ENTROPY AND ESCAPE OF MASS IN NON-COMPACT HOMOGENEOUS SPACES

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of the Ohio State University

By

Shirali Kadyrov, M.S., B.S.

Graduate Program in Mathematics

The Ohio State University

2010

Dissertation Committee:

Professor Nimish Shah, Advisor

Professor Vitaly Bergelson

Professor Alexander Leibman

Shirali Kadyrov

ABSTRACT

We study the limit of a sequence of probability measures on a non-compact homogeneous spaces invariant under diagonalizable flow. In this context the limit measure may not be probability. Our particular interest is to study how much mass could be left in the limit if we additionally assume that our measures have high entropy. This is a part of the project on generalizing a theorem of M. Einsiedler, E. Lindenstrauss, Ph. Michel, and A. Venkatesh. They prove that for any sequence (μ_i) of probability measure on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$ invariant under the time-one-map T of geodesic flow with entropies $h_{\mu_i}(T) \geq c$ one has that any weak* limit μ of (μ_i) has at least $\mu(X) \geq 2c-1$ mass left. We first consider the homogeneous space $\mathrm{SL}_3(\mathbb{Z})\backslash\mathrm{SL}_3(\mathbb{R})$ with an action T of a particular diagonal element $diag(e^{1/2}, e^{1/2}, e^{-1})$ and prove a generalization. Next, by constructing T-invariant probability measure with high entropy we show that our result is sharp. We also consider the Hilbert Modular space type quotient spaces and again obtain the a generalization by studying any diagonal element. As an application one can calculate an upper bound for the Hausdorff dimension of the set of points that lie on divergent trajectories with respect to the diagonal element considered, giving an alternative proof to a result of Y. Cheung. The work regarding $\mathrm{SL}_3(\mathbb{Z})\backslash\mathrm{SL}_3(\mathbb{R})$ is joint work with my co-adviser M. Einsiedler.

To My Family

ACKNOWLEDGMENTS

I would like to thank my advisor Nimish Shah and my co-adviser Manfred Einsiedler for allowing me to work on this thesis. I learned a great deal form my advisers. Their help throughout my graduate program and especially during my thesis preparation is appreciated.

I would also like to thank my other committee members, Professor Alexander Leibman, and Professor Vitaly Bergelson, for agreeing to serve on my committee.

Lastly, I would like to thank my family and all my friends in Columbus. I enjoyed living in Columbus.

VITA

1981.12	Born in Almaty, Kazakhstan
2004.3	B.S. in Mathematics, Bogazici University, Istanbul
2009.6	M.S. in Mathematics, The Ohio State University
2004.6-Present	Graduate Student, The Ohio State University

FIELDS OF STUDY

Major Field: Mathematics

Specialization: Ergodic Theory and Dynamical Systems

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Given a set of probability measures $\{\mu_i\}_{i=1}^{\infty}$ on a homogeneous space X, it is natural to ask what we can say about a weak* limit μ? Often one is interested in measures that are invariant under a transformation T acting on X, and in this case weak* limits are clearly also invariant under T. If X is non-compact, maybe the next question to ask is whether μ is a probability measure. If T acts on $X = \Gamma\backslash G$ by a unipotent element where G is a Lie group and Γ is a lattice, then it is known that μ is either the zero measure or a probability measure [MS]. This fact relies on the quantitative non-divergences estimates for unipotents due to works of S. G. Dani [Da] (further refined by G. A. Margulis and D. Kleinbock [MK]). On the other hand, if T acts on $X = \mathrm{SL}_n(\mathbb{Z})\backslash \mathrm{SL}_n(\mathbb{R})$ by a diagonal element, then $\mu(X)$ can be any value in the interval $[0, 1]$ due to softness of Anosov-flows, see § 4. However, as we will see there are constraints on $\mu(X)$ if we have additional information about the entropies $h_{\mu_i}(\mathrm{T})$ (cf. § 2.3 for a definition of entropy). This has been observed in [ELMV] for the action of geodesic flow on $\mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{SL}_2(\mathbb{R})$:

For a sequence of T-invariant probability measures μ_i with entropies $h_{\mu_i}(T) \geq c$ one has that any weak* limit μ of (μ_i) has at least $\mu(X) \geq 2c-1$ mass left where T is the time-one-map for the geodesic flow (cf. Theorem 1.2).

Our goal is to consider various generalizations of this result and in some cases show that the results we obtain are sharp. Moreover, we consider the probability measure without assumption on invariance and study the behavior of the measure under iterates of particular diagonal element. Now, we will state the results we obtain.

1.1 $\mathrm{SL}_3(\mathbb{Z})\backslash \mathrm{SL}_3(\mathbb{R})$

Let $G = \mathrm{SL}_3(\mathbb{R})$ and $\Gamma = \mathrm{SL}_3(\mathbb{Z})$. We consider the space $\Gamma\backslash G$ and a transformation T acting on it by the diagonal element $diag(e^{1/2}, e^{1/2}, e^{-1})$. We identify $X = \mathrm{SL}_n(\mathbb{Z})\backslash \mathrm{SL}_n(\mathbb{R})$ with the space of unimodular lattices in $\mathbb{R}^n$, see § 2.1. Using this correspondence in the case of $n = 3$ we can define the height function $\mathrm{ht}(x)$:

Definition 1.1. *For any 3-lattice $x \in \mathrm{SL}_3(\mathbb{Z})\backslash \mathrm{SL}_3(\mathbb{R})$ we define the* height $\mathrm{ht}(x)$ *to be the inverse of the minimum of the length of the shortest nonzero vector in x and the smallest covolume of planes w.r.t. x.*

Here, the length of a vector is given in terms of the Euclidean norm on $\mathbb{R}^3$. Also, if $n = 2$ then we consider the height $\mathrm{ht}(x)$ to be the inverse of the length of the shortest nonzero vector in x. Let

$$X_{\leq M} := \{x \in X \mid \mathrm{ht}(x) \leq M\} \text{ and } X_{\geq M} := \{x \in X \mid \mathrm{ht}(x) \geq M\}.$$

By Mahler's compactness criterion (see Theorem 2.3) $X_{\leq M}$ is compact and any compact subset of X is contained in some $X_{\leq M}$.

In [ELMV], M. Einsiedler, E. Lindenstrauss, Ph. Michel, and A. Venkatesh give the following theorem:

Theorem 1.2. *Let T be the time-one-map for the geodesic flow. There exists some $M_0 > 0$ with the property that*

$$h_\mu(T) \leq 1 + \frac{\log\log M}{\log M} - \frac{\mu(X_{\geq M})}{2}$$

for any invariant probability measure μ on $X = \mathrm{SL}(2,\mathbb{Z})\backslash \mathrm{SL}(2,\mathbb{R})$ for the geodesic flow and any $M \geq M_0$. In particular, for a sequence of T-invariant probability measures μ_i with entropies $h_{\mu_i}(T) \geq c$ we have that any weak limit μ has at least $\mu(X) \geq 2c - 1$ mass left.*

Here, μ is a weak* limit of the sequence $\{\mu_i\}_{i=1}^{\infty}$ if for some subsequence i_k and for all $f \in C_c(X)$ we have

$$\lim_{k\to\infty} \int_X f d\mu_{i_k} \to \int_X f d\mu.$$

The proof of Theorem 1.2 in [ELMV] makes use of the geometry of the upper half plane $\mathbb{H}$.

Now we let $X = \mathrm{SL}_3(\mathbb{Z})\backslash \mathrm{SL}_3(\mathbb{R})$ and as before let

$$\alpha = \begin{pmatrix} e^{1/2} & & \\ & e^{1/2} & \\ & & e^{-1} \end{pmatrix} \in \mathrm{SL}_3(\mathbb{R}).$$

We define the transformation $\mathrm{T} : X \to X$ via $\mathrm{T}(x) = x\alpha$. We have obtained in joint work [EK] with my co-adviser M. Einsiedler the following generalization of Theorem 1.2

Theorem 1.3. *There exists some $M_0 > 0$ such that*

$$h_\mu(\mathrm{T}) \leq 3 - \mu(X_{\geq M}) + O\Big(\frac{\log\log M}{\log M}\Big).$$

for any probability measure μ on X which is invariant under T *and any $M \geq M_0$.*

In this context we note that the maximum measure theoretic entropy, the entropy of T with respect to the Haar measure on X, is 3. This follows e.g. from [MT, Prop. 9.2].

As a consequence of Theorem 1.3 we have:

Corollary 1.4. *A sequence of* T*-invariant probability measures* $\{\mu_i\}_{i=1}^{\infty}$ *with entropy* $h_{\mu_i}(\mathrm{T}) \geq c$ *satisfies that any weak* limit* μ *has at least* $\mu(X) \geq c - 2$ *mass left.*

Theorem 1.3 and its corollary are obtained in § 3. Our next aim is to show that this result is sharp. In fact we will consider a more general setup.

1.2 Examples of escape of mass on $\mathrm{SL}_{d+1}(\mathbb{Z})\backslash \mathrm{SL}_{d+1}(\mathbb{R})$

The work done in this section has been submitted for publication [Ka]. Consider the homogeneous space $X = \mathrm{SL}_{d+1}(\mathbb{Z})\backslash \mathrm{SL}_{d+1}(\mathbb{R})$ and a transformation T defined by

$$\mathrm{T}(x) = xa$$

where $a = diag(e^{1/d}, e^{1/d}, ..., e^{1/d}, e^{-1}) \in \mathrm{SL}_{d+1}(\mathbb{R})$.

Theorem 1.5. *There exists a sequence of* T*-invariant probability measures* $(\mu_i)_{i\geq 1}$ *on* X *whose entropies satisfy* $\lim_{i\to\infty} h_{\mu_i}(\mathrm{T}) = d$ *but the weak* limit* μ *is the zero measure.*

In this context we note that the maximal entropy of T is $d+1$ (cf. [MT, Prop. 9.2]). One immediately obtains

Corollary 1.6. *For any* $c \in [0, 1]$ *there exists a sequence of* T*-invariant probability measures* $(\nu_i)_{i\geq 1}$ *on* X *whose entropies satisfy* $\lim_{i\to\infty} h_{\mu_i}(\mathrm{T}) = d + c$ *such that any weak* limit has mass* c*.*

§ 4 is devoted to prove Theorem 1.5 by constructing lattices on X.

1.3 $\mathrm{SL}_2(\mathcal{O})\backslash \prod_{n=1}^{r} \mathrm{SL}_2(\mathbb{R}) \times \prod_{m=1}^{s} \mathrm{SL}_2(\mathbb{C})$

In § 5 we generalize Theorem 1.2 to the following. Let F be an algebraic number field and $\mathcal{O}$ be its ring of integers. Let $S^\infty = \{\sigma_1, ..., \sigma_{r+s}\}$ be its archimedean places where $\{\sigma_1, ..., \sigma_r\}$ are the real places and the remaining ones are complex. Define

$$G := \prod_{n=1}^{r} \mathrm{SL}_2(\mathbb{R}) \times \prod_{m=1}^{s} \mathrm{SL}_2(\mathbb{C}) \text{ and } \Gamma := \mathrm{SL}_2(\mathcal{O}). \tag{1.3.1}$$

We have the natural embedding Γ into G via

$$\Delta : \gamma \to (\sigma_1(\gamma), \sigma_2(\gamma), \ldots, \sigma_{r+s}(\gamma))$$

where $\sigma_j(\gamma) = \begin{pmatrix} \sigma_j(a) & \sigma_j(b) \\ \sigma_j(c) & \sigma_j(d) \end{pmatrix}$ for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$. By Lemma 5.2 Γ becomes a lattice in G. It is in fact an irreducible lattice and the quetient space $X := \Gamma\backslash G$ is noncompact. We consider the space X as a subspace of $\mathcal{O}$-submodules Λ of $(\mathbb{R}^2)^r \times (\mathbb{C}^2)^s$ with the following properties:

1 . Λ is an $\mathcal{O}$-submodule generated by two vectors v, w of $(\mathbb{R}^2)^r \times (\mathbb{C}^2)^s$,

2 . $v = (v_1', v_1'') \times (v_2', v_2'') \times \cdots \times (v_{r+s}', v_{r+s}'')$ and $w = (w_1', w_1'') \times (w_2', w_2'') \times \cdots \times (w_{r+s}', w_{r+s}'')$ are such that $\det \begin{pmatrix} v_j' & v_j'' \\ w_j' & w_j'' \end{pmatrix} = 1$ for $j = 1, ..., r+s$.

Now, we define the *height* function $\mathrm{ht}(\cdot)$ from X to $\mathbb{R}^+$ as follows. For any $(v_j', v_j'') \in \mathbb{C}^2$ by the norm $|\cdot|$ we mean $|(v_j', v_j'')| = \max\{|v_j'|, |v_j''|\}$. For a vector $v = (v_1', v_1'') \times (v_2', v_2'') \times \cdots \times (v_{r+s}', v_{r+s}'')$ (this will be the standard notation for vectors $v \in (\mathbb{R}^2)^r \times (\mathbb{C}^2)^s$) in an $\mathcal{O}$-submodule $\Lambda \in X$ we define the norm by

$$\|v\| = |(v_1', v_1'')| \cdot |(v_2', v_2'')| \cdots |(v_{r+s}', v_{r+s}'')|.$$

Here, for any vector $v = (v_1', v_1'') \times (v_2', v_2'') \times \cdots \times (v_{r+s}', v_{r+s}'')$ the multiplication by $\lambda \in \mathcal{O}$ is defined as $\lambda \cdot v =$

$$(\sigma_1(\lambda)v_1', \sigma_1(\lambda)v_1'') \times (\sigma_2(\lambda)v_2', \sigma_2(\lambda)v_2'') \times \cdots \times (\sigma_{r+s}(\lambda)v_{r+s}', \sigma_{r+s}(\lambda)v_{r+s}'').$$

Now, we define the height of Λ:

$$\mathrm{ht}(\Lambda) := \max\{\|v\|^{-1} : v \in \Lambda - \{0\}\}.$$

Let α be any fixed diagonal element of G. Then there exist $a_j \in \mathbb{R}$ and $\theta_j \in [0, 2\pi]$ such that

$$\alpha = \begin{pmatrix} e^{i\theta_1} e^{a_1/2} & 0 \\ 0 & e^{-i\theta_1} e^{-a_1/2} \end{pmatrix} \times \cdots \times \begin{pmatrix} e^{i\theta_{r+s}} e^{a_{r+s}/2} & 0 \\ 0 & e^{-i\theta_{r+s}} e^{-a_{r+s}/2} \end{pmatrix}$$

with $\theta_1, \ldots, \theta_r = 0$. Now, we define the action of T on X by $\mathrm{T}(x) = x \cdot \alpha$.
Define $X_{<M} = \{x \in X : \mathrm{ht}(x) < M\}$ and similarly we define $X_{\geq M}$. We know from Lemma 5.2 that $X_{<M}$ has compact closure which motivates the definition of height. Now, we can state a second generalization of Theorem 1.2.
Let $|a_1| + \cdots + |a_r| = h_r$ and $|a_{r+1}| + \cdots + |a_{r+s}| = h_s$.

Theorem 1.7. *Let $M > \max\{e^{h_r+h_s}, e^{e(r+s)}\}$ be given. For any* T*-invariant probability measure μ one has*

$$h_\mu(\mathrm{T}) \leq h_r + 2h_s - \frac{(h_r + h_s)\mu(X_{\geq M})}{2} + O(\frac{\log\log M}{\log M})$$

In particular, for the sequence of T*-invariant probability measures μ_n with $h_{\mu_n}(\mathrm{T}) \geq h$ one has that any weak* limit μ_∞ has at least $\frac{2h-h_r-3h_s}{h_r+h_s}$ mass left.*

Note that the maximal metric entropy of T is $h_r + 2h_s$. Thus, whenever $2h \in (h_r + 3h_s, 2h_r + 4h_s]$ there will be some mass left in the limit. We think that if $s = 0$ then the theorem is sharp in the following sense: there exists a sequence of T-invariant probability measures with metric entropy equal to $\frac{h_r}{2}$ with limit being a 0 measure. On the other hand, if $r = 0$ then one can obtain from the arguments in the proof of Theorem 1.7 that for the sequence of T-invariant measures μ_n with $h_{\mu_n}(\mathrm{T}) \geq h$ one has that any weak* limit μ_∞ has at least $\frac{h-h_s}{h_s}$ mass left. In this case, we again think that this is sharp.

1.4 Measures with high local dimension

In the last chapter § 6 we study measures with high local dimension. The methods developed in § 3 and § 5 yields the following.

Let G be either $\mathrm{SL}_3(\mathbb{R})$ or as in (1.3.1) and Γ be either $\mathrm{SL}_3(\mathbb{Z})$ or as in (1.3.1), respectively. Also, let α be the corresponding diagonal element as before. Let us consider the following subgroups of G

$$U^+ = \{g \in G : \alpha^{-n} g \alpha^n \to 1 \text{ as } n \to -\infty\},$$

$$U^- = \{g \in G : \alpha^{-n} g \alpha^n \to 1 \text{ as } n \to \infty\},$$

$$C = \{g \in G : g\alpha = \alpha g\}.$$

For a group G' we define $B_\epsilon^{G'}(g)$ to be the open ball in G' of radius $\epsilon > 0$ centered at $g \in G'$. Let d be given and let us consider a probability measure ν on $X := \Gamma \backslash G$ with the following property. For any $\delta > 0$ there exists $\epsilon' > 0$ such that for any $\epsilon < \epsilon'$ one has

$$\nu(xB_\epsilon^{U^+} B_\eta^{U^-C}) \ll \epsilon^{d-\delta} \text{ for any } \eta \in (0,1) \text{ and for any } x \in X.$$

We say that ν has *dimension at least d in the unstable direction.*

Here $X \ll Z$ means that there exists a positive constant c such that $X \leq cZ$. Also, $X \ll_d Z$ means that the constant c depends on d.

Now, we consider the following sequence of measures μ_n defined by

$$\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} \mathrm{T}_*^i \, \nu$$

where $\mathrm{T}_*^i \, \nu$ is the push-forward of ν under T^i. In [EK] (see also § 6) we show the following connection between dimension and escape of mass.

Theorem 1.8. *Let $X = \mathrm{SL}_3(\mathbb{Z})\backslash \mathrm{SL}_3(\mathbb{R})$. For a fixed $d \in [0, 2]$ let ν be the probability measure of dimension at least d and let μ_n be as above. Then the sequence of probability measures $\{\mu_n\}_{n=1}^{\infty}$ satisfies that any weak* limit μ has at least $\mu(X) \geq \frac{3}{2}(d - \frac{4}{3})$ mass left.*

In particular, if $d = 2$ then the limit μ is a probability measure. In this case with a minor additional assumption on ν one in fact obtains the equidistribution result, that is, the limit measure μ is the Haar measure [Sh].

Another application of Theorem 1.8 is that it gives the sharp upper bound for the Hausdorff dimension of singular pairs. The exact calculation of Hausdorff dimension of singular pairs was achieved in [Ch]. We say that $\mathbf{r} \in \mathbb{R}^2$ is *singular* if for every $\delta > 0$ there exists $N_0 > 0$ such that for any $N > N_0$ the inequality

$$\|q\mathbf{r} - \mathbf{p}\| < \frac{\delta}{N^{1/2}}$$

admits an integer solution for $p \in \mathbb{Z}^2$ and for $q \in \mathbb{Z}$ with $0 < q < N$. From our results we obtain the precise upper bound that the set of singular pairs has Hausdorff dimension at most $\frac{4}{3}$, which gives an independent proof of the upper bound in [Ch]. Let $x \in \mathrm{SL}_3(\mathbb{Z})\backslash \mathrm{SL}_3(\mathbb{R})$. Then we say x is divergent if $T^n(x)$ diverges in $\mathrm{SL}_3(\mathbb{Z})\backslash \mathrm{SL}_3(\mathbb{R})$. We recall (e.g. from [Ch]) that $\mathbf{r}$ is singular if and only if

$$x_{\mathbf{r}} = \mathrm{SL}_3(\mathbb{Z}) \begin{pmatrix} 1 & & \\ & 1 & \\ r_1 & r_2 & 1 \end{pmatrix}$$

is divergent. An equivalent formulation[1] of the above Hausdorff dimension result (see [Ch]) is that the set of divergent points in $\mathrm{SL}_3(\mathbb{Z})\backslash \mathrm{SL}_3(\mathbb{R})$ has Hausdorff dimension $8 - \frac{2}{3} = \frac{4}{3} + 6$.

[1]Roughly speaking the additional 6 dimensions corresponding to U^-C are not as important as the 2 directions in the unstable horospherical subgroup U^+. The latter is parametrized by the unipotent matrix as in the definition of $x_{\mathbf{r}}$.

However, we can also strengthen this observation as follows. A weaker requirement on points (giving rise to a larger set) would be divergence on average, which we define as follows. A point x is *divergent on average* (under T) if the sequence of measures

$$\frac{1}{N}\sum_{n=0}^{N-1}\delta_{T^n(x)}$$

converges to zero in the weak* topology, i.e. if the mass of the orbit — but not necessarily the orbit itself — escapes to infinity.

Corollary 1.9. *The Hausdorff dimension of the set of points that are divergent on average is also* $\frac{4}{3}+6$.

We note that the nondivergence result [KLW, Theorem 3.3] is related to Theorem 1.8. In fact, [KLW, Theorem 3.3] implies that μ as in Theorem 1.8 is a probability measure if ν has additional regularity properties (ν is assumed to be friendly). However, to our knowledge these additional assumptions make it impossible to derive e.g. Corollary 1.9.

Theorem 1.10. *Let* $X = \Gamma\backslash G$ *where* G, Γ *are as in* (1.3.1). *For a fixed* d *let* ν *be the probability measure of dimension at least* d *and let* μ_n *be as above. Then the sequence of probability measures* $(\mu_n)_{n\geq 1}$ *satisfies that any weak* limit* μ_∞ *has at least* $\mu_\infty(X) \geq \frac{2d-h_r-3h_s}{h_r+h_s}$ *mass left.*

In particular, if ν has full dimension, that is if $d = h_r + 2h_s$, then the limit μ_∞ is a probability measure. In this case again with a minor additional assumption on ν one in fact obtains the equidistribution result, that is, the limit measure μ_∞ is the Haar measure [Sh].

Corollary 1.11. *The Hausdorff dimension of the points in* X *that lie on the forward divergent trajectories w.r.t* α *is* $\leq \frac{h_r+3h_s}{2} + 2(h_r + 2h_s)$.

CHAPTER 2

PRELIMINARIES

2.1 The space of lattices and Mahler's compactness criterion

In this section we will give a brief introduction to the space of unimodular lattices in $\mathbb{R}^n$.

Definition 2.1. $\Lambda \subset \mathbb{R}^n$ *is a* lattice *if it is a discrete subgroup and has a compact quotient* $\mathbb{R}^n/\Lambda$.

Note that this is equivalent to saying that $\Lambda = \langle v_1, v_2, \ldots, v_n \rangle_{\mathbb{Z}}$ where $v_1, v_2, \ldots, v_n$ are linearly independent vectors over $\mathbb{R}$.

Definition 2.2. *A lattice* $\Lambda = \langle v_1, v_2, \ldots, v_n \rangle_{\mathbb{Z}}$ *is said to be* unimodular *if it has covolume equal to* 1 *where the* covolume *is the absolute value of the determinant of the matrix with row vectors* $v_1, v_2, \ldots, v_n$.

Let $g \in \mathrm{SL}_n(\mathbb{R})$ and let $v_1, v_2, \ldots, v_n$ be its row vectors. We identify $\mathrm{SL}_n(\mathbb{Z})g \in X$ with the unimodular lattice in $\mathbb{R}^n$ generated by vectors $v_1, v_2, \ldots, v_n$. In particular $\mathrm{SL}_n(\mathbb{Z})$ corresponds to $\mathbb{Z}^n$.

Let $g, g' \in \mathrm{SL}_n(\mathbb{R})$ have row vectors $v_1, v_2, \ldots, v_n$ and $w_1, w_2, \ldots, w_n$ respectively. If $g' = hg$ where $h \in \mathrm{SL}_n(\mathbb{Z})$ then clearly

$$\langle v_1, v_2, \ldots, v_n \rangle_{\mathbb{Z}} = \langle w_1, w_2, \ldots, w_n \rangle_{\mathbb{Z}}$$

and hence they correspond to the the same unimodular lattice in $\mathbb{R}^n$.

Therefore, we can think of $\mathrm{SL}_n(\mathbb{Z})\backslash\mathrm{SL}_n(\mathbb{R})$ as a space of unimodular lattices in $\mathbb{R}^n$. Now we will state Mahler's compactness criterion which motivates the definition of the height function above.

Theorem 2.3 (Mahler's compactness criterion)**.** *A closed subset $K \subset \mathrm{SL}_n(\mathbb{Z})\backslash\mathrm{SL}_n(\mathbb{R})$ is compact if and only if there is a $\delta > 0$ such that no lattice in K contains a δ-small non-zero vector.*

For the proof the reader can refer to [Ra, Cor. 10.9].

2.2 Riemannian metric on $\Gamma\backslash G$

Let G be a closed linear group and Γ be a discrete subgroup of G. We fix a left-invariant Riemannian metric d_G on G and for any $x_1 = \Gamma g_1, x_2 = \Gamma g_2 \in \Gamma\backslash G$ we define

$$d_{\Gamma\backslash G}(x_1, x_2) = \inf_{\gamma\in\Gamma} d_G(g_1, \gamma g_2)$$

which gives us a left-invariant Riemannian metric $d_{\Gamma\backslash G}$ on $\Gamma\backslash G$. For more information about the Riemannian metric, we refer to [Sa, Chp. 2].

For any subgroup H of G let $B_r^H(x) := \{h \in H \mid d(h, x) < r\}$ and B_r^H is understood to be $B_r^H(1)$.

Lemma 2.4. *For any $x \in \Gamma\backslash G$ there is an injectivity radius $r > 0$ such that the map $g \mapsto xg$ from $B_r^G \to B_r^{\Gamma\backslash G}(x)$ is an isometry.*

Proof. Let $x = \Gamma h$ and let $g_1, g_2 \in B_r^G$, then we need

$$d_{\Gamma\backslash G}(\Gamma h g_1, \Gamma h g_2) = \inf_{\gamma\in\Gamma} d_G(hg_1, \gamma h g_2) = d_G(g_1, g_2).$$

Clearly

$$\inf_{\gamma\in\Gamma} d_G(hg_1, \gamma h g_2) = \inf_{\gamma\in\Gamma} d_G(g_1, h^{-1}\gamma h g_2) \leq d_G(g_1, g_2) < 2r.$$

Then, for $\gamma \in \Gamma$ by left-invariance and triangular inequality we have

$$d_G(e, h^{-1}\gamma h) \leq d_G(e, g_1) + d_G(g_1, h^{-1}\gamma h g_2) + d_G(h^{-1}\gamma h g_2, h^{-1}\gamma h) < 4r.$$

But, $h^{-1}\Gamma h$ is discrete since Γ is. Thus for small enough $r > 0$ we get

$$d_G(e, h^{-1}\gamma h) = 0$$

and hence γ must be e which gives the isometry

$$d_{\Gamma\backslash G}(\Gamma h g_1, \Gamma h g_2) = d_G(g_1, g_2).$$

□

Note that for G and Γ we consider in this paper since $X_{\leq M}$ is compact (cf. Theorem 2.3 and Lemma 5.2), we can choose $r > 0$ which is an injectivity radius for every point in $X_{\leq M}$. In this case, r is called *an injectivity radius of* $X_{\leq M}$.

2.3 Metric entropy

In this section we will give the definition of (metric) entropy. It is also known as measure theoretic entropy. For more information we refer to [Wa].

Let μ be a T-invariant probability measure and let $(X, \mathcal{B}, \mu)$ be a probability space where $\mathcal{B}$ is a Borel σ-algebra. A *partition* of $(X, \mathcal{B}, \mu)$ is a disjoint set of elements of $\mathcal{B}$ whose union is X. For two partitions $\xi = \{A_1, ..., A_l\}$ and $\beta = \{B_1, ..., B_m\}$ we can define their *join*

$$\xi \vee \beta = \{A_i \cap B_j : 1 \leq i \leq l, 1 \leq j \leq m\}.$$

First we define the *entropy of a partition* $\xi = \{A_1, ..., A_l\}$ by

$$H_\mu(\xi) = -\sum_{i=1}^{l} \mu(A_i) \log \mu(A_i)$$

with the convention $0 \log 0 = 0$. We note here that the strict convexity of the function $x \log x$ for $x > 0$ implies that $H_\mu(\xi) \leq \log |\xi|$ where $|\xi|$ is the number of partition elements of ξ.

In the second step we define the *entropy of* T *with respect to* ξ by

$$h(\mathrm{T},\xi) := \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} \mathrm{T}^{-i}\,\xi\right) = \lim_{n\to\infty} \frac{1}{2n-1} H_\mu\left(\bigvee_{i=-n+1}^{n-1} \mathrm{T}^{-i}\,\xi\right).$$

This limit exists and in fact, the sequence decreases to $h(\mathrm{T},\xi)$. Finally, we define the *entropy* of T by

$$h(\mathrm{T}) = \sup_{\xi} h(\mathrm{T},\xi)$$

where the supremum is taken over all finite partitions ξ of X.

We also want to recall the definition of conditional entropy and one of its properties that will be needed to prove Lemma 2.5 below. Conditional entropy is not necessary to define the entropy as it is seen above. However, it is useful to derive many properties of entropy. For any two partitions $\xi = \{A_1, ..., A_l\}$ and $\beta = \{B_1, ..., B_m\}$ we define *the entropy of* ξ *given* β to be

$$H_\mu(\xi \,|\, \beta) = -\sum_{i,j} \mu(A_i \cap B_j) \log \frac{\mu(A_i \cap B_j)}{\mu(B_j)}$$

where the j-terms omitted whenever $\mu(B_j) = 0$. Now it is easy to deduce that

$$H_\mu(\xi \vee \beta) = H_\mu(\beta) + H_\mu(\xi \,|\, \beta). \tag{2.3.1}$$

Let G be a closed linear group and Γ be a lattice in G, that is, a discrete subgroups with finite covolume and let $X = \Gamma\backslash G$. Fix $\eta > 0$ small enough so that B_η^G is an injective image under the exponential map of a neighborhood of 0 in the Lie algebra. Fix a diagonal element $\alpha \in G$ and define a Bowen N-ball to be the translate xB_N for some $x \in X$ of

$$B_N = \bigcap_{n=-N}^{N} \alpha^{-n} B_\eta^G \alpha^n.$$

Let T act on X as right multiplication by α. To obtain an upper estimate for entropy we need the following lemma which goes back to [BK].

Lemma 2.5. *Let μ be a* T*-invariant measure on X. For any $N \geq 1$ and $\epsilon > 0$ let $BC(N,\epsilon)$ be the minimal number of Bowen N-balls needed to cover any subset of X of measure bigger that $1-\epsilon$. Then*

$$h_\mu(\mathrm{T}) \leq \lim_{\epsilon \to 0} \liminf_{N \to \infty} \frac{\log BC(N,\epsilon)}{2N}.$$

Proof. We will follow the proof given in [ELMV]. For any $\delta > 0$, we claim that there exists a partition $\xi = \{A_0, A_1, \ldots, A_l\}$ with the following properties. The only unbounded element is A_0. The boundaries of partition elements are μ-null sets and there exists a constant $C > 0$ such that

$$\mu(\partial A_j B_t^G) < Ct$$

for any $t > 0$. Finally, $h_\mu(\mathrm{T}, \xi) > h_\mu(\mathrm{T}) - \delta$.

Since μ is a probability measure there cannot be uncountably many disjoint sets in X of positive measure. Thus, for any $x \in X$, $\mu(\partial B_r^G(x)) = 0$ for a.e. r. For any $x \in X$, $f(r) := \mu(B_r^G(x))$ is a decreasing function of r and hence it is differentiable a.e. If $\mu(B_r^G(x))$ is differentiable at r then we have

$$\lim_{t \to 0} \frac{\mu(B_{r+t}^G(x)) - \mu(B_{r-t}^G(x))}{2t} = f'(r).$$

Thus, $\mu((\partial B_r^G) B_t^G) \leq C_r t$ for sufficiently small $t > 0$ where $C_r = 2(|f'(r)| + 1)$.

Now, one can construct a fine partition with thin boundary as follows:

For any integer $M > 1$ let us cover the compact space $X_{\leq M}$ with finitely many balls $B_r^G(x_1), \ldots, B_r^G(x_l)$ of radius $r > 0$ where r is chosen so that it is less than the injectivity radius of $X_{\leq M}$ and each $B_r^G(x_j)$ has thin boundary. Also, assume that $r < 1/M$. Let $A_1 = B_r^G(x_1)$ and $A_{j+1} = B_r^G(x_{i+1}) \setminus \cup_{i=1}^{j} A_i$ for $j \in \{1, 2, \cdots, l\}$.

Finally we set $A_0 = X \backslash \cup_{j=1}^{l} A_j \subset X_{<M}$. Enlarging the constant C_r we still have that $\mu(\partial A_j B_t^G) < C_r t$ for any sufficiently small $t > 0$. This gives a partition $\mathcal{A}(M) = \{A_0, A_1, \ldots, A_l\}$. We have that except for A_0, $diam(A_j) \leq 2r < 1/M$ for all j and $A_0 \subset X_{>M}$. Thus, $\bigvee_{M=1}^{\infty} \mathcal{A}(M) = \mathcal{B}$ where $\mathcal{B}$ is the Borel σ-algebra. Inductively if $M' > M$ then we can assume that $\mathcal{A}(M) \subset \mathcal{A}(M')$. Thus, [Wa, Theorem 4.22] gives that $h(T) = \lim_{M\to\infty} h(T, \mathcal{A}(M))$. Hence, for sufficiently large $M > 1$ the finite σ-algebra $\mathcal{A}(M)$ gives the partition ξ with thin boundary such that $h_\mu(\mathrm{T}, \xi) > h_\mu(\mathrm{T}) - \delta$. Let $\xi^N = \bigvee_{j=-N}^{N} \mathrm{T}^j(\xi)$. We claim that except for a set of small measure ($\ll N^{-1}$), for any $x \in X$ the partition element in ξ^N containing x also contains the modified Bowen N-ball

$$yB'_N = y \bigcap_{n=-N}^{N} \alpha^{-n} B^G_{\eta N^{-2}} \alpha^n$$

whenever $x \in yB'_N$. Assume that for some $x \in X$ there exist y, y' such that $x \in yB'_N \cap y'B'_N$, but yh and $y'h'$ belong to different partition elements of ξ^N for some $h, h' \in B'_N$. Thus, there exists $n \in [-N, N]$ such that $yh\alpha^{-n}, y'h'\alpha^{-n}$ belong different partition elements of ξ. If we let $x = yg = y'g'$ for some $g, g' \in B'_N$ then

$$d(yh\alpha^{-n}, y'h'\alpha^{-n}) = d(xg^{-1}h\alpha^{-n}, x(g')^{-1}h'\alpha^{-n}) \leq d(\alpha^n (h')^{-1} g' g^{-1} h \alpha^n, 1) \leq 4\eta N^{-2}.$$

Thus, $yh\alpha^{-n}, y'h'\alpha^{-n} \in (\partial A_j) B^G_{4\eta N^{-2}}$ for some $A_j \in \xi$. On the other hand,

$$d(x\alpha^{-n}, yh\alpha^{-n}) \leq d(\alpha^n h^{-1} g \alpha^{-n}, 1) \leq 2\eta N^{-2}$$

so that $x\alpha^{-n} \in (\partial A_j) B^G_{6\eta N^{-2}}$. Hence,

$$x \in \bigcup_{n=-N}^{N} \mathrm{T}^n \Big(\bigcup_{j=0}^{l} (\partial A_j) B^G_{6\eta N^{-2}}\Big)$$

and this set has measure less than $(2N+1) l C 6\eta N^{-2} \ll N^{-1}$ which proves the claim. It is easy to see that B_N can be covered by $\ll N^k$ translates of B'_N where N^k depends on G and α.

Now, fix $f > \lim_{\epsilon\to 0} \liminf_{N\to\infty} \frac{\log BC(N,\epsilon)}{2N}$. Let $\epsilon > 0$ be given and let $N \geq 1$ be large enough so that $N^{-1} < \epsilon$ and $BC(N, \epsilon) \leq e^{2Nf}$. Since at most e^{2Nf} many Bowen N-balls needed to cover $1-\epsilon$ of the space, we need $\ll N^k e^{2Nf}$ many translates of B'_N to cover $1-\epsilon$ of the space. Thus, the above discussion yields that $1-2\epsilon$ of the space can be covered by $\ll N^k e^{2Nf}$ many partition elements of ξ. Now, we let A to be the union of these partition elements and let us consider the partition $\beta = \{A, X\backslash A\}$. Writing μ_D for the measure μ restricted to D we have

$$
\begin{aligned}
&H_\mu(\xi^N) = H_\mu(\xi^N \vee \beta) = H_\mu(\beta) + H_\mu(\xi^N \mid \beta) \\
&= H_\mu(\beta) - \sum_{P\in\xi^N} \mu(P\cap A) \log \frac{\mu(P\cap A)}{\mu(A)} - \sum_{P\in\xi^N} \mu(P\cap (X\backslash A)) \log \frac{\mu(P\cap (X\backslash A))}{\mu(X\backslash A)} \\
&\qquad = H_\mu(\beta) + \mu(A) H_{\mu_A}(\xi^N) + \mu(X\backslash A) H_{\mu_{X\backslash A}}(\xi^N).
\end{aligned}
$$

On the other hand, we know that $H_\mu(\zeta) \leq \log|\zeta|$ for any partition ζ of X. Thus,

$$H_\mu(\xi^N) \leq \log 2 + \log N^k e^{2Nf} + 2\epsilon \log(l+1)^{2N+1} \leq 2Nf + \log 2 + k\log N + 2\epsilon(2N+1)l.$$

Letting $N \to \infty$ we obtain

$$h_\mu(\mathrm{T}) - \delta < h_\mu(\mathrm{T}, \xi) \leq f + 2\epsilon l,$$

which completes the proof. □

2.4 Topological entropy and Variational principle

In this section we will briefly introduce topological entropy and its relation to metric entropy which is called the Variational principle. For details and proofs we refer to Chapter 7 and Chapter 8 of [Wa].

There are various definitions of topological entropy. Here, we will give the definition of topological entropy in terms of separated sets. Let (Y, d_0) be a compact metric space and let $\mathrm{T} : Y \to Y$ be a continuous map. Define a new metric d_n on Y by

$$d_n(x, y) = \max_{0 \le i \le n-1} d_0(\mathrm{T}^i(x), \mathrm{T}^i(y)).$$

For a given $\epsilon > 0$ and a natural number n, we say that the couple x, y is (n, ϵ)-*separated* if $d_n(x, y) \ge \epsilon$ and we say that the set E is (n, ϵ)-*separated* if any distinct $x, y \in E$ is (n, ϵ)-separated.

Now define $s_n(\epsilon, Y)$ to be the cardinality of the largest possible (n, ϵ)-separated set and let

$$s(\epsilon, Y) = \limsup_{n \to \infty} \frac{1}{n} \log s_n(\epsilon, Y).$$

Finally, we define the *topological entropy* of T with respect to Y by

$$h(\mathrm{T}) = \lim_{\epsilon \to 0} s(\epsilon, Y).$$

Here is the relation between the topological entropy and the metric entropy:

Theorem 2.6 (Variational principle)**.** *Topological entropy $h_{\mathrm{T}}(Y)$ of a* T*-invariant compact metric space Y is the supremum of metric entropies $h_\mu(Y)$ where supremum is taken over all* T*-invariant probability measures on the set Y.*

CHAPTER 3

THE DIAGONAL ACTION ON $\mathrm{SL}_3(\mathbb{Z})\backslash\mathrm{SL}_3(\mathbb{R})$ AND ESCAPE OF MASS

This chapter is joint work [EK] with my co-adviser M. Einsiedler. For this chapter we let $G = \mathrm{SL}_3(\mathbb{R}), \Gamma = \mathrm{SL}_3(\mathbb{Z})$ and $X = \Gamma\backslash G$ with a right action T of the diagonal element $\alpha = diag(e^{1/2}, e^{1/2}, e^{-1})$. Our goal for this chapter is to prove Theorem 1.3 and its Corollary 1.4. First, we will deduce Corollary 1.4 from Theorem 1.3.

Proof of Corollary 1.4. We need to approximate $1_{X_{\leq M}}$ by functions of compact support. So, let $f \in C_c(X)$ be such that

$$f(x) = \begin{cases} 1 & \text{for } x \in X_{\leq M} \\ 0 & \text{for } x \in X_{\geq (M+1)} \end{cases}$$

and $0 \leq f(x) \leq 1$ otherwise. This is possible by Urysohn's Lemma. Hence,

$$\int f\, d\mu_i \geq \int 1_{X_{\leq M}}\, d\mu_i = \mu_i(X_{\leq M}) \geq c - 2 - \epsilon(M)$$

where $\epsilon(M) = O(\frac{\log\log M}{\log M})$. Let μ be a weak* limit, then we have

$$\lim_{i_k \to \infty} \int f\, d\mu_{i_k} = \int f\, d\mu$$

and hence we deduce that

$$\int f\, d\mu \geq c - 2 - \epsilon(M).$$

Now, by definition of f we get $\int f\, d\mu \leq \mu(X_{<(M+1)})$. Thus,

$$\mu(X_{<(M+1)}) \geq c - 2 - \epsilon(M).$$

This is true for any $M \geq M_0$, so letting $M \to \infty$ finally we have

$$\mu(X) \geq c - 2$$

which completes the proof. $\square$

The rest of the chapter is devoted to prove Theorem 1.3.

3.1 Sets of labeled marked times

Since the function $\phi(x) = -\log x$ is convex in $(0, \infty)$ Jensen's inequality gives that for any partition ξ of X one has

$$H(\xi) \leq \log |\xi|$$

where $|\xi|$ is the number of elements of ξ. Hence, to obtain upper estimates of entropy it is useful to calculate the number of elements of partitions. In this section, we define the sets of labeled marked times which corresponds to a particular partitioning of X and we count the cardinality of this partition. By considering vectors and planes on a lattice in X we first characterize when the forward trajectory of x is above height M. However, we do not want to consider all vectors in x that are responsible for x being of height M at some time moment. Rather whenever there are two linearly independent primitive $1/M$-short vectors, our strategy is to consider a plane in x that contains both vectors. So, for a given lattice x we would like to associate a set of labeled marked times in $[-N, N]$ which tells us when a vector or a plane is getting resp. stops being $1/M$-short. Considering all such possible marked times for lattices in $X_{\leq M}$ we get a family $\mathcal{M}_N$ of sets of labeled marked times which will be defined in § 3.1. This will give rise to a partition of X, which will be helpful in the main estimates given in § 5.3.

Short lines and planes

Let $u, v \in \mathbb{R}^3$ be linearly independent. We recall that the covolume of the two-dimensional lattice $\mathbb{Z}u + \mathbb{Z}v$ in the plane $\mathbb{R}u + \mathbb{R}v$ equals $|u \wedge v|$. Here, $u \wedge v = (u_1, u_3, u_3) \wedge (v_1, v_2, v_3) = (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1)$. Below $u, v \in \mathbb{R}^3$ will always be such that $\mathbb{Z}u + \mathbb{Z}v = x \cap (\mathbb{R}u + \mathbb{R}v)$ for a lattice x. In this case we call $\mathbb{R}u + \mathbb{R}v$ *rational* w.r.t. x and will call $|u \wedge v|$ *the covolume of the plane* $\mathbb{R}u + \mathbb{R}v$ w.r.t. x.

We also note that the action of T extends to $\bigwedge^2 \mathbb{R}^2$ via

$$\begin{aligned} \mathrm{T}(u \wedge v) &= (u_1 e^{1/2}, u_2 e^{1/2}, u_3 e^{-1}) \wedge (v_1 e^{1/2}, v_2 e^{1/2}, v_3 e^{-1}) \\ &= ((u_2v_3 - u_3v_2)e^{-1/2}, (u_3v_1 - u_1v_3)e^{-1/2}, (u_1v_2 - u_2v_1)e^{1}). \end{aligned} \tag{3.1.1}$$

Let $\epsilon > 0$ be given. Fix $x \in X$, a vector v in x is *ϵ-short at time n* if $|\,\mathrm{T}^n(v)| \leq \epsilon$. We say that a nontrivial subspace $V \subset \mathbb{R}^3$ (i.e. a line or a plane) is *ϵ-short at time n* (w.r.t. x) if $\mathrm{T}^n(V)$ is rational w.r.t. $\mathrm{T}^n(x)$ and its covolume is $\leq \epsilon$.

(Labeled) Marked Times

Now, for a positive number N and a lattice $x \in \mathrm{T}^N(X_{\leq M})$ we explain which times will be marked in $[-N, N]$ and how they are labeled. The following lemma which is special to $\mathrm{SL}_3(\mathbb{Z})\backslash \mathrm{SL}_3(\mathbb{R})$ is crucial.

Lemma 3.1 (Minkowski)**.** *Let $\epsilon_1, \epsilon_2 \in (0,1)$ be given. If there are two linearly independent ϵ_1-short and ϵ_2-short vectors in a unimodular lattice in x, then there is a unique rational plane in x with covolume less than 1 which in fact is $\epsilon_1\epsilon_2$-short.*

If there are two different rational planes of covolumes ϵ_1 and ϵ_2 in a unimodular lattice x, then there is a unique primitive vector of length less than 1 which in fact is $\epsilon_1\epsilon_2$-short.

The first part of the lemma follows quickly from the assumption that x is unimodular. The second follows by considering the dual lattice to x. We will use these facts to mark and label certain times in an efficient manner so as to keep the total number of configurations as low as possible.

Some observations

Let us explain how we will use Lemma 3.1. Assume that we have the following situation: There are two linearly independent primitive vectors u, v in a unimodular lattice such that

$$|u| \leq 1/M \text{ and } |\,\mathrm{T}(v)| \leq 1/M.$$

Let $u = (u_1, u_2, u_3)$. It is easy to see that

$$|\,\mathrm{T}(u)| = |(e^{1/2}u_1, e^{1/2}u_2, e^{-1}u_3)| \leq \frac{e^{1/2}}{M}.$$

Assume $M \geq e^{1/2}$. From Lemma 3.1 we have that the plane containing both $\mathrm{T}(u), \mathrm{T}(v)$ has covolume at most $\frac{e^{1/2}}{M^2} \leq \frac{1}{M}$ and it is unique with this property.

The similar situation arises when we have two different planes P, P' which are rational for a unimodular lattice such that

$$|P| \leq 1/M \text{ and } |\,\mathrm{T}(P')| \leq 1/M$$

where $|\cdot|$ means the covolume. Assume $M \geq e$. One can see that $|\,\mathrm{T}(P)| \leq \frac{e}{M}$. Thus, we conclude from Lemma 3.1 that there is a unique vector of length at most $\frac{e}{M^2} \leq \frac{1}{M}$ contained in both planes $\mathrm{T}(P)$ and $\mathrm{T}(P')$.

Marked times

Let us consider a time interval $V = [a, b] \subset [-N, N]$ (for $a, b \in \mathbb{Z}$) with the following properties:

(a) either $a = -N$ (and so $\mathrm{ht}(T^a(x)) \leq M$) or $a > -N$ and $\mathrm{ht}(\mathrm{T}^{a-1}(x)) < M$,

(b) either $b = N$ or $\mathrm{ht}(T^{b+1}(x)) < M$, and

(b) $\mathrm{ht}(\mathrm{T}^n(x)) \geq M$ for all $n \in V$.

We first show how one should inductively pick the marked times for this interval V: We will successively choose vectors and planes in x and mark the time instances with particular labels when these vectors and planes get $1/M$-short on V and when they become big again. At time a we know that there is either a unique plane or a unique vector getting $1/M$-short. Here, uniqueness of either follows from Lemma 3.1. If we have both a unique $1/M$-short plane and vector then we consider whichever stays $1/M$-short longer (say with preference to vectors if again this gives no decision). Assume that we have a unique plane. The case where we start with a unique vector is similar. Mark a by p_1 which is the time when the plane is getting $1/M$-short, and also mark by p_1' the last time in $[a, b]$ when the same plane is still $1/M$-short. If $p_1' = b$ we stop marking. If not, then there is again by Lemma 3.1 a unique $1/M$-short plane or vector at $p_1' + 1$. If it is a $1/M$-short plane then at time $p_1' + 1$ we must have a unique $1/M$-short vector by the discussions in § 3.1. In either case, we have a unique $1/M$-short vector at time $p_1' + 1$. Let us mark by l_1 the instance in $[a, p_1' + 1]$ when this vector is getting $1/M$-short. Also, mark the last time in $[p_1' + 1, b]$ by l_1' for which this vector is still $1/M$-short. If $l_1' = b$ we stop, otherwise at time $l_1' + 1$ there must be a unique $1/M$-short plane or vector. If it is a short vector then we know that there must be a unique plane of covolume at most $1/M$ by the discussions in § 3.1. So, in either case there is a unique $1/M$-short plane at time $l_1' + 1$. So, there is an instance in $[a, l_1' + 1]$ which we mark by p_2 when for the first time this plane is $1/M$-short. Also, mark the instance in $[l_1' + 1, b]$ by p_2' when the plane is $1/M$-short the last time. If $p_2' = b$ we stop here, otherwise we repeat the arguments above and keep marking the time instances in V by l_i, l_i', p_j, p_j' until we hit time b.

Given a positive number N and a lattice $x \in \mathrm{T}^N(X_{\leq M})$ we first consider the disjoint intervals V_i of maximum length with the property as V above. Now start labeling some elements of the sets V_i as explained earlier starting with V_1 and continuing with V_2 etc. always increasing the indices of l_i, l_i', p_i, p_i'.

For any lattice x as above we construct in this way a set of labeled marked times in $[-N, N]$. We denote this set by

$$\mathcal{N}(x) = \mathcal{N}_{[-N,N]}(x) = (\mathcal{L}, \mathcal{L}', \mathcal{P}, \mathcal{P}').$$

Here $\mathcal{L}, \mathcal{L}', \mathcal{P}, \mathcal{P}'$ are subsets in $[-N, N]$ that contain all the labeled marked points l_i, l_i', p_j, p_j' for x respectively. Finally, we let

$$\mathcal{M}_N = \{\mathcal{N}(x) \,:\, x \in X\}$$

be the family of all sets of labeled marked times on the interval $[-N, N]$.

The Estimates

Lemma 3.2 (Noninclusion of marked intervals)**.** *Let $(\mathcal{L}, \mathcal{L}', \mathcal{P}, \mathcal{P}') \in \mathcal{M}_N$ be given. For any q in $\mathcal{L}$ or in $\mathcal{P}$ there is no r in $\mathcal{L}$ or in $\mathcal{P}$ with $q \leq r \leq r' \leq q'$.*

Proof. We have four cases to consider. Let us start with the case that $r = p_i, r' = p_i'$ and $q = p_j, q' = p_j'$ (where $j > i$ as it is in our construction only possible for a later marked interval $[q, q']$ to contain an earlier one). However, by construction the plane P_i that is $1/M$-short at that time we introduce the marked interval $[p_i, p_i']$ (which is either the beginning of the interval V or is the time the earlier short vector stops to be short) is the unique short plane at that time. Hence, it is impossible to have the stated inclusion as the plane P_j (responsible for $[p_j, p'_j]$) would otherwise also be short at that time. The case of two lines is completely similar.

Consider now the case $q = p_j \in \mathcal{P}$ and $r = l_i \in \mathcal{L}$ with $p_j \leq l_i \leq l'_i \leq p'_j$. If $l_i = a$ (and so also $l_i = p_j = a$) is the left end point of interval $V = [a, b]$ in the construction, then we would have marked either l_i, l'_i or p_j, p'_j but not both as we agreed to start by marking the end points of the longer interval (if there is a choice). Hence, we may assume $l_i > a$ and that times l_i, l'_i have been introduced after consideration of a plane with marked times p_k, p'_k satisfying $l_i \leq p_k + 1 \leq l'_i$, in particular $j \neq k$. We now treat two cases depending on whether $p_k \geq l_i$ or not. If $p_k \geq l_i$ then $p_j \leq p_k \leq p'_k \leq p'_j$ which is impossible by the first case. So, assume $p_k < l_i$ then we have two different planes that are $1/M$-short at time l_i. This implies that the vector responsible for the interval $[l_i, l'_i]$ is $1/M^2$-short by Lemma 3.1. However, this shows that the same vector is also $1/M$-short at time $l_i - 1$ for $M \geq e$, which contradicts the choice of l_i. The case of $q = l_i \in \mathcal{L}$ and $r = p_j \in \mathcal{P}$ is similar. $\square$

We would like to know that the cardinality of $\mathcal{M}_N$ can be made small (important in Lemma 2.5) with M large. In other words, for M large we would like to say that $\lim_{N\to\infty} \frac{\log \#\mathcal{M}_N}{2N}$ can be made close to zero. The proof is based on the geometric facts in Lemma 3.1.

Let $\mathcal{N} = (\mathcal{L}, \mathcal{L}', \mathcal{P}, \mathcal{P}') \in \mathcal{M}_N$ and let $\mathcal{L} = \{l_1, l_2, ..., l_m\}$ and $\mathcal{P} = \{p_1, p_2, ..., p_n\}$ be as in the construction of marked times. It is clear from the construction that $l'_i < l'_{i+1}$ for $l'_i, l'_{i+1} \in \mathcal{L}'$. Thus from Lemma 3.2 we conclude that $l_i \leq l_{i+1}$. Hence we have $\mathcal{L} = \{l_1 \leq l_2 \leq ... \leq l_m\}$. Similarly, we must have $\mathcal{P} = \{p_1 \leq p_2 \leq ... \leq p_n\}$. In fact, we have the following.

Lemma 3.3 (Separation of intervals)**.** *For any $i = 1, 2, ..., m-1$ and for any $j =$*

$1, 2, ..., n-1$ *we have*

$$l_{i+1} - l_i > \lfloor \log M \rfloor \text{ and } p_{j+1} - p_j > \lfloor \log M \rfloor.$$

Also,

$$l'_{i+1} - l'_i > \lfloor \log M \rfloor \text{ and } p'_{j+1} - p'_j > \lfloor \log M \rfloor.$$

Proof. For $1/M$-short vectors in $\mathbb{R}^3$, considering their forward trajectories under the action of diagonal flow $(e^{t/2}, e^{t/2}, e^{-t})$, we would like to know the minimum possible amount of time needed for the vector to reach size ≥ 1. Let $v = (v_1, v_2, v_3)$ be a vector of size $\leq 1/M$ which is of size ≥ 1 at time $t \geq 0$. We have

$$1 \leq v_1^2 e^t + v_2^2 e^t + v_3^2 e^{-2t} \leq (v_1^2 + v_2^2 + v_3^2) e^t \leq \frac{e^t}{M^2}.$$

So, we have

$$t \geq \log M^2.$$

Hence, it takes more than $2\lfloor \log M \rfloor$ steps for the vector to reach size ≥ 1. Similarly, for a vector $v = (v_1, v_2, v_3)$ of size ≥ 1, we calculate a lower bound for the time $t \geq 0$ when its trajectory reaches size $\leq 1/M$. We have

$$\frac{1}{M^2} \geq v_1^2 e^t + v_2^2 e^t + v_3^2 e^{-2t} \geq (v_1^2 + v_2^2 + v_3^2) e^{-2t} \geq e^{-2t}.$$

So, we must have $t \geq \log M$ and hence it takes at least $t = \lfloor \log M \rfloor$ steps for the vector to have size $\leq 1/M$.

Now, assume that $l_{i+1} - l_i \leq \lfloor \log M \rfloor$. Let u, v be the vectors in x that are responsible for l_i, l_{i+1} respectively. That is, u, v are $1/M$-short at times l_i, l_{i+1} respectively but not before. Then the above arguments imply that

$$|\,\mathrm{T}^{l_i}(v)| \leq 1 \text{ and } |\,\mathrm{T}^{l_{i+1}}(u)| \leq 1$$

so the plane P containing both u and v is $1/M$-short at times l_i and l_{i+1}. Thus, it is $1/M$-short in $[l_i, l_{i+1}]$ (and so l_i, l_{i+1} are constructed using the same V). From our construction we know that $l'_i < l'_{i+1}$. By Lemma 3.1 the same plane P is $1/M^2$-short on $[l_i, l'_i] \cap [l_{i+1}, l'_{i+1}] = [l_{i+1}, l'_i]$. Hence, P is also $1/M^2$-short at time $l'_i + 1$ (for $M \geq e$) which shows that it is the unique plane that is used to mark points, say p_k, p'_k, after marking l_i, l'_i. Therefore, $p_k \leq l_i \leq l'_i \leq p'_k$ which is a contradiction to Lemma 3.2. The proof of the remaining three cases are very similar to the arguments above and are left to the reader. $\square$

Let us consider the marked points of $\mathcal{L}$ in a subinterval of length $\lfloor \log M \rfloor$ then there could be at most 1 of them. Varying x while restricting ourselves to this interval of length $\lfloor \log M \rfloor$ we see that the number of possibilities to set the marked points in this interval is no more than $\lfloor \log M \rfloor + 1$. For M large, say $M \geq e^4$, we have

$$= \lfloor \log M \rfloor + 1 \leq \lfloor \log M \rfloor^{1.25}.$$

Therefore, there are

$$\leq \lfloor \log M \rfloor^{1.25(\left\lfloor \frac{2N}{\lfloor \log M \rfloor} \right\rfloor + 1)} \ll_M e^{\frac{2.5N \log \lfloor \log M \rfloor}{\lfloor \log M \rfloor}}$$

possible ways of choosing labeled marked points for $\mathcal{L}$ in $[-N, N]$. The same is true for $\mathcal{L}', \mathcal{P}, \mathcal{P}'$. Thus we have shown the following.

Lemma 3.4 (Estimate of $\mathcal{M}_N$). *For $M \geq e^4$ we have*

$$\#\mathcal{M}_N \ll_M e^{\frac{10N \log \lfloor \log M \rfloor}{\lfloor \log M \rfloor}}.$$

Configurations

Before we end this section, we need to point out another technical detail. For our purposes, we want to study a partition element in $X_{\leq M}$ corresponding to a particular set of labeled marked times. Since $X_{\leq M}$ is compact, it is sufficient for us to study an η-neighborhood of some x_0 in this partition. These are the close-by lattices which have the same set of labeled marked times. By knowing that x_0 and x share the same set of labeled marked times, that is $\mathcal{N}(x_0) = \mathcal{N}(x)$, we want to get some restrictions on the position of possible x's in the η-neighborhood of x_0 (see §3.2). However, just knowing that $\mathcal{N}(x_0) = \mathcal{N}(x)$ will not be sufficient for the later argument. Hence, we need to calculate how many possible ways (in terms of vectors and planes) we can have the same labeled marked times. For this purpose, we consider the following configurations.

Vectors

Let l be a marked time in $\mathcal{L} \in \mathcal{N}(x_0)$. Let v_0 be the vector in x_0 that is responsible for l in the construction of marked times for x_0. Let $y = \mathrm{T}^{l-1}(x)$ be in $\mathrm{T}^{l-1}(x_0)B_\eta^{\mathrm{SL}_3(\mathbb{R})}$ with $\mathcal{N}(x) = \mathcal{N}(x_0)$ and v in x that is responsible for l in the construction of marked times for x. Let $v' \in x_0$ be such that $\mathrm{T}^{l-1}(v')g = \mathrm{T}^{l-1}(v)$ for some $g \in B_\eta^{\mathrm{SL}_3(\mathbb{R})}$ with $y = \mathrm{T}^{l-1}(x_0)g$. We want to know how many choices for v' are realized by the various choices of x as above.

Lemma 3.5. *Let $\mathcal{N}(x_0)$ be given. Also, let $l \in \mathcal{L}$ and $v_0 \in x_0$ that is responsible for l. Let x be such that $\mathcal{N}(x) = \mathcal{N}(x_0)$ and $\mathrm{T}^{l-1}(x) = \mathrm{T}^{l-1}(x_0)g$ for $g \in B_\eta^{\mathrm{SL}_3(\mathbb{R})}$. Assume also that $v \in x$ is responsible for l.*
If l is the left end point of the interval V then we must have $\mathrm{T}^{l-1}(v) = \pm\,\mathrm{T}^{l-1}(v_0)g$. Otherwise, there are $p \in \mathcal{P}$ and $p' \in \mathcal{P}'$ with $p \leq l-1 \leq p'$. In this case, there are at most $\ll \min\{e^{(p'-l)}, e^{(l-p)/2}\}$ primitive vectors w' in x_0 for which we might have $\mathrm{T}^{l-1}(v) = \mathrm{T}^{l-1}(w')g$.

Proof. To simplify the notation below we set $w_0 = \mathrm{T}^{l-1}(v_0) \in \mathrm{T}^{l-1}(x_0)$, $w = \mathrm{T}^{l-1}(v) \in y$, and $w' = \mathrm{T}^{l-1}(v') = wg \in \mathrm{T}^{l-1}(x_0)$.

We have

$$\frac{1}{M} \leq |w| \leq \frac{e}{M},$$

and so

$$\begin{aligned}|w'| &\leq |w' - w| + |w| \\ &\leq |w|\|g^{-1} - 1\| + |w| \\ &\leq e(1 + 2\eta)/M.\end{aligned}$$

Also,

$$\begin{aligned}|w'| &\geq |w| - |w - w'| \\ &\geq (1 - 2\eta)/M.\end{aligned}$$

Together

$$\frac{1 - 2\eta}{M} \leq |w'| \leq \frac{e(1 + 2\eta)}{M}. \tag{3.1.2}$$

Assume first that $l = a$ is the left end point of the interval $V = [a, b]$ in the construction of marked times. In this case, w' and w_0 lie in the same line in $\mathbb{R}^3$. Otherwise, if they were linearly independent then the plane containing both would be $e^2(1 + 2\eta)/M^2$-short by Lemma 3.1. For $M \geq 3e^2$ this is a contradiction to the assumption that $l = a$. Since we only consider primitive vectors we only have the choice of $w' = \pm w_0$.

Now, assume that l is not the left end point of the interval V. Then, there is a plane P in x_0 associated to p, p' with $p \leq l - 1 \leq p'$ such that

$$\begin{aligned}&|\,\mathrm{T}^{p-1}(P)| \geq 1/M \text{ and } |\,\mathrm{T}^{p'+1}(P)| > 1/M \\ &|\,\mathrm{T}^{k}(P)| \leq 1/M \text{ for } k \in [p, p'].\end{aligned}$$

Let us calculate how many possibilities there are for $w' \in \mathrm{T}^{l-1}(x_0)$. By (3.1.2) w' is in the plane $\mathrm{T}^{l-1}(P)$ of covolume < 1 w.r.t. $\mathrm{T}^{l-1}(x_0)$ since $\mathrm{T}^{l-1}(x_0)$ is unimodular. Since

$$\frac{1}{M} < |\,\mathrm{T}^{p'+1}(P)| \text{ and } \frac{1}{M} \leq |\,\mathrm{T}^{p-1}(P)|,$$

we get

$$\max\left\{\frac{e^{-(p'-l+2)}}{M}, \frac{e^{-(l-p)/2}}{M}\right\} \leq |\,\mathrm{T}^{l-1}(P)|$$

(see § 3.1 for the action of T on planes). We note that the ball of radius r contains at most $\ll \max\{\frac{r^2}{A}, 1\}$ primitive vectors of a lattice in $\mathbb{R}^2$ of covolume A. This follows since in the case of r being smaller than the second successive minima we have at most 2 primitive vectors, and if r is bigger, then area considerations give $\ll \frac{r^2}{A}$ many lattice points in the r-ball.

We apply this for $A = |\,\mathrm{T}^{l-1}(P)| \geq \max\left\{\frac{e^{-(p'-l+2)}}{M}, \frac{e^{-(l-p)/2}}{M}\right\}$ and $r = \frac{(1+2\eta)e}{M}$ where

$$\frac{r^2}{A} = \frac{(1+2\eta)^2 e^2/M^2}{\max\left\{\frac{e^{-(p'-l+2)}}{M}, \frac{e^{-(l-p)/2}}{M}\right\}} \ll \min\{e^{(p'-l)}, e^{(l-p)/2}\},$$

which proves the lemma. □

Planes

Let p be a marked time in $\mathcal{P} \in \mathcal{N}(x_0)$. Let P_0 be a plane in $T^{p-1}(x_0)$ that is responsible for p in the construction of marked times for x_0. Let $y = \mathrm{T}^{l-1}(x)$ be in $\mathrm{T}^{l-1}(x_0)B_\eta^{\mathrm{SL}_3(\mathbb{R})}$ with $\mathcal{N}(x) = \mathcal{N}(x_0)$ and P in x that is responsible for l in the construction of marked times for x. Let P' be a plane that is rational w.r.t. x_0 such that $\mathrm{T}^{p-1}(P')g = \mathrm{T}^{p-1}(P)$ for some $g \in B_\eta^{\mathrm{SL}_3(\mathbb{R})}$ with $y = \mathrm{T}^{l-1}(x_0)g$. We want to know how many choices for P' are realized by the various choices of x as above. We have two cases.

Lemma 3.6. *Let $\mathcal{N}(x_0)$ be given. Also, let $p \in \mathcal{L}$ and $P_0 \in x_0$ that is responsible for p. Let x be such that $\mathcal{N}(x) = \mathcal{N}(x_0)$ and $\mathrm{T}^{p-1}(x) = \mathrm{T}^{p-1}(x_0)g$ for $g \in B_\eta^{\mathrm{SL}_3(\mathbb{R})}$. Assume also that P w.r.t. x that is responsible for p.*

If p is the left end point of the interval V then we must have $\mathrm{T}^{p-1}(P) = \mathrm{T}^{p-1}(P_0)g$. Otherwise, there are $l \in \mathcal{P}$ and $l' \in \mathcal{P}'$ with $l \leq p-1 \leq l'$. In this case, there are *at most* $\ll \min\{e^{(l'-p)/2}, e^{p-l}\}$ *rational planes* P' *w.r.t.* x_0 *for which we might have* $\mathrm{T}^{p-1}(P) = \mathrm{T}^{p-1}(P')g$.

Proof. Assume first that $p = a$ is the left end point of the interval $V = [a, b]$ in the construction of marked times. Arguing as above we can show that in this case there is no choice.

Now, assume that p is not the left end point of V. Then, there is a vector v in x_0 associated to marked times l, l' with $l \le p-1 \le l'$ such that

$$|\,\mathrm{T}^{l-1}\, v| \ge 1/M \text{ and } |\,\mathrm{T}^{l'+1}(v)| > 1/M$$
$$|\,\mathrm{T}^{k}(v)| \le 1/M \text{ for } k \in [l, l'].$$

On the space $\bigwedge^2 \mathbb{R}^3$, vectors correspond to planes in $\mathbb{R}^3$ and planes correspond to vectors in $\mathbb{R}^3$. Hence, we can reduce the current case to the case of a vector followed by a plane. However, we have a different action on $\bigwedge^2 \mathbb{R}^3$ (see § 3.1). Similar arguments as above show that there are

$$\ll \min\{e^{(l'-p)/2}, e^{p-l}\}$$

possibilities for P'. □

3.2 Main Proposition and Restrictions

We recall the Bowen N-balls defined earlier. Let $B_\eta^{\mathrm{SL}_3(\mathbb{R})}(1)$ be a ball in $\mathrm{SL}_3(\mathbb{R})$ of radius η with center at 1. Fix $\eta > 0$ small enough so that $B_\eta^{\mathrm{SL}_3(\mathbb{R})}(1)$ is an injective image under the exponential map of a neighborhood of 0 in the Lie algebra. Recall that a Bowen N-ball is the translate xB_N for some $x \in X$ of

$$B_N = \bigcap_{n=-N}^{N} \alpha^{-n} B_\eta^{\mathrm{SL}_3(\mathbb{R})} \alpha^n.$$

Fix a height $M \geq 1$. Let $N \geq 1$ and consider $\mathcal{N} = \mathcal{N}(x_0) \in \mathcal{M}_N$. Let V be the subset (not necessarily an interval) of $[-N, N]$ such that for any $n \in [-N, N]$, $n \in V$ if and only if there is a $1/M$-short plane or a $1/M$-short vector at time n. Define the set

$$Z(\mathcal{N}) := \{x \in \mathrm{T}^N(X_{\leq M}) \,|\, \mathcal{N}(x) = \mathcal{N}\}.$$

Now, we state the main proposition.

Proposition 3.7. *There exists a constant $c_0 > 0$, independent of M, such that the set $Z(\mathcal{N})$ can be covered by $\ll_M e^{6N-|V|} c_0^{\frac{2N}{\lfloor \log M \rfloor}}$ Bowen N-balls.*

In the proof of the main Theorem 1.3 we will consider

$$\lim_{N \to \infty} \frac{\log \# Z(\mathcal{N})}{2N}.$$

Thus, in this limit, the term arising from $c_0^{\frac{18N}{\lfloor \log M \rfloor}}$ can be made small for M large since c_0 does not depend on M. So, our main consideration is the $e^{6N-|V|}$ factor. On the other hand, it is easy to see that the set $Z(\mathcal{N})$ can be covered by $\ll e^{6N}$ many Bowen N-balls. But this does not give any meaningful conclusion. Therefore, $e^{-|V|}$ is the factor appearing in Proposition 3.7 that leads to the conclusion of the main Theorem 1.3.

In proving Proposition 3.7, we will make use of the lemmas below which give the restrictions needed in order to get the drop in the number of Bowen N-balls to cover the set $Z(\mathcal{N})$.

Restrictions of perturbations

Perturbations of vectors

Let $v = (v_1, v_2, v_3)$ be a vector in $\mathbb{R}^3$.

Lemma 3.8. *For a vector v of size $\geq 1/M$, if its trajectory stays $1/M$-short in the time interval $[1, S]$ then we must have $\frac{v_1^2+v_2^2}{v_3^2} < 2e^{-S}$.*

Proof. By assumption we have

$$v_1^2 + v_2^2 + v_3^2 \geq \frac{1}{M^2} \geq v_1^2 e^S + v_2^2 e^S + v_3^2 e^{-2S}.$$

This simplifies to

$$v_3^2(1 - e^{-2S}) > (v_1^2 + v_2^2)(e^S - 1).$$

Hence, $v_3 \neq 0$ and we have

$$\frac{v_1^2 + v_2^2}{v_3^2} \leq \frac{1 - e^{-2S}}{e^S - 1} < \frac{1}{e^S - 1} < 2e^{-S}$$

□

We would like to get restrictions for the vectors which are close to the vector v and whose trajectories behave as v on the time interval $[0, S]$. So, let $u = (u_1, u_2, u_3)$ be a vector in $\mathbb{R}^3$ with $u = vg$ for some $g \in B_\eta^{\mathrm{SL}_3(\mathbb{R})}$ such that $|u| \geq 1/M$ and that its forward trajectory stays $1/M$-short in the time interval $[1, S]$.

Let us first assume $g = \begin{pmatrix} 1 & & \\ & 1 & \\ -t_1 & -t_2 & 1 \end{pmatrix} \in B_\eta^{U^+}$ so that

$$\begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} \begin{pmatrix} 1 & & \\ & 1 & \\ -t_1 & -t_2 & 1 \end{pmatrix}.$$

From Lemma 3.8 we know that $\frac{u_1^2 + u_2^2}{u_3^2} < 2e^{-S}$. So,

$$\frac{(v_1 - v_3 t_1)^2 + (v_2 - v_3 t_2)^2}{v_3^2} < 2e^{-S}.$$

We are interested in possible restrictions on t_j's since they belong to the unstable horospherical subgroup of $\mathrm{SL}_3(\mathbb{R})$ under conjugation by $a = \mathrm{diag}(e^{1/2}, e^{1/2}, e^{-1})$. Simplifying the left hand side, we obtain

$$(\frac{v_1}{v_3} - t_1)^2 + (\frac{v_2}{v_3} - t_2)^2 < 2e^{-S}.$$

We also know $\frac{v_1^2}{v_3^2}+\frac{v_2^2}{v_3^2}<2e^{-S}$. Together with the triangular inequality, we get

$$t_1^2+t_2^2<(\sqrt{2e^{-S}}+\sqrt{2e^{-S}})^2=8e^{-S}.$$

In general, we have

$$g=\begin{pmatrix}1 & & \\ & 1 & \\ -t_1 & -t_2 & 1\end{pmatrix}\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ 0 & 0 & a_{33}\end{pmatrix}\in B_\eta^{\mathrm{SL}_3(\mathbb{R})}(1).$$

In this case, we still claim that

$$t_1^2+t_2^2<8e^{-S}.$$

Let

$$w=\begin{pmatrix}w_1 & w_2 & w_3\end{pmatrix}=\begin{pmatrix}v_1 & v_2 & v_3\end{pmatrix}\begin{pmatrix}1 & & \\ & 1 & \\ -t_1 & -t_2 & 1\end{pmatrix}$$

so that

$$u=vg=w\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ 0 & 0 & a_{33}\end{pmatrix}. \tag{3.2.1}$$

We observe

$$\mathrm{T}^S(u)=\mathrm{T}^S(w)\begin{pmatrix}a_{11} & a_{12} & a_{13}e^{-3S/2}\\ a_{21} & a_{22} & a_{23}e^{-3S/2}\\ 0 & 0 & a_{33}\end{pmatrix},$$

so that $\mathrm{T}^S(u)\in\mathrm{T}^S(w)B_\eta^{\mathrm{SL}_3(\mathbb{R})}(1)$ and $|\,\mathrm{T}^S(u)-\mathrm{T}^S(w)|<2\eta|\,\mathrm{T}^S(u)|$. Hence, $|\,\mathrm{T}^S(u)|<1/M$ implies

$$|\,\mathrm{T}^S(w)|\le|\,\mathrm{T}^S(u)|+|\,\mathrm{T}^S(u)-\mathrm{T}^S(w)|<\frac{1+2\eta}{M}.$$

On the other hand, since $g\in B_\eta^{\mathrm{SL}_3(\mathbb{R})}$ we have

$$|w|\ge|u|-|u-w|>\frac{1-2\eta}{M}$$

Together we get

$$\frac{|w|}{1-2\eta} > \frac{|\operatorname{T}^S(w)|}{1+2\eta}.$$

Now, arguing as in the proof of Lemma 3.8, for sufficiently small $\eta > 0$, we obtain

$$\frac{w_1^2 + w_2^2}{w_3^2} < 2e^{-S}.$$

Hence, we are in the previous case with u replaced by w. So, we have $t_1^2 + t_2^2 < 8e^{-S}$ which proves the claim. We have shown the following.

Lemma 3.9. *Let v, u be vectors in $\mathbb{R}^3$ with sizes $\geq 1/M$ whose trajectories in $[1, S]$ stay $1/M$-short. Assume that $u = vg$ with $g \in B_\eta^{\mathrm{SL}_3(\mathbb{R})}(1)$ and that the notation is as in* (3.2.1). *Then*

$$t_1^2 + t_2^2 \leq 8e^{-S}.$$

Lemma 3.10. *Consider the ball $t_1^2 + t_2^2 \leq 8e^{-S}$ on $[-2\eta, 2\eta]^2$ and let us divide $[-2\eta, 2\eta]^2$ into small squares of side length $\frac{1}{2}\eta e^{-3S'/2}$. Then there exists a constant $c > 0$, independent of M, so that there are $\ll \max\{1, e^{3S'-S}\}$ small squares that intersect with the ball $t_1^2 + t_2^2 \leq 8e^{-S}$.*

Proof. Note that $t_1^2 + t_2^2 \leq 8e^{-S}$ defines a ball with diameter $2\sqrt{8}e^{-S/2}$. If $\frac{1}{2}\eta e^{-3S'/2} \geq 2\sqrt{8}e^{-S/2}$ then there are 4 squares that intersects the ball. Otherwise (which makes $3S' - S$ bounded below), there can be at most $\ll \frac{(e^{-S/2})^2}{(e^{-3S'/2})^2} = e^{3S'-S}$ small squares that intersect with the given ball. $\square$

What Lemma 3.9 and Lemma 3.10 say is the following:

Consider a neighborhood $O = x_0 B_{\eta/2}^{U^+} B_{\eta/2}^{U^- C}$ of X where as before U^+, U^-, and C are unstable, stable, and centralizer subgroups of $\mathrm{SL}_3(\mathbb{R})$ with respect to α, respectively. If we partition the square of length 2η in $B_{\eta/2}^{U^+}(1)$ into small squares of side lengths $\eta e^{-3S'/2}$, then we have $\ll \lceil \frac{2\eta}{\eta e^{-3S'/2}} \rceil^2 \ll \lceil e^{3S'/2} \rceil^2$ many elements in this partition. Now, assume that there is a vector $v \in x_0$ with $|v| \geq 1/M$ that stays $1/M$-short in

$[1, S]$ and consider a set of lattices $x = x_0 g$ in O with the property that the vector $w = vg$ in x behaves as v in $[0, S]$. Then the above two lemmas say that this set is contained in $\leq c_0 e^{3S'-S}$ many partition elements (small squares). Hence, in the proof of Proposition 3.7, instead of $\leq c_0 \lceil e^{3S'/2} \rceil^2$ many Bowen balls we will only consider $\leq c_0 e^{3S'-S}$ many of them and this (together with the case below) will give us the drop in the exponent as appeared in Proposition 3.7.

Perturbations of planes

Assume that for a lattice $x \in X$ there is a plane P with

$$|P| \geq 1/M \text{ and } |\,\mathrm{T}^k(P)| \leq 1/M \text{ for } k \in [1, S].$$

Let u, v be generators of P with $|P| = |u \wedge v|$. So we have

$$|u \wedge v| \geq 1/M \geq |\,\mathrm{T}^S(u \wedge v)|.$$

Thus, substituting $a = u_2 v_3 - u_3 v_2, b = u_3 v_1 - u_1 v_3, c = u_1 v_2 - u_2 v_1$ (cf. 3.1.1) we obtain

$$a^2 + b^2 + c^2 \geq a^2 e^{-S} + b^2 e^{-S} + c^2 e^{2S},$$

which gives

$$\frac{c^2}{a^2 + b^2} \leq \frac{1 - e^{-S}}{e^{2S} - 1} = e^{-2S} \frac{1 - e^{-S}}{1 - e^{-2S}} = e^{-2S} \frac{1}{1 + e^{-S}} < e^{-2S}.$$

Assume $x' = xg$ for some $g \in B_\eta^{\mathrm{SL}_3(\mathbb{R})}$. For now, let us assume that

$$g = \begin{pmatrix} 1 & & \\ & 1 & \\ t_1 & t_2 & 1 \end{pmatrix}.$$

Let $u', v' \in x'$ be such that

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} u'_1 & u'_2 & u'_3 \\ v'_1 & v'_2 & v'_3 \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix} \begin{pmatrix} 1 & & \\ & 1 & \\ t_1 & t_2 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} u_1 + t_1 u_3 & u_2 + t_2 u_3 & u_3 \\ v_1 + t_1 v_3 & v_2 + t_2 v_3 & v_3 \end{pmatrix}.$$

We let $a' = u'_2 v'_3 - u'_3 v'_2 = (u_2 + t_2 u_3) v_3 - u_3 (v_2 + t_2 v_3)$ and hence $a' = a$. Similarly, $b' = u'_3 v'_1 - u'_1 v'_3 = b$ and let

$$c' = u'_1 v'_2 - u'_2 v'_1 =$$
$$(u_1 + t_1 u_3)(v_2 + t_2 v_3) - (u_2 + t_2 u_3)(v_1 + t_1 v_3) = c - a t_1 - b t_2.$$

Now, assume that

$$|u' \wedge v'| \geq 1/M \text{ and } |\operatorname{T}^k (u' \wedge v')| \leq 1/M \text{ for } k \in [1, S]$$

which by the above implies

$$\frac{c'^2}{a'^2 + b'^2} = \frac{(c - a t_1 - b t_2)^2}{a^2 + b^2} < e^{-2S}.$$

For a general $g \in B_\eta^{\mathrm{SL}_3(\mathbb{R})}$ we would like to obtain a similar equation. Let us write g as

$$g = \begin{pmatrix} 1 & & \\ & 1 & \\ t_1 & t_2 & 1 \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ 0 & 0 & g_{33} \end{pmatrix}. \tag{3.2.2}$$

Then we have

$$\operatorname{T}^l(x') = \operatorname{T}^l(xg) = \operatorname{T}^l \left(x \begin{pmatrix} 1 & & \\ & 1 & \\ t_1 & t_2 & 1 \end{pmatrix} \right) \begin{pmatrix} g_{11} & g_{12} & g_{13} e^{-\frac{3}{2} l} \\ g_{21} & g_{22} & g_{23} e^{-\frac{3}{2} l} \\ 0 & 0 & g_{33} \end{pmatrix}.$$

Hence the forward trajectories of x' and $x \begin{pmatrix} 1 & & \\ & 1 & \\ t_1 & t_2 & 1 \end{pmatrix}$ stay $\ll \eta$ close. Thus, we have

$$\frac{(c - at_1 - bt_2)^2}{a^2 + b^2} \ll e^{-2S}.$$

From the triangular inequality we obtain

$$\frac{(at_1 + bt_2)^2}{a^2 + b^2} \ll e^{-2S}.$$

Let $C > 0$ be the constant that appeared in the last inequality.

Lemma 3.11. *Let P, P' be planes in $\mathbb{R}^3$ with covolume $\geq 1/M$ whose trajectories in $[1, S]$ stay $1/M$-short and assume that $P' = Pg$ for some $g \in B_\eta^{\mathrm{SL}_3(\mathbb{R})}$, then for some a, b (dependent on P) we must have in the notation of* (3.2.2) *that*

$$\frac{(at_1 + bt_2)^2}{a^2 + b^2} \leq Ce^{-2S}.$$

We note that the inequality above describes a neighborhood of the line in $\mathbb{R}^2$ defined by the normal vector (a, b) of width $2\sqrt{C}e^{-s}$.

Lemma 3.12. *Consider the set defined by $\frac{(at_1+bt_2)^2}{a^2+b^2} \leq Ce^{-2S}$ on $[-2\eta, 2\eta]^2$ and let us divide $[-2\eta, 2\eta]^2$ into small squares of side length $\frac{1}{2}\eta e^{-3S'/2}$. Then there are $\ll$* $\max\{e^{3S'/2}, e^{3S'-S}\}$ *small squares that intersect with the region $\frac{(at_1+bt_2)^2}{a^2+b^2} \leq Ce^{-2S}$.*

Proof. The type of estimate depends on whether the side length $\frac{1}{2}\eta e^{-3S'/2}$ of the squares is smaller or bigger than the width $2\sqrt{C}e^{-S}$ of the neighborhood. We need to calculate the length and the area of the region R given by

$$|at_1 + bt_2| \leq \sqrt{C(a^2 + b^2)}e^{-S}$$

restricted to $[-2\eta, 2\eta]^2$. As mentioned earlier, the inequality above describes a $\sqrt{C}e^{-S}$-neighborhood of the line $at_1 + bt_2 = 0$. The length of the segment of this line in $[-2\eta, 2\eta]^2$ is at most $4\sqrt{2}\eta$, so that the area of R is $\leq 4\sqrt{2C}\eta e^{-S}$.

If $\sqrt{C}e^{-S} \leq \frac{1}{2}\eta e^{-3S'/2}$ then there are $\ll \frac{\eta}{\eta e^{-3S'/2}} = e^{3S'/2}$ many intersections. Otherwise, there are at most

$$\ll \frac{\sqrt{C}\eta e^{-S}}{\eta^2 e^{-3S'}} \ll e^{3S'-S}$$

small squares that intersect the region R. $\square$

Proof of Main Proposition

Proof of Proposition 3.7. By taking the images under a positive power of T it suffices to consider forward trajectories and the following reformulated problem:

Let $V \subset [0, N-1]$ and $x_0 \in X_{\leq M}$ be such that

$$n \in V \text{ if and only if } \mathrm{T}^n(x_0) \in X_{\geq M}.$$

Also let $\mathcal{N} = \mathcal{N}_{[0,N-1]}(x_0)$ be the marked times for x_0 (defined similarly to $\mathcal{N}_{[-N,N]}$ as in § 3.1).

We claim that

$$Z^+_{\leq M} = \{x \in X_{\leq M} : \mathcal{N}_{[0,N-1]}(x) = \mathcal{N}\}$$

can be covered by $\ll_M e^{3N-|V|} c_0^{\frac{9N}{[\log M]}}$ forward Bowen N-balls xB^+_N defined by

$$B^+_N = \bigcap_{n=0}^{N-1} \alpha^n B^{\mathrm{SL}_3(\mathbb{R})}_\eta \alpha^{-n}.$$

Since $X_{\leq M}$ is compact and since we allow the implicit constant above to depend on M it suffices to prove the following:

As before let $U^+ = \begin{pmatrix} 1 & & \\ & 1 & \\ * & * & 1 \end{pmatrix}$ and $U^- = \begin{pmatrix} 1 & & * \\ & 1 & * \\ & & 1 \end{pmatrix}$ be unstable and stable horospherical subgroups of $\mathrm{SL}_3(\mathbb{R})$ under the conjugation by α respectively, and let $C = \begin{pmatrix} * & * & \\ * & * & \\ & & * \end{pmatrix}$ be the centralizer of α in $\mathrm{SL}_3(\mathbb{R})$. Given $x_0 \in X_{\leq M}$ and a neighborhood

$$O = x_0 D_{\eta/2}^{U^+} B_{\eta/2}^{U^- C}$$

of x_0, where $D_{\eta/2}^{U^+}$ is the $\eta/2$-neighborhood of 1 in U^+ (identified with $\mathbb{R}^2$) w.r.t. maximum norm. Then we claim that the set

$$Z_O^+ = \{x \in O : \mathcal{N}_{[0,N-1]}(x) = \mathcal{N}\}$$

can be covered by $\ll e^{3N-|V|} c_0^{\frac{9N}{[\log M]}}$ forward Bowen N-balls.

If we apply T^n to O we get a neighborhood of $\mathrm{T}^n(x_0)$ for which the U^+-part is stretched by the factor $e^{3n/2}$, while the second part is still in $B_{\eta/2}^{U^- C}$. By breaking the U^+-part into $\lceil e^{3n/2} \rceil^2$ sets of the form $u_i^+ D_{\eta/2}^{U^+}$ for various $u_i^+ \in U^+$ we can write $\mathrm{T}^n(O)$ as a union of $\lceil e^{3n/2} \rceil^2$ sets of the form

$$\mathrm{T}^n(x_0) u^+ D_{\eta/2}^{U^+}(1) \alpha^{-n} B_{\eta/2}^{U^- C} \alpha^n.$$

Hence we got similar neighborhoods as before. If we take the pre-image under T^n of this set, we obtain the set

$$\mathrm{T}^{-n}(\mathrm{T}^n(x_0) u^+) \alpha^n D_{\eta/2}^{U^+} \alpha^{-n} B_{\eta/2}^{U^- C}(1).$$

Notice that $\mathrm{T}^{-n}(\mathrm{T}^n(x_0) u^+) \alpha^n D_{\eta/2}^{U^+} \alpha^{-n} B_{\eta/2}^{U^- C}(1)$ is contained in the forward Bowen n-ball $\mathrm{T}^{-n}(\mathrm{T}^n(x_0) u_i^+) B_n^+$. Indeed we may assume $D_\epsilon \subset B_\epsilon$ and so for $0 \leq k < n$ we have

$$\alpha^{-k}(\alpha^n D_{\eta/2}^{U^+} \alpha^{-n}) \alpha^k \;\subset\; \alpha^{n-k} B_{\eta/2}^{U^+} \alpha^{-(n-k)} \alpha^{-k} B_{\eta/2}^{U^- C} \alpha^k \;\subset\; B_{\eta/2}^{U^+} B_{\eta/2}^{U^- C} \;\subset\; B_\eta^{\mathrm{SL}_3(\mathbb{R})}.$$

We would like to reduce the number of u_i^+'s, so that we do not have to use all $\lceil e^{3n/2} \rceil^2$ forward Bowen n-balls to cover the set Z_O^+.

We can decompose V into disjoint intervals V_j's where $j \in \{1, 2, ..., m\}$ with m as small as possible. We note here that $m \leq |\mathcal{L}| + |\mathcal{P}|$ so that from Lemma 3.3 we obtain

$$m \leq \frac{N}{\lfloor \log M \rfloor} + 1 \tag{3.2.3}$$

Now, write $[0, N-1] \setminus V = W_1 \cup W_2 \cup ... \cup W_l$ where W_i's are maximal intervals. A bound similar to (3.2.3) also holds for l.

We will consider intervals V_j and W_i in their respective order in $[0, N-1]$. At each stage we will divide any of the sets obtained earlier into $\lceil e^{3|V_j|/2} \rceil^2$- or $\lceil e^{3|W_i|/2} \rceil^2$-many sets, and in the case of V_j show that we do not have to keep all of them. We inductively prove the following:

For $K \leq N$ such that $[0, K] = V_1 \cup V_2 \cup ... \cup V_n \cup W_1 \cup W_2 \cup ... \cup W_{n'}$ the set Z_O^+ can be covered by $\ll e^{3K} e^{-(|V_1| + ... + |V_n|)} c_0^{4\frac{|V_1| + ... + |V_n|}{\lfloor \log M \rfloor} + 4n + n'}$ many pre-images under T^K of sets of the form

$$\mathrm{T}^K(x_0) u^+ D_{\eta/2}^{U^+} \alpha^{-K} B_{\eta/2}^{U^- C} \alpha^K$$

and hence can be covered by $\ll e^{3K} e^{-(|V_1| + ... + |V_n|)} c_0^{4\frac{|V_1| + ... + |V_n|}{\lfloor \log M \rfloor} + 4n + n'}$ many forward Bowen K-balls. When $K = N$ we obtain the proposition.

For the inductive step, if the next interval is $W_{n'+1}$ then after dividing the set $\mathrm{T}^K(x_0) u^+ D_{\eta/2}^{U^+} \alpha^{-K} B_{\eta/2}^{U^- C} \alpha^K$ into $\lceil e^{3|W_{n'+1}|/2} \rceil^2 \leq 4 e^{3|W_{n'+1}|}$ many sets of the form

$$\mathrm{T}^{K + |W_{n'+1}|}(x_0) u^+ D_{\eta/2}^{U^+} \alpha^{-K - |W_{n'+1}|} B_{\eta/2}^{U^- C}(1) \alpha^{K + |W_{n'+1}|}$$

we just consider all of them, and hence have that Z_O^+ can be covered by

$$\ll e^{3(K + |W_{n'+1}|)} e^{-(|V_1| + ... + |V_n|)} c_0^{4\frac{|V_1| + ... + |V_n|}{\lfloor \log M \rfloor} + 4n + n' + 1}$$

many forward Bowen $K + |W_{n'+1}|$-balls (assuming $c_0 \geq 4$).

So, assume that the next time interval is $V_{n+1} = [K+1, K+R]$. Pick one of the sets obtained in an earlier step and denote it by

$$Y = \mathrm{T}^K(x_0) u^+ D_{\eta/2}^{U^+} \alpha^{-K} B_{\eta/2}^{U^- C} \alpha^K.$$

We are interested in lattices x in $Y \cap X_{\leq M}$ such that

$$\mathcal{N}_{[0,R]}(x) = \mathcal{N}_{[0,R]}(\mathrm{T}^K(x_0)) = \{\mathcal{L}, \mathcal{L}', \mathcal{P}, \mathcal{P}'\}.$$

We have

$$\mathcal{L} = \{l_1 < l_2 < ... < l_k\},\ \mathcal{L}' = \{l'_1 < l'_2 < ... < l'_k\}$$

and

$$\mathcal{P} = \{p_1 < p_2 < ... < p_{k'}\},\ \mathcal{P}' = \{p'_1 < p'_2 < ... < p'_{k'}\}$$

for some $k, k' \geq 0$. Without loss of generality we can assume that $K + 1 = l_1$. We note that

$$K + 1 = l_1 < p_1 < l_2 < p_2 < ... < \min\{l_k, p_{k'}\} < \max\{l_k, p_{k'}\}.$$

This easily follows from the construction of labeled marked times together with Lemma 3.2. So, we can divide the interval V_{n+1} into subintervals

$$[l_1, p_1], [p_1, l_2], ..., [\min\{l_k, p_{k'}\}, \max\{l_k, p_{k'}\}], [\max\{l_k, p_{k'}\}, K + R].$$

We consider each of the (overlapping) intervals in their respective order.

Let us define c_0 to be the maximum of the implicit constants that appeared in the conclusions of Lemma 3.5, Lemma 3.6, Lemma 3.10, and Lemma 3.12.

We would like to apply Lemma 3.10 and Lemma 3.12 to obtain a smaller number of forward Bowen $K + |V_{n+1}|$-balls to cover the set Y. Assume for example that there is a vector v in a lattice x that is getting $1/M$-short and staying short in some time interval, also assume that there is a vector u in a lattice xg for some $g \in B_\eta^{\mathrm{SL}_3(\mathbb{R})}$ which behaves the same as v. However, we can apply Lemma 3.10 only if we know that $u = xg$. Thus, it is necessary to know how many vectors w' there are in x for which $u = w'g$ for some g. This is handled by Lemma 3.5. Similar situation arises when we want to apply Lemma 3.12, and this case we first need to use Lemma 3.6.

Let us start with the interval $[l_1, p_1]$. Let us divide the set $Y \cap X_{\leq M}$ into $\lceil e^{3(p_1-l_1)/2}\rceil^2$ small sets by partitioning the set $D^{U^+}_{\eta/2}$ in the definition of Y as we did before. Since l_1 is the left end point of V_{n+1} we see that the assumptions of Lemma 3.9 are satisfied in the sense that if there is a lattice $\mathrm{T}^{l_1-1}(x_0)g$ which has the the same set of marked points as $\mathrm{T}^{l-1}(x_0)$ for some $g \in B^{\mathrm{SL}_3(\mathbb{R})}_{\eta}$, then there are unique vectors $v \in \mathrm{T}^{l_1-1}(x_0)$ and $u = vg \in \mathrm{T}^{l_1-1}(x_0)g$ which are of size $\leq 1/M$ and stay $1/M$-short in $[l_1, l_1']$. (cf. Lemma 3.5). Now, from Lemma 3.9 and Lemma 3.10 with $S' = p_1 - l_1$ and $S = l_1' - l_1$ we see that we only need to consider

$$\leq c_0 \max\{1, e^{3(p_1-l_1)-(l_1'-l_1)}\} =: N_1 \tag{3.2.4}$$

of these $\lceil e^{3(p_1-l_1)/2}\rceil^2$ sets (see the discussion at the end of § 3.2). Thus, we obtain sets of the form

$$\mathrm{T}^{p_1}(x_0)u^+ D^{U^+}_{\eta/2}\alpha^{-p_1} B^{U^-C}_{\eta/2}\alpha^{p_1}.$$

Now, let us consider the next interval $[p_1, l_2]$. Divide the sets obtained earlier into $\lceil e^{3(l_2-p_1)/2}\rceil^2$ subsets for which the U^+-component is of the from $u^+ D^{U^+}_{e^{-3(l_2-p_1)/2}\eta/2}$. We would like to apply Lemma 3.12. However, Lemma 3.12 concerns itself with the restrictions on g arising from common behaviors of two planes $P, P' = Pg$ and we only know the common behavior of the lattices. Moreover, if P_0 (resp. P) is the plane that is rational w.r.t. $\mathrm{T}^{p_1}(x_0)$ (resp. $\mathrm{T}^{p_1}(x_0)g$) which is responsible for the marking of $[p_1, p_1']$ then we do not necessarily know that $P = P_0 g$. On the other hand, we see from Lemma 3.6 that there are $\leq c_0 \min\{e^{(l_1'-p_1)/2}, e^{p_1-l_1}\}$ choices of planes P' that are rational w.r.t. $\mathrm{T}^{p_1}(x_0)$ for which we could possibly have $P = P'g$. For each choice we can apply Lemma 3.12 with $S' = l_2 - p_1$ and $S = p_1' - p_1$. Thus, for each choice we need to consider only $\leq c_0 \max\{e^{3(l_2-p_1)/2}, e^{3(l_2-p_1)-(p_1'-p_1)}\}$ of the $\lceil e^{3(l_2-p_1)/2}\rceil^2$ subsets. Thus, in total, we need to consider only

$$\leq c_0^2 \min\{e^{(l_1'-p_1)/2}, e^{p_1-l_1}\} \max\{e^{3(l_2-p_1)/2}, e^{3(l_2-p_1)-(p_1'-p_1)}\} =: N_2 \tag{3.2.5}$$

of these subsets.

Taking the images of these sets under $\mathrm{T}^{l_2-p_1}$ we obtain sets of the form

$$\mathrm{T}^{l_2}(x_0)u^+D^{U^+}_{\eta/2}\alpha^{-l_2}B^{U^-C}_{\eta/2}\alpha^{l_2}.$$

Now, let us consider the interval $[l_2, p_2]$ and let us divide the sets obtained earlier into $\lceil e^{3(p_2-l_2)/2}\rceil^2$ subsets of the form

$$\mathrm{T}^{p_2}(x_0)u^+D^{U^+}_{\eta/2}\alpha^{-p_2}B^{U^-C}_{\eta/2}\alpha^{p_2}.$$

From Lemma 3.5 we know that there are $\leq c_0\min\{e^{p_1'-l_2}, e^{(l_2-p_1)/2}\}$ many configurations and for each of them we can apply Lemma 3.10 with $S' = p_2 - l_2$ and $S = l_2' - l_2$. So, for each configuration we need only $\leq c_0\max\{1, e^{3(p_2-l_2)-(l_2'-l_2)}\}$ many of the subsets. Thus, we need

$$\leq c_0^2\min\{e^{p_1'-l_2}, e^{(l_2-p_1)/2}\}\max\{1, e^{3(p_2-l_2)-(l_2'-l_2)}\} =: N_3 \tag{3.2.6}$$

many of these subsets. Continuing in this way at the end of the inductive step we consider the interval $[\max\{l_k, p_{k'}\}, K+R]$. Assume that $\max\{l_k, p_{k'}\} = l_k$ so that $l_k' = K+R$ and $k' = k-1$ (the other case is similar and left to the reader). We have the sets of the form

$$\mathrm{T}^{l_k}(x_0)u^+D^{U^+}_{\eta/2}\alpha^{-l_k}B^{U^-C}_{\eta/2}\alpha^{l_k}$$

that are obtained in the previous step. Let us divide them into $\lceil e^{3(l_k'-l_k)/2}\rceil^2$ small sets. By Lemma 3.5 we have $\leq c_0\min\{e^{p_{k-1}'-l_k}, e^{(l_k-p_{k-1})/2}\}$ configurations and for each we apply Lemma 3.10 with $S' = S = l_k' - l_k$. Hence, we need to consider only

$$\leq c_0^2\min\{e^{p_{k-1}'-l_k}, e^{(l_k-p_{k-1})/2}\}e^{3(l_k'-l_l)-(l_k'-l_k)} =: N_{2k-1} \tag{3.2.7}$$

of them. Thus, in the inductive step we divided the sets obtained earlier into

$$\lceil e^{3(p_1-l_1)/2}\rceil^2\lceil e^{3(l_2-p_1)/2}\rceil^2\cdots\lceil e^{3(l_k'-l_k)/2}\rceil^2$$

many parts and deduced that we only need to take

$$\leq N_1 N_2 N_3 \cdots N_{2k-1} \tag{3.2.8}$$

many of them where each set is of the form

$$\mathrm{T}^{K+R}(x_0)u^+ D^{U^+}_{\eta/2}\alpha^{-K-R} B^{U^-C}_{\eta/2}\alpha^{K+R}.$$

On the other hand, let us multiply the max term of (3.2.4) with the min term of (3.2.5) to get

$$\max\{1, e^{3(p_1-l_1)-(l_1'-l_1)}\}\min\{e^{(l_1'-p_1)/2}, e^{p_1-l_1}\}.$$

If $\max\{1, e^{3(p_1-l_1)-(l_1'-l_1)}\} = e^{3(p_1-l_1)-(l_1'-l_1)}$ then clearly the multiplication above is $\leq e^{3(p_1-l_1)-(l_1'-l_1)}e^{(l_1'-p_1)/2} \leq e^{2(p_1-l_1)}$. Otherwise, it is $\leq e^{p_1-l_1}$. Thus, in either case we have

$$\leq e^{2(p_1-l_1)}.$$

Similarly, let us multiply the max term of (3.2.5) with the min term of (3.2.6)

$$\max\{e^{3(l_2-p_1)/2}, e^{3(l_2-p_1)-(p_1'-p_1)}\}\min\{e^{p_1'-l_2}, e^{(l_2-p_1)/2}\}.$$

If $\max\{e^{3(l_2-p_1)/2}, e^{3(l_2-p_1)-(p_1'-p_1)}\} = e^{3(l_2-p_1)-(p_1'-p_1)}$ then the above multiplication is $\leq e^{3(l_2-p_1)-(p_1'-p_1)}e^{p_1'-l_2} = e^{2(l_2-p_1)}$. Otherwise, it is

$$\leq e^{3(l_2-p_1)/2}e^{(l_2-p_1)/2} = e^{2(l_2-p_1)}.$$

Hence, in either case we have that the product is $\leq e^{2(l_2-p_1)}$.

We continue in this way until we have considered all max and min terms. Thus, we obtain that

$$\begin{aligned} N_1 N_2 N_3 \cdots N_{2k-1} &\leq c_0^{4k} e^{2(p_1-l_1)} e^{2(l_2-p_1)} \cdots e^{2(p_{k-1}-l_{k-1})} e^{2(l_k'-l_k)} \\ &= c_0^{4k} e^{2(p_1-l_1)+2(l_2-p_1)+\cdots+2(l_k'-l_k)} \\ &= c_0^{4k} e^{2|V_{n+1}|} \end{aligned}$$

We know that k is the number of elements of $\mathcal{L}$ restricted to the interval V_{n+1}. From Lemma 3.3 we have that $k \leq \frac{|V_{n+1}|}{\lfloor \log M \rfloor} + 1$. Therefore, for the inductive step $K + |V_{n+1}|$, we get that the set $Z_O^+(V)$ can be covered by

$$\begin{aligned} &\ll e^{3K} e^{-(|V_1|+\ldots+|V_n|)} c_0^{4\frac{|V_1|+\ldots+|V_n|}{\lfloor \log M \rfloor}+4n+n'} e^{2|V_{n+1}|} c_0^{4\frac{|V_{n+1}|}{\lfloor \log M \rfloor}+4} \\ &= e^{3(K+|V_{n+1}|)} e^{-(|V_1|+\ldots+|V_{n+1}|)} c_0^{4\frac{|V_1|+\ldots+|V_{n+1}|}{\lfloor \log M \rfloor}+4(n+1)+n'} \end{aligned}$$

many forward Bowen $K + |V_{n+1}|$-balls.

Hence, letting $K = N$ together with (3.2.3) we see that the set $Z_O^+(V)$ can be covered by $\leq e^{3N-|V|} c_0^{\frac{4|V|}{\lfloor \log M \rfloor}+\frac{5N}{\lfloor \log M \rfloor}} \leq e^{3N-|V|} c_0^{\frac{9N}{\lfloor \log M \rfloor}}$ many forward Bowen N-balls. Now, replacing c_0^9 by c_0 we obtain the proposition. □

3.3 Proof of Theorem 1.3

Now we are in a position to prove Theorem 1.3. We will use what we obtained in this chapter together with Lemma 2.5.

Proof of the Theorem 1.3. Note first that it suffices to consider ergodic measures. For if μ is not ergodic, we can write μ as an integral of its ergodic components $\mu = \int \mu_t d\tau(t)$ for some probability space (E, τ) by [EW, Theorem 6.2]. Therefore, we have $\mu(X_{\geq M}) = \int \mu_t(X_{\geq M}) d\tau(t)$, but also $h_\mu(\mathrm{T}) = \int h_{\mu_t}(\mathrm{T}) d\tau(t)$ by [Wa, Thm. 8.4], so that desired estimate follows from the ergodic case.

Suppose that μ is ergodic. We would like to apply Lemma 2.5, for this we need to find an upper bound for covering μ-most of the space X by Bowen N-balls. So, let $M \geq 100$ be such that $\mu(X_{\leq M}) > 0$. Thus, ergodicity of μ implies that $\mu(\bigcup_{k=0}^{\infty} \mathrm{T}^{-k} X_{\leq M}) = 1$. Hence, for every $\epsilon > 0$ there is a constant $K \geq 1$ such that $Y = \bigcup_{k=0}^{K-1} \mathrm{T}^{-k} X_{\leq M}$ satisfies $\mu(Y) > 1 - \epsilon$. Moreover, the pointwise ergodic theorem implies

$$\frac{1}{2N-1} \sum_{n=-N+1}^{N-1} 1_{X_{\geq M}}(\mathrm{T}^n(x)) \to \mu(X_{\geq M})$$

as $N \to \infty$ for a.e. $x \in X$. Thus, for $\epsilon > 0$ given there is N_0 such that for $N > N_0$ the average on the left will be bigger that $\mu(X_{\geq M}) - \epsilon$ for any $x \in X_1$ for some $X_1 \subset X$ with measure $\mu(X_1) > 1 - \epsilon$. Clearly, for any N we have $\mu(Z) > 1 - 3\epsilon$ where

$$Z = X_1 \cap \mathrm{T}^N Y \cap \mathrm{T}^{-N} Y.$$

Now, we would like to find an upper bound for the number of Bowen N-balls needed to cover the set Z. Here $N \to \infty$ while ϵ and hence K are fixed. Since $Y = \bigcup_{k=0}^{K-1} \mathrm{T}^{-k} X_{\leq M}$, we can decompose Z into K^2 sets of the form

$$Z' = X_1 \cap \mathrm{T}^{N-k_1} X_{\leq M} \cap \mathrm{T}^{-N-k_2} X_{\leq M}$$

but since K is fixed, it suffices to find an upper bound for the number of Bowen N-balls to cover one of these. Consider the set Z', and since $k_1, k_2 \leq K$ without loss of generality we can assume $k_1 = k_2 = 0$. Next we split Z' into the sets $Z(\mathcal{N})$ as in Proposition 3.7 for various subsets $\mathcal{N} \in \mathcal{M}_N$. By Lemma 3.4 we know that we need $\ll_M e^{\frac{10N \log \lfloor \log M \rfloor}{\lfloor \log M \rfloor}}$ many of these under the assumption that $M \geq 100 > e^4$. Moreover, by our assumption on X_1 we only need to look at sets $V \subset [-N+1, N-1]$ with $|V| \geq (\mu(X_{\geq M}) - \epsilon)(2N-1)$. On the other hand, Proposition 3.7 gives that each of those sets $Z(\mathcal{N})$ can be covered by $\leq e^{6N - |V|} c_0^{\frac{18N}{\lfloor \log M \rfloor}}$ Bowen N-balls for some constant $c_0 > 0$ that does not depend on M. Together we see that Z can be covered by

$$\ll_{M,K} e^{\frac{10N \log \lfloor \log M \rfloor}{\lfloor \log M \rfloor}} c_0^{\frac{18N}{\lfloor \log M \rfloor}} e^{6N - |V|}$$

many Bowen N-balls. Applying Lemma 2.5 we arrive at

$$\begin{aligned} h_\mu(\mathrm{T}) &\leq \lim_{\epsilon \to 0} \liminf_{N \to \infty} \frac{\log BC(N, \epsilon)}{2N} \\ &\leq \lim_{\epsilon \to 0} (3 - (\mu(X_{\geq M}) - \epsilon) + O(\frac{\log \log M}{\log M}) \\ &\leq 3 - \mu(X_{\geq M}) + O(\frac{\log \log M}{\log M}) \end{aligned}$$

which completes the proof for any sufficiently large M with $\mu(X_{\leq M}) > 0$. However, we claim that the same conclusion holds for any sufficiently large M independent of μ (which e.g. is crucial for proving Corollary 1.4).

If $\mu(X_{\leq 100}) > 0$ then the claim is true by the above discussion. So, assume that $\mu(X_{\leq 100}) = 0$ and let

$$M_\mu = \inf\{M > 100 : \mu(X_{\leq M}) > 0\}.$$

Since $\mu(X_{\leq M}) > 0$ for any $M > M_\mu \geq 100$ we have

$$h_\mu(\mathrm{T}) \leq 3 - \mu(X_{\geq M}) + O(\frac{\log\log M}{\log M}). \tag{3.3.1}$$

of the above.

If $\mu(X_{\leq M_\mu}) > 0$ then (3.3.1) also holds for $M = M_\mu$ by the above. If on the other hand, $\mu(X_{\leq M_\mu}) = 0$ then $\lim_{n\to\infty} \mu(X_{\geq M_\mu + \frac{1}{n}}) = \mu(X_{>M_\mu}) = \mu(X_{\geq M_\mu})$ and (3.3.1) for $M = M_\mu$ follows from (3.3.1) for $M = M_\mu + \frac{1}{n}$. Since $\mu(X_{\geq M_\mu}) = 1$ this simplifies to

$$h_\mu(\mathrm{T}) \leq 2 + O(\frac{\log\log M}{\log M}).$$

Since $\frac{\log\log M}{\log M}$ is a decreasing function for $M \geq 100$ and $\mu(X_{\geq M}) = 1$ for $M \leq M_\mu$ we obtain that (3.3.1) trivially also holds for any $M \in [100, M_\mu)$. $\square$

CHAPTER 4

POSITIVE ENTROPY INVARIANT MEASURES ON THE SPACE OF LATTICES WITH ESCAPE OF MASS

For this chapter we let $G = \mathrm{SL}_{d+1}(\mathbb{R})$, $\Gamma = \mathrm{SL}_{d+1}(\mathbb{Z})$ and transformation T acting on $X = \Gamma\backslash G$ as a right multiplication by the element $diag(e^{1/d}, \ldots, e^{1/d}, e^{-1}) \in G$. Our goal here is to prove Theorem 1.5 which gives the existence of a sequence of T-invariant probability measures with high entropy whose limit is 0.

Let $M > 0$ be given. For a lattice $x \in X$, define the height $\mathrm{ht}(x)$ to be the inverse of the length of the shortest nonzero vector in x. As before, we define the sets

$$X_{<M} = \{x \in X : \mathrm{ht}(x) < M\} \text{ and } X_{\geq M} = \{x \in X : \mathrm{ht}(x) \geq M\}.$$

We know by Mahler's compactness criterion $X_{<M}$ is pre-compact. Theorem 1.5 easily follows from the following.

Theorem 4.1. *For any $\epsilon > 0$ and $M \geq 1$ there exists a* T*-invariant probability measure μ with $h_\mu(\mathrm{T}) > d - \epsilon$ such that $\mu(X_{\geq M}) > 1 - \epsilon$.*

We will construct infinitely many points in $X_{<M}$ whose forward trajectories mostly stay above height M. Taking union of the sets of forward trajectories of these points, we will construct a T-invariant set S_N with topological entropy greater than $d-\epsilon$ (cf. Theorem 4.3). To construct the T-invariant probability measures we want, we will make use of the Variational Principle. Our main ingredient to prove Theorem 4.1 is

Theorem 4.3 below. In the next section, we prove Theorem 4.1 assuming Theorem 4.3. In the last two sections we prove Theorem 4.3.

4.1 The proof of Theorem 4.1

Before we start the construction, we would like to deduce Theorem 4.1 from Theorem 4.3 below.

Let $\delta > 0$ be an injectivity radius for $X_{<17M}$ with $\delta < \frac{1}{8M}$. Here is an easy lemma which will be used repeatedly in the last section.

Lemma 4.2. *There exists $N' > 0$ such that for any $x, y \in X_{<17M}$ there exists $z \in X_{<17M}$ such that $d(z, y) < \delta/5^9$ and $d(x, \mathrm{T}^{N'}(z)) < \delta/5^9$.*

Proof. Let λ be the Haar measure on X. Since $X_{<17M}$ is precompact we can cover it with open balls $\mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_k$ of radius $\delta/5^9$. They have positive measure with respect to the Haar measure. Since T is mixing with respect to the Haar measure, for any $i, j \in \{1, 2, ..., k\}$ there exists $N_{ij} \geq 0$ with $\lambda(\mathrm{T}^l(\mathcal{O}_j) \cap \mathcal{O}_i) > 0$ for any $l \geq N_{ij}$. Letting $N' = \max\{N_{ij} : i, j = 1, 2, ..., k\}$ we obtain the lemma. $\square$

For a given $M \geq 1$ we fix N' as in Lemma 4.2.

Theorem 4.3. *For any large N we let $K = \lfloor \frac{1}{13} e^{dN} \rfloor$. Then there exist a constant $M' > 1$ and a set S_N in $X_{<M}$ such that*

$$\mathrm{T}^l(x) \in X_{<M'} \text{ for all } x \in S_N \text{ and for all } l \geq 0.$$

Moreover, there exists a constant $s > 0$ such that for any $m \in \mathbb{N}$ there are subsets S_m of S_N with the following properties:

1 . *cardinality of S_m is K^m*

2 . *S_m is $(mN + (m-1)N', s)$-separated and*

3 . for any $x \in S_m$ we have $|\{l \in [0, mN + (m-1)N'] : \mathrm{T}^l(x) \in X_{\geq M}\}| \geq mN$.

Now we deduce Theorem 4.1 from Theorem 4.3.

Proof of the Theorem 4.1. Let $\epsilon > 0$ be given and let N' be as in Lemma 4.2. Choose N large enough so that

$$\frac{1}{N+N'} \log\lfloor 13e^{dN} \rfloor > d - \epsilon \text{ and } \frac{N'}{N+N'} < \epsilon$$

and let S_N be the set as in Thereom 4.3.

To obtain a T-invariant probability measure with high entropy we would like to make use of Variational Principle 2.6. For this, we need a compact subspace of X. We define

$$Y_{\leq M'} = \{x \in X_{\leq M'} \mid \mathrm{T}^l(x) \in X_{\leq M'}, \text{ for } l \geq 0\}.$$

Clearly, we obtain a T-invariant compact subspace containing $\mathrm{T}^l(S_N)$ for all $l \geq 0$. We have $h_{\mathrm{T}}(Y_{\leq M'}) > d - \epsilon$ since $Y_{\leq M'}$ contains the sets S_m which are $(mN + (m-1)N', s)$-separated by Theorem 4.3. Now, from Variational principle 2.6 we know that there is a T-invariant measure μ on $Y_{\leq M'}$, hence on X, with $h_\mu(\mathrm{T}) > d - \epsilon$. In order to obtain the theorem, we want to have $\mu(X_{\geq M}) > 1 - \epsilon$, but we do not get this from Variational principle itself. Thus, we need to look into the proof of Variational principle and see how the measures are constructed.

Let S_m be the subset of $Y_{\leq M'}$ as in Theorem 4.3. We have that S_m is $(mN + (m-1)N', s)$-separated and has cardinality K^m where $K = \lfloor \frac{1}{13}e^{dN} \rfloor$. Define a probability measure

$$\sigma_m = \frac{1}{K^m} \sum_{x \in S_m} \delta_x \text{ where } \delta_x(A) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ if } x \notin A \end{cases}.$$

Now, let a probability measure μ_m be defined by

$$\mu_m = \frac{1}{mN + (m-1)N'} \sum_{i=0}^{mN+(m-1)N'-1} \sigma_m \circ \mathrm{T}^{-i}$$

where $\sigma_m \circ \mathrm{T}^{-i}(A) = \sigma_m(\mathrm{T}^{-i}(A))$ for any measurable set A. We know that $\mathcal{M}(Y_{\leq M'})$, the space of Borel probability measures, is compact in the weak* topology [Wa, Theorem 6.5]. We obtained a set of measures $\mu_m \in \mathcal{M}(Y_{\leq M'})$. If necessary going into subsequence, we have that $\{\mu_m\}$ converges to some probability measure μ in $\mathcal{M}(Y_{\leq M'})$. The measure μ we obtained is T-invariant [Wa, Theorem 6.9]. From the proof of Variational Principle [Wa, Theorem 8.6], we know that μ has

$$\begin{aligned} h_\mu(\mathrm{T}_{|Y_{\leq M'}}) &\geq \lim_{m\to\infty} \frac{1}{mN+(m-1)N'} \log s_m(\epsilon, Y_{\leq M'}) \\ &\geq \lim_{m\to\infty} \frac{1}{mN+(m-1)N'} \log K^m \\ &= \frac{1}{N+N'} \log K. \end{aligned}$$

On the other hand, by assumption we have $\frac{1}{N+N'} \log K > d - \epsilon$ and hence we obtain

$$h_\mu(\mathrm{T}) \geq h_\mu(\mathrm{T}_{|Y_{\leq M'}}) > d - \epsilon.$$

Now, we claim that $\mu(X_{\geq M}) > 1 - \epsilon$ and this will complete the proof. To prove the claim we will show that $\mu(X_{<M}) < \epsilon$.

We need to approximate $1_{X_{<M}}$ by continuous functions of compact support. So, for any $\epsilon' > 0$ let $f \in C_c(X)$ be such that

$$f(x) = \begin{cases} 1 & \text{for } x \in X_{\leq (M-\epsilon')} \\ 0 & \text{for } x \in X_{\geq M} \end{cases}$$

and $0 \leq f(x) \leq 1$ otherwise. This is possible by Urysohn's Lemma. Hence,

$$\int f\, d\mu_m \leq \int 1_{X_{<M}}\, d\mu_m = \mu_m(X_{<M}).$$

We have $\mu_m(X_{<M}) = \frac{1}{mN+(m-1)N'} \sum_{i=0}^{mN+(m-1)N'-1} \sigma_m \circ \mathrm{T}^{-i}(X_{<M})$. Hence, from part (iii) of Theorem 4.3

$$\mu_m(X_{<M}) \leq \frac{(m-1)N'}{mN+(m-1)N'} < \frac{N'}{N+N'} < \epsilon.$$

Thus, $\int f\, d\mu_m < \epsilon$ so that $\mu(X_{\leq(M-\epsilon')}) \leq \int f\, d\mu \leq \epsilon$. Therefore,

$$\mu(X_{\geq M}) > 1 - \epsilon.$$

$\square$

4.2 Initial setup and shadowing lemma

In this section we will construct about e^{dN} lattices whose forward trajectories stay above height M in the time interval $[1, N]$ for some large number N. Later we prove the shadowing lemma 4.6, which will be used in the proof of Theorem 4.3 in the next section.

Throughout the paper the norms $\|\cdot\|$ on $\mathbb{R}^d$ and on $\mathbb{R}^{d+1}$ will be the maximum norms. Fix a height $M > 0$. Let $N \in \mathbb{N}$ be a given. For $t = (t_1, t_2, ..., t_d) \in [0, e^{-N/d}]^d$ consider the lattice $x_t = \Gamma g_t$ where

$$g_t = \begin{pmatrix} M^{1/d} & 0 & ... & 0 & 0 \\ 0 & M^{1/d} & ... & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & ... & M^{1/d} & 0 \\ \frac{t_1}{M} & \frac{t_2}{M} & ... & \frac{t_d}{M} & \frac{1}{M} \end{pmatrix}. \tag{4.2.1}$$

We would like to consider those lattices that stay above height M in $[1, N]$ and are in $X_{<16M}$ at time N. We start with first considering the set

$$A_N := \{t \in [0, e^{-N/d}]^d : \mathrm{T}^N(x_t) \in X_{<16M}\}.$$

We claim that A_N is significant in size.

Lemma 4.4. *For $d \geq 2$ let $m_{\mathbb{R}^d}$ be the Lebesgue measure on $\mathbb{R}^d$. Then*

$$m_{\mathbb{R}^d}(A_N) \geq (\frac{15^d}{16^d} - \frac{1}{4^d})e^{-N}.$$

The explicit constant $(\frac{15^d}{16^d} - \frac{1}{4^d})$ has no importance to us. All we need is that $m_{\mathbb{R}^d}(A_N) \gg e^{-N}$. However, the explicit constant simplifies the later work. We can think of A_N as a subset of the unstable subgroup U^+ in G w.r.t. a. Although A_N has small volume in $\mathbb{R}^d$, it gets expanded by T^N to a set of volume $\gg e^{dN}$ which will give us an (N, s)-separated set of cardinality $\gg e^{dN}$.

Proof. We will prove that $m_{\mathbb{R}^d}(A'_N) \geq (\frac{15^d}{16^d} - \frac{1}{4^d})e^{-N}$ where

$$A'_N = A_N \cap [\frac{1}{16}e^{-N/d}, e^{-N/d}]^d. \tag{4.2.2}$$

Assume that $\mathrm{ht}(\mathrm{T}^N(x_t)) > 16M$. So, for some nonzero $(p_1, p_2, ..., p_d, q) \in \mathbb{Z}^{d+1}$ with $\gcd(p_1, p_2, ..., p_d, q) = 1$ and $q > 0$ we must have

$$\begin{aligned} &\|(p_1, p_2, ..., p_d, q)g_t a^N\| \\ &= \|(p_1 M^{1/d} + q\frac{t_1}{M})e^{N/d}, (p_2 M^{1/d} + q\frac{t_2}{M})e^{N/d}, ..., (p_d M^{1/d} + q\frac{t_d}{M})e^{N/d}, q\frac{1}{M}e^{-N})\| \\ &< \frac{1}{16M}. \end{aligned}$$

So, letting $\epsilon = \frac{e^{-N/d}}{16M^{(d+1)/d}}$ we have

$$|p_i + q\frac{t_i}{M^{(d+1)/d}}| < \epsilon \text{ for all } i = 1, 2, ..., d \text{ and } q < \frac{e^N}{16}. \tag{4.2.3}$$

We have $t_i \in [\frac{1}{16}e^{-N/d}, e^{-N/d}]$. For a fixed q, we will calculate the Lebesgue measure of $(t_1, t_2, ..., t_d) \in [\frac{1}{16}e^{-N/d}, e^{-N/d}]^d$ for which (4.2.3) hold for some p_i's.

We have

$$q\frac{t_i}{M^{(d+1)/d}} \in [q\epsilon, 16q\epsilon].$$

If $16q\epsilon \leq \frac{1}{2}$ then $(p_1, p_2, ..., p_d) = 0$ and since we only need to consider the primitive vectors in x_t we have $q = 1$. In this case, $q\frac{t_i}{M^{(d+1)/d}} \in [\epsilon, 16\epsilon]$ and hence (4.2.3) does not hold. So, we can assume that

$$16q\epsilon > \frac{1}{2}.$$

We note that $q\frac{t_i}{M^{(d+1)/d}}$ must be in the ϵ-neighborhood of an integer point. If $16q\epsilon \in (1/2, 1)$ then $[q\epsilon, 16q\epsilon]$ does not contain any integers and only possible way for (4.2.3) to hold is when $q\frac{t_i}{M^{(d+1)/d}}$ is in $(1-\epsilon, 1+\epsilon)$ so that t_i must be in

$$(\frac{(1-\epsilon)M^{(d+1)/d}}{q}, \frac{(1+\epsilon)M^{(d+1)/d}}{q}).$$

Thus, for a fixed $q \in (\frac{1}{32\epsilon}, \frac{1}{16\epsilon})$ we have that the Lebesgue measure of points that satisfy (4.2.3) is

$$\leq \left(\frac{2\epsilon M^{(d+1)/d}}{q}\right)^d = \frac{2^d \epsilon^d M^{d+1}}{q^d}.$$

Now, for $16q\epsilon \geq 1$ we have that $[q\epsilon, 16q\epsilon]$ has at most $\leq 15q\epsilon + 1$ integer points. Thus, there could be $\leq 15q\epsilon + 2$ integers for which $q\frac{t_i}{M^{(d+1)/d}}$ can be ϵ-close for some t_i. Since $16q\epsilon \geq 1$ we have $15q\epsilon + 2 \leq 48q\epsilon$. Hence, arguing as in the previous case, for a fixed $q \geq \frac{1}{16\epsilon}$ we have that the Lebesgue measure of points satisfying (4.2.3) is

$$\leq \left((48q\epsilon)(2\epsilon)(\frac{M^{(d+1)/d}}{q})\right)^d = 96^d \epsilon^{2d} M^{d+1}.$$

Thus, we obtain that the Lebesgue measure of points for which (4.2.3) hold is

$$\leq \sum_{q=\lceil \frac{1}{32\epsilon} \rceil}^{\lfloor \frac{1}{16\epsilon} \rfloor} \frac{2^d \epsilon^d M^{d+1}}{q^d} + \sum_{q=\lceil \frac{1}{16\epsilon} \rceil}^{\lfloor \frac{e^N}{16} \rfloor} 96^d \epsilon^{2d} M^{d+1}.$$

Since $\epsilon^d = \frac{e^{-N}}{16^d M^{d+1}}$, the above inequality simplifies to

$$\leq e^{-N} \left(\sum_{q=\lceil \frac{1}{32\epsilon} \rceil}^{\lfloor \frac{1}{16\epsilon} \rfloor} \frac{2^d}{16^d q^d} + \sum_{q=\lceil \frac{1}{16\epsilon} \rceil}^{\lfloor \frac{e^N}{16} \rfloor} \frac{96^d e^{-N}}{16^{2d} M^{d+1}} \right). \tag{4.2.4}$$

We want to show that, independent of N, the term inside the parenthesis is strictly less than 1.

$$\sum_{q=\lceil \frac{1}{32\epsilon} \rceil}^{\lfloor \frac{1}{16\epsilon} \rfloor} \frac{2^d}{16^d q^d} \leq \sum_{q=\lceil \frac{1}{32\epsilon} \rceil}^{\lfloor \frac{1}{16\epsilon} \rfloor} \frac{2^d}{16^d q} \leq \frac{1}{8^d \frac{1}{32\epsilon}} (\lfloor \frac{1}{16\epsilon} \rfloor - \lceil \frac{1}{32\epsilon} \rceil) \leq \frac{1}{8^d}.$$

On the other hand,

$$\sum_{q=\lceil \frac{1}{16\epsilon} \rceil}^{\lfloor \frac{e^N}{16} \rfloor} \frac{96^d e^{-N}}{16^{2d} M^{d+1}} \leq \frac{96^d e^{-N}}{16^{2d} M^{d+1}} \frac{e^N}{16} < \frac{1}{2^{d+4} M^{d+1}}.$$

Together, we see that the inequality (4.2.4) is

$$< (\frac{1}{8^d} + \frac{1}{2^{d+4} M^{d+1}}) e^{-N} \leq \frac{e^{-N}}{4^d}.$$

Thus, we conclude that $m_{\mathbb{R}^d}(A_N) \geq m_{\mathbb{R}^d}(A'_N) > (\frac{15^d}{16^d} - \frac{1}{4}) e^{-N}$. $\square$

From the set A_N, in fact from A'_N as in (4.2.2), we want to pick about e^{dN} many elements which are not too close to each other so that within N iterations under T they get apart from each other. For this purpose, let us partition $[\frac{1}{16} e^{-N/d}, e^{-N/d}]^d$ into $\lfloor e^N \rfloor^d$ small d-cubes of side length $\frac{15}{16} e^{-N(d+1)/d}$.

Now, consider even smaller d-cubes of side length $\frac{13}{16} e^{-N(d+1)/d}$ each lying at the center of one of the small d-cubes. We need to find a lower bound for the number of these smaller d-cubes that intersect with the set A'_N. Each of these d-cubes has volume equal to $(\frac{13}{16})^d e^{-N(d+1)}$. Thus, there could be at most

$$\left\lceil \frac{(\frac{1}{4^d}) e^{-N}}{(\frac{13}{16})^d e^{-N(d+1)}} \right\rceil = \left\lceil \frac{4^d}{13^d} e^{dN} \right\rceil$$

many that do not intersect with A'_N. Therefore, for N large, at least

$$\lfloor e^N \rfloor^d - \left\lceil \frac{4^d}{13^d} e^{dN} \right\rceil \geq \frac{1}{13} e^{dN}$$

of these smaller d-cubes do intersect with A'_N.

Let us pick one element t from each of these smaller d-cubes that is also contained in A'_N and consider the set S'_1 of these lattices $x_t = \Gamma g_t$ where g_t is as in (4.2.1). To simplify notation we let

$$S'_1 = \{x_1, x_2, ..., x_K\} = \{\Gamma g_1, \Gamma g_2, ..., \Gamma g_K\} \tag{4.2.5}$$

where

$$K = \lfloor \frac{1}{13} e^{dN} \rfloor.$$

We note that for elements t, t' that are picked from different d-cubes one has

$$\frac{1}{4} e^{-N(d+1)/d} \leq \|t - t'\| < \frac{15}{16} e^{-N/d}. \tag{4.2.6}$$

Proposition 4.5. *For a given large N the set $S_1' = \{x_1, x_2, ..., x_K\}$ has the following properties:*

1 *.* $\mathrm{ht}(\mathrm{T}^l(x_i)) \geq M$ *for* $l \in [1, N]$

2 *.* $\mathrm{ht}(x_i) < M$ *and* $\mathrm{ht}(\mathrm{T}^N(x_i)) < 16M$*,*

3 *. for $i \neq j$ we have $d(g_i, g_j) < \frac{30}{16} e^{-N/d}$ and $d(\mathrm{T}^N(g_i), \mathrm{T}^N(g_j)) \geq \frac{1}{8M}$.*

Proof. Let $x_i = x_t = \Gamma g_t$ for some $t = (t_1, t_2, ..., t_d) \in [\frac{1}{16} e^{-N/d}, e^{-N/d}]^d$ (cf. (4.2.1)). It is easy to see that $x_t \in X_{<M}$. On the other hand, by construction $t \in A_N$ so that $\mathrm{T}^N(x_t) \in X_{<16M}$.

Now, consider the vector $v = (\frac{t_1}{M}, \frac{t_2}{M}, ..., \frac{t_d}{M}, \frac{1}{M}) \in x_t$. We have

$$\mathrm{T}(v) = (\frac{t_1 e^{1/d}}{M}, \frac{t_2 e^{1/d}}{M}, ..., \frac{t_d e^{1/d}}{M}, \frac{e^{-1}}{M})$$

so that

$$\| \mathrm{T}(v) \| \leq \max\{\frac{e^{-(N-1)/d}}{M}, \frac{e^{-1}}{M}\} < \frac{1}{M}.$$

Also,

$$\mathrm{T}^N(v) = (\frac{t_1 e^{N/d}}{M}, \frac{t_2 e^{N/d}}{M}, ..., \frac{t_d e^{N/d}}{M}, \frac{e^{-N}}{M})$$

which implies

$$\| \mathrm{T}^N(v) \| \leq \max\{\frac{1}{M}, \frac{e^{-N}}{M}\} \leq \frac{1}{M}.$$

Since the function $\mathrm{T}^l(v)$ in l has only one critical point we conclude that for $l = 1, 2, ..., N$

$$\mathrm{ht}(\mathrm{T}^l(x_t)) \geq M.$$

Let x_j be another element and let $t' \in [\frac{1}{16}e^{-N/d}, e^{-N/d}]^d$ be such that $x_j = x_{t'} = \Gamma g_{t'}$. From (4.2.6) together with left invariance of the metric we have

$$d(\mathrm{T}^N(g_t), \mathrm{T}^N(g_{t'})) = d(a^n a^{-n} g_t a^n, a^n a^{-n} g'_t a^n) \geq \frac{\|t-t'\|}{2M} e^{N(d+1)/d} \geq \frac{1}{8M}.$$

The fact that $d(g_i, g_j) < \frac{30}{16}e^{-N/d}$ follows from (4.2.6) also. □

Our main tool for the construction of lattices is the shadowing lemma:

Lemma 4.6 (Shadowing lemma). *Let $1/4 \geq \epsilon > 0$ be given. If $d(x_-, x_+) < \epsilon$ for some $x_-, x_+ \in X$ then there exists $y \in X$ such that*

1 . $d(\mathrm{T}^l(y), \mathrm{T}^l(x_-)) < 5\epsilon e^{l(d+1)/d}$ for all $l \leq 0$ and

2 . $d(\mathrm{T}^l(y), \mathrm{T}^l(x_+)) < 5\epsilon$ for all $l \geq 0$.

Moreover, there exists c in the centralizer C of a with $d(c, 1) < 5\epsilon$ such that $d(\mathrm{T}^l(y), \mathrm{T}^l(x_+c)) < 10\epsilon e^{-l(d+1)/d}$ for all $l \geq 0$.

Proof. We have $x_- = x_+ g$ for some $g = (g_{ij}) \in \mathrm{SL}(d+1, \mathbb{R})$ with $d(g, 1) < \epsilon$. Consider

$$u^+ = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ u_1 & u_2 & \dots & u_d & 1 \end{pmatrix}$$

and let $y = x_- u^+$. For $\|u_1, u_2, \ldots, u_d\| < 4\epsilon$ we have

$$\begin{aligned}
d(\mathrm{T}^l(y), \mathrm{T}^l(x_-)) &= d(x_- u^+ a^l, x_- a^l) \\
&= d(x_- a^l a^{-l} u^+ a^l, x_- a^l) \\
&\leq d\left(\begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \\ u_1 e^{l(d+1)/d} & u_2 e^{l(d+1)/d} & \ldots & u_d e^{l(d+1)/d} & 1 \end{pmatrix}, 1\right) \\
&< 5\epsilon e^{l(d+1)/d}.
\end{aligned}$$

This establishes part (i). Now, let $g' := gu^+$

$$= \begin{pmatrix} g_{11} + g_{1(d+1)} u_1 & \ldots & g_{1d} + g_{1(d+1)} u_d & g_{1(d+1)} \\ g_{21} + g_{2(d+1)} u_1 & \ldots & g_{2d} + g_{2(d+1)} u_d & g_{2(d+1)} \\ \vdots & & \vdots & \vdots \\ . & \ldots & . & . \\ g_{(d+1)1} + g_{(d+1)(d+1)} u_1 & \ldots & g_{(d+1)d} + g_{(d+1)(d+1)} u_d & g_{(d+1)(d+1)} \end{pmatrix}.$$

Hence, letting $u_i = -\frac{g_{(d+1)i}}{g_{(d+1)(d+1)}}$ for $i = 1, 2, ..., d$ we can make sure that the unstable part with respect to a is 0. We have $||u_1, u_2, \ldots, u_d|| < \frac{2\epsilon}{|g_{(d+1)(d+1)}|} < \frac{2\epsilon}{1-2\epsilon} \leq \frac{\epsilon}{1-1/2} = 4\epsilon$. So,

$$d(\mathrm{T}^l(y), \mathrm{T}^l(x_+)) = d(\mathrm{T}^l(x_+ g u^+), \mathrm{T}^l(x_+)) = d(x_+ a^l a^{-l} g' a^l, x_+ a^l) \leq d(a^{-l} g' a^l, 1).$$

Since unstable part of g' is 0, for $l \geq 0$ we obtain

$$d(\mathrm{T}^l(y), \mathrm{T}^l(x_+)) \leq d(g', 1) < 5\epsilon.$$

For the last part, let

$$c = \begin{pmatrix} g_{11} + g_{1(d+1)}u_1 & \dots & g_{1d} + g_{1(d+1)}u_d & 0 \\ g_{21} + g_{2(d+1)}u_1 & \dots & g_{2d} + g_{2(d+1)}u_d & 0 \\ \vdots & & \vdots & \vdots \\ . & \dots & . & 0 \\ 0 & \dots & 0 & g_{(d+1)(d+1)} \end{pmatrix},$$

then we have that $c \in C$ with $d(c, 1) < 5\epsilon$, and hence $d(c^{-1}, 1) < 5\epsilon$. On the other hand, $c^{-1}g' = u^-$ where $u^- = \begin{pmatrix} 1 & \dots & 0 & g'_{1(d+1)} \\ 0 & \dots & 0 & g'_{2(d+1)} \\ \vdots & & \vdots & \vdots \\ . & \dots & 1 & g'_{d(d+1)}. \\ 0 & \dots & 0 & 1 \end{pmatrix}$ is such that $d(u^-, 1) < 10\epsilon$. Thus, $d(\mathrm{T}^l(y), \mathrm{T}^l(x_+c)) = d(x_+gu^+a^l, x_+ca^l) = d(x_+g'a^l, x_+ca^l) \leq d(g'a^l, ca^l) = d(a^{-l}c^{-1}g'a^l, 1) = d(a^{-l}u^-a^l, 1) < 10\epsilon e^{-l(d+1)/d}$. □

4.3 The construction

In this section we construct the set S_N mentioned in the introduction with the properties as in Theorem 4.3. Repeatedly using both the shadowing lemma and K lattices constructed in the previous section we obtain more and more lattices that in the limit gives the set S_N.

Let $M' > 0$ be a height that depends on N such that for any $x_i \in S'_1$ and for any $l = 0, 1, ..., N$ we have $\mathrm{T}^l(x_i) \in X_{<M'}$. Now, let $\eta > 0$ be such that $\eta < \frac{1}{8M}$ and that 2η is an injectivity radius of $X_{<M'}$. Recall that $K = \lfloor \frac{1}{13} e^{dN} \rfloor$. We will prove Theorem 4.3 with choice of $c = \eta/e^2$ and with the choice of M' as defined above.

Theorem 4.3 follows from the following proposition. Recall that $\delta > 0$ is an injectivity radius for $X_{<17M}$ with $\delta < \frac{1}{8M}$.

Proposition 4.7. *For any positive integer n, there is a subset*

$$S'_n = \{x_{i_1 i_2 \dots i_n} : i_1, i_2, ..., i_n \in \{1, 2, ..., K\}\}$$

of $X_{<M}$ with the following properties:

1 . *for any $x \in S'_n$ and for any $m \leq n$ we have*

$$|\{l \in [0, mN + (m-1)N'] : \mathrm{T}^l(x) \in X_{\geq M}\}| \geq mN,$$

2 . *For any $x \in S'_n$ we have $\mathrm{T}^{nN+(n-1)N'}(x) \in X_{<17M}$,*

3 . *for any distinct $x_{i_1 i_2 \dots i_n}, x_{j_1 j_2 \dots j_n} \in S'_n$, say $i_m \neq j_m$, there exist $g, h \in G$ such that*

$$\mathrm{T}^{(m-1)(N+N')}(x_{i_1 i_2 \dots i_n}) = \Gamma g \text{ and } \mathrm{T}^{(m-1)(N+N')}(x_{j_1 j_2 \dots j_n}) = \Gamma h$$

with $d(\Gamma g, \Gamma h) = d(g, h)$ and that

$$d(\mathrm{T}^N(g), \mathrm{T}^N(h)) > \delta - \frac{\delta}{5^4} \text{ if } m = n \text{ and}$$
$$d(\mathrm{T}^N(g), \mathrm{T}^N(h)) > \delta - \delta \sum_{l=4}^{n-m+3} 5^{-l} \text{ if } m \in [1, n).$$

Moreover, we can make sure that for $x_{i_1 i_2 \dots i_n} \in S'_n$ and for $x_{i_1 i_2 \dots i_{n+1}} \in S'_{n+1}$ we have $d(x_{i_1 i_2 \dots i_n}, x_{i_1 i_2 \dots i_{n+1}}) < \delta e^{-n}$.

To derive Theorem 4.3 from Proposition 4.7 we need the lemma below which helps us to determine when two lattices get separated.

Lemma 4.8. *For $\Gamma g, \Gamma h \in X$ with $\mathrm{T}^l(\Gamma g), \mathrm{T}^l(\Gamma h) \in X_{<M'}$ in $[0, N]$ assume that $d(g, h) < \frac{\eta}{e^2}$ and $d(\mathrm{T}^N(g), \mathrm{T}^N(h)) \geq \frac{\eta}{e^2}$. Then $\Gamma g, \Gamma h$ is $(N, \frac{\eta}{e^2})$-separated, that is, there exists $l \in [1, N]$ with $d(\mathrm{T}^l(\Gamma g), \mathrm{T}^l(\Gamma h)) \geq \frac{\eta}{e^2}$.*

Proof. Since we have $d(g,h) < \frac{\eta}{e^2}$ and that $d(\mathrm{T}^N(g), \mathrm{T}^N(h)) > \frac{\eta}{e^2}$, there exists $l \in [1, N]$ such that

$$d(\mathrm{T}^{l-1}(g), \mathrm{T}^{l-1}(h)) < \frac{\eta}{e^2} \leq d(T^l(g), T^l(h)).$$

We have $d(\mathrm{T}(g), \mathrm{T}(h)) = d(a^{-1}h^{-1}ga, 1) = d(a^{-1}u^+aa^{-1}u^-ca, 1)$. On the other hand, we note that any two elements of the unstable subgroup with respect to a gets expanded at most by the factor of $e^{(d+1)/d}$ under the action of T. Together with triangle inequality we have

$$\begin{aligned}
d(a^{-1}u^+aa^{-1}u^-ca, 1) &\leq d(a^{-1}u^+aa^{-1}u^-ca, a^{-1}u^+a) + d(a^{-1}u^+a, 1) \\
&= d(a^{-1}u^-ca, 1) + d(a^{-1}u^+a, 1) \\
&\leq d(u^-c, 1) + e^{(d+1)/d}d(u^+, 1) \\
&\leq e^2(d(u^-c, 1) + d(u^+, 1)) \\
&\leq 2e^2 d(u^+u^-c, 1).
\end{aligned}$$

Thus, $d(\mathrm{T}^l(g), \mathrm{T}^l(h)) \leq 2e^2 d(\mathrm{T}^{l-1}(g), \mathrm{T}^{l-1}(h)) < 2\eta$. On the other hand, $\mathrm{T}^l(\Gamma g), \mathrm{T}^l(\Gamma h)$ are in $X_{<M'}$ and 2η is an injectivity radius of $X_{<M'}$. Hence, $d(\mathrm{T}^l(\Gamma g), \mathrm{T}^l(\Gamma h)) = d(\mathrm{T}^l(g), T^l(h)) \geq \frac{\eta}{e^2}$. □

Proof of Theorem 4.3. For any n let us pick a set

$$S'_n = \{x_{i_1 i_2 \ldots i_n} : i_1, i_2, ..., i_n \in \{1, 2, ..., K\}\}$$

as in Proposition 4.7. Also, assume for $x_{i_1 i_2 \ldots i_n} \in S'_n$ and for $x_{i_1 i_2 \ldots i_{n+1}} \in S'_{n+1}$ we have $d(x_{i_1 i_2 \ldots i_n}, x_{i_1 i_2 \ldots i_{n+1}}) < \delta e^{-n}$. If we fix a sequence $\{i_l\} \subset \{1, 2, ..., K\}^{\mathbb{N}}$, then the sequence $\{x_{i_1}, x_{i_1 i_2}, x_{i_1 i_2 i_3}, ...\}$ becomes a Chauchy sequence and hence converges. So, we let $x_{\{i_l\}} = \lim_{n\to\infty} x_{i_1 i_2 \ldots i_n}$. Varying the sequence $\{i_l\}$ we define the set

$$S_N = \left\{x_{\{i_l\}} : \{i_l\} \subset \{1, 2, ..., K\}^{\mathbb{N}}\right\}.$$

Also, define subsets S_m's of S_N

$$S_m = \left\{ x_{\{i_l\}} : \{i_l\} \subset \{1, 2, ..., K\}^{\mathbb{N}} \text{ with } i_l = 1 \text{ for all } l > m \right\}.$$

By definition of S_m, for any $x_{\{i_l\}} \in S_m$ with $\{i_l\} = \{i_1, i_2, ...\}$ we have

$$d(x_{\{i_l\}}, x_{i_1 i_2 \dots i_m}) < \eta.$$

Hence, from (i) of Proposition 4.7 we have

$$|\{l \in [0, mN + (m-1)N'] : \mathrm{T}^l(x) \in X_{\geq M/2}\}| \geq mN.$$

As for part (ii), again from the construction of the set S_m and from (iii) of Proposition 4.7 we conclude that for any distinct $x_{\{i_l\}}, x_{\{j_l\}} \in S_m$, say $i_n \neq j_n$, there exist $g, h \in G$ with $\mathrm{T}^{(n-1)(N+N')}(x_{\{i_l\}}) = \Gamma g$, $\mathrm{T}^{(n-1)(N+N')}(x_{\{j_l\}}) = \Gamma h$ and $d(\Gamma g, \Gamma h) = d(g, h)$ such that

$$d(\mathrm{T}^N(g), \mathrm{T}^N(h)) > \delta - \delta \sum_{l=3}^{\infty} 5^{-l} = \frac{99}{100}\delta.$$

If $d(\Gamma g, \Gamma h) \geq \frac{\eta}{e^2}$ then there is nothing to show, if not then from Lemma 4.8 for some $s \in [1, N]$ we conclude that $d(\mathrm{T}^s(\Gamma g), \mathrm{T}^s(\Gamma h)) \geq \frac{\eta}{e^2}$ since $\frac{\eta}{e^2} < \frac{99\eta}{100}$. Thus, for some $s \in [1, N]$ we have

$$d\left(\mathrm{T}^{(n-1)(N+N')+s}(x_{\{i_l\}}), \mathrm{T}^{(n-1)(N+N')+s}(x_{\{j_l\}})\right) \geq \frac{\eta}{e^2}$$

and hence the set S_m is $(mN+(m-1)N', \eta/e^2)$-separated since $n \leq m$. This concludes the proof. □

Now, we will make use of what we obtained in the previous section to prove Proposition 4.7.

Proof of Proposition 4.7. We prove the claim by induction. Fix some large N.

For $n = 1$ let $S_1' = \{x_1, x_2, ..., x_K\}$ be the set as in Proposition 4.5. It is clear that (i) and (ii) are satisfied. Let $x_i = \Gamma g_i, x_j = \Gamma g_j$ be distinct elements (cf. (4.2.5)). Then letting $g = g_i$ and $h = g_j$ we obtain (iii) since the part (iii) of Proposition 4.5 gives

$$d(\mathrm{T}^N(g_i), \mathrm{T}^N(g_j)) \geq \frac{1}{8M} = \delta.$$

Now, assume that the proposition holds for $n = k \geq 1$, we have the set $S_k' = \{x_{i_1 i_2 \ldots i_k} : i_1, i_2, ..., i_k = 1, ..., K\}$. Let us construct the set S_{k+1}'.

For any $x_{i_1 i_2 \ldots i_k} \in S_k'$, we have $\mathrm{T}^{kN+(k-1)N'}(x_{i_1 i_2 \ldots i_k}) \in X_{<16M}$. Hence, applying Lemma 4.2 we have that for x_j there exists z with

$$d(\mathrm{T}^{kN+(k-1)N'}(x_{i_1 i_2 \ldots i_k}), z) < \delta/5^9 \text{ and } d(x_j, \mathrm{T}^{N'}(y)) < \delta/5^9.$$

Now, apply shadowing lemma with $x_- = \mathrm{T}^{kN+(k-1)N'}(x_{i_1 i_2 \ldots i_k})$ and $x_+ = z$ and $\epsilon = \delta/5^9$. There exists y such that

$$d(\mathrm{T}^l(y), \mathrm{T}^l(T^{kN+(k-1)N'}(x_{i_1 i_2 \ldots i_k}))) < \frac{\delta}{5^8} e^{l(d+1)/d} \text{ for } l \leq 0 \text{ and} \tag{4.3.1}$$

$$d(\mathrm{T}^l(y), \mathrm{T}^l(z)) < \frac{\delta}{5^8} \text{ for } l \geq 0. \tag{4.3.2}$$

We have $d(x_j, \mathrm{T}^{N'}(y)) < d(x_j, \mathrm{T}^{N'}(z)) + d(\mathrm{T}^{N'}(z), \mathrm{T}^{N'}(y)) < \delta/5^9 + \delta/5^8 < \delta/5^7$. Apply shadowing lemma once more with $x_- = \mathrm{T}^{N'}(y)$ and $x_+ = x_j$ and $\epsilon = \delta/5^7$. There exists y' such that

$$d(\mathrm{T}^l(y'), \mathrm{T}^l(\mathrm{T}^{N'}(y))) < \frac{\delta}{5^6} e^{l(d+1)/d} \text{ for } l \leq 0 \text{ and} \tag{4.3.3}$$

$$d(\mathrm{T}^l(y'), \mathrm{T}^l(x_j)) < \frac{\delta}{5^6} \text{ for } l \geq 0 \tag{4.3.4}$$

Also, there exists $c_j \in C$ with $d(c_j, 1) < \frac{\delta}{5^6}$ such that

$$d(\mathrm{T}^l(y'), \mathrm{T}^l(x_j c_j)) < \frac{\delta}{5^5} e^{-l(d+1)/d} \text{ for } l \geq 0 \tag{4.3.5}$$

Now we let $x_{i_1 i_2 \ldots i_k j} = \mathrm{T}^{-k(N+N')}(y')$ and varying j we obtain the set

$$S_{k+1}' = \{x_{i_1 i_2 \ldots i_k j} : j \in \{1, 2, ..., K\}\}.$$

Part (i) of the proposition is clear from the construction since for any $j \in \{1, 2, ..., K\}$ we have that the forward trajectory of $x_{i_1 i_2 ... i_k j}$ stays close to $x_{i_1 i_2 ... i_k}$ in the time interval $[0, kN + (k-1)N']$ and then stays close to $\{\mathrm{T}^l(x_j) : l = 0, 1, ..., N\}$ in the time interval $[k(N + N'), (k+1)N + kN']$.

Now, let us justify part (ii). Let us fix some $j = 1, 2, ..., K$. Recalling that $x_{i_1 i_2 ... i_k j} = \mathrm{T}^{-k(N+N')}(y')$ we obtain from (4.3.4) with $l = N$ that

$$d(\mathrm{T}^{(k+1)N+kN'}(x_{i_1 i_2 ... i_k j}), \mathrm{T}^N(x_j)) < \frac{\delta}{5^6}.$$

Moreover, from Proposition 4.5 we have $\mathrm{T}^N(x_j) \in X_{<16M}$ so that

$$\mathrm{ht}(\mathrm{T}^{(k+1)N+kN'}(x_{i_1 i_2 ... i_k j})) \leq \frac{\mathrm{ht}(\mathrm{T}^N(x_j))}{1 - \frac{2\delta}{5^6}} < 17M.$$

Consider any distinct pairs $x_{i_1 i_2 ... i_k i_{k+1}}, x_{j_1 j_2 ... j_k j_{k+1}} \in S'_{k+1}$. First, assume that $i_{k+1} \neq j_{k+1}$ and let $g, h \in G$ be such that

$$\mathrm{T}^{k(N+N')}(x_{i_1 i_2 ... i_k i_{k+1}}) = \Gamma g, \;\; \mathrm{T}^{k(N+N')}(x_{j_1 j_2 ... j_k j_{k+1}}) = \Gamma h$$

with

$$\begin{aligned} &d(\mathrm{T}^{k(N+N')+N}(x_{i_1 i_2 ... i_k i_{k+1}} c_{i_{k+1}}), \mathrm{T}^N(x_{i_{k+1}})) \\ &\qquad = d(\mathrm{T}^N(g c_{i_{k+1}}), \mathrm{T}^N(g_{i_{k+1}})) < \frac{\delta}{5^5} e^{-N(d+1)/d} \text{ and} \end{aligned} \tag{4.3.6}$$

$$\begin{aligned} &d(\mathrm{T}^{k(N+N')+N}(x_{j_1 j_2 ... j_k j_{k+1}} c_{j_{k+1}}), \mathrm{T}^N(x_{j_{k+1}})) \\ &\qquad = d(\mathrm{T}^N(h c_{j_{k+1}}), \mathrm{T}^N(g_{j_{k+1}})) < \frac{\delta}{5^5} e^{-N(d+1)/d} \end{aligned} \tag{4.3.7}$$

for some $c_{i_{k+1}}, c_{j_{k+1}} \in C$ with $d(c_{i_{k+1}}, 1) < \frac{\delta}{5^6}$ and $d(c_{j_{k+1}}, 1) < \frac{\delta}{5^6}$ as in (4.3.5). Thus, we have

$$d(g_{i_{k+1}}, g c_{i_{k+1}}) < \frac{\delta}{5^5} \text{ and } d(g_{j_{k+1}}, h c_{j_{k+1}}) < \frac{\delta}{5^5}.$$

We also note from Proposition 4.5 that $d(g_{i_{k+1}}, g_{j_{k+1}}) < \frac{30}{16}e^{-N/d}$. Thus, for N large enough we get

$$\begin{aligned}
&d(g,h) \\
&< d(g, gc_{i_{k+1}}) + d(gc_{i_{k+1}}, g_{i_{k+1}}) + d(g_{i_{k+1}}, g_{j_{k+1}}) + d(g_{j_{k+1}}, hc_{j_{k+1}}) + d(hc_{j_{k+1}}, h) \\
&< \frac{\delta}{5^6} + \frac{\delta}{5^5} + \frac{30}{16}e^{-N/d} + \frac{\delta}{5^5} + \frac{\delta}{5^6} \\
&< \frac{\delta}{5^4}.
\end{aligned}$$

In particular, $d(\Gamma g, \Gamma h) = d(g,h)$ since δ is an injectivity radius for $X_{<17M}$. On the other hand, from Proposition 4.5 we know that

$$d(\mathrm{T}^N(g_{i_{k+1}}), \mathrm{T}^N(g_{j_{k+1}})) > \frac{1}{8M} > \delta.$$

So, together with (4.3.6) and (4.3.7) we conclude that

$$\begin{aligned}
&d(\mathrm{T}^N(g), \mathrm{T}^N(h)) \\
&> d(\mathrm{T}^N(g_{i_{k+1}}), \mathrm{T}^N(g_{j_{k+1}})) - d(\mathrm{T}^N(g_{i_{k+1}}), \mathrm{T}^N(g)) - d(\mathrm{T}^N(g_{j_{k+1}}), \mathrm{T}^N(h)) \\
&> \delta - \frac{\delta}{5^5}e^{-N(d+1)/d} - \frac{\delta}{5^6} - \frac{\delta}{5^5}e^{-N(d+1)/d} - \frac{\delta}{5^6} \\
&> \delta - \frac{\delta}{5^4}.
\end{aligned}$$

Now, assume that $i_m \neq j_m$ for some $m \leq k$. By replacing l in (4.3.1) by $l-(k-m)(N+N')$ we obtain

$$\begin{aligned}
&d(\mathrm{T}^{l-(k-m)(N+N')}(y), \mathrm{T}^{l+m(N+N')-N'}(x_{i_1 i_2 \dots i_k})) \\
&\qquad\qquad < \frac{\delta}{5^8}e^{(l-(k-m)(N+N'))(d+1)/d} \text{ for } l \leq 0. \qquad (4.3.8)
\end{aligned}$$

On the other hand, if we replace l in (4.3.3) by $l-(k-m)(N+N')-N'$ we get

$$\begin{aligned}
&d(\mathrm{T}^{l-(k-m)(N+N')-N'}(y'), \mathrm{T}^{l-(k-m)(N+N')}(y)) \\
&\qquad\qquad < \frac{\delta}{5^6}e^{(l-(k-m)(N+N')-N')(d+1)/d} \text{ for } l \leq 0. \qquad (4.3.9)
\end{aligned}$$

Thus, (4.3.8) and (4.3.9) together with the triangular inequality give

$$d(\mathrm{T}^{l-(k-m)(N+N')-N'}(y'),\mathrm{T}^{l+m(N+N')-N'}(x_{i_1i_2\dots i_k})) \;<\; \frac{\delta}{5^5}e^{(l-(k-m)(N+N')-N')(d+1)/d}$$

for $l\leq 0$ where $y'=\mathrm{T}^{-k(N+N')}(x_{i_1i_2\dots i_kj})$ for $j=1,2,...,K$. Thus, we have

$$d(\mathrm{T}^{m(N+N')-N'+l}(x_{i_1i_2\dots i_k}),\mathrm{T}^{m(N+N')-N'+l}(x_{i_1i_2\dots i_{k+1}})) < \frac{\delta}{5^5}e^{(l-(k-m)(N+N'))(d+1)/d} \tag{4.3.10}$$

and

$$d(\mathrm{T}^{m(N+N')-N'+l}(x_{j_1j_2\dots j_k}),\mathrm{T}^{m(N+N')-N'+l}(x_{j_1i_2\dots j_{k+1}})) < \frac{\delta}{5^5}e^{(l-(k-m)(N+N'))(d+1)/d}. \tag{4.3.11}$$

Now, from the induction hypothesis we have that there are g',h' with

$$\mathrm{T}^{m(N+N')}(x_{i_1i_2\dots i_k})=\Gamma g',\;\;\mathrm{T}^{m(N+N')}(x_{j_1j_2\dots j_k})=\Gamma h'$$

such that $d(\Gamma g',\Gamma h')=d(g',h')$ and that

$$d(\mathrm{T}^N(g'),\mathrm{T}^N(h'))>\delta-\frac{\delta}{5^4}\text{ if }m=k\text{ and}$$
$$d(\mathrm{T}^N(g'),\mathrm{T}^N(h'))>\delta-\delta\sum_{l=4}^{k-m+3}5^{-l}\text{ if }m\in[1,k).$$

Let $g,h\in G$ be such that

$$\mathrm{T}^{(m-1)(N+N')}(x_{i_1i_2\dots i_{k+1}})=\Gamma g\text{ and }\mathrm{T}^{(m-1)(N+N')}(x_{j_1j_2\dots j_{k+1}})=\Gamma h$$

with

$$d(g,g')<\frac{\delta}{5^5}e^{[-(k-m)(N+N')-N](d+1)/d},$$
$$d(h,h')<\frac{\delta}{5^5}e^{[-(k-m)(N+N')-N](d+1)/d}.$$

This can be done using (4.3.10) and (4.3.11) with $l = -N$. In particular,

$$d(\mathrm{T}^N(g), \mathrm{T}^N(g')) < \frac{\delta}{5^5} e^{-(k-m)(N+N')(d+1)/d},$$
$$d(\mathrm{T}^N(h), \mathrm{T}^N(h')) < \frac{\delta}{5^5} e^{-(k-m)(N+N')(d+1)/d}.$$

Also, since by construction

$$\mathrm{T}^{(m-1)(N+N')}(x_{i_1 i_2 \ldots i_{k+1}}), \mathrm{T}^{(m-1)(N+N')}(x_{j_1 j_2 \ldots j_{k+1}}) \in X_{<17M}$$

and since $\frac{\delta}{5^5} e^{[-(k-m)(N+N')-N](d+1)/d}$ is less than the injectivitiy radius δ for $X_{<17M}$ we have

$$d\left(\mathrm{T}^{(m-1)(N+N')}(x_{i_1 i_2 \ldots i_{k+1}}), \mathrm{T}^{(m-1)(N+N')}(x_{i_1 i_2 \ldots i_k})\right) = d(g, g') \text{ and}$$
$$d\left(\mathrm{T}^{(m-1)(N+N')}(x_{j_1 j_2 \ldots j_{k+1}}), \mathrm{T}^{(m-1)(N+N')}(x_{j_1 j_2 \ldots j_k})\right) = d(h, h').$$

Now, if $m = k$ then

$$\begin{aligned} d(\mathrm{T}^N(g), \mathrm{T}^N(h)) &\geq d(\mathrm{T}^N(g'), \mathrm{T}^N(h')) - d(\mathrm{T}^N(g'), \mathrm{T}^N(g)) - d(\mathrm{T}^N(h'), \mathrm{T}^N(h)) \\ &> \delta - \frac{\delta}{5^4} - \frac{\delta}{5^4} - \frac{\delta}{5^4} \\ &= \delta - \frac{\delta}{5^3} \\ &= \delta - \delta \sum_{l=3}^{k+1-m+2} 5^{-l}. \end{aligned}$$

Otherwise, if $m < k$ then

$$\begin{aligned} d(\mathrm{T}^N(g), \mathrm{T}^N(h)) &\geq d(\mathrm{T}^N(g'), \mathrm{T}^N(h')) - d(\mathrm{T}^N(g'), \mathrm{T}^N(g)) - d(\mathrm{T}^N(h'), \mathrm{T}^N(h)) \\ &> \delta - \delta \sum_{l=3}^{k-m+2} 5^{-l} - 2\frac{\delta}{5^5} e^{-(k-m)(N+N')(d+1)/d} \\ &> \delta - \delta \sum_{l=3}^{k-m+2} 5^{-l} - \delta \cdot 5^{-(k-m+3)} \\ &= \delta - \delta \sum_{l=3}^{k+1-m+2} 5^{-l}. \end{aligned}$$

Finally, from (4.3.10) with $m = 1$ and $l = -N$ we have

$$d(x_{i_1 i_2 \ldots i_k}, x_{i_1 i_2 \ldots i_{k+1}}) < \frac{\delta}{5^5} e^{(-N-(k-1)(N+N'))(d+1)/d} < \delta e^{-k}$$

which concludes the proof. $\square$

CHAPTER 5

ENTROPY AND ESCAPE OF MASS FOR HILBERT MODULAR SPACES

For this chapter we let $G = \prod_{n=1}^{r} \mathrm{SL}_2(\mathbb{R}) \times \prod_{m=1}^{s} \mathrm{SL}_2(\mathbb{C})$, $\Gamma = \mathrm{SL}_2(\mathcal{O})$ and consider the quotient space $X = \Gamma \backslash G$ and transformation T acting on X as a right multiplication by the diagonal element α introduced in § 1.3. Our goal is to prove Theorem 1.7.

In the next section we will consider some basic facts. In § 5.2 we introduce two partitions and count the number of elements in these partitions. In § 5.3 we obtain the main proposition and finally in § 5.4 we prove Theorem 1.7 using the partitions and the main proposition.

5.1 Basic facts

Definition. A vector v in a lattice Λ is said to be *primitive* if $\lambda u = v$ with $\lambda \in \mathcal{O}$ and $u \in \Lambda$ implies that λ is a unit.

Lemma 5.1. *Up to units, for any lattice $\Lambda \in X$ there can be at most one primitive (short) vector of norm < 1.*

Having only one short vector is crucial throughout the chapter. As we saw in chapter 3, it is more involved to deal with spaces that allow more than one primitive short vectors.

Proof. Assume by contradiction that there are two primitive vectors $u, v \in \Lambda$ such that $\|u\| < 1, \|v\| < 1$. Let $u = (u'_1, u''_1) \times (u'_2, u''_2) \times \cdots \times (u'_{r+s}, u''_{r+s})$ and $v = (v'_1, v''_1) \times (v'_2, v''_2) \times \cdots \times (v'_{r+s}, v''_{r+s})$. Let $w = (w'_1, w''_1) \times (w'_2, w''_2) \times \cdots \times (w'_{r+s}, w''_{r+s})$ be such that v, w generate Λ over $\mathcal{O}$ as a submodule and satisfy the property (ii). There are $\lambda_1, \lambda_2 \in \mathcal{O}$ such that $u = \lambda_1 v + \lambda_2 w$. Now, we have

$$\prod_{i=1}^{r+s} \det \begin{pmatrix} u'_j & u''_j \\ v'_j & v''_j \end{pmatrix} = \prod_{i=1}^{r+s} \det \begin{pmatrix} \sigma_j(\lambda_1)v'_j + \sigma_j(\lambda_2)w'_j & \sigma_j(\lambda_1)v''_j + \sigma_j(\lambda_2)w''_j \\ v'_j & v''_j \end{pmatrix}$$

where $\{\sigma_1, ..., \sigma_{r+s}\} = S^{\infty}$. Thus,

$$\prod_{j=1}^{r+s}(u'_j v''_j - u''_j v'_j) = \prod_{j=1}^{r+s} \sigma_j(\lambda_2)(w'_j v''_j - w''_j v'_j).$$

On the other hand, from property (ii) of the generators we have that

$$\prod_{j=1}^{r+s}(w'_j v''_j - w''_j v'_j) = (-1)^{r+s} = \pm 1.$$

Also, we have $\prod_{j=1}^{r+s} \sigma_j(\lambda_2) = N(\lambda_2) \geq 1$. Thus, we must have

$$|\prod_{j=1}^{r+s}(u'_j v''_j - u''_j v'_j)| \geq 1. \tag{5.1.1}$$

But, Cauchy-Schwartz inequality implies

$$|(u'_j v''_j - u''_j v'_j)| \leq |(u'_j, u''_j)| \cdot |(v'_j, v''_j)|.$$

Hence,

$$|\prod_{j=1}^{r+s}(u'_j v''_j - u''_j v'_j)| \leq \prod_{j=1}^{r+s} |(u'_j, u''_j)| \cdot \prod_{j=1}^{r+s} |(v'_j, v''_j)| = \|u\| \|v\| < 1.$$

Thus, we obtain a contradiction to (5.1.1). Therefore, up to units, there can be at most one primitive short vector of norm < 1. $\square$

Lemma 5.2. *Γ is a lattice in G and $X_{<M}$ is pre-compact.*

Proof. We would like to use Mahler's compactness criterion for the space of unimodular lattices. For this we will embed X into $\mathrm{SL}_{2d}(\mathbb{Z})\backslash\mathrm{SL}_{2d}(\mathbb{R})$ for $d = [F : \mathbb{Q}] = r + 2s$. There exists $\gamma \in F$ such that $1, \gamma, \gamma^2, \ldots, \gamma^{d-1}$ is a basis of F. Let us consider the following embedding of F onto $\mathcal{A}_1 \subset \mathrm{Mat}_d(\mathbb{Q})$. Any $k \in F$ gets sent to the matrix $[k] \in \mathcal{A}_1$ which describes multiplication by k in the basis $1, \gamma, \gamma^2, \ldots, \gamma^{d-1}$, that is, for any other element $l \in F$ we have $\phi(lk) = \phi(l)[k]$ where $\phi : F \to \mathbb{Q}^d$ is the map sending $l \in F$ to the row vector representing l w.r.t. the basis $1, \gamma, \ldots, \gamma^{d-1}$. $\mathcal{A}_1$ is a $\mathbb{Q}$-linear subspace and so there are linear equations that defines $\mathcal{A}_1$. Therefore, the image $\mathcal{A}_1$ is a d-dimensional rational space in $\mathrm{Mat}_d(\mathbb{Q})$ defined by rational linear equations and is a subalgebra of $\mathrm{Mat}_d(\mathbb{Q})$.

Now, we claim $\mathcal{A}_1$ is diagonalizable over $\mathbb{C}$ and for this all we need to show that the generator γ is diagonalizable. It is easy to see that

$$[\gamma] = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -c_0 & -c_1 & \cdots & -c_{d-2} & -c_{d-1} \end{pmatrix}$$

where $\chi_\gamma(t) = t^d + c_{d-1}t^{d-1} + \cdots + c_0$ is the minimal polynomial of γ. Now, it is easy to check that

$$[\gamma](1, \gamma, \gamma^2, \ldots, \gamma^{d-1})^T = \gamma(1, \gamma, \gamma^2, \ldots, \gamma^{d-1})^T$$

so that γ is an eigenvalue for the matrix $[\gamma]$ with eigenvector $(1, \gamma, \gamma^2, \ldots, \gamma^{d-1})^T$. On the other hand, we know that $\det([\gamma] - xI)$ is in $\mathbb{Q}[x]$. Together we see that $\sigma(\gamma)$ is an eigenvalue for $[\gamma]$ with the eigenvector $(1, \sigma(\gamma), \sigma(\gamma)^2, \ldots, \sigma(\gamma)^{d-1})^T$ for any $\sigma \in S^\infty$. Since γ is the generator and since there $r + 2s$ distinct embeddings of F we conclude that there exists $g_\infty \in \mathrm{SL}_d(\mathbb{C})$ such that $g_\infty^{-1}[\gamma]g_\infty$ is diagonal and hence $g_\infty^{-1}\mathcal{A}_1 g_\infty$ consists of diagonal matrices. We note here an important property that for any $k \in F$

the first row of $[k]$ is the element k itself written as a row vector in the given basis and hence the rest of $[k]$ is determined by $\mathbb{Q}$-linear combination of the first row. Thus, for any $n = 2, \ldots, d$ and for any $m = 1, \ldots, d$ there exists a $\mathbb{Q}$-linear function $q_{n,m}$ such that the (n, m)'th entry is $q_{n,m}(k)$ and these functions are the same for any element of F.

Let $\mathcal{A}_2 = \mathcal{A}_1 \otimes_{\mathbb{Q}} \mathrm{Mat}_2(\mathbb{Q}) \subset \mathrm{Mat}_{2d}(\mathbb{Q})$ then $\mathcal{A}_2$ is defined by rational equations in $\mathrm{Mat}_{2d}(\mathbb{Q})$. We will think of elements of $\mathcal{A}_2$ as matrices in $\mathrm{Mat}_d(\mathbb{Q})$ with entries in $\mathrm{Mat}_2(\mathbb{Q})$. In other words, for us any element of $\mathcal{A}_2$ is just a block matrix consisting of 2×2 matrices.

Let h_∞ be the $2d \times 2d$ matrix where each entry $g_{n,m}$ of g_∞ is replaced by the 2×2 matrix $\begin{pmatrix} g_{n,m} & 0 \\ 0 & g_{n,m} \end{pmatrix}$, that is $h_\infty = g_\infty \otimes I_2$. Let us identify any element $g = (g_1, g_2, \ldots, g_{r+s}) \in G$ by $diag(g_1, \ldots, g_{r+s}, \bar{g}_{r+1}, \ldots, \bar{g}_{r+s}) \in \mathrm{SL}_{2d}(\mathbb{C})$. Define

$$H = \{k \in \mathcal{A}_2 : h_\infty^{-1}\, k\, h_\infty \in G\}.$$

$\mathcal{A}_2^\times := \mathcal{A}_2 \cap \mathrm{GL}_{2d}$ is an algebraic group defined over $\mathbb{Q}$. Notice that for any $k \in \mathcal{A}_2$ since $h_\infty^{-1} k h_\infty$ is a matrix consisting of 2×2 blocks in the diagonal, if $k \in \mathcal{A}_2^\times$ then each 2×2 block in the diagonal of $h_\infty^{-1} k h_\infty$ is invertible. In particular, we have that $[\mathcal{A}_2^\times, \mathcal{A}_2^\times] \subset H$. Since G is semisimple we deduce that $H = [H, H] \subset [\mathcal{A}_2^\times, \mathcal{A}_2^\times]$ and hence $[\mathcal{A}_2^\times, \mathcal{A}_2^\times] = H$. Therefore, H is a semisimple algebraic group defined over $\mathbb{Q}$. Thus, from Proposition 10.15 in [Ra] we deduce that the natural map $H(\mathbb{Z})\backslash H(\mathbb{R}) \to \mathrm{SL}_{2d}(\mathbb{Z})\backslash \mathrm{SL}_{2d}(\mathbb{R})$ is proper. On the other hand, under the conjugation by h_∞ we note that $H(\mathbb{R})$ (resp. $H(\mathbb{Z})$) is sent to G (resp. Γ) so that $X = \Gamma\backslash G$ is isomorphic to $H(\mathbb{Z})\backslash H(\mathbb{R})$. In particular, Γ is a lattice in G. Now, under this isomorphism let $X_{<M}$ correspond to $K \subset H(\mathbb{Z})\backslash H(\mathbb{R})$.

We assume by contradiction that $X_{<M}$ is not pre-compact and hence K is not pre-compact in $H(\mathbb{Z})\backslash H(\mathbb{R})$. Then by Mahler's compactness criterion together with the

fact that the map $H(\mathbb{Z})\backslash H(\mathbb{R}) \to \mathrm{SL}_{2d}(\mathbb{Z})\backslash \mathrm{SL}_{2d}(\mathbb{R})$ is proper there exist a sequence $(H(\mathbb{Z})g_j)_{j\geq 0} \subset K$ such that the length of the shortest row vector in g_j tends to 0 as n increases. Then we claim that $h_\infty g_j h_\infty^{-1}$ has a short vector when considered as an element of G with respect to the norm $\|\cdot\|$ introduced in the introduction. For, as before we think of g_j as a $d \times d$ matrix with entries consisting of 2×2 blocks. So, let $g_j = [g_{n,m}^{(j)}]$ where $g_{n,m}^{(j)} \in \mathrm{Mat}_2(\mathbb{R})$. Then for any $l \in [1, d]$ the 2×2 block in the l'th diagonal entry of $h_\infty g_j h_\infty^{-1}$ is of the form $\sum_{n,m} c_{n,m}^{(l)} g_{n,m}^j$ where $c_{n,m}^{(l)}$ are coming from the matrix g_∞. We also note that the constants $c_{n,m}^{(l)}$ do not depend on the particular matrix g_j. Also, we recall that the first row blocks in g_j determines the rest of g_j by the linear functions $q_{n,m}$. Thus, if the first row vector of g_j is the short vector then the first row vectors of diagonal blocks in $h_\infty g_n h_\infty^{-1}$ are all small. In particular, the product of the length of the first $r+s$ vectors in the rows of the first $r+s$ diagonal blocks (which is the norm of a vector in G) is short. This proves the last claim. However, $X_{<M}$ cannot have very short vectors. A contradiction. $\square$

5.2 Partitions

In order to obtain Theorem 1.7 we need an upper estimate for the metric entropy. The formula to calculate the upper bound is given in Lemma 2.5. It roughly uses the counting argument, that is

$$H(\xi) \leq \log |\xi|$$

where $|\xi|$ is the number of elements of the partition ξ. In this section we introduce some partitions of X and calculate the upper estimate for their cardinality.

Recall the diagonal element α introduced in the introduction.:

$$\alpha = \begin{pmatrix} e^{i\theta_1}e^{a_1/2} & 0 \\ 0 & e^{-i\theta_1}e^{-a_1/2} \end{pmatrix} \times \cdots \times \begin{pmatrix} e^{i\theta_{r+s}}e^{a_{r+s}/2} & 0 \\ 0 & e^{-i\theta_{r+s}}e^{-a_{r+s}/2} \end{pmatrix} \in G.$$

From now on, for simplicity of notation, we assume that $a_j \geq 0$ for any $j \in [1, r+s]$.

The initial partition

For given $M, N > 0$ define a partition

$$Q_{M,N} := \bigvee_{n=-(N-1)}^{N-1} \mathrm{T}^{-n}\{X_{<M}, X_{\geq M}\}.$$

Lemma 5.3. *For $M \geq e^{h_r+h_s}$, the partition $Q_{M,N}$ has $\ll e^{O(\frac{\log\log M}{\log M})N}$ elements.*

Proof. For any x, the partition element of $Q_{M,N}$ containing x describes the time moments in $[-N+1, N-1]$ for which x stays above height M (and hence when it is below height M) under the action of T. So, we need to calculate the possible configurations of times in $[-N+1, N-1]$. Our main tool to calculate the upper bound for the possible configurations is Lemma 5.1. If there is a time when a lattice x (under the action of T) is above height M then there is a considerable gap until the next time (if any) when x reaches height M again. This is because the vectors in x can get short (under the action of T) at most once and for another vector in x to become short the earlier vector has to become of norm 1 at least. Now, we explicate the above discussion. Assume that for a vector $v \in (\mathbb{R}^2)^r \times (\mathbb{C}^2)^s$ we have $\|v\| = |(v_1', v_1'')| \cdot |(v_2', v_2'')| \cdots |(v_{r+s}', v_{r+s}'')| > 1$. We would like to know the soonest possible time n when this vector v reaches the norm $\leq 1/M$ under the action of T. It is easy to see that the best possible n occurs for example when $v_j' = 0$ for $j = 1, ..., r+s$. In this case, at time n we must have

$$|(0, v_1'' e^{-(i\theta_1 + a_1/2)n})| \cdot |(0, v_2'' e^{-(i\theta_2 + a_2/2)n})| \cdots |(0, v_{r+s}'' e^{-(i\theta_{r+s} + a_{r+s}/2)n})| \leq \frac{1}{M}.$$

Since $\|v\| = |v_1''| \cdots |v_{r+s}''| > 1$, we must have $e^{-(a_1+\cdots+a_{r+s})n/2} < 1/M$ which gives

$$n \geq \frac{2\log M}{h_r + h_s}.$$

Similarly, for a vector of norm at most $1/M$, under the action of T, the soonest possible time moment to become of norm greater than 1 is again $\geq \frac{2\log M}{h_r+h_s}$. We also note that for any vector v in x if the sequence $(\|\mathrm{T}^n(v)\|)_{n\geq 0}$ gets increased at some time then it becomes monotone increasing from that time moment. Thus, in a time interval of length $2\lfloor\frac{2\log M}{h_r+h_s}\rfloor$, for any lattice x in X there can be at most one time interval on which x stays above height M. Hence, $Q_{M,\lfloor\frac{2\log M}{h_r+h_s}\rfloor}$ has at most $\binom{2\lfloor\frac{2\log M}{h_r+h_s}\rfloor}{2} \ll \log^2 M$ many elements. On the other hand, to obtain $Q_{M,N}$ we need to take refinements of $\lfloor\frac{2N-1}{2\lfloor\frac{2\log M}{h_r+h_s}\rfloor-1}\rfloor$ many images and pre-images of $Q_{M,\lfloor\frac{2\log M}{h_r+h_s}\rfloor}$ and at most $2\lfloor\frac{2\log M}{h_r+h_s}\rfloor-1$ many of $\{X_{<M}, X_{\geq M}\}$. For $M \geq e^{h_r+h_s}$ we have

$$\left\lfloor\frac{2N-1}{2\lfloor\frac{2\log M}{h_r+h_s}\rfloor-1}\right\rfloor < \frac{2N}{\frac{4\log M}{h_r+h_s}-3} \leq \frac{N(h_r+h_s)}{\log M}$$

Hence, we obtain that $Q_{M,N}$ has $\ll (\log^2 M)^{\frac{N(h_r+h_s)}{\log M}} \leq e^{\frac{2(h_r+h_s)\log\log M}{\log M}N}$ elements. $\square$

The refined partition

Now, we would like to refine the partition $Q_{M,N}$ further by partitioning most of its elements. Let Q be one of its elements. Then there exists $V \subset [-N+1, N-1]$ such that

$$\begin{aligned} Q &:= Q(V) \\ &= \{x \in X : \text{ for all } n \in [-N+1, N-1], \mathrm{T}^n(x) \in X_{\geq M} \text{ if and only if } n \in V\}. \end{aligned} \tag{5.2.1}$$

We can decompose V into subintervals V_m of maximum possible length $V = V_1 \sqcup \ldots \sqcup V_k$. Let $V_m = [b, b+l]$ be one of them for some integers b, l. We know that for any $x \in Q$ there exists a unique primitive vector $v \in \mathrm{T}^{b-1}(x)$ such that

$$\|\mathrm{T}^n(v)\| \leq \frac{1}{M} \text{ for } n \in [1, l+1]. \tag{5.2.2}$$

For any $j \in [1, r+s]$ recall the fixed number a_j appeared in the definition of T. For $j \in [1, r+s]$ let us decompose the extended reals into the following $|V_m| + 2$ subintervals:

$$I_0^{(m,j)} = [-\infty, b], I_{l+1}^{(m,j)} = (b + la_j, \infty], I_j^{(m,j)} = (b + (k-1)a_j, b + ka_j] \text{ for } k \in [1, l]. \tag{5.2.3}$$

We let $\mathcal{I}^{(m,j)} = \{I_0^{(m,j)}, I_1^{(m,j)}, \ldots, I_{l+1}^{(m,j)}\}$. Now, for any $j \in [1, r+s]$ we pick one interval $J^{(m,j)}$ from the set $\mathcal{I}^{(m,j)}$ and consider the product set

$$J_m = J^{(m,1)} \times \cdots \times J^{(m,r+s)}. \tag{5.2.4}$$

Now, the set J_m defines a partition element, could be empty, in Q by

$$Q(J_m) := \{x \in Q : \exists v \in \mathrm{T}^{b-1}(x) \text{ such that (5.2.2) holds and}$$
$$|v_j''| = |v_j'|e^{s_j} \text{ for some } s_j \in J^{(m,j)} - b\}. \tag{5.2.5}$$

For any $m \in [1, k]$ let $Q(J_m)$ be one of the partitions as in (5.2.5) then if we consider their intersection we get a partition element $P(V)$ of X contained in Q :

$$P(V) = \bigcap_{m=1}^{k} Q(J_m). \tag{5.2.6}$$

In this way we obtain a refined partition $P_{M,N}$. The reason why this is the natural way to consider the refined partition elements is related to the action of T. For the motivation see § 5.3, in particular Lemma 5.6.

Lemma 5.4. *For $M > \max\{e^{h_r+h_s}, e^{e(r+s)}\}$, $P_{M,N}$ has at most $\ll e^{O(\frac{\log\log M}{\log M})N}$ many elements.*

Proof. Consider a partition element $Q(V)$ of $Q_{M,N}$ as in (5.2.1). Let $Q(J_m)$ be as in (5.2.5) and $P(V)$ be as in (5.2.6). There are at most $(l+2)^{r+s} = (|V_m|+2)^{r+s}$ many possible ways to choose J_m and hence $(|V_m|+2)^{r+s}$ possible ways to choose $Q(J_m)$

for a fixed $m \in [1,k]$. Thus, the number of partition elements of $P_{M,N}$ contained in $Q(V)$ is

$$(|V_1|+2)^{r+s}(|V_2|+2)^{r+s}...(|V_l|+2)^{r+s}$$
$$= e^{(r+s)\log(|V_1|+2)+\log(|V_2|+2)+\cdots+\log(|V_l|+2)}.$$

This is

$$\ll e^{(r+s)\log(|V_1||V_2|...|V_l|)}.$$

We have

$$|V_1||V_2|...|V_l| \leq \left(\frac{|V_1|+|V_2|+...+|V_l|}{l}\right)^l \leq \left(\frac{2N}{l}\right)^l.$$

Also, note that for the function $f(x) = (\frac{2N}{x})^x = (2N)^x e^{-x\log x}$ its derivative

$$\begin{aligned} f'(x) &= (2N)^x \log(2N) e^{-x\log x} + (2N)^x e^{-x\log x}(-\log x - 1) \\ &= (2N)^x e^{-x\log x}(\log(2N) - \log x - 1). \end{aligned}$$

Hence $f(x) = (\frac{2N}{x})^x$ is increasing on $[1, \frac{2N}{e}]$. On the other hand, from the proof of Lemma 5.3 we know that

$$l \leq \frac{2N-1}{2\lfloor\frac{2\log M}{r+s}\rfloor} + 1 \leq \frac{(r+s)N}{\log M} + 1.$$

If $l=1$ then $(\frac{2N}{l})^l = 2N$. If $l \leq \frac{2(r+s)N}{\log M}$ and for $M \geq e^{e(r+s)}$ we have

$$\left(\frac{2N}{l}\right)^l \leq \left(\frac{2N}{\frac{2(r+s)N}{\log M}}\right)^{\frac{2(r+s)N}{\log M}} = \left(\frac{\log M}{r+s}\right)^{\frac{2(r+s)N}{\log M}}.$$

Hence, the number of partitions contained in $Q(V)$ is $\ll e^{(r+s)\log(2N)}$ if $l=1$ and otherwise it is

$$\ll e^{(r+s)\log\left(\left(\frac{\log M}{r+s}\right)^{\frac{2(r+s)N}{\log M}}\right)} = e^{\frac{2(r+s)^2N}{\log M}(\log\log M - \log(r+s))}.$$

In any case, the number of partitions contained in $Q(V)$ is $\ll_M e^{O(\frac{\log\log M}{\log M})N}$. Thus, together with Lemma 5.3 we have that $P_{M,N}$ has $\ll_M e^{O(\frac{\log\log M}{\log M})N}$ elements. $\square$

5.3 Main proposition

The partition $P_{M,N}$ constructed in the previous section alone will not give us a meaningful conclusion since we do not know if $h_\mu(\mathrm{T}, P_{M,N})$ is close to $h_\mu(\mathrm{T})$ for M, N large. However, considering a further refined partition one can estimate $h_\mu(\mathrm{T})$. Since we only need an upper estimate for the entropy, instead we can consider covers of each partition element of $P_{M,N}$ by small "balls". The right way to do this is to consider the covers by Bowen balls (c.f. Lemma 2.5). We note that the Bowen balls are balls in a different topology. In this section we calculate the number of Bowen balls to cover each partition element of $P_{M,N}$.

Define a Bowen N-ball (of radius η) to be the translate xB_N for some $x \in X$ of

$$B_N = \bigcap_{n=-N+1}^{N-1} a^{-n} B_\eta^G a^n$$

where $\eta > 0$ is such that the log map from B_η^G to the lie algebra of G is injective.

Let $M > 1, N > 1$ be given. Let $P(V)$ be a partition element of $P_{M,N}$ as in (5.2.6) with additional property that $\mathrm{T}^{-N+1}(P(V)) \subset X_{<M}$. We recall that by definition $V \subset [-N+1, N-1]$ and for all $n \in [-N+1, N-1]$, $\mathrm{T}^n(x) \in X_{\geq M}$ if and only if $n \in V$. In particular, the additional restrictive property above equivalent to V being in $(-N+1, N-1]$.

Proposition 5.5. *The partition $P(V) \in P_{M,N}$ with $\mathrm{T}^{-N+1}(P(V)) \subset X_{<M}$ can be covered by $\ll_M c_0^{\frac{2(h_r+h_s)N}{\log M}} e^{2(h_r+2h_s)N - \frac{h_r+h_s}{2}|V|}$ Bowen N-balls for some universal constant $c_0 \geq 1$.*

Roughly, we note that since the number of elements of $P_{M,N}$ is slow exponential as $N \to \infty$, to calculate the entropy it is sufficient to consider the covers of each partition element $P_{M,N}$ by Bowen balls. Since we only need to count the number of covers of most of the space X (cf. Lemma 2.5) it is reasonable to consider only the partitions

$P(V) \in P_{M,N}$ with $\mathrm{T}^{-N+1}(P(V)) \subset X_{<M}$. Since the maximum entropy of T is $h_r + 2h_s$ it is not hard to show that each such partition element $P(V)$ can be covered by $\ll e^{2(h_e+2h_s)N}$ Bowen-N balls. Thus, the significant factor in Proposition 5.5 is $e^{-\frac{h_r+h_s}{2}|V|}$. Before we start proving Proposition 5.5 we need some preliminary preparations.

Restrictions of perturbations

If there are two points in $X_{<M}$ which are η-close to each other such that they both stay above height M for some time interval, then we would like to say that these points must be even closer to each other in the unstable direction U^+. This is not true in general. However, if additionally we know that they are in the same partition element of $P_{M,N}$ then we will show that this is possible.

As before let U^+, U^-, A be the unstable, stable, and centralizer subgroups of G w.r.t. a respectively. We naturally embed U^+ into $\mathbb{R}^r \times \mathbb{C}^s$. We let $u^+(\mathbf{t}) \in U^+$ be the element that corresponds to $\mathbf{t} = (t_1, t_2, \ldots, t_{r+s}) \in \mathbb{R}^r \times \mathbb{C}^s$. Let $P(V) \in P_{M,N}$ be given and let us decompose V into disjoint intervals V_j of maximum possible length. Let V_m be one of them and assume that $V_m = [b, b+l]$. As in (5.2.6) we have $P(V) = \bigcap_{m=1}^{k} Q(J_m)$ for some $Q(J_m)$ as in (5.2.5), namely

$$Q(J_m) := \{x \in Q : \exists v \in \mathrm{T}^{b-1}(x) \text{ such that (5.2.2) holds and}$$
$$|v_j''| = |v_j'|e^{s_j} \text{ for some } s_j \in J^{(m,j)} - b\}.$$

Lemma 5.6. *Let $x, y \in P(V) \cap \mathrm{T}^{N-1}(X_{<M})$ with $\mathrm{T}^{b-1}(y) \in \mathrm{T}^{b-1}(x)u^+(\mathbf{t})g$ for some $u^+(\mathbf{t}) \in B_{\eta/2}^{U^+}$ and $g \in B_{\eta/2}^{U^-A}$. Then for any $j \in \{1, 2, \ldots, r+s\}$ we have $|t_j| \ll e^{b-n_j}$ where n_j is the left end point of the interval $J^{(m,j)}$.*

Proof. If $J^{(m,j)} = [-\infty, b] = I_0^{(m,j)}$ then $n_j = -\infty$ and in this case the lemma is trivial. So, we can assume that $J^{(m,j)} \neq I_0^{(m,j)}$ so that $n_j \geq b$.

By maximality of V_m we know that $\mathrm{T}^{b-1}(x), \mathrm{T}^{b-1}(y) \in X_{<M}$ and

$$\mathrm{T}^n(\mathrm{T}^{b-1}(x)), \mathrm{T}^n(\mathrm{T}^{b-1}(y)) \in X_{\geq M} \text{ for any } n \in [1, l+1].$$

Thus there exist vectors $v \in \mathrm{T}^{b-1}(x)$ and $w \in \mathrm{T}^{b-1}(y)$ such that (5.2.2) holds. On the other hand, from (5.2.5) for v, w in the standard notation we know that

$$|v_j''| = |v_j'|e^{s_j} \text{ and } |w_j''| = |w_j'|e^{r_j} \text{ for some } s_j, r_j \in J^{(m,j)} - b.$$

We note that $v_j'' \neq 0 \neq w_j''$ since $(v_j', v_j''), (w_j', w_j'') \neq (0,0)$ (they are rows of matrices of determinant equal to 1) and $s_j, r_j \geq 0$. In particular, if n_j is the left end point of the interval $J^{(m,j)}$ then we have

$$\frac{|v_j'|}{|v_j''|} \leq e^{b-n_j} \text{ and } \frac{|w_j'|}{|w_j''|} \leq e^{b-n_j}. \tag{5.3.1}$$

Also, we know that $w = vu^+(\mathbf{t})g$. So, for $g = (g_1, \ldots, g_{r+s})$ we have $(w_j', w_j'') = (v_j', v_j'') \begin{pmatrix} 1 & 0 \\ t_j & 1 \end{pmatrix} g_j = (v_j' + t_j v_j'', v_j'')g_j$ (under the assumption that $a_j \geq 0$ where a_j is as in the definition of a). For $g_j = \begin{pmatrix} d & u \\ 0 & 1/d \end{pmatrix}$ we obtain that

$$(w_j', w_j'') = (d(v_j' + t_j v_j''), u(v_j' + t_j v_j'') + v_j''/d).$$

Now from (5.3.1) we get

$$e^{b-n_j} \geq \frac{|w_j'|}{|w_j''|} = \frac{|d(v_j' + t_j v_j'')|}{|u(v_j' + t_j v_j'') + v_j''/d|} \gg \frac{|v_j' + t_j v_j''|}{|v_j''|} = \left| \frac{v_j'}{v_j''} + t_j \right|$$

since d is close to 1 and u is close to 0. Together with (5.3.1) we deduce that

$$|t_j| \ll e^{b-n_j}.$$

$\square$

Lemma 5.6 alone does not tell us if x, y should be even closer to each other in the unstable direction since for example n_j could be equal to b. Even if $n_j > b$ we still do not know an effective lower bound for n_j. This is because we have only considered one part of the defining properties of $Q(J_m)$. We have not considered the fact that x, y stay above height M in $[1, b+1]$. In the next lemma we use this fact to obtain the relation among the intervals $J^{(m,j)}$.

As before, let $P(V) \in P_{M,N}$ be given with $\mathrm{T}^{-N+1}(P(V)) \subset X_{<M}$ and let us decompose V into disjoint intervals V_m of maximum possible length. Let $V_m = [b, b+l]$ be one of them. From (5.2.6) we have $P(V) = \bigcap_{m=1}^{k} Q(J_m)$.

Lemma 5.7. *Let J_m be as in* (5.2.4) *and consider $x \in Q(J_m)$ with $v \in \mathrm{T}^{b-1}(x)$ as in* (5.2.5)*. Let $S = \{s_1, ..., s_{r+s}\}$. Let $i_1, ..., i_L$ be the subset of S which are ≤ 0, let $j_1, ..., j_C$ be the subset of S such that $s_{j_i} \in (0, (l+1)a_{j_i})$, and let $k_1, ..., k_R$ be the subset of S such that $s_{k_i} > (l+1)a_{k_i}$. In particular, $L + C + R = r + s$. Then*

$$(l+1)(a_{i_1} + \cdots + a_{i_L} + a_{j_1} + \cdots + a_{j_C} - a_{k_1} - \cdots - a_{k_R}) < 2(s_{j_1} + \cdots + s_{j_C}).$$

Proof. Let us consider the j-th component vector (v'_j, v''_j) of v. T acts on v and hence it acts on each of its components and we have

$$\mathrm{T}^n((v'_j, v''_j)) = (v'_j e^{in\theta_j} e^{na_j/2}, v''_j e^{-in\theta_j} e^{-na_j/2})$$

where as before $\theta_j = 0$ if $j \leq r$, and $a_j \geq 0$ for any $j \in [1, r+s]$.Thus,

$$|\,\mathrm{T}^n((v'_j, v''_j))| = \max\{|v'_j e^{na_j/2}|, |v''_j e^{-na_j/2}|\} = \begin{cases} |v''_j| e^{-na_j/2} & \text{if } na_j < s_j \\ |v'_j| e^{na_j/2} & \text{if } na_j \geq s_j \end{cases}$$

since $|v'_j| e^{s_j/2} = |v''_j| e^{-s_j/2}$. We also note that

$$|(v'_j, v''_j)| = \begin{cases} |v'_j| & \text{if } s_j \leq 0 \\ |v''_j| & \text{if } s_j > 0 \end{cases}.$$

Together we get

$$\frac{|\,\mathrm{T}^{l+1}((v'_j, v''_j))|}{|(v'_j, v''_j)|} = \begin{cases} e^{\frac{(l+1)a_j}{2}} & \text{if } s_j \leq 0 \\ e^{\frac{(l+1)a_j}{2} - s_j} & \text{if } s_j \in (0, (l+1)a_j] \\ e^{-\frac{(l+1)a_j}{2}} & \text{if } s_j > (l+1)a_j. \end{cases} \tag{5.3.2}$$

By the assumption (5.2.2) on the vector $v \in (\mathbb{R}^2)^r \times (\mathbb{C}^2)^s$ we have

$$\|v\| > \frac{1}{M} \text{ and } \|\,\mathrm{T}^n(v)\| \leq \frac{1}{M} \text{ for } n \in [1, l+1].$$

In particular, this gives

$$\frac{\|\,\mathrm{T}^{l+1}(v)\|}{\|v\|} < 1. \tag{5.3.3}$$

Now, from (5.3.2) and (5.3.3) we get

$$\begin{aligned} &\frac{\prod_{j=1}^{r+s} |\,\mathrm{T}^{l+1}((v'_j, v''_j))|}{\prod_{j=1}^{r+s} |(v'_j, v''_j)|} \\ &\quad = \exp\left(\frac{(l+1)(a_{i_1} + \cdots + a_{i_L})}{2} + \frac{(l+1)(a_{j_1} + \cdots + a_{j_C})}{2} - s_{j_1} - \cdots - s_{j_C}\right) \times \\ &\qquad\qquad \times \exp\left(-\frac{(l+1)(a_{k_1} + \cdots + a_{k_R})}{2}\right) < 1. \end{aligned}$$

The exponent simplifies to

$$(l+1)(a_{i_1} + \cdots + a_{i_L} + a_{j_1} + \cdots + a_{j_C} - a_{k_1} - \cdots - a_{k_R}) < 2(s_{j_1} + \cdots + s_{j_C}).$$

□

The next lemma shows how we apply the above two lemmas. The reader can skip the lemma and come back when it is mentioned in the proof of Proposition 5.5. Recall the embedding of U^+ into $\mathbb{R}^r \times \mathbb{C}^s$.

Lemma 5.8. *Let $V_m = [b, b+l]$ and $Q(J_m)$ be as before and let C', C'' be given positive constants. Let us consider the set $D := \{u(\mathbf{t}) \in U^+ : |t_j| < C' \min\{\eta, e^{b-n_j}\}, j = 1, \ldots, r+s\}$ where n_j is the left end point of the interval $J^{(m,j)}$. Then the set D can* *be decomposed into* $\ll e^{(\frac{h_r+h_s}{2}+h_s)(l+1)}$ *many disjoint sets of the form* $E := \{u(\mathbf{t}) \in U^+ : |t_j| < C''\eta e^{-la_j}, j = 1, \ldots, r+s\}$.

Proof. For any $j = \{1, 2, \ldots, r+s\}$, let us consider the ball around 0 of radius $C' \cdot \min\{\eta, e^{b-n_j}\}$ in $\mathbb{R}$ or in $\mathbb{C}$ depending whether $j \leq r$ or not and decompose it into the small balls of radius $C''\eta e^{-l}$. If $n_j < b$ (in which case $n_j = -\infty$) then there are $\ll e^{la_j}$ small subintervals if $j \leq r$ and there are $\ll e^{2la_j}$ small balls if $j > r$. Suppose $n_j \geq b$. If $j \leq r$ then there are $\ll e^{la_j+b-n_j}$ small subintervals and if $j > r$ then there are $\ll e^{2(la_j+b-n_j)}$ small balls. We note that if $n_j \geq b + la_j$ (in which case $n_j = b + la_j$) then there are $\ll 1$ small subintervals or $\ll 1$ small balls depending on j. We have $i_1, ..., i_L, j_1, ..., j_C, k_1, ..., k_R$ as in Lemma 5.7. Now, let $i'_1, ..., i'_{L'}$ be the subset of $\{i_1, ..., i_L\}$ which are $\leq r$ and $i''_1, ..., i''_{L''}$ be the rest. Similarly, we consider the subsets $j'_1, ..., j'_{C'}$ and $j''_1, ..., j''_{C''}$ of $j_1, ..., j_C$.
Therefore, the set D can contain at most

$$\begin{aligned}
&\ll \exp(l(a_{i'_1} + \cdots + a_{i'_{L'}}) + 2l(a_{i''_1} + \cdots + a_{i''_{L''}}))(1)^R \times \\
&\quad \times \exp(l(a_{j'_1} + \cdots + a_{j'_{C'}}) + bC' - n_{j'_1} - \cdots - n_{j'_{C'}}) \times \\
&\quad \times \exp(2(l(a_{j''_1} + \cdots + a_{j''_{C''}}) + bC'' - n_{j''_1} - \cdots - n_{j''_{C''}})) \\
&= \exp((a_{i'_1} + \cdots + a_{i'_{L'}} + a_{j'_1} + \cdots + a_{j'_{C'}})l) \times \\
&\quad \times \exp(2(a_{i''_1} + \cdots + a_{i''_{L''}} + a_{j''_1} + \cdots + a_{j''_{C''}}))l) \times \\
&\quad \times \exp(b(C' + 2C'') - n_{j'_1} - \cdots - n_{j'_{C'}} - 2n_{j''_1} - \cdots - 2n_{j''_{C''}}) \\
&= \exp((a_{i_1} + \cdots + a_{i_L} + a_{j_1} + \cdots + a_{j_C})l - n_{j_1} - \cdots - n_{j_C} + bC) \times \\
&\quad \times \exp((a_{i''_1} + \cdots + a_{i''_{L''}} + a_{j''_1} + \cdots + a_{i''_{C''}})l - n_{j''_1} - \cdots - n_{j''_{C''}} + bC'')
\end{aligned}$$

many disjoint sets of the form E. On the other hand, Lemma 5.7 gives

$$(l+1)(a_{i_1} + \cdots + a_{i_L} + a_{j_1} + \cdots + a_{j_C} - a_{k_1} - \cdots - a_{k_R}) < 2(s_{j_1} + \cdots + s_{j_C}).$$

where $s_{j_k} \in J^{(m,j_k)} - b = (n_{j_k} - b, n_{j_k} + a_{j_k} - b]$. Thus,

$$(l+1)(a_{i_1} + \cdots + a_{i_L} + a_{j_1} + \cdots + a_{j_C} - a_{k_1} - \cdots - a_{k_R}) \\ < 2(n_{j_1} + \cdots + n_{j_C}) - 2bC + 2(a_{j_1} + \cdots + a_{j_C})$$

and since $a_{i_1} + \cdots + a_{i_L} + a_{j_1} + \cdots + a_{j_C} + a_{k_1} + \cdots + a_{k_R} = h_r + h_s$ we obtain

$$(l+1)(2(a_{i_1} + \cdots + a_{i_L} + a_{j_1} + \cdots + a_{j_C}) - h_r - h_s) \\ < 2(n_{j_1} + \cdots + n_{j_C}) - 2bC + 2(a_{j_1} + \cdots + a_{j_C}).$$

This gives

$$(a_{i_1} + \cdots + a_{i_L} + a_{j_1} + \cdots + a_{j_C})l - n_{j_1} - \cdots - n_{j_C} + bC \\ \leq \frac{-2(a_{i_1} + \cdots + a_{i_L}) + (h_r + h_s)(l+1)}{2}.$$

Hence, the set D can be decomposed into

$$\ll \exp(\frac{-2(a_{i_1} + \cdots + a_{i_L}) + (h_r + h_s)(l+1)}{2})\times \\ \times \exp((a_{i''_1} + \cdots + a_{i''_{L''}} + a_{j''_1} + \cdots + a_{i''_{C''}})l - n_{j''_1} - \cdots - n_{j''_{C''}} + bC'') \\ \ll e^{\frac{h_r+h_s}{2}(l+1)} \exp((a_{i''_1} + \cdots + a_{i''_{L''}} + a_{j''_1} + \cdots + a_{i''_{C''}})(l+1) - n_{j''_1} - \cdots - n_{j''_{C''}} + bC'')$$

many disjoint sets of the form E. Now, by definition of n_{j_k} we have $n_{j_k} \geq b$ for $k = 1, \ldots, C$ which implies that $bC'' - n_{j''_1} - \cdots - n_{j''_{C''}} \leq 0$. Also, $a_{i''_1} + \cdots + a_{i''_{L''}} + a_{j''_1} + \cdots + a_{i''_{C''}} \leq h_s$. Thus, D can be covered by

$$\ll e^{(\frac{h_r+h_s}{2} + h_s)(l+1)}$$

disjoint sets of the form E. □

If $r = 0$ one could get a better estimate as saying that the set D can be covered by $\ll e^{h_s l}$ (not just $\ll e^{3h_s l/2}$) sets of the form E. Hence, one could obtain a sharper result when $r = 0$ as was pointed out in the introduction.

The proof of Proposition 5.5

To simplify the proof, by taking images under T^{N-1}, we redefine the notions in § 5.2. For given $M, N > 0$ let $Q^+_{M,N} := \bigvee_{n=0}^{N-1} \mathrm{T}^{-n}\{X_{<M}, X_{\geq M}\}$. Also, define $P^+_{M,N}$ accordingly by restricting to the interval $[0, N-1]$ instead of $[-N+1, N-1]$ and consider the partition element $P^+(V)$ (a substitute for $P(V)$) of $P^+_{M,N}$ which is contained in $X_{<M}$ where $V \subset [0, N-1]$ (in fact, $V \subset (0, N-1])$. We have $P^+(V) = \cap_{j=1}^{k} Q(J_m)^+$ where $Q(J_m)^+$ is defined similar to $Q(J_m)$ with the difference that $Q(J_m)^+$ is contained in a partition element Q^+ of $Q^+_{M,N}$.

Also, since $X_{<M}$ is pre-compact it suffices to restrict ourselves to a neighborhood O of some $x_0 \in X_{<M} \cap P^+(V)$. Let $O = x_0 B^{U^+}_{\eta/2} B^{U^- A}_{\eta/2}$ be a neighborhood of $x_0 \in X_{<M}$. Then it suffices to prove that there exists a constant $c_0 > 0$ such that the set $P^+_O(V) = P^+(V) \cap O$ can be covered by $\ll c_0^{\frac{(h_r+h_s)N}{\log M}} e^{(h_r+2h_s)N - \frac{(h_r+h_s)|V|}{2}}$ many forward Bowen N-balls xB^+_N where

$$B^+_N = \bigcap_{n=0}^{N-1} a^n B^G_\eta a^{-n}.$$

Let us make some observations. If we consider the image of O under T^n we obtain the set

$$\mathrm{T}^n(O) = \mathrm{T}^n(x_0)(a^{-n} B^{U^+}_{\eta/2} a^n) a^{-n} B^{U^- A}_{\eta/2} a^n.$$

We see that the jth component of the U^+-part gets stretched by the factor e^{na_j}. Here again we naturally embed U^+ into $\mathbb{R}^r \times \mathbb{C}^s$. Under this identification, dividing $(a^{-n} B^{U^+}_{\eta/2} a^n)$ into $\prod_{j=1}^{r} \lceil e^{na_j} \rceil \prod_{j=r+1}^{r+s} (\lceil e^{na_j} \rceil)^2$ many small parts we obtain the sets of the form

$$\mathrm{T}^n(x_0) u^+ B^{U^+}_{\eta/2} a^{-n} B^{U^- A}_{\eta/2} a^n$$

for some $u^+ \in U^+$. Now, if we take the pre-image under T^n of these sets then we obtain the similar sets

$$\mathrm{T}^{-n}(\mathrm{T}^n(x_0) u^+) a^n B^{U^+}_{\eta/2} a^{-n} B^{U^- A}_{\eta/2}$$

as before. It is not hard to see that the set $\mathrm{T}^{-n}(\mathrm{T}^n(x_0)u^+)a^nB^{U^+}_{\eta/2}a^{-n}B^{U^-A}_{\eta/2}$ is contained in the forward Bowen n-ball $\mathrm{T}^{-n}(\mathrm{T}^n(x_0)u^+)B^+_n$. This in particular shows that O can be covered by $\ll e^{(h_r+2h_s)n}$ many forward Bowen n-balls which is the reason why the maximal entropy is h_r+2h_s. Here, we used the fact that $a_1+\cdots+a_r=h_r$ and $a_{r+1}+\cdots+a_{r+s}=h_s$. However, using Lemma 5.8 we will show that we need fewer Bowen balls to cover the set O.

Let us decompose V into disjoint ordered subintervals V_m of maximum length. So, we have

$$V = V_1 \cup V_2 \cup \ldots \cup V_k.$$

Now let $[0,N-1]\setminus V = W_1 \cup W_2 \cup \ldots \cup W_{k'}$ where W_m are maximal intervals. We inductively prove the following:

If $[0,b-1] = V_1 \cup V_2 \cup \ldots \cup V_{m-1} \cup W_1 \cup W_2 \cup \ldots \cup W_{n'}$ then for some constant c_0 the set $P^+_O(V)$ can be covered by $\ll c_0^{m-1+n'}\exp((h_r+2h_s)(b-1) - \frac{(h_r+h_s)(|V_1|+\cdots+|V_n|)}{2})$ many pre-images under T^{b-1} of sets of the form

$$\mathrm{T}^{b-1}(x_0)u^+B^{U^+}_{\eta/2}a^{-b+1}B^{U^-A}_{\eta/2}a^{b-1}. \tag{5.3.4}$$

For the interval $[0,0]$ the claim is obvious. Now, assume that the claim is true for the interval $[0,b-1]$ as above. In the inductive step, if the next interval is $W_{n'+1}$ then once we divide each set obtained earlier into $\prod_{j=1}^{r}\lceil e^{|W_{n'+1}|a_j}\rceil \prod_{j=r+1}^{r+s}(\lceil e^{|W_{n'+1}|a_j}\rceil)^2 \leq c_0e^{(h_r+2h_s)(|W_{n'+1}|)}$ small ones for some constant c_0, we just keep all of them. So, assume that the next interval is $V_m=[b,b+l]$. Let Y be one of the sets (5.3.4) obtained in the earlier step. We divide Y into $\prod_{j=1}^{r}\lceil e^{la_j}\rceil \prod_{j=r+1}^{r+s}(\lceil e^{la_j}\rceil)^2$ many sets of the form

$$\mathrm{T}^{b-1+l}(x_0)u^+(\mathbf{t})B^{U^+}_{\eta/2}a^{-b+1-l}B^{U^-A}_{\eta/2}a^{b-1+l} \tag{5.3.5}$$

for some $\mathbf{t}\in B^{\mathbb{R}^r\times\mathbb{C}^s}_{\eta/2}$. We are interested in the points $x\in Y$ for which $\mathrm{T}^{-b+1}(x)$ is in $Q(J_m)^+$. We know by assumption that x_0 is one of them. If $x\in Y$ is another one then by Lemma 5.6 there exists $\mathbf{t} \in B_{\eta/2}^{\mathbb{R}^r \times \mathbb{C}^s}$ such that $x = x_0 u^+(\mathbf{t}) g$ for some $g \in B_{\eta/2}^{U^- A}$ and for $j \in [1, r+s]$, $|t_j| \ll e^{b-n_j}$ where n_j is the left end point of the interval $J^{(m,j)}$. Hence the set we are interested in corresponds to the set D in Lemma 5.8 and each set as in (5.3.5) corresponds to the set E as in Lemma 5.8. Thus, Lemma 5.8 gives that once we divide Y into the sets of the form as in (5.3.5) we only need to keep $\leq c_0 e^{(\frac{h_r+h_s}{2}+h_s)(l+1)}$ many of them. Here, enlarging if necessary, we assume that $c_0 \geq 1$ is the implicit constant appeared in Lemma 5.8. Hence, we conclude that the set P_O^+ can be covered by

$$\begin{aligned}
&\leq c_0^{m+n'} \exp((h_r+2h_s)(b-1) - \frac{(h_r+h_s)(|V_1|+...+|V_n|)}{2} + (\frac{h_r+h_s}{2}+h_s)(l+1)) \\
&\leq c_0^{m+n'} \exp((h_r+2h_s)(b+l) - \frac{(h_r+h_s)(|V_1|+...+|V_n|+l)}{2})
\end{aligned}$$

many pre-images under T^{K+S} of the sets of the form

$$\mathrm{T}^{b+l}(x_0) u^+(t) B_{\eta/2}^{U^+} a^{-b-l} B_{\eta/2}^{U^- A} a^{b+l}.$$

Since $l = |V_m|$, this completes the inductive step.

Now, let $b = N$ then we see that the set P_O^+ can be covered by

$$\ll c_0^{k+k'} e^{(h_r+2h_s)N - \frac{(h_r+h_s)}{2}|V|}$$

many forward Bowen N-balls. On the other hand, the proof of Lemma 5.3 suggests that m and hence m' is bounded above by

$$\frac{N}{2\lfloor \frac{2\log M}{h_r+h_s} \rfloor} + 1 < \frac{(h_r+h_s)N}{2\log M}.$$

Thus, the set Z_O^+ can be covered by

$$\ll c_0^{\frac{(h_r+h_s)N}{\log M}} e^{(h_r+2h_s)N - \frac{(h_r+h_s)}{2}|V|}$$

many forward Bowen N-balls, which completes the proof.

$\square$

5.4 Proof of Theorem 1.7

Now we will apply Lemma 2.5 together what we obtained in this chapter to prove Thorem 1.7.

Proof of the Theorem 1.7. Note first that it suffices to consider ergodic measures. For if μ is not ergodic, we can write μ as an integral of its ergodic components $\mu = \int \mu_t d\tau(t)$ for some probability space (E, τ) by [EW, Theorem 6.2]. Therefore, we have $\mu(X_{\geq M}) = \int \mu_t(X_{\geq M}) d\tau(t)$, but also $h_\mu(\mathrm{T}) = \int h_{\mu_t}(\mathrm{T}) d\tau(t)$ by [Wa, Thm. 8.4], so that desired estimate follows from the ergodic case.

Suppose that μ is ergodic. We would like to apply Lemma 2.5. For this we need to find an upper bound for covering μ-most of the space X by Bowen N-balls. Except for the points that escape to the cusp, every forward trajectory visits $X_{<M}$ for $M \geq \max\{e^{h_r+h_s}, e^{e(r+s)}\}$ so that $\mu(X_{<M}) > 0$. Thus, ergodicity of μ implies that $\mu(\bigcup_{k=0}^{\infty} \mathrm{T}^{-k} X_{<M}) = 1$. Hence, for every $\epsilon > 0$ there is a constant $K \geq 1$ such that $Y = \displaystyle\bigcup_{k=0}^{K-1} \mathrm{T}^{-k} X_{<M}$ satisfies $\mu(Y) > 1 - \epsilon$. Moreover, the pointwise ergodic theorem implies

$$\frac{1}{2N-1} \sum_{n=-N+1}^{N-1} 1_{X \geq M}(\mathrm{T}^n(x)) \to \mu(X_{\geq M})$$

as $N \to \infty$ for a.e. $x \in X$. Thus, for $\epsilon > 0$ given there is N_0 such that for $N > N_0$ the average on the left will be bigger that $\mu(X_{\geq M}) - \epsilon$ for any $x \in X_1$ for some $X_1 \subset X$ with measure $\mu(X_1) > 1 - \epsilon$. Clearly, for any N we have $\mu(Z) > 1 - 3\epsilon$ where

$$Z = X_1 \cap \mathrm{T}^N Y \cap \mathrm{T}^{-N} Y.$$

Now, we would like to find an upper bound for the number of Bowen N-balls needed to cover the set Z. Here $N \to \infty$ while ϵ and hence K are fixed. Since $Y = \displaystyle\bigcup_{k=0}^{K-1} \mathrm{T}^{-k} X_{<M}$, we can decompose Z into K^2 sets of the form

$$Z' = X_1 \cap \mathrm{T}^{N-k_1} X_{<M} \cap \mathrm{T}^{-N-k_2} X_{<M}$$

but since K is fixed, it suffices to find an upper bound for the number of Bowen N-balls to cover one of these. Consider the set Z'. Since $k_1, k_2 \leq K$ without lost of generality we can assume $k_1 = k_2 = 0$. Next we split Z' into the sets $P(V)$ as in Proposition 5.5. By Lemma 5.4 we know that we need $\ll_M e^{O(\frac{\log\log M}{\log M})N}$ many of these. Moreover, by our assumption on X_1 we only need to look at sets $V \subset [-N+1, N-1]$ with $|V| \geq (\mu(X_{\geq M}) - \epsilon)(2N-1)$. On the other hand, Proposition 5.5 gives that each of those sets $P(V)$ can be covered by $\ll_M c_0^{\frac{2(h_r+h_s)}{\log M}N} e^{2(h_r+2h_s)N - \frac{(h_r+h_s)}{2}|V|}$ Bowen N-balls. Together we see that Z can be covered by

$$\ll_{M,K} e^{2(h_r+2h_s)N - \frac{(h_r+h_s)}{2}|V|} e^{O(\frac{\log\log M}{\log M})N} c_0^{\frac{2(h_r+h_s)}{\log M}N}$$

many Bowen N-balls. Applying Lemma 2.5 we arrive at

$$\begin{aligned} h_\mu(\mathrm{T}) &\leq \lim_{\epsilon\to 0}\liminf_{N\to\infty} \frac{\log BC(N,\epsilon)}{2N} \\ &\leq h_r + 2h_s - \frac{(h_r+h_s)(\mu(X_{\geq M}) - \epsilon)}{2} + O(\frac{\log\log M}{\log M}) + \frac{(h_r+h_s)\log c_0}{\log M} \\ &< h_r + 2h_s - \frac{(h_r+h_s)(\mu(X_{\geq M}) - \epsilon)}{2} + O(\frac{\log\log M}{\log M}). \end{aligned}$$

Since $\epsilon > 0$ was arbitrary, we get that

$$h_\mu(\mathrm{T}) \leq h_r + 2h_s - \frac{(h_r+h_s)\mu(X_{\geq M})}{2} + O(\frac{\log\log M}{\log M})$$

which completes the main part of the theorem. The last part is easily deduced from this and is left to the reader. $\square$

CHAPTER 6

MEASURES WITH HIGH LOCAL DIMENSION

This chapter is joint work [EK] with M. Einsiedler. In this chapter we prove Theorem 1.8 and Corollary 1.9. However, we will not prove Theorem 1.10 and its corollary since the proofs follow the same lines as the proofs of Theorem 1.8 and its corollary using Proposition 5.5 instead of Proposition 3.7 . Our main tool is a version of Proposition 3.7. Let $N, M > 0$ be given. For any x we define $V_x \in [0, N-1]$ to be the set of times $n \in [0, N-1]$ for which $\mathrm{T}^n(x) \in X_{\geq M}$. Now, Proposition 3.7 can be rephrased as follows.

Proposition 6.1. *For a fixed set $\mathcal{N} = \mathcal{N}_{[0,N-1]}(x_0)$ of labeled marked times in $[0, N-1]$ we have that the set*

$$Z^+(\mathcal{N}) = \{x \in X_{\leq M} : \mathcal{N}_{[0,N-1]}(x) = \mathcal{N}_{[0,N-1]}\}$$

can be covered by $\ll_M e^{3N-|V_{x_0}|} c_0^{\frac{9N}{[\log M]}}$ many sets of the form

$$\mathrm{T}^{-N}(\mathrm{T}^N(x)u^+) D^{U^+}_{\frac{\eta}{2}e^{-3N/2}} B^{U^-C}_{\frac{\eta}{2}}.$$

Proof. In the proof of Proposition 3.7 we inductively proved that the set

$$Z^+_O = \{x \in O : \mathcal{N}_{[0,N-1]}(x) = \mathcal{N}_{[0,N-1]}\}$$

can be covered by $e^{3N-|V_{x_0}|} c_0^{\frac{9N}{[\log M]}}$ many pre-images under T^N of sets of the form

$$\mathrm{T}^N(x_0) u^+ D^{U^+}_{\eta/2} \alpha^{-N} B^{U^-C}_{\eta/2} \alpha^N.$$

So, Z_O^+ can be covered by the sets of the form

$$\mathrm{T}^{-N}(\mathrm{T}^N(x_0)u^+)\alpha^N D^{U^+}_{\eta/2}\alpha^{-N}B^{U^-C}_{\eta/2}.$$

This completes the proof since we have $\alpha^N D^{U^+}_{\eta/2}\alpha^{-N} = D^{U^+}_{\frac{\eta}{2}e^{-3N/2}}$ and since $X_{\leq M}$ is compact. $\square$

For any $\kappa > 0$ small we are interested in the upper estimate for

$$\nu(\{x \in X_{<M} : |V_x| > \kappa N\}).$$

Proposition 6.1 together with Lemma 3.4 gives the following.

Lemma 6.2. *For any $N > 0$ large we have*

$$\nu(\{x \in X_{\leq M} : |V_x| > \kappa N\}) \ll_M e^{\frac{6-2\kappa-3d+3\delta}{2}N+\frac{9N\log(c_0\log M)}{\log M}}.$$

Proof. From Lemma 3.4 we know that the set $X_{<M}$ can be decomposed into

$$\ll_M e^{\frac{5N\log\lfloor\log M\rfloor}{\lfloor\log M\rfloor}}$$

many sets of the form $Z^+(\mathcal{N})$. We are only interested in those sets of marked times $\mathcal{N}_{[0,N-1]}(x)$ for which $|V_x| > \kappa N$. On the other hand, from Proposition 6.1 we know that such sets can be covered by $e^{(3-\kappa)N}c_0^{\frac{9N}{\lfloor\log M\rfloor}}$ many sets of the form

$$\mathrm{T}^{-N}(\mathrm{T}^N(x)u^+)D^{U^+}_{\frac{\eta}{2}e^{-3N/2}}B^{U^-C}_{\frac{\eta}{2}}.$$

However, from the assumption on dimension of the measure ν we have

$$\nu(\mathrm{T}^{-N}(\mathrm{T}^N(x)u^+)D^{U^+}_{\frac{\eta}{2}e^{-3N/2}}B^{U^-C}_{\frac{\eta}{2}}) \ll (\frac{\eta}{2}e^{-3N/2})^{d-\delta}$$

once N is sufficiently large. Thus,

$$\nu(\{x \in X_{\leq M} : |V_x| > \kappa N\}) \ll_M e^{\frac{5N\log\lfloor\log M\rfloor}{\lfloor\log M\rfloor}} e^{(3-\kappa)N}c_0^{\frac{9N}{\lfloor\log M\rfloor}}(\frac{\eta}{2}e^{-3N/2})^{d-\delta}.$$

This simplifies to

$$\nu(\{x \in X_{\leq M} : |V_x| > \kappa N\}) \ll_M e^{\frac{6-2\kappa-3d+3\delta}{2}N + \frac{9N\log(c_0 \log M)}{\log M}}.$$

□

Proof of Theorem 1.8. In order to prove Theorem 1.8 we need to estimate an upper bound for $\mu_N(X_{\geq M})$ for M, N large. Let us recall that

$$\mu_N = \frac{1}{N}\sum_{i=0}^{N-1} \mathrm{T}_*^i \nu.$$

Hence,

$$\begin{aligned}\mu_N(X_{\geq M}) &= \frac{1}{N}\sum_{n=0}^{N-1} \nu(\mathrm{T}^{-n}(X_{\geq M})) \\ &= \frac{1}{N}\sum_{n=0}^{N-1} \nu(X_{\leq M} \cap \mathrm{T}^{-n}(X_{\geq M})) + \frac{1}{N}\sum_{n=0}^{N-1} \nu(X_{>M} \cap \mathrm{T}^{-n}(X_{\geq M})).\end{aligned}$$

However, we have $\nu(X_{>M}) < \epsilon(M)$ where $\epsilon(M) \to 0$ as $M \to \infty$. Hence,

$$\mu_N(X_{\geq M}) \leq \epsilon(M) + \frac{1}{N}\sum_{n=0}^{N-1} \nu(X_{\leq M} \cap \mathrm{T}^{-n}(X_{\geq M})). \tag{6.0.1}$$

Thus, all we need to estimate is $\frac{1}{N}\sum_{n=0}^{N-1} \nu(X_{\leq M} \cap \mathrm{T}^{-n}(X_{\geq M}))$.

Now, recalling that $V_x = \{n \in [0, N-1] : \mathrm{T}^n(x) \in X_{\geq M}\}$ we note that

$$\begin{aligned}&\frac{1}{N}\sum_{n=0}^{N-1} \nu(X_{\leq M} \cap \mathrm{T}^{-n}(X_{\geq M})) \\ &= \frac{1}{N}\sum_{i=1}^{N} i\nu(\{x \in X_{\leq M} : |V_x| = i\}) \\ &= \frac{1}{N}\sum_{i=1}^{\lfloor \kappa N \rfloor} i\nu(\{x \in X_{\leq M} : |V_x| = i\}) + \frac{1}{N}\sum_{i=\lceil \kappa N \rceil}^{N} i\nu(\{x \in X_{\leq M} : |V_x| = i\}) \\ &\leq \frac{1}{N}\lfloor \kappa N \rfloor \nu(X_{\leq M}) + \frac{1}{N} N\nu(\{x \in X_{\leq M} : |V_x| > \kappa N\})\end{aligned}$$

Let $K(M) > 0$ be the implicit constant that appeared in Lemma 6.2. Then using Lemma 6.2 we obtain

$$\frac{1}{N}\sum_{n=0}^{N-1} \nu(X_{<M} \cap \mathrm{T}^{-n}(X_{\geq M})) \leq \kappa + K(M)e^{\frac{6-2\kappa-3d+3\delta}{2}N + \frac{9N\log(c_0 \log M)}{\log M}}.$$

Thus, together with (6.0.1) we get

$$\mu_N(X_{\geq M}) \leq \epsilon(M) + \kappa + K(M)e^{(\frac{6-2\kappa-3d+3\delta}{2}+\frac{9\log(c_0\log M)}{\log M})N}. \tag{6.0.2}$$

The theorem is only interesting when $d > \frac{4}{3}$. So, we fix some $d > \frac{4}{3}$ and let $\kappa > \frac{6-3d}{2}$. Now, we let $\delta > 0$ to be small enough so that

$$6 - 2\kappa - 3d + 3\delta < 0.$$

Let $\epsilon > 0$ be given. For M sufficiently large we can make sure that $\epsilon(M) < \epsilon/2$ and that $\frac{6-2\kappa-3d+3\delta}{2} + \frac{9\log(c_0\log M)}{\log M} < 0$. Thus,

$$K(M)e^{(\frac{6-2\kappa-3d+3\delta}{2}+\frac{9\log(c_0\log M)}{\log M})N} \to 0$$

as $N \to \infty$. So, we conclude that for N large enough we get

$$\mu_N(X_{\geq M}) \leq \kappa + \epsilon$$

which gives in the limit that $\mu(X) > 1 - \kappa$. This is true for any $\kappa > \frac{6-3d}{2}$. Thus,

$$\mu(X) \geq 1 - \frac{6-3d}{2} = \frac{3d-4}{2}.$$

□

Next, we prove Corollary 1.9. We need the following Corollary 4.12 from [Fa].

Theorem 6.3. *Let F be a Borel subset of $\mathbb{R}^n$ with $0 < \mathcal{H}^s(F) \leq \infty$. Then there is a compact set $E \subset F$ such that $0 < \mathcal{H}^s(E) < \infty$ and a constant b such that*

$$\mathcal{H}^s(E \cap B_\delta(\mathbf{r})) \leq b\delta^s$$

for all $\mathbf{r} \in \mathbb{R}^n$ and $\delta > 0$.

Proof of Corollary 1.9. As any divergent point is also divergent on average, we get from [Ch] that the set of points $F_0 \subset X$ that are divergent on average has at least dimension $\frac{4}{3}+6$. So assume now that the Hausdorff dimension of F_0 is greater than $\frac{4}{3}+6$. Then, by the behavior of Hausdorff dimension under countable unions, there is some subset $F \subset F_0$ with compact closure and small diameter for which the Hausdorff dimension is also bigger than $\frac{4}{3}+6$. Here we may assume that $F = F_0 \cap (x_0 D_\eta B_\eta^{U^-C})$ and that $x_0 D_\eta B_\eta^{U^-C}$ is the injective image of the corresponding set in $\mathrm{SL}_3(\mathbb{R})$. It then follows that $F = x_0 D' B_\eta^{U^-C}$ and that D' has Hausdorff dimension bigger than $\frac{4}{3}$. Thus, for sufficiently small $\epsilon > 0$ we have that $\mathcal{H}^{\frac{4}{3}+\epsilon}(D') = \infty$. We may identify U^+ with $\mathbb{R}^2$ and apply Theorem 6.3. Therefore, there exists a compact set $E \subset D'$ such that $0 < \mathcal{H}^{\frac{4}{3}+\epsilon}(E) < \infty$ and a constant b such that

$$\mathcal{H}^{\frac{4}{3}+\epsilon}(E \cap B_\delta(\mathbf{r})) \leq b\delta^{\frac{4}{3}+\epsilon}$$

for all $\mathbf{r} \in \mathbb{R}^2$ and $\delta > 0$. We define $\nu_0 = \frac{1}{\mathcal{H}^{\frac{4}{3}+\epsilon}(E)} \mathcal{H}^{\frac{4}{3}+\epsilon}_{|E}$ so that $\nu_0(U^+) = 1$. Let τ be the map from U^+ to X defined by $\tau(u) = x_0 u$. Now, we let $\nu = \tau_* \nu_0$ to be the push-forward of the measure ν_0 under the map τ. It follows that for any $\delta > 0$ and for any $x \in X$ we have

$$\nu(x B_\delta^{U^+} B_\eta^{U^-C}) \ll \delta^{\frac{4}{3}+\epsilon}.$$

Now, if we define μ_N as before then Theorem 1.8 implies that the limit measure μ has at least $\frac{3}{2}(\frac{4}{3}+\epsilon-\frac{4}{3})\frac{3\epsilon}{2} > 0$ mass left. However, the assumption on F_0 and dominated convergence applied to

$$\mu_N(X_{\leq M}) = \int \frac{1}{N} \sum_{n=0}^{N-1} \chi_{T^{-n} X_{\leq M}} d\nu$$

implies that $\mu_N(X_{\leq M}) \to 0$ as $N \to \infty$ for any fixed M. This gives a contradiction and the corollary. $\square$

BIBLIOGRAPHY

[ELMV] M. Einsiedler, E. Lindenstrauss, Ph. Michel, and A. Venkatesh, *Distribution Properties Of Closed Geodesics On The Modular Surface, Duke's Theorem.*, Preprint 2010.

[EK] M. Einsiedler, S. Kadyrov, *Escape of Mass and Entropy for* $\mathrm{SL}_3(\mathbb{Z})\backslash \mathrm{SL}_3(\mathbb{R})$., Submitted for publication 2009.

[Ka] S. Kadyrov, *Positive Entropy Invariant Measures On The Space Of Lattices With Escape Of Mass.*, Submitted for publication 2009.

[EW] M. Einsiedler and T. Ward, *Ergodic theory with a view towards number theory*, in preparation, chapters available online. http://www.mth. uea.ac.uk/ergodic/

[Ra] M.S. Raghunathan, *Discrete Subgroups of Lie Groups*, Springer-Verlag, Berlin USA, 1972.

[MT] G. A. Margulis and G. M. Tomanov, *Invariant measures for actions of unipotent groups over local fields on homogeneous spaces*, Invent. Math. **122(1-3)** (1994), 347-392.

[MS] S. Mozes and N. Shah, *On the space of ergodic invariant measures of unipotent flows*, Ergodic Theory Dynam. Systems **15** (1995), 149-159.

[Da] S. G. Dani, *On orbits of unipotent fows on homogeneous spaces*, Ergodic Theory Dynam. Systems **6** (1986), 167-182.

[MK] G. A. Margulis and D. Kleinbock, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Ann. Math. **148** (1998), 339-360.

[BK] M. Brin and A. Katok, *On local entropy*, Geometric dynamics **1007**(1983), 30-38.

[KLW] D. Kleinbock, E. Lindenstrauss, and B. Weiss, *On fractal measures and diophantine approximation*, Selecta Math. **10**(2004), 479-523.

[Sh] R. Shi, *Equidistribution of expanding measures with local maximal dimension and Diophantine Approximation*, Preprint 2010, http://arxiv.org/abs/0905.1152.

[Wa] P. Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics 79. Springer-Verlag, New York, 1982.

[Ch] Y. Cheung, *Hausdorff dimension of the set of Singular Pairs*, To appear in Annals of Math.

[Sa] T. Sakai, *Riemannian Geometry*, AMS Bookstore, 1996.

[Fa] K. Falconer, *Fractal Geometry*, Wiley, 2003.