ENTROPY AND ESCAPE OF MASS FOR $\mathrm{SL}_3(\mathbb{Z})\backslash\mathrm{SL}_3(\mathbb{R})$

MANFRED EINSIEDLER AND SHIRALI KADYROV

Abstract. We study the relation between measure theoretic entropy and escape of mass for the case of a singular diagonal flow on the moduli space of three-dimensional unimodular lattices.

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 [11]. This fact relies on the quantitative non-divergences estimates for unipotents due to works of S. G. Dani [4] (further refined by G. A. Margulis and D. Kleinbock [9]). On the other hand, if T acts on $X = \mathrm{SL}_d(\mathbb{Z})\backslash\mathrm{SL}_d(\mathbb{R})$ by a diagonal element, then $\mu(X)$ can be any value in the interval $[0, 1]$ due to softness of Anosov-flows, see for instance [8]. 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 [5] for the action of geodesic flow on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$, see Theorem 1.1. In this paper we will generalize this theorem to the space $\mathrm{SL}_3(\mathbb{Z})\backslash\mathrm{SL}_3(\mathbb{R})$ with the action of a particular diagonal element.

We identify $X = \mathrm{SL}_d(\mathbb{Z})\backslash\mathrm{SL}_d(\mathbb{R})$ with the space of unimodular lattices in $\mathbb{R}^d$, see § 2.1. Using this correspondence we can define for $d = 3$ the height function $\mathrm{ht}(x)$ of a lattice $x \in X$ 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 $\mathbb{R}^d$. Also, if $d = 2$ 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 (cf. Theorem 2.3) $X_{\leq M}$ is compact and any compact subset of X is contained in some $X_{\leq M}$.

In [5], M. E., E. Lindenstrauss, Ph. Michel, and A. Venkatesh give the following theorem:

M.E. acknowledges the support of the NSF from the grant 0554373, and both authors acknowledges support by the SNF (200021-127145).

Theorem 1.1. *Let T be the time-one-map for the geodesic flow. There exists some M_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 1.1 in [5] makes use of the geometry of the upper half plane $\mathbb{H}$.

From now on we let $X = \mathrm{SL}_3(\mathbb{Z})\backslash \mathrm{SL}_3(\mathbb{R})$ and 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 $\mathrm{T}(x) = x\alpha$. Now we can state the main theorem of the paper.

Theorem 1.2. *There exists some M_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 Haar measure on X, is 3. This follows e.g. from [10, Prop. 9.2].

As a consequence of Theorem 1.2 we have:

Corollary 1.3. *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.*

This result is sharp in the following sense. For any $c \in (2,3)$ one can construct a sequence of measures μ_i with $h_{\mu_i}(\mathrm{T}) \to c$ as $i \to \infty$ such that any weak* limit μ has precisely $c-2$ mass left, see [8].

Another interesting application of our method arises when we do not assume T-invariance of the measures we consider. In this case, instead of entropy consideration we assume that our measures have high dimension and study the behavior of the measure under iterates of T.

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 us consider the following subgroups of G

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

Let $d \in [0,2]$ be given and let us consider a probability measure ν on X with the following property. For any $\delta > 0$ there exists $\epsilon' > 0$ such that for any $\epsilon < \epsilon'$ one has

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

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

Theorem 1.4. *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 $\{\mu_n\}_{n=1}^\infty$ 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 [14].

Another application of Theorem 1.4 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 [2]. 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 [2]. 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 [2]) 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 [2]) is that the set of divergent points in $\mathrm{SL}_3(\mathbb{Z})\backslash\mathrm{SL}_3(\mathbb{R})$ has Hausdorff dimension $8 - \frac{2}{3} = \frac{4}{3} + 6$.

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.

[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 $x_\mathbf{r}$.

Corollary 1.5. *The Hausdorff dimension of the set of points that are divergent on average is also* $\frac{4}{3} + 6$*.*

We finally note that the nondivergence result [3, Theorem 3.3] is related to Theorem 1.4. In fact, [3, Theorem 3.3] implies that μ as in Theorem 1.4 is a probability measure if ν additional regularity properties (ν is assumed to be friendly). However, to our knowledge these additional assumptions make it impossible to derive e.g. Corollary 1.5.

The next section below has some basic definitions and facts. In § 3, we characterize what it means for a trajectory of a lattice to be above height M in some time interval. Using this we prove Theorem 1.2 in § 4-5. Theorem 1.4 and its corollary are discussed in § 6.

Acknowledgements: We would like to thank Jim Tseng for discussions and for pointing out the reference to [7].

2. Preliminaries

2.1. **The space of unimodular lattices.** In this section we will give a brief introduction to the space of unimodular lattices in $\mathbb{R}^3$.

Definition 2.1. $\Lambda \subset \mathbb{R}^3$ *is a* lattice *if it is a discrete subgroup and has a compact quotient* $\mathbb{R}^3/\Lambda$*.*

Note that this is equivalent to saying that $\Lambda = \langle v_1, v_2, v_3 \rangle_{\mathbb{Z}}$ where v_1, v_2, v_3 are linearly independent vectors over $\mathbb{R}$.

Definition 2.2. *A lattice* $\Lambda = \langle v_1, v_2, v_3 \rangle_{\mathbb{Z}}$ *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, v_3*.*

Let $g \in \mathrm{SL}_3(\mathbb{R})$ and let v_1, v_2, v_3 be its row vectors. We identify $\mathrm{SL}_3(\mathbb{Z})g \in X$ with the unimodular lattice in $\mathbb{R}^3$ generated by vectors v_1, v_2, v_3. In particular $\mathrm{SL}_3(\mathbb{Z})$ corresponds to $\mathbb{Z}^3$.

Let $g, g' \in \mathrm{SL}_3(\mathbb{R})$ have row vectors v_1, v_2, v_3 and w_1, w_2, w_3 respectively. If $g' = hg$ where $h \in \mathrm{SL}_3(\mathbb{Z})$ then clearly

$$\langle v_1, v_2, v_3 \rangle_{\mathbb{Z}} = \langle w_1, w_2, w_3 \rangle_{\mathbb{Z}}$$

and hence they correspond to the the same unimodular lattice in $\mathbb{R}^3$.

Therefore, we can think of $X = \mathrm{SL}_3(\mathbb{Z}) \backslash \mathrm{SL}_3(\mathbb{R})$ as a space of unimodular lattices in $\mathbb{R}^3$.

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* $K \subset X$ *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 [12, Cor. 10.9]. We will now deduce Corollary 1.3 from Theorem 1.2.

Proof. 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

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

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.

□

2.2. **Riemannian metric on X.** Let $G = \mathrm{SL}_3(\mathbb{R})$ and $\Gamma = \mathrm{SL}_3(\mathbb{Z})$. We fix a left-invariant Riemannian metric d_G on G and for any $x_1 = \Gamma g_1, x_2 = \Gamma g_2 \in X$ we define

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

which gives us a left-invariant Riemannian metric d_X on $X = \Gamma\backslash G$. For more information about the Riemannian metric, we refer to [13, Chp. 2].

2.2.1. *Injectivity radius.* Let $B_r^H(x) := \{h \in H \,|\, d(h, x) < r\}$ where d is a metric defined in H and B_r^H is understood to be $B_r^H(1)$.

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

Note that since $X_{\leq M}$ is compact, 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*. We refer to Proposition 9.14 in [6] for a proof of these claims.

2.3. **Entropy.** In this section we will give the definition of entropy. For more information we refer to [15].

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

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

with the convention $0 \log 0 = 0$.

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

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

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.

3. 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

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

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 $X_{\le M}$ we get a family $\mathcal{M}_N$ of sets of labeled marked times which will be defined in § 3.2. This will give rise to a partition of X, which will be helpful in the main estimates given in § 4.

3.1. **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

$$\begin{aligned}(3.1)\qquad \mathrm{T}(u\wedge v)&=(u_1e^{1/2},u_2e^{1/2},u_3e^{-1})\wedge(v_1e^{1/2},v_2e^{1/2},v_3e^{-1})\\&=((u_2v_3-u_3v_2)e^{-1/2},(u_3v_1-u_1v_3)e^{-1/2},(u_1v_2-u_2v_1)e^{1}).\end{aligned}$$

Let $\epsilon>0$ be given. Fix $x\in X$, a vector v in x is *ϵ-short at time n* if $|\,\mathrm{T}^n(v)|\le\epsilon$. 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 $\le\epsilon$.

3.2. **(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.

3.2.1. *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 $\mathrm{T}(P')$.

3.2.2. *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 § 3.2.1. 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 § 3.2.1. 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 $[-N, N]$.

3.2.3. *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. □

We would like to know that the cardinality of $\mathcal{M}_N$ can be made small (important in Lemma 5.1) 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

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

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. □

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

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

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*

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

3.3. **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 §4.1). 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.

3.3.1. *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

$$\frac{1-2\eta}{M} \leq |w'| \leq \frac{e(1+2\eta)}{M}. \tag{3.2}$$

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 (3.2) 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. □

3.3.2. *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 $l \le p-1 \le l'$. 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'. □

4. Main Proposition and Restrictions

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. Define a Bowen N-ball to be the translate xB_N for some $x \in X$ of

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

Fix a height $M \ge 1$. Let $N \ge 1$ 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

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

Now, we state the main proposition.

Proposition 4.1. *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 1.2 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 4.1 that leads to the conclusion of the main Theorem 1.2.

In proving Proposition 4.1, 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 $Z(\mathcal{N})$.

4.1. Restrictions of perturbations.

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

Lemma 4.2. *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 4.2 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

$$u = vg = w \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}. \tag{4.1}$$

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

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

Now, arguing as in the proof of Lemma 4.2, 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 4.3. *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 (4.1). Then*

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

Lemma 4.4. *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. □

What Lemma 4.3 and Lemma 4.4 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 4.1, 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 4.1.

4.1.2. *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 $a = u_2v_3 - u_3v_2, b = u_3v_1 - u_1v_3, c = u_1v_2 - u_2v_1$ (cf. 3.1) 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

$$c' = u'_1 v'_2 - u'_2 v'_1 = (u_1 + t_1 u_3)(v_2 + t_2 v_3) - (u_2 + t_2 u_3)(v_1 + t_1 v_3) = c - at_1 - bt_2.$$

Now, assume that

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

which by the above implies

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

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

$$g = \begin{pmatrix} 1 & & \\ & 1 & \\ t_1 & t_2 & 1 \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ 0 & 0 & g_{33} \end{pmatrix}. \tag{4.2}$$

Then we have

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

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 4.5. *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* (4.2) *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 4.6. *Consider the set defined by* $\frac{(at_1+bt_2)^2}{a^2+b^2} \le Ce^{-2S}$ *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* $\frac{(at_1+bt_2)^2}{a^2+b^2} \le Ce^{-2S}$.

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

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

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 $\le 4\sqrt{2C}\eta e^{-S}$.

If $\sqrt{C}e^{-S} \le \frac{1}{2}\eta e^{-3S'/2}$ 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. □

4.2. **Proof of Main Proposition.**

Proof of Proposition 4.1. 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 $x_0 \in X_{\le M}$ be such that

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

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 § 3.2.2).

We claim that

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

can be covered by $\ll_M e^{3N-|V|} c_0^{\frac{9N}{\lceil \log M \rceil}}$ forward Bowen N-balls xB_N^+ defined by

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

Since $X_{\le M}$ 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 $x_0 \in X_{\le M}$ and a neighborhood

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

of x_0, where $D^{U^+}_{\eta/2}$ 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 $\ll e^{3N-|V|}\, c_0^{\frac{9N}{\lfloor \log M \rfloor}}$ 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 $B^{U^-C}_{\eta/2}$. By breaking the U^+-part into $\lceil e^{3n/2} \rceil^2$ sets of the form $u_i^+ D^{U^+}_{\eta/2}$ 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

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

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

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

Notice that $\mathrm{T}^{-n}(\mathrm{T}^n(x_0)u^+) \alpha^n D^{U^+}_{\eta/2} \alpha^{-n} B^{U^-C}_{\eta/2}(1)$ 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 $0 \le k < n$ we have

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

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 $m \le |\mathcal{L}| + |\mathcal{P}|$ so that from Lemma 3.3 we obtain

$$m \le \frac{N}{\lfloor \log M \rfloor} + 1 \tag{4.3}$$

Now, write $[0, N-1] \setminus V = W_1 \cup W_2 \cup ... \cup W_l$ where W_i's are intervals. A bound similar to (4.3) 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 $K \le N$ 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

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

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 $\mathrm{T}^K(x_0) u^+ B^{U^+}_{\eta/2} \alpha^{-K} B^{U^-C}_{\eta/2} \alpha^K$ into $\lceil e^{3|W_{n'+1}|/2} \rceil^2 \le 4e^{3|W_{n'+1}|}$ many sets of the form

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

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

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

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

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

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

$$\mathcal{L} = \{l_1 < l_2 < \ldots < l_k\},\ \mathcal{L}' = \{l_1' < l_2' < \ldots < l_k'\}$$

and

$$\mathcal{P} = \{p_1 < p_2 < \ldots < p_{k'}\},\ \mathcal{P}' = \{p_1' < p_2' < \ldots < p_{k'}'\}$$

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

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

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

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

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 4.4, and Lemma 4.6.

We would like to apply Lemma 4.4 and Lemma 4.6 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 4.4 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 4.6, 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 $D_{\eta/2}^{U^+}$ 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 4.3 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 $g \in B_\eta^{\mathrm{SL}_3(\mathbb{R})}$, 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 4.3 and Lemma 4.4 with $S' = p_1 - l_1$ and $S = l_1' - l_1$ we see that we only need to consider

$$\leq c_0 \max\{1, e^{3(p_1-l_1)-(l_1'-l_1)}\} =: N_1 \tag{4.4}$$

of these $\lceil e^{3(p_1-l_1)/2}\rceil^2$ sets (see the discussion at the end of § 4.1.1). Thus, we obtain sets of the form

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

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 4.6. However, Lemma 4.6 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 $\le c_0 \min\{e^{(l_1'-p_1)/2}, e^{p_1-l_1}\}$ 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 4.6 with $S' = l_2 - p_1$ and $S = p_1' - p_1$. Thus, for each choice we need to consider only $\le c_0 \max\{e^{3(l_2-p_1)/2}, e^{3(l_2-p_1)-(p_1'-p_1)}\}$ of the $\lceil e^{3(l_2-p_1)/2}\rceil^2$ subsets. Thus, in total, we need to consider only

$$\le c_0^2 \min\{e^{(l_1'-p_1)/2}, e^{p_1-l_1}\} \max\{e^{3(l_2-p_1)/2}, e^{3(l_2-p_1)-(p_1'-p_1)}\} =: N_2 \tag{4.5}$$

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 $\le c_0 \min\{e^{p_1'-l_2}, e^{(l_2-p_1)/2}\}$ many configurations and for each of them we can apply Lemma 4.4 with $S' = p_2 - l_2$ and $S = l_2' - l_2$. So, for each configuration we need only $\le c_0 \max\{1, e^{3(p_2-l_2)-(l_2'-l_2)}\}$ many of the subsets. Thus, we need

$$\le c_0^2 \min\{e^{p_1'-l_2}, e^{(l_2-p_1)/2}\} \max\{1, e^{3(p_2-l_2)-(l_2'-l_2)}\} =: N_3 \tag{4.6}$$

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 $\le c_0 \min\{e^{p_{k-1}'-l_k}, e^{(l_k-p_{k-1})/2}\}$ configurations and for each we apply Lemma 4.4 with $S' = S = l_k' - l_k$. Hence, we need to consider only

$$\le c_0^2 \min\{e^{p_{k-1}'-l_k}, e^{(l_k-p_{k-1})/2}\} e^{3(l_k'-l_l)-(l_k'-l_k)} =: N_{2k-1} \tag{4.7}$$

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

$$\le N_1 N_2 N_3 \cdots N_{2k-1} \tag{4.8}$$

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 (4.4) with the min term of (4.5) 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 (4.5) with the min term of (4.6)

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

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

$$\begin{aligned} N_1N_2N_3\cdots N_{2k-1} &\leq c_0^{4k}e^{2(p_1-l_1)}e^{2(l_2-p_1)}\cdots e^{2(p_{k-1}-l_{k-1})}e^{2(l_k'-l_k)} \\ &= c_0^{4k}e^{2(p_1-l_1)+2(l_2-p_1)+\cdots+2(l_k'-l_k)} \\ &= c_0^{4k}e^{2|V_{n+1}|} \end{aligned}$$

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 (4.3) 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. □

5. Proof of Theorem 1.2

Before going to the proof of Theorem 1.2 we need Lemma 5.1 which gives an upper bound for entropy in terms of covers of Bowen balls.

Lemma 5.1. *Let μ be an* 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*

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

We omit the proof which is very similar to [5, Lemma 5.2] and goes back to [1].

Proof of the Theorem 1.2. 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 [6, 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 [15, Thm. 8.4], so that desired estimate follows from the ergodic case.

Suppose that μ is ergodic. We would like to apply Lemma 5.1, 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 lost of generality we can assume $k_1 = k_2 = 0$. Next we split Z' into the sets $Z(\mathcal{N})$ as in Proposition 4.1 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 4.1 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 5.1 we arrive at

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

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 1.3).

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

$$h_\mu(\mathrm{T})\leq 3-\mu(X_{\geq M})+O(\frac{\log\log M}{\log M}). \tag{5.1}$$

of the above.

If $\mu(X_{\leq M_\mu})>0$ then (5.1) 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 (5.1) for $M=M_\mu$ follows from (5.1) 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 (5.1) trivially also holds for any $M\in[100,M_\mu)$. □

6. Limits of measures with high dimension

In this section we prove Theorem 1.4 and Corollary 1.5. Our main tool is a version of Proposition 4.1. 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 4.1 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 $\ll_M e^{3N-|V_{x_0}|}c_0^{\frac{9N}{\lceil\log M\rceil}}$ 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 4.1 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 $e^{3N-|V_{x_0}|}c_0^{\frac{9N}{\lfloor \log M \rfloor}}$ many pre-images under T^N of sets of the form

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

So, Z_O^+ can be covered by the sets of the form

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

This completes the proof since we have $\alpha^N D_{\eta/2}^{U^+}\alpha^{-N} = D^{U^+}_{\frac{\eta}{2}e^{-3N/2}}$ and since $X_{\leq M}$ is compact. □

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,

$$\bar{\nu}(\{x \in X_{\leq M} : |V_x| > \kappa N\}) \ll_M e^{\frac{5N\log\lfloor \log M \rfloor}{\lfloor \log M \rfloor}} e^{(3-\kappa)N} c_0^{\frac{9N}{\lfloor \log M \rfloor}} (\frac{\eta}{2}e^{-3N/2})^{d-\delta}.$$

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 1.4. In order to prove Theorem 1.4 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,

$$\mu_N(X_{\geq M}) \leq \epsilon(M) + \frac{1}{N}\sum_{n=0}^{N-1} \nu(X_{\leq M} \cap \mathrm{T}^{-n}(X_{\geq M})). \tag{6.1}$$

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 (6.1) we get

$$\mu_N(X_{\geq M}) \leq \epsilon(M) + \kappa + K(M)e^{(\frac{6-2\kappa-3d+3\delta}{2} + \frac{9\log(c_0 \log M)}{\log M})N}. \tag{6.2}$$

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 1.5. We need the following Corollary 4.12 from [7].

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 1.5. As any divergent point is also divergent on average, we get from [2] 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

$$\mathcal{H}^{\frac{4}{3}+\epsilon}(E \cap B_\delta(\mathbf{r})) \le b\delta^{\frac{4}{3}+\epsilon}$$

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 1.4 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

$$\mu_N(X_{\le M}) = \int \frac{1}{N} \sum_{n=0}^{N-1} \chi_{T^{-n} X_{\le M}} d\nu$$

implies that $\mu_N(X_{\le M}) \to 0$ as $N \to \infty$ for any fixed M. This gives a contradiction and the corollary. □

References

[1] M. Brin and A. Katok. On local entropy. *Geometric dynamics*, 1007:30–38, 1983.

[2] Y. Cheung. Hausdorff dimension of the set of singular pairs. *Preprint*, 2009.

[3] E. Lindenstrauss D. Kleinbock and B. Weiss. On fractal measures and diophantine approximation. *Selecta Math.*, 10:479–523, 2004.

[4] S. G. Dani. On orbits of unipotent fows on homogeneous spaces. *Ergodic Theory Dynam. Systems*, 6:167–182, 1986.

[5] M. Einsiedler, E. Lindenstrauss, Ph. Michel, and A. Venkatesh. Distribution properties of closed geodesics on the modular surface, Duke's theorem. *preprint*.

[6] M. Einsiedler and T. Ward. *Ergodic Theory with a view towards Number Theory*. In preparation, 2009.

[7] K. Falconer. *Fractal Geometry*. Wiley, 2003.

[8] S. Kadyrov. A construction of invariant measures on the space of lattices. *Preprint*.

[9] G. A. Margulis and D. Kelinbock. Flows on homogeneous spaces and diophantine approximation on manifolds. *Ann. Math.*, 148:339–360, 1998.

[10] G. A. Margulis and G. M. Tomanov. Invariant measures for actions of unipotent groups over local fields on homogeneous spaces. *Invent. Math.*, 116(1-3):347–392, 1994.

[11] S. Mozes and N. Shah. On the space of ergodic invariant measures of unipotent flows. *Ergodic Theory Dynam. Systems*, 15:149–159, 1995.

[12] M.S. Raghunathan. *Discrete Subgroups of Lie Groups*. Springer-Verlag, 1972.

[13] Takashi Sakai. *Riemannian Geometry*. AMS Bookstore, 1996.

[14] R. Shi. Equidistribution of expanding measures with local maximal dimension and diophantine approximation. *Preprint*, 2009.

[15] Peter Walters. *An Introduction to Ergodic Theory*. Springer-Verlag, 2000.