On the dimension of limit sets on $\mathbb{P}(\mathbb{R}^3)$ via stationary measures: the theory and applications

Yuxiang Jiao, Jialun Li, Wenyu Pan and Disheng Xu

Abstract

Let ν be a probability measure on $\mathrm{SL}_3(\mathbb{R})$ whose support is finite and spans a Zariski dense subgroup. Let μ be the associated stationary measure for the action on $\mathbb{P}(\mathbb{R}^3)$. Under the exponential separation condition on ν, we prove that the Hausdorff dimension of μ equals its Lyapunov dimension. This allows us to study the limit sets of Anosov representations and Rauzy gasket. We show that their Hausdorff dimensions are equal to the affinity dimensions. In particular, there is a dimension gap in the Barbot component.

Contents

1 Introduction

The Hausdorff dimension of the limit set on the visual boundary $\partial\mathbb{H}^3$ of a discrete subgroup Γ of $\mathrm{Isom}_+(\mathbb{H}^3)$ is a significant quantity of the group. It has been extensively studied and is known to be related to the growth rate of elements in Γ, the bottom of the eigenvalue of the Laplace operator and the entropy of the geodesic flow on $\Gamma\backslash\mathbb{H}^3$, as established by Sullivan in his works [Sul79], [Sul84]. The present paper is concerned with the study of the Hausdorff dimension of limit sets in the real projective plane $\mathbb{P}(\mathbb{R}^3)$ of discrete subgroups of $\mathrm{SL}_3(\mathbb{R})$. We usually denote by $\dim\cdot$ for the Hausdorff dimension of a set and by $L(\cdot)$ for the limit set (see Definition 1.2) of projective actions of some discrete linear subgroups.

One major motivation in this paper is to generalize Bowen's dimension rigidity result [Bow79] to a more general setting. Specifically, let Γ be a surface group, i.e. the fundamental group of a closed connected surface of genus at least 2 and let $\eta_0 : \Gamma \to \mathrm{PSL}_2(\mathbb{R})$ be a faithful representation onto a cocompact Fuchsian group of $\mathrm{PSL}_2(\mathbb{R})$. We view $\mathrm{PSL}_2(\mathbb{R})$ as a subgroup of $\mathrm{PSL}_2(\mathbb{C})$ and let ι_0 be the natural embedding of $\mathrm{PSL}_2(\mathbb{R})$ in $\mathrm{PSL}_2(\mathbb{C})$, then $\iota_0 \circ \eta_0$ is a representation of Γ to $\mathrm{PSL}_2(\mathbb{C})$. The limit set of $\iota_0 \circ \eta_0(\Gamma)$ is the smooth circle $\mathbb{P}^1_{\mathbb{R}}$ in the sphere $\mathbb{P}^1_{\mathbb{C}}$. Let $\rho : \Gamma \to \mathrm{PSL}_2(\mathbb{C})$ be a representation which is a small perturbation of $\iota_0 \circ \eta_0$. The limit set $L(\rho(\Gamma))$ is known to be homeomorphic to a circle, which implies $\dim L(\rho(\Gamma)) \geqslant 1$. Bowen proved that $\dim L(\rho(\Gamma)) = 1$ if and only if $\rho(\Gamma)$ is contained in a conjugate of $\mathrm{SL}_2(\mathbb{R})$.[1]

Local dimension jump We study a problem which is similar to that of Bowen, but is about the action of $\mathrm{SL}_3(\mathbb{R})$ on $\mathbb{P}(\mathbb{R}^3)$. A key difference with Bowen's setting is the non-conformality of the action, which creates a lot of difficulties and also some new phenomena. One of our main results is a dimension jump phenomenon which leads to a local dimension rigidity as the following.

We consider $\mathrm{Hom}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$, the set of $\mathrm{SL}_3(\mathbb{R})$-representations of a surface group Γ, and we deform representations here in a way similar to Bowen's setting. More precisely, recall that $\eta_0 : \Gamma \to \mathrm{PSL}_2(\mathbb{R})$ is a faithful representation onto a cocompact Fuchsian group, and take an arbitrary lift $\rho_0 : \Gamma \to \mathrm{SL}_2(\mathbb{R})$ of η_0.[2] Denote by ι the embedding from $\mathrm{SL}_2(\mathbb{R})$ to the upper left corner of $\mathrm{SL}_3(\mathbb{R})$, then the composition $\rho_1 := \iota \circ \rho_0$ is an element in $\mathrm{Hom}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$. The action of $\rho_1(\Gamma)$ preserves $\mathbb{P}(\mathbb{R}^2)$ in $\mathbb{P}(\mathbb{R}^3)$, which is a smooth circle. Due to [Sul85], if ρ in $\mathrm{Hom}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$ is sufficiently close to ρ_1, i.e. its images on generators are close to that of ρ_1, then $\rho(\Gamma)$ preserves a unique topological circle close to $\mathbb{P}(\mathbb{R}^2)$ in $\mathbb{P}(\mathbb{R}^3)$, which turns out to be the limit set $L(\rho(\Gamma))$ (see Definition 1.2).

Theorem 1.1. *For every $\epsilon > 0$, there exists a small neighborhood O of ρ_1 in $\mathrm{Hom}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$ such that for any ρ in O we have*

- *either $\rho(\Gamma)$ acts reducibly, i.e. fixing a point or a projective line in $\mathbb{P}(\mathbb{R}^3)$;*
- *or $\rho(\Gamma)$ is irreducible and*
$$|\dim L(\rho(\Gamma)) - \frac{3}{2}| \leqslant \epsilon.$$

In other words, once we perturb ρ_1 to any irreducible ρ, there is a dimension jump with size about $1/2$, which does not occur in the conformal case. A generic perturbation ρ is irreducible. This is because reducible representations form a proper subvariety in $\mathrm{Hom}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$ (cf. [BCLS15]).

If ρ is reducible, the analysis becomes more subtle and we have several cases. If ρ preserves a projective line in $\mathbb{P}(\mathbb{R}^3)$, then $L(\rho(\Gamma))$ is this projective line. Hence $\dim L(\rho(\Gamma)) = 1$. If ρ does not preserve a projective line, then $L(\rho(\Gamma))$ is not Lipschitz [Bar10, Theorem 4.4]. And $L(\rho(\Gamma))$ is a graph-like set, and its dual representation ρ^* preserves a projective line which is actually $L(\rho^*(\Gamma))$. We hope to treat this case in a forthcoming paper and conjecture that a similar dimension jump phenomenon holds. We believe that one can distinguish different types of ρ in O using the Hausdorff dimension of limit sets in the following table. We conjecture that the entries that with question marks are true.

There are other settings where similar dimension jump occurs, for example, see [RS21] and references therein for results on graphs of Weierstrass-type functions.

[1]The actual result of Bowen is global, not only for small perturbation. The version stated here is a weaker local version.

[2]One can do it by choosing the images of generators in the double cover $\mathrm{SL}_2(\mathbb{R})$ of $\mathrm{PSL}_2(\mathbb{R})$.

Table 1: Hausdorff dimension of different types of representations

Reducibility type of $\rho \in O$	$\dim L(\rho(\Gamma))$	$\dim L(\rho^*(\Gamma))$
Semisimple	1	1
point-irreducible	1	$\approx 3/2$ (?)
line-irreducible	$\approx 3/2$ (?)	1
Irreducible (in this paper)	$\approx 3/2$	$\approx 3/2$

We follow the definition in [LLS21]. A reducible representation $\rho : \Gamma \to \mathrm{SL}_3(\mathbb{R})$ is called *point-irreducible* if it does not preserve any point in $\mathbb{P}(\mathbb{R}^3)$; *line-irreducible* if it does not preserve any projective line in $\mathbb{P}(\mathbb{R}^3)$; *semisimple* if it is a product of irreducible representations.

Anosov representations The perturbed representation ρ we considered in Theorem 1.1 is actually an example of *Anosov representations*, which is a generalization of convex-cocompact representations. The concept of Anosov representation was first introduced by Labourie in [Lab06] to study Hitchin components of the representations of surface groups. Here we use an equivalent definition which is due to [GW12], [KLP18], [GGKW17], [BPS19], etc.

Definition 1.2. Let Γ be a hyperbolic group [3]and ρ a representation from Γ to $\mathrm{SL}_3(\mathbb{R})$. Then ρ is called an Anosov representation if there exist $C, c > 0$ such that for any $\gamma \in \Gamma$, we have

$$\frac{\sigma_1(\rho(\gamma))}{\sigma_2(\rho(\gamma))} \geqslant \frac{1}{C} e^{c|\gamma|},$$

where $\sigma_i(\rho(\gamma))$ are the singular values of $\rho(\gamma)$ with decreasing order and $|\gamma|$ is the word length with respect to a fixed symmetric generating set. The limit set $L(\rho(\Gamma))$ of ρ in $\mathbb{P}(\mathbb{R}^3)$ is defined to be the closure of the set of attracting fixed points of elements in $\rho(\Gamma)$. We denote by $\mathrm{HA}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$ the subset of $\mathrm{Hom}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$ consisting of Anosov representations.

If $\rho \in \mathrm{HA}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$, then $\rho(\Gamma)$ is discrete in $\mathrm{SL}_3(\mathbb{R})$. And any Anosov representation is stable under small perturbations in $\mathrm{Hom}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$, hence $\mathrm{HA}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$ is open in $\mathrm{Hom}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$. It is not hard to check that ρ_1 in Theorem 1.1 is an Anosov representation, and hence without loss of generality, we can assume its small neighborhood $O \subset \mathrm{HA}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$.

Affinity dimension The core of the proof of Theorem 1.1 is to show that if ρ is an Anosov representation, then $\dim L(\rho(\Gamma))$ equals its *affinity dimension* (or *affinity exponent*).

The concept of affinity exponent of Anosov representations was introduced in [PSW22], which is essentially a natural extension of affinity dimension of self-affine fractals due to Falconer [Fal88].

To motivate the definition, let us first recall the classical critical exponent for Fuchsian groups and Kleinian groups. For a representation $\rho : \Gamma \to \mathrm{PSL}(2, \mathbb{R})$ (or $\mathrm{PSL}(2, \mathbb{C})$), the Poincaré series of the group $\rho(\Gamma)$ is defined by

$$P_\rho(s) = \sum_{\gamma \in \Gamma} \left(\frac{\sigma_2}{\sigma_1}\right)^s (\rho(\gamma)).$$

And the critical exponent $\delta(\rho(\Gamma))$ of $\rho(\Gamma)$ is defined to be the critical exponent of series $P_\rho(s)$, i.e.

$$\delta(\rho(\Gamma)) = \sup_{P_\rho(s)=\infty} s = \inf_{P_{\rho(s)}<\infty} s.$$

Sullivan [Sul79] established the beautiful equality between the critical exponent, a dynamical invariant, and the Hausdorff dimension of the conic limit set, a geometric invariant for a Kleinian group, extending the earlier work by Patterson for Fuchsian groups [Pat76].

[3]In this paper we always consider finitely generated hyperbolic groups.

For a representation $\rho : \Gamma \to \mathrm{SL}_3(\mathbb{R})$, we can similarly define a Poincaré series of the group $\rho(\Gamma)$ by

$$P_\rho(s) = \begin{cases} \sum_{\gamma\in\Gamma} \left(\frac{\sigma_2}{\sigma_1}\right)^s (\rho(\gamma)), & 0 < s \leqslant 1; \\ \sum_{\gamma\in\Gamma} \left(\frac{\sigma_2}{\sigma_1}\right)(\rho(\gamma))\left(\frac{\sigma_3}{\sigma_1}\right)^{s-1}(\rho(\gamma)), & 1 < s \leqslant 2. \end{cases} \tag{1.1}$$

We denote the critical exponent of $P_\rho(s)$ by $s_A(\rho)$ and call it the affinity exponent (or the affinity dimension). The affinity dimension $s_A(\rho)$ is always an upper bound of $\dim L(\rho(\Gamma)) \subset \mathbb{P}(\mathbb{R}^3)$ if $\rho : \Gamma \to \mathrm{SL}_3(\mathbb{R})$ is an Anosov representation [PSW22]. In this paper we show $s_A(\rho)$ is also an lower bound of $\dim L(\rho(\Gamma))$. Hence, we generalize Sullivan's theorem to Anosov representations in $\mathrm{SL}_3(\mathbb{R})$.

Theorem 1.3. *Let Γ be a hyperbolic group and $\rho : \Gamma \to \mathrm{SL}_3(\mathbb{R})$ be an irreducible Anosov representation, then the Hausdorff dimension of the limit set $L(\rho(\Gamma))$ in $\mathbb{P}(\mathbb{R}^3)$ equals the affinity exponent $s_A(\rho)$. Moreover, the exponent $s_A(\rho)$ is continuous on $\mathrm{HA}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$.*

The dimension jump phenomenon in Theorem 1.1, to some degree, is attributed to the fact that the Hausdorff dimension of the limit set is not equal to the affinity dimension when the representation preserves a plane. For higher dimensional Lie groups, we can similarly define Anosov representations (cf. [Can], etc.), associated limit sets, and affinity dimensions. We pose the following general question.

Question 1.4. *Let Γ be a hyperbolic group. Which $\rho \in \mathrm{HA}(\Gamma, \mathrm{SL}_n(\mathbb{R}))$ (can be reducible) satisfies*

$$\dim L(\rho(\Gamma)) = s_A(\rho)?$$

This question is in fact quite delicate even when one seeks a complete answer for $\mathrm{SL}_4(\mathbb{R})$.

The continuity of the affinity exponent $s_A(\rho)$ is essentially due to [PS17] and [BCLS15]. For completeness, we provide a proof in Section 10.4. In the setting of Bowen's result [Bow79], it is known that classical critical exponents depend on representations analytically [Rue82]. It is natural to ask the following question.

Question 1.5. *Does the affinity exponent $s_A(\rho)$ depend on $\rho \in \mathrm{HA}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$ analytically?*

When Γ is a surface group, if $\rho \in \mathrm{Hom}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$ is in the Hitchin component, that is, ρ can be continuously deformed to a discrete and faithful representation into $\mathrm{SO}(2,1)(\cong \mathrm{SL}_2(\mathbb{R}))$; then the limit set $L(\rho(\Gamma))$ is a C^1-circle [Lab06], and hence its Hausdorff dimension is 1. The representation ρ_1 in Theorem 1.1 is in the Barbot component of $\mathrm{Hom}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$. For an irreducible Anosov representation in the Barbot's component, Barbot [Bar10] proved that the limit set is not Lipschitz. Our result provides new knowledge for Anosov representations in Barbot's component. Bowen's result is a global result for all convex-cocompact representations. It is interesting to investigate the following question.

Question 1.6. *Do all irreducible Anosov representations in Barbot's component have affinity dimension greater than 1?*

In works like [Sul79], Patterson-Sullivan measures are central in computing the Hausdorff dimension of limit sets. For Anosov representations, we also have Patterson-Sullivan measures on the limit sets [Qui02], [PSW22]. Our current computation of the Hausdorff dimension of limit sets does not use these nice measures. It is interesting to explore the following question.

Question 1.7. *Is there a Patterson-Sullivan measure with the same Hausdorff dimension as the limit set?*

Related works on Hausdorff dimension of limit sets of Anosov representations include [PSW22], [PSW21], [GMT19], [DK22]. In [Duf17], Dufloux studied this problem for the non-conformal action of Schottky groups in $\mathrm{PU}(1,n)$ on $\partial\mathbb{H}^n_{\mathbb{C}}$, the boundary of the complex hyperbolic n-space. In [BHR19], Bárány-Hochman-Rapaport studied self-affine contracting IFS on $\mathbb{R}^2$. Our proof draws inspiration from their method. See also [HR21] and [Rap22] for further development on this setting. There were a lot of more classical results in this setting, we just name a few, [Fal88], [HL95]. See [RS21] for the Hausdorff dimension of the Weierstrass graph and [CP10] and [CPZ19] for more backgrounds about non-conformal repellers.

Rauzy gasket We also study the Rauzy gasket, a self-projective fractal set in $\mathbb{P}(\mathbb{R}^3)$, and we establish the identity between its Hausdorff dimension and its affinity dimension.

Let $\Delta := \{(x,y,z) : x,y,z \geqslant 0, x+y+z=1\}$. We view Δ as subset in $\mathbb{P}(\mathbb{R}^3)$. Let Γ be the semigroup generated by

$$A_1 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},\ A_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix},$$

and we call it the Rauzy semigroup. Then as $\Gamma \subset \mathrm{SL}_3(\mathbb{R})$, the semigroup Γ acts on $\mathbb{P}(\mathbb{R}^3)$. Due to the choice of Δ, the semigroup Γ preserves Δ. The Rauzy Gasket X is the unique attractor of the Rauzy semigroup which can be defined formally as

$$\bigcap_{n\to\infty} \bigcup_{i_j\in\{1,2,3\}} (A_{i_0}\cdots A_{i_{n-1}}\Delta).$$

The Rauzy gasket, depicted in Figure 1, was first introduced in 1991 by Arnoux and Rauzy [AR91] in the context of interval exchange transformations. They conjectured that the gasket has Lebesgue measure 0. Levitt [Lev93] rediscovered the gasket in 1993 and provided a proof, due to Yoccoz, of its 0 Lebesgue measure. Later, the Rauzy gasket was studied by Dynnikov and De Leo [DD09] in connection with Novikov's problem [Nov82] of plane sections of triply periodic surfaces. The Hausdorff dimension of the Rauzy gasket was estimated numerically in [DD09], with suggested lower and upper bounds of 1.7 and 1.8, respectively. Arnoux asked whether the Hausdorff dimension is less than or equal to 2 [AS13]. Avila, Hubert, and Skripchenko [AHS16] provided a positive answer to this question, and a lower bound was shown in [GRM20]. Recently, Pollicott and Sewell [PS21] used a renewal theoretical argument to show that the Hausdorff dimension of the Rauzy gasket is less than 1.7407. See also [Fou20] for a weaker upper bound.

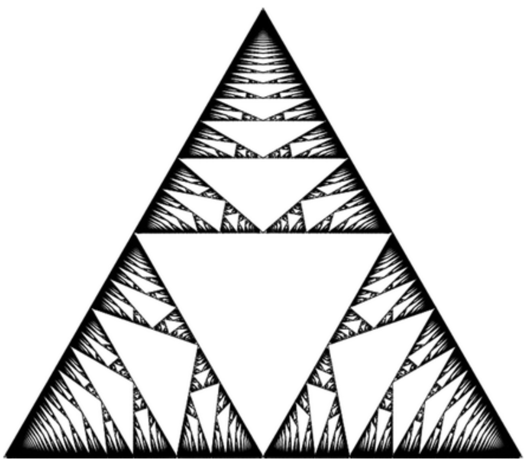

Figure 1: Figure in [AS13]

The affinity exponent (affinity dimension) $s_A(\Gamma)$ of the Rauzy Gasket is the critical exponent of the Poincaré series $P_\Gamma(s)$ defined in Eq. (1.1) using the Rauzy semigroup Γ. The following theorem confirms a folk-lore conjecture of the Hausdorff dimension of Rauzy Gasket.

Theorem 1.8. *Let X, Γ be the Rauzy Gasket and Rauzy semigroup respectively, $s_A(\Gamma)$ be the associated affinity exponent, then $\dim X = s_A(\Gamma)$.*

Remark 1.9. Recently Natalia Jurga and Ariel Rapaport also obtain Theorem 1.8 independently. They obtain the dimension of the Furstenberg measure in this special case with a similar method. But their approximation argument differs from ours (Section 11.3).

Hausdorff dimension and Lyapunov dimension of Furstenberg measures One key ingredient to prove Theorem 1.3 and Theorem 1.8 is to study the Hausdorff dimension and *Lyapunov dimension* of Furstenberg measures coming from random walks on $\mathbb{P}(\mathbb{R}^3)$, which is partially inspired by the methods in [BHR19] and [HS17], which are developed from [Hoc14].

Recall that the Hausdorff dimension of a Borel probability measure μ on a metric space X is defined as

$$\dim \mu := \inf_{A\subset X, \mu(A)=1} \dim A.$$

The Lyapunov dimension of measures is an analog of the affinity dimension of sets. For our purpose, we only define it for Furstenberg measures on $\mathbb{P}(\mathbb{R}^3)$ as the follows.

A measure supported on $\mathrm{SL}_3(\mathbb{R})$ is called *Zariski dense* if the group generated by $\operatorname{supp}\nu$ is Zariski dense in $\mathrm{SL}_3(\mathbb{R})$. Let ν be a finitely supported Zariski dense probability measure on $\mathrm{SL}_3(\mathbb{R})$, and μ be the unique Furstenberg measure on $\mathbb{P}(\mathbb{R}^3)$ (ν-stationary probability measure), that is, μ satisfies $\nu * \mu = \mu$, where the convolution is induced by the projective action of $\mathrm{SL}_3(\mathbb{R})$ on $\mathbb{P}(\mathbb{R}^3)$. Let $\lambda(\nu) = (\lambda_1, \lambda_2, \lambda_3)$ be the Lyapunov vector of ν with decreasing order and $\chi_i(\nu) = \lambda_1 - \lambda_{i+1}$. As ν is Zariski dense, we have $\chi_2(\nu) > \chi_1(\nu) > 0$. Then the Lyapunov dimension of μ can be defined as (following [KY79] and [DO80])

$$\dim_{\mathrm{LY}} \mu = \begin{cases} \dfrac{h_{\mathrm{F}}(\mu,\nu)}{\chi_1(\nu)}, & \text{if } h_{\mathrm{F}}(\mu,\nu) \leqslant \chi_1(\nu); \\ 1 + \dfrac{h_{\mathrm{F}}(\mu,\nu) - \chi_1(\nu)}{\chi_2(\nu)}, & \text{otherwise}, \end{cases}$$

where $h_{\mathrm{F}}(\mu,\nu)$ is the Furstenberg entropy (see Eq. (8.1) for definition).

In [LL23b] and [Rap21], the authors proved that $\dim_{\mathrm{LY}} \mu \geqslant \dim \mu$. In this paper we show the equality actually holds under a mild separation assumption as in [HS17], [BHR19], [Hoc14], etc. More precisely, a finitely supported measure ν on $\mathrm{SL}_3(\mathbb{R})$ is called satisfying the exponential separation condition, if there exist $C > 0$ and $N > 0$, for all $n \geqslant N$ and $(g_1, \cdots, g_n)$, $(g_1', \cdots, g_n')$ in $(\operatorname{supp}\nu)^{\times n}$ with $g_1 \cdots g_n \neq g_1' \cdots g_n'$, then

$$d(g_1 \cdots g_n, g_1' \cdots g_n') > C^{-n},$$

where d is a left $\mathrm{SL}_3(\mathbb{R})$-invariant and right $\mathrm{SO}_3(\mathbb{R})$-invariant Riemannian metric on $\mathrm{SL}_3(\mathbb{R})$. In particular, if $\operatorname{supp}\nu$ generates a discrete subgroup (for example, ν is supported on an Anosov representation), then ν satisfies the exponential separation condition.

Theorem 1.10. *Let ν be a Zariski dense, finitely supported probability measure on $\mathrm{SL}_3(\mathbb{R})$ that satisfies the exponential separation condition, and μ be its Furstenberg measure on $\mathbb{P}(\mathbb{R}^3)$. Then we have*

$$\dim \mu = \dim_{\mathrm{LY}} \mu.$$

Ledrappier-Lessa [LL23b] proved Theorem 1.10 for measures ν supported on Hitchin representations in $\mathrm{PSL}_n(\mathbb{R})$ for all $n \geqslant 2$. For a Hitchin representation, the associated limit set and hence the support of μ lies in a $C^{1+\beta}$-circle with $\beta > 0$, which helps them to obtain the dimension formula.

If we consider the action of $\mathrm{SL}_2(\mathbb{R})$ on $\mathbb{P}(\mathbb{R}^2)$, an early result is due to Ledrappier [Led83]. Later, Hochman-Solomayk advanced the knowledge in this setting [HS17]. In particular, recall that the random walk entropy $h_{\mathrm{RW}}(\nu)$ of ν is defined by

$$h_{\mathrm{RW}}(\nu) := \lim_{n\to\infty} H(\nu^{*n})/n \tag{1.2}$$

with $H(\nu^{*n})$ the Shannon entropy of a finite set. Hochman-Solomyak obtained Theorem 1.10 by replacing the Furtenberg entropy $h_{\mathrm{F}}(\mu,\nu)$ by the random walk entropy $h_{\mathrm{RW}}(\nu)$.

Dimension of the projection of Furstenberg measures Our strategy to show Theorem 1.10 is to combine a Ledrappier-Young formula in [Rap21] and [LL23b], and a central result of dimension of projection measures which is a projective version of that in [BHR19].

For any one-dimensional subspace $V \subset \mathbb{R}^3$, we consider the orthogonal projection from $\mathbb{R}^3 - V$ to $V^\perp$, which induces a smooth projection (along projective lines passing through $\mathbb{P}(V)$) from $\mathbb{P}(\mathbb{R}^3) - \{V\}$ to $\mathbb{P}(V^\perp)$. We denote the smooth projection by $\pi_{V^\perp}$. For any probability measure ν on $\mathrm{SL}_3(\mathbb{R})$, its inverse ν^- is defined by $\nu^-(A) := \nu(\{g^{-1} : g \in A\})$ with a Borel set A. In particular ν is Zariski dense if and only if so is ν^-.

Given ν a Zariski dense, finitely supported probability measure on $\mathrm{SL}_3(\mathbb{R})$, let μ^- be the Furstenberg measure of ν^- on $\mathbb{P}(\mathbb{R}^3)$. In [Rap21] and [LL23b], they proved a Ledrappier-Young formula, i.e. for μ^--a.e. $V \in \mathbb{P}(\mathbb{R}^3)$,

$$\dim \mu = \dim \pi_{V^\perp}\mu + \frac{h_{\mathrm{F}}(\mu,\nu) - \chi_1(\nu)\dim \pi_{V^\perp}\mu}{\chi_2(\nu)}. \tag{1.3}$$

Moreover, they established an upper bound: $\dim \pi_{V^\perp}\mu \leqslant \min\{1, h_{\mathrm{F}}(\mu,\nu)/\chi_1\}$.

In the following core theorem, we show that under the assumptions of Theorem 1.10, the upper bound above is actually also a lower bound. Then Theorem 1.10 is a direct corollary of the following theorem, (1.3) and the general fact that $h_{\mathrm{F}}(\mu,\nu) \leqslant h_{\mathrm{RW}}(\nu)$.

Theorem 1.11. *Let ν be a Zariski dense, finitely supported probability measure on $\mathrm{SL}_3(\mathbb{R})$ that satisfies the exponential separation condition, and μ be its Furstenberg measure on $\mathbb{P}(\mathbb{R}^3)$. Then for μ^--a.e. $V \in \mathbb{P}(\mathbb{R}^3)$, we have*

$$\dim \pi_{V^\perp}\mu = \min\{1, \frac{h_{\mathrm{RW}}(\nu)}{\chi_1}\}.$$

Approximation of affinity dimensions To complete the proof of Theorem 1.3, it remains to obtain $\dim L(\rho(\Gamma))$ from the dimension formula of stationary measures. We need a variational principle type result, i.e. show that the affinity exponent $s_A(\rho)$ of $\rho(\Gamma)$ can be approximated by the Lyapunov dimension of stationary measures.

Theorem 1.12. *Let Γ be a hyperbolic group and $\rho \in \mathrm{HA}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$ whose image is Zariski dense. For every $\epsilon > 0$, there exists a Zariski dense probability measure ν supported on $\rho(\Gamma)$ such that its Furstenberg measure μ satisfies $\dim_{\mathrm{LY}} \mu \geqslant s_A(\rho) - \epsilon$.*

We actually prove a more general version of Theorem 1.12 for higher dimensional (Borel) Anosov representations on projective spaces (Proposition 10.3) and the flag varieties (Proposition 10.4). The proof utilizes a geometric group theoretic argument to find the approximation (Sections 9 and 10).

Similar types of variational principles are proved in [MS19] and [MS23], which are used to obtain the equality between the affinity exponent and the Hausdorff dimension of attractors of self-affine IFSs in [BHR19] and [Rap22]. In [HJX23], a variational principle is proved and leads to an identity of the Hausdorff dimension of minimal sets on the circle.

Convention Through out the paper we use base-q logarithm, where q is a fixed integer large enough that will be defined in Section 2.9.

Organization of the paper After some preparations, the proof of Theorem 1.11 will occupy the most part the paper, Section 3-Section 7. In Sections 8 to 10, we will give the proof of Theorem 1.3 from Theorem 1.10. In Section 11, we will prove Theorem 1.8 from Theorem 1.10.

Acknowledgement The authors would like to thank in particular Andres Sambarino for helping discussions about the Barbot's component which motivates this work and the observation of dimension jump. The authors would also like to thank François Ledrappier for his courses on the Ledrappier-Young formula. We would like to thank Wenyuan Yang for very carefully explaining the basic ideas and arguments in [Yan19], which is very useful for Section 9. We would like to thank Cagri Sert, Federico Rodriguez-Hertz, Pablo Lessa for helpful discussions.

Notation

We summarize our main notation and conventions here.

- If A and B are two quantities, we write $A \ll B$ or $A = O(B)$ means that there exists some constant $C > 0$ (possibly depending on the ambient group and the random walk ν) such that $A \leqslant CB$.
- We write $A \ll_a B$ and $A = O_a(B)$, if the constant C depending on an extra parameter a.
- We $A \simeq B$, if $A \ll B$ and $B \ll A$. We write $A \simeq_a B$, if $A \ll_a B$ and $B \ll_a A$.

$\log$	Logarithm with base q.
q	a fixed large integer defined in Definition 2.31.
V_g^+	an attracting point in the projective space of $g \in \mathrm{SL}_n(\mathbb{R})$ defined in Lemma 2.4.
H_g^-	a repelling hyperplane in the projective space of $g \in \mathrm{SL}_n(\mathbb{R})$ defined in Lemma 2.4.
$b(g^-, \epsilon), B(g^+, \epsilon)$	repelling and attracting basins of $\mathbb{P}(\mathbb{R}^n)$ defined in Lemma 2.4.
$E_i, 1 \leqslant i \leqslant 3$	projectifications of standard basis of $\mathbb{R}^3$.
$\Pi_{V^\perp}, \Pi_{V,W}$	linear projections defined in Definition 2.9.
$\Pi(V, W, W') = \Pi_{V,W}\|_{W'}$	the restriction to W' of a linear projection $\Pi_{V,W}$ in Definition 2.9.
$\pi_{V^\perp}, \pi_{V,W}, \pi(V, W, W')$	projective transformations induced by $\Pi_{V^\perp}, \Pi_{V,W}, \Pi(V, W, W')$ in Definition 2.9.
$h_{V,g}$	the composition map $\pi_{V^\perp} \circ g\|_{V^\perp}$ where $g \in \mathrm{SL}_3(\mathbb{R})$.
U_V, L_V decomposition	a decomposition of $\mathrm{SL}_3(\mathbb{R})$, where the parameter V is an element in $\mathbb{P}(\mathbb{R}^3)$.
U, L decomposition	U_V, L_V decomposition with $V = E_1$.
π_{L_V}	the projection from $U_V L_V$ to L_V.
$b(f_\ell, r)$	the complement of the r-neighborhood of a hyperplane in $\mathbb{P}(\mathbb{R}^3)$ Definition 2.13.
$\mathcal{Q}_n$	q-adic decomposition of metric spaces as $\mathbb{R}$ and Lie groups.
$\mathbf{U}(n), \mathbf{I}(n)$	random words defined in Section 2.9.
$g\mu$	the pushforward of μ under the action of g.
$\nu * \mu$	the convolution of ν on $\mathrm{SL}_n(\mathbb{R})$ and μ on $\mathbb{P}(\mathbb{R}^n)$
$\theta * \tau$	the convolution of two measures on $\mathbb{R}$.
$[\theta.\mu]$	the projection $\pi_{E_1^\perp}$ of convolution of θ on $\mathrm{SL}_3(\mathbb{R})$ and μ on $\mathbb{P}(\mathbb{R}^3)$. Section 5.2
$h_{\mathrm{F}}(\mu, \nu)$	the Furstenberg entropy defined in Eq. (8.1)
$h_{\mathrm{RW}}(\nu)$	the random walk entropy of ν defined in Eq. (1.2).
$\sigma_i(g)$	Singular values of $g \in \mathrm{SL}_n(\mathbb{R})$.
$\chi_i(\nu)$	$\lambda_1(\nu) - \lambda_{i+1}(\nu)$, where $(\lambda_1, \ldots, \lambda_n)$ are the Lyapunov exponents of ν.
$\mathcal{F} = \mathcal{F}(\mathbb{R}^3)$	Flag variety in $\mathbb{R}^3$, cf. Definition 3.3.
C_L	constant defined in Eq. (2.29).
C_p	constant defined in Eq. (2.47).

2 Preliminaries

2.1 Action of $\mathrm{SL}_n(\mathbb{R})$ on $\mathbb{P}(\mathbb{R}^n)$

Consider the n-dimensional Euclidean space $\mathbb{R}^n$ and denote by $\|\cdot\|$ the Euclidean norm. By abuse of the notation, we denote by $\|\cdot\|$ the norm on $\wedge^2\mathbb{R}^n$ induced by the one in $\mathbb{R}$. Any element $g \in \mathrm{SL}_n(\mathbb{R})$ acts on $\mathbb{R}^n$ and we denote the operator norm by $\|g\|$. Let $e_1, \cdots, e_n$ be the standard orthonormal basis of $\mathbb{R}^n$ and $e_1^*, \cdots, e_n^*$ be the dual basis of $(\mathbb{R}^n)^*$. Let $E_i = \mathbb{R}e_i$ be the corresponding element in $\mathbb{P}(\mathbb{R}^n)$ with $i = 1, \cdots, n$. Throughout the paper, we always consider the following metric on the projective space $\mathbb{P}(\mathbb{R}^n)$, unless otherwise stated.

Definition 2.1. The distance d on $\mathbb{P}(\mathbb{R}^n)$ defined by

$$d(\mathbb{R}v, \mathbb{R}w) := \frac{\|v \wedge w\|}{\|v\|\|w\|} \text{ for any } \mathbb{R}v, \mathbb{R}w \in \mathbb{P}(\mathbb{R}^n). \tag{2.1}$$

For two subsets A, B of $\mathbb{P}(\mathbb{R}^n)$, we write $d(A, B)$ for their Hausdorff distance.

Remark 2.2. The metric d is bi-Lipschitz equivalent to the standard $\mathrm{SO}(n)$-invariant Riemannian metric d_R on $\mathbb{P}(\mathbb{R}^n)$ (coming from the double cover of $\mathbb{P}(\mathbb{R}^n)$ by the unit sphere). More precisely, we have $d(\mathbb{R}v, \mathbb{R}w) = \sin(d_R(\mathbb{R}v, \mathbb{R}w))$ for any $\mathbb{R}v, \mathbb{R}w \in \mathbb{P}(\mathbb{R}^n)$.

We will frequently consider the following subgroups of $\mathrm{SL}_n(\mathbb{R})$:

$$\begin{aligned} K :=& \mathrm{SO}_n(\mathbb{R}), \\ A :=& \{a = \mathrm{diag}(a_1, \dots, a_n) : a_i \neq 0, a_1 \cdots a_n = 1\}, \\ A^+ :=& \{a = \mathrm{diag}(a_1, \dots, a_n) \in A : a_1 \geqslant \dots \geqslant a_n\}. \end{aligned}$$

Definition 2.3. For $g \in \mathrm{SL}_n(\mathbb{R})$, let $\sigma_1(g) \geqslant \cdots \geqslant \sigma_n(g)$ be the singular values of g and $g = \tilde{k}_g a_g k_g \in KA^+K$ be the Cartan decomposition of g (singular value decomposition). Let $\epsilon > 0$. Set

- $\chi_i := \log(\sigma_1/\sigma_{i+1}), 1 \leqslant i \leqslant n-1$;
- $V_g^+ := \tilde{k}_g E_1 \in \mathbb{P}(\mathbb{R}^n)$, which is an attracting point of g.
- $H_g^- := k_g^{-1}(E_2 \oplus \cdots \oplus E_n) \subset \mathbb{P}(\mathbb{R}^d)$, which is a repelling hyperplane of g (if $\sigma_1 > \sigma_2$, then V_g^+ and H_g^- are uniquely defined);
- $b(g^-, \epsilon) := \{x \in \mathbb{P}(\mathbb{R}^n) : \ d(x, H_g^-) > \epsilon\}$ for any $\epsilon > 0$;
- $B(g^+, \epsilon) := \{x \in \mathbb{P}(\mathbb{R}^n) : \ d(x, V_g^+) \leqslant \epsilon\}$ for any $\epsilon > 0$.

The Cartan projection of g is defined to be $\kappa(g) := (\log \sigma_1(g), \cdots, \log \sigma_n(g))$.

We now state some basic contracting properties of the action of $\mathrm{SL}_n(\mathbb{R})$ on $\mathbb{P}(\mathbb{R}^n)$.

Lemma 2.4. *For any $g \in \mathrm{SL}_n(\mathbb{R})$ and any $\epsilon > 0$, we have*

- *g acts on $b(g^-, \epsilon)$ by contraction: for any $x \neq y$ in $b(g^-, \epsilon)$, we have $\frac{d(g(x), g(y))}{d(x,y)} \leqslant \frac{\sigma_2}{\epsilon^2 \sigma_1}$;*
- *$g(b(g^-, \epsilon)) \subset B(g^+, \frac{\sigma_2}{\epsilon^2 \sigma_1})$;*
- *when $g \in \mathrm{SL}_2(\mathbb{R})$, we have $\frac{\sigma_1(g)}{\sigma_2(g)} = \|g\|^2$.*

Please see [BQ16] or [Li22, Lemma 2.11] for the proof.
We state another useful lemma [BQ16, Lemma 14.2] for later use.

Lemma 2.5. *For any $g \in \mathrm{SL}_n(\mathbb{R})$ and $V = \mathbb{R}v \in \mathbb{P}(\mathbb{R}^n)$, we have*

$$d(V, H_g^-) \leqslant \frac{\|gv\|}{\|g\|\|v\|} \leqslant d(V, H_g^-) + q^{-\chi_1(g)}, \quad d(gV, V_g^+)d(V, H_g^-) \leqslant q^{-\chi_1(g)}.$$

Definition 2.6. For a bi-Lipschitz map f between two metric spaces (X, d_X) and (Y, d_Y), we say f scales by $u > 0$ with distortion $C > 1$ if for any $x \neq x' \in X$,

$$\frac{1}{C} \leqslant \frac{d_Y(f(x), f(x'))}{u d_X(x, x')} \leqslant C. \tag{2.2}$$

We fix an identification between $\mathbb{P}(\mathbb{R}^2)$ and $\mathbb{R}/\mathbb{Z}$:

$$\begin{aligned} \iota : \mathbb{P}(\mathbb{R}^2) &\to \mathbb{R}/\mathbb{Z} \\ \mathbb{R}(\cos\theta, \sin\theta) &\mapsto \frac{\theta}{\pi}. \end{aligned} \tag{2.3}$$

When an element $g \in \mathrm{SL}_2(\mathbb{R})$ acts on $\mathbb{P}(\mathbb{R}^2)$, we view g as a diffeomorphism of $\mathbb{R}/\mathbb{Z}$ by conjugating it by ι and denote by $|g'x|$ and $|g''x|$ the 1st and 2nd derivatives of $\iota g \iota^{-1}$ at ιx. The followings are estimates of $|g'x|$ and $|g''x|$ which can be found in [HS17, Page 826].

Lemma 2.7. *For $g \in \mathrm{SL}_2(\mathbb{R})$ and $x \in \mathbb{P}(\mathbb{R}^2)$, we have*

$$\|g\|^{-2} \leqslant |g'x| \leqslant \|g\|^2, \ |g''x| \leqslant 4\|g\|^2.$$

Lemma 2.8. *For all $0 < \epsilon < 1/3$, the following holds. Let g be an element in $\mathrm{SL}_2(\mathbb{R})$. Then for action of g on $b(g^-, \epsilon) \subset \mathbb{P}(\mathbb{R}^2)$, it scales by $\|g\|^{-2}$ with distortion $10\epsilon^{-2}$. Actually, we have a more precise estimate: for any $x \in b(g^-, \epsilon)$,*

$$\begin{aligned} -2\log\|g\| \leqslant \log|g'(x)| &\leqslant -2\log\|g\| - 2\log(\epsilon), \\ |g''(x)| &\leqslant \frac{10}{\epsilon^3\|g\|^2}. \end{aligned}$$

2.2 Projections in $\mathbb{R}^3$ and $\mathbb{P}(\mathbb{R}^3)$

Definition 2.9. For any line V and any hyperplanes W, W' of $\mathbb{R}^3$ such that $V \not\subset W, W'$, and any $g \in \mathrm{SL}_3(\mathbb{R})$ such that $g^{-1}V \not\subset V^\perp$, we denote

- the linear projection on $\mathbb{R}^3$ with kernel V and image W by $\Pi_{V,W}$;
- the orthogonal projection with kernel V by $\Pi_{V^\perp}(= \Pi_{V,V^\perp})$;
- the linear projection from W' to W along V (i.e. with kernel V) by $\Pi(V, W, W')$, i.e. $\Pi(V, W, W') = \Pi_{V,W}|_{W'}$;
- the projective transformations associated to $\Pi_{V,W}$, $\Pi_{V^\perp}$, $\Pi(V, W, W')$ by $\pi_{V,W}$, $\pi_{V^\perp}$, $\pi(V, W, W')$ respectively;
- the composition map $\pi_{V^\perp} \circ g|_{V^\perp}$ by $h_{V,g} : \mathbb{P}(V^\perp) \to \mathbb{P}(V^\perp)$.

One important observation is the following geometric lemma on a decomposition of the map $\pi_{V^\perp} \circ g$, which enables us to change the direction of projection and apply the ergodic theory on projections.

Lemma 2.10. *For any $g \in \mathrm{SL}_3(\mathbb{R})$, $V \in \mathbb{P}(\mathbb{R}^3)$ such that $g^{-1}V \not\subset V^\perp$, we have:*

$$\pi_{V^\perp} \circ g = h_{V,g} \circ \pi_{g^{-1}V, V^\perp}, \tag{2.4}$$

$$\pi_{V^\perp} \circ g = h_{V,g} \circ \pi(g^{-1}V, V^\perp, (g^{-1}V)^\perp) \circ \pi_{(g^{-1}V)^\perp}. \tag{2.5}$$

Proof. It suffices to show the corresponding equations for linear maps in $\mathbb{R}^3$. We have

$$\begin{aligned} & \Pi_{V^\perp} \circ g \\ = \; & (\Pi_{V^\perp} \circ g)|_{V^\perp} \circ \Pi_{g^{-1}V, V^\perp} \\ = \; & (\Pi_{V^\perp} \circ g)|_{V^\perp} \circ \Pi(g^{-1}V, V^\perp, (g^{-1}V)^\perp) \circ \Pi_{g^{-1}V, (g^{-1}V)^\perp}. \end{aligned}$$

The first equality holds because both $\Pi_{V^\perp} \circ g$ and $(\Pi_{V^\perp} \circ g)|_{V^\perp} \circ \Pi_{g^{-1}V, V^\perp}$ are linear maps from $\mathbb{R}^3$ to $V^\perp$. They have the same kernel which is $g^{-1}V$, and their restrictions to $V^\perp$, a complement to $g^{-1}V$, are the same.

The second equality is obtained similarly by analyzing the kernels and images of the linear maps. □

Remark 2.11. Compared to [BHR19], in self-affine case, the difficulties lie in non-linearity and the non-uniform contracting of the action. Non-uniform contracting forces us to use attracting decomposition and change the measure in step 3 of the proof of Theorem 5.7.

2.3 Decomposition of $\mathrm{SL}_3(\mathbb{R})$ for the composition maps $\pi_{V^\perp} g$

In this subsection, we introduce a decomposition of the Lie group $\mathrm{SL}_3(\mathbb{R})$, which describes the structure of the composition maps $\pi_{V^\perp} g$ with $V \in \mathbb{P}(\mathbb{R}^3)$ and $g \in \mathrm{SL}_3(\mathbb{R})$.

We start with the decomposition for $E_1 \in \mathbb{P}(\mathbb{R}^3)$. Let U, L be two closed Lie subgroups of $\mathrm{SL}_3(\mathbb{R})$ defined by

$$\begin{aligned} U &:= \left\{ \begin{pmatrix} \lambda^2 & x \;\; y \\ 0 & \lambda^{-1}\mathrm{Id}_2 \end{pmatrix} : \lambda \in \mathbb{R}^+, x, y \in \mathbb{R} \right\}, \\ L &:= \left\{ \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix} : n \in \mathbb{R}^2, h \in \mathrm{SL}_2(\mathbb{R}) \right\}. \end{aligned}$$

The group U is a solvable subgroup, and the group L is isomorphic to $\mathrm{SL}_2(\mathbb{R}) \ltimes \mathbb{R}^2$.

For a general $V \in \mathbb{P}(\mathbb{R}^3)$, we define $U_V := k^{-1}Uk$ and $L_V := k^{-1}Lk$, where k is any matrix in $\mathrm{SO}_3(\mathbb{R})$ satisfying $kV = E_1$. We also define the projection

$$\begin{aligned} \pi_{L_V} : U_V L_V &\to L_V \\ u\ell &\mapsto \ell. \end{aligned} \tag{2.6}$$

The following lemma shows that U_V, L_V, π_{L_V} are well-defined and explains the roles that U_V and L_V play in the composition maps $\pi_{V^\perp} g$ with $g \in \mathrm{SL}_3(\mathbb{R})$.

Lemma 2.12. *For any $V \in \mathbb{P}(\mathbb{R}^3)$ and any $k \in \mathrm{SO}_3(\mathbb{R})$ such that $kV = E_1$. The followings hold.*

1. *The groups U_V and L_V are well-defined, i.e., they are independent of the choice of $k \in \mathrm{SO}_3(\mathbb{R})$.*
2. *The product $U_V L_V$ equals the set*
$$\{g \in \mathrm{SL}_3(\mathbb{R}) : g^{-1}V \notin V^\perp\},$$
and it is Zariski open and Zariski dense in $\mathrm{SL}_3(\mathbb{R})$.
3. *The map π_{L_V} is well-defined, i.e., if $u\ell = u'\ell'$ in $U_V L_V$, then $\pi_{L_V}(u\ell) = \pi_{L_V}(u'\ell')$.*
4. *For the group U_V, we have*
$$\{g \in \mathrm{SL}_3(\mathbb{R}) : \pi_{V^\perp} g = \pi_{V^\perp} id\} = U_V \cup U_V k^{-1} \begin{pmatrix} 1 & 0 \\ 0 & -\mathrm{Id}_2 \end{pmatrix} k. \tag{2.7}$$

5. *For every $g \in U_V L_V$, write $g = k^{-1}u\ell k = k^{-1}u\begin{pmatrix}1 & 0\\ n & h\end{pmatrix}k$ with $u \in U$, $\ell \in L$, $n \in \mathbb{R}^2$ and $h \in \mathrm{SL}_2(\mathbb{R})$. We have*
$$h_{V,g} = h$$
as elements in $\mathrm{PSL}_2(\mathbb{R})$, where $h_{V,g}$ is given as in Definition 2.9.

6. *Consider the map*
$$\begin{aligned}\Phi : L_V &\to \{\text{maps from } \mathbb{P}(\mathbb{R}^3) \text{ to } \mathbb{P}(V^\perp)\}\\ \ell &\mapsto \pi_{V^\perp}\ell.\end{aligned}$$
If $\ell \neq \ell' \in L$ satisfy $\Phi(\ell) = \Phi(\ell')$, then $\ell^{-1}\ell' = k^{-1}\begin{pmatrix}1 & 0\\ 0 & -\mathrm{Id}_2\end{pmatrix}k$.

Proof. 1. For any $V \in \mathbb{P}(\mathbb{R}^3)$, let k_0 be a matrix in $\mathrm{SO}_3(\mathbb{R})$ satisfying $kV = E_1$. We have
$$\{k \in \mathrm{SO}_3(\mathbb{R}) : kV = E_1\} = \left\{\begin{pmatrix}1 & 0\\ 0 & k_1\end{pmatrix}k_0 : k_1 \in \mathrm{SO}_2(\mathbb{R})\right\}.$$

Any matrix of the form $\begin{pmatrix}1 & 0\\ 0 & k_1\end{pmatrix}$ with $k_1 \in \mathrm{SO}_2(\mathbb{R})$ normalizes the groups U and L. This shows that U_V and L_V are well-defined.

With the first statement available, it suffices to prove the remaining statements for the case $V = E_1$.

2. We prove
$$UL = \{g \in \mathrm{SL}_3(\mathbb{R}) : g^{-1}E_1 \notin E_1^\perp\}. \tag{2.8}$$
The direction $\subset$ can be checked by a straightforward computation using the definitions of U and L.

For the direction $\supset$, given any $g \in \mathrm{SL}_3(\mathbb{R})$ such that $g^{-1}E_1 \notin E_1^\perp$, the entry in the first row and first column of g^{-1} is non-zero. So the 2 by 2 submatrix g' of g in the lower right corner is non-degenerate. We can write $g' = \lambda^{-1}h$ with $\lambda > 0$ and $h \in \mathrm{SL}_2(\mathbb{R})$, and hence $g = u\ell$ with $u \in U$ and $\ell \in L$.

UL is Zariski open and Zariski dense because (2.8) shows that it is the complement of the proper Zariski closed subvariety $\{g \in \mathrm{SL}_3(\mathbb{R}) : g^{-1}E_1 \in E_1^\perp\}$.

3. Suppose $u\ell = u'\ell'$ in UL. Then $\ell'\ell^{-1} = u'^{-1}u \in U \cap L = \{\mathrm{Id}_3\}$.

4. We can check that the group U is contained in the group (2.7) by a straightforward computation using the definition of U.

Given any $g \in \mathrm{SL}_3(\mathbb{R})$ such that $\pi_{E_1^\perp}g = \pi_{E_1^\perp}\mathrm{id}$, we have $gE_1 = E_1$. So $g \in UL$ by (2.8) and we can write $g = u\ell$ with $u \in U$ and $\ell \in L$. The relation $E_1 = gE_1 = u\ell E_1$ gives $\ell E_1 = E_1$. As a result, we can write $\ell = \begin{pmatrix}1 & 0\\ 0 & h\end{pmatrix}$ with $h \in \mathrm{SL}_2(\mathbb{R})$. So
$$\pi_{E_1^\perp}\mathrm{id} = \pi_{E_1^\perp}g = \pi_{E_1{}^\perp}\begin{pmatrix}1 & 0\\ 0 & h\end{pmatrix} = h.$$

This implies $h = \mathrm{Id}_2$ or $-\mathrm{Id}_2$.

5. The fifth statement holds because
$$\pi_{E_1^\perp}g|_{E_1^\perp} = \pi_{E_1^\perp}\begin{pmatrix}1 & 0\\ n & h\end{pmatrix}\Big|_{E_1^\perp} = h.$$

6. The sixth statement follows from the third statement and the fact that $L \cap U = \{\mathrm{Id}_3\}$.

$\square$

Lemma A.1 fixes an identification between $\mathbb{P}(V^\perp)$ with $\mathbb{P}(\mathbb{R}^2)$ for $V \in \mathcal{C} = \mathbb{P}(\mathbb{R}^3) - E_3$. This identification provides a unique choice of $k \in \mathrm{SO}_3(\mathbb{R})$.

2.4 Linearize the projection of L-action

In this part, we consider the point $E_1 \in \mathbb{P}(\mathbb{R}^3)$ and the corresponding UL-decomposition. We fix a left L-invariant and right $\mathrm{SO}(2)$-invariant Riemannian metric d on L. For $\ell \in L$ and $x \in \mathbb{P}(\mathbb{R}^3)$, we introduce the following notation

$$[\ell(x)] := \pi_{E_1^\perp}\ell(x) \in \mathbb{P}(E_1^\perp) \cong \mathbb{P}(\mathbb{R}^2).$$

We will always write $\ell = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix}$ with $n \in \mathbb{R}^2$ and $h \in \mathrm{SL}_2(\mathbb{R})$ as the matrix representation of ℓ. Then

$$[\ell(x)] = \mathbb{R}(h(b,c)^t + an), \tag{2.9}$$

for any $x = \mathbb{R}(a,b,c)^t \in \mathbb{P}(\mathbb{R}^3)$.

The matrix h acts on $\mathbb{P}(E_1^\perp) \cong \mathbb{P}(\mathbb{R}^2)$. Let $h = \tilde{k}_h a_h k_h \in KA^+K$ be the Cartan decomposition of h, and let $H_h^- \in \mathbb{P}(E_1^\perp)$ be its repelling point given as in Definition 2.3.

Observe that by Eq. (2.4) and Lemma 2.12, for $\ell = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix}$, we have

$$\pi_{E_1{}^\perp}\ell = h \circ \pi_{\ell^{-1}E_1, E_1^\perp}. \tag{2.10}$$

(As in Figure 2, for any x, the projection $\pi_{\ell^{-1}E_1,E_1^\perp}x$ can be viewed as the intersection of the projective line joint by x and $\pi_{(\ell^{-1}E_1)^\perp}x$ with the projective line in red $\mathbb{P}(E_1^\perp)$, and we decompose $\pi_{E_1^\perp}\ell$ as the composition of h with the projection $\pi_{\ell^{-1}E_1,E_1^\perp}$) The contracting region of $\pi_{E_1^\perp}\ell$

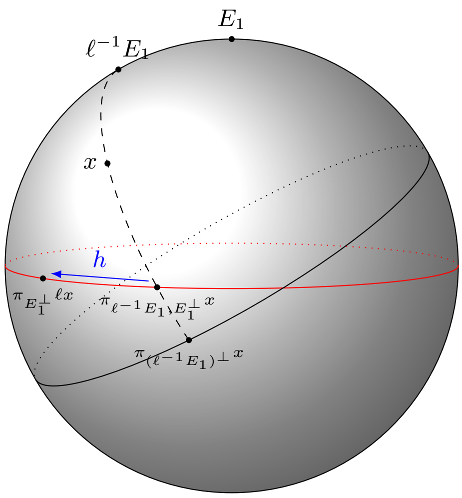

Figure 2: $\pi_{E_1{}^\perp}\ell = h \circ \pi_{\ell^{-1}E_1,E_1^\perp}$

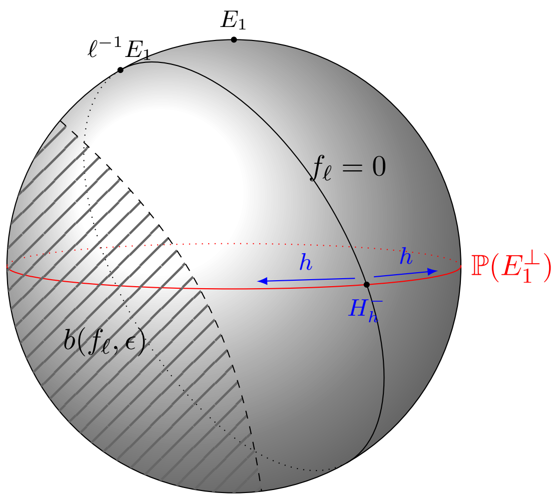

Figure 3: The $[\ell]$-attracting region

should be away from the projective plane spanned by $\ell^{-1}E_1$ and H_h^-. At the same time, note that we have the formula for the linear map

$$\begin{aligned} \Pi_{\ell^{-1}E_1,E_1^\perp} : \mathbb{R}^3 &\to \mathbb{R}^2 \\ (a,b,c)^t &\mapsto (b,c)^t + ah^{-1}n. \end{aligned}$$

This motivates the following construction.

Let f_h be a linear form on $E_1^{\perp} \cong \mathbb{R}^2$ with kernel H_h^- of norm 1. For any non-zero vector $v \in \mathbb{R}^2$, we have

$$|f_h(v)| = \|v\| d(\mathbb{R}v, H_h^-). \tag{2.11}$$

Definition 2.13. Let $\ell = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix} \in L$. We define the linear form f_ℓ on $\mathbb{R}^3$ by

$$f_\ell(v) := f_h(\Pi_{\ell^{-1}E_1, E_1^{\perp}}(v)) \text{ for } v \in \mathbb{R}^3.$$

The kernel of f_ℓ is the plane spanned by $\ell^{-1}E_1 = \mathbb{R}(1, -h^{-1}n)^t$ [4] and $H_h^- \subset E_1^{\perp} \subset \mathbb{R}^3$.

For $\epsilon > 0$, we define an $[\ell]$-attracting region by

$$b(f_\ell, \epsilon) := \left\{ x = \mathbb{R}v \in \mathbb{P}(\mathbb{R}^3) : |f_\ell(v)| \geqslant \epsilon \|f_\ell\| \|v\| \right\}. \tag{2.12}$$

Figure 3 shows the region $b(f_l, \epsilon)$, here the arrows in blue means h is repelling near $H_h^- \in \mathbb{P}(E_1^{\perp})$.

We summarize the structure of the rest of this subsection: in Lemma 2.14, we prove the contracting properties of $\pi_{E_1^{\perp}}\ell(\cdot) = [\ell(\cdot)]$ on the region $b(f_\ell, \epsilon)$; Lemma 2.15, Lemma 2.17 and Lemma 2.18 collect the continuity properties of ℓ, f_ℓ and $b(f_\ell, \epsilon)$; we also need an estimate of the diameter of a set acted by the projection of L-action (Lemma 2.19); Lemma 2.21 is the linearization of the map $\pi_{E_1^{\perp}}\ell$.

We start with some basic estimates which will be frequently used. Note that for any $\ell \in L$,

$$d(\ell^{-1}E_1, E_1^{\perp}) = d(\mathbb{R}(1, -h^{-1}n)^t, E_1^{\perp}) = 1/\|(1, -h^{-1}n)^t\|. \tag{2.13}$$

Hence

$$(1 + \|h^{-1}n\|^2)^{1/2} = 1/d(\ell^{-1}E_1, E_1^{\perp}). \tag{2.14}$$

For $x = \mathbb{R}v$ with $v = (a, b, c)^t$ and $\|v\| = 1$, we have

- $$\|\Pi_{\ell^{-1}E_1, E_1^{\perp}}(v)\| \leqslant 1/d(\ell^{-1}E_1, E_1^{\perp}); \tag{2.15}$$
- $$\|f_\ell\| \in [1, 1/d(\ell^{-1}E_1, E_1^{\perp})]; \tag{2.16}$$
- if $x \in b(f_\ell, \epsilon)$, then
$$\|\Pi_{\ell^{-1}E_1, E_1^{\perp}}(v)\| \geqslant \epsilon \|v\|, \tag{2.17}$$
$$d(x, \ell^{-1}E_1) \geqslant \epsilon d(\ell^{-1}E_1, {E_1}^{\perp}). \tag{2.18}$$

Eq. (2.15) holds because

$$\|\Pi_{\ell^{-1}E_1, E_1^{\perp}}(v)\| = \|(b, c)^t + a h^{-1}n\| \leqslant \|(1, -h^{-1}n)^t\|.$$

Eq. (2.17) holds because

$$\|\Pi_{\ell^{-1}E_1, E_1^{\perp}}(v)\| \geqslant |f_h(\Pi_{\ell^{-1}E_1, E_1^{\perp}}(v))| \geqslant \epsilon \|f_\ell\| \|v\| \geqslant \epsilon \|v\|.$$

Eq. (2.16) is obtained as follows. For any $v \in E_1^{\perp}$, $f_\ell(v) = f_h(v)$. As the norm of f_h is 1, $\|f_\ell\| \geqslant 1$. Meanwhile, by definition, we have $\|f_\ell\| \leqslant \|f_h\| \cdot \|\Pi_{\ell^{-1}E_1, E_1^{\perp}}\|$. So we use Eq. (2.15) to get the upper bound.

[4] We abuse the notation here and $(1, -h^{-1}n)$ is actually the row vector $(1, -(h^{-1}n)^t)$.

We obtain Eq. (2.18) as follows:

$$\begin{aligned} d(x, \ell^{-1}E_1) &= \frac{\|(a,b,c)^t \wedge (1, -h^{-1}n)^t\|}{\|(1,-h^{-1}n)^t\|} \\ &\geqslant d(\ell^{-1}E_1, E_1^{\perp})\|(b,c)^t + ah^{-1}n\| = d(\ell^{-1}E_1, E_1^{\perp})\|\Pi_{\ell^{-1}E_1, E_1^{\perp}}(v)\|, \end{aligned}$$

where for the inequality we project the vector $(a,b,c)^t \wedge (1, -h^{-1}n)^t$ to the subspace generated by $E_1 \wedge E_2, E_1 \wedge E_3$ in $\wedge^2\mathbb{R}^3$ and due to Eq. (2.13).

Lemma 2.14. *Let $C_1, C_2 > 2$. Let $\ell = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix} \in L$ be such that $d(\ell^{-1}E_1, E_1^{\perp}) > 1/C_1$. Then for $x = \mathbb{R}v$ with $\|v\| = 1$ in the $[\ell]$-attracting region $b(f_\ell, 1/C_2)$, we have*

$$\|\Pi_{E_1^{\perp}}\ell(v)\| = \|h \circ \Pi_{\ell^{-1}E_1, E_1^{\perp}}(v)\| \geqslant \|h\|/C_2, \tag{2.19}$$

$$d(\pi_{\ell^{-1}E_1, E_1^{\perp}}(x), H_h^-) \geqslant 1/C_1C_2. \tag{2.20}$$

Moreover, for the action of $\pi_{E_1^{\perp}}\ell$ on $b(f_\ell, 1/C_2)$, we have

$$\frac{d([\ell(x)], [\ell(x')])}{d(x,x')} \leqslant 2C_1C_2^2/\|h\|^2 \tag{2.21}$$

for any two distinct points x, x' in $b(f_\ell, 1/C_2)$, and the inclusion

$$[\ell(b(f_\ell, 1/C_2))] \subset B(h^+, C_1C_2/\|h\|^2). \tag{2.22}$$

Proof. We write $w = \Pi_{\ell^{-1}E_1, E_1^{\perp}}(v)$. Hence $\mathbb{R}w = \pi_{\ell^{-1}E_1, E_1^{\perp}}(x)$.

Eq. (2.19) holds because

$$\|hw\| \geqslant \|h\|\|w\| d(\mathbb{R}w, H_h^-) = \|h\||f_h(w)| = \|h\||f_\ell(v)| \geqslant \|h\|/C_2.$$

Here the first inequality follows from Lemma 2.5; the second one is by Eq. (2.11); the last inequality is due to $\|f_\ell\| \geqslant 1$ (Eq. (2.16)) and the definition of $b(f_\ell, 1/C_2)$.

As $\|v\| = 1$, by Eq. (2.15), we have

$$\|w\| = \|\Pi_{\ell^{-1}E_1, E_1^{\perp}}(v)\| \leqslant 1/d(\ell^{-1}E_1, E_1^{\perp}) < C_1. \tag{2.23}$$

Eq. (2.20) holds because

$$d(\mathbb{R}w, H_h^-) = |f_h(w)|/\|w\| \geqslant 1/C_1C_2.$$

For Eq. (2.22), note that Eq. (2.10) gives

$$[\ell(x)] = h(\mathbb{R}w).$$

With Eq. (2.20) available, we apply Lemma 2.4 to the action of h and obtain the inclusion relation.

It remains to prove Eq. (2.21). Take any point $x' \neq x$ in $b(f_\ell, 1/C_2)$. Write $x' = \mathbb{R}(a', b', c')^t$ with $\|(a', b', c')^t\| = 1$. We have

$$\begin{aligned} d([\ell x], [\ell x']) &= \frac{\|(h(b,c)^t + an) \wedge (h(b', c')^t + a'n)\|}{\|h(b,c)^t + an\|\|h(b',c')^t + a'n\|} \\ &\leqslant \left(\frac{C_2}{\|h\|}\right)^2 \|(h(b,c)^t + an) \wedge (h(b', c')^t + a'n)\|, \end{aligned} \tag{2.24}$$

where we use Eq. (2.19) to obtain the inequality.

Note that

$$d(x,x') = \|(a,b,c)^t \wedge (a',b',c')^t\| \geqslant \max\{\|(b,c)^t \wedge (b',c')^t\|, \|a(b',c')^t - a'(b,c)^t\|\}. \qquad (2.25)$$

Then the numerator of Eq. (2.24) is equal to

$$\begin{aligned}
&\|h(b,c)^t \wedge h(b',c')^t + n \wedge h(a(b',c')^t - a'(b,c)^t)\| \\
=&\|(b,c)^t \wedge (b',c')^t + h^{-1}n \wedge (a(b',c')^t - a'(b,c)^t)\| \\
\leqslant&\|(b,c)^t \wedge (b',c')^t\| + \|h^{-1}n\| \cdot \|a(b',c')^t - a'(b,c)^t\| \leqslant 2C_1 d(x,x'),
\end{aligned} \qquad (2.26)$$

where the second line is due to $h \in \mathrm{SL}_2(\mathbb{R})$ and the third line is due to $\|h^{-1}n\| \leqslant 1/d(\ell^{-1}E_1, E_1^{\perp}) \leqslant C_1$ (see Eq. (2.14)). This verifies Eq. (2.21). □

Lemma 2.15. *There exist $\epsilon_1 > 0$ and $C_3 > 1$ such that for $\ell = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix}$ and $\ell' = \begin{pmatrix} 1 & 0 \\ n' & h' \end{pmatrix}$ in L with $d(\ell,\ell') \leqslant \epsilon_1$, we have*

$$\begin{aligned}
&\|id - h^{-1}h'\| \leqslant C_3 d(\ell,\ell'), \ \|h^{-1}n' - h^{-1}n\| \leqslant C_3 d(\ell,\ell'), \\
&\|h^{-1}n - h'^{-1}n'\| \leqslant C_3 d(\ell,\ell')/d(\ell^{-1}E_1, E_1^{\perp}), \\
&\big|\log\|h'\| - \log\|h\|\big| \leqslant C_3 d(\ell,\ell'), \\
&d(\ell^{-1}E_1, (\ell')^{-1}E_1) \leqslant C_3 d(\ell,\ell')/d(\ell^{-1}E_1, E_1^{\perp}).
\end{aligned} \qquad (2.27)$$

Moreover, if ℓ, ℓ' satisfy $d(\ell',\ell) \leqslant \min\{\epsilon_1, 1/C_3\}$, then

$$d((\ell')^{-1}E_1, E_1^{\perp}) \geqslant d(\ell^{-1}E_1, E_1^{\perp})/2.$$

Proof. For any $\ell, \ell' \in L$, as d is a left L-invariant metric on L, we have

$$d(\ell,\ell') = d(id, \ell^{-1}\ell') = d\left(id, \begin{pmatrix} 1 & 0 \\ h^{-1}n' - h^{-1}n & h^{-1}h' \end{pmatrix}\right).$$

The metric d and the distance induced by the norm are locally bi-Lipschitz. Hence we obtain the two inequalities of the first line.

We also have

$$\begin{aligned}
&\|h^{-1}n - h'^{-1}n'\| = \|h^{-1}n - h^{-1}n' + h^{-1}n' - h'^{-1}n'\| \\
\ll& d(\ell,\ell') + \|(h^{-1}h' - id)h'^{-1}n'\| \ll d(\ell,\ell')(1 + \|h'^{-1}n'\|).
\end{aligned}$$

Using Eq. (2.14), we obtain Eq. (2.27).

For $d((\ell')^{-1}E_1, E_1^{\perp})$, using Eq. (2.14) and Eq. (2.27), we have

$$\begin{aligned}
1/d((\ell')^{-1}E_1, E_1^{\perp}) &= \|(1, -h'^{-1}n')\| \leqslant \|(1, -h^{-1}n)\| + \|h'^{-1}n' - h^{-1}n\| \\
&\leqslant 1/d(\ell^{-1}E_1, E_1^{\perp}) + C_3 d(\ell,\ell')/d(\ell^{-1}E_1, E_1^{\perp}) \leqslant 2/d(\ell^{-1}E_1, E_1^{\perp}).
\end{aligned}$$

For $\log\|h\|$, we may suppose $\|h\| \geqslant \|h'\|$. Then we have

$$\log\|h\| - \log\|h'\| = \log\left(1 + \frac{\|h\| - \|h'\|}{\|h'\|}\right) \leqslant \frac{\|h\| - \|h'\|}{\|h'\|} \leqslant \frac{\|h - h'\|}{\|h'\|} \leqslant \|h'^{-1}h - id\|.$$

For $d(\ell^{-1}E_1, \ell'^{-1}E_1)$, by definition and Eq. (2.27), we have

$$d(\ell^{-1}E_1, \ell'^{-1}E_1) = \frac{\|(1, -h^{-1}n)^t \wedge (1, -h'^{-1}n')^t\|}{\|(1, -h^{-1}n)^t\|\|(1, -h'^{-1}n')^t\|}$$

The numerator is bounded above by

$$\begin{aligned}
&\|h^{-1}n - h'^{-1}n'\| + \|h^{-1}n \wedge h'^{-1}n'\| \\
=&\|h^{-1}n - h'^{-1}n'\| + \|(h^{-1}n - h'^{-1}n') \wedge h'^{-1}n'\| \leqslant (1 + \|h'^{-1}n'\|)\|h^{-1}n - h'^{-1}n'\|,
\end{aligned}$$

which yields the estimate of $d(\ell^{-1}E_1, \ell'^{-1}E_1)$. □

We need a lemma about the continuity of Cartan decomposition. For any $\epsilon > 0$, let $\mathcal{O}_\epsilon$ be the neighborhood of the identity in $\mathrm{SL}_2(\mathbb{R})$ with radius ϵ. Recall that A is the diagonal subgroup. Let $\mathfrak{a}$ be its Lie algebra and $\mathfrak{a}^{++}$ its positive Weyl chamber. Let $\widetilde{A}^\delta = \{\exp(a),\ a \in \mathfrak{a}^{++},\ d(a, \partial\mathfrak{a}^{++}) \geqslant \delta\}$.

Lemma 2.16 (Effective Cartan decomposition, [GOS10], Theorem 1.6)**.** *Given $\delta > 0$, there exist $l_0, \epsilon_2 > 0$ such that the following holds. For any $\epsilon < \epsilon_2$ and any $g = k_1 a k_2 \in K\widetilde{A}^\delta K$, we have*

$$\mathcal{O}_\epsilon g \mathcal{O}_\epsilon \subset (\mathcal{O}_{l_0\epsilon} \cap K) k_1 (\mathcal{O}_{l_0\epsilon} \cap A) a k_2 (\mathcal{O}_{l_0\epsilon} \cap K).$$

Since we will mostly deal with $h \in \mathrm{SL}_2(\mathbb{R})$ with $\|h\|$ large, we fix $\delta = \log 2$ and the corresponding constants l_0, ϵ_2 in the last lemma.

Lemma 2.17. *There exists $C_4 > 1$ such that the following holds. Let $\ell = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix} \in L$ such that $\|h\| > 2$. The linear form f_ℓ is continuous with respect to ℓ: if $d(\ell, \ell') \leqslant \min\{\epsilon_1, \epsilon_2\}$, then*

$$\|f_\ell - f_{\ell'}\| \leqslant C_4 d(\ell, \ell') / d(\ell^{-1}E_1, E_1^\perp).$$

Proof. Write $\ell' = \begin{pmatrix} 1 & 0 \\ n' & h' \end{pmatrix} \in L$. When $d(\ell, \ell') \leqslant \epsilon_1$, Lemma 2.15 gives that

$$\|id - h^{-1}h'\| \ll d(\ell, \ell').$$

Let $h = \tilde{k}_h a_h k_h$ and $h' = \tilde{k}_{h'} a_{h'} k_{h'}$ be the Cartan decomposition of h and h' respectively. Then Lemma 2.16 gives

$$d(k_h, k_{h'}) \leqslant l_0 d(h, h').$$

Combining these two inequalities, we obtain

$$\|f_h - f_{h'}\| = \|e_1^* \circ k_h - e_1^* \circ k_{h'}\| \leqslant l_0 d(h, h') \ll d(\ell, \ell'). \tag{2.28}$$

For any $x = \mathbb{R}(a, b, c)^t$ with $\|(a, b, c)^t\| = 1$, by Definition 2.13, we obtain

$$\begin{aligned} &|f_\ell((a,b,c)^t) - f_{\ell'}((a,b,c)^t)| = |f_h((b,c)^t + ah^{-1}n) - f_{h'}((b,c)^t + a(h')^{-1}n')| \\ \leqslant &\|f_h - f_{h'}\| \cdot \|(b,c)^t + ah^{-1}n\| + |f_{h'}(ah^{-1}n - a(h')^{-1}n')|. \end{aligned}$$

To get the stated inequality, we use Eq. (2.28) and Eq. (2.15) to estimate the first term, and Eq. (2.27) for the second term. ☐

We fix a large constant $C_L > 0$ such that

$$C_L > \max\{1/\epsilon_1, C_3, 1/\epsilon_2, 4C_4\} \tag{2.29}$$

where the constants $\epsilon_1, \epsilon_2, C_3$ and C_4 are given as in Lemma 2.15, Lemma 2.16 and Lemma 2.17.

Lemma 2.18. *Fix any $C_1, C_2 > 2$. Let $\ell = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix} \in L$ be such that $\|h\| > 2$. If $d(\ell^{-1}E_1, {E_1}^\perp) > 1/C_1$, then for any $\ell' = \begin{pmatrix} 1 & 0 \\ n' & h' \end{pmatrix} \in B(\ell, 1/C_L C_1 C_2) \subset L$, we have*

$$\begin{gathered} \frac{1}{2}\|h\| \leqslant \|h'\| \leqslant 2\|h\|, \\ b(f_\ell, 1/C_2) \subset b(f_{\ell'}, 1/2C_2). \\ d((\ell')^{-1}E_1, E_1^\perp) > 1/(2C_1) \end{gathered}$$

Proof. The first line follows from Lemma 2.15.

For the second line, due to Lemma 2.17 and $\|f_\ell\| \geqslant 1$ (Eq. (2.16)), we have

$$| \|f_\ell\| - \|f_{\ell'}\| | \leqslant 1/4C_2 \leqslant \|f_\ell\|/4C_2.$$

Hence $\|f_\ell\| \geqslant \|f_{\ell'}\|/(1+1/4C_2)$. Again, due to Lemma 2.17, for any $\mathbb{R}v \in b(f_\ell, 1/C_2)$ with $\|v\| = 1$, we have

$$|f_{\ell'}(v)| \geqslant |f_\ell(v)| - 1/4C_2 \geqslant \|f_\ell\|/C_2 - 1/4C_2 \geqslant \|f_{\ell'}\|/(C_2+1/4) - \|f_{\ell'}\|/4C_2 \geqslant \|f_{\ell'}\|/2C_2.$$

The third line follows from Lemma 2.15. □

For sets $E \subset L$ and $F \subset \mathbb{P}(\mathbb{R}^3)$, let $[E.F] = \{[\ell(x)],\ \ell \in E,\ x \in F\}$ be a subset in $\mathbb{P}(\mathbb{R}^2)$.

Lemma 2.19. *Fix any $C_1, C_2 > 2$. Let E be a subset in L such that any $\ell = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix} \in E$ satifies $\|h\| > 2$ and $\operatorname{diam} E < 1/C_L C_1 C_2$. Suppose there exists $\ell \in E$ such that $d(\ell^{-1}E_1, E_1^\perp) > 1/C_1$. Let F be a subset in $[\ell]$-attracting region $b(f_\ell, 1/C_2) \subset \mathbb{P}(\mathbb{R}^3)$. Then*

$$\operatorname{diam}([E.F]) \leqslant 16 C_L C_1 C_2^2 (\operatorname{diam} E + \operatorname{diam} F)\|h\|^{-2}.$$

Proof. For any $\ell \in E$ and $x, x' \in F$, Lemma 2.14 gives $d([\ell x], [\ell x']) \leqslant 2C_1 C_2^2 d(x, x')\|h\|^{-2}$.

Let $\ell, \ell' \in E$ and $x \in F$. Write $\ell' = \begin{pmatrix} 1 & 0 \\ n' & h' \end{pmatrix}$ and $x = \mathbb{R}(a,b,c)^t$ with $\|(a,b,c)^t\| = 1$. We have

$$d([\ell x], [\ell' x]) = \frac{\|(h(b,c)^t + an) \wedge (h'(b,c)^t + an')\|}{\|h(b,c)^t + an\|\|h'(b,c)^t + an'\|}. \tag{2.30}$$

Due to $x \in b(f_\ell, 1/C_2)$, we obtain from Lemma 2.14 that

$$\|h(b,c)^t + an\| \geqslant \|h\|/C_2.$$

Lemma 2.18 gives that $x \in b(f_{\ell'}, 1/2C_2)$ and

$$\|h'(b,c)^t + an'\| \geqslant \|h'\|/2C_2 \geqslant \|h\|/4C_2,$$

For the numerator of Eq. (2.30), we have

$$\begin{aligned}
&\|(h(b,c)^t + an) \wedge (h'(b,c)^t + an')\| \\
=&\|(b,c)^t \wedge h^{-1}h'(b,c)^t + a^2 h^{-1}n \wedge h^{-1}n' + ah^{-1}n \wedge h^{-1}h'(b,c)^t - ah^{-1}n' \wedge (b,c)^t\| \\
\leqslant&\|(b,c)^t \wedge h^{-1}h'(b,c)^t\| + \|a^2 h^{-1}n \wedge h^{-1}n'\| + \|ah^{-1}n \wedge h^{-1}h'(b,c)^t - ah^{-1}n' \wedge (b,c)^t\|.
\end{aligned}$$

We use Lemma 2.15 to estimate the first two terms and get an upper bound $2C_L d(\ell, \ell')$. For the last term, we also use Lemma 2.15

$$\begin{aligned}
&\|ah^{-1}n \wedge h^{-1}h'(b,c)^t - ah^{-1}n \wedge (b,c)^t\| \\
&\leqslant \|h^{-1}n \wedge (h^{-1}h' - id)(b,c)^t + (h^{-1}n - h^{-1}n') \wedge (b,c)^t\| \\
&\leqslant C_1 C_L d(\ell', \ell) + C_L d(\ell', \ell).
\end{aligned}$$

Collecting all the terms, we obtain the lemma. □

Let $\ell = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix} \in L$. We are interested in the map

$$[\ell(\cdot)] = \pi_{E_1^\perp}\ell(\cdot) = h \circ \pi(\ell^{-1}E_1, E_1^\perp, (\ell^{-1}E_1)^\perp) \circ \pi_{(\ell^{-1}E_1)^\perp}(\cdot). \tag{2.31}$$

Now we are ready to explain how to linearize the map $[\ell(\cdot)]$ with $\ell \in L$. We will approximate it by a composition of translation and scaling map, which is a stronger estimate compared to the one in [BHR19].

Definition 2.20. For any $t \in \mathbb{R}$, we define the scaling map $S_t : \mathbb{R} \to \mathbb{R}$ by $S_t x = q^t x$.

We have the identifications $\mathbb{P}(V^\perp) \cong \mathbb{P}(\mathbb{R}^2)$ for $V \in \mathbb{P}(\mathbb{R}^3)$ in Lemma A.1 and $\iota : \mathbb{P}(\mathbb{R}^2) \cong \mathbb{R}/\mathbb{Z}$ (Eq. (2.3)). In the followings, for any $\star, \bullet \in \mathbb{P}(V^\perp)$, $d(\star, \bullet)$ means to view $\star, \bullet$ as elements in $\mathbb{P}(\mathbb{R}^2)$ and d is the distance given as in Definition 2.1; $\star - \bullet$ means to view $\star, \bullet$ as elements in $\mathbb{R}/\mathbb{Z}$ and $|\star - \bullet|$ is to compute their distance using the rotational invariant distance on $\mathbb{R}/\mathbb{Z}$. We have $d(\star, \bullet) = \sin(\pi|\star - \bullet|)$.

Lemma 2.21. *Fix any $C_1, C_2 > 20$. Let $\ell_0 = \begin{pmatrix} 1 & 0 \\ n_0 & h_0 \end{pmatrix} \in L$ be such that $\|h_0\| > 2$ and $d(\ell_0^{-1}E_1, E_1^\perp) > 1/C_1$. Then for any $\ell \in B(\ell_0, 1/C_L C_1 C_2) \subset L$, and x, x_0 in the $[\ell_0]$-attracting region $b(f_{\ell_0}, 1/C_2)$ with $d(x, x_0) \leqslant (1/C_1C_2)^{10}$, there exists a scaling map $S_{-t(\ell_0, x_0)} : \mathbb{R} \to \mathbb{R}$ with*

$$\left|t(\ell_0, x_0) - 2\log\|h_0\|\right| \leqslant 4\log(C_1C_2) \tag{2.32}$$

such that

$$\begin{aligned} &\left|[\ell(x)] - [\ell(x_0)] - S_{-t(\ell_0,x_0)}\left(\pi_{(\ell_0^{-1}E_1)^\perp} x - \pi_{(\ell_0^{-1}E_1)^\perp} x_0\right)\right| \\ \leqslant &C_L (C_2C_1)^9 (d(\ell, \ell_0) + d(x, x_0)) d(x, x_0)/\|h_0\|^2. \end{aligned} \tag{2.33}$$

Remark 2.22. Let us explain the notions in Eq. (2.33). Let ι' be the identification between $\mathbb{R}/\mathbb{Z}$ and $[-\frac{1}{2}, \frac{1}{2})$. Let $\pi_{\mathbb{R}/\mathbb{Z}} : \mathbb{R} \to \mathbb{R}/\mathbb{Z}$ be the projection map. It will be clear from the proof that $\left|\pi_{(\ell_0^{-1}E_1)^\perp} x - \pi_{(\ell_0^{-1}E_1)^\perp} x_0\right| < \frac{1}{2}$. By abusing the notation, we denote $\pi_{\mathbb{R}/\mathbb{Z}} \circ S_{-t(\ell_0,x_0)} \circ \iota'\left(\pi_{(\ell_0^{-1}E_1)^\perp} x - \pi_{(\ell_0^{-1}E_1)^\perp} x_0\right)$ by $S_{-t(\ell_0,x_0)}\left(\pi_{(\ell_0^{-1}E_1)^\perp} x - \pi_{(\ell_0^{-1}E_1)^\perp} x_0\right)$. Eq. (2.33) is to subtract elements in $\mathbb{R}/\mathbb{Z}$.

Proof. **Step 1:** Approximate $[\ell_0(x)] - [\ell_0(x_0)]$. We have

$$\begin{aligned} [\ell_0(x)] - [\ell_0(x_0)] =& h_0 \circ \pi(\ell_0^{-1}E_1, E_1^\perp, (\ell_0^{-1}E_1)^\perp) \circ \pi_{(\ell_0^{-1}E_1)^\perp} x \\ &- h_0 \circ \pi(\ell_0^{-1}E_1, E_1^\perp, (\ell_0^{-1}E_1)^\perp) \circ \pi_{(\ell_0^{-1}E_1)^\perp} x_0. \end{aligned}$$

We will use a basic inequality in analysis to approximate the difference gradually. For a C^2-map g on $\mathbb{R}$ and two points $z < y$ in $\mathbb{R}$, we have

$$|gz - gy - g'y(z - y)| \leqslant |z - y|^2 \sup\{|g''w| : z \leqslant w \leqslant y\}. \tag{2.34}$$

We want apply Eq. (2.34) to

$$\begin{aligned} g :=& h_0, \\ z :=& \pi(\ell_0^{-1}E_1, E_1^\perp, (\ell_0^{-1}E_1)^\perp) \circ \pi_{(\ell_0^{-1}E_1)^\perp}(x) = \pi_{\ell_0^{-1}E_1, E_1^\perp}(x), \\ y :=& \pi(\ell_0^{-1}E_1, E_1^\perp, (\ell_0^{-1}E_1)^\perp) \circ \pi_{(\ell_0^{-1}E_1)^\perp}(x_0) = \pi_{\ell_0^{-1}E_1, E_1^\perp}(x_0). \end{aligned}$$

Due to $x, x_0 \in b(f_{\ell_0}, 1/C_2)$, Eq. (2.20) in Lemma 2.14 implies that

$$y, z \in b(h_0^-, 1/C_1C_2). \tag{2.35}$$

h_0 acts on $b(h_0^-, 1/C_1C_2)$ by contraction and due to Lemma 2.8, we know

$$|\log h_0'(y) + 2\log\|h_0\|| \leqslant 2\log(C_1C_2). \tag{2.36}$$

Using Lemma 2.4 and Eq. (2.21), we obtain

$$\begin{aligned} d(z, y) =& \frac{d(h_0^{-1}[\ell_0(x_0)], h_0^{-1}[\ell_0(x)])}{d([\ell_0(x_0)], [\ell_0(x)])} \frac{d([\ell_0(x_0)], [\ell_0(x)])}{d(x_0, x)} d(x, x_0) \\ \leqslant& \|h_0\|^2 \frac{2C_1C_2^2}{\|h_0\|^2} d(x, x_0) = 2C_1C_2^2 d(x, x_0). \end{aligned} \tag{2.37}$$

We have $[z,y] \subset b(h^-, 1/C_1C_2)$Eq. (2.35), Eq. (2.37) and the hypothesis that $d(x,x_0) \leqslant 1/(C_1C_2)^{10}$. Hence, by locally identifying $\mathbb{P}(\ell_0^{-1}E_1)$ with $\mathbb{R}$, we can use Eq. (2.34) and the estimate for h_0'' (Lemma 2.8) to obtain

$$\begin{aligned}&\big|[\ell_0(x)] - [\ell_0(x_0)] - h_0'(y)(z-y)\big| \leqslant |z-y|^2 \sup\{|h_0''w| : z \leqslant w \leqslant y\}\\ \leqslant&|z-y|^2 10(C_1C_2)^3/\|h_0\|^2 \leqslant 40C_1^5C_2^7 d(x,x_0)^2/\|h_0\|^2.\end{aligned} \tag{2.38}$$

Then we let

$$\begin{aligned}\tilde{g} :=&\pi(\ell_0^{-1}E_1, E_1^\perp, (\ell_0^{-1}E_1)^\perp),\\ z' :=&\pi_{(\ell_0^{-1}E_1)^\perp}x,\\ y' :=&\pi_{(\ell_0^{-1}E_1)^\perp}x_0.\end{aligned}$$

We use Eq. (2.34) to approximate $\tilde{g}(z') - \tilde{g}(y')$. Due to Lemma B.1 and Lemma 2.7, we have for any $w \in \mathbb{P}(E_1^\perp)$,

$$|\tilde{g}''(w)| \leqslant 4/d(\ell_0^{-1}E_1, E_1^\perp) \leqslant 4C_1,\ 1/C_1 \leqslant |\tilde{g}'(w)| \leqslant C_1. \tag{2.39}$$

We apply Eq. (2.18) to x, x_0 and then Lemma B.2(2) yields

$$d(z',y') = d(\pi_{(\ell_0^{-1}E_1)^\perp}x, \pi_{(\ell_0^{-1}E_1)^\perp}x_0) \leqslant C_1C_2 d(x,x_0). \tag{2.40}$$

Therefore, we use Eq. (2.34) to obtain

$$|z - y - \tilde{g}'(y')(z'-y')| \leqslant |z'-y'|^2(4C_1) \leqslant 4C_1^3C_2^2 d(x,x_0)^2. \tag{2.41}$$

Combining with Eq. (2.38) and Eq. (2.36), we obtain

$$|[\ell_0(x)] - [\ell_0(x_0)] - S_{-t(\ell_0,x_0)}(\pi_{(\ell_0^{-1}E_1)^\perp}x - \pi_{(\ell_0^{-1}E_1)^\perp}x_0)| \leqslant 44(C_1C_2)^7 d(x,x_0)^2/\|h\|^2. \tag{2.42}$$

Here $S_{-t(\ell_0,x_0)} = h_0'(y)\tilde{g}'(y')$ is the contracting action with contracting rate of $h_0 \circ \tilde{g}$ at $y_0 := \pi_{(\ell_0^{-1}E_1)^\perp}x_0$, and we obtain the estimate of $t(\ell_0, x_0)$ (Eq. (2.32)) using Eq. (2.36) and (2.39).

Before proceeding to the general case, we would like to point out a corollary of Step 1: the sign of $[\ell_0(x)] - [\ell_0(x_0)]$ is the same with the sign of $z' - y'$ as elements in $[-\frac{1}{2}, \frac{1}{2})$. The reason is as follows. Note that $\tilde{g}$ scales $\mathbb{P}(\mathbb{R}^2)$ by 1 with distortion C_1 (Lemma 2.4). As $d(z', y')$ is small (Eq. (2.40)), we can deduce that the sign of $z - y$ is the same with that of $z' - y'$. As $[\ell_0(x)] = h_0(z)$, $[\ell_0(x_0)] = h_0(y)$, and h_0 contracts the minor arc $[z,y]$, we obtain the corollary.

Step 2: We need to distinguish two cases.

Case 1: Suppose $\ell^{-1}E_1$ and $\ell_0^{-1}E_1$ are in the same side of the arc x, x_0, see Fig. 4. We write $\ell_0 = \begin{pmatrix} 1 & 0 \\ n_0 & h_0 \end{pmatrix}$, $\ell = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix}$, $x_0 = \mathbb{R}(a_0, b_0, c_0)^t$, $x = \mathbb{R}(a,b,c)^t$. We have

$$\begin{aligned}d([\ell(x)], [\ell(x_0)]) - d([\ell_0(x)], [\ell_0(x_0)]) =&\frac{\|((b,c)^t \wedge (b_0,c_0)^t + h^{-1}n \wedge (a(b_0,c_0)^t - a_0(b,c)^t)\|}{\|h(b,c)^t + an\|\|h(b_0,c_0)^t + a_0 n\|}\\ &- \frac{\|((b,c)^t \wedge (b_0,c_0)^t + h_0^{-1}n_0 \wedge (a(b_0,c_0)^t - a_0(b,c)^t)\|}{\|h_0(b,c)^t + an_0\|\|h_0(b_0,c_0)^t + a_0 n_0\|}.\end{aligned}$$

The idea is to compare the difference term by term using the formula

$$\frac{A}{BC} - \frac{A_0}{B_0C_0} = \frac{AB_0C_0 - A_0BC}{BCB_0C_0} = \frac{(A-A_0)B_0C_0 + A_0(B_0-B)C_0 + A_0B(C_0-C)}{BCB_0C_0},$$

with

$$A = \|((b,c)^t \wedge (b_0,c_0)^t + h^{-1}n \wedge (a(b_0,c_0)^t - a_0(b,c)^t)\|,$$
$$B = \|h(b,c)^t + an\|, \;\; C = \|h(b_0,c_0)^t + a_0 n\|$$

and A_0, B_0, C_0 the corresponding one with h, n replaced by h_0, n_0. From Eq. (2.14), we obtain

$$A_0 \leqslant (1 + \|h_0^{-1}n_0\|)d(x,x_0) \leqslant 2d(x,x_0)/d(\ell_0^{-1}E_1, E_1^{\perp}) \leqslant 2C_1 d(x,x_0).$$

Similarly, we have $A_0 \leqslant 4C_1 d(x,x_0)$. From Lemma 2.18, we can apply Lemma 2.14 to obtain to all possible pairs between ℓ, ℓ_0 and x, x_0. Due to Eq. (2.19), we obtain

$$B, B_0, C, C_0 \geqslant \frac{1}{4C_2}\|h_0\|.$$

Then due to $\|h(b,c)^t + an\| \leqslant \|h\|(1 + \|h^{-1}n\|)$, using Eq. (2.14) we obtain

$$B, B_0, C, C_0 \leqslant 4C_1\|h_0\|.$$

For the differences, by triangle inequality, Eq. (2.27) and Eq. (2.25), we have

$$|A - A_0| \leqslant \|(h^{-1}n - h_0^{-1}n_0) \wedge (a(b_0,c_0)^t - a_0(b,c)^t)\| \leqslant \frac{C_L d(\ell,\ell_0) d(x,x_0)}{d(\ell_0^{-1}E_1, E_1^{\perp})} \leqslant C_L C_1 d(\ell,\ell_0) d(x,x_0).$$

Similarly, using triangle inequality and Lemma 2.15

$$|B - B_0|, |C - C_0| \leqslant \|h_0\|(\|id - h_0^{-1}h\| + \|h_0^{-1}n - h_0^{-1}n_0\|) \leqslant 2C_L\|h_0\| d(\ell,\ell_0).$$

In the end, we obtain from these computations that

$$|d([\ell(x)],[\ell(x_0)]) - d([\ell_0(x)],[\ell_0(x_0)])| \leqslant 2^{16} C_L (C_1C_2)^4 d(x,x_0) d(\ell,\ell_0)/\|h_0\|^2. \tag{2.43}$$

By the position of $\ell^{-1}E_1, \ell_0^{-1}E_1, x, x_0$ on $\mathbb{P}(\mathbb{R}^3)$ (see Fig. 4) and Eq. (2.40), $\pi_{\ell^{-1}E_1^{\perp}}x - \pi_{\ell^{-1}E_1^{\perp}}x_0$ and $\pi_{\ell_0^{-1}E_1^{\perp}}x - \pi_{\ell_0^{-1}E_1^{\perp}}x_0$ have the same sign as elements in $[-1/2, 1/2)$. The corollary of Step 1 holds for both $[\ell(x)] - [\ell(x_0)]$ and $[\ell_0(x)] - [\ell_0(x_0)]$. Then they have the same sign. We have the relation

$$|([\ell(x)] - [\ell(x_0)]) - ([\ell_0(x)] - [\ell_0(x_0)])| \leqslant \big||[\ell(x)] - [\ell(x_0)]| - |[\ell_0(x)] - [\ell_0(x_0)]|\big|. \tag{2.44}$$

Due to $d([\ell(x)],[\ell(x_0)]) \leqslant 1/2$ and $|a - b| \leqslant 2|\sin(a) - \sin(b)|$ for $a, b \in [0, \pi/3]$, combined with Eq. (2.43), we obtain

$$|([\ell(x)] - [\ell(x_0)]) - ([\ell_0(x)] - [\ell_0(x_0)])| \leqslant 2^{17} C_L (C_1C_2)^4 d(x,x_0) d(\ell,\ell_0)/\|h_0\|^2. \tag{2.45}$$

Combined with Eq. (2.42) for both ℓ and ℓ_0, we obtain the lemma due to C_1, C_2 large.

Case 2: Suppose $\ell^{-1}E_1$ and $\ell_0^{-1}E_1$ are in different sides of the arc x, x_0. By spherical trigonometry of the triangle $(x, \ell^{-1}E_1, \ell_0^{-1}E_1)$, we have

$$\sin \angle_x(\ell^{-1}E_1, \ell_0^{-1}E_1) = \frac{\sin \angle_{\ell^{-1}E_1}(x, \ell_0^{-1}E_1)}{d(x, \ell_0^{-1}E_1)} d(\ell^{-1}E_1, \ell_0^{-1}E_1) \leqslant \frac{d(\ell^{-1}E_1, \ell_0^{-1}E_1)}{d(x, \ell_0^{-1}E_1)}.$$

Due to Lemma 2.15 and Eq. (2.18), we have

$$\sin \angle_x(\ell^{-1}E_1, \ell_0^{-1}E_1) \leqslant \frac{C_L d(\ell,\ell_0)/d(\ell_0^{-1}E_1, {E_1}^{\perp})}{d(\ell_0^{-1}E_1, {E_1}^{\perp})/C_2} \leqslant C_L C_1^2 C_2 d(\ell,\ell_0). \tag{2.46}$$

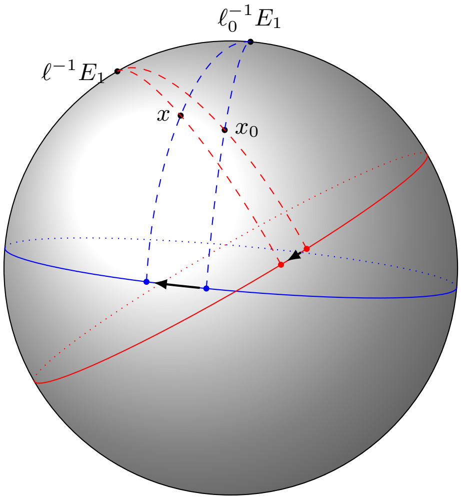

Figure 4: Case 1

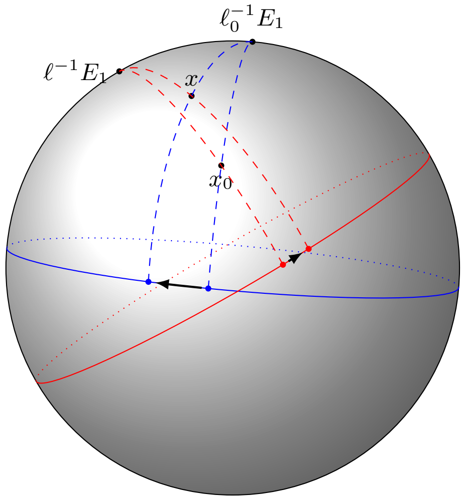

Figure 5: Case 2

Again by spherical trigonometry of the triangle $(x, x_0, \ell^{-1}E_1)$, we have

$$d(\pi_{(\ell^{-1}E_1)^\perp}x, \pi_{(\ell^{-1}E_1)^\perp}x_0) = \sin\angle_{\ell^{-1}E_1}(x, x_0) = d(x, x_0)\frac{\sin\angle_x(\ell^{-1}E_1, x_0)}{d(x_0, \ell^{-1}E_1)}$$

Due to the hypothesis that $\ell^{-1}E_1$ and $\ell_0^{-1}E_1$ are in different sides of the arc x, x_0, we obtain $\sin\angle_x(\ell^{-1}E_1, x_0) \leqslant \sin\angle_x(\ell^{-1}E_1, \ell_0^{-1}E_1)$. Combined with Eq. (2.46) and Eq. (2.18), we obtain

$$d(\pi_{(\ell^{-1}E_1)^\perp}x, \pi_{(\ell^{-1}E_1)^\perp}x_0) \leqslant d(x, x_0)\frac{\sin\angle_x(\ell^{-1}E_1, \ell_0^{-1}E_1)}{d(x_0, \ell^{-1}E_1)} \leqslant C_L C_1^3 C_2^2 d(\ell, \ell_0) d(x, x_0).$$

The same inequality also holds with ℓ replaced by ℓ_0. Combining with Eq. (2.42) for ℓ and the triangle inequality, using the fact that $\theta \leqslant 2\sin\theta$ for any $\theta \in [0, \pi/2]$, we obtain

$$\begin{aligned}
&\left|[\ell(x)] - [\ell(x_0)] - S_{-t(\ell_0,x_0)}\left(\pi_{(\ell_0^{-1}E_1)^\perp}x - \pi_{(\ell_0^{-1}E_1)^\perp}x_0\right)\right| \\
\leqslant &\left|[\ell(x)] - [\ell(x_0)] - S_{-t(\ell,x_0)}\left(\pi_{(\ell^{-1}E_1)^\perp}x - \pi_{(\ell^{-1}E_1)^\perp}x_0\right)\right| + 2q^{-t(\ell,x_0)} d(\pi_{(\ell^{-1}E_1)^\perp}x, \pi_{(\ell^{-1}E_1)^\perp}x_0) \\
+&2q^{-t(\ell_0,x_0)} d(\pi_{(\ell_0^{-1}E_1)^\perp}x, \pi_{(\ell_0^{-1}E_1)^\perp}x_0) \\
\leqslant &44(C_1C_2)^7 d(x, x_0)^2/\|h\|^2 + (q^{-t(\ell,x_0)} + q^{-t(\ell_0,x_0)}) C_L C_1^3 C_2^2 d(\ell, \ell_0) d(x, x_0) \\
\leqslant &C_L (C_1C_2)^7 (2d(\ell, \ell_0) + 88 d(x, x_0)) d(x, x_0)/\|h_0\|^2.
\end{aligned}$$

□

2.5 q-Adic partitions

The q-adic level-n partition for $n \in \mathbb{N}$ of $[0, 1)$ is defined by

$$\mathcal{Q}_n = \left\{ [\frac{k}{q^n}, \frac{k+1}{q^n}) : 0 \leqslant k < q^n,\ k \in \mathbb{Z} \right\}.$$

Using the canonical identification between $[0, 1)$ and the circle $\mathbb{P}(\mathbb{R}^2)$, we can regard $\mathcal{Q}_n$ as a partition of $\mathbb{P}(\mathbb{R}^2)$. We write $\mathcal{Q}_t = \mathcal{Q}_{[t]}$ when $t \in \mathbb{R}^+$ is not an integer.

We also need similar partitions of the group L defined in Section 2.3. By Theorem 2.1 of [KRS12], for q large enough, there exists a collection of Borel sets $\{Q_{n,i} \subset L : n \in \mathbb{N}, i \in \mathbb{N}\}$, having the following properties:

1. $L = \cup_{i \in \mathbb{N}} Q_{n,i}$ for every $n \in \mathbb{N}$;

2. $Q_{n,i} \cap Q_{m,j} = \emptyset$ or $Q_{n,i} \subset Q_{m,j}$ for $n, m \in \mathbb{N}$ with $n \geqslant m$ and any $i, j \in \mathbb{N}$;
3. There exists a constant $C_p > 1$ (independent of large q) such that for every $n \in \mathbb{N}$ and $i \in \mathbb{N}$ there exists $\ell \in Q_{n,i}$ with

$$B(\ell, C_p^{-1} q^{-n}) \subset Q_{n,i} \subset B(\ell, q^{-n}), \tag{2.47}$$

where $B(\ell, q^{-n})$ is measured using the left L-invariant metric on L.

For each $n \in \mathbb{N}$, denote by $\mathcal{Q}_n^L$ the partition $\{Q_{n,i} : i \in \mathbb{N}\}$ of L. We will usually omit the superscript when there is no ambiguity about which space to partition. Fix a left-invariant Haar measure on L. In view of (2.47), there exists a constant $C > 1$ independent of n such that

$$1/C \leqslant q^{5n} \times \text{ measure of any atom in } \mathcal{Q}_n^L \leqslant C.$$

Hence there exists a constant $C' > 0$ such that for every n and any atom $Q \in \mathcal{Q}_n^L$,

$$\#\{Q' \in \mathcal{Q}_{n+1}^L : Q' \subset Q\} \leqslant C'.$$

2.6 Component measures and random measures

For a probability measure θ on a measure space X and a measurable subset Y of X, define the normalized restriction measure by

$$\theta_Y = \begin{cases} \frac{1}{\theta(Y)} \theta|_Y, & \text{if } \theta(Y) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

For a probability measure θ on $\mathbb{P}(\mathbb{R}^2)$ (or L) equipped with a partition a partition $\mathcal{Q}_n$, we write $\mathcal{Q}_n(x)$ for the unique partition element containing x. We define measure valued random variables $\theta_{x,n}$ such that $\theta_{x,n} = \theta_{\mathcal{Q}_n(x)}$ with probability $\theta(\mathcal{Q}_n(x))$. For integers $n_2 \geqslant n_1$ and an event $\mathcal{U}$, we write

$$\mathbb{P}_{n_1 \leqslant i \leqslant n_2}(\theta_{x,i} \in \mathcal{U}) = \frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} \mathbb{P}(\theta_{x,n} \in \mathcal{U}).$$

2.7 Regularity of stationary measures

Let Λ be a finite set and let $\nu = \sum_{i \in \Lambda} p_i \delta_{g_i}$ be a finitely supported probability measure on $\mathrm{SL}_3(\mathbb{R})$ such that the group generated by the support of ν, $\langle \operatorname{supp} \nu \rangle$, is Zariski dense in $\mathrm{SL}_3(\mathbb{R})$. We write ν^{*n} for the convolution, that is, ν^{*n} is the law of the product $g_1 \cdots g_n$, with i.i.d. g_i following the law of ν.

Let μ be the unique associated stationary measure on $\mathbb{P}(\mathbb{R}^3)$ ([Fur73]), that is to say, we have

$$\mu = \nu * \mu := \sum_{i \in \Lambda} p_i (g_i)_* \mu,$$

where $(g_i)_*\mu$ is the pushforward measure defined by $(g_i)_*\mu(E) = \mu(g_i^{-1}E)$ for any measurable set $E \subset \mathbb{P}(\mathbb{R}^3)$.

There are 3 *Lyapunov exponents* $\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3$ of ν such that for any $i = 1, 2, 3$,

$$\lambda_i = \lim_{n \to \infty} \int \frac{1}{n} \log \sigma_i(g_{i_1} \cdots g_{i_n}) d\nu(g_{i_1}) \cdots d\nu(g_{i_n}).$$

Since $\langle \operatorname{supp} \nu \rangle$ is Zariski dense, we have $\lambda_1 > \lambda_2 > \lambda_3$ (see for example [GR85], [GM89]). Set $\chi_i = \lambda_1 - \lambda_{i+1}$ for $i = 1, 2$, and $\lambda(\nu) = (\lambda_1, \lambda_2, \lambda_3)$ to be the Lyapunov vector.

We will also consider the measure ν^- on $\mathrm{SL}_3(\mathbb{R})$ defined by $\nu^-(g) = \nu(g^{-1})$. Let μ^- be the unique associated stationary measure.

The results stated in this section and Section 2.8 hold under a more general assumption that $\operatorname{supp}\nu$ generates a Zariski dense subgroup and that ν is of finite exponential moment. But for the interest of the current paper, we will focus on the case when ν is finitely supported.

Classical references of products of random matrices include [Fur63] and [BL85]. We refer the reader to [BQ16]. We recall an equidistribution result for μ, μ^- ([GR85], [BL85, Chapter III, Theorem 4.3])

Lemma 2.23. *For any $x \in \mathbb{P}(\mathbb{R}^3)$, as $n \to \infty$ we have*

$$\nu^{*n} * \delta_x \to \mu, \ (\nu^-)^{*n} * \delta_x \to \mu^-,$$

and the convergence is uniform with respect to x.

We have the Guivarc'h regularity of stationary measures ([Gui90], [BQ16, Prop 14.1])

Lemma 2.24. *There exist $C > 0, \beta > 0$ such that for any $n \geqslant 1$ and hyperplane W in $\mathbb{P}(\mathbb{R}^3)$*

$$\mu\{x \in \mathbb{P}(\mathbb{R}^3) : \ d(W, x) \leqslant q^{-n}\} \leqslant Cq^{-\beta n}.$$

In particular, for any $V \in \mathbb{P}(\mathbb{R}^3)$ and $x \in \mathbb{P}(\mathbb{R}^2)$

$$(\pi_{V^\perp}\mu)\left(B(x, q^{-n})\right) \leqslant \mu\{y : d((x, V), y) \leqslant q^{-n}\} \leqslant Cq^{-\beta n},$$

where (x, V) is the projective plane spanned by x and V.

As a corollary, we have the following result.

Lemma 2.25. *Let $C_1 > 1$. For any $g \in \mathrm{SL}_3(\mathbb{R})$ and $V \in \mathbb{P}(\mathbb{R}^3)$ satisfy $d(g^{-1}V, V^\perp) > 1/C_1$, we have for any for any $x \in \mathbb{P}(V^\perp)$ and $r > 0$,*

$$(\pi_{g^{-1}V, V^\perp}\mu)(B(x, r)) \leqslant CC_1^\beta r^\beta.$$

where C, β are defined in Lemma 2.24.

Proof. It follows from the proof of Lemma 2.10 that

$$\pi_{g^{-1}V, V^\perp} = \pi(g^{-1}V, V^\perp, (g^{-1}V)^\perp) \circ \pi_{(g^{-1}V)^\perp}.$$

Lemma 2.24 holds for $\pi_{(g^{-1}V)^\perp}\mu$. Moreover, Lemma B.1 states that $\pi(g^{-1}V, V^\perp, (g^{-1}V)^\perp)$ scales by 1 with distortion C_1. These imply the lemma. $\square$

More generally, the Hölder regularity of stationary measures also holds for irreducible representations ([Gui90], [BQ16, Proposition 10.1]). Let p_2 be the map from $\mathbb{P}(\mathbb{R}^3)$ to the projective space of the second symmetric power $\mathbb{P}(Sym^2(\mathbb{R}^3))$, defined by sending $\mathbb{R}v$ to $\mathbb{R}v \otimes v$. It is $\mathrm{SL}_3(\mathbb{R})$-equivariant. Let $(p_2)_*\mu$ be the pushfoward measure on $\mathbb{P}(Sym^2(\mathbb{R}^3))$. Then $(p_2)_*\mu$ is ν-stationary and we have

Lemma 2.26. *There exist $C > 0, \beta > 0$ such that for any $n \geqslant 1$ and hyperplane W in $\mathbb{P}(Sym^2(\mathbb{R}^3))$, we have*

$$(p_2)_*\mu\{x \in \mathbb{P}(Sym^2\mathbb{R}^3) : \ d(W, x) \leqslant q^{-n}\} \leqslant Cq^{-\beta n}.$$

2.8 Large deviation estimates

We first introduce several LDP (Large Deviation Principle) estimates for random walks on $\mathrm{SL}_3(\mathbb{R})$. Set $G = \mathrm{SL}_3(\mathbb{R})$ in this subsection. The following lemma can be found in [LP82] and Proposition 14.3, Theorem 13.17 (iii), Theorem 13.11 (iii) in [BQ16]. Recall that ν is finitely supported and Zariski dense and μ is the ν-stationary measure on $\mathbb{P}(\mathbb{R}^3)$.

Lemma 2.27. *For any $c > 0$, there exist $C > 0$ and $\epsilon > 0$ such that for any $n \in \mathbb{N}$, we have for any hyperplane W in $\mathbb{P}(\mathbb{R}^3)$ and any $V = \mathbb{R}v \in \mathbb{P}(\mathbb{R}^3)$ with $\|v\| = 1$,*

$$\nu^{*n}\{g \in G : \ d(gV, W) \leqslant q^{-cn}\} \leqslant Cq^{-\epsilon n}, \tag{2.48}$$
$$\nu^{*n}\{g \in G : \ d(V_g^+, W) \leqslant q^{-cn}\} \leqslant Cq^{-\epsilon n}, \tag{2.49}$$
$$\nu^{*n}\{g \in G : \ d(H_g^-, V) \leqslant q^{-cn}\} \leqslant Cq^{-\epsilon n}, \tag{2.50}$$
$$\nu^{*n}\{g \in G : \ d(V_g^+, gV) \geqslant q^{-(\chi_1 - c)n}\} \leqslant Cq^{-\epsilon n}, \tag{2.51}$$
$$\nu^{*n}\{g \in G : \ \|\frac{1}{n}\kappa(g) - \lambda(\nu)\| \geqslant c\} \leqslant Cq^{-\epsilon n}, \tag{2.52}$$
$$\nu^{*n}\{g \in G : \ \frac{1}{n}|\log\|gv\| - \log\sigma_1(g)| \geqslant c\} \leqslant Cq^{-\epsilon n}. \tag{2.53}$$

As a corollary, we have the following lemma as in [BQ16, Lemma 14.11].

Lemma 2.28. *For any $c > 0$, there exist $C > 0$ and $\epsilon > 0$ such that for any $m \leqslant n$, we have for any hyperplane W in $\mathbb{P}(\mathbb{R}^3)$ and any $V \in \mathbb{P}(\mathbb{R}^3)$,*

$$\nu^{*n}\{g \in G : \ d(gV, W) \leqslant q^{-cm}\} \leqslant Cq^{-\epsilon m}, \tag{2.54}$$
$$\nu^{*n}\{g \in G : \ d(V_g^+, W) \leqslant q^{-cm}\} \leqslant Cq^{-\epsilon m}, \tag{2.55}$$
$$\nu^{*n}\{g \in G : \ d(V, H_g^-) \leqslant q^{-cm}\} \leqslant Cq^{-\epsilon m}. \tag{2.56}$$

Proof. For Eq. (2.54), when $m = n$, this is Eq. (2.48). Since

$$\nu^{*n}\{g \in G : d(gV, W) < q^{-cm}\} = \int_G \nu^{*m}\{g \in G : d(ghV, W) < q^{-cm}\} d\nu^{*(n-m)}(h),$$

then case $n \geqslant m$ follows.

For Eq. (2.55), we use triangle inequality

$$d(V_g^+, W) \geqslant d(gV, W) - d(V_g^+, gV).$$

Therefore

$$\begin{aligned} &\nu^{*n}\{g \in G : \ d(V_g^+, W) < q^{-cm}\} \\ \leqslant &\nu^{*n}\{g \in G : \ d(gV, W) < 2q^{-cm}\} + \nu^{*n}\{g \in G : \ d(V_g^+, gV) > q^{-cm}\}. \end{aligned}$$

We may assume $c < \chi_1/2$. Then $q^{-cm} \leqslant q^{-(\chi_1 - c)n}$ and the second term above can be estimated using Eq. (2.51). The first term can be estimated using Eq. (2.54). Hence we obtain Eq. (2.55).

For Eq. (2.56), due to $H_g^- = (V_{(g^t)}^+)^\perp$, it is equivalent to prove

$$\nu^{*n}\{g \in G : \ d(V_{(g^t)}^+, W) < q^{-cm}\} \leqslant Cq^{-\epsilon m}, \tag{2.57}$$

with $W = V^\perp$, which is similar to Eq. (2.55). Let ν^t be the probability measure on G defined by $\nu^t(g) = \nu(g^t)$. It is finitely generated and Zariski dense. Eq. (2.57) holds as we have Eq. (2.55) for ν^t. ☐

Next we use LDP of $\mathrm{SL}_3(\mathbb{R})$ to obtain LDP for the $\mathrm{SL}_2(\mathbb{R})$-part in the UL-decomposition. It will allow us to reduce the $\mathrm{SL}_3(\mathbb{R})$-action on $\mathbb{P}(\mathbb{R}^3)$ to the induced $\mathrm{SL}_2(\mathbb{R})$-action on $\mathbb{P}(\mathbb{R}^2)$ and obtain the convergence of entropy for the induced action.

Recall that we have the projection map $\pi_L : UL \to L$ (see Eq. (2.6)). In the following lemma, it is important that it holds for all $n \gg \log C$, which will be very useful in proving porosity in Section 4.

Lemma 2.29. *There exists $\beta > 0$ such that the following holds. For any $C > 1$, there exists $N_C = O(\log C) > 0$ such that we have for any $n > N_C$ and $V \in \mathbb{P}(\mathbb{R}^3)$*

$$\nu^{*n}\{g \in G : \ d(g^{-1}V, V^\perp) \leqslant 1/C\} \leqslant C^{-\beta}, \tag{2.58}$$

$$\nu^{*n}\{g \in G : \ |\chi_1(h) - \chi_1(g)| \geqslant \log C\} \leqslant C^{-\beta}, \tag{2.59}$$

where $h \in \mathrm{SL}_2(\mathbb{R})$ comes from the UL-decomposition of g: $\pi_L(g) = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix}$ with $n \in \mathbb{R}^2$.

Proof. Applying Eq. (2.54) to the measure ν^- and $W = V^\perp$, we obtain Eq. (2.58) .

For any $g \in UL$, write $g = \begin{pmatrix} \lambda^{-2} & x & y \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix}$. We define $\tilde{h} := \Pi_{E_1^\perp} g|_{E_1^\perp} = \lambda h \in \mathrm{GL}(E_1^\perp)$. We have

$$q^{\chi_1(h)} = \sup_{v \in E_1^\perp, \|v\|=1} \frac{\|\tilde{h}v\|^2}{\det \tilde{h}}.$$

The determinant is given by

$$\det \tilde{h} = \lambda^2 = |\langle g^{-1}e_1, e_1\rangle| = d(g^{-1}E_1, E_1^\perp)\|g^{-1}e_1\|.$$

The norm of $\|\tilde{h}v\|$ for $v \in E_1^\perp$ is given by

$$\|\tilde{h}v\| = \|\Pi_{E_1^\perp} gv\| = \|(gv) \wedge e_1\| = d(g\mathbb{R}v, E_1)\|gv\|. \tag{2.60}$$

Therefore, $\exists v \in E_1^\perp$ with $\|v\| = 1$ such that

$$q^{-\chi_1(h)} \frac{(d(g\mathbb{R}v, E_1)\|gv\|)^2}{d(g^{-1}E_1, E_1^\perp)\|g^{-1}e_1\|} \in (\frac{1}{2}, 1).$$

It follows from Lemma 2.5 that

$$d(\mathbb{R}v, H_g^-) \leqslant \frac{\|gv\|}{\|g\|} \leqslant 1, \ \ d(E_1, H_{g^{-1}}^-) \leqslant \frac{\|g^{-1}e_1\|}{\|g^{-1}\|} \leqslant 1. \tag{2.61}$$

We also have $\|g\|^2/\|g^{-1}\| = \sigma_1(g)^2/\sigma_1(g^{-1}) = \sigma_1(g)^2\sigma_3(g) = \sigma_1(g)/\sigma_2(g) = q^{\chi_1(g)}$.

Consider the elements $g \in G$ satisfying all the following inequalities:

(i) $d(\mathbb{R}v, H_g^-) \geqslant q^{-C/8}$;

(ii) $d(E_1, H_{g^{-1}}^-) \geqslant q^{-C/8}$;

(iii) $d(g^{-1}E_1, E_1^\perp) \geqslant q^{-3C/4}$;

(iv) $d(g\mathbb{R}v, E_1) \geqslant q^{-3C/8}$.

It follows from Eq. (2.54) and Eq. (2.56). Then ν^{*n}-measure of such g is greater than $1-C^{-\beta}$ for some $\beta > 0$. Using Eq. (2.61) and (i), (ii) and (iii), we have $q^{\chi_1(h)-\chi_1(g)} \leqslant q^C$; using Eq. (2.61) and (i), (ii) and (iv), we have $q^{\chi_1(g)-\chi_1(h)} \leqslant q^C$. This finishes the proof of Eq. (2.59). ☐

We need more LDP estimates of h.

Lemma 2.30. *For any $c > 0$, there exist $C > 0$ and $\epsilon > 0$ such that for any $V \in \mathbb{P}(\mathbb{R}^3)$, $x \in \mathbb{P}(V^\perp)$ and $n \in \mathbb{N}$, we have*

$$\nu^{*n}\{g \in G : \ \frac{1}{n}|\chi_1(h) - \chi_1 n| \geqslant c\} \leqslant Cq^{-\epsilon n}, \tag{2.62}$$

$$\nu^{*n}\{g \in G : \ d(H_h^-, x) \leqslant q^{-cn}\} \leqslant Cq^{-\epsilon n}, \tag{2.63}$$

$$\nu^{*n}\{g \in G : \ d(V_h^+, hx) \geqslant q^{(-\chi_1+c)n}\} \leqslant Cq^{-\epsilon n}, \tag{2.64}$$

where $h \in \mathrm{SL}_2(\mathbb{R})$ comes from the UL-decomposition of g: $\pi_L(g) = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix}$ with $n \in \mathbb{R}^2$.

Proof. For Eq. (2.62), we compare $\chi_1(h)$ with $\chi_1(g)$ using Eq. (2.59). Then use Eq. (2.52) for the LDP of $\chi_1(g)$.

For any $g \in UL$, write $g = \begin{pmatrix} \lambda^{-2} & x & y \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix}$. We define $\tilde{h} := \Pi_{E_1^\perp} g|_{E_1^\perp} = \lambda h \in \mathrm{GL}(E_1^\perp)$. If $x = E_1$, then it is immediate to have Eq. (2.63) as $H_h^- \in E_1^\perp$. Otherwise, write $x = \mathbb{R}w$ with $w' := \Pi_{E_1^\perp}(w)$ of norm 1. Due to Lemma 2.5, that is

$$d(H_h^-, x) \geqslant \frac{1}{2} d(H_h^-, \mathbb{R}w') \geqslant \frac{1}{2}\left(\frac{\|\tilde{h}w'\|}{\|\tilde{h}\|} - q^{-\chi_1(h)}\right).$$

To obtain Eq. (2.63), it suffices to estimate the ν^{*n}-measure of g such that

(i) either its corresponding $\tilde{h}$ satisfies $\|\tilde{h}w'\| \leqslant 4q^{-cn}\|\tilde{h}\|$,

(ii) or $\chi_1(h) \leqslant 2cn$.

We can bound the ν^{*n}-measure of g satisfying (ii) by applying Eq. (2.62) to another sufficiently small constant c. For (i), by Eq. (2.60), we have for any $v \in E_1^\perp$,

$$\|\tilde{h}v\| = d(g\mathbb{R}v, E_1)\|gv\|.$$

Hence

$$\|\tilde{h}\| = \sup_{v \in E_1^\perp, \|v\|=1} \|\tilde{h}v\| = \sup_{v \in E_1^\perp, \|v\|=1} d(g\mathbb{R}v, E_1)\|gv\| \leqslant \sigma_1(g).$$

So we have

$$d(g\mathbb{R}w', E_1)\frac{\|gw'\|}{\sigma_1(g)} \leqslant \frac{\|\tilde{h}w'\|}{\|\tilde{h}\|}.$$

We use Eq. (2.48) and Eq. (2.53) to estimate the first term and the second term on the left respectively. This yields the estimate of the ν^{*n}-measure of g satisfying (i), and hence we obtain Eq. (2.63).

Using Lemma 2.5, we have

$$d(V_h^+, hx) \leqslant \frac{q^{-\chi_1(h)}}{d(x, H_h^-)}.$$

Then Eq. (2.64) follows from Eq. (2.62) and Eq. (2.63). $\square$

At the end of the subsection, we introduce two kinds of good subsets in L:

- for any $n \in \mathbb{N}$ and $c > 0$, set

$$L(n, c) := \{\ell \in L : \ d(\ell^{-1}E_1, E_1^\perp) \geqslant q^{-cn}, \ \chi_1(h) \in [\chi_1 - c, \chi_1 + c]n\}; \tag{2.65}$$

- for any $t \geqslant 1$ and $C_1 > 0$, set

$$L(t, C_1) := \{\ell \in L : \ d(\ell^{-1}E_1, E_1^{\perp}) > 1/C_1, \ \|h\| \in q^{t/2}[1/C_p, C_p]\}, \tag{2.66}$$

where C_p is defined as in Eq. (2.47), and h is the $\mathrm{SL}_2(\mathbb{R})$ matrix of ℓ: we have $\ell = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix}$ with $n \in \mathbb{R}^2$. For $\ell \in L(t, C_1)$, we have $\|h\| \geqslant 2$. Moreover, it follows from Eq. (2.58) and Eq. (2.62) that

$$(\pi_L \nu^{*n})(L(n, c)^c) \leqslant Cq^{-\epsilon(c)n}. \tag{2.67}$$

2.9 Coding, partitions of symbolic space and random words

We consider the general coding scheme of the random walk. Consider the symbolic space $\Lambda^{\mathbb{N}}$, which is the set of one-sided infinite words over Λ. Let Λ^* be the set of finite words over Λ and for each $n \in \mathbb{N}$, Λ^n be the set of words of length n. For a word $\mathbf{i} = (i_1, ..., i_n) \in \Lambda^*$, we define $g_{\mathbf{i}} := g_{i_1} \circ \cdots \circ g_{i_n}$. Similarly, define $p_{\mathbf{i}} := p_{i_1} \cdots p_{i_n}$.

The space $\Lambda^{\mathbb{N}}$ comes with the natural partitions into level-n cylinder sets. But it will be more convenient to consider general partitions into cylinders of varying lengths.

The partition Ψ_n^q and the definition of q. We consider a collection of words

$$\Psi_n^q = \{(i_1, \cdots, i_m) \in \Lambda^* : \chi_1(g_{i_1 \ldots i_m}) \geqslant n > \chi_1(g_{i_1 \ldots i_j}) \ \text{ for all } j < m\}.$$

Different than the setting in [BHR19], the exponent χ_1 here is not monotone with respect to the length of the word m. However, note that the measure ν has compact support, and the Cartan projection is subadditive (see for example [Kas08, Lemma 2.3]). There exists a constant $c_0 > 0$, depending only on the support of ν, such that for every $n \in \mathbb{N}$ and for every $\mathbf{i} \in \Psi_n^q$,

$$c_0 q^{-n} \leqslant q^{-\chi_1(g_{\mathbf{i}})} \leqslant q^{-n}. \tag{2.68}$$

If we take $q > 1/c_0$, then for any $n \neq k$, we have

$$\Psi_n^q \cap \Psi_k^q = \emptyset. \tag{2.69}$$

By the definition of Ψ_n^q, no word in is a prefix of another word. And the law of large numbers holds for ν([Fur63]): we have for $\nu^{\otimes \mathbb{N}}$-a.e. $\mathbf{i} \in \Lambda^{\mathbb{N}}$,

$$\frac{\chi_1(g_{i_1 \ldots i_m})}{m} \to \chi_1 > 0. \tag{2.70}$$

Therefore, the set of $\mathbf{i} \in \Lambda^{\mathbb{N}}$ such that $\chi_1(g_{i_1 \ldots i_m}) < n$ for all m is of measure zero. Consider the map $\Psi_n^q : \Lambda^{\mathbb{N}} \to \Lambda^*$ sending $\mathbf{i}$ to its first $(m+1)$ coordinates with m the least nonnegative integer such that $\chi_1(g_{i_1 \ldots i_m}) \geqslant n$. It is defined $\nu^{\otimes \mathbb{N}}$- almost everywhere and it gives a measurable partition of $\Lambda^{\mathbb{N}}$.

Definition 2.31. From now on, we fix the choice of an integer q with $q > \max\{1/c_0, C_L, 4C_p^2\}$ (see Eq. (2.29) and Eq. (2.47)), and the logarithm is in base q, $\log = \log_q$.

Definition 2.32 (Random words $\mathbf{U}(n)$, $\mathbf{I}(n)$)**.** For any given $n \in \mathbb{N}$, we introduce the random words $\mathbf{I}(n)$ and $\mathbf{U}(n)$ taking values from subsets coming from measurable partitions of Λ^*:

- $\mathbf{I}(n)$ is the random word taking values in $\Psi_n (= \Psi_n^q)$ according to the probability $\nu^{\otimes \mathbb{N}}$. i.e.

$$\mathbb{P}(\mathbf{I}(n) = \mathbf{i}) = \begin{cases} p_{\mathbf{i}}, & \text{if } \mathbf{i} \in \Psi_n \\ 0, & \text{otherwise.} \end{cases}$$

- $\mathbf{U}(n)$ is the random word taking values from Λ^n according to the probability $\nu^{\otimes n}$.

For each $n \in \mathbb{N}$ and for any measurable set $\mathcal{U} \subset G$, we have

$$\nu^{*n}\{g \in G : g \in \mathcal{U}\} = \mathbb{P}_{i=n}\{\mathbf{U}(i) : g_{\mathbf{U}(i)} \in \mathcal{U}\}. \tag{2.71}$$

Due to the stationarity of the measure μ and that $\mathbf{I}(n)$ is defined by stopping time, using the martingale property, we have for any $n \in \mathbb{N}$

$$\mu = \mathbb{E}_{j=n}(g_{\mathbf{I}(j)}\mu). \tag{2.72}$$

The following lemma is the relation between $\mathbf{U}(i)$ and $\mathbf{I}(i)$.

Lemma 2.33. *There exist $C_I > 1$ and $\beta > 0$ such that the following holds. Let $\mathcal{U} \subset \Lambda^*$ be a set of words. Suppose that for some $\epsilon > 0$ and $n(\epsilon) \in \mathbb{N}$, we have for every $n \geqslant n(\epsilon)$,*

$$\mathbb{P}_{1 \leqslant i \leqslant n}\{\mathbf{U}(i) \in \mathcal{U}\} > 1 - \epsilon.$$

Then for every $n \geqslant n(\epsilon)$,

$$\mathbb{P}_{1 \leqslant i \leqslant n}\{\mathbf{I}(i) \in \mathcal{U}\} > 1 - C_I\epsilon/2 - O(q^{-\beta n}).$$

Proof. The subtlety of the proof lies in that $\chi_1(g_{i_1 \dots i_m})$ is not monotone with respect to the length of the word m.

For any $n \geqslant n(\epsilon)$, we want to bound

$$\mathbb{P}_{1 \leqslant i \leqslant n}\{\mathbf{I}(i) \in \mathcal{U}^c\}.$$

Set $C' > 2(1 - \log c_0)/\chi_1$ and $\mathcal{B} := \{\mathbf{i} \in \Psi_i : 1 \leqslant i \leqslant n\}$. We have $\Psi_i \cap \Psi_j = \emptyset$ for any $1 \leqslant i < j \leqslant C'n$ (Eq. (2.69)).

We consider two cases. For any $\mathbf{i} \in \mathcal{B}$ with the word length $l(\mathbf{i}) \leqslant C'n$, by definition, $\mathbf{i} \in \mathbf{U}(j)$ for some $j \leqslant C'n$. So we have

$$\mathbb{P}_{1 \leqslant i \leqslant n}(\mathbf{I}(i) : \mathbf{I}(i) \in \mathcal{U}^c : l(\mathbf{I}(i)) \leqslant C'n) \leqslant C'\mathbb{P}_{1 \leqslant i \leqslant C'n}(\mathbf{U}(i) : \mathbf{U}(i) \in \mathcal{U}^c) \leqslant C'\epsilon. \tag{2.73}$$

Consider any $\mathbf{i} \in \mathcal{B}$ with $l(\mathbf{i}) > C'n$. Take $\epsilon_1 = \min\{\epsilon, \chi_1/2\}$. It follows from Eq. (2.68) that

$$\chi_1(g_{\mathbf{i}}) \leqslant n - \log c_0 \leqslant C'n(\chi_1 - \epsilon_1) \leqslant l(\mathbf{i})(\chi_1 - \epsilon_1).$$

As $\mathbf{i} \in U(j)$ for some $j > C'n$, the large deviation estimate Eq. (2.52) yields

$$\begin{aligned}\mathbb{P}_{1 \leqslant i \leqslant n}\{\mathbf{I}(i) : l(\mathbf{I}(i)) > C'n\} \leqslant & \frac{1}{n}\sum_{j > C'n} \mathbb{P}\{\mathbf{U}(j) : \chi_1(g_{\mathbf{U}(j)}) \leqslant l(\mathbf{U}(j))(\chi_1 - \epsilon_1)\} \\ \leqslant & \frac{1}{n}\sum_{j \geqslant C'n} Cq^{-\beta(\epsilon_1)j}.\end{aligned} \tag{2.74}$$

Combining Eqs. (2.73) and (2.74), we finish the proof of the lemma. □

2.10 Entropy

For a probability measure μ on a metric space X and any countable measurable partitions $\mathcal{E}, \mathcal{E}'$ of X, define the entropy and conditional entropy by

$$\begin{aligned}H(\mu, \mathcal{E}) :=& \sum_{E \in \mathcal{E}} -\mu(E)\log\mu(E), \\ H(\mu, \mathcal{E}'|\mathcal{E}) :=& H(\mu, \mathcal{E}' \vee \mathcal{E}) - H(\mu, \mathcal{E}),\end{aligned}$$

where $\mathcal{E}' \vee \mathcal{E}$ denotes the common refinement of $\mathcal{E}'$ and $\mathcal{E}$.

We collect some basic properties of entropy.

Lemma 2.34. *Let $\mathcal{E}$ and $\mathcal{F}$ be two measurable partitions of X. If each atom in E intersects at most k atoms in $\mathcal{F}$, then*

$$H(\mu, \mathcal{F}) \leqslant H(\mu, \mathcal{E}) + O(\log k).$$

Suppose $X = \mathbb{P}(\mathbb{R}^2)$ or L and $\mathcal{Q}_n$ the q-adic partition on X. This lemma implies that for any $m, n \in \mathbb{N}$,

$$|H(\mu, \mathcal{Q}_n) - H(\mu, \mathcal{Q}_m)| = O(|m - n|). \tag{2.75}$$

Let f and g be two continuous maps on X, such that $\|f - g\|_\infty \leqslant q^{-n}$. Denote by $f\mu$ and $g\mu$ the pushforward measures. Then we have

$$|H(f\mu, \mathcal{Q}_n) - H(g\mu, \mathcal{Q}_n)| = O(1) \tag{2.76}$$

(see [BHR19, (2.14)]).

Lemma 2.35 (Lemma 4.3 in [HS17])**.** *Suppose f scales* $\operatorname{supp} \mu$ *by $u > 0$ with distortion $C > 1$ (see Definition 2.6 for definition). Then*

$$H(f\mu, \mathcal{Q}_{n - \log u}) = H(\mu, \mathcal{Q}_n) + O(\log C).$$

The following is about the concavity of entropy. Let $\delta \in [0, 1]$, and μ_1, μ_2 be probability measures on X and $\mathcal{Q}$ be a measurable partition. We have

$$\begin{aligned} \delta H(\mu_1, \mathcal{Q}) + (1-\delta) H(\mu_2, \mathcal{Q}) &\leqslant H(\delta\mu_1 + (1-\delta)\mu_2, \mathcal{Q}) \\ &\leqslant \delta H(\mu_1, \mathcal{Q}) + (1-\delta) H(\mu_2, \mathcal{Q}) + H(\delta), \end{aligned} \tag{2.77}$$

where $H(\delta) := -\delta \log(\delta) - (1-\delta)\log(1-\delta)$. The quantity $H(\delta)$ is small if δ is sufficiently close to 0 or 1.

Lemma 2.36. *For any $k, n \in \mathbb{N}$, any $\epsilon_1, \epsilon_2 \in (0,1)$ and any probability measures τ, τ_1, τ_2 on $\mathbb{P}(\mathbb{R}^2)$ that satisfy $\tau = (1-\epsilon_2)\tau_1 + \epsilon_2\tau_2$ and* $\operatorname{diam}(\operatorname{supp} \tau_1) = O(q^{-k+\epsilon_1 n})$*, we have*

$$\begin{aligned} \frac{1}{n} H\left(\tau, \mathcal{Q}_{n+k} | \mathcal{Q}_k\right) &\geqslant \frac{1}{n} H(\tau, \mathcal{Q}_{n+k}) - \left(\epsilon_2 + \frac{H(\epsilon_2)}{n} + O(\epsilon_1)\right), \\ \frac{1}{n} H\left(\tau, \mathcal{Q}_{n+k} | \mathcal{Q}_k\right) &\geqslant \frac{1}{n} H(\tau_1, \mathcal{Q}_{n+k}) - (\epsilon_2 + O(\epsilon_1)). \end{aligned}$$

Proof. Note that the support of the measure τ_1 intersects at most $O(q^{\epsilon_1 n})$ atoms of $\mathcal{Q}_k$. By Eq. (2.75) and Eq. (2.77), we have

$$\begin{aligned} &\frac{1}{n} H\left(\tau, \mathcal{Q}_{n+k} | \mathcal{Q}_k\right) \geqslant (1-\epsilon_2)\frac{1}{n} H(\tau_1, \mathcal{Q}_{n+k} | \mathcal{Q}_k) \\ \geqslant &(1-\epsilon_2)\frac{1}{n} H(\tau_1, \mathcal{Q}_{n+k}) - O(\epsilon_1) \geqslant \frac{1}{n} H(\tau, \mathcal{Q}_{n+k}) - (\epsilon_2 + H(\epsilon_2)/n + O(\epsilon_1)). \end{aligned}$$

For the second one, we have

$$\begin{aligned} &\frac{1}{n} H\left(\tau, \mathcal{Q}_{n+k} | \mathcal{Q}_k\right) \geqslant (1-\epsilon_2)\frac{1}{n} H(\tau_1, \mathcal{Q}_{n+k} | \mathcal{Q}_k) \\ \geqslant &\frac{1}{n} H(\tau_1, \mathcal{Q}_{n+k}) - O(\epsilon_1) - \epsilon_2 \frac{1}{n} H(\tau_1, \mathcal{Q}_{n+k} | \mathcal{Q}_k) \geqslant \frac{1}{n} H(\tau_1, \mathcal{Q}_{n+k}) - (\epsilon_2 + O(\epsilon_1)). \end{aligned}$$ □

2.11 Entropy dimension of projection measure

A probability measure μ on a metric space X is called *exact dimensional* if for μ-a.e. x, we have that

$$\lim_{r\to 0}\frac{\log\mu(B(x,r))}{\log r}$$

exists and equals to a constant independent of x. This constant is called the *exact dimension of the measure* μ. Young [You82] proved that if μ is exact dimensional, then its exact dimension coincides with $\dim\mu$, the Hausdorff dimension of μ, which we defined in the Introduction. Hence we also use $\dim\mu$ to denote the exact dimension of μ.

In the same paper [You82], Young proved a fundamental property of exact dimensional measures (see also [FLR02]).

Lemma 2.37. *Let μ be an exact dimensional probability measure on $\mathbb{R}$. Then*

$$\dim\mu=\lim_{n\to\infty}\frac{1}{n}H(\mu,\mathcal{Q}_n).$$

This lemma enables us to use the entropy to compute the dimension of the measure μ. The limit on the right-hand side is usually called the *entropy dimension*.

Let μ be the ν-stationary measure on $\mathbb{P}(\mathbb{R}^3)$ given as in Section 2.7. In [LL23b] and [Rap21], it is proved that for μ^--a.e. $V\in\mathbb{P}(\mathbb{R}^3)$ the projection measure $\pi_{V^\perp}\mu$ is exact dimensional. Thus we have

Lemma 2.38. *For μ^--a.e. $V\in\mathbb{P}(\mathbb{R}^3)$, we have*

$$\dim\pi_{V^\perp}\mu=\lim_{n}\frac{1}{n}H(\pi_{V^\perp}\mu,\mathcal{Q}_n),$$

and they are of the same value.

See [LL23b, Corollary 6.5] and [Rap21, Theorem 1.3] for the proof.

The following lemma is an application of "pseudo-continuity" property of fixed-scale entropy.

Lemma 2.39. *For any $\epsilon>0$, any $m\geqslant M(\epsilon)$ large enough, any $n\geqslant N(\epsilon,m)$, we have*

$$\inf_{V\in\mathbb{P}(\mathbb{R}^3)}\mathbb{P}\left\{\mathbf{U}(n):\alpha-\epsilon\leqslant\frac{1}{m}H(\pi_{(g_{\mathbf{U}(n)}^{-1}V)^\perp}\mu,\mathcal{Q}_m)\leqslant\alpha+\epsilon\right\}>1-\epsilon. \tag{2.78}$$

Proof. By Lemma 2.38, for μ^--almost every V, $\pi_{V^\perp}\mu$ has entropy dimension α. We have the equidistribution of random walks (Lemma 2.23), i.e. $\mathbb{E}(\delta_{A_{\mathbf{U}(n)}^{-1}V})$ weakly converges to μ^- as n goes to infinity and the convergence is uniform with respect to V. Then (2.78) can be shown immediately if the scale-m entropy were continuous as a function of measure. Since it is not, we need the following argument.

We can replace the entropy $\frac{1}{m}H(\cdot,\mathcal{Q}_m)$ by a continuous variant $F_m(\cdot)$ at the cost of $O(1/m)$, where F_m is continuous in the weak* topology. For a possible choice of F_m, see the end of Section 2.8 in [BHR19]. Due to Lemma 2.38, for large m, we can obtain an open set $E(m)$ with $\mu^-(E(m))>1-\epsilon/2$ such that for $V\in E(m)$, we have $|F_m(\pi_{V^\perp}\mu)-\alpha|<\epsilon/2$. Applying Lemma 2.23 to $E(m)$, for large $n>N_2$, we have Eq. (2.78) with $\frac{1}{m}H(\cdot,\mathcal{Q}_m)$ replaced $F_m(\cdot)$. Then replace F_m back to $\frac{1}{m}H(\cdot,\mathcal{Q}_m)$. The resulting $O(1/m)$ error can be absorbed in ϵ, which proves the claim. $\square$

3 Non-concentration on arithmetic sequences

One of the main ingredients to show Theorem 1.10 is to prove the *porosity* property of projections of the stationary measure, see Section 4. To obtain the porosity of measures, a key step which we show below is to get the *non-concentration at arithmetic sequences* property. We identify $\mathbb{P}(\mathbb{R}^2)$ with $\mathbb{R}/\mathbb{Z}$ (see Eq. (2.3)).

Definition 3.1. A probability measure τ on $\mathbb{P}(\mathbb{R}^2)$ is called *non-concentrated on arithmetic sequences* (across scales) if for any $\delta > 0$, there exist $l, k_0 \in \mathbb{N}$ such that for all $k > k_0$, we have

$$\tau\left(\bigcup_{0\leqslant n<q^k} B\left(\frac{n}{q^k}, \frac{1}{q^{k+l}}\right)\right) \leqslant \delta.$$

In [BHR19], a doubling property called *uniformly continuous across scales* (UCAS) was defined for measures on metric spaces, which plays an important role in the proof of porosity in their setting (Lemma 3.5 in [BHR19]). This property is quite strong and can imply the uniform Hölder regularity of measures. However, it is not clear whether it holds for a (projection of a) general stationary measure. In our paper, we apply the knowledge of Fourier decay of Furstenberg measures in [Li22] to obtain the weaker substitution of UCAS, i.e. non-concentration at arithmetic sequences, which implies Lemma 4.3 that plays a role as that of Lemma 3.5 in [BHR19].

Recall that the Fourier coefficients of $\tau \in \mathbf{P}(\mathbb{P}(\mathbb{R}^2))$ are defined as follows: for any $m \in \mathbb{Z}$,

$$\hat{\tau}(m) := \int_{x\in\mathbb{R}/\mathbb{Z}} e^{2\pi i m x} \mathrm{d}\tau(x).$$

We say τ is a Rajchman measure (or Fourier decay) if $\hat{\tau}(m) \to 0$ as $|m| \to \infty$.

Lemma 3.2. *Let $\tau \in \mathbf{P}(\mathbb{P}(\mathbb{R}^2))$ be a Rajchman measure. Then τ satisfies Definition 3.1. In particular, we can choose $l, k_0 \in \mathbb{N}$ in Definition 3.1 such that*

$$\hat{\tau}(jq^{k_0}) \leqslant \delta^3/(10^4 q^2) \quad \text{for any } j \neq 0 \text{ and } 10 < q^l\delta \leqslant 10q.$$

Proof. Let f be a smooth function supported on $B(0, 2/q^l) \subset \mathbb{R}/\mathbb{Z}$ such that it equals to 1 on $B(0, 1/q^l)$ and $\|f''\|_{L^1} \leqslant 4q^{2l}$. Define $F : \mathbb{R}/\mathbb{Z} \to \mathbb{R}$ by $F(x) = \sum_{0\leqslant n<q^k} f(xq^k - n)$. Then

$$\tau\left(\bigcup_{0\leqslant n<q^k} B\left(nq^{-k}, q^{-(k+l)}\right)\right) \leqslant \int F\mathrm{d}\tau = \sum_{m\in\mathbb{Z}} \hat{\tau}(m)\hat{F}(-m),$$

where $\hat{F}(m) = \int_{\mathbb{R}/\mathbb{Z}} F(x)e^{i2\pi mx}dx$. We have

$$\begin{aligned}\hat{F}(m) &= \int_{\mathbb{R}/\mathbb{Z}} \sum_{0\leqslant n<q^k} f(q^k x - n)e^{i2\pi mx}\mathrm{d}x = \frac{1}{q^k}\int_{\mathbb{R}/\mathbb{Z}} \sum_{1\leqslant n\leqslant q^k} f(y)e^{i2\pi m(y+n)/q^k}\mathrm{d}y \\ &= \frac{\sum_{1\leqslant n\leqslant q^k} e^{i2\pi mn/q^k}}{q^k} \int_{\mathbb{R}/\mathbb{Z}} f(y)e^{i2\pi my/q^k}\mathrm{d}y.\end{aligned}$$

It is non-zero only if $m = jq^k$. Therefore, we have $\hat{F}(jq^k) = \hat{f}(j)$ and

$$\tau\left(\cup_{0\leqslant n<q^k} B\left(nq^{-k}, q^{-(k+l)}\right)\right) \leqslant \sum_{j\in\mathbb{Z}} \hat{\tau}(q^k j)\hat{f}(-j).$$

When $j=0$, we have $\hat{f}(0)=\int f\leqslant 4q^{-l}$. For the sum of the remaining terms, we have

$$\left|\sum_{j\neq 0}\hat{\tau}(q^kj)\hat{f}(j)\right|\leqslant\left(\sup_{j\neq 0}|\hat{\tau}(q^kj)|\right)\sum_{j\neq 0}|\hat{f}(j)|.$$

Since f is smooth, we have

$$\sum_{j\neq 0}|\hat{f}(j)|\leqslant\sum_{j\neq 0}\frac{1}{j^2}\|f''\|_{L^1}\leqslant 4\|f''\|_{L^1}\leqslant 16q^{2l}<\infty.$$

Since τ is a Rajchman measure, we have for any k large enough, $\sup_{j\neq 0}|\hat{\tau}(q^kj)|\leqslant\delta^3/(10^4q^2)$. Choose $l\in\mathbb{N}$ such that $10<q^l\delta\leqslant 10q$. We have

$$\sum_{j\in\mathbb{Z}}\hat{\tau}(q^kj)\hat{f}(j)\leqslant\int f+\left(\sup_{j\neq 0}|\hat{\tau}(q^kj)|\right)\sum_{j\neq 0}|\hat{f}(j)|\leqslant 4/q^l+(\delta^3/10^4q^2)\times 16q^{2l}<\delta.$$

Therefore τ satisfies Definition 3.1. □

For higher dimensional cases, the Fourier decay of Furstenberg measures on flag varieties and projective spaces is known in [Li22], cf. Appendix E for 3-dimensional case. For our purpose we only consider the flag variety $\mathcal{F}(\mathbb{R}^3)$ here.

Definition 3.3. 1. (Definitions and notations) Let

$$\mathcal{F}(\mathbb{R}^3):=\left\{(V_2,V_2):V_1\subset V_2,V_i\text{ is a linear subspace of }\mathbb{R}^3\text{ of dimension }i\right\}$$

be the space of flags in $\mathbb{R}^3$, which admits a natural $\mathrm{SL}_3(\mathbb{R})$-action. We abbreviate $\mathcal{F}(\mathbb{R}^3)$ to $\mathcal{F}$ in later discussion.

2. (Metrics on $\mathcal{F}$) For $\eta=(V_1(\eta),V_2(\eta)),\eta'=(V_1(\eta'),V_2(\eta'))\in\mathcal{F}$, let

$$d_1(\eta,\eta')=d_{\mathbb{P}(\mathbb{R}^3)}(V_1(\eta),V_1(\eta'))\ \text{ and }\ d_2(\eta,\eta')=d_{\mathbb{P}(\mathbb{R}^3)}(V_2(\eta),V_2(\eta')).$$

Here $d_{\mathbb{P}(\mathbb{R}^3)}(V_2(\eta),V_2(\eta'))$ is the Hausdorff distance between two projective lines in $(\mathbb{P}(\mathbb{R}^3),d)$. Then the metric d on $\mathcal{F}$ can be defined as the following:

$$d(\eta,\eta'):=\max\{d_1(\eta,\eta'),d_2(\eta,\eta')\}. \tag{3.1}$$

3. (Lifting functions from $\mathbb{P}(\mathbb{R}^3)$)A function φ on $\mathcal{F}$ is called *lifted from* $\mathbb{P}(\mathbb{R}^3)$ if $\varphi(\eta)$ only depends on the first coordinate $V_1(\eta)$.

Recall that μ is the ν-stationary measure on $\mathbb{P}(\mathbb{R}^3)$. In the following, for any $V\in\mathbb{P}(\mathbb{R}^3)$, by identifying $\mathbb{P}(V^\perp)$ isometrically with $\mathbb{P}(\mathbb{R}^2)$, we view $\pi_{V^\perp}\mu$ as a measure on $\mathbb{P}(\mathbb{R}^2)$.

Proposition 3.4. *For any $\delta>0$, there exist $l,k_0\in\mathbb{N}$ such that for any $V\in\mathbb{P}(\mathbb{R}^3)$ and any $k\geqslant k_0$*

$$\pi_{V^\perp}\mu\Big(\bigcup_{0\leqslant n<q^k}B(nq^{-k},q^{-(k+l)})\Big)\leqslant\delta. \tag{3.2}$$

Proof. By the proof of Lemma 3.2, it is sufficient to prove $\widehat{\pi_{V^\perp}\mu}(m)\to 0$ uniformly for all V as $|m|\to\infty$. To obtain it our strategy is to apply Theorem E.2. Recall the constants $\epsilon_0,\epsilon_1>0$ from Theorem E.2. Let $\epsilon_2=\epsilon_0/8$. Fix an arbitrary $V\in\mathbb{P}(\mathbb{R}^3)$ and an integer $m\in\mathbb{Z}$ such that $|m|^{\epsilon_2}>10$. We identify $\mathbb{P}(V^\perp)$ with $\mathbb{R}/\mathbb{Z}$ isometrically.

Step 1: Define, lift and localize φ. Consider the map

$$\begin{aligned}\varphi : \mathbb{P}(\mathbb{R}^3) - \{V\} &\to \mathbb{P}(V^\perp), \\ x &\mapsto \pi_{V^\perp} x.\end{aligned} \tag{3.3}$$

Then φ can be lifted as a map from $\mathcal{F} - \{\eta = (V_1(\eta), V_2(\eta)) : V_1(\eta) \neq V\}$. From now to end of the proof of Proposition 3.4 we only consider the lifting of φ and we still denote by φ. By the definition, φ is neither a well-defined map on whole $\mathcal{F}$ nor a well-defined real valued function, to which we cannot apply Fourier decay (i.e. Theorem E.2) directly. To overcome this difficulty we localize φ through partition of unity.

Let $\rho_m(t)$ be a C^1-function on $[0, 1]$ such that $\rho_m(t) = 0$ if $t \leqslant 1/|m|^{\epsilon_2}$ and $\rho_m(x) = 1$ if $t \geqslant 2/|m|^{\epsilon_2}$. We further assume that the Lipchitz constant of ρ_m satisfies $\mathrm{Lip}(\rho_m) \leqslant 2|m|^{\epsilon_2}$. Let r_m be a cutoff function on $\mathcal{F}$ given by

$$r_m(\eta) = \rho_m(d_{\mathbb{P}(\mathbb{R}^3)}(V_2(\eta), V)). \tag{3.4}$$

Roughly speaking the support of r_m consists of points $\eta = (V_1(\eta), V_2(\eta))$ such that V is not too close to $V_2(\eta)$; in particular, V is also not too close to $V_1(\eta)$ and we have

$$d(V_1(\eta), V) \geqslant 1/|m|^{\epsilon_2} \text{ if } r_m(\eta) = 1. \tag{3.5}$$

Our idea to localize φ is to locally identify φ with smooth real valued functions φ_j on $\mathcal{F}$ such that $\varphi = \varphi_j \mod \mathbb{Z}$. Then the value of the oscillatory integral (i.e. Eq. (E.3)) remains the same. Here is how we proceed: take smooth functions p_1 and p_2 on $\mathbb{R}/\mathbb{Z}$ as a partition of unity with respect to open intervals $d(z, \mathbb{Z}) < 1/3$ and $d(z, \frac{1}{2} + \mathbb{Z}) < 1/3$. Let $r_{m,j}(\eta) = r_m(\eta) p_j(\pi_{V^\perp}(V_1(\eta)))$ with $\eta = (V_1(\eta), V_2(\eta))$, for $j = 1, 2$. Then on a neighborhood of the support of $r_{m,j}$, the map φ takes value in a proper subset of $\mathbb{R}/\mathbb{Z}$ which can be identified with a subset of $\mathbb{R}$ continuously. Therefore on the neighborhood of the support of $r_{m,j}$, φ can be identified with smooth real valued lifted functions φ_j, without changing the integral.

Step 2: Verify Definition E.3. Our plan is to apply Theorem E.2 to φ_j. Since both φ_j are functions lifted from $\mathbb{P}(\mathbb{R}^3)$, we only need to verify the conditions in Definition E.3. A remarkable fact for $\mathcal{F}$ is that for any $\eta \in \mathcal{F}$, its coordinate provides not only its projection $V_1(\eta)$ to $\mathbb{P}(\mathbb{R}^3)$, but also an important tangent direction $V_{\alpha_2} \subset T_\eta \mathcal{F}$ (see Appendix E).

1. (Verify Eq. (E.4) and Eq. (E.7) with $C = 64m^{4\epsilon_2}$). Using the spherical law of sines, we have

   $$\partial_1 \varphi_j(\eta) = \lim_{t \to 0} \frac{\varphi_j(V_1) - \varphi(y(t))}{t} = \lim_{t \to 0} \frac{\angle_V(V_1, y(t))}{d_R(V_1, y(t))} = \lim_{t \to 0} \frac{\sin(\angle_V(V_1, y(t)))}{d(V_1, y(t))} = \lim_{t \to 0} \frac{\sin(\angle_{V_1}(V, y(t)))}{d(y(t), V)},$$

   where $y(t)$ is a smooth curve on $\mathbb{P}(\mathbb{R}^3)$ which is contained in V_2 with $d_R(V_1, y(t)) = t$. The angle $\angle_{V_1}(V, y(t))$ is equal to angle between two hyperplane $V_2 = V_1 \oplus y(t)$ and $V_1 \oplus V$, so

   $$\partial_1 \varphi_j(\eta) = \frac{d(V_2, (V_1 \oplus V))}{d(V_1, V)} = \frac{d(V_2, V)}{d(V_1, V)^2},$$

   where we use again spherical law of sines to obtain $d(V_2, (V_1 \oplus V)) = d(V_2, V)/d(V_1, V)$[5] for the last equality. By the choice of $r_{m,j}$, on its support, due to Eq. (3.4) and Eq. (3.5), we have (See Fig. 6)

   $$\partial_1 \varphi_j(\eta) \in [1/|m|^{\epsilon_2}, |m|^{2\epsilon_2}],$$

   which implies Eq. (E.4) and Eq. (E.7) with $C = 64m^{4\epsilon_2}$.

[5]The Hausdorff distance $d(V_2, (V_1 \oplus V))$ equals $\sin(\angle_{V_1}(V, y(t)))$. Let V' be the point in V_2 such that $d(V, V') = d(V, V_2)$. Then $\angle_{V'}(V, V_1) = \pi/2$. Hence by spherical law of sines, $\sin(\angle_{V_1}(V, y(t)))/d(V, V_2) = \sin(\angle_{V_1}(V, y(t)))/d(V, V') = \sin(\pi/2)/d(V, V_1)$.

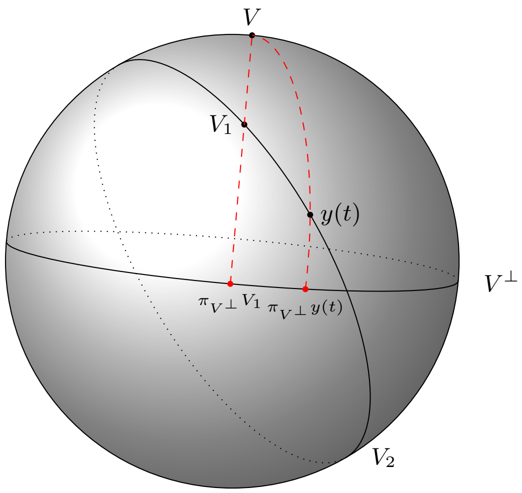

Figure 6: α_1-direction

2. (Verify Eq. (E.5)). Eq. (E.5) can be obtained directly from Lemma B.2.

3. (Verify Eq. (E.6)).

   we have

$$\partial_1 \varphi_j(\eta) - \partial_1 \varphi_j(\eta') = \frac{d(V_2, V)}{d(V_1, V)^2} - \frac{d(V_2', V)}{d(V_1', V)^2} = \frac{d(V_2, V) d(V_1', V)^2 - d(V_2', V) d(V_1, V)^2}{d(V_1, V)^2 d(V_1', V)^2}.$$

   where $V_i' := V_i(\eta')$. Due to the triangle inequality, the numerator is less than

$$|(d(V_2, V) - d(V_2', V)) d(V_1', V)^2 + d(V_2', V)(d(V_1', V)^2 - d(V_1, V)^2)| \leqslant 2(d(V_2, V_2') + d(V_1, V_1')).$$

   Due to Eq. (3.5), for η, η' in the $1/C$-neighborhood of the support of r, the denominator is greater than $1/(2|m|^{\epsilon_2})^4$. Hence, we have

$$|\partial_1 \varphi_j(\eta) - \partial_1 \varphi_j(\eta')| \leqslant 64 |m|^{4\epsilon_2} d(\eta, \eta')$$

   which is Eq. (E.6), where $d(\eta, \eta')$ is defined in Eq. (3.1).

4. (Estimate the Lipschitz constants of $r_{m,j}$). Recall a basic estimate of the Lipschitz constant of a product of two Lipchitz continuous functions f, g:

$$\mathrm{Lip}(fg) \leqslant \mathrm{Lip}(f)|g|_\infty + |f|_\infty \sup_{\eta \neq \eta', \eta, \eta' \in \operatorname{supp} f} \frac{|g(\eta) - g(\eta')|}{d(\eta, \eta')}.$$

   By this estimate and (2) in Lemma B.2,

$$\begin{aligned} \mathrm{Lip}(r_{m,j}) &\leqslant \mathrm{Lip}(r_m) + \sup_{\eta \neq \eta', \eta, \eta' \in \operatorname{supp} r_m} \frac{|p_j(\pi_{V^\perp}(V_1(\eta))) - p_j(\pi_{V^\perp}(V_1(\eta')))|}{d(\eta, \eta')} \\ &\leqslant |m|^{\epsilon_2} + \mathrm{Lip}(p_j)|m|^{\epsilon_2} \leqslant 4|m|^{\epsilon_2}. \end{aligned}$$

Step 3: Estimate of $\widehat{\pi_{V^\perp}\mu}(m)$. From Step 2, each φ_j is $(m^{\epsilon_0}, r_{m,j})$-good (due to the choice of ϵ_2). Then for each j and all m with $|m|$ large enough, by Theorem E.2 we have

$$|\int e^{2\pi i m \varphi_j(\eta)} r_{m,j}(\eta) \mathrm{d}\mu_{\mathcal{F}}(\eta)| \leqslant |m|^{-\epsilon_1},$$

where $\mu_{\mathcal{F}}$ is the ν-stationary measure on $\mathcal{F}$, whose projection on $\mathbb{P}(\mathbb{R}^3)$ is μ. Therefore, combined with Lemma 2.24 and Eq. (3.5), we obtain

$$\begin{aligned}
|\widehat{\pi_{V^\perp}\mu}(m)| &= |\int e^{2\pi i m\varphi(\eta)}\mathrm{d}\mu_{\mathcal{F}}(\eta)| \\
&\leqslant \left|\int e^{2\pi i m\varphi(\eta)}(r_{m,1}(\eta)+r_{m,2}(\eta))\mathrm{d}\mu_{\mathcal{F}}(\eta)\right| + \left|\int e^{2\pi i m\varphi(\eta)}(1-r_{m,1}(\eta)-r_{m,2}(\eta))\mathrm{d}\mu_{\mathcal{F}}(\eta)\right| \\
&\leqslant \sum_j \left|\int e^{2\pi i m\varphi_j(\eta)}r_{m,j}(\eta)\mathrm{d}\mu_{\mathcal{F}}(\eta)\right| + \mu_{\mathcal{F}}(\{r_m\neq 1\}) \\
&\leqslant \sum_j \left|\int e^{2\pi i m\varphi_j(\eta)}r_{m,j}(\eta)\mathrm{d}\mu_{\mathcal{F}}(\eta)\right| + \mu(\{V_1(\eta): d(V_1(\eta),V)<1/|m|^{\epsilon_2}\}) \ll |m|^{-\beta},
\end{aligned}$$

for some $\beta>0$ and uniform for all V. The proof is complete. ☐

Proposition 3.5. *For any $\delta>0$, there exist $l, k_0\in\mathbb{N}$ such that for any $V\in\mathbb{P}(\mathbb{R}^3)$, any $k\geqslant k_0$ and $g\in\mathrm{SL}_2(\mathbb{R})$, we have*

$$g\pi_{V^\perp}\mu\left(\bigcup_{0\leqslant n<q^k} B(nq^{-(k+[\chi_1(g)])}, q^{-(k+[\chi_1(g)]+l)})\right)\leqslant\delta. \tag{3.6}$$

Proof. Abbreviate $\pi_{V^\perp}\mu$ by τ. It follows from Lemma 2.24 that there exists $s\in(0,1)$ such that $\tau(b(g^-,s)^c)<\delta/10$ and s only depends on μ and δ. We decompose the measure τ as follows:

$$\tau = \lambda\cdot\tau_{b(g^-,s)^c} + (1-\lambda)\cdot\tau_{b(g^-,r)},$$

where $\lambda=\tau(b(g^-,s)^c)$. We denote $\tau_{b(g^-,s)^c}$ by τ_I to simply to notation. Set

$$t=[\chi_1(g)-2|\log s|-1]=[\chi_1(g)+2\log s-1].$$

Let $S_t:\mathbb{R}\to\mathbb{R}, x\mapsto q^t x$ be the scaling map given as in Definition 2.20. Lemma 2.4 gives that $gb(g^-,s)\subset B(g^+,q^{-\chi_1(g)}/s^2)$. If t is a non-negative integer, then S_t induces a map $\mathbb{R}/\mathbb{Z}\to\mathbb{R}/\mathbb{Z}$, which we still denote by S_t by abusing the notation. Lemma 2.4 gives that $gb(g^-,s)\subset B(g^+,q^{-\chi_1(g)}/s^2)$. Hence the restriction of S_t on $gb(g^-,s)$ is a diffeomorphism. If t is negative, then S_t is a contracting map. Identifying $\mathbb{R}/\mathbb{Z}$ (hence $\mathbb{P}(V^\perp)$) with $[-\frac12,\frac12)$ such that $V_g^+=0$, we have that $gb(g^-,s)\subset(-\frac12,\frac12)$ is an interval containing 0. Hence S_t restricted on $gb(g^-,s)$ is also a diffeomorphism to its image in $gb(g^-,s)$.

We have

$$\begin{aligned}
g\tau_{\mathbf{I}}\left(\cup_n B\left(\frac{n}{q^{[\chi_1(g)]+k}},\frac{1}{q^{[\chi_1(g)]+k+\ell}}\right)\right) &= S_t(g\tau_{\mathbf{I}})\left(S_t(\cup_n B\left(\frac{n}{q^{[\chi_1(g)]+k}},\frac{1}{q^{[\chi_1(g)]+k+\ell}})\right)\right) \\
&= S_t g\tau_{\mathbf{I}}\left(\cup_n B\left(\frac{n}{q^{[\chi_1(g)]+k-t}},\frac{1}{q^{[\chi_1(g)]+k-t+\ell}}\right)\right),
\end{aligned}$$

where we only sum up n's such that the interval $B\left(\frac{n}{q^{[\chi_1(g)]+k}},\frac{1}{q^{[\chi_1(g)]+k+\ell}}\right)$ intersects the support of $\tau_{\mathbf{I}}$.

We use Lemma 2.8 to obtain distortion estimates. For $y\in b(g^-,s)$, we have

$$|(S_tg)'y| = q^t|g'y|\in[s^2/q,1/q], |(S_tg)''(y)| = q^t|g''(y)|\leqslant 10/(qs). \tag{3.7}$$

Let p_1 be a cuffoff function which equals 1 on $b(g^-,s)$ and 0 outside $b(g^-,s/2)$. Let $p_1\tau$ be the measure on $\mathbb{P}(V^\perp)$ defined by $p_1\tau(f)=\int p_1\cdot f d\tau$ for any measurable function f on $\mathbb{P}(V^\perp)$. Consider the map $\psi:\mathbb{P}(\mathbb{R}^3)-\{V\}\to\mathbb{P}(V^\perp)$ defined by

$$\psi(x)=(S_tg)\pi_{V^\perp}(x)=S_tg\varphi(x)$$

with φ given as in Eq. (3.3). Then ψ can be lifted to a function defined on $\mathcal{F} - \{\eta = (V_1(\eta), V_2(\eta)) : V_1(\eta) \neq V\}$.

Take any integer m with $|m|^{\epsilon_2} > 4$, where $\epsilon_2 = \epsilon_0/8$ with ϵ_0 given as in Theorem E.2. Let $r_{m,1}(\eta) = r_m(\eta)p_1(\pi_{V^\perp}(\eta))$ where $\eta = (V_1(\eta), V_2(\eta))$ and r_m is defined as in Eq. (3.4). Here is the key observation: since $S_t g$ on $b(g^-, s)$ is a bi-Lipschitz map to its image with bounded C^2-norm (Eq. (3.7)), if φ is $(C, r_{m,1})$-good as in Definition E.3, then $(S_t g)\varphi$ is $(C', r_{m,1})$-good with $C' = (10q/s)^2 C$. As s is fixed, we can verify as in Proposition 3.4 that φ is $(64m^{\epsilon_0/2}, r_{m,1})$-good and hence $\psi = (S_t g)\varphi$ is $(m^{\epsilon_0}, r_{m,1})$-good for all integer m with $|m|$ sufficiently large. Then the same argument as in Proposition 3.4 yields that for all integer m with $|m|$ large,

$$|\widehat{S_t g(p_1\tau)}(m)| \leqslant |\int e^{2\pi i m\psi(\eta)} r_{m,1}(\eta) \mathrm{d}\mu_{\mathcal{F}}(\eta)| + \mu_{\mathcal{F}}(\{r_m = 1\}^c) \ll m^{-\beta}.$$

Lemma 3.2 implies that there exist $l, k_0 \in \mathbb{N}$ (depending on δ) such that for all $k \in N$ with $[\chi_1(g)] + k - t > k_0$,

$$\begin{aligned} &S_t g \tau_{\mathbf{I}} \left(\cup_n B \left(\frac{n}{q^{[\chi_1(g)]+k-t}}, \frac{1}{q^{[\chi_1(g)]+k-t+\ell}} \right) \right) \\ \leqslant &\frac{1}{1-\lambda} S_t g(p_1\tau) \left(\cup_n B \left(\frac{n}{q^{[\chi_1(g)]+k-t}}, \frac{1}{q^{[\chi_1(g)]+k-t+\ell}} \right) \right) < \delta/2. \end{aligned}$$

Note that $g\tau = (1-\lambda)\cdot g\tau_{\mathbf{I}} + \lambda \cdot g\tau_{b(g^-,s)^c}$ with $\lambda < \delta/10$. This proof finishes. □

4 Convolution and porosity

Definition 4.1. Let $\alpha, \epsilon > 0$ and $m, i_1, i_2 \in \mathbb{N}$ with $i_1 \leqslant i_2$. We call a probability measure τ on $\mathbb{R}$ or $\mathbb{P}(\mathbb{R}^2)$ is (α, ϵ, m)-*entropy porous* from scale i_1 to i_2 if

$$\mathbb{P}_{i_1 \leqslant j \leqslant i_2} \left\{ \frac{1}{m} H(\tau_{x,j}, \mathcal{Q}_{m+j}) < \alpha + \epsilon \right\} > 1 - \epsilon.$$

The main result of this part is the entropy porosity of projections of stationary measures, which is a key ingredient in entropy growth argument.

Proposition 4.2. *For every $\epsilon > 0$, $m \geqslant M(\epsilon)$, $k \geqslant K(\epsilon, m)$, $n \geqslant N(\epsilon, m, k)$,*

$$\inf_{V \in \mathbb{P}(\mathbb{R}^3)} \mathbb{P}_{1 \leqslant i \leqslant n} \left\{ \pi_{V^\perp} g_{\mathbf{I}(i)} \mu \text{ is } (\alpha, \epsilon, m)\text{-entropy porous from scale } i \text{ to } i+k \right\} > 1 - \epsilon. \quad (4.1)$$

4.1 General results on entropy porosity

In this section, we list some general lemmas in [BHR19, Section 3.1] which show that if a measure τ decomposes as a convex combination of measures τ_i whose support has a short length, then many properties of the τ_i, and specifically their entropies, are inherited by the q-adic components of τ.

In the following lemma, one fixes a scale k, a short scale $k + l_0$ for the length of the support of the measures τ_i which make up τ, and an even shorter scale $k + m$ at which the entropy appears. The dependence of the parameters is that l_0 is large depending on τ and δ, $m \gg l_0$, and k is arbitrary.

Lemma 4.3. *For every $\epsilon > 0$, there exists $\delta > 0$ with the following property. Let $\tau \in \mathbf{P}(\mathbb{P}(\mathbb{R}^2))$ be written as a convex combination $\tau = (\sum_{i=1}^N p_i \tau_i) + p_0 \tau_0$, with $p_0 < \delta$, and suppose that for $l_0 = l_0(\tau, \delta)$, some $m, k \in \mathbb{N}$ and $\alpha > 0$,*

(1) $\frac{1}{m} H(\tau_i, \mathcal{Q}_{k+m}) > \alpha$ for every $i = 1, ..., N$.

(2) $\operatorname{diam}(\operatorname{supp}\tau_i) \leqslant q^{-(k+l_0)}$ *for every* $i = 1, \ldots, N$.

(3)

$$\tau\left(\bigcup_n B\left(\frac{n}{q^k}, \frac{1}{q^{k+l_0}}\right)\right) \leqslant \delta.$$

Then

$$\mathbb{P}_{j=k}\left(\frac{1}{m}H(\tau_{x,j}, \mathcal{Q}_{j+m}) > \alpha - \epsilon\right) > 1 - \epsilon.$$

We view τ_0 and τ_i with $1 \leqslant i \leqslant N$ as the *bad* part and the *good* part of τ respectively, see Definition 4.6 in the next subsection. Lemma 4.3 can be proved as in [BHR19, Lemma 3.5]. We replace UCAS condition ([BHR19, Definition 3.11]) by the current weaker condition (3) (Definition 3.1), which is sufficient for the proof.

As in [BHR19, Lemma 3.7], we use the above lemma to deduce the entropy porosity.

Lemma 4.4. *For every* $\epsilon > 0$, *there exists* $\delta > 0$ *with the following property. Let* $m, l \in \mathbb{N}$ *and* $n > n(m, l)$ *be given, and suppose that* $\tau \in \mathbf{P}(\mathbb{P}(\mathbb{R}^2))$ *is a measure such that for* $(1-\delta)$*-fraction of* $1 \leqslant k \leqslant n$, *we can write* $\tau = \sum p_i \tau_i$ *so as to satisfy the conditions of the previous lemma for the given* δ, l_0, m *and* k. *Assume further that*

$$|\frac{1}{n}H(\tau, \mathcal{Q}_n) - \alpha| < \delta.$$

Then τ *is* (α, ϵ, m)*-entropy porous from scale* 1 *to* n.

Lemma 4.5. *For every* $\epsilon > 0$, *there exists* $\delta > 0$ *with the following property. Let* $\tau \in \mathbf{P}(\mathbb{P}(\mathbb{R}^2))$ *be written as a convex combination* $\tau = (\sum_{1 \leqslant i \leqslant N} p_i \tau_i) + p_0 \tau_0$, *with* $p_0 < \delta$, *and suppose that for* $l_1 = l_1(\tau, \delta)$, *some* $k, p \in \mathbb{N}$, $m > m(\epsilon, p)$ *and* $\beta > 0$,

(1) $\frac{1}{m}H(\tau_i, \mathcal{Q}_{k+m}) < \beta$ *for every* $i = 1, \ldots, N$.

(2) $\operatorname{diam}\operatorname{supp}\tau_i \leqslant q^{-(k+l_1)}$ *for every* $i = 1, \ldots, N$.

(3)

$$\tau\left(\bigcup_n B\left(\frac{n}{q^k}, \frac{1}{q^{k+l_1}}\right)\right) \leqslant \delta.$$

(4) Every interval of length q^{-k} *intersects the support of at most* p *of the measures* τ_i.

Then

$$\mathbb{P}_{j=k}\left\{\frac{1}{m}H(\tau_{x,j}, \mathcal{Q}_{j+m}) < \beta + \epsilon\right\} > 1 - \epsilon.$$

The proof is the same as that of [BHR19, Lemma 3.9], where they used almost convexity of entropy to replace the concavity of entropy in the proof of [BHR19, Lemma 3.5]. Notice that condition (3) is changed. See the discussion below Lemma 4.3.

4.2 Decomposition of a general measure

Lemma 4.3 motivates the following decomposition of measures for the projective action on $\mathbb{P}(\mathbb{R}^d)$. This decomposition is essential because the projective action is more subtle: it only contracts outside a neighborhood of a hyperplane (Lemma 2.4). This decomposition allows us to estimate the entropy of measures that are pushed forward by the projective action.

Definition 4.6. Let τ be a probability measure on a metric space X. For $s, \delta > 0$, a convex decomposition $\tau = \theta\tau_{\mathbf{II}} + (1-\theta)\tau_{\mathbf{I}}$ is called an (s,δ)-*decomposition* of τ, if $\theta \leqslant \delta$ and $\operatorname{diam}(\operatorname{supp}\tau_{\mathbf{I}}) \leqslant s$. We call $\tau_{\mathbf{I}}$ and $\tau_{\mathbf{II}}$ the *good* part and the *bad* part of τ, respectively.

Moreover, if the good part $\tau_{\mathbf{I}}$ satisfies $\frac{1}{m}H(\tau_{\mathbf{I}}, \mathcal{Q}_{m+l}) > \alpha$, then we call the decomposition an (s,δ)-*decomposition with* α-*entropy concentration at scale* (m,l).

We will be most interested in the following decomposition coming from the action of $\mathrm{SL}_d(\mathbb{R})$ on $\mathbb{P}(\mathbb{R}^d)$.

Definition 4.7. Let $d \geqslant 2$. Let $h \in \mathrm{SL}_d(\mathbb{R})$ be such that $\sigma_1(h) > \sigma_2(h)$ and τ be a probability measure on $\mathbb{P}(\mathbb{R}^d)$. The r-attracting decomposition of the pair (h,τ) is defined as follows:

$$h\tau = \theta \cdot h(\tau_{b(h^-,r)^c}) + (1-\theta)\cdot h(\tau_{b(h^-,r)}), \tag{4.2}$$

where $\theta = \tau(b(h^-,r)^c)$. We call $h(\tau_{b(h^-,r)^c})$ (resp. $h(\tau_{b(h^-,r)})$) the r-*repelling part* (resp. *the* r-*attracting part*) of (h,τ).

We need a lemma of a more precise estimate of the support.

Lemma 4.8. *For any $h \in \mathrm{SL}_2(\mathbb{R})$ with $\sigma_1(h) > \sigma_2(h)$ and any $r \in [q^{-\chi_1(h)}, 1/10)$, we have*

$$B(h^+, q^{-\chi_1(h)}/2r) \subset \operatorname{supp}\ h(\tau_{b(h^-,r)}) \subset B(h^+, q^{-\chi_1(h)}/r^2). \tag{4.3}$$

Proof. Due to Lemma 2.4, we obtain the second relation.

From some elementary computations, we can obtain a lower bound of the support. Without loss of generality, we may assume $h = \operatorname{diag}(q^{\chi_1(h)/2}, q^{-\chi_1(h)/2})$. In view of Remark 2.2, by a straightforward computation, we have that the diameter of $hb(h^-,r)$ is bounded above by $\sin\left(\arctan\left(\frac{q^{-\chi_1(h)}}{\tan(\arcsin(r))}\right)\right)$. □

The following lemma links the decomposition given in Definition 4.6 and the one given in Definition 4.7.

Lemma 4.9. *Let τ be a probability measure on $\mathbb{P}(\mathbb{R}^d)$. Suppose there exist $C, \beta > 0$ such that for any $R \geqslant 0$ and hyperplane W in $\mathbb{P}(\mathbb{R}^d)$, we have*

$$\tau(B(W,R)) \leqslant CR^\beta. \tag{4.4}$$

Then for any $g \in \mathrm{SL}_d(\mathbb{R})$ and $0 < r \leqslant 1/2$, the r-attracting decomposition of (g,τ) is an $(q^{-\chi_1(g)}/r^2, Cr^\beta)$-decomposition of the pushforward measure $g\tau$.

Proof. We have $\operatorname{diam}(B(g^+, q^{-\chi_1(g)}/r^2)) \geqslant \operatorname{diam}(gb(g^-,r))$ (see Lemma 2.4). And by Eq. (4.4), we have $\tau(b(g^-,r)^c) \leqslant Cr^\beta$ □

Lemma 4.10. *Let $g \in \mathrm{SL}_2(\mathbb{R})$ and τ be a probability measure that satisfies the assumption of Lemma 4.9. For any $\epsilon > 0$, $0 < r < r(\epsilon) < \frac{1}{2}$ and $m \geqslant M(\epsilon, r)$, if $\frac{1}{m}H(\tau, \mathcal{Q}_m) \geqslant \alpha$, then the r-attracting decomposition of (g,τ) is a $(q^{-\chi_1(g)}/r^2, Cr^\beta)$-decomposition of $g\tau$ with $(\alpha-\epsilon)$-entropy concentration at scale (m,l) with $l = \chi_1(g)$.*

Proof. We use (2.77) to obtain

$$\begin{aligned}\frac{1}{m}H\left(\tau_{b(g^-,r)}, \mathcal{Q}_m\right) &\geqslant \frac{1}{1-\theta}\left(\frac{1}{m}H(\tau,\mathcal{Q}_m) - \frac{1}{m}H(\theta) - \frac{\theta}{m}H\left(\tau_{b(g^-,r)^c}, \mathcal{Q}_m\right)\right) \\ &\geqslant \frac{1}{1-\theta}\left(\frac{1}{m}H(\tau,\mathcal{Q}_m) - \frac{1}{m}H(\theta) - \theta\right).\end{aligned}$$

with $\theta < Cr^\beta$ by the previous lemma. Due to Lemma 2.8 and Lemma 2.35, we have

$$\frac{1}{m}H\left(g(\tau_{b(g^-,r)}), \mathcal{Q}_{m+\chi_1(g)}\right) = \frac{1}{m}H\left(\tau_{b(g^-,r)}, \mathcal{Q}_m\right) + \frac{1}{m}O(|\log r|).$$

Combining the above two formulas, we finish the proof by taking Cr^β (hence θ) sufficiently small with respect to ϵ and m large with respect to $|\log r|$ and ϵ. □

4.3 Keeping porosity under projective transformation

Let μ be the ν-stationary measure on $\mathbb{P}(\mathbb{R}^3)$ given as in Section 2.7. In this section, we show the entropy properties of $\pi_{V^\perp}\mu$ are inherited by the pushforward measure $g\pi_{V^\perp}\mu$ with $V \in \mathbb{P}(\mathbb{R}^3)$ and $g \in \mathrm{SL}_2(\mathbb{R})$, which is a projective version of [BHR19, Lemma 3.10].

Proposition 4.11. *For any $\epsilon > 0$, there exists $\epsilon' > 0$ such that for any $m \geqslant m(\epsilon)$, $n \geqslant n(m, \epsilon)$ the following holds. For any $n_1 \in \mathbb{N}$ and $V \in \mathbb{P}(\mathbb{R}^3)$, if the measure $\pi_{V^\perp}\mu$ is (α, ϵ', m)-entropy porous from scale n_1 to $n_2 = n_1 + n$, then for any $g \in \mathrm{SL}_2(\mathbb{R})$, $g\pi_{V^\perp}\mu$ is (α, ϵ, m)-entropy porous from scale $n_1 + \chi_1(g)$ to $n_2 + \chi_1(g)$.*

Proof. Abbreviate $\pi_{V^\perp}\mu$ by τ. Fix $\epsilon > 0$ and $g \in \mathrm{SL}_2(\mathbb{R})$. Without loss of generality, we can assume ϵ small such that $\epsilon^2 < \frac{\epsilon}{100}$. Recall the uniform Hölder constant $C > 0$ and Hölder exponent of $\beta > 0$ given in Lemma 2.24.

We fix a few constants:

- let $\delta > 0$ be the constant given in Lemma 4.5 for ϵ^2;
- let $l_0, k_0 \in \mathbb{N}$ be the constants l, k_0 given in Proposition 3.5 with parameter $\delta/10$;
- fix an $l \gg l_0$ such that $q^{-(l-l_0)} \leqslant r^2$ with $r = (\delta/10C)^{1/\beta}/4$.

These constants actually depend only on ϵ. We will determine the constants $m, n \in \mathbb{N}$ later.

We will prove the proposition for $\epsilon' = \min\{\epsilon^4, \frac{\delta^2}{10}\}$. Assume that τ is (α, ϵ', m)-entropy porous from scale n_1 to n_2. By Markov inequality, there are at least $(1 - \sqrt{\epsilon'})$ fraction of levels $n_1 \leqslant j \leqslant n_2$ such that

$$\mathbb{P}\left\{\frac{1}{m}H(\tau_{j,x}, \mathcal{Q}_{j+m}) < \alpha + \epsilon'\right\} > 1 - \sqrt{\epsilon'}.$$

By taking n sufficiently large, which depends on k_0, l, ϵ', and hence only depends on ϵ, we have that for at least $(1 - 2\sqrt{\epsilon'})$ fraction of levels $n_1 \leqslant j + l \leqslant n_2$ such that

$$j \geqslant k_0 \ \text{ and } \ \mathbb{P}\left\{\frac{1}{m}H(\tau_{j+l,x}, \mathcal{Q}_{j+l+m}) < \alpha + \epsilon'\right\} > 1 - \sqrt{\epsilon'}. \tag{4.5}$$

We call j *typical* if j satisfies both inequalities in Eq. (4.5).

Recall $r = (\delta/10C)^{1/\beta}/4$. Applying Lemma 2.24 to τ, we have for any $x \in \mathbb{P}(\mathbb{R}^2)$,

$$\tau(B(x, 4r)) \leqslant \frac{\delta}{10}.$$

In particular, we have

$$\tau(b(g^-, 4r)^c) \leqslant \frac{\delta}{10}.$$

We have $q^{-(j+l)} \leqslant q^{-(l-l_0)} \leqslant r$, hence for typical j, we have

$$\mathbb{P}\left\{\begin{array}{l} x : \frac{1}{m}H(\tau_{j+l,x}, \mathcal{Q}_{m+j+l}) < \alpha + \epsilon', \\ \mathrm{supp}(\tau_{j+l,x}) \subset b(g^-, r) \end{array}\right\} > 1 - \sqrt{\epsilon'} - \frac{\delta}{10}. \tag{4.6}$$

We consider the action of g on $b(g^-, r)$. By Lemma 2.8, it scales by $q^{-\chi_1(g)}$ with distortion r^{-2}. Therefore for typical j and for x in the event given in Eq. (4.6), Lemma 2.35 yields

$$\frac{1}{m}H(g(\tau_{j+l,x}), \mathcal{Q}_{m+j+l+\chi_1(g)}) < \alpha + \epsilon' + \frac{O(|\log r|)}{m} \leqslant \alpha + 2\epsilon'. \tag{4.7}$$

The second inequality holds by taking m sufficiently large, depending on r, ϵ' and hence only depending on ϵ. Combining Eqs. (4.6) and (4.7), we have

$$\mathbb{P}\left\{\begin{array}{l} x : \frac{1}{m}H(g(\tau_{j+l,x}), \mathcal{Q}_{m+j+l+\chi_1(g)}) < \alpha + 2\epsilon', \\ \mathrm{supp}(\tau_{j+l,x}) \subset b(g^-, r) \end{array}\right\} > 1 - \sqrt{\epsilon'} - \frac{\delta}{10} > 1 - \frac{\delta}{2}, \tag{4.8}$$

where the second inequality is due to $\epsilon' \leqslant \delta^2/10$. We take out all $\tau_{j+l,x}$ in the event in (4.8). There are finitely many different measures $\tau_{j+l,x}$, and we denote these measures by $\tau_1, \dots, \tau_N$ and $\theta_i = \tau(\operatorname{supp} \tau_i)$ for $1 \leqslant i \leqslant N$. We will apply Lemma 4.5 to $g\tau$ in the level $\chi_1(g) + j$. The convex combination is given by

$$g\tau = \mathbb{E} g\tau_{j+l,x} = \sum_{1 \leqslant i \leqslant N} \theta_i \cdot g\tau_i + \theta_0 \cdot g\tau_0,$$

where τ_0 is the convex combinations of $\tau_{j+l,x}$ not in the event in (4.8). We have $\theta_0 < \delta/2$. We check the conditions of Lemma 4.5:

1. $\frac{1}{m} H(g\tau_i, \mathcal{Q}_{m+j+\chi_1(g)}) \leqslant \frac{1}{m} H(g\tau_i, \mathcal{Q}_{m+j+l+\chi_1(g)}) < \alpha + 2\epsilon' < \alpha + \epsilon^2$ for $1 \leqslant i \leqslant N$;
2. $\operatorname{diam}(\operatorname{supp}(g(\tau_i))) < q^{-\chi_1(g) - l_0 - j}$ for $1 \leqslant i \leqslant N$. Here we use the fact that the distortion of g on $b(g^-, r)$ is r^{-2} with $q^{-(l-l_0)} \leqslant r^2$;
3. $g\tau \left(\cup_n B \left(\frac{n}{q^{\chi_1(g)+j}}, \frac{1}{q^{\chi_1(g)+j+l_0}} \right) \right) < \delta/10$. This is due to Proposition 3.5 and typical $j \geqslant k_0$;
4. Each interval of length $q^{-\chi_1(g)-j}$ intersects at most $p = q^l$ of the support of $g\tau_i$, thanks to $\operatorname{diam}(g\tau_i) \geqslant q^{-\chi_1(g)-j-l}$.

Therefore by our choice of δ, it follows from Lemma 4.5 that for typical j, for $m \geqslant m(\epsilon, q^l)$ (hence only depends on ϵ),

$$\mathbb{P}\{\frac{1}{m} H((g\tau)_{j+\chi_1(g),x}, \mathcal{Q}_{j+\chi_1(g)+m}) \leqslant \alpha + 2\epsilon'\} > 1 - \epsilon^2.$$

Taking account for non-typical j, we get that

$$\mathbb{P}_{n_1 \leqslant j \leqslant n_2}\{\frac{1}{m} H((g\tau)_{j+\chi_1(g),x}, \mathcal{Q}_{j+\chi_1(g)+m}) \leqslant \alpha + 2\epsilon'\} > (1 - \epsilon^2)(1 - 2\sqrt{\epsilon'}).$$

By our choice of ϵ', we have that $\alpha + 2\epsilon' \leqslant \alpha + \epsilon$, $(1-\epsilon^2)(1 - 2\sqrt{\epsilon'}) > 1 - \epsilon$. Therefore, $g\tau$ is (α, ϵ, m)-entropy porous from scale $n_1 + \chi_1(g)$ to $n_2 + \chi_1(g)$. □

4.4 Decomposition of random projection measures

Let's recall the decomposition given in Eq. (2.5): for any $\spadesuit \in \{\mathbf{U}(n), \mathbf{I}(i)\}$ and $V \in \mathbb{P}(\mathbb{R}^3)$, if $g_\spadesuit^{-1} V \not\subset V^\perp$, then we have

$$\begin{aligned} &\pi_{V^\perp} \circ g_\spadesuit \\ =& h_{V,g_\spadesuit} \circ \pi_{g_\spadesuit^{-1}V, V^\perp} \\ =& h_{V,g_\spadesuit} \circ \pi(g_\spadesuit^{-1}V, V^\perp, (g_\spadesuit^{-1}V)^\perp) \circ \pi_{(g_\spadesuit^{-1}V)^\perp}, \end{aligned} \tag{4.9}$$

where $h_{V,g_\spadesuit}$ is a projective transformation of $\mathbb{P}(V^\perp)$ and we regard it as an element in $\mathrm{SL}_2(\mathbb{R})$ by identifying $\mathbb{P}(V^\perp)$ with $\mathbb{P}(\mathbb{R}^2)$. To simply the notations, we will denote $h_{V,g_\spadesuit}$ by $h_{V,\spadesuit}$ and $\pi_{g_\spadesuit^{-1}V,V^\perp}$ by $\pi_{V,\spadesuit}$ in the followings.

The focus of this subsection is to study the decomposition of a random projection measure $\pi_{V,\spadesuit}\mu$ defined as in Definition 4.6 and Definition 4.7. The following lemma is a corollary of the large deviation theorem of random walks on projective space.

Lemma 4.12. *For every $\epsilon > 0$, there exist $C_1 = C_1(\epsilon), N_1 = N_1(\epsilon) \geqslant 1$ such that for $n \geqslant N_1$, $V \in \mathbb{P}(\mathbb{R}^3)$*

$$\mathbb{P}\left\{ \begin{array}{l} \mathbf{U}(n) : d(g_{\mathbf{U}(n)}^{-1} V, V^\perp) > 1/C_1, \\ |\chi_1(h_{V,\mathbf{U}(n)}) - \chi_1(g_{\mathbf{U}(n)})| \leqslant \log C_1 \end{array} \right\} > 1 - \epsilon. \tag{4.10}$$

Similarly, we have a corresponding inequality for $\mathbf{I}(i)$

$$\inf_{V\in\mathbb{P}(\mathbb{R}^3)}\mathbb{P}_{1\leqslant i\leqslant n}\left\{\begin{array}{c}\mathbf{I}(i): d(g_{\mathbf{I}(i)}^{-1}V, V^\perp) > 1/C_1\\ |\chi_1(h_{V,\mathbf{I}(i)})-\chi_1(g_{\mathbf{I}(i)})|\leqslant \log C_1\end{array}\right\} > 1-\epsilon. \tag{4.11}$$

Proof. Recall the relation between the convolution measure ν^{*n} and the random walk $\mathbf{U}(n)$ Eq. (2.71). The first line for the statement of $\mathbf{U}(n)$ is due to Eq. (2.58). The second line is obtained by applying Eq. (2.59) to the random walk which is the conjugation of ν by $k\in\mathrm{SO}(3)$ with $kV=E_1$.

Then the statement of $\mathbf{I}(n)$ follows from Lemma 2.33 and the statement Eq. (4.10) for $\mathbf{U}(n)$ with $\epsilon=\epsilon/C_I$. □

To simplify the notation we denote by $X_{\mathbf{U},n,V}=X_{\mathbf{U},n,V}(C_1,\epsilon)$ the set of $\mathbf{U}(n)$ considered in (4.10). Then we have the following property for $X_{\mathbf{U},n,V}$.

Lemma 4.13. *For any $\epsilon>0$ and $\delta>0$, let $C_1=C_1(\epsilon)\geqslant 1$ and $N_1=N_1(\epsilon)$ be the constants given in Lemma 4.12. Then there exist constants $r=r(\epsilon,C_1,\delta)$ and $C_2=C_2(C_1,r)$ such that for any $V\in\mathbb{P}(\mathbb{R}^3)$ and any $n\geqslant N_1$, if $\mathbf{U}(n)\in X_{\mathbf{U},n,V}$, then the r-attracting decomposition of $(h_{V,\mathbf{U}(n)},\pi_{V,\mathbf{U}(n)}\mu)$ enjoys the following properties:*

1. *it is a $(C_2\cdot q^{-\chi_1(g_{\mathbf{U}(n)})},\delta)$-decomposition of $\pi_{V^\perp}g_{\mathbf{U}(n)}\mu$.*
2. *for any $m\geqslant M_1(\epsilon,C_1,r)$, if*

$$\frac{1}{m}H(\pi_{(g_{\mathbf{U}(n)}^{-1}V)^\perp}\mu,\mathcal{Q}_m)\geqslant\alpha,$$

then the decomposition is with $(\alpha-\epsilon)$-entropy concentration at the scale $(m,\chi_1(g_{\mathbf{U}(n)}))$.

Proof. Recall the decomposition of $\pi_{V^\perp}\circ g_{\mathbf{U}(n)}$ in Eq. (4.9).

Due to Eq. (4.10) and Lemma 2.25, the measure $\pi_{V,\mathbf{U}(n)}\mu$ is uniform Hölder regular: for any $x\in\mathbb{P}(V^\perp)$ and any $R>0$, we have

$$(\pi_{V,\mathbf{U}(n)}\mu)(B(x,R))\leqslant CC_1^\beta R^\beta.$$

Applying Lemma 4.9 to the measure $\pi_{V,\mathbf{U}(n)}\mu$, we have for any $0<r\leqslant\frac{1}{2}$, the r-attracting decomposition $(h_{V,\mathbf{U}(n)},\pi_{V,\mathbf{U}(n)}\mu)$ is a $(q^{-\chi_1(h_{V,\mathbf{U}(n)})}/r^2, CC_1^\beta r^\beta)$-decomposition of $\pi_{V^\perp}g_{\mathbf{U}(n)}\mu$. We choose r such that $CC_1^\beta r^\beta\leqslant\delta$ and and replace $\chi_1(h_{V,\mathbf{U}(n)})$ by $\chi_1(g_{\mathbf{U}(n)})$ with bounded loss which is due to Eq. (4.10). Therefore, the decomposition is a $(C_2\cdot q^{-\chi_1(g_{\mathbf{U}(n)})},\delta)$-decompostion of $\pi_{V,\mathbf{U}(n)}\mu$ with $C_2=C_2(C_1,r)$.

For the second statement, due to Eq. (4.10) and Lemma B.1, the map $\pi(g_{\mathbf{U}(n)}^{-1}V,V^\perp,(g_{\mathbf{U}(n)}^{-1}V)^\perp)$ has scale 1 with distortion C_1. Hence, by Lemma 2.35, we obtain

$$\frac{1}{m}H(\pi_{V,\mathbf{U}(n)}\mu,\mathcal{Q}_m)\geqslant\alpha-O(\frac{\log C_1}{m}).$$

By taking m large, we can use Lemma 4.10 to obtain the entropy concentration at scale $(m,\chi_1(h_{V,\mathbf{U}(n)}))$ and use $|\chi_1(h_{V,\mathbf{U}(n)})-\chi_1(g_{\mathbf{U}(n)})|\leqslant\log C_1$ to change the scale and conclude. □

Now we show that the good part of projections of typical cylinders have high entropy at smaller scales, i.e. the good part version of Lemma 3.15 of [BHR19].

Lemma 4.14. *For any $\epsilon>0$ and $\delta>0$, there exist $r_0=r_0(\epsilon,\delta)$, $C_0=C_0(\epsilon,\delta)\geqslant 1$ such that for $m\geqslant M_0(\epsilon,C_0)$, $n\geqslant N_0(\epsilon,m)$, we have*

$$\inf_{V\in\mathbb{P}(\mathbb{R}^3)}\mathbb{P}\left\{\begin{array}{l}\mathbf{U}(n):\ \textit{the } r_0\textit{-attracting decomposition of } (h_{V,\mathbf{U}(n)},\pi_{V,\mathbf{U}(n)}\mu)\\ \textit{is a } (C_0\cdot q^{-\chi(g_{\mathbf{U}(n)})},\delta)\textit{-decomposition of } \pi_{V^\perp}g_{\mathbf{U}(n)}\mu\\ \textit{with } (\alpha-\epsilon)\textit{-entropy concentration at the scale } (m,\chi_1(g_{\mathbf{U}(n)}))\end{array}\right\}>1-\epsilon.$$

Proof. To prove Lemma 4.14, it suffices to estimate the probability of $\mathbf{U}(n)$ satisfying the conditions in Lemma 4.13. We choose ϵ_0, δ_0 much smaller than given ϵ, δ. By Lemma 4.12, there exist $C_1(\epsilon_0)$ and $N_1(\epsilon_0)$ such that for any $n \geqslant N_1(\epsilon_0)$, the probability of $\mathbf{U}(n)$ belonging to the set $X_{\mathbf{U},n,V} = X_{\mathbf{U},n,V}(C_1(\epsilon_0), \epsilon_0)$, which is the probability of $\mathbf{U}(n)$ satisfying the conditions in (4.10), is greater than $1 - \epsilon_0$.

Moreover, Lemma 2.39 yields that for any $m \geqslant M_2(\epsilon_0)$ and any $n \geqslant N_2(\epsilon_0, m)$, the probability of $\mathbf{U}(n)$ satisfying

$$\frac{1}{m} H(\pi_{(g_{\mathbf{U}(n)}^{-1} V)^\perp} \mu, \mathcal{Q}_m) \geqslant \alpha - \epsilon_0 \tag{4.12}$$

is greater than $1 - \epsilon_0$. Consequently, letting $r = r(\epsilon_0, C_1, \delta_0)$ be the constant given in Lemma 4.13, for any $m \geqslant \max\{M_1(\epsilon_0, C_1, r), M_2(\epsilon_0)\}$ and any $n \geqslant \max\{N_1(\epsilon_0), N_2(\epsilon_0, m)\}$, the probability of $\mathbf{U}(n)$ satisfying the conditions of Lemma 4.13 is greater than $1 - 2\epsilon_0$. □

We can make a similar statement for $\pi_{V^\perp} g_{\mathbf{I}(n)} \mu$,

Lemma 4.15. *For any $\epsilon > 0$ and $\delta > 0$, there exist $r_1 = r(\epsilon, \delta), C_1 = C_1(\epsilon, \delta) \geqslant 1$ such that for $m \geqslant M_1(\epsilon, C_1)$, $n \geqslant N_1(\epsilon, m)$, we have*

$$\inf_{V \in \mathbb{P}(\mathbb{R}^3)} \mathbb{P}_{1 \leqslant i \leqslant n} \left\{ \begin{array}{l} \mathbf{I}(i) : \textit{ the } r_1\textit{-attracting decomposition of } (h_{V,\mathbf{I}(i)}, \pi_{V,\mathbf{I}(i)}\mu) \\ \textit{is a } (C_1 \cdot q^{-i}, \delta)\textit{-decomposition of } \pi_{V^\perp} g_{\mathbf{I}(i)} \mu \\ \textit{with } (\alpha - \epsilon)\textit{-entropy concentration at the scale } (m, i) \end{array} \right\} > 1 - \epsilon. \tag{4.13}$$

Proof. Recall $C_I > 1$ from Lemma 2.33. Given any $\epsilon > 0$ and $\delta > 0$, let $\epsilon_0 = \epsilon / C_I$, $r_0 = r_0(\epsilon_0/2, \delta)$, $C_0 = C_0(\epsilon_0/2, \delta)$ and $M_0(\epsilon_0/2, C_0)$ be the constants given in Lemma 4.14. For any $m \geqslant M_0$ and any $V \in \mathbb{P}(\mathbb{R}^3)$, we define the subset $\mathcal{U}_V$ of the set of words Λ^* as follows:

$$\mathcal{U}_V := \left\{ \begin{array}{l} \mathbf{i} \in \Lambda^* : \text{ the } r_0\text{-attracting decomposition of } (h_{V,\mathbf{i}}, \pi_{V,\mathbf{i}}\mu) \\ \text{is a } (C_0 \cdot q^{-\chi_1(g_\mathbf{i})}, \delta)\text{-decomposition of } \pi_{V^\perp} g_\mathbf{i} \mu \\ \text{with } (\alpha - \epsilon_0/2)\text{-entropy concentration at the scale } \ (m, \chi_1(g_\mathbf{i})) \end{array} \right\}.$$

By Lemma 4.14, for any $n \geqslant N_0(\epsilon_0/2, m)$, we have

$$\inf_{V \in \mathbb{P}(\mathbb{R}^3)} \mathbb{P}\{\mathbf{U}(n) : \mathbf{U}(n) \in \mathcal{U}_V\} > 1 - \epsilon_0/2.$$

Hence, it follows that for any sufficiently large n depending on ϵ_0,

$$\inf_{V \in \mathbb{P}(\mathbb{R}^3)} \mathbb{P}_{1 \leqslant i \leqslant n}\{\mathbf{U}(i) : \mathbf{U}(i) \in \mathcal{U}_V\} > 1 - \epsilon_0.$$

Lemma 2.33 implies for any $n \geqslant N(\epsilon_0, m)$,

$$\inf_{V \in \mathbb{P}(\mathbb{R}^3)} \mathbb{P}_{1 \leqslant i \leqslant n}\{\mathbf{I}(i) : \mathbf{I}(i) \in \mathcal{U}_V\} > 1 - C_I \epsilon_0 = 1 - \epsilon.$$

For any $V \in \mathbb{P}(\mathbb{R}^3)$ and any $1 \leqslant i \leqslant n$, if $\mathbf{I}(i) \in \mathcal{U}_V$, we actually have that the r_0-attracting decomposition of $(h_{V,\mathbf{I}(i)}, \pi_{V,\mathbf{I}(i)}\mu)$ is a $(C_1 \cdot q^{-i}, \delta)$-decomposition of $\pi_{V^\perp} g_{\mathbf{I}(i)} \mu$ with $(\alpha - \epsilon_0)$-entropy concentration at the scale (m, i). Here is how we obtain the last statement: Due to Eq. (2.68), we replace $C_0 \cdot q^{-\chi_1(g_{\mathbf{I}(i)})}$ by $C_1 \cdot q^{-i}$, and replace the scale $(m, \chi_1(g_{\mathbf{I}(i)}))$ by (m, i) by choosing m large enough if necessary. □

In the later proof, we need an immediate corollary of Lemma 4.15, which relaxes the condition on the parameter r.

Lemma 4.16. *For any $\epsilon > 0$ and $\delta > 0$, there exist $r_1 = r(\epsilon, \delta), C_1 = C_1(\epsilon, \delta) \geqslant 1$ such that for $m \geqslant M_1(\epsilon, C_1)$, $n \geqslant N_1(\epsilon, m)$ and $r < r_1$, we have*

$$\inf_{V \in \mathbb{P}(\mathbb{R}^3)} \mathbb{P}_{1 \leqslant i \leqslant n} \left\{ \begin{array}{l} \mathbf{I}(i) : \textit{ the } r\textit{-attracting part } \mathfrak{m}_r \textit{ of } (h_{V,\mathbf{I}(i)}, \pi_{V,\mathbf{I}(i)}\mu) \\ \textit{satisfies } \frac{1}{m} H(\mathfrak{m}_r, \mathcal{Q}_{m+i}) \geqslant \alpha - \epsilon - 2\delta \end{array} \right\} > 1 - \epsilon. \tag{4.14}$$

Proof. We have the following general result.

Claim. *Let $\beta, \delta \in (0,1)$ and $m, i \in \mathbb{N}$. Let $\tau_1, \tau_2 \in \mathbf{P}(\mathbb{P}(\mathbb{R}^2))$ such that τ_1 is a restriction of τ_2 with $\tau_2((\operatorname{supp}\tau_1)^c) \leqslant \delta$. If the measure τ_1 satisfies*

$$\frac{1}{m}H(\tau_1, \mathcal{Q}_{m+i}) \geqslant \beta,$$

then

$$\frac{1}{m}H(\tau_2, \mathcal{Q}_{m+i}) \geqslant \beta - \delta.$$

Due to the hypothesis, we can write $\tau_2 = (1-\eta)\tau_1 + \eta\tau'$ with $\eta \leqslant \delta$. Then the claim follows directly from the concavity of entropy Eq. (2.77).

For each $\mathbf{I}(i)$ belonging to the set in the formula Eq. (4.13), we consider $\tau_1 = \mathfrak{m}_{r_1}$ (the r_1-attracting part) and $\tau_2 = \mathfrak{m}_r$. Due to $r \leqslant r_1$, the measure τ_1 is a restriction of τ_2. Since the r_1-attracting decomposition of $(h_{V,\mathbf{I}(i)}, \pi_{V,\mathbf{I}(i)}\mu)$is an $(C_1 \cdot q^{-i}, \delta)$-decomposition of $\pi_{V^\perp} g_{\mathbf{I}(i)}\mu$, we have $\tau_2((\operatorname{supp}\tau_1)^c) \leqslant \frac{\tau(b(h^-,r)^c)}{\tau(b(h^-,r_1))} \leqslant \frac{\delta}{(1-\delta)} \leqslant 2\delta$. As the r_1-decomposition has $(\alpha-\epsilon)$-entropy concentration, that is $\frac{1}{m}H(\tau_1, \mathcal{Q}_{m+i}) \geqslant \alpha - \epsilon$, we can apply the claim to τ_1 and τ_2 to finish the proof. ☐

4.5 Proof of Proposition 4.2

We are ready to prove Proposition 4.2. In the following proposition, we fix a point $V \in \mathbb{P}(\mathbb{R}^3)$ such that the projection measure $\pi_{V^\perp}\mu$ has entropy dimension α.

Proposition 4.17. *For any $\epsilon > 0$, $m \geqslant M(\epsilon)$ and $n \geqslant N(\epsilon, m, V)$, the projection measure $\pi_{V^\perp}\mu$ is (α, ϵ, m)-entropy porous from scale 1 to n.*

Proof. Write $\tau = \pi_{V^\perp}\mu$. Let $\epsilon > 0$ be as given. We fix a few other constants:

- let $\delta'(\epsilon/2)$ be the constant given in Lemma 4.3 for $\epsilon/2$ and $\delta''(\epsilon)$ be the one given in Lemma 4.4 for ϵ. Fix $\delta < \min\{\delta'(\epsilon/2), \delta''(\epsilon)\}$.
- choose constants ϵ_0, δ_0 much smaller than ϵ, δ respectively.
- let $C_1 = C_1(\epsilon_0, \delta_0)$ be the constant given in Lemma 4.15. By Proposition 3.4, there exist $l_0 = l_0(\delta_0), k_0$ which are independent of V such that for any $k \geqslant k_0$ large we have

$$\tau\left(\cup_n B\left(\frac{n}{q^k}, \frac{1}{q^{k+l_0}}\right)\right) < \delta_0. \tag{4.15}$$

  We fix an l such that $C_1 \cdot q^{-l} \leqslant q^{-l_0}$.

We want to apply Lemma 4.4 to τ. We first construct convex combinations of τ which satisfy the conditions in Lemma 4.3.

By Lemma 4.15, for $m \geqslant M_1(\epsilon_0, C_1)$, $n \geqslant N_1(\epsilon_0, m)$, we have

$$\mathbb{P}_{1\leqslant i\leqslant n}\left\{\begin{array}{l}\mathbf{I}(i+l) : \pi_{V^\perp} g_{\mathbf{I}(i+l)}\mu \text{ has a } (C_1 \cdot q^{-(i+l)}, \delta_0)\text{-decomposition} \\ \text{with } (\alpha-\epsilon_0)\text{-entropy concentration at the scale } (m, i+l)\end{array}\right\} > 1 - 2\epsilon_0. \tag{4.16}$$

We replace the index i in Eq. (4.13) by $i+l$ in Eq. (4.16). This is possible by taking n large enough. Applying Markov's inequality, we find that for at least a $(1-\sqrt{2\epsilon_0})$-fraction of levels $1 \leqslant k \leqslant n$, we have

$$\mathbb{P}_{i=k}\left\{\begin{array}{l}\mathbf{I}(i+l) : \pi_{V^\perp} g_{\mathbf{I}(i+l)}\mu \text{ has a } (C_1 \cdot q^{-(i+l)}, \delta_0)\text{-decomposition} \\ \text{with } (\alpha-\epsilon_0)\text{-entropy concentration at the scale } (m, i+l)\end{array}\right\} > 1 - \sqrt{2\epsilon_0}. \tag{4.17}$$

Since we can take n arbitrarily large, we can assume for at least a $(1-2\sqrt{\epsilon_0})$ fraction of levels $1 \leqslant k \leqslant n$ satisfies both (4.17) and (4.15). In the following discussion, we fix any such k.

Note that we have $\tau = \mathbb{E}_{i=k}(\pi_{V^\perp} g_{\mathbf{I}(i+l)}\mu)$ due to Eq. (2.72). As a result, we can use Eq. (4.17) to construct a convex combination of τ:

$$\tau = \sum_{1\leqslant j\leqslant N} p_j \tau_j + p_0\tau_0,$$

where for $1 \leqslant j \leqslant N$, τ_j comes from the good part of $\pi_{V^\perp} g_{\mathbf{I}(k+l)}$ with $\mathbf{I}(k+l)$ belonging to the event in Eq. (4.17); τ_0 is the sum of the bad parts of $\pi_{V^\perp} g_{\mathbf{I}(k+l)}$ with $\mathbf{I}(k+l)$ belonging to the event in Eq. (4.17) and those $\pi_{V^\perp} g_{\mathbf{I}(k+l)}$ with $\mathbf{I}(k+l)$ not belonging to the event in Eq. (4.17). This construction gives $p_0 \leqslant \sqrt{2\epsilon_0} + \delta_0$. We continue to check that this convex combination satisfies the conditions of Lemma 4.3:

1. for $1 \leqslant j \leqslant N$, due to τ_j the good part of $\pi_{V^\perp} g_{\mathbf{I}(k+l)}$ with $\mathbf{I}(k+l)$ in Eq. (4.17), we have

$$\frac{1}{m} H(\tau_j, \mathcal{Q}_{k+m}) \geqslant \frac{1}{m} H(\tau_j, \mathcal{Q}_{k+l+m}) - O(\frac{l}{m}) \geqslant \alpha - 2\epsilon_0,$$

where the second inequality is possible by taking m large enough depending on l, ϵ_0.

2. for $1 \leqslant j \leqslant N$, we have $\mathrm{diam}(\mathrm{supp}\,\tau_j) \leqslant C_1 \cdot q^{-(k+l)} \leqslant q^{-(k+l_0)}$.

3. $\tau\left(\cup_n B\left(\frac{n}{q^k}, \frac{1}{q^{k+l_0}}\right)\right) < \delta_0$ by Eq. (4.15).

We can choose ϵ_0, δ_0 such that $\sqrt{2\epsilon_0} + \delta_0 < \delta$ and $\epsilon_0 < \frac{\epsilon}{2}$, etc. Then Lemma 4.3 implies that

$$\mathbb{P}_{i=k}\left\{\frac{1}{m} H(\tau_{x,i}, \mathcal{Q}_{i+m}) > \alpha - \epsilon\right\} > 1 - \epsilon.$$

Since τ has entropy dimension α, we can take n large such that $|\frac{1}{n} H(\tau, \mathcal{Q}_n) - \alpha| < \delta$ (in view of Lemma 2.38, here is the only place that n depends on V). Therefore, we can apply Lemma 4.4 to τ and conclude that for any large n, τ is (α, ϵ, m)-entropy porous from scale 1 to n. $\square$

Lemma 2.38 states that there exists a μ^--full measure set E_0 such that for every $V \in E_0$, $\pi_{V^\perp}\mu$ has entropy dimension α. Hence Proposition 4.17 holds for every $V \in E_0$. In the following corollary, we get rid of the dependency of n on $V \in \mathbb{P}(\mathbb{R}^3)$.

Corollary 4.18. *For every $\epsilon > 0, m > M_1(\epsilon)$ and $n > N_1(\epsilon, m)$, there exists an open set $E = E(\epsilon, m, n) \subset \mathbb{P}(\mathbb{R}^3)$ of measure $\mu^-(E) > 1 - \epsilon$ such that for every $V \in E$, the projection measure $\pi_{V^\perp}\mu$ is (α, ϵ, m)-entropy porous from scale 1 to n.*

Recall that in Lemma A.1, we fix point $V_0 \in \mathbb{P}(\mathbb{R}^3)$ and let $\mathcal{C} = \mathbb{P}(\mathbb{R}^3) - \{V_0\}$. For each $V \in \mathcal{C}$, the identification between $\mathbb{P}(V^\perp)$ and $\mathbb{P}(\mathbb{R}^2)$, denoted by $\mathcal{I}_V$, is to identify $\pi_{V^\perp}V_0$ with $\mathbb{R}(1,0)$. The maps $\mathcal{I}_V$ are continuous with respect to $V \in \mathcal{C}$, which is part of the reason why we can obtain Corollary 4.18. The proof of Corollary 4.18 relies on Lemma A.2 and Lemma A.3.

Proof of Corollary 4.18. Let $m > M_1(\epsilon) = \max\{M(\epsilon), M_2(\epsilon)\}$ with $M(\epsilon)$ and $M_2(\epsilon)$ given as in Proposition 4.17 and Lemma A.3 respectively. For $n > N(\epsilon, m)$ sufficiently large, the existence of measurable set $E_0(\epsilon, m, n)$ is a direct consequence of Proposition 4.17. It remains to show that we can find an open set $E(\epsilon, m, n)$ with large measure.

Let $V \in E_0$. We find a small open neighborhood $B_{\epsilon,m,n}(V)$ of V in $\mathbb{P}(\mathbb{R}^3)$ that has the following property: for any $W \in B_{\epsilon,m,n}(V)$, identifying $\mathbb{P}(W^\perp)$ with $\mathbb{P}(\mathbb{R}^2)$ by $\mathcal{I}_W$, we have any scale $1 \leqslant i \leqslant n$, any $x \in \mathbb{P}(\mathbb{R}^2)$ and any sub-interval I of $\mathcal{Q}_{i+m}$,

$$|(\pi_{V^\perp}\mu)_{x,i}(I) - (\pi_{W^\perp}\mu)_{x,i}(I)| < q^{-3m} \tag{4.18}$$

The existence of $B_{\epsilon,m,n}(V)$ is guaranteed by Lemma A.2. We continue to use Lemma A.3 to obtain that for any component measure $(\pi_{V^\perp}\mu)_{x,i}$, if $\frac{1}{m}H((\pi_{V^\perp}\mu)_{x,i},\mathcal{Q}_{i+m}) < \alpha+\epsilon$, then $\frac{1}{m}H((\pi_{W^\perp}\mu)_{x,i},\mathcal{Q}_{i+m}) < \alpha+2\epsilon$ for any $W\in B_{\epsilon,m,n}(V)$. Based on these, as $\pi_{V^\perp}\mu$ is (α,ϵ,m)-entropy porous from scale 1 to n, $\pi_{W^\perp}\mu$ is also $(\alpha,2\epsilon,m)$-entropy porous from scale 1 to n for any $W\in B_{\epsilon,m,n}(V)$.

As a result, for each $V\in E_0$, we can find an open neighborhood of V such that the entropy porosity still holds. Take the union as the set $E(\epsilon,m,n)$. $\square$

We are ready to prove the main proposition of this section.

Proof of Proposition 4.2. Fix $\epsilon>0$ be as given. Recall the decompositions given in Eq. (4.9) for any $\spadesuit\in\{\mathbf{U}(n),\mathbf{I}(i)\}$ and $V\in\mathbb{P}(\mathbb{R}^3)$ such that $g_{\spadesuit}^{-1}V\notin V^\perp$.

Let $\epsilon_0=\min\{\epsilon'(\epsilon/10),\epsilon/(2C_I+1)\}$, where $\epsilon'(\epsilon/10)$ is the constant given in Proposition 4.11 for $\epsilon/10$. For $m\geqslant M(\epsilon_0)$ and $k\geqslant K(m,\epsilon_0)$, let $E(\epsilon_0,m,k)$ be the open set in $\mathbb{P}(\mathbb{R}^3)$ given in Corollary 4.18. Using the equidistribution of $g_{\mathbf{U}(n)}^{-1}V$ (Lemma 2.23), we have that for any large n and every $V\in\mathbb{P}(\mathbb{R}^3)$, we have

$$\mathbb{P}\{\mathbf{U}(n):g_{\mathbf{U}(n)}^{-1}V\in E(\epsilon_0,m,k)\}>1-2\epsilon_0.$$

Using Corollary 4.18 and Lemma 2.33, we have for any $n\geqslant N(\epsilon_0,m,k)$

$$\inf_{V\in\mathbb{P}(\mathbb{R}^3)}\mathbb{P}_{1\leqslant i\leqslant n}\{\mathbf{I}(i):\pi_{(g_{\mathbf{I}(i)}^{-1}V)^\perp}\mu \text{ is } (\alpha,\epsilon_0,m)\text{-entropy porous from scale 1 to } k\}>1-2C_I\epsilon_0. \tag{4.19}$$

Meanwhile, Lemma 4.12 states that there exist $C_1(\epsilon_0)$ and $N_1(\epsilon_0)\geqslant 1$ such that for any $n\geqslant N_1(\epsilon_0)$, we have

$$\inf_{V\in\mathbb{P}(\mathbb{R}^3)}\mathbb{P}_{1\leqslant i\leqslant n}\left\{\begin{array}{l}\mathbf{I}(i):d(g_{\mathbf{I}(i)}^{-1}V,V^\perp)>1/C_1\\ |\chi_1(h_{V,\mathbf{I}(i)})-\chi_1(g_{\mathbf{I}(i)})|\leqslant\log C_1\end{array}\right\}>1-\epsilon_0. \tag{4.20}$$

Eqs. (4.19) and (4.20) give that for any $n\geqslant\max\{N(\epsilon_0,m,k),N_1(\epsilon_0)\}$,

$$\inf_{V\in\mathbb{P}(\mathbb{R}^3)}\mathbb{P}\{\mathbf{I}(i):\mathbf{I}(i)\text{ satisfies both the conditions in } Eq.\ (4.19)\text{ and } Eq.\ (4.20)\}>1-(C_I+1)\epsilon_0.$$

To continue, we will use the decomposition

$$\pi_{V^\perp}g_{\mathbf{I}(i)}=h_{V,\mathbf{I}(i)}\circ\pi(g_{\mathbf{I}(i)}^{-1}V,V^\perp,(g_{\mathbf{I}(i)}^{-1}V)^\perp)\circ\pi_{(g_{\mathbf{I}(i)}^{-1}V)^\perp}.$$

Due to $\epsilon_0\leqslant\epsilon'(\epsilon/10)$, it follows from Proposition 4.11 that for $m\geqslant m(\epsilon')$, $n\geqslant n(m,\epsilon')$, any $V\in\mathbb{P}(\mathbb{R}^3)$ and any $\tilde g\in\mathrm{SL}_2(\mathbb{R})$, if $\pi_{(g_{\mathbf{I}(i)}^{-1}V)^\perp}\mu$ is (α,ϵ_0,m)-entropy porous from scale 1 to k, then $\tilde g\pi_{(g_{\mathbf{I}(i)}^{-1}V)^\perp}\mu$ is $(\alpha,\epsilon/10,m)$-entropy porous from scale $\chi_1(\tilde g)+1$ to $\chi_1(\tilde g)+k$. In particular, we consider the words $\mathbf{I}(i)$ satisfying both the conditions in Eq. (4.19) and Eq. (4.20) and let

$$\tilde g=h_{V,\mathbf{I}(i)}\circ\pi(g_{\mathbf{I}(i)}^{-1}V,V^\perp,(g_{\mathbf{I}(i)}^{-1}V)^\perp).$$

We get $\pi_{V^\perp}g_{\mathbf{I}(i)}\mu$ is $(\alpha,\epsilon/10,m)$-entropy porous from scale $\chi_1(\tilde g)+1$ to $\chi_1(\tilde g)+k$ from the decomposition of $\pi_{V^\perp}g_{\mathbf{I}(i)}\mu$. To finish, we estimate $\chi_1(\tilde g)=\chi_1(h_{V,\mathbf{I}(i)}\circ\pi(g_{\mathbf{I}(i)}^{-1}V,V^\perp,(g_{\mathbf{I}(i)}^{-1}V)^\perp))$. Using Eq. (2.68), Eq. (4.20), and Lemma B.1, we have that $\chi_1(\tilde g)$ equals i, up to an additive constant $C_2=C_2(\epsilon_0)$. Therefore, $\pi_{V^\perp}g_{\mathbf{I}(i)}\mu$ is $(\alpha,\epsilon/10+O(C_2/k),m)$-entropy porous from scales i to $i+k$. By taking k large enough and due to $(2C_I+1)\epsilon_0\leqslant\epsilon$, we obtain the statement of Proposition 4.2. $\square$

5 Entropy growth

The main result of this section is Theorem 5.7, which is a projective version of [BHR19, Theorem 4.1]. The main difficulty in the current setting comes from the non-uniform contraction and non-linearity of the projective action. To deal with the non-uniform contraction of the action, we rely on the decomposition introduced in Definition 4.6, which makes it possible to analyze the change of the entropy. To deal with the non-linearity of the action, we repeatedly use Eq. (2.5).

5.1 Entropy growth under convolutions: Euclidean case

For $\alpha, \beta \in \mathbf{P}(\mathbb{R})$, we denote by $\alpha * \beta$ the additive convolution of α, β. For every bounded support $\alpha, \beta \in \mathbf{P}(\mathbb{R})$, we have the following lower bound from concavity of entropy (cf. [Hoc14, Corollary 4.10]):

$$\frac{1}{m} H(\alpha * \beta, \mathcal{Q}_m) \geqslant \frac{1}{m} H(\beta, \mathcal{Q}_m) - O(\frac{1}{m}). \tag{5.1}$$

One may expect that $\frac{1}{m} H(\alpha * \beta, \mathcal{Q}_m)$ is close to $\frac{1}{m} H(\alpha, \mathcal{Q}_m) + \frac{1}{m} H(\beta, \mathcal{Q}_m)$, but in general this is not the case. The following modification of [BHR19, Theorem 4.1] shows a sufficient condition which implies the presence of non-trivial entropy growths. See also Theorem 2.8 in [Hoc14].

Theorem 5.1. *For every $\epsilon \in (0, \frac{1}{10})$, there exists $\delta > 0$ such that for $m \geqslant m(\epsilon, \delta)$, $C \geqslant 1$ and $n > N(\epsilon, \delta, m) \log C$, the following holds.*

Let $k \in \mathbb{N}$ and $\theta, \tau \in \mathbf{P}(\mathbb{R})$, and suppose that

- *θ and τ are supported on intervals of length less than Cq^{-k},*
- *τ is $\left(1 - \frac{5\epsilon}{2}, \frac{\epsilon}{2}, m\right)$-entropy porous from scale k to $k + n$,*
- *$\frac{1}{n} H(\theta, \mathcal{Q}_{k+n}) > 3\epsilon$.*

Then we obtain the growth of the entropy of scale (k, n):

$$\frac{1}{n} H(\theta * \tau, \mathcal{Q}_{k+n}) \geqslant \frac{1}{n} H(\tau, \mathcal{Q}_{k+n}) + \delta.$$

Proof. [BHR19, Theorem 4.1] was stated for the case when $C = 1$. Set $k' = k - \log C$ and $\epsilon' = 2\epsilon$. Choose n large enough so that $\frac{2 \log C}{n} < \epsilon$. Then k' and θ, τ satisfy the hypothesis [BHR19, Theorem 4.1]. More precisely, we have

- θ and τ are supported on intervals of length less than $q^{-k'}$,
- τ is $\left(1 - \epsilon', \frac{\epsilon'}{2}, m\right)$-entropy porous from scale k' to $k' + n$,
- $\frac{1}{n} H(\theta, \mathcal{Q}_{k'+n}) > \epsilon'$.

Hence, there exists $\delta = \delta(\epsilon') > 0$ such that

$$\frac{1}{n} H(\theta * \tau, \mathcal{Q}_{k'+n}) \geqslant \frac{1}{n} H(\tau, \mathcal{Q}_{k'+n}) + \delta.$$

It follows that

$$\frac{1}{n} H(\theta * \tau, \mathcal{Q}_{k+n}) \geqslant \frac{1}{n} H(\theta * \tau, \mathcal{Q}_{k'+n}) \geqslant \frac{1}{n} H(\tau, \mathcal{Q}_{k'+n}) + \delta \geqslant \frac{1}{n} H(\tau, \mathcal{Q}_{k+n}) + \delta - \frac{\log C}{n}.$$

By taking n large enough, we obtain the result. $\square$

5.2 Convolution of measures on $\mathbb{P}(\mathbb{R}^3)$ and measures on L: preparation

Recall the subgroup $L = L_{E_1} \simeq \mathrm{SL}_2(\mathbb{R}) \ltimes \mathbb{R}^2$ in $\mathrm{SL}_3(\mathbb{R})$ defined in the UL-decomposition for $E_1 \in \mathbb{P}(\mathbb{R}^3)$ (see Section 2.3). We are interested in the maps $\pi_{E_1^\perp}\ell$ from $\mathbb{P}(\mathbb{R}^3)$ to $\mathbb{P}(\mathbb{R}^2)$ with $\ell \in L$. Note that for each $\ell \in L$, $\pi_{E_1^\perp}\ell$ is not defined at $\ell^{-1}E_1$. We define a bad locus of $L \times \mathbb{P}(\mathbb{R}^3)$ by

$$\mathcal{B} := \{(\ell, \ell^{-1}E_1) : \ell \in L\}, \tag{5.2}$$

which is a closed subset of $L \times \mathbb{P}(\mathbb{R}^3)$.

Let $\theta \in \mathbf{P}(L)$. Let $\tau \in \mathbf{P}(\mathbb{P}(\mathbb{R}^3))$ be atom-free (for example the stationary measure μ on $\mathbb{P}(\mathbb{R}^3)$ is atom-free). Let $[\theta.\tau]$ be the projection of the convolution measure on $\mathbb{P}(\mathbb{R}^2)$. More precisely, let $[\ell x] = \pi_{E_1^\perp}\ell x$, then the measure $[\theta.\tau]$ is defined as follows: for any continuous function f on $\mathbb{P}(\mathbb{R}^2)$,

$$\int_{\mathbb{P}(\mathbb{R}^2)} f(y)\mathrm{d}[\theta.\tau](y) := \int_{\mathbb{P}(\mathbb{R}^3)}\int_L f([\ell x])\mathrm{d}\theta(\ell)\mathrm{d}\tau(x).$$

Since we suppose that τ is atom-free, we have $\theta \otimes \tau(\mathcal{B}) = 0$. The convolution is well-defined.

5.2.1 Linearization of projection of convolution measures

We first need a linearization lemma. Due to non-linearity, for later application of the entropy growth argument, we need to locally replace convolution in the sense of L-action by the additive convolution on $\mathbb{R}$, without large cost on entropy.

Identifying $\mathbb{P}(\mathbb{R}^2)$ with $\mathbb{R}/\mathbb{Z}$, for $v \in \mathbb{P}(\mathbb{R}^2)$, let T_v be the translation on $\mathbb{P}(\mathbb{R}^2)$ given by $T_v(x) = x - v$. For any $\theta \in \mathbf{P}(L)$ and any $x_0 \in \mathbb{P}(\mathbb{R}^3)$, we set $[\theta.x_0] := [\theta.\delta_{x_0}]$, which is the projection of the convolution of θ with Dirac measure δ_{x_0}.

Lemma 5.2 (Linearization of measures)**.** *For any $\epsilon > 0$, $0 < r < 1/2$, $C_1 > 20$, $k > K(\epsilon)$, $t \geqslant 1$, we have the followings. Let $\ell_0 \in L(t, C_1)$ ($\subset L$ defined in Eq. (2.66)). For any $x_0 \in b(f_{\ell_0}, r/2)$, any $\rho < q^{-k}(r/C_1)^8$, any measure $\theta \in \mathbf{P}(L)$ with $\operatorname{supp}\theta \subset B(\ell_0, \rho)$ and $\tau \in \mathbf{P}(\mathbb{P}(\mathbb{R}^3))$ with $\operatorname{supp}\tau \subset B(x_0, \rho) \subset b(f_{\ell_0}, r)$, we have*

$$\begin{aligned}&\left|\frac{1}{k}H([\theta.\tau], \mathcal{Q}_{k+t-\log\rho}) - \frac{1}{k}H\left((S_{t(\ell_0,x_0)}T_{[\ell_0(x_0)]}[\theta.x_0]) * (\pi_{(\ell_0^{-1}E_1)^\perp}\tau), \mathcal{Q}_{k-\log\rho}\right)\right| \\ <&\epsilon + \frac{4\log(C_1/r) + O(1)}{k},\end{aligned} \tag{5.3}$$

with $t(\ell_0, x_0)$ the number given as in Lemma 2.21 and

$$|t(\ell_0, x_0) - t| \leqslant 8\log(C_1/r).$$

Proof. Let $C_2 = 1/r$ and $y = \pi_{(\ell_0^{-1}E_1)^\perp}x_0$. Write $\ell_0 = \begin{pmatrix} 1 & 0 \\ n_0 & h_0\end{pmatrix} \in L$.

Under the assumption of this lemma, the condition of Lemma 2.21 is satisfied. Using Lemma 2.21 and the definition of the set $L(t, C_1)$, we obtain

$$|t(\ell_0, x_0) - t| \leqslant |t(\ell_0, x_0) - 2\log\|h_0\|| + |2\log\|h_0\| - t| \leqslant 8\log(C_1/r).$$

Step 1. We first prove

$$|\frac{1}{k}H([\theta.\tau], \mathcal{Q}_{k+t-\log\rho}) - \frac{1}{k}H([\theta.x_0] * (S_{-t(\ell_0,x_0)}T_y\pi_{(\ell_0^{-1}E_1)^\perp}\tau), \mathcal{Q}_{k+t-\log\rho})| < \epsilon.$$

Let f, h_{z_0} with $z_0 = (\ell_0, x_0)$ be two maps from $L \times \mathbb{P}(\mathbb{R}^3) - \mathcal{B}$ to $\mathbb{P}(\mathbb{R}^2)$ given by

$$f(\ell, x) = [\ell(x)],\ h_{z_0}(\ell, x) = [\ell(x_0)] + S_{-t(\ell_0,x_0)}(\pi_{(\ell_0^{-1}E_1)^\perp}x - y).$$

The push-forward measure $h_{z_0}(\ell, x)(\theta, \tau)$ on $\mathbb{P}(\mathbb{R}^2)$ is exactly the convolution measure $[\theta.x_0] * (S_{-t(\ell_0,x_0)} T_y \pi_{(\ell_0^{-1}E_1)^\perp} \tau)$.

Due to Lemma 2.21 and $\rho \leqslant q^{-k}(r/C_1)^8$, we have

$$\left\| (f - h_{z_0})|_{B(\ell_0,\rho)\times B(x_0,\rho)} \right\| \leqslant 2C_L (C_1/r)^8 \rho^2 \|h_0\|^{-2} \leqslant 2C_L C_p^2 \rho q^{-k-t}.$$

Here C_L, C_p are large fixed constants and we take $K(\epsilon) \gg \log(C_L C_p^2)$. Using the continuity of entropy Eq. (2.76) and Eq. (2.75), we conclude Step 1.

Step 2. Consider the following measure on $\mathbb{P}(\mathbb{R}^2)$

$$\nu = T_{[\ell_0(x_0)]}[\theta.x_0] * S_{-t(\ell_0,x_0)} T_y (\pi_{(\ell_0^{-1}E_1)^\perp} \tau).$$

Observe that the support of $[\theta.x_0]$ is near $[\ell_0(x_0)]$. So the support of $T_{[\ell_0(x_0)]}[\theta.x_0]$ is near zero. Originally, S_t is a scaling map defined on $\mathbb{R}$. Comparing ν with the target measure that appears in Eq. (5.3), the strategy is to show that the support of ν is small enough. Then the restriction map $S_t : \text{supp } \nu \to S_t(\text{supp } \nu)$ is a diffeomorphism, and hence we can use the invariance of entropy under the scaling map to change the scale in the partition

$$H(\nu, \mathcal{Q}_{t+k}) = H(S_t \nu, \mathcal{Q}_k) + O(1). \tag{5.4}$$

We estimate the support of ν.

- ($\text{supp}[\theta.x_0]$) Observe that $\text{supp}[\theta.x_0] \subset [\text{supp}\,\theta.\{x_0\}]$. Let $E = \text{supp}\,\theta$ and $F = \{x_0\}$. Under the assumption of this lemma, we can apply Lemma 2.19 to $E = \text{supp}\,\theta$ and $F = \{x_0\}$ and obtain

$$\text{diam}(\text{supp}[\theta.x_0]) \leqslant 16 C_L \frac{C_1}{r^2} \text{diam}(\text{supp}\,\theta)/\|h_0\|^2 \leqslant 16 C_L C_p^2 \frac{C_1}{r^2} \rho q^{-t}. \tag{5.5}$$

- ($\text{supp}(\pi_{(\ell_0^{-1}E)^\perp}\tau)$) Note that $\text{supp}\,\tau \subset b(f_{\ell_0}, r)$. Then for any $y, z \in \text{supp}\,\tau$, using Eq. (2.18), we obtain $d(y, \ell_0^{-1}E_1) > r/C_1$, $\ d(z, \ell_0^{-1}E_1) > r/C_1$. So by Lemma B.2(2),

$$d(\pi_{(\ell_0^{-1}E)^\perp} y, \pi_{(\ell_0^{-1}E)^\perp} z) \leqslant \frac{C_1}{r} d(y, z).$$

  Therefore

$$\text{diam}(\text{supp}(\pi_{(\ell_0^{-1}E)^\perp}\tau)) \leqslant \text{diam}(\tau) C_1/r \leqslant C_1 \rho / r.$$

- We want to deduce from the above two diameter estimates that

$$\text{diam}\,\nu \leqslant q^{-t(\ell_0,x_0)}/2,$$

  which is sufficient to have

$$16 C_L C_p^2 \frac{C_1}{r^2} \rho q^{-t} + q^{-t(\ell_0,x_0)} C_1 \rho / r \leqslant q^{-t(\ell_0,x_0)}/2. \tag{5.6}$$

  This is possible because we have $|t(\ell_0, x_0) - t| \leqslant 4\log(C_1/r)$ and $\rho \leqslant q^{-k}(C_1/r)^8$. Using the assumption $k \geqslant K(\epsilon) \geqslant \log(64 C_L C_p^2)$, we get Eq. (5.6).

Therefore, we obtain the restriction of the scaling map $S_{t(\ell_0,x_0)} : \text{supp}\,\nu \to S_{t(\ell_0,x_0)}(\text{supp}\,\nu)$ is a diffeomorphism.

Recall that we use the convention $\mathcal{Q}_t = \mathcal{Q}_{[t]}$ for the partition. If $t(\ell_0, x_0)$ is an integer, we can apply directly Eq. (5.4) and obtain

$$H(S_{t(\ell_0,x_0)}\nu, \mathcal{Q}_{k-\log\rho+t-t(\ell_0,x_0)}) = H(\nu, \mathcal{Q}_{k-\log\rho+t}) + O(1). \tag{5.7}$$

Otherwise, we can use [HR21, (4.1)] and still obtain Eq. (5.7).

Comparing $S_{t(\ell_0,x_0)}\nu$ with the target measure $S_{t(\ell_0,x_0)}T_{[\ell_0(x_0)]}[\theta.x_0]*(\pi_{(\ell_0^{-1}E_1)^\perp}\tau)$ in Eq. (5.3), the only difference is the shift map T_y. But the shift map only changes the entropy by $O(1)$. So combining with Step 1, we obtain

$$\left|\frac{1}{k}H([\theta.\tau],\mathcal{Q}_{k+t-\log\rho})-\frac{1}{k}H\left((S_{t(\ell_0,x_0)}T_{[\ell_0(x_0)]}[\theta.x_0])*(\pi_{(\ell_0^{-1}E_1)^\perp}\tau),\mathcal{Q}_{k-\log\rho+t-t(\ell_0,x_0)}\right)\right|<\epsilon+\frac{O(1)}{k}.$$

To finish, we replace $\mathcal{Q}_{k-\log\rho+t-t(\ell_0,x_0)}$ by $\mathcal{Q}_{k-\log\rho}$, and it brings the error term $4\log(C_1/r)/k$ by Eq. (2.75). □

5.2.2 Decompose projection of convolution measures

Proposition 5.3 uses the concavity of entropy to decompose the projection of the convolution into measures with small support without losing too much entropy. It is called the multi-scale formula for entropy.

We recall and introduce a few notions.

- For each $j\in\mathbb{N}$, let $\mu_{\varphi(j,\mathbf{i})}$ be the random measure $g_{\mathbf{i}}\mu$ with $\mathbf{i}$ in the symbolic space $\mathbf{I}(j)$ given as in Definition 2.32.
- In Definition 4.7, we introduced the r-attracting decomposition of the pair $(g_{\mathbf{i}},\mu)$:

  $$\mu_{\varphi(j,\mathbf{i})}=g_{\mathbf{i}}\mu:=\epsilon\cdot g_{\mathbf{i}}(\mu_{b(g_{\mathbf{i}}^-,r)^c})+(1-\epsilon)\cdot g_{\mathbf{i}}(\mu_{b(g_{\mathbf{i}}^-,r)}),\tag{5.8}$$

  where $\epsilon=\mu(b(g_{\mathbf{i}}^-,r)^c)$. Denote

  $$(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}:=g_{\mathbf{i}}(\mu_{b(g_{\mathbf{i}}^-,r)}),\ \ (\mu_{\varphi(j,\mathbf{i})})_{\mathbf{II}}:=g_{\mathbf{i}}(\mu_{b(g_{\mathbf{i}}^-,r)^c}).$$

  Lemma 4.9 states that the decomposition of $\mu_{\varphi(j,\mathbf{i})}$ in Eq. (5.8) gives a decomposition satisfying Definition 4.6.
- Let

  $$\mathcal{E}_0:=\{(\ell,\mathbf{i})\in L\times\Lambda^*:\ V_{g_{\mathbf{i}}}^+\in b(f_\ell,r/4)\},\tag{5.9}$$

  where $b(f_\ell,r)\subset\mathbb{P}(\mathbb{R}^3)$ is given as in Definition 2.13.

The constant r is the parameter for both $\mathcal{E}_0$ where $b(f_\ell,r)$ is used to define $\mathcal{E}_0$, and r-attracting decomposition in Eq. (5.8).

Proposition 5.3. *For any $C_1>1$ and $t\geqslant 1$, let $\theta\in\mathbf{P}(L)$ be such that* $\operatorname{supp}(\theta)\subset L(t,C_1)$. *Then any pair of natural numbers $n\geqslant k$, we consider the r-attracting decomposition of $(g_{\mathbf{i}},\mu)$ with $r=q^{-\sqrt{k}/10}$. We have*

$$\frac{1}{n}H([\theta.\mu],\mathcal{Q}_{t+n})\geqslant\mathbb{E}_{1\leqslant j\leqslant n,\ell\sim\theta}\left(\frac{1}{k}H([\theta_{\ell,j}.(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}],\mathcal{Q}_{t+j+k}),\ \mathcal{E}_0\right)-O\left(r^\beta+\frac{k}{n}+\frac{\log(C_1/r^4)}{k}\right),$$

where $\mathbb{E}(\ ,\mathcal{E}_0)$ is the restriction to the subset $\mathcal{E}_0$ and $\beta>0$ is from Lemma 2.24.

Proof. The proof consists of two steps. Firstly, we use the identity

$$[\theta.\mu]=\mathbb{E}_{j=i}\left([\theta_{\ell,j}.\mu_{\varphi(j,\mathbf{i})}]\right),$$

which is valid for any $i \in \mathbb{N}$. Let s be the integer part of n/k. Similar to the proof of Lemma 4.3 in [BHR19], we have for for every residue $0 \leqslant p < k$,

$$
\begin{aligned}
H([\theta.\mu], \mathcal{Q}_{t+n}) &= \sum_{m=0}^{s-2} H\left([\theta.\mu], \mathcal{Q}_{t+(m+1)k+p} \mid \mathcal{Q}_{t+mk+p}\right) \\
&\quad + H([\theta.\mu], \mathcal{Q}_{t+p}) + H([\theta.\mu], \mathcal{Q}_{t+n}|\mathcal{Q}_{t+(s-1)k+p}) \\
&\geqslant \sum_{m=0}^{s-2} H\left([\theta.\mu], \mathcal{Q}_{t+(m+1)k+p} \mid \mathcal{Q}_{t+mk+p}\right) \\
&\geqslant \sum_{m=0}^{s-2} \mathbb{E}_{j=mk+p}\left(H\left([\theta_{\ell,j}.\mu_{\varphi(j,\mathbf{i})}], \mathcal{Q}_{t+k+j} \mid \mathcal{Q}_{t+j}\right)\right).
\end{aligned}
\tag{5.10}
$$

where in the last line we used concavity of entropy. Now by averaging over $0 \leq p \leq k-1$ and dividing by n, and recalling that $s/n \leqslant 1/k$, we get

$$
\begin{aligned}
\frac{1}{n}H([\theta.\mu], \mathcal{Q}_{t+n}) &\geqslant \frac{1}{n}\sum_{p=0}^{k-1}\sum_{m=0}^{s-2} \mathbb{E}_{j=mk+p}\left(\frac{1}{k}H\left([\theta_{\ell,j}.\mu_{\varphi(j,\mathbf{i})}], \mathcal{Q}_{t+k+j} \mid \mathcal{Q}_{t+j}\right)\right) - O(\frac{s}{n}) \\
&= \mathbb{E}_{1\leqslant j\leqslant n}\left(\frac{1}{k}H\left([\theta_{\ell,j}.\mu_{\varphi(j,\mathbf{i})}], \mathcal{Q}_{t+k+j} \mid \mathcal{Q}_{t+j}\right)\right) - O(\frac{k}{n}+\frac{1}{k}),
\end{aligned}
\tag{5.11}
$$

Secondly, we drop the conditional part $\mathcal{Q}_{t+j}$ using the support of measures $\theta_{\ell,j}.\mu_{\varphi(j,\mathbf{i})}$, which is similar to [HS17, Lemma 5.3]. We restrict to $\mathcal{E}_0$ and have

$$
\mathbb{E}_{1\leqslant j\leqslant n}\left(\frac{1}{k}H([\theta_{\ell,j}.\mu_{\varphi(j,\mathbf{i})}], \mathcal{Q}_{t+j+k}|\mathcal{Q}_{t+j})\right) \geqslant \mathbb{E}_{1\leqslant j\leqslant n}\left(\frac{1}{k}H([\theta_{\ell,j}.\mu_{\varphi(j,\mathbf{i})}], \mathcal{Q}_{t+j+k}|\mathcal{Q}_{t+j}),\ \mathcal{E}_0\right)
$$

Recall from Eq. (5.8) the r-attracting decomposition of $\mu_{\varphi(j,\mathbf{i})}$:

$$
\begin{aligned}
\mu_{\varphi(j,\mathbf{i})} &= \epsilon \cdot g_{\mathbf{i}}(\mu_{b(g_{\mathbf{i}}^-,r)^c}) + (1-\epsilon)\cdot g_{\mathbf{i}}(\mu_{b(g_{\mathbf{i}}^-,r)}) \\
&= \epsilon \cdot (\mu_{\varphi(j,\mathbf{i})})_{\mathbf{II}} + (1-\epsilon)\cdot(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}.
\end{aligned}
\tag{5.12}
$$

We have $\epsilon = \mu(b(g_{\mathbf{i}}^-, r)^c) \leqslant Cr^\beta$ with the inequality due to Lemma 2.24.

By the definition of $\mathbf{I}(j)$ (Eq. (2.68)), we have $\chi_1(g_{\mathbf{i}}) \geqslant j$ for $\mathbf{i} \in \mathbf{I}(j)$. Combining with Lemma 2.4, we obtain the following lemma.

Lemma 5.4. *For any $j \in \mathbb{N}$ and for pair $(\ell, \mathbf{i}) \in \mathcal{E}_0 \cap (L \times \mathbf{I}(j))$, we have*

$$
\operatorname{diam}(\operatorname{supp}((\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}})) \leqslant q^{-j}/r^2. \tag{5.13}
$$

Moreover, if $j \geqslant \log(4/r^3)$, then

$$
\operatorname{supp}((\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}) \subset B(V_{g_{\mathbf{i}}}^+, q^{-j}/r^2) \subset b(f_\ell, r/2). \tag{5.14}
$$

To continue, suppose $j \geqslant \log(2C_LC_1/r)$. We can apply Lemma 2.19 to $C_2 = 1/r$, $E := \operatorname{supp}\theta_{\ell,j}$ and $F := \operatorname{supp}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}$. This is because we have $\operatorname{diam} E = \operatorname{diam} \mathcal{Q}_j(\ell) \leqslant q^{-j} \leqslant r/(C_LC_1)$ and Lemma 5.4. For $\ell \in \operatorname{supp}\theta \subset L(t, C_1)$, write $\ell = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix}$. We have $\|h\| \leqslant C_p q^{-t}$. Hence, we obtain

$$
\begin{aligned}
\operatorname{diam}(\operatorname{supp}[\theta_{\ell,j}.(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}]) &\leqslant 16C_LC_1(\operatorname{diam}(\operatorname{supp}\theta_{\ell,j}) + \operatorname{diam}(\operatorname{supp}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}))/(r^2\|h\|^2) \\
&\leqslant 32C_LC_p^2C_1q^{-t-j}/r^4.
\end{aligned}
$$

Therefore, the support of $\mathrm{supp}[\theta_{\ell,j}.(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}]$ is of size $O(q^{-(t+j)+\log(C_1/r^4)})$. Applying Lemma 2.36 to $[\theta_{\ell,j}.\mu_{\varphi(j\mathbf{i})}]$ at the scale $(n,k)=(k,t+j)$, we obtain

$$\frac{1}{k}H([\theta_{\ell,j}.\mu_{\varphi(j,\mathbf{i})}],\mathcal{Q}_{t+j+k}|\mathcal{Q}_{t+j})\geqslant\frac{1}{k}H([\theta_{\ell,j}.(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}],\mathcal{Q}_{t+j+k})-\frac{1}{k}O(\log(C_1/r^3))-\epsilon. \tag{5.15}$$

Consequently, we have

$$\begin{aligned}&\mathbb{E}_{1\leqslant j\leqslant n}\left(\frac{1}{k}H([\theta_{\ell,j}.(\mu_{\varphi(j,\mathbf{i})})],\mathcal{Q}_{t+j+k}|\mathcal{Q}_{t+j}),\ \mathcal{E}_0\right)\\ \geqslant&\mathbb{E}_{1\leqslant j\leqslant n}\left(\frac{1}{k}H([\theta_{\ell,j}.(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}],\mathcal{Q}_{t+j+k}),\ \mathcal{E}_0\right)-Cr^\beta-(\frac{1}{k}+\frac{1}{n})O(\log(C_1/r^4)).\end{aligned}\tag{5.16}$$

Here is how we get the inequality: when $j\geqslant\max\{\log(2/r^2),\log(2C_LC_1/r)\}$, we use Eq. (5.15) to obtain a lower bound with an error term $O(\log(C_1/r^3))/k+\epsilon$, which is less than $O(\log(C_1/r^4))/k+Cr^\beta$; otherwise, we use the trivial bound of entropy $H([\theta_{\ell,j}.(\mu_{\varphi(j,\mathbf{i})})],\mathcal{Q}_{t+j+k}|\mathcal{Q}_{t+j})$ and the definition of $\mathbb{E}_{1\leqslant j\leqslant n}$ to obtain an error term $O(\log(C_1/r^4))/n$. □

By the same proof as in Lemma 4.4 in [BHR19], we get

Lemma 5.5. *For any $\theta\in\mathbf{P}(L)$ and any pair of natural numbers $n\geqslant k$, we have*

$$\frac{1}{n}H(\theta,\mathcal{Q}_n)=\mathbb{E}_{1\leqslant j\leqslant n}\left(\frac{1}{k}H(\theta_{\ell,j},\mathcal{Q}_{j+k})\right)+O\left(\frac{k}{n}+\frac{\log(1+R)}{n}\right),$$

where $R=\mathrm{diam}(\mathrm{supp}(\theta))$.

5.2.3 General projection of convolution of measures

Let $\theta\in\mathbf{P}(L)$ and $x\in\mathbb{P}(\mathbb{R}^3)$. We define the measure $\theta.x$ on $\mathbb{P}(\mathbb{R}^3)$ by $\theta.x:=\int_{\mathrm{SL}_3(\mathbb{R})}\delta_{gx}\mathrm{d}\theta(g)$. Previously, we treated the measure $[\theta.x]$ which is $\pi_{E_1^\perp}(\theta.x)$. More generally, for any $V\in\mathbb{P}(\mathbb{R}^3)$, we consider the measure $\pi_{V^\perp}(\theta.x)$.

Lemma 5.6. *For every $C_1>1$, there exists a compact neighborhood Z of the identity in L such that for any $V\in\mathbb{P}(\mathbb{R}^3)$ with $d(V,E_1^\perp)>1/C_1$, any $\theta\in\mathbf{P}(L)$ such that $\mathrm{supp}(\theta)\subset Z$, and any $k,i\in\mathbb{N}$, we have*

$$\mu\left\{x\in\mathbb{P}(\mathbb{R}^3):\ \frac{1}{k}H(\pi_{V^\perp}(\theta.x),\mathcal{Q}_{k+i})\geqslant\frac{1}{Ck}H(\theta,\mathcal{Q}_{k+i})-\frac{C}{k}\right\}\geqslant\frac{1}{C},$$

where $C>1$ is a constant depending on Z and C_1.

The proof is given in Appendix C.

5.3 Entropy growth under convolutions: projective case

Now we are ready to prove a projective version of Theorem 4.6 in [BHR19]. Recall that α is the value of $\dim\pi_{V^\perp}\mu$ for μ^--a.e. V defined in Lemma 2.38.

Theorem 5.7. *Suppose $0<\alpha<1$. For every $\epsilon>0$ and $C_1>1$, there exists $\delta(\alpha,\epsilon,C_1)>0$ such that the following holds for any $n\geqslant N(\alpha,\epsilon,C_1)$.*

Let $t\geqslant 1$ and let θ be a probability measure on L satisfies

- $\mathrm{diam}(\mathrm{supp}(\theta))<C_p$,
- $\mathrm{supp}(\theta)\subset L(t,C_1)$ *(Eq. (2.66)),*
- $\frac{1}{n}H(\theta,\mathcal{Q}_n)>\epsilon$.

Then
$$\frac{1}{n}H([\theta.\mu],\mathcal{Q}_{t+n})>\alpha+\delta.$$

We introduce a parameter k, a large integer less than n will be defined at the end of the proof of Theorem 5.7. We take r to be $\exp(-\sqrt{k}/10)$. Recall $\mathcal{E}_0$ from Eq. (5.9) with parameter r.

The idea of the proof is similar to [BHR19, Section 4] and [HS17]. In Section 5.3.1, we find measures with small support and positive entropy (Lemma 5.8). In Section 5.3.2, we apply the linearization technique to the convolution to obtain euclidean convolution (Eq. (5.23)). In Section 5.3.3, we continue to modify the small measures, which is to replace $\mathfrak{m}_1$ Eq. (5.25) by $\mathfrak{m}_2$ Eq. (5.26), upgrading Eq. (5.23) to Eq. (5.28). Finally, in Section 5.3.4, we use Proposition 4.2 to verify that $\mathfrak{m}_2$ is entropy porous, hence Theorem 5.1 implies entropy growth.

5.3.1 Find measures with small support that have positive entropy

By Lemma 5.5, we have

$$\mathbb{E}_{1\leqslant j\leqslant n}\left(\frac{1}{k}H(\theta_{\ell,j},\mathcal{Q}_{j+k})\right)=\frac{1}{n}H(\theta,\mathcal{Q}_n)-O(\frac{k}{n}+\frac{\log(1+C_p)}{n})>\epsilon/2,$$

where the second inequality holds under the assumption that k/n and $\log(1+C_p)/n$ are small compare to ϵ. As a direct consequence, we have

$$\mathbb{P}_{1\leqslant j\leqslant n}\left(\frac{1}{k}H(\theta_{\ell,j},\mathcal{Q}_{j+k})\geqslant\frac{\epsilon}{4}\right)\geqslant\frac{\epsilon}{4}. \tag{5.17}$$

Lemma 5.8. *There exists $C_5=C_5(C_1)>0$ such that for k large compared to C_1 and $1/\epsilon$,*

$$\mathbb{E}_{1\leqslant j\leqslant n}\left(\mu\left(x:\frac{1}{k}H([\theta_{\ell,j}.x],\mathcal{Q}_{j+k+t})\geqslant\frac{\epsilon}{C_5}\right)\right)\geqslant\frac{\epsilon}{C_5}.$$

Proof. Let $C_6=C_6(C_1)>1$ be the constant from Lemma 5.6.

Since Lemma 5.6 only works for measures supported in a neighbourhood of identity, we first translate the measures. Due to left-invariance of the metric on L and Eq. (2.47), we have

$$\frac{1}{k}H(\theta_{\ell,j},\mathcal{Q}_{j+k})=\frac{1}{k}H(\ell^{-1}\theta_{\ell,j},\mathcal{Q}_{j+k})+O(\frac{1}{k}).$$

Then $\ell^{-1}\theta_{\ell,j}$ is a measure supported on a neighbourhood of the identity of radius $O(q^{-j})$.

Suppose j is large enough. Then $\ell^{-1}\theta_{\ell,j}$ is supported on Z, the neighborhood of the identity given in Lemma 5.6. As $\operatorname{supp}\theta\subset L(t,c_1)$, we get $d(\ell^{-1}E_1,E_1^{\perp})\geqslant 1/C_1$. Hence we can apply Lemma 5.6 to $V=\ell^{-1}E_1$ and obtain

$$\mu\left(x:\ \frac{1}{k}H(\pi_{(\ell^{-1}E_1)^{\perp}}(\ell^{-1}\theta_{\ell,j}.x),\mathcal{Q}_{j+k})\geqslant\frac{1}{C_6k}H(\theta_{\ell,j},\mathcal{Q}_{j+k})-\frac{C_6}{k}\right)\geqslant\frac{1}{C_6}. \tag{5.18}$$

Using Lemma 2.10, we rewrite

$$[\ell(\ell^{-1}\theta_{\ell,j}.x)]=F(\pi_{(\ell^{-1}E_1)^{\perp}}(\ell^{-1}\theta_{\ell,j}.x)),$$

where we write $\ell=\begin{pmatrix}1&0\\n&h\end{pmatrix}$ and $F=h\pi(\ell^{-1}(E_1),E_1^{\perp},(\ell^{-1}E_1)^{\perp})$ is a diffeomorphism from $\mathbb{P}((\ell^{-1}E_1)^{\perp})$ to $\mathbb{P}(E_1^{\perp})$. The map F satisfies

$$\frac{d(F(y),F(y'))}{d(y,y')}\geqslant q^{-t}/(C_1C_p)^2 \tag{5.19}$$

for $y \neq y'$ in $\mathbb{P}((\ell^{-1}E_1)^\perp)$, because the part $\pi(\ell^{-1}(E_1), E_1^\perp, (\ell^{-1}E_1)^\perp)$ gives a contraction at most C_1 (Lemma B.1) and the part h at most $\|h\|^{-2}$. For $\mathcal{F} = \mathcal{Q}_{j+k}$ and $\mathcal{E} = F^{-1}\mathcal{Q}_{j+k+t}$, we have each atom of $\mathcal{E}$ intersects at most $O((C_1C_p)^2)$ atoms of $\mathcal{F}$ due to Eq. (5.19). We can apply Lemma 2.34. Therefore for x satisfying (5.18), we have

$$\begin{aligned}\frac{1}{k}H([\theta_{\ell,j}.x], \mathcal{Q}_{j+k+t}) &= \frac{1}{k}H(F\pi_{(\ell^{-1}E_1)^\perp}(\ell^{-1}\theta_{\ell,j}.x), \mathcal{Q}_{j+k+t}) \\ &\geqslant \frac{1}{k}H(\pi_{(\ell^{-1}E_1)^\perp}(\ell^{-1}\theta_{\ell,j}.x), \mathcal{Q}_{j+k}) - \frac{O(\log(C_1C_p))}{k} \\ &\geqslant \frac{1}{C_6k}H(\theta_{\ell,j}, \mathcal{Q}_{j+k}) - \frac{C_6 + O(\log(C_1C_p))}{k}.\end{aligned} \tag{5.20}$$

For each $\theta_{\ell,j}$ in the set of Eq. (5.17) and x satisfying Eq. (5.20), we have

$$\frac{1}{k}H([\theta_{\ell,j}.x], \mathcal{Q}_{j+k+t}) \geqslant \frac{1}{C_6k}H(\theta_{\ell,j}, \mathcal{Q}_{j+k}) - \frac{C_6 + O(\log C_1)}{k} \geqslant \frac{\epsilon}{4C_6} - \frac{C_6 + O(\log C_1)}{k} \geqslant \frac{\epsilon}{8C_6},$$

if k is large enough. Hence setting $C_5 = 32C_6$, we have

$$\mathbb{E}_j\left(\mu\left(x : \frac{1}{k}H([\theta_{\ell,j}.x], \mathcal{Q}_{j+k+t}) \geqslant \frac{\epsilon}{C_5}\right)\right) \geqslant \frac{1}{C_5}\mathbb{P}_j\left(\frac{1}{k}H(\theta_{\ell,j}, \mathcal{Q}_{j+k}) \geqslant \frac{\epsilon}{4}\right).$$

The proportion of j between 1 and n such that the ball $B(id, q^{-j})$ is not contained in Z goes to zero as $n \to \infty$. The proof is complete by Eq. (5.17). □

5.3.2 Apply linearization

We take a small constant $\epsilon_1 > 0$, which will be chosen small compared to fixed constants ϵ' and $\delta(\epsilon')$ defined later.

Let $K(\epsilon/1)$ be the constant given in Lemma 5.2. We may assume $k > K(\epsilon_1)$. Consider any $j \in \mathbb{N}$ such that

$$q^{-j}/r^2 \geqslant q^{-k}(r/C_1)^{10}$$

and set $\rho = q^{-j}/r$. We apply Lemma 5.2 to

- the constants ϵ_1, k and $r = q^{-\sqrt{k}/10}$;
- the convolution measure $[\theta_{\ell,j}.(\mu_{\varphi(j,\mathbf{j})})_{\mathbf{I}}]$ at any point $x \in \mathrm{supp}(\mu_{\varphi(j,\mathbf{j})})_{\mathbf{I}}$ with $(\ell, \mathbf{i}) \in \mathcal{E}_0$.

We check the conditions of Lemma 5.2:

- $\mathrm{supp}(\theta_{\ell,j}) \subset B(\ell, q^{-j}) \subset B(\ell, \rho)$;
- due to $j \geqslant \log(4/r^2)$, by Lemma 5.4, we have $\mathrm{supp}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}} \subset B(V^+_{g_{\mathbf{i}}}, \rho) \subset b(f_\ell, r/2)$. Hence, for any $x \in \mathrm{supp}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}$, we have $\mathrm{supp}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}} \subset B(x, 2\rho) \subset b(f_\ell, r)$.

Therefore, Lemma 5.2 yields

$$\begin{aligned}&\frac{1}{k}H([\theta_{\ell,j}.(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}], \mathcal{Q}_{k+t+j}) \geqslant \frac{1}{k}H([\theta_{\ell,j}.(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}], \mathcal{Q}_{k+t-\log(2\rho)}) + O\left(\frac{\log(1/r)}{k}\right) \\ \geqslant &\frac{1}{k}H(S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x] * \pi_{(\ell^{-1}E_1)^\perp}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}, \mathcal{Q}_{k+j}) - O\left(\frac{\log(C_1/r)}{k}\right) - \epsilon_1,\end{aligned} \tag{5.21}$$

with

$$|t(\ell, x) - t| \leqslant 8\log(C_1/r). \tag{5.22}$$

Hence, combining with Proposition 5.3, we have

$$\frac{1}{n}H([\theta.\mu],\mathcal{Q}_{t+n}) \geqslant \mathbb{E}_{1\leqslant j\leqslant n,\ell\sim\theta}\left(\frac{1}{k}H([\theta_{\ell,j}.(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}],\mathcal{Q}_{t+j+k}),\ \mathcal{E}_0\right) - \textit{error term}$$

$$\geqslant \mathbb{E}_{1\leqslant j\leqslant n,\ell\sim\theta}\left(\left(\int\frac{1}{k}H((S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x])*\pi_{(\ell^{-1}E_1)^\perp}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}},\mathcal{Q}_{j+k})\mathrm{d}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}(x)\right),\mathcal{E}_0\right)$$

$$-\textit{ error term}\ \ \text{(integrating Eq. (5.21) for } x \text{ with the measure } (\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}). \tag{5.23}$$

5.3.3 Replace measures by their good parts

For any $\mathbf{i}\in\mathbf{I}(j)$ and $V\in\mathbb{P}(\mathbb{R}^3)$, Eq. (4.9) states that

$$\pi_{V^\perp}\circ g_{\mathbf{i}} = h_{V,g_{\mathbf{i}}}\circ\pi_{g_{\mathbf{i}}^{-1}V,V^\perp}. \tag{5.24}$$

We denote $h_{V,g_{\mathbf{i}}}$ by $h_{V,\mathbf{i}}$ and $\pi_{g_{\mathbf{i}}^{-1}V,V^\perp}$ by $\pi_{V,\mathbf{i}}$ for simplification.

Recall that $\mu_{\varphi(j,\mathbf{i})}=g_{\mathbf{i}}\mu$. So far, we have considered the measure

$$\mathfrak{m}_1 := \pi_{(\ell^{-1}E_1)^\perp}((\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}), \tag{5.25}$$

which is the pushforward of the r-attracting part of $(g_{\mathbf{i}},\mu)$ by the projection $\pi_{(\ell^{-1}E_1)^\perp}$. To continue, we will switch the order and consider the following measure

$$\mathfrak{m}_2(\ell,j,\mathbf{i}) := r^5\text{-attracting part of the pair } (h_{\ell^{-1}E_1,\mathbf{i}},\pi_{\ell^{-1}E_1,\mathbf{i}}\mu) \tag{5.26}$$

(see Definition 4.7 for the construction of the attracting part). **The measure $\mathfrak{m}_2(\ell,j,\mathbf{i})$ is a random measure with respect to the random variables $\ell(\sim\theta)$ and $\mathbf{i}(\in\mathbf{I}(j))$.** We will frequently write $\mathfrak{m}_2(\ell,j,\mathbf{i})$ as $\mathfrak{m}_2$ to simplify the notation. The advantage of considering $\mathfrak{m}_2$ is that is $\mathfrak{m}_2$ is a restriction of $\pi_{(\ell^{-1}E_1)^\perp}\mu_{\varphi(j,\mathbf{i})}$ and we can apply to it the porosity result Proposition 4.2.

To relate $\mathfrak{m}_1$ and $\mathfrak{m}_2$, we introduce the following subset:

$$\mathcal{E}_0'(j)=\left\{\begin{array}{l}(\ell,\mathbf{i})\in\mathcal{E}_0\cap(L\times\mathbf{I}(j)):\ d(\pi_{(\ell^{-1}E_1)^\perp}V_{g_{\mathbf{i}}}^+,h_{\ell^{-1}E_1,\mathbf{i}}{}^+)\leqslant q^{-j}/r^2,\\ |\chi_1(h_{\ell^{-1}E_1,\mathbf{i}})-\chi_1(g_{\mathbf{i}})|\leqslant|\log r|\end{array}\right\} \tag{5.27}$$

We will upgrade Eq. (5.23) in Step 2 to the following version.

Proposition 5.9. *For any $n\geqslant k\geqslant K(1)$, we have*

$$\frac{1}{n}H([\theta.\mu],\mathcal{Q}_{t+n})$$
$$\geqslant\mathbb{E}_{1\leqslant j\leqslant n,\ell\sim\theta}\left(\int\frac{1}{k}H((S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x])*\mathfrak{m}_2(\ell,j,\mathbf{i}),\mathcal{Q}_{j+k})\mathrm{d}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}(x),\mathcal{E}_0'(j)\right) \tag{5.28}$$
$$-\epsilon_1-O\Big(\frac{k}{n}+r^\beta+\frac{\log(C_1/r)}{k}\Big),$$

where $r=q^{-\sqrt{k}/10}$.

We start the proof of Proposition 5.9 by comparing the support of $\mathfrak{m}_1$ and $\mathfrak{m}_2$.

Lemma 5.10. *For any $k\geqslant K(1)\geqslant 1$ and $(\ell,\mathbf{i})\in\mathcal{E}_0'(j)$, we have:*

1. *If $j\geqslant 6|\log r|$, then*

$$B(h_{\ell^{-1}E_1,\mathbf{i}}^+,c_0q^{-j}/2r^4)\subset\operatorname{supp}\mathfrak{m}_2\subset B(h_{\ell^{-1}E_1,\mathbf{i}}^+,q^{-j}/r^{11}), \tag{5.29}$$

where $0<c_0<1$ is the constant given as in Eq. (2.68).

2. *If $j \geqslant \log(4/r^3)$, we can write*

$$\mathfrak{m}_2 = (1-\delta)\mathfrak{m}_1 + \delta\mathfrak{m}', \tag{5.30}$$

with $\mathfrak{m}'$ a probability measure on $\mathbb{P}((\ell^{-1}E_1)^\perp)$ and $\delta < Cr^\beta$.

3. *We can write*

$$\pi_{(\ell^{-1}E_1)^\perp}\mu_{\varphi(j,\mathbf{i})} = \lambda\mathfrak{m}'' + (1-\lambda)\mathfrak{m}_2(\ell,j,\mathbf{i}),$$

where $\mathfrak{m}''$ is a probability measure on $\mathbb{P}((\ell^{-1}E_1)^\perp)$. It is a $(q^{-j}/r^{11}, Cr^\beta)$-decomposition of $\pi_{(\ell^{-1}E_1)^\perp}\mu_{\varphi(j,\mathbf{i})}$ (see Definition 4.6).

4. *For any pair $m, j \in \mathbb{N}$ with $m - j \geqslant |\log r|$, the equality of the component measures*

$$(\pi_{(\ell^{-1}E_1)^\perp}\mu_{\varphi(j,\mathbf{i})})_{x,m} = (\mathfrak{m}_2(\ell,j,\mathbf{i}))_{x,m} \tag{5.31}$$

holds at a set of x with $\pi_{(\ell^{-1}E_1)^\perp}\mu_{\varphi(j,\mathbf{i})}$-measure greater than $1 - O(r^\beta)$.

Proof. 1. Using the definition of the random variable $\mathbf{I}(j)$ (Eq. (2.68)), we have $\chi_1(g_\mathbf{i}) \in [j, j + |\log c_0|]$. Combined with Eq. (5.27) and $k \geqslant K(1)$, we have $\chi_1(h_{\ell^{-1}E_1,\mathbf{i}}) \in j + [-|\log r|, |\log r| + |\log c_0|]$. For $\mathfrak{m}_2$, it is the r^5-attracting part of the pair $(h_{\ell^{-1}E_1,\mathbf{i}}, \pi_{\ell^{-1}E_1,\mathbf{i}}\mu)$ and we also have $q^{-\chi_1(h_{\ell^{-1}E_1,\mathbf{i}})} \leqslant r^5$ as $j \geqslant 4|\log r|$. Hence by Eq. (4.3), we have

$$B({h_{\ell^{-1}E_1,\mathbf{i}}}^+, c_0q^{-j}/2r^4) \subset \operatorname{supp}\mathfrak{m}_2 \subset B({h_{\ell^{-1}E_1,\mathbf{i}}}^+, q^{-j}/r^{11}).$$

2. As $j \geqslant \log(4/r^3)$, we can apply Eq. (5.14) to $(\mu_{\varphi(j,\mathbf{i})})_\mathbf{I}$ and have

$$\operatorname{supp}\mathfrak{m}_1 = \pi_{(\ell^{-1}E_1)^\perp}\operatorname{supp}(\mu_{\varphi(j,\mathbf{i})})_\mathbf{I} \subset \pi_{(\ell^{-1}E_1)^\perp}B(V_{g_\mathbf{i}}^+, q^{-j}/r^2) \subset \pi_{(\ell^{-1}E_1)^\perp}b(f_\ell, r).$$

As $(\ell, \mathbf{i}) \in \mathcal{E}_0$, by the definition of $\mathcal{E}_0$ (Eq. (5.9)), we have $V_{g_\mathbf{i}}^+ \in b(f_\ell, r)$. Using Eq. (2.18), we have $d(V_{g_\mathbf{i}}^+, \ell^{-1}E_1) > r/C_1$. Applying (2) in Lemma B.2 to $B(V_{g_\mathbf{i}}^+, q^{-j}/r^2)$, we obtain

$$\operatorname{supp}\mathfrak{m}_1 \subset B(\pi_{(\ell^{-1}E_1)^\perp}V_{g_\mathbf{i}}^+, C_1q^{-j}/r^3).$$

It follows from the definition of $\mathcal{E}_0'(j)$ that $d({h_{\ell^{-1}E_1,\mathbf{i}}}^+, \pi_{(\ell^{-1}E_1)^\perp}V_{g_\mathbf{i}}^+) \leqslant q^{-j}/r^2$. Therefore, we have $\operatorname{supp}\mathfrak{m}_1 \subset \operatorname{supp}\mathfrak{m}_2$.

From the definitions of $\mathfrak{m}_2$ and Eq. (5.24), we know that $\mathfrak{m}_2$ is a restriction of $\pi_{(\ell^{-1}E_1)^\perp}(\mu_{\varphi(j,\mathbf{i})})$. So we can write $\mathfrak{m}_2 = \pi_{(\ell^{-1}E_1)^\perp}((g_\mathbf{i}\mu)_S)$ with $S = \pi^{-1}_{(\ell^{-1}E_1)^\perp}\operatorname{supp}\mathfrak{m}_2$. So the fact that $\operatorname{supp}\mathfrak{m}_1 \subset \operatorname{supp}\mathfrak{m}_2$ yields the inclusion $S \supset S'$ with $S' = \operatorname{supp}((\mu_{\varphi(j,\mathbf{i})})_\mathbf{I})$. Since the measure $(\mu_{\varphi(j,\mathbf{i})})_\mathbf{I}$ is the restriction of $\mu_{\varphi(j,\mathbf{i})}$ on S', we can write

$$(\mu_{\varphi(j,\mathbf{i})})_S = (1-\delta)(\mu_{\varphi(j,\mathbf{i})})_\mathbf{I} + \delta\mu', \tag{5.32}$$

with μ' a probability measure on $\mathbb{P}(\mathbb{R}^3)$, and

$$\delta = 1 - (\mu_{\varphi(j,\mathbf{i})})_S(S') \leqslant 1 - \mu_{\varphi(j,\mathbf{i})}(S') \leqslant Cr^\beta,$$

where the last inequality is due to Eq. (5.12). Applying $\pi_{(\ell^{-1}E_1)^\perp}$ to Eq. (5.32), we obtain Eq. (5.29).

3. We can apply Lemma 4.9 to the measure $\mathfrak{m}_2(\ell, j, \mathbf{i})$ and the estimate of $\chi_1(h_{\ell^{-1}E_1,\mathbf{i}})$ to obtain the third statement.

4. With the third statement available, it remains to show that the two atoms of $\mathcal{Q}_m$ containing the endpoints of the support of $\mathfrak{m}_2$ have small measure. As we denote $g_{\mathbf{i}}\mu$ by $\mu_{\varphi(j,\mathbf{i})}$, we have

$$\pi_{(\ell^{-1}E_1)^\perp}g_{\mathbf{i}}\mu(B(x,q^{-m})) = \mu(g_{\mathbf{i}}^{-1}B((x,\ell^{-1}E_1),q^{-m})) \leqslant \mu(B(W,q^{-(m-j)})) \leqslant Cr^\beta$$

where $(x,\ell^{-1}E_1)$ is the plane passing x, $\ell^{-1}E_1$, $W = g_{\mathbf{i}}^{-1}(x,\ell^{-1}E_1)$ and the last inequality is due to Lemma 2.24.

□

We will prove Proposition 5.9 in the remaining Step 3. Then in Step 4, we will estimate the integrand in the right-hand side of Eq. (5.28).

Proof of Proposition 5.9. We only consider $(\ell,\mathbf{i}) \in \mathcal{E}_0'(j)$, $x \in \operatorname{supp}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}$ and $j \geqslant 6|\log r|$. The part of expectation with $j \leqslant 6|\log r|$ is bounded by $\frac{6|\log r|}{n} \times \frac{k+6|\log r|}{k} = O(\frac{|\log r|}{n})$.

Due to Eq. (5.30) and Eq. (2.77) (inverse of the concavity of entropy), we obtain

$$\begin{aligned}\frac{1}{k}H(S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x] * \mathfrak{m}_1, \mathcal{Q}_{j+k}) \geqslant & \frac{1}{1-\delta}\Big(\frac{1}{k}H(S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x] * \mathfrak{m}_2, \mathcal{Q}_{j+k}) \\ & - 2\delta\frac{1}{k}H(S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x] * \mathfrak{m}', \mathcal{Q}_{j+k})\Big) - \frac{2}{k}H(\delta).\end{aligned}$$

We give an upper bound to the second term on the right-hand side. Note that by Lemma 2.34, for any $\tau \in \mathbb{P}(\mathbb{P}(\mathbb{R}^2))$, we have

$$H(\tau,\mathcal{Q}_{j+k}) = H(\tau,\mathcal{Q}_{j+k}|\mathcal{Q}_j) + H(\tau,\mathcal{Q}_j) \leqslant k + O(\log(\operatorname{diam}(\operatorname{supp}\tau)q^j)).$$

To apply this to the convolution measure, we estimate the diameter of the convolution measure. In Step 2, we already verified that we can apply Lemma 5.2 to $(\theta_{\ell,j},x)$ for any $j \in N$ satisfying $q^{-j}/r^2 \leqslant q^{-k}(r/C_1)^{10}$. In particular, Eq. (5.5) in Lemma 5.2 holds for $[\theta_{\ell,j}.x]$. Combining with Eq. (5.22), we have

$$\operatorname{diam}(\operatorname{supp} S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x]) \leqslant 16C_LC_p^2C_1^9q^{-j}/r^{10}. \tag{5.33}$$

The support of $\mathfrak{m}'$ is controlled by Eq. (5.29), that is $O(q^{-j}/r^7)$. So we have

$$\frac{1}{k}H(S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x] * \mathfrak{m}', \mathcal{Q}_{j+k})) \leqslant 1 + \frac{O(\log(C_1/r))}{k}.$$

Hence, we have

$$\frac{1}{k}H(S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x]*\mathfrak{m}_1, \mathcal{Q}_{j+k}) \geqslant \frac{1}{k}H(S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x]*\mathfrak{m}_2, \mathcal{Q}_{j+k}) - O(\delta\log(C_1/r))/k - 2H(\delta)/k,$$

which holds for every $(\ell,\mathbf{i}) \in \mathcal{E}_0'(j)$.

Notice that $\mathfrak{m}_1 = \pi_{(\ell^{-1}E_1)^\perp}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}$. We replace $\mathfrak{m}_1$ by $\mathfrak{m}_2$ in Eq. (5.23) and obtain

$$\frac{1}{n}H([\theta.\mu],\mathcal{Q}_{t+n}) \geqslant \mathbb{E}_{1\leqslant j\leqslant n,\ell\sim\theta}\left(\int \frac{1}{k}H((S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x]) * \mathfrak{m}_2(\ell,j,\mathbf{i}), \mathcal{Q}_{j+k})\mathrm{d}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}(x), \mathcal{E}_0'(j)\right)$$

$-$ *error term.* □

Let $\mathbb{P}_{1\leqslant j\leqslant n,\ell\sim\theta}$ be the product measure $\mathbb{P}_{1\leqslant j\leqslant n} \otimes \theta$ for the random variable $(\mathbf{i},\ell)$. For any $\ell \in L$ and any $j \in \mathbb{N}$, let $\mathcal{E}_0'(j)(\ell) = \{\mathbf{i} \in \mathbf{I} : (\ell,\mathbf{i}) \in \mathcal{E}_0'(j)\}$.

Lemma 5.11. *For every $\epsilon > 0$ and $n \geqslant k \geqslant K(\epsilon)$, we have for any $\ell \in \operatorname{supp}\theta$,*

$$\mathbb{P}_{1\leqslant j\leqslant n}(\mathcal{E}_0'(j)(\ell)) \geqslant 1-\epsilon$$

and hence

$$\mathbb{P}_{1\leqslant j\leqslant n,\ell\sim\theta}(\mathcal{E}_0'(j) \geqslant 1-\epsilon.$$

Proof. Fix any $\ell \in \operatorname{supp}\theta$ and let $V = \ell^{-1}E_1 \in \mathbb{P}(\mathbb{R}^3)$.

By Lemma D.1, we have

$$\begin{aligned}&\mathbb{P}_{1\leqslant j\leqslant n}(\mathcal{E}_0'(j)(\ell))\\ \geqslant&\mathbb{P}_{1\leqslant j\leqslant n}\left\{\mathbf{i}\in\mathbf{I}(j): V_{g_{\mathbf{i}}}^+\in b(f_\ell,2r),\ \ d(V^\perp,H_{g_{\mathbf{i}}}^-)>2r,\ \ |\chi_1(h_{V,\mathbf{i}})-\chi_1(g_{\mathbf{i}})|\leqslant|\log r|/2\right\}.\end{aligned}$$

Denote the set in the right hand side by $\mathcal{U}$. In view of Lemma 2.33, to estimate the probability $\mathbb{P}_{1\leqslant j\leqslant n}$ of the above set for $\mathbf{I}$, it suffices to prove for $\mathbf{U}$.

We apply the large deviation estimate Eq. (2.55) to $V_{g_{\mathbf{i}}}^+$ and the kernel of the linear form f_ℓ (Definition 2.13); apply Eq. (2.56) to $H_{g_{\mathbf{i}}}^-$ and any point $V_1 \in V^\perp$ (which is stronger than the Hausdorff distance $d(V^\perp, H_{g_{\mathbf{i}}}^-)$); for the third condition in $\mathcal{U}$, apply Eq. (2.59) (which is stated for $V = E_1$, but due to equivalence, it is also true for general V). Therefore, we obtain for $n \geqslant k$,

$$\begin{aligned}\mathbb{P}_{1\leqslant j\leqslant n}(\mathbf{U}(j)\in\mathcal{U}^c)&=\frac{|\log r|/c}{n}\mathbb{P}_{1\leqslant j<|\log r|/c}(\mathbf{U}(j)\in\mathcal{U}^c)+\frac{n-|\log r|/c}{n}\mathbb{P}_{|\log r|/c\leqslant j\leqslant n}(\mathbf{U}(j)\in\mathcal{U}^c)\\&<\frac{|\log r|/c}{n}+Cr^\beta\leqslant C(r^\beta+1/\sqrt{k}),\end{aligned}$$

where the constant $c>0$ comes from Eq. (2.55) and the last inequality is due to $r=\exp(-\sqrt{k}/10)$, $k \leqslant n$. Using Lemma 2.33, we have for $n \geqslant k$

$$\mathbb{P}_{1\leqslant j\leqslant n}(\mathbf{I}(j)\in\mathcal{U}^c)\ll r^\beta+1/\sqrt{k}+e^{-\beta n}. \tag{5.34}$$

The proof is complete by taking k large. □

5.3.4 Apply entropy growth of the Euclidean case

Let

$$\epsilon'=\min\{\frac{\epsilon}{10C_5},\frac{1-\alpha}{10}\}, \tag{5.35}$$

where C_5 comes from Lemma 5.8. Let $\delta_1=\delta(\epsilon'/2)$ from Theorem 5.1 and let $m=M(\epsilon'/4)$ from Proposition 4.2. Take $\epsilon_1>0$ to be a constant independent of ϵ',δ_1,m, which will be determined at the end of the proof. We introduce the following sets for $j,k\in\mathbb{N}$

$$\begin{aligned}\mathcal{E}_1(j)&=\left\{(\ell,\mathbf{i},x)\in\mathcal{E}_0'(j)\times\mathbb{P}(\mathbb{R}^3):\ \ x\in\operatorname{supp}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}},\ \ \frac{1}{k}H(S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x],\mathcal{Q}_{j+k})\geqslant\frac{\epsilon}{C_5}\right\},\\ \mathcal{E}_2(j)&=\left\{\begin{array}{l}(\ell,\mathbf{i},x)\in\mathcal{E}_0'(j)\times\mathbb{P}(\mathbb{R}^3):\mathfrak{m}_2(\ell,j,\mathbf{i})\text{ is }(1-\epsilon',\epsilon'/2,m)\text{-entropy porous}\\ \text{from scale } j \text{ to } j+k\end{array}\right\},\\ \mathcal{E}_3(j)&=\left\{\begin{array}{l}(\ell,\mathbf{i},x)\in\mathcal{E}_0'(j)\times\mathbb{P}(\mathbb{R}^3):\ \ x\in\operatorname{supp}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}},\\ \frac{1}{k}H(S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x]*\mathfrak{m}_2(\ell,j,\mathbf{i}),\mathcal{Q}_{j+k})\geqslant\frac{1}{k}H(\mathfrak{m}_2(\ell,j,\mathbf{i}),\mathcal{Q}_{j+k})+\delta_1/2\end{array}\right\},\\ \mathcal{E}_4(j)&=\left\{(\ell,\mathbf{i},x)\in\mathcal{E}_0'(j)\times\mathbb{P}(\mathbb{R}^3):\ \frac{1}{k}H(\mathfrak{m}_2(\ell,j,\mathbf{i}),\mathcal{Q}_{j+k})\geqslant\alpha-\epsilon_1\right\}.\end{aligned}$$

We denote by π_2 the projections from $L\times\Lambda^*\times\mathbb{P}(\mathbb{R}^3)$ to $L\times\Lambda^*$. We want to estimate

$$\mathbb{E}_{1\leqslant j\leqslant n,\ell\sim\theta}\left(\int\frac{1}{k}H((S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x])*\mathfrak{m}_2(\ell,j,\mathbf{i}),\mathcal{Q}_{j+k})\mathrm{d}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}(x),\mathcal{E}_0'(j)\right)$$

in Eq. (5.28). The rough idea is to estimate the measure of each set above and it is achieved in Claim 1 to Claim 4.

Claim 1. *For $k>K(\epsilon',\delta_1)$, we have*

$$\mathcal{E}_1(j)\cap\mathcal{E}_2(j)\subset\mathcal{E}_3(j).$$

Proof. For any $(\ell, \mathbf{i}, x) \in \mathcal{E}_1(j) \cap \mathcal{E}_2(j)$, we want to apply Theorem 5.1 to $S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x])$ and $\mathfrak{m}_2(\ell, j, \mathbf{i})$ to obtain growth of entropy at the scale $(k, n) = (j, k)$. In view of the definitions of $\mathcal{E}_1(j)$ and $\mathcal{E}_2(j)$, it remains to estimate the size of the support of these two measures.

It follows from Eq. (5.29) that

$$\operatorname{supp} \mathfrak{m}_2(\ell, j, \mathbf{i}) \subset B(h^+_{\ell^{-1}E_1,\mathbf{i}}, q^{-j}/r^{11}).$$

And Eq. (5.33) yields that

$$\operatorname{diam}\left(\operatorname{supp} S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x]\right) \leqslant 16C_LC_p^2C_1^9q^{-j}/r^{10}.$$

Take the constant C in Theorem 5.1 to be $16C_LC_p^2C_1^9/r^{12}$. Then

$$\frac{\log C}{k} = \frac{\log(16C_LC_p^2) + 9\log C_1 - 12\log r}{k} \leqslant \frac{\log(16C_LC_p^2) + 9\log C_1 + \sqrt{k}}{k},$$

which goes to 0 as k goes to infinity. So we have $k > N(\epsilon', \delta_1, m)\log C$. Therefore, Theorem 5.1 yields

$$\frac{1}{k}H(S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x] * \mathfrak{m}_2(\ell, j, \mathbf{i}), \mathcal{Q}_{j+k}) \geqslant \frac{1}{k}H(\mathfrak{m}_2(\ell, j, \mathbf{i}), \mathcal{Q}_{j+k}) + \delta_1.$$ □

Claim 2. *For $k \geqslant K(C_1, \epsilon)$, we have*

$$\mathbb{E}_{1\leqslant j\leqslant n,\ell\sim\theta}\left(\int 1_{\mathcal{E}_1(j)}(\ell, \mathbf{i}, x)\mathrm{d}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}(x)\right) \geqslant \frac{\epsilon}{4C_5}.$$

Proof. For any $j \in \mathbb{N}$, we have the equality $\mu = \mathbb{E}_j(\mu_{\varphi(j,\mathbf{i})})$. Hence, using (5.12), we have

$$\sup_{A \text{ Borel}} |\mu(A) - \mathbb{E}_j((\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}})(A)| \leqslant \mathbb{E}_j(|\mu_{\varphi(j,\mathbf{i})}(A) - (\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}(A)|) \ll r^\beta.$$

Let $A(\ell, j) = \{x : \frac{1}{k}H([\theta_{\ell,j}.x], \mathcal{Q}_{j+k+t}) \geqslant \frac{\epsilon}{C_5}\}$. Then for any $1 \leqslant j \leqslant n$,

$$\mathbb{E}_{j,\ell\sim\theta}\left((\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}(A(\ell, j))\right) = \mathbb{E}_{\ell\sim\theta}\mathbb{E}_j\left((\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}(A(\ell, j))\right) \geqslant \mathbb{E}_{\ell\sim\theta}\left(\mu(A(\ell, j))\right) - O(r^\beta),$$

where the inequality is due to the above inequality. Summing over $1 \leqslant j \leqslant n$ and combining with Lemma 5.8, we obtain

$$\begin{aligned}
&\mathbb{E}_{1\leqslant j\leqslant n,\ell\sim\theta}\left((\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}\left\{x : \frac{1}{k}H([\theta_{\ell,j}.x], \mathcal{Q}_{j+k+t}) \geqslant \frac{\epsilon}{C_5}\right\}\right) \\
&\geqslant \mathbb{E}_{1\leqslant j\leqslant n,\ell\sim\theta}\left(\mu\left\{x : \frac{1}{k}H([\theta_{\ell,j}.x], \mathcal{Q}_{j+k+t}) \geqslant \frac{\epsilon}{C_5}\right\}\right) - O(r^\beta) \\
&\geqslant \frac{\epsilon}{C_5} - O(r^\beta) \geqslant \frac{\epsilon}{2C_5}.
\end{aligned} \tag{5.36}$$

Then we can use Eq. (5.22), $|t(\ell, x) - t| \leqslant 8\log C_1/r$ (Eq. (5.22)) to replace $\frac{1}{k}H([\theta_{\ell,j}.x], \mathcal{Q}_{j+k+t})$ by $\frac{1}{k}H(S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x], \mathcal{Q}_{j+k})$, and it produces an error $O(\log C_1/r) \leqslant \epsilon'/2$.

Comparing Eq. (5.36) with Claim 2, we use Lemma 5.11 to add the restriction of $\mathcal{E}_0'(j)$. The proof is complete. □

Claim 3. *For $k > K(\epsilon')$ and $n > N(\epsilon', k)$, we have*

$$\mathbb{P}_{1\leqslant j\leqslant n,\ell\sim\theta}(\pi_2(\mathcal{E}_2(j))) \geqslant 1 - \epsilon'.$$

Proof. We fix any $\ell \in \operatorname{supp}\theta$, we will show that

$$\mathbb{P}_{1\leqslant j\leqslant n}\left\{\mathbf{i}\in\mathcal{E}_0'(j)(\ell):\ \mathfrak{m}_2(\ell,j,\mathbf{i}) \text{ is } (1-\epsilon',\epsilon'/2,m)\text{-entropy porous from scale } j \text{ to } j+k\right\}>1-\epsilon'.$$

Then the claim follows by integrating ℓ using the measure θ.

It is good to recall that that $\mathfrak{m}_2(\ell,j,\mathbf{i})$ is a restriction of $\pi_{\ell^{-1}E_1)^\perp} g_{\mathbf{i}}\mu$ from Eq. (5.26) and Lemma 5.10. Proposition 4.2 yields

$$\mathbb{P}_{1\leqslant j\leqslant n}\left\{\mathbf{i}\in\mathbf{I}(j):\pi_{(\ell^{-1}E_1)^\perp}g_{\mathbf{i}}\mu \text{ is } (\alpha,\epsilon'/4,m)\text{-entropy porous from scale } j \text{ to } j+k\right\}>1-\epsilon'/4. \tag{5.37}$$

We take k large enough such that for $r=\exp(-\sqrt{k}/10)$, we have $r^\beta<\epsilon'/4$ and $|\log r|/k = 1/\sqrt{k}\leqslant\epsilon'/4$.

For $\mathbf{i}\in\mathcal{E}_0'(j)(\ell)$, if $\mathbf{i}$ belongs to the set defined in Eq. (5.37), then Eq. (5.31) implies that $\mathfrak{m}_2(\ell,j,\mathbf{i})$ is $(\alpha,\epsilon'/2,m)$-entropy porous from scale j to $j+k$. (Actually, Eq. (5.31) only allows us to obtain the entropy porosity of $\mathfrak{m}_2(\ell,j,\mathbf{i})$ from scale $j+|\log r|$ to $j+k$. Since $|\log r|/k$ is sufficiently small, we obtain the result.)

Combined with Lemma 5.11 for the measure of $\mathcal{E}_0'(j)(\ell)$, we obtain the first inequality in the proof and finish the proof of the claim. ☐

Claim 4. *For $k>K(\epsilon_1)$ and $n>N(\epsilon_1,k)$*

$$\mathbb{P}_{1\leqslant j\leqslant n,\ell\sim\theta}(\pi_2(\mathcal{E}_4(j)))\geqslant 1-\epsilon_1.$$

Proof. Fix any $\ell\in\operatorname{supp}\theta$. Recall from Eq. (5.26) that $\mathfrak{m}_2(\ell,j,\mathbf{i})$ is the r^5-attracting part of the pair $(h_{\ell^{-1}E_1,\mathbf{i}},\pi_{\ell^{-1}E_1,\mathbf{i}}\mu)$. Let $\epsilon=\delta_1=\epsilon_1/8$. If $r^5\leqslant r_1(\epsilon_1)$, then we can apply Lemma 4.16 to $\mathfrak{m}_2$ with $V=\ell^{-1}E_1$ and $(m,i)=(k,j)$. This gives

$$\mathbb{P}_{1\leqslant j\leqslant n}\left(\frac{1}{k}H(\mathfrak{m}_2(\ell,j,\mathbf{i}),\mathcal{Q}_{k+j})\geqslant\alpha-\epsilon_1/2\right)\geqslant 1-\epsilon_1/4. \tag{5.38}$$

Observed that $\pi_2(\mathcal{E}_4(j))(\ell)=\{\mathbf{i}:\ (\ell,\mathbf{i})\in\pi_2(\mathcal{E}_4(j))\}$ is the intersection of the set of Eq. (5.38) with $\mathcal{E}_0'(j)(\ell)$. Combined with Lemma 5.11, we finishes the proof of the claim. ☐

Proof of Theorem 5.7. . With Proposition 5.9 available, it remains to estimate

$$\mathbb{E}_{1\leqslant j\leqslant n,\ell\sim\theta}\left(\int\frac{1}{k}H((S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x])*\mathfrak{m}_2(\ell,j,\mathbf{i}),\mathcal{Q}_{j+k})\mathrm{d}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}(x),\mathcal{E}_0'(j)\right).$$

Here we abuse the notation, for a set F in $L\times\Lambda^*\times\mathbb{P}(\mathbb{R}^3)$, we use $\mathbb{P}_{1\leqslant j\leqslant n}(F)$ to denote

$$\mathbb{P}_{1\leqslant j\leqslant n}(F)=\mathbb{E}_{1\leqslant j\leqslant n,\ell\sim\theta}\left(\int 1_F(\ell,\mathbf{i},x)\mathrm{d}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}(x)\right).$$

Under this notation, due to the sets $\mathcal{E}_2(j)$ and $\mathcal{E}_4(j)$ independent of x, we have $\mathbb{P}_{1\leqslant j\leqslant n}(\mathcal{E}_2(j))=\mathbb{P}_{1\leqslant j\leqslant n,\ell\sim\theta}(\pi_2(\mathcal{E}_2(j)))$, similarly for $\mathcal{E}_4(j)$. Claim 2, Claim 3 and Claim 4 are the estimates of $\mathbb{P}_{1\leqslant j\leqslant n}(\mathcal{E}_1(j))$, $\mathbb{P}_{1\leqslant j\leqslant n}(\mathcal{E}_2(j))$ and $\mathbb{P}_{1\leqslant j\leqslant n}(\mathcal{E}_4(j))$.

Therefore, we have

$$\begin{aligned}&\mathbb{E}_{1\leqslant j\leqslant n,\ell\sim\theta}\left(\int\frac{1}{k}H(S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x]*\mathfrak{m}_2(\ell,j,\mathbf{i}),\mathcal{Q}_{j+k})\mathrm{d}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}(x),\mathcal{E}_0'(j)\right)\\ &\geqslant(\delta_1/2+\alpha-\epsilon_1)\mathbb{P}_{1\leqslant j\leqslant n}(\mathcal{E}_3(j)\cap\mathcal{E}_4(j))+(\alpha-\epsilon_1-O(1/k))\mathbb{P}_{1\leqslant j\leqslant n}(\mathcal{E}_4(j)-\mathcal{E}_3(j))\\ &\geqslant(\delta_1/2)\mathbb{P}_{1\leqslant j\leqslant n}(\mathcal{E}_3(j)\cap\mathcal{E}_4(j))+(\alpha-\epsilon_1)\mathbb{P}_{1\leqslant j\leqslant n}(\mathcal{E}_4(j))-O(1/k)\end{aligned}$$

where on $\mathcal{E}_3(j)\cap\mathcal{E}_4(j)$, we use their definition; on $\mathcal{E}_4(j)-\mathcal{E}_3(j)$, we estimate the integrand using the trivial lower bound Eq. (5.1).

For $\mathbb{P}_{1\leqslant j\leqslant n}(\mathcal{E}_4(j))$, we use Claim 4. For $\mathbb{P}_{1\leqslant j\leqslant n}(\mathcal{E}_3(j)\cap\mathcal{E}_4(j))$, we use Claim 1, Claim 2, Claim 3 and Claim 4 to obtain

$$\mathbb{P}_{1\leqslant j\leqslant n}(\mathcal{E}_3(j)\cap\mathcal{E}_4(j))\geqslant\mathbb{P}_{1\leqslant j\leqslant n}((\mathcal{E}_1(j)\cap\mathcal{E}_2(j))\cap\mathcal{E}_4(j))\geqslant\frac{\epsilon}{4C_5}-\epsilon'-\epsilon_1\geqslant\frac{\epsilon}{10C_5}-\epsilon_1,$$

where $\epsilon'<\epsilon/10C_5$ due to Eq. (5.35). Therefore

$$\mathbb{E}_{1\leqslant j\leqslant n,\ell\sim\theta}\left(\int\frac{1}{k}H(S_{t(\ell,x)}T_{[\ell(x)]}[\theta_{\ell,j}.x]*\mathfrak{m}_2(\ell,j,\mathbf{i}),\mathcal{Q}_{j+k})\mathrm{d}(\mu_{\varphi(j,\mathbf{i})})_{\mathbf{I}}(x),\mathcal{E}_0'(j)\right)$$
$$\geqslant(\delta_1/2)(\frac{\epsilon}{10C_5}-\epsilon_1)+(\alpha-\epsilon_1)(1-\epsilon_1)-O(1/k)=\alpha+\frac{\delta_1\epsilon}{20C_5}-\epsilon_1(\alpha+\delta_1)-O(1/k).$$

Then from Eq. (5.28), we obtain

$$\frac{1}{n}H([\theta.\mu],\mathcal{Q}_{t+n})\geqslant\alpha+\delta_1\epsilon/5C-\epsilon_1-O(\frac{k}{n}+r^{\beta}+\frac{\log(C_1/r)}{k}),$$

where $\delta_1\epsilon/C$ is constant and other terms can be made arbitrarily small by taking ϵ_1 small, then k and n/k large and $1/r=\exp(\sqrt{k}/10)$. $\square$

6 Exponential separation

Recall the exponential separation condition for ν supported on $\mathrm{SL}_3(\mathbb{R})$: there exist $C>0$ and $n_C\in\mathbb{N}$ such that for any $n\geqslant n_C$ any $g\neq g'$ in $\operatorname{supp}\nu^{*n}$, we have

$$d(g,g')>1/C^n,$$

where d is the left-invariant Riemannian metric on $\mathrm{SL}_3(\mathbb{R})$. Notice that from Lemma 2.12, for any $g\in\mathrm{SL}_3(\mathbb{R})$ such that $g^{-1}V\notin V^{\perp}$, the action of $\pi_{V^{\perp}}g$ from $\mathbb{P}(\mathbb{R}^3)$ to $\mathbb{P}(V^{\perp})$ can be identified by an element ℓ in L_V.

Lemma 6.1. *For a fixed $g\in\mathrm{SL}_3(\mathbb{R})$, the set of V such that $g^{-1}V\in V^{\perp}$ is contained in a hyperplane in $\mathbb{P}(Sym^2\mathbb{R}^3)$ through the embedding p_2 (defined in Lemma 2.26).*

Proof. The set of such V is given by the zero locus of the equation

$$\langle g^{-1}v,v\rangle=0,$$

where v is a non-zero vector in V. For $v,w\in\mathbb{R}^3$ let $f(v,w)=\langle g^{-1}v,w\rangle+\langle g^{-1}w,v\rangle$, which is non-trivial. Then due to $f(w,v)=f(v,w)$, the linear form f on $\mathbb{R}^3\otimes\mathbb{R}^3$ induces a linear form on $Sym^2\mathbb{R}^3$. Since $\langle g^{-1}v,v\rangle=0$ is equivalent to $f(v,v)=0$, we obtain the lemma. $\square$

Recall that we have a left-invariant Riemannian distance on L_V. The action of $\mathrm{SL}_3(\mathbb{R})$ on row vectors W is by right multiplication. We say $W\in\mathbb{P}(V^{\perp})$ if $W\cdot V=0$, where $W\cdot V$ is the product of a row vector with a column vector. Recall from Eq. (2.6) that π_{L_V} is a map from a dense open set of $\mathrm{SL}_3(\mathbb{R})$ to L_V.

Lemma 6.2. *There exists a family of small neighbourhoods $Z_V=B(id,\epsilon_Z)$ of L_V for some $\epsilon_Z>0$ independent of V and $C(Z)>0$ such that for any V, any $u\in U_V$ and $\ell\in Z_V$*

$$d(\pi_{L_V}u\ell,\pi_{L_V}id)=d(\pi_{L_V}\ell,\pi_{L_V}id)\simeq_{C(Z)}\sup_{W\in\mathbb{P}(V^{\perp})}d(W\ell,W)=\sup_{W\in\mathbb{P}(V^{\perp})}d(Wu\ell,W).$$

Here the first two d are just the left-invariant distance in L_V, the last two d are the spherical distance on the projective space.

In particular, there exists a small neighbourhood B_ϵ of identity in $\mathrm{SL}_3(\mathbb{R})$, we have for $g\in B_\epsilon$

$$d(\pi_{L_V}g,\pi_{L_V}id)\simeq_{C(Z)}\sup_{W\in\mathbb{P}(V^{\perp})}d(Wg,W).$$

Proof. Without loss of generality we can assume $V = E_1$. Recall that the element $\ell \in L$ can be written as $\ell = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix}$. For $(0, w_1, w_2)$ a unit row vector in $W \in V^{\perp}$, we denote (w_1, w_2) by $\mathbf{w}$. Then we have

$$d(W\ell, W) = \frac{\|(0,\mathbf{w})\ell \wedge (0,\mathbf{w})\|}{\|(0,\mathbf{w})\ell\|\|(0,\mathbf{w})\|}.$$

Since ℓ is close to the identity, the vector $(0,\mathbf{w})\ell$ has length close to 1. So we have

$$\begin{aligned} d(W\ell, W) &\simeq \|(0,\mathbf{w})\ell \wedge (0,\mathbf{w})\| \simeq \|\mathbf{w}\cdot n\| + \|\mathbf{w}h \wedge \mathbf{w}\| \\ &\simeq \|\mathbf{w}\cdot n\| + d(\mathbb{P}(\mathbf{w}h), \mathbb{P}(\mathbf{w})). \end{aligned}$$

Notice that near the identity, the metric induced by the norm and the Riemannian distance are bi-Lipschitz equivalent. Therefore we have $d(W\ell, W) \leqslant \|n\| + \|h - id\| \simeq d(\ell, id)$.

On the other hand, we have

$$\sup_{\mathbf{w}} \|\mathbf{w}\cdot n\| = \|n\|.$$

For any triple of unit vectors $\mathbf{w}_i, i = 1,2,3$ which is $1/10$ separated,

$$d_{\mathrm{SL}_2(\mathbb{R})}(h, id) \ll \sup_{i=1,2,3}(d(\mathbb{P}(\mathbf{w}_i h), \mathbb{P}(\mathbf{w}_i))) \leqslant \sup_{\mathbf{w}} d(\mathbb{P}(\mathbf{w}h), \mathbb{P}(\mathbf{w})) \leqslant \sup_{W\in\mathbb{P}(V^{\perp})} d(W\ell, W).$$

Here we use the fact that the map $h \mapsto (\mathbf{w}_1 h, \mathbf{w}_2 h, \mathbf{w}_3 h), \mathrm{PSL}_2(\mathbb{R}) \to \mathbb{P}(\mathbb{R}^2)^3$ is a smooth injection and bi-Lipschitz to its image on a small neighborhood of identity, see for example page 827 in [HS17]. Therefore $d(\ell, id) \ll \sup_{W\in\mathbb{P}(V^{\perp})} d(W\ell, W)$ and

$$\sup_{W\in\mathbb{P}(V^{\perp})} d(W\ell, W) \gg d(h, id) + \|n\| \gg d(\ell, id).$$

We can add U_V-part since it acts trivially on the left on $V^{\perp}$.

The last statement is due to the product map from $U \times Z_V$ to $\mathrm{SL}_3(\mathbb{R})$ contains a neighborhood of identity. ☐

Lemma 6.3. *There exists constant $c > 0$ such that the following holds. For any $g \in \mathrm{SL}_3(\mathbb{R})$ and $\epsilon > 0$ which is sufficiently small, if $d(id, g) > \epsilon^{1/3}$ and $d(\pi_{L_V} id, \pi_{L_V} g) \leqslant \epsilon$, then for any V' with $d(V', V) \geqslant c\epsilon^{1/3}$, we have $d(\pi_{L_{V'}} id, \pi_{L_{V'}} g) > \epsilon$.*

Proof. Suppose that our lemma does not hold, i.e. $d(\pi_{L_{V'}} id, \pi_{L_{V'}} g) \leqslant \epsilon$ for some V' with $d(V', V) \geqslant c\epsilon^{1/3}$, the constant c will be defined at the end of the proof of the lemma. For simplicity we can just assume $V = E_1$. without loss of generality we could also assume $V' = \mathbb{R}(e_1 + \delta e_2)$ where $|\delta| < 1$ and $\delta \geqslant d(V, V') \geqslant c\epsilon^{1/3}$.

Consider the UL decomposition $g = u\ell$ of g. we write $u = \begin{pmatrix} \lambda^2 & x & y \\ 0 & \lambda^{-1} & 0 \\ 0 & 0 & \lambda^{-1} \end{pmatrix}$. **The idea is to use the condition $d(\pi_{L_{V'}} id, \pi_{L_{V'}} g) \leqslant \epsilon$ to obtain the u part is also small, which contradicts the hypothesis $d(id, g) > \epsilon^{1/3}$.**

We apply Lemma 6.2 to $d(\pi_{L_{V'}} id, \pi_{L_{V'}} g) \leqslant \epsilon$ and using $V' = \mathbb{R}(e_1 + \delta e_2)$, we obtain

$$\begin{aligned} d(\delta e_1^t - e_2^t, (\delta e_1^t - e_2^t)g) &\leqslant c_2\epsilon, \\ d(\delta e_1^t - e_2^t + e_3^t, (\delta e_1^t - e_2^t + e_3^t)g) &\leqslant c_2\epsilon, \end{aligned} \tag{6.1}$$

where these vectors are chosen in $(V')^{\perp}$ but not in $V^{\perp}$. Due to $d_L(\ell, id) \leqslant \epsilon$, which implies $d_{\mathrm{SL}_3(\mathbb{R})}(\ell, id) \leqslant O(\epsilon)$, hence ℓ moves any vector in $\mathbb{P}(\mathbb{R}^3)$ of distance at most $c_3\epsilon$ for some $c_3 > 0$

if ϵ is small enough. Therefore by triangle inequality and Eq. (6.1) we obtain

$$\begin{aligned}
& d(\delta e_1^t - e_2^t, \delta e_1^t + (-\lambda^{-3} + \delta x\lambda^{-2})e_2^t + \delta y\lambda^{-2}e_3^t) \\
= \;& d(\delta e_1^t - e_2^t, (\delta e_1^t - e_2^t)u) \\
\leqslant \;& d(\delta e_1^t - e_2^t, (\delta e_1^t - e_2^t)g) + d((\delta e_1^t - e_2^t)u\ell, (\delta e_1^t - e_2^t)u) \\
\leqslant \;& c_2\epsilon + c_3\epsilon \leqslant c_4\epsilon
\end{aligned} \tag{6.2}$$

for some constant $c_4 > 0$. Similarly we get

$$d(\delta e_1^t - e_2^t + e_3^t, \delta e_1^t + (-\lambda^{-3} + \delta x\lambda^{-2})e_2^t + (\delta y\lambda^{-2} + \lambda^{-3})e_3^t) \leqslant c_4\epsilon. \tag{6.3}$$

We consider the following elementary geometric lemma.

Lemma 6.4. *There exists $c_5 > 0$ such that for any $\delta \geqslant 6c_4\epsilon$ with ϵ small enough, if $w, w' \in \mathrm{Span}(e_2^t, e_3^t)$ and $1 \leqslant \|w\| \leqslant 2$, $d_{\mathbb{P}(\mathbb{R}^3)}(\delta e_1^t + w, \delta e_1^t + w') < c_4\epsilon$. Then*

$$\|w - w'\| \leqslant c_5 \cdot \frac{\epsilon}{\delta}.$$

Proof. Consider the triangle $A_1A_2A_3$ contained in $\mathbb{R}^3$, where $A_1 = 0 \in \mathbb{R}^3$, $A_2 = \delta e_1^t + w$, $A_3 = \delta e_1^t + w'$, then by the sine rule we have

$$\frac{\|A_2A_3\|}{\|A_1A_2\|} = \frac{|\sin\angle A_2A_1A_3|}{|\sin\angle A_1A_3A_2|}. \tag{6.4}$$

By our assumption of ϵ, δ,

$$1 \leqslant \|A_1A_2\| \leqslant 2, \quad \|A_2A_3\| = \|w - w'\|, \quad |\sin\angle A_2A_1A_3| \leqslant c_4\epsilon.$$

From the third one and $w, w' \in \mathrm{Span}(e_2^t, e_3^t)$ we have

$$\delta\|w - w'\| \leqslant \|(\delta e_1^t + w) \wedge (\delta e_1^t + w')\| \leqslant c_4\epsilon\|\delta e_1^t + w\|\|\delta e_1^t + w'\| \leqslant c_4\epsilon \cdot 3(1 + \|w'\|).$$

Therefore

$$(1 - \frac{3c_4\epsilon}{\delta})\|w'\| \leqslant \frac{3c_4\epsilon}{\delta} + 2.$$

This implies

$$\|w'\| \leqslant 6, \text{ and } |\sin\angle A_1A_3A_2| = \frac{\|(\delta e_1^t + w') \wedge (w - w')\|}{\|\delta e_1^t + w'\|\|w - w'\|} \geqslant \frac{\delta\|w - w'\|}{\|\delta e_1^t + w'\|\|w - w'\|} \geqslant \delta/7.$$

Therefore by Eq. (6.4)

$$\|w - w'\| \leqslant c_5 \cdot \frac{\epsilon}{\delta}$$

for some $c_5 > 0$. ◻

Applying Lemma 6.4 to (6.2),(6.3), and (w, w') to be $(-e_2^t, (-\lambda^{-3} + \delta x\lambda^{-2})e_2^t + \delta y\lambda^{-2}e_3^t)$ in (6.2) and $(-e_2^t + e_3^t, (-\lambda^{-3} + \delta x\lambda^{-2})e_2^t + (\delta y\lambda^{-2} + \lambda^{-3})e_3^t)$ in (6.3), we get

$$\|(1 - \lambda^{-3} + \delta x\lambda^{-2})e_2^t + \delta y\lambda^{-2}e_3^t\| \leqslant c_5\frac{\epsilon}{\delta}, \tag{6.5}$$

$$\|(1 - \lambda^{-3} + \delta x\lambda^{-2})e_2^t + (-1 + \delta y\lambda^{-2} + \lambda^{-3})e_3^t\| \leqslant c_5\frac{\epsilon}{\delta}. \tag{6.6}$$

Due to $\delta \geqslant c\epsilon^{1/3}$, by taking c large enough (compare to $100c_i, i = 2, \cdots, 5$) and comparing the coefficients of e_3^t, we obtain $|\delta y\lambda^{-2}|, |1 - \lambda^{-3}| \leqslant 2c_5\frac{\epsilon}{\delta}$, which implies $|\lambda - 1| \leqslant \epsilon^{2/3}/10$ and $|\delta y| \leqslant \epsilon^{2/3}/10$. From coefficients of e_2^t, we obtain $|\delta x| \leqslant \epsilon^{2/3}/10$. Therefore we obtain $|x|, |y| \leqslant \epsilon^{1/3}/10$ and $|\lambda - 1| \leqslant \epsilon^{1/3}/10$. Then we obtain a contradiction from $d(id, g) > \epsilon^{1/3}$ if ϵ is small enough. ◻

Lemma 6.5. *There exists $c' > 0$ such that for any $C > 0$ large enough, any $g \neq g'$ in $\mathrm{SL}_3(\mathbb{R})$ with $d(g, g') > 1/C$, the set*

$$\{V \in \mathbb{P}(\mathbb{R}^3) : \ d(\pi_{L_V} g, \pi_{L_V} g') < 1/C^4\} \tag{6.7}$$

has diameter less than $c'\|g\|^6/C$.

Proof. By left-invariance of d on $\mathrm{SL}_3(\mathbb{R})$, we have $d(g, g') = d(id, g^{-1}g')$. For any V in the set defined in (6.7), we plan to obtain information of $\pi_{L_{g^{-1}V}} g^{-1}g'$ and then apply Lemma 6.5. For such V, we consider the U_V, L_V decomposition of g, g' respectively, $g = u\ell$ and $g' = u'\ell'$. Then by (6.7) we have

$$d(\pi_{L_V} g, \pi_{L_V} g') = d_{L_V}(\ell, \ell') = d_{L_V}(id, \ell^{-1}\ell') \leqslant 1/C^4.$$

So the element $\ell^{-1}\ell'$ is close to id, also in the sense of $\mathrm{SL}_3(\mathbb{R})$. Since we assume that C is large enough, we can assume $\ell^{-1}\ell'$ moves point on $\mathbb{P}(\mathbb{R}^3)$ with distance not greater than $C^{-7/2}$. For any row vector $W \in (\ell^{-1}V)^{\perp}$, which means $0 = W \cdot (\ell^{-1}V) = (W\ell^{-1}) \cdot V$, we have $W\ell^{-1} \in V^{\perp}$. Due to $g^{-1}g' = \ell^{-1}u^{-1}u'\ell'$, the left action of U on column vectors in $\mathbb{P}(V^{\perp})$ being trivial, so

$$\sup_{W \in (\ell^{-1}V)^{\perp}} d(W, Wg^{-1}g') = \sup_{W \in (\ell^{-1}V)^{\perp}} d(W, W\ell^{-1}\ell') < 1/C^{7/2}.$$

This is the main advantage of considering row vectors. The previous estimate implies that $\pi_{L_{\ell^{-1}V}} g^{-1}g'$ is close to identity in $L_{\ell^{-1}V}$.[6] Then we can apply Lemma 6.2 to $g^{-1}g'$, for C sufficiently large, we have

$$d(\pi_{L_{\ell^{-1}V}} id, \pi_{L_{\ell^{-1}V}} g^{-1}g') \leqslant 1/C^3.$$

Then we can apply Lemma 6.3 to obtain that the set of $V' = \ell^{-1}V = g^{-1}V$ such that the previous inequality holds has radius less than c'/C, where $c' = 2c$ and c is the constant we get in Lemma 6.3. So such $V = gV'$ has radius less than $c'\|g\|^6/C$, since the Lipschitz constant of g on projective space is bounded by $\|g\|^6$ (See for example (13.2) in [BQ16]). □

Once we have this lemma, we can take similar arguments as in [BHR19], [HS17] and [Hoc14] to obtain the main proposition of this section.

Proposition 6.6. *Suppose that ν satisfies the exponential separation condition. Then there exist $C > 0$ and a subset $Y \subset \mathbb{P}(\mathbb{R}^3)$ with μ^- measure zero such that for every $V \in \mathbb{P}(\mathbb{R}^3) - Y$, there exists N_V and for $n \geqslant N_V$, we have for $g \neq g' \in \operatorname{supp} \nu^n$*

$$d(\pi_{L_V} g, \pi_{L_V} g') > C^{-n}.$$

Proof. For $\mathbf{i}, \mathbf{j}$ in Λ^n with $g_{\mathbf{i}} \neq g_{\mathbf{j}}$, let

$$E_{\mathbf{i},\mathbf{j}}(C) := \{V \in \mathbb{P}(\mathbb{R}^3), \ d(\pi_{L_V} g_{\mathbf{i}}, \pi_{L_V} g_{\mathbf{j}}) < 1/C^n\}.$$

We consider the set

$$E_1(C) := \bigcap_{m \in \mathbb{N}} \bigcup_{n \geqslant m} \bigcup_{\mathbf{i},\mathbf{j} \in \Lambda^n} E_{\mathbf{i},\mathbf{j}}(C).$$

By Lemma 6.1 we know that for any V not in $E_1(C)$ and the countable union of subvariety with μ^--measure 0 in Lemma 6.1, there exists $N(V) \geqslant 1$ such that for any $g \neq g' \in \operatorname{supp}(\nu^n)$, we

[6] Because the condition implies for any unit vector w in $(\ell^{-1}V)^{\perp}$

$$\|wh \wedge w\| + |w \cdot n| \leqslant (\|wh\| + |w \cdot n|)/C^{7/2}.$$

Hence $\sup_{W \in (\ell^{-1}V)^{\perp}} d(W, Wh) \leqslant 1/C^{7/2}$ for the h part in $\pi_{L_{\ell^{-1}V}} g^{-1}g'$. We can obtain the h-part close to identity: use $1 \simeq \frac{d(v_1h, v_2h)}{d(v_1, v_2)} = \frac{|v_1||v_2|}{|v_1h||v_2h|}$ for $d(v_1, v_2)$ not too small. Then take unit vectors v_1, v_2 such that $\|v_1h\| = \|h\|$. This implies $\|h\|$ close to 1. Then use Cartan decomposition to obtain h close to id. Finally, use the first formula to obtain that $\|n\|$ is close to zero.

know that $n \geqslant N(V)$, $\pi_{L_V} g, \pi_{L_V} g'$ are well-defined, and $d(\pi_{L_V} g, \pi_{L_V} g') \geqslant C^{-n}$. Hence we only need to show that $E_1(C)$ has μ^--measure 0 if C is large enough.

Notice that by Guivarc'h theorem (Lemma 2.24) on Hölder regularity of μ^- and Frostman lemma, there exists $s(\mu^-) > 0$ such that any subset of $\mathbb{P}(\mathbb{R}^3)$ with positive μ^- measure set has Hausdorff dimension at least $s(\mu^-)$. Then it is suffice to show that for any $s > 0$, we can choose C large enough such that $\dim(E_1(C)) \leqslant s$. We apply Lemma 6.5 to estimate the Hausdorff dimension of $E_1(C)$.

Fix an $s > 0$ small enough and write $\mathcal{H}^s_\infty$ for the s-dimensional Hausdorff content on $\mathbb{P}(\mathbb{R}^3)$. By exponential separation condition of ν, for n large, we have C' such that for any $g \neq g' \in \operatorname{supp} \nu^n$, $d(g, g') > C'^{-n}$. Then for $C \gg C'$ we get

$$\begin{aligned}
\mathcal{H}^s_\infty(E_1(C)) &\leqslant \lim_{N\to\infty} \sum_{n=N}^{\infty} \sum_{\mathbf{i},\mathbf{j}\in\Lambda^n} (\operatorname{diam}\{E_{\mathbf{i},\mathbf{j}}(C)\})^s \\
&\leqslant \lim_{N\to\infty} \sum_{n=N}^{\infty} \sum_{\mathbf{i},\mathbf{j}\in\Lambda^n} \left(\frac{\|g_{\mathbf{i}}\|^6 c'}{C^{n/4}}\right)^s \\
&\quad (\text{ apply Lemma 6.5 with } C \text{ large such that } d(g_{\mathbf{i}}, g_{\mathbf{j}}) > C^{-n/4}) \\
&\leqslant \lim_{N\to\infty} \sum_{n=N}^{\infty} \sum_{\mathbf{i},\mathbf{j}\in\Lambda^n} \left(\frac{b}{C^{1/4}}\right)^{sn} \ (\text{here } b := \sup_{g\in\operatorname{supp}\nu} \|g\|^6 \cdot c') \\
&\leqslant \lim_{N\to\infty} \sum_{n=N}^{\infty} D^{2n} \left(\frac{b}{C^{1/4}}\right)^{sn} \ (\text{here } D := \#\{\operatorname{supp}\nu\}) \\
&\leqslant 0 \ (\text{by taking } C \text{ large enough compare to } b, D).
\end{aligned}$$

Therefore the Hausdorff dimension of $E_1(C)$ is not greater than s for C sufficiently large. □

7 Proofs of Theorem 1.10 and Theorem 1.11

In this section, after preparations (Lemma 7.1, Lemma 7.2 and Lemma 7.4), we first apply Theorem 5.7 to show Theorem 1.11, then combine it with a Lerappier-Young formula recently shown in [LL23a] and [Rap21] to conclude Theorem 1.10.

7.1 Preparations: estimates for the entropy

Recall $\mathcal{Q}_n^{L_V}$ is the $q-$adic partition of the group L_V. In the case that there is no ambiguity, we may write it as $\mathcal{Q}_n$. Consider the set

$$R := \{V \in \mathbb{P}(\mathbb{R}^3) : g^{-1}V \notin V^\perp, \forall g \in \cup_{n\geqslant 0} \operatorname{supp} \nu^{*n}\}.$$

By Lemma 2.26 and Lemma 6.1, R is a μ^--full measure set. Recall that π_{L_V} is a projection from $U_V L_V$ to L_V defined in Eq. (2.6). Then for $\nu^{*n}, n \geqslant 1$ on $\mathrm{SL}_3(\mathbb{R})$ and $V \in R$, the measure $\pi_{L_V}\nu^{*n}$ is a well-defined measure on L_V. From the exponential separation condition, we have

Lemma 7.1. *There exists $C > 1$ such that for μ^--almost every $V \in \mathbb{P}(\mathbb{R}^3)$,*

$$\lim_{n\to\infty} \frac{1}{n} H(\pi_{L_V}\nu^{*n}, \mathcal{Q}_{Cn}^{L_V}) = h_{\mathrm{RW}}(\nu).$$

Proof. By Proposition 6.6, there exists $C > 0$ such that for V in a full μ^--measure set, distinct atoms of $\pi_{L_V}\nu^{*n}$ have distance at least $q^{-Cn/2}$ from each other for any n large enough. On the other hand, the size of atoms of the partition $\mathcal{Q}_{Cn}$ is about q^{-Cn}. So for n large, the atoms of $\pi_{L_V}\nu^{*n}$ are located in different atoms in $\mathcal{Q}_{Cn}$. Therefore

$$\lim_{n\to\infty} \frac{1}{n} H(\pi_{L_V}\nu^{*n}, \mathcal{Q}_{Cn}) = \lim_{n\to\infty} \frac{1}{n} H(\pi_{L_V}\nu^{*n}) = \lim_{n\to\infty} \frac{1}{n} H(\nu^{*n}) = h_{\mathrm{RW}}(\nu),$$

where the second equality is due to the injectivity of the map π_{L_V} on $\operatorname{supp}\nu^{*n}$ given in Proposition 6.6. $\square$

Lemma 7.2. *Let $C > 1$ and $\tau > 0$. Then for μ^--almost every $V \in \mathbb{P}(\mathbb{R}^3)$,*

$$\limsup_{n\to\infty} \pi_{L_V}\nu^{*n}\{\ell \in L_V : \frac{1}{Cn}H(\pi_{V^\perp}\ell\mu, \mathcal{Q}_{(C+\chi_1)n}) > \alpha - \tau\} = 1.$$

Proof. Recall that for $g \in U_V L_V$, $(\pi_{L_V} g)\mu = \pi_{V^\perp} g\mu$, and the measure ν^{*n} is the law of $\mathbf{U}(n)$. So it suffices to show that for any $\epsilon > 0$, for μ^--almost every V,

$$\limsup_{n\to\infty} \mathbb{P}\left\{\mathbf{U}(n) : \frac{1}{Cn}H(\pi_{V^\perp} g_{\mathbf{U}(n)}\mu, \mathcal{Q}_{(C+\chi_1)n}) > \alpha - \tau\right\} \geqslant 1 - \epsilon. \tag{7.1}$$

Fix a small $\epsilon > 0$ in (7.1). By Lemma 4.12, Lemma 4.13 (let δ in Lemma 4.13 to be $\tau/4$) and Eq. (2.52), there exists some small $r = r(\epsilon, \tau) > 0$ such that for any n large, we can construct there a set $X_{n,\epsilon}$ of $\mathbf{U}(n)$ such that

1. $\mathbb{P}(\mathbf{U}(n) \in X_{n,\epsilon}) > 1 - \epsilon/2$.
2. For any $\mathbf{U}(n) \in X_{n,\epsilon}$, $|\chi_1(g_{\mathbf{U}(n)}) - \chi_1 n| \leqslant \tau n/10$.
3. For any $\mathbf{U}(n) \in X_{n,\epsilon}, n \geqslant N(r)$, the r-attracting part $(\pi_{V^\perp} g_{\mathbf{U}(n)}\mu)_{\mathbf{I}}$ of the pair $(h_{V,\mathbf{U}(n)}, \pi_{V,\mathbf{U}(n)}\mu)$ satisfies

$$\frac{1}{Cn}H((\pi_{V^\perp} g_{\mathbf{U}(n)}\mu)_{\mathbf{I}}, \mathcal{Q}_{Cn+\chi_1(g_{\mathbf{U}(n)})}) \geqslant \frac{1}{Cn}H(\pi_{(g^{-1}_{\mathbf{U}(n)}V)^\perp}\mu, \mathcal{Q}_{Cn}) - \frac{\tau}{10}, \tag{7.2}$$

   with $\pi_{V^\perp} g_{\mathbf{U}(n)}\mu = (1-\tilde{\delta})(\pi_{V^\perp} g_{\mathbf{U}(n)}\mu)_{\mathbf{I}} + \tilde{\delta}(\pi_{V^\perp} g_{\mathbf{U}(n)}\mu)_{\mathbf{II}}$ and $\tilde{\delta} \leqslant \tau/4$.

By the 2nd condition of $X_{n,\epsilon}$ and Eq. (2.75), for $\mathbf{U}(n) \in X_{n,\epsilon}$, replacing $\chi_1(g_{\mathbf{U}(n)})$ by $\chi_1 n$ in LFS of Eq. (7.2) changes the value of LFS less than $\tau/5$. So we obtain

$$\frac{1}{Cn}H((\pi_{V^\perp} g_{\mathbf{U}(n)}\mu)_{\mathbf{I}}, \mathcal{Q}_{(C+\chi_1)n}) \geqslant \frac{1}{Cn}H(\pi_{(g^{-1}_{\mathbf{U}(n)}V)^\perp}\mu, \mathcal{Q}_{Cn}) - \frac{\tau}{2}.$$

Then by the concavity of entropy, for $\mathbf{U}(n) \in X_{n,\epsilon}$

$$\begin{aligned}\frac{1}{Cn}H(\pi_{V^\perp} g_{\mathbf{U}(n)}\mu, \mathcal{Q}_{(C+\chi_1)n}) &\geqslant (1-\tilde{\delta})\frac{1}{Cn}H((\pi_{V^\perp} g_{\mathbf{U}(n)}\mu)_{\mathbf{I}}, \mathcal{Q}_{(C+\chi_1)n}) \\ &\geqslant (1-\tilde{\delta})(\frac{1}{Cn}H(\pi_{(g^{-1}_{\mathbf{U}(n)}V)^\perp}\mu, \mathcal{Q}_{Cn}) - \frac{\tau}{2}) \\ &\geqslant \frac{1}{Cn}H(\pi_{(g^{-1}_{\mathbf{U}(n)}V)^\perp}\mu, \mathcal{Q}_{Cn}) - \frac{3\tau}{4},\end{aligned} \tag{7.3}$$

where the last inequality is due to $\tilde{\delta} \leqslant \tau/4$.

On the other hand, we can show the following lemma which is similar to Lemma 6.2 in [BHR19].

Lemma 7.3. *Let $\tau > 0$, $\epsilon > 0$. For μ^--almost every $V \in \mathbb{P}(\mathbb{R}^3)$, there exists a sequence $n_k \to \infty$ such that*

$$\mathbb{P}\left\{\mathbf{U}(n_k) : \frac{1}{Cn_k}H(\pi_{(g^{-1}_{\mathbf{U}(n_k)}V)^\perp}\mu, \mathcal{Q}_{Cn_k}) > \alpha - \tau/10\right\} > 1 - \epsilon/10.$$

Proof. Let $f_n(\omega, V)$ be a function on $\Lambda^{\mathbb{N}} \times \mathbb{P}(\mathbb{R}^3)$ defined by

$$f_n(\omega, V) = \begin{cases} 1, & \text{if } \frac{1}{Cn}H(\pi_{V^\perp}\mu, \mathcal{Q}_{Cn}) > \alpha - \tau/10 \\ 0, & \text{otherwise.}\end{cases}$$

By exact dimensionality of $\pi_{V^\perp}\mu$, f_n converges $\nu^{\otimes\mathbb{N}} \otimes \mu^-$-a.e. to 1 as $n \to \infty$ (Lemma 2.38).

We define a skew-product transformation T on $\Lambda^{\mathbb{N}} \times \mathbb{P}(\mathbb{R}^3)$ by

$$T(\omega, V) = (T\omega, g_{\omega_1}^{-1}V),$$

where $T\omega$ is the shift map given by $(T\omega)_j = \omega_{j+1}$. Due to stationarity of μ^- under ν^-, the measure $\nu^{\otimes\mathbb{N}} \otimes \mu^-$ is T-invariant. Moreover, because μ^- is the unique ν^--stationary measure, $\nu^{\otimes\mathbb{N}} \otimes \mu^-$ is also T-ergodic, see for example [BQ16, Proposition 2.23].

Therefore by Maker's ergodic theorem, for $\nu^{\otimes\mathbb{N}} \otimes \mu^-$-a.e. (ω, V), we have

$$\frac{1}{N}\sum_{n=1}^{N} f_n(T^n(\omega, V)) \to 1 \text{ as } N \to \infty.$$

Integrating the above equation over the ω-component, by dominated convergence theorem, we get for μ^--a.e. V,

$$\frac{1}{N}\sum_{n=1}^{N} \mathbb{P}\left\{\mathbf{U}(n) : \frac{1}{Cn} H(\pi_{(g_{\mathbf{U}(n)}^{-1}V)^\perp}\mu, \mathcal{Q}_{Cn}) > \alpha - \tau/10\right\} \to 1 \text{ as } N \to \infty.$$

Therefore we can find a sequence $n_k \to \infty$ such that

$$\mathbb{P}\left\{\mathbf{U}(n_k) : \frac{1}{Cn_k} H(\pi_{(g_{\mathbf{U}(n_k)}^{-1}V)^\perp}\mu, \mathcal{Q}_{Cn_k}) > \alpha - \tau/10\right\} > 1 - \epsilon/10. \qquad \square$$

Combine Lemma 7.3 with (7.3). Let $\{n_k\}$ be the infinite sequence given in Lemma 7.3. For all large n_k, $\frac{1}{Cn} H(\pi_{V^\perp} g_{\mathbf{U}(n)}\mu, \mathcal{Q}_{(C+\chi_1)n}) > \alpha - \tau$ holds for a set of $\mathbf{U}(n_k)$ with measure greater than $1 - \epsilon$. Hence

$$\mathbb{P}\left\{\mathbf{U}(n_k) : \frac{1}{Cn_k} H(\pi_{V^\perp} g_{\mathbf{U}(n_k)}\mu, \mathcal{Q}_{(C+\chi_1)n_k}) \geqslant \alpha - \tau\right\} > 1 - \epsilon. \qquad \square$$

Lemma 7.4. *For μ^--almost every $V \in \mathbb{P}(\mathbb{R}^3)$, we have*

$$\lim_{n\to\infty} \frac{1}{n} H(\pi_{L_V}\nu^{*n}, \mathcal{Q}_1^{L_V}) \leqslant \alpha\chi_1(\nu).$$

Proof. By Lemma 2.38, for μ^-- a.e. V we have

$$\lim_{n\to\infty} \frac{1}{\chi_1 n} H(\pi_{V^\perp}\mu, \mathcal{Q}_{\chi_1 n}) \to \alpha. \tag{7.4}$$

Fix any V for which Eq. (7.4) holds. By conjugating with an element in $\mathrm{SO}_3(\mathbb{R})$, we may assume $V = E_1$.

Let X_i be matrix-valued i.i.d. random variables with the law ν. Denote by $Z_n(\omega) := X_1(\omega)\cdots X_n(\omega), \omega \in \Lambda^{\times\mathbb{N}}$. By Furstenberg's theorem (see for example [BQ16, Proposition 4.7]), the following Furstenberg boundary map is well-defined for $\nu^{\otimes\mathbb{N}}$-almost everywhere.

$$\xi : (\Lambda^{\times\mathbb{N}}, \nu^{\otimes\mathbb{N}}) \to \mathbb{P}(\mathbb{R}^3), \xi(\omega) := \lim_{n\to\infty} Z_n(\omega)_*\mu, ,$$

and satisfies that $\mu = \mathbb{E}(\delta_{\xi(\omega)})$ and $\xi(\omega) = \omega_1\xi(T\omega), \nu^{\otimes\mathbb{N}} - a.e.\omega$, where $\mathbb{T} : \Lambda^{\times\mathbb{N}} \to \Lambda^{\times\mathbb{N}}$ is the shift map.

For each element $g \in \mathrm{SL}_3(\mathbb{R})$, we consider its UL-decomposition: $g = u\ell = u\begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix}$ with $u \in U$ and $\ell \in L$. Recall h^+ is the contracting point of the h part of ℓ in $\mathbb{P}({E_1}^\perp)$. Let $v_n = h(Z_n)^+$. Then we can state the following exponential convergence lemma.

Lemma 7.5. *For any $\epsilon > 0$, there exists $c(\epsilon) > 0$ such that*

$$\mathbb{P}\{\omega \in \Lambda^{\times \mathbb{N}} : d_{\mathbb{P}(\mathbb{R}^2)}(\pi_{E_1{}^\perp}\xi(\omega), v_n) < q^{-(\chi_1-\epsilon)n}\} = 1 - O(e^{-c(\epsilon)n}). \tag{7.5}$$

Proof. Fix any $x \in \mathbb{P}(E_1^\perp)$ and let $w_n(\omega) = Z_n(\omega)x$ in $\mathbb{P}(\mathbb{R}^3)$.

Claim.

$$\mathbb{P}\{\omega \in \Lambda^{\times \mathbb{N}} : d_{\mathbb{P}(\mathbb{R}^3)}(\xi(\omega), w_n) < q^{-(\chi_1-\epsilon/10)n}\} = 1 - O(e^{-c(\epsilon)n}).$$

Proof. Due to the equivariance of ξ, we have

$$d(\xi(\omega), w_n) = d(Z_n\xi(T^n\omega), Z_n x).$$

By Lemma 2.4, if we have

$$\chi_1(Z_n) > (\chi_1 - \epsilon/20)n,\ d(\xi(T^n\omega), H^-_{Z_n}) > q^{-\epsilon n/40} \text{ and } d(x, H^-_{Z_n}) > q^{-\epsilon n/40}, \tag{7.6}$$

then we have

$$d(Z_n\xi(T^n\omega), Z_n x) < q^{-\chi_1(Z_n)+\epsilon n/20} \leqslant q^{-(\chi_1-\epsilon/10)n}.$$

Notice that the random variable Z_n has the law ν^{*n}. Then for fixed $\xi(T^n w)$, by Eqs. (2.50) and (2.52) we know Eq. (7.6) holds with ν^{*n}-probability greater than $1 - O(e^{-cn})$. Since $T^n\omega$ and Z_n are independent, we obtain the claim by applying a Fubini argument. □

We back to the proof of Eq. (7.5). Observe that

$$d_{\mathbb{P}(\mathbb{R}^2)}(\pi_{E_1{}^\perp}\xi(\omega), v_n) \leqslant d_{\mathbb{P}(\mathbb{R}^2)}(\pi_{E_1{}^\perp}\xi(\omega), \pi_{E_1^\perp}w_n) + d_{\mathbb{P}(\mathbb{R}^2)}(\pi_{E_1^\perp}w_n, v_n). \tag{7.7}$$

It suffices to show both terms on RHS of (7.7) are less than $\frac{1}{2}q^{-(\chi_1-\epsilon)n}$ with probability at least $(1 - O(e^{-c(\epsilon)n}))$.

For the term $d_{\mathbb{P}(\mathbb{R}^2)}(\pi_{E_1{}^\perp}\xi(\omega), \pi_{E_1^\perp}w_n)$, since the distribution of $\xi(\omega)$ is μ, by Lemma 2.24 we know that with probability at least $(1 - O(e^{-c(\epsilon)n}))$ we have $d(\xi(\omega), E_1) > q^{-\epsilon n/10}$. Combining with Lemma B.2, we get that with probability greater than $(1 - O(e^{-c(\epsilon)n}))$,

$$d_{\mathbb{P}(\mathbb{R}^2)}(\pi_{E_1{}^\perp}\xi(\omega), \pi_{E_1^\perp}w_n) < q^{-(\chi_1-\epsilon/10)n} \cdot q^{\epsilon n/10} < \frac{1}{2}q^{-(\chi_1-\epsilon)n},\ \text{for } n \text{ large enough.}$$

For the term $d_{\mathbb{P}(\mathbb{R}^2)}(\pi_{E_1^\perp}w_n, v_n)$, notice that $v_n = h(Z_n)^+$ and $\pi_{E_1{}^\perp}w_n = \pi_{E_1^\perp}Z_n x = h(Z_n)x$ (use the UL-decomposition and the assumption $x \in \mathbb{P}(E_1^\perp)$). Then by Eq. (2.64), with probability greater than $(1 - O(e^{-c(\epsilon)n}))$,

$$d_{\mathbb{P}(\mathbb{R}^2)}(v_n, \pi_{E_1{}^\perp}w_n) < q^{-(\chi_1-\epsilon/10)n}) < \frac{1}{2}q^{-(\chi_1-\epsilon)n},\ \text{for } n \text{ large enough.}$$

which completes the proof of Lemma 7.5. □

Now we back to the proof of Lemma 7.4. By (7.4), Lemma 4.4 of [HS17] and Eq. (7.5), we have

$$|\frac{1}{\chi_1 n}H(\tau_n, \mathcal{Q}_{\chi_1 n}) - \alpha| \to 0, \tag{7.8}$$

here the measure τ_n is the distribution of v_n on $\mathbb{P}(E_1^\perp)$.

We define a map Φ from L to $\mathbb{P}(E_1^\perp) \cong \mathbb{P}(\mathbb{R}^2)$ by mapping $\ell = \begin{pmatrix} 1 & 0 \\ n & h \end{pmatrix}$ to h^+. (If $h \in \mathrm{SO}_2(\mathbb{R})$, we take $h^+ = E_2$.) Then $\Phi(\pi_L\nu^{*n})$ is the distribution of v_n and $\tau_n = \Phi(\pi_L\nu^{*n})$. Recall that the set $L(n,\epsilon) \subset L$ is defined in Eq. (2.65). Then we have

Lemma 7.6. *For any atom E in the partition $\mathcal{Q}_{\chi_1 n}$ of $\mathbb{P}(\mathbb{R}^2)$, $\Phi^{-1}(E) \cap L(n,\epsilon)$ intersects at most $O(q^{10\epsilon n})$ number of atoms in $\mathcal{Q}_1^L$.*

Proof. From the definition of $L(n,\epsilon)$, for an atom E in $\mathcal{Q}_{\chi_1 n}$, either $L(n,\epsilon)\cap\Phi^{-1}E=\emptyset$ or we could pick a block diagonal element $\ell=\begin{pmatrix}1 & 0\\ 0 & h\end{pmatrix}\in L(n,\epsilon)\cap\Phi^{-1}E$. Our lemma is trivial if $L(n,\epsilon)\cap\Phi^{-1}E=\emptyset$. Suppose $L(n,\epsilon)\cap\Phi^{-1}E\neq\emptyset$, then for any $\ell'=\begin{pmatrix}1 & 0\\ n' & h'\end{pmatrix}\in L(n,\epsilon)\cap\Phi^{-1}E$ we have

$$\ell^{-1}\ell'=\begin{pmatrix}1 & 0\\ h^{-1}n' & h^{-1}h'\end{pmatrix}.$$

Take the Cartan decomposition of h,h', $h=k_h a_h l_h$ and $h'=k_{h'}a_{h'}l_{h'}$. Since $h^+,(h')^+\in E$, we have

$$d(k_h^{-1}k_{h'},id)=d(h^+,(h')^+)\leqslant q^{-\chi_1 n}.$$

By definition of $L(n,\epsilon)$, $\chi_1(h),\chi_1(h')\leqslant(\chi_1+\epsilon)n$,

$$\|h^{-1}h'\|=\|l_h^{-1}a_h^{-1}(k_h^{-1}k_{h'})a_{h'}l_{h'}\|=\|a_h^{-1}(k_h^{-1}k_{h'})a_{h'}\|\leqslant q^{\frac{1}{2}(\chi_1+\epsilon)n}\cdot O(q^{-\chi_1 n})\cdot q^{\frac{1}{2}(\chi_1+\epsilon)n}\leqslant O(q^{2\epsilon n}).$$

Recall a basic property of left-invariant and right $SO(2)$-invariant Riemannian metric on $\mathrm{SL}_2(\mathbb{R})$ i.e. for diagonal matrix we have $d(\mathrm{diag}(a,a^{-1}),id)=2\log|a|$ [Bry]. Combine it with the Cartan decomposition, we get

$$d(h^{-1}h',id)\leqslant 1+2\log\|h^{-1}h'\|\leqslant 4\epsilon n.$$

We then have

$$\begin{aligned}d_L(\ell^{-1}\ell',id)&\leqslant d_L(\ell^{-1}\ell',h^{-1}h')+d_L(h^{-1}h',id)=d_L((h')^{-1}h\ell^{-1}\ell',id)+d_L(h^{-1}h',id)\\&=d_L(\begin{pmatrix}1 & 0\\ (h')^{-1}n' & id_2\end{pmatrix},id)+d_L(h^{-1}h',id)\ll\|(h')^{-1}n'\|+d(h^{-1}h',id)\leqslant O(q^{2\epsilon n}),\end{aligned}$$

where we used $\|(h')^{-1}n'\|\leqslant 1/d(\ell^{-1}E_1,E_1^{\perp})$ (Eq. (2.14)) and the definition of $L(n,\epsilon)$. Therefore by our choice of ℓ', $\ell^{-1}(L(n,\epsilon)\cap\Phi^{-1}E)$ is contained in a neighborhood of id in L with radius $O(q^{2\epsilon n})$. By left invariance of d_L and $\dim L=5$, we get $\ell^{-1}(L(n,\epsilon)\cap\Phi^{-1}E)$ intersects at most $O(q^{10\epsilon n})$ number of atoms in $\mathcal{Q}_1^L$. □

We are ready to prove Lemma 7.4. To simplify the notation, let $\nu_1=\pi_L\nu^{*n}|_{L(n,\epsilon)},\nu_2=\pi_L\nu^{*n}|_{L(n,\epsilon)^c}$. We take the decomposition $\pi_L\nu^{*n}=(1-\delta)\nu_1+\delta\nu_2$, with $\delta=\pi_L\nu^{*n}(L(n,\epsilon)^c)$. Denote $\Phi\nu_i$ by m_i, $i=1,2$, then $\tau_n=\Phi(\pi_L\nu^{*n})=(1-\delta)m_1+\delta m_2$. Since $\operatorname{supp}\nu_1\subset L(n,\epsilon)$, by Lemma 7.6 and Lemma 2.34 we obtain

$$\begin{aligned}\frac{1}{n}H(\nu_1,\mathcal{Q}_1^L)&\leqslant\frac{1}{n}H(\nu_1,\Phi^{-1}\mathcal{Q}_{\chi_1 n})+O(\epsilon)\\&=\frac{1}{n}H(\Phi\nu_1,\mathcal{Q}_{\chi_1 n})+O(\epsilon)=\frac{1}{n}H(m_1,\mathcal{Q}_{\chi_1 n})+O(\epsilon).\end{aligned}$$

Combined with Eq. (2.77),

$$\begin{aligned}H(\pi_L\nu^{*n},\mathcal{Q}_1^L)&\leqslant(1-\delta)H(\nu_1,\mathcal{Q}_1^L)+\delta H(\nu_2,\mathcal{Q}_1^L)+H(\delta)\\&\leqslant(1-\delta)H(m_1,\mathcal{Q}_{\chi_1 n})+O(n\epsilon)+\delta H(\nu_2,\mathcal{Q}_1^L)+H(\delta)\\&\leqslant H(\tau_n,\mathcal{Q}_{\chi_1 n})+O(n\epsilon)+\delta H(\nu_2,\mathcal{Q}_1^L)+H(\delta).\end{aligned}$$

The measure ν_2 is supported on a ball of radius $q^{C_\nu n}$ due to compact support of ν, so $H(\nu_2,\mathcal{Q}_1^L)\leqslant O(C_\nu n)$. Due to $\delta=O(q^{-c(\epsilon)n})$ (Eq. (2.67)), we obtain

$$\frac{1}{n}H(\pi_L\nu^{*n},\mathcal{Q}_1^L)\leqslant\frac{1}{n}H(\tau_n,\mathcal{Q}_{\chi_1 n})+O(\epsilon+\delta)\leqslant\frac{1}{n}H(\tau_n,\mathcal{Q}_{\chi_1 n})+O(\epsilon+q^{-c(\epsilon)n}).$$

The proof can be completed by Eq. (7.8) and using the fact that ϵ can be arbitrarily small. □

7.2 Proof of Theorem 1.11

We start with a contradiction hypothesis that for μ^--a.e. V we have

$$\alpha = \dim \pi_{V^\perp}\mu < \min\left\{1, \frac{h_{\mathrm{RW}}(\nu)}{\chi_1(\nu)}\right\}. \tag{7.9}$$

and $\pi_{V^\perp}\mu$ is of exact dimension (Lemma 2.38). We fix any $V \in \mathbb{P}(\mathbb{R}^3)$ satisfying Eq. (7.9), Lemma 7.1, Lemma 7.2 and Lemma 7.4. Without loss of generality, using conjugation with some element k in $SO(3)$, that is replacing ν by $k\nu k^{-1}$, we can assume $V = E_1$.

Lower bound of entropy: application of entropy growth Theorem 5.7 and contradiction hypothesis: Due to Lemma 7.1 and Lemma 7.4, we have that for any large n,

$$\begin{aligned}\frac{1}{n}H(\pi_L\nu^{*n}, \mathcal{Q}_{Cn}|\mathcal{Q}_1) =&\frac{1}{n}(H(\pi_L\nu^{*n}, \mathcal{Q}_{Cn}) - H(\pi_L\nu^{*n}, \mathcal{Q}_1)) \\ \geqslant&(h_{\mathrm{RW}}(\nu) - \alpha\chi_1(\nu))/2.\end{aligned}$$

Observe that

$$\frac{1}{n}H(\pi_L\nu^{*n}, \mathcal{Q}_{Cn}|\mathcal{Q}_1) = \mathbb{P}_{i=1}\left(\frac{1}{n}H((\pi_L\nu^{*n})_{\ell,i}, \mathcal{Q}_{Cn})\right).$$

By Chebyshev's inequality, there exists $\epsilon > 0$ such that for any n large enough

$$\mathbb{P}_{i=1}\left(\frac{1}{n}H((\pi_L\nu^{*n})_{\ell,i}, \mathcal{Q}_{Cn}) > \epsilon\right) > \epsilon. \tag{7.10}$$

We fix such ϵ in the proof of Theorem 1.11. We will see later Eq. (7.10) is actually the source of positive entropy to apply the entropy growth argument. Denote by $\theta = \pi_L\nu^{*n}$ and $\mathcal{U}_0 = \mathcal{U}_0(n) := \{\ell \in L : \frac{1}{n}H(\theta_{\ell,1}, \mathcal{Q}_{Cn}) > \epsilon\}$, then Eq. (7.10) is equivalent to $\pi_L\nu^{*n}(\mathcal{U}_0(n)) > \epsilon$ for n large.

Consider the subset $\mathcal{U}_0' = \mathcal{U}_0'(n, C_1')$ of $\mathcal{U}_0$ defined by

$$\mathcal{U}_0'(n, C_1') := \{\ell \in \mathcal{U}_0(n) : d(\ell^{-1}, E_1^\perp) > 1/C_1'\}$$

The constant C_1' is chosen such that $C_1'^{-\beta} < \epsilon/4$, where β is given in Eq. (2.58). Then by Eq. (2.58) and $\pi_L\nu^{*n}(\mathcal{U}_0(n)) > \epsilon$, for any n large enough we have

$$\pi_L\nu^{*n}(\mathcal{U}_0'(n, C_1')) > \epsilon/1.1 \tag{7.11}$$

Now we try to apply Theorem Theorem 5.7 using $\mathcal{U}_0'$. For any $\ell \in \mathcal{U}_0'(n, C_1')$, for the component measure $\theta_{\ell,1}$ we have

- $\mathrm{diam}(\theta_{\ell,1}) \leqslant \mathrm{diam}(\mathcal{Q}_1(\ell)) \leqslant 1 < C_p$;
- for any $\ell' \in \mathrm{supp}\,\theta_{\ell,1} \subset \mathcal{Q}_1(\ell)$, $|\chi_1(h(\ell')) - \chi_1(h(\ell))| \leqslant 1$ by our choice of q (and the choice of constants C_p, C_L), see Definition 2.31 and Lemma 2.15.
- for any $\ell' \in \mathrm{supp}\,\theta_{\ell,1} \subset \mathcal{Q}_1(\ell)$,
$$d(\ell'^{-1}E_1, E_1^\perp) \geqslant 1/2d(\ell^{-1}E_1, E_1^\perp) \geqslant 1/2C_1' \tag{7.12}$$
- $\frac{1}{n}H(\theta_{\ell,1}, \mathcal{Q}_{Cn}) > \epsilon$ since $\ell \in \mathcal{U}_0$.

Therefore for $\ell \in \mathcal{U}_0'$, we can apply Theorem 5.7 to $\theta_{\ell,1}$ and μ with $C_1 = 2C_1'$ and contracting rate $t = \chi_1(h(\ell))$: for n large enough,

$$\frac{1}{Cn}H([\theta_{\ell,1}.\mu], \mathcal{Q}_{Cn+\chi_1(h(\ell))}) > \alpha + \delta,$$

where the constant $\delta = \delta(\alpha, \epsilon, C_1)$ is given as in Theorem 5.7.

By Eq. (2.62) and (7.11), for any $\epsilon_1 > 0$, for n large enough, the subset $\mathcal{U}_0''(n, C_1', \epsilon) \subset \mathcal{U}_0'$ formed by those those $\ell \in \mathcal{U}_0'$ such that $|\chi_1(h(\ell)) - \chi_1 n| \leqslant \epsilon_1 n$ satisfies $\pi_L\nu^{*n}(\mathcal{U}_0''(n, C_1', \epsilon_1)) > \epsilon/1.2$. Fix a positive number ϵ_1 much smaller than δ, then by Eq. (2.75), for any $\ell \in \mathcal{U}_0''$ we have

$$\frac{1}{Cn}H([\theta_{\ell,1}.\mu], \mathcal{Q}_{Cn+\chi_1 n}) \geqslant \frac{1}{Cn}H([\theta_{\ell,1}.\mu], \mathcal{Q}_{Cn+\chi_1(h(\ell))}) - O(\frac{\epsilon_1}{C}) > \alpha + \delta - O(\frac{\epsilon_1}{C}) > \alpha + \delta/2. \tag{7.13}$$

We set $\mathcal{U}_1(n) = \mathcal{U}_1 := \{\ell \in L: \ \ell \text{ satisfies } Eq.\ (7.13)\}$. Then by the discussion above about $\mathcal{U}_0''$, for n large enough, the θ-measure of $\mathcal{U}_1$ greater than $\epsilon/2$, and $\mathcal{U}_1$ consists of atoms of $\mathcal{Q}_1$.

Due to Lemma 7.2, for any $\tau > 0$, there exists an infinite sequence $\{n_k\}$ and a sequence of subset $\mathcal{U}_2 = \mathcal{U}_2(n_k) \subset L$ such that for each k, $\mathcal{U}_2(n_k)$ has ν^{n_k}-probability greater than $1-\tau$, and for any $\ell \in \mathcal{U}_2$,

$$\frac{1}{Cn}H([\ell\mu], \mathcal{Q}_{(C+\chi_1)n}) > \alpha - \tau. \tag{7.14}$$

Take n large enough to be some n_k defined above. Consider

$$\begin{aligned}
&\mathbb{E}_{i=1}\left(\frac{1}{Cn}H\left([\theta_{\ell,i}.\mu], \mathcal{Q}_{(C+\chi_1)n}\right)\right)\\
&= \int \frac{1}{Cn}H\left([\theta_{\ell,1}.\mu], \mathcal{Q}_{(C+\chi_1)n}\right) \mathrm{d}\theta(\ell)\\
&= \int_{\mathcal{U}_1} \frac{1}{Cn}H\left([\theta_{\ell,1}.\mu], \mathcal{Q}_{(C+\chi_1)n}\right) \mathrm{d}\theta(\ell) + \int_{\mathcal{U}_1^c} \frac{1}{Cn}H\left([\theta_{\ell,1}.\mu], \mathcal{Q}_{(C+\chi_1)n}\right) \mathrm{d}\theta(\ell).
\end{aligned}$$

For the first term on RHS, we apply Eq. (7.13) to estimate it. For the second term, since $\mathcal{U}_1$ consists of atoms of $\mathcal{Q}_1$, by the concavity of entropy we obtain

$$\int_{\mathcal{U}_1^c} \frac{1}{Cn}H\left([\theta_{\ell,1}.\mu], \mathcal{Q}_{(C+\chi_1)n}\right) \mathrm{d}\theta(\ell) \geqslant \int_{\mathcal{U}_1^c} \frac{1}{Cn}H\left([\ell\mu], \mathcal{Q}_{(C+\chi_1)n}\right) \mathrm{d}\theta(\ell) \geqslant (\alpha - \tau)\theta(\mathcal{U}_2 - \mathcal{U}_1),$$

where the last inequality is due to Eq. (7.14). Therefore, we have

$$\mathbb{E}_{i=1}\left(\frac{1}{Cn}H\left([\theta_{\ell,i}.\mu], \mathcal{Q}_{(C+\chi_1)n}\right)\right) \geqslant \frac{\epsilon}{2}(\alpha + \frac{\delta}{2}) + (1 - \tau - \epsilon/2)(\alpha - \tau) > \frac{\epsilon\delta}{4} + (\alpha - \tau)(1 - \tau).$$

By taking τ small, we obtain

$$\limsup_{n\to\infty} \mathbb{E}_{i=1}\left(\frac{1}{Cn}H\left([\theta_{\ell,i}.\mu], \mathcal{Q}_{(C+\chi_1)n}\right)\right) > \alpha + \epsilon\delta/8. \tag{7.15}$$

Upper bound of entropy: application of exact dimension. Another consequence from the exact dimension of $\pi_{E_1^\perp}\mu$ is the following

$$\lim_{n\to\infty} \frac{1}{Cn}H(\pi_{E_1^\perp}\mu, \mathcal{Q}_{(C+\chi_1)n}|\mathcal{Q}_{\chi_1 n}) = \alpha.$$

Note that $\pi_{E_1^\perp}\mu = \pi_{E_1^\perp}(\nu^{*n} * \mu) = [(\pi_L\nu^{*n}).\mu] = [\theta.\mu]$, so $\pi_{E_1^\perp}\mu = \mathbb{E}_{i=1}([\theta_{\ell,i}.\mu])$. By the concavity of entropy we have

$$\limsup_{n\to\infty} \mathbb{E}_{i=1}\left(\frac{1}{Cn}H([\theta_{\ell,i}.\mu], \mathcal{Q}_{(C+\chi_1)n}|\mathcal{Q}_{\chi_1 n})\right) \leqslant \alpha. \tag{7.16}$$

We would like to get rid of the conditional entropy in Eq. (7.16). This will enable us to compare Eq. (7.16) with Eq. (7.15) and obtain a contradiction. The idea is that for most of ℓ in $\operatorname{supp}\theta$, we can decompose $[\theta_{\ell,1}.\mu]$ similar to the r-attracting decomposition (Definition 4.7) as the following claim.

Claim 5. *There exists $\epsilon_0 > 0$ such that for any positive $\epsilon_1 < \epsilon_0$, for $\ell \in L(n, \epsilon_1)$ defined in Eq. (2.65), the component $[\theta_{\ell,1}.\mu]$ can be decomposed as the sum $[\theta_{\ell,1}.\mu] = (1 - \epsilon_2)\tau_1 + \epsilon_2\tau_2$ such that*

$$\epsilon_2 \ll q^{-\beta\epsilon_1 n} \text{ and } \operatorname{diam}(\operatorname{supp}\tau_1) \leqslant q^{-(\chi_1-4\epsilon_1)n} \text{ for } n \text{ large enough.}$$

Proof of Claim 5. Consider $\tau = \tau(\ell) = \ell^{-1}\theta_{\ell,1}$ on L and $\tau * \mu$ on $\mathbb{P}(\mathbb{R}^3)$. Then

$$[\theta_{\ell,1}.\mu] = [\ell(\tau * \mu)].$$

And $\tau(\ell)$ is supported on a uniformly bounded neighborhood of the identity in L.

Let $m := \tau*\mu$, $m_1 := m_{b(f_\ell, q^{-\epsilon_1 n})}$ and $m_2 := m_{b(f_\ell, q^{-\epsilon_1 n})^c}$, where the good region $b(f_\ell, q^{-\epsilon_1 n})$ is introduced in Definition 2.13, whose complement is the $q^{-\epsilon_1 n}$ neighborhood of a hyperplane. Since τ is uniformly bounded supported, there exists $C' > 0$ such that for any $g \in \operatorname{supp}\tau$, we can apply the bound in Lemma 2.24 to obtain

$$g\mu(b(f_\ell, q^{-\epsilon_1 n})^c) \leqslant \mu(b(W, C'q^{-\epsilon_1 n})^c) \ll (q^{-\epsilon_1 n})^\beta,$$

for some hyperplane W. Hence for ϵ_2 from the decomposition of $m = (1 - \epsilon_2)m_1 + \epsilon_2 m_2$

$$\epsilon_2 = \tau * \mu(b(f_\ell, q^{-\epsilon_1 n})^c) = \int g\mu(b(f_\ell, q^{-\epsilon_1 n})^c)\mathrm{d}\tau(g) \ll q^{-\beta\epsilon_1 n}. \tag{7.17}$$

Apply Lemma 2.14 we obtain

$$\operatorname{diam}([\ell m_1]) \leqslant q^{-(\chi_1-4\epsilon_1)n}, \text{ for } n \text{ large enough.}$$

By letting $\tau_1 = [\ell m_1]$ and $\tau_2 = [\ell m_2]$, we finish the proof of the claim. $\square$

Due to Claim 5, we can apply Lemma 2.36 to $\ell \in L(n, \epsilon_1)$ and obtain

$$\begin{aligned}
&\mathbb{E}_{i=1}\left(\frac{1}{Cn}H\left([\theta_{\ell,i}.\mu], \mathcal{Q}_{(C+\chi_1)n}|\mathcal{Q}_{\chi_1 n}\right)\right)\\
\geqslant&\int_{L(n,\epsilon_1)}\frac{1}{Cn}H([\theta_{\ell,1}.\mu], \mathcal{Q}_{(C+\chi_1)n}|\mathcal{Q}_{\chi_1 n})\mathrm{d}\theta(\ell)\\
\geqslant&\int_{L(n,\epsilon_1)}\frac{1}{Cn}H([\theta_{\ell,1}.\mu], \mathcal{Q}_{(C+\chi_1)n}) - \frac{1}{Cn}(O(\epsilon_2 n) + H(\epsilon_2) + O(\epsilon_1 n))d\theta(\ell)\\
\geqslant&\mathbb{E}_{i=1}\left(\frac{1}{Cn}H\left([\theta_{\ell,i}.\mu], \mathcal{Q}_{(C+\chi_1)n}\right)\right) - \frac{1}{Cn}(O(\epsilon_2 n) + H(\epsilon_2) + O(\epsilon_1 n) + O(q^{-c(\epsilon_1)n})),
\end{aligned}$$

where the last inequality is due to Eq. (2.67). By taking ϵ_1 sufficiently small (and hence ϵ_2 small by Eq. (7.17)), as n going to infinity, combining with Eq. (7.16), we obtain a contradiction to (7.15).

Hence we conclude $\dim \pi_{E_1^\perp}\mu = \min\{1, h_{\mathrm{RW}}(\nu)/\chi_1(\nu)\}$.

7.3 A Ledrappier-Young formula and the proof of Theorem 1.10 from Theorem 1.11.

We first recall the Ledrappier-Young type formulas of Furstenberg measures recently shown in [LL23b, Corollary 6.5] and [Rap21]. (See [LL23a] for more general case.)

Proposition 7.7. *Let ν be a finitely supported probability measure such that* $\operatorname{supp}\nu$ *generates a Zariski dense subgroup in* $\mathrm{SL}_3(\mathbb{R})$*. Then there exist $\gamma_1, \gamma_2 \geqslant 0$ such that for μ^- a.e. V we have*

1. *the projection measure $\pi_{V^\perp}\mu$ is of exact dimension γ_1;*
2. *for $\pi_{V^\perp}\mu$ a.e. $x \in \mathbb{P}(V^\perp)$, the disintegration μ_V^x of the projection $\pi_{V^\perp}$ is of exact dimension γ_2;*
3. *We have*
$$h_{\mathrm{F}}(\mu,\nu) = \gamma_1\chi_1 + \gamma_2\chi_2,\ \dim\mu = \gamma_1 + \gamma_2, \tag{7.18}$$
where $h_{\mathrm{F}}(\mu,\nu)$ is the Furstenberg entropy defined by (8.1).

Proof. Proposition 7.7 is actually stated in Corollary 6.5 of [LL23b] in slightly a different language. To explain it we recall some notions from [LL23b]. For the set $\{1,2,3\}$, we have topopogies $T^2 = (\{1\},\{2,3\},\{3\}) \prec T^1 = (\{1,3\},\{2,3\},\{3\}) \prec T^0 = (\{1,2,3\},\{2,3\},\{3\})$. For each topology $T = T^i$, and two flags $f = (\{0\} = U_0 \subset U_1 \subset U_2 \subset U_3 = \mathbb{R}^3)$ and $f' = (\{0\} = U_0' \subset U_1' \subset U_2' \subset U_3' = \mathbb{R}^3)$ in general position, that is $U_1 \oplus U_2' = U_1' \oplus U_2 = \mathbb{R}^3$, we define the map for $I \in T$
$$[F_T(f,f')]_I = \oplus_{i\in I}(U_i \cap U_{4-i}').$$
For example let T be $T^0 = (\{1,2,3\},\{2,3\},\{3\})$ and consider the map $[F_{T^0}(f,f')]_I, I \in T^0$

- for $I = \{1,2,3\}$, due to flags in general position, we obtain $\mathbb{R}^3$ and there is no information;
- for $I = \{2,3\}$, we obtain $(U_2 \cap U_2') \oplus U_1' = U_2'$;
- for $I = \{3\}$, we obtain U_1'.

In summary we get (U_2', U_1') from $[F_{T^0}(f,f')]$, similarly from $[F_{T^2}(f,f')]$ and $[F_{T^1}(f,f')]$ we obtain (U_1, U_2', U_1') and $(U_1 \oplus U_1', U_2', U_1')$, therefore projections (in the sense of [LL23b])
$$(U_1, U_2', U_1') \to (U_1 \oplus U_1', U_2', U_1') \to (U_2', U_1')$$
with one-dimensional fibre. Then the conditional measure $\mu_{T^1,T^0}^{U_1'}$ in [LL23b] is the fiber measure with fixed U_1', and its fiber $U_1 \oplus U_1'$ is in the space of two-planes containing U_1'. So by projection to $(U_1')^\perp$ it can be identified with $\pi_{V^\perp}\mu$ for $V = U_1'$ following the law μ^-. Similarly, the conditional measure $\mu_{T^2,T^1}^{U_1\oplus U_1', U_1'}$ is the distribution of U_1 with $U_1 \oplus U_1'$ and U_1' fixed. Denote by $V_2 = U_1 \oplus U_1'$ and $x = \pi_{V^\perp}V_2$. We can identify $\mu_{T^2,T^1}^{U_1\oplus U_1', U_1'}$ with μ_V^x the disintegration from the projection $\mu \to \pi_{V^\perp}\mu$.

With all these identification above, [LL23b, Corollary 6.5] can be viewed as in Eq. (7.18). $\square$

Proof of Theorem 1.10 from Theorem 1.11. By the exponential separation assumption in Theorem 1.10, Theorem 1.11 implies that $\gamma_1 = \min\{1, \frac{h_{\mathrm{RW}}(\nu)}{\chi_1}\}$.

If $\gamma_1 = 1$, Theorem 1.10 directly follows the definition of $\dim_{\mathrm{LY}}(\nu)$ and Eq. (7.18). For the case that $\gamma_1 = \frac{h_{\mathrm{RW}}(\nu)}{\chi_1} < 1$, we use the general fact that $h_{\mathrm{F}}(\mu,\nu) \leqslant h_{\mathrm{RW}}(\nu)$, cf. [KV83] and also [Fur02, Theorem 2.31]. These result are stated for the discrete groups. If $G_\nu = \langle \operatorname{supp}\nu \rangle$ is not discrete in $\mathrm{SL}_3(\mathbb{R})$, we can give a discrete topology on G_ν as an abstract group, denoted by G_ν' and corresponding measure by ν'. Then the space $(\mathbb{P}(\mathbb{R}^3), \mu)$ is still a (G_ν', ν')-space. So $h_{\mathrm{F}}(\mu,\nu) \leqslant h_{\mathrm{RW}}(\nu') = h_{\mathrm{RW}}(\nu)$. Then we obtain $\dim_{\mathrm{LY}}\mu \leqslant \dim\mu$. Since $\dim_{\mathrm{LY}}\mu$ is always an upper bound of $\dim\mu$, we finish the proof. $\square$

8 Different notions of entropy

In Sections 8, 9 and 10, we will deal with general $\mathrm{SL}_n(\mathbb{R})$ actions for $n \geqslant 3$. Let us first recall some notion.

Let Γ be a hyperbolic group with a fixed word norm $|\cdot|$. A representation $\rho : \Gamma \to \mathrm{SL}_n(\mathbb{R})$ is called Borel Anosov if there exists $c > 0$ such that for every $1 \leqslant p \leqslant n-1$,

$$\frac{\sigma_p(\rho(\gamma))}{\sigma_{p+1}(\rho(\gamma))} \geqslant ce^{c|\gamma|}, \quad \forall \gamma \in \Gamma$$

where $\sigma_1 \geqslant \cdots \geqslant \sigma_n$ are singular values. For the case of $n = 3$, an Anosov representation is automatically a Borel Anosov representation. We say ρ is Zariski dense if its image is Zariski dense in $\mathrm{SL}_n(\mathbb{R})$.

Recall that $\mathbb{P}(\mathbb{R}^n)$ is the projective space. We also consider the flag variety of $\mathbb{R}^n$ as

$$\mathcal{F}(\mathbb{R}^n) \coloneqq \left\{ \xi = (\xi^1 \subset \xi^2 \subset \cdots \subset \xi^i \subset \cdots \subset \xi^{n-1}) : \xi^i \text{ is a linear subspace of } \mathbb{R}^n \text{ of dimension } i \right\}$$

and that $\mathrm{SL}_n(\mathbb{R})$ acts on $\mathcal{F}(\mathbb{R}^n)$ canonically. We abbreviate $\mathcal{F}(\mathbb{R}^3)$ to $\mathcal{F}$.

Let ν be a finitely supported probability measure on $\mathrm{SL}_n(\mathbb{R})$. We denote G_ν to be the group generated by $\operatorname{supp}\nu$. If we further assume that G_ν is Zariski dense in $\mathrm{SL}(n, \mathbb{R})$. Then the Lyapunov spectrum $\lambda(\nu) = \{\lambda_1(\nu) \geqslant \cdots \geqslant \lambda_n(\nu)\}$ is simple. Moreover, the random walks induced by ν on the projective space $\mathbb{P}(\mathbb{R}^n)$ and the flag variety $\mathcal{F}(\mathbb{R}^n)$ both have a unique stationary measure ([Fur63, GR85, GM89], see also [BQ16, Chapter 10]).

Now we recall two notions of the entropy associated to the random walk.

Definition 8.1. 1. Let μ be a ν-stationary measure on $\mathbb{P}(\mathbb{R}^n)$ or $\mathcal{F}(\mathbb{R}^n)$, the *Furstenberg entropy* is given by

$$h_{\mathrm{F}}(\mu, \nu) = \int \log \frac{\mathrm{d}g\mu}{\mathrm{d}\mu}(\xi) \left(\frac{\mathrm{d}g\mu}{\mathrm{d}\mu}(\xi) \right) \mathrm{d}\nu(g)\mathrm{d}\mu(\xi). \tag{8.1}$$

2. The *random walk entropy* of ν is

$$h_{\mathrm{RW}}(\nu) = \lim_{k \to \infty} \frac{1}{k} H(\nu^{*k}).$$

Remark 8.2. In [BHR19] and [Rap21], the notion of the entropy they used is the following: for a discrete measure $\nu = \sum p_i \delta_{g_i}$, $H(\nu) \coloneqq H(p) = -\sum p_i \log p_i$. This notion works well in many settings of IFSs. For general semigroup actions, the random walk entropy $h_{\mathrm{RW}}(\nu)$ as in [HS17] is more precise. In particular, if $\operatorname{supp}\nu$ freely generates a free semigroup then $H(\nu) = h_{\mathrm{RW}}(\nu)$. This kind of entropy was first studied by Avez in [Ave72] to study the structure of the group action on the boundary.

The Furstenberg entropy is mysterious and might be difficult to compute. To obtain a more calculable dimension formula, it is expected to show that the Furstenberg entropy in Theorem 1.10 is equal to the random walk entropy. In the following proposition we will see that for some concrete examples, we do have this equality.

Proposition 8.3. *Let Γ be a hyperbolic group and $\rho : \Gamma \to \mathrm{SL}_n(\mathbb{R})$ be a Zariski dense Borel Anosov representation. Let ν be a finitely supported probability measure on $\rho(\Gamma)$ such that G_ν is Zariski dense. Then the unique ν-stationary measure μ on $\mathbb{P}(\mathbb{R}^n)$ satisfies*

$$h_{\mathrm{F}}(\mu, \nu) = h_{\mathrm{RW}}(\nu).$$

To show Propostion 8.3, we may consider the random walk on the flag variety.

Proposition 8.4. *Let ν be a probability measure on $\mathrm{SL}_n(\mathbb{R})$ such that G_ν is a Zariski dense discrete subgroup. Let $\mu_{\mathcal{F}} = \mu_{\mathcal{F}}(\nu)$ and $\mu = \mu(\nu)$ be the unique ν-stationary measure on $\mathcal{F}(\mathbb{R}^n)$ and $\mathbb{P}(\mathbb{R}^n)$ respectively. Then we have*

$$h_{\mathrm{RW}}(\nu) = h_{\mathrm{F}}(\mu_{\mathcal{F}}, \nu) \geqslant h_{\mathrm{F}}(\mu, \nu).$$

Proof. Proposition 8.4 follows from [Fur02, Theorem 2.31] (originally in [KV83, Theorem 3.2], [Led85, Section 3.2]). In order to apply their result, we need to use [Fur02, Theorem 2.21] (originally in [KV83], [Led85]) to obtain that $(\operatorname{supp} \mu_{\mathcal{F}}, \mu_{\mathcal{F}})$ is the Poisson boundary of (G_ν, ν). Then the Furstenberg entropy of the Poisson boundary is exactly the random walk entropy for discrete G_ν due to [Fur02, Theorem 2.31]. □

Proof of Proposition 8.3. By Proposition 8.4 and that the image of an Anosov representation is discrete, it remains to prove that $h_{\mathrm{F}}(\mu_{\mathcal{F}}, \nu) = h_{\mathrm{F}}(\mu, \nu)$.

Consider the canonical projection

$$\pi : \mathcal{F}(\mathbb{R}^n) \to \mathbb{P}(\mathbb{R}^n), \xi = (\xi^1 \subset \xi^2 \subset \cdots \subset \xi^{n-1}) \mapsto \xi^1.$$

Due to uniqueness of the Furstenberg measure on $\mathbb{P}(\mathbb{R}^n)$, we know that $\pi_*(\mu_{\mathcal{F}}) = \mu$. By classical Rokhlin's disintegration theorem, we can define a desintegration $\{\mu^\xi\}$ of the measure $\mu_{\mathcal{F}}$ over μ, where μ^ξ is a well-defined probabilty measure on $\pi^{-1}(\xi)$ for μ a.e. ξ.

Let $L_{\mathcal{F}}(\rho(\Gamma))$ and $L(\rho(\Gamma))$ be the limit sets on $\mathcal{F}$ and $\mathbb{P}(\mathbb{R}^n)$ (the closure of attracting fixed points of proximal elements) respectively. [BQ16, Lemma 9.5] tells us that $L_{\mathcal{F}}(\rho(\Gamma))$ (resp. $L(\rho(\Gamma))$) is the unique $\rho(\Gamma)$-minimal closed invariant subset on $\mathcal{F}(\mathbb{R}^n)$ (resp. $\mathbb{P}(\mathbb{R}^n)$). Hence $\mu_{\mathcal{F}}$ (resp. μ) is supported on the limit set $L_{\mathcal{F}}(\rho(\Gamma))$ (resp. $L(\rho(\Gamma))$). Before continuing, we need a lemma about the structure of the limit sets.

Lemma 8.5. *Let $\rho : \Gamma \to \mathrm{SL}_n(\mathbb{R})$ be a Borel Anosov representation. Then the canonical projection $\pi : L_{\mathcal{F}}(\rho(\Gamma)) \to L(\rho(\Gamma))$ has a trivial fibre.*

Proof. Because $\rho : \Gamma \to \mathrm{SL}_n(\mathbb{R})$ is Borel Anosov, we can apply [Can, Theorem 31.1]. Then there exists a continuous ρ equivariant map (ξ^1, ξ') from $\partial\Gamma$ to $\mathbb{P}(\mathbb{R}^n) \times \mathrm{Grass}(n-1, \mathbb{R}^n)$, such that ξ^1 satisfies the Cartan property in [Can, Chapter 30] and $\xi^1(x) \subset \xi'(x)$,

$$\xi^1(x) \oplus \xi'(y) = \mathbb{R}^n, \quad \forall x \neq y.$$

It follows that ξ^1 is injective from $\partial\Gamma$ to $\mathbb{P}(\mathbb{R}^n)$. Moreover, the image of ξ^1 is exactly $L(\rho(\Gamma))$.

From Borel Anosov property, ξ^1 can be extended to a limit map $\xi : \partial\Gamma \to \mathcal{F}$ given by $\xi(x) = (\xi^1(x) \subset \xi^2(x) \subset \cdots \subset \xi^{n-1}(x))$. [7] The image of ξ is $L_{\mathcal{F}}(\rho(\Gamma)) \subset \mathcal{F}$. The map ξ^1 being injective implies that the natural projection of $L_{\mathcal{F}}(\rho(\Gamma))$ to $L(\rho(\Gamma))$ has trivial fiber. □

By Lemma 8.5, we know that the fiber is trivial for all $\xi \in L(\rho(\Gamma))$. Hence we have the relation of *relative measure-preserving* of $(\mathcal{F}(\mathbb{R}^3), \mu_{\mathcal{F}}) \to (\mathbb{P}(\mathbb{R}^3), \mu)$, that is for ν a.e. g and μ a.e. ξ

$$g\mu^\xi = g\delta_{\pi^{-1}\xi} = \delta_{\pi^{-1}(g\xi)} = \mu^{g\xi}.$$

By [Fur02, Proposition 2.25] ([KV83]), we obtain

$$h_{\mathrm{F}}(\mu_{\mathcal{F}}, \nu) = h_{\mathrm{F}}(\mu, \nu)$$

Then by Proposition 8.4, the proof is complete. □

[7] The existence of the limit map $\xi^k(x)$ is from P_k-Anosov and the consistence condition $\xi^k(x) \subset \xi^{k+1}(x)$ can be deduced from the Cartan property of the limit map (see for example Section 30 and 31 in [Can]), that is $\xi^k(x) = \lim_{n\to\infty} U_k(\gamma_n)$. Here γ_n is a sequence converges to the boundary point x and $U_k(\gamma_n) = k_{\gamma_n}(E_1 \oplus \cdots \oplus E_k)$ from the Cartan decomposition $\gamma_n = k_{\gamma_n} a_{\gamma_n} k'_{\gamma_n}$.

9 Free sub-semigroups in hyperbolic groups

Let Γ be a finitely generated group and $\mathcal{S}$ be a fixed symmetric generating set. For every $g \in \Gamma$, let $|g|$ denote the word length of g with respect to $\mathcal{S}$. For a positive integer L, let $A(L) := \{g \in \Gamma : |g| = L\}$ refer to an annulus. The following is the main proposition of this section.

Proposition 9.1. *Let Γ be a non-cyclic torsion-free hyperbolic group and $\rho : \Gamma \to \mathrm{SL}_n(\mathbb{R})$ be a faithful representation. Let $\mathbf{G}$ be the Zariski closure of $\rho(\Gamma)$ which is assumed to be Zariski connected. Then there exists a finite subset $F \subset \Gamma$ with $\#F \geqslant 3$, constants $C_7, C_8, L_0 > 0$ and $m \in \mathbb{Z}_+$ such that the following holds.*

For every subset $S \subset A(L)$ for some $L \geqslant L_0$ there exists a subset $S' \subset S$ with $\#S' \geqslant C_7^{-1}\#S$ and $F' \subset F$ with $\#F' = \#F - 2$ satisfying

(1) $\{\rho(f)^m : f \in F'\}$ generates a semigroup whose Zarski closure is $\mathbf{G}$.

(2) $\widetilde{S} := \{sf^\varsigma : s \in S', f \in F', \varsigma = m, 2m\} \subset \Gamma$ freely generates a free semigroup.

(3) For any sequence of elements $\widetilde{s}_1, \cdots, \widetilde{s}_k \in \widetilde{S}$, we have

$$|\widetilde{s}_1 \cdots \widetilde{s}_k| \geqslant \sum_{i=1}^{k} |\widetilde{s}_i| - kC_8. \tag{9.1}$$

9.1 Preliminaries on geometric group theory

Recall that Γ is a finitely generated group and $\mathcal{S}$ is a fixed symmetric generating set. Let $X = \mathrm{Cay}(\Gamma, \mathcal{S})$ be the Cayley graph of Γ with respect to $\mathcal{S}$. Endow X with the graph metric $d = d_{\mathcal{S}}$, which makes X a proper geodesic space. Then Γ has a natural left action on (X, d) by isometries. Letting o be the point in X corresponding to the identity in Γ, we will fix o as the base point. Then the word length $|g|$ is equal to $d(o, go)$.

For every subset $Y \subset X$ and $r > 0$, we use $\mathcal{N}_r(Y)$ to denote the r-neighborhood of Y. For two subsets $Y_1, Y_2 \subset X$, we use $d_{\mathrm{H}}(Y_1, Y_2)$ to denote the Hausdorff distance between Y_1, Y_2 with respect to $d_{\mathcal{S}}$, given by

$$d_{\mathrm{H}}(Y_1, Y_2) := \inf \{0 < r \leqslant \infty : Y_1 \subset \mathcal{N}_r(Y_2) \text{ and } Y_2 \subset \mathcal{N}_r(Y_1)\}.$$

For a subset $Y \subset X$ and a point $x \in X$, we denote $\pi_Y(x)$ to be the projection of x on Y, that is $\pi_Y(x) := \{y \in Y : d(x, y) = d(x, Y)\}$. For another subset $Z \subset X$, the projection of Z on Y is $\pi_Y(Z) := \cup_{z \in Z} \pi_Y(z)$.

For a rectifiable path $\alpha \subset X$, we denote α_- and α_+ to be the initial and terminal points of α, respectively. The length of α is denoted by $l(\alpha)$. For every pair $x, y \in X$, we denote $[x, y]$ to be a choice of a geodesic between x and y.

A path α is called a *c-quasi-geodesic* for $c \geqslant 1$ if

$$l(\beta) \leqslant c \cdot d(\beta_-, \beta_+) + c$$

for every rectifiable subpath $\beta \subset \alpha$. Morse lemma states that every c-quasi-geodesic α is contained in the c'-neighborhood of $[\alpha_-, \alpha_+]$ in a δ-hyperbolic space, where c' only depends on c and δ. A subset $Y \subset X$ is called *c-quasi-convex* for $c \geqslant 0$ if for every $x_1, x_2 \in Y$, we have $[x_1, x_2] \subset \mathcal{N}_c(Y)$. A quasi-geodesic can also be interpreted as a quasi-convex subset.

We say Γ is a (δ-)hyperbolic group if X is a (δ-)hyperbolic space. That is, for every geodesic triangle in X, every edge is contained in the δ-neighborhood of the other two edges. In this section, we always assume that Γ is a finitely generated δ-hyperbolic group. An infinite order element $g \in \Gamma$ is called loxodromic. For every loxodromic element g, the map $n \mapsto g^n o$ is a quasi-isometric embedding from $\mathbb{Z}$ to (X,d). That is, there exists $c_1,c_2,c_3,c_4>0$ such that for every $m,n\in\mathbb{Z}$,

$$c_1|m-n|-c_2\leqslant d(g^mo,g^no)\leqslant c_3|m-n|+c_4.$$

For every loxodromic element $g\in\Gamma$, the set

$$E(g):=\{h\in\Gamma: d_{\mathrm{H}}(h\langle g\rangle o,\langle g\rangle o)<\infty\}$$

is a subgroup of Γ satisfying $[E(g):\langle g\rangle]<\infty$. We use $\mathrm{Ax}(g)=E(g)o\subset X$ to denote the axis corresponding to g, which is a quasi convex subset. We also remark that if $g_1\in E(g_2)$ is also loxodromic then $E(g_1)=E(g_2)$. Two loxodromic elements g_1,g_2 are called *independent*[8] if $E(g_1)\neq E(g_2)$, namely $E(g_1)\cap E(g_2)$ is finite. Moreover, we have the following bounded intersection property.

Lemma 9.2. *Let $\mathrm{Ax}(g_1),\mathrm{Ax}(g_2)\subset X$ be different axes, then for every $r>0$, $\mathcal{N}_r(\mathrm{Ax}(g_1))\cap\mathcal{N}_r(\mathrm{Ax}(g_2))$ is bounded.*

Proof. Since $d_{\mathrm{H}}(\mathrm{Ax}(g_i),\langle g_i\rangle o)<\infty$, it suffices to show that $\mathcal{N}_r(\langle g_1\rangle o))\cap\mathcal{N}_r(\langle g_2\rangle o)$ is bounded. Otherwise, there exists infinitely many pairs of integers (n_1,n_2) such that $d(g_1^{n_1}o,g_2^{n_2}o)\leqslant 2r$. Then there exists $(n_1,n_2)\neq(n_1',n_2')$ such that

$$g_2^{-n_2}g_1^{n_1}=g_2^{-n_2'}g_1^{n_1'}.$$

This implies that $g_1^{n_1'-n_1}=g_2^{n_2'-n_2}$, which contradicts $\mathrm{Ax}(g_1)\neq\mathrm{Ax}(g_2)$. $\square$

In the case of Γ is torsion-free, every nontrivial element is loxodromic. By a classification of virtually cyclic group [Hem04, Lemma 11.4], $E(g)$ is cyclic for every nontrivial element $g\in\Gamma$. Moreover, g_1 and g_2 are independent if and only if $E(g_1)\cap E(g_2)=\{\mathrm{id}\}$.

9.2 An extension lemma

Now we give the main technical tool for showing Proposition 9.1. It is a variant of [Yan19, Lemma 2.19], which states more generally for the group with contracting elements. A direct proof for the case of hyperbolic groups is given in this section, which is also inspired by the work of W. Yang.

For a subset $S\subset\Gamma$, we say S is R-separated for $R>0$ if $\{so:s\in S\}$ is R-separated in X.

Proposition 9.3. *Let Γ be a non-elementary hyperbolic group. Let $F\subset\Gamma$ be a finite subset of pairwise independent loxodromic elements with $\#F\geqslant 3$. Then there exist $m_0,R,C_7,L_0>0$ such that for every R-separated subset $S\subset A(L)$ for some $L\geqslant L_0$ there exists a subset $S'\subset S$ with $\#S'\geqslant C_7^{-1}\#S$ and $F'\subset F$ with $\#F'=\#F-2$ such that for every $m\geqslant m_0$,*

$$\widetilde{S}:=\left\{sf^{\varsigma}:s\in S',f\in F',\varsigma=m,2m\right\}\subset\Gamma$$

freely generates a free semigroup. Furthermore, there exists $C_8>0$ only depends on F such that for any sequence of elements $\widetilde{s}_1,\cdots,\widetilde{s}_k\in\widetilde{S}$, we have

$$|\widetilde{s}_1\cdots\widetilde{s}_k|\geqslant\sum_{i=1}^{k}|\widetilde{s}_i|-kC_8. \tag{9.2}$$

Before going through the proof, we need some preparation on hyperbolic groups. Recall that X is a δ-hyperbolic space. Let $\langle x,y\rangle_z=(d(x,z)+d(y,z)-d(x,y))/2$ be the Gromov product, then $d(z,[x,y])-10\delta\leqslant\langle x,y\rangle_z\leqslant d(z,[x,y])$.

[8] This definition of independence is different with the one in [Yan19]. But this is enough for our applications for hyperbolic groups.

Definition 9.4. A (τ, D)-chain is a sequence of points $x_0, x_1, \cdots, x_n \in X$ such that

- $\langle x_{i-1}, x_{i+1} \rangle_{x_i} \leqslant \tau$ for every $1 \leqslant i \leqslant n-1$, and
- $d(x_{i-1}, x_i) \geqslant D$ for every $1 \leqslant i \leqslant n$.

By connecting two consecutive points in a (τ, D), we obtain a quasi-geodesic. The following lemma shows such path is indeed a uniform quasi-geodesic. See also [Gou22, Section 3B].

Lemma 9.5. *For every $\tau > 0$, there exists $D = D(\tau), c = c(\tau) > 0$ such that for every (τ, D)-chain $x_0, x_1, \cdots, x_n$, the path $\bigcup_{i=1}^{n} [x_{i-1}, x_i]$ is a c-quasi-geodesic.*

The following it a bounded projection property for different axes. A general version for spaces with contracting elements can be found in [Yan19, Lemma 2.17].

Lemma 9.6. *Let $f_1, f_2 \in \Gamma$ be independent loxodromic elements. There exists $\tau_1 > 0$ such that for every $g \in \Gamma$, we have*

$$\min \left\{ d(o, \pi_{\mathrm{Ax}(f_1)}(go)), d(o, \pi_{\mathrm{Ax}(f_2)}(go)) \right\} \leqslant \tau_1.$$

Proof. Let $z_i \in \pi_{\mathrm{Ax}(f_i)}(go)$, we first show that $[o, z_i][z_i, go]$ is a c_1-quasi-geodesic for some $c_1 > 0$ only depends on $\mathrm{Ax}(f_1)$ and $\mathrm{Ax}(f_2)$. Note that there exists $c_2 > 0$ such that $\mathrm{Ax}(f_i)$ is c_2-quasi-convex. Fixing an $i \in \{1, 2\}$, for every $x \in [o, z_i]$ and $y \in [z_i, go]$, we have

$$d(y, x) \geqslant d(y, \mathrm{Ax}(f_i)) - c_2 = d(y, z_i) - c_2.$$

Hence $d(x, z_i) + d(z_i, y) \leqslant d(x, y) + 2d(z_i, y) \leqslant 3d(x, y) + 2c_2$ is a $(3 + 2c_2)$-quasi-geodesic.

By Morse lemma, $[o, z_i] \subset \mathcal{N}_c([o, go])$ where $c = c(c_1) > 0$. If both $d(z_1, o)$ and $d(z_2, o)$ are larger than τ_1, we can choose $z_i' \in [o, z_i]$ with $d(o, z_i') = \tau_1$. Let $w_i \in [o, go]$ with $d(w_i, z_i') \leqslant c$, then $|d(w_i, o) - \tau_1| \leqslant c$. Hence $w_1, w_2 \in \mathcal{N}_{3c}(\mathrm{Ax}(f_1)) \cap \mathcal{N}_{3c}(\mathrm{Ax}(f_2))$. Since $d(w_1, o) \geqslant \tau_1 - c$, this contradicts Lemma 9.2 for a sufficiently large τ_1. $\square$

Definition 9.7. Let $g \in \Gamma$ and $\tau > 0$, we call a loxodromic element $f \in \Gamma$ is τ-contracting for g if for every $h \in E(f)$, we have $\langle go, ho \rangle_o \leqslant \tau$ and $\langle g^{-1}o, ho \rangle_o \leqslant \tau$.

Lemma 9.8. *Let $F \subset \Gamma$ be a finite subset of pairwise independent loxodromic elements with $\#F \geqslant 3$. There exists $\tau > 0$ such that for every $g \in \Gamma$, there exists $F_g \subset F$ with $\#F_g = \#F - 2$ such that every element in F_g is τ-contracting for g.*

Proof. By the previous lemma, there exists $\tau_1 > 0$ such that for every $f_1, f_2 \in F$ and $g \in \Gamma$, we have $\min \left\{ d(o, \pi_{\mathrm{Ax}(f_1)}(go)), d(o, \pi_{\mathrm{Ax}(f_2)}(go)) \right\} \leqslant \tau_1$. Note that there exists $c_1 > 0$ such that for every $f \in F$, $\mathrm{Ax}(f)$ is c_1-quasi-convex. We take $\tau = \tau_1 + c_1 + 10\delta$.

For each $g \in \Gamma$ and $f \in F$ satisfying $d(o, \pi_{\mathrm{Ax}(f)}(go)) \leqslant \tau_1$. For every $h \in E(f)$, we have $[o, ho] \in \mathcal{N}_{c_1}(\mathrm{Ax}(f))$. Hence

$$d(go, [o, ho]) \geqslant d(go, \mathrm{Ax}(f)) - c_1 \geqslant d(go, o) - d(o, \pi_{\mathrm{Ax}(f)}(go)) - c_1 \geqslant d(o, go) - \tau_1 - c_1.$$

Then $\langle o, ho \rangle_{go} \geqslant d(go, [o, ho]) - 10\delta \geqslant d(o, go) - \tau_1 - c_1 - 10\delta = d(o, go) - \tau$. We obtain

$$\langle go, ho \rangle_o = d(o, go) - \langle o, ho \rangle_{go} \leqslant \tau.$$

To complete the proof, we apply the argument to both g and g^{-1}. By the previous lemma, there are at least $\#F - 2$ elements in F which are τ-contracting for g. $\square$

Now we are at the stage of proving Proposition 9.3.

Proof of Proposition 9.3. We apply the previous lemma to F and obtain a constant $\tau > 0$. Let $C_7 = \#F(\#F-1)/2$. By the pigeonhole principle, there exists $F' \subset F$ with $\#F' = \#F - 2$ such that

$$S' = \{s \in S : F_s = F'\}$$

has the cardinality at least $C_7^{-1}\#S$, where $F_s \subset F$ is the subset given by the previous lemma consisting of τ-contracting elements.

By Lemma 9.5, there exists $D = D(\tau) > 0$ and $c_1 = c_1(\tau) > 0$ such that every (τ, D)-chain forms a c_1-quasi-geodesic. By Morse Lemma, every c-quasi-geodesic α is contained in $\mathcal{N}_C([\alpha_-, \alpha_+])$ for some $C > 0$. Now we take $C_8 = 2C, R = 4C + 1, L_0 = D + 4C$. The choice of m_0 will be given later. Let $c_2 > 0$ such that every axis $\mathrm{Ax}(f)$ is c_2-quasi-convex for $f \in F$. Now we show that $\widetilde{S}$ freely generates a free semigroup.

Let $k \geqslant 1$ and $s_i \in S', f_i \in F', \varsigma_i \in \{m, 2m\}$ for $1 \leqslant i \leqslant k$. We consider

$$x_i = s_1 f_1^{\varsigma_1} \cdots s_i f_i^{\varsigma_i} o, \quad y_i = s_1 f_1^{\varsigma_1} \cdots s_i o.$$

We claim that $x_0, y_1, x_1, \cdots, y_k, x_k$ is a (τ, D)-chain. Assume that m_0 is large enough guaranteeing $|f^m| > D$ for every $f \in F$ and $m \geqslant m_0$. Since $|s_i| = L \geqslant D + 4C$ and $|f_i^{\varsigma_i}| \geqslant D$ for each i, the second condition in Definition 9.4 is verified. Besides, for each i, we have

$$\langle x_{i-1}, x_i \rangle_{y_i} = \langle s_i^{-1} o, f_i^{\varsigma_i} o \rangle_o \leqslant \tau, \quad \langle y_i, y_{i+1} \rangle_{x_i} = \langle f_i^{-\varsigma_i} o, s_{i+1} o \rangle_o \leqslant \tau$$

by the τ-contracting property.

Hence $\alpha = [x_0, y_1][y_1, x_1] \cdots [x_{k-1}, y_k][y_k, x_k]$ is a c_1-quasi geodesic, which is contained in the C-neighborhood of $[\alpha_-, \alpha_+] = [o, x_k]$. Then we have two estimates on the length of $d(o, x_k)$. Firstly, since $|s_i| \geqslant D + 4C$ and $|f_i^{\varsigma_i}| \geqslant D$, we have

$$d(o, x_k) \geqslant \sum_{i=1}^{k} (d(x_{i-1}, y_i) + d(y_i, x_i)) - 4kC \geqslant k(2D + 2C) - 4kC > 0. \tag{9.3}$$

This implies that $x_k \neq o$. Besides, we have

$$d(o, x_k) \geqslant \sum_{i=1}^{k} d(x_{i-1}, x_i) - 2kC = \sum_{i=1}^{k} d(x_{i-1}, x_i) - kC_8,$$

which gives the desired estimate (9.2).

Checking the freeness. We consider two such sequences s_i, f_i, ς_i for $1 \leqslant i \leqslant k$ and s'_j, f'_j, ς'_j for $1 \leqslant j \leqslant l$. We get two sequences of points x_i, y_i and x'_j, y'_j in X. Assume that $x_k = x'_l$, we are going to show that $k = l$ and $s_i = s'_i, f_i = f'_i$, $\varsigma_i = \varsigma'_i$. By an inductive argument, it suffices to check for $i = 1$. Let α be a fixed geodesic connecting o and $x_k = x'_l$. Then there exist $z_1, z'_1 \in \alpha$ such that $d(y_1, z_1), d(y'_1, z'_1) \leqslant C$. Since $d(y_1, o) = d(y'_1, o) = L$, we have $d(z_1, z'_1) = |d(o, z_1) - d(o, z'_1)| \leqslant 2C$. This implies that $d(y_1, y'_1) \leqslant 4C$. By the R-separation, $y_1 = y'_1$.

Now we consider the geodesic β between $y_1 = y'_1$ and $x_k = x'_l$. Without loss of generality, we assume that $d(y_1, x_1) \leqslant d(y'_1, x'_1)$. Then there exists $w \in \beta$ such that $d(x_1, w) \leqslant C$. Let $t = d(y_1, x_1) = |f_1^{\varsigma_1}|$, then $d(w, y_1) \geqslant t - C$. Let $x' \in [y'_1, x'_1]$ such that $d(y'_1, x'_1) = t$. Note that $x' \in \mathcal{N}_C(\beta)$, we find that $d(x', w) \leqslant 3C$. This implies $w \in \mathcal{N}_{3C}([y_1, x_1]) \cap \mathcal{N}_{3C}([y'_1, x'_1])$. Note that $y_1, x_1 \in s_1 \mathrm{Ax}(f_1)$, hence $[y_1, x_1] \subset \mathcal{N}_{c_2}(s_1 \mathrm{Ax}(f_1))$ by the quasi-convexity of the axis. Therefore, both y_1 and w are contained in $s_1(\mathcal{N}_{3C+c_2}(\mathrm{Ax}(f_1)) \cap \mathcal{N}_{3C+c_2}(\mathrm{Ax}(f_1)))$. Since $d(y_1, w) \geqslant t - C$, we conclude that

$$\operatorname{diam} \mathcal{N}_{3C+c_2}(\mathrm{Ax}(f_1)) \cap \mathcal{N}_{3C+c_2}(\mathrm{Ax}(f'_1)) \geqslant d(y_1, w) \geqslant t - C.$$

By Lemma 9.2, the $(3C+c_2)$-neighborhoods of axes of different elements in F have uniformly bounded intersections. We can take m_0 sufficiently large at beginning such that $|f^m|-C$ is strictly larger than diameters of all such intersections for every $f\in F, m\geqslant m_0$, this forces $f_1=f_1'$.

Finally, it suffices to show $\varsigma_1=\varsigma_1'$. Otherwise, we assume that $\varsigma_1=m$ and $\varsigma_1'=2m$. In this case, we have

$$y_2=s_1f_1^m s_2 o,\quad x_1=s_1f_1^m o,\quad x_1'=s_1f_1^{2m}o,\quad y_2'=s_1f_1^{2m}s_2'o.$$

By the τ-contracting property, these points form a (τ,D)-chain. This leads to that

$$x_k,y_k,\cdots,y_2,x_1,x_1',y_2',\cdots,y_l',x_l'.$$

is also a (τ,D)-chain. Applying the estimate in (9.3), this contradicts $x_k=x_l'$. ☐

9.3 Proof of Proposition 9.1.

To prove Proposition 9.1, we need to find a finite subset consisting of pairwise independent elements satisfying the desired condition on Zariski closures.

Lemma 9.9. *Let Γ be a non-cyclic torsion-free hyperbolic group and $\rho:\Gamma\to \mathrm{SL}_n(\mathbb{R})$ be a faithful representation. Let $\mathbf{G}$ be the Zariski closure of $\rho(\Gamma)$ and assume $\mathbf{G}$ is Zariski connected. For every $k\geqslant 0$, there exists a finite subset $F\subset\Gamma$ consisting of pairwise independent nontrivial elements such that for every subset $F'\subset F$ with $\#F'\geqslant F-k$, $\rho(F')$ generates a semigroup whose Zarski closure is $\mathbf{G}$.*

Proof. We will prove this lemma by an induction on k. We will find $F=F_k$ for each $k\geqslant 0$.

For the case of $k=0$, applying [MS23, Lemma 3.6], we can find a finite subset $F_0\subset\Gamma$ such that $\rho(F_0)$ generates a semigroup whose Zariski closure is $\mathbf{G}$. Furthermore, we can assume that elements in F_0 are pairwise independent by the following process. If $f_1,f_2\in F_0$ are not independent, then there exists $f\in\Gamma$ and $m_1,m_2\in\mathbb{Z}$ such that $f_i=f^{m_i}$ since Γ is torsion free. We can remove f_1,f_2 and add f in F_0.

Assume that F_k is found. Recall that for every $f\in\Gamma$, $E(f)$ is a cyclic group containing all elements which are not independent of f. Note that Zariski closure of $\rho(E(f))$ is commutative. On the other hand, applying Tits alternative for hyperbolic groups, we obtain that Γ contains a nonabelian free subgroup. This implies that $\rho(E(f))$ is not Zariski dense in $\mathbf{G}$. By the connectivity of $\mathbf{G}$, the set $S=\Gamma\setminus\bigcup_{f\in F_k}E(f)$ is Zariski dense in $\mathbf{G}$. Using the argument for $k=0$ case, we can find a finite subset $\widetilde{F}\subset S$ consisting of pairwise independent elements such that $\rho(\widetilde{F})$ generates a semigroup whose Zariski closure is $\mathbf{G}$. We take $F_{k+1}=F_k\cup\widetilde{F}$, which is a desired construction for $(k+1)$-case. ☐

Proof of Proposition 9.1. Applying Lemma 9.9 to the case of $k=2$, we can find a finite subset $F\subset\Gamma$ consisting of pairwise independent elements such that for every $F'\subset F$ with $\#F'=\#F-2$, $\rho(F')$ generates a semigroup whose Zariski closure is $\mathbf{G}$. We apply Proposition 9.3 to F. Let m_0 be the constant given by this proposition. In order to find $m\geqslant m_0$ satisfying the first condition in the proposition, we need the following lemma.

Lemma 9.10. *For every $g\in\mathrm{SL}_n(\mathbb{R})$, there exists a positive integer $l=l(g)$ such that for every m coprime with l, the Zariski closure of $\langle g^m\rangle$ is equal to the Zariski closure of $\langle g\rangle$.*

Proof. Let $\mathbf{H}$ be the Zariski closure of $\langle g\rangle$. Let l be the number of connected components of $\mathbf{H}$. Let $\mathbf{H}'$ be the Zariski closure of $\langle g^m\rangle$ where m is coprime with l. Then $\mathbf{H}'$ is a finite index algebraic subgroup of $\mathbf{H}$ and hence a union of some connected components of $\mathbf{H}$. Note that the action of $\langle g\rangle$ given by left translation is transitive among connected components of $\mathbf{H}$ and so does $\langle g^m\rangle$, due to m coprime with l. Hence $\mathbf{H}'=\mathbf{H}$. ☐

We take $m \geqslant m_0$ to be a sufficiently large prime number such that for every $f \in F$, the Zariski closure of $\langle \rho(f)^m \rangle$ equals to the Zariski closure of $\langle \rho(f) \rangle$. Then for every $F' \subset F$ with $\# F' = \# F - 2$, the Zariski closure of the semigroup generated by $\{\rho(f)^m : f \in F'\}$ equals to that of $\rho(F)$, which is $\mathbf{G}$. The first condition holds.

Note that for every finite subset $S \subset \Gamma$, it contains an R-separated subset of cardinality at least $(\#\mathcal{S}+1)^{-R}\# S$, where $\mathcal{S}$ is the symmetric generating set of Γ. Then the last two conditions follow from Proposition 9.3 by enlarging C_7 suitably. □

10 Variational principle for Anosov representations and applications: the proof of Theorem 1.3.

In this section, we will show a variational principle for dimensions of limit sets of an Anosov representation. Let Γ be a finitely generated hyperbolic group with a fixed finite symmetric generating set $\mathcal{S} \subset \Gamma$. Recall $|\gamma|$ denotes the word norm of $\gamma \in \Gamma$ with respect to $\mathcal{S}$. Let $\rho : \Gamma \to \mathrm{SL}_n(\mathbb{R})$ be a Borel Anosov representation. Recall Definition 2.3, $\kappa(\rho(\gamma)) = (\log \sigma_1(\rho(\gamma)), \cdots, \log \sigma_n(\rho(\gamma)))$ is the Cartan projection of $\rho(\gamma)$.

10.1 Key proposition

Let ψ be a linear functional on the Lie algebra $\mathfrak{a} = \{\lambda = (\lambda_1, \cdots, \lambda_n) | \lambda_i \in \mathbb{R},\ \sum \lambda_i = 0\}$, which is positive with respect to simple roots $\alpha_i(\lambda) = \lambda_i - \lambda_{i+1}$. That is, there are not all zero constants $a_1, \cdots, a_{n-1} \geqslant 0$ such that for $\lambda \in \mathfrak{a}$

$$\psi(\lambda) = \sum_{i=1}^{n-1} a_i \cdot \alpha_i(\lambda).$$

These simple roots α_i are non-negative on the positive Weyl chamber $\mathfrak{a}^+$. Recall that for a finitely supported probability measure ν on $\mathrm{SL}_n(\mathbb{R})$, $\lambda(\nu) = (\lambda_1(\nu), \cdots, \lambda_n(\nu))$ denotes the Lyapunov spectrum of ν, which is also a vector in the positive Weyl chamber $\mathfrak{a}^+$. Then $\psi(\lambda(\nu))$ is a nonnegative number.

Throughout this subsection, we further assume that Γ is a torsion-free hyperbolic group and ρ is a faithful representation. Note that Γ is non-cyclic since ρ is irreducible. Recall that q is a large integer which is the base of logarithm in this article.

Proposition 10.1. *Let $\mathbf{G}$ be the Zariski closure of $\rho(\Gamma)$ which is assumed to be Zariski connected. Let ψ be a linear functional as above. If the series $\sum_{\gamma \in \Gamma} q^{-\psi(\kappa(\rho(\gamma)))}$ diverges, then there exists $c > 0$ such that the following holds. For every $\epsilon > 0$, there exists infinitely many positive integers N with a finitely supported probability measure ν on $\rho(\Gamma)$ such that*

- *G_ν is Zariski dense in $\mathbf{G}$.*
- *$\lambda_p(\nu) - \lambda_{p+1}(\nu) \geqslant cN$ for every $p = 1, \cdots, n-1$.*
- *$h_{\mathrm{RW}}(\nu) \geqslant (1-\epsilon)N$ and $\psi(\lambda(\nu)) \leqslant (1+\epsilon)N$.*

We first recall an estimate on the lost of singular values under composition. The following lemma is a direct consequence of combining Lemmas 2.5 and A.7 in [BPS19].

Lemma 10.2. *Let $1 \leqslant p \leqslant n$. Given $c > 0$, then there exists $\delta > 0$ such that the following holds. Let $(\gamma_k)_{k \in \mathbb{N}} \in \mathcal{S}^{\mathbb{N}}$ satisfying for every $l \leqslant m$,*

$$\frac{\sigma_p(\rho(\gamma_{l+1} \cdots \gamma_m))}{\sigma_{p+1}(\rho(\gamma_{l+1} \cdots \gamma_m))} \geqslant c \cdot e^{c(m-l)}.$$

Then for every $l \leqslant k \leqslant m$, we have

$$\sigma_p(\rho(\gamma_{l+1}\cdots\gamma_m)) \geqslant \delta \cdot \sigma_p(\rho(\gamma_{l+1}\cdots\gamma_k))\sigma_p(\rho(\gamma_{k+1}\cdots\gamma_m)),$$

$$\sigma_{p+1}(\rho(\gamma_{l+1}\cdots\gamma_m)) \leqslant \delta^{-1} \cdot \sigma_{p+1}(\rho(\gamma_{l+1}\cdots\gamma_k))\sigma_{p+1}(\rho(\gamma_{k+1}\cdots\gamma_m)).$$

Proof of Proposition 10.1. Applying Proposition 9.1, we obtain a finite subset $F \subset \Gamma$ with $\#F \geqslant 3$, constants $C_7, C_8, L_0 > 0$ and a positive integer m. For every $\epsilon > 0$ sufficiently small, there are infinitely many integers N such that

$$S_1 = \{\gamma \in \Gamma : \psi(\kappa(\rho(\gamma))) \leqslant N\}$$

has cardinality at least $q^{(1-\epsilon)N}$ due to the divergence of the series. Since ψ is positive, we have $\psi(\kappa(\rho(\gamma))) \geqslant c_1(\log\sigma_p(\rho(\gamma)) - \log\sigma_{p+1}(\rho(\gamma)))$ for some $1 \leqslant p \leqslant n-1$ and $c_1 > 0$. Because ρ is Borel Anosov, there exists $c_2 > 0$ such that for $|\gamma|$ large enough,

$$\psi(\kappa(\rho(\gamma))) \geqslant c_1(\log\sigma_p(\rho(\gamma)) - \log\sigma_{p+1}(\rho(\gamma))) \geqslant c_2|\gamma|.$$

Hence $S_1 \subset \left\{\gamma \in \Gamma : |\gamma| \leqslant c_2^{-1}N\right\}$. Let $c_3 = (2\#\log\mathcal{S})^{-1}$, then

$$\#\left\{\gamma \in \Gamma : |\gamma| \leqslant c_3N\right\} \leqslant 2(\#\mathcal{S})^{c_3N} \leqslant \frac{1}{2}q^{(1-\epsilon)N}$$

providing ϵ small and N large. Hence there exists $c_3N \leqslant L \leqslant c_2^{-1}N$ such that

$$S_2 := \{\gamma \in S_1 : |\gamma| = L\}$$

has cardinality at least $c_2(2N)^{-1}q^{(1-\epsilon)N} \geqslant q^{(1-2\epsilon)N}$ assuming N large. By Proposition 9.1 , there exists $S_3 \subset S_2$ and $F' \subset F$ with $\#S_3 \geqslant C_7^{-1}\#S_2$ and $\#F' = \#F - 2$ such that

$$\widetilde{S} := \left\{sf^{\varsigma} : s \in S_3, f \in F', \varsigma = m, 2m\right\}$$

freely generates a free semigroup. Assuming N large, we have $\#\widetilde{S} \geqslant \#S_3 \geqslant q^{(1-3\epsilon)N}$. Letting ν be the uniform measure on $\rho(\widetilde{S})$, we now verify that this is a desired construction.

- Note that $\rho(f)^m = \rho(sf^m)^{-1}\rho(sf^{2m}) \in G_\nu$ for every $f \in F'$. By the first condition in 9.1, we have G_ν is Zariski dense in $\mathbf{G}$.

- Take $c_4 > 0$ such that for every $1 \leqslant p \leqslant n-1$ and $|\gamma|$ large enough, $\log\sigma_p(\rho(\gamma)) - \log\sigma_{p+1}(\rho(\gamma)) \geqslant c_4|\gamma|$. Assuming N is large enough only depends on m, for every $\widetilde{s} \in \widetilde{S}$ we have

  $$|\widetilde{s}| \geqslant c_3N - 2m\max_{f\in F}|f| \geqslant c_3'N + C_8 \quad \text{and} \quad |\widetilde{s}| \leqslant c_2^{-1}N + 2m\max_{f\in F}|f| \leqslant c_2'^{-1}N$$

  where $c_2' = c_2/2, c_3' = c_3/2$ and C_8 is the constant given by Proposition 9.1. Then for every $\widetilde{s}_1, \cdots \widetilde{s}_k \in \widetilde{S}$, due to Eq. (9.2), we have $|\widetilde{s}_1\cdots\widetilde{s}_k| \geqslant c_3'kN$. Hence for every $1 \leqslant p \leqslant n-1$, we have

  $$(\log\sigma_p - \log\sigma_{p+1})(\rho(\widetilde{s}_1\cdots\widetilde{s}_k)) \geqslant c_4c_3'kN \geqslant c_2'c_3'c_4\sum_{i=1}^{k}|\widetilde{s}_i|. \tag{10.1}$$

  Recall that $\operatorname{supp}\nu = \widetilde{S}$. By the first inequality in Eq. (10.1) we have

  $$\lambda_p(\nu) - \lambda_{p+1}(\nu) = \lim_{k\to\infty}\frac{1}{k}\int[\log\sigma_p(\rho(\gamma)) - \log\sigma_{p+1}(\rho(\gamma))]\mathrm{d}\nu^{*k}(\gamma) \geqslant c_4c_3'N.$$

  Taking $c = c_4c_3'$, we obtain the conclusion.

- Since $\widetilde{S}$ freely generates a free semigroup and ν is the uniform measure on $\widetilde{S}$, we have

$$h_{\mathrm{RW}}(\nu) = \log \# \widetilde{S} \geqslant (1-3\epsilon)N.$$

  Shrinking $\epsilon > 0$, we obtain the first estimate.

  In order to estimate $\psi(\lambda(\nu))$, we need an almost additivity property of $\log \sigma_p$. Recall the second inequality in (10.1). By applying Lemma 10.2, there exists $\delta > 0$ only depending on C_8 and $c_2' c_3' c_4$ such that

$$\left| \log \sigma_p(\rho(\widetilde{s}_1 \cdots \widetilde{s}_k)) - \sum_{i=1}^{k} \log \sigma_p(\widetilde{s}_i) \right| \leqslant -k \log \delta$$

  for every $\widetilde{s}_1, \cdots, \widetilde{s}_k \in \widetilde{S}$ and $1 \leqslant p \leqslant n$. Since ψ is a linear functional, we have

$$\left| \psi(\kappa(\rho(\widetilde{s}_1 \cdots \widetilde{s}_k))) - \sum_{i=1}^{k} \psi(\kappa(\rho(\widetilde{s}_i))) \right| \leqslant -k C_3 \log \delta,$$

  where $C_3 > 0$ only depends on ψ. For each $\widetilde{s} \in \widetilde{S}$, write $\widetilde{s} = s f^{\varsigma}$, where $s \in S_3$ and $f \in F, \varsigma \in \{m, 2m\}$. Then there exists $C_4 > 0$ only depending on F and m such that $|\log \sigma_p(\rho \widetilde{s}) - \log \sigma_p(\rho s)| \leqslant C_4$ for every p. Hence $\psi(\kappa(\rho \widetilde{s})) \leqslant N + C_3 C_4$ since $s \in S_3 \subset S_1$.

  Then for sufficiently large N, we have

$$\psi(\kappa(\rho(\widetilde{s}_1 \cdots \widetilde{s}_k))) \leqslant kN + kC_7 C_2 - k \log \delta \leqslant (1+\epsilon)kN.$$

  This implies that

$$\psi(\lambda(\nu)) = \lim_{k \to \infty} \frac{1}{k} \int \psi(\kappa(\rho(\gamma))) \mathrm{d}\nu^{*k}(\gamma) \leqslant (1+\epsilon)N. \qquad \square$$

10.2 The variational principle of dimensions

Let $\rho : \Gamma \to \mathrm{SL}_n(\mathbb{R})$ be a Borel Anosov representation and $\mathbf{G}$ be the Zariski closure of $\rho(\Gamma)$. Recall $s_A(\rho)$ is the affinity dimension of the limit set of $\rho(A)$ on $\mathbb{P}(\mathbb{R}^n)$, which is given by

$$s_A(\rho) := \sup \left\{ s : \sum_{\gamma \in \Gamma} q^{-\psi_s(\rho(\gamma))} = \infty \right\}$$

where

$$\psi_s(g) := \sum_{1 \leqslant i \leqslant \lfloor s \rfloor} (\log \sigma_1(g) - \log \sigma_{i+1}(g)) + (s - \lfloor s \rfloor)(\log \sigma_1(g) - \log \sigma_{\lfloor s \rfloor + 2}(g)), \quad g \in \mathrm{SL}_n(\mathbb{R}).$$

Notice that in order to simplify the notation, we abbreviate $\psi_s(\kappa(g))$ to $\psi_s(g)$.

Let $\mathscr{P}_{\mathrm{f.s.}}(\rho)$ be the family of finitely supported probability measures on $\rho(\Gamma)$. Let $\mathscr{P}^{\mathbf{G}}_{\mathrm{f.s.}}(\rho)$ be the family of $\nu \in \mathscr{P}_{\mathrm{f.s.}}(\rho)$ satisfying G_ν is Zariksi dense in $\mathbf{G}$. We will also consider $\mathscr{P}^{\mathbf{G}^0}_{\mathrm{f.s.}}(\rho)$ where $\mathbf{G}^0$ denotes the identity (Zariski-)component of $\mathbf{G}$. For $\nu \in \mathscr{P}_{\mathrm{f.s}}(\rho)$, let μ be an ergodic stationary measure of ν on $\mathbb{P}(\mathbb{R}^d)$. Then the Lyapunov dimension of μ is given by

$$\dim_{\mathrm{LY}} \mu = d + \frac{h_{\mathrm{F}}(\mu, \nu) - (\chi_1(\nu) + \cdots + \chi_d(\nu))}{\chi_{d+1}(\nu)},$$

where d is the maximal integer such that $\chi_1(\nu) + \cdots + \chi_d(\nu) \leqslant h_{\mathrm{F}}(\mu, \nu)$.

Proposition 10.3. *Let $\rho : \Gamma \to \mathrm{SL}_n(\mathbb{R})$ be a Zariski dense Borel Anosov representation. Then*

$$\begin{aligned} s_A(\rho) &\leqslant \sup\left\{\dim_{\mathrm{LY}} \mu : \mu \textit{ is the unique stationary measure on } \mathbb{P}(\mathbb{R}^n) \textit{ of some } \nu \in \mathscr{P}^{\mathbf{G}}_{\mathrm{f.s.}}(\rho)\right\} \\ &\leqslant \sup\left\{\dim_{\mathrm{LY}} \mu : \mu \textit{ is an ergodic stationary measure on } \mathbb{P}(\mathbb{R}^n) \textit{ of some } \nu \in \mathscr{P}_{\mathrm{f.s.}}(\rho)\right\}. \end{aligned}$$

Besides, we also have a variational principle for dimensions on the flag variety. We consider

$$\psi_{F,s}(g) \coloneqq \inf\left\{\sum_{1\leqslant i<j\leqslant n} a_{ij}\log\frac{\sigma_i(g)}{\sigma_j(g)} : 0 \leqslant a_{ij} \leqslant 1, \sum_{1\leqslant i<j\leqslant n} a_{ij} = s\right\}.$$

Note that $\psi_{F,s}(g)$ is increasing with respect to s. The affinity dimension on the flag variety is given by

$$s_{A,F}(\rho) \coloneqq \sup\left\{s : \sum_{\gamma\in\Gamma} q^{-\psi_{F,s}(\rho(\gamma))} = \infty\right\}.$$

For instance, in the case of $n = 3$, the function $\psi_{F,s}$ is given by

$$\psi_{F,s}(g) = \begin{cases} s\min\left\{\log\dfrac{\sigma_1(g)}{\sigma_2(g)}, \log\dfrac{\sigma_2(g)}{\sigma_3(g)}\right\}, & s \leqslant 1; \\ (s-1)\left(\log\dfrac{\sigma_1(g)}{\sigma_2(g)} + \log\dfrac{\sigma_2(g)}{\sigma_3(g)}\right) + (2-s)\min\left\{\log\dfrac{\sigma_1(g)}{\sigma_2(g)}, \log\dfrac{\sigma_2(g)}{\sigma_3(g)}\right\}, & 1 < s \leqslant 2; \\ \log\dfrac{\sigma_1(g)}{\sigma_2(g)} + \log\dfrac{\sigma_2(g)}{\sigma_3(g)} + (s-2)\log\dfrac{\sigma_1(g)}{\sigma_3(g)}, & 2 < s \leqslant 3. \end{cases}$$

In general, $\psi_{F,s}(g)$ is always a minimum of finitely many positive linear functionals on $\log\sigma_i(g) - \log\sigma_{i+1}(g)$. Specifically, $\psi_{F,s}(g) = \min\{\psi_{F,k}(g) : k \in \mathcal{K}_s\}$, where each $\psi_{F,k}(g)$ is of the form

$$\sum_{1\leqslant i<j\leqslant n} a_{ij}(\log\sigma_i(g) - \log\sigma_j(g))$$

satisfying $0 \leqslant a_{ij} \leqslant 1, \Sigma_{1\leqslant i<j\leqslant n} a_{ij} = s$ and at most one of a_{ij} is neither 0 nor 1. For given s and n, there are at most finitely many such linear functionals. Hence $\#\mathcal{K}_s$ is finite.

For $\nu \in \mathscr{P}_{\mathrm{f.s}}(\rho)$, let μ be an ergodic stationary measure of ν on $\mathcal{F}(\mathbb{R}^n)$. Then the Lyapunov dimension of μ is given by

$$\dim_{\mathrm{LY}} \mu = \sup\left\{\sum_{1\leqslant i<j\leqslant n} d_{ij} : 0 \leqslant d_{ij} \leqslant 1, \sum_{1\leqslant i<j\leqslant n} d_{ij}(\lambda_i(\nu) - \lambda_j(\nu)) = h_{\mathrm{F}}(\mu,\nu)\right\}.$$

Proposition 10.4. *Let $\rho : \Gamma \to \mathrm{SL}_n(\mathbb{R})$ be a Borel Anosov representation. Then*

$$\begin{aligned} s_{A,F}(\rho) &\leqslant \sup\left\{\dim_{\mathrm{LY}} \mu : \mu \textit{ is an ergodic stationary measure on } \mathcal{F}(\mathbb{R}^n) \textit{ of some } \nu \in \mathscr{P}^{\mathbf{G}^0}_{\mathrm{f.s.}}(\rho)\right\} \\ &\leqslant \sup\left\{\dim_{\mathrm{LY}} \mu : \mu \textit{ is an ergodic stationary measure on } \mathcal{F}(\mathbb{R}^n) \textit{ of some } \nu \in \mathscr{P}_{\mathrm{f.s.}}(\rho)\right\}. \end{aligned}$$

We will prove the case on flag varieties first. The case on the projective space is very similar and we will only indicate where we need to modify during the proof.

Proof of Proposition 10.4. Since ρ is Anosov, $\ker\rho$ is finite. Replacing Γ by $\Gamma/\ker\rho$, we can assume that ρ is faithful. Besides, by Selberg's lemma, $\rho(\Gamma)$ is virtually torsion-free. Replacing Γ by a finite index torsion-free subgroup, we may assume that Γ is torsion free and the Zariski closure of $\rho(\Gamma)$ is Zariski connected. This process does not affect the value of affinity dimension and the Zariski closure of $\rho(\Gamma)$ is indeed the identity component of the original one. Then the proposition is a direct consequence of the following lemma.

Lemma 10.5. *Let Γ be a non-cyclic torsion-free hyperbolic group and $\rho : \Gamma \to \mathrm{SL}_n(\mathbb{R})$ be faithful such that the Zariski closure $\mathbf{G}$ of $\rho(\Gamma)$ is Zariski connected. Let $s > 0$ such that the series $\sum_{\gamma\in\Gamma} q^{-\psi_{F,s}(\rho(\gamma))}$ diverges. Then for every $\epsilon > 0$ sufficiently small, there exists $\nu \in \mathscr{P}^{\mathbf{G}}_{\mathrm{f.s.}}(\rho)$ and an ergodic ν-stationary measure μ on $\mathcal{F}(\mathbb{R}^n)$ satisfying $\dim_{\mathrm{LY}} \mu \geqslant s - \epsilon$.*

Proof. Recall that $\psi_{F,s}(\rho(\gamma)) = \min\{\psi_{F,k}(\rho(\gamma)) : k \in \mathcal{K}_s\}$ is a minimum of finitely many linear functionals. Then

$$\sum_{k\in\mathcal{K}_s}\sum_{\gamma\in\Gamma} q^{-\psi_{F,k}(\rho(\gamma))} \geqslant \sum_{\gamma\in\Gamma} q^{-\psi_{F,s}(\rho(\gamma))} = \infty.$$

This implies that there exists $k \in \mathcal{K}_s$ satisfying $\sum_{\gamma\in\Gamma} q^{-\psi_{F,k}(\rho(\gamma))}$ diverges. We fix a such $\psi_{F,k}$ in latter discussions and assume that

$$\psi_{F,k}(g) = \sum_{1\leqslant i<j\leqslant n} a_{ij}(\log\sigma_i(g) - \log\sigma_j(g)), \quad \forall g \in \mathrm{SL}_n(\mathbb{R}).$$

Now we apply Proposition 10.1 to $\psi_{F,k}$. We obtain a constant $c > 0$ and there exists a positive integer N and $\nu \in \mathscr{P}_{\mathrm{f.s.}}(\rho)$ satisfying

- $\operatorname{supp}\nu$ generates a semigroup whose Zariski closure is $\mathbf{G}$.
- $\lambda_p(\nu) - \lambda_{p+1}(\nu) \geqslant cN$ for every $p = 1, \cdots, n-1$.
- $h_{\mathrm{RW}}(\nu) \geqslant (1 - \frac{1}{2}c\epsilon)N$ and $\psi_{F,k}(\lambda(\nu)) \leqslant (1 + \frac{1}{2}c\epsilon)N$.

Since ν has a simple Lyapunov spectrum, there exists a ν-stationary measure μ on $\mathcal{F}(\mathbb{R}^n)$ which corresponds to the distribution of Oseledec's splitting. Furthermore, $(\operatorname{supp}\mu, \mu)$ is the Poisson boundary for (Γ_ν, ν) by [Fur02, Theorem 2.21], where Γ_ν is the group generated by $\operatorname{supp}\nu$. By [Fur02, Theorem 2.31], $h_{\mathrm{F}}(\mu,\nu) = h_{\mathrm{RW}}(\nu) \geqslant (1 - \frac{1}{2}c\epsilon)N$. Now we estimate $\dim_{\mathrm{LY}}\mu$.

Note that

$$\psi_{F,k}(\lambda(\nu)) = \sum_{1\leqslant i<j\leqslant n} a_{ij}(\lambda_i(\nu) - \lambda_j(\nu)) \leqslant (1 + \frac{1}{2}c\epsilon)N.$$

Assuming $a_{i_0 j_0} > 0$ for some $1 \leqslant i_0 < j_0 \leqslant n$, we take

$$a'_{ij} = \begin{cases} a_{ij} - \epsilon, & i = i_0, j = j_0; \\ a_{ij}, & \text{otherwise.} \end{cases}$$

Then

$$\sum_{1\leqslant i<j\leqslant n} a'_{ij}(\lambda_i(\nu) - \lambda_j(\nu)) \leqslant (1 + \frac{1}{2}c\epsilon)N - \epsilon(\lambda_{i_0}(\nu) - \lambda_{j_0}(\nu)) \leqslant (1 - \frac{1}{2}c\epsilon)N \leqslant h_{\mathrm{F}}(\mu,\nu).$$

Hence

$$\dim_{\mathrm{LY}}\mu \geqslant \sum_{1\leqslant i<j\leqslant n} a'_{ij} = \sum_{1\leqslant i<j\leqslant n} a_{ij} - \epsilon = s - \epsilon.$$

We obtain the desired conclusion. □

□

Proof of Proposition 10.3. The only difference is replacing Furstenberg entropy by the random walk entropy. Since we assume that $\rho(\Gamma)$ is Zariski dense in $\mathrm{SL}_n(\mathbb{R})$, this follows from Proposition 8.3. □

10.3 Proof of Theorem 1.3

A direct application of the dimension variation is to compute the dimension of the minimal sets. Once we establish the identity between the Lyapunov dimension and the exact dimension of stationary measures, we can obtain the identity between the affinity dimension and the Hausdorff dimension of minimal sets.

Proof of Theorem 1.3. Recall that $\mathbf{G}$ is the Zariski closure of $\rho(\Gamma)$. We first consider the case $\mathbf{G} = \mathrm{SL}_3(\mathbb{R})$. Note that $\rho(\Gamma)$ is a discrete subgroup of $\mathrm{SL}_3(\mathbb{R})$. Then for every $\nu \in \mathscr{P}^{\mathbf{G}}_{\text{f.s.}}(\rho(\Gamma))$, ν is a Zariski dense finitly supported probability measure with exponential separation. Hence the unique ν-stationary measure μ on $\mathbb{P}(\mathbb{R}^3)$ satisfying $\dim_{\mathrm{LY}} \mu = \dim \mu$ by Theorem 1.10. Moreover, μ is supported on the limit set $L(\rho(\Gamma)) \subset \mathbb{P}(\mathbb{R}^3)$. By [You82], $\dim \mu = \dim_{\mathrm{H}} \mu \leqslant \dim_{\mathrm{H}} L(\rho(\Gamma))$. Combining with the upper bound estimate of Hausdorff dimension in [PSW21], we obtain

$$\begin{aligned}\dim_{\mathrm{H}} L(\rho(\Gamma)) \leqslant s_A(\rho) &\leqslant \sup\left\{\dim_{\mathrm{LY}} \mu : \nu \in \mathscr{P}^{\mathbf{G}}_{\text{f.s.}}(\rho)\right\} \\ &= \sup\left\{\dim \mu : \nu \in \mathscr{P}^{\mathbf{G}}_{\text{f.s.}}(\rho)\right\} \leqslant \dim_{\mathrm{H}} L(\rho(\Gamma)).\end{aligned}$$

Then all unequal signs are equal. In particular, $\dim_{\mathrm{H}} L(\rho(\Gamma)) = s_A(\rho)$.

Secondly, we consider the case $\mathbf{G} \neq \mathrm{SL}_3(\mathbb{R})$. Since ρ is an irreducible Anosov representation, due to [BCLS15, Corollary 2.20][9], the Zariski closure of $\rho(\Gamma)$ is semi-simple without compact factor and its centralizer is contained in $\{\pm \mathrm{Id}\}$. So $\rho(\Gamma)$ is the image of an irreducible representation of $\mathrm{SL}_2(\mathbb{R})$. We can write ρ as the composition of $\rho_0 : \Gamma \to \mathrm{SL}_2(\mathbb{R})$ and an irreducible representation $\iota_1 : \mathrm{SL}_2(\mathbb{R}) \to \mathrm{SL}_3(\mathbb{R})$. The representation ι_1 induces an algebraic map from $\mathbb{P}(\mathbb{R}^2)$ to $\mathbb{P}(\mathbb{R}^3)$. We only need to compute the Hausdorff dimension of $L(\rho_0(\Gamma))$. Due to classical result of Patterson, Bowen and Sullivan, the Hausdorff dimension of the limit set is also equal to critical exponent, which is equal to affinity exponent in this case.

The proof of the continuity of $s_A(\rho)$ is postponed to the next subsection. □

10.4 Continuity and discontinuity of dimensions: proof of Theorem 1.1 and the continuity part of Theorem 1.3.

Recall that $\mathrm{HA}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$ is the space of Anosov represesntations of Γ inside $\mathrm{Hom}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$, which is an open subset. In the following of this section, we always assume that ρ is a representation in $\mathrm{HA}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$. Recall that the Lie algebra $\mathfrak{a} = \{\lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3) : \lambda_i \in \mathbb{R},\ \sum \lambda_i = 0\}$ and the positive Weyl chamber $\mathfrak{a}^+ = \{\lambda \in \mathfrak{a} : \lambda_1 \geqslant \lambda_2 \geqslant \lambda_3\}$. For an element $g \in \mathrm{SL}_3(\mathbb{R})$ let $\lambda(g) \in \mathfrak{a}^+$ be its Jordan projection. Let α_1 and α_2 be two simple roots on $\mathfrak{a}$, given by $\alpha_i(\lambda) = \lambda_i - \lambda_{i+1}$, which are nonnegative on $\mathfrak{a}^+$. For any nonzero $\psi \in \{a\alpha_1 + b\alpha_2 :\ a, b \in \mathbb{R}_{\geqslant 0}\} \subset \mathfrak{a}^*$, we define

$$h^{\psi}_{\rho} = \limsup_{T \to \infty} \frac{1}{T} \log \#\{[\gamma] \in [\Gamma] :\ \text{non torsion},\ \psi(\lambda(\rho(\gamma))) \leqslant T\},$$

where $[\Gamma]$ is the set of conjugacy classes and $[\gamma]$ is the conjugacy class of γ in Γ. For such ψ, we have $h^{\psi}_{\rho} \in (0, \infty)$ since $\rho \in \mathrm{HA}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$. By [PS17, Cor. 4.9] or [BCLS15], we know that h^{ψ}_{ρ} is analytic with respect to ρ in $\mathrm{HA}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$.

In order to study the affinity exponent, we consider the following function ψ_s on $\mathfrak{a}$ as

$$\psi_s = \begin{cases} s\alpha_1, & 0 < s \leqslant 1; \\ \alpha_1 + (s-1)(\alpha_1 + \alpha_2), & 1 < s \leqslant 2. \end{cases} \tag{10.2}$$

Since ρ is an Anosov representation, we obtain that $h^{\psi_s}_{\rho} \in (0, \infty)$ for every $0 < s \leqslant 2$.

Lemma 10.6. *(1) For any $0 < t < 1$, we have $h^{\psi_{ts}}_{\rho} > t h^{\psi_{ts}}_{\rho} \geqslant h^{\psi_s}_{\rho}$;*

[9] The statement of Corollary is for projective Anosov representation. In $\mathrm{SL}_3(\mathbb{R})$, projective Anosov is equivalent to Anosov, so we can freely use this Corollary.

(2) *For* $0 < t < 1, 0 < s \leqslant 1$, *we have* $th_\rho^{\psi_{ts}} = h_\rho^{\psi_s}$;

(3) *For* $0 < t < 1, s > 1$ *and* $ts > 1$, *we have*

$$h_\rho^{\psi_s} \geqslant \frac{ts-1}{s-1} h_\rho^{\psi_{ts}}.$$

Proof. The first two items can be checked by definition. The third one follows from the inequality $\psi_s \leqslant \frac{s-1}{ts-1}\psi_{ts}$ when $0 < t < 1, s > 1$ and $ts > 1$. □

As a corollary, we have

Lemma 10.7. $h_\rho^{\psi_s}$ *is a strictly decreasing continuous function on* s.

Lemma 10.8. *The affinity exponent* $s_A(\rho)$ *is the unique* s *such that* $h_\rho^{\psi_s} = 1$.

Proof. In [Sam14, Corollary 4.4], it is proved that for Anosov representation ρ, the exponent $h_\rho^{\psi_s}$ is also equal to

$$\limsup_{T\to\infty} \frac{1}{T} \log \#\{\gamma \in \Gamma : \ \psi_s(\kappa(\rho\gamma)) \leqslant T\},$$

which equals the critical exponent of the series

$$\sum_{\gamma\in\Gamma} q^{-t\psi_s\kappa(\rho\gamma)}$$

with respect to $t \in \mathbb{R}$. Here q is a large integer which is the base of logarithm in this article. Hence, if $h_\rho^{\psi_s} > 1$, then the series $\sum_{\gamma\in\Gamma} q^{-\psi_s\kappa(\rho\gamma)}$ diverges; if $h_\rho^{\psi_s} < 1$, then the series $\sum_{\gamma\in\Gamma} q^{-\psi_s\kappa(\rho\gamma)}$ converges. By the definition of affinity exponent, this completes the proof. □

We are able to prove the continuity part in Theorem 1.3.

Proposition 10.9. *The affinity exponent* $s_A(\rho)$ *is continuous on* $\mathrm{HA}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$.

Proof. Consider the $\mathbb{R}^+$-value map $h_\rho^{\psi_s}$ that maps (ρ, s) to $h_\rho^{\psi_s}$. For a fixed s, the aforementioned result ([PS17, Cor. 4.9] and [BCLS15]) implies that $h_\rho^{\psi_s}$ is an analytic function on $\rho \in \mathrm{HA}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$. For a fixed ρ, $h_\rho^{\psi_s}$ is a strictly decreasing continuous function (Lemma 10.7). A basic analysis lemma says that a two-variable function which is continuous on ρ and strictly monotonic continuous on s satisfies implicit function theorem. We obtain that $s_A(\rho)$, the solution of $h_\rho^{\psi_s} = 1$ (Lemma 10.8), is a continuous function on $\mathrm{HA}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$. □

We prove Theorem 1.1 in a more general setting, that is starting from any convex cocompact representation $\rho_0 : \Gamma \to \mathrm{SL}_2(\mathbb{R})$. Recall that ι and $\rho_1 = \iota \circ \rho_0$ are defined in the introduction.

Theorem 10.10. *For every* $\epsilon > 0$, *there exists a small neighborhood* O *of* ρ_1 *in* $\mathrm{Hom}(\Gamma, \mathrm{SL}_3(\mathbb{R}))$ *such that for any* ρ *in* O *we have*

- *either* $\rho(\Gamma)$ *acts reducibly on* $\mathbb{R}^3$, *i.e. fixing a line or a plane in* $\mathbb{R}^3$;
- *or* $\rho(\Gamma)$ *is irreducible and*

$$|\dim L(\rho(\Gamma)) - (\dim L(\rho_1(\Gamma)) + \min\{\dim L(\rho_1(\Gamma)), 1/2\})| \leqslant \epsilon.$$

If we start with a surface group Γ, then the limit set $L(\rho_1(\Gamma))$ is the full circle and of dimension one. Then we obtain $3/2$ in Theorem 1.1.

Proof. By the definition of the embedding ι, for any $\gamma \in \Gamma$ we have $\sigma_1(\rho_1(\gamma)) = 1/\sigma_3(\rho_1(\gamma)) = \mu_1(\rho_0(\gamma)) = 1/\mu_2(\rho_0(\gamma))$ and $\sigma_2(\rho_1(\gamma)) = 1$ (In order to distinguish the notations, we use $\mu_1(\rho_0(\gamma)) \geqslant \mu_2(\rho_0(\gamma))$ to denote the singular values of $\rho_0(\gamma) \in \mathrm{SL}(2, \mathbb{R})$). Then

$$\sum_{\gamma\in\Gamma} \left(\frac{\mu_2(\rho_0\gamma)}{\mu_1(\rho_0\gamma)}\right)^s = \sum_{\gamma\in\Gamma} \left(\frac{\sigma_2(\rho_1\gamma)}{\sigma_1(\rho_1\gamma)}\right)^{2s} = \sum_{\gamma\in\Gamma} \frac{\sigma_2(\rho_1\gamma)}{\sigma_1(\rho_1\gamma)} \left(\frac{\sigma_3(\rho_1\gamma)}{\sigma_1(\rho_1\gamma)}\right)^{s-1/2}. \tag{10.3}$$

The map ι induces an isometric embedding of the limit set $L(\rho_0(\Gamma))$ to $L(\rho_1(\Gamma))$, in particular, they have the same Hausdorff dimension. A classic result in the theory of Fuchsian groups tells us that $\dim L(\rho_0(\Gamma))$ is the critical exponent of first series in (10.3). Hence

$$\dim L(\rho_1(\Gamma)) = \max\{s_A(\rho_1)/2, s_A(\rho_1) - 1/2\}.$$

Therefore

$$s_A(\rho_1) = \min\{2 \dim L(\rho_1(\Gamma)), \dim L(\rho_1(\Gamma)) + 1/2\}. \tag{10.4}$$

Due to Proposition 10.9, for any irreducible ρ in the neighborhood O of ρ_1, by Theorem 1.3. and Eq. (10.4) we have

$$\dim L(\rho(\Gamma)) = s_A(\rho) \geqslant \min\{2 \dim L(\rho_1(\Gamma)), \dim L(\rho_1(\Gamma)) + 1/2\} - \epsilon.$$ □

11 Hausdorff dimension of the Rauzy Gasket

11.1 Preliminaries and Notation

Let $\tilde{\Delta} = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 1, x, y, z \geqslant 0\}$ and Δ be the projectization of $\tilde{\Delta}$ on $\mathbb{P}(\mathbb{R}^3)$. The euclidean distance d_E of $\tilde{\Delta}$ from $\mathbb{R}^3$ is bi-Lipschitz equivalent to the projective distance d coming from Δ. Since Lipschitz constant not affect the statements of lemmas, we do not distinguish the euclidean metric and the projective metric. The points on Δ will be naturally regarded as points on $\tilde{\Delta}$. The area of a subset of Δ will be understood as the area of the corresponding subset of $\tilde{\Delta}$.

Recall that the Rauzy gasket $X \subset \mathbb{P}(\mathbb{R}^3)$ is a projective fractal set defined in the introduction. We may consider the classical coding of X by infinite words as the case of IFSs. Let $\Lambda := \{1, 2, 3\}$ be the set of symbols. We have the following basic fact [AS13, Lemma 3].

Lemma 11.1. *For every $\mathbf{i} = (i_0, i_1, \cdots) \in \Lambda^{\mathbb{N}}$, we have $\lim_{n\to\infty} \operatorname{diam} A_{i_0} \cdots A_{i_{n-1}} \Delta = 0$.*

This fact allows us to define the coding map

$$\Phi : \Lambda^{\mathbb{N}} \to \Delta, \quad \mathbf{i} = (i_0, i_1, \cdots) \mapsto \cap_{n\in\mathbb{N}} A_{i_0} \cdots A_{i_{n-1}} \Delta.$$

Then the image of Φ is exactly the Rauzy Gasket X.

For any $\gamma \in \Gamma$ we denote by Δ_γ the image of $\gamma \cdot \Delta$. We also use $|\gamma|$ for the word length of γ with respect to the standard generator set $\{A_i\}$ of Γ. For later use we consider the following notations. Recall that by freeness of Γ, any $\gamma \in \Gamma$ can be decomposed uniquely as the following $\gamma = A_{i_1} \cdots A_{i_{|\gamma|}}$. For any $n \leqslant |\gamma|$, we denote by $\gamma_n := A_{i_1} \cdots A_{i_n}$. We say that the last n digits of γ are not the same for some $n \leqslant |\gamma|$, if in the decomposition, $i_{|\gamma|-n+1}, \dots, i_{|\gamma|}$ are not the same. Recall $s_A(\Gamma)$ defined in the introduction is the affinity dimension of Γ.

We will also consider the transpose action of Γ. For any $\gamma = A_{i_1} \cdots A_{i_{|\gamma|}} \in \Gamma$, the transpose action $\gamma^t : \mathbb{P}(\mathbb{R}^3) \to \mathbb{P}(\mathbb{R}^3)$ of γ is defined by $\gamma^t := A^t_{i_{|\gamma|}} \cdots A^t_{i_1}$, where the transposes of A_i's are

$$A_1^t = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \ A_2^t = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \ A_3^t = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

It is not hard to check that for any $\gamma \in \Gamma$, the transpose action γ^t preserves Δ since the entries of the matrix presentation of γ^t are all non-negative. For any $n \leqslant |\gamma|$, we denote by $\gamma^t_n := A^t_{i_{|\gamma|}} \cdots A^t_{i_{|\gamma|-n+1}}$. Notice that $\gamma^t_n \neq (\gamma_n)^t$, it should be identified as $(\gamma^t)_n$.

11.2 The upper bound of the Hausdorff dimension

The goal of this section is to show the upper bound of the Hausdorff dimension of the Rauzy Gasket, that is $\dim_{\mathrm{H}}(X) \leqslant s_A(\Gamma)$. The following elementary lemma in linear algebra is the key observation, which plays a crucial role on estimating the upper bound of $\dim_{\mathrm{H}}(X)$.

Lemma 11.2. *For every $n \in \mathbb{N}$, there exists $\epsilon_n > 0$ such that for any $\gamma \in \Gamma$, if the last n digits of γ are not the same, then for $i = 1, 2, 3$*

$$\|\gamma e_i\| \geqslant \epsilon_n \sigma_1(\gamma).$$

Proof. The idea to show the lemma is to consider the transpose action of Γ. We list some basic facts on the action of A_i^t on Δ.

Lemma 11.3. *For every $i \neq j \in \{1, 2, 3\}$, the following hold:*

(1) A_i^t preserves Δ.

*(2) A_i^t preserves Δ', where Δ' is the **open** projective triangle in $\mathbb{P}(\mathbb{R}^3)$ with vertices* $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

(3) $A_i^t E_j = E_j$.

(4) $A_i^t E_i = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \Delta'$.

Combining (2)(3)(4) in the lemma above, we obtain

Lemma 11.4. *For any $i \in \{1, 2, 3\}$, for any $\gamma = A_{i_1} \cdots A_{i_{|\gamma|}} \in \Gamma$, if there exists some $i_j(\gamma) = i$, then $\gamma^t E_i \in \Delta'$.*

We back to the proof of Lemma 11.2. Let $\gamma \in \Gamma$ be an element such that the last n digits of γ are not the same. Due to Lemma 2.5, we have

$$\|\gamma e_i\| \geqslant \|\gamma\| d(E_i, H_{\gamma^-}). \tag{11.1}$$

Recall the relation $(V_{(\gamma^t)^+})^\perp = H_{\gamma^-}$. So it is sufficient to show the angle between e_i and $V_{(\gamma^t)^+}$ is bounded away from $\pi/2$. Note that Δ has a special geometry property. That is, $e_i^\perp$ is the span $\{e_j : j \neq i\}$, which corresponds to an edge of Δ. In order to show the angle between $V_{(\gamma^t)^+}$ and e_i is bounded away from $\pi/2$, it suffices to show that $d(V_{(\gamma^t)^+}, \partial\Delta)$ is lower bounded by a positive constant only depending on n. To determine the position of $V_{(\gamma^t)^+}$, we use the fact that it is the attracting fixed point of $\gamma^t\gamma$ on the projective plane. Hence $V_{(\gamma^t)^+} \in \gamma^t\gamma\Delta$.

Lemma 11.5. $\gamma^t\gamma\Delta \subset \gamma_n^t\Delta \cap \overline{\Delta'}$.

Proof. Since A_i and A_i^t preserve Δ for $i = 1, 2, 3$, we obtain $\gamma^t\gamma\Delta \subset \gamma_n^t\Delta$. Now we show $\gamma^t\gamma\Delta \subset \overline{\Delta'}$. There are two possible cases. If all of A_1, A_2, A_3 occur in γ, then $\gamma^t E_i \in \Delta'$ for each $i = 1, 2, 3$ by Lemma 11.4. Therefore $\gamma^t\gamma\Delta \subset \gamma^t\Delta \subset \Delta'$.

Otherwise, there are at most two of A_i occur in the γ. Recall the assumption that the last n digits of γ are not the same. Without loss of generality, we can assume that both A_1 and A_2 occur in the last n digits of γ. Now we consider the region

$$\nabla_z := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \Delta : z \leqslant x + y \right\}.$$

Then A_1, A_2, A_1^t, A_2^t preserve ∇_z. Moreover, $A_1\Delta \subset \nabla_z$ and $A_2\Delta \subset \nabla_z$. Since A_3 does not occur in γ, we have $\gamma^t\gamma\Delta \subset \nabla_z$. Finally, we notice that the vertices of $\gamma_n^t\Delta$ satisfy $\gamma_n^t E_1, \gamma_n^t E_2 \in \Delta'$ and $\gamma_n^t E_3 = E_3$ by Lemma 11.4. This gives $\gamma^t\gamma\Delta \subset \gamma_n^t\Delta \cap \nabla_z = \gamma_n^t\Delta \cap \overline{\Delta'}$. $\square$

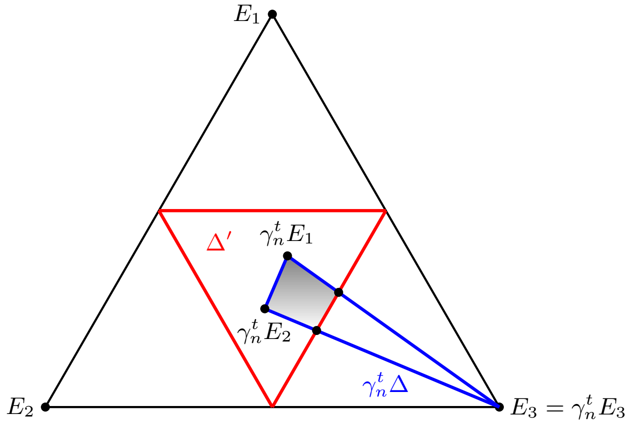

Figure 7: The pattern of $\gamma_n^t \Delta \cap \Delta'$.

Now we complete the proof of Lemma 11.2. Write $\gamma = A_{i_1} \cdots A_{i_{|\gamma|}}$. By an inductive argument on n, we can assume that there are exactly two of $\{1, 2, 3\}$ occur in $i_{|\gamma|-n+1}, \cdots i_{|\gamma|}$. Without loss of generality, we assume these two digits are 1 and 2. By Lemma 11.4, the vertices of $\gamma_n^t \Delta$ satisfy

$$\gamma_n^t E_1 \in \Delta', \quad \gamma_n^t E_2 \in \Delta', \quad \gamma_n^t E_3 = E_3.$$

Notice that set $\gamma_n^t \Delta \cap \overline{\Delta'}$ is a closed quadrilateral (see Figure 7), which does not intersect the boundary of Δ. Thus for every such γ_n, we have $d(\partial \Delta, \gamma_n^t \Delta \cap \overline{\Delta'}) > 0$. Since there are only finitely many such γ_n for a given positive integer n, there exists $d_n > 0$ only depending on n such that

$$d(\partial \Delta, \gamma_n^t \Delta \cap \overline{\Delta'}) > d_n$$

for all such γ_n^t. Recalling $V_{(\gamma^t)^+} \in \gamma^t \gamma \Delta \subset \gamma_n^t \Delta \cap \Delta'$, we have $d(V_{(\gamma^t)^+}, \partial \Delta) > d_n$. $\square$

The following lemma shows some basic estimates of the diameter and the area of Δ_γ.

Lemma 11.6. *There exists $C_2 > 1$ such that if the last 2 digits of γ are not the same, then*

(1) $\operatorname{diam}(\Delta_\gamma) \leqslant C_2 \cdot \frac{\sigma_2(\gamma)}{\sigma_1(\gamma)}$*;*

(2) $\operatorname{Area}(\Delta_\gamma) \leqslant C_2 \cdot \sigma_1(\gamma)^{-3}$.

Proof. (1) It suffices to show $d_E(\gamma E_i, \gamma E_j) \ll \frac{\sigma_2(\gamma)}{\sigma_1(\gamma)}$. Due to bi-Lipschitz property, we only need to show this for projective distance. Then

$$\begin{aligned} d(\gamma E_i, \gamma E_j) &= \frac{\|\gamma e_i \wedge \gamma e_j\|}{\|\gamma e_i\| \cdot \|\gamma e_j\|} \leqslant \frac{\sigma_1(\gamma)\sigma_2(\gamma)\|e_i \wedge e_j\|}{\|\gamma e_i\| \cdot \|\gamma e_j\|} \\ &\leqslant \frac{\sigma_1(\gamma)\sigma_2(\gamma)}{\epsilon_2^2 \sigma_1(\gamma)^2} \text{ (by Lemma 11.2)} \\ &= \epsilon_2^{-2} \cdot \frac{\sigma_2(\gamma)}{\sigma_1(\gamma)} \end{aligned}$$

(2) We use the following elementary geometric fact: Let x, y, z be three points in $\mathbb{R}^3 - \{o\}$, then the area of the triangle formed by x, y, z (which we denote by (xyz)) is equal to

$$\operatorname{Area}(xyz) = \frac{\|x \wedge y \wedge z\|}{2 d_E(o, (xyz))},$$

where $d_E(o, (xyz))$ is the distance from the origin o to the two plane of (xyz). This is because the numerator gives the volume of the polyhedron of $oxyz$.

For $\gamma \in \Gamma$, recall that the area of Δ_γ is understood in the area of the corresponding subset of $\tilde{\Delta}$ in the euclidean space. Let $x_\gamma, y_\gamma, z_\gamma$ be the corresponding points in $\tilde{\Delta}$ of vertices of Δ_γ. Then

$$\mathrm{Area}(x_\gamma y_\gamma z_\gamma) = \frac{\|x_\gamma \wedge y_\gamma \wedge z_\gamma\|}{2d(o,(x_\gamma y_\gamma z_\gamma))} = \frac{\|x_\gamma \wedge y_\gamma \wedge z_\gamma\|}{2d(o,\tilde{\Delta})}.$$

We know that

$$x_\gamma = \gamma e_1/\varphi(\gamma e_1),$$

with $\varphi(v) = v_1 + v_2 + v_3$, similarily for y_γ, z_γ. Therefore,

$$\|x_\gamma \wedge y_\gamma \wedge z_\gamma\| = \|\gamma e_1 \wedge \gamma e_2 \wedge \gamma e_3\| / \prod_{1 \leqslant i \leqslant 3} \varphi(\gamma e_i)$$

We actually have $\|\gamma e_1 \wedge \gamma e_2 \wedge \gamma e_3\| = \|e_1 \wedge e_2 \wedge e_3\| = 1$. Due to $\gamma E_i \in \Delta$, we also have

$$\varphi(\gamma e_i) \geqslant \|\gamma e_i\|$$

Then by Lemma 11.2 we get the proof of (2). $\square$

The following geometric lemma is essentially proved in Lemma 4.1 of [PS21].

Lemma 11.7. *For any $\delta > 0$, there exists $c_\delta > 0$ such that for any $\gamma \in \Gamma$, there exists a finite open cover $\{D_i(\gamma) : i = 1, \cdots, k\}$ of Δ_γ with $\operatorname{diam} D_i(\gamma) \leqslant \operatorname{diam} \Delta_\gamma$ such that*

$$\sum_i \operatorname{diam}^{1+\delta} D_i(\gamma) \leqslant c_\delta \cdot \operatorname{diam}^{1-\delta} \Delta_\gamma \cdot \mathrm{Area}^\delta \Delta_\gamma.$$

Proof. As in the proof of Lemma 4.1 of [PS21], we know that every Δ_γ can be covered by $O(\operatorname{diam}^2(\Delta_\gamma)/\mathrm{Area}(\Delta_\gamma))$ disks $\{D_i : i = 1, \cdots, k\}$ of diameter $O(\mathrm{Area}(\Delta_\gamma)/\operatorname{diam}(\Delta_\gamma))$, then we get the proof. $\square$

Now we back to the proof of $\dim_{\mathrm{H}}(X) \leqslant s_A(\Gamma)$. Recall $P_\Gamma(s) = \sum_\gamma \varphi_s(\gamma)$ is the Poincaré series of Γ, where $\varphi_s(\gamma)$ is defined by

$$\varphi_s(\gamma) = \begin{cases} \left(\frac{\sigma_2}{\sigma_1}\right)^s(\gamma), & 0 < s \leqslant 1; \\ \left(\frac{\sigma_2}{\sigma_1}\right)(\gamma)\left(\frac{\sigma_3}{\sigma_1}\right)^{s-1}(\gamma), & 1 < s \leqslant 2. \end{cases}$$

By definition of s_A, it suffices to show if $P_\Gamma(s) < \infty$ then for any $\epsilon > 0$, $\dim_{\mathrm{H}}(X) \leqslant s + \epsilon$.

Definition 11.8. For $x \in X$, we say x is *nice* if every $(i_0(x), i_1(x), \cdots) \in \psi^{-1}(x) \in \Lambda^{\mathbb{N}}$ is not ending by a single element in $\Lambda = \{1, 2, 3\}$.

We remark that if x is not uniquely coding, then x is not nice. This is because the only possibility of $\Phi(\mathbf{i}) = \Phi(\mathbf{i}')$ is

$$\mathbf{i} = (w, j_1, j_2, j_2, \cdots) \text{ and } \mathbf{i}' = (w, j_2, j_1, j_1, \cdots),$$

where $w \in \Lambda^N$ for some $N \in \mathbb{N}$ and $j_1 \neq j_2 \in \Lambda$.

Then the Rauzy Gasket X can be decomposed as the set of nice points which we denote it by X_{nice} and a countable set. So we only need to show the Hausdorff outer measure $H^{s+\epsilon}(X_{\text{nice}})$ is 0.

We consider

$$\Gamma_m := \{\gamma \in \Gamma : \text{the last 2 digits of } \gamma \text{ are not the same and } \operatorname{diam} \Delta_\gamma \leqslant 1/m\}.$$

Now we construct two families of covers $\mathcal{U}_m$ and $\mathcal{U}'_m$ for $m \in \mathbb{N}$ as

$$\mathcal{U}_m := \{D_i(\gamma) : \gamma \in \Gamma_m\}, \quad \mathcal{U}'_m := \{\Delta_\gamma : \gamma \in \Gamma_m\},$$

where $\{D_i(\gamma)\}$ is the finite open cover of Δ_γ we obtained from Lemma 11.7. Then the sequence of covers $\mathcal{U}_m$ is a family of Vitali covers of the set $Y := \cap_{m=1}^\infty (\cup_{U \in \mathcal{U}_m} U)$. [10] Notice that Y contains $Y' := \cap_{m=1}^\infty (\cup_{U \in \mathcal{U}'_m} U)$, where Y' is just formed by the points which can be covered by infinitely many Δ_γ such that the last 2 digits of γ are not the same.

Lemma 11.9. $Y \supset X_{\text{nice}}$.

Proof. We claim that for any $\mathbf{i} \in \{1,2,3\}^{\mathbb{N}}$ which is not ending by a single element, the point $x = \Phi(\mathbf{i})$ is contained in Y. Since $\mathbf{i} = (i_0(x), \cdots)$ is not ending by a single element in $\{1,2,3\}$, there exists infinitely many ℓ such that $i_{\ell-1}(x), i_\ell(x)$ are not the same (otherwise it will be ended by only one element). Collect all such ℓ and consider all the elements $\gamma = A_{i_1} \cdots A_{i_\ell}$. Then we get there are infinitely many $\gamma \in \Gamma$ such that the last two digits of γ are not the same and $x \in \Delta_\gamma$. By Lemma 11.1, $\operatorname{diam} \Delta_\gamma \to 0$ as ℓ tending to infinity. Therefore $x \in Y' \subset Y$. $\square$

As a consequence, it suffices to show the Hausdorff outer measure

$$H^{s+\epsilon}(Y) = \lim_{\delta \to 0} H_\delta^{s+\epsilon}(Y) = 0.$$

Recall that $\operatorname{diam} U \leqslant 1/m$ for every $U \in \mathcal{U}_m$. Then for $s \geqslant 1$, we have

$$\begin{aligned}
H^{s+\epsilon}(Y) &\leqslant \limsup_{\delta \to 0} H_\delta^{s+\epsilon}(Y) \leqslant \limsup_{m \to \infty} \sum_{U \in \mathcal{U}_m} \operatorname{diam}(U)^{s+\epsilon} \\
&= \limsup_{m \to \infty} \sum_{\gamma \in \Gamma_m} \sum_i \operatorname{diam}(D_i(\gamma))^{s+\epsilon} \\
&\leqslant \limsup_{m \to \infty} c_{s+\epsilon-1} \sum_{\gamma \in \Gamma_m} \operatorname{diam}^{2-s-\epsilon} \Delta_\gamma \cdot \operatorname{Area}^{s+\epsilon-1} \Delta_\gamma \qquad \text{(by Lemma 11.7)} \\
&\leqslant \limsup_{m \to \infty} c_{s+\epsilon-1} C_2 \sum_{\gamma \in \Gamma_m} \varphi_{s+\epsilon}(\gamma) \qquad \text{(by Lemma 11.6 and the definition of } \varphi_s).
\end{aligned}$$

Since $\sum \varphi_{s+\epsilon}(\gamma) < \infty$, the right hand side of the last inequality is just 0 when $m \to \infty$, which is exactly what we want. For the case $s < 1$ (we actually know $s \sim 1.72$ by numerical test), we do not need to use Lemma 11.7 and just consider Y' instead of Y by the same argument we still get the proof.

11.3 The lower bound of the Hausdorff dimension: from the dimension formula

In this section, we will show that $\dim_{\mathrm{H}} X \geqslant s_A(\Gamma)$. Note that X contains the boundary of Δ, which has the Hausdorff dimension equal to 1. Therefore we have an apriori estimate that $\dim_{\mathrm{H}} X \geqslant 1$. This allows us to assume without loss of generality that $s_A(\Gamma) > 1$. Equivalently, there exists $s > 1$ such that the series $\sum_\gamma \varphi_s(\gamma)$ diverges. Recall the fact that for every exact dimensional probability measure μ supported on X, we have $\dim \mu \leqslant \dim_{\mathrm{H}} X$. The inequality $\dim_{\mathrm{H}} X \geqslant s_A(\Gamma)$ is a direct consequence of the following lemma.

Lemma 11.10. *Let $s > 1$ such that the series $\sum_\gamma \varphi_s(\gamma)$ diverges, then for every $\epsilon > 0$, there exists a finitely supported measure ν supported on Γ and a ν-stationary measure μ supported on X satisfying* $\dim \nu \geqslant s - \epsilon$.

[10] Recall that a Vitali cover $\mathcal{V}$ of a set E is a family of set so that, for every $\delta > 0$ and every $x \in E$, there is some U in the family $\mathcal{V}$ with $\operatorname{diam} U < \delta$ and $x \in U$.

Our idea is a stopping time argument which is partially inspired by the study of the variational principle of iterated functional systems [FH09] and self-affine measures [MS23]. Specifically, our strategy is to combine our dimension formula of stationary measures with a modification of the proof of Theorem 3.1 in [MS23]. The main difference of our proof from [MS23] is that due to the presences of unipotent elements, the word lengths of elements may loss control. So instead of considering the word length, we only consider $\kappa(\gamma)$. This is enough to estimate the Lyapunov exponents. Then we use a combinatorial argument to make elements generating a free semigroup. This generates a large enough random walk entropy. Therefore, we can find a stationary measure with a large dimension approximating the affinity dimension. A similar argument to address the issue of words with uncontrolled length also appears in [HJX23].

In order to apply the dimension formula (Theorem 1.10), we need to show the Zariski density of Γ at first.

Lemma 11.11. *The semigroup Γ is Zariski dense in $\mathrm{SL}_3(\mathbb{R})$.*

Proof. Let $\mathbf{H}$ be the Zariski closure of semigroup Γ in $\mathrm{SL}_3(\mathbb{R})$, which is a real algebraic group. Let $\mathfrak{h}$ be the Lie algebra of the connected component of $\mathbf{H}$. By [Bor91, Page 106, Remark], the fact that A_1 is unipotent implies that the one-parameter unipotent group $\left\{s\in\mathbb{R},\begin{pmatrix}1&s&s\\0&1&0\\0&0&1\end{pmatrix}\right\}$ is in $\mathbf{H}$ and $X_1=\begin{pmatrix}0&1&1\\0&0&0\\0&0&0\end{pmatrix}$ is in $\mathfrak{h}$. Similarly, the nilpotent elements $X_2=\begin{pmatrix}0&0&0\\1&0&1\\0&0&0\end{pmatrix}$, $X_3=\begin{pmatrix}0&0&0\\0&0&0\\1&1&0\end{pmatrix}$ are also in $\mathfrak{h}$. Then we can play with these elements in the Lie algebra $\mathfrak{h}$ and prove that they generate all $\mathfrak{sl}_3$. We have $Y_3=[X_1,X_2]=\begin{pmatrix}1&0&1\\0&-1&-1\\0&0&0\end{pmatrix}$, $Y_1=[X_2,X_3]=\begin{pmatrix}0&0&0\\1&1&0\\-1&0&-1\end{pmatrix}$, $Y_2=[X_3,X_1]=\begin{pmatrix}-1&-1&0\\0&0&0\\0&1&1\end{pmatrix}$ also in $\mathfrak{h}$. Then $[Y_1,X_1]+Y_2+Y_3+2X_1=\begin{pmatrix}0&0&4\\0&0&0\\0&0&0\end{pmatrix}$, $[Y_2,X_2]+Y_3+Y_1+2X_2=\begin{pmatrix}0&0&0\\4&0&0\\0&0&0\end{pmatrix}$ also in $\mathfrak{h}$. From these, we can obtain the whole $\mathfrak{sl}(3,\mathbb{R})$ and then $\mathbf{H}$ must be the whole group $\mathrm{SL}_3(\mathbb{R})$. $\square$

Recall that for $s>1$, the linear functional ψ_s on $\mathfrak{a}$ is given by

$$\psi_s(\kappa(g))=(\log\sigma_1(g)-\log\sigma_2(g))+(s-1)(\log\sigma_1(g)-\log\sigma_3(g)),\quad \forall g\in\mathrm{SL}_3(\mathbb{R}).$$

Now we construct a good set coming from the divergent series $\sum_\gamma\varphi_s(\gamma)=\sum_\gamma q^{-\psi_s(\kappa(\gamma))}$.

Definition 11.12. For $\beta>0,x\in\mathfrak{a}^+$ and $n\in\mathbb{N}$, a subset $S\subset\Gamma$ is called (n,β,x)*-approximate* if $\#S\geqslant q^{(1-\beta)n}$ and for every $\gamma\in S$, $\|\frac{1}{n}\kappa(\gamma)-x\|\leqslant\beta$.

Lemma 11.13. *For any $\beta>0$, there exists $x\in\mathfrak{a}^+$ satisfying $|\psi_s(x)-1|\leqslant\beta$ and infinitely many $n\in\mathbb{N}$ such that there exists a (n,β,x)-approximate subset $S\subset\Gamma$.*

Proof. Note that $\psi_s(\kappa(\gamma))\geqslant(s-1)\log\sigma_1(\gamma)=(s-1)\log\|\gamma\|$. Hence for every $t>0$, the set $\{\gamma\in\Gamma:\psi_s(\kappa(\gamma))<t\}$ is finite by the discreteness of Γ. Combining with the hypothesis that the series $\sum_\gamma q^{-\psi_s(\kappa(\gamma))}$ diverges, there are infinitely many n such that

$$\{\gamma:\psi_s(\kappa(\gamma))\in[(1-\beta/10)n,(1+\beta/10)n]\}$$

contains at least $q^{(1-\beta/10)n}$ elements.

Since we have $\psi_s(\kappa(\gamma)) \geqslant (s-1)\log\sigma_1(\gamma)$ and $\psi_s(\kappa(\gamma)) \geqslant -(s-1)\log\sigma_3(\gamma)$, the set $\psi_s^{-1}([1-\beta/10, 1+\beta/10]) \cap \mathfrak{a}^+$ is compact. Therefore we can cover it by finitely many balls of radius $\beta/100$. By a pigeonhole principle, there exists a center of some ball, say $x \in \mathfrak{a}^+$, such that for infinitely many n,

$$\left\{\gamma : |\frac{\psi_s(\kappa(\gamma))}{n} - 1| \leqslant \frac{\beta}{10},\ \|\frac{\kappa(\gamma)}{n} - x\| \leqslant \frac{\beta}{100}\right\}$$

contains at least $q^{(1-\beta)n}$ elements. Using the fact that $|\psi_s(\kappa(\gamma))| \leqslant 10\|\kappa(\gamma)\|$, we obtain $|\psi_s(x) - 1| \leqslant \beta$, which is exactly what we want. □

We will modify an (n,β,x)-approximate subset to a set which freely generates a free semigroup. The idea is to avoid the "prefix relations". We consider the following concepts.

Definition 11.14. (1) An element $j_1 \in \Gamma$ is called *starting with* $j_2 \in \Gamma$ if there is $j_3 \in \Gamma - \{\mathrm{id}\}$ such that $j_1 = j_2 j_3$. We also say j_2 is a *prefix* of j_1 if j_1 is starting with j_2.

(2) An element $j_1 \in \Gamma$ is called *ending with* $j_2 \in \Gamma - \{\mathrm{id}\}$ if there is $j_3 \in \Gamma$ such that $j_1 = j_3 j_2$.

(3) An element j is called *minimal* in a subset S of Γ if there is no element $j' \in S$ such that j is starting with j'.

Within a set S, a minimal element of S is never a prefix of another minimal element. So the set of minimal elements $S_{\min}$ of S will freely generate a free semigroup. Moreover, the subset $S_{\min}^{*l} \subset \Gamma$ satisfies that there is no pair of elements such that one is a prefix of the other. In the following lemma, using hyperbolicity and discreteness we get a lower bound of minimal elements in a set with approximately the same sizes of Cartan projections.

Lemma 11.15. *There exists $C \geqslant 1$ such that the following hold. For every $\beta > 0$, $n \in \mathbb{N}$ and $x \in \mathfrak{a}^+$. For any element j in the set*

$$\mathcal{J} = \{\gamma \in \Gamma :\ \|\frac{\kappa(\gamma)}{n} - x\| \leqslant \beta,\quad \gamma \text{ is ending with } A_1A_2A_3\},$$

there are at most $q^{C\beta n}$ elements in $\mathcal{J}$ which is starting with j.

Proof. For every unit vector $v \in V \in \Delta$, since all the entries of γe_i are non-negative, we have

$$\|\gamma v\| \geqslant \max\{v_i\|\gamma e_i\|\} \geqslant \frac{1}{2}\min\{\|\gamma e_i\|\}.$$

Take $n = 2$ in Lemma 11.2, for any element γ with last two digits different, we have $\|\gamma e_i\| \geqslant \epsilon_2\sigma_1(\gamma)$, where ϵ_2 is defined in Lemma 11.2. Hence

$$\|\gamma v\| \geqslant \frac{\epsilon_2}{2}\sigma_1(\gamma).$$

Then for any $j, jj'' \in \mathcal{J}$ with $j'' \in \Gamma - \{\mathrm{id}\}$ and any unit $v \in V \in \Delta$, since $j''V \in \Delta$ and j, j'' end with $A_1A_2A_3$, we obtain

$$\sigma_1(jj'') \geqslant \|jj''v\| = \frac{\|j(j''v)\|}{\|j''v\|}\frac{\|j''v\|}{\|v\|} \geqslant \frac{\epsilon_2^2}{4}\sigma_1(j)\sigma_1(j''). \tag{11.2}$$

Since $\|\frac{1}{n}\kappa(jj'') - x\| \leqslant \beta$ and $\|\frac{1}{n}\kappa(j) - x\| \leqslant \beta$, we have

$$\log\sigma_1(jj'') \leqslant 2\beta n + \log\sigma_1(j). \tag{11.3}$$

Combining (11.2),(11.3) we get

$$\sigma_1(j'') \leqslant 4q^{2\beta n}/\epsilon_2^2.$$

Since the semigroup Γ is discrete in $\mathrm{SL}_3(\mathbb{R})$, up to a constant the number

$$\#\{\gamma \in \Gamma,\ \|\gamma\| \leqslant t\}$$

is bounded by the volume of the set $\{g \in \mathrm{SL}_3(\mathbb{R}),\ \|g\| \leqslant t\}$, which grows at most polynomially on t. Then the possible number of j'' is bounded by $q^{C\beta n}$ for some constant $C > 0$. $\square$

In order to estimate the Lyapunov exponents of the constructing measure, we need to estimate the Cartan projection of products. Let us recall some notions.

Definition 11.16. (1) For an element $g \in \mathrm{SL}_3(\mathbb{R})$, we call it (r,ϵ)-*loxodromic* for $r, \epsilon > 0$, if $\sigma_i(g)/\sigma_{i+1}(g) \geqslant 1/\epsilon$ and

$$d(V_g^+, H_g^-) > r,\ d(V^+_{g,\wedge^2\mathbb{R}^3}, H^-_{g,\wedge^2\mathbb{R}^3}) > r,$$

where $\wedge^2\mathbb{R}^3$ is the wedge representation of $\mathrm{SL}_3(\mathbb{R})$ and $V^+_{g,\wedge^2\mathbb{R}^3} = k_g(E_1 \wedge E_2)$ and $H^-_{g,\wedge^2\mathbb{R}^3} = l_g^{-1}(E_2 \wedge E_3 \oplus E_3 \wedge E_1)$ are the corresponding attracting point and repelling hyperplane in the projective space $\mathbb{P}(\wedge^2\mathbb{R}^3)$.

(2) Let F be a subset in $\mathrm{SL}_3(\mathbb{R})$. We call F a (r,ϵ)-*Schottky family* if every element in F is (r,ϵ)-loxodromic and for any pair (g,h) in F, we have

$$d(V_g^+, H_h^-) > 6r,\ d(V^+_{g,\wedge^2\mathbb{R}^3}, H^-_{h,\wedge^2\mathbb{R}^3}) > 6r.$$

(3) A set F of $\mathrm{SL}_3(\mathbb{R})$ is called η-*narrow* if for g in F, the attracting points V_g^+ (resp. $V^+_{g,\wedge^2\mathbb{R}^3}$) are within η-distance of one another and the repelling hyperplanes H_g^- (resp. $H^-_{g,\wedge^2\mathbb{R}^3}$) are within η Hausdorff distance of one another.

(4) A set F is η-narrow around h if for g in F, the attracting points V_g^+ (resp. $V^+_{g,\wedge^2\mathbb{R}^3}$) are within $\eta/2$-distance to V_h^+(resp. $V^+_{h,\wedge^2\mathbb{R}^3}$) and the repelling hyperplanes H_g^- (resp. $H^-_{g,\wedge^2\mathbb{R}^3}$) are within $\eta/2$ Hausdorff distance to H_h^- (resp. $H^-_{h,\wedge^2\mathbb{R}^3}$).

These definitions and properties are originally due to Benoist [Ben97]. We borrow them from [MS23, Corollary 2.16, 2.17]

Lemma 11.17. *(1) For $r > 4\epsilon > 0$, if F is a (r,ϵ)-Schottky family, then the semigroup generated by F is a $(r/2, 2\epsilon)$-Schottky family.*

(2) Let E be a η-narrow collection of (r,ϵ)-loxodromic elements with $r > 4\max\{\epsilon,\eta\}$. Then, E is a $(r/4,\epsilon)$-Schottky family.

(3) If F is a (r,ϵ)-Schottky family, then there exists $C_r > 0$ only depending on r such that for any $g_1, \cdots, g_l$ in F, we have

$$\|\kappa(g_1 \cdots g_l) - \sum_{1 \leqslant i \leqslant l} \kappa(g_i)\| \leqslant lC_r.$$

Now we state the main construction, which gives a good set to support a desired random walk.

Lemma 11.18. *For every $\beta > 0$. There exists $N \in \mathbb{N}$, $x \in \mathfrak{a}^+, \mathcal{J} \subset \Gamma$, such that*

(1) $|\psi_s(x) - 1| \leqslant \beta$.

(2) The semigroup generated by $\mathcal{J}$ is Zariski dense.

(3) For every $k \in \mathbb{N}$ and $j \in \mathcal{J}^{*k} := \{j_1 \cdots j_k : j_i \in \mathcal{J}\}$, $\|\frac{1}{kN}\kappa(j) - x\| \leqslant 10\beta$.

(4) The set $\mathcal{J}$ contains a subset $\mathcal{J}_1$ satisfies $\#\mathcal{J}_1 \geqslant q^{(1-10C\beta)N}$ and no element in $\mathcal{J}_1$ is a prefix of another one, where $C \geqslant 1$ is the absolute constant given by Lemma 11.15.

Proof. Fix an $x \in \mathfrak{a}^+$ be given by Lemma 11.13. Let S be a (n, β, x)-approximate subset for some sufficiently large n. Now we construct the set $\mathcal{J}$ by modifying S.

Step 1. For every $\gamma \in S$, we add $A_1A_2A_3$ at the end and denote the new set by W_2. Since $\|\kappa(\gamma A_1A_2A_3) - \kappa(\gamma)\| \leqslant \|\kappa(A_1A_2A_3)\|$, the set W_2 is $(n, 2\beta, x)$-approximate for n large enough.

Step 2. We apply a theorem of Abels-Margulis-Soifer [AMS95], see also [MS23, Theorem 3.2].

Theorem 11.19 (Abels-Margulis-Soifer)**.** *Let G be a Zariski-connected real reductive group and Γ be a Zariski-dense subsemigroup. Then there exists $0 < r = r(\Gamma)$ such that for all $0 < \epsilon \leqslant r$, there exists a finite subset $F = F(r, \epsilon, \Gamma) \subset \Gamma$ with the property that for every $g \in G$, there exists $f \in F$ such that fg is (r, ϵ)-loxodromic in G.*

We can fix $r_0, \epsilon_0 > 0$ sufficiently small comparing to β with $r_0 > 100\epsilon_0$. By Theorem 11.19, we could find a finite subset $F_1 = F(r_0, \epsilon_0, \Gamma)$ of Γ. Therefore for every element $\gamma \in W_2$, there exists $f \in F_1$ such that $f\gamma$ is (r_0, ϵ_0)-loxodromic. By the pigeonhole principle, we can find an $f \in F_1$ such that for at least $(\#F_1)^{-1}$ proportion of γ in W_2, the product $f\gamma$ is (r_0, ϵ_0)-loxodromic. Fix this $f \in F_1$ and let

$$W_3 = \{f\gamma : \gamma \in W_2, f\gamma \text{ is } (r_0, \epsilon_0)\text{-loxodromic }\}.$$

Then W_3 is is $(n, 3\beta, x)$-approximate assuming n large enough.

Step 3. By compactness, we can cover $\prod_{1 \leqslant i \leqslant 2} \mathbb{P}(V_i) \times \mathbb{P}(V_i^*)$ with $O(\epsilon_0^{-8})$ balls of radius ϵ_0, where $V_1 = \mathbb{R}^3$ and $V_2 = \wedge^2\mathbb{R}^3$. By the pigeonhole principle, there exists a subset $W_4 \subset W_3$, such that $\#W_4 \gg \epsilon_0^8 \cdot \#W_3$ and W_4 is an ϵ_0-narrow set of (r_0, ϵ_0)-loxodromic elements. For n sufficiently large comparing to ϵ_0, W_4 is $(n, 4\beta, x)$-approximate.

Step 4. Before making the next modification, we recall the following lemma from [Ben97] and [MS23].

Lemma 11.20 ([MS23, Lemma 3.4])**.** *There exists $r_1 > 0$ depending only on Γ such that the following hold. For every loxodromic element $g \in G$ and $0 < \epsilon < r_1$, there exists a Zariski dense (r_1, ϵ)-Schottky subgroup of Γ which is ϵ-narrow around g.*

Since r_1 is determined by Γ, we can assume at first that $r_0 < r_1$. Now we fix an element $g \in W_4$. Then we can find a Zariski dense (r_0, ϵ_0)-Schottky subgroup Γ_1 of Γ which is ϵ_0-narrow around g. By the proof of $k = 0$ in Lemma 9.9 (see also [MS23, Lemma 3.6]), we can find a finite subset $\{\theta_i : i = 1, \cdots, p\} \subset \Gamma_1$, which generates Zariski dense sub-semigroup. Let $W' = W_4 \cup \{\theta_i : i = 1, \cdots, p\}$, which consists of (r_0, ϵ_0)-loxodromic elements. Moreover, every element in W' is ϵ_0-narrow around g. Hence W' is $2\epsilon_0$-narrow. Therefore, W' is a $(r_0/4, \epsilon_0)$-Schottky family and the semigroup it generates is a $(r_0/8, 2\epsilon_0)$-Shottky family, by Lemma 11.17.

Take an $m \in \mathbb{N}$ large enough depending on θ_i, β and ϵ_0. Let

$$W_5 := W_4^{*m} \cup \{\theta_i g^m, \ i = 1, \cdots, p\}$$

Now we verify that the set $\mathcal{J} := W_5$ satisfies the condition for $N = nm$.

(1) This is because x is given by Lemma 11.13.

(2) Let $\mathbf{H}$ be the Zariski closure of the semigroup generated by W_5, which is an algebraic subgroup of $\mathrm{SL}_3(\mathbb{R})$. Note that $g^m \in W_4^{*m} \subset W_5$ and $\theta_i g^m \in W_5$ for every i. We obtain $\theta_i \in \mathbf{H}$. Since $\{\theta_i : i = 1, \cdots, p\}$ generates a Zariski dense subgroup, we have $\mathbf{H} = \mathrm{SL}_3(\mathbb{R})$.

(3) Recall that the semigroup generated by W' is a (r, ϵ)-Schottky family, where $r = r_0/8$ and $\epsilon = 2\epsilon_0$. Also recall that W_4 is $(n, 4\beta, x)$-approximate. By Lemma 11.17, for every $h \in W_4^{*m}$, we have

$$\|\frac{\kappa(h)}{N} - x\| = \|\frac{\kappa(h)}{nm} - x\| \leqslant \frac{C_r}{n} + 4\beta.$$

For $h \in \{\theta_i g^m, i = 1, \ldots, p\}$, we have

$$\|\frac{\kappa(h)}{N} - x\| = \|\frac{\kappa(h)}{nm} - x\| \leqslant \frac{C_r}{n} + 4\beta + \frac{\kappa(\theta_i)}{nm}.$$

By taking n large enough and then taking m large enough, we can find that the set $\mathcal{J} = W_5$ is $(N, 5\beta, x)$-approximate. Since the semigroup generated by W_5 is a (r, ϵ)- Schottky family, for every $j \in \mathcal{J}^{*k}$, we have

$$\|\frac{\kappa(j)}{kN} - x\| \leqslant \frac{C_r}{N} + 5\beta \leqslant 10\beta.$$

(4) We take minimal elements $(W_4)_{\min}$ in W_4. Let $\mathcal{J}_1 = ((W_4)_{\min})^{*m}$. Then there is no element in $\mathcal{J}_1$ which is a prefix of another element. Note that W_4 is $(n, 4\beta, x)$-approximate and every element in W_4 is ending with $A_1A_2A_3$. By Lemma 11.15, we have

$$\#\mathcal{J}_1 = (\#(W_4)_{\min})^m \geqslant (\#W_4/q^{4C\beta n})^m \geqslant q^{(1-4\beta-4C\beta)nm} \geqslant q^{(1-10C\beta)N}. \qquad \square$$

Finally, we will construct the random walk and estimate the dimension of the stationary measure. This part is to complete the proof of Lemma 11.10.

Let $\mathcal{J}, \mathcal{J}_1$ be the sets given by Lemma 11.18. We take $\nu = (1-\beta)\nu_1 + \beta\nu_2$, where ν_1 is the uniform measure on $\mathcal{J}_1$ and ν_2 is the uniform measure on $\mathcal{J} - \mathcal{J}_1$. Then by Lemma 11.18 the support of ν generates a Zariski dense subgroup in $\mathrm{SL}_3(\mathbb{R})$ and the associated Lyapunov vector is close to Nx. Hence, by Lemma 11.18(3),

$$|\frac{\psi_s(\lambda(\nu))}{N} - 1| \leqslant |\psi_s(x) - 1| + 5 \cdot \|\frac{\lambda(\nu)}{N} - x\| \leqslant 100\beta. \tag{11.4}$$

Now we should estimate the random walk entropy of ν. Firstly, note that $\mathcal{J}_1$ freely generates a free semigroup, we have

$$h_{\mathrm{RW}}(\nu_1) = H(\nu_1) \geqslant (1 - 10C\beta)N.$$

Moreover, since the support of ν_1 satisfies that no element is a prefix of another one, by certain "continuity" property of the random walk entropy and freeness of Γ, we have

Lemma 11.21. *Let ν, ν_1, ν_2 be probability measures which supported on Γ such that $\nu = (1-\beta)\nu_1 + \beta\nu_2$. If the support of ν_1 is a minimal set (i.e. no element is a prefix of another), then*

$$h_{\mathrm{RW}}(\nu) \geqslant (1-\beta)h_{\mathrm{RW}}(\nu_1). \tag{11.5}$$

Proof. By the concavity of the entropy function H,

$$H(\nu_1^{*l_1} * \nu_2^{*j_1} * \nu_1^{*l_2} * \cdots * \nu_2^{*j_k}) \geqslant \int H(\nu_1^{*l_1} * \delta_{g_1} * \nu_1^{*l_2} * \cdots * \delta_{g_k}) \mathrm{d}\nu_2^{*j_1}(g_1) \cdots \mathrm{d}\nu_2^{*j_k}(g_k).$$

We claim that $H(\nu_1^{*l_1} * \delta_{g_1} * \nu_1^{*l_2} * \cdots * \delta_{g_k}) = H(\nu_1^{*(l_1+\cdots+l_k)})$. It suffices to show that all the elements in the support of LHS are distinct. Otherwise suppose that two elements $(h_1, \cdots, h_k)$ and $(h_1', \cdots, h_k')$ in $\operatorname{supp}\nu_1^{*l_1} \times \cdots \times \operatorname{supp}\nu_1^{*l_k}$ satisfy $h_1g_1 \cdots h_kg_k = h_1'g_1 \cdots h_k'g_k$. Since no element in $\operatorname{supp}\nu_1$ is a prefix of another one, so does $\operatorname{supp}\nu_1^{*l_1}$. Therefore by freeness of Γ, we obtain that $h_1 = h_1'$. By induction we get $h_2 = h_2', \cdots, h_k = h_k'$. Therefore

$$H(\nu_1^{*l_1} * \nu_2^{*j_1} * \nu_1^{*l_2} * \cdots * \nu_2^{*j_k}) \geqslant \int H(\nu_1^{*(l_1+\cdots+l_k)})\mathrm{d}\nu_2^{*j_1}(g_1)\cdots d\nu_2^{*j_k}(g_k) = H(\nu_1^{*(l_1+\cdots+l_k)}).$$

Combining with the concavity of H, we get

$$\begin{aligned} h_{\mathrm{RW}}((1-\beta)\nu_1+\beta\nu_2) &= \lim_{l\to\infty}\frac{1}{l}H(((1-\beta)\nu_1+\beta\nu_2)^{*l}) \\ &\geqslant \lim\frac{1}{l}\sum_j\binom{l}{j}(1-\beta)^{l-j}\beta^j H(\nu_1^{*(l-j)}) \\ &= \lim\frac{1}{l}\sum_j\binom{l}{j}(1-\beta)^{l-j}\beta^j(l-j)h_{\mathrm{RW}}(\nu_1) \\ &\geqslant \lim(1-\beta)\sum_j\binom{l-1}{j}(1-\beta)^{l-j-1}\beta^j h_{\mathrm{RW}}(\nu_1) \\ &= (1-\beta)h_{\mathrm{RW}}(\nu_1). \end{aligned}$$

□

As a consequence $h_{\mathrm{RW}}(\nu) \geqslant (1-\beta)h_{\mathrm{RW}}(\nu_1) \geqslant (1-\beta)(1-10C\beta)N$. Moreover, the group generated by $\operatorname{supp}\nu$ is Zariski dense in $\mathrm{SL}_3(\mathbb{R})$ and satisfies exponential separation property (actually discreteness). Let μ be the unique ν-stationary measure. By the dimension formula, i.e. Theorem 1.10, we have $\dim\mu = \dim_{\mathrm{LY}}\mu$.

We admit the fact that $h_{\mathrm{F}}(\mu,\nu) = h_{\mathrm{RW}}(\nu)$ and postpone the proof to the end of this section. We give an estimate on $\dim_{\mathrm{LY}}\mu$. Since $\psi_s(\lambda(\nu)) \geqslant (1-100\beta)N$, we obtain

$$s(\lambda_1(\nu)-\lambda_3(\nu)) \geqslant \psi_s(\lambda(\nu)) \geqslant (1-100\beta)N.$$

Then $\lambda_1(\nu)-\lambda_3(\nu) \geqslant \frac{1}{2s}N$ assuming β small enough. For a given $0<\epsilon<s-1$, we have

$$\psi_{s-\epsilon}(\lambda(\nu)) = \psi_s(\lambda(\nu)) - \epsilon(\lambda_1(\nu)-\lambda_3(\nu)) \leqslant (1+100\beta)N - \frac{\epsilon}{2s}N.$$

Now we take $\beta>0$ sufficiently small comparing to ϵ, we obtain

$$\psi_{s-\epsilon}(\lambda(\nu)) \leqslant (1+100\beta)N - \frac{\epsilon}{2s}N \leqslant (1-\beta)(1-10C\beta)N \leqslant h_{\mathrm{RW}}(\nu).$$

Therefore, $\dim\mu = \dim_{\mathrm{LY}}\mu \geqslant s-\epsilon$.

To complete the proof, it remains to show the identity between the Furstenberg entropy and the random walk entropy.

Lemma 11.22. *Let ν be a finitely supported probability measure on Γ such that $\langle\operatorname{supp}\nu\rangle$ is Zariski dense. Let μ be the unique ν-stationary measure on $\mathbb{P}(\mathbb{R}^3)$, then $h_{\mathrm{F}}(\mu,\nu) = h_{\mathrm{RW}}(\nu)$.*

Proof. As discussions after Definition 11.8, $X - X_{\mathrm{nice}}$ is a countable set and hence a μ-null set. Recall the notation $B = \mathrm{SL}_3(\mathbb{R})^{\times\mathbb{N}}$ and its element $b = (b_0, b_1, \cdots)$. For $\nu^{\otimes\mathbb{N}}$ almost every b, we consider the Furstenberg boundary μ_b. Here μ_b is defined as the weak limit

$$\lim_{n\to\infty}(b_0b_1\cdots b_{n-1})_*\mu,$$

which is a Dirac measure [BQ16, Lemma 4.5] and denoted by $\delta_{\xi(b)}$. Because $\mu = \int\delta_{\xi(b)}\mathrm{d}\nu^{\otimes\mathbb{N}}(b)$ and $\mu(X - X_{\mathrm{nice}}) = 0$, there exists a conull set $\Omega\subset B$ such that $\xi(b)\in X_{\mathrm{nice}}$ for every $b\in\Omega$.

Note that ν also induces a random walk on the flag variety $\mathcal{F} = \mathcal{F}(\mathbb{R}^3)$ with a unique stationary measure $\nu_{\mathcal{F}}$. We can also consider the Furstenberg boundary on the flag variety. Then for almost every $b \in B$, we can associate a Dirac measure $(\mu_{\mathcal{F}})_b = \delta_{\xi_{\mathcal{F}}(b)}$. We will show that $\xi_{\mathcal{F}}(b)$ is uniquely determined by $\xi(b)$ for every $b \in B$.

Suppose we have $b \neq b'$ with $\xi(b') = \xi(b) \in X_{\text{nice}}$. We know that $b' = (b'_0, b'_1, \cdots)$ and $b = (b_0, b_1, \cdots)$ are different partitions of a same infinite word $A_{i_0}A_{i_1}\cdots A_{i_n}\cdots$. Denote $\xi_{\mathcal{F}}(b) = (\xi(b), \xi_2(b)) \in \mathcal{F}(\mathbb{R}^3)$, where $\xi_2(b)$ is a two dimensional subspace in $\mathbb{R}^3$. Then the geometrical meaning $\xi_2(b)$ is the image of any nonzero limit of $\lambda_n b_0 b_1 \cdots b_{n-1}$ in $\mathrm{End}(\wedge^2\mathbb{R}^3)$ with $\lambda_n \in \mathbb{R}$. Since the support of the measure ν is finite, for any convergent sequence $\lambda_n b_0 b_1 \cdots b_{n-1}$, we can find m_n such that

$$l(b_0 b_1 \cdots b_{n-1}) \leqslant l(b'_0 b'_1 \cdots b'_{m_n-1}) \leqslant l(b_0 b_1 \cdots b_{n-1}) + N_\nu,$$

where $l(\cdot)$ denotes the length with respect to the decomposition of the infinite word $A_{i_0}A_{i_1}\cdots$. Passing to a subsequence, we can assume that $(b'_0 \cdots b'_{m_n-1})^{-1} b_0 \cdots b_{n-1}$ is a fixed element in $\mathrm{SL}_3(\mathbb{R})$. Then the limit of $\lambda_n b'_0 \cdots b'_{m_n-1}$ has the same image with the limit of $\lambda_n b_0 \cdots b_{n-1}$ in $\mathrm{End}(\wedge^2\mathbb{R}^3)$. Thus we have

$$\xi_{\mathcal{F}}(b) = \xi_{\mathcal{F}}(b'). \tag{11.6}$$

Now we consider a conditional measure at a full measure set $\xi(\Omega)$. For any $x \in \xi(\Omega)$, choose a $b \in \Omega$ with $\xi(b) = \xi \in \mathbb{P}(\mathbb{R}^3)$ and let

$$\mu^\xi = \delta_{\xi_{\mathcal{F}}(b)}.$$

Due to Eq. (11.6), we have

$$\mu_{\mathcal{F}} = \int_\Omega \delta_{\xi_{\mathcal{F}}(b)} \mathrm{d}\nu^{\otimes \mathbb{N}}(b) = \int_{\mathbb{P}(\mathbb{R}^3)} \mu^\xi \mathrm{d}\mu(\xi).$$

So the family of measures μ^ξ is actually the disintegration of $\mu_{\mathcal{F}}$ with respect to the natural projection $\pi : \mathcal{F} \to \mathbb{P}(\mathbb{R}^3)$. Hence we obtain the trivial-fiber property. Now we can apply a same argument as in the end of the proof Proposition 8.3 by using relative measure preserving property. We obtain the desired equality between the Furstenberg entropy and the random walk entropy. □

A Auxiliary results

Identification between $\mathbb{P}(V^\perp)$ with $\mathbb{P}(\mathbb{R}^2)$ We will identify $\mathbb{P}(V^\perp)$ with $\mathbb{P}(\mathbb{R}^2)$ through rotation invariant metrics for general $V \in \mathbb{P}(\mathbb{R}^3)$. A priori there is no canonical choice of such identification, but different identifications only differ by rotations. In order to apply certain continuity arguments, we fix a one-dimensional subspace $V_0 = E_3$ and let $\mathcal{C} = \mathbb{P}(\mathbb{R}^3) - \{V_0\}$. Since most of our results are on measure-theoretical properties, we can freely delete this single *bad point* $V_0 \in \mathbb{P}(\mathbb{R}^3)$.

Lemma A.1. *For each point $V \in \mathcal{C}$, we have an identification between $\mathbb{P}(V^\perp)$ and $\mathbb{P}(\mathbb{R}^2)$ and the identification is continuous with respect to V.*

Proof. For any $V \neq V_0$, the projection $\pi_{V^\perp}$ is well-defined at V_0. We take the isometric identification between $\mathbb{P}(V^\perp)$ and $\mathbb{P}(\mathbb{R}^2)$ such that $\pi_{V^\perp} V_0$ is identified with $\mathbb{R}(1,0)$. The orientation of $\mathbb{P}(V^\perp)$ comes from the identification between $V^\perp$ and $E_2 \oplus E_3$ by elements in $\mathrm{SO}(3)$. □

Under the assumption of this lemma, by identifying $\mathbb{P}(V^\perp)$ with $\mathbb{P}(\mathbb{R}^2)$, we can regard the linear map $h_{V,g}$ in (2.9) as an element in $\mathrm{PSL}_2(\mathbb{R})$.

Weakly convergence of projections of a measure

Lemma A.2. *For any $V \in \mathcal{C}$ and any interval $I \subset \mathbb{P}(\mathbb{R}^2)$, we have*

1. *$\pi_{W^\perp}(\mu)(I)$ converges to $\pi_{V^\perp}(\mu)(I)$ if W converges to V.*
2. *$(\pi_{W^\perp}\mu)_I \to (\pi_{V^\perp}\mu)_I$ weakly if W converges to V.*

Proof. Let $I \subset \mathbb{P}(\mathbb{R}^2)$ be any interval. For any $W \in \mathbb{P}(\mathbb{R}^3)$, denote the region $(\pi_{W^\perp})^{-1}(\mathcal{I}_W^{-1}(I))$ by $S(I, W)$. So we have

$$\pi_{W^\perp}(\mu)(I) = \mu(S(I, W)).$$

For the first statement, it suffices to show $\mu(S(I, W)) \to \mu(S(I, V))$ when $W \to V$. Due to the definition of $\mathcal{I}_V$, for any $\rho > 0$, there exists $\delta > 0$ such that if $d(W, V) < \delta$, then for any $x \in \mathbb{P}(\mathbb{R}^2)$,

$$d(\mathcal{I}_V^{-1}(x), \mathcal{I}_W^{-1}(x)) < \rho.$$

Therefore, the symmetric difference $S(I, W)\Delta S(I, V)$ of $S(I, W)$ and $S(I, V)$ can be covered by the ρ neighborhood of $\partial S(I, V)$, that is $\pi_{V^\perp}^{-1}\mathcal{I}_V^{-1}(\partial I)$. We this denote by $(\partial S(I, V))^{(\rho)}$. Notice that $\partial S(I, V)$ is the union of two great circles. Therefore by Lemma 2.24, we know that $\mu(S(I, W)\Delta S(I, V))$ can be arbitrarily small when $W \to V$.

We prove the second statement. For any sub-interval $J \subset I$, $\pi_{W^\perp}(\mu)(J)$ converges to $\pi_{V^\perp}(\mu)(J)$. Therefore for any open set $U \subset I$, say $U = \cup_{i=1}^{\infty} J_i$ which is a countable disjoint union of open intervals, we have that for any n, $\pi_{W^\perp}(\mu)(\cup_{i=1}^{n} J_i) \to \pi_{V^\perp}(\mu)(\cup_{i=1}^{n} J_i)$. Hence $\liminf_{W\to V} \pi_{W^\perp}(\mu)(U) \geqslant \pi_{V^\perp}(\mu)(U)$, then by (1) and equivalent condition of weak convergence of measure we get (2). □

Sequences of numbers

Lemma A.3. *For any $\epsilon > 0$ and any $m \geqslant M(\epsilon)$, the following holds. Given any two finite sequence of numbers $\{a_i\}_{i=1}^{q^m}$ and $\{b_i\}_{i=1}^{q^m}$ in $[0, 1]$ that satisfy*

$$\sum a_i = \sum b_i = 1 \text{ and } |a_i - b_i| \leqslant q^{-3m} \text{ for each } 1 \leqslant i \leqslant q^m,$$

we have

$$\frac{1}{m}\left|\sum a_i \log a_i - \sum b_i \log b_i\right| \leqslant \epsilon$$

Proof. We estimate $|a_i \log a_i - b_i \log b_i|$ in two ways:

1. If $a_i > q^{-(1.5m)}$, since $|a_i - b_i| \leqslant q^{-3m}$, we can assume that $b_i > q^{-1.6m}$. Then by Lagrange mean value theorem we have $|\log a_i - \log b_i| \leqslant q^{1.6m} \cdot q^{-3m} = q^{-1.4m}$. Therefore

$$\begin{aligned} |a_i \log a_i - b_i \log b_i| &\leqslant |a_i - b_i| \cdot |\log a_i| + |b_i| \cdot |\log a_i - \log b_i| \\ &\leqslant q^{-3m} \cdot (1.5m) + q^{-1.4m} \leqslant q^{-1.3m} \end{aligned}$$

   if m is large enough.

2. If $a_i \leqslant q^{-1.5m}$, since $|a_i - b_i| \leqslant q^{-3m}$, we can assume that $b_i < q^{-1.4m}$ and m large enough such that q^{-m} less than the unique extreme point of $x \log x$ in $[0, 1]$. Therefore we have $|a_i \log a_i - b_i \log b_i| \leqslant 2q^{-1.4m} \cdot (1.4m) \leqslant q^{-1.3m}$ if m large enough.

Therefore combine two estimates above we get $\frac{1}{m}|\sum a_i \log a_i - \sum b_i \log b_i| \leqslant \frac{1}{m}q^{-0.3m} \leqslant \epsilon$ for any m large enough. □

B Spherical geometry

Lemma B.1. *Let $C > 0$. If $d(V, (E_1)^\perp) \geqslant 1/C$, then the map $\pi(V, E_1^\perp, V^\perp)$ has scale $u = 1$ and distortion bounded by C.*

Moreover, by the identification in Lemma A.1, for $V_1 \not\subset V_2^\perp$, the action $\pi(V_1, V_2^\perp, V_1^\perp)$ can be identified with some element in $\mathrm{PSL}_2(\mathbb{R})$ of norm $|1/d(V_1, V_2^\perp)|^{1/2}$.

Proof. Without loss of generality, we can assume that $V = \mathbb{R}v, v = \frac{e_1+\lambda e_3}{\sqrt{1+\lambda^2}}$ with $\sqrt{1+\lambda^2} = 1/d(V,(E_1)^\perp)$. Let $V' = \mathbb{R}v', v' = \frac{-\lambda e_1 + e_3}{\sqrt{1+\lambda^2}}$, then (v, v', e_2) forms an orthonormal basis of $\mathbb{R}^3$. The map $\pi(V, E_1^\perp, V^\perp)$ projects a vector $ae_2 + bv'$ to $E_1^\perp$, along kernel V. So the image is

$$(ae_2 + bv') + b\lambda v = ae_2 + b\sqrt{\lambda^2+1}e_3.$$

Therefore with basis e_2, v' and e_2, e_3, the transformation $\Pi(V, E_1^\perp, V^\perp)$ can be written as a matrix $\begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1+\lambda^2} \end{pmatrix}$. Then the distortion estimate of $\pi(V, E_1^\perp, V^\perp)$ follows from basic estimates of $\mathrm{PSL}_2(\mathbb{R})$ (Lemma 2.7).

We can find $k \in \mathrm{SO}(3)$ such that $V_2 = kE_1$, then

$$\pi(V_1, V_2^\perp, V_1^\perp) = k\pi(k^{-1}V_1, E_1^\perp, (k^{-1}V_1)^\perp)k^{-1}.$$

Then the statement follows from the computation of the first part by noticing $d(k^{-1}V_1, E_1^\perp) = d(V_1, V_2^\perp)$. $\square$

We will use the following classical result in spherical geometry later, for completeness we provide a proof.

Lemma B.2. *Let $x \neq y, V \in \mathbb{P}(\mathbb{R}^3)$ and $C > 0$. If we denote by (xy) the projective line in $\mathbb{P}(\mathbb{R}^3)$ through x, y, then we have*

1. *If $d(V, (xy)) > 1/C$, then $d(\pi_{V^\perp}x, \pi_{V^\perp}y) \geqslant d(x,y)/C$;*
2. *If either $d(V, x) \geqslant 1/C$ or $d(V, y) \geqslant 1/C$, then $d(\pi_{V^\perp}x, \pi_{V^\perp}y) \leqslant Cd(x,y)$.*

Proof. Let $z = V$. Then $d(\pi_{V^\perp}x, \pi_{V^\perp}y) = \sin(\angle_z(x,y))$ (the angle between two projective lines $(xz), (yz)$ at vertex z). By the spherical law of sines, the definition of d on $\mathbb{P}(\mathbb{R}^3)$ and the assumption of the lemma, we have

$$\frac{d(x,y)}{\sin\angle_z(x,y)} = \frac{d(y,z)}{\sin\angle_x(y,z)} \text{ and } \frac{d(z,(xy))}{\sin\angle_x(y,z)} = \frac{d(x,z)}{\sin(\pi/2)}.$$

Therefore we have

$$\frac{d(x,y)}{\sin\angle_z(x,y)} = \frac{d(y,z)d(x,z)}{d(z,(xy))} \leqslant C.$$

For the second statement, without loss of generality we assume that $d(V,y) = d(z,y) \geqslant 1/C$. By the spherical law of sines again, we have

$$\frac{d(x,y)}{\sin\angle_z(x,y)} = \frac{d(y,z)}{\sin\angle_x(y,z)} \geqslant 1/C,$$

which implies $d(\pi_{V^\perp}x, \pi_{V^\perp}y) = \sin\angle_z(x,y) \leqslant Cd(x,y)$. $\square$

C Proof of Lemma 5.6

We consider the following definition which is a quadratic analogue of σ-independence in [BHR19], (also occured in [Hoc]).

Definition C.1. Recall the map p_2 from Lemma 2.26. For six points $V_1, \cdots, V_6$ in $\mathbb{P}(\mathbb{R}^3)$, we call them ρ-quadratic independent if $p_2(V_i)$ has distance at least ρ to the subspace generated by $p_2(V_j),\ j \neq i$ in $\mathbb{P}(Sym^2\mathbb{R}^3)$.

By the following lemma, we can always find (ρ, V)-quadratic independent sets in a set with certain μ measure.

Lemma C.2. *For any $\epsilon > 0$, any $V \in \mathbb{P}(\mathbb{R}^3)$, there exists $\rho > 0$ such that for any $A \subset \mathbb{P}(\mathbb{R}^3), \mu(A) > (100\epsilon)^{1/5}$, there exists $\{V_1, \dots, V_5\}$ such that $V_1, \cdots, V_5$ and V are ρ-quadratic independent.*

Proof. The proof is basically the same as that of [Hoc]. But we first need a lemma saying that in general μ cannot have big weight on a small neighborhood of a subvariety. For a subset W of a metric space write $W^{(\epsilon)}$ for its $\epsilon-$neighborhood. Then Lemma 2.26 reads as

Lemma C.3. *For any $\epsilon > 0$, there exists $\rho(\epsilon) > 0$ such that for any proper subspace $W \subset \mathbb{P}(Sym^2\mathbb{R}^3)$, $(p_2)_*\mu(W^{(\rho)}) < \epsilon$.*

We back to the proof of Lemma C.2. Let $\mathcal{C}(V_1, V_2, \dots V_5)$ be the subspace generated by $p_2(V_1), \cdots, p_2(V_5)$ in $\mathbb{P}(Sym^2\mathbb{R}^3)$. Now we fix $V \in \mathbb{P}(\mathbb{R}^3)$, $\epsilon > 0$, and $\rho(\epsilon)$ we get from last lemma. Let $X_1, \dots, X_5$ be independent $\mathbb{P}(\mathbb{R}^3)$-valued random variables, each distributed according to μ. Therefore by Lemma C.3, if $\mu(A) > (100\epsilon)^{1/5}$,

$$\begin{aligned} & \mathbb{P}(X_i \in A, X_i \notin \mathcal{C}(V, X_j, j \neq i)^{(\rho)}) \geqslant \mathbb{P}(X_i \in A, \forall i) - 100\epsilon \\ > \quad & \mu(A)^5 - 100\epsilon > 0 \end{aligned}$$

Any realization $X_1, \dots, X_5$ from the event above is ρ-quadratic independent. □

Recall that $L \simeq \mathrm{SL}_2(\mathbb{R}) \ltimes \mathbb{R}^2$ is the group formed by $\{\begin{pmatrix} 1 & \\ n & h \end{pmatrix}, n \in \mathbb{R}^2, h \in \mathrm{SL}_2(\mathbb{R})\}$. By the assumption of Lemma 5.6, we can assume $V = \mathbb{R} \cdot (v_1, v_2, v_3)^t$ with $\sum v_i^2 = 1$ and $|v_1| \geqslant 1/2C_1$, which is equivalent to $d(V, e_1^{\perp}) \geqslant 1/2C_1$. Fix $\epsilon = 10^{-7}$ and $\rho = \rho(\epsilon)$ is decided by Lemma C.2. Denote by

$$\begin{aligned} I(\rho, V) &:= \{(V_1, \dots, V_5) \in \mathbb{P}(\mathbb{R}^3)^5, (V_1, \dots, V_5, V) \text{ is } \rho\text{-quadratic independent}\} \\ I(\rho) &:= \{(V_1, \cdots, V_5, V) \in \mathbb{P}(\mathbb{R}^3)^6,\ (V_1, \dots, V_5) \in I(\rho, V),\ d(V, e_1^{\perp}) \geqslant 1/2C_1\} \end{aligned}$$

then $I(\rho, V)$ and $I(\rho)$ are closed and hence compact.

For any (ρ, V)-quadratic independent point set $\{V_1, \dots, V_5\}$, we consider the evaluation map from L to $\mathbb{P}(\mathbb{R}^3)^5$,

$$E_{V_1, \dots V_5} : L \to \mathbb{P}(\mathbb{R}^3), \quad E_{V_1, \dots, V_5}(g) := (g(V_1), \dots, g(V_5))$$

Define the map

$$\Phi^V_{V_1, \dots, V_5} : L \to \mathbb{P}(\mathbb{R}^2)^5, \quad \Phi^V_{V_1, \dots, V_5}(g) := \pi^5_{V^{\perp}} \circ E_{V_1, \dots, V_5}(g),$$

here we denoting by $\pi^5_{V^{\perp}}$ the map $(\pi_{V^{\perp}}, \dots, \pi_{V^{\perp}}) : \mathbb{P}(\mathbb{R}^3)^5 \to \mathbb{P}(\mathbb{R}^2)^5$.

Notice that even though $\pi^5_{V^{\perp}}$ is not well-defined everywhere, the restriction of $\pi^5_{V^{\perp}}$ on $((\mathbb{P}(\mathbb{R}^3) - B(V, \rho/2))^5$ is C^1, uniformly in V. Then by compacticity, we can let

$$C_1' = C_1'(\rho, V) = \sup_{V_i \notin B(V, \rho/2)} \|D\pi^5_{V^{\perp}}\| < \infty$$

which is finite.

By continuity, there exists a compact neighborhood Z_1 of $\mathrm{id} \in L$ such that for any $g \in Z_1$ and $V_0 \in \mathbb{P}(\mathbb{R}^3)$, we have $d_{\mathbb{P}(\mathbb{R}^3)}(gV_0, V_0) \leqslant \rho/2$. Since the action $E_{V_1,\ldots,V_5}$ is C^∞ on L and all the derivatives continuously depend on $\{V_1, \ldots, V_5\}$, hence by compacticity of $I(\rho, V)$,

$$C_2' := \sup_{(V_1,\ldots,V_5)\in I(\rho,V), g_0\in Z_1} \|D_g E_{V_1,\ldots,V_5}|_{g=g_0}\|$$

is finite. The definition of ρ-quadratic independence also implies $d(V, V_i) \geqslant \rho$. Therefore by our choice of Z_1 we know that

$$C_3' := \sup_{(V_1,\ldots,V_5,V)\in I(\rho), g_0\in Z_1} \|D_g \Phi^V_{V_1,\ldots,V_5}|_{g=g_0}\| \leqslant C_1' C_2' \tag{C.1}$$

is finite. The following lemma explains the reason that we consider quadratic independence.

Lemma C.4. *There exist a compact neighborhood $Z \subset Z_1$ of $\mathrm{id} \in L$ and a constant $C_4' > 0$ such that*

$$\inf_{(V_1,\ldots,V_5,V)\in I(\rho), g_0\in Z} |\mathrm{Jac}(D_g \Phi^V_{V_1,\ldots,V_5})|_{g=g_0}| \geqslant C_4' \tag{C.2}$$

Proof. Notice that the function $|\mathrm{Jac}(Dg|_{g=g_0} \Phi_{V_1,\ldots,V_5})|$ continuously depends on $g_0 \in Z_1$ and $(V_1, \ldots, V_5, V) \in I(\rho)$. Therefore by compacticity of Z_1 and $I(\rho, V)$, to find Z and C_4' we only need to show that for all $(V_1, \ldots, V_5, V) \in I(\rho)$

$$|\mathrm{Jac}(D_g \Phi^V_{V_1,\ldots,V_5})|_{g=\mathrm{id}}| > 0. \tag{C.3}$$

We show (C.3) in a quite straight forward way. Apparently any vector in $\mathrm{Lie}(L)$ can be written as $X_{\mathfrak{n},\mathfrak{h}} = \begin{pmatrix} n_1 & h_{11} & h_{12} \\ n_2 & h_{21} & -h_{11} \end{pmatrix}$ and we assume that $V_i = \mathbb{R} \cdot (a_i, b_i, c_i)^t$. Then if (C.3) does not holds then there exists $(V_1, \ldots, V_5, V) \in I(\rho)$ and $X_{\mathfrak{n},\mathfrak{h}} \in \mathrm{Lie}(L) - \{0\}$ such that

$$(D_g \Phi^V_{V_1,\ldots,V_5})|_{g=\mathrm{id}}(X_{\mathfrak{n},\mathfrak{h}}) = 0. \tag{C.4}$$

Notice that for any $V_i \neq V$, the kernel $K_{V_i} \subset T_{V_i}\mathbb{P}(\mathbb{R}^3)$ of the tangent map $D_x \pi_{V^\perp}|_{x=V_i} : \mathbb{P}(\mathbb{R}^3) \to \mathbb{P}(\mathbb{R}^2)$ is the tangent space at V_i of the project line (VV_i) passing through V, V_i. Therefore if (C.4) holds, then

$$D_g E_{V_1,\ldots,V_5}|_{g=\mathrm{id}}(X_{\mathfrak{n},\mathfrak{h}}) \in K_{V_1} \oplus \cdots \oplus K_{V_5}$$

By the separable form of the evaluation map and the canonical correspondence between Lie algebras and the one-parameter subgroup of Lie groups, we get that the last equation is equivalent to

$$\frac{d}{dt}|_{t=0}(e^{tX_{\mathfrak{n},\mathfrak{h}}} \cdot V_i) \in K_{V_i} = T_{V_i}(VV_i), \tag{C.5}$$

for $i = 1, \cdots, 5$. If we consider the corresponding equation of (C.5) in linear space (rather than the equation in the tangent space of projective space), we get (here the action is linear action): for $i = 1, \cdots, 5$

$$\frac{d}{dt}|_{t=0}(e^{tX_{\mathfrak{n},\mathfrak{h}}} \cdot \begin{pmatrix} a_i \\ b_i \\ c_i \end{pmatrix}) \in \left\{ l \cdot \begin{pmatrix} a_i \\ b_i \\ c_i \end{pmatrix} + k \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, (k, l) \in \mathbb{R}^2 \right\}.$$

By direct calculation, we have

$$\frac{d}{dt}|_{t=0}(e^{tX_{\mathfrak{n},\mathfrak{h}}} \cdot \begin{pmatrix} a_i \\ b_i \\ c_i \end{pmatrix}) = \begin{pmatrix} n_1 & h_{11} & h_{12} \\ n_2 & h_{21} & -h_{11} \end{pmatrix} \cdot \begin{pmatrix} a_i \\ b_i \\ c_i \end{pmatrix} = \begin{pmatrix} a_i n_1 + b_i h_{11} + c_i h_{12} \\ a_i n_2 + b_i h_{21} - c_i h_{11} \end{pmatrix}.$$

Therefore

$$\det \begin{pmatrix} 0 & a_i & v_1 \\ a_i n_1 + b_i h_{11} + c_i h_{12} & b_i & v_2 \\ a_i n_2 + b_i h_{21} - c_i h_{11} & c_i & v_3 \end{pmatrix} = 0. \tag{C.6}$$

Consider the linear form φ on $Sym^2\mathbb{R}^3$

$$\begin{aligned} & (n_2 v_2 - n_1 v_3)(e_1e_1)^* + (h_{21}v_2 - h_{11}v_3 - n_2 v_1)(e_1e_2)^* \\ + \quad & (n_1 v_1 - h_{12}v_3 - h_{11}v_2)(e_1e_3)^* - (h_{21}v_1)(e_2e_2)^* + 2(h_{11}v_1)(e_2e_3)^* + (h_{12}v_1)(e_3e_3)^*. \end{aligned}$$

Recall that e_1, e_2, e_3 is a basis of $\mathbb{R}^3$ and $e_ie_j,\ i \leqslant j$ is a basis of $Sym^2\mathbb{R}^3$. The set $(e_ie_j)^*, i \leqslant j$ is the the dual basis of $e_ie_j, i \leqslant j$. Then Eq. (C.6) is just for $i = 1, \cdots, 5$

$$\varphi(u_iu_i) = 0,\ \ u_i = (a_i, b_i, c_i)^t. \tag{C.7}$$

Then $W = \mathbb{P}(\ker\varphi)$ in $\mathbb{P}(Sym^2\mathbb{R}^3)$ (unless all the coefficients of the equations are 0, but since $v_1 \neq 0$, we can easily see that if $X_{\mathfrak{n},\mathfrak{h}} \neq 0$, then at least one coefficient of the equation is non-zero).

By (C.7), W passes through $p_2(V_i)$ for $i = 1, \cdots, 5$. Observe that in (C.6) if let $(a_i, b_i, c_i) = (v_1, v_2, v_3)$ then the determinant is 0, therefore W also passes through $p_2(V)$ for $V = \mathbb{R} \cdot (v_1, v_2, v_3)^t$, contradicts with the assumption that $(V_1, \ldots, V_5, V)$ is ρ-quadratic independent. □

Remark. In Lemma C.4 and the next lemma, the condition $v_1 \neq 0$, i.e. $d(V, \mathbb{P}(E_1^\perp)) > 0$ is a crucial and necessary condition. Otherwise, the group U_V may intersect L non-trivially, which means there is a non-trivial subgroup S of L such that $\pi_{V^\perp}S = \pi_{V^\perp}id$. In this case, we cannot get bi-Lipschitz map.

As a corollary of Lemma C.4, we have

Lemma C.5. *There exists a compact neighborhood Z of identity of L and $C_6' > 0$ such that for any $(V_1, \ldots, V_5, V) \in I(\rho)$, $\Phi^V_{V_1,\ldots,V_5}|_Z$ is bi-Lipchitz with its image, with Lipchitz constant C_6'.*

Proof. By Lemma C.4, we get a compact neighborhood $Z \subset Z_1$ of identity in L and constants $C_3', C_4' > 0$ such that

$$\sup_{(V_1,\ldots,V_5,V)\in I(\rho), g_0\in Z} \|D_g\Phi^V_{V_1,\ldots,V_5}|_{g=g_0}\| \leqslant C_3', \inf_{(V_1,\ldots,V_5,V)\in I(\rho), g_0\in Z} |\mathrm{Jac}(D_g\Phi^V_{V_1,\ldots,V_5})|_{g=g_0}| \geqslant C_4'$$

which implies that there exists $C_5' > 0$

$$\sup_{(V_1,\ldots,V_5,V)\in I(\rho), g_0\in Z} \|(D_g\Phi^V_{V_1,\ldots,V_5}|_{g=g_0})^{-1}\| \leqslant C_5'.$$

Since the map $\Phi^V_{V_1,\ldots,V_5}$ is smooth on g and also on parameters $(V_1, \cdots, V_5, V)$ in a compact set $I(\rho)$, by the inverse function theorem, we obtain the lemma. □

Notice that C_6' depends on $Z, I(\rho)$ and $I(\rho)$ by definition depending on C_1, so we emphasize that $C_6' = C_6'(Z, C_1)$. As a corollary, we have

Lemma C.6. *There exists a constant $C' = C'(Z, C_1) > 0$ such that if $(V_1, \ldots, V_5, V) \in I(\rho)$ and θ supported on Z, then for any $n \in \mathbb{N}$*

$$H(\theta, \mathcal{Q}_n^L) \leqslant \sum_{j=1}^{5} H(\pi_{V^\perp}\theta \cdot V_j, \mathcal{Q}_n) + C'. \tag{C.8}$$

Proof. Let $\pi_j : \mathbb{P}(\mathbb{R}^2)^5 \to \mathbb{P}(\mathbb{R}^2)$ be the coordinate projections. Consider $\Phi^V_{V_1,\dots,V_5}$ restricted on Z in Lemma C.5, then $\Phi^V_{V_1,\dots,V_5}$ is a diffeomorphism on Z and be bi-Lipschitz to its image $\Phi^V_{V_1,\dots,V_5}(Z)$ with constant C_6'. Thus by Lemma 2.35 there is a constant C' depending on Z , $I(\rho)$ (hence C_1) such that

$$|H(\theta, \mathcal{Q}_n^L) - H(\Phi^V_{V_1,\dots,V_5}\theta, \mathcal{Q}_n^{\mathbb{P}(\mathbb{R}^2)^5})| \leqslant C'.$$

Since $\mathcal{Q}_n^{\mathbb{P}(\mathbb{R}^2)^5} = \bigvee_{j=1}^5 \pi_j^{-1}\mathcal{Q}_n^{\mathbb{P}(\mathbb{R}^2)}$, this is the same as

$$|H(\theta, \mathcal{Q}_n^L) - H(\Phi^V_{V_1,\dots,V_5}\theta, \bigvee_{j=1}^5 \pi_j^{-1}\mathcal{Q}_n^{\mathbb{P}(\mathbb{R}^2)})| \leqslant C'.$$

The statement now follows by

$$H(\Phi^V_{V_1,\dots,V_5}\theta, \bigvee_{j=1}^5 \pi_j^{-1}\mathcal{Q}_n^{\mathbb{P}(\mathbb{R}^2)}) \leqslant \sum_{j=1}^5 H(\Phi^V_{V_1,\dots,V_5}\theta, \pi_j^{-1}\mathcal{Q}_n^{\mathbb{P}(\mathbb{R}^2)}) = \sum_{i=1}^5 H(\pi_{V^\perp}\theta.V_j, \mathcal{Q}_n).$$

□

Proof of Lemma 5.6. Let C' be as in the last lemma and set

$$B = \{x \in \mathbb{P}(\mathbb{R}^3) : \frac{1}{k}H(\pi_{V^\perp}\theta.x, \mathcal{Q}_{i+k}) > \frac{1}{5k}H(\theta, \mathcal{Q}_{i+k}) - \frac{C'}{k}\}$$

Then we claim that $\mu(B) \geqslant 1 - (100\epsilon)^{1/5} = 0.9$, otherwise by Lemma C.2, for any V such that $d(V, \mathbb{P}(e_1^\perp)) \geqslant \frac{1}{2C_1}$, there exists $(V_1, \dots, V_5)$ such that $V_i \notin B$, and $(V_1, \dots, V_5, V) \in I(\rho)$. Hence by applying Lemma C.6, we get

$$H(\theta, \mathcal{Q}_{i+k}^L) \leqslant \sum_{i=1}^5 H(\pi_{V^\perp}\theta.V_i, \mathcal{Q}_{i+k}) + C' \leqslant H(\theta, \mathcal{Q}_{i+k}^L) - 4C'$$

which is a contradiction.

Therefore the lemma follows by letting $C = \max\{5, C'(Z, C_1)\}$. □

D Results in linear algebra

Lemma D.1. *For any $1/4 > r > 0$, for $g \in \mathrm{SL}_3(\mathbb{R})$ and $V \in \mathbb{P}(\mathbb{R}^3)$ with $d(V_g^+, V) > 2r$, $d(V^\perp, H_g^-) > 2r$ (in the sense of Hausdorff distance) and $|\chi_1(h_{V,g}) - \chi_1(g)| \leqslant |\log r|/2$, then*

$$d(\pi_{V^\perp}V_g^+, h_{V,g}{}^+) \leqslant q^{-\chi_1(g)}/r^2.$$

Proof. From the condition $d(V^\perp, H_g^-) > 2r$, we obtain

$$\exists x \in \mathbb{P}(V^\perp), d(x, H_g^-),\ d(x, h_{V,g}^-) > r.$$

Due to Lemma 2.5, combined with $\chi_1(h_{V,g}) \geqslant \chi_1(g) - |\log r|/2$, we have

$$d(V_g^+, gx) \leqslant \frac{q^{-\chi_1(g)}}{d(x, H_g^-)} \leqslant q^{-\chi_1(g)}/r,\ d(h_{V,g}^+, h_{V,g}x) \leqslant \frac{q^{-\chi_1(h_{V,g})}}{d(x, h_{V,g}^-)} \leqslant q^{-\chi_1(g)}/r^{3/2}. \tag{D.1}$$

Due to Lemma B.2 and $d(V_g^+, V) > r$ from the hypothesis, we obtain from the first inequality of Eq. (D.1) that

$$d(\pi_{V^\perp}V_g^+, \pi_{V^\perp}gx) \leqslant q^{-\chi_1(g)}/2r^2.$$

Combined with $\pi_{V^\perp}gx = h_{V,g}x$ and the second inequality of Eq. (D.1), we obtain the claim if $r < 1/4$. □

E Fourier decay property of stationary measures on flag variety

Recall that the flag variety $\mathcal{F} = \mathcal{F}(\mathbb{R}^3)$ is the space of flags

$$\mathcal{F}(\mathbb{R}^3) := \left\{(V_1, V_2) : V_1 \subset V_2, V_i \text{ is a linear subspace of } \mathbb{R}^3 \text{ of dimension } i\right\}.$$

In this section, we recall and explain the result in [Li22] on the Fourier decay property of stationary measures on $\mathcal{F}$. Roughly speaking, the Fourier decay property of the stationary measure μ (see Theorem E.2 for details)

$$\int_{\mathcal{F}} e^{i\xi\varphi(\eta)} r(\eta) \mathrm{d}\mu(\eta), \tag{E.1}$$

as the real parameter ξ tends to infinity, where φ is a C^2-function and r is a C^1-function.

The classical van der Corput lemma is stated for $([0,1], \mathrm{Leb})$ instead of $(\mathcal{F}, \mu)$ in Eq. (E.1), and it requires a non-degenerate condition of φ'. In our higher dimensional case, a similar non-degeneracy assumption (Definition E.1, Definition E.3) is also introduced to obtain Fourier decay.

Due to the group action of $\mathrm{SL}_3(\mathbb{R})$ and the structure of $\mathcal{F}$, not all the tangent directions on $\mathcal{F}$ play the same role. For this reason, we introduce the *simple root directions* of the tangent space.

Simple root directions of $T\mathcal{F}$. Recall α_1, α_2 are the simple roots of the root system of $\mathrm{SL}_3(\mathbb{R})$, that is for λ in the Lie algebra $\mathfrak{a} = \{\lambda = (\lambda_1, \lambda_2, \lambda_3) : \lambda_i \in \mathbb{R}, \sum \lambda_i = 0\}$, we have $\alpha_i(\lambda) = \lambda_i - \lambda_{i+1}$ for $i = 1, 2$.

We define the line bundle V_{α_1} as the sub-bundle of the tangent bundle $T\mathcal{F}$, whose fiber at any $\eta = (V_1(\eta), V_2(\eta)) \in \mathcal{F}$ is the tangent space of the circle $\{(V_1', V_2(\eta)) : V_1' \subset V_2(\eta), \dim V_1' = 1\} \subset \mathcal{F}$. Similarly, V_{α_2} is defined as the sub-line bundle of $T\mathcal{F}$ over $\mathcal{F}$, whose fiber at any $\eta \in T_\eta\mathcal{F}$ is the tangent space of the circle $\{(V_1(\eta), V_2') : V_1(\eta) \subset V_2', \dim V_2' = 2\} \subset \mathcal{F}$. From the definition, both $V_{\alpha_i}, i = 1, 2$ are $\mathrm{SL}_3(\mathbb{R})$-invariant sub-bundles of $T\mathcal{F}$.

For each $\eta \in \mathcal{F}$, let $V_{\alpha_1+\alpha_2}(\eta)$ be the one-dimensional subspace of $T_\eta\mathcal{F}$ orthogonal to $V_{\alpha_1}(\eta)$ and $V_{\alpha_2}(\eta)$. For any $g \in \mathrm{SL}_3(\mathbb{R})$, the tangent map $Dg : T\mathcal{F} \to T\mathcal{F}$ of the action $g : \mathcal{F} \to \mathcal{F}$ contracts V_{α_i} with a speed can be computed using α_i for $i = 1, 2$. For a non-zero vector v in the direction $V_{\alpha_1+\alpha_2}(\eta)$, for most of g the image vector $Dg(v)$ will have non-trivial component in $V_{\alpha_1}(g\eta) \oplus V_{\alpha_2}(g\eta)$ dominating the whole vector. See Section 2.4 in [Li22] for more details of these directions and computations of the contraction rates.

Let $Y_{\alpha_i}(\eta)$ be a unit vector in $V_{\alpha_i}(\eta)$ at η, for $i = 1, 2$, where the metric is from the $\mathrm{SO}(3)$-invariant Riemannian metric on $\mathcal{F}$. We denote by $\partial_i\varphi(\eta) = \partial_{Y_{\alpha_i}}\varphi(\eta)$ the directional derivative at a point η.[11]

(C, r)-good functions and Fourier decay of stationary measures. We introduce the (C, r)-good condition of φ for $C \geqslant 1, r \in C^1(\mathcal{F}), \varphi \in C^2(\mathcal{F})$. Recall the metrics $d, d_i, i = 1, 2$ on $\mathcal{F}$ defined in Definition 3.3.

Definition E.1 (Definition 4.1 in [Li22])**.** Let $C \geqslant 1$, $\varphi \in C^2(\mathcal{F})$ and $r \in C^1(\mathcal{F})$. Denote by $v_i := \sup_{\eta \in \operatorname{supp} r} |\partial_i\varphi|$. Let J be the $1/C$-neighborhood of $\operatorname{supp} r$. The function φ is called (C, r)-good if:

1. For $\eta \in \operatorname{supp} r$ and $i = 1, 2$
$$|\partial_i\varphi(\eta)| \geqslant v_i/C.$$

[11] The unit vector $Y_{\alpha_1}(\eta)$ cannot be continuously extended to the full $\mathcal{F}$, due to the fact that $\mathcal{F}$ is non-orientable. But we can always extend it locally and will be used in Eq. (E.6)

2. For $\eta, \eta' \in J$ and $d(\eta, \eta') < 1/C$,
$$|\varphi(\eta) - \varphi(\eta')| \leqslant C \sum_{1\leqslant i\leqslant 2} d_i(\eta, \eta')v_i.$$
3. For $\eta, \eta' \in J$ and $d(\eta, \eta') < 1/C$, $i = 1, 2$
$$|\partial_i\varphi(\eta) - \partial_i\varphi(\eta')| \leqslant Cd(\eta, \eta')v_i.$$
4.
$$\sup_{1\leqslant i\leqslant 2} v_i \in [1/C, C].$$

The third condition serves as a similar role to the second derivative. The second condition is due to that we only use two directions V_{α_1} and V_{α_2}, and we need to use the derivations on these two directions to control the Lipschitz property.

For Lipschitz-continuous function r on a metric space (X, d_X), its Lipchitz constant is defined by
$$\mathrm{Lip}(r) = \sup_{x\neq x'} \frac{|r(x) - r(x')|}{d_X(x, x')}. \tag{E.2}$$
The main result on Fourier decay is as follows.

Theorem E.2 (Theorem 1.7 in [Li22]). *Consider the action of* $\mathrm{SL}_3(\mathbb{R})$ *on* $\mathcal{F}$. *Let* ν *be a Zariski dense finitely supported probability measure on* $\mathrm{SL}_3(\mathbb{R})$, *and* $\mu_{\mathcal{F}}$ *be the* ν*-stationary measure on* $\mathcal{F}$. *Then there exist* $\epsilon_0, \epsilon_1 > 0$ *only depending on* ν *such that the following holds.*

Let $\xi > 1$ *large enough. For any pair of real functions* $\varphi \in C^2(\mathcal{F})$, $r \in C^1(\mathcal{F})$ *if* φ *is* (ξ^{ϵ_0}, r)*-good,* $\|r\|_\infty \leqslant 1$ *and* $\mathrm{Lip}(r) \leqslant \xi^{\epsilon_0}$, *then*
$$\left|\int_{\mathcal{F}} e^{i\xi\varphi(\eta)} r(\eta) \mathrm{d}\mu_{\mathcal{F}}(\eta)\right| \leqslant \xi^{-\epsilon_1}. \tag{E.3}$$

For its application in Section 3, it suffices to consider $\varphi \in C^2(\mathcal{F})$ lifted from $\mathbb{P}(\mathbb{R}^3)$ (cf. Definition 3.3). For such a function φ, (C, r)-good condition can be simplified as follows.

Definition E.3. Let $C \geqslant 1$, $r \in C^1(\mathcal{F})$ and $\varphi \in C^2(\mathcal{F})$ be a function lifted from $\mathbb{P}(\mathbb{R}^3)$. Let $v_1 = \sup_{\eta\in\operatorname{supp} r} |\partial_1\varphi|$. Let J be the $1/C$-neighborhood of $\operatorname{supp} r$. We say that φ is (C, r)-good if:

1. For $\eta \in \operatorname{supp} r$,
$$|\partial_1\varphi(\eta)| \geqslant v_1/C. \tag{E.4}$$
2. For $\eta, \eta' \in J$ and $d(\eta, \eta') < 1/C$,
$$|\varphi(\eta) - \varphi(\eta')| \leqslant Cd(V_1(\eta), V_1(\eta'))v_1. \tag{E.5}$$
3. For $\eta, \eta' \in J$ and $d(\eta, \eta') < 1/C$
$$|\partial_1\varphi(\eta) - \partial_1\varphi(\eta')| \leqslant Cd(\eta, \eta')v_1. \tag{E.6}$$
4.
$$v_1 \in [1/C, C]. \tag{E.7}$$

Claim: If a lifted function $\varphi \in C^2(\mathcal{F})$ is (C, r)-good in the sense of Definition E.3, then it is (C, r)-good in the sense of Definition E.1.

Proof of the claim. Since φ is lifted from $\mathbb{P}(\mathbb{R}^3)$, the function $\varphi(\eta) = \varphi(V_1(\eta), V_2(\eta))$ is independent of $V_2(\eta)$. Note that $Y_{\alpha_2}(\eta)$ is a tangent vector to the circle $\{(V_1(\eta), V_2') : V_1(\eta) \subset V_2', \dim V_2' = 2\}$ at η, and φ is constant on this circle. Therefore $\partial_2\varphi(\eta) = \partial_{Y_{\alpha_2}}\varphi(\eta) = 0$. Then Definition E.1 can be simplified to Definition E.3. $\square$

References

[AHS16] Artur Avila, Pascal Hubert, and Alexandra Skripchenko. On the Hausdorff dimension of the Rauzy gasket. *Bulletin de la Société Mathématique de France*, 144(3):539–568, 2016.

[AMS95] Herbert Abels, Grigorij A Margulis, and Grigorij A Soifer. Semigroups containing proximal linear maps. *Israel journal of mathematics*, 91:1–30, 1995.

[AR91] Pierre Arnoux and Gérard Rauzy. Geometric representation of sequences of complexity $(2n+1)$. *Bulletin de la Société Mathématique de France*, 119(2):199–215, 1991.

[AS13] Pierre Arnoux and Štěpán Starosta. The Rauzy Gasket. In Julien Barral and Stéphane Seuret, editors, *Further Developments in Fractals and Related Fields: Mathematical Foundations and Connections*, Trends in Mathematics, pages 1–23. Birkhäuser Boston, Boston, 2013.

[Ave72] Andre Avez. Entropie des groupes de type fini. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences, Série A*, 275:1363–1366, 1972.

[Bar10] Thierry Barbot. Three-dimensional Anosov flag manifolds. *Geometry and Topology*, 14(1):153–191, 2010.

[BCLS15] Martin Bridgeman, Richard Canary, François Labourie, and Andres Sambarino. The pressure metric for Anosov representations. *Geometric and Functional Analysis*, 25(4):1089–1179, July 2015.

[Ben97] Y. Benoist. Propriétés Asymptotiques des Groupes Linéaires:. *Geometric and Functional Analysis*, 7(1):1–47, March 1997.

[BHR19] Balázs Bárány, Michael Hochman, and Ariel Rapaport. Hausdorff dimension of planar self-affine sets and measures. *Inventiones Mathematicae*, 216(3):601–659, 2019.

[BL85] Philippe Bougerol and Jean Lacroix. *Products of Random Matrices with Applications to Schrödinger Operators.* Birkhäuser Boston, Boston, MA, 1985.

[Bor91] Armand Borel. *Linear algebraic groups*, volume 126 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1991.

[Bow79] Rufus Bowen. Hausdorff dimension of quasi-circles. *Publications mathématiques de l'IHÉS*, 50(1):11–25, December 1979.

[BPS19] Jairo Bochi, Rafael Potrie, and Andrés Sambarino. Anosov representations and dominated splittings. *Journal of the European Mathematical Society*, 21(11):3343–3414, July 2019.

[BQ16] Yves Benoist and Jean-François Quint. *Random Walks on Reductive Groups*, volume 62. Springer, 2016.

[Bry] Robert L. Bryant. Notes on geodesics on lie groups.

[Can] Richard Canary. Anosov representations: Informal lecture notes.

[CP10] Jianyu Chen and Yakov Pesin. Dimension of non-conformal repellers: a survey. *Nonlinearity*, 23(4):R93, February 2010.

[CPZ19] Yongluo Cao, Yakov Pesin, and Yun Zhao. Dimension Estimates for Non-conformal Repellers and Continuity of Sub-additive Topological Pressure. *Geometric and Functional Analysis*, 29(5):1325–1368, October 2019.

[DD09] Roberto DeLeo and Ivan A Dynnikov. Geometry of plane sections of the infinite regular skew polyhedron $\{4, 6|4\}$. *Geometriae Dedicata*, 138:51–67, 2009.

[DK22] Subhadip Dey and Michael Kapovich. Patterson-Sullivan theory for Anosov subgroups. *Transactions of the American Mathematical Society*, 375(12):8687–8737, September 2022.

[DO80] Adrien Douady and Joseph Oesterlé. Dimension de Hausdorff des attracteurs. *C. R. Acad. Sci. Paris Sér. A-B*, 290(24):A1135–A1138, 1980.

[Duf17] Laurent Dufloux. Hausdorff dimension of limit sets. *Geometriae Dedicata*, 191(1):1–35, December 2017.

[Fal88] K. J. Falconer. The Hausdorff dimension of self-affine fractals. *Mathematical Proceedings of the Cambridge Philosophical Society*, 103(2):339–350, March 1988.

[FH09] De-Jun Feng and Huyi Hu. Dimension theory of iterated function systems. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 62(11):1435–1500, 2009.

[FLR02] Ai-Hua Fan, Ka-Sing Lau, and Hui Rao. Relationships between different dimensions of a measure. *Monatshefte für Mathematik*, 135(3):191–201, 2002.

[Fou20] Charles Fougeron. Dynamical properties of simplicial systems and continued fraction algorithms. *arXiv preprint arXiv:2001.01367*, 2020.

[Fur63] H. Furstenberg. Noncommuting random products. *Transactions of the American Mathematical Society*, 108:377–428, 1963.

[Fur73] Harry Furstenberg. Boundary theory and stochastic processes on homogeneous spaces, 1973. Published: Harmonic Analysis homog. Spaces, Proc. Sympos. Pure Math. 26, Williamstown 1972, 193-229 (1973).

[Fur02] Alex Furman. Random walks on groups and random transformations. In B. Hasselblatt and A. Katok, editors, *Handbook of Dynamical Systems*, volume 1, pages 931–1014. Elsevier Science, January 2002.

[GGKW17] François Guéritaud, Olivier Guichard, Fanny Kassel, and Anna Wienhard. Anosov representations and proper actions. *Geometry & Topology*, 21(1):485–584, January 2017. Publisher: MSP.

[GM89] I Ya Goldsheid and G.A. Margulis. Lyapunov exponents of random matrices product. *Usp. Mat. Nauk*, 44:13–60, 1989.

[GMT19] Olivier Glorieux, Daniel Monclair, and Nicolas Tholozan. Hausdorff dimension of limit sets for projective Anosov representations, February 2019. arXiv:1902.01844 [math].

[GOS10] A. Gorodnik, H. Oh, and N. Shah. Strong wavefront lemma and counting lattice points in sectors. *Israel Journal of Mathematics*, 176:419–444, 2010.

[Gou22] Sébastien Gouëzel. Exponential bounds for random walks on hyperbolic spaces without moment conditions. *Tunis. J. Math.*, 4(4):635–671, 2022.

[GR85] Y. Guivarc'h and A. Raugi. Frontière de Furstenberg, propriétés de contraction et théorèmes de convergence. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 69(2):187–242, 1985.

[GRM20] Rodolfo Gutierrez-Romo and Carlos Matheus. Lower bounds on the dimension of the Rauzy gasket. *Bulletin de la Société Mathématique de France*, 148(2):321–327, 2020.

[Gui90] Yves Guivarc'h. Produits de matrices aléatoires et applications aux propriétés géometriques des sous-groupes du groupe linéaire. *Ergodic Theory and Dynamical Systems*, 10(03), September 1990.

[GW12] Olivier Guichard and Anna Wienhard. Anosov representations: Domains of discontinuity and applications. *Inventiones mathematicae*, 190(2):357–438, November 2012.

[Hem04] John Hempel. *3-manifolds*. AMS Chelsea Publishing, Providence, RI, 2004. Reprint of the 1976 original.

[HJX23] Weikun He, Yuxiang Jiao, and Disheng Xu. On dimension theory of random walks and group actions by circle diffeomorphisms. *arXiv preprint arXiv:2304.08372*, 2023.

[HL95] Irene Hueter and Steven P. Lalley. Falconer's formula for the Hausdorff dimension of a self-affine set in $\mathbb{R}^2$. *Ergodic Theory and Dynamical Systems*, 15(1):77–97, February 1995.

[Hoc] Michael Hochman. On self-similar sets with overlaps and inverse theorems for entropy in $\mathbb{R}^d$. *to appear in Memoires of the AMS*.

[Hoc14] Michael Hochman. On self-similar sets with overlaps and inverse theorems for entropy. *Annals of Mathematics*, 180(2):773–822, September 2014.

[HR21] Michael Hochman and Ariel Rapaport. Hausdorff dimension of planar self-affine sets and measures with overlaps. *Journal of the European Mathematical Society*, 24(7):2361–2441, October 2021.

[HS17] Michael Hochman and Boris Solomyak. On the dimension of Furstenberg measure for $\mathrm{SL}_2(\mathbb{R})$-random matrix products. *Inventiones mathematicae*, 210(3):815–875, December 2017.

[Kas08] F. Kassel. Proper actions on corank-one reductive homogeneous spaces. *J. Lie Theory*, 18(4):961–978, 2008.

[KLP18] Michael Kapovich, Bernhard Leeb, and Joan Porti. A Morse lemma for quasigeodesics in symmetric spaces and euclidean buildings. *Geometry & Topology*, 22(7):3827–3923, January 2018.

[KRS12] Antti Käenmäki, Tapio Rajala, and Ville Suomala. Existence of Doubling Measures Via Generalised Nested Cubes. *Proceedings of the American Mathematical Society*, 140(9):3275–3281, 2012.

[KV83] V. A. Kaĭmanovich and A. M. Vershik. Random walks on discrete groups: boundary and entropy. *The Annals of Probability*, 11(3):457–490, 1983.

[KY79] James L. Kaplan and James A. Yorke. Chaotic behavior of multidimensional difference equations. In *Functional differential equations and approximation of fixed points (Proc. Summer School and Conf., Univ. Bonn, Bonn, 1978)*, volume 730 of *Lecture Notes in Math.*, pages 204–227. Springer, Berlin, 1979.

[Lab06] François Labourie. Anosov flows, surface groups and curves in projective space. *Inventiones Mathematicae*, 165(1):51–114, 2006.

[Led83] F Ledrappier. Une relation entre entropie, dimension et exposant pour certaines marches aléatoires. *C. R. Acad. Sci. Paris Sér. I Math. 296(8), 369–372*, 1983.

[Led85] François Ledrappier. Poisson boundaries of discrete groups of matrices. *Israel Journal of Mathematics*, 50(4):319–336, December 1985.

[Lev93] Gilbert Levitt. La dynamique des pseudogroupes de rotations. *Inventiones mathematicae*, 113:633–670, 1993.

[Li22] Jialun Li. Fourier decay, renewal theorem and spectral gaps for random walks on split semisimple Lie groups. *Annales scientifiques de l'École Normale Supérieure*, 55(6):1613–1686, 2022.

[LL23a] François Ledrappier and Pablo Lessa. Exact dimension of dynamical stationary measures, March 2023. arXiv:2303.13341 [math].

[LL23b] François Ledrappier and Pablo Lessa. Exact dimension of Furstenberg measures. *Geometric and Functional Analysis*, 33(1):245–298, February 2023.

[LLS21] Gye-Seon Lee, Jaejeong Lee, and Florian Stecker. Anosov triangle reflection groups in $\mathrm{SL}(3,\mathbb{R})$, June 2021. arXiv:2106.11349 [math].

[LP82] Emile Le Page. Théoremes limites pour les produits de matrices aléatoires, 1982. Lect. Notes Math. 928, 258-303 (1982).

[MS19] Ian D. Morris and Pablo Shmerkin. On equality of Hausdorff and affinity dimensions, via self-affine measures on positive subsystems. *Transactions of the American Mathematical Society*, 371(3):1547–1582, 2019.

[MS23] Ian D. Morris and Cagri Sert. A variational principle relating self-affine measures to self-affine sets, March 2023. arXiv:2303.03437 [math].

[Nov82] Sergei Petrovich Novikov. The Hamiltonian formalism and a many-valued analogue of Morse theory. *Russian mathematical surveys*, 37(5):1, 1982.

[Pat76] S.J. Patterson. The limit set of a Fuchsian group. *Acta Mathematica*, 136:241–273, 1976.

[PS17] Rafael Potrie and Andrés Sambarino. Eigenvalues and entropy of a Hitchin representation. *Inventiones Mathematicae*, 209(3):885–925, 2017.

[PS21] Mark Pollicott and Benedict Sewell. An upper bound on the dimension of the Rauzy gasket, October 2021. arXiv:2110.07264 [math].

[PSW21] Maria Beatrice Pozzetti, Andrés Sambarino, and Anna Wienhard. Conformality for a robust class of non-conformal attractors. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2021(774):1–51, May 2021.

[PSW22] Beatrice Pozzetti, Andres Sambarino, and Anna Wienhard. Anosov representations with Lipschitz limit set. *Geometry and Topology*, 2022.

[Qui02] J.-F. Quint. Mesures de Patterson-Sullivan en rang supérieur. *Geometric and Functional Analysis*, 12(4):776–809, 2002.

[Rap21] Ariel Rapaport. Exact dimensionality and Ledrappier-Young formula for the Furstenberg measure. *Transactions of the American Mathematical Society*, 374(7):5225–5268, April 2021.

[Rap22] Ariel Rapaport. On self-affine measures associated to strongly irreducible and proximal systems. *arXiv.org*, December 2022.

[RS21] Haojie Ren and Weixiao Shen. A Dichotomy for the Weierstrass-type functions. *Inventiones mathematicae*, 226(3):1057–1100, December 2021.

[Rue82] David Ruelle. Repellers for real analytic maps. *Ergodic Theory and Dynamical Systems*, 2(1):99–107, March 1982. Publisher: Cambridge University Press.

[Sam14] Andres Sambarino. Hyperconvex representations and exponential growth. *Ergodic Theory and Dynamical Systems*, 34(3):986–1010, June 2014. arXiv:1203.0272 [math].

[Sul79] Dennis Sullivan. The density at infinity of a discrete group of hyperbolic motions. *Publications mathématiques de l'IHÉS*, 50(1):171–202, December 1979.

[Sul84] Dennis Sullivan. Entropy, Hausdorff measures old and new, and limit sets of geometrically finite Kleinian groups. *Acta Mathematica*, 153(0):259–277, 1984.

[Sul85] Dennis Sullivan. Quasiconformal homeomorphisms and dynamics II: Structural stability implies hyperbolicity for Kleinian groups. *Acta Mathematica*, 155(0):243–260, 1985.

[Yan19] Wen-yuan Yang. Statistically convex-cocompact actions of groups with contracting elements. *Int. Math. Res. Not. IMRN*, (23):7259–7323, 2019.

[You82] Lai-Sang Young. Dimension, entropy and Lyapunov exponents. *Ergodic Theory and Dynamical Systems*, 2(1):109–124, March 1982.

Yuxiang Jiao. *Peking University, No.5 Yiheyuan Road, Haidian District, Beijing, China.*
email: `ajorda@pku.edu.cn`

Jialun Li. *CNRS-Centre de Mathématiques Laurent Schwartz, École Polytechnique, Palaiseau, France.*
email: `jialun.li@cnrs.fr`

Wenyu Pan. *University of Toronto, 40 St. George St., Toronto, ON, M5S 2E4, Canada.*
email: `wenyup.pan@utoronto.ca`

Disheng Xu. *Great Bay University, Songshanhu International Community, Dongguan, Guangdong, 523000, China.*
email: `xudisheng@gbu.edu.cn`