LIMITING DISTRIBUTIONS OF TRANSLATES OF DIVERGENT DIAGONAL ORBITS

URI SHAPIRA AND CHENG ZHENG

ABSTRACT. We define a natural topology on the collection of (equivalence classes up to scaling of) locally finite measures on a homogeneous space and prove that in this topology, pushforwards of certain infinite volume orbits equidistribute in the ambient space. As an application of our results we prove an asymptotic formula for the number of integral points in a ball on some varieties as the radius goes to infinity.

1. INTRODUCTION

This paper deals with the study of the possible limits of periodic orbits in homogeneous spaces. Before explaining what we mean by this we start by motivating this study. In many instances arithmetic properties of an object are captured by periodicity of a corresponding orbit in some dynamical system. A simple instance of this phenomenon is that $\alpha \in \mathbb{R}$ is rational if and only if its decimal expansion is eventually periodic. In dynamical terms this is expressed by the fact that the orbit of α modulo 1 on the torus $\mathbb{R}/\mathbb{Z}$ under multiplication by 10 (modulo 1) is eventually periodic. Furthermore, from knowing distributional information regarding the periodic orbit one can draw meaningful arithmetical conclusions. In the above example this means that if the orbit is very close to being evenly distributed on the circle then the frequency of appearance of say the digit 3 in the period of the decimal expansion is roughly $\frac{1}{10}$. This naive scheme has far reaching analogous manifestations capturing deep arithmetic concepts in dynamical terms. More elaborate instances are for example the following:

- Similarly to the above example regarding decimal expansion, periodic geodesics on the modular surface correspond to continued fraction expansions of quadratic numbers and distributional properties of the former implies statistical information regarding the latter (see [AS] where this was used).
- Representing an integral quadratic form by another is related to periodic orbits of orthogonal groups (see [EV08]).
- Class groups of number fields correspond to adelic torus orbits (see [ELMV11]).
- Values of rational quadratic forms are governed by the volume of periodic orbits of orthogonal groups (see [EMV09, Theorem1.1])

- Asymptotic formulas for counting integer and rational points on varieties are encoded by distributional properties of periodic orbits (see [DRS93, EM93, EMS96, GMO08] for example).

In all the above examples the orbits that are considered are of *finite volume*. Recently in [KK] and [OS14] this barrier was crossed and particular instances of the above principle were used for infinite volume orbits in order to obtain asymptotic estimates for counting integral points on some varieties and weighted second moments of GL(2) automorphic L-functions.

At this point let us make more precise our terminology. Let X be a locally compact second countable Hausdorff space and let H be a unimodular topological group acting on X continuously. We say that an orbit Hx is *periodic* if it supports an H-invariant *locally finite* Borel measure. In such a case the orbit is necessarily closed and this measure is unique up to scaling and is obtained by restricting the Haar measure of H to a fundamental domain of $\mathrm{Stab}_H(x)$ in H which is identified with the orbit via $h \mapsto hx$. We say that such an orbit is *of finite volume* if the total mass of the orbit is finite. It is then customary to normalize the total mass of the orbit to 1. We remark that in some texts the term periodic orbit is reserved for finite volume ones but we wish to extend the terminology as above. If Hx is a periodic orbit we denote by μ_{Hx} a choice of such a measure, which in the finite volume case is assumed to be normalized to a probability measure.

Given a sequence of periodic orbits Hx_i, it makes sense to ask if they converge in some sense to a limiting object. When the orbits are of finite volume the common definition is that of weak* convergence; each orbit is represented by the probability measure μ_{Hx_i} and one equips the space of probability measures $\mathcal{P}(X)$ with the weak* topology coming from identifying $\mathcal{P}(X)$ as a subset of the unit sphere in the dual of the Banach space of continuous functions on X vanishing at infinity $C_0(X)$. The starting point of this paper is to challenge this and propose a slight modification which will allow to bring into the picture periodic orbits of infinite volume. For that we will shortly concern ourselves with topologizing the space of equivalence classes $[\mu]$ of locally finite measures μ on X.

This approach has several advantages over the classical weak* convergence approach. As said above it allows to discuss limiting distributions of infinite volume orbits but also it allows to detect in some cases information which is invisible for the weak* topology: In the classical discussion, it is common that a sequence of periodic probability measures μ_{Hx_i} converges to the zero measure (phenomenon known as full escape of mass). Nevertheless it sometimes happens that the orbits themselves do converge to a limiting object but this information was lost because the measures along the sequence were not scaled properly. This phenomenon happens for example in [Sha] which inspired us to define the notion of convergence to be defined below.

Although the results we will prove are rather specialized we wish to present the framework in which our discussion takes place in some generality. Let G be a Lie group[1] and let $\Gamma < G$ be a lattice.

Question 1.1. Let $X = G/\Gamma$ and let $H_i x_i$ be a sequence of periodic orbits. Under which conditions the following holds:

(1) The sequence $[\mu_{H_i x_i}]$ has a converging subsequence?
(2) The accumulation points of $[\mu_{H_i x_i}]$ are themselves (homothety classes of) periodic measures?

2. Basic definitions and results

2.1. **Topologies.** Now we make our discussion in the introduction more rigorous. Let X be a locally compact second countable Hausdorff space and $\mathcal{M}(X)$ the space of locally finite measures on X. We say that two locally finite measures μ and ν in $\mathcal{M}(X)$ are equivalent if there exists a constant $\lambda > 0$ such that $\mu = \lambda\nu$. This forms an equivalence relation and we denote the equivalence class of μ by $[\mu]$. We denote by $\mathbb{P}\mathcal{M}(X)$ the set of all equivalence classes of nonzero locally finite measures on X.

We topologize $\mathcal{M}(X)$ and $\mathbb{P}\mathcal{M}(X)$ as follows. Let $C_c(X)$ be the space of compactly supported continuous functions on X. For any $\rho \in C_c(X)$, define a map

$$i_\rho : \mathcal{M}(X) \to C_0(X)^*$$

by sending $d\mu \in \mathcal{M}(X)$ to $\rho d\mu \in C_0(X)^*$. Here $C_0(X)$ is the space of continuous functions on X vanishing at infinity equipped with the supremum norm, and $C_0(X)^*$ denotes its dual space. The weak* topology on $C_0(X)^*$ then induces a topology τ_ρ on $\mathcal{M}(X)$ via the map i_ρ. We will denote by τ_X the topology on $\mathcal{M}(X)$ generated by $(\mathcal{M}(X), \tau_\rho)$ $(\rho \in C_c(X))$. Equivalently, τ_X is the smallest topology on $\mathcal{M}(X)$ such that for any $f \in C_c(X)$ the map

$$\mu \mapsto \int f d\mu$$

is a continuous map from $\mathcal{M}(X)$ to $\mathbb{R}$.

Definition 2.1. Let π_P be the natural projection map from $\mathcal{M}(X) \setminus \{0\}$ to $\mathbb{P}\mathcal{M}(X)$. The topology τ_P on $\mathbb{P}\mathcal{M}(X)$ is then defined to be the quotient topology induced by τ_X on $\mathcal{M}(X)$ via π_P. In other words, U is an open subset in $\mathbb{P}\mathcal{M}(X)$ if and only if $\pi_P^{-1}(U)$ is open in $\mathcal{M}(X) \setminus \{0\}$. In this way, we obtain a topological space $(\mathbb{P}\mathcal{M}(X), \tau_P)$.

[1] One could (and should) develop this discussion in the S-arithmetic and adelic settings as well.

2.2. **Main results.** Now let $G = \mathrm{SL}(n, \mathbb{R})$, $\Gamma = \mathrm{SL}(n, \mathbb{Z})$ and $X = G/\Gamma = \mathrm{SL}(n, \mathbb{R})/\mathrm{SL}(n, \mathbb{Z})$. Denote by m_X the unique G-invariant probability measure on X and by Ad the adjoint representation of G. We write

$$A = \{\mathrm{diag}(e^{t_1}, e^{t_2}, \dots, e^{t_{n-1}}, e^{t_n}) : t_1 + t_2 + \cdots + t_n = 0\}$$

for the connected component of the full diagonal group in G. In this paper, we address Question 1.1 in the space $X = \mathrm{SL}(n, \mathbb{R})/\mathrm{SL}(n, \mathbb{Z})$ with certain sequences $\{H_i x_i\}$ and prove the convergence of $[\mu_{H_i x_i}]$ with respect to $(\tau_P, \mathbb{P}\mathcal{M}(X))$ in the sense of Definition 2.1. As a simple exercise, and to motivate such a statement, the reader can show that if $[\mu_{H_i x_i}] \to [m_X]$ for example, then the orbits $H_i x_i$ become dense in X. In many cases our results imply that indeed the limit homothety class is the class of the uniform measure m_X.

Before stating our theorems, we need some notations. For a Lie subgroup $H < G$, let H^0 denote the connected component of identity of H, and $\mathrm{Lie}(H)$ its Lie algebra. Denote by $C_G(H)$ (resp. $C_G(\mathrm{Lie}(H))$) the centralizer of H (resp. $\mathrm{Lie}(H)$) in G. For G, we write $\mathfrak{g} = \mathrm{Lie}(G) = \mathfrak{sl}(n, \mathbb{R})$ and

$$\exp : \mathfrak{sl}(n, \mathbb{R}) \to \mathrm{SL}(n, \mathbb{R})$$

the exponential map from $\mathfrak{g}$ to G. For any $g \in G$ and measure μ on X, define the measure $g_*\mu$ by

$$g_*\mu(E) = \mu(g^{-1}E)$$

for any Borel subset $E \subset X$. An A-orbit Ax in X is called divergent if the map $a \mapsto ax$ from A to X is proper.

Definition 2.2. Let $\{g_k\}$ be a sequence in G. For any subgroup $S \subset A$, we define

$$\mathcal{A}(S, g_k) = \{Y \in \mathrm{Lie}(S) : \{\mathrm{Ad}(g_k)Y\} \text{ is bounded in } \mathfrak{g}\}.$$

This is a subalgebra in $\mathrm{Lie}(S)$.

Remark 2.3. By passing to a subsequence, we can always assume that for any $Y \in \mathrm{Lie}(S) \setminus \mathcal{A}(S, g_k)$, the sequence $\mathrm{Ad}(g_k)Y \to \infty$. Indeed, observe that for two vectors v_1 and $v_2 \in \mathrm{Lie}(S)$, if $\{\mathrm{Ad}(g_k)v_1\}$ and $\{\mathrm{Ad}(g_k)v_2\}$ are bounded, then for any v in the linear span of v_1 and v_2, $\{\mathrm{Ad}(g_k)v\}$ is also bounded. Because of this, one can collect vectors v with $\{\mathrm{Ad}(g_k)v\}$ bounded by passing to subsequences of $\{g_k\}$, and due to the finite dimension of $\mathrm{Lie}(S)$, this process would stop at some point. Then $\mathcal{A}(S, g_k)$ is the set of the vectors collected in this process, and for any vector Y which is not collected, the sequence $\mathrm{Ad}(g_k)Y \to \infty$.

The following theorem answers Question 1.1 for sequences of translates of a divergent diagonal orbit in G/Γ. Moreover, it gives a description of all accumulation points.

Theorem 2.4. *Let Ax be a divergent orbit in X. Then for any $\{g_k\}_{k\in\mathbb{N}}$ in G, the sequence $[(g_k)_*\mu_{Ax}]$ has a subsequence converging to an equivalence class of a periodic measure on $\mathrm{SL}(n,\mathbb{R})/\mathrm{SL}(n,\mathbb{Z})$.*

Furthermore, assume that for any $Y \in \mathrm{Lie}(A) \setminus \mathcal{A}(A, g_k)$ the sequence $\{\mathrm{Ad}(g_k)Y\}$ diverges (see Remark (2.3)). Then we have the following description of the limit points of $[(g_k)_*\mu_{Ax}]$. The subgroup $\exp(\mathcal{A}(A, g_k))$ is the connected component of the center of the reductive group $H = C_G(\mathcal{A}(A, g_k))$, and any limit point of the sequence $[(g_k)_*\mu_{Ax}]$ is a translate of the equivalence class $[\mu_{H^0x}]$. In particular, if the subspace $\mathcal{A}(A, g_k) = \{0\}$, then the sequence $[(g_k)_*\mu_{Ax}]$ converges to the equivalence class of the Haar measure on $\mathrm{SL}(n,\mathbb{R})/\mathrm{SL}(n,\mathbb{Z})$.

We will also deduce the following theorem from Theorem 2.4, which answers Question 1.1 for translates of an orbit of a connected reductive group H containing A. Such a reductive group is known as the connected component of $C_G(S)$ where S is a subtorus in A. We will see by Lemma 10.2 in section 10 that for $x \in \mathrm{SL}(n,\mathbb{R})/\mathrm{SL}(n,\mathbb{Z})$ with Ax divergent, Hx is closed for any reductive group H containing A.

Theorem 2.5. *Let Ax be a divergent orbit in X and let H be a connected reductive group containing A. Then for any $\{g_k\}_{k\in\mathbb{N}}$ in G, the sequence $[(g_k)_*\mu_{Hx}]$ has a subsequence converging to an equivalence class of a periodic measure on $\mathrm{SL}(n,\mathbb{R})/\mathrm{SL}(n,\mathbb{Z})$.*

Furthermore, let S be the center of H and assume that for any $Y \in \mathrm{Lie}(S)\setminus\mathcal{A}(S, g_k)$ the sequence $\{\mathrm{Ad}(g_k)Y\}$ diverges. Then we have the following description of the limit points of $[(g_k)_*\mu_{Hx}]$. The subgroup $\exp(\mathcal{A}(S, g_k))$ is the connected component of the center of the reductive group $C_G(\mathcal{A}(S, g_k))$ and any limit point of the sequence $[(g_k)_*\mu_{Hx}]$ is a translate of the equivalence class $[\mu_{C_G(\mathcal{A}(S,g_k))^0x}]$. In particular, if the subspace $\mathcal{A}(S, g_k) = \{0\}$, then the sequence $[(g_k)_*\mu_{Hx}]$ converges to the equivalence class of the Haar measure on $\mathrm{SL}(n,\mathbb{R})/\mathrm{SL}(n,\mathbb{Z})$.

Remark 2.6. The proof of Theorem 2.4 also gives a criterion on the convergence of $[(g_k)_*\mu_{Ax}]$. Similar criterion on the convergence of $[(g_k)_*\mu_{Hx}]$ for a connected reductive group H containing A could be obtained from the proof of Theorem 2.5.

We give several examples to illustrate Theorem 2.4 and Theorem 2.5.

(1) Let $G = \mathrm{SL}(3,\mathbb{R})$ and $\Gamma = \mathrm{SL}(3,\mathbb{Z})$. Pick the initial point $x = \mathbb{Z}^n \in X$ and the sequence $g_k = \begin{pmatrix} 1 & k & k^2/2 \\ 0 & 1 & k \\ 0 & 0 & 1 \end{pmatrix}$. In this case one can show that the subalgebra $\mathcal{A}(A, g_k) = \{0\}$ and $C_G(\mathcal{A}(A, g_k)) = \mathrm{SL}(3,\mathbb{R})$. Theorem 2.4 then says that $[(g_k)_*\mu_{Ax}]$ converges to $[\mu_{\mathrm{SL}(3,\mathbb{R})x}] = [m_X]$, i.e. the class of the uniform measure m_X.

(2) Fix G, Γ, x and g_k as in example (1). Let H be the connected component of the reductive subgroup

$$\begin{pmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & * \end{pmatrix} \cap \mathrm{SL}(3,\mathbb{R})$$

in $\mathrm{SL}(3,\mathbb{R})$. Then the center S of H is equal to $\left\{ \left(\begin{smallmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a^{-2} \end{smallmatrix}\right) : a \neq 0 \right\}$, and it is easy to see that the subalgebra $\mathcal{A}(S, g_k) = \{0\}$ and $C_G(\mathcal{A}(S, g_k)) = \mathrm{SL}(3,\mathbb{R})$. Then Theorem 2.5 implies that the sequence $[(g_k)_*\mu_{Hx}]$ converges to $[\mu_{\mathrm{SL}(3,\mathbb{R})x}] = [m_X]$.

(3) Let $G = \mathrm{SL}(4,\mathbb{R})$ and $\Gamma = \mathrm{SL}(4,\mathbb{Z})$. Pick the initial point $x = \mathbb{Z}^n \in X$ and the sequence $g_k = \left(\begin{smallmatrix} 1 & k & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & k \\ 0 & 0 & 0 & 1 \end{smallmatrix}\right)$. In this case one can show that the subalgebra

$$\mathcal{A}(A, g_k) = \left\{ \begin{pmatrix} t & 0 & 0 & 0 \\ 0 & t & 0 & 0 \\ 0 & 0 & -t & 0 \\ 0 & 0 & 0 & -t \end{pmatrix} : t \in \mathbb{R} \right\}$$

and

$$C_G(\mathcal{A}(A, g_k)) = \begin{pmatrix} * & * & 0 & 0 \\ * & * & 0 & 0 \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{pmatrix} \cap \mathrm{SL}(4,\mathbb{R}).$$

Theorem 2.4 then says that any limit point of the sequence $[(g_k)_*\mu_{Ax}]$ is a translate $[\mu_{C_G(\mathcal{A}(A,g_k))^0 x}]$. In fact, we will see in the proof of Theorem 2.4 that in this particular example, the sequence $[(g_k)_*\mu_{Ax}]$ actually converges to $[\mu_{C_G(\mathcal{A}(A,g_k))^0 x}]$.

(4) Fix G, Γ and x as in example (3), and pick the sequence $g_k = \left(\begin{smallmatrix} 1 & k & k^2/2 & 0 \\ 0 & 1 & k & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{smallmatrix}\right)$. Let H be the connected component of the reductive subgroup

$$\begin{pmatrix} * & * & 0 & 0 \\ * & * & 0 & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & * \end{pmatrix} \cap \mathrm{SL}(4,\mathbb{R})$$

in $\mathrm{SL}(4,\mathbb{R})$. Then the center S of H is equal to $\left\{ \left(\begin{smallmatrix} a & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & b & 0 \\ 0 & 0 & 0 & c \end{smallmatrix}\right) : a^2bc = 1 \right\}$, and it is easy to see that the subalgebra

$$\mathcal{A}(S, g_k) = \mathcal{A}(S, g_k) = \left\{ \begin{pmatrix} s & 0 & 0 & 0 \\ 0 & s & 0 & 0 \\ 0 & 0 & s & 0 \\ 0 & 0 & 0 & -3s \end{pmatrix} : s \in \mathbb{R} \right\}$$

and

$$C_G(\mathcal{A}(S, g_k)) = \begin{pmatrix} * & * & * & 0 \\ * & * & * & 0 \\ * & * & * & 0 \\ 0 & 0 & 0 & * \end{pmatrix} \cap \mathrm{SL}(4, \mathbb{R})$$

in $\mathrm{SL}(4, \mathbb{R})$. In this case, Theorem 2.5 tells that any limit point of the sequence $[(g_k)_*\mu_{Hx}]$ is a translate $[\mu_{C_G(\mathcal{A}(S,g_k))^0 x}]$, and the proof of Theorem 2.5 would imply that $[(g_k)_*\mu_{Hx}]$ actually converges to $[\mu_{C_G(\mathcal{A}(S,g_k))^0 x}]$ for this sequence $\{g_k\}$.

By comparing example (1) and (3) (resp. (2) and (4)), one can see that the subalgebra $\mathcal{A}(A, g_k)$ (resp. $\mathcal{A}(S, g_k)$) plays an important role in indicating what kinds of limit points the sequence $[(g_k)_*\mu_{Ax}]$ (resp. $[(g_k)_*\mu_{Hx}]$) could converge to. In example (1), we have $\mathcal{A}(A, g_k) = \{0\}$. By pushing the orbit Ax with g_k, the sequence $\{g_k Ax\}$ starts snaking in the space $\mathrm{SL}(3, \mathbb{R})/\mathrm{SL}(3, \mathbb{Z})$, and eventually fills up the entire space. In example (3), $\mathcal{A}(A, g_k)$ is a 1-dimensional subalgebra in $\mathrm{Lie}(A)$ which commutes with g_k, and it corresponds to the part of the orbit Ax which stays still and is not affected when we push μ_{Ax} by g_k. This would result in the limit orbit having this part as the 'central direction', and the 'orthogonal' part in Ax would be pushed by g_k and fill up the sub-homogeneous space $\begin{pmatrix} \mathrm{SL}(2, \mathbb{R}) & 0 \\ 0 & \mathrm{SL}(2, \mathbb{R}) \end{pmatrix} x$ in $\mathrm{SL}(4, \mathbb{R})/\mathrm{SL}(4, \mathbb{Z})$.

From the characterization of convergence given in Proposition 3.3, Theorem 2.4 and Theorem 2.5 can be restated in the form of the following

Theorem 2.7. *Let Ax be a divergent orbit and $\{g_k\}$ be any sequence in G with $[(g_k)_*\mu_{Ax}]$ converging to an equivalence class of locally finite algebraic measures $[\nu]$ as in Theorem 2.4. Then there exists a sequence $\lambda_k > 0$ such that*

$$\lambda_k (g_k)_*\mu_{Ax} \to \nu$$

with respect to the topology τ_X. In particular, for any $F_1, F_2 \in C_c(\mathrm{SL}(n, \mathbb{R})/\mathrm{SL}(n, \mathbb{Z}))$ we have

$$\frac{\int F_2 d(g_k)_*\mu_{Ax}}{\int F_1 d(g_k)_*\mu_{Ax}} \to \frac{\int F_2 d\nu}{\int F_1 d\nu}$$

whenever $\int F_1 d\nu \neq 0$. The same results hold if A is replaced by any connected reductive group H containing A.

Remark 2.8. From the proof of Theorem 2.4, we will see directly that in the case $\mathcal{A}(A, g_k) = \{0\}$, the numbers λ_k's in Theorem 2.7 are actually related to volumes of convex polytopes of a special type in $\mathrm{Lie}(A)$ (see Definition 4.1 and Corollary 10.1). We remark here that in view of Theorem 2.7, the λ_k's in this case can also be calculated by a function $F_1 \in C_c(\mathrm{SL}(n, \mathbb{R})/\mathrm{SL}(n, \mathbb{Z}))$ with its support being a large compact subset. This makes Theorem 2.7 practical in other problems.

2.3. **Applications.** As an application of our results, we give one example of a counting problem. More details about such counting problem could be found in [DRS93], [EM93], [EMS96] and [Sha00].

Let $M(n,\mathbb{R})$ be the space of $n\times n$ matrices with the norm

$$\|M\|^2 = \mathrm{Tr}(M^t M) = \sum_{1\le i,j\le n} x_{ij}^2$$

for $M=(x_{ij})_{1\le i,j\le n}\in M(n,\mathbb{R})$. Denote by B_T the ball of radius T centered at 0 in $M(n,\mathbb{R})$. Fix a monic polynomial $p_0(\lambda)$ in $\mathbb{Z}[\lambda]$ which splits completely over $\mathbb{Q}$. By Gauss Lemma, the roots α_i of $p(\lambda)$ are integers. We assume that the α_i's are distinct and nonzero. Let

$$M_\alpha = \mathrm{diag}(\alpha_1,\alpha_2,\dots,\alpha_n)\in M(n,\mathbb{Z}).$$

For $M\in M(n,\mathbb{R})$, denote by $p_M(\lambda)$ the characteristic polynomial of M. We define

$$V(\mathbb{R}) := \{M\in M(n,\mathbb{R}) : p_M(\lambda)=p_0(\lambda)\}$$

the variety of matrices M with characteristic polynomial $p_M(\lambda)$ equal to $p_0(\lambda)$, and

$$V(\mathbb{Z}) = \{M\in M(n,\mathbb{Z}) : p_M(\lambda)=p_0(\lambda)\}$$

the integer points in the variety $V(\mathbb{R})$.

There is a natural volume form on the variety $V(\mathbb{R})$ inherited from $G=\mathrm{SL}(n,\mathbb{R})$. Specifically, the orbit map

$$G\to V(\mathbb{R})$$

defined by $g\mapsto \mathrm{Ad}(g)M_\alpha$ gives an isomorphism between the quotient space G/A and the variety $V(\mathbb{R})$, and the volume form is defined to be the G-invariant measure on G/A. The existence of such a measure is well-known, and the proof of it could be found, for example, in [Rag72]. With this volume form, one can compute (see Proposition 11.7) that for any T, the volume of $V(\mathbb{R})\cap B_T$

$$\mathrm{Vol}(V(\mathbb{R})\cap B_T) = \frac{\mathrm{Vol}(B_1)}{\prod_{j>i}|\alpha_j-\alpha_i|}T^{n(n-1)/2}.$$

The following theorem concerns the asymptotic formula for the number of integer points in $V(\mathbb{Z})\cap B_T$. We will see that the set $V(\mathbb{Z})\cap B_T$ behaves differently from $V(\mathbb{R})\cap B_T$, with an extra natural log term.

By a well-known theorem of Borel and Harish-Chandra [BHC62], the subset $V(\mathbb{Z})$ is a finite disjoint union of $\mathrm{Ad}(\Gamma)$-orbits where $\Gamma=\mathrm{SL}(n,\mathbb{Z})$. One can write this disjoint union as

$$V(\mathbb{Z}) = \bigcup_{i=1}^{h_0}\mathrm{Ad}(\Gamma)M_i$$

for some $h_0\in\mathbb{N}$ and $M_i\in V(\mathbb{Z})$ ($1\le i\le h_0$). Note that for each M_i, the stabilizer Γ_{M_i} of M_i is finite. Also the number of the orbits h_0 is equal to the number of equivalence classes of nonsingular ideals in the subring in $M(n,\mathbb{R})$ generated by M_α, for which readers may refer to [BHC62] and

[LM33]. In the following theorem, to ease the notation, we write $\mathbf{t}$ for a vector $(t_1, t_2, \ldots, t_n) \in \mathbb{R}^n$.

Theorem 2.9. *We have*

$$|V(\mathbb{Z}) \cap B_T| \sim \left(\sum_{i=1}^{h_0} \frac{1}{|\Gamma_{M_i}|}\right) \frac{c_0 \operatorname{Vol}(B_1)}{\prod_{j>i} |\alpha_j - \alpha_i|} T^{n(n-1)/2} (\ln T)^{n-1}$$

where $\operatorname{Vol}(B_1)$ *is the volume of the ball of radius one in* $\mathbb{R}^{n(n-1)/2}$ *and* c_0 *is the volume of the* $(n-1)$*-convex polytope*

$$\left\{\mathbf{t} \in \mathbb{R}^n : \sum_{i=1}^n t_i = 0, \sum_{j=1}^l t_{i_j} \geq \sum_{j=1}^l (j - i_j), \forall 1 \leq l \leq n, \forall i_1 < \cdots < i_l \right\}$$

with respect to the natural measure induced by the Lebesgue measure on $\mathbb{R}^n$.

In the sequel, we will mainly focus on Theorem 2.4 as all the other theorems will be corollaries of it. In the course of the proof of Theorem 2.4, the case $\mathcal{A}(A, g_k) = \{0\}$ plays an important role, and other cases could be proved by induction. Therefore, most of our arguments in this paper would work for the case $\mathcal{A}(A, g_k) = \{0\}$. We remark here that our proof is inspired by [OS14], where Hee Oh and Nimish Shah deal with the case $G = \mathrm{SL}(2, \mathbb{R})$ by applying exponential mixing and obtain an error estimate. This effective result is improved recently in [KK] by Dubi Kelmer and Alex Kontorovich.

When we showed an earlier draft of the manuscript to Nimish Shah, he pointed out to us that similar results to those appearing in this paper were established by him at the beginning of this century, but were never published.

The paper is organized as follows:

- We start our work in section 3 by studying the topology τ_P on $\mathbb{P}\mathcal{M}(X)$ for a locally compact second countable Hausdorff space X. In particular, a characterization of convergence in $\mathbb{P}\mathcal{M}(X)$ is given, and Theorem 2.7 is obtained as a natural corollary, if Theorem 2.4 and Theorem 2.5 are presumed.
- In section 4, a special type of convex polytopes in $\operatorname{Lie}(A)$ is introduced. Such a convex polytope is related to non-divergence of an orbit $g_k Ax$. To analyze these convex polytopes, we define graphs associated to them and prove some auxiliary results concerning the graphs in section 5. With the assumption $\mathcal{A}(A, g_k) = \{0\}$, these auxiliary results imply some properties of the convex polytopes, which we prove in section 6.
- Keeping the assumption $\mathcal{A}(A, g_k) = \{0\}$ in section 7, we prove a statement on the non-divergence of the sequence of $[(g_k)_* \mu_{Ax}]$ and show that $[(g_k)_* \mu_{Ax}]$ converges to $[\nu]$ for some probability measure ν invariant under a unipotent subgroup. Then we translate section 7 in terms of adjoint representation in section 8. The linearization

technique and the measure classification theorem for unipotent actions on homogeneous spaces are discussed in section 9, which enable us to study the measure rigidity in our setting.

- We complete the proofs of Theorem 2.4 and Theorem 2.5 in section 10. The proof of Theorem 2.9 is given in section 11.

Acknowledgments. We would like to express our gratitude to Nimish Shah for insightful comments and support on this work. We are grateful to Barak Weiss for valuable communications and Roy Meshulam for teaching us Lemma 4.4. We also thank Ofir David, Asaf Katz, Rene Ruhr, Oliver Sargent, Lei Yang, Pengyu Yang and Runlin Zhang for helpful discussions and support. Finally, the authors acknowledge the support of ISF grant numbers 871/17 and 357/13, and the second author is in part supported at the Technion by a Fine Fellowship.

3. Topology on $\mathbb{P}\mathcal{M}(X)$

In this section, we study the topology τ_P on $\mathbb{P}\mathcal{M}(X)$ for any locally compact second countable Hausdorff space X. We will give a description of the convergence of a sequence $[\mu_k]$ in $\mathbb{P}\mathcal{M}(X)$ (Proposition 3.3). This will help us study the convergence of the sequence $[(g_k)_*\mu_{Ax}]$ in Theorem 2.4 (resp. $[(g_k)_*\mu_{Hx}]$ in Theorem 2.5).

Before proving Proposition 3.3, we need some preparations.

Proposition 3.1. *The topology $(\tau_P, \mathbb{P}\mathcal{M}(X))$ is Hausdorff. In particular, any convergent sequence in $\mathbb{P}\mathcal{M}(X)$ has a unique limit.*

Proof. Let $[\mu]$ and $[\nu]$ be two distinct elements in $\mathbb{P}\mathcal{M}(X)$. We choose $f \in C_c(X)$ and representatives μ and ν such that

$$\int f d\mu = \int f d\nu = 1.$$

Since $[\mu] \neq [\nu]$, there exists a nonnegative function $g \in C_c(X)$ such that

$$\frac{\int g d\mu}{\int g d\nu} \neq 1.$$

We define neighborhoods of μ and ν in $\mathcal{M}(X)$ by

$$V(\mu; f, g, \epsilon) = \left\{\lambda : \left|\int g d\lambda - \int g d\mu\right| < \epsilon, \left|\int f d\lambda - \int f d\mu\right| < \epsilon\right\}$$

$$V(\nu; f, g, \epsilon) = \left\{\lambda : \left|\int g d\lambda - \int g d\nu\right| < \epsilon, \left|\int f d\lambda - \int f d\nu\right| < \epsilon\right\}.$$

Since $\pi_P : \mathcal{M}(X) \to \mathbb{P}\mathcal{M}(X)$ is an open map, $\pi_P(V(\mu; f, g, \epsilon))$ and $\pi_P(V(\nu; f, g, \epsilon))$ are open neighborhoods of $[\mu]$ and $[\nu]$ in $\mathbb{P}\mathcal{M}(X)$. We prove that for sufficiently small $\epsilon > 0$

$$\pi_P(V(\mu; f, g, \epsilon)) \cap \pi_P(V(\nu; f, g, \epsilon)) = \emptyset.$$

Suppose, on the contrary, that $[\lambda] \in \pi_P(V(\mu; f, g, \epsilon)) \cap \pi_P(V(\nu; f, g, \epsilon))$. Then there exist constants $\alpha, \beta > 0$ such that

$$\left|\alpha \int g d\lambda - \int g d\mu\right| < \epsilon, \left|\alpha \int f d\lambda - \int f d\mu\right| < \epsilon$$

$$\left|\beta \int g d\lambda - \int g d\nu\right| < \epsilon, \left|\beta \int f d\lambda - \int f d\nu\right| < \epsilon.$$

This implies that

$$\frac{\int g d\mu - \epsilon}{\int g d\nu + \epsilon} < \frac{\alpha}{\beta} < \frac{\int g d\mu + \epsilon}{\int g d\nu - \epsilon}$$
$$\frac{\int f d\mu - \epsilon}{\int f d\nu + \epsilon} < \frac{\alpha}{\beta} < \frac{\int f d\mu + \epsilon}{\int f d\nu - \epsilon}$$

and we get a contradiction for sufficiently small $\epsilon > 0$. □

Proposition 3.2. *A sequence $[\mu_k] \in \mathbb{P}\mathcal{M}(X)$ converges to $[\nu]$ if and only if for each $k \in \mathbb{N}$ there exists a representative μ_k' in $[\mu_k]$ and for $[\nu]$ a representative $\nu' \in [\nu]$ such that μ_k' converges to ν' in $\mathcal{M}(X)$.*

Proof. Suppose that $[\mu_k] \to [\nu]$. We choose $f \in C_c(X)$ and representatives μ_k' and ν' such that

$$\int f d\mu_k' = \int f d\nu' = 1.$$

Suppose that $\mu_k' \not\to \nu'$ in $\mathcal{M}(X)$. Then there exists $g \in C_c(X)$ such that after passing to a subsequence

$$\left|\int g d\mu_k' - \int g d\nu'\right| \geq \delta$$

for some $\delta > 0$. Here one may assume $\int g d\nu' \neq 0$. Then by the same arguments as in Proposition 3.1, we can find a neighborhood $\pi_P(V(\nu; f, g, \epsilon))$ of $[\nu]$ in $\mathbb{P}\mathcal{M}(X)$ such that

$$[\mu_k] \notin \pi_P(V(\nu; f, g, \epsilon))$$

which contradicts the condition $[\mu_k] \to [\nu]$. The other direction follows from Definition 2.1. □

Now we are in the position to prove the following proposition, which provides a characterization of the convergence of a sequence $[\mu_k]$ in $\mathbb{P}\mathcal{M}(X)$. This will help us study the convergence of equivalence classes of locally finite measures on $\mathrm{SL}(n, \mathbb{R})/\mathrm{SL}(n, \mathbb{Z})$ in the rest of the paper.

Proposition 3.3. *(1) Let μ_k be a sequence in $\mathcal{M}(X)$. Then $[\mu_k]$ converges to $[\nu]$ in $\mathbb{P}\mathcal{M}(X)$ if and only if there exists a sequence $\{\lambda_k\}$ of positive numbers such that $\lambda_k \mu_k$ converges to ν in $\mathcal{M}(X)$. If there exists another sequence $\{\lambda_k'\}$ with $\lambda_k' \mu_k \to \nu' \neq 0$ in $\mathcal{M}(X)$, then*

$$[\nu'] = [\nu]$$

and $\lim_k \lambda_k'/\lambda_k$ exists.

(2) The sequence $[\mu_k]$ *converges to* $[\nu]$ *if and only if for any* $f, g \in C_c(X)$ *with* $\int g d\nu \neq 0$*, we have* $\int g d\mu_k \neq 0$ *for sufficiently large* k *and*

$$\frac{\int f d\mu_k}{\int g d\mu_k} \to \frac{\int f d\nu}{\int g d\nu}.$$

Proof. The first statement follows from Proposition 3.1 and Proposition 3.2. For $\lim_k \lambda'_k/\lambda_k$, we choose $f \in C_c(X)$ with $\int f d\nu \neq 0$, and we have

$$\frac{\lambda'_k}{\lambda_k} = \frac{\lambda'_k \int f d\mu_k}{\lambda_k \int f d\mu_k} \to \frac{\int f d\nu'}{\int f d\nu}.$$

For the second statement, if $[\mu_k] \to [\nu]$, then there exists a sequence $\lambda_k > 0$ such that $\lambda_k \mu_k \to \nu \neq 0$. For any $f, g \in C_c(X)$ with $\int g d\nu \neq 0$ we have

$$\lambda_k \int g d\mu_k \neq 0$$

for sufficiently large k and

$$\frac{\int f d\mu_k}{\int g d\mu_k} = \frac{\int f d(\lambda_k \mu_k)}{\int g d(\lambda_k \mu_k)} \to \frac{\int f d\nu}{\int g d\nu}.$$

Conversely, let $g \in C_c(X)$ with $\int g d\nu \neq 0$ and

$$\lambda_k = \frac{\int g d\nu}{\int g d\mu_k}.$$

Then we have $\lambda_k \mu_k \to \nu$ and $[\mu_k] \to [\nu]$. □

Remark 3.4. This shows that Theorem 2.7 is equivalent to Theorem 2.4 and Theorem 2.5.

From the discussions in this section, we know that to prove Theorem 2.4, one needs to find a sequence of $\lambda_k > 0$ such that $\lambda_k (g_k)_* \mu_{Ax}$ converges to a locally finite measure ν, and then prove that ν is a periodic measure. From section 4 to section 6, we will construct the sequence λ_k in an explicit way.

4. Convex polytopes

Let $\{e_1, e_2, \ldots, e_n\}$ be the standard basis of $\mathbb{R}^n$. We write $K = \mathrm{SO}(n, \mathbb{R})$ the maximal compact subgroup in G, the connected component of the full diagonal subgroup

$$A = \left\{ \mathrm{diag}(e^{t_1}, e^{t_2}, \ldots, e^{t_{n-1}}, e^{t_n}) : t_1 + t_2 + \cdots + t_n = 0 \right\}$$

and the upper triangular unipotent subgroup

$$N = \{(u_{ij}) : u_{ii} = 1, u_{ij} = 0 (i > j)\}.$$

In this section, we will construct a special type of convex polytopes in $\mathrm{Lie}(A)$. These convex polytopes will play an important role in our proof.

By Theorem 1.4 in [TW03], Ax is divergent if and only if $x \in A \cdot \mathrm{SL}(n, \mathbb{Q})\Gamma$. Note that for any $q \in \mathrm{SL}(n, \mathbb{Q})$ the lattice $q\Gamma q^{-1}$ is commensurable with Γ, and all results in this paper would hold if Γ is replaced by $q\Gamma q^{-1}$.

Therefore, without loss of generality, we may assume that the initial point $x = e\,\mathrm{SL}(n,\mathbb{Z})$.

To ease the notations, we will write $\mathbf{t}$ for a vector $(t_1, t_2, \ldots, t_n)$ in a n-dimensional space, and $[n]$ will denote the index set $\{1, 2, \ldots, n\}$. We write $\mathcal{I}_n$ for the collection of all multi-index subsets of $[n]$, and $\mathcal{I}_n^l$ for the collection of the index subsets of cardinality l in $\mathcal{I}_n$. For $\mathbb{R}^n$ with the basis $\{e_1, e_2, \ldots, e_n\}$, and for any index subset $I = \{i_1 < i_2 \cdots < i_l\}$ of $[n]$, we denote by

$$e_I := e_{i_1} \wedge \cdots \wedge e_{i_l}$$

the wedge product of the vectors in $\{e_1, \ldots, e_n\}$ indexed by I. We will use $\omega_I(\mathbf{t})$ ($\mathbf{t} = (t_1, t_2, \ldots, t_n) \in \mathbb{R}^n$) for the linear functional $\sum_{i \in I} t_i$ on $\mathbb{R}^n$.

Let $g \in \mathrm{SL}(n,\mathbb{R})$ and $\delta > 0$. We define a region $\Omega_{g,\delta}$ in $\mathrm{Lie}(A)$ as follows. Let $\mathbf{t} = (t_1, t_2, \ldots, t_n) \in \mathrm{Lie}(A)$. For each $e_i \in \mathbb{R}^n$, the vector

$$g\exp(\mathbf{t})e_i = e^{t_i} g e_i \notin B_\delta$$

if and only if

$$t_i \geq \ln\delta - \ln\|ge_i\|.$$

Here we denote by B_δ the ball of radius $\delta > 0$ around 0 in $\mathbb{R}^n$ with the standard Euclidean norm $\|\cdot\|$. We also consider the wedge product e_I for any nonempty subset $I \in \mathcal{I}_n$ and

$$g\exp(\mathbf{t})e_I = e^{\omega_I(\mathbf{t})} g e_I \notin B_\delta$$

if and only if

$$\omega_I(\mathbf{t}) \geq \ln\delta - \ln\|ge_I\|.$$

This leads to the following

Definition 4.1. For any $g \in \mathrm{SL}(n,\mathbb{R})$ and $\delta > 0$, we define in $\mathrm{Lie}(A)$

$$\Omega_{g,\delta} = \{\mathbf{t} \in \mathrm{Lie}(A) : \omega_I(\mathbf{t}) \geq \ln\delta - \ln\|ge_I\| \text{ for any nonempty } I \in \mathcal{I}_n\}.$$

Remark 4.2. By the construction above, for any $\mathbf{t} \in \mathrm{Lie}(A) \setminus \Omega_{g,\delta}$, the lattice $g\exp(\mathbf{t})\mathbb{Z}^n$ has a short nonzero vector with the length depending on $\delta > 0$, and hence by Mahler's compactness criterion, the point $g\exp(\mathbf{t})\Gamma \in gA\Gamma$ is close to infinity. By this reason, we will mainly study $gA\Gamma$ inside $\Omega_{g,\delta}$.

Lemma 4.3. *The region $\Omega_{g,\delta}$ is a bounded convex polytope in* $\mathrm{Lie}(A)$ *for any $\delta > 0$.*

Proof. Since the region $\Omega_{g,\delta}$ is defined by various linear functionals on $\mathrm{Lie}(A)$, $\Omega_{g,\delta}$ is a convex polygon. Now by definition, $\Omega_{g,\delta}$ is contained in the following region

$$\left\{\mathbf{t} \in \mathbb{R}^n : \sum_{i=1}^n t_i = 0, t_i \geq \ln\delta - \ln\|ge_i\|, \forall i \in [n]\right\}$$

which is bounded. The boundedness of $\Omega_{g,\delta}$ then follows. $\square$

In section 6, we will closely study the convex polytope $\Omega_{g,\delta}$. We list here some properties of convex polytopes which will be used later. The following lemma is well known. We learnt it from Roy Meshulam.

Lemma 4.4. *Let Ω be a convex subset in $\mathbb{R}^d$. Suppose that Ω contains a ball of radius $r > 0$. Then we have*

$$\frac{\mathrm{Vol}(\partial\Omega)}{\mathrm{Vol}(\Omega)} \leq \frac{d}{r}.$$

Proof. Let $B_r(0)$ denote the ball of radius r centered at 0 in $\mathbb{R}^d$ and we may assume, without loss of generality, that $B_r(0) \subset \Omega$. We have

$$\begin{aligned}
\mathrm{Vol}(\partial\Omega) &= \lim_{\epsilon\to 0} \frac{\mathrm{Vol}(\Omega + \epsilon B_1(0)) - \mathrm{Vol}(\Omega)}{\epsilon} \\
&= \lim_{\epsilon\to 0} \frac{\mathrm{Vol}(\Omega + (\epsilon/r) B_r(0)) - \mathrm{Vol}(\Omega)}{\epsilon} \\
&\leq \lim_{\epsilon\to 0} \frac{\mathrm{Vol}(\Omega + (\epsilon/r)\Omega) - \mathrm{Vol}(\Omega)}{\epsilon} \\
&= \lim_{\epsilon\to 0} \frac{(1 + (\epsilon/r))^d - 1}{\epsilon} \mathrm{Vol}(\Omega) \\
&= \frac{d}{r} \mathrm{Vol}(\Omega).
\end{aligned}$$

□

Lemma 4.5. *Let $R \subset \Omega$ be two bounded d-dimensional convex polytopes in $\mathbb{R}^d$. Suppose that Ω contains a ball of radius $r > 0$ and*

$$\frac{\mathrm{Vol}(R)}{\mathrm{Vol}(\Omega)} \geq c$$

for some constant $c > 0$. Then R contains a ball of radius rc/d.

Proof. Let ρ be the largest number such that R contains a ball of radius ρ. It suffices to show that $\rho \geq rc/d$. First, we claim

$$\mathrm{Vol}(R) \leq \rho \, \mathrm{Vol}(\partial R).$$

Proof of the claim. Let $\{f_i\}$ be the collection of the facets of R, and denote by P_i the hyperplane determined by f_i. For each f_i, let B_i be the unique cylinder with the following properties:

(1) the base of B_i is f_i, and the height of B_i is equal to ρ.
(2) B_i and R lie in the same half-space determined by P_i.

The maximality of ρ then implies

$$R \subset \bigcup_i B_i;$$

otherwise, one would find a point $x \in R$ such that for each f_i, the distance between x and f_i is strictly larger than ρ. Now we have

$$\mathrm{Vol}(R) \leq \sum_i \mathrm{Vol}(B_i) = \rho \sum \mathrm{Vol}(f_i) = \rho \,\mathrm{Vol}(\partial R)$$

and the claim follows. □

Now we can finish the proof of the lemma. By Lemma 4.4 and the claim above, we have

$$\rho \geq \frac{\mathrm{Vol}(R)}{\mathrm{Vol}(\partial R)} \geq \frac{c\,\mathrm{Vol}(\Omega)}{\mathrm{Vol}(\partial\Omega)} \geq \frac{cr}{d}.$$

Here we use the fact that $\mathrm{Vol}(\partial R) \leq \mathrm{Vol}(\partial \Omega)$ for any two convex polytopes $R \subset \Omega$. □

By Iwasawa decomposition, for each element $g \in \mathrm{SL}(n,\mathbb{R})$ we can write

$$g = kua$$

where $k \in K = \mathrm{SO}(n,\mathbb{R})$, $u \in N$ and $a \in A$. Note that μ_A is A-invariant, and we have

$$g_* \mu_A = (ku)_* \mu_A.$$

Because of this and since we will consider all the possible limits of $\{(g_k)_* \mu_A\}$ for $g_k \in G$, it is harmless to assume that all g_k belong to the upper triangular unipotent group N. In other words, we have

$$g_k = (u_{ij}(k))_{1 \leq i,j \leq n}$$

where $u_{ij}(k) = 0$ $(i > j)$ and $u_{ii}(k) = 1$. Moreover, using Gauss elimination and by the same reason, we can assume, after passing to a subsequence, the following dichotomy for each entry $u_{ij}(k)$ $(i \neq j)$ as $k \to \infty$:

$$\text{either } u_{ij}(k) \to \infty \text{ or } u_{ij}(k) = 0.$$

Unless something else is specified, we will work under these assumptions on $\{g_k\}$ in the rest of the paper.

5. Auxiliary results in graph theory

In this section, we will study a special class of graphs and prove some properties of these graphs (Proposition 5.5 and Lemma 5.8), which will be crucial to our study in convex polytopes in section 6. We continue the assumptions on $\{g_k\}$ at the end of section 4, and work in the homogeneous space $X = \mathrm{SL}(n,\mathbb{R})/\mathrm{SL}(n,\mathbb{Z})$.

In order to prove Proposition 5.5, we will need some lemmas involving complex calculations which will guarantee the validity of the arguments in the proof of Proposition 5.5. Here we introduce the following notation. For any $g \in \mathrm{SL}(n,\mathbb{R})$ and any $1 \leq l \leq n$, we will denote by $(g)_{l\times l}$ the $l \times l$ submatrix in the upper left corner of g. Note that if $g, h \in \mathrm{SL}(n,\mathbb{R})$ are upper triangular, then $(gh)_{l\times l} = (g)_{l\times l}(h)_{l\times l}$.

Lemma 5.1. *For any $a \in A$ and any $1 \leq l \leq n$, we have either $(g_k)_{l\times l} = (a^{-1}g_k a)_{l\times l}$ for all sufficiently large k or $(g_k)_{l\times l} \neq (a^{-1}g_k a)_{l\times l}$ for all sufficiently large k*

Proof. Write $a = (a_1, a_2, \ldots, a_n) \in A$ and $g_k = (u_{ij}(k))_{n\times n}$. By definition, we have

$$(g_k)_{l\times l} = (u_{ij}(k))_{1\leq i,j\leq l}$$

and

$$(a^{-1}g_k a)_{l\times l} = (a_i^{-1}a_j u_{ij}(k))_{1\leq i,j\leq l}.$$

The equation $(g_k)_{l\times l} = (a^{-1}g_k a)_{l\times l}$ then yields

$$\text{either } u_{ij}(k) = 0 \text{ or } a_i = a_j, \quad \forall 1 \leq i,j \leq l.$$

Now the lemma follows from our dichotomy assumption on the entries of g_k. □

Lemma 5.2. *Let $a \in A$. Suppose the sequence $\{g_k a g_k^{-1} : k \in \mathbb{N}\}$ is bounded in $\mathrm{SL}(n,\mathbb{R})$. Then for sufficiently large k, g_k commutes with a.*

Proof. Suppose not. Then by Lemma 5.1 with $l = n$, for sufficiently large k, we have

$$g_k \neq a^{-1}g_k a.$$

In this case, we would like to find a contradiction.

Let l_0 be the minimum of the integers $0 \leq l \leq n-1$ with the property

$$(g_k)_{(l+1)\times(l+1)} \neq (a^{-1}g_k a)_{(l+1)\times(l+1)}$$

for sufficiently large k. Again the existence of such l_0 is guaranteed by Lemma 5.1. In other words, l_0 is the maximum of $0 \leq l \leq n-1$ such that $(g_k)_{l\times l}$ commutes with $(a)_{l\times l}$ for sufficiently large k.

We write $a = \mathrm{diag}(a_1, a_2, \ldots, a_n) \in A$. Then for any $1 \leq l \leq n$

$$(a)_{l\times l} = \mathrm{diag}(a_1, a_2, \ldots, a_l).$$

We also write g_k as

$$g_k = \begin{pmatrix} (g_k)_{l_0\times l_0} & \mathbf{v}_k & \ldots \\ 0 & 1 & \ldots \\ \vdots & \vdots & \vdots \\ 0 & 0 & \ldots \end{pmatrix} \in \mathrm{SL}(n,\mathbb{R})$$

where $\mathbf{v}_k$ is the l_0-dimensional column vector next to $(g_k)_{l_0\times l_0}$ in g_k. Note that according to this expression, $(g_k)_{(l_0+1)\times(l_0+1)}$ could be written as

$$(g_k)_{(l_0+1)\times(l_0+1)} = \begin{pmatrix} (g_k)_{l_0\times l_0} & \mathbf{v}_k \\ 0 & 1 \end{pmatrix}.$$

Since $(g_k)_{l_0\times l_0}$ commutes with $(a)_{l_0\times l_0}$, one can compute

$$\begin{aligned} &(a^{-1}g_k a)_{(l_0+1)\times(l_0+1)} \\ =&(a^{-1})_{(l_0+1)\times(l_0+1)}(g_k)_{(l_0+1)\times(l_0+1)}(a)_{(l_0+1)\times(l_0+1)} \end{aligned}$$

$$= \begin{pmatrix} (g_k)_{l_0\times l_0} & a_{l_0+1}(a^{-1})_{l_0\times l_0}\mathbf{v}_k \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} (g_k)_{l_0\times l_0} & \mathbf{w}_k \\ 0 & 1 \end{pmatrix}$$

where

$$\mathbf{w}_k := a_{l_0+1}(a^{-1})_{l_0\times l_0}\mathbf{v}_k.$$

As $(g_k)_{(l_0+1)\times(l_0+1)}$ does not commute with $(a)_{(l_0+1)\times(l_0+1)}$ for sufficiently large k, we have

$$\mathbf{v}_k \neq \mathbf{w}_k$$

for sufficiently large k. From this and our dichotomy assumption on the entries of g_k, one can then easily deduce that $\mathbf{v}_k \neq \mathbf{0}$, $\mathbf{v}_k \to \infty$ and

$$\mathbf{w}_k - \mathbf{v}_k = \left(a_{l_0+1}(a^{-1})_{l_0\times l_0} - \mathrm{I}_{l_0}\right)\mathbf{v}_k \to \infty$$

as $k \to \infty$. Here I_{l_0} is the $l_0 \times l_0$ identity matrix.

Now one can easily compute

$$\begin{aligned} & a^{-1}g_k a g_k^{-1} = (a^{-1}g_k a)g_k^{-1} \\ = & \begin{pmatrix} (g_k)_{l_0\times l_0} & \mathbf{w_k} & \dots \\ 0 & 1 & \dots \\ \vdots & \vdots & \vdots \\ 0 & 0 & \dots \end{pmatrix} \begin{pmatrix} (g_k)_{l_0\times l_0} & \mathbf{v}_k & \dots \\ 0 & 1 & \dots \\ \vdots & \vdots & \vdots \\ 0 & 0 & \dots \end{pmatrix}^{-1} \\ = & \begin{pmatrix} \mathrm{I}_{l_0} & \mathbf{w}_k - \mathbf{v}_k & \dots \\ 0 & 1 & \dots \\ \vdots & \vdots & \vdots \\ 0 & 0 & \dots \end{pmatrix}. \end{aligned}$$

Since $\mathbf{w}_k - \mathbf{v}_k \to \infty$ as $k \to \infty$, the formula above implies that $\{a^{-1}g_k a g_k^{-1}\}$ diverges, which contradicts the boundedness of $g_k a g_k^{-1}$. This completes the proof of the lemma. □

The following is an immediate corollary of Lemma 5.1 and Lemma 5.2.

Corollary 5.3. *Let $S \subset A$ be a subgroup in A. Then for any $\mathbf{t} \in \mathrm{Lie}(S)$, either $\mathrm{Ad}(g_k)\mathbf{t} \to \infty$ as $k \to \infty$ or $\mathrm{Ad}(g_k)\mathbf{t} = \mathbf{t}$ for sufficiently large k. In particular, if the subalgebra $\mathcal{A}(S, g_k)$ of $\mathrm{Lie}(A)$ is not trivial, then there exists an element $\mathbf{t} \in \mathrm{Lie}(S)$ such that each g_k commutes with $\mathbf{t}$ for sufficiently large k.*

Proof. Apply Lemma 5.1 and Lemma 5.2 with $a = \exp(\mathbf{t})$. □

Definition 5.4. We define a graph from $\{g_k\}$ as follows. The vertex set V is the index set $[n] = \{1, 2, \dots, n\}$, and two vertices $i \neq j$ are connected by an edge $e \in E$, which we denote by $i \sim j$, if $u_{ij}(k) \to \infty$ as $k \to \infty$. In this way, we obtain a graph $G(g_k) = (V, E)$ associated to $\{g_k\}$.

Now we can prove our first result in this section.

Proposition 5.5. *The subalgebra $\mathcal{A}(A, g_k)$ of* Lie(A) *(as defined in Definition 2.2) is trivial if and only if the graph $G(g_k)$ associated to $\{g_k\}$ is connected.*

Proof. Suppose that the graph $G(g_k)$ associated to g_k is not connected. Let $G_i = (V_i, E_i)$ $(1 \le i \le m)$ be the connected components of $G(g_k)$. We pick $x_i \in \mathbb{R} \setminus \{0\}$ such that $\sum_{i=1}^m |V_i| x_i = 0$. For any vertex $j \in V_i$, we assign $t_j = x_i$. In this way we obtain an element $\mathbf{t} = (t_j) \in \mathrm{Lie}(A) \setminus \{0\}$. Note that $\mathbf{t}$ is invertible. We show that

$$g_k \mathbf{t} = \mathbf{t} g_k.$$

Indeed, since $\mathbf{t}$ is invertible, we compute

$$\mathbf{t} g_k \mathbf{t}^{-1} = (t_i t_j^{-1} u_{ij}(k)).$$

For $u_{ij}(k) \neq 0$, by the definition of the graph $G(g_k)$, the vertices i and j are in the same connected components. Hence we have $t_i = t_j$ and

$$\mathbf{t} g_k \mathbf{t}^{-1} = (t_i t_j^{-1} u_{ij}(k)) = (u_{ij}(k)) = g_k$$

as desired. This implies that $\mathrm{Ad}(g_k)$ fixes $\mathbf{t}$, and by definition $\mathbf{t} \in \mathcal{A}(A, g_k) \neq \{0\}$.

Now assume that the graph $G(g_k)$ is connected. Suppose that the subalgebra $\mathcal{A}(A, g_k)$ is not trivial. Then there exists an element $\mathbf{t} \in \mathrm{Lie}\, A \setminus \{0\}$ such that $\mathrm{Ad}(g_k)\mathbf{t}$ is bounded as $k \to \infty$.

Let $a = \exp \mathbf{t} \in A \setminus \{e\}$. Then $\{g_k a g_k^{-1}\}$ is bounded in $\mathrm{SL}(n, \mathbb{R})$. By lemma 5.2, g_k commutes with a. If we write $a = \mathrm{diag}(a_1, a_2, \ldots, a_n)$, then the equation $g_k = a g_k a^{-1}$ yields

$$(u_{ij}(k))_{1 \le i,j \le n} = (a_i a_j^{-1} u_{ij}(k))_{1 \le i,j \le n}$$

and hence $a_i = a_j$ whenever $u_{ij}(k) \neq 0$. The connectedness of the graph G then implies that all a_i's are equal and $a = e$, which contradicts $a \in A \setminus \{e\}$. This completes the proof of the proposition. □

Definition 5.6. Let $G(V, E)$ be a graph consisting of the set of vertices V and the set of edges E. Here we assume $V = \{v_1, v_2, \ldots, v_n\}$ is an ordered set with the ordering $\prec$, and we denote by $v_i \sim v_j$ if v_i and v_j are adjacent by an edge in E. A subset $S \subset V$ is called UDS (uniquely determined by successors) if it satisfies the following property: for any $v_i \in V$

$$v_i \in S \implies v_j \in S \text{ for all } j \prec i \text{ with } v_j \sim v_i \tag{1}$$

For our purpose, we will consider UDS subsets of $[n]$ in the graph $G(g_k)$ associated to $\{g_k\}$. The ordering of $[n]$ inherits the natural ordering on $\mathbb{N}$. The following proposition will be needed in our computations later.

Proposition 5.7. *For any $1 \le l \le n$ and any nonempty $I \in \mathcal{I}_n^l$, the sequence $\{g_k e_I\} \subset \wedge^l \mathbb{R}^n$ is bounded if and only if I is UDS in the vertex set $[n]$ of $G(g_k)$. If this case happens, then we have $g_k e_I = e_I$.*

Proof. Write $I = \{i_1 < i_2 < \cdots < i_l\}$. Suppose that $\{g_k e_I\}$ is bounded. We show that I is UDS in $[n]$. If not, let i_0 be the minimum in $I = \{i_1, \ldots, i_l\}$ such that the property (1) in Definition 5.6 does not hold for i_0. Then there is $j_0 < i_0$ with $j_0 \sim i_0$ but $j_0 \notin I$. By the minimality of i_0, for any $i \in I = \{i_1, i_2, \ldots, i_l\}$ with $j_0 < i < i_0$, we have $j_0 \not\sim i$; otherwise $j_0 \in I$. This implies $u_{j_0,i}(k) = 0$ for all $i \in \{i_1, i_2, \ldots, i_l\}$ with $i < i_0$. Note that $u_{j_0,i_0}(k) \to \infty$ as $k \to 0$ by our assumption on the entries of g_k.

Now we compute $g_k e_I$. In particular, by expanding $g_k e_I$ in terms of the standard basis $\{e_J : J \in \mathcal{I}_n^l\}$ in $\wedge^l \mathbb{R}^n$, we are interested in the coefficient in the e_{J_0}-coordinate, where $J_0 = \{i \in I : i \neq i_0\} \cup \{j_0\}$. As $u_{j_0,i}(k) = 0$ for all $i \in \{i_1, i_2, \ldots, i_l\}$ with $i < i_0$, one can easily compute

$$g_k e_I = u_{j_0,i_0}(k)(\wedge_{i\in I, i<i_0} e_i) \wedge e_{j_0} \wedge (\wedge_{i\in I, i>i_0} e_i) + \sum_{J\neq J_0} c_J e_J$$

for some $c_J \in \mathbb{R}$ $(J \neq J_0)$. The divergence of $u_{j_0,i_0}(k)$ then contradicts the boundedness of $g_k e_I$. This proves that I is UDS.

Conversely, suppose that I is a UDS subset in $[n]$. In this case, we will show inductively that for any $1 \leq j \leq l$

$$g_k(e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_j}) = e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_j}$$

and hence obtain that $g_k e_I = g_k(e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_l})$ remains fixed. For $j = 1$, since $\{i_1, \ldots, i_l\}$ is UDS, this implies that $u_{i,i_1} = 0$ for all $i < i_1$ and $g_k e_{i_1} = e_{i_1}$. Now assume that the formula holds for j. For $j+1$, we know that

$$g_k e_{i_{j+1}} = e_{i_{j+1}} + \sum_{i\in\{i_1,\ldots,i_j\}} u_{i,i_{j+1}} e_i$$

and hence

$$\begin{aligned} & g_k(e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_j} \wedge e_{i_{j+1}}) = e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_j} \wedge (g_k e_{i_{j+1}}) \\ = \ & e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_j} \wedge e_{i_{j+1}}. \end{aligned}$$

This concludes the proof of the proposition. □

Finally, we will show our second result in this section, which will be crucial in our study of convex polytopes.

Lemma 5.8. *Let $G(V, E)$ be a connected graph, where $V = \{v_1, v_2, \ldots, v_n\}$ is an ordered set with the ordering $\prec$. Then we can assign values $x_1, x_2, \ldots, x_n$ to the vertices $v_1, v_2, \ldots, v_n$ such that*

(1) $\sum_{v_i\in V} x_i = 0$
(2) For any proper UDS subset $S \subset V$, $\sum_{v_i\in S} x_i > 0$.

Proof. We use induction on the number of vertices in $G(V, E)$. There is nothing to prove for $n = 1$. Now suppose we have $n + 1$ vertices. Assume without loss of generality that v_1 is the smallest according to the ordering $\prec$ on V. We remove the vertex v_1 and all the edges adjacent to v_1 from the graph G. This yields a new graph G' which has m connected components $G'_1 = (V'_1, E'_1), \dots, G'_m = (V'_m, E'_m)$ for some $m \in \mathbb{N}$. Since $|V'_j| \leq n$ $(1 \leq j \leq m)$ and V'_j inherits the ordering from V, we can apply the induction hypothesis on each $G'_j = (V'_j, E'_i)$. In particular, we obtain a vector $(x'_2, \dots, x'_{n+1}) \in \mathbb{R}^n$ such that the value assignment

$$v_i \mapsto x'_i, \quad 2 \leq i \leq n+1$$

satisfies conditions (1) and (2) for each of the graphs G'_j $(1 \leq j \leq m)$.

Now we pick a sufficiently small positive number $\epsilon > 0$ such that the new value assignment $x_i = x'_i - \epsilon$ $(2 \leq i \leq n+1)$ still satisfies condition (2) for each $G'_j = (V'_j, E'_j)$, and let $x_1 = n\epsilon$. We show that this value assignment

$$v_i \mapsto x_i, \quad 1 \leq i \leq n+1$$

meets our requirements for $G(V, E)$. The sum of x_i is zero by induction hypothesis. For a proper UDS subset $S \subset V$, if $v_1 \notin S$, then

$$S = \bigcup_{j=1}^{m} S'_j$$

where S'_j is a subset in $G'_j = (V'_j, E'_j)$, and either S'_j is a proper UDS subset in $G'_j = (V'_j, E'_j)$ or $S'_j = V'_j$. Since $v_1 \notin S$, by the connectedness of $G(V, E)$ and the UDS property of S, there is some j with $S'_j \neq V'_j$ and hence by taking ϵ sufficiently small,

$$\sum_{v_i \in S} x_i = \sum_{j=1}^{m} \sum_{v_i \in S'_j} x_i > 0.$$

If $S = \{v_1\}$, then condition (2) holds automatically. If $v_1 \in S$ and $S \neq \{v_1\}$, then

$$S \setminus \{v_1\} = \bigcup_{j=1}^{m} S'_j$$

where S'_j is a subset in $G'_j = (V'_j, E'_j)$, and either S'_j is a proper UDS subset in $G'_j = (V'_j, E'_j)$ or $S'_j = V'_j$. Since S is proper in V, there is some j with $S'_j \neq V'_j$ and hence we have

$$\sum_{v_i \in S} x_i = \sum_{j=1}^{m} \sum_{v_i \in S'_j} x_i + x_1 > (-n\epsilon) + n\epsilon = 0.$$

This completes the proof of the lemma. $\square$

6. Revisit convex polytopes

In this section, we will study the convex polytopes $\Omega_{g_k,\delta}$ where $\{g_k\}$ is a sequence in $\mathrm{SL}(n, \mathbb{R})$ satisfying the assumptions at the end of section 4. Our aim in this section is Proposition 6.3, which shows a crucial property of $\Omega_{g_k,\delta}$. This property will play an important role in various places of this paper.

In the proof of Theorem 2.4, the case of $\mathcal{A}(A, g_k) = \{0\}$ plays a central role, and other cases can be deduced from this case. We remark here that $\mathcal{A}(A, g_k) = \{0\}$ if and only if the limit points of $\{\mathrm{Ad}(g_k)\,\mathrm{Lie}(A)\}$ in the Grassmanian manifold of $\mathfrak{sl}(n, \mathbb{R})$ are all nilpotent subalgebras. So starting from this section to section 9, we will make additional assumptions on $\{g_k\}$, namely, that $\mathcal{A}(A, g_k) = \{0\}$, and by passing to a subsequence, $\mathrm{Ad}(g_k)\,\mathrm{Lie}(A)$ converges to a subalgebra consisting of nilpotent elements in the Grassmanian manifold of $\mathfrak{sl}(n, \mathbb{R})$. We will write $\lim_{k\to\infty} \mathrm{Ad}(g_k)\,\mathrm{Lie}(A)$ for the limit nilpotent subalgebra and $\lim_{k\to\infty} \mathrm{Ad}(g_k)A$ for the corresponding limit unipotent subgroup.

Following Definition 5.4, we write $G(g_k) = (V, E)$ for the graph associated to $\{g_k\}$.

Lemma 6.1. *For any $0 < \delta < 1$, the region*

$$\{\mathbf{t} \in \mathrm{Lie}(A) : \omega_I(\mathbf{t}) \geq \ln \delta, \forall \textit{ nonempty proper UDS } I \in \mathcal{I}_n\}$$

is a convex subset in $\mathrm{Lie}(A)$ which contains an unbound open cone.

Proof. It suffices to prove the lemma for the region

$$\{\mathbf{t} \in \mathrm{Lie}(A) : \omega_I(\mathbf{t}) \geq 0, \forall \text{ nonempty proper UDS } I \in \mathcal{I}_n\}.$$

By our assumption on $\{g_k\}$ and Proposition 5.5, the graph $G(g_k)$ associated to $\{g_k\}$ is connected. Now by applying Lemma 5.8 with the graph $G(g_k)$, one can find $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathrm{Lie}(A)$ such that

$$\mathbf{x} \in \{\mathbf{t} \in \mathrm{Lie}(A) : \omega_I(\mathbf{t}) > 0, \forall \text{ nonempty proper UDS } I \in \mathcal{I}_n\}.$$

Then by linearity, for any $\lambda > 0$

$$\lambda\mathbf{x} \in \{\mathbf{t} \in \mathrm{Lie}(A) : \omega_I(\mathbf{t}) > 0, \forall \text{ nonempty proper UDS } I \in \mathcal{I}_n\}.$$

This implies that there exists an unbounded open cone around the axis $\{\lambda\mathbf{x}, \lambda > 0\}$, which is contained in

$$\{\mathbf{t} \in \mathrm{Lie}(A) : \omega_I(\mathbf{t}) \geq 0, \forall \text{ nonempty proper UDS } I \in \mathcal{I}_n\}.$$

This completes the proof of the lemma. $\square$

Lemma 6.2. *For every $k \in \mathbb{N}$ the region $\Omega_{g_k,\delta}$ contains a ball B_k of radius r_k such that $r_k \to \infty$ as $k \to \infty$.*

Proof. By definition, we know that

$$\Omega_{g_k,\delta} = \bigcap_{I \in \mathcal{I}_n} \left\{\mathbf{t} \in \mathrm{Lie}(A) : \omega_I(\mathbf{t}) \geq \ln \delta - \ln \|g_k e_I\|\right\}.$$

Note that the origin belongs to $\Omega_{g_k,\delta}$ by Proposition 5.7 for sufficiently large $k > 0$. Now we can write

$$\Omega_{g_k,\delta} = \bigcap_{I \text{ UDS}} \{\omega_I(\mathbf{t}) \geq \ln(\delta/\|g_k e_I\|)\} \cap \bigcap_{I \text{ not UDS}} \{\omega_I(\mathbf{t}) \geq \ln(\delta/\|g_k e_I\|)\}$$

$$= \bigcap_{I \text{ UDS}} \{\omega_I(\mathbf{t}) \geq \ln \delta\} \cap \bigcap_{I \text{ not UDS}} \{\omega_I(\mathbf{t}) \geq \ln(\delta/\|g_k e_I\|)\}$$

as $g_k e_I = e_I$ for any UDS I by Proposition 5.7. For I not UDS, we have $g_k e_I \to \infty$.

Since $g_k e_I \to \infty$ for any I not UDS, the region

$$\bigcap_{I \text{ not UDS}} \{\omega_I(\mathbf{t}) \geq \ln(\delta/\|g_k e_I\|)\}$$

contains a large ball S_k around the origin for sufficiently large k. By Lemma 6.1, the region

$$\bigcap_{I \text{ UDS}} \{\omega_I(\mathbf{t}) \geq \ln \delta\}$$

contains an unbounded cone C (which does not depend on k) with cusp at the origin. This implies that

$$\Omega_{g_k,\delta} \supset S_k \cap C$$

and $\Omega_{g_k,\delta}$ contains a large ball B_k of radius r_k with $r_k \to \infty$ as $k \to \infty$. □

Proposition 6.3. *For any $0 < \delta < 1$, we have*

$$\lim_{k\to\infty} \frac{\mathrm{Vol}(\partial\Omega_{g_k,\delta})}{\mathrm{Vol}(\Omega_{g_k,\delta})} = 0.$$

Proof. The proposition follows from Lemma 4.4 and Lemma 6.2. □

Actually, we will apply the following variant of Proposition 6.3 in our arguments later.

Corollary 6.4. *Let $0 < \delta_1 < \delta_2 < 1$. Then*

$$\lim_{k\to\infty} \frac{\mathrm{Vol}(\Omega_{g_k,\delta_2})}{\mathrm{Vol}(\Omega_{g_k,\delta_1})} = 1.$$

Proof. By definition, we know that $\Omega_{g_k,\delta_2} \subset \Omega_{g_k,\delta_1}$. Let $\{f_i\}$ be the collection of the facets of Ω_{g_k,δ_1}, and denote by P_i the hyperplane determined by f_i. For each f_i, let B_i be the unique cylinder with the following properties:

(1) the base of B_i is f_i, and the height of B_i is equal to $\delta_2 - \delta_1$.
(2) B_i and Ω_{g_k,δ_1} lie in the same half-space determined by P_i.

Then one has

$$\Omega_{g_k,\delta_1} = \bigcup_i B_i \cup \Omega_{g_k,\delta_2}$$

and

$$\mathrm{Vol}(\Omega_{g_k,\delta_1}) \leq \sum_i \mathrm{Vol}(B_i) + \mathrm{Vol}(\Omega_{g_k,\delta_2}) = (\delta_2 - \delta_1)\,\mathrm{Vol}(\partial\Omega_{g_k,\delta_1}) + \mathrm{Vol}(\Omega_{g_k,\delta_2})$$

Now the corollary follows from Proposition 6.3. □

Now for each $k \in \mathbb{N}$ we choose the representative

$$\frac{1}{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})}(g_k)_*\mu_{Ax}$$

in $[(g_k)_*\mu_{Ax}]$ and we will show in the following section that these representatives converge to a locally finite measure ν. From now on, we will fix a $\delta > 0$ for $\Omega_{g_k,\delta}$ unless otherwise specified. We will also denote by

$$\mu_{Ax}|_{\Omega_{g_k,\delta}}$$

the restriction of μ_{Ax} on $\exp(\Omega_{g_k,\delta})x$ ($x = e\,\mathrm{SL}(n,\mathbb{Z})$).

7. Nondivergence

In this section, as mentioned above, we will study the nondivergence of the sequence

$$\frac{1}{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})}(g_k)_*\mu_{Ax}.$$

The study relies on a growth property of a special class of functions studied by Eskin, Mozes and Shah [EMS97], and a non-divergence theorem proved by Kleinbock and Margulis [KM98, Kle10]. As a corollary we will deduce that these measures actually converge to a probability measure, which is invariant under a unipotent subgroup. This is where Ratner's theorem will come into play in section 9 and help us prove the measure rigidity. Our ultimate goal in this section is to prove Proposition 7.7.

First, we need the following definition of a class of functions, which is introduced in [EMS97].

Definition 7.1 ([EMS97] Definition 2.1)**.** Let $d \in \mathbb{N}$ and $\lambda > 0$ be given. Define by $E(d, \lambda)$ the set of functions $f : \mathbb{R} \to \mathbb{C}$ of the form

$$f(t) = \sum_{i=1}^{d}\sum_{l=0}^{d-1} a_{i,l}t^l e^{\lambda_i t} \quad (\forall t \in \mathbb{R})$$

where $a_{i,l} \in \mathbb{C}$ and $\lambda_i \in \mathbb{C}$ with $|\lambda_i| \leq \lambda$.

The following proposition describes the growth property of functions in $E(d, \lambda)$. We will denote by $m_{\mathbb{R}^k}$ the Lebesgue measure on $\mathbb{R}^k$.

Proposition 7.2 ([EMS97] Cor.2.10)**.** *For any $d \in \mathbb{N}$ and $\lambda > 0$, there exists a constant $\delta_0 = \delta_0(d, \lambda)$ satisfying the following: given $\epsilon > 0$, there exists $M > 0$ such that for any $f \in E(d, \lambda)$ and any interval Ξ of length at most δ_0*

$$m_{\mathbb{R}}(\{t \in \Xi : |f(t)| < (1/M)\sup_{t\in\Xi}|f(t)|\}) \leq \epsilon m_{\mathbb{R}}(\Xi). \tag{2}$$

The following theorem is essentially proved in [Kle10] and [KM98].

Theorem 7.3 (Cf.[Kle10] Theorem 3.4, [KM98])**.** *Let $d \in \mathbb{N}$ and $\Lambda > 0$. Let $\delta_0 = \delta_0(d, \Lambda)$ be as in Proposition 7.2. Suppose an interval $\Xi \subset \mathbb{R}$ of length at most δ_0, $0 < \rho < 1$ and a continuous map $h : \Xi \to \mathrm{SL}(n, \mathbb{R})$ are given. Assume that for any discrete subgroup Δ in $\mathbb{Z}^n$ we have*

(1) the function $x \to \|h(x)\Delta\|$ on Ξ belongs to $E(d, \lambda)$ and

(2) $\sup_{x\in\Xi} \|h(x)\Delta\| \geq \rho$.

Then for any $\epsilon < \rho$, there exists a constant $\delta(\epsilon) > 0$ depending only on $E(d, \lambda)$ such that

$$m_{\mathbb{R}}(\{x \in \Xi : h(x)\mathbb{Z}^n \cap B_{\delta(\epsilon)} \neq \{0\}\}) \leq \epsilon m_{\mathbb{R}}(\Xi).$$

Proof. The proof is the same as in [KM98], but the inequality (2) is used instead of the analogue property of (C, α)-good. □

Lemma 7.4. *Let E be a normed vector space, and α_i ($1 \leq i \leq m$) different linear functionals on E. Then for any $r > 0$, we can find m vectors $x_1, x_2, \ldots, x_m \in B_r(0)$ such that*

$$\det\left(\left(e^{\alpha_i(x_j)}\right)_{1\leq i\leq m, 1\leq j\leq m}\right) \neq 0.$$

Here $B_r(0)$ is the ball of radius $r > 0$ in E.

Proof. We can find a line L through the origin such that $\alpha_i|L$ are different functionals defined on L. This could be achieved by picking a line which avoids all the kernels of $\alpha_i - \alpha_j$. Hence it suffices to prove the lemma for $\dim E = 1$.

Let $E = \mathbb{R}$ and $\alpha_i(x) = \lambda_i x$ for different λ_i's. We will show inductively that for any $r > 0$ there exist $x_1, x_2, \ldots, x_m \in (-r, r)$ such that

$$\det\left(\left(e^{\lambda_i x_j}\right)_{1\leq i\leq m, 1\leq j\leq m}\right) \neq 0.$$

It is easy to verify for $m = 1$. Now for $m + 1$ different λ_i's, we compute

$$\begin{aligned} &\det\left(\left(e^{\lambda_i x_j}\right)_{1\leq i\leq m+1, 1\leq j\leq m+1}\right) \\ = \quad & e^{\lambda_1 x_{m+1}} A_1 + e^{\lambda_2 x_{m+1}} A_2 + \cdots + e^{\lambda_{m+1} x_{m+1}} A_{m+1} \end{aligned}$$

where $A_{m+1} = \det\left(\left(e^{\lambda_i x_j}\right)_{1\leq i\leq m, 1\leq j\leq m}\right)$. By induction hypothesis, we can find $x_1, x_2, \ldots, x_m \in (-r, r)$ such that $A_{m+1} \neq 0$. By the fact that $e^{\lambda_i x}$ are linearly independent functions and with this choice of $x_1, x_2, \ldots, x_m$, the function $\det\left(\left(e^{\lambda_i x_j}\right)_{1\leq i\leq m+1, 1\leq j\leq m+1}\right)$ is a nonzero analytic function in x_{m+1}. Since zeros of any analytic function are isolated, this implies that there exists a $x_{m+1} \in (-r, r)$ such that $\det\left(\left(e^{\lambda_i x_j}\right)_{1\leq i\leq m+1, 1\leq j\leq m+1}\right) \neq 0$. □

The following proposition describes the supremum of a special function. We will need this proposition to verify the assumption (2) in Theorem 7.3.

Proposition 7.5. *Let E and V be normed vector spaces and $v_i \in V$ $(1 \leq i \leq m)$. Let f be a map from E to V defined by*

$$f(x) = \sum_{i=1}^{m} e^{\alpha_i(x)} v_i$$

where the α_i's $(1 \leq i \leq m)$ are different linear functionals on E. Suppose that on an open ball $R \subset E$ of radius $r > 0$ we have

$$e^{\alpha_i(x)} \|v_i\| \geq M \quad (\forall x \in R)$$

for some $M > 0$. Then there exists a constant $c > 0$ which only depends on the α_i's and r such that

$$\sup_{x \in R} \|f(x)\| \geq cM.$$

Proof. Let x_0 be the center of R. By Lemma 7.4, we can find $y_i \in B_r(0)$ such that $\det(e^{\alpha_i(y_j)}) \neq 0$. We fix this choice of y_i's which only depends on α_i's and r. Let $x_i = x_0 + y_i \in R$. We have

$$\left(e^{\alpha_i(y_j)}\right) \left(e^{\alpha_i(x_0)} v_i\right) = (f(x_i))$$

$$\left(e^{\alpha_i(x_0)} v_i\right) = \left(e^{\alpha_i(y_j)}\right)^{-1} (f(x_i)) \,.$$

Let $C = \left\| \left(e^{\alpha_i(y_j)}\right)^{-1} \right\|$. Since $e^{\alpha_i(x)} \|v_i\| \geq M$, this implies that one of $\|f(x_i)\|$ is greater than or equal to $M/(mC)$, and hence so is $\sup_{x \in R} \|f(x)\|$. $\square$

With the help of Theorem 7.3 and Proposition 7.5, we can now study the nondivergence of the sequence $\frac{1}{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})}(g_k)_* \mu_{Ax}$. In what follows, we will denote by

$$\mathcal{K}_r := \{g\Gamma \in G/\Gamma : \text{ every nonzero vector in } g\mathbb{Z}^n \text{ has norm} \geq r\}.$$

By Mahler's compactness criterion, this is a compact subset in G/Γ. The following proposition is crucial in our proof of Proposition 7.7.

Proposition 7.6. *For any $\epsilon > 0$, there exists a constant $\delta(\epsilon) > 0$ such that*

$$m_{\mathbb{R}^{n-1}}(\{\mathbf{t} \in \Omega_{g_k,\delta} : g_k \exp(\mathbf{t})\mathbb{Z}^n \notin \mathcal{K}_{\delta(\epsilon)}\}) \leq \epsilon m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta}).$$

Proof. Fix a vector of norm one $\vec{v} \in \mathrm{Lie}(A)$ such that the values

$$\{\omega_I(\vec{v}) : I \in \mathcal{I}_n\}$$

are all different . Let $d \in \mathbb{N}$ and $\lambda > 0$ such that for any $x_0 \in \mathrm{Lie}(A)$, any $l \in \mathbb{N}$ and any $w \in \wedge^l \mathbb{R}^n$, the function

$$\|g_k \exp(x_0 + t\vec{v}) \cdot w\|, \quad t \in \mathbb{R}$$

belongs to $E(d, \lambda)$ as defined in Definition 7.1. Here we denote by $\|\cdot\|$ the standard norm on $\wedge^l \mathbb{R}^n$, and will write δ_0 the constant $\delta_0(d, \lambda)$ defined in Proposition 7.2.

We can cut the region $\Omega_{g_k,\delta}$ into countably many disjoint small boxes of diameter at most δ_0 such that each box has one side parallel to $\vec{v}$. In other words, each box B is of the form

$$B = \{x_0 + t\vec{v} : x_0 \in S,\ t \in \Xi\}$$

where S is the base of B perpendicular to $\vec{v}$, and $\Xi = [0, a]$ is an interval for some $0 < a \leq \delta_0$. In order to prove the proposition, it suffices to show that for each such box B of diameter at most δ_0, we have

$$m_{\mathbb{R}^{n-1}}(\{\mathbf{t} \in B : g_k \exp(\mathbf{t})\mathbb{Z}^n \notin \mathcal{K}_{\delta(\epsilon)}\}) \leq \epsilon m_{\mathbb{R}^{n-1}}(B).$$

We will apply Theorem 7.3. Let Δ be a discrete subgroup of rank l in $\mathbb{Z}^n$. The point in $\wedge^l\mathbb{R}$ corresponding to Δ can be written as

$$\sum_{I \in \mathcal{I}_n^l} a_I e_I$$

where $a_I \in \mathbb{Z}$. We construct a map from B to $\wedge^l\mathbb{R}$ by

$$f(\mathbf{t})|_B \quad = \quad \wedge^l \operatorname{Ad}(g_k \exp \mathbf{t})\Delta = \sum_{I \in \mathcal{I}_n, I \in \mathcal{I}_n^l} a_I e^{\omega_I(\mathbf{t})} g_k e_I.$$

Since $B \subset \Omega_{g_k,\delta}$, by our construction of $\Omega_{g_k,\delta}$, we have

$$\|e^{\omega_I(\mathbf{t})} g_k e_I\| \geq \delta, \quad \forall \mathbf{t} \in B$$

for any $I \in \mathcal{I}_n$ with $I \in \mathcal{I}_n^l$. For each $x_0 \in S$, we consider the map

$$t \mapsto f(x_0 + t\vec{v})$$

from $\Xi = [0, a]$ to $\wedge^l\mathbb{R}$. By Proposition 7.5, we have

$$\sup_{t \in \Xi} \|f(x_0 + t\vec{v})\| \geq c\delta.$$

Note that by Proposition 7.5, this inequality holds with a uniform constant $c > 0$ for any $\Delta \subset \mathbb{Z}^n$. Also by definition, the map $\|f(x_0 + t\vec{v})\|$ is a function in $E(d, \lambda)$. Hence we can apply Theorem 7.3 and obtain that

$$m_{\mathbb{R}}(\{t \in \Xi : g_k \exp(x_0 + t\vec{v})\mathbb{Z}^n \notin \mathcal{K}_{\delta(\epsilon)}\}) \leq \epsilon m_{\mathbb{R}}(\Xi)$$

for some constant $\delta(\epsilon) > 0$ and for any $x_0 \in S$. Now by integrating the inequality above over the region $x_0 \in S$, we have

$$m_{\mathbb{R}^{n-1}}(\{\mathbf{t} \in B : g_k \exp(\mathbf{t})\mathbb{Z}^n \notin \mathcal{K}_{\delta(\epsilon)}\}) \leq \epsilon m_{\mathbb{R}^{n-1}}(B)$$

and then the proposition follows. □

Finally, we can prove our central result in this section.

Proposition 7.7. *The sequence* $\frac{1}{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})}(g_k)_*(\mu_{Ax}|_{\Omega_{g_k,\delta}})$ *has a subsequence converging to a probability measure* ν*. Furthermore, we have*

$$\frac{1}{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})}(g_k)_*\mu_{Ax} \to \nu$$

and hence the sequence $[(g_k)_*\mu_{Ax}]$ *converges to* $[\nu]$*. Here the probability measure* ν *is invariant under the action of the unipotent subgroup* $\lim_{n\to\infty} \operatorname{Ad}(g_k)A$*.*

Proof. The first claim follows from Proposition 7.6. For the second claim, we will show that

$$\frac{1}{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})}(g_k)_*\mu_{Ax} - \frac{1}{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})}(g_k)_*(\mu_{Ax}|_{\Omega_{g_k,\delta}}) \to 0.$$

Let $f \in C_c(X)$. Then by definition, there exists a small number $\delta' < \delta$ such that

$$\int f(g_k A) d\mu_{Ax} = \int_{\Omega_{g_k,\delta'}} f(g_k A) d\mu_{Ax}.$$

By Corollary 6.4, we have

$$\begin{aligned}
& \left| \frac{1}{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})} \int f(g_k A) d\mu_{Ax} - \frac{1}{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})} \int_{\Omega_{g_k,\delta}} f(g_k A) d\mu_{Ax} \right| \\
= & \left| \frac{1}{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})} \int_{\Omega_{g_k,\delta'}} f(g_k A) d\mu_{Ax} - \frac{1}{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})} \int_{\Omega_{g_k,\delta}} f(g_k A) d\mu_{Ax} \right| \\
= & \left| \frac{1}{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})} \int_{\Omega_{g_k,\delta'} \setminus \Omega_{g_k,\delta}} f(g_k A) d\mu_{Ax} \right| \\
\leq & \|f\|_\infty \frac{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta'} \setminus \Omega_{g_k,\delta})}{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})} \to 0.
\end{aligned}$$

Since $g_k \mu_{Ax}$ is invariant under the action of $\mathrm{Ad}(g_k)A$, the probability measure ν is invariant under the action of $\lim_{n\to\infty} \mathrm{Ad}(g_k)A$, which is a unipotent subgroup by our assumptions on $\{g_k\}$. □

8. Nondivergence in terms of adjoint representations

In this section, we rewrite section 7 in terms of adjoint representations. The advantage of doing so is that we can then apply Ratner's theorem for unipotent actions on homogeneous spaces.

Let $\mathrm{Ad} : \mathrm{SL}(n, \mathbb{R}) \to \mathrm{SL}(\mathfrak{g})$ be the adjoint representation of $\mathrm{SL}(n, \mathbb{R})$. The Lie algebra $\mathfrak{sl}(n, \mathbb{R})$ has a $\mathbb{Q}$-basis $\{E_{ij} : 1 \leq i \neq j \leq n\} \cup \{E_{ii} : 1 \leq i \leq n-1\}$, where E_{ij} ($i \neq j$) is the matrix with only nonzero entry 1 in the ith row and the jth column, and E_{ii} ($1 \leq i \leq n-1$) is the diagonal matrix with 1 in the (i,i)-entry and -1 in the $(i+1, i+1)$-entry. We will also consider the representations $\wedge^l \mathrm{Ad} : \mathrm{SL}(n, \mathbb{R}) \to \mathrm{SL}(\wedge^l \mathfrak{g})$ for $1 \leq l \leq \dim \mathfrak{g}$. A $\mathbb{Q}$-basis of $\wedge^l \mathfrak{g}$ is then $\{\wedge^l E_{ij} : 1 \leq i, j \leq n\}$.

Let $1 \leq l \leq \dim \mathfrak{g}$. For $\wedge^l \mathfrak{g}$, its decomposition with respect to the action of $\wedge^l \mathrm{Ad}\, A$ is given by

$$\wedge^l \mathfrak{g} = \sum_\chi \mathfrak{g}_\chi$$

where each χ is a linear functional on $\mathrm{Lie}(A)$ such that for any $Y \in \mathrm{Lie}(A)$ we have

$$\wedge^l \mathrm{Ad}(\exp(Y))v = \exp(\chi(Y))v, \quad \forall v \in \mathfrak{g}_\chi.$$

We know that each $\mathfrak{g}_\chi$ has a $\mathbb{Q}$-basis from $\{\wedge^l E_{ij} : 1 \le i, j \le n\}$. We denote by $\mathfrak{g}_\chi(\mathbb{Z})$ the subset of integer vectors with respect to this basis, and by $\mathcal{W}(\mathfrak{g})$ the collection of such χ's for all $1 \le l \le \dim \mathfrak{g}$.

Let $g \in \mathrm{SL}(n, \mathbb{R})$. We will define for $gA\Gamma$ a convex polytope in $\mathrm{Lie}(A)$ in terms of adjoint representations, which is, in a way, similar to the region $\Omega_{g,\delta}$ in section 4. Let $v \in \mathfrak{g}_\chi(\mathbb{Z}) \setminus \{0\}$. Then for $\mathbf{t} \in \mathrm{Lie}(A)$, the vector

$$\wedge^l \mathrm{Ad}(g \exp(\mathbf{t}))v = e^{\chi(\mathbf{t})} \wedge^l \mathrm{Ad}(g)v \notin B_\delta$$

if and only if

$$\chi(\mathbf{t}) \ge \ln \delta - \ln \| \wedge^l \mathrm{Ad}(g)v\|.$$

Here we denote by B_δ the ball of radius $\delta > 0$ around 0 with the norm $\| \cdot \|$ induced by a norm on $\mathfrak{g}$. In this way, we give the following

Definition 8.1. For any $g \in \mathrm{SL}(n, \mathbb{R})$ and $\delta > 0$, define a region $R_{g,\delta}$ in the Lie algebra $\mathrm{Lie}(A)$ by

$$R_{g,\delta} = \{\mathbf{t} \in \mathrm{Lie}(A) : \chi(\mathbf{t}) \ge \ln \delta - \ln \|\wedge^l \mathrm{Ad}(g)v\|, \forall v \in \mathfrak{g}_\chi(\mathbb{Z})\setminus\{0\} \text{ and } \forall \chi \in \mathcal{W}(\mathfrak{g})\}.$$

We list some properties about the convex polytopes $R_{g_k,\delta}$ for $\{g_k\}$, which are parallel to those in section 6. The proof of the following proposition is similar to that in Lemma 4.3.

Proposition 8.2. *The region $R_{g,\delta}$ is a bounded convex subset in $\mathrm{Lie}(A)$ for any $g \in \mathrm{SL}(n, \mathbb{R})$.*

Proposition 8.3. *Let $\delta > 0$. We have*

(1) For any $\epsilon > 0$ there exists $\delta(\epsilon) > 0$ such that

$$m_{\mathbb{R}^{n-1}}(R_{g_k,\delta(\epsilon)} \cap \Omega_{g_k,\delta}) \ge (1-\epsilon) m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta}).$$

(2) For each $k \in \mathbb{N}$, $R_{g_k,\delta}$ contains a ball of radius r_k such that $r_k \to \infty$ as $k \to \infty$.

Proof. For any $\epsilon > 0$, let $\delta(\epsilon)$ be as in Proposition 7.6. By applying Mahler's compactness criterion on $\mathrm{SL}(\wedge^l \mathfrak{g})$ $(1 \le l \le \dim \mathfrak{g})$, we can find a $\delta'(\epsilon) > 0$ such that

$$\{\mathbf{t} \in \Omega_{g_k,\delta} : g_k \exp(\mathbf{t})\mathbb{Z}^n \in \mathcal{K}_{\delta(\epsilon)}\} \subset R_{g_k,\delta'(\epsilon)} \cap \Omega_{g_k,\delta}.$$

Now the first part of the proposition follows from Proposition 7.6.

By Lemma 4.5 and Lemma 6.2, for each $k \in \mathbb{N}$, the convex polytope $R_{g_k,\delta(\epsilon)} \cap \Omega_{g_k,\delta}$ contains a ball of radius r_k and $r_k \to \infty$ as $k \to \infty$, and hence so does $R_{g_k,\delta}$ for any $\delta > 0$. □

Proposition 8.4. *For any $\delta > 0$, we have*

$$\lim_{k\to\infty} \frac{\mathrm{Vol}(\partial R_{g_k,\delta})}{\mathrm{Vol}(R_{g_k,\delta})} = 0.$$

Proof. The proof is identical to that in Proposition 6.3. □

Proposition 8.5. *Let $\delta > 0$. For any $\epsilon > 0$, there exists a constant $\delta(\epsilon) > 0$ such that*

$$m_{\mathbb{R}^{n-1}}(\{\mathbf{t} \in R_{g_k,\delta} : g_k \exp(\mathbf{t})\mathbb{Z}^n \notin K_{\delta(\epsilon)}\}) \leq \epsilon m_{\mathbb{R}^{n-1}}(R_{g_k,\delta}).$$

Proof. It is similar to Proposition 7.6, except that we replace the linear functionals $\sum_{j=1}^{l} t_{i_j}$ by χ's in $\mathcal{W}(\mathfrak{g})$. □

Proposition 8.6. *Let $\delta > 0$. The sequence $\frac{1}{m_{\mathbb{R}^{n-1}}(R_{g_k,\delta})} g_k(\mu_{Ax}|_{R_{g_k,\delta}})$ has a subsequence converging to a probability measure ν. We also have*

$$\frac{1}{m_{\mathbb{R}^{n-1}}(R_{g_k,\delta})} g_k \mu_{Ax} \to \nu$$

and hence the sequence $[g_k \mu_{Ax}]$ converges to $[\nu]$. Furthermore, the probability measure ν is invariant under the action of the unipotent subgroup $\lim_{n\to\infty} \mathrm{Ad}(g_k)A$.

Proof. It is identical to Proposition 7.7 with $\Omega_{g_k,\delta}$ replaced by $R_{g_k,\delta}$. □

The following is an immediate corollary of Proposition 3.3, Proposition 7.7 and Proposition 8.6.

Corollary 8.7. *For any $\delta > 0$, we have*

$$\lim_{k\to\infty} \frac{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})}{m_{\mathbb{R}^{n-1}}(R_{g_k,\delta})} = 1.$$

A single convex polytope $R_{g_k,\delta}$ for each g_k will suffice in our arguments below. So we will fix a $\delta > 0$ for $R_{g_k,\delta}$ in the rest of the paper, unless otherwise specified.

9. Ratner's theorem and linearization

Thanks to Proposition 8.6, we can apply measure classification theorem for unipotent actions on homogeneous spaces. It was first conjectured by Raghunathan and Dani [Dan81], and later proved by Ratner in her seminal work [Rat90a, Rat90b, Rat91]. Here we will proceed by following the framework of [EMS96] and [MS95]. Readers may refer to [Sha91] and [DM93] for more details. This section is the final step of preparation for the proof of Theorem 2.4, and is devoted to proving Proposition 9.9.

9.1. **Prerequisites.** We start by recalling some well-known results, which will be needed later in this section. Let $\mathcal{H}$ be the collection of all closed connected subgroups H of G such that $H \cap \Gamma$ is a lattice in H and the subgroup generated by all the unipotent one-parameter subgroups of G contained in H acts ergodically on $H\Gamma/\Gamma$ with respect to the H-invariant probability measure.

Theorem 9.1 ([Rat91] Theorem 1.1, [DM93] Proposition 2.1)**.** *$\mathcal{H}$ is a countable collection.*

Let W be a subgroup of G which is generated by unipotent one-parameter subgroups of G contained in W. For $H \in \mathcal{H}$, define

$$N(H, W) = \{g \in G : W \subset gHg^{-1}\}$$

$$S(H, W) = \bigcup_{H' \in \mathcal{H}, H' \subset H, H' \neq H} N(H', W)$$

and

$$T_H(W) = \pi(N(H, W) \backslash S(H, W)) = \pi(N(H, W)) \backslash \pi(S(H, W))$$

where $\pi : G \to G/\Gamma$ is the natural projection. We have for any $H_1, H_2 \in \mathcal{H}$

$$T_{H_1}(W) \cap T_{H_2}(W) \neq \emptyset \iff T_{H_1}(W) = T_{H_2}(W)$$

Theorem 9.2 ([Rat91], [MS95] Theorem 2.2)**.** *Let μ be a W-invariant probability measure on X. For every $H \in \mathcal{H}$, let μ_H denote the restriction of μ on $T_H(W)$. Then one has the following:*

(1) For all Borel measurable subsets $F \subset X$,

$$\mu(F) = \sum_{H \in \mathcal{H}^*} \mu_H(F)$$

where $\mathcal{H}^* \subset \mathcal{H}$ is a countable set consisting of representatives from Γ-conjugacy classes in $\mathcal{H}$.

(2) Each μ_H is W-invariant. For any W-ergodic component ν of μ_H, there exists $g \in N(H, W)$ such that ν is the unique gHg^{-1}-invariant probability measure on the closed orbit $gH\Gamma/\Gamma$.

Now in our case, the subgroup W will be $\lim_{n\to\infty} \mathrm{Ad}(g_k)A$. By our assumptions, W is a unipotent subgroup of G. In the following, we will fix a subgroup $H \in \mathcal{H}$ ($H \neq G$). Let $\mathfrak{g}$ denote the Lie algebra of G and let $\mathfrak{h}$ denote the Lie subalgebra of H. For $d = \dim \mathfrak{h}$, put $V_H = \wedge^d \mathfrak{g}$, the d-th exterior power, and consider the linear G-action on V_H via the representation $\wedge^d \mathrm{Ad}$, the d-th exterior of the adjoint representation of G on $\mathfrak{g}$. Since H is a $\mathbb{Q}$-group, we can find an integral point $p_H \in \wedge^d \mathfrak{h} \setminus \{0\}$. We fix this p_H and let $\eta_H : G \to V_H$ be the map defined by

$$\eta_H(g) = g \cdot p_H = (\wedge^d \mathrm{Ad}\, g) p_H$$

for all $g \in G$. Note that

$$\eta_H^{-1}(p_H) = \{g \in N(H) : \det(\mathrm{Ad}\, g|_{\mathfrak{h}}) = 1\}$$

where $N(H)$ denotes the normalizer of H in G.

Put $\Gamma_H = N(H) \cap \Gamma$. Then for any $\chi \in \Gamma_H$, we have $\chi(H\Gamma/\Gamma) = H\Gamma/\Gamma$ and χ preserves the volume of $H\Gamma/\Gamma$. Therefore $|\det(\mathrm{Ad}\, \chi|_{\mathfrak{h}})| = 1$ and $\chi \cdot p_H = \pm p_H$.

In view of this we define $\overline{V}_H = V_H/\{1, -1\}$ if $\Gamma_H \cdot p_H = \{p_H, -p_H\}$ and define $\overline{V}_H = V_H$ if $\Gamma_H \cdot p_H = p_H$. Define the orbit map of G on $\overline{V}_H$

$$\overline{\eta}_H : G \to \overline{V}_H$$

by $\overline{\eta}_H(g) = g \cdot \overline{p}_H$ where $\overline{p}_H$ denotes the image of p_H in $\overline{V}_H$. We denote by L_H the Zariski closure of $\overline{\eta}_H(N(H,W))$ in $\overline{V}_H$.

Theorem 9.3 ([DM93] Theorem 3.4). *The orbit $\Gamma \cdot \overline{p}_H$ is discrete in $\overline{V}_H$.*

Proposition 9.4 ([DM93] Proposition 3.2). *Let L_H denote the Zariski closure of $\overline{\eta}_H(N(H,W))$ in $\overline{V}_H$. Then*

$$\overline{\eta}_H^{-1}(L_H) = N(H,W).$$

Proposition 9.5 ([MS95] Proposition 3.2). *Let D be a compact subset of L_H. Define*

$$S(D) = \{g \in \overline{\eta}_H^{-1}(D) : g\chi \in \overline{\eta}_H^{-1}(D) \text{ for some } \chi \in \Gamma \setminus \Gamma_H\}.$$

Then we have the following:

(1) $S(D) \subset S(H,W)$.
(2) $\pi(S(D))$ is closed in X.
(3) For any compact set $\mathcal{K} \subset X \setminus \pi(S(D))$, there exists a neighbourhood Φ of D in $\overline{V}_H$ such that every $y \in \pi(\overline{\eta}_H^{-1}(\Phi)) \cap \mathcal{K}$ has a unique representation in Φ; that is, the set $\overline{\eta}_H(\pi^{-1}(y)) \cap \Phi$ consists of a single element.

9.2. **Proof of Proposition 9.9.** Now we begin our journey towards Proposition 9.9. Let $\{f_1, f_2, \dots, f_m\}$ be a set of polynomials defining L_H in $\overline{V}_H$. In the rest of the section, we will fix a vector of norm one $\vec{v} \in \mathrm{Lie}(A)$ such that all the linear functionals $\chi \in \mathcal{W}(\mathfrak{g})$ are different on $\vec{v}$. Also, one can find $d \in \mathbb{N}$, and $\lambda > 0$ such that for any $x_0 \in \mathrm{Lie}(A)$, the functions of $t \in \mathbb{R}$

$$\|(g_k \exp(x_0 + t\vec{v}) \cdot w\|^2, \quad f_j(g_k \exp(x_0 + t\vec{v}) \cdot w), \quad 1 \le j \le m$$

belong to $E(d,\lambda)$ as defined in Definition 7.1. Here the norm $\|\cdot\|$ on $\overline{V}_H$ is induced by a norm on $\mathfrak{g}$. We will write δ_0 for the constant $\delta_0(d,\lambda)$ defined in Proposition 7.2.

Proposition 9.6 (Cf.[DM93] Proposition 4.2). *Let a compact set $C \subset L_H$ and $\epsilon > 0$ be given. Then there exists a compact set $D \subset L_H$ with $C \subset D$ such that for any neighborhood Φ of D in $\overline{V}_H$, there exists a neighborhood Ψ of C in $\overline{V}_H$ with the following property. For any $x_0 \in \mathrm{Lie}(A)$, $w \in \overline{V}_H$ and any interval $\Xi \subset [0,\delta_0]$, if $\{g_k \exp(x_0 + t\vec{v}) \cdot w : t \in \Xi\} \not\subset \Phi$, then we have*

$$\begin{aligned} & m_{\mathbb{R}}(\{t \in \Xi : g_k \exp(x_0 + t\vec{v}) \cdot w \in \Psi\}) \\ \le \quad & \epsilon m_{\mathbb{R}}(\{t \in \Xi : g_k \exp(x_0 + t\vec{v}) \cdot w \in \Phi\}). \end{aligned}$$

Proof. Let d and λ be defined as above. We choose a ball $B_0(r)$ of radius $r > 0$ centered at 0 in $\overline{V}_H$ such that the closure $\overline{C} \subset B_0(r)$. Now for a given $\epsilon > 0$, let $M > 0$ be the constant as in Proposition 7.2. Denote by $B_0(M^{\frac{1}{2}} r)$ the ball of radius $M^{\frac{1}{2}} r > 0$ centered at 0. Then we take

$$D := B_0(M^{\frac{1}{2}} r) \cap L_H,$$

and we will prove the proposition for this D.

Indeed, for any neighborhood Φ of D in $\overline{V}_H$, one can find $\alpha > 0$ such that

$$\{u \in \overline{V}_H : \|u\| \leq M^{\frac{1}{2}} r,\ |f_j(u)| \leq \alpha\ (1 \leq j \leq m)\} \subset \Phi.$$

Define

$$\Psi := \{u \in \overline{V}_H : \|u\| < r,\ |f_j(u)| < \alpha/M\}$$

which is a neighborhood of C in $\overline{V}_H$, and contained in Φ. We show that Φ and Ψ satisfy the desired property.

Suppose

$$\{g_k \exp(x_0 + t\vec{v}) \cdot w : t \in \Xi\} \not\subset \Phi$$

for $x_0 \in \mathrm{Lie}(A)$, $w \in \overline{V}_H$ and $\Xi \subset [0, \delta_0]$. Denote by $\mathfrak{I}$ the following closed subset

$$\{t \in \Xi : \|g_k \exp(x_0 + t\vec{v}) \cdot w\| \leq M^{\frac{1}{2}} r, |f_j(g_k \exp(x_0 + t\vec{v}) \cdot w)| \leq \alpha\ (1 \leq j \leq m)\}.$$

One can write $\mathfrak{I}$ as a disjoint union of the connected components I_i of $\mathfrak{I}$, namely

$$\mathfrak{I} = \bigcup I_i.$$

On each I_i, we have either

$$\sup_{t \in I_i} \|g_k \exp(x_0 + t\vec{v}) \cdot w\|^2 = Mr^2$$

or

$$\sup_{t \in I_i} |f_j(g_k \exp(x_0 + t\vec{v}) \cdot w)| = \alpha$$

for some $1 \leq j \leq m$. Since $\|g_k \exp(x_0 + t\vec{v}) \cdot w\|^2$ and $f_j(g_k \exp(x_0 + t\vec{v}) \cdot w)$ $(1 \leq j \leq m)$ belong to $E(d, \lambda)$, by Proposition 7.2 and the definition of Ψ, we obtain

$$m_{\mathbb{R}}(\{t \in I_i : g_k \exp(x_0 + t\vec{v}) \cdot w \in \Psi\}) \leq \epsilon m_{\mathbb{R}}(I_i).$$

Now we compute

$$\begin{aligned} & m_{\mathbb{R}}(\{t \in \Xi : g_k \exp(x_0 + t\vec{v}) \cdot w \in \Psi\}) \\ = & m_{\mathbb{R}}(\{t \in \mathfrak{I} : g_k \exp(x_0 + t\vec{v}) \cdot w \in \Psi\}) \\ \leq & \sum_i m_{\mathbb{R}}(\{t \in I_i : g_k \exp(x_0 + t\vec{v}) \cdot w \in \Psi\}) \\ \leq & \sum_i \epsilon m_{\mathbb{R}}(I_i) = \epsilon m_{\mathbb{R}}(\mathfrak{I}) \\ \leq & \epsilon m_{\mathbb{R}}(\{t \in \Xi : g_k \exp(x_0 + t\vec{v}) \cdot w \in \Phi\}), \end{aligned}$$

and this concludes the proof of the proposition. □

For our purpose, we introduce a subregion of $R_{g_k,\delta}$ as follows. By Proposition 8.4, we know that

$$\lim_{k \to \infty} \frac{\mathrm{Vol}(\partial R_{g_k,\delta})}{\mathrm{Vol}(R_{g_k,\delta})} = 0.$$

Therefore, for each $k \in \mathbb{N}$, we can find a constant $d_k > 0$ such that

$$\lim_{k\to\infty} d_k = \infty \text{ and } \lim_{k\to\infty} \frac{d_k \operatorname{Vol}(\partial R_{g_k,\delta})}{\operatorname{Vol}(R_{g_k,\delta})} = 0.$$

We define a subregion in $R_{g_k,\delta}$ by

$$R'_{g_k,\delta} = \{\mathbf{t} \in \operatorname{Lie}(A) : \chi(\mathbf{t}) \geq \ln\delta + d_k - \ln|\wedge^l \operatorname{Ad}(g_k)v|, \forall v \in \mathfrak{g}_\chi(\mathbb{Z})\backslash\{0\} \text{ and } \forall \chi \in \mathcal{W}(\mathfrak{g})\}.$$

Here we list some properties about $R'_{g_k,\delta}$.

Lemma 9.7. *Let d_k and $R'_{g_k\delta}$ be as above.*

(1) We have

$$\lim_{k\to\infty} \frac{\operatorname{Vol}(R'_{g_k,\delta})}{\operatorname{Vol}(R_{g_k,\delta})} = 1.$$

(2) For any functional $\chi \in \mathcal{W}(\mathfrak{g})$ and any integer vector $v \in \mathfrak{g}_\chi$ we have

$$\|e^{\chi(\mathbf{t})}(\wedge^l \operatorname{Ad}(g_k)v)\| \geq \delta e^{d_k} \ (\forall \mathbf{t} \in R'_{g_k,\delta}).$$

(3) For $x_0 \in R'_{g_k,\delta}$ and the interval $\Xi = [0, \delta_0] \subset \mathbb{R}$, there exists a constant $c > 0$ which only depends on the linear functionals χ's and δ_0, such that for any integer vector $w \in \wedge^l \mathfrak{g}$ $(1 \leq l \leq \dim \mathfrak{g})$, one has

$$\sup_{t\in\Xi} \| \wedge^l (\operatorname{Ad}(g_k \exp(x_0 + t\vec{v}))) \cdot w\| \geq c\delta e^{d_k}.$$

Proof. The proof of the first claim is similar to that of Corollary 6.4. Indeed, let $\{f_i\}$ be the collection of the facets of $\operatorname{Vol}(R_{g_k,\delta})$, and denote by P_i the hyperplane determined by f_i. For each f_i, let B_i be the unique cylinder with the following properties:

(a) the base of B_i is f_i, and the height of B_i is equal to d_k.
(b) B_i and $R_{g_k,\delta}$ lie in the same half-space determined by P_i.

Then one has

$$\operatorname{Vol}(R_{g_k,\delta}) = \bigcup_i B_i \cup \operatorname{Vol}(R'_{g_k,\delta})$$

and

$$\operatorname{Vol}(R_{g_k,\delta}) \leq \sum_i \operatorname{Vol}(B_i) + \operatorname{Vol}(R'_{g_k,\delta})) = d_k \operatorname{Vol}(\partial R_{g_k,\delta}) + \operatorname{Vol}(R'_{g_k,\delta}).$$

Now the first claim follows from our choice of d_k.

The second claim follows easily from the definition of $R'_{g_k,\delta}$. To prove the last statement, we write for any integer vector $w \in \wedge^l \mathfrak{g}$

$$w = \sum_\chi v_\chi$$

where $v_\chi \in \mathfrak{g}_\chi$ is the integral $\mathfrak{g}_\chi$-coordinate of w. One can compute

$$(\wedge^l \operatorname{Ad}(g_k \exp(\mathbf{t}))) \cdot w = \sum_\chi e^{\chi(\mathbf{t})} \wedge^l \operatorname{Ad}(g_k)v_\chi.$$

Now the last claim follows from the second claim and Proposition 7.5. $\square$

The following proposition is an important step towards Proposition 9.9.

Proposition 9.8 (Cf.[MS95] Proposition 3.4)**.** *Let a compact set $C \subset L_H$ and $0 < \epsilon < 1$ be given. Then there exists a closed subset $\mathcal{S}$ of X contained in $\pi(S(H,W))$ with the following property: for a given compact set $\mathcal{K} \subset X \setminus \mathcal{S}$, there exists a neighbourhood Ψ of C in $\overline{V}_H$ such that for sufficiently large k, for any $x_0 \in R'_{g_k,\delta}$ and $\Xi \subset [0, \delta_0]$, we have*

$$m_{\mathbb{R}}(\{t \in \Xi : g_k \exp(x_0 + t\vec{v})\mathbb{Z}^n \in \mathcal{K} \cap \pi(\overline{\eta}_H^{-1}(\Psi))\}) \leq \epsilon m_{\mathbb{R}}(\Xi).$$

Proof. For the given C and ϵ, we obtain a compact set $D \subset L_H$ as in Proposition 9.6. For this D, we apply Proposition 9.5 and obtain a closed subset $\mathcal{S} = \pi(S(D))$ of X contained in $\pi(S(H,W))$. Now let $\mathcal{K}$ be any compact subset of $X \setminus \mathcal{S}$ and let Φ be an open neighborhood of D in $\overline{V}_H$ as in Proposition 9.5. Finally let Ψ be a neighborhood of C in $\overline{V}_H$ such that the inequality in Proposition 9.6 is satisfied.

By the choice of x_0 and Lemma 9.7, for any integer vector $w \in \wedge^d \mathfrak{g}$ we have

$$\sup_{t \in \Xi} \|g_k \exp(x_0 + t\vec{v}) \cdot w\| \geq c\delta e^{d_k}$$

for some $c > 0$ and hence

$$\{g_k \exp(x_0 + t\vec{v}) \cdot w : t \in \Xi\} \not\subset \Phi$$

for sufficiently large k.

Now for any $s \in \Xi$ with

$$g_k \exp(x_0 + s\vec{v})\mathbb{Z}^n \in \mathcal{K} \cap \pi(\overline{\eta}_H^{-1}(\Psi)),$$

by Proposition 9.5, there is a unique element w_s in $\overline{\eta}_H(\Gamma)$ such that

$$g_k \exp(x_0 + s\vec{v}) \cdot w_s \in \Psi,$$

and let $I_s = [a_s, b_s]$ be the largest closed interval in Ξ containing s such that

(1) for any $t \in I_s$, we have

$$g_k \exp(x_0 + t\vec{v}) \cdot w_s \in \overline{\Phi}$$

(2) either $g_k \exp(x_0 + a_s\vec{v}) \cdot w_s$ or $g_k \exp(x_0 + b_s\vec{v}) \cdot w_s \in \overline{\Phi} \setminus \Phi$.

We denote by $\mathcal{F}$ the collection of all these intervals I_s as s runs over Ξ with

$$g_k \exp(x_0 + s\vec{v})\mathbb{Z}^n \in \mathcal{K} \cap \pi(\overline{\eta}_H^{-1}(\Psi)).$$

By Proposition 9.5 property (3), we know that the intervals in $\mathcal{F}$ cover Ξ at most twice. Also by Proposition 9.6, we have

$$\begin{aligned} & m_{\mathbb{R}}(t \in I_s : g_k \exp(x_0 + t\vec{v}) \cdot w_s \in \Psi) \\ \leq \quad & \epsilon m_{\mathbb{R}}(t \in I_s : g_k \exp(x_0 + t\vec{v}) \cdot w_s \in \Phi) \\ \leq \quad & \epsilon m_{\mathbb{R}}(I_s). \end{aligned}$$

Therefore

$$m_{\mathbb{R}}(t \in \Xi : g_k \exp(x_0 + t\vec{v})\mathbb{Z}^n \in \mathcal{K} \cap \pi(\overline{\eta}_H^{-1}(\Psi)))$$

$$\begin{aligned} &\leq \sum_{I_s \in \mathcal{F}} m_{\mathbb{R}}(t \in I_s : g_k \exp(x_0 + t\vec{v}) \cdot w_s \in \Psi) \\ &\leq \epsilon \sum_{I_s \in \mathcal{F}} m_{\mathbb{R}}(I_s) \leq 2\epsilon m_{\mathbb{R}}(\Xi). \end{aligned}$$

This completes the proof of the proposition. □

Finally, we reach the following

Proposition 9.9. *Let a compact set $C \subset L_H$ and $0 < \epsilon < 1$ be given. Then there exists a closed subset $\mathcal{S}$ of X contained in $\pi(S(H,W))$ with the following property: for a given compact set $\mathcal{K} \subset X \setminus \mathcal{S}$, there exists a neighbourhood Ψ of C in $\overline{V}_H$ such that for sufficiently large $k > 0$ we have*

$$m_{\mathbb{R}^{n-1}}(\{\mathbf{t} \in R_{g_k,\delta} : g_k \exp(\mathbf{t})\mathbb{Z}^n \in \mathcal{K} \cap \pi(\overline{\eta}_H^{-1}(\Psi))\}) \leq \epsilon m_{\mathbb{R}^{n-1}}(R_{g_k,\delta}).$$

Proof. By Lemma 9.7, let k be sufficiently large such that

$$\frac{\mathrm{Vol}(R_{g_k,\delta} \setminus R'_{g_k,\delta})}{\mathrm{Vol}(R_{g_k,\delta})} \leq \frac{\epsilon}{2}.$$

We cut the region $R'_{g_k,\delta}$ into countably many disjoint small boxes of diameter at most δ_0 such that each box has one side parallel to $\vec{v}$. In other words, each box B is of the form

$$B = \{x_0 + t\vec{v} : x_0 \in S \text{ and } t \in \Xi\}$$

where S is the base of B perpendicular to $\vec{v}$, and $\Xi = [0, a]$ is an interval for some $0 < a \leq \delta_0$. Denote by $\mathcal{F}$ the collection of these boxes B.

For any $B \in \mathcal{F}$, and for each $x_0 \in S$ (S the base of B), by Proposition 9.8 we obtain that

$$m_{\mathbb{R}}(\{t \in \Xi : g_k \exp(x_0 + t\vec{v})\mathbb{Z}^n \in \mathcal{K} \cap \pi(\overline{\eta}_H^{-1}(\Psi))\}) \leq \frac{\epsilon}{2} m_{\mathbb{R}}(\Xi)$$

for sufficiently large k. By integrating the inequality above over the region $x_0 \in S$, one has

$$m_{\mathbb{R}^{n-1}}(\{\mathbf{t} \in B : g_k \exp(\mathbf{t})\mathbb{Z}^n \in \mathcal{K} \cap \pi(\overline{\eta}_H^{-1}(\Psi))\}) \leq \frac{\epsilon}{2} m_{\mathbb{R}^{n-1}}(B).$$

Now we compute

$$\begin{aligned} &m_{\mathbb{R}^{n-1}}(\{\mathbf{t} \in R_{g_k,\delta} : g_k \exp(\mathbf{t})\mathbb{Z}^n \in \mathcal{K} \cap \pi(\overline{\eta}_H^{-1}(\Psi))\}) \\ \leq &m_{\mathbb{R}^{n-1}}(\{\mathbf{t} \in R_{g_k,\delta} \setminus R'_{g_k,\delta} : g_k \exp(\mathbf{t})\mathbb{Z}^n \in \mathcal{K} \cap \pi(\overline{\eta}_H^{-1}(\Psi))\}) \\ &+ \sum_{B \in \mathcal{F}} m_{\mathbb{R}^{n-1}}(\{\mathbf{t} \in B : g_k \exp(\mathbf{t})\mathbb{Z}^n \in \mathcal{K} \cap \pi(\overline{\eta}_H^{-1}(\Psi))\}) \\ \leq &\frac{\epsilon}{2} m_{\mathbb{R}^{n-1}}(R_{g_k,\delta}) + \sum_{B \in \mathcal{F}} \frac{\epsilon}{2} m_{\mathbb{R}^{n-1}}(B) \leq \epsilon m_{\mathbb{R}^{n-1}}(R_{g_k,\delta}) \end{aligned}$$

and then the proposition follows. □

10. Proofs of Theorem 2.4 and Theorem 2.5

Proof of Theorem 2.4. We will prove the theorem by induction. Let g_k be a sequence in $\mathrm{SL}(n,\mathbb{R})$. As explained at the end of section 4, without loss of generality, we may assume that g_k's are in the upper triangular unipotent subgroups of $\mathrm{SL}(n,\mathbb{R})$, and each entry of g_k is either zero or diverges to infinity.

Suppose for a start that $\mathcal{A}(A, g_k) = \{0\}$. By passing to a subsequence, we may further assume that $\operatorname{Ad} g_k(\operatorname{Lie} A)$ converges to a subalgebra consisting of nilpotent elements in $\mathfrak{g}$, in the space of Grassmanian of $\mathfrak{g}$. Then by Proposition 8.6, after passing to a subsequence, $[(g_k)_*\mu_{Ax}]$ converges to $[\nu]$ for a probability measure ν. Furthermore, we have

$$\frac{1}{m_{\mathbb{R}^{n-1}}(R_{g_k,\delta})}(g_k)_*(\mu_{Ax}|_{R_{g_k,\delta}}) \to v$$

and ν is invariant under the unipotent subgroup $\exp(\lim_{n\to\infty} \operatorname{Ad} g_k(\operatorname{Lie} A))$.

We will apply Ratner's theorem and the technique of linearization to prove that ν is the Haar measure on $\mathrm{SL}(n,\mathbb{R})/\mathrm{SL}(n,\mathbb{Z})$. According to Theorem 9.2, suppose by way of contradiction that for some $H \in \mathcal{H}^*$ $(H \neq G)$ we have $\nu(T_H(W)) > 0$. Then we can find a compact subset $C \subset T_H(W)$ such that

$$\nu(C) = \alpha > 0.$$

Now let $0 < \epsilon < \alpha$, $C_1 = \overline{\eta}_H(C)$ and $\mathcal{S}$ be the closed subset of X as in Proposition 9.9. Since $C \cap \mathcal{S} = \emptyset$, we can pick a compact neighborhood $\mathcal{K} \subset X \setminus S$ of C. Then by Proposition 9.9, there exists a neighborhood Ψ of C in $\overline{V}_H$ such that for sufficiently large $k > 0$

$$m_{\mathbb{R}^{n-1}}(\{\mathbf{t} \in R_{g_k,\delta} : g_k \exp(\mathbf{t})\mathbb{Z}^n \in K \cap \pi(\overline{\eta}_H^{-1}(\Psi))\}) \leq \epsilon m_{\mathbb{R}^{n-1}}(R_{g_k,\delta})$$

and

$$C \subset K \cap \pi(\overline{\eta}_H^{-1}(\Psi)).$$

This implies that

$$\nu(C) \leq \epsilon < \alpha$$

which contradicts the equation above. Hence ν is the Haar measure on $\mathrm{SL}(n,\mathbb{R})/\mathrm{SL}(n,\mathbb{Z})$.

Now suppose that $\mathcal{A}(A, g_k) \neq \{0\}$. Then by Corollary 5.3, there exists an element $a \in A$ such that $\operatorname{Ad} g_k(a) = a$. This implies that all elements in A and g_k belong to $C_G(a)$ where $C_G(a)$ denotes the centralizer of a in $\mathrm{SL}(n,\mathbb{R})$. Moreover, we have

$$C_G(a) \cong S \times H$$

where S is the center of $C_G(a)$, H is the semisimple component of $C_G(a)$ and H is isomorphic to the product of various $\mathrm{SL}(n_i,\mathbb{R})$ with $n_i < n$, i.e.

$$H \cong \prod \mathrm{SL}(n_i,\mathbb{R}).$$

Let $A_i = A \cap \mathrm{SL}(n_i, \mathbb{R})$ be the connected component of the full diagonal subgroup in $\mathrm{SL}(n_i, \mathbb{R})$, and we have

$$A = S^0 \times \prod A_i.$$

Since $g_k \in N$ is unipotent ($\forall k \in \mathbb{N}$), one has $g_k \in H$. Then we can write $g_k = \prod g_{i,k} \in \prod \mathrm{SL}(n_i, \mathbb{R})$.

The above discussions tell us that our problem now can be reduced to the following setting (recall that $x = e\,\mathrm{SL}(n, \mathbb{Z})$):

(1) the measure μ_{Ax} is supported in the homogeneous space $C_G(a)/(\Gamma \cap C_G(a))$, where one has

$$\begin{aligned} &C_G(a)/(\Gamma \cap C_G(a)) \\ =&S/(\Gamma \cap S) \times H/(\Gamma \cap H) \\ =&S/(\Gamma \cap S) \times \prod (\mathrm{SL}(n_i, \mathbb{R})/\mathrm{SL}(n_i, \mathbb{Z})). \end{aligned}$$

(2) the measure μ_{Ax} can be decomposed, according to the decomposition of $C_G(a)/(\Gamma \cap C_G(a))$, as

$$\mu_{Ax} = \mu_{S^0} \times \prod \mu_{A_i x_i}.$$

Here μ_{S^0} denotes the S^0-invariant measure on $S^0/(\Gamma \cap S^0) = S^0 \cong S/(\Gamma \cap S)$. For each i, $x_i = e\,\mathrm{SL}(k_i, \mathbb{Z})$ is the identity coset in $\mathrm{SL}(n_i, \mathbb{R})/\mathrm{SL}(n_i, \mathbb{Z})$, and $\mu_{A_i x_i}$ denotes the A_i-invariant measure on $A_i x_i$ in $\mathrm{SL}(n_i, \mathbb{R})/\mathrm{SL}(n_i, \mathbb{Z})$.

(3) one pushes μ_{Ax} by the sequence $\{g_k\}$ in the space $C_G(a)/(\Gamma \cap C_G(a))$ in the following manner:

$$(g_k)_* \mu_{Ax} = \mu_{S^0} \times \prod (g_{i,k})_* \mu_{A_i x_i}.$$

Since $n_i < n$, we can now apply the induction hypothesis to each $(g_{i,k})_* \mu_{A_i x_i}$, and obtain that $[g_{i,k} \mu_{A_i x_i}]$ converges to an equivalence class of a periodic measure $[\mu_{G_i y_i}]$ on $\mathrm{SL}(n_i, \mathbb{R})/\mathrm{SL}(n_i, \mathbb{Z})$. Now the first paragraph of the theorem follows by grouping all the measures $[\mu_{G_i y_i}]$ and μ_S^0 back together in the space $\mathrm{SL}(n, \mathbb{R})/\mathrm{SL}(n, \mathbb{Z})$.

As for the second paragraph of the theorem, it essentially follows from the inductive steps (especially, the induction hypothesis) above. Indeed, one can repeat the induction hypothesis several times until we reach the following setting (here we still assume that g_k's are in the upper triangular unipotent subgroups of $\mathrm{SL}(n, \mathbb{R})$, and each entry of g_k is either zero or diverges to infinity.): there is a connected subgroup $S \subset A$ such that

(1) if we denote by $C_G(S)$ the centralizer of S in $G = \mathrm{SL}(n, \mathbb{R})$, then $C_G(S) = S \times H$, where H is the semisimple component of $C_G(S)$ and H is isomorphic to the product of various $\mathrm{SL}(n_i, \mathbb{R})$ with $n_i < n$. Also $g_k \in C_G(S)$.

(2) the measure μ_{Ax} is supported in the homogeneous space $C_G(S)/(\Gamma \cap C_G(S))$, where one has

$$\begin{aligned}&C_G(S)/(\Gamma \cap C_G(a))\\=&S/(\Gamma \cap S) \times H/(\Gamma \cap H)\\=&S/(\Gamma \cap S) \times \prod(\mathrm{SL}(n_i, \mathbb{R})/\mathrm{SL}(n_i, \mathbb{Z})).\end{aligned}$$

(3) the measure μ_{Ax} can be decomposed, according to the decomposition of $C_G(S)/(\Gamma \cap C_G(S))$, as

$$\mu_{Ax} = \mu_{S^0} \times \prod \mu_{A_i x_i}.$$

Here μ_{S^0} denotes the S^0-invariant measure on $S^0/(\Gamma \cap S^0) = S^0 \cong S/(\Gamma \cap S)$. For each i, $x_i = e\,\mathrm{SL}(k_i, \mathbb{Z})$ is the identity coset in $\mathrm{SL}(n_i, \mathbb{R})/\mathrm{SL}(n_i, \mathbb{Z})$, $A_i = A \cap \mathrm{SL}(n_i, \mathbb{R})$ and $\mu_{A_i x_i}$ denotes the A_i-invariant measure on $A_i x_i$ in $\mathrm{SL}(n_i, \mathbb{R})/\mathrm{SL}(n_i, \mathbb{Z})$.

(4) one pushes μ_{Ax} by the sequence $\{g_k\}$ in the space $C_G(S)/(\Gamma \cap C_G(S))$ in the following manner:

$$(g_k)_* \mu_{Ax} = \mu_{S^0} \times \prod (g_{i,k})_* \mu_{A_i x_i}.$$

(5) futhermore, for each $A_i x_i$ in $\mathrm{SL}(n_i, \mathbb{R})/\mathrm{SL}(n_i, \mathbb{Z})$, one has $\mathcal{A}(A_i, g_{i,k}) = \{0\}$.

Here one can see that $S = \mathcal{A}(A, g_k)$. Now we can apply the starting step of the induction (the case $\mathcal{A}(A, g_k) = \{0\}$) to each $\mu_{A_i x_i}$ and obtain that $[(g_{i,k})_* \mu_{A_i x_i}]$ converges to the equivalence class of the Haar measure on $\mathrm{SL}(n_i, \mathbb{R})/\mathrm{SL}(n_i, \mathbb{Z})$. So by the decompositions in (2) and (3), $[(g_K)_* \mu_{Ax}]$ converges to the equivalence class of the periodic measure $[\mu_{C_G(S)^0 x}]$.

If we allow $\{g_k\}$ to be arbitrary, then any limit point of the sequence $[(g_k)_* \mu_{Ax}]$ is a translate of the equivalence class $[\mu_{C_G(S)^0 x}]$. This completes the proof of the theorem. □

The following is an immediate corollary from the proof of Theorem 2.4, which gives an example of λ_k's in Theorem 2.7 for $\mathcal{A}(A, g_k) = \{0\}$. This also generalizes the result in [OS14]. We will apply this special case of Theorem 2.7 in the counting problem in section 11.

Corollary 10.1 (Cf. Theorem 2.7)**.** *For a sequence $g_k \in KN$ with $\mathcal{A}(A, g_k) = \{0\}$, we have*

$$\frac{1}{m_{\mathbb{R}^{n-1}}(\Omega_{g_k,\delta})}(g_k)_* \mu_{Ax} \to \mu_{G/\Gamma}$$

where $\mu_{G/\Gamma}$ is the Haar measure on G/Γ.

In the rest of this section, we will prove Theorem 2.5. As before, let m_X denote the Haar measure on $X = G/\Gamma$. Let H be a connected reductive group containing A. It is known that up to conjugation by an element in the Weyl group of G, H consists of diagonal blocks with each block isomorphic to $GL(k, \mathbb{R})$ with $k < n$. We will assume, for convenience, that H has the form of diagonal blocks, since conjugations by Weyl elements do not affect the theorem.

The following lemma clarifies an assumption in Theorem 2.5.

Lemma 10.2. *Let Ax be a divergent orbit in X and let H be a connected reductive group containing A. Then Hx is closed in X.*

Proof. By the classification of divergent A-orbits of Margulis which appears in the appendix of [TW03], we may assume without loss of generality that x is commensurable to $\mathbb{Z}^n$. Thus, it is enough to prove the lemma for $x = \mathbb{Z}^n$. Then the lemma follows easily for any reductive group H under consideration. □

By reasoning in the same way as at the beginning of section 4, it is harmless to assume $x = e\,\mathrm{SL}(n,\mathbb{Z})$ in the proof of Theorem 2.5. So in the sequel, x will always denote the identity coset in $\mathrm{SL}(n,\mathbb{R})/\,\mathrm{SL}(n,\mathbb{Z})$

Let P be the standard $\mathbb{Q}$-parabolic subgroup in G having H as (the connected component of) a Levi component. Let $U \subset N$ be the unipotent radical of P. For any element $g \in G$, we can write

$$g = kuh$$

where $k \in K = \mathrm{SO}(n,\mathbb{R})$, $u \in U$ and $h \in H$. Since

$$g_*\mu_H = (ku)_*\mu_H,$$

we may assume that $g_k \in U$ in the theorem. We write

$$H = S \times H_{ss}$$

where S is the connected component of the center of H, and H_{ss} is the semisimple component of H. We will denote by A_{ss} the connected component of the full diagonal group in H_{ss}. Note that we have

$$A = S \times A_{ss}.$$

By Theorem 2.4, we can find a sequence of upper triangular unipotent matrices $h_k \in H$ such that

$$[(h_k)_*\mu_{Ax}] \to [\mu_{Hx}]$$

and this happens when $\mathrm{Ad}(h_k)\,\mathrm{Lie}(A_{ss})$ converges to a nilpotent subalgebra in $\mathfrak{g} = \mathfrak{sl}(n,\mathbb{R})$. We will fix such a sequence $\{h_k\}$.

In what follows, we will keep the assumption on the sequence $\{g_k\}$ that g_k's are in the upper triangular unipotent subgroup U, and each entry of g_k either equals 0 or diverges to infinity.

Proposition 10.3. *If the subalgebra $\mathcal{A}(S, g_k)$ of $\mathrm{Lie}(A)$ equals $\{0\}$, then for any subsequence $\{g_{m_k}\}$ of $\{g_m\}$ and any subsequence $\{h_{n_k}\}$ of $\{h_n\}$, the subalgebra $\mathcal{A}(A, g_{m_k}h_{n_k})$ of $\mathrm{Lie}(A)$ equals $\{0\}$.*

Proof. It suffices to show that every element $Y \in \mathrm{Lie}(A)$ diverges to ∞ under the adjoint action of $g_{m_k}h_{n_k}$. Let

$$Y = Y_1 + Y_2$$

where $Y_1 \in \mathrm{Lie}(S)$ and $Y_2 \in \mathrm{Lie}(A_{ss})$. If $Y_2 = 0$, then Y diverges to ∞ by the condition $\mathcal{A}(S, g_k) = \{0\}$. If $Y_2 \neq 0$, then we have

$$\begin{aligned} & \mathrm{Ad}(g_{m_k}h_{n_k})(Y) \\ = \;& \mathrm{Ad}(g_{n_k})(Y_1 + \mathrm{Ad}(h_{n_k})Y_2) \\ = \;& (\mathrm{Ad}(g_{n_k})(Y_1 + \mathrm{Ad}(h_{n_k})Y_2) - (Y_1 + \mathrm{Ad}(h_{n_k})Y_2)) \\ & \quad + (Y_1 + \mathrm{Ad}(h_{n_k})Y_2). \end{aligned}$$

Since $g_n \in U$ and $Y_1 + \mathrm{Ad}(h_{n_k})Y_2 \in \mathrm{Lie}(H)$, we know that $\mathrm{Ad}(g_{n_k})(Y_1 + \mathrm{Ad}(h_{n_k})Y_2) - (Y_1 + \mathrm{Ad}(h_{n_k})Y_2) \in \mathrm{Lie}(U)$. Also $\mathrm{Ad}(h_{n_k})Y_2 \in \mathrm{Lie}(H)$ and $\mathrm{Ad}(h_{n_k})Y_2 \to \infty$ by our choice of $\{h_n\}$. Hence $\mathrm{Ad}(g_{m_k}h_{n_k})(Y)$ diverges to ∞. □

We will fix a nonnegative function $f_0 \in C_c(X)$ such that $\mathrm{supp}(f_0)$ contains the compact orbit $N\mathbb{Z}^n$ in X. This implies that for any $g \in N$ we have

$$\int f_0 dg_* \mu_{Ax} > 0.$$

Proposition 10.4. *Suppose that the subalgebra $\mathcal{A}(S, g_k)$ of $\mathrm{Lie}(A)$ equals $\{0\}$. Let $f \in C_c(X)$. Then for any $\epsilon > 0$, there exists $N > 0$ such that for any $m, n > N$*

$$\left| \frac{\int f d(g_m h_n)\mu_{Ax}}{\int f_0 d(g_m h_n)\mu_{Ax}} - \frac{\int f dm_X}{\int f_0 dm_X} \right| \leq \epsilon.$$

Proof. Suppose that there exists $\epsilon > 0$ such that for any $k > 0$ there exist $m_k, n_k > k$ with the condition

$$\left| \frac{\int f d(g_{m_k} h_{n_k})\mu_{Ax}}{\int f_0 d(g_{m_k} h_{n_k})\mu_{Ax}} - \frac{\int f dm_X}{\int f_0 dm_X} \right| \geq \epsilon.$$

By Proposition 10.3, we know that the subalgebra $\mathcal{A}(A, g_{m_k}h_{n_k}) = \{0\}$. Hence by Theorem 2.4, we have

$$[(g_{m_k}h_{n_k})\mu_{Ax}] \to [m_X]$$

which contradicts the inequality above. This completes the proof of the proposition. □

Proof of Theorem 2.5. We will prove the theorem by induction. Let g_k be a sequence in G and without loss of generality, we may assume that g_k's are in the upper triangular unipotent subgroup U, and each entry of g_k either equals 0 or diverges to infinity.

Suppose that $\mathcal{A}(S, g_k) = \{0\}$. Let $f \in C_c(X)$. By Proposition 10.4, for any $\epsilon > 0$ there exists $N > 0$ such that for any $m, n > N$

$$\left| \frac{\int f d(g_m h_n)\mu_{Ax}}{\int f_0 d(g_m h_n)\mu_{Ax}} - \frac{\int f dm_X}{\int f_0 dm_X} \right| \leq \epsilon.$$

Now we fix m, let $n \to \infty$ and obtain

$$\left| \frac{\int f dg_m \mu_{Hx}}{\int f_0 dg_m \mu_{Hx}} - \frac{\int f dm_X}{\int f_0 dm_X} \right| \le \epsilon.$$

This implies that $[g_k \mu_{Hx}] \to [m_X]$.

Now suppose that $\mathcal{A}(S, g_k)$ of $\mathrm{Lie}(A)$ is not trivial. By Proposition 5.3, there exists $a \in \mathrm{Lie}(S)$ such that g_k commutes with a. This implies that all elements of H and g_k belong to $C_G(a)$. Moreover, we have

$$C_G(a) \cong S' \times H'$$

where S' is the center of $C_G(a)$, H' is the semisimple component of $C_G(a)$ and H' is isomorphic to the product of various $\mathrm{SL}(n_i, \mathbb{R})$ with $n_i < n$, i.e.

$$H' \cong \prod_i \mathrm{SL}(n_i, \mathbb{R}).$$

Let H_i be the reductive subgroup $H \cap \mathrm{SL}(n_i, \mathbb{R})$ in $\mathrm{SL}(n_i, \mathbb{R})$, and we have

$$H = S'^0 \times \prod_i H_i.$$

Since $g_k \in N$ is unipotent ($\forall k \in \mathbb{N}$), one has $g_k \in H'$. Then we can write $g_k = \prod_i g_{i,k}$ ($g_{i,k} \in \mathrm{SL}(n_i, \mathbb{R})$).

Similar to the arguments in the proof of Theorem 2.4, the above discussions imply that the problem is in the following setting ($x = e\,\mathrm{SL}(n, \mathbb{Z})$):

(1) the measure μ_{Hx} is supported in the homogeneous space $C_G(a)/(\Gamma \cap C_G(a))$, where one has

$$\begin{aligned} & C_G(a)/(\Gamma \cap C_G(a)) \\ =& S'/(\Gamma \cap S') \times H'/(\Gamma \cap H') \\ =& S'/(\Gamma \cap S') \times \prod (\mathrm{SL}(n_i, \mathbb{R})/\mathrm{SL}(n_i, \mathbb{Z})). \end{aligned}$$

(2) the measure μ_{Hx} can be decomposed, according to the decomposition of $C_G(a)/(\Gamma \cap C_G(a))$, as

$$\mu_{Hx} = \mu_{S'^0} \times \prod \mu_{H_i x_i}.$$

Here $\mu_{S'^0}$ denotes the S'^0-invariant measure on $S^0/(\Gamma \cap S'^0) = S'^0 \cong S'/(\Gamma \cap S')$. For each i, $x_i = e\,\mathrm{SL}(k_i, \mathbb{Z})$ is the identity coset in $\mathrm{SL}(n_i, \mathbb{R})/\mathrm{SL}(n_i, \mathbb{Z})$, and $\mu_{H_i x_i}$ denotes the H_i-invariant measure on $H_i x_i$ in $\mathrm{SL}(n_i, \mathbb{R})/\mathrm{SL}(n_i, \mathbb{Z})$.

(3) one pushes μ_{Hx} by the sequence $\{g_k\}$ in the space $C_G(a)/(\Gamma \cap C_G(a))$:

$$(g_k)_* \mu_{Hx} = \mu_{S'^0} \times \prod (g_{i,k})_* \mu_{H_i x_i}.$$

Since $n_i < n$, we can now apply the induction hypothesis to each $(g_{i,k})_* \mu_{H_i x_i}$, and obtain that $[(g_{i,k})_* \mu_{H_i x_i}]$ converges to an equivalence class of a periodic measure $[\mu_{G_i y_i}]$ on $\mathrm{SL}(n_i, \mathbb{R})/\mathrm{SL}(n_i, \mathbb{Z})$. Now the first paragraph of the theorem follows by gluing all the measures $[\mu_{G_i y_i}]$ and μ_S^0 back together in the space $\mathrm{SL}(n, \mathbb{R})/\mathrm{SL}(n, \mathbb{Z})$.

As for the second paragraph of the theorem, the proof is very similar to Theorem 2.4. One just needs to replace A in the proof of Theorem 2.4 by the reductive group H. This completes the proof of Theorem 2.5. $\square$

11. An application to counting problem

In this section, we will prove Theorem 2.9. Let $p_0(\lambda)$ be a monic polynomial in $\mathbb{Z}[x]$ such that $p_0(\lambda)$ splits completely in $\mathbb{Q}$. Then by Gauss lemma, we have $p(\lambda) = (\lambda - \alpha_1)(\lambda - \alpha_2)\cdots(\lambda - \alpha_n)$ for $\alpha_i \in \mathbb{Z}$. We assume that α_i are distinct and nonzero. Let $M(n,\mathbb{R})$ be the space of $n \times n$ matrices with the norm

$$\|M\|^2 = \mathrm{Tr}(M^t M) = \sum_{1\le i,j\le n} x_{ij}^2$$

for $M = (x_{ij})_{1\le i,j\le n}$. Note that this norm is $\mathrm{Ad}(K)$-invariant. We will denote by B_T the ball of radius T centered at 0 in $M(n,\mathbb{R})$. We denote by

$$M_\alpha = \mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n) \in M(n,\mathbb{Z}).$$

For $M \in M(n,\mathbb{R})$, we denote by $p_M(\lambda)$ the characteristic polynomial of M. We consider

$$V(\mathbb{R}) = \{M \in M(n,\mathbb{R}) : p_M(\lambda) = p_0(\lambda)\}$$

and its subset of integral points

$$V(\mathbb{Z}) = \{M \in M(n,\mathbb{Z}) : p_M(\lambda) = p_0(\lambda)\}.$$

We would like to get an asymptotic formula for

$$\#|V(\mathbb{Z}) \cap B_T| = \#|\{M \in M(n,\mathbb{Z}) : p_M(\lambda) = p_0(\lambda),\ \|M\| \le T\}|.$$

We begin with the following proposition which is a corollary of [BHC62] and [LM33].

Proposition 11.1. *We have*

$$\mathrm{Ad}(\mathrm{SL}(n,\mathbb{R}))M_\alpha = V(\mathbb{R})$$

and there are finitely many $\mathrm{SL}(n,\mathbb{Z})$*-orbits in* $V(\mathbb{Z})$*. The number of the* $\mathrm{SL}(n,\mathbb{Z})$*-orbits in* $V(\mathbb{Z})$ *is equal to the number of classes of nonsingular ideals in the ring* $\mathbb{Z}[M_\alpha]$.

By Proposition 11.1, it suffices to compute the integral points of an $\mathrm{SL}(n,\mathbb{Z})$-orbit. In what follows, we will consider the $\mathrm{SL}(n,\mathbb{Z})$-orbit of M_α. We will apply Theorem 2.7 (more precisely, Corollary 10.1) with initial point $x = e\Gamma$ to count

$$\#|\mathrm{Ad}(\mathrm{SL}(n,\mathbb{Z}))M_\alpha \cap B_T|.$$

For any other $\mathrm{SL}(n,\mathbb{Z})$-orbit of $M' \in V(\mathbb{Z})$, there exists $M_q \in \mathrm{SL}(n,\mathbb{Q})$ such that

$$\mathrm{Ad}(M_q)M' = M_\alpha$$

and the treatment for $\mathrm{Ad}(\mathrm{SL}(n,\mathbb{Z}))M'$ would be similar, just with a change of initial point from $e\Gamma$ to $x_q = M_q\Gamma$. See also the beginning of section 4.

Now let $h = (u_{ij}) \in N$ and write

$$\mathrm{Ad}(h)M_\alpha = hM_\alpha h^{-1} = (x_{ij})$$

where $x_{ii} = \alpha_i$ and $u_{ij} = 0$ $(i > j)$. We have

$$hM_\alpha = (x_{ij})h$$

and

$$\alpha_j u_{ij} = \sum_k x_{ik} u_{kj}$$

$$(\alpha_j - \alpha_i)u_{ij} = \sum_{k \neq i} x_{ik} u_{kj}.$$

Let

$$q_i(x) = \prod_{k=1}^{i} (x - \alpha_k).$$

The following lemmas (Lemma 11.2 and Lemma 11.3) describe the relation between u_{ij} and x_{ij}.

Lemma 11.2. *For $j > i$, we have*

$$u_{ij} = \frac{1}{\alpha_j - \alpha_i} x_{ij} + f_{ij}(x)$$

where f_{ij} is a polynomial in variables x_{pq} with $0 < q - p < j - i$, and $f_{ij} = 0$ for $j - i = 1$. In particular, we have the change of coordinates of the Haar measure on N

$$\prod_{j>i} du_{ij} = \frac{1}{\prod_{j>i} |\alpha_j - \alpha_i|} \prod_{j>i} dx_{ij}.$$

Proof. It is easy to see that $u_{ij} = x_{ij} = 0$ $(i > j)$ and $u_{ii} = 1$. We prove the proposition by induction on $j - i$. For $j - i = 1$, we have

$$u_{ij} = u_{j-1,j} = \frac{1}{\alpha_j - \alpha_{j-1}} \sum_{k \neq j-1} x_{j-1,k} u_{kj} = \frac{1}{\alpha_j - \alpha_{j-1}} x_{j-1,j}.$$

Now we have

$$(\alpha_j - \alpha_i)u_{ij} = \sum_{k \neq i} x_{ik} u_{kj} = \sum_{i<k<j} x_{ik} u_{kj} + x_{ij}$$

where $j - k < j - i$. We complete the proof by applying the induction hypothesis on u_{kj}. □

Lemma 11.3. *For $j > i$, we have*

$$u_{ij} = \prod_{k=i}^{j-1} \frac{x_{k,k+1}}{\alpha_j - \alpha_k} + f_{ij}(x) = \frac{q_{i-1}(\alpha_j)}{q_{j-1}(\alpha_j)} \prod_{k=i}^{j-1} x_{k,k+1} + f_{ij}(x)$$

where $f_{ij}(x)$ is a polynomial in variables x_{pq} $(p < q)$ of degree less than $j - i$.

Proof. We prove the proposition by induction on $j-i$. For $j-i=1$, we have

$$(\alpha_j-\alpha_i)u_{ij}=(\alpha_j-\alpha_{j-1})u_{j-1,j}=\sum_{k\neq j-1}x_{j-1,k}u_{kj}=x_{j-1,j}.$$

Now we have

$$(\alpha_j-\alpha_i)u_{ij}=\sum_{k\neq i}x_{ik}u_{kj}=\sum_{i<k\leq j}x_{ik}u_{kj}$$

where $j-k<j-i$. By applying the induction hypothesis on u_{kj} we have

$$\begin{aligned}&(\alpha_j-\alpha_i)u_{ij}\\ =\ &\sum_{i<k\leq j}x_{ik}\prod_{p=k}^{j-1}\frac{x_{p,p+1}}{\alpha_j-\alpha_p}+\ldots\\ =\ &x_{i,i+1}\prod_{p=i+1}^{j-1}\frac{x_{p,p+1}}{\alpha_j-\alpha_p}+\ldots\end{aligned}$$

Here we omit the terms of degree less than $j-i$. This completes the proof of the proposition. $\square$

Lemma 11.4. *For any $1\leq l\leq n$ and $1\leq i_1<i_2<\cdots<i_l\leq n$ we have*

$$c(i_1,i_2,\ldots,i_l):=\det\left(\frac{q_{k-1}(\alpha_{i_j})}{q_{i_j-1}(\alpha_{i_j})}\right)_{1\leq k\leq l,1\leq j\leq l}\neq 0.$$

Proof. By algebraic manipulations, we can rewrite the determinant above as

$$\prod_{j=1}^{l}\frac{1}{q_{i_j-1}(\alpha_{i_j})}\begin{pmatrix}1&1&\cdots&1\\ q_1(\alpha_{i_1})&q_1(\alpha_{i_2})&\cdots&q_1(\alpha_{i_l})\\ \vdots&\vdots&\cdots&\vdots\\ q_{l-1}(\alpha_{i_1})&q_{l-1}(\alpha_{i_2})&\cdots&q_{l-1}(\alpha_{i_l})\end{pmatrix}.$$

Since $\deg q_i=i$, by row reductions we have

$$\begin{aligned}&\det\begin{pmatrix}1&1&\cdots&1\\ q_1(\alpha_{i_1})&q_1(\alpha_{i_2})&\cdots&q_1(\alpha_{i_l})\\ \vdots&\vdots&\cdots&\vdots\\ q_{l-1}(\alpha_{i_1})&q_{l-1}(\alpha_{i_2})&\cdots&q_{l-1}(\alpha_{i_l})\end{pmatrix}\\ =\ &\det\begin{pmatrix}1&1&\cdots&1\\ \alpha_{i_1}&\alpha_{i_2}&\cdots&\alpha_{i_l}\\ \vdots&\vdots&\cdots&\vdots\\ \alpha_{i_1}^{l-1}&\alpha_{i_2}^{l-1}&\cdots&\alpha_{i_l}^{l-1}\end{pmatrix}\neq 0.\end{aligned}$$

$\square$

Proposition 11.5. *For any $h\in N$ (recall $\mathrm{Ad}(h)M_\alpha=(x_{ij})$), we have*

$$h(e_{i_1}\wedge e_{i_2}\wedge\cdots\wedge e_{i_l})$$

$$= \quad c(i_1, i_2, \ldots, i_l) \prod_{j=1}^{l} \prod_{p=j}^{i_j-1} x_{p,p+1}(e_1 \wedge e_2 \wedge \cdots \wedge e_l) + \ldots$$

Here $c(i_1, i_2, \ldots, i_l)$ is the number in Lemma 11.4 and we omit the terms of polynomials in variables x_{pq} $(p < q)$ of degrees smaller than $\sum_{j=1}^{l}(i_j - j)$.

Proof. By Lemma 11.3, we know that u_{ij} is a polynomial of degree $j - i$. This implies that the term in $h(e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_l})$ corresponding to the $e_{j_1} \wedge e_{j_2} \wedge \cdots \wedge e_{j_l}$-coordinate has degree at most $i_1 + i_2 + \cdots + i_l - j_1 - j_2 - \cdots - j_l$. To prove the proposition, it suffices to prove that the term corresponding to $e_1 \wedge e_2 \wedge \cdots \wedge e_l$ is a polynomial with its leading term

$$c(i_1, i_2, \ldots, i_l) \prod_{j=1}^{l} \prod_{p=j}^{i_j-1} x_{p,p+1}$$

of degree $i_1 + i_2 + \cdots + i_l - 1 - 2 - \cdots - l$.

We know that the coefficient of $e_1 \wedge e_2 \wedge \cdots \wedge e_l$ is equal to

$$\det(u_{k,i_j})_{1 \le k \le l, 1 \le j \le l}$$

and by Lemma 11.3 we know that the leading term of this coefficient is equal to

$$\det \left(\frac{q_{k-1}(\alpha_{i_j})}{q_{i_j-1}(\alpha_{i_j})} \prod_{p=k}^{i_j-1} x_{p,p+1} \right)_{1 \le k \le l, 1 \le j \le l}.$$

The expansion formula of determinant then gives

$$\sum_{\sigma \in S_l} (-1)^{\mathrm{sign}(\sigma)} \prod_{j=1}^{l} \frac{q_{\sigma(j)-1}(\alpha_{i_j})}{q_{i_j-1}(\alpha_{i_j})} \prod_{p=\sigma(j)}^{i_j-1} x_{p,p+1}$$

where σ runs over all the permutations in the symmetric group S_l. Note that we have

$$\begin{aligned} \prod_{j=1}^{l} \prod_{p=\sigma(j)}^{i_j-1} x_{p,p+1} &= \prod_{j=1}^{l} \frac{\prod_{p=1}^{i_j-1} x_{p,p+1}}{\prod_{p=1}^{\sigma(j)-1} x_{p,p+1}} = \frac{\prod_{j=1}^{l} \prod_{p=1}^{i_j-1} x_{p,p+1}}{\prod_{j=1}^{l} \prod_{p=1}^{j-1} x_{p,p+1}} \\ &= \prod_{j=1}^{l} \prod_{p=j}^{i_j-1} x_{p,p+1}. \end{aligned}$$

This implies that

$$\begin{aligned} &\det \left(\frac{q_{k-1}(\alpha_{i_j})}{q_{i_j-1}(\alpha_{i_j})} \prod_{p=k}^{i_j-1} x_{p,p+1} \right)_{1 \le k \le l, 1 \le j \le l} \\ = &\left(\sum_{\sigma \in S_l} (-1)^{\mathrm{sign}(\sigma)} \prod_{j=1}^{l} \frac{q_{\sigma(j)-1}(\alpha_{i_j})}{q_{i_j-1}(\alpha_{i_j})} \right) \prod_{j=1}^{l} \prod_{p=j}^{i_j-1} x_{p,p+1} \end{aligned}$$

$$= \quad c(i_1, i_2, \ldots, i_l) \prod_{j=1}^{l} \prod_{p=j}^{i_j - 1} x_{p,p+1}$$

where $c(i_1, i_2, \ldots, i_l)$ is the number as in Lemma 11.4. $\square$

We define the regions

$$N(T) = \{h \in N : \mathrm{Ad}(h) M_\alpha = (x_{ij}) \in B_T\}$$

$$N(\epsilon, T) = \{h \in N : \mathrm{Ad}(h) M_\alpha \in B_T, |x_{i,i+1}| \geq \epsilon T\}.$$

Lemma 11.6. *We have*

$$\mu_N(N(T)) = \frac{\mathrm{Vol}(B_1)}{\prod_{j>i} |\alpha_j - \alpha_i|} T^{n(n-1)/2}$$

$$\mu_N(N \setminus N(\epsilon, T)) = O(\epsilon T^{n(n-1)/2}).$$

Here μ_N denotes the Haar measure on N.

Proof. This follows immediately from Lemma 11.2. $\square$

In the following, we compute the volume $\mathrm{Vol}(V(\mathbb{R}) \cap B_T)$ with respect to the volume form $d\mu_{V(\mathbb{R})}$ on $V(\mathbb{R})$ induced by the G-invariant measure on G/A. ($G = \mathrm{SL}(n, \mathbb{R})$) By Iwasawa decomposition, one has

$$V(\mathbb{R}) \cong G/A \cong KN$$

and it is well-known that for any $f \in C_c(G/A)$

$$\int_{G/A} f d\mu_{V(\mathbb{R})} = \int_K \int_N f(kh) d\mu_K(k) d\mu_N(h)$$

via this isomorphism.

Proposition 11.7. *We have*

$$\mathrm{Vol}(V(\mathbb{R}) \cap B_T) = \frac{\mathrm{Vol}(B_1)}{\prod_{j>i} |\alpha_j - \alpha_i|} T^{n(n-1)/2}.$$

Proof. Note that by the remark above, one has

$$\mu_{V(\mathbb{R})}(V(\mathbb{R}) \cap B_T) = \mu_K \times \mu_N(\{kh : \mathrm{Ad}(kh) M_\alpha \in B_T\}).$$

By Lemma 11.6 and the $\mathrm{Ad}(K)$-invariance of the norm on $M(n, \mathbb{R})$, we compute

$$\begin{aligned} & \mu_K \times \mu_N(\{kh : \mathrm{Ad}(kh) M_\alpha \in B_T\}) \\ = & \mu_N(\{h : \mathrm{Ad}(h) M_\alpha \in B_T\}) \\ = & \frac{\mathrm{Vol}(B_1)}{\prod_{j>i} |\alpha_j - \alpha_i|} T^{n(n-1)/2}. \end{aligned}$$

This completes the proof of the proposition. $\square$

Proposition 11.8. *For any $k \in K$ and $h \in N(T)$ we have*

$$\mathrm{Vol}(\Omega_{kh,\delta}) = O((\ln T)^{n-1})$$

where the implicit constant depends only on δ and M_α. Furthermore, for $h \in N(\epsilon, T)$ we have

$$\mathrm{Vol}(\Omega_{kh,\delta}) = (c_0 + o(1))(\ln T)^{n-1}$$

where the implicit constant depends on $\epsilon, \delta, M_\alpha$, and c_0 equals the volume of

$$\left\{ \mathbf{t} \in \mathrm{Lie}(A) : \sum_{j=1}^{l} t_{i_j} \geq \sum_{j=1}^{l} (j - i_j), \forall 1 \leq l \leq n, \forall i_1 < \cdots < i_l \right\}.$$

Proof. From the definition of $\Omega_{kh,\delta}$, we know that

$$\Omega_{kh,\delta} = \{\mathbf{t} \in \mathrm{Lie}(A) : \sum_{j=1}^{l} t_{i_j} \geq \ln \delta - \ln \|khe_I\| \text{ for any nonempty } I \in \mathcal{I}_n\}.$$

Since $k \in \mathrm{SO}(n, \mathbb{R})$, by Proposition 11.5, for any $i_1 < i_2 < \cdots < i_l$ we have

$$\begin{aligned} & \ln \delta - \ln \|kh(e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_l})\| \\ \geq \; & O(1) - (i_1 + i_2 + \cdots + i_l - 1 - 2 - \cdots - l)(\ln T) \end{aligned}$$

where the implicit constant depends only on δ and M_α. Moreover if $h \in N(\epsilon, T)$ then we have

$$\begin{aligned} & \ln \delta - \ln \|kh(e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_l})\| \\ = \; & O(1) - (i_1 + i_2 + \cdots + i_l - 1 - 2 - \cdots - l)(\ln T) \end{aligned}$$

where the implicit constant depends only on ϵ, δ and M_α. The proposition now follows from these equations. $\square$

Define

$$F_T(g) = \sum_{\gamma \in \Gamma / \Gamma_{M_\alpha}} \chi_T(\mathrm{Ad}(g\gamma) M_\alpha)$$

where χ_T is the characteristic function of B_T in $M(n, \mathbb{R})$ and Γ_{M_α} is the stabilizer of M_α in $\Gamma = \mathrm{SL}(n, \mathbb{Z})$. This defines a function on G/Γ. Note that χ_T is $\mathrm{Ad}(K)$-invariant and Γ_{M_α} is finite. In the following proposition, we will denote by

$$(f, \phi) := \int_{G/\Gamma} f(g)\phi(g) d\mu_{G/\Gamma}(g)$$

for any two functions f, ϕ on G/Γ, whenever this integral is valid. We will also write μ_H for the Haar measure of a subgroup H in G

Proposition 11.9. *For any $\psi \in C_c(G/\Gamma)$, we have*

$$\left(\frac{|\Gamma_{M_\alpha}|}{n_0 T^{n(n-1)/2} (\ln T)^{n-1}} F_T, \psi \right) \to (1, \psi).$$

Here
$$n_0 = \frac{c_0 \operatorname{Vol}(B_1)}{\prod_{j>i} |\alpha_j - \alpha_i|}$$
and c_0 is the number as in Proposition 11.8.

Proof. We have
$$\begin{aligned}
(F_T, \psi) &= \frac{1}{|\Gamma_{M_\alpha}|} \int_{G/\Gamma} \sum_{\gamma\in\Gamma} \chi_T(\operatorname{Ad}(g\gamma)M_\alpha)\psi(g) dm_X \\
&= \frac{1}{|\Gamma_{M_\alpha}|} \int_{G} \chi_T(\operatorname{Ad}(g)M_\alpha)\psi(g) d\mu_G(g) \\
&= \frac{1}{|\Gamma_{M_\alpha}|} \int_{KN} \int_A \chi_T(\operatorname{Ad}(kha)M_\alpha)\psi(kha) d\mu_K d\mu_N d\mu_A \\
&= \frac{1}{|\Gamma_{M_\alpha}|} \int_{K} \int_N \chi_T(\operatorname{Ad}(h)M_\alpha) d\mu_N d\mu_K \int_A \psi(kha) d\mu_A.
\end{aligned}$$
Now fix $\epsilon > 0$ and by Corollary 10.1, we proceed
$$\begin{aligned}
= & \frac{1}{|\Gamma_{M_\alpha}|} \int_{K} \int_{N(\epsilon,T)} \chi_T(\operatorname{Ad}(h)M_\alpha) \operatorname{Vol}(\Omega_{kh,\delta}) d\mu_N d\mu_K \frac{1}{\operatorname{Vol}(\Omega_{kh,\delta})} \int_A \psi(kha) d\mu_A \\
& + \frac{1}{|\Gamma_{M_\alpha}|} \int_{K} \int_{N\setminus N(\epsilon,T)} \chi_T(\operatorname{Ad}(h)M_\alpha) d\mu_N d\mu_K \int_A \psi(kha) d\mu_A \\
= & \frac{1}{|\Gamma_{M_\alpha}|} \int_{K} \int_{N(\epsilon,T)} \chi_T(\operatorname{Ad}(h)M_\alpha) \operatorname{Vol}(\Omega_{kh,\delta}) d\mu_N d\mu_K \left(\int_{G/\Gamma} \psi dm_X + o_\epsilon(1) \right) \\
& + \frac{1}{|\Gamma_{M_\alpha}|} \int_{K} \int_{N\setminus N(\epsilon,T)} \chi_T(\operatorname{Ad}(h)M_\alpha) d\mu_N d\mu_K \int_A \psi(kha) d\mu_A.
\end{aligned}$$
Note that since $\psi \in C_c(G/\Gamma)$ we can find $\delta_\psi > 0$ such that
$$\int_A \psi(kha) d\mu_A = \int_{\Omega_{kh,\delta_\psi}} \psi(kha) d\mu_A.$$
So by Lemma 11.6 and Proposition 11.8, we proceed
$$\begin{aligned}
= \quad & \frac{1}{|\Gamma_{M_\alpha}|} \int_{K} \int_{N(\epsilon,T)} \chi_T(\operatorname{Ad}(h)M_\alpha) \operatorname{Vol}(\Omega_{kh,\delta}) d\mu_N d\mu_K \int_{G/\Gamma} \psi dm_X \\
& + o_\epsilon(T^{n(n-1)/2} (\ln T)^{n-1}) + O_\psi(\epsilon T^{n(n-1)/2} (\ln T)^{n-1}) \\
= \quad & \frac{n_0 T^{n(n-1)/2} (\ln T)^{n-1}}{|\Gamma_{M_\alpha}|} \int_{G/\Gamma} \psi dm_X \\
& + o_{\epsilon,\delta}(T^{n(n-1)/2} (\ln T)^{n-1}) + O_\psi(\epsilon T^{n(n-1)/2} (\ln T)^{n-1}).
\end{aligned}$$
This implies that
$$\limsup_{T\to\infty} \left| \left(\frac{|\Gamma_{M_\alpha}|}{n_0 T^{n(n-1)/2} (\ln T)^{n-1}} F_T, \psi \right) - (1, \psi) \right| \leq O_\psi(\epsilon).$$
We complete the proof by letting $\epsilon \to 0$.

□

Proof of Theorem 2.9. We follow the same proofs as in [DRS93] and [EMS96], and by combing Lemma 11.6 and Proposition 11.9, we conclude that

$$\frac{|\Gamma_{M_\alpha}|}{n_0 T^{n(n-1)/2}(\ln T)^{n-1}} F_T \to 1.$$

Now Theorem 2.9 follows from this equation and Proposition 11.1 □

References

[AS] M. Aka and U. Shapira, *On the evolution of continued fractions in a fixed quadratic field*, to appear in J. Analyse Math.

[BHC62] A. Borel and Harish-Chandra, *Arithmetic subgroups of algebraic groups*, Ann. of Math. (2) **75** (1962), 485–535. MR0147566

[Dan81] S. G. Dani, *Invariant measures and minimal sets of horospherical flows*, Invent. Math. **64** (1981), no. 2, 357–385. MR629475

[DM93] S. G. Dani and G. A. Margulis, *Limit distributions of orbits of unipotent flows and values of quadratic forms*, I. M. Gel′fand Seminar, 1993, pp. 91–137. MR1237827

[DRS93] W. Duke, Z. Rudnick, and P. Sarnak, *Density of integer points on affine homogeneous varieties*, Duke Math. J. **71** (1993), no. 1, 143–179. MR1230289

[ELMV11] M. Einsiedler, E. Lindenstrauss, P. Michel, and A. Venkatesh, *Distribution of periodic torus orbits and Duke's theorem for cubic fields*, Ann. of Math. (2) **173** (2011), no. 2, 815–885. MR2776363

[EM93] A. Eskin and C. McMullen, *Mixing, counting, and equidistribution in Lie groups*, Duke Math. J. **71** (1993), no. 1, 181–209. MR1230290

[EMS96] A. Eskin, S. Mozes, and N. Shah, *Unipotent flows and counting lattice points on homogeneous varieties*, Ann. of Math. (2) **143** (1996), no. 2, 253–299. MR1381987

[EMS97] A. Eskin, S. Mozes, and N. Shah, *Non-divergence of translates of certain algebraic measures*, Geom. Funct. Anal. **7** (1997), no. 1, 48–80. MR1437473

[EMV09] M. Einsiedler, G. Margulis, and A. Venkatesh, *Effective equidistribution for closed orbits of semisimple groups on homogeneous spaces*, Invent. Math. **177** (2009), no. 1, 137–212. MR2507639

[EV08] J. S. Ellenberg and A. Venkatesh, *Local-global principles for representations of quadratic forms*, Invent. Math. **171** (2008), no. 2, 257–279. MR2367020

[GMO08] A. Gorodnik, F. Maucourant, and H. Oh, *Manin's and Peyre's conjectures on rational points and adelic mixing*, Ann. Sci. Éc. Norm. Supér. (4) **41** (2008), no. 3, 383–435. MR2482443

[KK] D. Kelmer and A. Kontorovich, *Effective equidistribution of shears and applications*, to appear in Math. Ann.

[Kle10] D. Kleinbock, *Quantitative nondivergence and its Diophantine applications*, Homogeneous flows, moduli spaces and arithmetic, 2010, pp. 131–153. MR2648694

[KM98] D. Y. Kleinbock and G. A. Margulis, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Ann. of Math. (2) **148** (1998), no. 1, 339–360. MR1652916

[LM33] C. G. Latimer and C. C. MacDuffee, *A correspondence between classes of ideals and classes of matrices*, Ann. of Math. (2) **34** (1933), no. 2, 313–316. MR1503108

[MS95] S. Mozes and N. Shah, *On the space of ergodic invariant measures of unipotent flows*, Ergodic Theory Dynam. Systems **15** (1995), no. 1, 149–159. MR1314973

[OS14] H. Oh and N. A. Shah, *Limits of translates of divergent geodesics and integral points on one-sheeted hyperboloids*, Israel J. Math. **199** (2014), no. 2, 915–931. MR3219562

[Rag72] M. S. Raghunathan, *Discrete subgroups of Lie groups*, Springer-Verlag, New York-Heidelberg, 1972. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 68. MR0507234

[Rat90a] M. Ratner, *On measure rigidity of unipotent subgroups of semisimple groups*, Acta Math. **165** (1990), no. 3-4, 229–309. MR1075042

[Rat90b] M. Ratner, *Strict measure rigidity for unipotent subgroups of solvable groups*, Invent. Math. **101** (1990), no. 2, 449–482. MR1062971

[Rat91] M. Ratner, *On Raghunathan's measure conjecture*, Ann. of Math. (2) **134** (1991), no. 3, 545–607. MR1135878

[Sha00] N. A. Shah, *Counting integral matrices with a given characteristic polynomial*, Sankhyā Ser. A **62** (2000), no. 3, 386–412. Ergodic theory and harmonic analysis (Mumbai, 1999). MR1803465

[Sha91] N. A. Shah, *Uniformly distributed orbits of certain flows on homogeneous spaces*, Math. Ann. **289** (1991), no. 2, 315–334. MR1092178

[Sha] U. Shapira, *Full escape of mass for the diagonal group*, to appear in Int. Math. Res. Not.

[TW03] G. Tomanov and B. Weiss, *Closed orbits for actions of maximal tori on homogeneous spaces*, Duke Math. J. **119** (2003), no. 2, 367–392. MR1997950

Department of Mathematics, Technion, Haifa, Israel
E-mail address: `ushapira@tx.technion.ac.il`
E-mail address: `cheng.zheng@tx.technion.ac.il`