Annals of Mathematics, **143** (1996), 253–299

Unipotent flows and counting lattice points on homogeneous varieties

By ALEX ESKIN,* SHAHAR MOZES,† and NIMISH SHAH‡

1. Introduction

In this paper, using ergodic properties of subgroup actions on homogeneous spaces of Lie groups, we study the asymptotic behavior of the number of lattice points on certain affine varieties. Consider for instance the following:

Example 1. Let $p(\lambda)$ be a monic polynomial of degree $n \geq 2$ with integer coefficients and irreducible over $\mathbb{Q}$. Let $\mathrm{M}_n(\mathbb{Z})$ denote the set of $n \times n$ integer matrices, and put

$$V_p(\mathbb{Z}) = \{A \in \mathrm{M}_n(\mathbb{Z}) \ : \ \det(\lambda I - A) = p(\lambda)\}.$$

Hence $V_p(\mathbb{Z})$ is the set of integral matrices with characteristic polynomial $p(\lambda)$. Consider the norm on $n \times n$ real matrices given by $\|(x_{ij})\| = \sqrt{\sum_{ij} x_{ij}^2}$, and let $N(T, V_p)$ denote the number of elements of $V_p(\mathbb{Z})$ with norm less than T.

THEOREM 1.1. *Suppose further that $p(\lambda)$ splits over $\mathbb{R}$, and for a root α of $p(\lambda)$ the ring of algebraic integers in $\mathbb{Q}(\alpha)$ is $\mathbb{Z}[\alpha]$. Then, asymptotically as $T \to \infty$,*

$$N(T,V_p) \sim \frac{2^{n-1} hR\omega_n}{\sqrt{D} \cdot \prod_{k=2}^{n} \Lambda(k/2)} T^{n(n-1)/2}$$

where h is the class number of $\mathbb{Z}[\alpha]$, R is the regulator of $\mathbb{Q}(\alpha)$, D is the discriminant of $p(\lambda)$, ω_n is the volume of the unit ball in $\mathbb{R}^{n(n-1)/2}$, and $\Lambda(s) = \pi^{-s}\Gamma(s)\zeta(2s)$.

Example 1 is a special case of the following counting problem which was first studied in [DRS] and [EM]:

---

*Research partially supported by the NSF and the Alfred P. Sloan Foundation.

†Sponsored in part by the Edmund Landau Center for Research in Mathematical Analysis, and supported by the Minerva Foundation (Germany).

‡Research at MSRI supported by NSF grant DMS-8505550.

The counting problem. Let W be a real finite dimensional vector space with a $\mathbb{Q}$ structure and V a Zariski closed real subvariety of W defined over $\mathbb{Q}$. Let G be a reductive real algebraic group defined over $\mathbb{Q}$, which acts on W via a $\mathbb{Q}$-representation $\rho\colon G \to \mathrm{GL}(W)$. Suppose that G acts transitively on V. Let $\|\cdot\|$ denote a Euclidean norm on W. Let B_T denote the ball of radius $T > 0$ in W around the origin, and define

$$N(T, V) = |V \cap B_T \cap \mathbb{Z}^n|,$$

the number of integral points on V with norm less than T. We are interested in the asymptotics of $N(T, V)$ as $T \to \infty$. We use the rich theory of unipotent flows on homogeneous spaces developed in [Mar1], [DM1], [Rat1], [Rat2], [Rat3], [Rat4], [Sha1] and [DM3] to obtain results in this direction.

Let Γ be a subgroup of finite index in $G(\mathbb{Z})$ such that $W(\mathbb{Z})\Gamma \subset W(\mathbb{Z})$. By a theorem of Borel and Harish-Chandra [BH-C], $V(\mathbb{Z})$ is a union of finitely many Γ-orbits. Therefore, to compute the asymptotics of $N(T, V)$, it is enough to consider each Γ-orbit, say $\mathcal{O}$, separately and compute the asymptotics of

$$N(T, V, \mathcal{O}) = |\mathcal{O} \cap B_T|.$$

Suppose that $\mathcal{O} = v_0 \cdot \Gamma$ for some $v_0 \in V(\mathbb{Z})$. Then the stabilizer $H = \{g \in G\colon\ v_0 g = v_0\}$ is a reductive real algebraic $\mathbb{Q}$-subgroup, and $V \cong H\backslash G$. Define

$$R_T = \{Hg \in H\backslash G\colon\ v_0 g \in B_T\},$$

the pullback of the ball B_T to $H\backslash G$.

Assume that G^0 and H^0 do not admit nontrivial $\mathbb{Q}$-characters. Then by the theorem of Borel and Harish-Chandra, $\Gamma\backslash G$ admits a G-invariant (Borel) probability measure, say μ_G, and $(\Gamma \cap H)\backslash H$ admits an H-invariant probability measure, say μ_H. Now the natural inclusion $(\Gamma \cap H)\backslash H \hookrightarrow \Gamma\backslash G$ is an H-equivariant proper map. Let $\pi\colon G \to \Gamma\backslash G$ be the natural quotient map. Then the orbit $\pi(H)$ is closed, $(\Gamma \cap H)\backslash H \cong \pi(H)$, and μ_H can be treated as a measure on $\Gamma\backslash G$ supported on $\pi(H)$. Such finite invariant measures supported on closed orbits of subgroups are called *homogeneous measures*. Let $\lambda_{H\backslash G}$ denote the (unique) G-invariant measure on $H\backslash G$ induced by the normalization of the Haar measures on G and H.

The following result was proved in [DRS]; subsequently, a simpler proof appeared in [EM].

THEOREM 1.2. *Suppose that V is affine symmetric and Γ is irreducible (equivalently, H is the set of fixed points of an involution of G, and G is $\mathbb{Q}$-simple). Then asymptotically as $T \to \infty$,*

$$N(T,V,\mathcal{O}) \sim \lambda_{H\backslash G}(R_T).$$

Translates of homogeneous measures. For any $g \in G$, let $\mu_H g$ denote the translated measure defined as

$$\mu_H g(E) = \mu_H(Eg^{-1}), \qquad \text{for all Borel sets } E \subset \Gamma\backslash G.$$

Note that $\mu_H g$ is supported on $\pi(H)g$. A key ingredient in the proofs of Theorem 1.2 in [DRS] and [EM] is showing that if H is the set of fixed points of an involution of G, then for any sequence $\{g_i\} \subset G$ such that $\{Hg_i\}$ has no convergent subsequence in $H\backslash G$, the translated measures $\mu_H g_i$ get 'equidistributed' on $\Gamma\backslash G$ as $i \to \infty$; that is, the sequence $\{\mu_H g_i\}$ weakly converges to μ_G. The method of [DRS] uses spectral analysis on $\Gamma\backslash G$, while the argument of [EM] uses the mixing property of the geodesic flow. However, both methods seem limited essentially to the affine symmetric case. It should be remarked that for the proof of Theorem 1.2 one needs only certain averages of translates of the form $\mu_H g$ to become equidistributed.

Motivated by this approach to the counting problem, we study the limit distributions of translates of homogeneous measures. We show that under certain conditions if for some sequence $\{g_i\}$ we have $\lim \mu_H g_i = \nu$, then the measure ν is again homogeneous. We give exact algebraic conditions on the sequence $\{g_i\}$ relating it to the limit measure ν. Using this analysis, we show that the counting estimates as in Theorem 1.2 hold for a large class of homogeneous varieties. The following particular cases of homogeneous varieties, which are not affine symmetric, are of interest. We first place Example 1 in this context.

Example 1, *continued.* Note that $V_p(\mathbb{Z})$ is the set of integral points on the real subvariety $V_p = \{A \in \mathrm{M}_n(\mathbb{R})\colon \det(\lambda I - A) = p(\lambda)\}$ contained in the vector space $W = \mathrm{M}_n(\mathbb{R})$. Let $G = \{g \in \mathrm{GL}_n(\mathbb{R})\colon \det g = \pm 1\}$. Then G acts on W via conjugations, and V_p is a closed orbit of G (see [New, Th. III.7]). Put $\Gamma = G(\mathbb{Z}) = \mathrm{GL}_n(\mathbb{Z})$. The companion matrix of $p(\lambda)$ is

$$v_0 = \begin{pmatrix} 0 & & 0 & -a_n \\ 1 & & 0 & -a_{n-1} \\ 0 & \cdots & \vdots & \vdots \\ \vdots & & 0 & \\ 0 & & 1 & -a_1 \end{pmatrix} \in V_p(\mathbb{Z}). \tag{1}$$

The centralizer H of v_0 is a maximal $\mathbb{Q}$-torus and H^0 has no nontrivial $\mathbb{Q}$-characters. We emphasize that H is not the set of fixed points of an involution, and the variety $V_p = H\backslash G$ is *not* affine symmetric. Nevertheless, we show that $N(T, V_p, \Gamma v_0) \sim \lambda_{H\backslash G}(R_T)$. By computing the volumes, we obtain the following estimate:

THEOREM 1.3. *Let $N(T,V_p)$ be the number of points on $V_p(\mathbb{Z})$ of norm less than T. Then asymptotically as $T \to \infty$,*

$$N(T,V_p) \sim c_p T^{n(n-1)/2},$$

where $c_p > 0$ is an explicitly computable constant.

We obtain a 'formula' for calculating c_p; for the sake of simplicity, we calculate it explicitly only under the additional assumptions on $p(\lambda)$ of Theorem 1.1. See [BR] for some deeper consequences of the above result.

Example 2. Let A be a nondegenerate indefinite integral quadratic form in $n \geq 3$ variables and of signature (p,q), where $p \geq q$, and B a definite integral quadratic form in $m \leq p$ variables. Let $W = \mathrm{M}_{m\times n}(\mathbb{R})$ be the space of $m \times n$ matrices. Consider the norm on W given by $\|(x_{ij})\| = \sqrt{\sum_{i,j} x_{ij}^2}$. Define

$$V_{A,B} = \{X \in \mathrm{M}_{m\times n}(\mathbb{R}) \;:\; XA\,{}^{t}X = B\}.$$

Thus a point on $V_{A,B}(\mathbb{Z})$ corresponds to a way of representing B by A over $\mathbb{Z}$. We assume that $V_{A,B}(\mathbb{Z})$ is not empty.

The group $G = \mathrm{SO}(A)$ acts on W via right multiplication, and the action is transitive on $V_{A,B}$. The stabilizer of a point $\xi \in V_{A,B}$ is an orthogonal group H_ξ in $n-m$ variables. Let $\Gamma = G(\mathbb{Z})$. Then the number of Γ-orbits on $V_{A,B}(\mathbb{Z})$ is finite. Let $\xi_1, \dots, \xi_h$ be the representatives for the orbits.

THEOREM 1.4. *Let $N(T,V_{A,B})$ denote the number of points on $V_{A,B}(\mathbb{Z})$ with norm less than T. Then asymptotically as $T \to \infty$,*

$$N(T,V_{A,B}) \sim \sum_{i=1}^{h} \frac{\mathrm{vol}(\Gamma \cap H_{\xi_i} \backslash H_{\xi_i})}{\mathrm{vol}(\Gamma \backslash G)} c_{A,B} T^{r(n-r-1)}$$

where $r = \min(m,q)$, and $c_{A,B} > 0$ is an explicitly computable constant (*see equation* 40).

Remark 1.5. In some ranges of p,q,m,n this formula may be proved by the Hardy-Littlewood circle method, or by Θ-function techniques. Using our method one also obtains asymptotic formulas for the number of points in the individual orbits $\Gamma\xi_i$.

Remark 1.6. In the case $m > q$, the asymptotics of the number of integer points does not agree with the heuristic of the Hardy-Littlewood circle method, even if the number of variables mn is very large compared to the number of quadratic equations $m(m+1)/2$. The discrepancy occurs because the null locus $\{X\colon\ XA\,{}^{t}X = 0\}$ does not contain a nonsingular real point (cf. [Bir, Th. 1]) and so the 'singular integral' vanishes.

Limiting distributions of translates of homogeneous measures. The following is the main result of this paper, which allows us to investigate the counting problems. The result is also of general interest, especially from the viewpoint of ergodic theory on homogeneous spaces of Lie groups.

THEOREM 1.7. *Let G be a connected real algebraic group defined over $\mathbb{Q}$, $\Gamma \subset G(\mathbb{Q})$ an arithmetic lattice in G with respect to the $\mathbb{Q}$-structure on G, and $\pi\colon G \to \Gamma\backslash G$ the natural quotient map. Let $H \subset G$ be a connected real algebraic $\mathbb{Q}$-subgroup admitting no nontrivial $\mathbb{Q}$-characters. Let μ_H denote the H-invariant probability measure on the closed orbit $\pi(H)$. For a sequence $\{g_i\} \subset G$, suppose that the translated measures $\mu_H g_i$ converge to a probability measure μ on $\Gamma\backslash G$. Then there exists a connected real algebraic $\mathbb{Q}$-subgroup L of G containing H such that the following holds:*

(i) *There exists $c_0 \in G$ such that μ is a $c_0{}^{-1}Lc_0$-invariant measure supported on $\pi(L)c_0$. In particular, μ is a homogeneous measure.*

(ii) *There exist sequences $\{\gamma_i\} \subset \Gamma$ and $c_i \to c_0$ in G such that $\gamma_i{}^{-1}H\gamma_i \subset L$ and $Hg_i = H\gamma_i c_i$ for all but finitely many $i \in \mathbb{N}$.*

Our proof of this theorem is based on the following observation:

PROPOSITION 1.8. *Let the notation be as in Theorem 1.7. Then either there exists a sequence $c_i \to c$ in G such that $\mu_i = \mu_H c_i$ for all $i \in \mathbb{N}$ (in which case $\mu = \mu_H c$), or μ is invariant under the action of a nontrivial unipotent one-parameter subgroup of G.*

Thanks to this proposition, we can apply the well-developed techniques of unipotent flows to study limit distributions of translates of homogeneous measures, and eventually to the problem of counting lattice points on homogeneous varieties. It is of interest to note that much of the motivation for the extensive study of the behavior of unipotent flows on homogeneous spaces came from number theory. The celebrated Oppenheim conjecture, which was proved by G. A. Margulis in [Mar1], concerning values of irrational indefinite quadratic forms at integer vectors, had led M. S. Raghunathan to formulate his conjecture concerning the closures of unipotent orbits. The Raghunathan conjecture as well as a measure theoretic version of it (conjectured by S. G. Dani and by G. A. Margulis) were proved in M. Ratner's seminal work [Rat1–4]. We refer the reader to the ICM addresses of Margulis [Mar2], Ratner [Rat5], and S. G. Dani [Dan2], and a survey article of Borel [Bor3] for a discussion of these and related questions.

In order to be able to apply Theorem 1.7 to the problem of counting, we need to know some conditions under which the sequence $\{\mu_H g_i\}$ of probability measures does not escape to infinity. Suppose further that G and H are reductive. Let $Z(H)$ be the centralizer of H in G. By rationality $\pi(Z(H))$ is closed in $\Gamma\backslash G$. Now if $\pi(Z(H))$ is noncompact, there exists a sequence $\{z_i\} \subset Z(H)$ such that $\{\pi(z_i)\}$ is divergent; that is, it has no convergent subsequence. Then $\mu_H z_i$ escapes to the infinity; that is, $\mu_H z_i(K) \to 0$ for any compact set $K \subset \Gamma\backslash G$. The condition that $\pi(Z(H))$ is noncompact is equivalent to the condition that H is contained in a proper parabolic $\mathbb{Q}$-subgroup of G. In the converse direction, we have the following (see [EMS]):

THEOREM 1.9. *Let G be a connected real reductive algebraic group defined over $\mathbb{Q}$, and H a connected real reductive $\mathbb{Q}$-subgroup of G, both admitting no nontrivial $\mathbb{Q}$-characters. Suppose that H is not contained in any proper parabolic $\mathbb{Q}$-subgroup of G defined over $\mathbb{Q}$. Let $\Gamma \subset G(\mathbb{Q})$ be an arithmetic lattice in G and π: $G \to \Gamma\backslash G$ the natural quotient map. Let μ_H denote the H-invariant probability measure on $\pi(H)$. Then given an $\varepsilon > 0$ there exists a compact set $K \subset \Gamma\backslash G$ such that $\mu_H g(K) > 1 - \varepsilon$, for all $g \in G$.*

The proof of this result uses generalizations of some results of Dani and Margulis [DM2]. Combining this theorem with Theorem 1.7, we deduce the following consequences:

COROLLARY 1.10. *Suppose that H is reductive and a proper maximal connected real algebraic $\mathbb{Q}$-subgroup of G. Then for any sequence $\{g_i\} \subset G$, if the sequence $\{Hg_i\}$ is divergent (that is, it has no convergent subsequence) in $H\backslash G$, then the sequence $\{\mu_H g_i\}$ gets equidistributed with respect to μ_G as $i \to \infty$ (that is, $\mu_H g_i \to \mu_G$ weakly).*

From this consequence, we obtain the following estimate regarding the counting problem stated in the beginning of the introduction:

THEOREM 1.11. *Let G and H be as in the counting problem. Suppose that H^0 is reductive and a maximal proper connected real algebraic $\mathbb{Q}$-subgroup of G, where H^0 denotes the connected component of identity in H. Then asymptotically as $T \to \infty$*

$$N(T,V,\mathcal{O}) \sim \lambda_{H\backslash G}(R_T).$$

Remark 1.12. Suppose that H is the set of fixed points of an involution of G. Let L be a connected real reductive $\mathbb{Q}$-subgroup of G containing H^0. Then there exists a normal $\mathbb{Q}$-subgroup N of G such that $L = H^0 N$. Now if G is $\mathbb{Q}$-simple, then H^0 is a maximal proper connected $\mathbb{Q}$-subgroup of G (see [Bor2, Lemma 8.0]); hence, Theorem 1.2 follows from Theorem 1.11.

In the general case, we obtain the following analogue of Corollary 1.10. We note that the condition that H is not contained in any proper $\mathbb{Q}$-parabolic subgroup of G, is also equivalent to saying that any real algebraic $\mathbb{Q}$-subgroup L of G containing H is reductive.

COROLLARY 1.13. *Let G be a connected real reductive algebraic group defined over $\mathbb{Q}$, and H a connected real reductive $\mathbb{Q}$-subgroup of G not contained in any proper parabolic $\mathbb{Q}$-subgroup of G. Let $\Gamma \subset G(\mathbb{Q})$ be an arithmetic lattice in G. Suppose that a sequence $\{g_i\} \subset G$ is such that the sequence $\{\mu_H \cdot g_i\}$ does not converge to the G-invariant probability measure. Then after passing to a subsequence, there exists a proper connected real reductive $\mathbb{Q}$-subgroup L of G containing H, and a compact set $C \subset G$ such that*

$$\{g_i\} \subset (Z(H) \cap \Gamma)LC.$$

For applying this result to the counting problem, we need to know that averages of translates of the measure μ_H along the sets R_T become equidistributed as T tends to infinity; i.e., we want the set of 'singular sequences', for which the limit measure is not G-invariant, to have negligible 'measure' in the sets R_T as $T \to \infty$. This does not hold when the sets R_T are 'focused' along $H\backslash L$ ($\subset H\backslash G$):

Definition 1.14. Let G and H be as in the counting problem. For a sequence $T_n \to \infty$, the sequence $\{R_{T_n}\}$ of open sets in $H\backslash G$ is said to be *focused*, if there exists a proper connected reductive real algebraic $\mathbb{Q}$-subgroup L of G containing H^0 and a compact set $C \subset G$ such that

$$\limsup_{n\to\infty} \frac{\lambda_{H\backslash G}(q_H((Z(H^0) \cap \Gamma)LC) \cap R_{T_n})}{\lambda_{H\backslash G}(R_{T_n})} > 0,$$

where $q_H \colon G \to H\backslash G$ is the natural quotient map.

Note that since L is reductive and defined over $\mathbb{Q}$, we have that $\pi(L)$ is closed in $\Gamma\backslash G$. In particular, $(Z(H^0) \cap \Gamma)L$ is closed in G. Also $H^0 zL = zL$ for any $z \in Z(H^0)$. Now since C is compact, the set $q_H((Z(H^0) \cap \Gamma)LC)$ is closed in $H\backslash G$.

Now if the focusing of $\{R_{T_n}\}$ does not occur, then using Corollary 1.13 we can obtain the following analogue of Corollary 1.10:

COROLLARY 1.15. *Let G and H be as in the counting problem. Suppose that H^0 is not contained in any proper $\mathbb{Q}$-parabolic subgroup of G^0, and for some sequence $T_n \to \infty$, the sequence $\{R_{T_n}\}$ is not focused. Then given $\varepsilon > 0$ there exists an open set $\mathcal{A} \subset H\backslash G$ with the following properties:*

$$\liminf_{n\to\infty} \frac{\lambda_{H\backslash G}(\mathcal{A} \cap R_{T_n})}{\lambda_{H\backslash G}(R_{T_n})} > 1 - \varepsilon \tag{2}$$

and given any sequence $\{g_i\} \subset {q_H}^{-1}(\mathcal{A})$, if the sequence $\{q_H(g_i)\}$ is divergent in $H\backslash G$, then the sequence $\{\mu_H g_i\}$ converges to μ_G.

This corollary allows us to obtain the counting estimates as in Theorems 1.2 and 1.11 for a large class of homogeneous varieties.

THEOREM 1.16. *Let G and H be as in the counting problem. Suppose that H^0 is not contained in any proper $\mathbb{Q}$-parabolic subgroup of G^0 (equivalently, $(Z(H) \cap \Gamma)\backslash Z(H)$ is compact), and for some sequence $T_n \to \infty$ with bounded gaps, the sequence $\{R_{T_n}\}$ is not focused. Then asymptotically*

$$N(T,V,\mathcal{O}) \sim \lambda_{H\backslash G}(R_T).$$

From the proof of this theorem we also obtain the following version on counting points on closed Γ-orbits on homogeneous varieties in place of integral points:

THEOREM 1.17. *Let G be a reductive real algebraic group defined over $\mathbb{Q}$ with no nontrivial $\mathbb{Q}$-characters and Γ an arithmetic lattice with respect to the $\mathbb{Q}$-structure on G. Let H be a reductive real algebraic $\mathbb{Q}$-subgroup of G such that $(Z(H^0) \cap \Gamma)\backslash Z(H^0)$ is compact. Suppose that G acts linearly on a Euclidean vector space W and there exists a point $p \in W$ such that the orbit $p \cdot G$ is closed and $H = \{g \in G\colon\ pg = p\}$. Define $B_T \subset W$ and $R_T \subset H\backslash G$ as in the counting problem. Suppose further that for some sequence $T_n \to \infty$ with bounded gaps, the sequence $\{R_{T_n}\}$ is not focused. Then asymptotically as $T \to \infty$,*

$$N(T,p\Gamma) = |p\Gamma \cap B_T| \sim \lambda_{H\backslash G}(R_T).$$

For Examples 1 and 2 considered above, in Section 6 we will illustrate the use of Theorem 1.16 by verifying the nonfocusing of $\{R_{T_n}\}$. We note that the nonfocusing assumption in Theorem 1.16 is not vacuous. In Section 7, we consider an example, where a sequence $\{R_{T_n}\}$ is focused and the counting estimate is different from what is predicted by Theorem 1.17.

In the above setup one is required to verify the condition of nonfocusing in Theorem 1.16 separately for each application of the result. From our examples, it seems that the process of computing volumes of R_T itself shows how to verify the focusing condition.

The paper is organized as follows: In Section 2 we prove Proposition 1.8, and describe the results of M. Ratner classifying measures invariant under unipotent flows. In Section 3, we loosely follow some ideas and methods developed by S. G. Dani and G. A. Margulis to study behavior of the translated orbits $\pi(H)g$ in $\Gamma\backslash G$ near images of certain algebraic subvarieties of G. The main result of this section is Proposition 3.13, which gives a condition in terms of an appropriate representation space which holds when the limit measure μ is not the G-invariant measure. In the course of completion of the proof of Theorem 1.7 in Section 4, we develop a general method to derive ergodic theoretic information from the above mentioned condition. The counting Theorem 1.16 is proved in Section 5.

We would like to thank M. Burger, S. G. Dani, G. A. Margulis, G. Prasad, M. Ratner, Z. Rudnick, and P. Sarnak for useful discussions. We also thank the referee for valuable suggestions leading to significant improvement in the readability of the paper. Parts of this work were carried at MSRI, University of Chicago, Princeton University, Institute for Advanced Study, and Stanford University. We would like to acknowledge their hospitality.

2. Invariance under unipotents and Ratner's theorem

First we make some reductions regarding the proof of Theorem 1.7: Let $c \in G$ be such that $\pi(c) \in \mathrm{supp}(\mu)$. Since $\mu_H g_i \to \mu$, there exist sequences $\{h_i\} \subset H$ and $\{\gamma_i\} \subset \Gamma$ such that ${\gamma_i}^{-1}{h_i}^{-1}g_i \to c$ as $i \to \infty$. Therefore, to prove the theorem we may assume, without loss of generality, that $g_i = \gamma_i$ for all $i \in \mathbb{N}$. Let L be a minimal connected real algebraic $\mathbb{Q}$-subgroup of G containing ${\gamma_i}^{-1}H\gamma_i$ for infinitely many $i \in \mathbb{N}$. By passing to a subsequence we may assume that ${\gamma_i}^{-1}H\gamma_i \subset L$ for all $i \in \mathbb{N}$. Again, without loss of generality, we can replace H by ${\gamma_1}^{-1}H\gamma_1$ and γ_i by ${\gamma_1}^{-1}\gamma_i$ for all $i \in \mathbb{N}$. Thus $H \subset L$. At this stage it is enough to prove the following result for $G = L$ and the homomorphisms ρ_i: $H \to L$ given by $\rho_i(h) = {\gamma_i}^{-1}h\gamma_i$ ($\forall h \in H$):

THEOREM 2.1. *Let G and H be connected real algebraic groups defined over $\mathbb{Q}$ and with no nontrivial $\mathbb{Q}$-characters. Let $\Gamma \subset G(\mathbb{Q})$ and $\Lambda \subset H(\mathbb{Q})$ be arithmetic lattices in G and H respectively. Let ρ_i: $H \to G$ ($i \in \mathbb{N}$) be a sequence of $\mathbb{Q}$-homomorphisms with the following properties:*

(1) *No proper $\mathbb{Q}$-subgroup of G contains $\rho_i(H)$ for infinitely many $i \in \mathbb{N}$.*

(2) *For every $h \in H(\mathbb{Q})$, there exists $k \in \mathbb{N}$ such that $\{\rho_i(h)\colon i \in \mathbb{N}\} \subset G(\frac{1}{k}\mathbb{Z})$.*

(3) *For any sequence $h_i \to e$ in H, all the eigenvalues for the action of $\mathrm{Ad}(\rho_i(h_i))$ on the Lie algebra of G tend to 1 as $i \to \infty$.*

(4) *For any regular algebraic function f on G, the functions $f \circ \rho_i$ span a finite dimensional space of functions on H.*

(5) *For all $i \in \mathbb{N}$, $\rho_i(\Lambda) \subset \Gamma$.*

Let π: $G \to \Gamma\backslash G$ be the natural quotient map and μ_G denote the G-invariant probability measure on $\Gamma\backslash G$. For each $i \in \mathbb{N}$, let μ_i denote the $\rho_i(H)$-invariant probability measure on $\pi(\rho_i(H))$. Then $\mu_i \to \mu_G$ weakly as $i \to \infty$.

Remark 2.1$'$. By property (5), the map $\pi \circ \rho_i\colon H \to \Gamma\backslash G$ factors through the canonical map $\pi_i\colon \Lambda\backslash H \to \Gamma\backslash G$ for all $i \in \mathbb{N}$. Let μ_H denote the H-invariant probability measure on $\Lambda\backslash H$. Then $\mu_i = (\pi_i)_*(\mu_H)$, the pushforward of μ_H under π_i for all $i \in N$.

It is straightforward to verify that the maps $\rho_i(h) = {\gamma_i}^{-1}h\gamma_i$ satisfy the properties (1)–(5); use [Bor2, Cor. 1.9, p. 54] to verify property (4).

Invariance under a unipotent flow. A basic observation in the proof of Theorem 2.1 is the following version of Proposition 1.8.

PROPOSITION 2.2. *Let the notation be as in Theorem* 2.1. *Assume that* $\rho_i(H) \neq G$ *for infinitely many* $i \in \mathbb{N}$. *Then* μ *is invariant under a nontrivial unipotent one-parameter subgroup of* G.

Proof. Let $\mathfrak{g}$ and $\mathfrak{h}$ denote the Lie algebras of G and H respectively. Let $D\rho_i\colon \mathfrak{h} \to \mathfrak{g}$ denote the differential of ρ_i at the identity. First suppose that, after passing to a subsequence, there exists a sequence $\{X_i\} \subset \mathfrak{h}$ and $Y \in \mathfrak{g}\backslash\{0\}$ such that as $i \to \infty$,

$$X_i \to 0 \qquad \text{and} \qquad D\rho_i(X_i) \to Y. \tag{3}$$

Let $t \in \mathbb{R}$. By property (3) of the maps ρ_i, $\mathrm{Ad}(\exp tY)$ is a unipotent transformation. Put $h_i = \exp(tX_i)$, for all $i \in \mathbb{N}$. Then as $i \to \infty$,

$$\mu_i = \mu_i\rho_i(h_i) = \mu_i \cdot \exp(D\rho_i(tX_i)) \to \mu \cdot \exp(tY).$$

Thus $\mu = \mu \exp(tY)$, and the conclusion of the proposition holds (cf. [Moz]).

If the condition in equation (3) does not hold, the set $\{D\rho_i(X)\colon\ i \in \mathbb{N}\}$ is relatively compact in $\mathfrak{g}$ for all $X \in \mathfrak{h}$.

Since H is a connected real algebraic group defined over $\mathbb{Q}$, by weak approximation (see [PR, Th. 7.7]), $H(\mathbb{Q})$ is dense in H. Hence there exists a finite set $S \subset \mathfrak{h}$ and the subgroup generated by $\exp(S)$ is Zariski dense in H.

Take any $h \in \exp(S)$. Then the set $D = \{\rho_i(h)\colon\ i \in \mathbb{N}\}$ is relatively compact, and by property (2) of the maps ρ_i, there exists $k \in \mathbb{N}$ such that $D \subset G(\frac{1}{k}\mathbb{Z})$. Since $G(\frac{1}{k}\mathbb{Z})$ is discrete, D is finite. Hence, by passing to a subsequence, we get that $\rho_i(h) = \rho_j(h)$ for all $i, j \in \mathbb{N}$ and $h \in \exp(S)$. Since the group generated by $\exp(S)$ is Zariski dense in H, this shows that $\rho_i = \rho_j$ for all i, j. Now by property (1) of the maps ρ_i, we have $\rho_i(H) = G$ for all $i \in \mathbb{N}$. This contradicts the hypothesis of the proposition, and the proof is complete. □

This proposition allows us to use the nice algebraic behavior of unipotent flows in our investigation.

Ratner's theorem on measure rigidity of unipotent flows. To study the measure μ as in Proposition 2.2, we reformulate Ratner's description [Rat3] of finite ergodic invariant measures for the actions of unipotent subgroups on homogeneous spaces of Lie groups.

Let G be a Lie group, Γ a discrete subgroup of G, and $\pi\colon G \to \Gamma\backslash G$ the natural quotient map. Let $\mathcal{H}$ be the collection of all closed subgroups F of G such that $F\cap\Gamma$ is a lattice in F and the subgroup generated by unipotent one-parameter subgroups of G contained in F acts ergodically on $\pi(F) \cong (F\cap\Gamma)\backslash F$ with respect to the F-invariant probability measure.

Note that for every $F \in \mathcal{H}$, $\mathrm{Ad}(F \cap \Gamma)$ is Zariski dense in $\mathrm{Ad}(F)$, where Ad denotes the Adjoint representation of G (see [MS, Prop. 2.1]).

Remark 2.3. Suppose that G is an algebraic $\mathbb{Q}$-group and $\Gamma \subset G(\mathbb{Q})$ is an arithmetic lattice in G. Then every $F \in \mathcal{H}$ is a $\mathbb{Q}$-subgroup of G and the radical of F is unipotent (see [Sha1, Prop. 3.2]).

PROPOSITION 2.4. *The collection $\mathcal{H}$ is countable.*

Proof. See [Rat3, Th. 1.1] or [DM3, Prop. 2.1] for different proofs of this result. □

Let U be a unipotent one-parameter subgroup of G and $F \in \mathcal{H}$. Define

$$\begin{aligned} N(F,U) &= \{g \in G\colon U \subset g^{-1}Fg\} \\ S(F,U) &= \bigcup\{N(F',U)\colon F' \in \mathcal{H},\ F' \subset F,\ \dim F' < \dim F\}. \end{aligned}$$

LEMMA 2.5 ([MS, Lemma 2.4]). *Let $g \in G$ and $F \in \mathcal{H}$. Then $g \in N(F,U)\backslash S(F,U)$ if and only if the group $g^{-1}Fg$ is the smallest closed subgroup of G which contains U and whose orbit through $\pi(g)$ is closed in $\Gamma\backslash G$. Moreover in this case the action of U on $\pi(F)g$ is ergodic with respect to a finite $g^{-1}Fg$-invariant measure.*

As a consequence of this lemma,

$$\pi(N(F,U) \setminus S(F,U)) = \pi(N(F,U)) \setminus \pi(S(F,U)), \qquad \text{for all } F \in \mathcal{H}. \tag{4}$$

Remark 2.6. Let G and Γ be as in Remark 2.3, and $g \in G(\mathbb{Q})$. Then $g \in N(F,U) \setminus S(F,U)$ if and only if the smallest $\mathbb{Q}$-subgroup of G containing U is $g^{-1}Fg$ (see [Sha1, Prop. 3.2]).

Ratner's theorem [Rat3] states that given any U-ergodic invariant probability measure on $\Gamma\backslash G$, there exists $F \in \mathcal{H}$ and $g \in G$ such that μ is $g^{-1}Fg$-invariant and $\mu(\pi(F)g) = 1$. Now decomposing any finite invariant measure into its ergodic component, and using Lemma 2.5, we obtain the following description for any U-invariant probability measure on $\Gamma\backslash G$ (see [MS, Th. 2.2]):

THEOREM 2.7 (Ratner). *Let U be a unipotent one-parameter subgroup of G and μ be a finite U-invariant measure on $\Gamma\backslash G$. For every $F \in \mathcal{H}$, let μ_F denote the restriction of μ on $\pi(N(F,U)\setminus S(F,U))$. Then μ_F is U-invariant and any U-ergodic component of μ_F is a $g^{-1}Fg$-invariant measure on the closed* *orbit* $\pi(F)g$ *for some* $g \in N(F,U) \setminus S(F,U)$. *In particular, for all Borel measurable subsets* A *of* $\Gamma\backslash G$,

$$\mu(A) = \sum_{F \in \mathcal{H}^*} \mu_F(A),$$

where $\mathcal{H}^* \subset \mathcal{H}$ *is a countable set consisting of one representative from each* Γ*-conjugacy class of elements in* $\mathcal{H}$.

3. Local behavior of translates of H-orbits

In this section we study the following situation: we have a relatively compact open set $\Omega \subset H$ with a probability measure ν such that the measure $\rho\nu$ on Ω extends to a Haar measure on H, where ρ is a continuous function which is bounded above and below by strictly positive constants on Ω. Given a sequence $\{g_i\} \subset G$, we have maps ϕ_i: $\Omega \to \Gamma\backslash G$ given by $\phi_i(h) = \pi(hg_i)$, for all $h \in \Omega$. For each $i \in \mathbb{N}$, let $\nu_i \in \mathcal{P}(\Gamma\backslash G)$ be such that $\nu_i(E) = \nu({\phi_i}^{-1}(E))$ for all Borel sets $E \subset \Gamma\backslash G$. Suppose that $\nu_i \to \mu$ in $\mathcal{P}(\Gamma\backslash G)$ and μ is invariant under some nontrivial unipotent one-parameter subgroup U of G.

We want to analyze the case when the limit measure μ is not the G-invariant measure. By Ratner's description of μ as in Theorem 2.7, there exists a proper subgroup $F \in \mathcal{H}$, $\varepsilon_0 > 0$, and a compact set $C_1 \subset N(F,U) \setminus S(F,U)$ such that $\mu(\pi(C_1)) > \varepsilon_0$. Thus for any neighborhood Φ of $\pi(C_1)$, we have $\nu(\phi_i^{-1}(\Phi)) > \varepsilon_0$ for all large $i \in \mathbb{N}$; that is, the image of Ω 'spends a fixed proportion of time in Φ'. To investigate the behavior of the maps ϕ_i near $\pi(N(F,U) \setminus S(F,U))$, we follow the methods of Dani and Margulis as in [DM3] (cf. [Sha1]). This involves constructing a linear representation of G associated with F as follows.

Linearization of neighborhoods of singular subsets. Let $F \in \mathcal{H}$. Let $\mathfrak{g}$ denote the Lie algebra of G and let $\mathfrak{f}$ denote its Lie subalgebra associated to F. For $d = \dim \mathfrak{f}$, put $V_F = \wedge^d \mathfrak{g}$, the d-th exterior power, and consider the linear G-action on V_F via the representation $\wedge^d \operatorname{Ad}$, the d-th exterior power of the Adjoint representation of G on $\mathfrak{g}$. Fix $p_F \in \wedge^d \mathfrak{f} \setminus \{0\}$, and let η_F: $G \to V_F$ be the map defined by $\eta_F(g) = p_F \cdot g = p_F \cdot (\wedge^d \operatorname{Ad} g)$ for all $g \in G$. Note that

$${\eta_F}^{-1}(p_F) = \{g \in N_G(F)\colon \det(\operatorname{Ad} g|_{\mathfrak{f}}) = 1\}.$$

PROPOSITION 3.1 ([DM3, Th. 3.4]). *The orbit* $p_F \cdot \Gamma$ *is discrete in* V_F.

Remark 3.2. In the arithmetic case described in Remark 2.3, the above proposition is immediate.

Now let U be a unipotent one-parameter subgroup of G. Let the notation be as in the previous section. First we recall a version of [DM3, Prop. 3.2].

PROPOSITION 3.3. *Let A_F be the linear span of $\eta_F(N(F,U))$ in V_F. Then*

$$\eta_F{}^{-1}(A_F) = N(F,U).$$

Put $\Gamma_F = N_G(F) \cap \Gamma$. Then for any $\gamma \in \Gamma_F$, we have $\pi(F)\gamma = \pi(F)$, and hence γ preserves the volume of $\pi(F)$. Therefore $|\det(\mathrm{Ad}\,\gamma|_{\mathfrak{f}})| = 1$; hence $p_F \cdot \gamma = \pm p_F$. Now define

$$\bar{V}_F = \begin{cases} V_F/\{\mathrm{Id},\text{-}\mathrm{Id}\} & \text{if } p_F \cdot \Gamma_F = \{p_F, -p_F\} \\ V_F & \text{if } p_F \cdot \Gamma_F = p_F. \end{cases}$$

The action of G factors through the quotient map of V_F onto $\bar{V}_F$. Let $\bar{p}_F$ denote the image of p_F in $\bar{V}_F$, and define $\bar{\eta}_F$: $G \to \bar{V}_F$ as $\bar{\eta}_F(g) = \bar{p}_F \cdot g$ for all $g \in G$. Then $\Gamma_F = \bar{\eta}_F{}^{-1}(\bar{p}_F) \cap \Gamma$. Let $\mathcal{A}_F$ denote the image of A_F in $\bar{V}_F$. Note that the inverse image of $\mathcal{A}_F$ in V_F is A_F.

For every $x \in \Gamma\backslash G$, define the set of representatives of x in $\bar{V}_F$ to be

$$\mathrm{Rep}(x) = \bar{\eta}_F(\pi^{-1}(x)) \subset \bar{V}_F.$$

Using Lemma 2.5 and Proposition 3.3, we get

$$\mathrm{Rep}(\pi(g)) \cap \mathcal{A}_F = \bar{p}_F \cdot \Gamma g \cap \mathcal{A}_F = \bar{p}_F \cdot g, \qquad \text{for all } g \in N(F,U) \setminus S(F,U).$$

We extend this observation in the following result (cf. [Sha1, Prop. 6.5]).

PROPOSITION 3.4 ([DM3, Cor. 3.5]). *Let D be a compact subset of $\mathcal{A}_F$. Then for any compact set $K \subset X \setminus \pi(S(F,U))$, there exists a neighborhood Φ of D in $\bar{V}_F$ such that any $x \in K$ has at most one representative in Φ.*

Using this proposition, we can uniquely represent in Φ the parts of the trajectories $\phi_i(\Omega)$ lying in K. In order to understand the behavior of these trajectories in Φ, we need to study certain growth properties of the following class of functions:

Certain growth properties for a class of functions.

Definition. For any $n \in \mathbb{N}$ and $\Lambda > 0$, let $E(n,\Lambda)$ be the collection of functions ϕ: $\mathbb{R} \to \mathbb{C}$ of the form

$$\phi(t) = \sum_{i=1}^{n} \sum_{l=0}^{n-1} a_{il} t^l e^{\lambda_i t}, \qquad \text{for all } t \in \mathbb{R},$$

where $a_{il} \in \mathbb{C}$ and $\lambda_i \in \mathbb{C}$ with $|\lambda_i| \leq \Lambda$ for all i.

Let $E_G(n,\Lambda)$ be the collection of functions θ: $\mathbb{R} \to G$ such that the following holds: for any $1 \leq d \leq \dim \mathfrak{g}$ and any $\mathbf{v} \in \wedge^d \mathfrak{g}$, if we define $\phi(t) = f(\mathbf{v} \cdot \theta(t))$

($\forall t \in \mathbb{R}$), where f is a linear functional on $\wedge^d \mathfrak{g}$, or if we define $\phi(t) = \|\mathbf{v} \cdot \theta(t)\|^2$ ($\forall t \in \mathbb{R}$), where $\|\cdot\|$ denotes a Euclidean norm, then $\phi \in E(n, \Lambda)$.

Take $m \in \mathbb{N}$. Let $E_G(m, n, \Lambda)$ be the set of functions Θ: $\mathbb{R}^m \to G$ such that for any $\mathbf{x} \in \mathbb{R}^m$ with $\|\mathbf{x}\| = 1$ and any $\mathbf{y} \in \mathbb{R}^m$, if we define $\theta(t) = \Theta(\mathbf{y} + t\mathbf{x})$, for all $t \in \mathbb{R}$, then $\theta \in E_G(n, \Lambda)$.

Remark 3.5. These functions arise in our study in the following context: Let $m \in \mathbb{N}$, and $\{X_1, \dots, X_m\} \subset G$. Define Θ: $\mathbb{R}^m \to G$ by

$$\Theta(\underline{t}) = \exp(t_1 X_1) \cdots \exp(t_m X_m) \qquad \text{for all } \underline{t} = (t_1, \dots, t_m) \in \mathbb{R}^m.$$

Then there exist $\Lambda > 0$ and $n \in \mathbb{N}$ such that for any $g \in G$, the map $\mathbf{t} \mapsto \Theta(\mathbf{t})g$ belongs to $E_G(m, n, \Lambda)$. Here Λ depends on the choice of $(X_i)_{1 \le i \le m}$ only up to an upper bound on the absolute values of eigenvalues of $\{\operatorname{Ad}\exp(X_i)\colon i = 1, \dots, m\}$; while n is independent of the choice.

The following growth property of functions in $E(n, \Lambda)$ plays an important role in the sequel; see [EMS, Cor. 2.10] for a proof.

PROPOSITION 3.6. *For any $n \in \mathbb{N}$ and $\Lambda > 0$, there exists a constant $\delta_0 = \delta_0(n,\Lambda) > 0$ such that the following holds: given $\varepsilon > 0$, there exists $M > 0$ such that for any $\phi \in E(n,\Lambda)$ and any interval I of length at most δ_0,*

$$|\{t \in I\colon\ |f(t)| < (1/M) \sup_{t \in I} |f(t)|\}| \le \varepsilon \cdot |I|.$$

Such growth properties for polynomials of bounded degrees were used in many works on dynamics of unipotent flows. The next result is a basic tool for our analysis.

Case of one-dimensional trajectories.

PROPOSITION 3.7. *Let $n \in \mathbb{N}$, $\Lambda \geq 0$, a compact set $C \subset \mathcal{A}_F$ and an $\varepsilon > 0$ be given. Then there exists a (larger) compact set $D \subset \mathcal{A}_F$ with the following property: For any neighborhood Φ of D in $\bar{V}_F$ there exists a neighborhood Ψ of C in $\bar{V}_F$ (with $\Psi \subset \Phi$) such that for any $\theta \in E_G(n,\Lambda)$, any $\mathbf{v} \in \bar{V}_F$, and any interval I of length at most $\delta_0(n,\Lambda)$, if $\overline{\mathbf{v} \cdot \theta(I)} \not\subset \Phi$, then*

$$|\{t \in I\colon\ \mathbf{v} \cdot \theta(t) \in \Psi\}| \le \varepsilon \cdot |\{t \in I\colon\ \mathbf{v} \cdot \theta(t) \in \Phi\}|. \tag{5}$$

Proof. Let $\mathcal{C}$ be a finite collection of linear functionals on V_F such that

$$A_F = \cap_{f \in \mathcal{C}} f^{-1}(0).$$

By Proposition 3.6, there exists $M > 0$ such that for any interval J of length at most $\delta_0(n, \Lambda)$ and $\psi \in E(n, \Lambda)$,

$$|\{t \in J\colon\ |\psi(t)| \le (1/M) \sup_{t \in J} |\psi(t)|\}| \ge \varepsilon \cdot |J|.$$

For $R > 0$, define $B(R) = \{\mathbf{v} \in \bar{V}_F\colon \|\mathbf{v}\|^2 < R\}$. Let $R > 0$ be such that $C \subset B(R)$. Put

$$D = \mathcal{A}_F \cap \overline{B(M \cdot R)}.$$

For $c > 0$, let $Z_c(\mathcal{C})$ be the image of $\{\mathbf{v} \in V_F\colon |f(\mathbf{v})| < c, \forall f \in \mathcal{C}\}$ in $\bar{V}_F$. Now given a neighborhood Φ of D, there exists $c > 0$ such that $Z_c(\mathcal{C}) \cap \overline{B(M \cdot R)} \subset \Phi$. Let

$$\Psi = Z_{c/M}(\mathcal{C}) \cap B(R).$$

Then Ψ is a neighborhood of C contained in Φ.

Fix any $\mathbf{v} \in V_F$, let $\bar{\mathbf{v}}$ be its image in $\bar{V}_F$. Let J be any connected component of $\{t \in I\colon \bar{\mathbf{v}} \cdot \theta(t) \in \Phi\}$. Suppose that $\overline{\bar{\mathbf{v}} \cdot \theta(I)} \not\subset \Phi$. Then there exists $a_1 \in \overline{J}$ such that $\bar{\mathbf{v}} \cdot \theta(a_1) \notin \Phi$. Therefore either $|f_0(\mathbf{v} \cdot \theta(a_1))| \geq c$ for some $f_0 \in \mathcal{C}$ or $\|\mathbf{v} \cdot \theta(a_1)\|^2 \geq M \cdot R$. Hence by the choice of $M > 0$ and since $\theta \in E_G(n, \Lambda)$,

$$\left|\{t \in J\colon |f(\mathbf{v} \cdot \theta(t))| < c/M \text{ and } \|\mathbf{v} \cdot \theta(t)\|^2 < R\}\right| \leq \varepsilon \cdot |J|.$$

From this, equation (5) follows. □

The following result is one of the main components of our proof of Theorem 1.7. Similar results for unipotent trajectories were obtained in [DM1], [Sha1], [DM3], and [MS].

PROPOSITION 3.8. *Let $n \in \mathbb{N}$, $\Lambda \geq 0$, $\varepsilon > 0$, a compact set $K \subset \Gamma\backslash G \setminus \pi(S(F,U))$, and a compact set $C \subset \mathcal{A}_F$ be given. Then there exists a (larger) compact set $D \subset \mathcal{A}_F$ with the following property: for any neighborhood Φ of D in $\bar{V}_F$, there exists a neighborhood Ψ of C in $\bar{V}_F$ (with $\Psi \subset \Phi$) such that for any $\theta \in E_G(n,\Lambda)$ and any interval $I \subset \mathbb{R}$ of length at most $\delta_0(n,\Lambda)$, one of the following conditions is satisfied:*

1. *There exists $\gamma \in \Gamma$ such that $\bar{p}_F \cdot \gamma\theta(I) \subset \Phi$.*

2. $|\{t \in I\colon \pi(\theta(t)) \in K$ *and* $\bar{p}_F \cdot \Gamma \cdot \theta(t) \cap \Psi \neq \emptyset\}| < \varepsilon \cdot |I|$.

Proof. Let a compact set $D \subset \mathcal{A}_F$ be as in Proposition 3.7. Let Φ be a given neighborhood of D in $\bar{V}_F$. We replace Φ by a smaller neighborhood of D, and by Proposition 3.4, the set $\mathrm{Rep}(x) \cap \Phi$ contains at most one element for all $x \in K$. By the choice of D there exists a neighborhood Ψ of C contained in Φ such that equation (5) holds.

Now put

$$E = \{t \in I\colon \pi(\theta(t)) \in K \text{ and } \bar{p}_F \cdot \Gamma \cdot \theta(t) \cap \Psi \neq \emptyset\}. \tag{6}$$

Let $t \in E$. By the choice of Φ, there exists a unique $\mathbf{v}_t \in \bar{p}_F \cdot \Gamma$ such that $\mathbf{v}_t \cdot \theta(t) \in \Phi$. Now suppose that condition 1 does not hold. Then for every $t \in E$, there exists a largest open interval $I(t) \subset I$ containing t such that

$$\mathbf{v}_t \cdot \theta(I(t)) \subset \Phi \text{ and } \overline{\mathbf{v}_t \cdot \theta(I(t))} \not\subset \Phi. \tag{7}$$

Put $\mathcal{I} = \{I(t)\colon\ t \in E\}$. Then for any $I_1 \in \mathcal{I}$ and $s \in I_1 \cap E$, we have $I(s) = I_1$. Therefore for any $t_1, t_2 \in E$, if $t_1 < t_2$ then either $I(t_1) = I(t_2)$ or $I(t_1) \cap I(t_2) \subset (t_1, t_2)$. Hence any $t \in I$ is contained in at most two distinct elements of $\mathcal{I}$. Thus

$$\sum_{I_1 \in \mathcal{I}} |I_1| \leq 2|I|. \tag{8}$$

Now by equations (5) and (7), for any $t \in E$,

$$|\{s \in I(t)\colon\ \mathbf{v}_t \cdot \theta(s) \in \Psi\}| < \varepsilon \cdot |I(t)|. \tag{9}$$

Therefore by equations (8) and (9), we get

$$|E| \leq \varepsilon \cdot \sum_{I_1 \in \mathcal{I}} |I_1| \leq (2\varepsilon)|I|,$$

which is condition 2 for 2ε in place of ε. □

Case of higher dimensional trajectories. To obtain a higher dimensional analogue of Proposition 3.8, we need the following results:

Let $m \in \mathbb{N}$. Let S be the unit sphere in $\mathbb{R}^m$ centered at 0 and let σ be a rotation invariant measure on S.

LEMMA 3.9. *Given $\varepsilon > 0$ there exists $\varepsilon_1 > 0$ such that for any measurable subset $A \subset S$ with $\sigma(A) < \varepsilon_1 \cdot \sigma(S)$, any ball $B \subset \mathbb{R}^m$, and any $\mathbf{b}_0 \in B$,*

$$|\{\mathbf{b}_0 + t \cdot \mathbf{x} \in B\colon\ \mathbf{x} \in A, t \geq 0\}| < \varepsilon \cdot |B|.$$

LEMMA 3.10. *Given $\varepsilon > 0$ there exists $\kappa \in (0,1)$ such that the following holds: For a ball $B \subset \mathbb{R}^m$ and a measurable subset $E \subset B$, if $|E| \geq \kappa \cdot |B|$ then for any $\mathbf{b}_0 \in B$,*

$$\sigma(\{\mathbf{x} \in S\colon\ (\mathbf{b}_0 + \mathbb{R} \cdot \mathbf{x}) \cap E \neq \emptyset\}) \geq (1 - \varepsilon)\sigma(S).$$

Proofs of Lemmas 3.9 *and* 3.10. Observe that it is enough to prove the results for a unit ball B, for which the assertions are easily verified.

The following result is a generalization of Lemma 3.10:

LEMMA 3.11. *Given $\varepsilon > 0$ there exists $\eta = \eta(m,\varepsilon) \in (0,1)$ such that the following holds: For a ball B in $\mathbb{R}^m$ and nonempty measurable subsets*

E,$F \subset B$, if $|E \cup F| \geq (1-\eta)|B|$ and $|E| \geq |F|$, then there exists $\mathbf{b}_0 \in F$ such that

$$\sigma(\{\mathbf{x} \in S\colon\ (\mathbf{b}_0 + \mathbb{R} \cdot \mathbf{x}) \cap E \neq \emptyset\}) \geq (1-\varepsilon)\sigma(S). \tag{10}$$

Proof. Let $\kappa \in (0,1)$ be as in Lemma 3.10 for the given $\varepsilon > 0$. Put $\eta = 3^{-(m+1)}(1-\kappa)$. Let E and F be nonempty measurable subsets of B such that $|E \cup F| \geq (1-\eta)|B|$ and $|E| \geq |F|$.

If $|E|/|B| \geq \kappa$, then by Lemma 3.10, equation (10) follows for any $\mathbf{t}_0 \in F$. Therefore we can assume that $|E|/|B| \leq \kappa$. Recall that if m_E denotes the Lebesgue measure of $\mathbb{R}^m$ restricted to E then there exists a measurable subset $E^* \subset E$ such that $|E \setminus E^*| = 0$ and the Radon-Nikodym derivative of m_E with respect to the Lebesgue measure is 1 at all $\mathbf{t} \in E^*$ (see [Rud, Th. 8.6]). Thus for a sufficiently small ball B' centered at a point of E^*, $|E \cap B'| > \kappa|B'|$. However, for sufficiently large balls $B' \subseteq B$, $|E \cap B'| \leq \kappa|B'|$, because $|E| \leq \kappa|B|$. Therefore there exists a covering $\mathcal{C}$ of E^* consisting of balls $B' \subset B$ such that

$$|E \cap B'|/|B'| = \kappa. \tag{11}$$

First suppose that $F \cap B' \neq \emptyset$ for some $B' \in \mathcal{C}$. Then for any $\mathbf{t}_0 \in F \cap B'$, equation (10) follows from equation (11) and Lemma 3.10, applied to B' in place of B.

Now suppose that $F \cap B' = \emptyset$, for all $B' \in \mathcal{C}$. Put $D = B \setminus (E \cup F)$. Then due to equation (11),

$$|D \cap B'|/|B'| = 1-\kappa, \qquad \text{for all } B' \in \mathcal{C}. \tag{12}$$

By [Rud, Lemma 8.4], there exists a finite subcollection $\mathcal{C}' \subset \mathcal{C}$ consisting of disjoint balls such that

$$\sum_{B' \in \mathcal{C}'} |B'| > 3^{-m} \cdot |E|.$$

Hence using equation (12), since $|E| \geq ((1-\eta)/2)|B|$, we get

$$|D| > (1-\kappa)3^{-m} \cdot |E| \geq (1-\kappa)3^{-(m+1)} \cdot |B| = \eta \cdot |B| \geq |B| - |E \cup F| = |D|,$$

which is a contradiction and the proof is complete. □

In the next proposition we obtain the basic property regarding the dynamical behavior of $E_G(m,n,\Lambda)$-type trajectories (cf. [Sha2, Prop. 5.4] in the case of polynomials of several variables).

PROPOSITION 3.12. *Let $m \in \mathbb{N}$, $n \in \mathbb{N}$, $\Lambda \geq 0$, $\varepsilon > 0$, a compact set $K \subset \Gamma\backslash G \setminus \pi(S(F,U))$, and a compact set $C \subset \mathcal{A}_F$ be given. Then there exists a (larger) compact set $D \subset \mathcal{A}_F$, and there exists $\eta > 0$ depending on ε and m, so that for any neighborhood Φ of D in $\bar{V}_F$, there exists a neighborhood Ψ*

of C contained in Φ such that for any $\Theta \in E_G(m,n,\Lambda)$ and a ball $B \subset \mathbb{R}^m$ of diameter at most $\delta_0(n,\Lambda)$ (see Prop. 3.6), one of the following possibilities holds:

1. *Let $B_K = \{\mathbf{t} \in B\colon \pi(\Theta(\mathbf{t})) \in K\}$. Then $|B_K| < (1-\eta)\cdot|B|$.*
2. *There exists $\gamma \in \Gamma$ such that $\bar{p}_F \cdot \gamma\Theta(B) \subset \Phi$.*
3. $|\{\mathbf{t} \in B\colon \operatorname{Rep}(\pi(\Theta(\mathbf{t}))) \cap \Psi \neq \emptyset\}| < \varepsilon \cdot |B|$.

Proof. For the given compact sets C and K, let D be a compact subset of $\mathcal{A}_F$ as in Proposition 3.8 for $\varepsilon/(4m)$ in place of ε. Let Φ be a given neighborhood of D. We replace Φ by a smaller neighborhood of D; thus by Proposition 3.4, the set $\operatorname{Rep}(x) \cap \Phi$ contains at most one element for any $x \in K$. Let $\Psi \subset \Phi$ be a neighborhood of C in $\bar{V}_F$ such that the conclusion of Proposition 3.8 holds for $\varepsilon/(4m)$ in place of ε.

Let $\varepsilon_1 > 0$ be such that the conclusion of Lemma 3.9 holds for $\varepsilon/2 > 0$ in place of ε. Let $\eta \in (0, \varepsilon/4)$ be such that the conclusion of Lemma 3.11 holds for ε_1 in place of ε.

Take a $\Theta \in E_G(m,n,\Lambda)$ and a ball $B \subset \mathbb{R}^m$ of diameter at most $\delta_0(n,\Lambda)$ (see Prop. 3.6). For every $\gamma \in \Gamma$,

$$A_\gamma = \{\mathbf{t} \in B_K\colon \bar{p}_F \cdot \gamma\Theta(\mathbf{t}) \in \Phi\}.$$

Note that for $\gamma_1, \gamma_2 \in \Gamma$ if $\bar{p}_F \cdot \gamma_1 \neq \bar{p}_F \cdot \gamma_2$ then $A_{\gamma_1} \cap A_{\gamma_2} = \emptyset$.

There are three cases:

(1) There exists some $\mathbf{t}_0 \in B_K \setminus \cup_{\gamma\in\Gamma} A_\gamma$.

(2) $A_{\gamma_0} = B_K$ for some $\gamma_0 \in \Gamma$.

(3) There exist $\gamma_1, \gamma_2 \in \Gamma$ such that $\bar{p}_F\gamma_1 \neq \bar{p}_F\gamma_2$, $A_{\gamma_i} \neq \emptyset$ and $|A_{\gamma_1}| \leq |A_{\gamma_2}|$.

Define

$$F = \begin{cases} \{\mathbf{t}_0\}, & \text{in case (1)} \\ \{\mathbf{t} \in B\colon \bar{p}_F \cdot \gamma_0\Theta(\mathbf{t}) \notin \Phi\}, & \text{in case (2)} \\ A_{\gamma_1}, & \text{in case (3).} \end{cases}$$

Put $E = B_K \setminus F$. Note that $B_K \cap F = \emptyset$ in Case (2). Now suppose that possibilities 1 and 2 of the proposition do not hold. Then

$$|E \cup F| \geq |B_K| \geq (1-\eta)|B|$$

and $F \neq \emptyset$. Also $|F| \leq |E|$. Therefore by Lemma 3.11, there exists $\mathbf{t}_0 \in F$ such that if we define

$$\Sigma = \{\mathbf{x} \in S\colon (\mathbf{t}_0 + \mathbb{R}\cdot\mathbf{x}) \cap E \neq \emptyset\},$$

then

$$|\Sigma| \geq (1-\varepsilon_1)|S|. \tag{13}$$

For every $\mathbf{x} \in S$, define $B_{\mathbf{x}} = \{t \in \mathbb{R}\colon t{\cdot}\mathbf{x}{+}\mathbf{t}_0 \in B\}$. Then by equation (13) and Lemma 3.9,

$$\Big|\bigcup_{\mathbf{x}\in S\setminus\Sigma} (\mathbf{t}_0 + B_{\mathbf{x}}\cdot\mathbf{x})\Big| \leq (\varepsilon/2)|B|. \tag{14}$$

Take any $\mathbf{x} \in \Sigma$. Apply Proposition 3.8 for the function $\theta(t) = \Theta(t{\cdot}\mathbf{x}{+}\mathbf{t}_0)$, for all $t \in \mathbb{R}$ and $I = B_{\mathbf{x}}$. Then by the definition of Σ, in each of the three cases we conclude that possibility 1 of Proposition 3.8 does not hold. Therefore according to possibility 2 of Proposition 3.8, if we define

$$A_{\mathbf{x}} = |\{t \in B_{\mathbf{x}}\colon \pi(\theta(t)) \in K \text{ and } \mathrm{Rep}(\pi(\theta(t))) \cap \Psi \neq \emptyset\}|,$$

then $|A_{\mathbf{x}}|/|B_{\mathbf{x}}| < \varepsilon/(4m)$. Put

$$A = \bigcup_{\mathbf{x}\in\Sigma} (\mathbf{t}_0 + A_{\mathbf{x}}\cdot\mathbf{x}).$$

Then A is a measurable subset of B. By the polar decomposition of B with $\mathbf{t}_0$ as a pole,

$$|A| \leq (\varepsilon/4)|B|. \tag{15}$$

Now possibility 3 follows from equations (14) and (15) and the assumption that possibility 1 of the proposition does not hold. This completes the proof. □

In the proof of Theorem 2.1, we need the following consequence of Proposition 3.12.

PROPOSITION 3.13. *Let $m,n \in \mathbb{N}$ and $\Lambda > 0$ be given. Let B be a ball of diameter at most $\delta_0 = \delta_0(n,\Lambda)$ (see Prop. 3.6) in $\mathbb{R}^m$ around 0. Let $\Theta_i\colon \mathbb{R}^m \to G$ ($i \in \mathbb{N}$) be mappings in $E(m,n,\Lambda)$ and λ_i be the probability measure on $\pi(\Theta_i(B))$ which is the pushforward under $\pi\circ\Theta_i$ of the normalized Lebesgue measure on B. Suppose that $\lambda_i \to \lambda$ weakly in the space of probability measures on $\Gamma\backslash G$. Suppose there exist a unipotent one-parameter subgroup U of G and $F \in \mathcal{H}$ such that $\lambda(\pi(N(F,U))) > 0$ and $\lambda(\pi(S(F,U)))$. Then there exists a compact set $D \subset \mathcal{A}_F$ such that the following holds: For any sequence of neighborhoods $\{\Phi_i\}$ of D in $\bar{V}_F$, there exists a sequence $\{\gamma_i\} \subset \Gamma$ such that for all large $i \in \mathbb{N}$,*

$$\bar{p}_F \cdot \gamma_i\Theta_i(B) \subset \Phi_i. \tag{16}$$

Proof. Choose a compact set $C_1 \subset N(F,U) \setminus S(F,U)$ and an $\varepsilon > 0$ such that $\lambda(\pi(C_1)) > \varepsilon$. Let $\eta > 0$ be as in the statement of Proposition 3.12. There exists a compact set $K \subset X \setminus \pi(S(F,U))$ such that

$$\lambda(K) > 1 - \eta/2.$$

Put $C = \bar{\eta}_F(C_1) \subset \mathcal{A}_F$. Then there exists a compact set D of $\mathcal{A}_F$ containing C with the property that given a neighborhood Φ of D in $\bar{V}_F$ there exists a neighborhood Ψ of C in $\bar{V}_F$ contained in Φ such that the conclusion of Proposition 3.12 holds.

By the choice of K, there exists $i_0 \in \mathbb{N}$ such that possibility 1 of Proposition 3.12 does not hold for Θ_i, for all $i \geq i_0$. Now possibility 3 of Proposition 3.12 applied to Θ_i, for any $i \geq i_0$, says that

$$|\{\mathbf{t} \in B\colon\ \pi(\Theta_i(\mathbf{t})) \in \pi({\bar{\eta}_F}^{-1}(\Psi))\}| < \varepsilon.$$

Since $\pi({\bar{\eta}_F}^{-1}(\Psi))$ is a neighborhood of $\pi(C_1)$, this means $\lambda_i(\pi(C_1)) \leq \varepsilon$, which is a contradiction for all large $i \in \mathbb{N}$.

Thus for any decreasing sequence $\{\Phi_i\}$ of neighborhoods of D in $\bar{V}_F$ such that $\cap_{i\in\mathbb{N}}\Phi_i = D$, the possibility 2 of Proposition 3.12 holds for Θ_i and Φ_i for all large $i \in \mathbb{N}$. Hence there exists a sequence $\{\gamma_i\} \subset \Gamma$ such that equation (16) holds. □

Thus we can reduce the study of limits of translates of the measure μ_Ω to a simpler situation in terms of finite dimensional linear representations. We shall carry out further analysis of the condition given by equation (16) in the next section.

4. Proof of Theorem 2.1

Let $m = \dim(H)$ and $\{X_1, \ldots, X_m\}$ be a basis of $\mathfrak{h}$. Define a map $\Theta\colon \mathbb{R}^m \to H$ by

$$\Theta(t_1, \ldots, t_m) = \exp(t_1 X_1) \cdots \exp(t_m X_m), \qquad \text{for all } (t_1, \ldots, t_m) \in \mathbb{R}^m.$$

For any $h \in H$ and $i \in \mathbb{N}$, define $\Theta_i^h\colon \mathbb{R}^m \to G$ as $\Theta_i^h(\mathbf{t}) = \rho_i(h\Theta(\mathbf{t}))$, for all $\mathbf{t} \in \mathbb{R}$. By property (3) of the maps ρ_i, it follows that absolute values of all the eigenvalues of $\operatorname{Ad}\exp(D\rho_i X_k) = \operatorname{Ad}\rho_i(\exp X_k)$ are bounded above by a positive constant independent of $i \in \mathbb{N}$ and $1 \leq k \leq m$. Therefore by Remark 3.5, there exist $n \in \mathbb{N}$ and $\Lambda > 0$ such that $\Theta_i^h \in E_G(m, n, \Lambda)$, for all $i \in \mathbb{N}$. Let B be a ball of diameter at most $\delta_0(n, \Lambda)$ in $\mathbb{R}^m$ around 0 (see Prop. 3.6). Let λ_i^h denote the probability measure on $\pi(\Theta_i^h(B))$ which is the pushforward under $\pi \circ \Theta_i^h$ of a multiple of the Lebesgue measure restricted to B.

It will be convenient to recall a result from [EMS, Th. 3.5] which is deduced from [DM3, Th. 6.1].

THEOREM. *Given $m \in \mathbb{N}$, $n \in \mathbb{N}$, $\Lambda > 0$, a compact set $K \subset \Gamma\backslash G$, $R > 0$, and $\varepsilon > 0$, there exists a larger compact set $K' \subset \Gamma\backslash G$ such that for* *any* $\phi \in E_G(m,n,\Lambda)$ *and a ball* B *of diameter at most* R *in* $\mathbb{R}^m$, *one of the following holds*:

1. $\pi(\phi(B)) \cap K = \emptyset$.
2. $|\{\mathbf{t} \in B\colon\ \pi(\phi(\mathbf{t})) \in K'\}| \geq (1-\varepsilon)|B|$.

Note that for each $h \in H$, there exists a finite sequence $\{h_j\}_{j=0}^{r} \subset H$ such that $h_0 = e$, $h_r = h$, and $|\{\mathbf{t} \in B\colon\ h_{j-1}\Theta(\mathbf{t}) \in h_j\Theta(B)\}| > \varepsilon|B|$ for $1 \leq j \leq r$. Now for each $i \in \mathbb{N}$, we apply the above Theorem for $\phi = \Theta_i^{h_j}$ and $K = K_j$ to obtain $K_{j+1} = K'$ such that $\lambda_i^{h_j}(K_{j+1}) > 1-\varepsilon$, where $K_0 = \{\pi(e)\}$ and $0 \leq j \leq r$. In particular, $\lambda_i^h(K_{r+1}) > 1-\varepsilon$ for all $i \in \mathbb{N}$.

Any compact subset of H is covered by finitely many sets of the form $h\Theta(B)$, where $h \in H$. Therefore by Remark 2.1′ there exists a compact set $K' \subset \Gamma\backslash G$ such that $\mu_i(K') > 1-\varepsilon$ for all $i \in \mathbb{N}$. Now by passing to a subsequence, there exists a probability measure μ on $\Gamma\backslash G$ such that $\mu_i \to \mu$ weakly as $i \to \infty$.

By Proposition 2.2, μ is invariant under a nontrivial unipotent one-parameter subgroup, say U, of G. By a version of Ratner's Theorem as in Theorem 2.7, there exists $F \in \mathcal{H}$ such that $\mu(\pi(N(F,U))) > 0$ and $\mu(\pi(S(F,U))) = 0$.

Since $H(\mathbb{Q})$ is dense in H, a given compact subset of H can be covered by finitely many open sets of the form $h\Theta(B)$ with $h \in H(\mathbb{Q})$. Therefore by passing to a subsequence, there exists $h \in H(\mathbb{Q})$ such that $\{\lambda_i^h\}$ converges weakly to a probability measure λ^h and $\lambda^h(\pi(N(F,U))) > 0$.

There exists a constant $c > 0$ such that for any Borel measurable set $A \subset \Gamma\backslash G$, we have $\lambda_i^h(A) < c\mu_i(A)$ ($\forall i \in \mathbb{N}$). Hence λ^h is absolutely continuous with respect to μ. Therefore $\lambda^h(\pi(S(F,U))) = 0$.

Now by Proposition 3.13, there exists a compact set $D \subset \mathcal{A}_F$, a decreasing sequence of relatively compact neighborhoods Φ_i of D in $\bar{V}_F$ with $\bigcap_{i=1}^{\infty}\Phi_i = D$, and a sequence $\{\gamma_i\} \subset \Gamma$ such that for all $i \in \mathbb{N}$,

$$\bar{p}_F \cdot \gamma_i\Theta_i^h(B) \subset \Phi_i.$$

Let $\Omega' = h\Theta(B) \cap H(\mathbb{Q})$. Let $\mathcal{S}$ be the collection of all real valued functions on Ω' of the form

$$\omega \mapsto f(p_F \cdot \gamma\rho_i(\omega)),$$

where $\gamma \in G$, $i \in \mathbb{N}$, and f is a real linear functional on V_F which factors through $V_F \to \bar{V}_F$. By property (4) of the maps ρ_i, $\mathcal{S}$ spans a finite dimensional space of functions on Ω'. Therefore there exists a finite set $\Sigma \subset \Omega'$ such that for any $\phi \in \mathcal{S}$, if $\phi(\Sigma) = 0$ then $\phi(\Omega') = 0$.

Now for any $s \in \Sigma$ and $i \in \mathbb{N}$,

$$\bar{p}_F \cdot \gamma_i\rho_i(s) \in \bar{p}_F \cdot \gamma_i\Theta_i^h(B) \subset \Phi_1.$$

By property (2) of the maps ρ_i, there exists $k \in \mathbb{N}$ such that $\{\rho_i(s)\colon i \in \mathbb{N}, s \in \Sigma\} \subset G(\frac{1}{k}\mathbb{Z})$. Since $\bar{p}_F\gamma \cdot G(\frac{1}{k}\mathbb{Z})$ is a discrete subset of $\bar{V}_F$ and Φ_1 is bounded, by passing to a subsequence, we may assume that

$$\bar{p}_F \cdot \gamma_i\rho_i(s) = \bar{p}_F\gamma_1 \cdot \rho_1(s), \qquad \text{for all } i \in \mathbb{N},\ s \in \Sigma.$$

Therefore by the choice of Σ, for all $\omega \in \Omega'$,

$$\bar{p}_F\gamma_i \cdot \rho_i(\omega) = \bar{p}_F\gamma_1 \cdot \rho_1(\omega), \qquad \text{for all } i \in \mathbb{N}. \tag{17}$$

Since $H(\mathbb{Q})$ is dense in H and $h\Theta(B)$ is open in H, equation (17) holds for all $\omega \in h\Theta(B)$.

Now putting $\omega = e$ in equation (17), we get $\bar{p}_F\gamma_i = \bar{p}_F\gamma_1$. Thus

$$\bar{p}_F\gamma_1 \cdot \rho_i(g) = \bar{p}_F\gamma_1 \cdot \rho_1(g) \subset \Phi_i, \text{ for all } i \in \mathbb{N}, g \in h\Theta(B). \tag{18}$$

Since $\cap_{i=1}^{\infty}\Phi_i = D$, we have that

$$\bar{p}_F\gamma_1 \cdot \rho_i(h\Theta(B)) \subset \mathcal{A}_F, \qquad \text{for all } i \in \mathbb{N}.$$

Therefore by Zariski density of $h\Theta(B)$ in H and Proposition 3.3,

$$\gamma_1\rho_i(H) \subset N(F,U), \qquad \text{for all } i \in \mathbb{N}.$$

Replacing F by $\gamma_1{}^{-1}F\gamma_1$, we obtain that

$$\rho_i(H) \subset N(F,U), \qquad \text{for all } i \in \mathbb{N}. \tag{19}$$

Take any $i \in \mathbb{N}$. Let $L \in \mathcal{H}$ be of minimal dimension such that $L \subset F$ and $\rho_i(H) \subset N(L,U)$. By definition

$$S(L,U) = \bigcup\{N(L_1,U)\colon L_1 \in \mathcal{H},\ L_1 \subset L,\ \dim L_1 < \dim L\}.$$

Since H is a connected real algebraic group, each $\rho_i(H)$ is an irreducible real subvariety of G. Therefore if $\rho_i(H) \subset S(L,U)$, then $\rho_i(H) \subset N(L_1,U)$ for some $L_1 \in \mathcal{H}$ with $L_1 \subset L$ and $\dim L_1 < \dim L$. This contradicts the choice of L. Hence $\rho_i(H) \not\subset S(L,U)$. Thus $\rho_i(H(\mathbb{Q}))\setminus S(L,U)$ is Zariski dense in $\rho_i(H)$.

Let $w \in \rho_i(H(\mathbb{Q}))\setminus S(L,U)$. By Remark 2.6, the smallest $\mathbb{Q}$-subgroup of G containing U is $\omega^{-1}L\omega$. Since $U \subset L$, by dimension considerations $w^{-1}Lw \subset L$. Thus $w \in N(L)$. Now by Zariski density,

$$\rho_i(H) \subset N(L). \tag{20}$$

Since L is the smallest $\mathbb{Q}$-subgroup of G containing U, it is determined independent of the choice of i. Therefore equation (20) holds for all $i \in \mathbb{N}$. Therefore by property (1) of the maps ρ_i, we have $G = N(L)$. In particular, $N(L,U) = G$. Since $L \subset F$ and $N(F,U) \setminus S(F,U) \neq \emptyset$, we have $L = F$. Now since $\mu(\pi(S(F,U)) = 0$ by Theorem 2.7, every U-ergodic component of μ is F-invariant. Hence μ is F-invariant.

Now we project everything onto $F\backslash G$, and use induction on $\dim(G)$ as follows. Put $\bar{G} = F\backslash G$. Then $\bar{G}$ is a connected real reductive group defined over $\mathbb{Q}$, and the quotient homomorphism ϕ: $G \to \bar{G}$ is defined over $\mathbb{Q}$. Put $\bar{\Gamma} = \phi(\Gamma)$. Then $\bar{\Gamma} \subset \bar{G}(\mathbb{Q})$ is an arithmetic lattice in $\bar{G}$. Put $\bar{\rho}_i = \phi \circ \rho_i$ for all $i \in \mathbb{N}$. Note that the properties (1)–(5) continue to hold for the maps $\bar{\rho}_i$. Let q: $\Gamma\backslash G \to \bar{\Gamma}\backslash\bar{G}$ and $\bar{\pi}$: $\bar{G} \to \bar{\Gamma}\backslash\bar{G}$ be the natural quotient maps. Then $q \circ \pi = \bar{\pi} \circ \phi$.

Let q_*: $\mathcal{P}(\Gamma\backslash G) \to \mathcal{P}(\bar{\Gamma}\backslash\bar{G})$ be the pushforward map induced by q from the space of (Borel) probability measures on $\Gamma\backslash G$ to that on $\bar{\Gamma}\backslash\bar{G}$. This map is continuous with respect to the topologies of weak-star convergence on both the spaces. Let $\bar{\mu}_i = q_*(\mu_i)$ and $\bar{\mu} = q_*(\mu)$. Then $\bar{\mu}_i$ is the $\bar{\rho}_i(H)$-invariant probability measure supported on $\bar{\pi}(\bar{\rho}_i(H))$, and $\bar{\mu}_i \to \bar{\mu}$.

Since $U \subset F$ and $\dim U > 0$, we have that $\dim \bar{G} < \dim G$. Therefore to prove the theorem by induction on $\dim G$, we can assume the validity of the theorem for $\bar{G}$ in place of G. Therefore $\bar{\mu}$ is the $\bar{G}$-invariant probability measure on $\bar{\Gamma}\backslash\bar{G}$. Since $F = \ker\phi$ and μ is F-invariant, we have that μ is G-invariant. This completes the proof of the theorem. $\square$

Remark 4.1. From the above proof, it follows that G contains a normal subgroup $F \in \mathcal{H}$ such that $G = HF$.

5. Applications to the counting problem

In this section we deduce the consequences of Theorems 1.7 and 1.9, and prove the counting Theorem 1.16. First we need some lemmas.

Let G, H, and L be connected real algebraic groups such that $H \subset L \subset G$. Define

$$Z(H,L) = \{g \in G\colon\ g^{-1}Hg \subset L\}.$$

In view of Theorem 1.7, we want to understand the set $Z(H,L) \cap \Gamma$, when G, H and L are defined over $\mathbb{Q}$.

LEMMA 5.1. *Suppose that at least one of G, L and H is reductive. Then $Z(H,L)$ is a union of finitely many closed double cosets of the form $Z(H)\cdot g \cdot L$, where $g \in Z(H,L)$.*

Proof. Observe that $Z(H,L)$ is a real subvariety of G. Therefore it has finitely many closed connected components. Take any $g \in Z(H,L)$. Put $H_1 = g^{-1}Hg$. Then $H_1 \subset L$ and $Z(H,L) = g \cdot Z(H_1,L)$. It is enough to show that $Z(H_1)L$ contains a neighborhood of e in $Z(H_1,L)$.

Let $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$ denote the Lie algebras of G, H_1, and L, respectively. Let $X \in \mathfrak{g}$ be a vector tangent to $Z(H_1,L)$ at e. Then

$$[X,\mathfrak{h}] \subset \mathfrak{l}. \tag{21}$$

If either G or L is reductive, let $\mathfrak{l}^{\perp}$ be the ortho-complement of $\mathfrak{l}$ in $\mathfrak{g}$ with respect to the Killing form. Otherwise, if H_1 is reductive, let $\mathfrak{l}^{\perp}$ be an $\mathrm{Ad}(H_1)$-invariant complementary subspace of $\mathfrak{l}$ in $\mathfrak{g}$. In both the cases,

$$\mathfrak{g} = \mathfrak{l} \oplus \mathfrak{l}^{\perp} \qquad \text{and} \qquad \mathrm{Ad}(H_1)(\mathfrak{l}^{\perp}) = \mathfrak{l}^{\perp}.$$

Therefore,

$$[\mathfrak{l}, \mathfrak{h}] \subset \mathfrak{l} \qquad \text{and} \qquad [\mathfrak{l}^{\perp}, \mathfrak{h}] \subset \mathfrak{l}^{\perp}. \tag{22}$$

Write $X = X_1 + X_2$, where $X_1 \in \mathfrak{l}$ and $X_2 \in \mathfrak{l}^{\perp}$. Then by equations (21) and (22),

$$[X_2, \mathfrak{h}] = 0.$$

This shows that $X \in \mathfrak{l} + \mathfrak{z}$, where $\mathfrak{z}$ is the Lie algebra of $Z(H_1)$. Thus at the identity, the tangent space of $Z(H_1, L)$ is contained in the tangent space of $Z(H_1) \cdot L$. By definition $Z(H_1)L \subset Z(H_1, L)$. Therefore $Z(H_1)L$ contains a neighborhood of e in $Z(H_1, L)$. This completes the proof. □

LEMMA 5.2. *Let G, H, and L be connected reductive real algebraic groups defined over $\mathbb{Q}$ such that $H \subset L \subset G$. Let $\Gamma \subset G(\mathbb{Q})$ be an arithmetic lattice in G. Then there exists a finite set $D \subset Z(H,L) \cap \Gamma$ such that*

$$Z(H,L) \cap \Gamma = (Z(H) \cap \Gamma) \cdot D \cdot (L \cap \Gamma).$$

Proof. By [Bor1, Prop. 7.7], there exists a vector space V with a $\mathbb{Q}$-structure, a point $\mathbf{p} \in V(\mathbb{Q})$, and a rational representation of G on V such that the orbit $G \cdot \mathbf{p}$ is closed in V and $L = \{g \in G\colon\ g \cdot \mathbf{p} = \mathbf{p}\}$. By Lemma 5.1, $A = Z(H, L)\mathbf{p}$ is a union of finitely many closed $Z(H)$-orbits. Since $Z(H)$ is a reductive real algebraic group defined over $\mathbb{Q}$, by [Bor1, Th. 9.11], the set $A \cap \Gamma \cdot \mathbf{p}$ is a finite union of $Z(H) \cap \Gamma$-orbits. Hence there exists a finite set $D \subset \Gamma$ such that

$$Z(H, L) \cap \Gamma L = (Z(H) \cap \Gamma)DL.$$

Now the result follows. □

Proof of Corollary 1.13. By Theorem 1.9, after passing to a subsequence, $\{\mu_H g_i\}$ converges weakly to a probability measure μ on $\Gamma\backslash G$. By Theorem 1.7, after passing to a subsequence, there exist a $\mathbb{Q}$-subgroup L' of G containing H and sequences $\{\gamma_i'\} \subset \Gamma \cap Z(H, L')$ and $c_i' \to c_0'$ such that $Hg_i = H\gamma_i'c_i'$, for all $i \in \mathbb{N}$. Since H is not contained in any proper $\mathbb{Q}$-parabolic subgroup of G, we have that L' is reductive. By Lemma 5.2, after passing to a subsequence, there exist $\gamma \in Z(H, L) \cap \Gamma$ and a sequence $\{\gamma_i\} \subset Z(H) \cap \Gamma$ such that $\gamma_i' = \gamma_i\gamma L'$ for all $i \in \mathbb{N}$. Put $L = \gamma L'\gamma^{-1}$, $c_0 = \gamma c_0'$ and $c_i = \gamma c_i'$ for all $i \in \mathbb{N}$. Then $H \subset L$,

$c_i \to c_0$, and

$$H\gamma_i{}^{-1}g_i = \gamma_i{}^{-1}Hg_i = \gamma_i{}^{-1}H\gamma_i'c_i' \subset H\gamma L'c_i' = Lc_i,$$

for all $i \in \mathbb{N}$. Hence $g_i \in \gamma_i L c_i$, for all $i \in \mathbb{N}$. □

Proof of Corollary 1.15. In view of Remark 4.1, let $\mathcal{L}$ be the collection of subgroups of G of the form HF, where $F \in \mathcal{H}$ and $H \subset N(F)$. Since $\mathcal{H}$ is countable, we can write $\mathcal{L} = \{L_i\}_{i\in\mathbb{N}}$. Let $\mathcal{C} = \{C_i\}_{i\in\mathbb{N}}$ be an increasing sequence of compact subsets of G such that $G = \cup_{i\in\mathbb{N}} C_i$. For every $k \in \mathbb{N}$, put

$$B_k = q_H\left(\bigcup_{i=1}^{k}(Z(H)\cap\Gamma)L_iC_i\right).$$

Since $\{R_{T_n}\}_{n\in\mathbb{N}}$ is not focused, for every $k \in \mathbb{N}$ there exists $n_k \in \mathbb{N}$ such that

$$\lambda_{H\backslash G}(B_k \cap R_{T_n})/\lambda_{H\backslash G}(R_{T_n}) < \varepsilon, \qquad \text{for all } n \geq n_k.$$

Therefore the set

$$\mathcal{A} = \cup_{k=1}^{\infty}(R_{n_{k+1}} \setminus B_k)$$

satisfies equation (2).

Let a sequence $\{g_i\} \subset q_H{}^{-1}(\mathcal{A})$ be such that $\{q_H(g_i)\}$ is divergent in $H\backslash G$. Then for every $n \in \mathbb{N}$, there exists $i_n \in \mathbb{N}$ such that $q_H(g_i) \notin R_{T_n}$ for every $i \geq i_n$. Therefore for any compact set $C \subset G$ and any $L \in \mathcal{L}'$,

$$\{g_i\}_{i\in\mathbb{N}} \not\subset (Z(H)\cap\Gamma)LC. \tag{23}$$

Hence by Corollary 1.11, μ is G-invariant. □

Correspondence between counting and translates of measures. We recall some observations from [DRS, Sect. 2]; see also [EM]. Let the notation be as in the counting problem stated in the introduction. For $T > 0$, define a function F_T on G by

$$F_T(g) = \sum_{\gamma\in(H\cap\Gamma)\backslash\Gamma} \chi_T(v_0 \cdot \gamma g),$$

where χ_T is the characteristic function of B_T. By construction F_T is left Γ-invariant, and hence it will be treated as a function on $\Gamma\backslash G$. Note that

$$F_T(e) = \sum_{\gamma\in(H\cap\Gamma)\backslash\Gamma} \chi_T(v_0 \cdot \gamma) = N(T, V, \mathcal{O}).$$

Since we expect, as in Theorem 1.2, that

$$N(T, V, \mathcal{O}) \sim \lambda_{H\backslash G}(R_T),$$

we define

$$\hat{F}_T(g) = \frac{1}{\lambda_{H\backslash G}(R_T)}F_T(g).$$

Thus Theorem 1.16 is the assertion

$$\hat{F}_T(e) \to 1 \qquad \text{as } T \to \infty.$$

The connection between Theorem 1.16 and Corollary 1.15 is the following proposition:

PROPOSITION 5.3. *Let the notation and conditions be as in Theorem* 1.16. *Then* $\hat{F}_{T_n} \to 1$ *in the weak-star topology on* $L^\infty(\Gamma\backslash G,\mu_G)$; *that is,* $\langle \hat{F}_{T_n},\psi\rangle \to \langle 1,\psi\rangle$ *for any compactly supported continuous function* ψ *on* $\Gamma\backslash G$.

Proof. As in [DRS, Sect. 2],

$$\langle \hat{F}_T, \psi\rangle = \frac{1}{\lambda_{H\backslash G}(R_T)} \int_{R_T} \overline{\psi^H}\, d\lambda_{H\backslash G},$$

where

$$\psi^H(Hg) = \int_{\Gamma\backslash\Gamma H} \psi(\Gamma hg)\, d\mu_H(\Gamma h) = \int_{\Gamma\backslash G} \psi\, d(\mu_H g)$$

is a function on $H\backslash G$.

Let $\varepsilon > 0$ be given. Since the sequence $\{R_{T_n}\}$ is not focused, we obtain a set $\mathcal{A} \subset H\backslash G$ as in Corollary 1.15. Break up the integral over R_{T_n} into the integrals over $R_{T_n} \cap \mathcal{A}$ and $R_{T_n} \setminus \mathcal{A}$. By equation (2) and the boundedness of ψ, the second integral is $O(\varepsilon)$. By Corollary 1.15, for any sequence $\{g_i\} \subset {q_H}^{-1}(\mathcal{A})$, if $\{q_H(g_i)\}$ has no convergent subsequence in $H\backslash G$, then $\mu_H{\cdot}g_i \to \mu_G$. Hence

$$\psi^H(Hg_i) \to \int_{\Gamma\backslash G} \psi\, d\mu_G = \langle \psi, 1\rangle.$$

We use dominated convergence theorem to justify the interchange of limits. Now

$$\begin{aligned}
\lim_{n\to\infty} \langle \hat{F}_{T_n}, \psi\rangle &= \lim_{n\to\infty} \frac{1}{\lambda_{H\backslash G}(R_{T_n})} \int_{R_{T_n}\cap\mathcal{A}} \overline{\psi^H}\, d\lambda_{H\backslash G} + O(\varepsilon)\\
&= \lim_{n\to\infty} \frac{1}{\lambda_{H\backslash G}(R_{T_n})} \int_{R_{T_n}\cap\mathcal{A}} \overline{\langle\psi,1\rangle}\, d\lambda_{H\backslash G} + O(\varepsilon)\\
&= \lim_{n\to\infty} \frac{\lambda_{H\backslash G}(R_{T_n}\cap\mathcal{A})}{\lambda_{H\backslash G}(R_{T_n})} \langle 1,\psi\rangle + O(\varepsilon)\\
&= \langle 1,\psi\rangle + O(\varepsilon).
\end{aligned}$$

Since ε is arbitrary, the proof is complete. □

PROPOSITION 5.4. *There are constants* $a(\delta)$ *and* $b(\delta)$ *tending to* 1 *as* $\delta \to 0$ *such that*

$$b(\delta) \le \liminf_{T\to\infty} \frac{\lambda_{H\backslash G}(R_{(1-\delta)T})}{\lambda_{H\backslash G}(R_T)} \le \limsup_{T\to\infty} \frac{\lambda_{H\backslash G}(R_{(1+\delta)T})}{\lambda_{H\backslash G}(R_T)} \le a(\delta).$$

For a proof, see Appendix A.

Proof of Theorem 1.16. Combining Propositions 5.3 and 5.4 exactly as in [DRS, Lemma 2.3], we obtain that $\hat{F}_{T_i} \to 1$ pointwise on $\Gamma\backslash G$ as $i \to \infty$. As we observed before, this completes the proof. □

6. Examples

Given a "counting problem", i.e., the problem of estimating the number of lattice points on a homogeneous variety $V = pG$ as described in the introduction, in order to apply Theorem 1.16 we need to:

1. Verify that the centralizer $Z(H^0)$ of the component of the identity of the stabilizer is anisotropic over $\mathbb{Q}$; that is, $\pi(Z(H^0))$ is compact.

2. Verify that for any sequence $T_i \to \infty$, the sequence $\{R_{T_i}\}$ in $H\backslash G$ of pull backs of balls of radii T_i around the origin is not focused (see Definition 1.14).

In this section we verify these conditions for the examples discussed in the introduction and complete the proofs of the counting estimates given there.

Example 1.

Proof of Theorem 1.3. Let D be the centralizer of v_0 in $\mathrm{M}_n(\mathbb{R})$. Since the eigenvalues of v_0 are distinct, D is an abelian algebra of dimension n. Also D has a natural $\mathbb{Q}$-structure, and hence $\mathbb{Q}$-dim$(D(\mathbb{Q})) = n$.

Let α be a root of $p(\lambda)$. Then $\{1, \alpha, \ldots, \alpha^{n-1}\}$ is a $\mathbb{Q}$-basis of the field $\mathbb{Q}(\alpha)$. The multiplication by any $x \in \mathbb{Q}(\alpha)$ can be expressed as a matrix C_x with respect to this basis. Since $C_\alpha = v_0$, we have that $C_x \in D(\mathbb{Q})$ for all $x \in \mathbb{Q}(\alpha)$. By dimension considerations, the linear map $C\colon \mathbb{Q}(\alpha) \to D(\mathbb{Q})$ is an isomorphism of algebras over $\mathbb{Q}$. Let Norm: $\mathbb{Q}(\alpha) \to \mathbb{Q}$ be the norm map. Then $\mathrm{Norm}(x) = \det(C_x)$ for all $x \in Q(\alpha)$.

Let H be the stabilizer of v_0 in G. Then $H = D \cap G$. Hence $H(\mathbb{Q}) = \ker(|\mathrm{Norm}|)$. Also $C^{-1}(H(\mathbb{Z}))$ is a subgroup of finite index in the group of units in $\mathbb{Q}(\alpha)$. Since the units in $\mathbb{Q}(\alpha)$ form a lattice in the group of unit norm elements, we have that $H(\mathbb{Z})\backslash H$ is compact. Hence H^0 is $\mathbb{Q}$-anisotropic. Since the dimension of H is $n-1$, we have that H^0 is a maximal connected real torus in G defined over $\mathbb{Q}$. Since a maximal $\mathbb{Q}$-torus of a proper $\mathbb{Q}$-parabolic subgroup admits nontrivial $\mathbb{Q}$-characters, H^0 is not contained in a proper $\mathbb{Q}$-parabolic subgroup of G.

Thus in order to apply Theorem 1.16 to complete to proof of the present theorem, it remains to compute the volume growth of the pullback sets R_T, and to verify that they are not focused. □

Reduction to the diagonal case. Suppose that $p(\lambda)$ has $r \geq 0$ pairs of complex conjugate roots, say $\lambda_1, \bar{\lambda}_1, \ldots, \lambda_r, \bar{\lambda}_r$, and $s \geq 0$ real roots, say $\mu_1, \ldots, \mu_s$. Then $2r + s = n$. Suppose that $\lambda_k = x_k + iy_k$, where $x_k, y_k \in \mathbb{R}$, $1 \leq k \leq r$. There exists $g_0 \in G$ such that

$$g_0^{-1} v_0 g_0 = v_1 = \operatorname{diag}\left(\begin{pmatrix} x_1 & y_1 \\ -y_1 & x_1 \end{pmatrix}, \ldots, \begin{pmatrix} x_r & y_r \\ -y_r & x_r \end{pmatrix}, \mu_1, \ldots, \mu_s\right). \tag{24}$$

The stabilizer H_1 of v_1 in G consists of elements of the form

$$h = \operatorname{diag}\left(t_1 \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{pmatrix}, \ldots, t_r \begin{pmatrix} \cos\theta_r & \sin\theta_r \\ -\sin\theta_r & \cos\theta_r \end{pmatrix}, t_{r+1}, \ldots, t_{r+s}\right),$$

where $\prod_{j=1}^{r+s} t_i = \pm 1$, $t_j \in \mathbb{R}^\times$, and $\theta_k \in \mathbb{R}$. Note that the map ϕ: $H_1\backslash G \to H\backslash G$, defined by $\phi(H_1 g) = H(g_0 g)$, for all $g \in G$, is G-equivariant. For $T > 0$, define

$$R_T^1 = \{H_1 g \in H_1\backslash G\colon\ \|g^{-1} v_1 g\| < T\}.$$

Then $R_T = \phi(R_T^1)$ and $\lambda_{H\backslash G}(R_T) = \operatorname{vol}_{H_1\backslash G}(R_T^1)$. Note that conjugation g_0 preserves the Haar measures on H and H_1.

Haar integral on $H_1\backslash G$. Let N be the group of elements of the form $n = (n_{ij})_{i,j=1,\ldots,r+s} \in \mathrm{SL}_n(\mathbb{R})$, where

$$n_{ij} \in \begin{cases} \mathrm{M}_{2\times 2}(\mathbb{R}) & \text{if } i \leq r,\ j \leq r \\ \mathrm{M}_{1\times 2}(\mathbb{R}) & \text{if } i > r,\ j \leq r \\ \mathrm{M}_{2\times 1}(\mathbb{R}) & \text{if } i \leq r,\ j > r \\ \mathrm{M}_{1\times 1}(\mathbb{R}) & \text{if } i > r,\ j > r, \end{cases} \tag{25}$$

$n_{ij} = 0$ if $i < j$ and n_{ii} is the identity matrix for each i.

Let

$$M = \{\operatorname{diag}(M_1, \ldots, M_r, t_1, \ldots, t_s) \in G\colon\ M_i \in \mathrm{GL}_2(\mathbb{R}),\ t_i \in \mathbb{R}^\times\}.$$

Then MN is a standard parabolic subgroup of G. Hence $G = MNK$, where $K = O(n)$.

Let A be the maximal $\mathbb{R}$-split torus in H_1. Then $H_1 = A(K \cap H_1)$ and $M \cong A \times \mathrm{SL}_2(\mathbb{R})^r$. Let

$$B = \{\operatorname{diag}(a_1, a_1^{-1}, \ldots, a_r, a_r^{-1}, 1, \ldots, 1) \in M\colon\ a_i > 0\}. \tag{26}$$

By Cartan decomposition of $\mathrm{SL}_2(\mathbb{R})$, we have $M = A(K \cap M)B(K \cap M)$. Since $K \cap M = K \cap H_1$, we get $M = H_1 B(K \cap M)$. Let dm be a Haar integral on M such that for any $f \in \mathrm{C_c}(M)$,

$$\int_M f\, dm = \int_{(h,b,k_1)\in H_1\times B\times(K\cap M)} f(hbk_1)\left(\prod_{i=1}^r a_i^2\right) dh\, db\, dk_1 \tag{27}$$

where dh, db, and dk_1 are Haar integrals on H_1, B and $K \cap M$, respectively, and b as in equation (26).

Since M normalizes N, we have that $G = H_1BNK$. Now let $\lambda_{H_1\backslash G}$ be the G-invariant measure on $H_1\backslash G$ such that for any $f \in \mathrm{C_c}(H_1\backslash G)$,

$$\int_{H_1\backslash G} f\, d\lambda_{H_1\backslash G} = \int_{B\times N\times K} f(H_1bnk)) \left(\prod_{i=1}^{r} a_i\right) da_1 \cdots da_r\, dn\, dk, \tag{28}$$

where dn and dk are Haar integrals on N and K, respectively, b as in equation (26), and each da_i is the Lebesgue integral on $\mathbb{R}$.

Coordinate description of R_T^1. Note that R_T^1 corresponds to the set

$$\{(b,n,k) \in B \times N \times K\colon\ \|n^{-1}b^{-1}v_1bn\| < T\}.$$

If we express

$$b^{-1}v_1b = \mathrm{diag}(z_1, \ldots, z_{r+s}),$$

then by equation (24),

$$z_i = \begin{cases} \begin{pmatrix} a_i{}^{-1} & 0 \\ 0 & a_i \end{pmatrix} \begin{pmatrix} x_i & y_i \\ -y_i & x_i \end{pmatrix} \begin{pmatrix} a_i & 0 \\ 0 & a_i{}^{-1} \end{pmatrix} & \text{for } 1 \le i \le r, \\ \mu_k & \text{for } i = r+k, \end{cases}$$

where $b = (a_1, a_1{}^{-1}, \ldots, a_r, a_r{}^{-1}, 1, \ldots, 1)$.

As in equation (25), for $n = (n_{ij}) \in N$, if we express $n^{-1} = (n'_{ij}) \in N$, then

$$n'_{ij} = -n_{ij} + f_{ij}, \qquad (i < j),$$

where f_{ij} depends only on $\{n_{kl}\colon\ k < l \text{ and } l - k < j - i\}$.

Now if we express $n^{-1}(b^{-1}v_1b)n = (w_{ij})_{i,j=1,\ldots,r+s}$, then

$$w_{ij} = \begin{cases} 0 & \text{if } i > j \\ z_i & \text{if } i = j \\ -n_{ij}z_j + z_in_{ij} + A_{ij} & \text{if } i < j, \end{cases}$$

where

(29)

$$A_{ij} \text{ depends only on } \{z_k\colon\ 1 \le k \le r+s\} \cup \{n_{kl}\colon\ k < l \text{ and } l-k < j-i\}.$$

Thus R_T^1 corresponds to the set

$$\{(b,n,k) \in B \times N \times K\colon \sum_{1\le i\le r+s} \|z_i\|^2 + \sum_{1\le i<j\le r+s} \|w_{ij}\|^2 < T^2\}. \tag{30}$$

Asymptotics of $\lambda_{H_1\backslash G}(R_T^1)$. For $i<j$, let L_{ij} denote the linear transformation which takes the matrix n_{ij} to the matrix $\tilde{n}_{ij} = -n_{ij}z_j + z_i n_{ij}$. Then the eigenvalues of L_{ij} are

$$\begin{array}{ll} \lambda_i-\lambda_j,\ \lambda_i-\bar{\lambda}_j,\ \bar{\lambda}_i-\lambda_j,\ \bar{\lambda}_i-\bar{\lambda}_j & \text{if } 1\le i<j\le r \\ \lambda_i-\mu_j,\ \bar{\lambda}_i-\mu_j & \text{if } i\le r<j \\ \mu_i-\mu_j & \text{if } r+1\le i<j. \end{array}$$

Hence the Jacobian of the transformation on N which sends $(n_{ij})_{i<j}$ to $(\tilde{n}_{ij})_{i<j}$ is

$$J=\prod_{1\le i<j\le r}|\lambda_i-\lambda_j|^2|\lambda_i-\bar{\lambda}_j|^2\cdot\prod_{i\le r<j}|\lambda_i-\mu_j|^2\cdot\prod_{r+1\le i<j}|\mu_i-\mu_j|.$$

Now $w_{ij}=\tilde{n}_{ij}+A_{ij}$ for all $i<j$. By (29), the Jacobian of the transformation which takes $(n_{ij})_{i<j}$ to $(w_{ij})_{i<j}$ is the constant J. Note that $\|z_i\|\sim|y_i|a_i^2$ as $a_i\to\infty$ for $1\le i\le r$ and $\|z_{r+k}\|=|\mu_k|$ for $1\le k\le s$. Now using (30) and the change of variables $(n_{ij})_{i<j}\mapsto(w_{ij})_{i<j}$, we get that asymptotically

$$\lambda_{H_1\backslash G}(R_T^1)\sim\frac{\omega_n}{J}\cdot T^{\dim B+\dim N},$$

where ω_n is the volume of the unit ball in $\mathbb{R}^{n(n-1)/2}$.

Nonfocusing of the sets R_T. Now let L be the component of identity of the $\mathbb{R}$-points of a proper connected (reductive) $\mathbb{Q}$-subgroup of G containing H^0. Put $L_1={g_0}^{-1}Lg_0$. Then $L_1\supset H_1$. From the structure of H_1, one verifies that each irreducible subspace for the adjoint action of H_1 on the Lie algebra of L_1 is in fact invariant under the adjoint action of M. Put $L_1'=M^0L_1$. Then L_1' is a proper reductive subgroup of G. Since M^0 contains the full diagonal subgroup of $G^0=\mathrm{SL}_n(\mathbb{R})$, we have that L_1' is the Levi part of a proper parabolic subgroup of $\mathrm{SL}_n(\mathbb{R})$. Hence L_1' is the Levi part of a standard parabolic subgroup associated to a permutation of the standard basis of $\mathbb{R}^n$. Therefore the simple components of L_1' are $\mathrm{SL}_m(\mathbb{R})$, where $m<n$, and H_1 intersects each one of them in a maximal torus. Therefore by the computation as above for each m in place of n, we conclude that for any compact set $C\subset G$,

$$\lambda_{H_1\backslash G}(L_1'C\cap R_T^1)=O(T^{\dim(B)+\dim(N\cap L_1)}).$$

Since $\dim L_1'<\dim(\mathrm{SL}_n(\mathbb{R}))$ and $B\subset L_1'$, we have that $N\not\subset L_1$; hence,

$$\dim(N\cap L_1)<\dim(N).$$

Therefore

$$\lim_{T\to\infty}\lambda_{H_1\backslash G}(L_1'C\cap R_T^1)/\lambda_{H_1\backslash G}(R_T^1)=0.$$

Hence for any compact set $C\subset G$,

$$\lim_{T\to\infty}\lambda_{H\backslash G}(LC\cap R_T)/\lambda_{H\backslash G}(R_T)=0.$$

Thus the sets R_T are not focused along any $\mathbb{Q}$-subgroup L of G containing H.

Applying the counting theorem. Now we can apply Theorem 1.16 to conclude that

$$N(T, V_p, \Gamma v_0) \sim \frac{\mathrm{vol}(H \cap \Gamma\backslash H)}{\mathrm{vol}(\Gamma\backslash G)} \mathrm{vol}(R_T).$$

Hence

$$N(T, V_p) \sim c_p T^{n(n-1)/2},$$

where $c_p > 0$ is an explicitly computable constant. This completes the proof of Theorem 1.3. □

As we mentioned earlier in the introduction, for the sake of simplicity we assume that all roots of $p(\lambda)$ are real and $\mathbb{Z}[\alpha]$ is the ring of integers in $\mathbb{Q}(\alpha)$, where $p(\alpha) = 0$, and give a formula for c_p.

Proof of Theorem 1.1. Let λ_G and λ_H be any Haar measures on G and H, respectively. Let $\lambda_{H\backslash G}$ be the G-invariant measure on $H\backslash G$ such that for any $f \in \mathrm{C_c}(G)$, if we put $\bar{f}(Hg) = \int_H f(hg)d\lambda_H(h)$ for all $Hg \in H\backslash G$, then $\bar{f} \in \mathrm{C_c}(H\backslash G)$ and

$$\int_G f\, d\lambda_G = \int_{H\backslash G} \bar{f} d\lambda_{H\backslash G}.$$

Similarly, the choice of λ_G and λ_H determines volume forms on $\Gamma\backslash G$ and $H(\mathbb{Z})\backslash H$. Let c_G and c_H denote the volumes of $\Gamma\backslash G$ and $H(\mathbb{Z})\backslash H$, respectively. In view of the normalizations of the Haar measures on G and H, as chosen in the counting Theorem 1.16, we obtain that asymptotically

$$N(T, V_p, v_0\Gamma) \sim (c_H {c_G}^{-1})\lambda_{H\backslash G}(R_T). \tag{31}$$

Choice of λ_G and λ_H. Let D denote the full diagonal subgroup of $\mathrm{GL}_n(\mathbb{R})$. Fix a Haar integral such that $dx = \prod_{i=1}^n dx_i/x_i$ on D, where $x = (x_1, \dots, x_n)$ and each dx_i is the Lebesgue integral on $\mathbb{R}$. Put $H_1 = D \cap G$. Fix a Haar integral da on H_1 such that $dx = t^{-1}dtda$, where $x = t^{1/n}a$, $t = |\det x|$, and $a \in H_1$. Let $K = O(n)$ and N be the group of upper triangular unipotent matrices in $\mathrm{GL}_n(\mathbb{R})$. Let dk be the Haar integral on K such that $\int_K 1dk = 1$. Let $dn = \prod_{i<j} dn_{ij}$ be the Haar integral on N, where $n = (n_{ij})$ and dn_{ij} is the Lebesgue integral on $\mathbb{R}$. In view of the Iwasawa decomposition $G = KNA$, the integral $dg = dkdnda$ is a Haar integral on G. We choose λ_G such that $d\lambda_G(g) = dg$ on G. Let $g_0 \in G$ such that $H_1 = {g_0}^{-1}Hg_0$. We choose λ_H such that under the $\mathrm{Ad}\, g_0$ action the Haar integral $d\lambda_H$ on H maps to the integral da on H_1.

Volume computations. Using [Terras, Ex. 21, Eqn. 4.40, Th. 4], we deduce that

$$c_G = 2^{-(n-1)} \prod_{k=2}^{n} \Lambda(k/2), \qquad \text{where } \Lambda(s) = \pi^{-s}\Gamma(s)\zeta(2s). \tag{32}$$

Since we assume that $\mathbb{Z}[\alpha]$ is the ring of integers in $\mathbb{Q}(\alpha)$, we have that $C^{-1}(H(\mathbb{Z}))$ is the group of units in $\mathbb{Q}(\alpha)$. By [Lang, Sect. V.1, p. 110], if R is the regulator of $\mathbb{Q}(\alpha)$, then

$$c_H = 2^n R. \tag{33}$$

Since all roots of $p(\lambda)$ are real, we follow the computations in the proof of Theorem 1.3, and obtain asymptotically as $T \to 0$

$$\lambda_{H\backslash G}(R_T) \sim 2^{-n}\omega_n D^{-1/2} \cdot T^{n(n-2)/2}, \tag{34}$$

where $D = \prod_{i \neq j} |\mu_i - \mu_j|$ is the discriminant of $p(\lambda)$ and $\{\mu_1, \dots, \mu_n\}$ are the roots of $p(\lambda)$.

By [New, Th. III.14], the number of distinct Γ-orbits in $V_p(\mathbb{Z})$ is the class number, say h, of $\mathbb{Z}[\alpha]$.

Thus by equations (31), (32), (33), and (34), asymptotically

$$N(T, V_p) \quad \sim \quad \frac{2^{n-1} h R \omega_n}{\sqrt{D} \prod_{k=2}^{n} \Lambda(k/2)} \cdot T^{n(n-1)/2}. \qquad \square$$

Example 2.

Proof of Theorem 1.4. Fix $\xi \in V_{A,B}(\mathbb{Z})$. Let W_ξ be the m-dimensional subspace of $\mathbb{R}^n$ spanned by the rows of ξ. Since $\xi A\, {}^t\xi = B$ is definite, the restriction of A to W_ξ is definite. Choose an orthonormal basis for W_ξ with respect to A, and extend it to a basis of $\mathbb{R}^n$ such that the matrix of A takes the form

$$A = \begin{pmatrix} I & 0 \\ 0 & J_0 \end{pmatrix},$$

where J_0 is a $(n-m) \times (n-m)$ diagonal matrix whose diagonal entries are ± 1's. In this basis $\xi = (\xi_0, 0)$, where $\xi_0 \in \mathrm{M}_{m\times m}(\mathbb{R})$. Let H be the stabilizer of ξ in G and $M = Z(H^0)$. Then

$$H = \begin{pmatrix} I & 0 \\ 0 & H_0 \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} M_0 & 0 \\ 0 & I \end{pmatrix}, \tag{35}$$

where $H_0 \in \mathrm{SO}(J_0)$ and $M_0 \in \mathrm{SO}(m)$. In particular $\pi(Z(H^0))$ is compact.

In order to apply Theorem 1.16, it remains to show that the pullback sets $\{R_T\}$ are not focused, and to calculate their volume growth.

Let θ denote the Cartan involution $\theta(g) = {}^t g^{-1}$. Let $K = \{g \in G\colon \theta(g) = g\}$. Then K is a maximal compact subgroup of G and $M \subset K$. Let σ be the involution of $G = \mathrm{SO}(A)$ defined by

$$\sigma(g) = \begin{pmatrix} I & 0 \\ 0 & -J_0 \end{pmatrix} \theta(g) \begin{pmatrix} I & 0 \\ 0 & -J_0 \end{pmatrix}.$$

Note that $\theta\sigma = \sigma\theta$. Also the fixed point set of σ is $\tilde{H} = HM$. Thus $\tilde{H}$ is an affine symmetric subgroup of G. There exists an $\mathbb{R}$-split torus $\tilde{A}$ such that $\sigma(a) = a^{-1} = \theta(a)$ ($\forall a \in \tilde{A}$) and $G = \tilde{H}\tilde{A}K$ (see [Sch, Ch. 7]). Thus $H\backslash G/K \cong M \times \tilde{A}$. As we shall see, the nonfocusing of the sets R_T will follow from the fact that they grow at the same rate in all directions of $\tilde{A}$.

$\tilde{A}$-root-space decomposition of Lie(G). Let $r = \min(m, q)$. Write matrix of the quadratic form A as

$$A = \begin{pmatrix} I & 0 & 0 \\ 0 & J_1 & 0 \\ 0 & 0 & -I \end{pmatrix},$$

where J_1 is a diagonal matrix with ± 1 in each diagonal entry and I is the identity matrix. Note that in the above matrix, and all that follows, the first and the third columns are r elements wide, and the second column is $n - 2r$ elements wide. Then Lie$(\tilde{A})$ is

$$\tilde{\mathfrak{a}} = \left\{ \tilde{\Lambda}_{\mathbf{t}} = \begin{pmatrix} 0 & 0 & \Lambda_{\mathbf{t}} E \\ 0 & 0 & 0 \\ E\Lambda_{\mathbf{t}} & 0 & 0 \end{pmatrix} : \mathbf{t} = (t_1, \ldots, t_r) \in \mathbb{R}^r,\ \Lambda_{\mathbf{t}} = \operatorname{diag}(t_1, \ldots, t_r) \right\},$$

where E is a matrix with 1 in each anti-diagonal entries and 0 elsewhere. Put

$$J = \begin{pmatrix} I & 0 & E \\ 0 & I & 0 \\ E & 0 & -I \end{pmatrix}.$$

Then

$$J^{-1}AJ = \begin{pmatrix} 0 & 0 & E \\ 0 & J_1 & 0 \\ E & 0 & 0 \end{pmatrix} \quad \text{and} \quad J^{-1}\tilde{\Lambda}_{\mathbf{t}}J = \begin{pmatrix} \Lambda_t & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -E\Lambda_{\mathbf{t}}E \end{pmatrix}.$$

Put $\mathfrak{g} = \mathrm{Lie}(G) = \mathfrak{so}(A)$, then

$$J^{-1}\mathfrak{g}J = \begin{pmatrix} X & Y_1 & Z_1 \\ Y_2 & W & -J_1\,{}^{\mathrm{t}}Y_1 E \\ Z_2 & -E\,{}^{\mathrm{t}}Y_2 J_1 & -E\,{}^{\mathrm{t}}XE \end{pmatrix},$$

where X, and the Y_i are arbitrary, the $Z_i \in \mathfrak{so}(E)$, and $W \in \mathfrak{so}(J_1)$. Thus the positive roots $\Sigma^+(\mathfrak{g}, \tilde{\mathfrak{a}})$ obtained by diagonalizing $\tilde{\mathfrak{a}}$ on $\mathfrak{g}$ are given by:

$$\beta(\tilde{\Lambda}_{\mathbf{t}}) = \begin{cases} t_i - t_j & (\dim \mathfrak{g}^\beta = 1) \quad (1 \le i < j \le r) \\ t_i + t_j & (\dim \mathfrak{g}^\beta = 1) \quad (1 \le i < j \le r) \\ t_i & (\dim \mathfrak{g}^\beta = n - 2r) \quad (1 \le i \le r). \end{cases} \tag{36}$$

Therefore the half sum of the positive roots is

$$\rho(\tilde{\Lambda}_{\mathbf{t}}) = (1/2)\sum_{i=1}^{r}(n - 2i)t_i. \tag{37}$$

Haar measure on $H\backslash G$. We will use a formula for a G-invariant integral $d(\tilde{H}g)$ on the symmetric space $\tilde{H}\backslash G$ (see [Sch, Th. 8.1.1]):

$$\int_{\tilde{H}\backslash G} f(\tilde{H}g)d(\tilde{H}g) = \int_{\tilde{\mathfrak{a}}_+}\int_K f(\tilde{H}\exp(\tilde{\Lambda}_{\mathbf{t}})k)\delta(\tilde{\Lambda}_{\mathbf{t}})\,dt_1\cdots dt_r\,dk,$$

where $\tilde{\mathfrak{a}}_+$ is the positive (closed) Weyl chamber in $\tilde{\mathfrak{a}}$,

$$\delta(\tilde{\Lambda}_{\mathbf{t}}) = \prod_\beta \left(\sinh\beta(\tilde{\Lambda}_{\mathbf{t}})\right)^{p_\beta}\left(\cosh\beta(\tilde{\Lambda}_{\mathbf{t}})\right)^{q_\beta},$$

β runs over the positive roots $\Sigma^+(\mathfrak{g},\mathfrak{a})$, and $\dim\mathfrak{g}^\beta = p_\beta + q_\beta$.

Since $\tilde{H} = HM$, using the above formula a G-invariant integral $d(Hg)$ on $H\backslash G$ is given by

$$\int_{H\backslash G} f(Hg)d(Hg) = \int_M\int_{\tilde{\mathfrak{a}}_+}\int_K f(Hm\exp(\tilde{\Lambda}_{\mathbf{t}})k)\delta(\tilde{\Lambda}_{\mathbf{t}})\,dm\,dt_1\cdots dt_r\,dk.$$

As $t\to+\infty$, both $\sinh t$ and $\cosh t$ are asymptotic to $(1/2)\exp t$. Therefore in the interior of $\tilde{\mathfrak{a}}_+$, asymptotically

$$\delta(\tilde{\Lambda}_{\mathbf{t}}) \sim \exp(2\rho(\tilde{\Lambda}_{\mathbf{t}})). \tag{38}$$

Coordinate description of R_T. Put $a_i = \exp(t_i)$, and express $(a_1,\ldots,a_r) = \eta\omega$, where $\omega\in S^{r-1}$. We will use the notation $D_{\eta\omega} = \exp(\tilde{\Lambda}_{\mathbf{t}})$.

Claim 1. The limit

$$\lambda(m,\omega) = \lim_{\eta\to\infty}(1/\eta)\|\xi\cdot mD_{\eta\omega}k\| \tag{39}$$

exists, it is independent of $k\in K$, and the convergence is uniform in the parameters $m\in M$ and $\omega\in S_+^{r-1}$, where $S_+^{r-1} = S^{r-1}\cap\exp\tilde{\mathfrak{a}}_+$. Moreover the function $\lambda(m,\omega)$ is continuous and bounded away from 0.

To prove the claim, write

$$\begin{aligned}\|\xi\cdot mD_{\eta\omega}k\| &= \|(\xi_0 m_0, 0)\cdot JJ^{-1}D_{\eta\omega}\|\\ &= \|(\xi_0 m_0,\psi)\cdot J^{-1}D_{\eta\omega}\|\\ &= \eta\|(\xi_0 m_0)_1\cdot J^{-1}D_\omega\| + O(1),\end{aligned}$$

where $m_0\in M_0$, $\psi\in \mathrm{M}_{m\times(n-m)}(\mathbb{R})$ is bounded, and $(\xi_0 m_0)_1$ is the matrix obtained by replacing the last $n-r$ columns of $(\xi_0 m_0, 0)$ by 0. Thus

$$\lim_{\eta\to\infty}(1/\eta)\|\xi\cdot mD_{\eta\omega}k\| = \|(\xi_0 m_0)_1\cdot J^{-1}D_\omega\|$$

and the claim follows.

Claim 2. Asymptotically, $\mathrm{vol}(R_T)\sim c_{A,B}T^{r(n-r-1)}$, where

$$c_{A,B} = \frac{1}{r(n-r-1)}\int_M\int_{S_+^{r-1}}\frac{\prod_1^r\omega_j^{n-2j-1}}{\lambda(m,\omega)^{r(n-r-1)}}\,dm\,d\omega > 0. \tag{40}$$

To prove the claim we express

$$\begin{aligned}\operatorname{vol}(R_T) &= \int_M \int_{\{\mathbf{t}\in\mathbb{R}^r:\ \|\xi_0\cdot m\exp(\tilde{\Lambda}_{\mathbf{t}})k\|<T\}} \delta(\tilde{\Lambda}_{\mathbf{t}})\,dm\,d\mathbf{t}\,dk \\ &= \int_M \int_K \int_{S_+^{r-1}} \int_{\{\eta>0:\ \|\xi_0\cdot mD_{\eta\omega}k\|<T\}} \delta(D_{\eta\omega})\frac{\eta^{-1}}{\omega_1\cdots\omega_r}\,dm\,dk\,d\omega\,d\eta.\end{aligned}$$

By Claim 1, and equations (38) and (37), asymptotically

$$\begin{aligned}\operatorname{vol}(R_T) &\sim \int_M \int_K \int_{S_+^{r-1}} \int_0^{T/\lambda(m,\omega)} \eta^{r(n-r-1)-1} \prod_{j=1}^{r} \omega_j^{n-2j-1}\,dm\,dk\,d\omega\,d\eta \\ &= c_{A,B}\cdot T^{r(n-r-1)}.\end{aligned}$$

This proves the claim.

Nonfocusing of the sets R_T.

Claim 3. Let L be a connected reductive real algebraic subgroup of G containing H. Then $\theta(L) = L$, (recall that $\theta(g) = {}^{\mathrm{t}}g^{-1}$).

To prove the claim, let θ_1 be a Cartan involution of G stabilizing L. Since $H \subset L$ is reductive, there exists $l \in L$ such that $\theta_2 = i_l^{-1}\theta_1 i_l$ stabilizes H, where i_l is conjugation by l. Since θ restricted to H is a Cartan involution of H, there exists $h \in H$, such that $\theta' = i_h^{-1}\theta_2 i_h$ agrees with θ on H. Note that $\theta'(L) = L$. Since θ and θ' are Cartan involutions of G, $\theta' = i_g\theta i_g^{-1}$ for some $g \in G$. Let $g_1 = g^{-1}\theta(g)$. Then $\theta' = i_{g_1}\theta$. Since θ and θ' agree on H, $g_1 \in Z(H) = M$.

Let $\mathcal{P} = \{p \in G\colon\ \theta(p) = p^{-1}\}$. By Cartan decomposition, $G = \mathcal{P}K$ and $K \cap \mathcal{P} = \{e\}$. Write $g = pk$ for $p \in \mathcal{P}$ and $k \in K$. Then $g_1 = p^2 \in \mathcal{P}$.

But $M \cap \mathcal{P} = \{e\}$. Therefore $g_1 = e$, and hence $\theta = \theta'$. Thus $\theta(L) = L$ and the claim holds.

Claim 4. The sequence $\{R_{T_i}\}$ is not focused for any sequence $T_i \to \infty$.

Now suppose that $\{R_{T_i}\}$ is focused along a proper connected reductive real algebraic subgroup L of G containing H. Let $\mathfrak{l} = \operatorname{Lie}(L)$, $\mathfrak{h} = \operatorname{Lie}(H)$ and $\mathfrak{k} = \operatorname{Lie}(K)$. With respect to the Killing form on $\mathfrak{g}$, let $\mathfrak{q} = \mathfrak{h}^{\perp}$, $\mathfrak{p} = \mathfrak{k}^{\perp}$, and $\mathfrak{q}_L = \mathfrak{l}^{\perp} \cap \mathfrak{p} \cap \mathfrak{q}$. Since $\theta(L) = L$, by Theorem A.1, $L = H\exp(\mathfrak{p}\cap\mathfrak{q}\cap\mathfrak{l})K_L$ and $G = L\exp\mathfrak{q}_L K$.

Now let $C \subset G$ be a compact set. Then there exists a compact set $\mathfrak{c} \subset \mathfrak{q}_L$, such that

$$LC \subset H\exp(\mathfrak{q}\cap\mathfrak{p}\cap\mathfrak{l}+\mathfrak{c})K. \tag{41}$$

Put

$$\Delta = \{(m,\omega) \in M \times S_+^{r-1}\colon\ mD_\omega m^{-1} \in L\}.$$

Then

$$LK = \bigcup\{HmD_{\eta\omega}K\colon\ (m,\omega)\in\Delta,\ \eta\geq 0\}.$$

For $\eta > 0$, put

$$\mathcal{U}_\eta = \{(m,\omega)\in M\times S^{r-1}\colon\ HmD_{\eta\omega}K\cap LC\neq\emptyset\}.$$

Then by equation (41), $\mathcal{U}_\eta \downarrow \Delta$ as $\eta\to\infty$. Note that measure of Δ is zero with respect to the integral $dmdk$. Now by Claim 1 and the computations in Claim 2, we obtain that $\mathrm{vol}(R_T\cap LC) = o(\mathrm{vol}(R_T))$. This completes the proof of the claim. Now applying Theorem 1.16, we obtain Theorem 1.4.

7. Effect of focusing on counting estimates

We present an example where the sets R_T ($T\gg 0$) are focused along an intermediate subgroup $H\subset L\subset G$. Unlike in Theorem 1.17, here $N(T,V,\mathcal{O})\not\sim \lambda_{H\backslash G}(R_T)$ as $T\to\infty$. This example was presented without proof in [EM].

Let $G=\mathrm{SL}_2(\mathbb{C})$ and $H=\{\mathrm{diag}(a,a^{-1})\colon\ a\in\mathbb{R}^\times\}$. Let $\Gamma = g^{-1}\mathrm{SL}_2(\mathbb{Z}[i])g$, where $g\in\mathrm{SL}_2(\mathbb{R})$ conjugates a hyperbolic element of $\mathrm{SL}_2(\mathbb{Z})$ into H. Then $\Gamma\cap H\backslash H$ is compact. Here we consider the $\mathbb{Q}$-structure on G such that $G(\mathbb{Q}) = g^{-1}\mathrm{SL}_2(\mathbb{Q}[i])g$.

We will consider the following representation of G, with a closed orbit isomorphic to $H\backslash G$. Fix $N\geq 4$, and let (z_1,z_2) be coordinates on $\mathbb{C}^2$. Let W be the vector space of polynomial functions $v(z_1,z_2,\bar{z}_1,\bar{z}_2)$ on $\mathbb{C}^2$ which are homogeneous of degree N in z_1 and z_2, and also in $\bar{z}_1$ and $\bar{z}_2$. The monomials $z_1^m z_2^{N-m}\bar{z}_1{}^n\bar{z}_2{}^{N-n}$, where $0\leq m,n\leq N$, form a basis for W. Since G acts linearly on $\mathbb{C}^2$, it acts linearly on W by substitution; i.e., for $g\in G$ and $w\in W$, we have $[w\cdot g](z_1,z_2) = w((z_1,z_2)g^{-1})$ ($\forall (z_1,z_2)\in\mathbb{C}^2$).

Consider the polynomial

$$p(z_1,z_2) \;=\; \left[\frac{z_1\bar{z}_2+\bar{z}_1z_2}{2}\right]^2\left[\frac{-z_1\bar{z}_2+\bar{z}_1z_2}{2i}\right]^{N-2}.$$

Then $\mathrm{Stab}_G(p)=H$. Let $V=p\cdot G$. Put $U=\{u(w)=\left(\begin{smallmatrix}1 & w\\ 0 & 1\end{smallmatrix}\right)\colon\ w\in\mathbb{C}\}$ and $K=\mathrm{SU}(2)$. Then $G=HUK$. Since U is unipotent and K is compact, V is closed in the Hausdorff topology. Now since the G-action on W is linear, V is Zariski closed. Thus the variety V is naturally identified with $H\backslash G$.

Next we consider the distribution of the subset $p\cdot\Gamma\subset V$. Let $\|\cdot\|$ be a K-invariant norm on W. Then for $T>0$ we have $R_T=\{Hg\in H\backslash G\colon\ \|p{\cdot}g\|<T\}$. Put $L=\mathrm{SL}_2(\mathbb{R})$. Then L is a $\mathbb{Q}$-subgroup of G containing H with no nontrivial $\mathbb{Q}$-characters. We will show that for any sequence $T_n\to\infty$, the sequence $\{R_{T_n}\}$ is focused along L.

Coordinate description of R_T. Since R_T is K-invariant, it can be treated as a subset of $H\backslash G/K$. Since $G = HUK$, we can use $U \cong \mathbb{C}$ as the coordinate space for $H\backslash G/K$. We identify U with $\mathbb{R}^2$ by the mapping $(x, y) \mapsto u(x + iy)$. Thus we can identify R_T with a subset of $\mathbb{R}^2$.

LEMMA 7.1. *In the $\mathbb{R}^2$-coordinates $R_T = \{(x,y)\colon\ f(x,y) < T^2\}$, where*

$$f(x,y) = c_1x^2 + c_2y^{2M} + c_3x^2y^{2M} + O(x^2y^{2M-2}) + O(y^{2M-2}).$$

Here $M = N - 2 \geq 2$, $c_i > 0$ $(i = 1,2,3)$, and the big O-terms in the definition of f are sums of squares of monomials.

Proof. Since $(z_1, z_2) \cdot u(w)^{-1} = (z_1, z_2 - wz_1)$, if $w = x + iy$, then

$$\begin{aligned}
&[p \cdot u(w)](z_1, z_2) \\
&\quad = \left[\frac{z_1(\bar{z}_2 - \bar{w}\bar{z}_1) + \bar{z}_1(z_2 - wz_1)}{2}\right]^2 \left[\frac{-z_1(\bar{z}_2 - \bar{w}\bar{z}_1) + \bar{z}_1(z_2 - wz_1)}{2i}\right]^{N-2} \\
&\quad = \left[\frac{z_1\bar{z}_2 + z_2\bar{z}_1 - 2x|z_1|^2}{2}\right]^2 \left[\frac{\bar{z}_1z_2 - z_1\bar{z}_2 - 2iy|z_1|^2}{2i}\right]^{N-2}. \qquad (42)
\end{aligned}$$

Since the given norm on W is K-invariant, the distinct weight spaces are orthogonal. Therefore the norm can be expressed in the form

$$\left\|\sum_{m,n} c_{mn} z_1^m z_2^{N-m} \bar{z}_1{}^n \bar{z}_2{}^{N-n}\right\| = \sum_{m,n} \lambda_{mn}|c_{mn}|^2, \qquad (43)$$

where the λ_{mn}'s are positive constants. The conclusion of the lemma follows from equations (42) and (43). □

In view of the last lemma, by ignoring the lower order terms, we define

$$R'_T = \{(x, y)\colon\ c_1x^2 + c_2y^{2M} + c_3x^2y^{2M} < T^2\}.$$

Focusing of $\{R_T\}$ along the x-axis.

LEMMA 7.2. *For $T \gg 1$, $\mathrm{vol}(R'_T) = O(T)$. Also for any $\varepsilon > 0$, there exists $c_\varepsilon > 0$ such that*

$$\limsup_{T\to\infty} \frac{\mathrm{vol}(R'_T \cap \{|y| > c_\varepsilon\})}{\mathrm{vol}(R'_T)} \leq \varepsilon.$$

Proof. We have $(x, y) \in R'_T \Longrightarrow y^{2M}(c_2 + c_3x^2) < T^2 - c_1x^2$. Thus

$$\frac{1}{T}\mathrm{vol}(R'_T) \quad = \quad \frac{4}{T}\int_0^{T/\sqrt{c_1}} \left[\frac{T^2 - c_1x^2}{c_2 + c_3x^2}\right]^{1/(2M)} dx$$

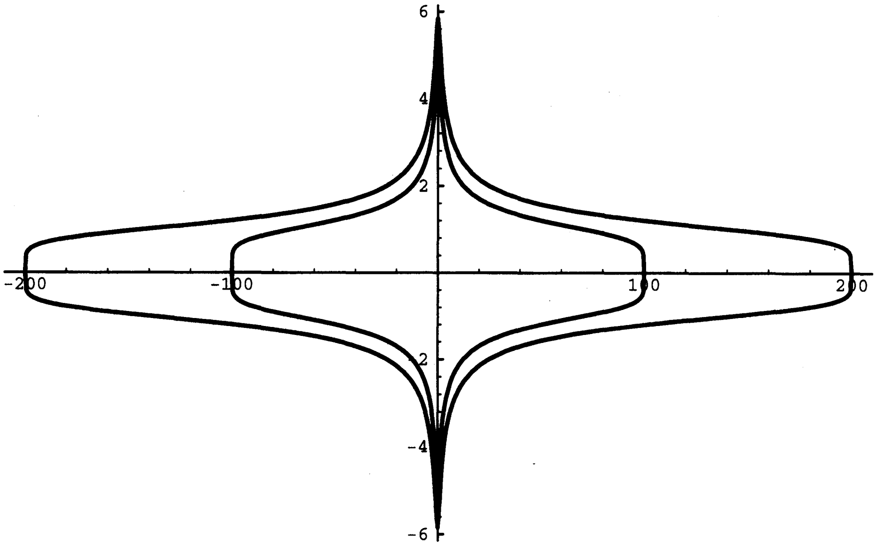

FIGURE 1. The sets R'_T for $N = 5$ and $T = 100, 200$.

$$= \frac{4}{\sqrt{c_1}} \int_0^1 \left[\frac{1-u^2}{c_2/T^2 + (c_3/c_1)u^2} \right]^{1/(2M)} du$$

$$\to \frac{4}{\sqrt{c_1}} \left(\frac{c_1}{c_3} \right)^{1/(2M)} \int_0^1 \left[\frac{1-u^2}{u^2} \right]^{1/(2M)} du$$

as $T \to \infty$, and the limit is finite because $M \geq 2$. □

In terms of the $U \times K$-coordinates on $H\backslash G$, the G-invariant integral is given by $d(Hg) = du\, dk$, where du is the Lebesgue integral on $U \cong \mathbb{R}^2$ and dk is a Haar integral on K.

LEMMA 7.3. *For $T \gg 1$, $\mathrm{vol}_{H\backslash G}(R_T) = O(T)$. Also for every $\varepsilon > 0$, there exists $c_\varepsilon > 0$ such that*

$$\limsup_{T\to\infty} \frac{\mathrm{vol}_{H\backslash G}(R_T \cap \{|y| > c_\varepsilon\})}{\mathrm{vol}_{H\backslash G}(R_T)} \leq \varepsilon.$$

Proof. For $|y| \gg 1$, y^{2M-2} is dominated by y^{2M}, and $x^2 y^{2M-2}$ is dominated by $x^2 y^{2M}$. Thus, in this region, R'_T is an excellent approximation to R_T. Hence the conclusion follows from Lemma 7.2. □

PROPOSITION 7.4. *The sets R_T are focused along $L = \mathrm{SL}_2(\mathbb{R})$. In fact, for any $\varepsilon > 0$ there exists a compact subset C of G such that*

$$\liminf_{T\to\infty} \frac{\mathrm{vol}_{H\backslash G}(LC \cap R_T)}{\mathrm{vol}_{H\backslash G}(R_T)} > 1 - \varepsilon.$$

Proof. For any $c > 0$ there exists a compact set $C \subset G$, such that $R_T \cap LC \supset R_T \cap \{|y| < c\}$. Now the conclusion of the proposition follows from Lemma 7.3. $\square$

Focusing function. Since L is affine symmetric, $G = LBK$, where $B = \{b(t) = \begin{pmatrix} \cosh t & i\sinh t \\ -i\sinh t & \cosh t \end{pmatrix}: t \in \mathbb{R}\}$. For $c \geq 0$, express $u(ic) = lb(t)k$, where $l \in L$, $t \geq 0$ and $k \in K$. By multiplying the equation on the left by the inverse of its complex conjugate, and then taking the Hilbert-Schmidt norm of both sides, we get $2 + 4c^2 = e^{4t} + e^{-4t}$, i.e. $c = \sinh 2t$. In particular, we have the decomposition, $G = Lu(i\mathbb{R})K$.

By [Sch, Sect. 8.4.2], a Haar integral on G in the $L \times B \times K$-coordinates is $\cosh^2(2t)\, dl\, dt\, dk$. Hence in the $L \times u(i\mathbb{R}) \times K$-coordinates, the integral $j(c)\, dl\, dc\, dk$ is G-invariant, where $j(c) = \sqrt{1+c^2}$. Thus using the $u(i\mathbb{R}) \times K$-coordinates on $L\backslash G$, the G-invariant integral on $L\backslash G$ is $d(Lg) = j(c)\, dc\, dk$.

For $c \in \mathbb{R}$, define $X_T(c) = \{x \in \mathbb{R}\colon (x, c) \in R_T\}$. By Lemma 7.1, $|X_T(c)| \sim O(c^{-M})T$ for $T \gg 1$. Now define

$$\lambda(c) = \lim_{T\to\infty} \frac{|X_T(c)|}{\mathrm{vol}_{H\backslash G}(R_T)}; \tag{44}$$

the limit exists because $\mathrm{vol}_{H\backslash G}(R_T) = O(T)$, by Lemma 7.3. Also, $\lambda(c)$ is continuous, and

$$\lambda(c) = \lambda(-c) = O(c^{-M}) \qquad \text{as } c \to \infty. \tag{45}$$

For $g = lu(ic)k$, where $l \in L$, $c \in \mathbb{R}$, and $k \in K$, define $\tilde{\lambda}(g) = \lambda(c)/j(c)$. Then $\tilde{\lambda}$ is a left L-invariant and right K-invariant continuous function on G. We can also treat $\tilde{\lambda}$ as a function on $L\backslash G$. Since $M \geq 2$, by equation (45),

$$\int_{L\backslash G} \tilde{\lambda}(Lg)\, d(Lg) = \int_{-\infty}^{\infty} \lambda(c)\, dc = O\left(\int_1^{\infty} c^{-M} dc\right) < \infty. \tag{46}$$

Define the 'focusing map' on $\Gamma\backslash G$ by

$$\Lambda(\Gamma g) = \sum_{\gamma \in (L\cap\Gamma)\backslash\Gamma} \tilde{\lambda}(\gamma g).$$

Then

$$\begin{aligned} \langle \Lambda, \psi \rangle &= \int_{\Gamma\backslash G} d(\Gamma g) \sum_{\gamma \in (L\cap\Gamma)\backslash\Gamma} \tilde{\lambda}(\gamma g)\overline{\psi}(\Gamma g) \\ &= \int_{L\backslash G} d(Lg) \int_{(L\cap\Gamma)\backslash L} \tilde{\lambda}(Lg)\overline{\psi}(\Gamma lg)\, d((L\cap\Gamma)l) \\ &= \int_{L\backslash G} d(Lg)\, \tilde{\lambda}(Lg) \left(\int_{\Gamma\backslash G} \overline{\psi}\, d\mu_{Lg}\right), \end{aligned} \tag{47}$$

which is finite by equation (46).

Counting estimate. Let ψ be a compactly supported continuous function on $\Gamma\backslash G$. Then for any $c \in \mathbb{R}$, by Theorem 1.7,

$$\lim_{x\to\infty} \int_{\Gamma\backslash G} \overline{\psi}\, d[\mu_H \cdot u(x+ic)] = \int_{\Gamma\backslash G} \overline{\psi}\, d[\mu_L \cdot u(ic)]. \tag{48}$$

As in Proposition 5.3, define the 'counting function' on $\Gamma\backslash G$ by

$$\hat{F}_T(\Gamma g) = \frac{1}{\mathrm{vol}_{H\backslash G}(R_T)} \sum_{\dot\gamma\in(H\cap\Gamma)\backslash\Gamma} \chi_{R_T}(p\cdot\gamma g).$$

Then

$$\begin{aligned}
\lim_{T\to\infty} \langle \hat{F}_T, \psi\rangle &= \lim_{T\to\infty} \frac{1}{\mathrm{vol}_{H\backslash G}(R_T)} \int_{R_T} \left(\int_{\Gamma\backslash G} \overline{\psi}\, d\mu_{Hg} \right) d(Hg) \\
&= \lim_{T\to\infty} \frac{1}{\mathrm{vol}_{H\backslash G}(R_T)} \int_K dk \int_{-\infty}^{\infty} dc \int_{X_T(c)} dx \\
&\quad \times \left(\int_{\Gamma\backslash G} \overline{\psi}\, d[\mu_H \cdot u(x+ic)k] \right) \\
&= \lim_{T\to\infty} \frac{1}{\mathrm{vol}_{H\backslash G}(R_T)} \int_K dk \int_{-\infty}^{\infty} |X_T(c)| \qquad (49)\\
&\quad \times \left(\int_{\Gamma\backslash G} \overline{\psi}\, d[\mu_L \cdot u(ic)k] \right) dc \\
&= \int_K dk \int_{-\infty}^{\infty} \lambda(c) \left(\int_{\Gamma\backslash G} \overline{\psi}\, d[\mu_L \cdot u(ic)k] \right) dc \\
&= \int_{L\backslash G} \tilde{\lambda}(Lg) \left(\int_{\Gamma\backslash G} \overline{\psi}\, d\mu_{Lg} \right) d(Lg) \qquad (50)\\
&= \langle \Lambda, \psi\rangle; \qquad (51)
\end{aligned}$$

the equations (49), (50), (51), and (52) follow from equations (48), (44), (46), and (47), respectively. Hence as in the proof of Theorem 1.16 given in Section 5, we obtain the following:

THEOREM 7.5. *There exists a* nonconstant *continuous positive function* $\Lambda(g)$ *on* $\Gamma\backslash G$ *such that asymptotically as* $T \to \infty$,

$$F_T(g) \sim \Lambda(g)\, \mathrm{vol}_{H\backslash G}(R_T).$$

Appendix A. Asymptotics of volume growth of $\mathbf{R_T}$

In this appendix, we prove Proposition 5.4. Our argument is more involved than that of [DRS, Appendix 1] because we do not assume that H is an affine symmetric subgroup, and we allow dilations of arbitrary symmetric convex sets B_T, not only K-invariant balls.

The following theorem is proved in [Mos]; see also [Hel, Ch. IV, exercise A.2].

THEOREM A.1. *Let G be an algebraic reductive group, H a reductive subgroup. Let K be a maximal compact subgroup of G, associated to a Cartan involution stabilizing H. Let $\mathfrak{g}$, $\mathfrak{k}$ and $\mathfrak{h}$ denote the Lie algebras of G, H and K, respectively. Let $\mathfrak{p} = \mathfrak{k}^\perp$ and $\mathfrak{q} = \mathfrak{h}^\perp$ with respect to the Killing form on $\mathfrak{g}$. Then*

$$G = H\exp(\mathfrak{p}\cap\mathfrak{q})K.$$

For $g \in G$, consider the corresponding decomposition $g = hqk$, where $h \in H$, $q \in \exp(\mathfrak{q}\cap\mathfrak{p})$, and $k \in K$. Then h, q, and k are uniquely determined by g up to the transformation $(h,q,k) \mapsto (hm, \operatorname{Ad} m(q), m^{-1}k)$, where $m \in H \cap K$.

Let $M = K\cap H$ and let π: $K \to M\backslash K$ be the natural quotient map. There exists a submanifold K' of K such that $e \in K'$, π is a local diffeomorphism on K', and $\pi(K')$ has full K-invariant measure on $M\backslash K$.

Let Φ: $(\mathfrak{p}\cap\mathfrak{q}) \times K' \to H\backslash G$ be the map given by $\Phi(Y,k) = H\exp(Y)k$. We will relate the invariant measure on $H\backslash G$ to the Euclidean measure on $\mathfrak{p}\cap\mathfrak{q}$, say dY, and the pullback of the K-invariant measure on $M\backslash K$ to K', say dk. Let $\mathrm{J}\,\Phi(Y,k)$ denote the Jacobian of Φ at (Y,k). Then the Jacobian is independent of $k \in K'$. We write $\mathrm{J}\,\Phi(Y,k) = \delta(Y)$.

Now a G-invariant integral $d(Hg)$ on $H\backslash G$ can be given by

$$\int_{H\backslash G} f(Hg)\,d(Hg) = \int_{\mathfrak{p}\cap\mathfrak{q}}\int_{K'} f(H\exp(Y)k)\delta(Y)\,dY\,dk. \tag{53}$$

Using the polar coordinates on $\mathfrak{p}\cap\mathfrak{q}$, we write $Y = r\omega$, where $r = \|Y\|$ and $\omega = Y/\|Y\|$. We also write $\delta(Y) = \delta(r,\omega,k)$.

LEMMA A.2. *There exists $n \in \mathbb{N}$ and $\Lambda > 0$ such that as a function of r, we have $r^l\delta(r,\omega) \in E(n,\Lambda)$ for all $\omega \in p\cap q$ with $\|\omega\| = 1$, where $l = \dim(\mathfrak{p}\cap\mathfrak{q})$.*

Proof (cf. [Sch, Thm. 8.1.1]). Put $q = \exp(Y)$. For $Z \in \mathfrak{p}\cap\mathfrak{q}$ and $X \in T_e(K')$,

$$\Phi(\exp(Z+Y), \exp Xk') = H\exp(Z+Y)\exp(-Y)\exp(\operatorname{Ad}(q^{-1})X)qk'.$$

Therefore

$$d\Phi_{(Y,k')}(Z,X) = \mathrm{Pr}_{\mathfrak{q}}((d\exp)_Y(Z) + \operatorname{Ad}(q^{-1})(X)),$$

where $\mathrm{Pr}_{\mathfrak{q}}$: $\mathfrak{g} \to \mathfrak{q}$ is the projection parallel to $\mathfrak{h}$.

By [Var, Th. 2.14.3],

$$(d\exp)_Y(Z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!}(\operatorname{ad} Y)^n(Z). \tag{54}$$

Write $Y = r\omega$ in polar coordinates, where $r > 0$ and $\omega \in S^{l-1}$, the unit sphere in $\mathfrak{p} \cap \mathfrak{q}$. For $\omega \in S^{l-1}$, let $\nu_i(\omega)$ and $W_i(\omega)$ be the eigenvalues and eigenvectors, respectively, of $\operatorname{ad}\omega \in \operatorname{End}(\mathfrak{g})$. Since all vectors in $\mathfrak{p} \cap \mathfrak{q}$ are semi-simple, we can choose the ν_i and W_i to depend continuously on ω. By equation (54),

$$(d\exp)_Y(Z) + \operatorname{Ad}_q(X) = \sum_{i=1}^{\dim \mathfrak{g}} \left(\frac{1 - e^{-r\nu_i(\omega)}}{r\nu_i(\omega)} \langle Z, W_i \rangle + e^{r\nu_i(\omega)} \langle X, W_i \rangle \right) W_i.$$

Thus as a function of r, each matrix entry of $\Pr_{\mathfrak{q}}(r((d\exp)_Y(Z) + \operatorname{Ad}_q(X)))$ belongs to some $E(n', \Lambda')$, where n' and Λ' can be chosen independent of ω. Therefore r^l times the Jacobian determinant $\delta(r, \omega)$ belongs to $E(n, \Lambda)$, for some n and Λ independent of ω.

Note that in the polar coordinates, the integral $d(Hg)$ on $H\backslash G$ as in equation (53) can be expressed as

$$(55) \qquad \int_{H\backslash G} f(Hg)\, d(Hg) = \int_{r>0} \int_{S^{l-1}} \int_{K'} f(H \exp(r\omega)k) \delta(r, \omega) r^{l-1}\, dr\, d\omega.$$

LEMMA A.3. *There exists a constant C, independent of ω, such that for any $r > 1$,*

$$(56) \qquad \delta(r,\omega) r^{l-1} \le C \int_0^r \delta(s,\omega) s^{l-1}\, ds.$$

Proof. In view of Lemma A.2, it is enough to show the following: Let δ_0 be as in Proposition 3.6. Then for all $A \le r \le A + \delta_0$ and $f \in E(n, \Lambda)$,

$$(57) \qquad f(r) \le C \int_A^{A+\delta_0} f(s)\, ds.$$

Since we can make the change of variable $x \to x + A$, we may assume $A = 0$. We may also assume that f is the maximum at r on $[0, \delta_0]$. Now equation (57) follows from Proposition 3.6. $\square$

Shape of R_T in (r, ω, k)-coordinates. We are given a linear representation of G on a vector space $\mathbb{R}^N$, a vector $v_0 \in \mathbb{R}^N$ whose stabilizer is H, and a norm $\|\cdot\|$ on $\mathbb{R}^N$ which may not be K-invariant.

Let the notation be as above. For any $\omega \in S^{n-1}$, choose a basis consisting of eigenvectors $\{v_i(\omega)\}_{i=1}^N$ for $\exp(r\omega)$ such that

$$v_i \cdot \exp(r\omega) = e^{r\lambda_i(\omega)} v_i\ .$$

The $v_i(\omega)$ can be made to depend continuously on ω. Then if $g = h\exp(r\omega)k$,

$$v_0 \cdot g = \sum_{i=1}^N e^{r\lambda_i(\omega)} v_i(\omega) \cdot k,$$

and hence

$$\text{(58)} \qquad \operatorname{vol}(R_T) = \int_K \int_{S_T(k)} \delta(r,\omega)\, r^{l-1}\, dr\, d\omega\, dk = \int_K m(T,k)\, dk,$$

where

$$\text{(59)} \qquad S_T(k) = \{(r,\omega) \in (\mathfrak{p} \cap \mathfrak{q}) \ : \ \|\sum e^{r\lambda_i(\omega)} v_i(\omega) \cdot k\| \leq T\},$$

and

$$\text{(60)} \qquad m(T,k) = \int_{S_T(k)} \delta(r,\omega) r^{l-1}\, dr\, d\omega.$$

LEMMA A.4. *For $T \gg 0$ the following holds:*

(i) *$S_T(k)$ is star-shaped. In particular, for each $\omega \in S^{n-1}$ and $k \in K$, there exists a unique $s > 0$, denoted by $r(T,\omega,k)$, such that*

$$\left\|\sum_{i=1}^{N} e^{s\lambda_i(\omega)} v_i(\omega) \cdot k\right\| = T.$$

(ii) *$r(\kappa T,\omega,k,) \leq r(T,\omega,k) + \beta(\kappa)$, where $\beta(\kappa) \to 1$ as $\kappa \to 1$. The function $\beta(\kappa)$ is independent of ω and k.*

First assuming this lemma, we prove the following:

PROPOSITION A.5. *For some functions $a(\kappa) \to 1$ and $b(\kappa) \to 1$ as $\kappa \to 1$,*

$$\text{(61)} \qquad a(\kappa) \leq \liminf_{T\to\infty} \frac{m(\kappa T,k)}{m(T,k)} \leq \limsup_{T\to\infty} \frac{m(\kappa T,k)}{m(T,k)} \leq b(\kappa).$$

Proof. Because $S_T(k)$ is star-shaped, we can do the integration as in equation (60) in polar coordinates:

$$m(T,\omega,k) = \int_0^{r(\omega,k,T)} \delta(r,\omega) r^{l-1}\, dr,$$

and

$$m(T,k) = \int_\omega m(T,\omega,k)\, d\omega.$$

Therefore

$$\begin{aligned}
\log \frac{m(\kappa T,\omega,k)}{m(T,\omega,k)} &= \log \int_0^{r(\kappa T,\omega,k)} \delta(s,\omega) s^{l-1}\, ds \\
&\quad - \log \int_0^{r(T,\omega,k)} \delta(s,\omega) s^{l-1}\, ds \\
\text{(62)} \qquad &\leq \log \int_0^{r(T,\omega,k)+\beta(\kappa)} \delta(s,\omega) s^{l-1}\, ds \\
&\quad - \log \int_0^{r(T,\omega,k)} \delta(s,\omega) s^{l-1}\, ds \\
\text{(63)} \qquad &= \beta(\kappa) \frac{\delta(\sigma,\omega)\sigma^{l-1}}{\int_0^\sigma \delta(s,\omega) s^{l-1}\, ds} \leq C\beta(\kappa),
\end{aligned}$$

where equation (62) follows from Lemma A.4 and equation (63) follows from the mean value theorem, for some

$$\sigma \in [r(T, \omega, k), r(T, \omega, k) + \beta(\kappa)].$$

Thus

$$\frac{m(\kappa T, \omega, k)}{m(T, \omega, k)} \leq \exp(C\beta(\kappa)) = b(\kappa)$$

with $b(\kappa) \to 1$ as $\kappa \to 1$. From this, equation (61) follows. □

Proof of Proposition 5.4. The result is an immediate consequence of Proposition A.5 and equation (58). □

Proof of Lemma A.4. We will prove both the statements simultaneously. First we need to define $\lambda_{\max}(\omega) = \max\{\lambda_i(\omega)\}$, ($\lambda_i$ are as in equation (59)). Since all norms are equivalent, there exist constants c_1, c_2 such that

$$c_1 e^{s\lambda_{\max}(\omega)} \leq \left\| \sum e^{s\lambda_i(\omega)} v_i(\omega) \cdot k \right\| = T \leq c_2 e^{s\lambda_{\max}(\omega)}.$$

This implies that if the region $S_T(k)$ is *not* star-shaped, say a ray hits the boundary $\partial S_T(k)$ at two points

$$p_j = \sum e^{s_j \lambda_i(\omega)} v_i(\omega) \cdot k \qquad (j = 1, 2),$$

then

$$|s_1 - s_2| < (\log c_2 - \log c_1)/\lambda^- = c_3, \tag{64}$$

where $\lambda^+ = \min_\omega \lambda_{\max}(\omega) > 0$. □

Claim. For $\beta > 0$ and T sufficiently large,

$$\left\| \sum e^{(r(\omega,k,T)+\beta)\lambda_i(\omega)} k v_i \right\| \geq T g(\beta) \tag{65}$$

with $g(0) = 1$, $g'(0) > 0$, and $g(\beta) > 1$ if $0 < \beta < c_3$ (c_3 is defined in equation (64)).

Assuming this for now, we immediately see that $S_T(k)$ is star-shaped for sufficiently large T: since by equation (65)

$$\left\| \sum e^{(r(\omega,k,T)+\beta)\lambda_i(\omega)} v_i(\omega) \cdot k \right\| > T,$$

if $0 < \beta < c_3$, while by equation (64) this is the only range where coincidence may occur. Also equation (65) shows that, because $S_T(k)$ is star-shaped,

$$r(\omega, k, T) + \beta \geq r(\omega, k, g(\beta)T).$$

Since g is increasing near $\beta = 0$, for κ near 1 we can determine β by requiring $g(\beta) = \kappa$. Thus for sufficiently large T, and for κ near 1,

$$r(\omega, k, T) + \beta(\kappa) \geq r(\omega, k, \kappa T).$$

This proves Lemma A.4, assuming the claim.

Proof of the claim. Let $\varepsilon > 0$ be a small parameter to be chosen later. Put $s = r(T, \omega, k)$, $X_i = e^{s\lambda_i(\omega)} v_i(\omega) \cdot k$, and $\delta_\varepsilon(\omega) = \lambda_{\max}(\omega) - \varepsilon$. Let $\Delta^+(\omega) = \{i\colon \lambda_i(\omega) > \delta_\varepsilon(\omega)\}$ and $\Delta^-(\omega) = \{i\colon \lambda_i(\omega) < \delta_\varepsilon(\omega)\}$. For $\beta > 0$,

$$\begin{aligned}
&\left\| \sum_i e^{(s+\beta)\lambda_i(\omega)} v_i(\omega) \cdot k \right\| \\
&= \|e^{\beta(\lambda_{\max}(\omega)-\varepsilon)} \sum_i X_i + \sum_i (e^{\beta\lambda_i(\omega)} - e^{\beta(\lambda_{\max}(\omega)-\varepsilon)}) X_i\| \\
&\geq e^{\beta\delta_\varepsilon(\omega)} \|\sum_i X_i\| - \sum_i |e^{\beta\delta_\varepsilon(\omega)} - e^{\beta\lambda_i(\omega)}| \, \|X_i\| \\
&= e^{\beta\delta_\varepsilon(\omega)} T - \sum_{i\in\Delta^+(\omega)} (e^{\beta\lambda_i(\omega)} - e^{\beta\delta_\varepsilon(\omega)}) \|X_i\| - \sum_{i\in\Delta^-(\omega)} (e^{\beta\delta_\varepsilon(\omega)} - e^{\beta\lambda_i(\omega)}) \|X_i\| \\
&\geq e^{\beta\delta_\varepsilon(\omega)} T - c_4 T \sum_{i\in\Delta^+(\omega)} (e^{\beta\lambda_i(\omega)} - e^{\beta\delta_\varepsilon(\omega)}) \\
&\quad - c_5 \sum_{i\in\Delta^-(\omega)} (e^{\beta\delta_\varepsilon(\omega)} - e^{\beta\lambda_i(\omega)}) T^{\frac{\delta_\varepsilon(\omega)}{\lambda_{\max}(\omega)}} \\
&\geq T \left[e^{\beta(\lambda^- - \varepsilon)} (1 - c_4 \sum_{i\in\Delta^+(\omega)} (e^{\varepsilon\beta} - 1)) - c_5 \sum_{i\in\Delta^-(\omega)} (e^{\beta(\lambda^+ - \varepsilon)} - e^{\beta\nu})\, T^{-\varepsilon/\lambda^+} \right],
\end{aligned}$$

where $\lambda^- = \min_\omega \lambda_{\max}(\omega) > 0$, $\lambda^+ = \max_\omega \lambda_{\max}(\omega) > 0$, and $\nu = \min_{i,\omega} \lambda_i(\omega)$.

$$\begin{aligned}
&= T[e^{\beta(\lambda^- - \varepsilon)} (1 - c_6(e^{\varepsilon\beta} - 1)) - c_7(e^{\beta(\lambda^+ - \varepsilon)} - e^{\beta\nu}) T^{-\varepsilon/\lambda^+}] \\
&= T(g_1(\beta) - g_2(\beta) T^{-\varepsilon/\lambda^+}).
\end{aligned}$$

We can choose ε small enough so that

$$g_1(\beta) = e^{\beta(\lambda^- - \varepsilon)} (1 - c_6(e^{\varepsilon\beta} - 1))$$

has positive derivative at $\beta = 0$, and $g_1(\beta) > 1$ for $0 < \beta < c_3$. Then

$$\|\sum e^{(s+\beta)\lambda_i(\omega)} v_i(\omega) \cdot k\| \geq T(g_1(\beta) - g_2(\beta) T^{-\varepsilon/\lambda^+}).$$

Since $g_2(\beta) = c_7(e^{\beta(\lambda^+ - \varepsilon)} - e^{\beta\nu})$ vanishes at $\beta = 0$, we can adjust $\varepsilon > 0$ so that for sufficiently large T,

$$g_1(\beta) - g_2(\beta) T^{-\varepsilon/\lambda^+} \geq g(\beta)$$

with $g(0) = 1$, $g'(0) > 0$, and $g(\beta) > 1$ for $0 < \beta < c_3$. This proves the claim. □

UNIVERSITY OF CHICAGO, CHICAGO, IL
ESKIN@MATH.UCHICAGO.EDU
INSTITUTE OF MATHEMATICS, HEBREW UNIVERSITY, JERUSALEM, ISRAEL
MOZES@MATH.HUJI.AC.IL
SCHOOL OF MATHEMATICS, TATA INSTITUTE OF FUNDAMENTAL RESEARCH, BOMBAY, INDIA
NIMISH@TIFRVAX.TIFR.RES.IN

REFERENCES

[Bir] B. J. Birch, Forms in many variables, Proc. Roy. Soc. London **265** (1962), 245–263.

[Bor1] A. Borel, *Introduction aux groupes arithmétiques*, Publ. de l'Inst. Math. de l'Univ. de Strasbourg XV, Hermann, Paris, 1969.

[Bor2] ———, *Linear Algebraic Groups*, Springer-Verlag, Berlin, 2nd edition, 1991.

[Bor3] ———, Values of indefinite quadratic forms at integral points and flows on spaces of lattices, Bull. AMS. (N.S.) **32** (1995), 184–204.

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

[BR] M. Borovoi and Z. Rudnick, Hardy-Littlewood varieties and semisimple groups, Invent. Math., **119** (1995), 37–66.

[Dan1] S. G. Dani, Invariant measures of horospherical flows on noncompact homogeneous spaces, Invent. Math. **47** (1978), 101–138.

[Dan2] ———, Flows on homogeneous spaces and diophantine approximation, In: *Proc. of I.C.M.* (1994), to appear.

[DM1] S. G. Dani and G. A. Margulis, Orbit closures of generic unipotent flows on homogeneous spaces of $SL(3,\mathbb{R})$, Math. Ann. **286** (1990), 101–128.

[DM2] ———, Asymptotic behaviour of trajectories of unipotent flows on homogeneous spaces, Proc. Indian. Acad. Sci. Math. Sci. **101** (1991), 1–17.

[DM3] ———, Limit distributions of orbits of unipotent flows and values of quadratic forms, Adv. Soviet Math. **6** (1993), 91–137.

[DRS] W. Duke, Z. Rudnick, and P. Sarnak, Density of integer points on affine homogeneous varieties, Duke Math. J. **71** (1993), 143–179.

[EM] A. Eskin and C. McMullen, Mixing, counting and equidistribution in Lie groups, Duke Math. J. **71** (1993), 181–209.

[EMS] A. Eskin, S. Mozes, and N. Shah, Nondivergence of translates of certain algebraic measures, preprint.

[Hel] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, New York, 1962.

[Lang] S. Lang, *Algebraic Number Theory*, Addison-Wesley, Reading, Mass., 1970.

[Mar1] G. A. Margulis, Discrete subgroups and ergodic theory, In: *Number Theory, Trace Formulas and Discrete Groups* (symposium in honour of A. Selberg, Oslo), Academic Press, San Diego, Calif., 1989, 377–398.

[Mar2] ———, Dynamical and ergodic properties of subgroup actions on homogeneous spaces with applications to number theory, In: *Proc. of I.C.M.* (Kyoto, 1990). Math. Soc. of Japan, Tokyo, 1991, 193–215.

[Mos] G. D. Mostow, Some new decomposition theorems for semi-simple groups, Mem. Amer. Math. Soc., no. 14 (1955), 31–54.

[Moz] S. Mozes, Mixing of all orders of Lie group actions, Invent. Math. **107** (1992), 235–241.

[MS] S. Mozes and N. A. Shah, On the space of ergodic invariant measures of unipotent flows, Ergodic Th. and Dynam. Sys. **15** (1995), 149–159.

[New] M. Newman, *Integral Matrices*, Academic Press, New York, 1972.

[PR] V. Platonov, and A. Rapinchuk, *Algebraic Groups*, Academic Press, New York, 1994.

[Rat1] M. Ratner, Strict measure rigidity for unipotent subgroups of solvable groups, Invent. Math. **101** (1990), 449–482.

[Rat2] ———, On measure rigidity of unipotent subgroups of semisimple groups, Acta. Math. **165** (1990), 229–309.

[Rat3] ———, On Ragunathan's measure conjecture, Ann. of Math. **134** (1991), 545–607.

[Rat4] ———, Raghunathan's topological conjecture and distributions of unipotent flows, Duke Math. J. **63** (1991), 235–290.

[Rat5] ———, Interactions between ergodic theory, Lie groups and number theory, *Proc. of I.C.M.* 94, to appear.

[Rud] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, 2nd edition, 1974.

[Sch] H. Schlichtkrull, *Hyperfunctions and Harmonic Analysis on Symmetric Spaces*, Birkhäuser, Boston, 1984.

[Sha1] N. A. Shah, Uniformly distributed orbits of certain flows on homogeneous spaces, Math. Ann. **289** (1991), 315–334.

[Sha2] ———, Limit distributions of polynomial trajectories on homogeneous spaces, Duke Math. J. **75** (1994), 711–732.

[Terras] A. Terras, *Harmonic Analysis on Symmetric Spaces and Applications* II, Springer-Verlag, Berlin, 1988.

[Var] V. S. Varadarajan, *Lie Groups, Lie Algebras, and their Representations*, Springer-Verlag, Berlin, 1974.

(Received February 1, 1994)