S. Lim *et al.* (2018) "Dimension Bound for Badly Approximable Grids,"
International Mathematics Research Notices, Vol. 00, No. 0, pp. 1–30
Advance Access publication on Month XX, 2018
doi:10.1093/imrn/rnx330

Dimension Bound for Badly Approximable Grids

Seonhee Lim[1], Nicolas de Saxcé[2], and Uri Shapira[3,*]
[1]Department of Mathematical Sciences, Seoul National University, [2]C.N.R.S. – Laboratoire Analyse, Géométrie et Applications, Université Paris 13, Villetaneuse, and [3]Department of Mathematics, Technion, Haifa, Israel

[*]*Correspondence to be sent to: e-mail: ushapira@tx.technion.ac.il*

We show that there exists a subset of full Lebesgue measure $V \subset \mathbb{R}^n$ such that for every $\epsilon > 0$ there exists $\delta > 0$ such that for any $v \in V$ the dimension of the set of vectors w satisfying

$$\liminf_{k\to\infty} k^{1/n}\langle kv - w\rangle \geqslant \epsilon$$

(where $\langle\cdot\rangle$ denotes the distance from the nearest integer) is bounded above by $n - \delta$. This result is obtained as a corollary of a discussion in homogeneous dynamics and the main tool in the proof is a relative version of the principle of uniqueness of measures with maximal entropy.

1 The Main Result and Its Applications

1.1 Geometry of numbers

A general theme in the geometry of numbers is to fix a domain $S \subset \mathbb{R}^d$ and study the intersection of it with sets possessing an algebraic structure such as lattices or their cosets. One is usually interested in bounding the cardinality of such an intersection

Received July 2, 2017; Revised December 1, 2017; Accepted December 6, 2017

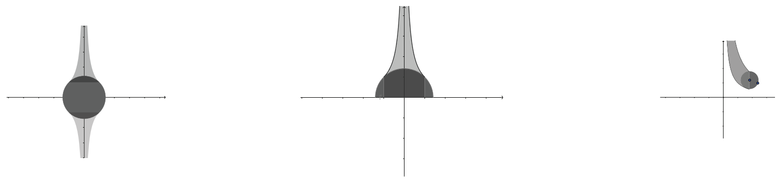

FIG. 1. Spikes for $a_t = \operatorname{diag}\left(e^t, e^{-t}\right)$ of a ball around 0, a half-ball around 0, a ball around (3,2)

and this will be the case in our discussion as well. We begin by describing the domains we will consider and which we refer to as *spikes*. Throughout, we fix a dimension $d \geq 2$ and a diagonal flow

$$a_t = \operatorname{diag}\left(e^{c_1 t}, \ldots, e^{c_d t}\right);$$

where c_i are fixed *nonzero* numbers such that $\sum c_j = 0$. Given a bounded open set $\mathcal{O} \subset \mathbb{R}^d$, we define the *positive spike of* $\mathcal{O}$ *with respect to* a_t to be the set (One could work with two-sided spikes taking the union over $t \in \mathbb{R}$ in (1.1) but our results are stronger as the domain decreases so we will concentrate on the one-sided case.)

$$S^+(a_t, \mathcal{O}) = S^+(\mathcal{O}) \overset{\text{def}}{=} \bigcup_{t>0} a_t^{-1}\mathcal{O}. \tag{1.1}$$

As a_t will be fixed throughout our discussion we omit it from the notation. The space of unimodular lattices (i.e., of covolume 1) in $\mathbb{R}^d$ will be denoted by X. By a unimodular *grid* y in $\mathbb{R}^d$, we mean a coset $x + \mathbf{w}$ of a lattice $x \in X$ where $\mathbf{w} \in \mathbb{R}^d$. We denote by Y the space of unimodular grids in $\mathbb{R}^d$ and by $\pi : Y \to X$ the natural projection. Note that for $x \in X$, the fiber $\pi^{-1}(x)$ is simply the torus $\mathbb{R}^d/x$. For $x \in X$ and an open set $\mathcal{O} \subset \mathbb{R}^d$, we set

$$F_{S^+(\mathcal{O})} \overset{\text{def}}{=} \left\{y \in Y : y \cap S^+(\mathcal{O}) \text{ is finite}\right\};$$
$$F_{S^+(\mathcal{O})}(x) \overset{\text{def}}{=} F_{S^+(\mathcal{O})} \cap \pi^{-1}(x). \tag{1.2}$$

Our main result, Theorem 1.3 below, says that under a mild dynamical assumption on the forward a_t-orbit of x, the set $F_{S^+(\mathcal{O})}(x)$ cannot have maximal Hausdorff dimension in $\pi^{-1}(x)$. Let us introduce this dynamical assumption. The standard action of $G_0 \overset{\text{def}}{=} \mathrm{SL}_d(\mathbb{R})$ on $\mathbb{R}^d$ induces a transitive action of G_0 on X, and since for $x_0 = \mathbb{Z}^d$ we have $\operatorname{Stab} x_0 = G_0(\mathbb{Z})$, we may identify $X \simeq G_0/G_0(\mathbb{Z})$. This endows X with a smooth manifold structure and with a unique G_0-invariant Borel probability measure, which we denote by m_X.

For a locally compact second countable Hausdorff space Z we will denote by $\mathscr{P}(Z)$ the space of Borel probability measures on Z and endow it with the weak* topology by identifying $\mathscr{P}(Z)$ with a subset of the unit sphere in the dual of $C_0(Z)$. For $\mu \in \mathscr{P}(Z)$ and $f \in C_0(X)$ we alternate between the notations $\mu(f)$ and $\int f \, d\mu$. We will denote by $\delta_z \in \mathscr{P}(Z)$ the Dirac probability measure at z. If a map $g : Z \to Z$ is fixed we denote for $z \in Z$ and $T \in \mathbb{Z}_+$, $\delta_z^T \overset{\text{def}}{=} \frac{1}{T}\sum_{i=0}^{T-1} \delta_{g^i z} \in \mathscr{P}(Z)$. Note that by the Banach–Alaoglu theorem $\{\alpha\mu : \alpha \in [0,1], \mu \in \mathscr{P}(Z)\}$ is compact in the weak* topology and thus for any $z \in Z$, the sequence $\delta_z^T \in \mathscr{P}(Z)$ has accumulation points of the form $\alpha\mu$ with $\alpha \in [0,1]$ and $\mu \in \mathscr{P}(Z)$. The following is concerned with the situation where δ_z^T can accumulate on a *probability* measure. Throughout we use the notation δ_x^T for the transformation $a = a_1 : X \to X$.

Definition 1.1. (Heavy lattice)

(1) A lattice $x \in X$ is called *heavy* (for a_t in positive time) if

$$\left\{\delta_x^T : T \in \mathbb{Z}_+\right\}^{\omega} \cap \mathscr{P}(X) \neq \varnothing,$$

where F^{ω} is the set of accumulation points of F.

(2) We fix once and for all a sequence of compactly supported functions $\psi_i \in C_c(X)$ such that $0 \leqslant \psi_i \leqslant 1$, and $\psi_i^{-1}(1)$ is an increasing sequence of compact sets that covers X. Given a sequence of nonnegative numbers $\eta_i \to 0$, we define

$$\mathscr{P}(X, (\eta_i)) \overset{\text{def}}{=} \{\mu \in \mathscr{P}(X) : \forall i,\ \mu(\psi_i) \geqslant 1 - \eta_i\}.$$

(3) Given a sequence of nonnegative numbers $\eta_i \to 0$, we define

$$\mathcal{H}(\eta_i) \overset{\text{def}}{=} \left\{x \in X : \left\{\delta_x^T : T \in \mathbb{Z}_+\right\}^{\omega} \cap \mathscr{P}(X, (\eta_i)) \neq \varnothing\right\}.$$

As the following lemma shows, any heavy lattice belongs to some $\mathcal{H}(\eta_i)$. The point in defining $\mathscr{P}(X, (\eta_i))$ and $\mathcal{H}(\eta_i)$ is that our results about heavy lattices will be uniform on $\mathcal{H}(\eta_i)$.

Lemma 1.2.

(1) $\mathscr{P}(X) = \bigcup \mathscr{P}(X, (\eta_i))$ where the union is taken over all sequences of nonnegative numbers $\eta_i \to 0$.

(2) The set of heavy lattices equals $\bigcup \mathcal{H}(\eta_i)$ where the union is taken over all sequences of nonnegative numbers $\eta_i \to 0$.

(3) $\mathscr{P}(X, (\eta_i))$ is compact.

Proof.

(1) For $\mu \in \mathscr{P}(X)$, $\mu \in \mathscr{P}(X, (\eta_i))$ for $\eta_i \stackrel{\text{def}}{=} 1 - \mu(\psi_i)$. Note that $\eta_i \to 0$ because $\psi_i^{-1}(1)$ is an increasing cover of X.

(2) Let x be a heavy lattice and let $\mu \in \left\{\delta_x^T : T \in \mathbb{Z}^+\right\}^\omega \cap \mathscr{P}(X)$. By part (1.2) $\mu \in \mathscr{P}(X, (\eta_i))$ for some sequence (η_i). By definition we then have that $x \in \mathcal{H}(\eta_i)$.

(3) Let $\mu_m \in \mathscr{P}(X, (\eta_i))$ be a sequence and let μ be a weak* accumulation point of it. For each i we have $\mu(X) \geqslant \mu(\psi_i) = \lim_m \mu_m(\psi_i) \geqslant 1 - \eta_i$. Letting $i \to \infty$ we obtain $\mu(X) = 1$ and $\mu(\psi_i) \geqslant 1 - \eta_i$ so that $\mu \in \mathscr{P}(X, (\eta_i))$ by definition. ∎

Our main result is as follows. Here, $\dim_H$ denotes Hausdorff dimension with respect to the Euclidean metric on $\pi^{-1}(x) \simeq \mathbb{R}^d/x$.

Theorem 1.3 (Heavy lattices have few bad grids). For any bounded open set $\mathcal{O} \subset \mathbb{R}^d$, if x is heavy (for a_t in positive time), then

$$\dim_H F_{S^+(\mathcal{O})}(x) < d.$$

In fact, for a given $\eta_i \to 0$, there exists $\delta = \delta(\mathcal{O}, (\eta_i)) > 0$ such that for any $x \in \mathcal{H}(\eta_i)$,

$$\dim_H F_{S^+(\mathcal{O})}(x) < d - \delta.$$

By Lemma 1.2(2), results stated for lattices in $\mathcal{H}(\eta_i)$ for arbitrary (η_i) automatically hold for heavy lattices. Thus, the second part of Theorem 1.3 implies the first and demonstrates the uniformity gained by exhausting the set of heavy lattices by the sets $\mathcal{H}(\eta_i)$. The following corollary shows that this uniformity survives if one is only interested in an almost sure statement with respect to the smooth measure m_X.

Corollary 1.4 (Random lattices have few bad grids). For any bounded open set $\mathcal{O} \subset \mathbb{R}^d$, there exists $\delta > 0$ such that for m_X-almost any lattice x, $\dim_H F_{S^+(\mathcal{O})}(x) \leqslant d - \delta$.

Proof. By the ergodicity of the action of a_1 on X, for almost any x, $\delta_x^T \xrightarrow{\text{w}^*} m_X$ which trivially implies $m_X \in \left\{\delta_x^T : T > 0\right\}^\omega$. By Lemma 1.2(1), $m_X \in \mathscr{P}(X, (\eta_i))$ for a suitable sequence (η_i). Thus by definition $x \in \mathcal{H}(\eta_i)$. The result then follows from Theorem 1.3. ∎

1.2 An application to Diophantine approximation

For a vector $v \in \mathbb{R}^n$, we are interested in the behavior of the sequence

$$\left\{kv \bmod \mathbb{Z}^n \;;\; k \in \mathbb{N}\right\} \subset \mathbb{R}^n/\mathbb{Z}^n.$$

If v does not belong to a rational subspace, this sequence is dense and even equidistributed so that for any *target* $w \in \mathbb{R}^n$, we have that $\inf_{k\geq 1}\langle kv-w\rangle = 0$, where $\langle t\rangle$ denotes the distance from t to $\mathbb{Z}^n$. A more subtle question is whether $\liminf_{k\to\infty}\psi(k)\langle kv-w\rangle = 0$ for some prescribed function $\psi \nearrow \infty$ on $\mathbb{N}$. One may visualize this as a *shrinking target* problem where one asks if for any $\epsilon > 0$ and for arbitrarily large k, the point $kv \bmod \mathbb{Z}^n$ on the n-torus is inside the ball of radius $\psi(k)^{-1}\epsilon$ centered at w (which is the shrinking target). The case that we consider here is $\psi(k) = k^{1/n}$. We call w ϵ*-bad* for v if

$$\liminf_{k\to\infty} k^{1/n}\langle kv - w\rangle \geqslant \epsilon, \tag{1.3}$$

and denote

$$\mathbf{Bad}^{\epsilon}(v) \stackrel{\text{def}}{=} \left\{w \in \mathbb{R}^n : w \text{is } \epsilon\text{-bad for } v\right\},$$

$$\mathbf{Bad}(v) \stackrel{\text{def}}{=} \bigcup_{\epsilon>0} \mathbf{Bad}^{\epsilon}(v).$$

Our main application is the following.

Theorem 1.5. For any $\epsilon > 0$ there exists $\delta > 0$ such that for Lebesgue almost every $v \in \mathbb{R}^n$, $\dim_H \mathbf{Bad}^{\epsilon}(v) < n - \delta$.

To put this in context we mention that it follows from [2] that for any $v \in \mathbb{R}^n$, $\dim_H \mathbf{Bad}(v) = n$. Later, it was shown in [6] that $\mathbf{Bad}(v)$ is a winning set.

Note also that the conclusion of the theorem cannot hold for every $v \in \mathbb{R}^n$. Indeed, if v lies in a rational subspace, then for small enough ϵ, the set $\mathbf{Bad}^{\epsilon}(v)$ has nonempty interior and thus obviously has dimension n. It is clear that when the vector v is rational, it is in particular *singular*, that is, for any $\delta > 0$, there exists T_δ such that for any $T > T_\delta$, there exists $k < T$ such that $T^{1/n}\langle kv\rangle < \delta$.

On the other hand, in Section 6 we construct non-singular vectors which violate the conclusion of the theorem and satisfy $\dim_H \mathbf{Bad}^{\epsilon}(v) = n$ for a positive ϵ.

As we will see in Section 4, Theorem 1.5 holds not only for almost every v, but for any heavy vector v (although the constant δ in Theorem 1.5 might depend on v), see Definition 5.1 and Theorem 5.3. It follows from [12] that for non-singular vectors $\lambda(\mathbf{Bad}(v)) = 0$, where λ is the Lebesgue measure on $\mathbb{R}^n$. Thus the above theorem is an upgrade of the result in [12] just mentioned under the stronger assumption of heaviness. We refer the reader to Section 4 for other examples of applications of Theorem 1.3, such as Diophantine approximation of affine subspaces of $\mathbb{R}^n$.

1.3 Outline of the proof of Theorem 1.3

We briefly describe our strategy for the proof of the first part of Theorem 1.3 which is similar in spirit to the idea given in [1, Remark 2.2] and which could be described in a nutshell as rigidity of measures with maximal entropy. Assuming by way of contradiction that $\dim F_{S^+(\mathcal{O})}(x) = d$ for a heavy lattice x, we construct a sequence of probability measures μ_k defined by taking the uniform measures ν_k on large finite sets $S_k \subset F_{S^+(\mathcal{O})}(x) \subset \pi^{-1}(x)$ and averaging them along the a_1-orbit; $\mu_k = n_k^{-1} \sum_{i=1}^{n_k} (a_1)_*^i \nu_k$, for a suitable n_k (defined in the proof of Proposition 2.3). The heaviness assumption allows us to take a weak-* limit μ of μ_k that is a probability measure on Y and moreover, the maximal dimension assumption translates into the maximality of the relative entropy of μ with respect to a_1 relative to the factor X. We then prove that maximality of the relative entropy implies invariance of μ under the whole subgroup of translations in $\mathbb{R}^d$. This leads us to a contradiction because by construction, μ is supported on the accumulation points of forward a_t-orbits of the points in $F_{S^+(\mathcal{O})}$ and in particular, must be supported in the closed a_t-invariant set $\{y \in Y : \ 0 \in y \text{ or } \forall t \in \mathbb{R},\ a_t \cdot y \cap \mathcal{O} = \varnothing\}$.

1.4 Plan of the paper

Apart from this introduction, this paper consists of five parts. In Section 2, we show that a set of large dimension in the space of grids, contained in a single fiber, can be used to construct an a_t-invariant measure on Y with large entropy with respect to the factor X. Then, in Section 3, we study a_t-invariant measures with maximal relative entropy with respect to X; using work of Einsiedler and Lindenstrauss, we show that they are invariant under the unstable horospherical subgroup U^+ associated to a_t. In Section 4, using the two previous sections, we finish the proof of Theorem 1.3, before explaining the applications to Diophantine approximation in Section 5. We conclude the paper with Section 6, detailing the construction of some lattices with non-divergent a_t-orbits that do not satisfy the conclusion of Theorem 1.3.

2 Measures With Large Entropy

Given a heavy lattice x for a_t and a set $S \subset \pi^{-1}(x)$ of grids lying above x, we explain how to construct a measure μ on Y with large entropy relative to the factor X, which is supported in the closure of the forward orbit of S under the diagonal flow a_t.

2.1 Action on the space of grids

Just like for the space X of lattices in $\mathbb{R}^d$, one can view the space Y of unimodular grids as a homogeneous space. Indeed, the natural action on $\mathbb{R}^d$ of the group $G = \mathrm{ASL}_d(\mathbb{R})$ of area-preserving affine transformations induces a transitive action on Y, with $\mathrm{Stab}\, y_0 = G(\mathbb{Z})$ if $y_0 = \mathbb{Z}^d$, so that $Y \simeq G/G(\mathbb{Z})$ has a natural smooth manifold structure, and carries a unique G-invariant probability measure m_Y.

We denote by U the unipotent radical of G, which consists of all translations on $\mathbb{R}^d$. It is clear that U acts simply transitively on $\mathbb{R}^d$. One naturally identifies $G_0 = \mathrm{SL}_d(\mathbb{R})$ with G/U, and then, the canonical projection $\pi : Y \to X$ intertwines the actions of G and G_0 on Y and X, respectively, in the sense that for any $g \in G$ and $y \in Y$,

$$\pi(g \cdot y) = \bar{g} \cdot \pi(y),$$

where $\bar{g}$ denotes the projection of g to $G_0 \simeq G/U$. Clearly, if y is a grid with underlying lattice x, then

$$\pi^{-1}(x) = Uy.$$

It is sometimes convenient to view G as a subgroup of $\mathrm{SL}_{d+1}(\mathbb{R})$ by

$$G = \left\{ \begin{pmatrix} A & u \\ 0 & 1 \end{pmatrix};\ A \in \mathrm{SL}_d(\mathbb{R}),\ u \in \mathbb{R}^d \right\},$$

in which case the unipotent radical is $U = \left\{ \begin{pmatrix} I_d & u \\ 0 & 1 \end{pmatrix} : u \in \mathbb{R}^d \right\}$.

Recall that $a_t = \mathrm{diag}\left(e^{c_1 t}, \ldots, e^{c_d t}\right)$ is a one-parameter diagonal subgroup in $G_0 = \mathrm{SL}_d(\mathbb{R})$. We take a lift of this one-parameter group to $G \subset \mathrm{SL}_{d+1}(\mathbb{R})$ given by $\begin{pmatrix} a_t & 0 \\ 0 & 1 \end{pmatrix}$ and by abuse of notation we denote it again by a_t. It will be convenient to normalize the flow a_t so that

$$\max_{1 \le i \le d} c_i = 1. \tag{2.1}$$

We let $a \stackrel{\text{def}}{=} a_1$ be the time-one map for the diagonal flow and denote by G^+ the unstable horospherical subgroup for a in G, that is,

$$G^+ \stackrel{\text{def}}{=} \left\{g \in G \mid a_t g a_t^{-1} \to e \text{ as } t \to -\infty\right\}.$$

We let $U^+ \stackrel{\text{def}}{=} U \cap G^+$ so that if $J^+ = \{i \in \{1, \dots, d\} \mid c_i > 0\}$, then we have

$$U^+ = \left\{\begin{pmatrix} I_d & u \\ 0 & 1 \end{pmatrix};\ u = {}^t(u_1, \dots, u_d) \in \mathbb{R}^d \text{ with } u_i = 0\ \forall i \notin J^+\right\}.$$

On U, we will use the Euclidean distance d_E inherited from $\mathbb{R}^d$. This metric induces a metric, still denoted by d_E, on the fiber $\pi^{-1}(x) \simeq \mathbb{R}^d/x$ of all grids lying above $x \in X$. Given a grid y, we define the *fiber injectivity radius* at y to be the maximal number $r_y > 0$ such that the orbit map $u \mapsto uy$ is injective on the open ball $B^{U,d_E}_{2r_y}(0)$ of radius $2r_y$ in U for the Euclidean metric d_E, therefore isometric on $B^{U,d_E}_{r_y}(0)$. Note that the fiber injectivity radius is constant along the fiber and is bounded away from zero on compact sets in Y.

On U^+, we will also make use of another metric, or rather, quasi-metric, more adapted to the action of a_t. We define the quasi-norm associated to a by $|u|_a = \max_{j \in J^+} |u_j|^{1/c_j}$. The function on $U^+ \times U^+$, given by $d_a(u, v) = |u - v|_a$, is a quasi-metric: it is symmetric, is positive definite, and satisfies, for some constant C depending on the c_i, for all u, v, w in U^+,

$$d_a(u, w) \leqslant C\,(d_a(u, v) + d_a(v, w)).$$

The ball $B^{U^+,d_a}_\delta(u)$ of radius δ around u for d_a is simply the set of $v \in U^+$ such that $d_a(u, v) < \delta$.

Remark 2.1. We observe two things:

(1) A ball B^{U^+,d_a}_δ in U^+ is simply a box with side-lengths $2\delta^{c_j}$, $j \in J^+$ with respect to d_E.

(2) The action of a_t on U^+ is a dilation by a factor of e^t for the quasi-metric d_a; that is, for all u, v in U^+ and $t \in \mathbb{R}$, we have

$$d_a(a_t u, a_t v) = e^t d_a(u, v). \tag{2.2}$$

We let W_y^+ be the image of y under the action of $B_{r_y}^{U^+,d_E}$. We call W_y^+ *the injective unstable leaf at y in the fiber*. By definition of r_y, the orbit map identifies $(B_{r_y}^{U^+,d_E}, d_E)$ and (W_y^+, d_E) isometrically. Of course, the quasi-metric d_a on U^+ also induces a quasi-metric on W_y^+, which we again denote by d_a. Since we will use both distances d_E and d_a on W_y^+ (which are far from being equivalent if $c_j < 1$ for some $j \in J^+$), we will indicate the metric in the superscript when necessary.

2.2 Dimensions

Let X be a space endowed with a quasi-metric d. For a bounded subset $S \subset X$ we will consider its lower Minkowski dimension (or lower box dimension) $\underline{\dim}_d S$ for the quasi-metric d, defined by

$$\underline{\dim}_d S \stackrel{\text{def}}{=} \liminf_{\delta \to 0} \frac{\log N_d(S, \delta)}{\log \frac{1}{\delta}},$$

where $N_d(S, \delta)$ is the maximal cardinality of a δ-separated subset of S for the quasi-metric d. If S is unbounded, we let $\underline{\dim}_d S = \sup\{\underline{\dim}_d S \cap K\ ;\ K \text{ compact}\}$.

In particular, for a set $S \subset W_y^+$, we will consider its lower Minkowski dimensions

$$\underline{\dim}_a S \stackrel{\text{def}}{=} \underline{\dim}_{d_a} S, \qquad \underline{\dim}_M S \stackrel{\text{def}}{=} \underline{\dim}_{d_E} S$$

for the quasi-metric d_a and the Euclidean metric d_E, respectively. We will also consider the Hausdorff dimension $\dim_H S$, always defined with respect to the Euclidean metric. We refer the reader to [7] for general properties of Minkowski or Hausdorff dimensions, such as the inequality

$$\underline{\dim}_M S \geq \dim_H S.$$

We introduce $\underline{\dim}_a$ in order to relate dimension $\underline{\dim}_M$ to entropy, and further to Hausdorff dimension. The following simple observation gives a relation between $\underline{\dim}_M S$ and $\underline{\dim}_a S$. Let

$$h_a = \sum_{i \in J^+} c_i.$$

Lemma 2.2. We have $\underline{\dim}_a U^+ = h_a$, and moreover, for any set $S \subset U^+$,

$$\underline{\dim}_a S \geqslant \underline{\dim}_M S + h_a - \dim U^+.$$

Proof. A δ-ball for d_a is a Euclidean box with side lengths $2\delta^{c_i}$, so that any bounded set in U^+ can be covered by $O(\delta^{-h_a})$ balls of radius δ for d_a. Conversely, for volume reasons, one needs at least $O(\delta^{-h_a})\delta$-balls for d_a to cover any nonempty open set in U^+. This shows the first equality.

For the general inequality, let $u = \dim U^+$. Each d_a-ball of radius δ can be covered by at most δ^{h_a-u} boxes of side lengths $2\ \delta$ which in turn can be covered by the same number of Euclidean balls (up to a multiplicative constant, say C). Thus $N_{d_a}(S,\delta) \geq C\delta^{h_a-u}N_{d_E}(S,\delta)$. Taking logarithms, dividing by $\log 1/\delta$, and taking $\delta \to 0$ give the result. ∎

2.3 Constructing a measure of large entropy

We refer the reader to [5, §2.2] for the definition of relative entropy with respect to an infinite countably generated σ-algebra; in particular, if $\mathcal{P}$ is any countable partition of Y, then $H_\mu(\mathcal{P}|X)$ will denote the relative entropy of $\mathcal{P}$ with respect to the σ-algebra $\pi^{-1}(\mathcal{B}_X)$, that is, the inverse image under π of the Borel σ-algebra $\mathcal{B}_X$ on X. Finally, $h_\mu(a|X)$ denotes the relative entropy of the transformation $a = a_1 : Y \to Y$ for the measure μ (relative to X), that is,

$$h_\mu(a|X) \stackrel{\text{def}}{=} \sup_{\mathcal{P}} \inf_{q\in\mathbb{Z}^+} \frac{1}{q} H_\mu\left(\mathcal{P}^{(q)}|X\right),$$

where the supremum runs over countable partitions $\mathcal{P}$ with $H_\mu(\mathcal{P}|X) < \infty$, and $\mathcal{P}^{(q)} = \bigvee_{i=0}^{q-1} a^{-i}\mathcal{P}$ denotes the join of the preimages $a^{-i}\mathcal{P}$. We can now state the main result of this section. The reader might benefit from reviewing Definition 1.1 first.

Proposition 2.3 (Lower bound on the entropy). Let x be a lattice in $\mathcal{H}(\eta_i)$ for some sequence $\eta_i \to 0$ and let $y \in \pi^{-1}(x)$. Furthermore, let

$$\mu^0 \in \left\{\delta_x^T : T \in \mathbb{Z}_+\right\}^\omega \cap \mathscr{P}(X, (\eta_i)).$$

For any $S \subset W_y^+$, where $W_y^+ \subset \pi^{-1}(x)$ is the injective unstable leaf at y in the fiber, there exists an a-invariant $\mu \in \mathscr{P}(Y)$ satisfying:

(1) $\pi_*\mu = \mu^0$,

(2) $\operatorname{Supp}\mu \subset \bigcap_{s\in\mathbb{Z}^+} \overline{\bigcup_{t\geq s} a_t S}$,

(3) $h_\mu(a|X) \geqslant \underline{\dim}_a S \geqslant \underline{\dim}_M S + h_a - \dim U^+$.

Furthermore, if $\mathcal{P}$ is any finite partition of Y satisfying:

- $\mathcal{P}$ contains an atom P_∞ of the form $\pi^{-1}(P_\infty^0)$, where $X \setminus P_\infty^0$ has compact closure and $\psi_i|_{P_\infty^0} \equiv 0$ for some i (ψ_i is as in Definition 1.1).

- $\forall P \in \mathcal{P} \setminus \{P_\infty\}$, $\operatorname{diam} P < r$, with $r \in (0, \frac{1}{2})$ such that any d_a-ball of radius $3r$ has Euclidean diameter smaller than the fiber injectivity radius on $Y \setminus P_\infty$,
- $\forall P \in \mathcal{P}$, $\mu(\partial P) = 0$,

then, for all $q \geqslant 1$, $\frac{1}{q} H_\mu(\mathcal{P}^{(q)}|X) \geqslant \underline{\dim}_a S - D\eta_i$, where D is a constant depending only on the c_i's and the dimension d.

The proof of Proposition 2.3 will follow the strategy used to derive the variational principle for the topological entropy, as in [5, §5.3.3], but there is a slight complication here, because the space Y of grids is not compact. To solve this problem, we will need Lemma 2.4 below, which is inspired by [4, Lemma 4.5].

Lemma 2.4. Let $P_\infty^0 \subset X$ be such that $X \setminus P_\infty^0$ has compact closure. Set $P_\infty = \pi^{-1}(P_\infty^0)$ and fix $0 < r < 1$ such that any d_a-ball of radius $3r$ has Euclidean diameter smaller than the fiber injectivity radius on $Y \setminus P_\infty$. Let $y \in Y \setminus P_\infty$ and set $I = \{t \in \mathbb{Z}^+ \mid a_t y \in P_\infty\}$. For any nonnegative integer T, let

$$E_{y,T} = \left\{ z \in W_y^+ \mid \forall t \in \{1, \dots, T\} \setminus I,\ d_E(a_t y, a_t z) \leqslant r \right\}.$$

Then one can cover $E_{y,T}$ by $Ce^{D|I\cap\{1,\dots,T\}|}$ d_a-balls of radius $\delta_T = re^{-T}$, where C is a constant depending on y, r, and a, and D is a constant depending on a and the dimension d. In particular, they are independent of T.

Proof. Before we start the proof, we make the following observation: for $y', z' \notin P_\infty$ in the same fiber, the intersection $B_r^{Uy', d_E}(y') \cap W_{z'}^+$ is contained in $B_r^{W_{z'}^+, d_E}(z'')$ for some $z'' \in W_{z'}^+$ since the Euclidean distance d_E on the fiber $Uy' = \pi^{-1}(\pi(y'))$ restricts to a Euclidean distance on the injective unstable leaf at z'. Moreover, since $r < 1$ and by (2.1), Euclidean r-balls are contained in d_a-balls and so we conclude that

$$B_r^{Uy', d_E}(y') \cap W_{z'}^+ \subset B_r^{W_{z'}^+, d_a}(z''). \tag{2.3}$$

We prove the lemma by induction on T.

$\underline{T = 0}$: By (1) of Remark 2.1, the number of balls of radius $\delta_0 = r$ for the metric d_a needed to cover W_y^+ is bounded by a integer constant C depending on a, r, and y, so that the lemma holds in this case.

$\underline{T - 1 \to T}$: Choose D such that any d_a-ball of radius δ on U^+ can be covered by e^D d_a-balls of radius δ/e. Assume for clarity that e^D is an integer. By the induction hypothesis, $E_{y,T-1}$ can be covered by $N_{T-1} \stackrel{\text{def}}{=} Ce^{D|I\cap\{1,\dots,T-1\}|}$ d_a-balls of radius $\delta_{T-1} = re^{-T+1}$.

If $T \in I$, we simply cover each δ_{T-1}-ball in W_y^+ by e^D balls of radius $\delta_T = re^{-T}$ for d_a, and get a cover $E_{y,T}$ with cardinality $N_T = e^D N_{T-1}$ by d_a-balls.

If $T \notin I$, we need to cover $E_{y,T}$ by $N_T = N_{T-1} d_a$-balls of radius δ_T. Denote the above cover of $E_{y,T-1}$ by $\{B_{\delta_{T-1}}^{W_y^+,d_a}(z_i); i = 1, \dots, N_{T-1}\}$. As $E_{y,T} \subset E_{y,T-1}$, the set $\{E_{y,T} \cap B_{\delta_{T-1}}^{W_y^+,d_a}(z_i); i = 1, \dots, N_{T-1}\}$ covers $E_{y,T}$. We claim that for each z_i, there exists some p_i with

$$E_{y,T} \cap B_{\delta_{T-1}}^{W_y^+,d_a}(z_i) \subset B_{\delta_T}^{W_z^+,d_a}(p_i),$$

so that $E_{y,T}$ is actually covered by $\{B_{\delta_T}^{W_z^+,d_a}(p_i)\}$ that is, by $N_T = N_{T-1} d_a$-balls of radius δ_T. Observe that

$$E_{y,T} \cap B_{\delta_{T-1}}^{W_z^+,d_a}(z) \subset a_T^{-1}\left(B_r^{Ua_T y,d_E}(a_T y) \cap a_T B_{\delta_{T-1}}^{W_z^+,d_a}(z)\right).$$

By our choice of r, the fact that $a_T y \notin P_\infty$ (and hence $a_T z \notin P_\infty$) and (2.2), the map a_T scales d_a by a factor of e^T and we conclude that

$$a_T B_{\delta_{T-1}}^{W_z^+,d_a}(z) \cap B_r^{Ua_T y,d_E}(a_T y) = B_{er}^{W_{a_T z}^+,d_a}(a_T z) \cap B_r^{Ua_T y,d_E}(a_T y) \tag{2.4}$$

which is contained in a single d_a-ball of radius r by the observation (2.3) (with $z' = a_T z$, $y' = a_T y$). Thus $E_{y,T} \cap B_{\delta_{T-1}}^{W_z^+,d_a}(z)$ is contained in a single d_a-ball of radius re^{-T}. This concludes the inductive step. ■

Now we can prove Proposition 2.3.

Proof of Proposition 2.3. The assumption that $\mu^0 \in \left\{\delta_x^T : T \in \mathbb{Z}_+\right\}^\omega \cap \mathscr{P}(X, (\eta_i))$ means that we may fix an increasing sequence of integers (n_k) such that

$$\mu_k^0 \stackrel{\text{def}}{=} \frac{1}{n_k} \sum_{n=0}^{n_k-1} \delta_{a^n x} \xrightarrow{\text{w}^*} \mu^0 \in \mathscr{P}(X, (\eta_i)).$$

Then, for each $k \geqslant 1$, let S_k be a maximal ρ_k-separated subset of S, for the metric d_a, where $\rho_k \stackrel{\text{def}}{=} e^{-n_k}$. Let $\nu_k \stackrel{\text{def}}{=} \frac{1}{|S_k|} \sum_{y \in S_k} \delta_y$ be the normalized counting measure on S_k and

$$\mu_k \stackrel{\text{def}}{=} \frac{1}{n_k} \sum_{n=0}^{n_k-1} a_*^n \nu_k.$$

Since $\pi : Y \to X$ is proper, and the measures $\pi_* \mu_k = \mu_k^0$ converge to a probability measure μ^0 on X, we conclude that the sequence of measures (μ_k) is tight, that is, that any weak-* converging subsequence of it converges to a probability measure. Extracting a subsequence if necessary, we may assume without loss of generality that (μ_k) converges weak-* to some probability measure, which we denote by μ. By continuity of π_* we obtain

$$\pi_*\mu = \pi_* \lim_k \mu_k = \lim_k \mu_k^0 = \mu^0$$

which is the item (1) in the proposition.

By construction, Suppμ is contained in the set of accumulation points of the forward orbit of S under (a_t) which establishes item (2) in the proposition.

The right inequality in item (3) follows directly from Lemma 2.2. For simplicity of notation, let $\beta \stackrel{\text{def}}{=} \underline{\dim}_a S$, so that

$$\liminf_{k\to\infty} \frac{\log |S_k|}{n_k} \geqslant \beta. \tag{2.5}$$

To prove that μ also satisfies the the left inequality in item (3) of the proposition saying that $h_\mu(a|X) \geqslant \beta$ we proceed as follows. Given i we construct a partition $\mathcal{P}$ of Y for which

$$\frac{1}{q} H_\mu \left(\mathcal{P}^{(q)}|X\right) \geqslant \beta - D\eta_i, \tag{2.6}$$

where D is as in Lemma 2.4. This implies that

$$h_\mu(a|X) \geqslant h_\mu\left(a, \mathcal{P}|X\right) = \lim_q \frac{1}{q} H_\mu \left(\mathcal{P}^{(q)}|X\right) \geqslant \beta - D\eta_i,$$

and letting $i \to \infty$ we obtain the desired inequality $h_\mu(a|X) \geqslant \beta$.

To this end fix i_0 and consider $\eta = \eta_{i_0}$. Choose a set $P_\infty^0 \subset X$ such that $X \setminus P_\infty^0$ is compact, $\mu^0(\partial P_\infty^0) = 0$, and such that $\psi_{i_0}|_{P_0^\infty} \equiv 0$. Since $\mu^0 \in \mathscr{P}(X, (\eta_i))$ it follows that $\mu^0(P_\infty^0) \leqslant \eta_{i_0} = \eta$ and in turn, $P_\infty = \pi^{-1}(P_\infty^0)$ satisfies $\mu(P_\infty) \leqslant \eta$ and $\mu(\partial P_\infty) = 0$. Since (μ_k) converges to μ, we have for any $y \in \pi^{-1}(x)$, with $I = \{t \in \mathbb{Z}^+ \mid a_t y \in P_\infty\}$ (note that I depends only on x),

$$\limsup_{T\to\infty} \frac{1}{n_k} |I \cap \{1, \ldots, n_k\}| \leqslant \eta. \tag{2.7}$$

Then, let $r \in (0, \frac{1}{2})$ be as in Lemma 2.4 and complement P_∞ to a finite partition $\mathcal{P} = \{P_\infty, P_1, \ldots, P_\ell\}$ of Y such that for every atom $P_i \neq P_\infty$ and every $x \in X$, the Euclidean diameter of $P_i \cap \pi^{-1}(x)$ is at most r, and such that for each $P \in \mathcal{P}$, $\mu(\partial P) = 0$, where ∂P

denotes the boundary of P. To build such $\mathcal{P}$, observe that around each point z in $Y \setminus P_\infty$, there exists a ball B_z around z in Y such that $\mu(\partial B_z) = 0$, and

$$\forall x \in X,\ \operatorname{diam}_E(B_z \cap \pi^{-1}(x)) < r. \tag{2.8}$$

A finite cover of $Y \setminus P_\infty$ by such balls generates the desired partition by a simple disjointification procedure.

For $q \geqslant 1$, let $\mathcal{P}^{(q)} = \bigvee_{p=0}^{q-1} a^{-p}\mathcal{P}$. For n_k large, write the Euclidean division of $n_k - 1$ by q

$$n_k - 1 = qn' + s,\ \text{with } s \in \{0, \dots, q-1\}.$$

By subadditivity of the entropy with respect to the partition, for each $p \in \{0, \dots, q-1\}$,

$$H_{\nu_k}\left(\mathcal{P}^{(n_k)}|X\right) \leqslant H_{a^p\nu_k}\left(\mathcal{P}^{(q)}|X\right) + H_{a^{p+q}\nu_k}\left(\mathcal{P}^{(q)}|X\right) + \cdots + H_{a^{p+qn'}\nu_k}\left(\mathcal{P}^{(q)}|X\right) + 2q\log|\mathcal{P}|.$$

Summing those inequalities for $p = 0, \dots, q-1$, and using the fact that entropy is a concave function of the measure, we obtain

$$\begin{aligned} qH_{\nu_k}\left(\mathcal{P}^{(n_k)}|X\right) &\leqslant \sum_{n=0}^{n_k-1} H_{a^n\nu_k}\left(\mathcal{P}^{(q)}|X\right) + 2q^2\log|\mathcal{P}| \\ &\leqslant n_k H_{\mu_k}\left(\mathcal{P}^{(q)}|X\right) + 2q^2\log|\mathcal{P}| \end{aligned}$$

and therefore

$$\frac{1}{q}H_{\mu_k}\left(\mathcal{P}^{(q)}|X\right) \geqslant \frac{1}{n_k}H_{\nu_k}\left(\mathcal{P}^{(n_k)}|X\right) - \frac{2q\log|\mathcal{P}|}{n_k}. \tag{2.9}$$

Now since ν_k is supported on a single atom of the σ-algebra $\pi^{-1}(\mathcal{B}_X)$, we have $H_{\nu_k}(\mathcal{P}^{(n_k)}|X) = H_{\nu_k}(\mathcal{P}^{(n_k)})$. Moreover, we claim that

$$H_{\nu_k}\left(\mathcal{P}^{(n_k)}\right) \geqslant \log|S_k| - D|I \cap \{1, \dots, n_k\}| - D - \log C, \tag{2.10}$$

where C, D are the constants given by Lemma 2.4. To see this, it suffices to show that an atom of $\mathcal{P}^{(n_k)}$ contains at most $Ce^{D(|I\cap\{1,\dots,n_k\}|+1)}$ points of $S_k = \operatorname{Supp}\nu_k$. This follows from Lemma 2.4. Indeed, Equation (2.8) implies that if P is any nonempty atom of $\mathcal{P}^{(n_k)}$, fixing any $y \in P$,

$$S_k \cap P = S_k \cap [y]_{\mathcal{P}^{(n_k)}} \subset E_{y,n_k-1}$$

can be covered by $Ce^{D(|I\cap\{1,\dots,n_k\}|+1)}$ many re^{-n_k}-balls for d_a. Since S_k is $\rho_k = e^{-n_k}$-separated with respect to d_a and $r < \frac{1}{2}$, we get

$$\operatorname{card}\left(S_k \cap [y]_{\mathcal{P}^{(n_k)}}\right) \leqslant Ce^{D(|I\cap\{1,\dots,n_k\}|+1)}.$$

Going back to (2.9), we find

$$\frac{1}{q}H_{\mu_k}\left(\mathcal{P}^{(q)}|X\right) \geqslant \frac{1}{n_k}\left(\log|S_k| - D|I\cap\{1,\dots,n_k\}| - D - \log C - 2q^2\log|\mathcal{P}|\right).$$

Now the atoms of $\mathcal{P}$ – and hence of $\mathcal{P}^{(q)}$ – satisfy $\mu(\partial P) = 0$, so we may let k go to infinity to obtain equation (2.6) using (2.5) and (2.7).

Finally the second part of the proposition regarding partitions satisfying the bullet-requirements follows by reviewing the proof of (2.6) and noting that the only properties of the constructed partition $\mathcal{P}$ we used are those in the bullet list. ■

3 Maximal Entropy and Invariance

In this section we recall some concepts and results from [3] and explain how they imply the following proposition, which is essential for the proof of Theorem 1.3.

Proposition 3.1 (Maximal entropy implies U-invariance). Let μ be an a-invariant probability measure on Y. Then

$$h_\mu(a|X) \leqslant h_a$$

with equality if and only if μ is U-invariant.

To prove Proposition 3.1 we relate the "dynamical" relative entropy $h_\mu(a|X)$ to some relative "static" entropy $H_\mu(\mathcal{A}_1|\mathcal{A}_2)$ where the $\mathcal{A}_i$ are countably generated σ-algebras that encode the dynamics. For definitions and elementary properties of relative entropies of σ-algebras we refer the reader to [5, Chapter 2].

Definition 3.2 (7.25. of [3]**).** Let $G^- \overset{\text{def}}{=} \{g \in G \mid a_t g a_t^{-1} \to e \text{ as } t \to \infty\}$ be the stable horospherical subgroup associated to a and let $U^- = U \cap G^-$. Let μ be an a-invariant measure on Y and $U < G^-$ a closed a-normalized subgroup.

(1) We say that a countably generated σ-algebra $\mathcal{A}$ is *subordinate to U* (mod μ) if for μ-a.e. y, there exists $\delta > 0$ such that

$$B^U_\delta \cdot y \subset [y]_{\mathcal{A}} \subset B^U_{\delta^{-1}} \cdot y.$$

(2) We say that $\mathcal{A}$ is *a-descending* if $a^{-1}\mathcal{A} \subset \mathcal{A}$.

Theorem 3.3 (Einsiedler–Lindenstrauss). Let μ be an a-invariant probability measure on Y. If $\mathcal{A}$ is a countably generated sub-σ-algebra of the Borel σ-algebra which is a-descending and U^--subordinate then $H_\mu(\mathcal{A}|a^{-1}\mathcal{A}) \leqslant h_a$ with equality if and only if μ is U^--invariant.

Proof. By considering the ergodic decomposition one sees that it is enough to prove the theorem in the case μ is ergodic. Under the ergodicity assumption the statement follows directly from combining [3, Proposition 7.34] and [3, Theorem 7.9]. ■

The following lemma furnishes the link between the relative dynamical entropy $h_\mu(a|X)$ and the relative entropy $H_\mu(\mathcal{A}_1|\mathcal{A}_2)$ for suitable $\mathcal{A}_i$. If $\mathcal{P}$ is a partition of Y, we write for any integer m,

$$\mathcal{P}_m^\infty = \bigvee_{k=m}^{\infty} a^{-k}\mathcal{P}.$$

Recall also that a partition $\mathcal{P}$ is said to be a two-sided generator for a with respect to an a-invariant probability measure if the Borel σ-algebra is generated up to null sets by the union of all partitions $\bigvee_{-m}^{m} a^{-k}\mathcal{P}$, $m \geqslant 1$.

Lemma 3.4. Assume that μ is an a-invariant probability measure on Y and $\mathcal{P}$ is a countable partition that is a two-sided generator for a with respect to μ. Let $\mathcal{A}$ be the σ-algebra generated by $\mathcal{P}_0^\infty \vee \pi^{-1}(\mathcal{B}_X)$. Then

$$h_\mu(a|X) = H_\mu(\mathcal{A}|a^{-1}\mathcal{A}).$$

Proof. Let $\mathcal{P}$ be a countable partition that is a two-sided generator for a. By [5, Proposition 2.19 and Theorem 2.20], writing $\mathcal{C}$ for $\pi^{-1}(\mathcal{B}_X)$, we have

$$h_\mu(a|X) = H_\mu(\mathcal{P}|\mathcal{P}_1^\infty \vee \mathcal{C}) = H_\mu(\mathcal{P}_0^\infty \vee \mathcal{C}|\mathcal{P}_1^\infty \vee \mathcal{C}).$$

Since $\mathcal{C}$ is a-invariant, we indeed find

$$h_\mu(a|X) = H_\mu(\mathcal{A}|a^{-1}\mathcal{A}).$$

■

We now prove Proposition 3.1 by constructing (almost by citation) a two-sided generator $\mathcal{P}$ for a modulo μ with the property that $\mathcal{P}_0^\infty \vee \pi^{-1}(\mathcal{B}_X)$ is U^--subordinate, and then the statement follows by combining Theorem 3.3 with Lemma 3.4.

Proof of Proposition 3.1. Writing the ergodic decomposition $\mu = \int \mu_y^{\mathcal{E}}\,\mathrm{d}\mu(y)$, we have

$$h_\mu(a|X) = \int h_{\mu_y^{\mathcal{E}}}(a|X)\,\mathrm{d}\mu(y),$$

so it is enough to prove the proposition for μ ergodic.

We can then use [3, Proposition 7.44] to obtain a countable partition $\mathcal{P}$ that is a generator for a mod μ, and such that $\mathcal{P}_0^\infty$ is a-descending and subordinate to G^-. Let

$$\mathcal{C} \overset{\text{def}}{=} \pi^{-1}(\mathcal{B}_X) \quad \text{and} \quad \mathcal{A} \overset{\text{def}}{=} \mathcal{C} \vee \mathcal{P}_0^\infty.$$

Then $\mathcal{A}$ is clearly countably generated and a-descending, and we claim that it is also U^--subordinate. Indeed, since $[y]_{\mathcal{C}} = \pi^{-1}(\pi(\{y\}))$ is equal to the orbit of y under the unipotent radical U, we get $[y]_{\mathcal{A}} = U\cdot y \cap [y]_{\mathcal{P}_0^\infty}$. But $\mathcal{P}_0^\infty$ is subordinate to G^-, and for $\delta > 0$,

$$B_{\delta_1}^{U^-}\cdot y \subset B_\delta^{G^-}\cdot y \cap U\cdot y \quad \text{and} \quad B_{\delta^{-1}}^{G^-}\cdot y \cap U\cdot y \subset B_{\delta_1^{-1}}^{U^-}\cdot y, \tag{3.1}$$

for some constant δ_1 depending on δ and y, because $G^-y \cap Uy = U^-y$. This shows that, for almost every y, for some $\delta > 0$,

$$B_\delta^{U^-}\cdot y \subset [y]_{\mathcal{A}} \subset B_{\delta^{-1}}^{U^-}\cdot y,$$

that is, that $\mathcal{A}$ is subordinate to U^-.

By Lemma 3.4 we have $h_\mu(a|X) = H_\mu(\mathcal{A}|a^{-1}\mathcal{A})$, and so Theorem 3.3 shows that $h_\mu(a|X) \leqslant h_a$ with equality if and only if μ is U^--invariant. Moreover, if $h_\mu(a|X) = h_a$ then

$$h_\mu(a^{-1}|X) = h_\mu(a|X) = h_a = h_{a^{-1}},$$

where the last equality follows from the fact that $\sum_{j=1}^d c_j = 0$. We then apply the same reasoning to a^{-1} and find that μ is also U^+-invariant. This completes the proof since U is generated by U^+ and U^-, because we assume that for any $1 \leqslant j \leqslant d$, $c_j \neq 0$. ■

4 Proof of the Main Theorem

Using the results of the preceding two sections, we now state and prove Theorem 4.2, which is the main result of this article. We deduce Theorem 1.3 as a corollary.

4.1 $\mathcal{O}$-avoiding grids

Before we can state the more precise version of Theorem 1.3 that we will derive here, we need to set up some notation.

For a bounded open set $\mathcal{O} \subset \mathbb{R}^d$ and a heavy lattice x, our goal is to bound the dimension of the set

$$F_{S^+(\mathcal{O})}(x) = \left\{ y \in \pi^{-1}(x) : y \cap (\cup_{t\geq 0} a_{-t}\mathcal{O}) \text{ is finite} \right\}.$$

For an interval $I \subset \mathbb{R}$, let

$$E_I^{\mathcal{O}} = \{ y \in Y : y \cap (\cup_{t\in I} a_{-t}\mathcal{O}) = \varnothing \} = \{ y \in Y : \forall t \in I, a_t y \cap \mathcal{O} = \varnothing \}$$

and

$$E_I^{\mathcal{O}}(x) = E_I^{\mathcal{O}} \cap \pi^{-1}(x).$$

It is more natural from the dynamical point of view to work with $E_{\mathbb{R}}^{\mathcal{O}}$ than with $F_{S^+(\mathcal{O})}$ because it is a_t-invariant and closed. However, we insist on working with finite intersections instead of empty intersection to obtain lim inf statements in our applications rather than inf. Note that for a vector $v \in \mathbb{R}^d$, either $a_t v \to \infty$ or $v_i = 0$ for all $i \in J^+$ in which case $a_t v \to 0$. Since all c_j are nonzero and $\mathcal{O}$ is bounded, we deduce that

$$F_{S^+(\mathcal{O})}(x) \subset \bigcup_{r\geq 0} E_{[r,\infty)}^{\mathcal{O}}(x) \cup \left\{ y \in \pi^{-1}(x) : \exists v \in y, \forall i \in J^+, v_i = 0 \right\}. \tag{4.1}$$

Since

$$\dim_H \left\{ y \in \pi^{-1}(x) : \exists v \in y, \forall i \in J^+, v_i = 0 \right\} = d - |J^+| \leqslant d - 1, \tag{4.2}$$

we will focus on bounding the dimension of each $E_{[r,\infty)}^{\mathcal{O}}(x)$. Finally, a nice feature of the set $E_{[r,\infty)}^{\mathcal{O}}$ is the following simple observation which can be verified by the reader.

Lemma 4.1. Let $y \in E_{[r,\infty)}^{\mathcal{O}}$ and suppose z is an accumulation point of the forward orbit $\{a_t y\}_{t>0}$. Then $z \in E_{\mathbb{R}}^{\mathcal{O}}$. In particular, any measure μ obtained by averaging along the forward trajectory of y is supported in $E_{\mathbb{R}}^{\mathcal{O}}$.

We are now in a position to state and prove our main results.

Theorem 4.2 (Heavy lattices have few bad grids). Fix a sequence $\eta_i \to 0$, and a bounded open set $\mathcal{O}$ in $\mathbb{R}^d$. Then, there exists $\delta > 0$ such that for any $r > 0$ and any $x \in \mathcal{H}(\eta_i)$,

$$\underline{\dim}_M E^{\mathcal{O}}_{[r,\infty)}(x) \leqslant d - \delta.$$

Proof of Theorem 1.3. Using (4.1) and (4.2), the theorem follows at once from Theorem 4.2, because $E^{\mathcal{O}}_{[r,\infty)}$ is increasing in r. Note that as opposed to $\underline{\dim}_M$, the Hausdorff dimension of a countable union of sets is bounded above by any bound on the individual dimensions, hence our passage to Hausdorff dimension (see also Remark 4.3 below). ■

Remark 4.3. The set $F_{S^+(\mathcal{O})}(x)$ is dense is $\pi^{-1}(x)$, so that $\underline{\dim}_M F_{S^+(\mathcal{O})}(x) = d$. But its Hausdorff dimension is strictly less than d.

Proof of Theorem (4.2). Fix $\epsilon > 0$. We argue by contradiction and assume the following:

$$\forall m \geq 1, \exists r_m > 0, \exists x_m \in \mathcal{H}(\eta_i) : \ \underline{\dim}_M E^{\mathcal{O}}_{[r_m,\infty)}(x_m) \geqslant d - \frac{1}{m}. \tag{4.3}$$

We may further assume that $r_m \to \infty$ as $m \to \infty$. Fix a smaller open set $\mathcal{O}'$ whose closure is in $\mathcal{O}$. Set $u = \dim U^+$.

Claim 1. *For any large enough m, there exists a grid $y_m \in \pi^{-1}(x_m)$ such that the injective unstable leaf $W^+_{y_m}$ satisfies*

$$\underline{\dim}_M \left(W^+_{y_m} \cap E^{\mathcal{O}'}_{[r_m,\infty)}\right) \geqslant u - \frac{1}{m}.$$

Recall that $U^- = U \cap G^- = U \cap \{g \in G \mid a_t g a_t^{-1} \to e \text{ as } t \to \infty\}$. For any grid $y \in Y$, any $v \in U^-$ of norm $\leqslant 1$ and $t > 0$, the two grids $a_t y$ and $a_t v y$ differ by a translation in the direction of U^- which is of norm $\leqslant e^{-\alpha t}$, where $\alpha = \min\left\{|c_i| : i \in J^-\right\} > 0$. We deduce that if $a_t y \cap \mathcal{O} = \varnothing$ and t is large enough, then $a_t v y \cap \mathcal{O}' = \varnothing$. In particular, for large enough m and and all $v \in U^-$ of norm $\leqslant 1$, we have $vE^{\mathcal{O}}_{[r_m,\infty)} \subset E^{\mathcal{O}'}_{[r_m,\infty)}$. For such m and for any $y_m \in E^{\mathcal{O}}_{[r_m,\infty)}$,

$$\underline{\dim}_M \left(W^+_{y_m} \cap E^{\mathcal{O}'}_{[r_m,\infty)}\right) + \dim U^- \geqslant d - \frac{1}{m}.$$

Since $c_i \neq 0$ for all i , $d - \dim U^- = u$ and Claim 1 follows.

The first part of Proposition 2.3 yields an a-invariant probability measure μ_m with

(1) $\pi_*\mu_m \in \mathscr{P}(X, (\eta_i))$.
(2) $\operatorname{Supp}\mu_m \subset E_{\mathbb{R}}^{\mathcal{O}'}$ (by Proposition 2.3 and Lemma 4.1)
(3) $h_{\mu_m}(a|X) \geqslant h_a - \frac{1}{m}$.

Lemma 1.2(3) says that $\mathscr{P}(X, (\eta_i))$ is compact. Together with the fact that $\pi : Y \to X$ is proper, after taking a subsequence of μ_m, we may assume that μ_m converges to some a-invariant probability measure μ such that $\pi_*\mu \in \mathscr{P}(X, (\eta_i))$. Since $\operatorname{Supp}\mu_m \subset E_{\mathbb{R}}^{\mathcal{O}'}$ and $E_{\mathbb{R}}^{\mathcal{O}'}$ is closed, we conclude that μ is supported in $E_{\mathbb{R}}^{\mathcal{O}'}$. Our next goal is to show the following.

Claim 2. *The measure μ is U-invariant.*

To prove Claim 2 we apply the second part of Proposition 2.3 to the measures μ_m simultaneously in the following manner. Fix i_0 and let $\eta = \eta_{i_0}$. Let $\mathcal{P}$ be a finite partition of Y satisfying:

- $\mathcal{P}$ contains a single unbounded atom P_∞ and it is of the form $\pi^{-1}(P_\infty^0)$, where $\psi_{i_0}|_{P_\infty^0} \equiv 0$ (ψ_i is as in Definition 1.1).
- $\forall P \in \mathcal{P} \setminus \{P_\infty\}$, $\operatorname{diam} P < r$, with $r \in (0, \frac{1}{2})$ such that any d_a-ball of radius $3r$ has Euclidean diameter smaller than the fiber injectivity radius on $Y \setminus P_\infty$,
- $\forall P \in \mathcal{P}$, $\nu(\partial P) = 0$ for $\nu \in \{\mu, \mu_m : m \in \mathbb{N}\}$.

A similar partition was built in the proof of Proposition 2.3. The only difference here is that in the third bullet here we demand that the boundaries of the atoms of $\mathcal{P}$ will be simultaneous null sets for more than one measure. Since we are only requiring this for a countable collection of measures this is easily achieved.

From the second part of Proposition 2.3 we deduce that for any m, for all $q \geqslant 1$,

$$\frac{1}{q} H_{\mu_m}\left(\mathcal{P}^{(q)}|X\right) \geqslant h_a - \frac{1}{m} - D\eta_{i_0}. \tag{4.4}$$

Since the boundary of the atoms of $\mathcal{P}^{(q)}$ is μ-null, we can pass to the limit as $m \to \infty$ in (4.4) and deduce that for any q, $\frac{1}{q} H_\mu(\mathcal{P}^{(q)}|X) \geqslant h_a - D\eta_{i_0}$. Taking $q \to \infty$ and $i_0 \to \infty$ we deduce that

$$h_\mu(a|X) \geqslant \lim \frac{1}{q} H_\mu\left(\mathcal{P}^{(q)}|X\right) \geqslant h_a.$$

By Proposition 3.1 we deduce that equality holds and that μ is U-invariant, as claimed.

We arrive at the desired contradiction because μ is supported in $E_{\mathbb{R}}^{\mathcal{O}'}$, which cannot contain a full fiber: given a lattice x, the grids of x which contain points in $\mathcal{O}'$ cannot be in $E_{\mathbb{R}}^{\mathcal{O}'}$. ■

5 Diophantine Approximation

In this section we prove Theorem 1.5, which, in fact, will follow from a sharper result in the spirit of Theorem 4.2. We also reformulate and generalize the result in terms of approximation of affine subspaces of $\mathbb{R}^n$ by integer points.

5.1 Inhomogeneous Diophantine approximation of vectors in $\mathbb{R}^n$

Fix a dimension $n \geqslant 1$ and let $d \overset{\text{def}}{=} n+1$. For clarity of exposition, we start with the diagonal flow $a_t = \operatorname{diag}\left(e^t, \ldots, e^t, e^{-nt}\right)$. Recall that given a vector $v \in \mathbb{R}^n$, we let

$$\mathbf{Bad}^{\epsilon}(v) = \left\{ w \in \mathbb{R}^n \ : \ \liminf_{k \to \infty} k^{1/n} \langle kv - w \rangle \geqslant \epsilon \right\}.$$

Given a vecitor $v \in \mathbb{R}^n$ we let $x_v \overset{\text{def}}{=} \left(\begin{smallmatrix} I_n & v \\ 0 & 1 \end{smallmatrix}\right) x_0 \in X$, where x_0 denotes the identity coset, which represents the standard lattice $\mathbb{Z}^d$. The Diophantine properties of the vector v are usually captured by the dynamics of the lattice x_v. For example, singularity of v is equivalent to the divergence of the orbit $(a_t x_v)_{t>0}$. In analogy with Definition 1.1 we make the following.

Definition 5.1. A vector $v \in \mathbb{R}^n$ is said to be *heavy* if the lattice x_v is heavy according to Definition 1.1.

A nice exercise is the following characterization of heaviness of a number $\alpha \in \mathbb{R}$ in terms of the continued fraction expansion of α.

Exercise 5.2. Show that a number $\alpha = [a_0; a_1, a_2, \ldots]$ is heavy if and only if

$$\forall \delta > 0 \ \exists \epsilon > 0 \text{ such that } \liminf_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \max\left\{\log \epsilon a_k, 0\right\} \leqslant \delta.$$

We prove the following result, which will easily imply Theorem 1.5.

Theorem 5.3 (Heavy vectors have few badly approximable points). If $v \in \mathbb{R}^n$ is heavy then for any $\epsilon > 0$, $\dim_H(\mathbf{Bad}^{\epsilon}(v)) < n$. In fact, if $\eta_i \to 0$ is a sequence of nonnegative numbers, then for any $\epsilon > 0$ there exists $\delta = \delta(\epsilon, (\eta_i)) > 0$ such that for any $v \in \mathbb{R}^n$ for which $x_v \in \mathcal{H}(\eta_i)$, $\dim_H(\mathbf{Bad}^\epsilon(v)) \leqslant n - \delta$.

Proof. We write vectors in $\mathbb{R}^d = \mathbb{R}^n \times \mathbb{R}$ as $w_s \overset{\text{def}}{=} \binom{w}{s}$ with $w \in \mathbb{R}^n, s \in \mathbb{R}$. Let (η_i) be as in the statement and let $v \in \mathbb{R}^n$ be such that $x_v \in \mathcal{H}(\eta_i)$. Let $\epsilon > 0$ and let $\mathcal{O} \overset{\text{def}}{=} \left\{w_s \in \mathbb{R}^d : s \in (0,1), \|w\| < \frac{\epsilon}{2}\right\}$. Note that

$$S^+(\mathcal{O}) = \left\{w_s : s \geqslant 1, s^{1/n}\|w\| < \frac{\epsilon}{2}\right\} \cup \mathcal{O}. \tag{5.1}$$

We know by Theorem 1.3 that there exists $\delta = \delta(\epsilon, (\eta_i)) > 0$ such that $\dim_H F_{S^+(\mathcal{O})}(x_v) \leqslant d - \delta$. We will show that for any $w \in \mathbf{Bad}^\epsilon(v)$ and for any $s \in [0, 1]$, the grid $x_v - w_s$ belongs to $F_{S^+(\mathcal{O})}(x_v)$. This will finish the proof.

To this end, let $w \in \mathbf{Bad}^\epsilon(v)$ and $s \in [0, 1]$. Note that

$$x_v - w_s = \bigcup_{k\in\mathbb{Z}} \left\{ \begin{pmatrix} \vec{m} + kv - w \\ k - s \end{pmatrix} : \vec{m} \in \mathbb{Z}^n \right\}.$$

We call k the *layer parameter* of $\left\{ \begin{pmatrix} \vec{m} + kv - w \\ k - s \end{pmatrix} : \vec{m} \in \mathbb{Z}^n \right\}$. Note that the set of vectors in each layer is discrete. Therefore if we suppose that $x_v - w_s$ intersects $S^+(\mathcal{O})$ in infinitely many points, then we conclude that $S^+(\mathcal{O})$ must contain points in arbitrarily high layers (i.e., with k arbitrarily large). In particular, the description of $S_+(\mathcal{O})$ given in (5.1) implies that there exist arbitrarily large $k > 0$ and vectors $\vec{m} \in \mathbb{Z}^n$ such that

$$(k - s)^{1/n}\|\vec{m} + kv - w\| \leqslant \frac{\epsilon}{2}.$$

In particular, $\liminf_{k\to\infty} k^{1/n}\langle kv - w\rangle < \epsilon$ and so $w \notin \mathbf{Bad}^\epsilon(v)$ contradicting our assumption. We deduce that $\#\left\{x_v - w_s \cap S^+(\mathcal{O})\right\} < \infty$, that is, $x_v - w_s \in F_{S^+(\mathcal{O})}(x_v)$ as claimed. ∎

As a corollary, we now derive Theorem 1.5 from the introduction, which we recall here for convenience.

Corollary 5.4. For any $\epsilon > 0$, there exists $\delta > 0$ such that for almost every $v \in \mathbb{R}^n$, $\dim_H \mathbf{Bad}^\epsilon(v) < n - \delta$.

Proof. The proof is similar to the proof of Corollary 1.4. It is well known that

$$\Omega = \left\{ v \in \mathbb{R}^n : \delta_{x_v}^T \xrightarrow{\mathrm{w}^*} m_X \right\}$$

has full Lebesgue measure (see for example [8] which covers also the weighted case). Using Lemma 1.2(1) we choose (η_i) so that $m_X \in \mathscr{P}(X, (\eta_i))$. We then have by definition that $\{x_v : v \in \Omega\} \subset \mathcal{H}(\eta_i)$. The theorem thus follows from Theorem 5.3. ∎

Theorem 5.3 can be generalized in several ways, using Theorem 4.2 for more general flows (a_t). For example, if $(i_1, \ldots, i_n)$ is an n-tuple of real numbers such that

$$\forall \ell,\ i_\ell \in (0,1) \quad \text{and} \quad \sum_{\ell=1}^{n} i_\ell = 1,$$

we can define, for any vector $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$,

$$\mathbf{Bad}^{\epsilon}_{(i_1,\ldots,i_n)}(v) \overset{\text{def}}{=} \left\{ w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{R}^n \ : \ \forall \ell,\ \liminf_{k\to\infty} k^{i_\ell} \langle k v_\ell - w_\ell \rangle \geqslant \epsilon \right\},$$

and

$$\mathbf{Bad}_{(i_1,\ldots,i_n)}(v) = \bigcup_{\epsilon>0} \mathbf{Bad}^{\epsilon}_{(i_1,\ldots,i_n)}(v).$$

It is known (see for example [9–11]) that for any $v \in \mathbb{R}^n$, $\dim_H \mathbf{Bad}_{(i_1,\ldots,i_n)}(v) = n$. Theorem 4.2 applied with the flow $a_t = \operatorname{diag}\left(e^{i_1 t}, \ldots, e^{i_n t}, e^{-t}\right)$ yields the following.

Theorem 5.5 (Heavy vectors for weighted approximation). Let $v \in \mathbb{R}^n$ be heavy for $a_t = \operatorname{diag}\left(e^{i_1 t}, \ldots, e^{i_n t}, e^{-t}\right)$. Then, for all $\epsilon > 0$,

$$\dim_H \mathbf{Bad}^{\epsilon}_{(i_1,\ldots,i_n)}(v) < n.$$

Moreover, for any $\epsilon > 0$, there exists $\delta > 0$ such that for a.e. $v \in \mathbb{R}^n$,

$$\dim_H \mathbf{Bad}^{\epsilon}_{(i_1,\ldots,i_n)}(v) < n - \delta.$$

Being very similar to that of Theorem 5.3, the proof of Theorem 5.5 is left to the reader.

5.2 Approximation of affine subspaces

We obtain another natural generalization of Theorem 5.3 by replacing the vector v with a matrix. We choose to present this generalization in a *projective* manner, that is, in the context of Diophantine approximation of affine subspaces of $\mathbb{R}^d$ by points in $\mathbb{Z}^d$, which is not common but which we find very natural. The case of Theorem 5.3 corresponding to the subspace being a line (see Remark 5.6).

Let $\mathrm{Grass}(\ell, d)$ be the Grassmannian of ℓ-dimensional linear subspaces of $\mathbb{R}^d$. Recall that by Minkowski's first theorem on convex bodies, for every $W_0 \in \mathrm{Grass}(\ell, d)$, the inequality

$$d(\mathbf{k}, W_0) \leqslant 2^d \cdot \|\mathbf{k}\|^{\frac{-\ell}{d-\ell}}$$

has infinitely many solutions $\mathbf{k} \in \mathbb{Z}^d$, where $\|\mathbf{k}\|$ denotes the Euclidean norm of $\mathbf{k}$. It is therefore natural to say that an affine subspace W of dimension ℓ in $\mathbb{R}^d$ is ϵ-*badly approximable* if it satisfies

$$\liminf_{\substack{\mathbf{k}\to\infty \\ \mathbf{k}\in\mathbb{Z}^d}} \|\mathbf{k}\|^{\frac{\ell}{d-\ell}} d(\mathbf{k}, W) \geqslant \epsilon.$$

Let $\mathrm{Grass}_A(\ell, d)$ denote the Grassmannian of ℓ-dimensional affine subspaces of $\mathbb{R}^d$ and $\pi : \mathrm{Grass}_A(\ell, d) \to \mathrm{Grass}(\ell, d)$ the natural projection, mapping an affine subspace to its linear part. For a linear subspace $W_0 \in \mathrm{Grass}(\ell, d)$ of $\mathbb{R}^d$, we want to study the set

$$\mathbf{Bad}^{\epsilon}_{\ell,d}(W_0) \stackrel{\text{def}}{=} \left\{ W \in \pi^{-1}(W_0) \mid \liminf_{\substack{\mathbf{k}\to\infty \\ \mathbf{k}\in\mathbb{Z}^d}} \|\mathbf{k}\|^{\frac{\ell}{d-\ell}} d(\mathbf{k}, W) \geqslant \epsilon \right\}$$

of ϵ-badly approximable affine subspaces $W \leqslant \mathbb{R}^d$ with linear part W_0. It is known that [6]

$$\dim_H \left(\mathbf{Bad}_{\ell,d}(W_0)\right) = d - \ell,$$

where $\mathbf{Bad}_{\ell,d}(W_0) \stackrel{\text{def}}{=} \bigcup_{\epsilon>0} \mathbf{Bad}^{\epsilon}_{\ell,d}(W_0)$.

Remark 5.6. Let $n = d - 1$. For $v \in \mathbb{R}^n$, consider the line $W_0 \in \mathbb{R}^d$ spanned by the vector $\tilde{v} = \binom{1}{v} \in \mathbb{R}^d$. Then a vector $w \in \mathbb{R}^n$ is in $\mathbf{Bad}(v)$ if and only if $\tilde{w} + W_0$ is a badly approximable line in $\mathbb{R}^d$, so that the setting of the previous subsection corresponds to Diophantine approximation of lines in $\mathbb{R}^d$.

Theorem 5.7 (Approximation of affine subspaces). For all $\epsilon > 0$, there exists $\delta > 0$ such that for almost every $W_0 \in \mathrm{Grass}(\ell, d)$,

$$\dim_H \mathbf{Bad}^{\epsilon}_{\ell,d}(W_0) \leqslant d - \ell - \delta.$$

Proof. Since the proof is very similar to that of Theorem 5.3, we keep it terse. We apply Theorem 4.2 with flow

$$a_t = \mathrm{diag}\left(e^t, \ldots, e^t, e^{-\frac{\ell t}{d-\ell}}, \ldots, e^{-\frac{\ell t}{d-\ell}}\right).$$

Let $W_0 \in \mathrm{Grass}(\ell, d)$, and choose $g_{W_0} \in G_0 = \mathrm{SL}_d(\mathbb{R})$ such that $g_{W_0} \cdot W_0 = \mathrm{Span}(e_1, \ldots, e_\ell)$. For almost every W_0, the orbit $(a_t g_{W_0} \mathbb{Z}^d)_{t>0}$ equidistributes in X [8] (note that this property does not depend on our choice of g_{W_0}). Taking $\mathcal{O} = B_{\frac{\epsilon}{2}}$ to be the open ball of radius $\epsilon/2$ in $\mathbb{R}^d$, Theorem 4.2 shows that there exists $\delta > 0$ such that for almost every W_0,

$$\dim_H F_{S^+(B_{\frac{\epsilon}{2}})}(g_{W_0}\mathbb{Z}^d) \leqslant d - \delta.$$

Assume now that W is an affine subspace with linear part W_0, and choose $g_W \in G = \mathrm{ASL}_d(\mathbb{R})$ such that $g_W \cdot W = \mathrm{Span}(e_1, \ldots, e_\ell)$ (as an affine subspace). It is a simple computation to check that if $W \in \mathbf{Bad}^{\epsilon}_{\ell,d}(W_0)$, then the grid $g_W \mathbb{Z}^d$ lies in $F_{S^+(B_{\frac{\epsilon}{2}})}(g_{W_0}\mathbb{Z}^d)$, independently of our choice of g_W. The above bound on the Hausdorff dimension of $F_{S^+(B_{\frac{\epsilon}{2}})}(g_{W_0}\mathbb{Z}^d)$ therefore implies

$$\dim_H \mathbf{Bad}^{\epsilon}_{\ell,d}(W_0) \leqslant d - \ell - \delta.$$

■

6 Examples

In this section, to justify the necessity of some non-escape-of-mass assumption on x in Theorem 4.2, we construct non-singular lattices in $\mathbb{R}^2$ with lots of bad grids: x with non-divergent orbits but for which there exists $\epsilon > 0$ such that $F_{S^+(\mathcal{O})}(x)$ has full Hausdorff dimension in $\pi^{-1}(x)$ for a suitable choice of $\mathcal{O}$.

Proposition 6.1 (Lattices with lots of bad grids). There exists a non-singular unimodular lattice x and an open bounded set $\mathcal{O}$ such that

$$\dim_H F_{S^+(\mathcal{O})}(x) = 2.$$

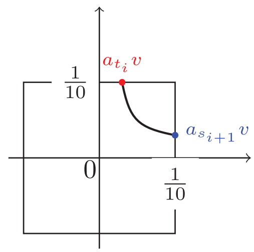

FIG. 2. $a_t v$, for $t \in (t_i, s_{i+1})$

Fix a lattice x in $\mathbb{R}^2$, assume $\lambda_1(x) \geqslant \frac{1}{10}$, where $\lambda_1(x)$ denotes the shortest nonzero vector in x with respect to the supremum norm, which we denote by $\|\cdot\|$, and let $E = \{t > 0 \mid \lambda_1(a_t x) \geqslant \frac{1}{10}\}$. Assume that E can be written as a disjoint union of closed intervals

$$E = [s_1, t_1] \cup [s_2, t_2] \cup \ldots$$

where the reals $s_1, s_2, \ldots$ and $t_1, t_2, \ldots$ are defined inductively by $s_1 = 0$, and, for $i \geqslant 1$,

$$t_i = \inf\left\{t > s_i \mid \lambda_1(a_t x) \leqslant \frac{1}{10}\right\}, \quad \text{and} \quad s_{i+1} = \inf\left\{s > t_i \mid \lambda_1(a_t x) \geqslant \frac{1}{10}\right\}.$$

Now, for each i, choose a nonzero vector $v = \binom{v_1}{v_2}$ in x such that $\lambda_1(a_{t_i}x) = \|a_{t_i}v\|$. One readily checks the following

(1) $\forall t \in (t_i, s_{i+1}), \lambda_1(a_t x) = \|a_t v\|$
(2) $e^{t_i}|v_1| = e^{-s_{i+1}}|v_2| \leqslant \frac{1}{10}$ and $e^{-t_i}|v_2| = e^{s_{i+1}}|v_1| = \frac{1}{10}$,

using the fact that in dimension 2, if $\|a_t v\| < 1$ (and v is primitive), then $\lambda_1(a_t x) = \|a_t v\|$ (see Figure 2).

From the fact that $a_{t_i}v$ has norm $\frac{1}{10}$ and makes an angle of at least $\frac{\pi}{4}$ with the first coordinate axis, we see that the translates of the line $\mathbb{R}a_{t_i}v$ by vectors of $a_{t_i}x$ intersect the horizontal axis in a discrete subgroup $\ell_i\mathbb{Z}$, with $\ell_i \in [5, 50]$. Let

$$B_i = \left\{\gamma \in \mathbb{R} \mid d(\gamma, e^{-t_i}\ell_i\mathbb{Z}) > 2e^{-t_i}\right\}.$$

For a grid y, we let $\sigma(y) \overset{\text{def}}{=} \min\{\|v\|\,;\, v \in y\}$ denote the norm of the shortest vector in y. We claim that

$$\forall \gamma \in B_i, \forall t \in [t_i, s_{i+1}], \ \sigma\left(a_t\left(x + \begin{pmatrix}\gamma \\ 0\end{pmatrix}\right)\right) \geqslant 1. \tag{6.1}$$

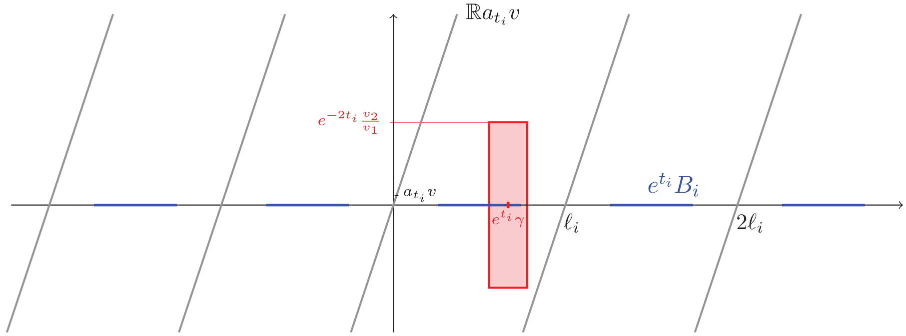

FIG. 3. Translates of $\mathbb{R}a_{t_i}v$, $e^{t_i}B_i$, and the box around $e^{t_i}\gamma$

To see this, observe that if $\gamma \in B_i$, then the box

$$\begin{pmatrix} e^{t_i}\gamma \\ 0 \end{pmatrix} + [-1,1] \times \left[-e^{-2t_i}\frac{v_2}{v_1}, e^{-2t_i}\frac{v_2}{v_1}\right]$$

around $a_{t_i}\binom{\gamma}{0}$ does not intersect $a_{t_i}x$ (see Figure 3).

Therefore, if $t \in [t_i, s_{i+1}]$, then the box around the origin

$$\begin{aligned}[-e^{t-t_i}, e^{t-t_i}] \times \left[-e^{-t_i-t}\frac{v_2}{v_1}, e^{-t_i-t}\frac{v_2}{v_1}\right] &\supset [-1,1] \times \left[-e^{-t_i-s_{i+1}}\frac{v_2}{v_1}, e^{-t_i-s_{i+1}}\frac{v_2}{v_1}\right] \\ &= [-1,1] \times [-1,1].\end{aligned}$$

This proves our claim.

To prove Proposition 6.1, we use the following elementary Hausdorff dimension estimate.

Lemma 6.2. With the above notation, suppose $\lim_{i\to\infty} \frac{t_i}{i} = \infty$. Then

$$\dim_H \bigcap_{i\geqslant 1} B_i = 1.$$

Proof. By the mass distribution principle [7, §4.2 page 60], it suffices to construct on $B = \bigcap B_i$, for $\epsilon > 0$ arbitrarily small, a probability measure μ satisfying, for all x and all $r > 0$ sufficiently small,

$$\mu(B(x,r)) \leqslant r^{1-\epsilon}.$$

For this, we define μ_1 to be the Lebesgue measure on each interval of B_1 included in [0, 1], normalized to be a probability measure. Let N_1 be the number of intervals of B_1 included in [0, 1]. Within bounded multiplicative constants, we have

$$N_1 \asymp \frac{e^{t_1}}{\ell_1} \asymp e^{t_1}.$$

Then, we let μ_2 be the probability measure compatible with μ_1 (in the sense that the μ_2-mass of a B_1-interval is equal to its μ_1-mass) and equal to the appropriately normalized Lebesgue measure on each B_2-interval. The number of B_2-intervals inside a B_1 interval is

$$N_2 \asymp \frac{e^{-t_1}\ell_1}{e^{-t_2}\ell_2} \asymp e^{t_2-t_1}.$$

Iterating this procedure, we obtain a sequence of probability measures μ_n supported on $\bigcap_{i=1}^n B_i$; then, we let μ be a weak-* limit of the sequence (μ_n). Note that, by our construction, if I is a B_i-interval then for all $n \geqslant i$, $\mu(I) = \mu_n(I)$.

For $r > 0$ sufficiently small, find i such that $e^{-t_{i-1}}\ell_{i-1} > r \geqslant e^{-t_i}\ell_i$. Since B_i-intervals are separated by a distance $e^{-t_i}\ell_i$, the number of B_i-intervals intersecting $B(x, r)$ is at most $\frac{r}{e^{-t_i}\ell_i} \asymp re^{t_i}$, and the μ-mass of a B_i-interval is $\asymp (N_1 \dots N_i)^{-1} \leqslant C^i e^{-t_i}$, where C is some positive constant independent of i, so that

$$\mu(B(x,r)) \leqslant re^{t_i}C^i e^{-t_i} \leqslant rC^i.$$

Using that $r \ll e^{-t_{i-1}}$ and that $\lim \frac{i}{t_{i-1}} = 0$, we find that, given any $\epsilon > 0$, for i large enough (i.e., r small enough), $C^i = e^{t_{i-1}\frac{i\log C}{t_{i-1}}} \leqslant r^{-\epsilon}$. Thus, for sufficiently small $r > 0$ (depending on ϵ)

$$\mu(B(x,r)) \leqslant r^{1-\epsilon}.$$

■

We can now derive Proposition 6.1.

Proof of Proposition 6.1. Let $\alpha \in [0, 1]$ be an irrational number with continued fraction expansion $\alpha = [n_1, n_2, \dots]$ such that $\lim n_i = \infty$, and set

$$x = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \mathbb{Z}^2.$$

The set $E = \{t \geqslant 0 \mid \lambda_1(a_t x) \geqslant \frac{1}{10}\}$ can be written as a union of disjoint intervals

$$E = [s_1, t_1] \cup [s_2, t_2] \cup \ldots$$

and for some constant C, for all i, $t_i - s_i \leqslant C$, and $\lim \frac{t_i}{i} = \infty$. By Lemma 6.2, the set $B = \bigcap_{i \geq 1} B_i$ has Hausdorff dimension 1. Moreover, by (6.1), for any $\gamma \in B$, for all t not in any interval $[s_i, t_i]$,

$$\sigma\left(a_t\left(x + \begin{pmatrix} \gamma \\ 0 \end{pmatrix}\right)\right) \geqslant 1.$$

Since the intervals $[s_i, t_i]$ are disjoint and have length at most C, we find that for all $\gamma \in B$ and all $t \geqslant 0$,

$$\sigma\left(a_t\left(x + \begin{pmatrix} \gamma \\ 0 \end{pmatrix}\right)\right) \geqslant e^{-C}.$$

This shows that for $\mathcal{O}$ being the $\frac{1}{2}e^{-C}$-ball around the origin with respect to the sup-norm, we have that the image of the set $B \times [-1, 1]$ in the torus $\mathbb{R}^2/x$ is contained in $F_{S^+(\mathcal{O})}(x)$ (note that translating in the stable direction does not affect the asymptotic properties of the a_t-orbit), which implies in particular that

$$\dim_H F_{S^+(\mathcal{O})}(x) = 2.$$

■

Acknowledgments

We would like to thank Manfred Einsiedler and Elon Lindenstrauss for valuable discussions. Seonhee Lim acknowledges the support of the Samsung Science and Technology Foundation under Project No. SSTF-BA1601-03. Nicolas de Saxcé acknowledges the warm hospitality of the mathematics department at the Technion. Uri Shapira acknowledges the support of ISF grant 357/13.

References

[1] Berend, D. and W. Moran. "The inhomogeneous minimum of binary quadratic forms." *Math. Proc. Cambridge Philos. Soc.* 112, no. 1 (1992): 7–19. MR 1162928

[2] Bugeaud, Y., S. Harrap, S. Kristensen, and S. Velani. "On shrinking targets for Z^m actions on tori." *Mathematika* 56 (2010): 193–202.

[3] Einsiedler, M. and E. Lindenstrauss. "Diagonal actions on locally homogeneous spaces." In *Homogeneous Flows, Moduli Spaces and Arithmetic, Clay Math. Proc.* Providence, RI: Amer. Math. Soc. 10 (2010): 155–241. MR 2648695

[4] Einsiedler, M., E. Lindenstrauss, P. Michel, and A. Venkatesh. “The distribution of closed geodesics on the modular surface, and Duke’s theorem.” *Enseign. Math. (2)* 58, no. 3–4 (2012): 249–313. MR 3058601

[5] Einsiedler, M., E. Lindenstrauss, and T. Ward. Entropy in ergodic theory and homogeneous dynamics. Preprint available at http://www.personal.leeds.ac.uk/ mattbw/entropy.

[6] Einsiedler, M. and J. Tseng. “Badly approximable systems of affine forms, fractals, and Schmidt games.” *J. Reine Angew. Math.* 660 (2011): 83–97. MR 2855820

[7] Falconer, K. *Fractal Geometry, second ed. Mathematical foundations and applications*. Hoboken, NJ: John Wiley & Sons, Inc., 2003. MR 2118797.

[8] Kleinbock, D., R. Shi, and B. Weiss. “Pointwise equidistribution with an error rate and with respect to unbounded functions.” *Math. Ann.* 2 (2016): 1–23.

[9] Kleinbock, D. and B. Weiss. “Modified Schmidt games and Diophantine approximation with weights.” *Adv. Math.* 223, no. 4 (2010): 1276–98. MR 2581371

[10] Kristensen, S., R. Thorn, and S. Velani. “Diophantine approximation and badly approximable sets.” *Adv. Math.* 203, no. 1 (2006): 132–69. MR 2231044

[11] Pollington, A. and S. Velani. “On simultaneously badly approximable numbers.” *J. Lond. Math. Soc.* 66, no. 1 (2002): 29–40. MR 1911218

[12] Shapira, U. “Grids with dense values.” *Comment. Math. Helv.* 88, no. 2 (2013): 485–506. MR 3048195