HAL open science

On emergence and complexity of ergodic decompositions

Pierre Berger, Jairo Bochi

To cite this version:

Pierre Berger, Jairo Bochi. On emergence and complexity of ergodic decompositions. Advances in Mathematics, 2021, 390, pp.107904. 10.1016/j.aim.2021.107904 . hal-03520600

HAL Id: hal-03520600
https://hal.sorbonne-universite.fr/hal-03520600

Submitted on 11 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ON EMERGENCE AND COMPLEXITY OF ERGODIC DECOMPOSITIONS

PIERRE BERGER* AND JAIRO BOCHI**

Abstract. A concept of *emergence* was recently introduced in [Be2] in order to quantify the richness of possible statistical behaviors of orbits of a given dynamical system. In this paper, we develop this concept and provide several new definitions, results, and examples. We introduce the notion of *topological emergence* of a dynamical system, which essentially evaluates how big the set of all its ergodic probability measures is. On the other hand, the *metric emergence* of a particular reference measure (usually Lebesgue) quantifies how non-ergodic this measure is. We prove fundamental properties of these two emergences, relating them with classical concepts such as Kolmogorov's ϵ-entropy of metric spaces and quantization of measures. We also relate the two types of emergences by means of a variational principle. Furthermore, we provide several examples of dynamics with high emergence. First, we show that the topological emergence of some standard classes of hyperbolic dynamical systems is essentially the maximal one allowed by the ambient. Secondly, we construct examples of smooth area-preserving diffeomorphisms that are extremely non-ergodic in the sense that the metric emergence of the Lebesgue measure is essentially maximal. These examples confirm that super-polynomial emergence indeed exists, as conjectured in [Be2]. Finally, we prove that such examples are locally generic among smooth diffeomorphisms.

Contents

Date: January, 2019 (first version); June, 2021 (final version).

1991 *Mathematics Subject Classification.* 37A35; 37C05, 37C45, 37C40, 37J40.

Key words and phrases. Emergence, ergodic decomposition, Wassertein space, covering number, quantization number, dimension, entropy, KAM theorem.

* Partially supported by the ERC project 818737 *Emergence of wild differentiable dynamical systems.* 818737 Emergence of wild differentiable dynamicalsystems. ** Partially supported by project Fondecyt 1180371 and project Conicyt PIA ACT172001 *New trends in ergodic theory.*

Introduction

An unsophisticated but fruitful way of quantifying the size of a compact metric space X goes as follows: one counts how many points can be distinguished up to error $\epsilon > 0$, and then studies the behavior of this number $N(\epsilon)$ as the resolution ϵ tends to zero. For example, if $N(\epsilon)$ is of the order of ϵ^{-d}, for some $d > 0$, then we say that X has (box-counting) dimension d. This dimension, when it exists, is a geometric invariant of X: it is preserved under bi-Lipschitz maps.

A similar idea can be used to define invariants of dynamical systems. One considers how many orbits can be distinguished up to time $t > 0$ and up to a fine resolution; if this number is roughly $\exp(h \cdot t)$ then the dynamics has topological entropy h. The metric entropy (also called Kolmogorov–Sinai entropy) of an invariant probability measure can be characterized similarly: in this case we are allowed to disregard a set of orbits of small probability.

This discretization paradigm can be used to quantify the complexity of a dynamical system in another way, called *emergence*, which was recently introduced in [Be2]. Emergence is significant when a finite number of statistics is not enough to describe the behavior of the orbit of almost every point. In this paper, we carry out a more detailed study of the concept of emergence, sometimes guided by analogies with the concept of entropy. Furthermore, we provide examples of topologically generic dynamics with high emergence, substantiating a conjecture from [Be2].

Let us note that the word *emergence* is used with several different meanings in the scientific literature. Our use is compatible with MacKay's viewpoint, according to whom "emergence means non-unique statistical behaviour" [MK]. He elaborates on this as follows:

> "Note that emergence is very different from chaos, in which sensitive dependence produces highly non-unique trajectories according to their initial conditions. Indeed, the nicest forms of chaos produce unique statistical behaviors in the basin of the attractor. The distinction is like that between the weather and the climate. For weather we care about individual realizations; for climate we care about statistical averages." [MK]

Given a continuous self-map f of a compact metric space X, emergence distinguishes only the statistical behavior of orbits of f. So it does not matter *when* a segment of orbit $(f^i(x))_{i=0}^{n-1}$ visits a certain region of the phase space X, but only *how often*. This can be quantified by a probability, the n^{th} *empirical measure associated to* x:

$$\mathbf{e}_n^f(x) := \frac{1}{n}\sum_{i=0}^{n-1} \delta_{f^i(x)} \ .$$

In the paradigm of ergodic theory, one focuses on the probability measures μ which are *invariant*: $f_*\mu = \mu$. We denote by $\mathcal{M}_f(X)$ the convex, closed subset of such measures. Then, by Birkhoff ergodic theorem, for μ-a.e. $x \in X$, the sequence $(\mathbf{e}_n^f(x))_n$ converges to a unique measure:

$$\mathbf{e}^f(x) := \lim_{n\to\infty} \mathbf{e}_n^f(x) \ ,$$

called the *empirical measure associated to* x. Furthermore, this measure is almost surely ergodic, by the ergodic decomposition theorem. We recall that a measure μ is *ergodic* if and only the *empirical function* $x \mapsto \mathbf{e}^f(x)$ is μ-a.e. constant. We denote by $\mathcal{M}_f^{\text{erg}}(X) \subset \mathcal{M}_f(X)$ the subset of ergodic probability measures.

It is natural to study how many ergodic statistical behaviors a dynamical system admits up to resolution ϵ (in a sense to be made precise). We are interested in the behavior of this number as ϵ tends to zero. This leads us to introduce the following notion:

Definition 0.1 (Topological Emergence)**.** Let X be a compact metric space, let f be a continuous self-map of X, and let d be a distance on the space of probabilities $\mathcal{M}(X)$ of X so that $(\mathcal{M}(X), \mathsf{d})$ is compact.

The *topological emergence* $\mathscr{E}_{\text{top}}(f)(\epsilon)$ of f is the function which associates to $\epsilon > 0$ the minimal number of ϵ-balls of $\mathcal{M}(X)$ whose union covers $\mathcal{M}_f^{\text{erg}}(X)$.

Of course this definition depends on how the space of measures is metrized. There are basically two classical types of distances on the space of probabilities $\mathcal{M}(X)$ which define the same weak topology (which is the most relevant one in ergodic theory): the Lévy–Prokhorov distance LP, and Wasserstein distances W_p, which depend on a parameter $p \in [1, \infty)$. We will recall their definitions in Section 1.2. For the rest of this introduction, we fix any distance $\mathsf{d} \in \{\mathsf{W}_p : 1 \leqslant p < \infty\} \cup \{\mathsf{LP}\}$.

In Section 2, we will give examples of open sets of mappings with essentially maximal topological emergence:

Theorem A. *Let f be $C^{1+\alpha}$-mapping of a manifold which admits a basic hyperbolic set K with box-counting dimension d. Assume that f is conformal expanding or that f is a conservative surface diffeomorphism. Then the topological emergence of $f|K$ is stretched exponential with exponent d:*

$$\lim_{\epsilon\to 0} \frac{\log\log \mathscr{E}_{\text{top}}(f|K)(\epsilon)}{-\log\epsilon} = d \ .$$

The emergence exponent is indeed maximal, since for any such a compact set K, the covering number of the space of probability measures $\mathcal{M}(K)$ is stretched exponential with exponent d, both for the Lévy–Prokhorov metric LP and the Wasserstein metrics W_p; see Section 1.3 for details.

The concept of topological emergence is linked to classical ideas of size of functional spaces developed by the Kolmogorov school (and emanating from Hilbert's 13th problem): see Section 1 for more details. Let us also note that the set $\mathcal{M}_f(X)$ has been investigated from several (topological, convex-analytic, ...) points of view for various classes of maps f, from older works [Sig, Dow] to recent ones [GoP, GeR, BBG, BoZ, DGMR, DGR]. The study of topological emergence expands this theme of research by imparting a more quantitative aspect to it.

We may be only interested in physically relevant statistics, and so we are allowed to disregard statistics that correspond to a set of orbits of zero Lebesgue measure. This led to the following concept [Be2], initially introduced for X a manifold and μ the Lebesgue measure:

Definition 0.2 (Metric Emergence)**.** Let (X, μ) be a compact metric space X endowed with a probability measure μ, and let f be a continuous self-map of X (not necessarily μ-preserving).

The *metric emergence* $\mathscr{E}_\mu(f)$ is the function that associates to $\epsilon > 0$ the minimal number $\mathscr{E}_\mu(f)(\epsilon) = N$ of probability measures $\mu_1, \dots, \mu_N$ so that:

$$\limsup_{n\to\infty} \int \min_{1\leqslant i\leqslant N} \mathsf{d}(\mathbf{e}^f_n(x), \mu_i)\, d\mu(x) \leqslant \epsilon \,. \tag{0.1}$$

Let us note that when μ is f-invariant, then $(\mathbf{e}^f_n(x))_n$ converges to $\mathbf{e}^f(x)$ μ-a.e. and so (0.1) can be replaced by:

$$\int \min_{1\leqslant i\leqslant N} \mathsf{d}(\mathbf{e}^f(x), \mu_i)\, d\mu(x) \leqslant \epsilon \,. \tag{0.2}$$

As we will explain in Section 3, when the measure μ is f-invariant, metric emergence becomes a particular case of the classic problem of quantization (or discretization) of a measure [GrL].

Let us recall some examples of metric emergence. By definition, if (f, μ) is ergodic then $\mathbf{e}^f(x) = \mu$ for μ-a.e. x and so its metric emergence $\mathscr{E}_\mu$ is identically 1 (i.e. minimal).

When X is a compact manifold M, the metric emergence will be canonically considered for $\mu = \mathrm{Leb}$, the Lebesgue measure of M (that is, the probability measure corresponding to a fixed normalized smooth positive volume form). The map f is called *conservative* if it leaves the Lebesgue measure invariant. The group of conservative C^r-diffeomorphisms is denoted by $\mathit{Diff}^{\,r}_{\mathrm{Leb}}(M)$.

There are well-studied subsets of $\mathit{Diff}^{\,r}_{\mathrm{Leb}}(M)$ consisting of ergodic diffeomorphisms: uniformly hyperbolic dynamics, quasi-periodic mappings (e.g. minimal translations of tori), and many classes of partially hyperbolic dynamics [BuW, ACW, Ob].

For a while, Boltzmann's ergodic hypothesis prevailed and typical Hamiltonian dynamical systems were believed to be ergodic [BiK, Dum]. However, KAM (Kolmogorov–Arnold–Moser) theory revealed that every perturbation of certain integrable systems displays infinitely many invariant tori filling a set of positive Lebesgue measure, in each of which the dynamics is an ergodic rotation. Thus the ergodic hypothesis was refuted. This phenomenon also showed that a typical symplectic diffeomorphism is in general not ergodic, since nearby its totally elliptic periodic points it is Hamiltonian and smoothly approximable by an integrable system. As we will explain later (see Corollary 5.7), the metric emergence of systems displaying KAM phenomenon is at least polynomial:

$$(\geqslant P) \qquad \liminf_{\epsilon\to 0} \frac{\log \mathscr{E}_{\mathrm{Leb}}(f)(\epsilon)}{-\log \epsilon} \geqslant 1\,.$$

Another phenomenon, discovered by Newhouse [Ne1], is the co-existence of infinitely many invariant open sets, each of which having an asymptotically constant empirical function, so that the corresponding probability measures can approximate any invariant ergodic measure supported on a certain non-trivial hyperbolic compact set. This phenomenon occurs generically in many categories of dynamical systems: see [Ne2, Dua, Buz, BoD, DNP, Bie]. The Newhouse phenomenon has been recently shown to be typical in the sense of Kolmogorov: see [Be1, Be2].

In view of Theorem A, one might believe that the metric emergence of systems displaying Newhouse phenomenon can have super-polynomial growth. In the paper [Be2], it is actually conjectured that super-polynomial growth is typical in open sets of many categories of dynamical systems. We prove one step toward this conjecture by showing (in Section 4) the existence of a smooth (that is, C^∞) conservative flow with stretched exponential metric emergence:

Theorem B. *There exists a smooth conservative flow $(\Phi^t)_t$ on the annulus $\mathbb{A} := \mathbb{R}/\mathbb{Z} \times [0,1]$ such that for every $t \neq 0$ the emergence of $f = \Phi^t$ is stretched exponential with (maximal) exponent $d = 2$:*

$$(S\exp^d) \qquad \lim_{\epsilon\to 0} \frac{\log\log \mathscr{E}_{\mathrm{Leb}}(f)(\epsilon)}{-\log \epsilon} = d\,.$$

Remark 0.3. We recall that surface flows have zero topological entropy. Thus the previous theorem provides an example of smooth conservative dynamics with high emergence but zero topological entropy.

Question 0.4. Is there a smooth conservative surface map f such that:

$$\liminf_{\epsilon\to 0} \epsilon^2 \log \mathscr{E}_{\mathrm{Leb}}(f)(\epsilon) > 0\,?$$

Actually the proof of this theorem can be adapted to show the existence of smooth conservative flow $(\Phi^t)_t$ of a compact manifold M of any dimension $d \geqslant 2$ such that the metric emergence of $f = \Phi_t$, $t \neq 0$ satisfies ($S\exp^d$).

We recall that a conservative map is C^r_{Leb}-*weakly stable* if every conservative mapping in a C^r neighborhood has only hyperbolic periodic points (i.e. points $x = f^p(x)$ for which the eigenvalues of $Df^p(x)$ have moduli different than 1). Such mappings are conjecturally uniformly hyperbolic (and so structurally stable): see [BeT], [Ma1, Conj. 2]. In that paper, it was shown that any conservative surface diffeomorphism which is not weakly stable can be C^∞-approximated by one with positive metric entropy. Here we obtain a stronger emergence counterpart of this result:

Theorem C. *A C^∞-generic, conservative, surface diffeomorphism f either is weakly stable or has a metric emergence with* lim sup *stretched exponential with exponent $d = 2$:*

$$(\overline{S}\exp^d) \qquad \limsup_{\epsilon \to 0} \frac{\log\log \mathscr{E}_{\mathrm{Leb}}(f)(\epsilon)}{-\log \epsilon} = d\,.$$

Remark 0.5. Theorem C certainly requires some appropriate degree of differentiability, and is completely false for the C^0 category – indeed, generic volume-preserving homeomorphisms of a compact manifold are ergodic [OxU] and therefore have minimal metric emergence.

With a relatively simple modification of the proof of the previous theorem, we also obtain its dissipative (i.e. non conservative) counterpart:

Theorem D. *For every $r \in [1, \infty]$ and for every surface M, there exists a non-empty open set $\mathcal{U} \subset \mathit{Diff}^r(M)$, such that a generic map $f \in \mathcal{U}$ has metric emergence $\mathscr{E}_{\mathrm{Leb}}(f)$ that satisfies ($\overline{S}\exp^d$) with $d = 2$.*

These results prove a weak version of Conjecture A of [Be2] for the classes of smooth conservative and non-conservative surface diffeomorphisms. This conjecture posits the existence of many open classes of dynamics for which super-polynomial emergence is typical in many senses (including Kolmogorov's). In this regard, let us note that it is an open question whether Newhouse phenomenon implies typically high emergence.

Our results also make it clear that emergence and entropy are completely unrelated. On one hand, a uniformly hyperbolic, conservative map has positive metric entropy but minimal metric emergence (identically equal to 1), since the volume measure is ergodic. Furthermore, a construction of Rees and Béguin–Crovisier–Le Roux [BCLR] yields a homeomorphism which is uniquely ergodic (and so has minimal topological and metric emergences) but has positive topological entropy. On the other hand, Theorem B gives an example of conservative dynamics with stretched exponential emergence but (as noted in Remark 0.3) with zero topological entropy and in particular (by the entropy variational principle) with zero metric entropy.

As we will show in Section 3, the metric emergence of any invariant measure is at most the topological emergence (see Proposition 3.14). Furthermore, we will prove that the latter upper bound is asymptotically attained, therefore obtaining the following statement that mirrors the entropy variational principle:

Theorem E (Variational Principle for Emergence)**.** *For every continuous self-map f of a compact metric space X, there exists an invariant probability measure μ such that:*

$$\begin{cases} \limsup\limits_{\epsilon\to 0} \dfrac{\log\log \mathscr{E}_\mu(f)(\epsilon)}{-\log\epsilon} = \limsup\limits_{\epsilon\to 0} \dfrac{\log\log \mathscr{E}_{\mathrm{top}}(f)(\epsilon)}{-\log\epsilon} \,, \\ \liminf\limits_{\epsilon\to 0} \dfrac{\log\log \mathscr{E}_\mu(f)(\epsilon)}{-\log\epsilon} = \liminf\limits_{\epsilon\to 0} \dfrac{\log\log \mathscr{E}_{\mathrm{top}}(f)(\epsilon)}{-\log\epsilon} \,. \end{cases}$$

Question 0.6. Can we find an invariant measure μ such that $\mathscr{E}_\mu(f) \sim \mathscr{E}_{\mathrm{top}}(f)$ (that is, these two functions of ϵ are asymptotic as $\epsilon \to 0$)?

Organization of the paper. In Section 1 we discuss covering numbers and the related concepts of box-counting dimension (the exponent when covering numbers obey a power law) and metric order (the exponent when covering numbers obey a stretched exponential law); furthermore, we define precisely the Wasserstein and Lévy–Prokhorov metrics on space $\mathcal{M}(X)$ of probability measures on a compact metric space X, and show that the metric order of $\mathcal{M}(X)$ coincides with the box-counting dimension of X (Theorem 1.3). Since topological emergence of a dynamical system is defined as the covering number function of its set of invariant probability measures, we obtain a simple upper bound for the growth rate of topological emergence of a dynamical system in terms of the dimension of the phase space. In Section 2 we exhibit classes of examples where this upper bound is attained: this is the content of Theorem A. The proof uses elementary properties of Gibbs measures.

In Section 3 we recall the notion of quantization of probability measures. We define quantization numbers, which express how efficiently a measure can be discretized, and are bounded from above by the covering numbers of the ambient space. We show that metric emergence of an invariant measure amounts to the quantization number function of its ergodic decomposition. We prove Theorem E, which says that one can always find an invariant probability measure with essentially maximal metric emergence; actually this is deduced from a more abstract result (Theorem 3.9) on the existence of measures with essentially maximal quantization numbers.

In Section 4 we construct an example of smooth conservative surface diffeomorphism such that Lebesgue measure has essentially maximal metric emergence; more precisely, we prove Theorem B. In Section 5 we prove our results on genericity of high emergence for conservative and dissipative surface diffeomorphisms, Theorems C and D; these proofs are relatively short because we make use of an intermediate result used to obtain Theorem B, namely Proposition 4.2. The proofs also use elementary versions of the KAM theorem and the persistence of normally contracted submanifolds.

In Appendix A we recall a characterization of metric entropy due to Katok [Ka], and show how it provides a characterization of metric entropy in terms of quantization numbers. Katok's theorem is used in Section 2; that is because the proof of Theorem A uses the measure of maximal dimension as an auxiliary device in the construction of large sufficiently separated sets of periodic orbits.

Notation. We employ the usual notations:

- $f \sim g$ means $f/g \to 1$;
- $f \asymp g$ means $f = O(g)$ and $g = O(f)$.

Acknowledgements. We are grateful to Rémi Peyre for meticulous discussions on the covering numbers of spaces of measures, and for allowing us to include the resulting proof of Theorem 1.3 in this paper. We also thank Viviane Baladi, Abed Bounemoura, Sylvain Crovisier, Bassam Fayad, Godofredo Iommi, François Ledrappier, Enrique Pujals, and Michael Shub for insightful comments. Finally, we thank the referee for corrections and suggestions.

1. Metric orders and spaces of measures

1.1. **Dimension and metric order of a compact metric space.** Let X be a totally bounded space, and let $\epsilon > 0$. A subset $F \subset X$ is called:

- ϵ-*dense* if X is covered by the closed balls of radius ϵ and centers in F;
- ϵ-*separated* if the distance between any two distinct points of F is greater than ϵ.

Then we define the following numbers (which are finite by total boundedness):

- the *covering number* $D_X(\epsilon) = D(X, \epsilon)$ is the minimum cardinality of an ϵ-dense set;
- the *packing number* $S_X(\epsilon) = S(X, \epsilon)$ is the maximum cardinality of an ϵ-separated set;

Precise computation of these numbers is seldom possible (see the classic [Rog] for problems of this nature). However we are only interested in the asymptotics of these numbers as ϵ tends to 0, and so moderately fine estimates will suffice.

Covering and packing numbers can be compared as follows:

$$S_X(2\epsilon) \leqslant D_X(\epsilon) \leqslant S_X(\epsilon)\,; \tag{1.1}$$

indeed the first inequality follows from the observation that a 2ϵ-separated set of cardinality n cannot be covered by less than n closed balls of radius ϵ, while the second inequality follows from the fact that every maximal ϵ-separated set is ϵ-dense.

The *upper box-counting dimension* of X is defined as:

$$\overline{\dim}(X) := \limsup_{\epsilon \to 0} \frac{\log D_X(\epsilon)}{-\log \epsilon} \in [0, \infty]\,.$$

Note that by inequalities (1.1), it makes no difference if $D_X(\epsilon)$ is replaced by $S_X(\epsilon)$ in the definition above. We define the *lower box-counting dimension* $\underline{\dim}(X)$ by taking $\liminf$ instead of $\limsup$. If these two quantities coincide, they are called the *box-counting dimension* of X and denoted by $\dim X$. The term *Minkowski dimension* is also used. For an elementary introduction and more information, see [Fa].

The dimensions defined above are infinite when the numbers $D_X(\epsilon)$ and $S_X(\epsilon)$ are super-polynomial with respect to ϵ^{-1}. However, these functions are often comparable to stretched exponentials; indeed many examples of functional spaces with this property are studied in the classic work by Kolmogorov and Tihomirov [KoT][1]. The corresponding exponent

$$\mathrm{mo}(X) := \lim_{\epsilon \to 0} \frac{\log \log D_X(\epsilon)}{-\log \epsilon} = \lim_{\epsilon \to 0} \frac{\log \log S_X(\epsilon)}{-\log \epsilon}\,,$$

if it exists, is called the *metric order* of X, following [KoT, p. 298]. In general we define *lower and upper metric orders* $\underline{\mathrm{mo}}(X) \leqslant \overline{\mathrm{mo}}(X)$ by taking $\liminf$ and $\limsup$.

Remark 1.1. A concept similar to metric order, called *critical parameter for the power-exponential scale*, was introduced and studied by Kloeckner [Kl1, Kl2]. Its definition is more akin to the Hausdorff dimension.

Remark 1.2. If Y is a subset of X then define the *relative covering number* $D_X(Y, \epsilon)$ as the minimal number of closed ϵ-balls in X whose union covers Y. Note that:

$$D_X(Y, \epsilon) \leqslant D(Y, \epsilon) \leqslant D_X(Y, \epsilon/2)\,.$$

Therefore dimension and metric order of subsets of X can be also computed using relative covering numbers.

1.2. **Spaces of measures.** Let (X, d) be a compact metric space. Let $\mathcal{M}(X)$ be the space of Borel probability measures on X, endowed with the weak topology and therefore compact. There are many different ways of metrizing the weak topology. We will consider two types of metrics in $\mathcal{M}(X)$: the Wasserstein distances and the Lévy–Prokhorov distance (defined below). These metrics respect the original metric on X, in the sense that the map $x \mapsto \delta_x$ (where δ_x is the Dirac probability measure concentrated at the point x) becomes an *isometric* embedding of X into $\mathcal{M}(X)$.

Given two measures μ, $\nu \in \mathcal{M}(X)$, a *transport plan* (or *coupling*) from μ to ν is a probability measure π on the product $X \times X$ such that $(p_1)_* \pi = \mu$ and $(p_2)_* \pi = \nu$, where p_1, $p_2 \colon X \times X \to X$ are the canonical projections. (We say that μ and ν are the *marginals* of π.) Such transport plans form

[1] See also [CLN, § 8.2.6] for historical context.

a closed and therefore compact subset $\Pi(\mu,\nu)$ of $\mathcal{M}(X\times X)$. For any real number $p \geqslant 1$, the *p-Wasserstein distance* between μ and ν is defined as:

$$\mathsf{W}_p(\mu,\nu) := \inf_{\pi\in\Pi(\mu,\nu)} \left(\int_{X\times X} [\mathsf{d}(x,y)]^p d\pi(x,y) \right)^{1/p},$$

(The integral in this formula is called the *cost* of the transport plan π with respect to the *cost function* d^p. The infimum is always attained, i.e., an *optimal* transport plan always exists.) It can be shown that W_p is a metric on $\mathcal{M}(X)$ which induces the weak topology: see e.g. [Vil, Theorems 7.3 and 7.12].

The *Lévy–Prokhorov distance* between two measures μ, $\nu \in \mathcal{M}(X)$ is denoted $\mathsf{LP}(\mu,\nu)$ and is defined as the infimum of $\epsilon > 0$ such that for every Borel set $E \subset X$, if $V_\epsilon(E)$ denotes the ϵ-neighborhood of E, then:

$$\nu(E) \leqslant \mu(V_\epsilon(E)) + \epsilon \quad \text{and} \quad \mu(E) \leqslant \nu(V_\epsilon(E)) + \epsilon\,.$$

For a proof that LP is a metric on $\mathcal{M}(X)$ and that induces the weak topology, see [Bil, p. 72].

The Lévy–Prokhorov distance can also be characterized in terms of transport plans: it equals the infimum of $\epsilon > 0$ such that for some $\pi \in \Pi(\mu,\nu)$, the set $\{(x,y) \in X \times X \ ; \ \mathsf{d}(x,y) > \epsilon\}$ has π-measure less than ϵ; this is Strassen's theorem: see [Bil, p. 74] or [Vil, p. 44].

The family of Wasserstein metrics are not Lipschitz-equivalent to one another nor to the Lévy–Prokhorov metric. On the other hand, the following Hölder comparisons hold:

$$\mathsf{W}_q \leqslant \mathsf{W}_p \leqslant (\operatorname{diam} X)^{1-\frac{q}{p}} \mathsf{W}_q^{\frac{q}{p}} \qquad \text{if } 1 \leqslant q \leqslant p; \tag{1.2}$$

$$\mathsf{LP}^{1+\frac{1}{p}} \leqslant \mathsf{W}_p \leqslant (1 + (\operatorname{diam} X)^p)^{\frac{1}{p}} \mathsf{LP}^{\frac{1}{p}}\,; \tag{1.3}$$

see [Vil, p. 210], [GiS, Theorem 2].

1.3. **Metric order of spaces of measures.** The following result relates the lower and upper metric orders of Wasserstein space with the lower and upper box-counting dimensions of the underlying metric space:

Theorem 1.3. *For any compact metric space X and any $p \geqslant 1$, we have:*

$$\underline{\dim}(X) \leqslant \underline{\mathrm{mo}}(\mathcal{M}(X),\mathsf{W}_p) \leqslant \overline{\mathrm{mo}}(\mathcal{M}(X),\mathsf{W}_p) \leqslant \overline{\dim}(X)\,.$$

In particular the metric order $\mathrm{mo}(\mathcal{M}(X),\mathsf{W}_p)$ *exists and equals the box-counting dimension* $\dim X$ *whenever the latter exists.*

Actually, the rightmost inequality in the theorem is a consequence of a more precise result of Bolley–Guillin–Villani [BGV] (details will be provided below), while a variation of the leftmost inequality was obtained by Kloeckner [Kl2, Theorem 1.3]. Here we will present a proof of the leftmost inequality which was obtained jointly with Rémi Peyre.

Remark 1.4. The exact same statement also holds for the Lévy–Prokhorov metric, as a consequence of [KuZ, Lemmas 1 and A1].

Remark 1.5. Other examples where the metric order of a functional space equals the dimension of the underlying space can be found in [KoT], namely uniformly bounded uniformly Lipschitz functions on an interval [KoT, p. 288], or on more general sets [KoT, p. 307].

Proof of the rightmost inequality in Theorem 1.3. By [BGV, Theorem A.1][2], there exists $C > 0$ such that:

$$D\big((\mathcal{M}(X), \mathsf{W}_p), \epsilon\big) \leqslant (C/\epsilon)^{pD_X(\epsilon/2)} .$$

So:

$$\frac{\log \log D\big((\mathcal{M}(X), \mathsf{W}_p), \epsilon\big)}{\log \epsilon^{-1}} \leqslant \frac{\log(\log C + \log \epsilon^{-1})}{\log \epsilon^{-1}} + \frac{\log 2 + \log \epsilon^{-1}}{\log \epsilon^{-1}} \, \frac{\log D_X(\epsilon/2) + \log p}{\log(2/\epsilon)} .$$

Taking $\limsup$ as $\epsilon \to 0$ we obtain $\overline{\mathrm{mo}}(\mathcal{M}(X), \mathsf{W}_p) \leqslant \overline{\dim}(X)$. □

The remaining part of Theorem 1.3 will be obtained as a consequence of a more general result that allows us to estimate the lower metric order of other spaces of measures.

Let us say that two probability measures μ, ν on X are *ϵ-apart* if their supports are ϵ-apart in the following sense:

$$\min\{\mathsf{d}(x, y) \mid x \in \operatorname{supp} \mu,\ y \in \operatorname{supp} \nu\} \geqslant \epsilon .$$

Theorem 1.6. *Let X be a compact metric space. Let $\mathcal{C}$ be a convex subset of $\mathcal{M}(X)$. For each $\epsilon > 0$, let $A(\mathcal{C}, \epsilon)$ denote the maximal number of pairwise ϵ-apart measures in $\mathcal{C}$. Then, for any $p \geqslant 1$,*

$$\underline{\mathrm{mo}}(\mathcal{C}, \mathsf{W}_p) \geqslant \liminf_{\epsilon \to 0} \frac{\log A(\mathcal{C}, \epsilon)}{-\log \epsilon} .$$

The same inequality holds for the distance LP.

Proof of the leftmost inequality in Theorem 1.3. We apply Theorem 1.6 with $\mathcal{C} = \mathcal{M}(X)$. If $\{x_1, \dots, x_N\}$ is an ϵ-separated subset of X then the Dirac measures δ_{x_1}, …, δ_{x_N} are pairwise ϵ-apart. This observation shows that $A(\mathcal{M}(X), \epsilon) \geqslant S(X, \epsilon)$. The result follows. □

To prove Theorem 1.6, we will need the following elementary large-deviations estimate (see e.g. [GrS, p. 32] for a proof):

Lemma 1.7 (Bernstein inequality)**.** *Let H_n (a random variable) be the number of heads on n tosses of a fair coin. Then for any $\delta > 0$,*

$$\operatorname{Prob}\left[\frac{H_n}{n} \leqslant \frac{1}{2} - \delta\right] \leqslant e^{-\frac{\pi}{4}\delta^2 n} .$$

[2] See [Ng, Lemma 4] for a related result.

Proof of Theorem 1.6 (with Rémi Peyre). Fix $\epsilon > 0$, and let $N := 8\lfloor A(\mathcal{C}, \epsilon)/8 \rfloor$. Observe that $A(\mathcal{C}, \epsilon) - 7 \leqslant N \leqslant A(\mathcal{C}, \epsilon)$, and so we can find measures $\nu_1, \dots, \nu_N \in \mathcal{C}$ that are pairwise ϵ-separated.

Denote

$$F := \left\{ f \colon \{1, \dots, N\} \to \{0,1\} \;\middle|\; \sum_{i=1}^{N} f(y) = \frac{N}{2} \right\}. \tag{1.4}$$

We endow F with the Hamming distance:

$$\mathsf{Hamm}(f,g) := \#\{i \in \{1, \dots, N\} |\; f(i) \neq g(i)\}$$

(which is always an even number between 0 and N). Let us estimate the cardinality of a ball B of radius $N/4$ in F and centered at some f. If g is an element of B, that is, $k := \frac{1}{2}\mathsf{Hamm}(f,g) \leqslant N/8$, then there are exactly k elements of $f^{-1}(\{0\})$ and k elements of $f^{-1}(\{1\})$ at which g differs from f. As both sets $f^{-1}(\{0\})$ and $f^{-1}(\{1\})$ have cardinality $N/2$, we obtain:

$$\# B = \sum_{k=0}^{N/8} \binom{N/2}{k}^2 \leqslant \left[\sum_{k=0}^{N/8} \binom{N/2}{k} \right]^2 .$$

The quantity between square brackets equals $2^{N/2}$ times the probability of obtaining at most $N/8$ heads on $N/2$ tosses of a fair coin. By Lemma 1.7, this probability is at most $e^{-\frac{\pi}{4}\left(\frac{1}{4}\right)^2 \frac{N}{2}}$. So

$$\# B \leqslant 2^N e^{-\frac{\pi}{4} \cdot \frac{N}{4^2}} . \tag{1.5}$$

Choose a maximal $N/4$-separated subset F' of F. Then F' is $N/4$-dense, that is, the balls of radius $N/4$ with centers in F' form a covering of F. The cardinality of F itself is $\binom{N}{N/2} \geqslant (2N)^{-1/2} 2^N$ (by Stirling's formula). Using (1.5), we conclude that

$$\# F' \geqslant \frac{\# F}{\# B} \geqslant (2N)^{-1/2} e^{\pi N/4^3} . \tag{1.6}$$

Now, for each $f \in F'$, consider the measure:

$$\mu_f := \frac{2}{N} \sum_{i=1}^{N} f(i) \nu_i ,$$

which by convexity belongs to $\mathcal{C}$. Consider the subset $\mathcal{F} := \{\mu_f \mid f \in F'\}$ of $\mathcal{C}$, which has the same cardinality as F'. This set has the following property, whose proof will be given later:

Claim 1.8. The set $\mathcal{F}$ is $4^{-1/p}\epsilon$-separated with respect to the Wasserstein distance W_p.

In particular, $S((\mathcal{C}, \mathsf{W}_p), 4^{-1/p}\epsilon) \geqslant \# F'$. On the other hand, it follows from (1.6) that $\# F' \geqslant e^{cN}$ for all sufficiently large N, where $c > 0$ is a constant. So:

$$\frac{\log\log S((\mathcal{C},\mathsf{W}_p),4^{-1/p}\epsilon)}{-\log(4^{-1/p}\epsilon)} \geqslant \frac{\log N+\log c}{-\log\epsilon+\frac{1}{p}\log 4}$$

Since $N \geqslant A(\mathcal{C},\epsilon)-7$, taking $\liminf$ as $\epsilon\to 0$ we obtain the conclusion of the theorem for the Wasserstein distance W_p.

As regards the Lévy–Prokhorov distance LP, inequalities (1.3) allows us to compare it with the W_1 distance, and so Claim 1.8 implies that $\mathcal{F}$ is $(4(1+\operatorname{diam} X))^{-1}\epsilon$-separated with respect to LP, which allows us to conclude as before.

This completes the proof of Theorem 1.3, modulo the claim. □

Proof of Claim 1.8. Fix two distinct elements f, g of F', and let us estimate $\mathsf{W}_p(\mu_f,\mu_g)$. Let S_f and S_g be the supports of μ_f and μ_g, respectively.

We claim that:

$$(x,y)\in(S_f\smallsetminus S_g)\times S_g \;\Rightarrow\; \mathsf{d}(x,y)\geqslant\epsilon\,.$$

Indeed, if $y\in S_g$ then $y\in\operatorname{supp}\nu_j$ for some $j\in\{1,\dots,N\}$ such that $g(j)=1$, while if $x\in S_f\smallsetminus S_g$ then $x\in\operatorname{supp}\nu_i$ for some $i\in\{1,\dots,N\}$ such that $f(i)=1$ and $g(i)=0$; in particular, $i\neq j$. So ν_i and ν_j are ϵ-apart, which guarantees that $\mathsf{d}(x,y)\geqslant\epsilon$, as claimed.

Also note that:

$$\mu_f(S_f\smallsetminus S_g)=\frac{2}{N}\,\#\big\{i\in\{1,\dots,N\}\mid f(i)=1,g(i)=0\big\}=\frac{\mathsf{Hamm}(f,g)}{N}\geqslant\frac{1}{4}\,,$$

since F is $N/4$-separated.

For any transport plan π from μ_f to μ_g, using the remarks above we can estimate:

$$\int_{X\times X}[\mathsf{d}(x,y)]^p d\pi(x,y)=\int_{S_f\times S_g}(\cdots)\geqslant\int_{(S_f\smallsetminus S_g)\times S_g}(\cdots)$$
$$\geqslant\epsilon^p\pi\big((S_f\smallsetminus S_g)\times S_g\big)=\epsilon^p\mu_f(S_f\smallsetminus S_g)\geqslant\frac{\epsilon^p}{4}\,.$$

So, by definition of the Wasserstein distance, we have $\mathsf{W}_p(\mu_f,\mu_g)\geqslant\frac{\epsilon}{4^{1/p}}$, completing the proof of the claim. □

2. Examples of dynamics with high topological emergence

Let f be a continuous self-map of a compact metric space X. We recall that $\mathcal{M}_f^{\mathrm{erg}}(X)$ denotes the space of invariant ergodic probability measures.

As explained in the introduction, the *topological emergence* of f is the relative covering number of $\mathcal{M}_f^{\mathrm{erg}}(X)$ (defined in Remark 1.2) endowed either with a Wasserstein distance W_p, $1\leqslant p<\infty$, or the Lévy-Prokhorov distance LP, that is:

$$\mathscr{E}_{\mathrm{top}}(f)(\epsilon):=D_{\mathcal{M}(X)}(\mathcal{M}_f^{\mathrm{erg}},\epsilon)\,. \tag{2.1}$$

We are concerned with the asymptotic behavior of this function for small ϵ. Since $\mathcal{M}_f^{\mathrm{erg}}(X)$ is included in $\mathcal{M}(X)$, by Theorem 1.3 and Remark 1.4 we have:

$$\limsup_{\epsilon\to 0} \frac{\log\log \mathscr{E}_{\mathrm{top}}(f)(\epsilon)}{-\log\epsilon} = \overline{\mathrm{mo}}(\mathcal{M}_f^{\mathrm{erg}}(X)) \leqslant \overline{\mathrm{mo}}(\mathcal{M}(X)) \leqslant \overline{\dim}(X). \tag{2.2}$$

Sometimes, this bound is far from being optimal. For instance, when f is uniquely ergodic, then $\mathscr{E}_{\mathrm{top}}(f)(\epsilon) = 1$ does not grow at all. If f is the identity of X, then $\mathcal{M}_f^{\mathrm{erg}}(X)$ is isometric to X, and so $\mathscr{E}_{\mathrm{top}}(f)(\epsilon)$ is comparable to ϵ^{-d} if X has well defined box-counting dimension d.

On the other hand, Theorem A gives examples of hyperbolic compact sets for which the above bound is optimal. Let us explain and prove them.

2.1. **Conformal expanding repellers.** Let M be a Riemannian manifold, U an open subset of M and $f\colon U \to M$ be $C^{1+\alpha}$ map which leaves invariant a compact subset K of U (i.e. $f^{-1}(K) = K$). We say that (K, f) is a *conformal expanding repeller* if f is conformal and expanding at K: for each $x \in K$, the derivative $Df(x)$ expands the Riemannian metric by a scalar factor greater than 1. Then its box-counting dimension $\dim(K)$ is well-defined, and it equals the Hausdorff dimension: see [PrU, Corol. 9.1.7].

Theorem 2.1. *Let (K, f) be a conformal expanding repeller of dimension d. Then the topological emergence of $f|K$ is stretched exponential with exponent d:*

$$\lim_{\epsilon\to 0} \frac{\log\log \mathscr{E}_{\mathrm{top}}(f|K)(\epsilon)}{-\log\epsilon} = d\,.$$

Proof. First, we can assume that K is transitive since it is always a finite disjoint union of transitive sets; moreover, up to taking an iterate of f, we can suppose that $f|K$ is topologically mixing – see [PrU, Thm. 3.3.8].

By standard results [PrU, §9.1], there exists an invariant ergodic probability measure μ supported on K of *maximal dimension*. The Lyapunov exponent $\chi_\mu := \int \log \|Df\|\, d\mu$ and metric entropy h_μ are related as follows:

$$\chi_\mu \cdot d = h_\mu\,, \quad \text{where } d = \dim K\,. \tag{2.3}$$

Let $\rho_0 > 0$ be such that U contains the ρ_0-neighborhood of K. Reducing ρ_0 if necessary, there exists $\lambda > 1$ such that f is λ-expanding on the ρ_0-neighborhood of K, in the sense that $\|Df^{-1}\|^{-1} \geqslant \lambda$. Then we have the following property [PrU, §4.1]: for all $x \in K$ and all $n \geqslant 1$, the connected component V_x^n of x in the preimage by f^n of the (Riemannian) ball $B(f^n(x), \rho_0)$ is included in $B(x, \lambda^{-n}\rho_0)$. Moreover V_x^n is sent by f^n diffeomorphically onto $B(f^n(x), \rho_0)$. Note that ρ_0 is an expansiveness constant for $f|K$, in the sense that if $x \neq y$ then there exists $n \geqslant 0$ such that $\mathsf{d}(f^n(x), f^n(y)) \geqslant \rho_0$.

Let d be the metric on M induced by the Riemannian structure, and for each $n \geqslant 1$, let d_n denote the time-n Bowen metric on K, defined by:

$$\mathsf{d}_n(x, y) := \max_{0 \leqslant i < n} \mathsf{d}(f^i(x), f^i(y))\,.$$

By the bounded distortion property [PrU, Lemma 4.4.2], there exists a constant $C_0 > 1$ such that for any $n \geqslant 0$, if a pair of points $(x, y) \in K \times U$ satisfies $\mathsf{d}_n(x, y) < \rho_0$ then $\|Df^n(y)\| \leqslant C_0 \|Df^n(x)\|$.

Reducing ρ_0 if necessary, we assume that every pair of points $(x, y) \in K \times U$ such that $\mathsf{d}(x, y) < \rho_0$ can be joined by a unique geodesic segment of minimal length, denoted $[x, y]$.

Claim 2.2. If $n \geqslant 1$ and $(x, y) \in K \times U$ are such that $\mathsf{d}_{n+1}(x, y) < \rho_1 := C_0^{-2}\rho_0$ then:

$$\frac{\mathsf{d}(f^n(x), f^n(y))}{\mathsf{d}(x, y)} \leqslant C_0 \|Df^n(x)\| \,.$$

Proof. Fix $x \in K$ and $n \geqslant 1$. As explained above, f^n maps V_x^n diffeomorphically onto $B(f^n(x), \rho_0)$; let $f_x^{-n} := (f^n|V_x^n)^{-1}$ be its inverse. Note that V_x^n is exactly the d_{n+1}-ball of center x and radius ρ_0. Now consider $y \in U$ such that $\mathsf{d}_{n+1}(x, y) < \rho_1 := C_0^{-2}\rho_0$. We have $\mathsf{d}(f^n(x), f^n(y)) < \rho_1$, by definition of the Bowen metric. Consider the geodesic segment $S := [f^n(x), f^n(y)]$. Since S is contained in $B(f^n(x), \rho_1) \subset B(f^n(x), \rho_0)$, the curve $f_x^{-n}(S)$ is well-defined and is contained in V_x^n. Since this curve joins x and y, we have:

$$\begin{aligned}
\mathsf{d}(x, y) \leqslant \operatorname{len}(f_x^{-n}(S)) &\leqslant C_0 \|Df^n(x)\|^{-1} \operatorname{len}(S) && (2.4)\\
&< C_0 \|Df^n(x)\|^{-1} \rho_1 \\
&\leqslant C_0^{-1} \|Df^n(x)\|^{-1} \rho_0 \,, && (2.5)
\end{aligned}$$

where the estimate (2.4) follows from the bounded distortion property and conformality of the derivatives, and (2.5) follows from the definition of ρ_1.

We claim that the geodesic segment $[x, y]$ is contained in the interior of V_x^n. Indeed, if that is not the case, there exists a subsegment $[x, z] \subset V_x^n$ such that $z \in \partial V_x^n$. On one hand, $f^n(z) \in f^n(\partial V_x^n) \subset \partial B(f^n(x), \rho_0)$; on the other hand, using bounded distortion again,

$$\begin{aligned}
\mathsf{d}(f^n(x), f^n(z)) \leqslant \operatorname{len}(f^n([x, z])) &\leqslant C_0 \|Df^n(x)\| \mathsf{d}(x, z) && (2.6)\\
&\leqslant C_0 \|Df^n(x)\| \mathsf{d}(x, y) \\
&< \rho_0 \qquad \text{(by (2.5))},
\end{aligned}$$

a contradiction. This confirms that $[x, y]$ is contained in the interior of V_x^n.

We are now allowed to apply estimate (2.6) with $z = y$ and therefore conclude the validity of Claim 2.2. □

Fix a small $\delta > 0$. By Katok's Theorem A.2 (see the appendix), there exists a positive number $\rho < \rho_1$ such that for all sufficiently large n, the least number $N_\mu(n, \rho, 1/2)$ of balls of radii ρ in the d_n metric necessary to cover a set of μ-measure $\geqslant 1/2$ satisfies:

$$N_\mu(n, \rho, 1/2) > e^{(h_\mu - \delta)n} \,.$$

For each $n \geqslant 1$, let B_n be the set of points $x \in K$ such that $\|Df^n(x)\| \leqslant e^{(\chi_\mu + \delta)n}$. By Birkhoff theorem, if n is large enough then $\mu(B_n) > 1/2$. Take a (d_n, ρ)-separated set $F_n \subset B_n$ of maximal cardinality. Then the balls of radii ρ and centered at points in F_n cover B_n. Therefore:

$$\# F_n \geqslant N_\mu(n, \rho, 1/2) > e^{(h_\mu - \delta)n}, \tag{2.7}$$

provided n is large enough.

By the specification property of topologically mixing repellers (see e.g. [ViO, Prop. 11.3.1]), there exists an integer $n_0 \geqslant 0$ (depending on ρ) such that for every n, each point $x \in F_n$ is shadowed by an $(n+n_0)$-periodic point $y \in K$ in such a way that $\mathsf{d}_n(x, y) < \rho/2$. Let G_n be the set of periodic points y obtained in this way. Note that G_n has the same cardinality as F_n. Also note that, by bounded distortion, $\|Df^n(y)\| \leqslant C_0 \|Df^n(x)\| \leqslant C_0 e^{(\chi_\mu + \delta)n}$ and so, if n is large enough,

$$\|Df^{n+n_0}(y)\| \leqslant e^{(\chi_\mu + 2\delta)n}. \tag{2.8}$$

Let $\Pi_n := \bigcup_{k \geqslant 0} f^k(G_n)$ be the union of the orbits of the points in G_n. By periodicity, the points $y \in \Pi_n$ satisfy the same estimate (2.8).

Claim 2.3. The set Π_n is (d, ϵ_n)-separated with $\epsilon_n := e^{-(\chi_\mu + 3\delta)(n+1)}$, provided n is large enough.

Proof. Take a pair of distinct points $y, z \in \Pi_n$, and let us prove that $\mathsf{d}(y, z) > \epsilon_n$. Both points are fixed by f^{n+n_0}, so, by expansiveness, there exists k in the range $0 \leqslant k < n + n_0$ such that $\mathsf{d}(f^k(y), f^k(z)) \geqslant \rho_1$ since $\rho_1 \leqslant \rho_0$. Assume that k is minimal. If $k = 0$ then the desired estimate is trivial, so consider $k > 0$. Then $\mathsf{d}_k(y, z) < \rho_1$ and the following estimates hold:

$$\begin{aligned}
\mathsf{d}(y, z) &\geqslant C_0^{-1} \mathsf{d}(f^{k-1}(y), f^{k-1}(z)) \|Df^{k-1}(y)\|^{-1} && \text{(by Claim 2.2)} \\
&\geqslant C_1^{-1} \rho_1 \|Df^{k-1}(y)\|^{-1} && \text{with } C_1 := C_0 \cdot \|Df\| \\
&\geqslant C_1^{-1} \rho_1 \|Df^{n+n_0}(y)\|^{-1} && \text{(since } f \text{ is expanding)} \\
&> C_1^{-1} \rho_1 \cdot e^{-(\chi_\mu + 2\delta)n} && \text{(by (2.8))}.
\end{aligned}$$

This implies the sough inequality when n is large enough. ☐

So any two distinct ergodic measures supported in the finite invariant set Π_n are ϵ_n-apart (in the sense defined in Section 1.3). The number A_n of such ergodic measures satisfies:

$$A_n \geqslant \frac{\# G_n}{n + n_0} = \frac{\# F_n}{n + n_0} \geqslant e^{(h_\mu - 2\delta)n}$$

if n is sufficiently large (by (2.7)).

Now, given $\epsilon > 0$ sufficiently small, take n such that $\epsilon_n \leqslant \epsilon < \epsilon_{n-1}$. Consider the convex set $\mathcal{C} := \mathcal{M}_f(K)$ of all f-invariant measures; then, in the notation of Theorem 1.6, we have $A(\mathcal{C}, \epsilon) \geqslant A(\mathcal{C}, \epsilon_n) \geqslant A_n$ and so

$$\frac{\log A(\mathcal{C}, \epsilon)}{-\log \epsilon} \geqslant \frac{\log A_n}{-\log \epsilon_{n-1}} \geqslant \frac{h_\mu - 2\delta}{\chi_\mu + 3\delta}.$$

So Theorem 1.6 yields $\underline{\mathrm{mo}}(\mathcal{M}_f(K)) \geqslant (h_\mu - 2\delta)/(\chi_\mu + 3\delta)$. As δ is arbitrarily close to 0, we conclude that $\underline{\mathrm{mo}}(\mathcal{M}_f(K))$ is at least h_μ/χ_μ, which by (2.3) equals $d = \dim K$.

As a consequence of specification (see [ViO, Thrm. 11.3.4]), the closure of $\mathcal{M}_f^{\mathrm{erg}}(K)$ equals $\mathcal{M}_f(K)$. Therefore:

$$\underline{\mathrm{mo}}(\mathcal{M}_f^{\mathrm{erg}}(K)) = \underline{\mathrm{mo}}(\mathcal{M}_f(K)) \geqslant d\,.$$

On the other hand, $\overline{\mathrm{mo}}(\mathcal{M}_f^{\mathrm{erg}}(K)) \leqslant d$ by (2.2). So $\mathrm{mo}(\mathcal{M}_f^{\mathrm{erg}}(K)) = d$, as we wanted to show. □

2.2. **Hyperbolic sets of conservative surface diffeomorphisms.** Let M be a surface and let $f\colon M \to M$ be a $C^{1+\alpha}$ diffeomorphism. Let $K \subset M$ be a *hyperbolic* set for f. This means that K is an invariant compact set K and there exists an invariant splitting $E^{\mathrm{s}} \oplus E^{\mathrm{u}}$ of the tangent bundle TM of M restricted to K such that the line bundles E^{s} and E^{u} are respectively contracted and expanded. In other words, there exists $\lambda > 1$ such that for every $z \in K$:

$$\begin{cases} D_z f(E^{\mathrm{s}}_z) = E^{\mathrm{s}}_{f(z)} & \& \quad \|D_z f|E^{\mathrm{s}}\|^{-1} > \lambda\,, \\ Df_z(E^{\mathrm{u}}) = E^{\mathrm{u}}_{f(z)} & \& \quad \|D_z f|E^{\mathrm{u}}\| > \lambda\,. \end{cases}$$

Let us assume moreover that the compact set K is *locally maximal*, that is, it admits a neighborhood U such that $K = \bigcap_{n\in\mathbb{Z}} f^n(U)$.

Theorem 2.4. *If f is conservative then the topological emergence of $f|K$ is stretched exponential with exponent $d := \dim(K)$:*

$$\lim_{\epsilon\to 0} \frac{\log\log \mathscr{E}_{\mathrm{top}}(f|K)(\epsilon)}{-\log\epsilon} = d\,.$$

Proof. First, we can assume that K is transitive since it is always a finite disjoint union of such sets; moreover, up to taking an iterate of f, we can consider that $f|K$ is topologically mixing – see [KaH, Thm. 18.3.1, p. 574].

From standard results on dimension theory of hyperbolic sets (see e.g. [Pe, Thrm. 22.2]), the box-counting dimension $d := \dim K$ is well defined, and it equals $d^{\mathrm{s}} + d^{\mathrm{u}}$, where d^{s} (resp. d^{u}) is the box-counting dimension of K intersected with any local stable (resp. unstable) manifold. Moreover, for every $\star \in \{\mathrm{u}, \mathrm{s}\}$, there exists an invariant ergodic probability measure $\mu^\star$ supported on K of *maximal $\star$-dimension.* The Lyapunov exponent $\chi_{\mu^\star} := \int \log\|Df|E^\star\|\, d\mu$ and the metric entropy $h_{\mu^\star}$ are related as follows:

$$\chi_{\mu^\star}\cdot d^\star = h_{\mu^\star}\,. \tag{2.9}$$

Those measures are obtained as the unique equilibrium states for the functions:

$$\varphi^{\mathrm{s}}(x) := -\log\|Df(x)|E^{\mathrm{s}}_x\|\,, \qquad \varphi^{\mathrm{u}}(x) := \log\|Df(x)|E^{\mathrm{u}}_x\|\,.$$

The dynamics being consevative, the functions φ^{s} and φ^{u} are cohomologous. Thus by uniqueness of equilibria:

$$\mu^{\mathrm{s}} = \mu^{\mathrm{u}} =: \mu \quad \text{and} \quad -\chi_{\mu^{\mathrm{s}}} = \chi_{\mu^{\mathrm{u}}} =: \chi_\mu \,,$$

and so by (2.9) and using $d = d^{\mathrm{u}} + d^{\mathrm{s}}$:

$$\chi_\mu \cdot \frac{d}{2} = h_\mu \,. \tag{2.10}$$

Let us fix continuous families of local stable and unstable manifolds $(W^{\mathrm{s}}_{\mathrm{loc}}(x))_{x\in K}$ and $(W^{\mathrm{u}}_{\mathrm{loc}}(x))_{x\in K}$, small enough to be λ^{-1}-contracted by respectively f and f^{-1}. Furthermore, whenever x and $y \in K$ are close enough, then $W^{\mathrm{u}}_{\mathrm{loc}}(x)$ intersects $W^{\mathrm{s}}_{\mathrm{loc}}(y)$ at a unique point, called the *bracket* of x and y and denoted $[x, y]$. By local maximality of K, the point $[x, y]$ belongs to K.

Let $\tilde{\mathsf{d}}_n$ denote the bilateral Bowen metric on M, defined by:

$$\tilde{\mathsf{d}}_n(x, y) := \max_{-n<i<n} \mathsf{d}(f^i(x), f^i(y)) \,.$$

We denote by d^{u} (resp. d^{s}) the distance along the local unstable (resp. stable) manifolds. Using the contraction along the local stable and unstable manifolds by f and f^{-1}, we obtain:

Claim 2.5. There exists $\rho_0 > 0$ small and $c > 0$ such that for any $x \neq y \in K$ which are ρ_0-close, there exists $k \geqslant 1$ such that $\tilde{\mathsf{d}}_k(x, y) < \rho_0 \leqslant \tilde{\mathsf{d}}_{k+1}(x, y)$ and:

$$\mathsf{d}^{\mathrm{u}}(f^k(x), f^k([x, y])) > c \cdot \rho_0 \quad \text{or} \quad \mathsf{d}^{\mathrm{s}}(f^{-k}(x), f^{-k}([y, x])) > c \cdot \rho_0 \,.$$

By the bounded distortion property [Pe, Prop. 22.1], there exists a constant $C_0 > 1$ such that for any $n \geqslant 0$ and $x \in K$, the following estimates hold for every $y \in M$ such that $\tilde{\mathsf{d}}_n(x, y) < \rho_0$:

$$\begin{array}{lcr} y \in W^{\mathrm{u}}_{\mathrm{loc}}(x) & \Rightarrow & \|Df^n | T_y W^{\mathrm{u}}_{\mathrm{loc}}(x)\| \leqslant C_0 \|Df^n | E^{\mathrm{u}}_x\| \,, \\ y \in W^{\mathrm{u}}_{\mathrm{loc}}(x) \cap K & \Rightarrow & \|Df^{-n} | E^{\mathrm{s}}_y\| \leqslant C_0 \|Df^{-n} | E^{\mathrm{s}}_x\| \,, \\ y \in W^{\mathrm{s}}_{\mathrm{loc}}(x) & \Rightarrow & \|Df^{-n} | T_y W^{\mathrm{s}}_{\mathrm{loc}}(x)\| \leqslant C_0 \|Df^{-n} | E^{\mathrm{s}}_x\| \,, \\ y \in W^{\mathrm{s}}_{\mathrm{loc}}(x) \cap K & \Rightarrow & \|Df^n | E^{\mathrm{u}}_x\| \leqslant C_0 \|Df^n | E^{\mathrm{u}}_x\| \,. \end{array} \tag{2.11}$$

Using the bracket, it follows for every $x, y \in K$ such that $\tilde{\mathsf{d}}_n(x, y) < \rho_0$:

$$\|Df^n | E^{\mathrm{u}}_y\| \leqslant C_0^2 \|Df^n | E^{\mathrm{u}}_x\| \quad \text{and} \quad \|Df^{-n} | E^{\mathrm{s}}_y\| \leqslant C_0^2 \|Df^{-n} | E^{\mathrm{s}}_x\| \,. \tag{2.12}$$

For each $n \geqslant 0$, let B_n be the set of points $x \in K$ such that

$$\|D_x f^{-n} | E^{\mathrm{s}}\| \geqslant e^{(-\chi_s - \delta)n} \quad \text{and} \quad \|D_x f^n | E^{\mathrm{u}}\| \leqslant e^{(\chi_u + \delta)n} \,.$$

Again, for n large enough, by the Birkhoff ergodic Theorem we have $\mu(B_n) > 1/2$. By the same argument as in the proof of Theorem 2.1 (using Corollary A.3 instead of Theorem A.2), there exists a positive number $\rho < \rho_0$ such that for all sufficiently large n we can find a $(\tilde{\mathsf{d}}_n, \rho)$-separated set $F_n \subset B_n$ of cardinality at least $e^{(h_\mu - \delta)2n}$. As before, we use specification [KaH, Thrm. 18.3.9] to shadow each $x \in F_n$ by a periodic point $y = f^{2n+2n_0}(y)$ in such a way that $\tilde{\mathsf{d}}_n(x,y) < \rho/2$, where $n_0 \geqslant 0$ is independent of n. Let G_n be the set of periodic points y obtained in this way; it has the same cardinality as F_n. Since $x \in B_n$, it follows from (2.12) that:

$$\|Df^{-n-n_0}|E^{\mathrm{s}}_y\| \geqslant e^{(-\chi_s-2\delta)n} \quad \text{and} \quad \|Df^{n+n_0}|E^{\mathrm{u}}_y\| \leqslant e^{(\chi_u+2\delta)n} \, . \tag{2.13}$$

provided n is large enough.

Let Π_n be the union of the orbits of the points in G_n.

Claim 2.6. If n is large enough then the set Π_n is (d, ϵ_n)-separated with $\epsilon_n := e^{-(\chi_\mu+3\delta)(n+1)}$.

Proof. Take a pair of distinct points $x, y \in \Pi_n$, and let us prove that $\mathsf{d}(x,y) > \epsilon_n$. If $\mathsf{d}(x,y) > \rho_0$ then there is nothing to prove. Otherwise, as both points are fixed by f^{2n+2n_0}, so, by Claim 2.5, there exists k with $1 \leqslant k \leqslant n+n_0$ such that $\tilde{\mathsf{d}}_k(x,y) < \rho_0$ and:

$$\mathsf{d}^{\mathrm{u}}(f^k(x), f^k([x,y])) > c \cdot \rho_0 \quad \text{or} \quad \mathsf{d}^{\mathrm{s}}(f^{-k}(x), f^{-k}([y,x])) > c \cdot \rho_0 \, .$$

Let us consider the case where the first inequality holds; the other case is similar. Putting $z := [x,y]$, we have:

$$\begin{aligned}
\mathsf{d}^{\mathrm{u}}(x,z) &\geqslant C_0^{-1} \mathsf{d}^{\mathrm{u}}(f^k(x), f^k(z)) \cdot \|Df^k|E^{\mathrm{u}}_x\|^{-1} && \text{(by (2.11))} \\
&\geqslant C_0^{-1} \cdot c \cdot \rho_0 \cdot \|Df^{n+n_0}|E^{\mathrm{u}}_x\|^{-1} && \text{(since $Df|E^{\mathrm{u}}$ is expanding)} \\
&> C_0^{-1} \cdot c \cdot \rho_0 \cdot e^{-(\chi_\mu+2\delta)n} && \text{(by (2.13)).}
\end{aligned}$$

Since local stable and unstable manifolds are uniformly transverse, there exists a constant $C_1 > 0$ such that $\mathsf{d}(x,y) \geqslant C_1 \cdot \mathsf{d}^{\mathrm{u}}(x,z)$. This implies:

$$\mathsf{d}(x,y) \geqslant C_1 \cdot C_0^{-1} \cdot c \cdot \rho_0 \cdot e^{-(\chi_\mu+2\delta)n} \, .$$

It follows that $\mathsf{d}(x,y) > \epsilon_n$, for n uniformly sufficiently large. ☐

The same argument as in the proof of Theorem 2.1 (based on Theorem 1.6 again) yields that $\underline{\mathrm{mo}}(\mathcal{M}_f(K)) \geqslant 2(h_\mu - 2\delta)/(\chi_\mu + 3\delta)$. As δ is arbitrarily close to 0, we conclude that $\underline{\mathrm{mo}}(\mathcal{M}_f(K))$ is at least $2h_\mu/\chi_\mu = d = \dim K$ by (2.10). It follows that $\mathrm{mo}(\mathcal{M}_f^{\mathrm{erg}}(K)) = d$. This completes the proof of Theorem 2.1. ☐

3. Metric emergence and quantization of measures

3.1. **Quantization of measures.** The problem of *quantization of measures* consists in approximating efficiently a given measure by another measure with finite support: see [GrL].

Let (Y, d) be a compact metric space. Consider the set of probability measures $\mathcal{M}(Y)$ endowed with a metric also denoted d, which can be either a q-Wasserstein metric W_q, $q \in [1, \infty)$ or the Lévy–Prokhorov metric LP.

Definition 3.1. The *quantization number* of a measure $\mu \in \mathcal{M}(Y)$ at a scale (or resolution) $\epsilon > 0$, denoted $Q_\mu(\epsilon)$, is defined as the least integer N such that there exists a probability measure ν with $\mathsf{d}(\mu, \nu) \leqslant \epsilon$ and supported on a set of cardinality N.

Here is a reformulation of the definition when a Wassertein metric is used:

Proposition 3.2. *The quantization number $Q_\mu(\epsilon)$ for the q-Wasserstein metric W_q is the minimal cardinality N of a set $F = \{x_1, \dots, x_N\}$ so that:*

$$\int_Y (\mathsf{d}(x, F))^q \, d\mu(x) \leqslant \epsilon^q \, .$$

Proof. Fix $\epsilon > 0$ and let $F \subset Y$ be a set of minimal cardinality N such that $\int [\mathsf{d}(x, F)]^q d\mu(x) \leqslant \epsilon^q$.

Take a measurable map $h \colon Y \to F$ that associates to each element in Y a closest element in F (w.r.t. the d metric). Let $\nu := h_* \mu \in \mathcal{M}(Y)$; this is a measure supported on F. We claim that $\mathsf{W}_q(\mu, \nu) \leqslant \epsilon$. Indeed, $\pi := (\mathrm{id} \times h)_*(\mu)$ is a transport plan from μ to ν with cost

$$\int [\mathsf{d}(x, h(x))]^q d\mu(x) = \int [\mathsf{d}(x, F)]^q d\mu(x) \leqslant \epsilon^q \, .$$

We have shown that $Q_\mu(\epsilon) \leqslant N$.

Let us prove the reverse inequality. Let $\nu \in \mathcal{M}(Y)$ be a measure whose support $F' \subset Y$ has cardinality $Q_\mu(\epsilon)$ and such that $\mathsf{W}_q(\mu, \nu) \leqslant \epsilon$. This means that there is a transport plan $\pi \in \mathcal{M}(Y \times Y)$ from μ to ν with cost at most ϵ^q. Consider a disintegration of π, that is, a family (ν_ξ) of elements of $\mathcal{M}(Y)$, defined for μ-almost every $\xi \in Y$, such that $\pi = \int \delta_\xi \otimes \nu_\xi d\mu(\xi)$. As the second marginal of π equals ν, whose support is the finite set F', it follows that $\operatorname{supp} \nu_\xi \subset F'$ for μ-almost every ξ. Therefore:

$$\epsilon^q \geqslant \mathrm{cost}(\pi) = \iint [\mathsf{d}(\xi, \eta)]^q d\nu_\xi(\eta) d\mu(\xi) \geqslant \int [\mathsf{d}(\xi, F')]^q d\mu(\xi) \, .$$

This shows that $N \leqslant \# F' = Q_\mu(\epsilon)$. ☐

Here is a similar characterization of the quantization number for the case of the Lévy–Prokhorov metric:

Proposition 3.3. *The quantization number $Q_\mu(\epsilon)$ for the* LP *metric is the least number of closed balls of radius ϵ that cover a set of μ-measure at least $1 - \epsilon$.*

Proof. Straightforward. ☐

Similarly to the definition of the lower and upper box-counting dimensions, following [GrL, p. 155] the *lower* and *upper quantization dimensions* of $\mu \in \mathcal{M}(Y)$ are defined as:

$$\underline{\dim}(\mu) := \liminf_{\epsilon \to 0} \frac{\log Q_\mu(\epsilon)}{-\log \epsilon} \quad \text{and} \quad \overline{\dim}(\mu) := \limsup_{\epsilon \to 0} \frac{\log Q_\mu(\epsilon)}{-\log \epsilon} \, .$$

If these numbers coincide then they are denoted by $\dim(\mu)$ and called *quantization dimension*. Furthermore, the *lower* and *upper quantization orders* are defined as:

$$\underline{\mathrm{qo}}(\mu) := \liminf_{\epsilon \to 0} \frac{\log \log Q_\mu(\epsilon)}{-\log \epsilon} \quad \text{and} \quad \overline{\mathrm{qo}}(\mu) := \limsup_{\epsilon \to 0} \frac{\log \log Q_\mu(\epsilon)}{-\log \epsilon} \,.$$

If these numbers coincide then they are denoted by $\mathrm{qo}(\mu)$ and called *quantization order*.

Proposition 3.4. *For any resolution $\epsilon > 0$, the quantization number of any $\mu \in \mathcal{M}(Y)$ is bounded from above by the covering number of Y, that is:*

$$Q_\mu(\epsilon) \leqslant D_Y(\epsilon) \,.$$

In particular,

$$\begin{aligned} \underline{\dim}(\mu) \leqslant \underline{\dim}(Y) \quad &\textit{and} \quad \overline{\dim}(\mu) \leqslant \overline{\dim}(Y) \,, \\ \underline{\mathrm{qo}}(\mu) \leqslant \underline{\mathrm{mo}}(Y) \quad &\textit{and} \quad \overline{\mathrm{qo}}(\mu) \leqslant \overline{\mathrm{mo}}(Y) \,. \end{aligned}$$

Proof. Given an ϵ-dense set F of cardinality N, we can transport any measure $\mu \in \mathcal{M}(Y)$ to a measure supported on F with cost $\leqslant \epsilon^q$ with respect to the cost function d^q. This shows that $Q_\mu(\epsilon) \leqslant D_Y(\epsilon)$ with respect to the W_q distance. In view of Proposition 3.3, the same statement is also immediate for the LP distance. Then it follows that quantization dimensions are bounded by box-counting dimensions, and quantization orders are bounded by metric orders. □

Example 3.5. Consider $Y = [0,1]$ with the usual metric, and endow the space $\mathcal{M}([0,1])$ with the metric W_q. Consider the Lebesgue measure on $[0,1]$; its quantization number is:

$$Q_{\mathrm{Leb}}(\epsilon) = \left\lceil \frac{1}{2(q+1)^{1/q}\,\epsilon} \right\rceil \,,$$

and in particular the quantization dimension is 1. Indeed, given $N \geqslant 1$, the probability measure on $[0,1]$ supported on N points which is W_q-closest to Lebesgue is:

$$\nu_N := \frac{1}{N} \sum_{j=1}^{N} \delta_{\frac{2j-1}{2N}} \,, \quad \text{for which} \quad \mathsf{W}_q(\nu_N, \mathrm{Leb}) = \frac{1}{2(q+1)^{1/q}\,N}$$

(see [GrL, p. 69]), so the asserted formula for $Q_{\mathrm{Leb}}(\epsilon)$ follows.

Example 3.6. If μ is a compactly supported measure on $\mathbb{R}^d$ which is absolutely continuous with respect to Lebesgue measure then $\dim(\mu) = d$; actually there is a precise asymptotic formula for the quantization number $Q_\mu(\epsilon)$ with respect to the W_q distance: see [GrL, p. 78, p. 52].

Example 3.7. See the paper [LiM] for the computation of the quantization dimension of certain self-similar measures (F-conformal measures) supported on fractal sets defined by conformal iterated function systems; let us note that the answer depends on the exponent q.

Example 3.8. The metric entropy of an ergodic measure can be described in terms of quantization numbers: see Appendix A.3.

In this paper, we are mostly interested in the situation where the quantization orders are positive, and so the quantization dimensions are infinite.

In view of Proposition 3.4, the next result yields measures with maximal quantization order:

Theorem 3.9. *Let Y be a Borel subset of a compact metric space Z. Then there exists a probability measure $\mu \in \mathcal{M}(Y)$ such that:*

$$\underline{\mathrm{qo}}(\mu) = \underline{\mathrm{mo}}(Y) \quad \textit{and} \quad \overline{\mathrm{qo}}(\mu) = \overline{\mathrm{mo}}(Y)\,.$$

The proof is given in Section 3.5.

3.2. **Ergodic decomposition.** Let X be a compact metric space and let $f\colon X \to X$ be a continuous map. Recall that the empirical measure at a point $x \in X$ is defined as $\mathbf{e}^f(x) := \lim \frac{1}{n} \sum_{i=0}^{n-1} \delta_{f^i x}$, when this limit exists. By the ergodic decomposition theorem (see [DGS, § 13] or [Ma2, § II.6]), there exists a Borel set $X_0 \subset X$ with full probability (that is, $\mu(X_0) = 1$ for every $\mu \in \mathcal{M}_f(X)$) such that for every $x \in X_0$, the empirical measure $\mathbf{e}^f(x)$ is f-invariant and ergodic. So for any $\mu \in \mathcal{M}_f(X)$, the measure $\mathbf{e}^f_*\mu \in \mathcal{M}(\mathcal{M}(X))$ gives full weight to the set $\mathcal{M}_f^{\mathrm{erg}}(X) \subset \mathcal{M}_f(X)$ of ergodic measures, and its barycenter $\mathrm{bar}(\mathbf{e}^f_*\mu) := \int \nu \, d(\mathbf{e}^f_*\mu)(\nu)$ is μ. The probability measure $\mathbf{e}^f_*\mu$ is called the *ergodic decomposition* of μ. There is a canonical bijection $\mathcal{M}_f(X) \to \mathcal{M}(\mathcal{M}_f^{\mathrm{erg}}(X))$, namely $\mu \mapsto \mathbf{e}^f_*\mu$.

Remark 3.10. Generic conservative diffeomorphisms (in any topology) constitute continuity points of the ergodic decomposition of Lebesgue measure: see [AB, Thrm. B]. We will see later in Section 5.2 non-trivial examples of continuity points w.r.t. the C^∞ topology.

Let us note the following property for later use:

Lemma 3.11 (Factors and ergodic decompositions)**.** *Suppose X and Y are compact metric spaces and let $f\colon X \to X$ and $g\colon Y \to Y$ be continuous maps which are semi-conjugate ($g \circ \varphi = \varphi \circ f$) via a continuous $\varphi\colon X \to Y$. Let $\Phi\colon \mathcal{M}(X) \to \mathcal{M}(Y)$ be the map $\mu \mapsto \varphi_*\mu$. Then $\Phi(\mathcal{M}_f(X)) \subset \mathcal{M}_g(Y)$, $\Phi(\mathcal{M}_f^{\mathrm{erg}}(X)) \subset \mathcal{M}_g^{\mathrm{erg}}(Y)$, and:*

$$\forall \mu \in \mathcal{M}_f(X), \quad \mathbf{e}^g_*(\varphi_*\mu) = \Phi_*(\mathbf{e}^f_*(\mu))\,.$$

When no confusion arises, we will write φ_* instead of Φ, so the last equation becomes $\mathbf{e}^g_*(\varphi_*\mu) = \varphi_{**}(\mathbf{e}^f_*(\mu))$.

Proof. Let $\mu \in \mathcal{M}_f(X)$ and let $\nu := \varphi_*(\mu)$. Then $g_*\nu = (g \circ \varphi)_*(\mu) = (\varphi \circ f)_*(\mu) = \nu$, that is, $\nu \in \mathcal{M}_g(Y)$, proving the first assertion.

Note that that if $B \subset Y$ is a g-invariant Borel set then $\varphi^{-1}(B)$ is f-invariant; it follows that ν is ergodic if μ is, proving the second assertion.

Let $\hat{\mu} := \mathbf{e}^f_*(\mu)$ and $\hat{\nu} := \mathbf{e}^g_*(\nu)$ be the corresponding ergodic decompositions. For every Borel set $B \subset Y$, we have:

$$\nu(B) = \mu(\varphi^{-1}(B)) = \int_{\mathcal{M}(X)} \eta(\varphi^{-1}(B))\, d\hat{\mu}(\eta)$$
$$= \int_{\mathcal{M}(X)} (\Phi(\eta))(B)\, d\hat{\mu}(\eta) = \int_{\mathcal{M}(Y)} \xi(B)\, d(\Phi_*(\hat{\mu}))(\xi)\,.$$

This means that ν is the barycenter of $\Phi_*(\hat{\mu})$. Since $\hat{\mu}$ gives full weight to $\mathcal{M}^{\mathrm{erg}}_f(X)$, the measure $\Phi_*(\hat{\mu})$ gives full weight to $\mathcal{M}^{\mathrm{erg}}_g(Y)$, and by uniqueness of the ergodic decomposition, it follows that $\Phi_*(\hat{\mu})$ equals $\hat{\nu}$, the ergodic decomposition of ν. ☐

3.3. **Metric emergence.** Given a continuous self-map $f\colon X \to X$ of a compact metric space X, we consider the set $\mathcal{M}(X)$ with a metric $\mathsf{d} \in \{\mathsf{W}_p\,;\, 1 \leqslant p < \infty\} \cup \{\mathsf{LP}\}$. We have introduced in Definition 0.2 the *metric emergence* of a measure $\mu \in \mathcal{M}(X)$. In the case μ is invariant, we have the following characterization of metric emergence:

Proposition 3.12. *For every dynamics $f\colon X \to X$, the metric emergence of any invariant measure $\mu \in \mathcal{M}_f(X)$ equals the quantization number of the ergodic decomposition $\hat{\mu} := \mathbf{e}^f_*\mu$ (considered as a measure on $\mathcal{M}(X)$):*

$$\mathscr{E}_\mu(f)(\epsilon) = Q_{\hat{\mu}}(\epsilon)\,,$$

where $Q_{\hat{\mu}}$ is the quantization number of $\hat{\mu}$ for the metric W_1 of $\mathcal{M}(\mathcal{M}(X))$.

Proof. Combine Definition 0.2 and Proposition 3.2. ☐

Remark 3.13. Given a parameter $q \geqslant 1$, we may define the *q-emergence* of an f-invariant measure μ at scale $\epsilon > 0$ as:

$$\mathscr{E}^{(q)}_\mu(\epsilon) := \min\left\{ N\,;\, \exists F \subset \mathcal{M}(X) \text{ with } \# F \leqslant N,\ \int \mathsf{d}(\mathbf{e}^f(x), F)^q d\mu(x) \leqslant \epsilon^q \right\}.$$

By Proposition 3.2, q-emergence is the quantization number of the ergodic decomposition with respect to the metric W_q on $\mathcal{M}(\mathcal{M}(X))$. For simplicity we will focus our study on $q = 1$.

Metric and topological emergences may be compared as follows:

Proposition 3.14. *For every dynamics $f\colon X \to X$, the metric emergence of any invariant measure $\mu \in \mathcal{M}_f(X)$ is at most the topological emergence:*

$$\mathscr{E}_\mu(f)(\epsilon) \leqslant \mathscr{E}_{\mathrm{top}}(f)(\epsilon)\,, \quad \forall \epsilon > 0\,,$$

provided both emergences are computed using the same metric W_p or LP on $\mathcal{M}(X)$.

Proof. By Proposition 3.12, the metric emergence $\mathscr{E}_\mu(f)(\epsilon)$ equals the quantization number $Q_{\hat{\mu}}(\epsilon)$ of the ergodic decomposition $\hat{\mu} := \mathbf{e}^f_*\mu$. Note that $\hat{\mu}$ is a measure on $Y := \mathcal{M}^{\mathrm{erg}}_f(X)$ which is a Borel subset of $Z := \mathcal{M}(X)$.

By Proposition 3.4, $Q_{\hat{\mu}}(\epsilon)$ is at most the relative covering number $D_Z(Y,\epsilon)$, which equals the topological emergence $\mathscr{E}_{\mathrm{top}}(f)(\epsilon)$ by its own definition (2.1). $\square$

We are now able to deduce the variational principle for emergence announced at the introduction:

Proof of Theorem E. Applying Theorem 3.9 with $Y := \mathcal{M}_f^{\mathrm{erg}}(X)$, $Z := \mathcal{M}(X)$ and $q = 1$, we obtain a probability measure $\nu \in \mathcal{M}(\mathcal{M}_f^{\mathrm{erg}}(X))$ such that:

$$\underline{\mathrm{qo}}(\mu) = \underline{\mathrm{mo}}(\mathcal{M}_f^{\mathrm{erg}}(X)) \quad \text{and} \quad \overline{\mathrm{qo}}(\mu) = \overline{\mathrm{mo}}(\mathcal{M}_f^{\mathrm{erg}}(X))\,.$$

Let $\mu := \int_{\mathcal{M}(X)} \eta d\nu(\eta)$. Since ν gives full weight to $\mathcal{M}_f^{\mathrm{erg}}(X)$, the measure μ is invariant and its ergodic decomposition is ν. Bearing in mind Proposition 3.12 and the definitions of lower and upper quantizations orders and metric orders, we obtain the equalities stated in Theorem E. $\square$

3.4. **Some properties of quantization numbers.** In this section we prove a few general properties about quantization numbers that will be needed later. To simplify matters, *all quantization numbers in this section are computed w.r.t. the W_1 metric.*

Lemma 3.15. *For all μ_1, $\mu_2 \in \mathcal{M}(Y)$,*

$$\mathsf{W}_1(\mu_1,\mu_2) \leqslant \epsilon \quad \Rightarrow \quad Q_{\mu_2}(2\epsilon) \leqslant Q_{\mu_1}(\epsilon)\,.$$

Proof. Immediate. $\square$

The next two lemmas deal with pushing forward a measure under a Lipschitz map, and the effect of this operation on the quantization numbers:

Lemma 3.16. *Let $f\colon (Y,\mathsf{d}) \to (Z,\mathsf{d})$ be a κ-Lipschitz map between compact metric spaces. Let $F\colon (\mathcal{M}(Y),\mathsf{W}_1) \to (\mathcal{M}(Z),\mathsf{W}_1)$ be the map $\mu \mapsto f_*\mu$. Then F is κ-Lispchitz.*

Proof. Given μ_1, $\mu_2 \in \mathcal{M}(Y)$, consider a transport plan $\pi \in \mathcal{M}(Y \times Y)$. Then $\tilde{\pi} := (f \times f)_*(\pi)$ is a transport plan from $f_*\mu_1$ to $f_*\mu_2$ with:

$$\mathrm{cost}(\tilde{\pi}) = \int \mathsf{d}(f(x),f(y))\, d\pi(x,y) \leqslant \kappa \int \mathsf{d}(x,y)\, d\pi(x,y) = \kappa\,\mathrm{cost}(\pi)\,.$$

So $\mathsf{W}_1(f_*\mu_1, f_*\mu_2) \leqslant \kappa \mathsf{W}_1(\mu_1,\mu_2)$. $\square$

Lemma 3.17. *Let (Y,d) and (Z,d) be compact metric spaces. Let $f\colon Y \to Z$ be a κ-Lipschitz map. Given a measure $\mu \in \mathcal{M}(Y)$, consider its push-forward $\nu := f_*\mu \in \mathcal{M}(Z)$. Then for every $\epsilon > 0$, we have:*

$$Q_\mu(\epsilon) \geqslant Q_\nu(\kappa\epsilon)\,.$$

Proof. Given $\mu \in \mathcal{M}(Y)$ and $\epsilon > 0$, let $\tilde{\mu} \in \mathcal{M}(Y)$ be a measure supported on $n := Q_\mu(\epsilon)$ points with $\mathsf{W}_1(\mu,\tilde{\mu}) \leqslant \epsilon$. By Lemma 3.16, the measures $\nu := f_*\mu$ and $\tilde{\nu} := f_*\tilde{\mu}$ satisfy $\mathsf{W}_1(\nu,\tilde{\nu}) \leqslant \kappa\epsilon$. Since $\tilde{\nu}$ is supported on at most n points, we conclude that $Q_\nu(\kappa\epsilon) \leqslant n$. $\square$

The next two lemmas will be used several times, in particular in the proof of Theorem 3.9:

Lemma 3.18. *Let μ, $\mu_1 \in \mathcal{M}(Y)$ be such that $\mu \geqslant t\mu_1$, for some $t > 0$. Then:*

$$Q_\mu(t\epsilon) \geqslant Q_{\mu_1}(\epsilon)\,.$$

Proof. Let $\tilde{\epsilon} := t\epsilon$. Let ν be a measure supported on a set of cardinality $\ell := Q_\mu(\tilde{\epsilon})$ and such that $\mathsf{W}_1(\mu,\nu) \leqslant \tilde{\epsilon}$. Let π be a transport plan from μ to ν with cost (w.r.t. d) not greater than $\tilde{\epsilon}$.

The Radon–Nikodym derivative $f := \frac{d\mu_1}{d\mu}$ is well-defined and satisfies $0 \leqslant f \leqslant t^{-1}$ at μ-a.e. point. Consider the measure $\tilde{\pi}$ on $Y \times Y$ defined by:

$$d\tilde{\pi}(x,y) = f(x)d\pi(x,y)\,.$$

Then $\tilde{\pi}$ is a probability, its first marginal is μ_1, and its second marginal is some measure $\tilde{\nu}$ which is absolutely continuous with respect to ν and therefore supported on a set of cardinality at most ℓ. We have:

$$\begin{aligned}\mathsf{W}_1(\mu_1,\tilde{\nu}) \leqslant \operatorname{cost}(\tilde{\pi}) = \int \mathsf{d}(x,y)\,d\tilde{\pi}(x,y) &= \int \mathsf{d}(x,y)\,f(x)\,d\pi(x,y)\\ &\leqslant t^{-1}\int \mathsf{d}(x,y)\,d\pi(x,y) = t^{-1}\operatorname{cost}(\pi) \leqslant t^{-1}\tilde{\epsilon}\,.\end{aligned}$$

That is, $\mathsf{W}_1(\mu_1,\tilde{\nu}) \leqslant t^{-1}\tilde{\epsilon} = \epsilon$. It follows that $\#\operatorname{supp}\tilde{\nu} \geqslant Q_{\mu_1}(\epsilon)$, and so $\ell \geqslant Q_{\mu_1}(\epsilon)$, as claimed. □

Lemma 3.19. *Let $\epsilon > 0$ and let $F \subset Y$ be an ϵ-separated set. Let $n := \# F$ and let μ be the equidistributed probability measure with support F. Let ν be any probability measure whose support has cardinality $m < n$. Then:*

$$\mathsf{W}_1(\mu,\nu) \geqslant \frac{n-m+1}{n}\cdot\frac{\epsilon}{2}\,.$$

Proof. Let $\operatorname{supp}\mu = \{x_1,\dots,x_n\}$ and $\operatorname{supp}\nu = \{y_1,\dots,y_m\}$. Since μ is equidistributed, transport plans from μ to ν take the form:

$$\pi = \pi_A = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}\delta_{(x_i,y_j)}\,,$$

where $A = (a_{ij})$ is a row-stochastic $n \times m$ matrix (that is, each a_{ij} is non-negative and $\sum_{j=1}^{m} a_{ij} = 1$ for every i). The cost of π_A is:

$$\operatorname{cost}(\pi_A) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}\mathsf{d}(x_i,y_j)\,,$$

which can be viewed as an affine function on the set of row-stochastic matrices. This set is compact and convex, and its extremal points consist on the matrices that contain exactly one entry equal to 1 on each row. So it is sufficient to consider matrices of this type in order to find a lower bound for the cost. Thus consider a row-stochastic matrix $A = A_T$ whose nonzero entries are $a_{i,T(i)} = 1$ for some map $T\colon \{1,\dots,n\} \to \{1,\dots,m\}$.

Claim 3.20. For every $j \in \{1, \dots, m\}$ such that $s := \# T^{-1}(j) \geqslant 2$, the following holds:

$$\sum_{i \in T^{-1}(j)} \mathsf{d}(x_i, y_j) \geqslant \frac{s \cdot \epsilon}{2}\,. \tag{3.1}$$

Proof of the claim. Indeed, write $T^{-1}(j) = \{i_1, \dots, i_s\}$; then the left hand side of (3.1) equals:

$$\sum_{k=1}^{s} \mathsf{d}(x_{i_k}, y_j) = \frac{1}{s-1} \sum_{1 \leqslant k < \ell \leqslant s} \left[\mathsf{d}(x_{i_k}, y_j) + \mathsf{d}(x_{i_\ell}, y_j)\right].$$

For every $1 \leqslant k < \ell \leqslant s$, since F is ϵ-separated, it hold:

$$\mathsf{d}(x_{i_k}, y_j) + \mathsf{d}(x_{i_\ell}, y_j) \geqslant \mathsf{d}(x_{i_k}, x_{i_\ell}) \geqslant \epsilon\,.$$

So we obtain:

$$\sum_{k=1}^{s} \mathsf{d}(x_{i_k}, y_j) \geqslant \frac{1}{s-1} \cdot \frac{s(s-1)}{2} \cdot \epsilon = \frac{s \cdot \epsilon}{2}\,,$$

as claimed. □

Using (3.1), we estimate:

$$\operatorname{cost}(\pi_{A_T}) = \frac{1}{n} \sum_{i=1}^{n} \mathsf{d}(x_i, y_{T(i)}) = \frac{1}{n} \sum_{j=1}^{m} \sum_{i \in T^{-1}(j)} \mathsf{d}(x_i, y_j) \geqslant \frac{n_*}{n} \frac{\epsilon}{2}\,,$$

where

$$\begin{aligned} n_* := \sum_{\substack{j \in \{1,\dots,m\},\\ \# T^{-1}(j) > 1}} \# T^{-1}(j) &= n - \sum_{\substack{j \in \{1,\dots,m\},\\ \# T^{-1}(j) \leqslant 1}} \# T^{-1}(j) \\ &= n - \# \left\{ j \in \{1, \dots, m\} \; ; \; \# T^{-1}(j) = 1 \right\} \\ &= n - m + \# \left\{ j \in \{1, \dots, m\} \; ; \; \# T^{-1}(j) \neq 1 \right\} \\ &\geqslant n - m + 1\,, \end{aligned}$$

since $m < n$. We conclude that $\operatorname{cost}(\pi_A)$ is at least $\frac{n-m+1}{n} \frac{\epsilon}{2}$ for every matrix A of type A_T, and therefore for every row-stochastic matrix A. The lemma follows. □

3.5. **Existence of a measure with essentially maximal quantization numbers.** In this subsection we prove Theorem 3.9, which was used to deduce Theorem E.

Proof of Theorem 3.9. It is sufficient to prove the theorem assuming that $\mathcal{M}(Y)$ is metrized with the W_1 distance. Indeed, by the first inequality in (1.2) (see p. 10), if the exponent q is reduced then the metric W_q does not increase, and so neither do quantization numbers and orders. Furthermore, by the second inequality in (1.3), the metric LP is bounded from below by a constant factor of the metric W_1, and so quantization numbers and orders with respect to LP are bounded from below by the corresponding quantities with respect to W_1. So from now on we assume that $\mathcal{M}(Y)$ is metrized with the W_1 distance.

By Proposition 3.4, it is sufficient to show the existence of a measure $\mu \in \mathcal{M}(Y)$ such that:

$$\underline{\mathrm{qo}}(\mu) \geqslant \underline{\mathrm{mo}}(Y) \quad \text{and} \quad \overline{\mathrm{qo}}(\mu) \geqslant \overline{\mathrm{mo}}(Y)\,. \tag{3.2}$$

Recall that, given $\epsilon > 0$, the corresponding packing number is denoted by $S_Y(\epsilon)$. We set $\epsilon_i := 2^{-i^2}$ for every $i \geqslant 1$. Let $F_i \subset Y$ be a $4\epsilon_i$-separated set of cardinality $n_i := S_Y(4\epsilon_i)$, and let $\mu_i \in \mathcal{M}(Y)$ be the equidistributed probability measure with support F_i. By Lemma 3.19, if ν is a probability measure whose support has cardinality at most $m_i := \lceil n_i/2 \rceil$ then

$$\mathsf{W}_1(\mu_i, \nu) \geqslant \epsilon_i\,.$$

That is, in terms of quantization number:

$$Q_{\mu_i}(\epsilon_i) \geqslant m_i\,. \tag{3.3}$$

Now consider the following probability measure:

$$\mu := \sum_{i=1}^{\infty} t_i \mu_i\,, \quad \text{where} \quad t_i := 2^{-i}\,.$$

By Lemma 3.18, for every $i \geqslant 1$ we have $Q_\mu(\tilde\epsilon_i) \geqslant Q_{\mu_i}(\epsilon_i)$, where $\tilde\epsilon_i := t_i \epsilon_i$. Using (3.3) we obtain:

$$\frac{\log\log Q_\mu(\tilde\epsilon_i)}{-\log \tilde\epsilon_i} \geqslant \frac{\log\log m_i}{-\log \tilde\epsilon_i} \underset{i\to\infty}{\sim} \frac{\log\log n_i}{-\log(4\epsilon_i)} := \frac{\log\log S_Y(4\epsilon_i)}{-\log(4\epsilon_i)}\,. \tag{3.4}$$

Claim 3.21. The following equalities hold:

$$\liminf_{i\to\infty} \frac{\log\log Q_\mu(\tilde\epsilon_i)}{-\log\tilde\epsilon_i} = \underline{\mathrm{qo}}(\mu)\,, \qquad \limsup_{i\to\infty} \frac{\log\log Q_\mu(\tilde\epsilon_i)}{-\log\tilde\epsilon_i} = \overline{\mathrm{qo}}(\mu)\,, \tag{3.5}$$

$$\liminf_{i\to\infty} \frac{\log\log S_Y(4\epsilon_i)}{-\log(4\epsilon_i)} = \underline{\mathrm{mo}}(Y)\,, \quad \limsup_{i\to\infty} \frac{\log\log S_Y(4\epsilon_i)}{-\log(4\epsilon_i)} = \overline{\mathrm{mo}}(Y)\,. \tag{3.6}$$

Proof of the claim. Let us prove (3.6); the proof of (3.5) is essentially the same. Given $\epsilon > 0$, let i be such that $\epsilon \in [4\epsilon_{i+1}, 4\epsilon_i]$. We have $S_Y(4\epsilon_i) \leqslant S_Y(\epsilon) \leqslant S_Y(4\epsilon_{i+1})$ and so:

$$\frac{\log\log S_Y(4\epsilon_{i+1})}{-\log(4\epsilon_i)} \geqslant \frac{\log\log S_Y(\epsilon)}{-\log\epsilon} \geqslant \frac{\log\log S_Y(4\epsilon_i)}{-\log(4\epsilon_{i+1})}\,.$$

Since $\log(4\epsilon_i) \sim \log(4\epsilon_{i+1})$ as $i \to \infty$, inequalities (3.6) follow. □

Combining (3.4) with Claim 3.21 we obtain inequality (3.2) and the theorem. □

4. Examples of conservative dynamics with high metric emergence

We are going to study the emergence of dynamics on the annulus $\mathbb{A}$:

$$\mathbb{A} := \mathbb{T} \times [0,1] \quad \text{with } \mathbb{T} := \mathbb{R}/\mathbb{Z}\,.$$

Lebesgue measure on either of theses sets is denoted by Leb.

The *horizontal flow* associated to a C^∞ function $\omega\colon [0,1] \to \mathbb{R}$ is defined as:

$$R^t_\omega\colon (\theta,\rho) \in \mathbb{A} \mapsto (\theta + \omega(\rho)t, \rho) \in \mathbb{A}\,. \tag{4.1}$$

So $(R^t_\omega)_t$ is a conservative smooth flow on the annulus. Assume that ω has no critical points. Then, for every fixed $t \neq 0$, Lebesgue almost every $\rho \in [0,1]$ has the property that $\omega(\rho) \cdot t$ is irrational, and therefore for every $\theta \in \mathbb{T}$, the empirical measure $\mathbf{e}^{R^t_\omega}(\theta,\rho)$ equals:

$$\lambda_\rho := \mathrm{Leb}_\mathbb{T} \otimes \delta_\rho \quad \text{(Lebesgue measure on the circle } \mathbb{T} \times \{\rho\}).$$

Hence the ergodic decomposition of the Lebesgue measure with respect to the time t map R^t_ω does not depend on $t \neq 0$ and is given by:

$$\mathbf{e}^{R^t_\omega}_*(\mathrm{Leb}) = \int_0^1 \delta_{\lambda_\rho}\, d\rho\,. \tag{4.2}$$

4.1. Robust examples of at least polynomial emergence.

Proposition 4.1. *Suppose $\omega\colon [0,1] \to \mathbb{R}$ is a smooth function without critical points and let $(R^t_\omega)_t$ be the corresponding horizontal flow. For every $t \neq 0$, the metric emergence of the time t map R^t_ω with respect to the Wasserstein metric W_1 is:*

$$\mathscr{E}_{\mathrm{Leb}}(R^t_\omega)(\epsilon) = \lceil (4\epsilon)^{-1} \rceil\,.$$

Proof. As seen in (4.2), the ergodic decomposition $\hat\mu := \mathbf{e}^{R^t_\omega}_*(\mathrm{Leb})$ is equidistributed on the curve $\{\lambda_\rho : \rho \in [0,1]\}$. This curve endowed with the Wasserstein metric W_1 is isometric to the unit interval $[0,1]$ endowed its usual distance; the isometry sends the measure $\hat\mu$ to the Lebesgue measure on $[0,1]$. Thus $Q_{\hat\mu}(\epsilon) = Q_{\mathrm{Leb}|[0,1]}(\epsilon) = \lceil (4\epsilon)^{-1} \rceil$, by Example 3.5 with $q = 1$. Using Proposition 3.12 we conclude. □

KAM theory ensures that most of the invariant circles of R^t_ω persist for any conservative C^∞ perturbation. As a consequence, we obtain C^∞-open sets of conservative surface diffeomorphisms whose metric emergence is at least of the order of ϵ^{-1}: see Section 5.2, more specifically Corollary 5.7.

4.2. Construction of a smooth conservative flow with high emergence at a given scale.

The heart of the proof of Theorem B is the following result:

Proposition 4.2. *There exists $C > 0$ such that for every $\epsilon_* > 0$, there exists a smooth conservative diffeomorphism h of $\mathbb{A}$ satisfying the following property. For every function $\omega \in C^\infty([0,1],\mathbb{R})$ without critical points and for every $t \neq 0$, the map $\Psi^t := h \circ R^t_\omega \circ h^{-1}$ satisfies:*

$$\mathscr{E}_{\mathrm{Leb}}(\Psi^t)(\epsilon_*) \geqslant \exp(C\epsilon_*^{-2}),$$

where the emergence is computed with respect to the Wasserstein metric W_1. Furthermore, h equals identity on a neighborhood of the boundary of $\mathbb{A}$.

The proof of the proposition will occupy the rest of this subsection.

Proof. We will actually construct a sequence h_n of diffeomorphisms such that the corresponding flows $\Psi^t_n := h_n^{-1} \circ R^t_\omega \circ h_n$ have high emergence at a certain scale ϵ_n; then we will show that for every $\epsilon_* > 0$ we can choose an appropriate $h = h_n$ and obtain the conclusion of Proposition 4.2. The proof is divided into several steps.

Zeroth step. Let $n \geqslant 3$ be an arbitrary integer. We will fix several numbers depending on n. Let $N := 32 \cdot n^2$. Let $M = m \cdot n$ be the multiple of n as big as possible such that:

$$M \leqslant (2N)^{-1/2} e^{\pi N/4^3}\,. \tag{4.3}$$

It is clear from this definition that:

$$\log M \asymp n^2\,. \tag{4.4}$$

Finally, let $\eta := 1/(1000n)$ and $\kappa := 1 - \eta$.

First step. The real proof begins with the construction of certain families of boxes in the annulus $\mathbb{A}$. An *$a \times b$-box* is a set of the form $I \times J$ where $I \subset \mathbb{T}$ and $J \subset [0,1]$ are closed intervals of respective lengths a (the *width* of the box) and b (the *height* of the box). An *a-square* is an $a \times a$-box. A *$k \times \ell$-family* is a disjoint collection of boxes of the form $I_i \times J_j$ where $1 \leqslant i \leqslant k$, $1 \leqslant j \leqslant \ell$. Such a family can be partitioned (in the obvious way) into k subfamilies called *columns* and into ℓ subfamilies called *rows*.

Let $\mathcal{G}$ be a $8n \times 4n$-family of $\frac{1}{10n}$-squares contained in the lower half-annulus $\mathbb{T} \times [0, \frac{1}{2}]$ and such that the gaps between rows and between columns is $\frac{1}{40n}$.

Inside each square G from the family $\mathcal{G}$ we take a $n \times m$-family $\mathcal{L}_G$ of $\frac{2\kappa}{N} \times \frac{1}{11M}$-boxes; it is possible to construct such a family since:

$$\max\left\{n \cdot \frac{2\kappa}{N},\, m \cdot \frac{1}{11M}\right\} = \max\left\{\frac{\kappa}{16n}, \frac{1}{11n}\right\} < \frac{1}{10n} = \text{width of } G.$$

Let $\mathcal{L} := \bigsqcup_{G \in \mathcal{G}} \mathcal{L}_G$; this is a family composed of NM boxes.

Let $\mathcal{U}$ be a $\frac{N}{2} \times M$-family of $\frac{2\kappa}{N} \times \frac{1}{11M}$-boxes contained in the upper half-annulus $\mathbb{T} \times [\frac{1}{2}, 1]$.

Second step. We will need some auxiliary combinatorial data, namely certain coloring of our boxes. We start by painting each $\mathcal{G}$-square with a different color, and then we paint each $\mathcal{L}_G$-box with the same color as G. We claim that it is possible to paint each $\mathcal{U}$-box with one of the N previously chosen colors so that the following properties hold:

- no row contains repeated colors (that is, exactly $N/2$ different colors appear in each row), and
- for any pair of distinct rows, there are at least $N/4$ colors that appear in one row but not in the other.

Indeed, if each choice of $N/2$ among N colors can be identified with a function $f\colon \{1,\dots,N\} \to \{0,1\}$ such that $\sum_{k=1}^{N} f(k) = \frac{N}{2}$. The set F of such functions was considered previously in the proof of Theorem 1.6, where we have shown the existence of a set $F' \subset F$ which is $N/4$-separated w.r.t. the Hamming distance and has cardinality at least $(2N)^{-1/2} e^{\pi N/4^3}$: see estimate (1.6). Thus, by (4.3), we can select M distinct elements of the set F'. Each of these specifies a way of coloring a row of the family $\mathcal{U}$; the order of the colors inside each row being arbitrary. This gives the desired coloring of the family $\mathcal{U}$.

Third step. We will find a smooth conservative diffeomorphism h of the annulus that maps each $\mathcal{U}$-box to a $\mathcal{L}$-box of the same color by means of a translation, and which equals the identity near the boundary of the annulus. Essentially, this diffeomorphism exists because for each color k, there are *at most* M $\mathcal{U}$-boxes of color k (at most one box for each row), while there are *exactly* $M = m \cdot n$ $\mathcal{L}$-boxes of color k. Let us construct h precisely.

We index the members of the family $\mathcal{U}$ as U_1, U_2, $\dots$, $U_{NM/2}$ in such a way that U_1, $\dots$, $U_{N/2}$ form the bottom row, $U_{N/2+1}$, $\dots$, U_N form the second from bottom row, and so on. Then we select distinct $\mathcal{L}$-boxes L_1, L_2, $\dots$, $L_{NM/2}$ in such a way that each L_i has the same color as U_i, and whenever L_i and L_j have the same color and $i < j$ then L_i is not above L_j.

For each $i = 1,\dots,NM/2$, we will choose a smooth path $u_i\colon [0,1] \to \mathbb{R}^2$ starting from $u_i(0) = 0$ such that $t \in [0,1] \mapsto B_i(t) := U_i + u_i(t)$ is a well-defined path of boxes in $\mathbb{A}$, starting at $B_i(0) = U_i$ and finishing at $U_i(1) = L_i$. We require the path of boxes $P_i := \bigcup_{t\in[0,1]} B_i(t)$ to be disjoint from the set

$$\partial\mathbb{A} \cup \bigcup_{j<i} L_j \cup \bigcup_{j>i} U_j \,. \tag{4.5}$$

These paths can be taken as follows: we start with the box U_i and move it always either directly downwards or horizontally (like a Tetris piece). Note that $\frac{1}{40n} > \frac{2\kappa}{N}$ (since $n \geqslant 3$); this means that the gaps between the squares of $\mathcal{G}$ are greater than the width of the box. Therefore it is possible to move between gaps and reach the destination L_i avoiding the obstacle set (4.5): see Fig. 1.

Let $\varphi_i\colon \mathbb{A} \to [0,1]$ be a smooth function that equals 1 on the set P_i and equals 0 outside a small neighborhood of it (which is still disjoint from the set (4.5)). Now, writing $u_i(t) =: (v_i(t), w_i(t))$, define a (non-autonomous)

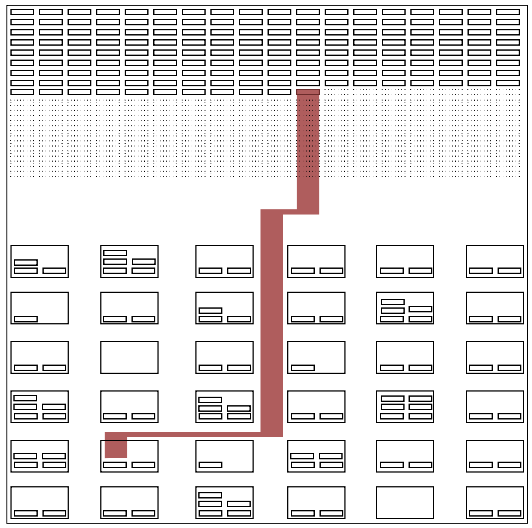

FIGURE 1. A path P_i that avoids the obstacle set (4.5).

Hamiltonian $H_i\colon \mathbb{A} \times [0,1] \to \mathbb{R}$ by:

$$H_i(\theta,\rho,t) := \varphi_i(\theta,\rho)\left[w_i'(t)\rho - v_i'(t)\theta\right].$$

Let $f_i \in \mathit{Diff}^{\infty}_{\mathrm{Leb}}(\mathbb{A})$ be the time one map of the associated Hamiltonian flow. Then f_i translates the box U_i to the box L_i, and equals the identity on the set (4.5). It follows that the diffeomorphism

$$h := f_{NM/2} \circ \cdots \circ f_2 \circ f_1$$

translates each box U_i to the corresponding L_i, and equals the identity on a neighborhood of $\partial\mathbb{A}$.

Fourth step. Let $\omega\colon [0,1] \to \mathbb{R}$ be any smooth function without critical points. Consider the conservative flow $\Psi^t := h \circ R^t_\omega \circ h^{-1}$ (see Fig. 2). We will estimate the emergence of the time t maps from below, at an appropriate scale.

For each $\rho \in [0,1]$, let λ_ρ denote Lebesgue measure on the circle $\mathbb{T} \times \{\rho\}$. Recall from (4.2) that for every $t \neq 0$, the ergodic decomposition of Lebesgue measure on $\mathbb{A}$ with respect to R^t_ω is $\mathbf{e}^{R^t}_*(\mathrm{Leb}) = \int_0^1 \delta_{\lambda_\rho}\, d\rho$, and in particular it is independent of t. It follows that the ergodic decomposition of Lebesgue with respect to $\Psi^t := h \circ R^t_\omega \circ h^{-1}$ is:

$$\hat{\mu} := \mathbf{e}^{\Psi^t}_*(\mathrm{Leb}) = \int_0^1 \delta_{\tilde{\lambda}_\rho}\, d\rho \quad \text{where} \quad \tilde{\lambda}_\rho := h_*(\lambda_\rho)\,. \tag{4.6}$$

We need to estimate the quantization number of this measure.

Let J be the set of $\rho \in [0,1]$ such that the circle $\mathbb{T} \times \{\rho\}$ intersects the boxes of the family $\mathcal{U}$; then J is a disjoint union of intervals $J_1, \ldots, J_M$, each of them of length $\frac{1}{11M}$.

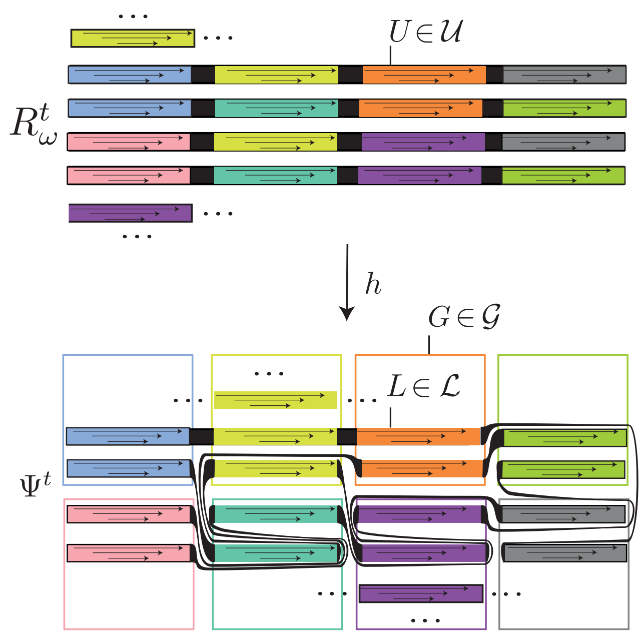

FIGURE 2. The flow $(\Psi^t)_t$.

Claim 4.3. If ρ, $\rho' \in J$ belong to the same interval J_i then:

$$\mathsf{W}_1(\tilde{\lambda}_\rho, \tilde{\lambda}_{\rho'}) \leqslant \frac{1}{11M} + 2\eta\,.$$

Proof of the claim. Fix ρ, $\rho' \in J_i$. We will use the following bound:

$$\mathsf{W}_1(\tilde{\lambda}_\rho, \tilde{\lambda}_{\rho'}) \leqslant \int_{\mathbb{T}} \mathsf{d}\big(h(\theta,\rho), h(\theta,\rho')\big)\, d\theta\,.$$

Indeed, the right hand side is the cost of transporting each point $h(\theta,\rho)$ to $h(\theta,\rho')$. Let $I \subset \mathbb{T}$ be the union of the projections of the $\mathcal{U}$-boxes on the first coordinate; this is a union of $\frac{N}{2}$ intervals of length $\frac{2\kappa}{N}$. Note that:

$$\mathsf{d}\big(h(\theta,\rho), h(\theta,\rho')\big) \leqslant \begin{cases} 1/(11M) & \text{if } \theta \in I, \\ \operatorname{diam} \mathbb{A} \leqslant 2 & \text{otherwise;} \end{cases}$$

indeed if $\theta \in I$ then both points (θ,ρ) and (θ,ρ') belong to the same $\mathcal{U}$-box U, which has height $\frac{1}{11M}$ and furthermore $h|_U$ is an isometry. Finally, using the fact that $\mathrm{Leb}(I^\mathsf{c}) = 1 - \kappa = \eta$, we obtain the asserted upper bound for the Wasserstein distance. $\square$

Claim 4.4. If $\rho \in J_i$, $\rho' \in J_j$ with $i \neq j$ then:

$$\mathsf{W}_1(\tilde{\lambda}_\rho, \tilde{\lambda}_{\rho'}) \geqslant \frac{1-3\eta}{80n}\,.$$

Proof of the claim. Fix $\rho \in J_i$, $\rho' \in J_j$ with $i \neq j$. Let $\mathcal{R}'$ be the family of $\mathcal{U}$-boxes that intersect the circle $\mathbb{T} \times \{\rho'\}$ (that is, a row of boxes). Let

$\mathcal{R}$ be the family of $\mathcal{U}$-boxes that intersect the circle $\mathbb{T} \times \{\rho\}$ and whose colors are distinct from those of the $\mathcal{R}'$-boxes. By construction, the family $\mathcal{R}$ contains at least $N/4$ boxes; let E be their union. Since $\lambda_\rho(U) = 2\kappa/N$ for each $\mathcal{R}$-box U, we have $\lambda_\rho(E) \geqslant \kappa/2 = (1-\eta)/2$. So $F := h(E)$ satisfies $\tilde{\lambda}_\rho(F) \geqslant (1-\eta)/2$. The set F is contained in the union of the $\mathcal{G}$-squares whose colors appear in the family $\mathcal{R}$. Recall that $\frac{1}{40n}$ is the minimal separation between $\mathcal{G}$-squares, so if V is the open $\frac{1}{40n}$-neighborhood of F then V does not intersect any $\mathcal{G}$-square with other colors. In particular, $\mathcal{R}'$-boxes are disjoint from $h^{-1}(V)$. Since the union of $\mathcal{R}'$-boxes has $\lambda_{\rho'}$ measure equal to $\kappa = 1-\eta$, it follows that $\tilde{\lambda}_{\rho'}(V) = \lambda_{\rho'}(h^{-1}(V)) \leqslant \eta$.

Consider an arbitrary transport plan π from $\tilde{\lambda}_\rho$ to $\tilde{\lambda}_{\rho'}$. Then:

$$\pi(F \times V^{\mathsf{c}}) \geqslant \pi(F \times \mathbb{A}) - \pi(\mathbb{A} \times V) = \tilde{\lambda}_\rho(F) - \tilde{\lambda}_{\rho'}(V) \geqslant \frac{1-\eta}{2} - \eta = \frac{1-3\eta}{2},$$

and so:

$$\operatorname{cost}(\pi) = \int_{\mathbb{A}\times\mathbb{A}} \mathsf{d}(x,y)\, d\pi(x,y) \geqslant \int_{F\times V^{\mathsf{c}}} (\cdots) \geqslant \frac{1}{40n}\pi(F \times V^{\mathsf{c}}) \geqslant \frac{1-3\eta}{80n}.$$

Since this estimate holds for all transport plans π, we obtain the asserted lower bound for the Wasserstein distance. □

For every $i \in \{1, \dots, M\}$, let us fix a point ρ_i in the interval J_i. The following measures $\hat{\mu}_1, \hat{\nu} \in \mathcal{M}(\mathcal{M}(\mathbb{A}))$ correspond respectively to the ergodic decomposition of $\Psi^t | h(\mathbb{T} \times J)$ and to the probability measure equidistributed on the set $\{\tilde{\lambda}_{\rho_i} \,;\, 1 \leqslant i \leqslant M\}$:

$$\hat{\mu}_1 := 11 \int_J \delta_{\tilde{\lambda}_\rho}\, d\rho \in \mathcal{M}(\mathcal{M}(\mathbb{A})) \quad \text{and} \quad \hat{\nu} := \frac{1}{M}\sum_{i=1}^{M} \delta_{\tilde{\lambda}_{\rho_i}}. \tag{4.7}$$

It follows from Claim 4.3 that $\mathsf{W}_1(\hat{\nu}, \hat{\mu}_1) \leqslant \frac{1}{11M} + 2\eta$; indeed each $\delta_{\tilde{\lambda}_\rho}$ with $\rho \in J_i$ can be transported to $\delta_{\tilde{\lambda}_{\rho_i}}$ at a cost no greater than $\frac{1}{11M} + 2\eta$.

On the other hand, by Claim 4.4, the measure $\hat{\nu}$ is equidistributed on a $\frac{1-3\eta}{80n}$-separated set of cardinality M. So, by Lemma 3.19, the W_1-distance from $\hat{\nu}$ to any probability measure supported on $M/2$ points is bigger than $\frac{1-3\eta}{320n}$. Therefore the W_1-distance from $\hat{\mu}_1$ to any probability measure supported on $M/2$ points is bigger than:

$$\frac{1-3\eta}{320n} - \left(\frac{1}{11M} + 2\eta\right) =: 11\epsilon. \tag{4.8}$$

In other words, $Q_{\hat{\mu}_1}(11\epsilon) \geqslant \frac{M}{2}$. By definitions (4.6), (4.7), we have $\hat{\mu} \geqslant \frac{1}{11}\hat{\mu}_1$, and so Lemma 3.18 yields $Q_{\hat{\mu}}(\epsilon) \geqslant \frac{M}{2}$. This quantization number is the metric emergence (by Proposition 3.12), so we obtain:

$$\mathscr{E}_{\mathrm{Leb}}(\Psi^t)(\epsilon) \geqslant \frac{M}{2} \quad \text{for all } t \neq 0. \tag{4.9}$$

Conclusion. If $\epsilon = \epsilon_n$ is defined by (4.8) then $\epsilon_n \asymp n^{-1} \asymp \epsilon_{n+1}$. For every sufficiently small number $\epsilon_* > 0$, we can find $n \geqslant 3$ such that $\epsilon_{n+1} < \epsilon_* \leqslant \epsilon_n$. If $(\Phi^t)_t = (\Phi^t_n)_t$ is the flow constructed above, then for every $t \neq 0$ we have:

$$\begin{aligned} \log \mathscr{E}_{\mathrm{Leb}}(\Psi^t)(\epsilon_*) \geqslant \log \mathscr{E}_{\mathrm{Leb}}(\Psi^t)(\epsilon_n) &\geqslant \log \frac{M}{2} && \text{(by (4.9))} \\ &\asymp n^2 && \text{(by (4.4))} \\ &\asymp \epsilon_n^{-2} \asymp \epsilon_*^{-2} . \end{aligned}$$

This ends the proof of Proposition 4.2. □

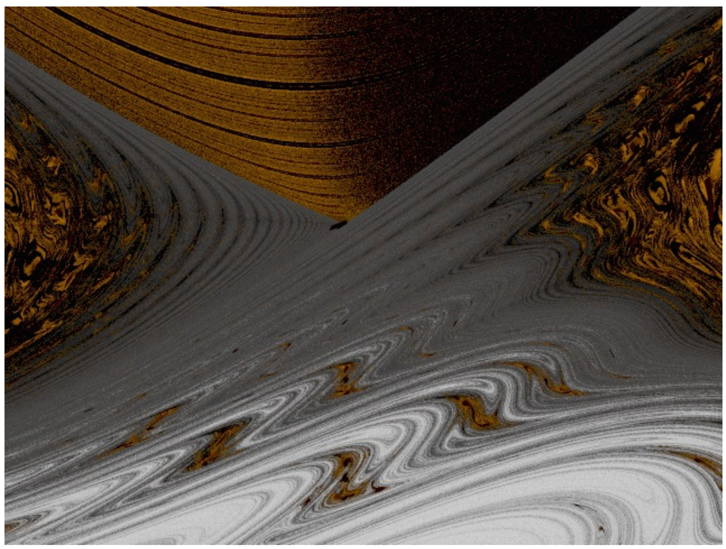

FIGURE 3. Interestingly, the geometry of the construction depicted at the bottom of Fig. 2 looks like the pattern obtained at the boundary of the so-called stochastic sea of the standard map for a certain parameter (depicted in gray in the above numerical experimentation).

4.3. **Construction of a smooth conservative flow with high emergence at every scale.** In this subsection, we will prove Theorem B. The main ingredient is Proposition 4.2.

Proof of Theorem B. Let us assume that the space $\mathcal{M}(\mathbb{A})$ is metrized with Wasserstein metric W_1. We will construct a conservative flow $(\Phi^t)_t$ on the annulus $\mathbb{A}$ whose metric emergence with respect to W_1 is stretched exponential with exponent 2. In the end of the proof we will see that the same holds if emergence is computed with respect to other Wasserstein metrics W_p or the Lévy–Prokhorov metric LP.

For each $i \geqslant 1$, let $h_i \in \mathit{Diff}^{\infty}_{\mathrm{Leb}}(\mathbb{A})$ be given by Proposition 4.2 for $\epsilon_* = \epsilon_i := 2^{-i^2}$. We define a smooth diffeomorphism between the annuli $\mathbb{A}$ and

$\mathbb{A}_i := \mathbb{T} \times [2^{-i}, 2^{-i+1}]$ as follows:

$$g_i \colon (\theta, \rho) \in \mathbb{A} \mapsto (\theta, 2^{-i}(\rho + 1)) \in \mathbb{A}_i \, .$$

Let h be the homeomorphism of the annulus $\mathbb{A}$ such that for each $i \geqslant 1$,

$$h(\mathbb{A}_i) = \mathbb{A}_i \quad \text{and} \quad h|\mathbb{A}_i := g_i \circ h_i \circ g_i^{-1} \, .$$

Since each g_i has constant jacobian, h is conservative. Furthermore, h equals the identity on the boundary of the annulus and is smooth on its interior. Next, fix a smooth function $r \colon \mathbb{R} \to \mathbb{R}$ that vanishes on $(-\infty, 0] \cup [1, +\infty)$ and is positive on the interval $(0, 1)$. Let (η_i) be a sequence of positive numbers converging very rapidly to 0. Define a function $\zeta \colon [0, 1] \to \mathbb{R}$ by:

$$\zeta(\rho) := \begin{cases} \eta_i \cdot r(2^i \rho - 1) & \text{if } \rho \in [2^{-i}, 2^{-i+1}]; \\ 0 & \text{if } \rho = 0. \end{cases}$$

If η_i tends to 0 sufficiently rapidly then ζ becomes a smooth function. Furthermore, it is positive Lebesgue a.e. Let $\omega \colon [0, 1] \to \mathbb{R}$ be the smooth function such that $\omega(0) = \omega'(0) = 0$ and $\omega'' = \zeta$. Observe that ω' is strictly positive on $(0, 1]$. Furthermore, the C^i-norm of $\omega|[2^{-i}, 2^{-i+1}]$ is small when $(\eta_j)_{j \geqslant i}$ is small. Hence we can choose inductively η_i sufficiently small so that the push forward of the vector field $\partial_t R^t_\omega$ by h, namely

$$h_* \partial_t R^t_\omega : (\theta, \rho) \mapsto \left[\omega \circ p_2 \circ h^{-1}(\theta, \rho)\right] \cdot \partial_\theta h \circ h^{-1}(\theta, \rho) \, ,$$

has C^i-norm restricted to each $\mathbb{A}_i$ smaller than 1 for every i.

Thus this vector field and its flow $\Phi^t = h \circ R^t_\omega \circ h^{-1}$ are smooth. The construction of the smooth conservative flow $(\Phi^t)_t$ is completed, and is depicted in Fig. 4. We are left to show that the flow has stretched exponential emergence with exponent 2. Of course, the homeomorphism h cannot be smooth on the whole annulus, because otherwise the flow would have polynomial emergence (by Proposition 4.1 and lemmas from Section 3.4).

Let $\hat{\mu} \in \mathcal{M}(\mathcal{M}(\mathbb{A}))$ denote the ergodic decomposition of Leb with respect to Φ^t (which is indeed independent of $t \neq 0$, since it is given by formula (4.6)). Recalling Proposition 3.12, we have $\mathscr{E}_{\mathrm{Leb}}(\Phi^t)(\epsilon) = Q_{\hat{\mu}}(\epsilon)$ for every ϵ.

Claim 4.5. For every $i \geqslant 1$ and $t \neq 0$ we have $\mathscr{E}_{\mathrm{Leb}}(\Phi^t)(\epsilon_{i+1}) \geqslant \exp(C\epsilon_i^{-2})$, where $C > 0$ is a constant.

Proof of the claim. Let $\mu_i := 2^i \,\mathrm{Leb}|\mathbb{A}_i = g_{i*}(\mathrm{Leb})$; this is a Φ^t-invariant probability measure. Its ergodic decomposition $\hat{\mu}_i$ is bounded from above by $2^i \hat{\mu}$ and so, by Lemma 3.18, for all $\epsilon > 0$ we have:

$$Q_{\hat{\mu}}(\epsilon) \geqslant Q_{\hat{\mu}_i}(2^i \epsilon).$$

Let $\omega_i := \omega \circ g_i$, and consider the conservative flow $\Psi^t_i := h_i \circ R^t_{\omega_i} \circ h_i^{-1}$ and its ergodic decomposition $\hat{\nu}_i := \mathbf{e}_*^{\Psi^t_i}(\mathrm{Leb})$ (for $t \neq 0$). Note that there exists a 1-Lipschitz retraction $p_i \colon \mathbb{A} \to \mathbb{A}_i$. Let $q_i \colon \mathbb{A} \to \mathbb{A}$ be the map $q_i := g_i^{-1} \circ p_i$. By Lemma 3.11 we have $\hat{\nu}_i = q_{i**}(\hat{\mu}_i)$, that is, $\hat{\nu}_i$ is the push-forward of

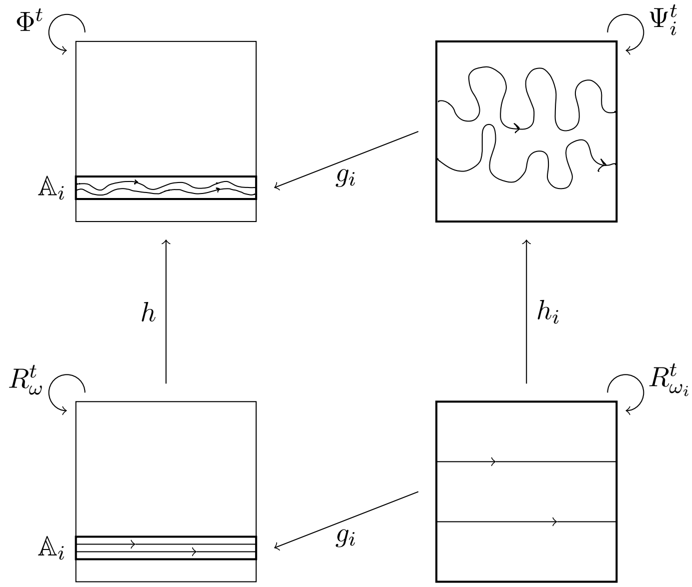

FIGURE 4. Proof of Theorem B.

$\hat{\mu}_i$ under the map $q_{i*}\colon \mathcal{M}(\mathbb{A}) \to \mathcal{M}(\mathbb{A})$. Since q_i is 2^i-Lipschitz, so is q_{i*}. Hence, by Lemma 3.17,

$$Q_{\hat{\mu}_i}(2^i\epsilon) \geqslant Q_{\hat{\nu}_i}(2^{2i}\epsilon)\,.$$

In summary, we have shown that $Q_{\hat{\mu}}(\epsilon) \geqslant Q_{\hat{\nu}_i}(2^{2i}\epsilon)$, that is,

$$\mathscr{E}_{\mathrm{Leb}}(\Phi^t)(\epsilon) \geqslant \mathscr{E}_{\mathrm{Leb}}(\Psi_i^t)(2^{2i}\epsilon)\,.$$

Taking $\epsilon = 2\epsilon_{i+1}$ and noting that $2^{2i}\epsilon = \epsilon_i$, we obtain:

$$\mathscr{E}_{\mathrm{Leb}}(\Phi^t)(\epsilon_{i+1}) \geqslant \mathscr{E}_{\mathrm{Leb}}(\Phi^t)(2\epsilon_{i+1}) \geqslant \mathscr{E}_{\mathrm{Leb}}(\Psi_i^t)(\epsilon_i) \geqslant \exp(C\epsilon_i^{-2})\,,$$

where the last inequality is the main property of the flow $(\Psi_i^t)_t$, coming from Proposition 4.2. This proves the claim. $\square$

Next, we claim that:

$$\liminf_{\epsilon\to 0} \frac{\log\log Q_{\hat{\mu}}(\epsilon)}{-\log\epsilon} \geqslant 2\,. \tag{4.10}$$

Indeed, given a small $\epsilon > 0$, let i be such that $\epsilon \in [\epsilon_{i+2}, \epsilon_{i+1}]$. We have $Q_{\hat{\mu}}(\epsilon) \leqslant Q_{\hat{\mu}}(\epsilon_{i+1})$ and so, using Claim 4.5,

$$\frac{\log\log Q_{\hat{\mu}}(\epsilon)}{-\log\epsilon} \geqslant \frac{\log\log Q_{\hat{\mu}}(\epsilon_{i+1})}{-\log\epsilon_{i+2}} \geqslant \frac{\log C - 2\log\epsilon_i}{-\log\epsilon_{i+2}}\,.$$

The right-hand side tends to 2 as $i \to \infty$, so (4.10) follows.

Inequality (4.10) means that the lower quantization order of $\hat{\mu}$ is at least 2, that is, $\underline{\mathrm{qo}}(\hat{\mu}) \geqslant 2$. Up to this moment we were assuming that $\mathcal{M}(\mathbb{A})$ is metrized with Wasserstein metric W_1, but now let us use any Wasserstein metric W_p, $1 \leqslant p < \infty$, or the Lévy–Prokhorov metric LP. By inequalities

(1.2), (1.3) from p. 10, we still have $\underline{\mathrm{qo}}(\hat{\mu}) \geqslant 2$ with respect to the other metrics. On the other hand, by Proposition 3.4, Theorem 1.3, and Remark 1.5, we have:

$$\overline{\mathrm{qo}}(\hat{\mu}) \leqslant \overline{\mathrm{mo}}(\mathcal{M}(\mathbb{A})) \leqslant \overline{\dim}(\mathbb{A}) = 2\,.$$

Therefore $\mathrm{qo}(\hat{\mu}) = 2$. This means exactly that:

$$\lim_{\epsilon \to 0} \frac{\log\log \mathscr{E}_{\mathrm{Leb}}(\Phi^t)(\epsilon)}{-\log \epsilon} = 2\,,$$

which completes the proof of Theorem B. ☐

5. Genericity of high emergence

In this section, we will prove our Theorems C and D on the genericity of high emergence among surface diffeomorphisms. Both proofs are based on Proposition 4.2. Another fundamental tool is the creation of *periodic spots*; we recall the relevant results in Section 5.1. Furthermore, for the conservative Theorem C, we also need a KAM theorem, which is discussed along with some of its consequences in Section 5.2.

Throughout this section, let (M, Leb) *be a compact surface endowed with a normalized smooth volume (i.e. area) measure.*

5.1. **Creation of periodic spots.**

Theorems C and D are proved using the following concept:

Definition 5.1 (Periodic spot)**.** An open subset $O \subset M$ is a *periodic spot* for a continuous self-map f of M if there exists $p \geqslant 1$ such that $f^p(O) = O$ and the restriction $f^p|O$ is the identity map on O.

Diffeomorphisms displaying a periodic spot appears densely in many open subsets of dynamical systems. In the conservative setting we have:

Theorem 5.2 ([GTS, Thrm. 5] and [GeT, Thrm. 1])**.** *For every* $r \in [1, \infty]$*, if* $\mathcal{U}$ *is the open subset of* $\mathit{Diff}^r_{\mathrm{Leb}}(M)$ *formed by dynamics having an elliptic periodic point, then there exists a dense subset* $\mathcal{D} \subset \mathcal{U}$ *formed by dynamics displaying a periodic spot.*

The statement above follows from the combination of [GTS, Thrm. 5] and [GeT, Thrm. 1].

In the dissipative setting we have:

Theorem 5.3 (Turaev, [Tu, Lemma 2])**.** *For every* $r \in [2, \infty]$*, there exists a non-empty open set* $\mathcal{U} \subset \mathit{Diff}^r(M)$ *and a dense set* $\mathcal{D} \subset \mathcal{U}$ *formed by dynamics displaying a periodic spot.*

We add that the set $\mathcal{U}$ in Theorem 5.3 contains the absolute Newhouse domain: it is formed by diffeomorphisms displaying a horseshoe having a robust homoclinic tangency, a volume expanding periodic point, and a volume contracting periodic point.

5.2. **KAM and stability of high emergence.** A *twist map* is a conservative diffeomorphism $f_0 \colon \mathbb{A} \to \mathbb{A}$ of the form $f_0 = R^t_\omega$, where $\omega \colon [0,1] \to \mathbb{R}$ is a smooth function without critical points, and $t \neq 0$.

Theorem 5.4 (Moser–Pöschel's twist mapping theorem [Mo], [Pö], [BrS, § 3.2.1])**.** *Let f be a twist map. Fix a number $\eta > 0$ a neighborhood $\mathcal{U}$ of the identity map in $\mathit{Diff}^\infty(\mathbb{A})$. Then there exist a closed subset $D \subset (0,1)$ with Lebesgue measure at least $1 - \eta$ and a neighborhood $\mathcal{V}$ of f in $\mathit{Diff}^\infty_{\mathrm{Leb}}(\mathbb{A})$ such that for every $g \in \mathcal{V}$, there exists $h \in \mathcal{U}$ such that the map $f|\mathbb{T} \times D$ is conjugate to $g|h(\mathbb{T} \times D)$ via h:*

$$g \circ h(z) = h \circ f(z), \quad \forall z \in \mathbb{T} \times D\,.$$

As a corollary of Theorem 5.4, we will prove below:

Corollary 5.5. *Let $f \in \mathit{Diff}^\infty_{\mathrm{Leb}}(M)$ be a conservative surface diffeomorphism that acts as a twist map on a embedded annulus $A \subset M$; more precisely, assume that there exist a smooth embedding $h_1 \colon \mathbb{A} \to M$ with constant jacobian and image $h_1(\mathbb{A}) = A$ and a twist map $f_0 \colon \mathbb{A} \to \mathbb{A}$ such that $f \circ h_1 = h_1 \circ f_0$. Then for every $\epsilon_1 > 0$, for every $g \in \mathit{Diff}^\infty_{\mathrm{Leb}}(M)$ sufficiently close to f, there exists a g-invariant embedded sub-annulus $B \subset A$ such that* $\mathrm{Leb}(A \smallsetminus B) < \epsilon_1$ *and*

$$\mathsf{W}_1 \left(\mathbf{e}^g_*(\mu_B), \mathbf{e}^f_*(\mu_A) \right) < \epsilon_1\,,$$

where μ_A and μ_B are the normalized Lebesgue measures on A and B, respectively.

In the case that f itself is a twist map (so $M = A = \mathbb{A}$, $f = f_0$, and $h_1 = \mathrm{id}$), we can actually take $B = A$. In particular, f becomes a continuity point for the ergodic decomposition of Lebesgue measure (c.f. Remark 3.10).

We will use the following general estimate:

Lemma 5.6. *Let (X, d) be a compact metric space and μ, $\nu \in \mathcal{M}(X)$. If ν is absolutely continuous w.r.t. μ, with density $r := \frac{d\nu}{d\mu}$, then:*

$$\mathsf{W}_1(\nu, \mu) \leqslant (\operatorname{diam} X) \|r - 1\|_{L^1(\mu)}\,.$$

Proof. This is an immediate consequence of the Kantorovich–Rubinstein duality formula [Vil, p. 207]. ☐

Proof of Corollary 5.5. Let us first consider the simpler case where $f = f_0$ is a twist map on $M = A = \mathbb{A}$, and so $h_1 = \mathrm{id}$. Let $\eta > 0$ be small, and let $\mathcal{U}$ be a small neighborhood of the identity map in $\mathit{Diff}^\infty(\mathbb{A})$. We apply Theorem 5.4, obtaining a set $D \subset [0,1]$ and a neighborhood $\mathcal{V}$ of the twist map f in $\mathit{Diff}^\infty_{\mathrm{Leb}}(\mathbb{A})$. Take an arbitrary $g \in \mathcal{V}$. We need to prove that the ergodic decompositions of Lebesgue measure with respect to f and g are approximately the same.

Consider the (non-conservative) diffeomorphism $\tilde{f} := h^{-1} \circ g \circ h$; by construction it equals f on $\mathbb{T} \times D$. The measure $\nu := h^{-1}_*(\mathrm{Leb})$ is $\tilde{f}$-invariant.

Let $c := \nu(\mathbb{T} \times D)$ (a number close to 1), and let $\nu_1 := c^{-1} \cdot \nu|\mathbb{T} \times D$. Then ν_1 is $\tilde{f}$-invariant. Since $\tilde{f}$ equals f on $\operatorname{supp} \nu_1 = \mathbb{T} \times D$, this measure is also f-invariant. We will prove that the four ergodic decompositions below are close to each other:

$$①:= \mathbf{e}^g_*(\mathrm{Leb}), \quad ②:= \mathbf{e}^{\tilde{f}}_*(\nu), \quad ③:= \mathbf{e}^{\tilde{f}}_*(\nu_1) = \mathbf{e}^f_*(\nu_1), \quad ④:= \mathbf{e}^f_*(\mathrm{Leb}).$$

Since h is close to identity, the distances $\mathsf{d}(h(x), x)$ are uniformly bounded by a small constant ϵ_2. Therefore:

$$\forall \xi \in \mathcal{M}(\mathbb{A}), \quad \mathsf{W}_1(h_*\xi, \xi) \leqslant \int \mathsf{d}(h(x), x)\, d\xi(x) \leqslant \epsilon_2 \,,$$

That is, $h_* \colon (\mathcal{M}(\mathbb{A}), \mathsf{W}_1) \to (\mathcal{M}(\mathbb{A}), \mathsf{W}_1)$ is ϵ_2-close to the identity map. Repeating the argument, we see that h_{**} is also close to the identity. By Lemma 3.11, $h_{**}(②) = ①$; this proves that the measures ① and ② are close.

The Radon–Nikodym derivative $r := \frac{d\nu}{d\mathrm{Leb}}$ is smooth and uniformly close to 1. We have:

$$\frac{d\nu_1}{d\nu} = c^{-1}\, \mathbf{1}_{\mathbb{T}\times D} \,.$$

(Here $\mathbf{1}$ denotes characteristic function.) On the other hand, for all $(\theta, \rho) \in \mathbb{A}$, the empirical measure $\mathbf{e}^f(\theta, \rho)$ is Lebesgue on the circle $\mathbb{T} \times \{\rho\}$, denoted λ_ρ. It follows that the ergodic decomposition of Leb and ν_1 with respect to f are:

$$④ = \mathbf{e}^f_*(\mathrm{Leb}) = \int_0^1 \delta_{\lambda_\rho}\, d\rho \,,$$

and

$$③ = \mathbf{e}^f_*(\nu_1) = c^{-1} \int_D \bar{r}(\rho)\, \delta_{\lambda_\rho}\, d\rho \,,$$

where $\bar{r}(\rho) := \int r \, d\lambda_\rho$. So ③ is absolutely continuous with respect to ④, with density:

$$\frac{d③}{d④}(\delta_{\lambda_\rho}) = c^{-1}\bar{r}(\rho)\mathbf{1}_D(\rho) \,.$$

This function is close to 1 in $L^1(④)$; so Lemma 5.6 implies that the measures ③ and ④ are close. Finally, we have:

$$\frac{d③}{d②} = c^{-1}\, \mathbf{1}_S \,, \quad \text{where } S := \operatorname{supp} ③ = \{\delta_{\lambda_\rho} \,;\, \rho \in D\} \,.$$

This function is close to 1 in $L^1(②)$, so Lemma 5.6 implies that the measures ② and ③ are close. The upshot is that $① = \mathbf{e}^g_*(\mathrm{Leb})$ and $④ = \mathbf{e}^f_*(\mathrm{Leb})$ are close. This completes the proof of the corollary in the case f is a twist map.

The general situation can be reduced to the previous case. Indeed, Theorem 5.4 also ensures that if $f \colon M \to M$ acts as a twist map on a annulus A, then any perturbation g of f admits a g-invariant sub-annulus $B \subset A$ which is close to A. Then the proof is verbatim the same by substituting the measures λ_ρ by their pushforward by h_1, g by $h_1 \circ g \circ h_1^{-1}$, and h by $h_1 \circ h \circ h_1^{-1}$. $\square$

As another corollary of Theorem 5.4, we obtain open sets with at least polynomial emergence, so justifying an assertion made in Section 4.1. (Readers anxious to see the proof of Theorem C may skip this.)

Corollary 5.7. *Under the same hypotheses as Corollary 5.5, there exists* $C > 0$ *such that for every* $g \in \mathit{Diff}^{\infty}_{\mathrm{Leb}}(M)$ *sufficiently close to* f*, its emergence of* g *with respect to the* W_1 *metric satisfies:*

$$\mathscr{E}_{\mathrm{Leb}}(g)(\epsilon) \geqslant C\epsilon^{-1}\,, \quad \forall \epsilon > 0\,,$$

Note that this is *not* a consequence of Corollary 5.5 by itself, since we bound the emergence of the perturbations at *every* scale.

Proof. We will provide a proof in the case that f itself is a twist map (so $M = A = B = \mathbb{A}$, $f = f_0$, and $h_1 = \mathrm{id}$), leaving for the reader to adapt the proof for the general situation.

Let $g \in \mathit{Diff}^{\infty}_{\mathrm{Leb}}(\mathbb{A})$ be a perturbation of f. Applying Theorem 5.4, we obtain $h \in \mathit{Diff}^{\infty}(\mathbb{A})$ close to identity such that $g \circ h = h \circ f$ on $\mathbb{T} \times D$, where $D \subset (0,1)$ is a closed set with almost full measure; say at least $1/2$. We can assume that $h^{\pm 1}$ are 2-Lipschitz and have jacobian at most 2. As in the proof of Corollary 5.5, let $\tilde f := h^{-1} \circ g \circ h$ and $\nu := h_*^{-1}(\mathrm{Leb})$; then ν is $\tilde f$-invariant.

Consider the ergodic decompositions $\hat\mu := \mathbf{e}^g_*(\mathrm{Leb})$ and $\hat\nu := \mathbf{e}^{\tilde f}_*(\nu)$. Then, for arbitrary $\epsilon > 0$,

$$\begin{aligned}
\mathscr{E}_{\mathrm{Leb}}(g)(\epsilon) &= Q_{\hat\mu}(\epsilon) && \text{(by Proposition 3.12)}\\
&\geqslant Q_{h^{-1}_{**}(\hat\mu)}(\mathrm{Lip}(h_*^{-1})\epsilon) && \text{(by Lemma 3.17)}\\
&\geqslant Q_{\hat\nu}(2\epsilon) && \text{(by Lemmas 3.11 and 3.16)}\,.
\end{aligned}$$

Note that $\nu \geqslant \frac{1}{2}\mathrm{Leb}$, by the bound on the jacobian. Let μ_1 be the normalized Lebesgue measure on $\mathbb{T} \times D$. Since $\mathrm{Leb}(D) \geqslant \frac{1}{2}$ we have $\mu_1 \leqslant 2\mathrm{Leb}$, and so $\nu \geqslant \frac{1}{4}\mu_1$. Furthermore, μ_1 is also $\tilde f$-invariant so the ergodic decomposition $\hat\mu_1 := \mathbf{e}^{\tilde f}_*(\nu)$ is well-defined. We have $\hat\nu \geqslant \frac{1}{4}\hat\mu_1$ and so, by Lemma 3.18,

$$\mathscr{E}_{\mathrm{Leb}}(\epsilon) \geqslant Q_{\hat\nu}(2\epsilon) \geqslant Q_{\hat\mu_1}(8\epsilon)\,.$$

Since $\tilde f$ equals f on the support of μ_1, we have:

$$\hat\mu_1 = \frac{1}{\mathrm{Leb}(D)} \int_D \delta_{\lambda_\rho}\, d\rho\,,$$

where λ_ρ denotes Lebesgue measure on the circle $\mathbb{T} \times \{\rho\}$. Similarly to the proof of Proposition 4.1, the measure $\hat\mu_1$ is supported on a set which is isometric to D under the isometry $\lambda_\rho \mapsto \rho$; moreover, the isometry carries $\hat\mu_1$ to the normalized Lebesgue measure on D (call it λ). Therefore $Q_{\hat\mu_1}(8\epsilon) = Q_\lambda(8\epsilon)$.

We are left to estimate the quantization number of the measure λ. Consider its distribution function $F\colon [0,1] \to [0,1]$ defined by $F(x) := \lambda([0,x])$.

Since $\lambda \leqslant 2\mathrm{Leb}$, the function F is 2-Lipschitz. Furthermore, $F_*(\lambda) = \mathrm{Leb}$. So, by Lemma 3.17,

$$Q_\lambda(8\epsilon) \geqslant Q_{\mathrm{Leb}}(16\epsilon)\,.$$

We have seen in Example 3.5 that quantization number of 1-dimensional Lebesgue measure is $Q_{\mathrm{Leb}}(\epsilon) \asymp \epsilon^{-1}$. We conclude that $\mathscr{E}_\mu(g)(\epsilon)$ is at least of the order of ϵ^{-1}, as we wanted to show. □

5.3. **Genericity of high emergence: conservative setting.** Here is a consequence of Proposition 4.2, combined with Corollary 5.5:

Lemma 5.8. *Suppose that $f \in \mathit{Diff}^\infty_{\mathrm{Leb}}(M)$ admits a periodic spot O. Let $\epsilon_0 > 0$, and let $\mathcal{U} \subset \mathit{Diff}^\infty_{\mathrm{Leb}}(M)$ be a neighborhood of f. Then there exists a nonempty open set $\mathcal{V} \subset \mathcal{U}$ such that for every $g \in \mathcal{V}$, its metric emergence w.r.t. W_1 metric satisfies:*

$$\sup_{\epsilon<\epsilon_0} \frac{\log\log \mathscr{E}_{\mathrm{Leb}}(g)(\epsilon)}{\log \epsilon} \geqslant 2 - \epsilon_0\,.$$

Proof. Assume that f has a periodic spot O. For simplicity of writing, let us assume that O consists of fixed points.

Let $\hat{\mathbb{A}} := \mathbb{T} \times [-1, 2]$. Take a smooth embedding $h_1 : \hat{\mathbb{A}} \to O$ with constant Jacobian J. Fix a small $\epsilon_* > 0$; how small it needs to be will become apparent at the end.

Let $\omega : \mathbb{R} \to \mathbb{R}$ be a smooth function that has no critical points in $[0, 1]$ and vanishes outside $[-1, 2]$. By Proposition 4.2, we can find $h \in \mathit{Diff}^\infty_{\mathrm{Leb}}(\mathbb{A})$ that equals the identity on the neighborhood of the boundary such that the maps $\Psi^t := h \circ R^t_\omega \circ h^{-1}$ has high emergence at scale ϵ_*:

$$\mathscr{E}(\Psi^t)(\epsilon_*) \geqslant \exp(C\epsilon_*^{-2})\,, \quad \forall t \neq 0\,,$$

where C is a constant. We fix $t \neq 0$ very close to 0 and write $\psi := \Psi^t$. We can extend h and ψ to smooth conservative diffeomorphisms $\hat{h}$ and $\hat{\psi}$ of the bigger annulus $\hat{\mathbb{A}}$, putting $\hat{h}(\theta, \rho) := (\theta, \rho)$ and $\hat{\psi}(\theta, \rho) := (\theta + t\omega(\rho), \rho)$ for $(\theta, \rho) \in \hat{\mathbb{A}} \smallsetminus \mathbb{A}$. Define $\tilde{f} : M \to M$ by:

$$\tilde{f}(x) := \begin{cases} h_1 \circ \hat{\psi} \circ h_1^{-1}(x) & \text{if } x \in h_1(\hat{\mathbb{A}}); \\ f(x) & \text{otherwise.} \end{cases}$$

Then $\tilde{f}$ is a smooth conservative diffeomorphism, and it is C^∞-close to f (since t is close to 0). So we can assume that $\tilde{f}$ belongs to the given neighborhood $\mathcal{U}$ of f. Note that $\tilde{f}$ acts as a twist map on the embedded annulus $A := h_1(\mathbb{A})$, which has measure J (the jacobian of h_1).

Let $g \in \mathit{Diff}^\infty_{\mathrm{Leb}}(M)$ be a small perturbation of $\tilde{f}$. By Corollary 5.5, g admits an invariant sub-annulus $B \subset A$ such that:

$$\mathrm{Leb}(A \smallsetminus B) < \tfrac{J}{2} \quad \text{and} \quad \mathsf{W}_1\left(\mathbf{e}^g_*(\mu_B), \mathbf{e}^{\tilde{f}}_*(\mu_A)\right) < L^{-1}\epsilon_*\,,$$

where μ_A (resp. μ_B) is the normalized Lebesgue measure on A (resp. B). Since $\mu_A = h_{1*}(\mathrm{Leb}_{\mathbb{A}})$, by Lemma 3.11, $\mathbf{e}_*^{\tilde{f}}(\mu_A) = h_{1**}(\mathbf{e}_*^{\psi}(\mathrm{Leb}_{\mathbb{A}}))$. Let L be the Lipschitz constant of h_1; then, by Lemmas 3.16 and 3.17,

$$Q_{\mathbf{e}_*^{\tilde{f}}(\mu_A)}(L^{-1}\epsilon_*) \geqslant Q_{\mathbf{e}_*^{\psi}(\mathrm{Leb}_{\mathbb{A}})}(\epsilon_*) \geqslant \exp(C\epsilon_*^{-2})\,.$$

It follows from Lemma 3.15 that:

$$Q_{\mathbf{e}_*^{g}(\mu_B)}(2L^{-1}\epsilon_*) \geqslant Q_{\mathbf{e}_*^{\tilde{f}}(\mu_A)}(L^{-1}\epsilon_*) \geqslant \exp(C\epsilon_*^{-2})\,.$$

Since $\mathrm{Leb}(B) \geqslant \frac{J}{2}$, we have $\mu_B \leqslant 2J^{-1}\mathrm{Leb}$ and so $\mathbf{e}_*^{g}(\mathrm{Leb}) \leqslant 2J^{-1}\mathbf{e}_*^{g}(\mu_B)$. It follows from Lemma 3.18 that:

$$Q_{\mathbf{e}_*^{g}(\mathrm{Leb})}(4L^{-1}J^{-1}\epsilon_*) \geqslant Q_{\mathbf{e}_*^{g}(\mu_B)}(2L^{-1}\epsilon_*) \geqslant \exp(C\epsilon_*^{-2})\,.$$

Let $\epsilon := 4L^{-1}J^{-1}\epsilon_*$. Since ϵ_* is very small, we conclude that $\epsilon < \epsilon_0$ and

$$\frac{\log\log \mathscr{E}_{\mathrm{Leb}}(g)(\epsilon)}{\log \epsilon} \geqslant 2 - \epsilon_0\,.$$

Therefore the neighborhood $\mathcal{V}$ of $\tilde{f}$ formed by the perturbations g has the required properties. $\square$

Proof of Theorem C. Consider the following two subsets of $\mathit{Diff}^{\infty}_{\mathrm{Leb}}(M)$:

- $\mathcal{W}$ is the set of weakly stable diffeomorphisms, i.e., those that robustly have only hyperbolic periodic points (if any);
- $\mathcal{U}$ is the set of diffeomorphisms that admit at least one elliptic periodic point.

These two sets are open and disjoint. Furthermore, since the periodic points of a generic area preserving map are either hyperbolic or elliptic, the union $\mathcal{W} \cup \mathcal{U}$ is dense in $\mathit{Diff}^{\infty}_{\mathrm{Leb}}(M)$.

By Theorem 5.2, there is a dense subset $\mathcal{D} \subset \mathcal{U}$ formed by diffeomorphisms displaying a periodic spot. For each $f \in \mathcal{D}$, let $(\mathcal{U}_{f,n})$ be a neighborhood basis for f. By Lemma 5.8, there exists a nonempty open subset $\mathcal{V}_{f,n} \subset \mathcal{U}_{f,n}$ such that:

$$\forall g \in \mathcal{V}_{f,n}\,, \quad \sup_{\epsilon < 1/n} \frac{\log\log \mathscr{E}_{\mathrm{Leb}}(g)(\epsilon)}{\log \epsilon} \geqslant 2 - \frac{1}{n}\,.$$

The set $\mathcal{O}_n := \bigcup_{f\in\mathcal{D}} \mathcal{V}_{f,n}$ is open and dense in $\mathcal{U}$. Then $\mathcal{R} := \bigcap_n \mathcal{O}_n$ is a residual subset of $\mathcal{U}$ which satisfies:

$$\forall g \in \mathcal{R}\,, \quad \limsup_{\epsilon\to 0} \frac{\log\log \mathscr{E}_{\mathrm{Leb}}(g)(\epsilon)}{\log \epsilon} \geqslant 2\,.$$

Thus $\mathcal{W} \cup \mathcal{R}$ is a residual subset of $\mathit{Diff}^{\infty}_{\mathrm{Leb}}(M)$ formed by diffeomorphisms that are either weakly stable or have lim sup stretched exponential emergence with exponent 2. $\square$

5.4. **Genericity of high emergence: dissipative setting.** Here is another consequence of Proposition 4.2:

Lemma 5.9. *Let $r \in [1, \infty]$. Suppose that $f \in \mathit{Diff}^r(M)$ admits a periodic spot O. Let $\epsilon_0 > 0$, and let $\mathcal{U} \subset \mathit{Diff}^\infty(M)$ be a neighborhood of f. Then there exists a nonempty open set $\mathcal{V} \subset \mathcal{U}$ such that for every $g \in \mathcal{V}$, its metric emergence w.r.t.* W_1 *metric satisfies:*

$$\sup_{\epsilon < \epsilon_0} \frac{\log \log \mathscr{E}_{\mathrm{Leb}}(g)(\epsilon)}{\log \epsilon} \geqslant 2 - \epsilon_0 \, .$$

Proof. Let us first consider the simpler case where M is the annulus and f is the identity map. Let $\epsilon_0 > 0$ be given. Fix a positive $\epsilon_* < \epsilon_0$ small enough such that:

$$\frac{\log C - 2 \log \epsilon_*}{\log 4 - \log \epsilon_*} \geqslant 2 - \epsilon_0 \, ,$$

where $C > 0$ is the constant from Proposition 4.2. Choose and fix a smooth function $\omega \colon [0, 1] \to \mathbb{R}$ without critical points. Applying Proposition 4.2, we obtain a smooth conservative diffeomorphism $h \colon \mathbb{A} \to \mathbb{A}$ that equals identity on a neighborhood of the boundary of the annulus, such that the flow $\Psi^t := h \circ R^t_\omega \circ h^{-1}$ has the following property:

$$\forall t \neq 0, \quad \mathscr{E}_{\mathrm{Leb}}(\Psi^t)(\epsilon_*) \geqslant \exp(C \epsilon_*^{-2}) \, .$$

For each $\rho \in [0, 1]$, let λ_ρ denote Lebesgue measure on the circle $\mathbb{T} \times \{\rho\}$, and let $\tilde{\lambda}_\rho := h_*(\lambda_\rho)$ be its push-forward under h. So $\tilde{\lambda}_\rho$ is supported on the curve $\mathcal{C}_\rho := h(\mathbb{T} \times \{\rho\})$. Consider the following sequence of elements of $\mathcal{M}(\mathcal{M}(\mathbb{A}))$:

$$\hat{\mu}_n := \frac{1}{n} \sum_{i=0}^{n-1} \delta_{\tilde{\lambda}_{(i+.5)/n}} \, .$$

Note that the sequence $(\hat{\mu}_n)$ tends to the measure $\hat{\mu}$ defined by (4.6), which is exactly the ergodic decomposition of any Ψ^t ($t \neq 0$). By Lemma 3.15, if n is large enough then $Q_{\hat{\mu}_n}(\epsilon_*/2)$ is at least $Q_{\hat{\mu}}(\epsilon_*)$, which by construction is at least $\exp(C\epsilon_*^{-2})$.

Let $((\Psi^t_n)_t)_n$ be a sequence of flows on the annulus $\mathbb{A}$ converging to $(\Psi^t)_t$ and such that, for each $n \geqslant 1$, the flow $(\Psi^t_n)_t$ satisfies:

- for every $i \in \{0, 1, \ldots, n\}$, the curve $\mathcal{C}_{i/n}$ is invariant and exponentially repelling;
- for every $i \in \{0, 1, \ldots, n-1\}$, the curve $\mathcal{C}_{(i+.5)/n}$ is invariant and exponentially attracting, with basin $h(\mathbb{T} \times (i/n, (i+1)/n))$;

Let $f_n := \Psi^{t_n}_n$, where (t_n) is a sequence of non-zero numbers tending to zero. Then f_n converges $f = \mathrm{id}$ in the C^∞ topology. Tweaking the sequence (t_n) if necessary, we can assume that each f_n acts as an irrational rotation on each attracting cycle $\mathcal{C}_{(i+.5)/n}$, $i \in \{0, 1, \ldots, n-1\}$. Then every point in the basin of $\mathcal{C}_{(i+.5)/n}$ has a well-defined empirical measure with respect to f_n, which is $\tilde{\lambda}_{(i+.5)/n}$. Each of these basins has Lebesgue measure $1/n$, so the measure

$\mathbf{e}_*^{f_n}(\mathrm{Leb})$, (which with some abuse of terminology we will call the ergodic decomposition of f_n) is well defined and equals $\hat{\mu}_n$. So for large enough n, the diffeomorphism f_n displays high emergence at scale $\epsilon_*/2$:

$$\mathscr{E}_{\mathrm{Leb}}(f_n)(\epsilon_*/2) = Q_{\hat{\mu}_n}(\epsilon_*/2) \geqslant \exp(C\epsilon_*^{-2}).$$

(Strictly speaking, Proposition 3.12 does not apply since Leb measure is not f_n-invariant, but it still works since the empirical measures are Leb-a.e. well defined and ergodic.) For the remainder of the proof, we fix a large n such that f_n has the above properties, and moreover belongs to the given neighborhood $\mathcal{U}$ of $f = \mathrm{id}$.

Now, if g is a small C^1-perturbation of f_n then by persistence of normally contracting submanifolds (see e.g. [BeB, Thm. 2.1]), g has n attracting curves C^1-close to the curves $\mathcal{C}_{(i+.5)/n}$, and their basins are bounded by repelling curves that are C^1-close to the curves $\mathcal{C}_{i/n}$. The rotation numbers along these attracting curves are either irrational or rational with a large denominator, so every point in the union of the basins has a well-defined empirical measure with respect to g, which is close to $\tilde{\lambda}_{(i+.5)/n}$. Thus g has a well-defined ergodic decomposition, which is close to $\hat{\mu}_n$. It follows from Lemma 3.15 that:

$$\mathscr{E}_{\mathrm{Leb}}(g)(\epsilon_*/4) = Q_{\mathbf{e}_*^g(\mathrm{Leb})}(\epsilon_*/4) \geqslant Q_{\hat{\mu}_n}(\epsilon_*/2) \geqslant \exp(C\epsilon_*^{-2}).$$

So it follows from the definition of ϵ_* that:

$$\epsilon := \frac{\epsilon_*}{4} \quad \Rightarrow \quad \frac{\log\log \mathscr{E}_{\mathrm{Leb}}(g)(\epsilon)}{\log \epsilon} \geqslant 2 - \epsilon_0 \,.$$

Letting $\mathcal{V}$ be a C^1-neighborhood of the diffeomorphism f_n where such estimates hold, we conclude the proof of the lemma in the case $M = \mathbb{A}$, $f = \mathrm{id}$.

If f is an arbitrary surface diffeomorphism admitting a periodic spot O, then we embed an annulus in O and reproduce the construction above. Emergences can be estimated from below similarly. Details are left for the reader. □

Proof of Theorem D. The proof is entirely analogous to the proof of Theorem C, using Theorem 5.3 instead of Theorem 5.2 and Lemma 5.9 instead of Lemma 5.8. □

Appendix A. Entropy

A.1. **Entropy in terms of covering numbers.** Let us explain how entropies are related to covering numbers. We use these relations in Section 2.

Let $f\colon X \to X$ be a continuous self-map of a compact metric space (X, d). For each integer $n \geqslant 1$, define the *Bowen metric*:

$$\mathsf{d}_n(x,y) := \max_{0 \leqslant i < n} \mathsf{d}(f^i(x), f^i(y))\,. \tag{A.1}$$

Let $N(n,\epsilon) := D_{\mathsf{d}_n}(\epsilon)$ denote the least number of balls of radii ϵ in the d_n-metric necessary to cover X. We recall the following:

Definition A.1. The *topological entropy* of f is:

$$h_{\mathrm{top}}(f) := \lim_{\epsilon \to 0} \lim_{n \to \infty} \frac{1}{n} \log N(n, \epsilon).$$

Fix an invariant measure $\mu \in \mathcal{M}_f(X)$. Given $n \geqslant 1$, $\epsilon > 0$, and $0 < \delta < 1$, let $N_\mu(n, \epsilon, \delta)$ denote the least number of balls of radii ϵ in the d_n-metric necessary to cover a set of μ-measure at least $1 - \delta$.

Though metric entropy is most commonly defined in terms of measurable partitions, the following result by Katok allows us to define it in terms of covering numbers:

Theorem A.2 (Katok [Ka], Theorem I.I)**.** *If μ is ergodic then for every δ in the range $0 < \delta < 1$,*

$$h_\mu(f) = \lim_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log N_\mu(n, \epsilon, \delta) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log N_\mu(n, \epsilon, \delta)\,. \tag{A.2}$$

When f is a homeomorphism, let $\tilde{N}_\mu(n, \epsilon, \delta)$ denote the least number of balls necessary to cover a set of μ-measure at least $1 - \delta$ of radii ϵ in the following metric:

$$\tilde{\mathsf{d}}_n(x, y) := \max_{-n < i < n} \mathsf{d}(f^i(x), f^i(y))\,. \tag{A.3}$$

We note that $\mathsf{d}_{2n}(f^{-n}(x), f^{-n}(y)) = \tilde{\mathsf{d}}_n(x, y)$ and so $\tilde{N}_\mu(n, \epsilon, \delta) = N_\mu(2n, \epsilon, \delta)$. So we obtain:

Corollary A.3. *If μ is ergodic then for every δ in the range $0 < \delta < 1$,*

$$h_\mu(f) = \lim_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{2n} \log \tilde{N}_\mu(n, \epsilon, \delta) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{2n} \log \tilde{N}_\mu(n, \epsilon, \delta)\,. \tag{A.4}$$

A.2. **Variational principle for entropy.** If a measurable self-map f of a measurable space X preserves a probability measure μ, then $h_\mu(f)$ denotes the corresponding metric entropy.

Theorem A.4 (Variational Principle for Entropy)**.** *If X is compact and f is continuous, then the topological entropy $h_{\mathrm{top}}(f)$ equals the supremum of $h_\mu(f)$ where μ runs over all the invariant Borel probability measures.*

Details can be found in the standard textbooks [DGS, Ma2, KaH, PrU, ViO].

A.3. **Metric entropy in terms of quantization numbers.** Let (X, d) be a compact metric space, and let W_p and LP denote the induced Wasserstein and Lévy–Prokhorov metrics on the space $\mathcal{M}(X)$. If $\mu \in \mathcal{M}(X)$, then let $Q_{\mu,\mathsf{W}_p}(\cdot)$ and $Q_{\mu,\mathsf{LP}}(\cdot)$ and denote the corresponding quantization numbers. They can be compared as follows:

Lemma A.5. *For every $\epsilon > 0$,*

$$Q_{\mu,\mathsf{LP}}\left(\epsilon^{\frac{p}{p+1}}\right) \leqslant Q_{\mu,\mathsf{W}_p}(\epsilon) \leqslant Q_{\mu,\mathsf{LP}}\left(\frac{\epsilon^p}{1 + (\operatorname{diam} X)^p}\right)\,.$$

Proof. This is an immediate consequence of inequalities (1.3). ☐

Given a continuous map $f \colon X \to X$ on the compact metric space (X, d) and an integer $n \geqslant 1$, the corresponding Bowen metric d_n induces Wasserstein and Lévy–Prokhorov metrics on the space $\mathcal{M}(X)$, which we respectively denote by $\mathsf{W}_{p,n}$ and LP_n. Now, given an invariant measure $\mu \in \mathcal{M}_f(X)$, we consider its quantization numbers with respect these two metrics. This relates to the entropy as follows:

Theorem A.6 (Reformulation of Katok's entropy theorem)**.** *If μ is ergodic then:*

$$h_\mu(f) = \lim_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} Q_{\mu, \mathsf{LP}_n}(\epsilon) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} Q_{\mu, \mathsf{LP}_n}(\epsilon)$$
$$= \lim_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} Q_{\mu, \mathsf{W}_{p,n}}(\epsilon) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} Q_{\mu, \mathsf{W}_{p,n}}(\epsilon) \, .$$

Proof. Note that existence of limits as $\epsilon \to 0$ is automatic by monotonicity.

In view of Lemma A.5, it is sufficient to consider the Lévy–Prokhorov metrics. By Lemma A.5, $Q_{\mu, \mathsf{LP}_n}(\epsilon) = N_\mu(n, \epsilon, \epsilon)$ (in the notation of Appendix A.1).

In the paper [Ka] (see inequality (I.I)), Katok proves that:

$$\forall \epsilon > 0, \ \forall \delta > 0, \quad \limsup_{n \to \infty} \frac{1}{n} \log N(n, \epsilon, \delta) \leqslant h_\mu(f) \, .$$

(This is actually the "easy part" of the proof of Theorem A.2, and a simple consequence of Shannon–MacMillan–Breiman's theorem.) Taking $\delta = \epsilon$ and then taking $\epsilon \to 0$, we obtain:

$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log N(n, \epsilon, \epsilon) \leqslant h_\mu(f) \, .$$

On the other hand, if $0 < \epsilon \leqslant \delta < 1$ then $N(n, \epsilon, \epsilon) \geqslant N(n, \epsilon, \delta)$, so Theorem A.2 implies that:

$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log N(n, \epsilon, \epsilon) \geqslant h_\mu(f) \, .$$

This concludes the proof. ☐

The reader will notice a certain parallelism between the notions of topological/metric entropies and topological/metric emergences: compare Definition A.1 with Definition 0.2, Theorem A.6 with Proposition 3.12, and Theorem A.4 with Theorem E.

References

[AB] Avila, A.; Bochi, J. – Nonuniform hyperbolicity, global dominated splittings and generic properties of volume-preserving diffeomorphisms. *Trans. Amer. Math. Soc.* 364 (2012), no. 6, 2883–2907. DOI:10.1090/S0002-9947-2012-05423-7 (Cited on page 22.)

[ACW] Avila, A.; Crovisier, S.; Wilkinson, A. – C^1 density of stable ergodicity. *Adv. Math.* 379 (2021), Article ID 107496, 69 p. DOI:10.1016/j.aim.2020.107496 (Cited on page 4.)

[BCLR] Béguin, F.; Crovisier, S.; Le Roux, F. – Construction of curious minimal uniquely ergodic homeomorphisms on manifolds: the Denjoy-Rees technique. *Ann. Sci. École Norm. Sup.* 40 (2007), no. 2, 251–308. DOI:10.1016/j.ansens.2007.01.001 (Cited on page 6.)

[Be1] Berger, P. – Generic family with robustly infinitely many sinks. *Invent. Math.* 205 (2016), no. 1, 121–172. DOI:10.1007/s00222-015-0632-6 (Cited on page 5.)

[Be2] Berger, P. – Emergence and non-typicality of the finiteness of the attractors in many topologies. *Proc. Steklov Inst. Math.* 297 (2017), no. 1, 1–27. DOI:10.1134/S0081543817040010 (Cited on pages 1, 2, 4, 5, and 6.)

[BeB] Berger, P.; Bounemoura, A. – A geometrical proof of the persistence of normally hyperbolic submanifolds. *Dyn. Syst.* 28 (2013), no. 4, 567–581. DOI:10.1080/14689367.2013.835386 (Cited on page 44.)

[BeT] Berger, P.; Turaev, D. – On Herman's positive entropy conjecture. *Adv. Math.* 349 (2019), 1234–1288. DOI:10.1016/j.aim.2019.04.002 (Cited on page 6.)

[Bie] Biebler, S. – Newhouse phenomenon for automorphisms of low degree in $\mathbb{C}^3$. *Adv. Math.* 361 (2020), Article ID 106952, 39 pp. DOI:10.1016/j.aim.2019.106952 (Cited on page 5.)

[Bil] Billingsley, P. – *Convergence of probability measures.* 2nd edition. Wiley Series in Probability and Statistics. John Wiley & Sons, New York, 1999. (Cited on page 10.)

[BiK] Birkhoff, G.D.; Koopman B.O. – Recent contributions to the ergodic theory. *Proc. Natl. Acad. Sci. USA* 18 (1932), no. 3, 279–282. (Cited on page 5.)

[BBG] Bochi, J.; Bonatti, C.; Gelfert, K. – Dominated Pesin theory: convex sum of hyperbolic measures. *Israel J. Math.* 226 (2018), no. 1, 387–417. DOI:10.1007/s11856-018-1699-8 (Cited on page 4.)

[BGV] Bolley, F.; Guillin, A.; Villani, C. – Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Related Fields* 137 (2007), no. 3–4, 541–593. DOI:10.1007/s00440-006-0004-7 (Cited on pages 10 and 11.)

[BoD] Bonatti, C.; Díaz, L.J. – Connexions hétéroclines et généricité d'une infinité de puits et de sources. *Ann. Sci. École Norm. Sup.* 32 (1999) no. 1, 135–150. DOI:10.1016/S0012-9593(99)80012-3 (Cited on page 5.)

[BoZ] Bonatti, C.; Zhang, J. – Periodic measures and partially hyperbolic homoclinic classes. *Trans. Amer. Math. Soc.* 372 (2019), no. 2, 755–802. (Cited on page 4.)

[BrS] Broer, H.W.; Sevryuk, M.B. – KAM Theory: quasi-periodicity in dynamical systems, in *Handbook of dynamical systems. Vol. 3,* H.W. Broer, B. Hasselblatt and F. Takens, Eds., North-Holland (Elsevier), Amsterdam, 2010, Chapter 6, pp. 249–344. (Cited on page 38.)

[BuW] Burns, K.; Wilkinson, A. – On the ergodicity of partially hyperbolic systems. *Ann. of Math.* 171 (2010), no. 1, 451–489. DOI:10.4007/annals.2010.171.451 (Cited on page 4.)

[Buz] Buzzard, G.T. – Infinitely many periodic attractors for holomorphic maps of 2 variables, *Ann. of Math.* 145, (1997) no. 2, 389–417. DOI:10.2307/2951819 (Cited on page 5.)

[CLN] Charpentier, E.; Lesne, A.; Nikolski, N.K. (eds.) – *Kolmogorov's heritage in mathematics.* Translated from the 2004 French original. Springer, Berlin, 2007. (Cited on page 9.)

[DGS] Denker, M.; Grillenberger, C.; Sigmund, K. – *Ergodic theory on compact spaces.* Lecture Notes in Mathematics, Vol. 527. Springer-Verlag, Berlin-New York, 1976. (Cited on pages 22 and 45.)

[DGMR] Díaz, L.J.; Gelfert, K.; Marcarini, T.; Rams, M. – The structure of the space of ergodic measures of transitive partially hyperbolic sets. *Monatsh. Math.* 190 (2019), no. 3, 441–479. DOI:10.1007/s00605-019-01325-2 (Cited on page 4.)

[DGR] Díaz, L.J.; Gelfert, K.; Rams, M. – Entropy spectrum of Lyapunov exponents for nonhyperbolic step skew-products and elliptic cocycles. *Comm. Math. Phys.* 367 (2019), no. 2, 351–416. DOI:10.1007/s00220-019-03412-9 (Cited on page 4.)

[DNP] Díaz, L.J.; Nogueira, A.; Pujals, E.R. – Heterodimensional tangencies. *Nonlinearity* 19 (2006), no. 11, 2543–2566. DOI:10.1088/0951-7715/19/11/003 (Cited on page 5.)

[Dow] Downarowicz, T. – The Choquet simplex of invariant measures for minimal flows. *Israel J. Math.* 74 (1991), no. 2-3, 241–256. DOI:10.1007/BF02775789 (Cited on page 4.)

[Dua] Duarte, P. – Plenty of elliptic islands for the standard family of area preserving maps. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 11 (1994) no.4, 359–409. DOI:10.1016/S0294-1449(16)30180-9 (Cited on page 5.)

[Dum] Dumas, H.S. – *The KAM story.* World Scientific, Hackensack, NJ, 2014. (Cited on page 5.)

[Fa] Falconer, K. – *Fractal geometry: Mathematical foundations and applications.* 3rd edition. John Wiley & Sons, Chichester, 2014. (Cited on page 9.)

[GeR] Gelfert, K.; Kwietniak, D. – On density of ergodic measures and generic points. *Ergodic Theory Dynam. Systems* 38 (2018), no. 5, 1745–1767. DOI:10.1017/etds.2016.97 (Cited on page 4.)

[GeT] Gelfreich, V.; Turaev, D. – Universal dynamics in a neighborhood of a generic elliptic periodic point. *Regul. Chaotic Dyn.* 15 (2010), no. 2–3, 159–164. DOI:10.1134/S156035471002005X (Cited on page 37.)

[GiS] Gibbs, A.L.; Su, F.E. – On choosing and bounding probability metrics. *Internat. Statist. Rev.* 70 (2002), no. 3, 419–435. DOI:10.1111/j.1751-5823.2002.tb00178.x (Cited on page 10.)

[GTS] Gonchenko, S.; Turaev, D.; Shilnikov, L. – Homoclinic tangencies of arbitrarily high orders in conservative and dissipative two-dimensional maps. *Nonlinearity* 20 (2007), no. 2, 241–275. DOI:10.1088/0951-7715/20/2/002 (Cited on page 37.)

[GoP] Gorodetski, A.; Pesin, Ya. – Path connectedness and entropy density of the space of hyperbolic ergodic measures. *Modern theory of dynamical systems*, 111–121, *Contemp. Math.*, 692, Amer. Math. Soc., Providence, RI, 2017. (Cited on page 4.)

[GrL] Graf, S.; Luschgy, H. – *Foundations of quantization for probability distributions.* Lecture Notes in Mathematics, 1730. Springer-Verlag, Berlin, 2000. (Cited on pages 4, 19, 20, and 21.)

[GrS] Grimmett, G.R.; Stirzaker, D.R. – *Probability and random processes.* 3rd edition. Oxford University Press, New York, 2001. (Cited on page 11.)

[Ka] Katok, A. – Lyapunov exponents, entropy and periodic orbits for diffeomorphisms. *Inst. Hautes Études Sci. Publ. Math.* 51 (1980), 137–173. (Cited on pages 8, 45, and 46.)

[KaH] Katok, A.; Hasselblatt, B. – *Introduction to the modern theory of dynamical systems.* Encyclopedia of Mathematics and its Applications, 54. Cambridge University Press, Cambridge, 1995. (Cited on pages 17, 19, and 45.)

[Kl1] Kloeckner, B. – A generalization of Hausdorff dimension applied to Hilbert cubes and Wasserstein spaces. *J. Topol. Anal.* 4 (2012), no. 2, 203–235. DOI:10.1142/S1793525312500094 (Cited on page 9.)

[Kl2] Kloeckner, B. – A geometric study of Wasserstein spaces: ultrametrics. *Mathematika* 61 (2015), no. 1, 162–178. DOI:10.1112/S0025579314000059 (Cited on pages 9 and 10.)

[KoT] Kolmogorov, A.N.; Tihomirov, V.M. – ϵ-entropy and ϵ-capacity of sets in functional space. *Amer. Math. Soc. Transl.* (2) 17 (1961), 277–364. (Cited on pages 9 and 11.)

[KuZ] Kulkarni, S.R.; Zeitouni, O. – A general classification rule for probability measures. *Ann. Statist.* 23 (1995), no. 4, 1393–1407. DOI:10.1214/aos/1176324714 (Cited on page 10.)

[LiM] Lindsay, L.J.; Mauldin, R.D. – Quantization dimension for conformal iterated function systems. *Nonlinearity* 15 (2002), no. 1, 189–199. DOI:10.1088/0951-7715/15/1/309 (Cited on page 21.)

[MK] MacKay, R.S. – Nonlinearity in complexity science. *Nonlinearity* 21 (2008), no. 12, T273–T281. DOI:10.1088/0951-7715/21/12/T03 (Cited on page 2.)

[Ma1] Mañé, R. – An ergodic closing lemma. *Ann. of Math.* 116 (1982), no. 3, 503–540. DOI:10.2307/2007021 (Cited on page 6.)

[Ma2] Mañé, R. – *Ergodic theory and differentiable dynamics.* Translated from the Portuguese by Silvio Levy. Ergebnisse der Mathematik und ihrer Grenzgebiete, 8. Springer–Verlag, Berlin, 1987. (Cited on pages 22 and 45.)

[Mo] Moser, J. – On invariant curves of area-preserving mappings of an annulus. *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II* 1962 (1962), 1–20. (Cited on page 38.)

[Ne1] Newhouse, S. – Diffeomorphisms with infinitely many sinks. *Topology* 13 (1974), 9–18. DOI:10.1016/0040-9383(74)90034-2 (Cited on page 5.)

[Ne2] Newhouse, S. – The abundance of wild hyperbolic sets and nonsmooth stable sets for diffeomorphisms. *Inst. Hautes Études Sci. Publ. Math.* 50 (1979), 101–151. (Cited on page 5.)

[Ng] Nguyen, XuanLong. – Convergence of latent mixing measures in finite and infinite mixture models. *Ann. Statist.* 41 (2013), no. 1, 370–400. DOI:10.1214/12-AOS1065 (Cited on page 11.)

[Ob] Obata, D. – On the stable ergodicity of Berger–Carrasco's example. *Erg. Theory Dyn. Systems* 40 (2020), no. 4, 1008–1056. DOI:10.1017/etds.2018.65 (Cited on page 4.)

[OxU] Oxtoby, J.C.; Ulam, S.M. – Measure-preserving homeomorphisms and metrical transitivity. *Ann. of Math.* 42 (1941), 874–920. DOI:10.2307/1968772 (Cited on page 6.)

[Pe] Pesin, Ya. B. – *Dimension theory in dynamical systems: Contemporary views and applications.* Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1997. (Cited on pages 17 and 18.)

[Pö] Pöschel, J. Integrability of Hamiltonian systems on Cantor sets. *Comm. Pure Appl. Math.* 35 (1982), no. 5, 653–696. DOI:10.1002/cpa.3160350504 (Cited on page 38.)

[PrU] Przytycki, F.; Urbański, M. – *Conformal fractals: ergodic theory methods.* London Mathematical Society Lecture Note Series, 371. Cambridge University Press, Cambridge, 2010. (Cited on pages 14, 15, and 45.)

[Rog] Rogers, C.A. – *Packing and covering.* Cambridge Tracts in Mathematics and Mathematical Physics, No. 54. Cambridge University Press, New York, 1964. (Cited on page 8.)

[Sig] Sigmund, K. – On the connectedness of ergodic systems. *Manuscripta Math.* 22 (1977), no. 1, 27–32. DOI:10.1007/BF01182064 (Cited on page 4.)

[Tu] Turaev, D. – Maps close to identity and universal maps in the Newhouse domain. *Comm. Math. Phys.* 335 (2015), no. 3, 1235–1277. DOI:10.1007/s00220-015-2338-4 (Cited on page 37.)

[ViO] Viana, M.; Oliveira, K. – *Foundations of ergodic theory.* Cambridge Studies in Advanced Mathematics, 151. Cambridge University Press, Cambridge, 2016. (Cited on pages 16, 17, and 45.)

[Vil] Villani, C. – *Topics in optimal transportation.* Graduate Studies in Mathematics, 58. American Mathematical Society, Providence, RI, 2003. (Cited on pages 10 and 38.)

Pierre Berger (berger@math.univ-paris13.fr), Université Paris 13, Sorbonne Paris Cité, LAGA, CNRS (UMR 7539), F-93430, Villetaneuse, France.

Jairo Bochi (jairo.bochi@mat.uc.cl), Facultad de Matemáticas, Pontificia Universidad Católica de Chile, Avenida Vicuña Mackenna 4860 Santiago, Chile