 0$, let $\ell = \ell(\varepsilon)$ be a positive integer such that $2^{-\ell} \operatorname{diam} X < \varepsilon$ and $\mathscr{U} = \{U_1, ..., U_{M(\varepsilon)}\}$ be any open cover of X satisfying $\operatorname{diam}(\mathscr{U}, d) < \varepsilon$. Given an open cover α of a set $A \subset Z$ such that $\operatorname{diam}(\alpha, d_n^{\mathbb{N}}) \leq \varepsilon$, then

$$\beta = \{... \times X \times X \times U_{j_{-\ell}} \times ... \times U_{j_0} \times V : V \in \alpha, \forall j_i = 1, ..., M(\varepsilon)\}.$$

is an open cover of $\pi^{-1}(A) \subset \Gamma_{\mathbb{Z}}$ satisfying $\operatorname{diam}(\beta, d_n^{\mathbb{Z}}) \leq \varepsilon$ and $\#\beta = \big(M(\varepsilon)\big)^{\ell+1}\#\alpha$. Thus, for every $A \subset Z$ we have

$$S(\pi^{-1}(A), d_n^{\mathbb{Z}}, \varepsilon) \leq \big(M(\varepsilon)\big)^{\ell+1} S(A, d_n^{\mathbb{N}}, \varepsilon).$$

Combining this information with (4.6) we get

$$\begin{aligned} N^{\mathbb{Z}}_{\mu}(\Gamma_{\mathbb{Z}}, d_{\mathbb{Z}}, \sigma_{\mathbb{Z}}, n, \varepsilon, \delta) &\leq \inf_{\substack{A \subset Z \\ \pi_*\mu(A) > 1-\delta}} S(\pi^{-1}(A), d^{\mathbb{Z}}_n, \varepsilon) \\ &\leq \big(M(\varepsilon)\big)^{\ell+1} \inf_{\substack{A \subset Z \\ \pi_*\mu(A) > 1-\delta}} S(A, d^{\mathbb{N}}_n, \varepsilon) \\ &= \big(M(\varepsilon)\big)^{\ell+1} N^{\mathbb{N}}_{\pi_*\mu}(Z, d_{\mathbb{N}}, \sigma_{\mathbb{N}}, n, \varepsilon, \delta). \end{aligned}$$

Hence, for every ergodic probability measure $\mu \in \mathscr{E}_{\sigma_{\mathbb{Z}}}(\Gamma_{\mathbb{Z}})$ we have

$$h^K_{\mu}(\Gamma_{\mathbb{Z}}, d_{\mathbb{Z}}, \sigma_{\mathbb{Z}}, \varepsilon, \delta) \leq h^K_{\pi_*\mu}(Z, d_{\mathbb{N}}, \sigma_{\mathbb{N}}, \varepsilon, \delta)$$

and therefore

$$\sup_{\mu \in \mathscr{E}_{\sigma_{\mathbb{Z}}}(\Gamma_{\mathbb{Z}})} h^K_{\mu}(\Gamma_{\mathbb{Z}}, d_{\mathbb{Z}}, \sigma_{\mathbb{Z}}, \varepsilon, \delta) \leq \sup_{\nu \in \mathscr{E}_{\sigma_{\mathbb{N}}}(Z)} h^K_{\nu}(Z, d_{\mathbb{N}}, \sigma_{\mathbb{N}}, \varepsilon, \delta).$$

Applying now Theorem 2.2.1, we obtain

$$\mathrm{mdim}_M(\Gamma_{\mathbb{Z}}, d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}) \leq \mathrm{mdim}_M(Z, d^{\mathbb{N}}, \sigma_{\mathbb{N}})$$

and the proof of the proposition is complete. $\square$

4.3 Free semigroup actions

In this section we recall Friedland's topologial entropy of a free semigroup action of continuous maps on a compact metric space X, which inspires a corresponding notion of metric mean dimension. Afterwards, we present the proposal in [12] for the metric mean dimension of a free semigroup action with respect to a fixed random walk. We proceed by establishing a variational principle connecting these concepts, from which Theorem F is a direct consequence.

Let (X,d) be a compact metric space and $G_1 = \{g_1, \cdots, g_p\}$ be a family of continuous maps. Denote by G the free semigroup having G_1 as a generator, where the semigroup operation $\circ$ is the composition of maps. Let $\mathbb{S}$ be the induced free semigroup action

$$\begin{aligned} \mathbb{S} : G \times X &\to X \\ (g,x) &\mapsto g(x). \end{aligned}$$

Denoting the index set of G_1 by $Y = \{1, \cdots, p\}$, we endow the product space $Y^{\mathbb{N}}$ with the metric

$$d'(w, \omega) = 2^{-\min\{j \geq 1 \,:\, w_j \neq \omega_j\}}$$

and consider the skew-product T_G associated to the action $\mathbb{S}$:

$$\begin{aligned} T_G \colon \; Y^{\mathbb{N}} \times X &\to Y^{\mathbb{N}} \times X \\ (w,x) &\mapsto (\sigma(w), g_{w_1}(x)), \end{aligned}$$

where $w = (w_1, w_2, \cdots)$.

4.3.1 Friedland's approach

For a finite set of continuous maps $G_1 = \{g_1, \cdots, g_p\}$, we consider the transition set

$$\Gamma(G_1) = \bigcup_{i=1}^{p} \mathrm{graph}(g_i)$$

and the associated subshift of compact type

$$\Gamma(G_1)_{\mathbb{N}} = \big\{(x, g_{w_1}(x), \cdots, g_{w_k} \circ \ldots \circ g_{w_1}(x), \cdots) : w \in Y^{\mathbb{N}}\big\}.$$

In what follows, the composition operation between two maps will be denoted by their concatenation.

Definition 4.3.1 *The Friedland topological entropy of* $\mathbb{S}$ *with respect to* G_1 *is defined as*

$$h^F_{top}(X, \mathbb{S}, G_1) = h(\Gamma(G_1)_{\mathbb{N}}, \sigma).$$

Actually, Friedland defined the entropy of $\mathbb{S}$ as the infimum of $h^F_{top}(X, \mathbb{S}, G_1)$ over all finite sets of generators G_1 of G. As both definitions that we consider are expressed in terms of a particular set of generators, we keep this dependence and omit G_1 from the notation. We now define the corresponding notion for the metric mean dimension.

Definition 4.3.2 *The upper/lower Friedland metric mean dimension of* $(X, d, \mathbb{S})$ *with respect to the set of generators* G_1 *is given by*

$$\begin{aligned} \overline{\mathrm{mdim}}^F_M(X, d, \mathbb{S}) &= \overline{\mathrm{mdim}}_M(\Gamma(G_1)_{\mathbb{N}}, d^{\mathbb{N}}, \sigma) \\ \underline{\mathrm{mdim}}^F_M(X, d, \mathbb{S}) &= \underline{\mathrm{mdim}}_M(\Gamma(G_1)_{\mathbb{N}}, d^{\mathbb{N}}, \sigma). \end{aligned} \tag{4.7}$$

By Theorem E, the previous definition is independent of whether $\mathbb{K} = \mathbb{N}$ or $\mathbb{K} = \mathbb{Z}$. The aim of Theorem F is to ensure that it coincides with Lindenstrauss-Weiss' one in the case when $G_1 = \{T\}$.

4.3.2 Carvalho, Rodrigues and Varandas' approach

Now we present the definition of metric mean dimension introduced in [14]. This perspective is inspired by Bufetov (see [9]), where one selects randomly which element of G_1 will be used to evolve time on each step.

Following [14], we shall code different concatenations of elements in G_1 by distinct sequences of symbols. Even though an element of the group may be generated by different combinations of generators, we will simply consider different concatenations instead of the elements in the group they create, since we do not make use of the group structure.

For every $w \in Y^{\mathbb{N}}$ and $n \geq 1$, consider the metric in X given by

$$d_{w,n}(x, y) = \max_{0 \leq j \leq n} d(g_{w_j} g_{w_{j-1}} \ldots g_{w_1}(x), g_{w_j} g_{w_{j-1}} \ldots g_{w_1}(y)).$$

Fix a *random walk*, that is, a shift invariant Borel probability measure $\mathbb{P} \in \mathscr{P}_\sigma(Y^{\mathbb{N}})$. The $\varepsilon-$*entropy* of $(X,d,\mathbb{S},\mathbb{P})$ is given by

$$h_\varepsilon(X,d,\mathbb{S},\mathbb{P}) = \limsup_{n\to+\infty} \frac{1}{n} \log \int_{Y^{\mathbb{N}}} S(X,d_{w,n},\varepsilon) d\mathbb{P}(w).$$

The *topological entropy of* $\mathbb{S}$ *with respect to the set of generators* G_1 (cf. [14]) is given by

$$h_{top}(X,\mathbb{S}) = \sup_{\mathbb{P}} \lim_{\varepsilon\to 0^+} h_\varepsilon(X,d,\mathbb{S},\mathbb{P}).$$

Definition 4.3.3 *The upper/lower metric mean dimension of* $(X,d,\mathbb{S})$ *with respect to the set of generators* G_1 *and the random walk* $\mathbb{P}$ *are given, respectively, by*

$$\begin{aligned} \overline{\mathrm{mdim}}_M(X,d,\mathbb{S},\mathbb{P}) &= \limsup_{\varepsilon\to 0^+} \frac{h_\varepsilon(X,d,\mathbb{S},\mathbb{P})}{\log(1/\varepsilon)} \\ \underline{\mathrm{mdim}}_M(X,d,\mathbb{S},\mathbb{P}) &= \liminf_{\varepsilon\to 0^+} \frac{h_\varepsilon(X,d,\mathbb{S},\mathbb{P})}{\log(1/\varepsilon)}. \end{aligned}$$

Once again, one can replace S in the definition by either S_1 or S_2 as in (2.2).

Definition 4.3.4 *Given a compact metric space M, we say that* $\nu \in \mathscr{P}(M)$ *is homogeneous if it is a fully supported doubling measure, namely,*

$$\exists L>0\colon \qquad \nu(B(x,\varepsilon)) \le L\,\nu(B(y,\varepsilon/2)), \qquad \forall x,y\in M,\ \forall \varepsilon>0.$$

Note that if a random walk $\mathbb{P} \in \mathscr{P}_\sigma(Y^{\mathbb{N}})$ is homogeneous, then

$$\mathbb{P}(B(w,2^{-k+1})) \ge (1/L)^k\, \mathbb{P}(B(w,1/2)) = (1/L)^k,$$

for every $w \in Y^{\mathbb{N}}$ and $k \ge 1$, where

$$B(w,2^{-k+1}) = \{\omega \in Y^{\mathbb{N}}\colon\ \omega_i = w_i, i=1,...,k\}.$$

In what follows, we adopt the notation $\overline{w_1...w_k} = B(w,2^{-k+1})$.

4.3.3 Variational principle for free semigroup actions

Our main result in this section states that the two aforementioned notions of metric mean dimension of finitely generated free semigroup actions are equal. Theorem F corresponds to the particular case of $G_1 = \{T\}$.

Theorem 4.3.5 *Let* (X,d) *be a compact metric space and* $G_1 = \{g_1,\cdots,g_p\}$ *be a collection of continuous maps in* X*. Then,*

$$\begin{aligned} \overline{\mathrm{mdim}}^F_M(X,d,\mathbb{S}) &= \max_{\mathbb{P}} \overline{\mathrm{mdim}}_M(X,d,\mathbb{S},\mathbb{P}) \\ \underline{\mathrm{mdim}}^F_M(X,d,\mathbb{S}) &= \max_{\mathbb{P}} \underline{\mathrm{mdim}}_M(X,d,\mathbb{S},\mathbb{P}). \end{aligned} \tag{4.8}$$

Moreover, the previous maxima are attained at every *homogeneous random walk* $\mathbb{P}$ *and we have*

$$\overline{\mathrm{mdim}}_M^F(X,d,\mathbb{S}) \;=\; \overline{\mathrm{mdim}}_M(X\times I^{\mathbb{N}},\, d\times d',\, T_G). \tag{4.9}$$

The equality (4.9) is a direct consequence of the fact that the maximum in (4.8) is attained by any fully supported Bernoulli measure $\mathbb{P}$ - which is homogeneous - combined with [14, Corollary III]. In fact, the choice of metric in $Y^{\mathbb{N}}$ considered in [14] is different from our choice d', but since both metrics are uniformly equivalent so are their products with d and therefore their corresponding metric mean dimensions of the skew-product coincide.

Remark 4.3.6 *We observe that a statement analogous to Theorem 4.3.5 for the topological entropy of a semigroup action does* not *hold. Friedland showed [25, Lemma 3.6] that if T is a non-involutive homeomorphism ($T^2 \neq Id$) of a compact metric space, then $h^F_{top}(\mathbb{S}) \geq \log 2$. If we consider the phase space to be the unit circle, both T and T^{-1} have zero topological entropy, and since compositions of them only represent delay in time, we have $h_{top}(\mathbb{S},\mathbb{P}) = 0$ for every $\mathbb{P}$. So*

$$h^F_{top}(\mathbb{S}) > \max_{\mathbb{P}} h_{top}(\mathbb{S},\mathbb{P}).$$

Actually, the Friedland subshift is a Lipschitz factor of the skew-product map - in particular,

$$h_\varepsilon(\Gamma_{\mathbb{N}}, d^{\mathbb{N}}, \sigma) \leq h_\varepsilon(T_G)$$

- and the entropy of the action relates to the one of the skew product by Bufetov's formula (cf. [9]): if $\mathbb{P}_p$ is the symmetric Bernoulli measure in the p symbols of Y, then

$$h_{top}(\mathbb{S},\mathbb{P}_p) = h_{top}(T_G) - \log p.$$

Whereas, in the case of metric mean dimension, the normalization $\log(1/\varepsilon)$ *in its definition makes the second term in the right hand side vanish and we obtain* (4.8) *and* (4.9).

Proof of Theorem 4.3.5: We start by proving that the Friedland metric mean dimension is an upper bound for the one induced by any random walk. Let $w \in Y^{\mathbb{N}}$. For each $n \geq 1$, consider $E \subset X$ a maximal $\varepsilon-$separated subset with respect to the metric $d_{w,n}$. That is, for every $x, y \in E$ we have

$$\varepsilon < \max\left\{d(x,y), d(g_{w_1}(x), g_{w_1}(y)), \cdots, d(g_{w_n}\cdots g_{w_1}(x), g_{w_n}\cdots g_{w_1}(y))\right\}.$$

Let us denote simply $\Gamma = \Gamma(G_1)$. Then, the elements of $\Gamma_{\mathbb{N}}$

$$\big(x, g_{w_1}(x), \cdots, g_{w_n}\cdots g_{w_1}(x), \text{any admissible tail}\big)$$
$$\big(y, g_{w_1}(y), \cdots, g_{w_n}\cdots g_{w_1}(y), \text{any admissible tail}\big)$$

are (σ,n,ε) separated. Hence, for every $\mathbb{P}\in\mathscr{P}_\sigma(Y^{\mathbb{N}})$ we have

$$\begin{aligned}
h_\varepsilon(X,d,\mathbb{S},\mathbb{P}) &= \limsup_{n\to+\infty}\frac{1}{n}\log\int_{Y^{\mathbb{N}}} S_1(X,d_{w,n},\varepsilon)\,d\mathbb{P}(w)\\
&\le \limsup_{n\to+\infty}\frac{1}{n}\log S_1(\Gamma_{\mathbb{N}},d_n^{\mathbb{N}},\varepsilon)\\
&= h_\varepsilon(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma)
\end{aligned}$$

Now let us show that the maximum is attained by any homogeneous $\mathbb{P}\in\mathscr{P}_\sigma(Y^{\mathbb{N}})$. Take one such $\mathbb{P}$. Then, for every $w\in Y^{\mathbb{N}}$ and $k\ge 1$, we have

$$\mathbb{P}(\overline{w_1\ldots w_k}) \ge (1/L)^k.$$

Given $n\ge 1$ and $\varepsilon>0$, let $E=\{x_1,x_2,\ldots,x_N\}$ be a maximal $(\sigma,n,\varepsilon)-$separated subset in $\Gamma_{\mathbb{N}}$, that is, $N=S_1(\Gamma_{\mathbb{N}},d_n^{\mathbb{N}},\varepsilon)$. Take $\ell=\ell(\varepsilon)$ such that $2^{-\ell}\operatorname{diam}(X)\le\varepsilon$. By the Pigeonhole Principle, there exists a subset $A\subset E$ such that

- all the transitions until the coordinate $n+\ell$ are given by the same $(n+\ell)-$tuple, which we denote by $(g_{w_1},g_{w_2},\ldots,g_{w_{n+\ell}})$;
- $a:=|A|\ge\frac{N}{p^{n+\ell}}$.

Note that, by definition of ℓ, the property of being separated only depends on the first $n+\ell$ coordinates of the elements of E. In particular, the first coordinate of all the elements in A are necessarily different from each other. Reordering if necessary, we denote the elements of A by

$$\begin{aligned}
x_1 &= \big(x^1,\, g_{w_1}(x^1),\, g_{w_2}g_{w_1}(x^1),\,\cdots,g_{w_{n+\ell}}\cdots g_{w_1}(x^1),\text{any admissible tail}\big)\\
x_2 &= \big(x^2,\, g_{w_1}(x^2),\, g_{w_2}g_{w_1}(x^2),\,\cdots,g_{w_{n+\ell}}\cdots g_{w_1}(x^2),\text{any admissible tail}\big)\\
&\;\;\vdots\\
x_a &= \big(x^a,\, g_{w_1}(x^a),\, g_{w_2}g_{w_1}(x^a),\,\cdots,g_{w_{n+\ell}}\cdots g_{w_1}(x^a),\text{any admissible tail}\big).
\end{aligned}$$

Then, the set $\{x^1,\ldots,x^a\}\subset X$ is $\varepsilon-$separated with respect to the metric $d_{\omega,n+\ell}$ for every $\omega\in\overline{w_1\ldots w_{n+\ell}}\subset Y^{\mathbb{N}}$. In particular, the balls of radius $\varepsilon/3$ with respect to $d_{\omega,n+\ell}$ centered at the elements of A are pairwise disjoint. Hence,

$$S_2(X,d_{\omega,n},\varepsilon/3)\ge a\ge\frac{N}{p^{n+\ell}}$$

for every $\omega\in\overline{w_1\ldots w_{n+\ell}}$. So,

$$\begin{aligned}
\int_{Y^{\mathbb{N}}} S_2(X,d_{\omega,n},\varepsilon/3)\,d\mathbb{P}(\omega) &\ge \int_{\overline{w_1\ldots w_{n+\ell}}} S_2(X,d_{\omega,n},\varepsilon/3)\,d\mathbb{P}(\omega)\\
&\ge \frac{N}{p^{n+\ell}}\,\mathbb{P}(\overline{w_1\ldots w_{n+\ell}})\\
&\ge \frac{N}{p^{n+\ell}}\,(1/L)^{n+\ell}.
\end{aligned}$$

Therefore,

$$\begin{aligned} h_{\varepsilon/3}(X,d,\mathbb{S},\mathbb{P}) &= \limsup_{n\to+\infty} \frac{1}{n} \log \int_{Y^{\mathbb{N}}} S_2(X,d_{\omega,n},\varepsilon/3)\, d\mathbb{P}(\omega) \\ &\geq h_\varepsilon(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma) - \log(pL). \end{aligned}$$

We conclude after dividing by $\log(1/\varepsilon)$ and making $\varepsilon \to 0^+$.

$\square$

4.3.4 Zero complexity maps may generate positive metric mean dimension

In this subsection we will make use of Theorem 4.3.5 to show that two maps with zero metric mean dimension may generate a free semigroup action with positive metric mean dimension. Let $X = [0,1]$ endowed with the Euclidean distance, which to ease notation in the proofs will be denoted by d.

Proposition 4.3.7 *There exist two zero entropy continuous maps $g_1, g_2\colon [0,1] \to [0,1]$ satisfying*

$$\mathrm{mdim}_M^F([0,1],d,\mathbb{S}) = \frac{1}{2}.$$

In particular, the topological entropy of the action is infinite.

Proof: Let $T\colon [0,1] \to [0,1]$ be the map $T_{a,b}$ defined in Subsection 2.4.2 with parameters $a = \left(\sum_{n=1}^{k} 6/\pi n^2\right)_k$ and $b = (3^{k-1})_k$. For such parameters, we have

$$\mathrm{mdim}_M([0,1],|\cdot|,T) = 1.$$

We consider the maps $g_1, g_2\colon [0,1] \to [0,1]$

$$g_1(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq \frac{1}{2} \\ \frac{T(2x-1)}{2} & \text{if } \frac{1}{2} \leq x \leq 1 \end{cases} \qquad g_2(x) = \begin{cases} x+\frac{1}{2} & \text{if } 0 \leq x \leq \frac{1}{2} \\ 1 & \text{if } \frac{1}{2} \leq x \leq 1 \end{cases}.$$

The transition set Γ which generates the Friedland subshift is illustrated in Figure 4.1.

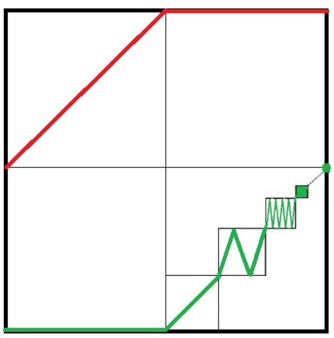

Fig. 4.1 $\Gamma = \mathrm{graph}(g_1) \cup \mathrm{graph}(g_2)$.

Since $g_1^2 \equiv 0$ and $g_2^2 \equiv 1$, it is clear that their topological entropy is zero, and so both maps have zero metric mean dimension. We also observe that - denoting the Bowen metric of time k with respect to a map ϕ by d_k^{ϕ} - we have (see Figure 4.2)

$$S_1([0,1], d_n^{g_1 \circ g_2}, \varepsilon) = S_1([0,1], d_n^{g_2 \circ g_1}, \varepsilon) = S_1([0,1], d_n^T, 2\varepsilon) + \left\lfloor \frac{1}{2\varepsilon} \right\rfloor.$$

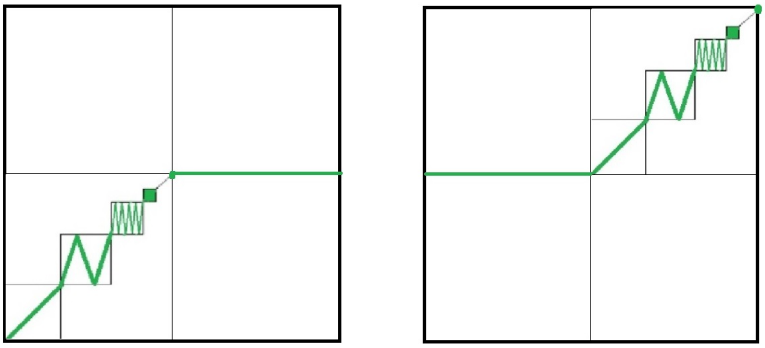

Fig. 4.2 Graphs of $g_1 \circ g_2$ and $g_2 \circ g_1$, respectively.

In order to prove the lower bound

$$\mathrm{mdim}_M^F([0,1], d, \mathbb{S}) \geq \frac{1}{2},$$

we consider the random walk $\mathbb{P} = \frac{1}{2}\big(\delta_{(12)^{\mathbb{N}}} + \delta_{(21)^{\mathbb{N}}}\big)$. Fix $w = (12)^{\mathbb{N}}$ (the case $w = (21)^{\mathbb{N}}$ is analogous). Since

$$d_{w,2n}(x_0, y_0) \geq d_n^{g_2 \circ g_1}(x_0, y_0), \qquad \forall x_0, y_0 \in [0,1],\ \forall n \geq 1,$$

we have

$$S_1([0,1], d_{w,2n}, \varepsilon) \geq S_1([0,1], d_n^{g_2 \circ g_1}, \varepsilon) = S_1([0,1], d_n^T, 2\varepsilon) + \left\lfloor \frac{1}{2\varepsilon} \right\rfloor$$

and so

$$h_\varepsilon([0,1], d, \mathbb{S}, \mathbb{P}) = \limsup_{n \to +\infty} \frac{1}{2n} \log \int_{\{0,1\}^{\mathbb{N}}} S_1([0,1], d_{w,2n}, \varepsilon)\, d\mathbb{P}(w) \geq \frac{h_{2\varepsilon}([0,1], d, T)}{2},$$

which implies that

$$\frac{1}{2} \leq \underline{\mathrm{mdim}}_M([0,1], d, \mathbb{S}, \mathbb{P}) \leq \underline{\mathrm{mdim}}_M^F([0,1], d, \mathbb{S}).$$

Now we proceed to prove the converse inequality. By [58, Theorem 1.1], for every $\delta > 0$ we have

$$\overline{\mathrm{mdim}}_M(\Gamma_{\mathbb{Z}}, d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}) = \limsup_{\varepsilon \to 0^+} \frac{1}{\log(1/\varepsilon)} \sup_{x \in \Gamma_{\mathbb{Z}}} h_\varepsilon(B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}}), d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}), \tag{4.10}$$

where $x = (x_j)_{j \in \mathbb{Z}}$

$$B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}}) = \{y \in \Gamma_{\mathbb{Z}} : d^{\mathbb{Z}}(\sigma^j(x), \sigma^j(y)) \leq \delta,\ \forall j \in \mathbb{Z}\} = \{y \in \Gamma_{\mathbb{Z}} : d(x_j, y_j) \leq \delta,\ \forall j \in \mathbb{Z}\}.$$

Fix $0 < \delta < |J_1|/6$. Since $\delta \ll (d \times d)(\mathrm{graph}(g_1), \mathrm{graph}(g_2))$, every $y \in B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}})$ must be prescribed by the same transitions as x, namely, there exists $w(x) \in \{0,1\}^{\mathbb{Z}}$ such that $y_{k+1} = g_{w(x)_k}(y_k)$ for every $k \in \mathbb{Z}$ and $y \in B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}})$. Let us estimate $h_\varepsilon(B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}}), d^{\mathbb{Z}}, \sigma_{\mathbb{Z}})$ separately for each type of $w(x)$.

Case 1: $(21)^{\mathbb{Z}} \neq w(x) \neq (12)^{\mathbb{Z}}$

Let $k_0 \in \mathbb{Z}$ be such that $w(x)_{k_0} = w(x)_{k_0-1}$. Then we have $y_k = x_k$ for every $k \geq k_0$ and $y \in B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}})$. Then,

$$h_\varepsilon(B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}}), d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}) = 0. \tag{4.11}$$

Case 2: $w(x) = (\cdots, 1, 2; 1, 2, 1, 2, \cdots)$

(a) $x_0 \in [0, 1/2 + 2\delta]$.
In this case, $\pi_0\Big(B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}})\Big) \subset [0, 1/2 + 3\delta]$. Given $y \in B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}})$, we have two possibilities

- If $y_0 \in [0, \frac{1}{2}]$, then $y_1 = g_1(y_0) = 0$. This implies that $y_{2k-1} = 0$ and $y_{2k} = 1/2$, for every $k \geq 1$. Then, $\pi_{\mathbb{N}}\Big(\pi_0^{-1}[0,\, 1/2] \cap B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}})\Big) = (0, 1/2, 0, 1/2, \ldots)$.
- If $y_0 \in (1/2, 1/2 + 3\delta]$, then $y_1 = g_1(y_0) = y_0 - 1/2$ and $y_2 = g_2(y_1) = y_0$ This implies that $y_{2k-1} = y_0 - 1/2$ and $y_{2k} = y_0$, for every $k \geq 1$. Hence,

$$\pi_{\mathbb{N}}\Big(\pi_0^{-1}(1/2, 1/2 + 3\delta] \cap B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}})\Big) \subset \{(y_0 - 1/2, y_0, y_0 - 1/2, y_0, \ldots) : y_0 \in (1/2, 1/2 + 3\delta]\}.$$

Thus,

$$\pi_{\mathbb{N}}\Big(B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}})\Big) \subset \{(y_0 - 1/2,\, y_0,\, y_0 - 1/2,\, y_0, \ldots) : y_0 \in [1/2,\, 1/2 + 2\delta]\},$$

which implies that

$$h_\varepsilon(B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}}), d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}) = 0.$$

(b) $x_0 \in [1/2 + 2\delta, 1]$.

Take $\ell = \ell(\varepsilon)$ such that $2^{-\ell} \leq \varepsilon$, assuming without loss that ℓ is even. Given $y, z \in B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}})$, we have $d(y_{2k}, z_{2k}) = d(y_{2k-1}, z_{2k-1})$ for every $k \in \mathbb{Z}$. If $d_{2n}^{\mathbb{Z}}(y, z) > \varepsilon$, we get $d_{n+\ell+1}^{g_2 \circ g_1}(y_{-\ell}, z_{-\ell}) > \varepsilon$, since the separation between y and z must occur between the coordinates $-\ell$ and $\ell + 2n$ and we may restrict to even times. In particular, this implies that

$$S_1(B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}}), d_{2n}^{\mathbb{Z}}, \varepsilon) \leq S_1([0,1], d_{n+\ell+1}^{g_2 \circ g_1}, \varepsilon) = S_1([0,1], d_{n+\ell+1}^{T}, 2\varepsilon) + \left\lfloor \frac{1}{2\varepsilon} \right\rfloor, \qquad \forall n \geq 1.$$

Hence

$$h_\varepsilon(B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}}), d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}) \leq \frac{h_{2\varepsilon}([0,1], d, T)}{2}.$$

Bringing together sub-cases (a) and (b) we conclude that for every $x \in \Gamma_{\mathbb{Z}}$ such that $w(x) = (\cdots, 1, 2; 1, 2, 1, 2, \cdots)$ one has

$$h_\varepsilon(B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}}), d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}) \leq \frac{h_{2\varepsilon}([0,1], d, T)}{2}. \tag{4.12}$$

Case 3: $w(x) = (\cdots, 1, 2, 1; 2, 1, 2, \cdots)$
Since $\sigma_{\mathbb{Z}}$ is a homeomorphism, $\sigma(w(x)) = w(\sigma_{\mathbb{Z}}(x))$ and $B_\delta(\sigma_{\mathbb{Z}}(x), d_{\mathbb{Z}}^{\mathbb{Z}}) = \sigma_{\mathbb{Z}}\Big(B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}})\Big)$, we have

$$\begin{aligned} h_\varepsilon(B_\delta(x, d_{\mathbb{Z}}^{\mathbb{Z}}), d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}) &= h_\varepsilon\Big(\sigma_{\mathbb{Z}}^{-1}(B_\delta(\sigma_{\mathbb{Z}}(x), d_{\mathbb{Z}}^{\mathbb{Z}}), d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}\Big) \\ &= h_\varepsilon\Big(B_\delta(\sigma_{\mathbb{Z}}(x), d_{\mathbb{Z}}^{\mathbb{Z}}), d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}\Big) \\ &\leq \frac{h_{2\varepsilon}([0,1], d, T)}{2}, \end{aligned} \tag{4.13}$$

where the inequality follows from Case 2. Combining (4.10), (4.11),(4.12) and (4.13) we deduce that

$$\overline{\mathrm{mdim}}_M^F([0,1], d, \mathbb{S}) = \overline{\mathrm{mdim}}_M(\Gamma_{\mathbb{Z}}, d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}) \leq \frac{1}{2}.$$

This finishes the proof of Proposition 4.3.7. □

4.4 Local metric mean dimension for non-invertible dynamics

Under appropriate adjustments, Theorem F allows us to extend properties valid for invertible maps to non-invertible ones. Let's make use of it to complete the proof of Theorem D, which was only proved under the assumption that the map was a homeomorphism. Given a continuous map T acting on a compact metric space (X,d), recall that the *local metric mean dimension function* is defined as

$$\begin{aligned} \mathscr{D}: X &\to \mathbb{R} \\ x &\mapsto \inf\Big\{\overline{\mathrm{mdim}}_M(U, d, T) \colon U \text{ is an open neighbourhood of } x\Big\}. \end{aligned} \tag{4.14}$$

The map $\mathscr{D}$ is upper semi-continuous (Lemma 3.5.1) and constant $\mu-$almost everywhere with respect to any ergodic probability measure (Lemma 3.5.2). To complete the proof of Theorem D, we can assume that it holds for homeomorphisms, since this was already proved in Section 3.5.

Proposition 4.4.1 *Let (X,d) be a compact metric space and $T: X \to X$ be a continuous map such that $\overline{\mathrm{mdim}}_M(X,d,T) < +\infty$. Then*

$$\overline{\mathrm{mdim}}_M(X,d,T) = \max_{\mu \in \mathscr{P}_T(X)} \int \mathscr{D}(x)\, d\mu(x) = \max_{\mu \in \mathscr{E}_T(X)} \int \mathscr{D}(x)\, d\mu(x). \tag{4.15}$$

Proof: Recall that, given $p \in \Gamma_{\mathbb{Z}}$ and $q \in X$,

$$\begin{aligned} \mathscr{D}_\Gamma(p) &= \inf\Big\{\overline{\mathrm{mdim}}_M(U, d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}) \colon U \subset \Gamma_{\mathbb{Z}} \text{ is an open neighbourhood of p}\Big\} \\ \mathscr{D}(q) &= \inf\Big\{\overline{\mathrm{mdim}}_M(U, d, T) \colon U \subset X \text{ is an open neighbourhood of } q\Big\} \end{aligned}$$

Let $\mu \in \mathscr{P}_{\sigma_{\mathbb{Z}}}(\Gamma_{\mathbb{Z}})$ be such that every $x = (x_i)_{i\in\mathbb{Z}} \in \mathrm{supp}(\mu)$ satisfies

$$\mathscr{D}_\Gamma(x) = \overline{\mathrm{mdim}}_M(\Gamma_{\mathbb{Z}}, d^{\mathbb{Z}}, \sigma_{\mathbb{Z}})$$

which always exists by Theorem D (since $\sigma_{\mathbb{Z}}$ is a homeomorphism). Fix one such x. Then, for any $\gamma > 0$, we have

$$\overline{\text{mdim}}_M(U_\gamma, d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}) = \overline{\text{mdim}}_M(\Gamma_{\mathbb{Z}}, d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}),$$

where $U_\gamma = \{y \in \Gamma_{\mathbb{Z}} : y_1 \in B_\gamma(x_1)\}$. Given $\varepsilon > 0$, let $\ell = \ell(\varepsilon)$ be a positive integer satisfying $2^{-\ell}\,\text{diam}(X,d) \leq \varepsilon$. For $n \geq 1$, consider a maximal $(n,\varepsilon)-$separated subset $E \subset U_\gamma$. Then, there exist $z^1, \ldots, z^a \in E$ such that

- $d(z_j^{m_1}, z_j^{m_2}) \leq \varepsilon$, for every $j = -\ell, \ldots, 0$ and $m_1 \neq m_2$
- $a \geq \#E \,/\, S(X,d,\varepsilon)^{\ell+1}$.

Hence

$$S_1(B_\gamma(x_1), d_n, \varepsilon) \geq \#E \,/\, S(X,d,\varepsilon)^{\ell+1} = S_1(U_\gamma, d_n^{\mathbb{Z}}, \varepsilon) \,/\, S(X,d,\varepsilon)^{\ell+1},$$

and so

$$\mathscr{D}(x_1) \geq \lim_{\gamma \to 0^+} \overline{\text{mdim}}_M(U_\gamma, d, T) \geq \overline{\text{mdim}}_M(\Gamma_{\mathbb{Z}}, d^{\mathbb{Z}}, \sigma_{\mathbb{Z}}) = \overline{\text{mdim}}_M(X, d, T),$$

where the last equality is due to Theorem F. On the other hand, by definition, the reverse inequality always holds. Hence,

$$\mathscr{D}(x_1) = \overline{\text{mdim}}_M(X, d, T).$$

Now, since $\text{supp}(\mu) \subset \Gamma_{\mathbb{Z}}$ is $\sigma_{\mathbb{Z}}-$invariant, we may carry on the same reasoning with x_1 replaced by $x_{1+j} = (\sigma_j(x))_1$, obtaining

$$\mathscr{D}(T^j(x_1)) = \mathscr{D}(x_{1+j}) = \overline{\text{mdim}}_M(X, d, T) \qquad \forall j \geq 0. \tag{4.16}$$

Let us consider the empirical measures along the $T-$orbit of x_1. For each $k \geq 1$, denote $\nu_k = \frac{1}{k}\sum_{j=0}^{k-1} \delta_{T^j(x_1)}$, which by (4.16) satisfies $\int \mathscr{D}\,d\nu_k = \overline{\text{mdim}}_M(X,d,T)$. Due to the upper-semicontinuity of $\mathscr{D}$, any weak* limit ν of $(\nu_k)_k$ is $T-$invariant and satisfies

$$\int \mathscr{D}\,d\nu = \overline{\text{mdim}}_M(X, d, T).$$

We have constructed an invariant measure attaining the maximum in (4.15). Since $\mu \mapsto \int \mathscr{D}\,d\mu$ is affine, it is, in particular, convex. Thus the maximum must also be attained by an ergodic measure.

□

4.5 Spectral radius bounds for subshifts of compact type

In what follows, we recall from [25] an upper bound for the $\varepsilon-$entropy of subshifts of compact type in terms of the spectral radius of a suitable matrix. Fix a compact metric space (X,d) and a subset $\Gamma \subset X \times X$.

Given $\varepsilon > 0$, let $\mathscr{U}(\varepsilon) = \{U_1, \ldots, U_{M(\varepsilon)}\}$ be an open $\varepsilon-$cover of X. Consider the $M(\varepsilon) \times M(\varepsilon)$ matrix Γ^ε given by

$$\Gamma^\varepsilon_{i,j} = \begin{cases} 1 & \text{if } \big(U_i \times U_j\big) \cap \Gamma \neq \emptyset \\ 0 & \text{otherwise.} \end{cases} \tag{4.17}$$

By counting arguments similar to the ones used for subshifts of finite type, we know that $||(\Gamma^{\varepsilon})^k||$ many cylinders composed by elements of $\mathscr{U}(\varepsilon)$ of size k are enough to cover $\Gamma_{\mathbb{N}}$, where $||\cdot||$ denotes the norm $||(a_{ij})_{ij}|| = \sum_{i,j} |a_{ij}|$ in the space of $M(\varepsilon) \times M(\varepsilon)$ matrices.

sum norm in the space of $M(\varepsilon) \times M(\varepsilon)$ matrices. This estimate yields the following equivalent formulation of [25, Lemma 4.1], where $r(A)$ stands for the spectral radius of the matrix A.

Lemma 4.5.1 *Let (X,d) be a compact metric space and $\Gamma \subset X \times X$. Then*

$$h_{\varepsilon}(\Gamma_{\mathbb{N}}, d^{\mathbb{N}}, \sigma_{\mathbb{N}}) \ \leq \ \log r(\Gamma^{\varepsilon}).$$

In the case $X = [0,1]$ (and also in higher dimensions), the matrix Γ^{ε} can be replaced by another one constructed by using intersections with an $\varepsilon-$grid instead of an open cover. More precisely, given $\varepsilon > 0$ and $i, j = 1, ..., \lceil 1/\varepsilon \rceil$, we define

$$\Gamma^{\varepsilon}_{i,j} = \begin{cases} 1 & \text{if } \big([(i-1)\varepsilon, i\varepsilon] \times [(j-1)\varepsilon, j\varepsilon]\big) \cap \Gamma \neq \emptyset \\ 0 & \text{otherwise.} \end{cases} \tag{4.18}$$

We also have the following reformulation of Lemma 4.5.1.

Lemma 4.5.2 *Let d be the Euclidean distance in $[0,1]$ and $\Gamma \subset [0,1]^2$. Then*

$$h_{\varepsilon}(\Gamma_{\mathbb{N}}, d^{\mathbb{N}}, \sigma_{\mathbb{N}}) \ \leq \ \log r(\Gamma^{\varepsilon}).$$

The previous lemmas will be very useful tools to estimate the metric mean dimension when combined with the following property of complex matrices.

Theorem 4.5.3 (Gershgorin Circle Theorem, [26]) *Let $A = (a_{ij})_{i,j}$ be a complex $n \times n$ matrix. Then, every eigenvalue of A is contained in the union*

$$\bigcup_{i=1}^{n} B^{\mathbb{C}}_{R_i}(a_{ii}),$$

where $B^{\mathbb{C}}_{\delta}(z) = \{\tilde{z} \in \mathbb{C} : |z - \tilde{z}| \leq \delta\}$ and $R_i = \sum_{j \neq i} |a_{ij}|$ for every $i = 1, ..., n$.

In particular, if we apply Theorem 4.5.3 to a $0-1$ matrix A and its transpose, we get

$$r(A) \ \leq \ \max_i \#\{j : a_{ij} = 1\} \tag{4.19}$$

$$r(A) = \ r(A^T) \ \leq \ \max_j \#\{i : a_{ij} = 1\}. \tag{4.20}$$

The latter estimates will be used in two different instances in this text. In order to prove the upper bound in Theorem G, we estimate the number of 1's in any given *column* of Γ^{ε}, and apply (4.20). In Chapter 5, we will relate the metric mean dimension with the Hölder regularity of the map (see Theorem H) by estimating the *rows* of Γ^{ε} and using (4.19). In the following example we will apply the estimates above to compute the metric mean dimension on a class of subshifts of compact type.

4.5.1 Delayed full shifts

Let us consider a family of maps which are prescribed by a closed set and a collection of finite transitions which connect any pair of points in the set. We will use the information of the previous subsection with a slight modification: the only difference in the method is that we code different elements of the cover by the same symbol as long as they behave in a similar manner.

Proposition 4.5.4 *Let (X,d) be a compact metric space. Fix a positive integer $k \geq 2$ and select distinct points $a_1,...,a_{k-1} \in X$ and a closed subset $F \subset X$. Consider $\Gamma \subset X \times X$ given by*

$$\Gamma = (F \times \{a_1\}) \cup \{(a_1,a_2),(a_2,a_3),...,(a_{k-2},a_{k-1})\} \cup (\{a_{k-1}\} \times F).$$

Then,

$$\overline{\mathrm{mdim}}_M(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma_{\mathbb{N}}) = \frac{\overline{\dim}_B(F,d)}{k} \qquad \text{and} \qquad \underline{\mathrm{mdim}}_M(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma_{\mathbb{N}}) = \frac{\underline{\dim}_B(F,d)}{k}.$$

Proof: The computations below do not depend on whether we take upper or lower limits in ε, so we do not distinguish them by the notation. We will make use of the parameter $\rho > 1$ in the definition of the metric (2.4), thus denoting it simply by d_ρ, as we will always be dealing with unilateral sequences.

Since the map $h\colon (F^{\mathbb{N}},d_{2^k}) \to (\Gamma_{\mathbb{N}},d_2)$ defined as

$$h(x_1,x_2,...) = (x_1,a_1,a_2,...,a_{k-1},x_2,a_1,a_2,...,a_{k-1},x_3,...)$$

is an isometry over its image and $h \circ \sigma_{\mathbb{N}} = \sigma_{\mathbb{N}}^k \circ h$, by Proposition 2.3.1 and the formula (2.5) we have

$$\begin{aligned} \dim_B(F,d) &= \mathrm{mdim}_M(F^{\mathbb{N}},d_{2^k},\sigma_{\mathbb{N}}) = \mathrm{mdim}_M(h(F^{\mathbb{N}}),d_2,\sigma_{\mathbb{N}}^k) \\ &\leq \mathrm{mdim}_M(\Gamma_{\mathbb{N}},d_2,\sigma_{\mathbb{N}}^k) \ \leq k\ \mathrm{mdim}_M(\Gamma_{\mathbb{N}},d_2,\sigma_{\mathbb{N}}). \end{aligned}$$

For the converse inequality, we may assume without loss of generality that $a_j \in F$ for every $j = 1,...,k-1$. Otherwise, the set Γ would be contained in Γ' corresponding to $F' = F \cup \{a_1,...,a_{k-1}\}$, whose upper and lower box dimension are equal to the ones of F.

Given $\varepsilon > 0$ we consider a minimal open $\varepsilon-$cover $U_1,...,U_N$ of F, that is, $N = S(F,d,\varepsilon)$. We can assume that $a_j \in U_j$ uniquely. For each $n \geq 1$, consider the following open cover of $\Gamma_{\mathbb{N}}$ by cylinders (see (4.5))

$$\alpha_{n,\varepsilon} = \big\{C(\Gamma_{\mathbb{N}}\,;1,\cdots,n\,;U_{i_1},\cdots,U_{i_n}) \,:\, 1 \leq i_1,...,i_n \leq N\big\}$$

In order to estimate the cardinality of $\alpha_{n,\varepsilon}$, we partition its elements in the following way: for each $j = 1,...,k-1$, let

$$\begin{aligned} x_n^0 &= \#\big\{C(\Gamma_{\mathbb{N}}\,;1,\cdots,n\,;U_{i_1},\cdots,U_{i_n}) \in \alpha_{n,\varepsilon} \,:\, i_n \geq k\big\} \\ x_n^j &= \#\big\{C(\Gamma_{\mathbb{N}}\,;1,\cdots,n\,;U_{i_1},\cdots,U_{i_n}) \in \alpha_{n,\varepsilon} \,:\, i_n = j\big\}. \end{aligned}$$

Note that

$$\begin{cases} x^0_{n+1} & = \; x^{k-1}_n(N-k+1) \\ x^1_{n+1} & = \; x^0_n + \ldots + x^{k-1}_n \\ x^j_{n+1} & = \; x^{j-1}_n + x^{k-1}_n \qquad\quad \forall\, j \in \{2,\ldots,k-1\}, \end{cases}$$

or equivalently,

$$\begin{pmatrix} x^0_{n+1} \\ x^1_{n+1} \\ \vdots \\ x^{k-1}_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & N-k+1 \\ 1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & 0 & \ddots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x^0_n \\ x^1_n \\ \vdots \\ x^{k-1}_n \end{pmatrix}$$

We denote the above equality by $v_{n+1} = Av_n$, where $A = A_\varepsilon$ and $v_1 = (N-k+1,1,1,\ldots,1)$. Then,

$$\#\alpha_{n,\varepsilon} = x^1_{n+1} \le |v_{n+1}| \le ||A^n||\,|v_1|.$$

Let $\ell = \ell(\varepsilon) \ge 1$ be a positive integer satisfying $2^{-\ell}\operatorname{diam}(X,d) \le \varepsilon$. Then $\operatorname{diam}(\alpha_{n+\ell(\varepsilon),\varepsilon},(d_2)_n) \le \varepsilon$ and we get

$$h_\varepsilon(\Gamma_{\mathbb{N}},d_2,\sigma_{\mathbb{N}}) \le \lim_{n\to+\infty}\frac{1}{n}\log\#\alpha_{n+\ell(\varepsilon),\varepsilon} = \lim_{n\to+\infty}\frac{1}{n}\log|x^1_{\ell+n+1}| \le \log r(A_\varepsilon). \tag{4.21}$$

On the other hand, it is easy to see that the entries of A^k are polynomials in N of degree at most 1. Consequently, there exists $C \ge 1$ independent of N satisfying $||A^k|| \le CN$. Combining this information with (4.21) we obtain

$$h_\varepsilon(\Gamma_{\mathbb{N}},d_2,\sigma_{\mathbb{N}}) \le \log r(A_\varepsilon) \le \log||A^k_\varepsilon||^{1/k} \le \frac{1}{k}\Big(\log C + \log S(F,d,\varepsilon)\Big).$$

We finish the proof by dividing by $\log(1/\varepsilon)$ and making $\varepsilon \to 0^+$. $\square$

4.6 Gauss-like maps

In this section we fix a compact interval I, a compact subset $A \subset I$ with zero Lebesgue measure and the associated family of cut-out sets $(J_k)_{k\ge1}$, ordered non-increasingly in diameter. For each $k \ge 1$, let $\varepsilon_k = |J_k|$ and denote the Euclidean distance in I by d. From now on we will assume that $\varepsilon_k > 0$ for every $k \ge 1$. If, otherwise, there is k_0 such that $J_{k_0} = \emptyset$, then T_A has finitely many branches and so, by (4.2), its entropy is finite; therefore its metric mean dimension is zero, trivially satisfying (4.4).

For the sets A under consideration, the following dimension estimates via cut-out sets are provided by [18, Propositions 3.6 and 3.7].

Lemma 4.6.1 *Let $A \subset I$ and $(\varepsilon_k)_{k\ge1}$ be as above. Then*

$$\liminf_{k\to+\infty}\frac{\log k}{\log(1/\varepsilon_k)} \le \underline{\dim}_B(A,d) \le \overline{\dim}_B(A,d) = \limsup_{k\to+\infty}\frac{\log k}{\log(1/\varepsilon_k)}.$$

Moreover, the box dimension of A exists if, and only if, the outer limits coincide.

Theorem G is a consequence of the previous lemma together with the following inequalities, which are the content of Propositions 4.6.2 and 4.6.3:

$$\begin{aligned}\limsup_{k\to+\infty}\frac{\log k}{\log(1/\varepsilon_k)} &\leq \overline{\mathrm{mdim}}_M(I,d,T_A) \leq \overline{\dim}_B(A,d)\\ \liminf_{k\to+\infty}\frac{\log k}{\log(1/\varepsilon_k)} &\leq \underline{\mathrm{mdim}}_M(I,d,T_A) \leq \underline{\dim}_B(A,d).\end{aligned}$$

To prove the following proposition, the differentiability assumption (H3) in the statement of the theorem is not needed.

Proposition 4.6.2 *Let $T_A\colon I\to I$ be a map satisfying conditions* (H1) *and* (H2). *Then*

$$\limsup_{k\to+\infty}\frac{\log k}{\log(1/\varepsilon_k)} \leq \overline{\mathrm{mdim}}_M(I,d,T_A) \qquad \textit{and} \qquad \liminf_{k\to+\infty}\frac{\log k}{\log(1/\varepsilon_k)} \leq \underline{\mathrm{mdim}}_M(I,d,T_A)).$$

Proof: By Definition 4.1.1 of metric mean dimension, we must consider $\Gamma=\mathrm{graph}(T_A)$ and the subset $\Gamma_{\mathbb{N}}\subset I^{\mathbb{N}}$ endowed with the metric $d^{\mathbb{N}}$.

Given $k\geq 1$, each bijectivity domain $J_1,J_2,...,J_{2k}$ have length at least ε_{2k}. Let $i_1,...,i_{2k}$ be a reordering of $\{1,...,2k\}$ so that $J_{i_1},...,J_{i_{2k}}$ are ordered from left to right. Then the sets $J_{i_1},J_{i_3},...,J_{i_{2k-1}}$ are pairwise separated by at least ε_{2k}. Since condition (H2) ensures that each cylinder composed by $J_{i_1},J_{i_3},...,J_{i_{2k-1}}$ is nonempty, for every $n\geq 1$

$$S_1(\Gamma_{\mathbb{N}},d_n^{\mathbb{N}},\varepsilon_{2k}) \geq k^n.$$

This implies that

$$h_{\varepsilon_{2k+1}}(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma_{\mathbb{N}}) \geq h_{\varepsilon_{2k}}(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma_{\mathbb{N}}) \geq \log k \qquad \forall k\geq 1.$$

Thus,

$$h_{\varepsilon_m}(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma_{\mathbb{N}}) \geq \log\big(\frac{m}{2}-1\big) \qquad \forall m\geq 1.$$

Given $\varepsilon>0$, take $m=m(\varepsilon)$ satisfying $\varepsilon_{m+1}<\varepsilon\leq\varepsilon_m$. Thus,

$$\frac{h_\varepsilon(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma_{\mathbb{N}})}{\log(1/\varepsilon)} \geq \frac{h_{\varepsilon_m}(\Gamma_{\mathbb{N}},d^{\mathbb{N}},\sigma_{\mathbb{N}})}{\log(1/\varepsilon_{m+1})} \geq \frac{\log(m+1)}{\log(1/\varepsilon_{m+1})}\,\frac{\log(\frac{m}{2}-1)}{\log(m+1)}.$$

We finish the proof by making $\varepsilon\to 0^+$ (and consequently $m\to+\infty$). □

Proposition 4.6.3 *Let $T_A\colon I\to I$ be a map satisfying conditions* (H1)-(H3). *Then*

$$\begin{aligned}\overline{\mathrm{mdim}}_M(I,d,T_A) &\leq \overline{\dim}_B(A,d)\\ \underline{\mathrm{mdim}}_M(I,d,T_A) &\leq \underline{\dim}_B(A,d).\end{aligned}$$

Proof: Assume for simplicity that $I = [0,1]$ and fix $\varepsilon > 0$. By Lemma 4.5.2 and (4.20) we have

$$h_\varepsilon(\Gamma_{\mathbb{N}}, d^{\mathbb{N}}, \sigma_{\mathbb{N}}) \leq \log r(\Gamma^\varepsilon) \leq \log \max_{1\leq j\leq \lceil 1/\varepsilon \rceil} \#\{i : \Gamma^\varepsilon_{ij} = 1\}, \tag{4.22}$$

where Γ^ε is defined as in (4.18). For each $j = 1, ..., \lceil 1/\varepsilon \rceil$ fix any point $x_j \in \big(\varepsilon(j-1), \varepsilon j\big) \cap [0,1]$ and observe that $T_A^{-1}(x_j) = \{a_k(j)\}_k$ is a sequence that intersects each J_k exactly once, by conditions (H1) and (H2).

Let M be a positive integer such that $M^{-1} < \eta$, where η is as in (H3). The graph of $T_A|_{J_k}$ can never intersect more than $M+2$ boxes in the $\varepsilon-$grid per line. Otherwise, by the Mean Value Theorem there would exist some point z in the interior of J_k such that $|T'(z)| < M^{-1}$. Hence, for every $j = 1, ..., \lceil 1/\varepsilon \rceil$ we have

$$\begin{aligned}
\#\{i : \Gamma^\varepsilon_{ij} = 1\} &= \#\{i : \big([(i-1)\varepsilon, i\varepsilon] \times [(j-1)\varepsilon, j\varepsilon]\big) \cap \Gamma \neq \emptyset\} \\
&\leq 2(M+2)\ \#\{i : [\varepsilon(i-1), \varepsilon i] \cap T_A^{-1}(x_j) \neq \emptyset\} \\
&\leq 4(M+2)\ S(T_A^{-1}(x_j), d, \varepsilon) \\
&\leq 8(M+2)\ S(A, d, \varepsilon),
\end{aligned} \tag{4.23}$$

where the second inequality is due to the fact that every open interval of diameter ε can intersect at most 2 intervals of the form $[(i-1)\varepsilon, i\varepsilon]$, and the last inequality is a consequence of the next lemma.

Lemma 4.6.4 *Let $(a_k)_{k\geq 1}$ be a sequence of positive real numbers such that $a_k \in J_k$, for every $k \geq 1$. Then for every $\varepsilon > 0$,*

$$S(\{a_k \colon k \geq 1\}, d, \varepsilon) \leq 2\, S(A, d, \varepsilon).$$

Proof: Given a minimal collection of intervals of diameter at most ε that covers A, there can be at most one element of $\{a_k\}_k$ in between each pair of consecutive intervals not yet covered. Adding one extra interval for each of these (at most $S(A,d,\varepsilon)$) points, we get an $\varepsilon-$cover of $\{a_k\}_k$ with cardinality $2S(A,d,\varepsilon)$. □

Combining (4.22) and (4.23), we get

$$h_\varepsilon(\Gamma_{\mathbb{N}}, d^{\mathbb{N}}, \sigma_{\mathbb{N}}) \leq \log\big(8(M+2)\big) + \log S(A, d, \varepsilon),$$

and we end the proof of Proposition 4.6.3 after dividing by $\log(1/\varepsilon)$ and making $\varepsilon \to 0^+$. □

4.6.1 Large derivative assumption

Condition (H3) was used in the previous reasoning to obtain a uniform upper bound for the number of entries equal to 1 in every row of Γ^ε in terms of the $\varepsilon-$covering number of A. However, it is not strictly necessary as we will illustrate.

Example 4.6.5 *Let us compute the metric mean dimension, with respect to the Euclidean distance d, of the following (discontinuous) interval map*

$$\begin{array}{rccc} T\colon & [-1,1] & \longrightarrow & [-1,1] \\ & x \neq 0 & \mapsto & \sin\frac{1}{|x|} \\ & 0 & \mapsto & 0 \end{array}.$$

We start by observing that the map T does not satisfy either of the conditions (H2) (since T restricted to the left and rightmost intervals is not full branch) and (H3) (since there are infinitely many differentiable maxima and minima). In order to show that (H3) is not strictly necessary for Theorem G to be valid, we must consider a map satisfying (H2). After computing the metric mean dimension of T, we will consider a slightly modified map satisfying (H2) and such that all the estimates still hold.

The map T can be seen as T_A satisfying condition (H1) and, up to the left and rightmost intervals, also (H2), where $A = \left\{\frac{2}{(2k+1)\pi}, \frac{-2}{(2k+1)\pi} : k \geq 0\right\} \cup \{0\}$. *Let us prove that*

$$\mathrm{mdim}_M([-1,1],d,T) = \frac{1}{2} = \dim_B(A). \tag{4.24}$$

Firstly, an argument similar to the one in the proof of Proposition 4.6.2 yields

$$\underline{\mathrm{mdim}}_M([-1,1],d,T) \geq 1/2.$$

For the converse inequality, we cannot directly apply Proposition 4.6.3 since the map T does not satisfy (H3). The main idea is still to make use of Lemma 4.5.2 and (4.20)*. However, we will estimate the number of* 1*'s in a column by making use of the Second Order Mean Value Theorem.*[1]

For each $n \geq 1$, fix the scale $\varepsilon_n = \left\lceil \frac{3\pi(4n+1)(4n+5)}{8} \right\rceil^{-1}$ *and consider an ε_n−grid. To estimate the spectral radius of Γ^{ε_n} - defined in* (4.18) *- we restrict our attention to the first quadrant, as the remaining ones can be dealt in an analogous way.*

Consider the positive maximum points $x_k = 2/(4k+1)$, $k \geq 0$. Given $y \in (0,1]$ and $L \in (0,1)$ such that $T(y) = \sin(1/y) \geq L$, which implies $|\cos(1/y)| \leq \sqrt{1-L^2}$, we have two possibilities:

*(*i*)* $y \in (0, x_0]$

Let $k \geq 1$ be such that $x_{k+1} \leq y \leq x_{k-1}$. Then,

$$\begin{aligned} |f''(y)| &= \left|y^{-4}\sin(1/y) + 2y^{-3}\cos(1/y)\right| \\ &\geq y^{-4}L - 2y^{-3}\sqrt{1-L^2} \\ &\geq x_{k-1}^{-4}L - 2x_{k+1}^{-3}\sqrt{1-L^2} \\ &= k^4\,\frac{\pi^3}{8}\left[\frac{\pi L}{2}(4-\frac{3}{k})^4 - 2\sqrt{1-L^2}\left(\frac{4}{k^{1/3}} + \frac{5}{k^{4/3}}\right)^3\right] \\ &\geq k^4\,\frac{\pi^3}{8}\left[\frac{\pi L}{2} - 2\sqrt{1-L^2}\,9^3\right] \end{aligned}$$

Thus, $|f''(y)| \geq k^4$ if L is close enough to 1*.*

[1] If $f\colon [a,b] \to \mathbb{R}$ is twice differentiable, then $f(b) + f(a) - 2f(\frac{a+b}{2}) = \frac{1}{4}(b-a)^2 f''(c)$ for some $c \in (a,b)$.

(ii) $y \in (x_0, 1]$

$$\begin{aligned} |f''(y)| &= \left|y^{-4}\sin(1/y) + 2y^{-3}\cos(1/y)\right| \\ &\geq y^{-4}L - 2y^{-3}\sqrt{1-L^2} \\ &\geq L - 2x_1^{-3}\sqrt{1-L^2} \\ &= L - 2\Big(\frac{5\pi}{2}\Big)^3\sqrt{1-L^2}. \end{aligned}$$

Thus, $|f''(y)| \geq 1/2$ *if* L *is close enough to* 1.

From now on, we fix $L \in (0,1)$ *close enough to* 1 *so that the two previous estimates hold. Take* $n \geq 1$ *big enough so that* $1-\varepsilon_n \gg L$. *Let us bound the spectral radius of* Γ^{ε_n} *from above by estimating the number of* 1*'s. To do so, we recall from* (4.20) *that*

$$r(\Gamma^{\varepsilon_n}) \leq \max_j \#\{i : (\Gamma^{\varepsilon_n})_{ij} = 1\}$$

and separate j *in* 3 *classes. First, note that* $\varepsilon_n < \frac{x_n - x_{n+1}}{3}$ *for every* n. *This implies that every interval* $[(i-1)\varepsilon_n, i\varepsilon_n]$, $i = 1, \ldots, \lceil \frac{x_n}{\varepsilon_n} \rceil$, *contains at least two* x_k*'s. Therefore,*

$$(\Gamma^{\varepsilon_n})_{i,j} = 1 \qquad \forall i = 1, \ldots, \left\lceil \frac{x_n}{\varepsilon_n} \right\rceil, \forall j = 1, \ldots, \frac{1}{\varepsilon_n}. \tag{4.25}$$

Case 1: $j = 1/\varepsilon_n$.
For each $k \geq 0$, *consider the nearest points* $a_k < x_k < b_k$ *such that* $T(a_k) = T(b_k) = 1 - \varepsilon_n$. *Then*

$$\{x \in (0,1) : T(x) \geq 1 - \varepsilon_n\} = \bigcup_{k \geq 0} [a_k, b_k].$$

By the Second Order Mean Value Theorem, for every $k \geq 0$ *there exists* $c_k \in (a_k, b_k)$ *such that* $|b_k - a_k|^2 \leq \frac{8\varepsilon_n}{|f''(c_k)|}$. *Moreover, by the choice of* L *we have* $|b_0 - a_0| \leq 4\sqrt{\varepsilon_n}$ *and* $|b_k - a_k| \leq \frac{\sqrt{8}\sqrt{\varepsilon_n}}{k^2}$, *for every* $k \geq 1$. *Combining this information with* (4.25), *we obtain*

$$\begin{aligned} \#\{i : (\Gamma^{\varepsilon_n})_{i,\frac{1}{\varepsilon_n}} = 1\} &\leq \left\lceil \frac{x_n}{\varepsilon_n} \right\rceil + \sum_{k=0}^{n} \left(\left\lfloor \frac{|b_k - a_k|}{\varepsilon_n} \right\rfloor + 2 \right) \\ &\leq \left\lceil \frac{x_n}{\varepsilon_n} \right\rceil + 2(n+1) + \frac{1}{\sqrt{\varepsilon_n}} \Big(4 + \sqrt{8} \sum_{k=1}^{\infty} k^{-2} \Big) \\ &\leq Cn, \end{aligned} \tag{4.26}$$

where $C > 0$ *is a constant independent of* n.

Case 2: $j = \lceil L/\varepsilon_n \rceil + 1, \ldots, 1/\varepsilon_n - 1$.
For each $k \geq 0$, *consider the nearest points* $a_k^1 < b_k^1 < x_k < a_k^2 < b_k^2$ *such that* $T(a_k^1) = T(b_k^2) = (j-1)\varepsilon_n$ *and* $T(a_k^2) = T(b_k^1) = j\varepsilon_n$. *Then*

$$\{x \in (0,1) : (j-1)\varepsilon_n \leq T(x) \leq i\varepsilon_n\} = \bigcup_{k \geq 0} [a_k, b_k].$$

By the Second Order Mean Value Theorem, for every $k \geq 0$ and $s = 1,2$ there exists $c_k^s \in (a_k^s, b_k^s)$ such that $|b_k^s - a_k^s|^2 \leq \frac{8\varepsilon_n}{|f''(c_k^s)|}$. Again, by the choice of L we have $|b_0^s - a_0^s| \leq 4\sqrt{\varepsilon_n}$ and $|b_k^s - a_k^s| \leq \frac{\sqrt{8}\sqrt{\varepsilon_n}}{k^2}$, for every $k \geq 1$. Joining this information with (4.25), *we obtain*

$$\begin{aligned}\#\big\{i : (\Gamma^{\varepsilon_n})_{i,j} = 1\big\} &\leq \left\lceil \frac{x_n}{\varepsilon_n} \right\rceil + \sum_{s=1}^{2}\sum_{k=0}^{n}\Big(\left\lfloor \frac{|b_k^s - a_k^s|}{\varepsilon_n} \right\rfloor + 2\Big) \\ &\leq \left\lceil \frac{x_n}{\varepsilon_n} \right\rceil + 4(n+1) + \frac{2}{\sqrt{\varepsilon_n}}\Big(4 + \sqrt{8}\sum_{k=1}^{\infty} k^{-2}\Big) \\ &\leq Cn, \end{aligned} \tag{4.27}$$

for every $j = \lceil L/\varepsilon_n \rceil + 1, ..., \frac{1}{\varepsilon_n} - 1$, where $C > 0$ is a constant independent of n.

Case 3: $j = 1, ..., \lceil L/\varepsilon_n \rceil$.

By taking n large enough, we may assume that $L + \varepsilon_n < \frac{1+L}{2}$. Thus, $T(x) = \sin(1/x) \leq \frac{1+L}{2}$, for every $x \in (0,1)$ such that $T(x) \in [(j-1)\varepsilon_n, j\varepsilon_n]$ for some $j = 1, ..., \lceil L/\varepsilon_n \rceil$. Hence, there exists some $M = M(L) \in \mathbb{N}$ such that, for every such x, we have

$$|T'(x)| = \frac{|\cos(1/x)|}{x^2} \geq \sqrt{1 - \Big(\frac{1+L}{2}\Big)^2} \geq M^{-1}.$$

By the proof of Proposition 4.6.3 we obtain

$$\#\big\{i : (\Gamma^{\varepsilon_n})_{i,j} = 1\big\} \leq 8(M+2)n, \tag{4.28}$$

for every $j = 1, ..., \lceil L/\varepsilon_n \rceil$.

Bringing together (4.26), (4.27) *and* (4.28), *and proceeding analogously in the other quadrants, we conclude that there exists a constant $C > 0$, which depends only on L, such that*

$$r(\Gamma^{\varepsilon_n}) \leq \max_j \#\{i : (\Gamma^{\varepsilon_n})_{ij} = 1\} \leq Cn,$$

for every large enough $n \in \mathbb{N}$.

We are ready to complete the computation of the upper bound of the metric mean dimension of T. Given $\varepsilon > 0$, take $n = n(\varepsilon)$ such that $\varepsilon_n \leq \varepsilon < \varepsilon_{n-1}$. Then

$$\frac{h_\varepsilon(\Gamma_{\mathbb{N}}, d^{\mathbb{N}}, \sigma_{\mathbb{N}})}{\log(1/\varepsilon)} \leq \frac{h_{\varepsilon_n}(\Gamma_{\mathbb{N}}, d^{\mathbb{N}}, \sigma_{\mathbb{N}})}{\log(1/\varepsilon_{n-1})} \leq \frac{r(\Gamma^{\varepsilon_n})}{\log(1/\varepsilon_{n-1})} \leq \frac{\log(Cn)}{\log(1/\varepsilon_{n-1})},$$

and so

$$\overline{\mathrm{mdim}}_M([-1,1], d, T) = \overline{\mathrm{mdim}}_M(\Gamma_{\mathbb{N}}, d^{\mathbb{N}}, \sigma_{\mathbb{N}}) \leq \limsup_{n \to +\infty} \frac{\log(Cn)}{\log(1/\varepsilon_{n-1})} = \frac{1}{2}.$$

We finish this section by observing that all arguments used to compute the metric mean dimension of the map T remain valid if we consider a smooth perturbation $\tilde{T}$ of T in small neighbourhoods $(1-\delta, 1]$ and $[-1, -1+\delta)$, so that $\tilde{T}(1) = \tilde{T}(-1) = -1$ and $|\tilde{T}'(x)| > M^{-1}$ for every $x \in [-1, -1+\delta) \cup (1-\delta, 1]$. In this case, we obtain a map satisfying (H1) and (H2) such that the equality (4.4) in Theorem G still holds, even though the assumption (H3) does not.

4.6.2 On the definition of metric mean dimension of discontinuous maps

Given a discontinuous map T, its associated transition set $\Gamma = \mathrm{graph}(T)$ is not closed. In order to associate a subshift of compact type to T, we could have followed two possible paths: one option would be to take *the closure of the transition set* and afterwards consider its associated sequence space $(\overline{\Gamma})_{\mathbb{N}}$; another way (which is the one we chose) is to consider *the closure of the sequence space* $\overline{(\Gamma_{\mathbb{N}})}$. At first glance, the former might appear more satisfactory since, by Theorem E, we would obtain a notion of metric mean dimension that could be expressed both by unilateral and bilateral sequences. Let us exemplify why it would not yield a *meaningful* notion of metric mean dimension.

Consider the set $A = \{e^{-n} : n \geq 0\} \cup \{0\}$ and let $T_A\colon [0,1] \to [0,1]$ be any map satisfying conditions (H1)-(H3). The arguments in the proof of Theorem G do not depend whether we define $\mathrm{mdim}_M([0,1],d,T_A)$ in terms of $\overline{(\Gamma(T_A)_{\mathbb{N}})}$ or $\big(\overline{\Gamma(T_A)}\big)_{\mathbb{N}}$. This indicates that

$$\mathrm{mdim}_M([0,1],d,T_A) \;=\; \dim_B(A,d) \;=\; 0 \tag{4.29}$$

is the right value for it. Yet, if we consider an interval map T such that $T|_{[0,e^{-1})} = (T_A)|_{[0,e^{-1})}$ and $T|_{[e^{-1},1]} \equiv 0$, then any meaningful notion of metric mean dimension compatible with (4.29) should also be zero. However, since we have

$$\overline{\Gamma(T)} \supset \big([e^{-1},1] \times \{0\}\big) \cup \big(\{0\} \times [e^{-1},1]\big),$$

then Proposition 4.5.4 implies that

$$\mathrm{mdim}_M\big((\overline{\Gamma(T)})_{\mathbb{N}}, d^{\mathbb{N}}, \sigma_{\mathbb{N}}\big) \;\geq\; \frac{1}{2}.$$

This shows that taking the closure of the transition set prior to generating the sequences may create too many new orbits that are not present in the original dynamics.

4.7 Open questions

In Subsection 4.3.4, we constructed two maps with zero dynamical complexity, whose joint action has positive metric mean dimension and, in particular, infinite entropy. Proposition 4.3.7 seems optimal: if two maps have zero complexity (in terms of metric mean dimension) but the associated action does not, the richness of the orbit structure can only be revealed after a composition of the two maps. This indicates that we need twice the time to see this maximal value, which is always bounded by the dimension of the phase space. This motivates the following

Question 3 *Let (X,d) be a compact metric space and $g_1, g_2\colon X \to X$ be continuous maps with zero metric mean dimension. Is it true that*

$$\overline{\mathrm{mdim}}_M^F(X,d,\mathbb{S}) \leq \frac{\overline{\dim}_B(X,d)}{2} \;\; ?$$

In Theorem G we consider a family of interval maps exhibiting infinitely many full branches along a partition by intervals $(J_k)_k$, similar to the Gauss map. An essential aspect in the proof of the upper bound (Proposition 4.6.3) is the fact that the maps are defined at all J_k's. Let us now consider a more general setting permitting holes in the domain: we start with a map T_A satisfying (H1)-(H3) in (4.3) and delete all branches but the ones indexed by a sequence $(k_n)_n$, obtaining a map $\tilde{T}_A$.

Question 4 *Is it true that*

$$\overline{\mathrm{mdim}}_M(I, |\cdot|, \tilde{T}_A) \;=\; \limsup_{n\to\infty} \frac{\log n}{\log |J_{k_n}|}?$$

Chapter 5

Hölder regularity and Assouad spectrum

The results in this chapter are based in the paper

[4] A.T. Baraviera, M. Carvalho and G. Pessil: *Metric mean dimension, Hölder regularity and Assouad spectrum.* arXiv:2407.15774v1. (Accepted for publication in Journal of Fractal Geometry)

5.1 Preamble

In this chapter, we investigate the interplay between the metric mean dimension, the geometry of the underlying space and the regularity of the map.

Lipschitz maps on compact metric spaces of finite dimension have finite topological entropy (see (5.1)), hence their metric mean dimension is zero. In [32], Hazard presented a family of Hölder continuous interval maps with infinite entropy. Thus, it is worthwhile studying the metric mean dimension within this family, and we wonder if and how it is related to the Hölder exponents of the maps. We answer to this question positively for the $2-$parameter family of interval maps $T_{a,b}$, presented in Subsection 2.4.2, which comprises Hazard's examples (see Subsection 5.1.2 and Theorem I). The core of the proof of this result is a formula for the metric mean dimension in terms of a sequence of horseshoes and the scales at which their dynamics can be detected, resembling Misiurewicz formula for the entropy (cf. [46]). We deduce this formula by developing a dynamical analogue of the Minkowski-Bouligand for subshifts on Ahlfors regular alphabets, which also provides an entropy formula in terms of the size of the set of admissible words, generalizing the classical result for subshifts on finite alphabets (see Section 5.3).

Taking into account that the metric mean dimension has both a dynamical and a geometrical imprint, in a more general framework, where the phase space is not as homogeneous as the interval, we need to deal with more refined geometric structure, which may not be detected if one chooses an unfit notion of dimension. It turns out that the more recent concept of Assouad spectrum (cf. [24]) captures in a precise way the impact of the Hölder regularity on the metric mean dimension (see Section 5.2 and Theorem H). The essence of our approach is an application of the spectral bounds, from Section 4.5 for the metric mean dimension in terms of matrices whose entries are ruled by the Hölder regularity of the map.

5.1.1 Hölder regularity

Let (X,d) be a compact metric space and $T\colon X\to X$ be a continuous map. We say that T is *Lipschitz* if there is some $C>0$ such that

$$d(T(x),T(y)) \le C\,d(x,y) \qquad \forall x,y\in X.$$

It is known that the topological entropy of a Lipschitz map T has a natural upper bound in terms of the Lipschitz constant and the dimension of the phase space (cf. [35, Theorem 3.2.9]), namely

$$h_{top}(T) \le \max\{0,\log C\}\,\overline{\dim}_B(X,d). \tag{5.1}$$

However, if one is interested in maps satisfying weaker forms of regularity, the phenomenon of infinite entropy may occur. Given $\alpha\in(0,1)$, we say that T is *$\alpha-$Hölder* if there exists a constant $C_\alpha>0$ such that

$$d(T(x),T(y)) \le C_\alpha\, d(x,y)^\alpha \qquad \forall x,y\in X.$$

We refer to α as the *Hölder exponent of T*.

The following result is an analogue of (5.1) for the upper metric mean dimension $\overline{\mathrm{mdim}}_M(X,d,T)$ on spaces with finite Assouad dimension $\dim_A(X,d)$ in terms of the Hölder exponent α and the Assouad spectrum $\dim_A^\alpha(X,d)$ of the phase space at the parameter α. See the precise definitions in Section 5.2.

Theorem H *Let (X,d) be a compact metric space with finite Assouad dimension. If $\alpha\in(0,1)$ and $T\colon X\to X$ is $\alpha-$Hölder, then*

$$\overline{\mathrm{mdim}}_M(X,d,T) \le (1-\alpha)\dim_A^\alpha(X,d).$$

If, in addition, $\dim_A(X,d)>0$ *then*

$$\sup\{\alpha\in(0,1)\,:\,T\ is\ \alpha-Hölder\} \le 1-\frac{\overline{\mathrm{mdim}}_M(X,d,T)}{\dim_A(X,d)}.$$

The previous inequalities are sharp: in Subsection 5.1.2 we show that the equalities are attained for the family of maps $T_{a,b}$, defined in Subsection 2.4.2. Theorem H answers positively the questions raised in [48].

Recall that, if a compact metric space admits a Lipschitz map with infinite entropy, then by (5.1) its box dimension is necessarily infinite. Taking into account the connection between the box dimension and the Assouad spectrum (see (5.7)), we deduce from Theorem H that the existence of a Hölder endomorphism with full metric mean dimension also imposes constraints on the fractal structure of the underlying space.

Corollary A *Let (X,d) be a compact metric space with finite Assouad dimension. If there exists a continuous map $T\colon X\to X$ such that $\overline{\mathrm{mdim}}_M(X,d,T)=\overline{\dim}_B(X,d)$, then*

$$\dim_A^\alpha(X,d) = \frac{\overline{\dim}_B(X,d)}{1-\alpha} \qquad \forall\alpha\in\mathscr{H}(T),$$

where $\mathscr{H}(T) = \{\alpha \in (0,1) : T \text{ is } \alpha-\text{Hölder}\} \cup \{0\}$.

In particular, if the space is sufficiently homogeneous so that the upper box and Assouad dimensions coincide, then it does not admit a Hölder endomorphism with full upper metric mean dimension. This is the case when the space is Ahlfors regular (see Definition 5.3.2).

Corollary B *Let (X,d) be a compact metric space such that*

$$0 < \overline{\dim}_B(X,d) = \dim_A(X,d) < +\infty.$$

If a continuous map $T\colon X \to X$ satisfies $\overline{\mathrm{mdim}}_M(X,d,T) = \overline{\dim}_B(X,d)$, then it is not $\alpha-$Hölder for any $\alpha \in (0,1)$.

Since a C^0-Baire generic homeomorphism on a compact manifold of dimension greater than 1 has full metric mean dimension (cf. [12]), the previous corollary provides an alternative proof of the well-known fact that a C^0-generic homeomorphism is not $\alpha-$Hölder for any $\alpha \in (0,1)$. A similar conclusion holds for a C^0-generic endomorphism of the interval.

Here are two concrete examples on which we may apply Theorem H. Consider the interval $[0,1]$ with the Euclidean distance, which we denote by $|\cdot|$.

Example 5.1.1 (Weierstrass functions) *Fix positive real numbers $b > 1$ and $1/b < a < 1$ and consider the map $W_{a,b}\colon [-c,c] \to [-c,c]$, where $c = (1-a)^{-1}$, defined by*

$$W_{a,b}(x) = \sum_{n=0}^{\infty} a^n \cos(2\pi b^n x).$$

It is known that $W_{a,b}$ is $\beta-$Hölder, where $\beta = \frac{\log(1/a)}{\log b}$ (cf. [31]). Moreover, $W_{a,b}$ is $\beta-$hypersensitive, that is, $\exists \delta, C > 0$ such that for every interval $J \subset [-c,c]$ with diameter smaller than δ, we have $\mathrm{diam}\big(W_{a,b}(J)\big) \geq C\,\mathrm{diam}(J)^\beta$ (cf. [3]). Combining this information with Theorem H and [37, Theorem 4.1], we conclude that

$$mdim_M([-c,c],|\cdot|,W_{a,b}) = 1-\beta = 1 + \frac{\log a}{\log b}.$$

Example 5.1.2 (Hypersensitive zipper maps) *Let us consider the parameterized family of maps defined in [37, Section 1.1]. Denote by $\mathscr{C}_0^0$ the space of continuous functions $f\colon [0,1] \to [0,1]$ fixing 0 and 1, with the supremum norm. Consider a pair of points $p = \big((x_1,y_1),(x_2,y_2)\big) \in (0,1)^4$ such that $x_2 > x_1$ and $y_2 < y_1$. Define the operator $\Phi_p\colon \mathscr{C}_0^0 \to \mathscr{C}_0^0$ by*

$$\Phi_p f(x) = \begin{cases} y_1\, f(\frac{x}{x_1}) & \text{if } x \in [0,x_1] \\ y_1 - (y_1-y_2) f\big(\frac{x-x_1}{x_2-x_1}\big) & \text{if } x \in [x_1,x_2] \\ y_2 + (1-y_2) f\big(\frac{x-x_2}{1-x_2}\big) & \text{if } x \in [x_2,1]. \end{cases}$$

Then Φ_p is a contraction and thus has a unique fixed point $Z_p\colon [0,1] \to [0,1]$, which is called the zipper map of parameter p. The point p brings forward 4 additional parameters, denoted by $\lambda_{\min}(p) \leq \lambda_{\max}(p)$ and $0 < h_{min}(p) \leq h_{max}(p) < 1$, which describe geometric properties of the *operator and impart dynamical information about Z_p. For instance, in [37, Proposition H] it was shown that Z_p is $\beta-$hypersensitive if and only if $\lambda_{min} > 1$, where $\beta = 1 - \frac{\log \lambda_{min}}{|\log h_{min}|}$. Consequently (see [37, Theorem D]),*

$$\frac{\log \lambda_{min}}{|\log h_{min}|} \leq \underline{mdim}_M([0,1], |\cdot|, Z_p).$$

Besides, Z_p is $\alpha-$Hölder for $\alpha = 1 - \frac{\log \lambda_{max}}{|\log h_{max}|}$ (cf. [37, Proposition 2.2]). Combining this information with Theorem H, we obtain

$$\frac{\log \lambda_{min}}{|\log h_{min}|} \leq \underline{mdim}_M([0,1], |\cdot|, Z_p) \leq \overline{mdim}_M([0,1], |\cdot|, Z_p) \leq \frac{\log \lambda_{max}}{|\log h_{max}|}.$$

5.1.2 Optimal bounds

In this subsection we will see that Theorem H is sharp. We will provide an explicit formula for the metric mean dimension of the maps $T_{a,b}$ (see Subsection 2.4.2) in terms of the parameters and it turns out that the obtained expression can be easily reformulated in terms of the Hölder exponents. Similar maps were studied by Hazard in [32], though without any mention to metric mean dimension. Actually, the main focus there was to construct interval maps which are $\alpha-$Hölder, for $\alpha \in (0,1)$, and whose topological entropy is infinite.

A classical result of Misiurewicz [46] states that if $T\colon [0,1] \to [0,1]$ is a continuous map then for every $n \in \mathbb{N}$ there exists

- an interval J_n
- a collection $\mathscr{F}_n$ of pairwise disjoint subintervals of J_n
- $k_n \in \mathbb{N}$

such that $J_n \subset T^{k_n}(I)$ for every $I \in \mathscr{F}_n$ and

$$h_{top}(T) = \limsup_{n \to +\infty} \frac{\log \# \mathscr{F}_n}{k_n}. \tag{5.2}$$

Observe that Misiurewicz formula (5.2) for the maps $T_{a,b}$ defined in (2.8) becomes

$$h_{top}(T_{a,b}) = \limsup_{k \to +\infty} \log b_k = +\infty.$$

We will obtain a formula similar to (5.2) for the metric mean dimension, with respect to the Euclidean distance, within the family $T_{a,b}$. As a consequence, we also establish a connection between the upper metric mean dimension and the Hölder regularity of $T_{a,b}$, which shows that the estimates in Theorem H are sharp.

Theorem I *Let $a = (a_k)_{k\geq 0}$ and $b = (b_k)_{k\geq 1}$ be two sequences satisfying the conditions* (C1)-(C3) *in* (2.7) *and $T_{a,b}$ be the corresponding interval map. Then*

$$\overline{\mathrm{mdim}}_M([0,1],|\cdot|,T_{a,b}) \;=\; \limsup_{k\to+\infty} \frac{\log b_k}{\log(1/\varepsilon_k)}$$
$$\underline{\mathrm{mdim}}_M([0,1],|\cdot|,T_{a,b}) \;=\; \liminf_{k\to+\infty} \frac{\log b_k}{\log(1/\varepsilon_k)}.$$

Moreover,

$$\sup \mathscr{H}(T) \;=\; 1 - \overline{\mathrm{mdim}}_M([0,1],|\cdot|,T_{a,b}),$$

where $\mathscr{H}(T) = \{\alpha \in (0,1) : T \text{ is } \alpha-\text{Hölder}\} \cup \{0\}$.

For example, by Theorem I:

(i) If $a_k = \sum_{n=1}^{k} 6/\pi n^2$ and $b_k = 3^k$, then $\mathrm{mdim}_M([0,1],|\cdot|,T_{a,b}) = 1$ and $T_{a,b}$ is not $\alpha-$Hölder for any $\alpha \in (0,1)$.

(ii) Given $0 < \beta < 1$, if $a_k = \sum_{n=1}^{k} C(\beta)\, 3^{n(1-\frac{1}{\beta})}$, where $C(\beta) = \left(\sum_{n=1}^{\infty} 3^{n(1-\frac{1}{\beta})}\right)^{-1}$, and $b_k = 3^k$, then $\mathrm{mdim}_M([0,1],|\cdot|,T_{a,b}) = \beta$ and $T_{a,b}$ is $\alpha-$Hölder for every $\alpha \in (0,1-\beta)$.

(iii) If $a_k = 1/2^k$ and $b_k = 2k+1$, then $\mathrm{mdim}_M([0,1],|\cdot|,T_{a,b}) = 0$ and $T_{a,b}$ is $\alpha-$Hölder for every $\alpha \in (0,1)$.

We stress that the lack of regularity represented by the value of the metric mean dimension in Theorem I is of a different nature from the one studied by Hazard in [32]. The maps in [32] are defined like we did for $T_{a,b}$, considering $a = (2^{-k})_k$ and $b = (2k+1)_k$, but replacing each f_k by $g_k = \phi_\alpha \circ f_k \circ \phi_\alpha^{-1}$, where $\phi_\alpha(x) = x^\alpha$. Thus, these maps have zero metric mean dimension, and their irregularity is due to the appearance of cusps. On the other hand, we consider maps composed by sequences of piecewise *linear* horseshoes whose lack of regularity comes from the distortion ratio between the lengths of the injectivity domains of the horseshoes and of their images; or equivalently, the ratio between the number of branches and the scales at which they separate. This is precisely what metric mean dimension quantifies.

5.1.3 An entropy formula

It is known [60, Theorem 7.13] that the topological entropy of a subshift on finitely many symbols $Y \subset \{1,\cdots,q\}^{\mathbb{N}}$ can be expressed in terms of the exponential growth of the number of $n-$admissible words in Y, namely,

$$h_{top}(Y,\sigma_{\mathbb{N}}) \;=\; \lim_{n\to+\infty} \frac{1}{n} \log \Theta_n(Y), \tag{5.3}$$

where, for each $n \in \mathbb{N}$, $\Theta_n(Y)$ is defined as the number of $n-$tuples $[i_1,...,i_n]$ such that the set $\{(y_k)_{k\in\mathbb{N}} \in Y : y_1 = i_1,\cdots,y_n = i_n\}$ is nonempty. We will extend this formula for subshifts on more general alphabets.

A compact metric space (X,d) is *Ahlfors regular* if there exists a Borel probability measure $\mu \in \mathscr{P}(X)$ satisfying

$$\exists \varepsilon_0, C_1, C_2 > 0: \qquad C_1\, \varepsilon^{\dim_B X} \leq \mu(B_\varepsilon(x)) \leq C_2\, \varepsilon^{\dim_B X} \qquad \forall x \in X, \forall 0 < \varepsilon < \varepsilon_0.$$

For instance, a finite set with the normalized counting measure or the Euclidean cube with the Lebesgue measure are Ahlfors regular. In this setting, we obtain sharp estimates for the $\varepsilon-$entropy of an arbitrary subshift $Y \subset X^{\mathbb{N}}$ on an Ahlfors regular alphabet X, namely

$$h_\varepsilon(Y, d_{\mathbb{N}}, \sigma_{\mathbb{N}}) \;=\; \lim_{n\to+\infty} \frac{1}{n} \log\Big(\frac{\mu^n\big(Y(n,\varepsilon)\big)}{\varepsilon^{n\,\dim_B X}}\Big) \;+\; o(1),$$

where $Y(n,\varepsilon)$ is the set $B_\varepsilon(\pi_n(Y)) \subset X^n$ with respect to the supremum norm $||\cdot||_\infty$ and $\pi_n \colon X^{\mathbb{N}} \to X^n$ is the projection on the first n coordinates.

Theorem J *Let (X,d,μ) be an Ahlfors regular space. For every subshift $Y \subset X^{\mathbb{N}}$, there exists $\lambda = \lambda(Y) > 0$, which is unique if $h_{top}(Y,\sigma_{\mathbb{N}}) < +\infty$, such that*

$$h_{top}(Y,\sigma_{\mathbb{N}}) \;=\; \limsup_{\varepsilon\to 0^+} \lim_{n\to+\infty} \frac{1}{n} \log\Big(\frac{\mu^n\big(Y(n,\varepsilon)\big)}{\lambda^n\, \varepsilon^{n\,\dim X}}\Big). \tag{5.4}$$

In addition, independently of λ,

$$\overline{\mathrm{mdim}}_M(Y, d^{\mathbb{N}}, \sigma_{\mathbb{N}}) \;=\; \dim(X,d) + \limsup_{\varepsilon\to 0^+} \frac{1}{\log(1/\varepsilon)} \lim_{n\to+\infty} \frac{1}{n} \log \mu^n\big(Y(n,\varepsilon)\big) \tag{5.5}$$

$$\underline{\mathrm{mdim}}_M(Y, d^{\mathbb{N}}, \sigma_{\mathbb{N}}) \;=\; \dim(X,d) + \liminf_{\varepsilon\to 0^+} \frac{1}{\log(1/\varepsilon)} \lim_{n\to+\infty} \frac{1}{n} \log \mu^n\big(Y(n,\varepsilon)\big).$$

In particular, if X is finite we recover (5.3). Indeed, if $X = \{1,\cdots,q\}$ then it has zero dimension and is an Ahlfors regular space with respect to the normalized counting measure μ, whose constants are $C_1 = C_2 = 1/q$. The proof of Theorem J yields $\lambda(Y) = 1/q$ for every Y, we obtain

$$h_{top}(Y,\sigma_{\mathbb{N}}) \;=\; \limsup_{\varepsilon\to 0^+} \lim_{n\to+\infty} \frac{1}{n} \log\Big(q^n \cdot \mu^n\big(Y(n,\varepsilon)\big)\Big) \;=\; \lim_{n\to+\infty} \frac{1}{n} \log \Theta_n(Y).$$

Recall that, by Theorem F, one can compute the metric mean dimension of a continuous map T acting on a compact metric space X by considering the graph of the map as a transition set and the corresponding subshift of compact type

$$Y \;=\; \big\{(x, T(x), T^2(x), \cdots) : x \in X\big\}.$$

As an application of Theorem J, we obtain an alternative formula for the metric mean dimension of T whenever X is an Ahlfors regular space. This has the advantage of providing an interpretation of the difference between the metric mean dimension of the map and the dimension of the underlying space.

Corollary 5.1.3 *Let (X,d,μ) be an Ahlfors regular space and $T : X \to X$ be a continuous map. Then, there exists $\lambda = \lambda(T) > 0$, which is unique if $h_{top}(X,T) < +\infty$, such that*

$$h_{top}(X,T) \;=\; \limsup_{\varepsilon\to 0^+} \lim_{n\to+\infty} \frac{1}{n} \log\Big(\frac{\mu^n\big(Y(T,n,\varepsilon)\big)}{\lambda^n\, \varepsilon^{n\,\dim X}}\Big).$$

In addition, independently of λ,

$$\overline{\mathrm{mdim}}_M(X,d,T) \;=\; \dim(X,d) + \limsup_{\varepsilon\to 0^+} \frac{1}{\log(1/\varepsilon)} \lim_{n\to+\infty} \frac{1}{n} \log \mu^n\big(Y(T,n,\varepsilon)\big)$$

$$\underline{\mathrm{mdim}}_M(X,d,T) \;=\; \dim(X,d) + \liminf_{\varepsilon\to 0^+} \frac{1}{\log(1/\varepsilon)} \lim_{n\to+\infty} \frac{1}{n} \log \mu^n\big(Y(T,n,\varepsilon)\big)$$

where

$$Y(T,n,\varepsilon) \;=\; \bigcup_{x\in X} B_\varepsilon(x)\times B_\varepsilon(T(x))\times\cdots\times B_\varepsilon(T^{n-1}(x)).$$

The previous corollary turns out to be a useful tool to explicitly compute the metric mean dimension on Ahlfors regular spaces. See, for instance, Proposition 5.5.1.

It is well known that the metric mean dimension of iterates of a continuous map behaves differently from the topological entropy. In fact, we have

$$\overline{\mathrm{mdim}}_M(X,d,T^k) \;\le\; k\;\overline{\mathrm{mdim}}_M(X,d,T), \quad \forall k\ge 1,$$

whereas this inequality can be strict (see Remark 2.4.1). As a consequence of Corollary 5.1.3, we obtain an optimal converse in this setting.

Corollary 5.1.4 *Let (X,d,μ) be an Ahlfors regular space and $T : X \to X$ be a continuous map. Then, for every $k\ge 1$,*

$$\begin{aligned}
\overline{\mathrm{mdim}}_M(X,d,T^k) &\;\ge\; k\;\overline{\mathrm{mdim}}_M(X,d,T) \;-\; (k-1)\dim X\\
\underline{\mathrm{mdim}}_M(X,d,T^k) &\;\ge\; k\;\underline{\mathrm{mdim}}_M(X,d,T) \;-\; (k-1)\dim X.
\end{aligned}$$

In particular, if T has full metric mean dimension, so do all of its iterates.

5.2 Assouad dimension and spectrum

In this section we will recall the definitions of the dimensional quantities in Theorem H.

Definition 5.2.1 *The Assouad dimension of (X,d) is defined as*

$$\dim_A(X,d) \;=\; \inf\left\{ s : \begin{array}{c} \exists C,\varepsilon_0>0 \;:\; \forall 0<r<R<\varepsilon_0\; \forall x\in X \\ S\big(B_R(x),d,r\big) \,\le C\left(\frac{R}{r}\right)^s \end{array} \right\}.$$

The Assouad dimension quantifies the maximal local complexity of a fractal, since it provides the maximal exponential growth rate of $\sup_{x\in X} S\big(B_R(x),d,r\big)$ as the parameters $0<r<R$ vary

independently. In particular, it is always an upper bound to the box dimension, which is a more averaged measurement of complexity over the whole space. Fraser and Yu proposed in [24] an interpolation between these notions of complexity, obtained by fixing the relation between the scales r and R in the definition of Assouad dimension. More precisely, for each fixed $\theta \in (0,1)$ the scales are chosen to satisfy $R = r^{\theta}$, yielding a $1-$parameter family of dimensions.

Definition 5.2.2 *The Assouad spectrum of* (X,d) *is the map*

$$\theta \in (0,1) \mapsto \dim_A^{\theta}(X,d) \,=\, \inf\left\{ s : \begin{array}{c} \exists C, \varepsilon_0 > 0 \,:\, \forall 0 < R < \varepsilon_0 \; \forall x \in X \\ S\big(B_R(x), d, R^{1/\theta}\big) \,\le\, C\left(\frac{R}{R^{1/\theta}}\right)^s \end{array} \right\}.$$

Since the Assouad spectrum is defined in terms of a unique decreasing scale, it can also be expressed as a limit. In order to maintain a unified notation we write $\varepsilon \,=\, r \,=\, R^{1/\theta}$, obtaining

$$\dim_A^{\theta}(X,d) \,=\, \limsup_{\varepsilon\to 0^+} \sup_{x\in X} \frac{\log S\big(B_{\varepsilon^{\theta}}(x), d, \varepsilon\big)}{(1-\theta)\log(1/\varepsilon)}. \tag{5.6}$$

The aforementioned notions of dimension are related by the inequalities established in [24, Proposition 3.1]: for every $\theta \in (0,1)$,

$$\overline{\dim}_B(X,d) \,\le\, \dim_A^{\theta}(X,d) \,\le\, \min\left\{ \frac{\overline{\dim}_B(X,d)}{1-\theta}, \dim_A(X,d) \right\}. \tag{5.7}$$

Moreover, the Assouad spectrum varies continuously in $(0,1)$ and can be extended to $\theta = 0$ by $\dim_A^0(X,d) = \overline{\dim}_B(X,d)$. It is also known that, as $\theta \to 1^-$, the spectrum converges to the so called *quasi-Assouad dimension*, which is always finite whenever the Assouad dimension is (see [22]). We observe that the Assouad dimension is finite if and only if (X,d) is *doubling*, that is, there exists some $K > 0$, which we refer to as the *doubling constant of* (X,d), such that any ball of radius ε can be covered by K balls of radius $\varepsilon/2$, for every $\varepsilon > 0$ (see [22, Theorem 13.1.1]).

5.3 Ahlfors regular spaces

As a consequence of Proposition 3.4.5 we know that the local box dimension structure of an alphabet is closely related to the metric mean dimension of the corresponding full shift. In particular, when the space (X,d) is box-dimensionally homogeneous[1], by Theorem C, we have

$$\overline{\mathrm{mdim}}_M(X^{\mathbb{N}}, d^{\mathbb{N}}, \sigma_{\mathbb{N}}, \varphi) \,=\, \dim_B(X,d) + \max_{\mu \in \mathscr{E}_{\sigma_{\mathbb{N}}}(X^{\mathbb{N}})} \int \varphi \, d\mu \qquad \forall \varphi \in C^0(X).$$

In what follows, we will approach dimensional homogeneity through a measure-theoretic perspective, which provides a stronger notion of homogeneity, yielding a new formulation for the metric mean dimension.

[1]We say that (X,d) is *box-dimensionally homogeneous* if $\dim_B(B_\varepsilon(x)) \,=\, \dim_B(X)$, for every $x \in X$ and $\varepsilon > 0$.

Definition 5.3.1 Given a compact metric space (X,d), we say that $\mu \in \mathscr{P}(X)$ is Ahlfors regular if

$$\exists \delta \geq 0,\ \exists \varepsilon_0, C_1, C_2 > 0: \qquad C_1\,\varepsilon^{\delta} \leq \mu(B_\varepsilon(x)) \leq C_2\,\varepsilon^{\delta} \qquad \forall x \in \mathrm{supp}(\mu), \forall 0 < \varepsilon < \varepsilon_0.$$

Whenever (X,d) admits an Ahlfors regular probability measure μ, we will restrict our study to $\mathrm{supp}(\mu)$. Therefore, without loss of generality, in what follows we will assume $\mathrm{supp}(\mu) = X$.

The existence of an Ahlfors regular measure on a compact metric space (X,d) reflects the ultimate form of dimensional homogeneity. Namely, it ensures that the Assouad and the lower dimension, which quantify, respectively, the highest and the lowest local complexity in terms of dimension, coincide and are equal to the exponent δ (see [22, Theorem 6.4.1]). In particular, all the intermediate notions of dimension such as lower spectrum, packing dimension, Hausdorff dimension, intermediate dimension, box dimension and Assouad spectrum coincide (see [19],[20],[22] and references therein). Thus, we will always denote $\delta = \dim(X,d)$, abbreviating into $\dim X$ if the metric is clear.

Definition 5.3.2 *A triplet* (X,d,μ) *is an* Ahlfors regular space *if*

$$\exists \varepsilon_0, C_1, C_2 > 0: \qquad C_1\,\varepsilon^{\dim X} \leq \mu(B_\varepsilon(x)) \leq C_2\,\varepsilon^{\dim X} \qquad \forall x \in X, \forall 0 < \varepsilon < \varepsilon_0. \tag{5.8}$$

The main examples we have in mind are the Euclidean cube $[0,1]^D$ with the Lebesgue measure, a compact Riemannian manifold with the volume measure and (finitely generated) self-similar/self-conformal sets satisfying the open-set-condition with the right dimensional Hausdorff measure (cf. [22, Corollary 6.4.4]). In fact, the classical notion of Ahlfors regularity is defined with the Hausdorff measure, though it may be formulated for a more general μ satisfying (5.8). In particular, the only zero dimensional compact Ahlfors regular spaces are the finite ones.

Every Ahlfors regular space is box-dimensionally homogeneous. However, not every box-dimensionally homogeneous metric space admits an Ahlfors regular measure. An example is the Bedford-McMullen carpet, whose Hausdorff and box dimensions are distinct (cf. [5],[45]).

5.3.1 Minkowski–Bouligand dimension

The box dimension for compact subsets of the Euclidean space admits an equivalent formulation, the so called Minkowski–Bouligand dimension (see [19, Proposition 2.4]). We briefly recall it here.

Let $D \in \mathbb{N}$ and $F \subset \mathbb{R}^D$ be a compact subset of zero Lebesgue measure (otherwise its box dimension is equal to D) and, for $\varepsilon > 0$, consider its $\varepsilon-$neighbourhood

$$B_\varepsilon(F) = \{x \in \mathbb{R}^D : \mathrm{dist}(x,F) < \varepsilon\},$$

where $\mathrm{dist}(x,F) = \inf_{a \in F} |x-a|$. This is an open set and it is natural to ask how fast its Lebesgue measure decreases as $\varepsilon \to 0^+$. It turns out that this rate provides information on the dimension, namely,

$$\overline{\dim}_B(F,|\cdot|) = D + \limsup_{\varepsilon \to 0^+} \frac{\log Leb(B_\varepsilon(F))}{\log(1/\varepsilon)} \quad \text{and} \quad \underline{\dim}_B(F,|\cdot|) = D + \liminf_{\varepsilon \to 0^+} \frac{\log Leb(B_\varepsilon(F))}{\log(1/\varepsilon)}.$$

We are heading to an analogous result in the dynamical setting. To this end, we need an ambient product space where we can embed and thicken, using $\varepsilon-$neighbourhoods, the space of time n orbits. Afterwards, we evaluate the exponential rate of the $n-$dimensional volume as time increases, and then compute the decay of this quantity as the scale ε goes to zero. Subshifts of the Euclidean cube are natural candidates for this procedure. In what follows, we consider more general Ahlfors regular alphabets.

5.3.2 Dynamical Minkowski–Bouligand dimension

The main result in this section is an equivalent formulation of the metric mean dimension of subshifts which makes use of the homogeneous structure of the alphabet. The proof is an adaptation of the one in [19, Proposition 2.4]. Let (X,d,μ) be an Ahlfors regular space. For each $n \geq 1$, denote the product measure by $\mu^n \in \mathscr{P}(X^n)$. Given an arbitrary subshift $Y \subset X^{\mathbb{N}}$ and $\varepsilon > 0$, let

$$Y(n,\varepsilon) := \bigcup_{y=(y_i)_{i\in\mathbb{N}} \in Y} B_\varepsilon(y_1) \times B_\varepsilon(y_2) \times \cdots \times B_\varepsilon(y_n) \subset X^n.$$

We note that $Y(n,\varepsilon)$ is precisely the set $B_\varepsilon(\pi_n(Y))$ with respect to the supremum norm $||\cdot||_\infty$, where $\pi_n \colon X^{\mathbb{N}} \to X^n$ is the projection on the first n coordinates.

Recall that a dynamical ball in the shift space $(X^{\mathbb{N}}, d^{\mathbb{N}}, \sigma_{\mathbb{N}})$ is given by

$$B_\varepsilon(x,d_n) = B_\varepsilon(x_1) \times B_\varepsilon(x_2) \times \cdots B_\varepsilon(x_n) \times B_{2\varepsilon}(x_{n+1}) \times \cdots \times B_{2^\ell\varepsilon}(x_{n+\ell}) \times X \times X \times \cdots$$

where $\ell = \ell(\varepsilon)$ is the smallest positive integer satisfying $2^\ell\varepsilon \geq \mathrm{diam}(X)$. Observe that, for every $\varepsilon > 0$ and $n \geq 1$, the measure $\mu^{\mathbb{N}}$ satisfies

$$\begin{aligned}
\mu^{n+\ell}\big(Y(n+\ell,\varepsilon)\big) &= \mu^{n+\ell}\Big(\bigcup_{y\in Y} B_\varepsilon(y_1) \times \cdots \times B_\varepsilon(y_{n+\ell})\Big) \\
&\leq \mu^{\mathbb{N}}\Big(\bigcup_{y\in Y} B_\varepsilon(y,d_n)\Big) \\
&\leq \mu^{n}\Big(\bigcup_{y\in Y} B_\varepsilon(y_1) \times \cdots \times B_\varepsilon(y_{n})\Big) = \mu^n\big(Y(n,\varepsilon)\big).
\end{aligned}$$

In particular, for the proof we can replace $\mu^n\big(Y(n,\varepsilon)\big)$ by $\mu^{\mathbb{N}}\big(Y_{n,\varepsilon}\big)$, where

$$Y_{n,\varepsilon} = \big\{x \in X^{\mathbb{N}} : d_n(x,Y) \leq \varepsilon\big\}.$$

Lemma 5.3.3 *Let (X,d,μ) be an Ahlfors regular space and $Y \subset X^{\mathbb{N}}$ be a subshift. For every $\varepsilon > 0$, the sequence $\Big(\log\mu^n\big(Y(n,\varepsilon)\big)\Big)_n$ is subadditive. In particular, the following limits exist*

$$\lim_{n\to+\infty} \frac{1}{n}\log\mu^n\big(Y(n,\varepsilon)\big) = \lim_{n\to+\infty} \frac{1}{n}\log\mu^{\mathbb{N}}\big(Y_{n,\varepsilon}\big).$$

Proof: Fix $\varepsilon > 0$. For every $y = (y_1, y_2, ...) \in Y$ and $n, k \in \mathbb{N}$,

$$\begin{aligned} B_\varepsilon(y_1) \times \cdots \times B_\varepsilon(y_{n+k}) &\subset B_\varepsilon(y_1) \times \cdots \times B_\varepsilon(y_k) \times \bigcup_{z=(z_i)_i \in Y} B_\varepsilon(z_{k+1}) \times \cdots \times B_\varepsilon(z_{n+k}) \\ &\subset B_\varepsilon(y_1) \times \cdots \times B_\varepsilon(y_k) \times \bigcup_{w=(w_i)_i \in Y} B_\varepsilon(w_1) \times \cdots \times B_\varepsilon(w_n) \\ &\subset B_\varepsilon(y_1) \times \cdots \times B_\varepsilon(y_k) \times Y(n,\varepsilon). \end{aligned}$$

Therefore,

$$\begin{aligned} \mu^{n+k}\Big(Y(n+k,\varepsilon)\Big) &\leq \mu^{n+k}\Big(\bigcup_{y=(y_i)_i \in Y} B_\varepsilon(y_1) \times \cdots \times B_\varepsilon(y_k) \times Y(n,\varepsilon)\Big) \\ &= \mu^{n+k}\Big(Y(k,\varepsilon) \times Y(n,\varepsilon)\Big) \\ &= \mu^k\Big(Y(k,\varepsilon)\Big) \cdot \mu^n\Big(Y(n,\varepsilon)\Big). \end{aligned}$$

□

Proof of Theorem J: Fix $\varepsilon > 0$ and $n \geq 1$. If Y can be covered by $N(n,\varepsilon)$ distinct $(n,\varepsilon)-$dynamical balls, then $Y_{n,\varepsilon}$ can be covered by the concentric dynamical balls of radius 2ε. Then

$$\mu^{\mathbb{N}}\big(Y_{n,\varepsilon}\big) \leq N(n,\varepsilon)\big(C_2(2\varepsilon)^{\dim X}\big)^n,$$

and therefore

$$\lim_{n\to+\infty} \frac{1}{n}\log\mu^{\mathbb{N}}\big(Y_{n,\varepsilon}\big) \leq \limsup_{n\to+\infty} \frac{1}{n}\log N(n,\varepsilon) + \log\big(C_2(2\varepsilon)^{\dim X}\big). \tag{5.9}$$

Hence,

$$\limsup_{\varepsilon\to 0^+} \frac{1}{\log(1/\varepsilon)} \lim_{n\to+\infty} \frac{1}{n}\log\mu^{\mathbb{N}}\big(Y_{n,\varepsilon}\big) \leq \overline{\mathrm{mdim}}_M(Y, d^{\mathbb{N}}, \sigma_{\mathbb{N}}) - \dim X.$$

For the converse inequality, if there are $S_1(n,\varepsilon)$ distinct $(n,\varepsilon)-$separated points $y_1, \cdots, y_{S_1(n,\varepsilon)}$ in Y, then the $(n,\varepsilon/3)-$dynamical balls centered at these points are disjoint. Then

$$\mu^{\mathbb{N}}\big(Y_{n,\varepsilon}\big) \geq \sum_{i=1}^{S_1(n,\varepsilon)} \mu^{\mathbb{N}}\big(B_{\varepsilon/3}(y_i, d_n)\big) \geq S_1(n,\varepsilon)\big(C_1(\varepsilon/3)^{\dim X}\big)^{n+\ell(\varepsilon)},$$

so

$$\lim_{n\to+\infty} \frac{1}{n}\log\mu^{\mathbb{N}}\big(Y_{n,\varepsilon}\big) \geq \limsup_{n\to+\infty} \frac{1}{n}\log S_1(n,\varepsilon) + \log(C_1(\varepsilon/3)^{\dim X}). \tag{5.10}$$

Thus,

$$\limsup_{\varepsilon\to 0^+} \frac{1}{\log(1/\varepsilon)} \lim_{n\to+\infty} \frac{1}{n}\log\mu^{\mathbb{N}}\big(Y_{n,\varepsilon}\big) \geq \overline{\mathrm{mdim}}_M(Y, d^{\mathbb{N}}, \sigma_{\mathbb{N}}) - \dim X.$$

This concludes the proof of (5.5). The statement for the lower metric mean dimension is proved analogously.

Regarding the entropy, taking the limsup as $\varepsilon \to 0^+$ in (5.9) and (5.10), we obtain

$$h_1 \leq h_{top}(Y, \sigma_{\mathbb{N}}) \leq h_2 \tag{5.11}$$

where

$$\begin{aligned} h_1 &= \limsup_{\varepsilon \to 0^+} \lim_{n \to +\infty} \frac{1}{n} \log\Big(\frac{\mu^n\big(Y(n,\varepsilon)\big)}{(C_2\, 2^{dimX})^n\, \varepsilon^{n\,\dim X}}\Big) \\ h_2 &= \limsup_{\varepsilon \to 0^+} \lim_{n \to +\infty} \frac{1}{n} \log\Big(\frac{\mu^n\big(Y(n,\varepsilon)\big)}{(C_1\, 3^{-\dim X})^n\, \varepsilon^{n\,\dim X}}\Big). \end{aligned}$$

When $\dim X = 0$, X is a finite set, say $X = \{x_1, ..., x_q\}$. If we consider the normalized counting measure μ, then $C_1 = C_2 = 1/q$ and (5.4), in this setting, is an immediate consequence of (5.11), with $\lambda = 1/q$. On the other hand, if $\dim X > 0$, we have two possibilities. If $h_{top}(Y, \sigma_{\mathbb{N}}) = +\infty$, then (5.4) holds with $\lambda = C_1\, 3^{-\dim X}$. If $h_{top}(Y, \sigma_{\mathbb{N}}) < +\infty$, let us show that

$$-\infty \;<\; \limsup_{\varepsilon \to 0^+} \lim_{n \to +\infty} \frac{1}{n} \log\Big(\frac{\mu^n\big(Y(n,\varepsilon)\big)}{\varepsilon^{n\,\dim X}}\Big) \;<\; +\infty.$$

Indeed, by Definition 5.3.2, for any $y \in Y$ one has

$$\mu^n\big(Y(n,\varepsilon)\big) \;\geq\; \mu^n\big(B_\varepsilon(y_1) \times ... \times B_\varepsilon(y_n)\big) \;\geq (C^1\, \varepsilon^{\dim X})^n$$

and therefore

$$\limsup_{\varepsilon \to 0^+} \lim_{n \to +\infty} \frac{1}{n} \log\Big(\frac{\mu^n\big(Y(n,\varepsilon)\big)}{\varepsilon^{n\,\dim X}}\Big) \;\geq\; \log C_1.$$

Moreover,

$$\limsup_{\varepsilon \to 0^+} \lim_{n \to +\infty} \frac{1}{n} \log\Big(\frac{\mu^n\big(Y(n,\varepsilon)\big)}{\varepsilon^{n\,\dim X}}\Big) \;\leq\; h_{top}(Y, \sigma_{\mathbb{N}}) \;+\; \log\big(C_2\, 2^{\dim X}\big) \;<\; +\infty.$$

Now, consider the strictly decreasing map $G\colon\,]0, +\infty[\to \mathbb{R}$ given by

$$G(t) \;=\; \limsup_{\varepsilon \to 0^+} \lim_{n \to +\infty} \frac{1}{n} \log\Big(\frac{\mu^n\big(Y(n,\varepsilon)\big)}{t^n\, \varepsilon^{n\,\dim X}}\Big) \;=\; \limsup_{\varepsilon \to 0^+} \lim_{n \to +\infty} \frac{1}{n} \log\Big(\frac{\mu^n\big(Y(n,\varepsilon)\big)}{\varepsilon^{n\,\dim X}}\Big) \;-\; \log t.$$

Observe that G is continuous, $G(C_2\, 2^{\dim X}) = h_1$ and $G(C_1\, (1/3)^{\dim X}) = h_2$. Thus, due to (5.11) and the intermediate value theorem, there exists a unique $\lambda \in [C_1\, (1/3)^{\dim X}, C_2\, 2^{\dim X}]$ such that $G(\lambda) = h_{top}(Y, \sigma_{\mathbb{N}})$, as claimed. The proof of Theorem J is complete.

□

We end this section by proving Corollary 5.1.4.

Proof of Corollary 5.1.4: For a map $F\colon X \to X$, $n \in \mathbb{N}$ and $\varepsilon > 0$, we denote

$$Y(F,n,\varepsilon) \;=\; \bigcup_{x \in X} B_\varepsilon(x) \times B_\varepsilon(F(x)) \times \cdots \times B_\varepsilon(F^{n-1}(x)).$$

Fix $k \geq 1$ and observe that, for every $\varepsilon > 0$ and large enough $n \geq 1$, we have

$$\begin{aligned} Y(T,kn,\varepsilon) &= \bigcup_{x\in X} B_\varepsilon(x) \times B_\varepsilon(T(x)) \times \cdots \times B_\varepsilon(T^{kn-1}(x)) \\ &\subset \bigcup_{x\in X} B_\varepsilon(x) \times X \times \ldots \times X \\ &\qquad \times B_\varepsilon(T^k(x)) \times X \times \ldots \times X \\ &\qquad \vdots \\ &\qquad \times B_\varepsilon(T^{(n-1)k}(x)) \times X \times \ldots \times X \\ &:= \ A. \end{aligned}$$

Then

$$\mu^{kn}\big(Y(T,kn,\varepsilon)\big) \ \leq \ \mu^{kn}\big(A\big) \ = \ \mu^{n}\big(Y(T^k,n,\varepsilon)\big),$$

and therefore

$$\begin{aligned} \lim_{n\to+\infty} \frac{1}{n} \log \mu^n\big(Y(T^k,n,\varepsilon)\big) \ &\geq \ k \lim_{n\to+\infty} \frac{1}{kn} \log \mu^{kn}\big(Y(T,kn,\varepsilon)\big) \\ &= \ k \lim_{n\to+\infty} \frac{1}{n} \log \mu^{n}\big(Y(T,n,\varepsilon)\big). \end{aligned}$$

By Corollary 5.1.3, we conclude that

$$\begin{aligned} \overline{\mathrm{mdim}}_M(X,d,T^k) \ &= \ \dim X \ + \ \limsup_{\varepsilon\to 0^+} \frac{1}{\log(1/\varepsilon)} \lim_{n\to+\infty} \frac{1}{n} \log \mu^n\big(Y(T^k,n,\varepsilon)\big) \\ &\geq \ \dim X \ + \ k \ \limsup_{\varepsilon\to 0^+} \frac{1}{\log(1/\varepsilon)} \lim_{n\to+\infty} \frac{1}{n} \log \mu^n\big(Y(T,n,\varepsilon)\big) \\ &= \ k \ \overline{\mathrm{mdim}}_M(X,d,T) \ - \ (k-1) \dim X. \end{aligned}$$

The statement for the lower metric mean dimension is proved analogously.

□

5.4 Hölder regularity bounds the metric mean dimension

Let (X,d) be a compact metric space with finite Assouad dimension and $T\colon X \to X$ be $\alpha-$Hölder for some $\alpha \in (0,1)$. We will make use of the information in Section 4.5 to estimate the metric mean dimension of T, by relating its Hölder exponent, the transition matrix (4.17) and the Assouad spectrum of parameter $\theta = \alpha$ (see Definition 5.2.2).

Fix $\varepsilon > 0$. Let $x_1,\ldots,x_N \in X$ be a maximal collection of $\varepsilon-$ separated points and consider the open cover $\mathscr{U}(2\varepsilon) = \{B_\varepsilon(x_1),\ldots,B_\varepsilon(x_N)\}$ of X, whose diameter is 2ε. Let us estimate the number of elements in this cover that can intersect a given ball.

Lemma 5.4.1 *For every $z \in X$ and $R > 2\varepsilon > 0$,*

$$\#\{i : B_\varepsilon(x_i) \cap B_R(z) \neq \emptyset\} \leq K \sup_{x \in X} S\big(B_R(x), d, \varepsilon\big),$$

where K is the doubling constant of (X,d).

Proof: Recall that the balls $\{B_{\varepsilon/2}(x_i) : i = 1, \ldots, N\}$ are pairwise disjoint. Reordering if necessary, assume that the elements of $\mathscr{U}(2\varepsilon)$ that intersect $B_R(z)$ are $B_\varepsilon(x_1), \ldots, B_\varepsilon(x_q)$. Then $\bigcup_{i=1}^q B_{\varepsilon/2}(x_i) \subset B_{R+2\varepsilon}(z)$, where the union is disjoint. Hence

$$q \leq S\big(B_{R+2\varepsilon}(z), d, \varepsilon\big) \leq S\big(B_{2R}(z), d, \varepsilon\big) \leq K \sup_{x \in X} S\big(B_R(x), d, \varepsilon\big).$$

□

Define the matrix $\Gamma^{2\varepsilon}$ associated to the cover $\mathscr{U}(2\varepsilon)$ as in (4.17). Since the map T is $\alpha-$Hölder, for every $i = 1, \ldots, N$ there exists some $z_i \in X$ such that $T\big(B_\varepsilon(x_i)\big) \subset B_{C_\alpha \varepsilon^\alpha}(z_i)$. By the previous Lemma, $T\big(B_\varepsilon(x_i)\big)$ can intersect at most

$$K \sup_{x \in X} S\big(B_{C_\alpha \varepsilon^\alpha}(x), d, \varepsilon\big) \leq K^{1+M} \sup_{x \in X} S\big(B_{\varepsilon^\alpha}(x), d, \varepsilon\big) \tag{5.12}$$

elements of $\mathscr{U}(2\varepsilon)$ for every $i = 1, \ldots, N$, where $M = \left\lceil \frac{\log 2}{\log C_\alpha} \right\rceil$ and C_α is the Hölder constant of T. This means that (5.12) provides an upper bound for the number of 1's on each row of $\Gamma^{2\varepsilon}$. By Lemma 4.5.1 and (4.19) we have

$$h_{2\varepsilon}(\Gamma_{\mathbb{N}}, d^{\mathbb{N}}, \sigma_{\mathbb{N}}) \leq \log r(\Gamma^{2\varepsilon}) \leq (1+M)\log K + \sup_{x \in X} \log S\big(B_{\varepsilon^\alpha}(x), d, \varepsilon\big),$$

where $\Gamma_{\mathbb{N}}$ is the subshift of compact type induced by $\Gamma = \mathrm{graph}(T)$. Combining the above estimate with Theorem F, we conclude that

$$\overline{\mathrm{mdim}}_M(X, d, T) = \overline{\mathrm{mdim}}_M(\Gamma_{\mathbb{N}}, d^{\mathbb{N}}, \sigma_{\mathbb{N}}) \leq (1-\alpha)\, \dim_A^\alpha(X, d).$$

This ends the proof of the first part of Theorem H. The second part is a straightforward consequence of the last inequality when $\dim_A(X,d) > 0$ and the right inequality in (5.7).

□

5.5 Scaled Misiurewicz formula for $T_{a,b}$

We will split the statement of Theorem I in two. We start by applying Corollary 5.1.3 to compute the upper and lower metric mean dimensions of $T_{a,b}$ (defined in Subsection 5.1.2) with respect to the Euclidean metric.

Proposition 5.5.1 *Let $T_{a,b}$ be the map satisfying the conditions (C1)-(C3) in* (2.7). *Then*

$$\begin{aligned}\overline{\mathrm{mdim}}_M([0,1],|\cdot|,T_{a,b}) &= \limsup_{k\to+\infty}\frac{\log b_k}{\log(1/\varepsilon_k)}\\ \underline{\mathrm{mdim}}_M([0,1],|\cdot|,T_{a,b}) &= \liminf_{k\to+\infty}\frac{\log b_k}{\log(1/\varepsilon_k)}.\end{aligned}$$

Proof: We start recalling that, for every $s\geq 1$ and $n\geq 1$, we can partition J_s in b_s^n subintervals $J_s(i_1,\cdots,i_n)$, for $i_1,...,i_n\in\{1,\cdots,b_s\}$, satisfying

$$T^j_{a,b}(J_s(i_1,\cdots,i_n)) = J_s(i_{j+1},\cdots,i_n) \qquad \forall j\in\{0,\cdots,n-1\}.$$

Given $\varepsilon>0$, let $k=k(\varepsilon)$ be the unique positive integer satisfying $\varepsilon_{k+1}\leq\varepsilon<\varepsilon_k$. Consider a positive integer $M=M(\varepsilon)\gg k$ such that $a_{M-1}>1-\varepsilon$. Then, for every $n\geq 1$,

$$\begin{aligned}Y(n,\varepsilon) &= \bigcup_{x\in[0,1]} B_\varepsilon(x)\times B_\varepsilon(T(x))\times\cdots\times B_\varepsilon(T^{n-1}(x))\\ &= \bigcup_{s=1}^{M}\left(\bigcup_{x\in J_s} B_\varepsilon(x)\times B_\varepsilon(T(x))\times\cdots\times B_\varepsilon(T^{n-1}(x))\right)\\ &:= Y([0,a_k],n,\varepsilon)\cup Y([a_k,a_M],n,\varepsilon),\end{aligned}$$

where

$$\begin{aligned}Y([0,a_k],n,\varepsilon) &= \bigcup_{s=1}^{k}\left(\bigcup_{x\in J_s} B_\varepsilon(x)\times B_\varepsilon(T(x))\times\cdots\times B_\varepsilon(T^{n-1}(x))\right)\\ Y([a_k,a_M],n,\varepsilon) &= \bigcup_{s=k+1}^{M}\left(\bigcup_{x\in J_s} B_\varepsilon(x)\times B_\varepsilon(T(x))\times\cdots\times B_\varepsilon(T^{n-1}(x))\right).\end{aligned}$$

Let us estimate the $n-$th Lebesgue measure of each of these sets separately. Consider a minimal $(n,\varepsilon)-$spanning set $x_1,\cdots,x_N$ in $[0,a_k]$. Then

$$Y([0,a_k],n,\varepsilon)\subset\bigcup_{i=1}^{N}B_{2\varepsilon}(x_i)\times B_{2\varepsilon}(T(x_i))\times\cdots\times B_{2\varepsilon}(T^{n-1}(x_i)),$$

and therefore

$$\begin{aligned}\limsup_{n\to+\infty}\frac{1}{n}\log\mathrm{Leb}^n\left(Y([0,a_k],n,\varepsilon)\right) &\leq h_\varepsilon([0,a_k])+\log(4\varepsilon)\\ &\leq \log b_k + \log\varepsilon_k+\log 4\\ &= \log|J_k| + \log 4. \end{aligned}\tag{5.13}$$

Besides, by invariance of each J_s, we can consider the coarse estimate

$$Y([a_k,a_M],n,\varepsilon)\subset\bigcup_{s=k+1}^{M}\left(B_\varepsilon(J_s)\right)^n.$$

As $(|J_s|)_s$ is decreasing, we get

$$\begin{aligned} \mathrm{Leb}^n\big(Y([a_k,a_M],n,\varepsilon)\big) &\le \sum_{s=k+1}^{M}\big(|J_s|+2\varepsilon\big)^n \\ &\le \big(M(\varepsilon)-k(\varepsilon)\big)\cdot\big(|J_k|+2\varepsilon_k\big)^n \\ &= \big(M(\varepsilon)-k(\varepsilon)\big)\cdot|J_k|^n\cdot\big(1+\frac{2}{b_k}\big)^n, \end{aligned}$$

so

$$\limsup_{n\to+\infty}\frac{1}{n}\log\mathrm{Leb}^n\big(Y([a_k,a_M],n,\varepsilon)\big) \le \log|J_k| + \log\big(1+\frac{2}{b_k}\big) \le \log|J_k| + \log 4. \tag{5.14}$$

Bringing together (5.13), (5.14) and Lemma 5.3.3 we obtain

$$\lim_{n\to+\infty}\frac{1}{n}\log\mathrm{Leb}^n\big(Y(n,\varepsilon)\big) \le \log|J_k| + \log 4. \tag{5.15}$$

On the other hand, it is easy to see that for every $n\ge 1$, $J_{k+1}^n \subset Y(n,\varepsilon)$. Indeed, given $(x_1,\cdots,x_n)\in J_{k+1}^n$ there exists i_j such that $x_j\in J_{k+1}(i_j)$ for every $j=1,\dots,n$. Then, any element $y\in J_{k+1}(i_1,\cdots,i_n)$ $\varepsilon_{k+1}-$shadows this trajectory along the first n iterates. Therefore,

$$(x_1,\cdots,x_n)\in B_\varepsilon(y)\times B_\varepsilon(T(y))\times\cdots\times B_\varepsilon(T^{n-1}(x)) \subset Y(n,\varepsilon).$$

Combining this information with (5.15) we get

$$\log|J_{k+1}| \le \lim_{n\to+\infty}\frac{1}{n}\log\mathrm{Leb}^n\big(Y(n,\varepsilon)\big) \le \log|J_k|+\log 4,$$

thus

$$\frac{\log|J_{k+1}|}{\log(1/\varepsilon_{k+1})} \le \frac{1}{\log(1/\varepsilon)}\lim_{n\to+\infty}\frac{1}{n}\log\mathrm{Leb}^n\big(Y(n,\varepsilon)\big) \le \frac{\log|J_k|+\log 4}{\log(1/\varepsilon_k)}.$$

The result now follows from Corollary 5.1.3 and the fact that

$$1+\frac{\log|J_s|}{\log(1/\varepsilon_s)} = \frac{\log b_s}{\log(1/\varepsilon_s)} \qquad \forall s\ge 1.$$

□

The previous reasoning does not depend on the affine structure of the f_k's. In fact, not even continuity was used: as long as, for every $k\ge 1$, $\tilde f_k$ satisfies both the geometric constraint (equidistribution of critical points) and the topological constraint (strictly monotone full branches), the maps $\tilde T_{a,b}$ and $T_{a,b}$ have the same upper and lower metric mean dimensions.

To end the proof of Theorem I we are left to link the upper metric mean dimension and the Hölder regularity of the map $T_{a,b}$.

Proposition 5.5.2 *Let $T_{a,b}$ be the map satisfying the conditions (C1)-(C3) in* (2.7). *Then,*

$$\sup\mathscr{H}(T_{a,b}) = 1 - \overline{\mathrm{mdim}}_M([0,1],|\cdot|,T_{a,b}),$$

where $\mathscr{H}(T_{a,b}) = \{\alpha \in (0,1) : T_{a,b} \textit{ is } \alpha-\textit{Hölder}\} \cup \{0\}$.

Proof: Let $T = T_{a,b}$. By Theorem H we already know that

$$\sup \mathscr{H}(T) \leq 1 - \overline{\mathrm{mdim}}_M([0,1], |\cdot|, T).$$

For the converse inequality, we may assume that $\overline{\mathrm{mdim}}_M([0,1], |\cdot|, T) < 1$, otherwise there is nothing else to prove. Let $\alpha \in (0,\, 1 - \overline{\mathrm{mdim}}_M([0,1], |\cdot|, T))$. Our aim is to bound

$$\sup_{x \neq y \in [0,1]} \frac{|T(x) - T(y)|}{|x-y|^\alpha}.$$

To this end we will consider the possible different location of the points x, y with respect to the decomposition provided by $(J_k)_k$.

Case 1: $x, y \in J_k$ for some $k \geq 1$.

Let

$$H_k(\alpha) = \sup_{x \neq y \in J_k} \frac{|T(x) - T(y)|}{|x-y|^\alpha} = \frac{|T(a_{k-1}) - T(a_{k-1} + \varepsilon_k)|}{\varepsilon_k^\alpha} = \frac{|J_k|}{\varepsilon_k^\alpha} = \frac{b_k^\alpha}{|J_k|^{\alpha - 1}}.$$

Observe that

$$\begin{aligned}
\log H_k(\alpha) &= \alpha \log b_k - (\alpha - 1) \log |J_k| \\
&= \log b_k \Big((\alpha - 1)\big(1 - \frac{\log |J_k|}{\log b_k}\big) + 1 \Big) \\
&= \log b_k \Big((\alpha - 1)\big(\frac{\log(1/\varepsilon_k)}{\log b_k}\big) + 1 \Big). \qquad (5.16)
\end{aligned}$$

The maximal distortion inside J_k is given precisely by $H_k(\alpha)$. By choice of α, we know that

$$\limsup_{k \to +\infty} \Big((\alpha - 1)\big(\frac{\log(1/\varepsilon_k)}{\log b_k}\big) + 1 \Big) < 0.$$

In particular, combining this information with (5.16) we obtain that $H_k(\alpha)$ is strictly smaller than 1 for large enough k. Hence, $(H_k(\alpha))_k$ is uniformly bounded and we obtain

$$\sup_{k \geq 1} \sup_{x \neq y \in J_k} \frac{|T(x) - T(y)|}{|x-y|^\alpha} = \sup_{k \geq 1} H_k(\alpha) < +\infty. \qquad (5.17)$$

Case 2: $x \in J_k$ for some $k \geq 1$ and $y = 1$.

The maximal distortion between 1 and some element in $x \in J_k$ is attained at $x = a_k - \varepsilon_k$, namely

$$\sup_{x \in J_k} \frac{|1 - T(x)|}{|1-x|^\alpha} = \frac{1 - a_{k-1}}{(1 - a_k + \varepsilon_k)^\alpha}.$$

By concavity of the map $t \mapsto t^{\alpha}$, we have

$$\begin{aligned}
\frac{1-a_{k-1}}{(1-a_k+\varepsilon_k)^{\alpha}} &= \frac{(1-a_k)+|J_k|}{(1-a_k+\varepsilon_k)^{\alpha}} \\
&\leq 2^{1-\alpha}\Big(\frac{(1-a_k)+|J_k|}{(1-a_k)^{\alpha}+\varepsilon_k^{\alpha}}\Big) \\
&\leq 2^{1-\alpha}\Big((1-a_k)^{1-\alpha}+\frac{|J_k|}{\varepsilon_k^{\alpha}}\Big) \\
&\leq 2^{1-\alpha}\Big(1+H_k(\alpha)\Big).
\end{aligned}$$

Thus,

$$\sup_{k\geq 1}\sup_{x\in J_k}\frac{|1-T(x)|}{|1-x|^{\alpha}} < +\infty. \tag{5.18}$$

Case 3: $x \in J_k$ and $y \in J_{k+n}$ for some $k,n \geq 1$.

The maximal distortion between $x \in J_k$ and $y \in J_{k+n}$ is attained at $x = a_k - \varepsilon_k$ and $y = a_{k+n-1}+\varepsilon_{k+n}$, namely

$$\sup_{x\in J_k, y\in J_{k+n}}\frac{|T(y)-T(x)|}{|y-x|^{\alpha}} = \frac{a_{k+n}-a_{k-1}}{(a_{k+n-1}+\varepsilon_{k+n}-a_k+\varepsilon_k)^{\alpha}} = \frac{\sum_{i=k}^{k+n}|J_i|}{(\varepsilon_k+\sum_{i=k+1}^{k+n-1}|J_i|+\varepsilon_{k+n})^{\alpha}},$$

where the denominator is $(\varepsilon_k + \varepsilon_{k+n})^{\alpha}$ if $n = 1$. Again by concavity, we obtain

$$\begin{aligned}
\frac{\sum_{i=k}^{k+n}|J_i|}{(\varepsilon_k+\sum_{i=k+1}^{k+n-1}|J_i|+\varepsilon_{k+n})^{\alpha}} &\leq 3^{1-\alpha}\Big(\frac{|J_k|+\sum_{i=k+1}^{k+n-1}|J_i|+|J_{k+n}|}{\varepsilon_k^{\alpha}+(\sum_{i=k+1}^{k+n-1}|J_i|)^{\alpha}+\varepsilon_{k+n}^{\alpha}}\Big) \\
&\leq 3^{1-\alpha}\Big(\frac{|J_k|}{\varepsilon_k^{\alpha}}+(\sum_{i=k+1}^{k+n-1}|J_i|)^{1-\alpha}+\frac{|J_{k+n}|}{\varepsilon_{k+n}^{\alpha}}\Big) \\
&\leq 3^{1-\alpha}\Big(H_k(\alpha)+1+H_{k+n}(\alpha)\Big).
\end{aligned}$$

Hence,

$$\sup_{k,n\geq 1}\sup_{x\in J_k, y\in J_{k+n}}\frac{|T(y)-T(x)|}{|y-x|^{\alpha}} < +\infty. \tag{5.19}$$

Bringing (5.17), (5.18) and (5.19) together we conclude that

$$\sup_{x\neq y\in[0,1]}\frac{|T(y)-T(x)|}{|y-x|^{\alpha}} < +\infty,$$

which means that T is $\alpha-$Hölder as claimed.

□

5.6 Open questions

In [17] the authors showed that, given a compact manifold X of dimension greater than one, the closure of the space of bi-Lipschitz homeomorphisms of X with respect to the Hölder topology contains a residual subset of homeomorphisms with infinite topological entropy. Theorem H suggests that we ask:

Question 5 *Given a compact smooth manifold* (X,d) *with dimension* $\dim X \geq 2$ *and* $\alpha \in (0,1)$*, let* $\mathscr{L}_\alpha$ *be the closure of the space of bi-Lipschitz homeomorphisms of X with respect to the* $\alpha-$*Hölder norm*

$$||T||_\alpha = ||T||_\infty + \sup_{x \neq y \in X} \frac{d(T(x),T(y))}{d(x,y)^\alpha}.$$

Is it true that a generic map $T \in \mathscr{L}_\alpha$ *satisfies*

$$\overline{\mathrm{mdim}}_M(X,d,T) = (1-\alpha)\dim X?$$

In Theorem I we express the metric mean dimension of a family of interval maps in terms of the data of a sequence of horseshoes in a manner similar to Misiurewicz formula for the entropy. The main difference is the relevance of the scales at which the dynamics of each horseshoe can be detected. This is not surprising since the metric mean dimension, contrary to the entropy, is not purely topological, depending also on the geometry of the underlying space.

Question 6 *Given a continuous interval endomorphism with positive metric mean dimension, can we find a sequence of horseshoes for iterates of the map and a sequence of scales upon which we may build a formula for the metric mean dimension similar to the one of Misiurewicz?*

For finite entropy maps acting on Ahlfors regular spaces, the computation of the entropy provided by Theorem J brings forth a constant λ, whose nature is worthwhile exploring.

Question 7 *Is there a dynamical or geometrical interpretation of this constant?*

References

[1] Acevedo, J. M. (2022). Genericity of continuous maps with positive metric mean dimension. *Results Math.*, 77(1):Paper No. 2, 30.

[2] Alves, J. F. (2020). *Nonuniformly Hyperbolic Attractors: Geometric and Probabilistic Aspects*. Springer Monographs in Mathematics.

[3] Bandt, C., Falconer, K., Zähle, M., et al. (2015). *Fractal geometry and stochastics V*, volume 70. Springer.

[4] Baraviera, A., Carvalho, M., and Pessil, G. (2024). Metric mean dimension, hölder regularity and assouad spectrum. *arXiv:2407.15774 (Accepted for publication in Journal of Fractal Geometry).*

[5] Bedford, T. (1984). *Crinkly curves, Markov partitions and box dimensions in self-similar sets.* PhD thesis, Thesis (Ph. D.)–The University of Warwick.

[6] Biś, A., Carvalho, M., Mendes, M., and Varandas, P. (2022). A convex analysis approach to entropy functions, variational principles and equilibrium states. *Comm. Math. Phys.*, 394(1):215–256.

[7] Biś, A., Carvalho, M., Mendes, M., Varandas, P., and Zhong, X. (2023). Correction: A convex analysis approach to entropy functions, variational principles and equilibrium states. *Comm. Math. Phys.*, 401(3):3335–3342.

[8] Bowen, R. (1970). Topological entropy and axiom A. In *Global Analysis (Proc. Sympos. Pure Math., Vols. XIV, XV, XVI, Berkeley, Calif., 1968)*, volume XIV-XVI of *Proc. Sympos. Pure Math.*, pages 23–41. Amer. Math. Soc., Providence, RI.

[9] Bufetov, A. (1999). Topological entropy of free semigroup actions and skew-product transformations. *J. Dynam. Control Systems*, 5(1):137–143.

[10] Burguet, D. and Shi, R. (2024). Mean dimension of continuous cellular automata. *Israel J. Math.*, 259(1):311–346.

[11] Carvalho, M., Pessil, G., and Varandas, P. (2024). A convex analysis approach to the metric mean dimension: limits of scaled pressures and variational principles. *Adv. Math.*, 436:Paper No. 109407, 54 pp.

[12] Carvalho, M., Rodrigues, F. B., and Varandas, P. (2018). A variational principle for free semigroup actions. *Adv. Math.*, 334:450–487.

[13] Carvalho, M., Rodrigues, F. B., and Varandas, P. (2022a). Generic homeomorphisms have full metric mean dimension. *Ergodic Theory Dynam. Systems*, 42(1):40–64.

[14] Carvalho, M., Rodrigues, F. B., and Varandas, P. (2022b). A variational principle for the metric mean dimension of free semigroup actions. *Ergodic Theory Dynam. Systems*, 42(1):65–85.

[15] Chen, H., Cheng, D., and Li, Z. (2022). Upper metric mean dimensions with potential. *Results Math.*, 77(1):Paper No. 54, 26.

[16] Cheng, D. and Li, Z. (2023). Scaled pressure of dynamical systems. *J. Differential Equations*, 342:441–471.

[17] de Faria, E., Hazard, P., and Tresser, C. (2017). Infinite entropy is generic in Hölder and Sobolev spaces. *C. R. Math. Acad. Sci. Paris*, 355(11):1185–1189.

[18] Falconer, K. (1997). *Techniques in fractal geometry*. John Wiley & Sons, Ltd., Chichester.

[19] Falconer, K. (2014). *Fractal geometry: mathematical foundations and applications*. John Wiley & Sons.

[20] Falconer, K. J., Fraser, J. M., and Kempton, T. (2020). Intermediate dimensions. *Math. Z.*, 296(1-2):813–830.

[21] Feng, D.-J. and Huang, W. (2016). Variational principle for weighted topological pressure. *J. Math. Pures Appl. (9)*, 106(3):411–452.

[22] Fraser, J. M. (2021). *Assouad dimension and fractal geometry*, volume 222 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge.

[23] Fraser, J. M. and Yu, H. (2018a). Assouad-type spectra for some fractal families. *Indiana Univ. Math. J.*, 67(5):2005–2043.

[24] Fraser, J. M. and Yu, H. (2018b). New dimension spectra: finer information on scaling and homogeneity. *Adv. Math.*, 329:273–328.

[25] Friedland, S. (1996). Entropy of graphs, semigroups and groups. In *Ergodic theory of $\mathbf{Z}^d$ actions (Warwick, 1993–1994)*, volume 228 of *London Math. Soc. Lecture Note Ser.*, pages 319–343. Cambridge Univ. Press, Cambridge.

[26] Gershgorin, S. A. (1931). Uber die abgrenzung der eigenwerte einer matrix. . , (6):749–754.

[27] Gromov, M. (1999). Topological invariants of dynamical systems and spaces of holomorphic maps. I. *Math. Phys. Anal. Geom.*, 2(4):323–415.

[28] Gutman, Y. and Śpiewak, A. (2020). Metric mean dimension and analog compression. *IEEE Trans. Inform. Theory*, 66(11):6977–6998.

[29] Gutman, Y. and Śpiewak, A. (2021). Around the variational principle for metric mean dimension. *Studia Math.*, 261(3):345–360.

[30] Gutman, Y. and Tsukamoto, M. (2020). Embedding minimal dynamical systems into Hilbert cubes. *Invent. Math.*, 221(1):113–166.

[31] Hardy, G. H. (1916). Weierstrass's non-differentiable function. *Trans. Amer. Math. Soc.*, 17(3):301–325.

[32] Hazard, P. (2020). Maps in dimension one with infinite entropy. *Ark. Mat.*, 58(1):95–119.

[33] Iommi, G. and Velozo, A. (2021). Pressure, Poincaré series and box dimension of the boundary. *Nonlinearity*, 34(6):3936–3952.

[34] Katok, A. (1980). Lyapunov exponents, entropy and periodic orbits for diffeomorphisms. *Inst. Hautes Études Sci. Publ. Math.*, (51):137–173.

[35] Katok, A. and Hasselblatt, B. (1995). *Introduction to the modern theory of dynamical systems*, volume 54 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge. With a supplementary chapter by Katok and Leonardo Mendoza.

[36] Kawabata, T. and Dembo, A. (1994). The rate-distortion dimension of sets and measures. *IEEE Trans. Inform. Theory*, 40(5):1564–1572.

[37] Kloeckner, B. and Mihalache, N. (2022). An invitation to rough dynamics: zipper maps. *arXiv:2210.13038v1*.

[38] Kolyada, S. i. and Snoha, v. L. (1996). Topological entropy of nonautonomous dynamical systems. *Random Comput. Dynam.*, 4(2-3):205–233.

[39] Lacerda, G. and Romaña, S. (2024). Typical Conservative Homeomorphisms Have Total Metric Mean Dimension. *IEEE Trans. Inform. Theory*, 70(11):7664–7672.

[40] Lindenstrauss, E. (1999). Mean dimension, small entropy factors and an embedding theorem. *Inst. Hautes Études Sci. Publ. Math.*, (89):227–262.

[41] Lindenstrauss, E. and Tsukamoto, M. (2018). From rate distortion theory to metric mean dimension: variational principle. *IEEE Trans. Inform. Theory*, 64(5):3590–3609.

[42] Lindenstrauss, E. and Tsukamoto, M. (2019). Double variational principle for mean dimension. *Geom. Funct. Anal.*, 29(4):1048–1109.

[43] Lindenstrauss, E. and Weiss, B. (2000). Mean topological dimension. *Israel Journal of Mathematics*, 115:1–24.

[44] Lopes, A., Mengue, J., and Oliveira, E. (2024). Idempotent approach to level-2 variational principles in thermodynamical formalism. *arXiv:2409.19203*.

[45] McMullen, C. (1984). The Hausdorff dimension of general Sierpiński carpets. *Nagoya Math. J.*, 96:1–9.

[46] Misiurewicz, M. (1980). Horseshoes for continuous mappings of an interval. In *Dynamical systems (Bressanone, 1978)*, pages 125–135. Liguori, Naples.

[47] Muentes, J., Becker, A. J., Baraviera, A. T., and Scopel, E. (2024). Metric Mean Dimension and Mean Hausdorff Dimension Varying the Metric. *Qual. Theory Dyn. Syst.*, 23:Paper No. 261.

[48] Muentes Acevedo, J. d. J., Romaña Ibarra, S. A., and Arias Cantillo, R. d. J. (2023). Hölder continuous maps on the interval with positive metric mean dimension. *Rev. Colombiana Mat.*, 57:57–76.

[49] Pessil, G. (2021). Dimensão métrica média de deslocamentos de tipo finito em alfabetos compactos. Master's thesis, Universidade Federal do Rio Grande do Sul - UFRGS, Brazil.

[50] Pessil, G. (2025). Metric mean dimension via subshifts of compact type. *Nonlinearity*, 38(4):Paper No. 045018, 29 pp.

[51] Pollicott, M. and Weiss, H. (1999). Multifractal analysis of Lyapunov exponent for continued fraction and Manneville-Pomeau transformations and applications to Diophantine approximation. *Comm. Math. Phys.*, 207(1):145–171.

[52] Shen, W. (2018). Hausdorff dimension of the graphs of the classical weierstrass functions. *Mathematische Zeitschrift*, 289:223–266.

[53] Shi, R. (2022). On variational principles for metric mean dimension. *IEEE Trans. Inform. Theory*, 68(7):4282–4288.

[54] Sigmund, K. (1974). On dynamical systems with the specification property. *Trans. Amer. Math. Soc.*, 190:285–299.

[55] Tsukamoto, M. (2018). Mean dimension of the dynamical system of Brody curves. *Invent. Math.*, 211(3):935–968.

[56] Tsukamoto, M. (2019). Mean dimension of full shifts. *Israel J. Math.*, 230(1):183–193.

[57] Tsukamoto, M. (2020). Double variational principle for mean dimension with potential. *Adv. Math.*, 361:106935, 53.

[58] Tsukamoto, M. (2022). Remark on the local nature of metric mean dimension. *Kyushu J. Math.*, 76(1):143–162.

[59] Velozo, A. and Velozo, R. (2017). Rate distortion theory, metric mean dimension and measure theoretic entropy. *arXiv:1707.05762*.

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

[61] Yang, R., Chen, E., and Zhou, X. (2022). On variational principle for upper metric mean dimension with potential. *arXiv: 2207.01901*.

[62] Yano, K. (1980). A remark on the topological entropy of homeomorphisms. *Invent. Math.*, 59(3):215–220.

[63] Ye, X. and Zhang, G. (2007). Entropy points and applications. *Transactions of the American Mathematical Society*, 359(12):6167–6186.