EFFECTIVE ARGUMENTS IN UNIPOTENT DYNAMICS

M. EINSIEDLER AND A. MOHAMMADI

Dedicated to Gregory Margulis

ABSTRACT. We survey effective arguments concerning unipotent flows on locally homogeneous spaces.

1. Introduction

In the mid 1980's Margulis resolved the long standing Oppenheim conjecture by establishing a special case of Raghunathan's conjecture. Further works by Dani and Margulis in the context of the Oppenheim conjecture and Ratner's full resolution of Raghunathan's conjectures have become a corner stone for many exciting applications in dynamics and number theory.

Let us briefly recall the setup. Let G be a connected Lie group and let $\Gamma \subset G$ be a lattice (i.e. a discrete subgroup with finite covolume) and $X = G/\Gamma$. Let $H \subset G$ be a closed subgroup of G. This algebraic setup gives hope for the following fundamental dynamical problem.

Describe the behavior of the orbit Hx for every point $x \in X$.

However, without further restrictions on H this question cannot have any meaningful answer, e.g. if G is semisimple and H is a one parameter \mathbb{R} -diagonalizable subgroup of G, then the time one map is partially hyperbolic (and in fact has positive entropy and is a Bernoulli automorphism) and the behavior of orbits can be rather wild giving rise to fractal orbit closures, see e.g. [47]. There is however a very satisfying answer when H is generated by unipotent subgroups, e.g. when H is a unipotent or a connected semisimple subgroup; in these cases Ratner's theorems imply that closure of all orbits are properly embedded manifolds, see §5.

These results, however, are not effective, e.g. they do not provide a rate at which the orbit fills out its closure. As it is already stated by Margulis in his ICM lecture [60], it is much anticipated and quite a challenging problem to give effective versions of these theorems. It is worth noting that except for uniquely ergodic systems, such a rate would generally depend on delicate properties of the point x and the acting group H. Already for an irrational rotation of a circle, the *Diophanine properties* of the rotation enters the picture. The purpose of this article is to provide an overview of effective results in this context of unipotently generated subgroups.

M.E. acknowledges support by the SNF (grant 178958).

A.M. acknowledges support by the NSF.

Let us further mention that there has been fantastic developments both for other choices of H and also beyond the homogeneous setting. In fact the papers [53, 29, 30] give partial solutions to the conjectures by Margulis [58] concerning higher rank diagonalizable flows, the papers [10, 5, 6, 7] (inspired by the methods of Eskin, Margulis and Mozes [32]) concern the classification of stationary measures, and [35, 36] concern the $SL_2(\mathbb{R})$ action on moduli spaces and apply also the method developed for stationary measures. These works, with the exception of [10], are all qualitative and an effective account of these would be very intriguing. This article however will focus on the case where H is generated by unipotent elements.

We note that good effective bounds for equidistribution of unipotent orbits can have far reaching consequences. Indeed the Riemann hypothesis is equivalent to giving an error term of the form $O_{\epsilon}(y^{\frac{3}{4}+\epsilon})$ for equidistribution of periodic horocycles of period 1/y on the modular surface [82, 73].

Given the impact of Margulis' work for the above research directions and especially the research concerning effective unipotent dynamics on homogeneous spaces portrait here, but more importantly our personal interests, it is a great pleasure to dedicate this survey to Gregory Margulis.

1.1. **Periodic orbits and a notion of volume.** Suppose $x \in X$ is so that Hx is dense in X. As will be evident in the following exposition, and was alluded to above, the orbit Hx may fill up the space very slowly, e.g. x may be very close to an H-invariant manifold of lower dimension. To have any effective account, we first need a measure of complexity for these intermediate behaviors.

We will always denote the G-invariant probability measure on X by vol_X . Let $L \subset G$ be a closed subgroup. A point $x \in X$ will be called L-periodic if

$$\operatorname{stab}_{L}(x) = \{ g \in L : gx = x \}$$

is a lattice in L. Similarly, a periodic L-orbit is a set Lx where x is an L-periodic point; given an L-periodic point x we let μ_{Lx} denote the probability L-invariant measure on Lx. By a homogeneous measure on X we always mean μ_{Lx} for some L and x. Sometimes we refer to the support of a homogeneous measure, which is an L-periodic set for some L, as a homogeneous set.

Fix some open neighborhood Ω of the identity in G with compact closure. For any L-periodic point $x \in X$ define

(1.1)
$$\operatorname{vol}(Lx) = \frac{m_L(Lx)}{m_L(\Omega)}$$

where m_L is any Haar measure on L. We will use this notion of volume as a measure of the complexity of the periodic orbit.

Evidently this notion depends on Ω , but the notions arising from two different choices of Ω are comparable to each other, in the sense that their ratio is bounded above and below. Consequently, we drop the dependence

on Ω in the notation. See [24, §2.3] for a discussion of basic properties of the above definition.

The general theme of statements will be a dichotomy of the following nature. Unless there is an explicit obstruction with *low complexity*, the orbit Hx fills up the whole space – the statements also provide rates for this density or equidistribution whenever available.

We have tried to arrange the results roughly in their chronological order.

2. Horospherical subgroups

Let **G** be a semisimple \mathbb{R} -group and let G denote the connected component of the identity in the Lie group $\mathbf{G}(\mathbb{R})$.

A subgroup $U \subset G$ is called a horospherical subgroup if there exists an (\mathbb{R} -diagonalizable) element $a \in G$ so that

$$U = W^+(a) := \{ g \in G : a^n g a^{-n} \to e \text{ as } n \to -\infty \};$$
 put $W^-(a) = W^+(a^{-1}).$

Horospherical subgroups are always unipotent, however, not necessarily vice versa, e.g. for $d \geq 3$, a one parameter unipotent subgroup in $\operatorname{SL}_d(\mathbb{R})$ is never horospherical. In a sense, horospherical subgroups are large unipotent subgroups. E.g. if $U \subset G$ is a horospherical subgroup, then $G/\operatorname{N}_G(U)$ is compact, where $\operatorname{N}_G(U)$ denotes the normalizer of U in G.

Let $\Gamma \subset G$ be a lattice and $X = G/\Gamma$. We fix a horospherical subgroup $U = W^+(a)$ for the rest of the discussion. The action of U on X has been the subject of extensive investigations by several authors – when $G = \mathrm{SL}_2(\mathbb{R})$ or more generally G has \mathbb{R} -rank one, this action induces the horocycle flow when $G = \mathrm{SL}_2(\mathbb{R})$, or horospherical flow in the general setting of rank one groups.

Various rigidity results in this context are known thanks to the works of Hedlund, Furstenberg, Margulis, Veech, Dani, Sarnak, Burger and others [44, 40, 61, 80, 14, 15, 18, 73, 12]. Many of these results and subsequent works use techniques which in addition to proving strong rigidity results, do so with a *polynomially* strong error term, e.g. the methods in [73, 12, 75, 78] relying on harmonic analysis or the more dynamical arguments in [61, 48, 81]; see Theorems 2.1 and 2.2 below for some examples.

Let $U_0 \subset U$ be a fixed neighborhood of the identity in U with smooth boundary and compact closure, e.g. one can take U_0 to be the image under the exponential map of a ball around 0 in Lie(U) with respect to some Euclidean norm on Lie(U). For every $k \in \mathbb{N}$, put $U_k = a^k U_0 a^{-k}$.

We normalize the Haar measure, σ , on U so that $\sigma(U_0) = 1$.

2.1. **Theorem.** Assume X is compact. There exists some $\delta > 0$, depending on Γ , so that the following holds. Let $f \in C^{\infty}(X)$, then for any $x \in X$ we have

$$\left| \frac{1}{\sigma(U_k)} \int_{U_k} f(ux) \, \mathrm{d}\sigma(u) - \int_X f \, \mathrm{d}\mathrm{vol}_X \right| \ll_X \mathcal{S}(f) e^{-\delta k},$$

where S(f) denotes a certain Sobolev norm.

The constant δ depends on the rate of decay for matrix coefficients corresponding to smooth vectors in $L^2(X, \operatorname{vol}_X)$, in other words, on the rate of mixing for the action of a on X. In particular, if G has property (T) or Γ is a congruence lattice, then δ can be taken to depend only on dim G.

As mentioned above there are different approaches to prove Theorem 2.1. We highlight a dynamical approach which is based on the mixing property of the action of a on X via the so called *thickening* or *banana* technique; this idea is already present in [61], the exposition here is taken from [48].

Making a change of variable, and using $\sigma(U_0) = 1$, one has

$$\frac{1}{\sigma(U_k)} \int_{U_k} f(ux) \, \mathrm{d}\sigma(u) = \int_{U_0} f(a^k uy) \, \mathrm{d}\sigma(u)$$

where $y = a_{-k}x$.

The key observation now is that the translation of U_0 by a^k is quite well approximated by the translation of a *thickening* of U by a^k . To be more precise, let B be an open neighborhood of the identity so that

$$B = (B \cap W^{-}(a))(B \cap Z_G(a))U_0.$$

Then since $a^k(B \cap W^-(a))a^{-k} \to e$ in the Hausdorff topology, we see that a^kU_0y and a^kBy stay near each other. This, in view of the fact that y stays in the compact set X, reduces the problem to the study of the correlation

$$\int_X \mathbb{1}_B(z) f(a^k z) \, \mathrm{dvol}_X;$$

which can be controlled using the mixing rate for the action of a on X.

As the above sketch indicates, compactness of X is essential for this unique ergodicity result (with a uniform rate) to hold. If X is not compact, there are intermediate behaviors which make the analysis more involved – for instance, if x lies on a closed orbit of U, or is very close to such an orbit, Theorem 2.1 as stated cannot hold. We state a possible formulation in a concrete setting, see [75, 78] for different formulations.

- 2.2. **Theorem.** Let $G = \mathrm{SL}_d(\mathbb{R})$ and $\Gamma = \mathrm{SL}_d(\mathbb{Z})$. There exists some $\delta > 0$ so that the following holds. For any $x = g\Gamma \in X$ and $n, k \in \mathbb{N}$ with k > n at least one of the following holds.
 - (1) For any $f \in C_c^{\infty}(X)$ we have

$$\left| \frac{1}{\sigma(U_k)} \int_{U_k} f(ux) d\sigma(u) - \int_X f dvol_X \right| \ll_d \mathcal{S}(f) e^{-\delta n},$$

where S(f) denotes a certain Sobolev norm.

(2) There exists a rational subspace $W \subset \mathbb{R}^d$ of dimension

$$m \in \{1, \dots, d-1\}$$

so that

$$||ug\mathbf{w}|| \ll_d e^n$$
 for all $u \in U_k$,

where $\mathbf{w} = w_1 \wedge \ldots \wedge w_m$ for a \mathbb{Z} -basis $\{w_1, \ldots, w_m\}$ of $W \cap \mathbb{Z}^n$, and $\|\cdot\|$ is a fixed norm on $\bigwedge^m \mathbb{R}^d$.

Similar results hold for any semisimple group G. In the more general setting, Theorem 2.2(2) needs to be stated using conjugacy classes of a finite collection of parabolic subgroups of G which describe the non-compactness (roughly speaking the cusp) of X.

The proof of Theorem 2.2 combines results on quantitative non-divergence of unipotent flows [55, 16, 17, 21, 49], together with the above sketch of the proof of Theorem 2.1, see [50] and [63].

Recall that a subgroup $H \subset G$ is called symmetric if H is the set of fixed points of an involution τ on G, e.g. H = SO(p, n-p) in $G = SL_n(\mathbb{R})$. Translations of closed orbits of symmetric subgroups presents another (closely related) setting where effective equidistribution results, with polynomial error rates, are available. In this case, as well, the so called wave front lemma [34, Thm. 3.1], asserts that translations of an H-orbit stay near translations of a thickening of it. Therefore, one may again utilize mixing; see, e.g. [34, 4]. Analytic methods also are applicable in this setting, see [23].

We end this section by discussing a case which is beyond the horospherical case, but is closely related. Let $\hat{G} = G \ltimes W$ where G is a semisimple group as above and W is the unipotent radical of \hat{G} . Let $\hat{\Gamma} \subset \hat{G}$ be a lattice and put $\hat{X} = \hat{G}/\hat{\Gamma}$. Let $\pi: \hat{G} \to G$ be the natural projection.

2.3. **Problem.** Let $U \subset \hat{G}$ be a unipotent subgroup so that $\pi(U)$ is a horospherical subgroup of G. Prove analogues of Theorem 2.2 for the action of U on \hat{X} .

Strömbergsson [79], used analytic methods to settle a special case of this problem, i.e., $G = \mathrm{SL}_2(\mathbb{R}) \ltimes \mathbb{R}^2$ with the standard action of $\mathrm{SL}_2(\mathbb{R})$ on \mathbb{R}^2 , $\Gamma = \mathrm{SL}_2(\mathbb{Z}) \ltimes \mathbb{Z}^2$, and U the group of unipotent upper triangular matrices in $\mathrm{SL}_2(\mathbb{R})$; his method has also been used to tackle some other cases.

3. Effective equidistribution theorems for nilflows

In this section we assume G is a unipotent group. That is: we may assume G is a closed connected subgroup of the group of strictly upper triangular $d \times d$ matrices. Let $\Gamma \subset G$ be a lattice and $X = G/\Gamma$, i.e., X is a nilmanifold.

Rigidity results in this setting have been known for quite some time thanks to works of Weyl, Kronecker, L. Green, and Parry [3, 67], and more recently Leibman [51].

Quantitative results, with a polynomial error rate, have also been established in this context and beyond the abelian case, see [39, 42]. The complete solution was given by B. Green and Tao [42]; here we present a special case from loc. cit. in this setting describing the equidistribution properties of pieces of trajectories.

- 3.1. **Theorem** ([42]). Let $X = G/\Gamma$ be a nilmanifold as above. There exists some $A \geq 1$ depending on dim G so that the following holds. Let $x \in X$, let $\{u(t): t \in \mathbb{R}\}$ be a one parameter subgroup of G, let $0 < \eta < 1/2$, and let T > 0. Then at least one of the following holds for the partial trajectory $\{u(t)x: t \in [0,T]\}$.
 - (1) For every $f \in C^{\infty}(X)$ we have

$$\left| \frac{1}{T} \int_0^T f(u(t)x) \, \mathrm{d}t - \int_X f \, \mathrm{d}\mathrm{vol}_X \right| \ll_{X,f} \eta$$

where the dependence on f is given using a certain Lipschitz norm.

(2) For every $0 \le t_0 \le T$ there exists some $g \in G$ and some $H \subsetneq G$ so that $H\Gamma/\Gamma$ is closed with $\operatorname{vol}(gH\Gamma/\Gamma) \ll_X \eta^{-A}$ and for $t \in [0,T]$ we have

$$|t-t_0| \le \eta^A T \implies \operatorname{dist}(u(t)x, gH\Gamma/\Gamma) \ll_X \eta$$

where dist is a metric on X induced from a right invariant Riemannian metric on G.

This is a consequence of a special case of a more general effective equidistribution result for polynomial trajectories on nilmanifolds [42, Thm. 2.9], as we now explicate. Since T>0 is arbitrary, we may assume that $u(t)=\exp(tz)$ for some z in the Lie algebra of G of norm one. We note that for $T\leq \eta^{-O(1)}$ the above is trivial. In fact, as is visible in the maximal abelian torus quotient every point belongs to an orbit $gH\Gamma/\Gamma$ of bounded volume of a proper subgroup $H\subsetneq G$ and now (2) follows by the continuity properties of the one-parameter subgroup if $T\leq \eta^{-O(1)}$. Hence we will assume in the following $T>\eta^{-O(1)}$ for a constant O(1) which will be optimized.

For every $\frac{1}{2} \leq \tau \leq 1$, put $B_{\tau} = \{u(n\tau) : n = 0, 1, \dots, N_{\tau}\}$ where $N_{\tau} = \lfloor T/\tau \rfloor$. We now apply [42, Thm. 2.9] for the sequence (discrete trajectory) B_{τ} . Assume first that $B_{\tau}x$ is η -equidistributed, for some $\tau \in [\frac{1}{2}, 1]$. That is:

(3.1)
$$\left| \frac{1}{N_{\tau}} \sum_{n=0}^{N_{\tau}-1} h(u(n\tau)x) - \int_{X} h \operatorname{dvol}_{X} \right| \ll_{X,h} \eta$$

for all $h \in C^{\infty}(X)$. In this case Theorem 3.1(1) holds. Indeed

$$\frac{1}{N_{\tau}\tau} \int_0^{N_{\tau}\tau} f(u(t)x) dt = \frac{1}{\tau} \int_0^{\tau} \frac{1}{N_{\tau}} \sum_{n=0}^{N_{\tau}-1} f(u(s)u(n\tau)x) ds,$$

so the claim in (1) follows from (3.1) applied with $h(\cdot) = f(u(s)\cdot)$ for all $0 \le s \le \tau$.

The alternative in [42, Thm. 2.9] to η -equidistribution as above is an obstruction to equidistribution in the form of a slowly varying character of G/Γ . To make a precise statement we need some notation. Fix a rational basis for Lie(G). Using this basis we put coordinates (also known as coordinates of the second kind) on G; the standing assumption is that Γ corresponds to elements with integral coordinates, see [42, Def. 2.1 and 2.4]

– the estimates¹ will depend on the *complexity* of the structural constants for group multiplication written in this basis (which we assume to be fixed). Following [42] we denote the coordinates of $g \in G$ by $\psi(g)$. In this notation, given a character $\chi: G \to \mathbb{R}/\mathbb{Z}$ with $\Gamma \subset \ker(\chi)$, there exists a unique $k_{\chi} \in \mathbb{Z}^{\dim G}$ so that $\chi(g) = k_{\chi} \cdot \psi(g) + \mathbb{Z}$, see [42, Def. 2.6].

Assume (3.1) fails for all $\tau \in [\frac{1}{2}, 1]$. Then, by [42, Thm. 2.9], see also [42, Lemma 2.8], we have: there are constants $A_0, A_1 > 1$, and for every τ there is a character $\chi_{\tau} : G \to \mathbb{R}/\mathbb{Z}$ with $\Gamma \subset \ker(\chi_{\tau})$ so that the following two conditions hold.

- (a) Let $k_{\tau} \in \mathbb{Z}^{\dim G}$ be so that $\chi_{\tau}(g) = k_{\tau} \cdot \psi(g) + \mathbb{Z}$, then we have the bound $||k_{\tau}|| \ll_{G,\Gamma} \eta^{-A_0}$,
- (b) $\|\chi_{\tau}(u(\tau))\|_{\mathbb{R}/\mathbb{Z}} \ll_{G,\Gamma} \eta^{-A_1}/T$, where $\|x\|_{\mathbb{R}/\mathbb{Z}} = \operatorname{dist}(x,\mathbb{Z})$.

Let H_{τ} denote the connected component of the identity in $\ker(\chi_{\tau})$. Informally, (a) tells us that χ_{τ} defines a closed orbit $H_{\tau}\Gamma/\Gamma$ of not too large volume – indeed the latter covolume is bounded by $||k_{\tau}||$. Moreover, $||k_{\tau}||$ controls the continuity properties of χ_{τ} . On the other hand, (b) tells us that the character changes its values very slowly along the discrete trajectory (since we are allowed to think of a large T). We wish to combine these for various $\tau \in [\frac{1}{2}, 1]$ to obtain (2).

To that end, note that the number of characters χ_{τ} so that (a) holds is $\leq \eta^{-O(1)}$ for some O(1) depending on A_0 . Moreover, (b) implies that there exist some $C = C(G, \Gamma)$, some A_2 depending on A_0 and A_1 , and for every $1/2 \leq \tau \leq 1$ some rational vector v_{τ} with $||v_{\tau}|| \ll 1$ and denominator bounded by $O(\eta^{-O(1)})$ so that the distance of $\psi(u(\tau))$ to $v_{\tau} + \psi(H_{\tau})$ is $< C\eta^{-A_2}/T$. For every $\frac{1}{2} \leq \tau \leq 1$ let I_{τ} be the maximal (relatively open) interval so that for all $s \in I_{\tau}$ the distance of $\psi(u(s))$ to $v_{\tau} + \psi(H_{\tau})$ is $< C\eta^{-A_2}/T$. This gives a covering of $[\frac{1}{2},1]$ with $\eta^{-O(1)}$ -many intervals. Therefore, at least one of these intervals, say $I_0 = (a_0,b_0)$ defined by τ_0 , has length $b_0 - a_0 \gg \eta^{O(1)}$. Let $\chi = \chi_{\tau_0}$. Then for any $\tau \in I_0$ we have that the distance of $\psi(u(\tau))$ to $v_{\tau_0} + \psi(H_{\tau_0})$ is $< C\eta^{-A_2}/T$. Since $b_0 - a_0 \gg \eta^{O(1)}$, we get that the distance of $\psi(u(\tau))$ to $v_{\tau_0} + \psi(H_{\tau_0})$ is $< C\eta^{-A_2}/T$. Since $b_0 - a_0 \gg \eta^{O(1)}$, we get that the distance of $\psi(u(\tau))$ to $v_{\tau_0} + \psi(H_{\tau_0})$ is $< \eta^{-O(1)}/T$ for all $0 \leq \tau \leq 1$. Hence we obtain the character estimate $||\chi(u(\tau))||_{\mathbb{R}/\mathbb{Z}} \ll_X \eta^{-O(1)}/T$ for all $0 \leq \tau \leq 1$.

Let $g \in G$ and $\gamma \in \Gamma$ be so that $u(t_0) = g\gamma$ and $||\psi(g)|| \ll_X 1$. Let $H = H_{\tau_0}$. Then since $\gamma H\Gamma = H\Gamma$, the claim in (2) holds with g and H if we choose A large enough.

We now highlight some elements involved in the proof of [42, Thm. 2.9] for our simplified setting of a *linear* sequence, i.e. a discrete trajectory – the reader may also refer to $[42, \S 5]$ where a concrete example is worked out.

Let $\tau = 1$, i.e., consider $\{u(n) : n = 0, 1, \dots, N-1\}$ on X. The goal is to show that either (3.1) holds or (a) and (b) above must hold for a character χ ; note first that replacing $\{u(t)\}$ by a conjugate we will assume x in (3.1)

¹Indeed the estimates in part (1) also depends on this basis.

is the identity coset, Γ , for the rest of discussion. The proof is based on an inductive argument² which aims at decreasing the nilpotency degree of G. For the base of the induction, i.e., when G is abelian, one may use Fourier analysis to deduce the result, e.g. see [42, Prop. 3.1].

The following corollary, see [42, Cor. 4.2], of the van der Corput trick plays an important role in the argument. Let $\{a_n:n=0,1,\ldots,N-1\}$ be a sequence of complex numbers so that $\frac{1}{N}\sum_{n=0}^{N-1}a_n\geq\eta$. Then for at least $\eta^2N/8$ values of $k\in\{0,1,\ldots,N-1\}$, we have $\frac{1}{N}\sum_{n=0}^{N-1}a_{n+k}\bar{a}_n\geq\eta^2/8$ where we put $a_n:=0$ for $n\notin\{0,1,\ldots,N-1\}$.

Let $a_n = h(u(n)\Gamma)$, and suppose that (3.1) fails. One may further restrict to the case where h is an eigenfunction for the action of the center of G corresponding to a character ξ whose complexity is controlled by $\eta^{-O(1)}$, see [42, Lemma 3.7]. Note that if ξ is trivial, then h is Z(G)-invariant; thus we have already reduced the problem to G/Z(G), i.e, a group with smaller nilpotency degree. Hence, assume that ξ is nontrivial, in consequence $\int h \, d\text{vol}_X = 0$. Now by the aforementioned corollary of the van der Corput trick, there are at least $\eta^2 N/8$ many choices of $0 \le k \le N$ so that

(3.2)
$$\frac{1}{N} \sum_{n=0}^{N-1} h(u(n+k)\Gamma) \overline{h(u(n)\Gamma)} \gg \eta^2/8.$$

Fix one such k and write $u(k)\Gamma = v\Gamma$ for an element v in the fundamental domain of Γ ; note that $\{(v^{-1}gv,g):g\in G\}\subset\{(g_1,g_2)\in G\times G:g_1g_2^{-1}\in [G,G]\}=:G'$. Similarly, define Γ' . Two observations are in order.

- (3.2) implies that $\frac{1}{N} \sum_{n=0}^{N-1} \tilde{h}(w_n \Gamma') \gg \eta^2/8$ where \tilde{h} is the restriction of $\hat{h}(y,z) = h(vy) h(z)$ to G' and $w_n = (v^{-1}u(n)v, u(n))$.
- since h is an eigenfunction for the center of G, we have \tilde{h} is invariant under $\Delta(\mathrm{Z}(G)) = \{(g,g) : g \in \mathrm{Z}(G)\}$, moreover, \tilde{h} has mean zero.

The above observations reduce (3.2) modulo $\Delta(Z(G))$, i.e., to the group $G'/\Delta(Z(G))$ which has smaller nilpotency degree, see [42, Prop. 7.2]. There is still work to be done, e.g. one needs to combine the information obtained for different values of k to prove (a) and (b); but this reduction, in a sense, is the heart of the argument.

4. Periodic orbits of semisimple groups

Beyond the settings which were discussed in §2 and §3, little was known until roughly a decade ago. The situation drastically changed thanks to the work of Einsiedler, Margulis, and Venkatesh [25], where a *polynomially effective* equidistribution result was established for closed orbits of semisimple groups.

We need some notation in order to state the main result. Let **G** be a connected, semisimple algebraic \mathbb{Q} -group, and let G be the connected component of the identity in the Lie group $\mathbf{G}(\mathbb{R})$. Let $\Gamma \subset \mathbf{G}(\mathbb{Q})$ be a

²The reader may also see the argument in [67, §3].

congruence lattice in G and put $X = G/\Gamma$. Suppose $H \subset G$ is a semisimple subgroup without any compact factors which has a finite centralizer in G.

The following is the main equidistribution theorem proved in [25].

4.1. **Theorem** ([25]). There exists some $\delta = \delta(G, H)$ so that the following holds. Let Hx be a periodic H-orbit. For every V > 1 there exists a subgroup $H \subset S \subset G$ so that x is S-periodic, $vol(Sx) \leq V$, and

$$\left| \int_X f \, \mathrm{d}\mu_{Hx} - \int_X f \, \mathrm{d}\mu_{Sx} \right| \ll_{G,\Gamma,H} \mathcal{S}(f) V^{-\delta} \quad \text{for all } f \in C_c^{\infty}(X),$$

where S(f) denotes a certain Sobolev norm.

Theorem 4.1 is an effective version (of a special case) of a theorem by Mozes and Shah [65]. The general strategy of the proof is based on effectively acquiring extra almost invariance properties for the measure $\mu = \mu_{Hx}$. This general strategy (in qualitative form) was used by in the topological context by Margulis [56], Dani and Margulis [19] and also by Ratner in her measure classification theorem [69, 70].

The polynomial nature of the error term, i.e., a (negative) power of V, in Theorem 4.1 is quite remarkable – effective dynamical arguments often yield worse rates, see §5. A crucial input in the proof of Theorem 4.1, which is responsible for the quality of the error, is a uniform spectral gap for congruence quotients.

The proof of Ratner's Measure Classification Theorem for the action of a semisimple group H is substantially simpler; a simplified proof in this case was given by Einsiedler [28], see also [25, §2]. This is due to complete reducibility of the adjoint action of H on Lie(G) as we now explicate. Let $\{u(t):t\in\mathbb{R}\}$ be a one parameter unipotent subgroup in H, and let L be a subgroup which contains H. Then one can show that the orbits u(t)y and u(t)z of two nearby points in general position diverge in a direction transversal to L. This observation goes a long way in the proof. Indeed starting from an H-invariant ergodic measure μ , one may use arguments like this to show that unless there is an algebraic obstruction, one can increase the dimension of the group which leaves μ invariant. In a sense, the argument in [25] is an effective version of this argument – a different argument which is directly based on the mixing property of an \mathbb{R} -diagonalizable subgroup in H was given by Margulis, see [64].

Let us elaborate on a possible effectivization of the above idea. The divergence of two nearby points u(t)y and u(t)z is governed by a polynomial function; e.g, in the setting at hand, write $y = \exp(v)z$ for some $v \in \text{Lie}(G)$, then this divergence is controlled by Ad(u(t))v. In consequence, one has a rather good quantitative control on this divergence.

However, the size of T so that the piece of orbit, $\{u(t)y : 0 \le t \le T\}$, approximates the measure μ depends on y. More precisely, suppose μ is

 $\{u(t)\}$ -ergodic³. Then it follows from Birkhoff's ergodic theorem that μ -a.e. y and all $f \in C_c^{\infty}(X)$ we have $\frac{1}{T} \int_0^T f(u(t)y) dt \to \int_X f d\mu_{Hx}$. However, for a given $\epsilon > 0$ the size of T so that

$$\left| \frac{1}{T} \int_0^T f(u_t y) \, \mathrm{d}t - \int_X f \, \mathrm{d}\mu_{Hx} \right| \ll_f \epsilon \quad \text{where } f \in C_c^\infty(X)$$

depends on delicate properties of the point y, e.g. y may be too close to a $\{u(t)\}$ -invariant submanifold in support of μ .

One of the remarkable innovations in [25] is the use of uniform spectral gap in order to obtain an effective version of the pointwise ergodic theorem. The required uniform spectral gap has been obtained in a series of papers [46, 66, 76, 45, 11, 13, 41]. This is then used to define an effective notion of generic points where the parameters ϵ and T above are polynomially related to each other.

If H is a maximal subgroup, one can use the above (combined with bounded generation of G by conjugates of H and spectral gap for vol_X) to finish the proof. However, Theorem 4.1 is more general and allows for (finitely many) intermediate subgroups. The main ingredient in [25] to deal with possible intermediate subgroups is an *effective closing lemma* that is proved in [25, §13]. In addition to being crucial for the argument in [25], this result is of independent interest – it is worth mentioning that the proof of [25, Prop. 13.1] also uses spectral gap.

Theorem 4.1 imposes some assumptions that are restrictive for some applications: H is not allowed to vary, moreover, H has a finite centralizer. The condition that H is assumed not to have any compact factors is a splitting condition at the infinite place; this too is restrictive in some applications.

4.2. **Adelic periods.** In a subsequent work by Einsiedler, Margulis, Mohammadi, and Venkatesh [24], the subgroup H in Theorem 4.1 is allowed to vary. Moreover, the need for a splitting conditions is also eliminated. The main theorem in [24] is best stated using the language of adeles; the reader may also see [26] for a more concrete setting.

Let **G** be a connected, semisimple, algebraic \mathbb{Q} -group⁴ and set $X = \mathbf{G}(\mathbb{A})/\mathbf{G}(\mathbb{Q})$ where \mathbb{A} denotes the ring of adeles. Then X admits an action of the locally compact group $\mathbf{G}(\mathbb{A})$ preserving the probability measure vol_X .

Let **H** be a semisimple, simply connected, algebraic \mathbb{Q} -group, and let $q \in \mathbf{G}(\mathbb{A})$. Fix also an algebraic homomorphism

$$\iota: \mathbf{H} \to \mathbf{G}$$

³By the generalized Mautner phenomenon, when H has no compact factors, one can always choose $\{u(t)\}$ so that μ is $\{u(t)\}$ -ergodic

⁴The paper [24] allows for any number field, F, but unless X is compact, δ in Theorem 4.3 will depend on dim \mathbf{G} and $[F:\mathbb{Q}]$.

defined over \mathbb{Q} with finite central kernel. E.g. we could have $\mathbf{G} = \mathrm{SL}_d$ and $\mathbf{H} = \mathrm{Spin}(Q)$ for an integral quadratic form Q in d variables.

To this algebraic data, we associate a homogeneous set

$$Y := g\iota(\mathbf{H}(\mathbb{A})/\mathbf{H}(\mathbb{Q})) \subset X$$

and a homogeneous measure μ ; recall that we always assume $\mu(Y) = 1$. The following is a special case of the main theorem in [24].

4.3. **Theorem** ([24]). Assume further that G is simply connected. There exists some $\delta > 0$, depending only on dim G, so that the following holds. Let Y be a homogeneous set and assume that $\iota(H) \subset G$ is maximal. Then

$$\left| \int_X f \, \mathrm{d}\mu - \int_X f \, \mathrm{d}\mathrm{vol}_X \right| \ll_{\mathbf{G}} \mathcal{S}(f) \mathrm{vol}(Y)^{-\delta} \quad \text{ for all } f \in C_c^{\infty}(X),$$

where S(f) is a certain adelic Sobolev norm.

As we alluded to above Theorem 4.3 allows H to vary, it also assumes no splitting conditions⁵ on \mathbf{H} ; this feature is also crucial for applications, e.g, $\mathbf{H}(\mathbb{R})$ is compact in an application to quadratic forms which will be discussed momentarily. These liberties are made possible thanks to Prasad's volume formula [68], and the seminal work of Borel and Prasad [8]. Roughly speaking, the argument in [24, §5] uses [68] and [8] to show that if at a prime p the group $\mathbf{H}(\mathbb{Q}_p)$ is either compact or too distorted, then there is at least a factor p contribution to $\operatorname{vol}(Y)$. Thus one can find a small prime p (compared to $\operatorname{vol}(Y)$) where H is not distorted.

The dynamical argument uses unipotent flows as described above – the source of a polynomially effective rate is again the uniform spectral gap.

Let us highlight two corollaries from Theorem 4.3. The method in [24] relies on uniform spectral gap. However, it provides an independent proof of property (τ) except for groups of type A_1 – i.e., if we only suppose property (τ) for groups of type A_1 , we can deduce property (τ) in all other cases as well as our theorem. In particular it gives an alternative proof of the main result of Clozel in [13] but with weaker exponents, see [24, §4].

Another application is an analogue of Duke's theorem for positive definite integral quadratic forms in $d \geq 3$ variables as we now explicate. Even in a qualitative form this result is new in dimensions 3 and 4 since the splitting condition prevented applying unipotent dynamics before.

Let $\mathcal{Q}_d = \operatorname{PO}_d(\mathbb{R}) \backslash \operatorname{PGL}_d(\mathbb{R}) / \operatorname{PGL}_d(\mathbb{Z})$ be the space of positive definite quadratic forms on \mathbb{R}^d up to the equivalence relation defined by scaling and equivalence over \mathbb{Z} . We equip \mathcal{Q}_d with the push-forward of the normalized Haar measure on $\operatorname{PGL}_d(\mathbb{R}) / \operatorname{PGL}_d(\mathbb{Z})$.

Let Q be a positive definite integral quadratic form on \mathbb{Z}^d , and let genus(Q) (resp. spin genus(Q)) be its *genus* (resp. spin genus).

⁵Compare this to the assumption that H has no compact factors in Theorem 4.1.

4.4. **Theorem** ([24]). Suppose $\{Q_n\}$ varies through any sequence of pairwise inequivalent, integral, positive definite quadratic forms. Then the genus (and also the spin genus) of Q_n , considered as a subset of Q_d , equidistributes as $n \to \infty$ (with speed determined by a power of $|\operatorname{genus}(Q)|$).

In the statement of Theorem 4.3 we made a simplifying assumption that G is simply connected; PGL_d , however, is not simply connected. Indeed the proof of Theorem 4.4 utilizes the more general [24, Thm. 1.5]. In addition one uses the fact that

$$\operatorname{PGL}_d(\mathbb{A}) = \operatorname{PGL}_d(\mathbb{R}) K \operatorname{PGL}_d(\mathbb{Q})$$
 where $K = \prod_p \operatorname{PGL}_d(\mathbb{Z}_p)$

to identify $L^2(\operatorname{PGL}_d(\mathbb{R})/\operatorname{PGL}_d(\mathbb{Z}))$ with the space of K-invariant functions in $L^2(\operatorname{PGL}_d(\mathbb{A})/\operatorname{PGL}_d(\mathbb{Q}))$.

Similar theorems have been proved elsewhere (see, e.g. [38] where the splitting condition is made at the archimedean place). What is novel here, besides the speed of convergence, is the absence of any type of splitting condition on the $\{Q_n\}$ – this is where the effective Theorem 4.3 becomes useful.

Theorem 4.3 assumed $\iota(\mathbf{H}) \subset \mathbf{G}$ is maximal. This is used in several places in the argument. This assumption is too restrictive for some applications; see, e.g. [31] where $\iota(\mathbf{H}) \subset \mathbf{G}$ has infinite centralizer. The following is a much desired generalization.

4.5. **Problem.** Prove an analogue of Theorem 4.3 allowing $\iota(\mathbf{H})$ to have arbitrary centralizer in \mathbf{G} .

The uniform spectral gap, which is the source of a polynomially effective error term, is still available in this case. However, in the presence of an infinite centralizer closed orbits come in families; moreover, there is an abundance of intermediate subgroups. These features introduce several technical difficulties.

See [27] in this volume and also [26, 1] for some progress toward this problem.

5. The action of unipotent subgroups

We now turn to the general case of unipotent trajectories. Let \mathbf{G} be an \mathbb{R} -group and let G be the connected component of the identity in the Lie group $\mathbf{G}(\mathbb{R})$; U will denote a unipotent subgroup of G. Let $\Gamma \subset G$ be a lattice and $X = G/\Gamma$.

Let us recall the following theorems of Ratner which resolved conjectures of Raghunathan and Dani.

- 5.1. **Theorem** ([69, 70, 71]). (1) Every U-invariant and ergodic probability measure on X is homogeneous.
 - (2) For every $x \in X$ the orbit closure \overline{Ux} is a homogeneous set.

The above actually holds for any group which is generated by unipotent elements. In the case of a unipotent one-parameter subroup U more can be said. Suppose $\overline{Ux} = Lx$ as in Theorem 5.1(2), Ratner [71] actually proved that the orbit Ux is equidistributed with respect to the L-invariant measure on Lx.

Prior to Ratner's work, some important special cases were studied by Margulis [56], and Dani and Margulis [19, 20]. The setup they considered was motivated by Margulis' solution to the Oppenheim conjecture – unlike Ratner's work, their method is topological and does not utilize measures. Let $G = \mathrm{SL}_3(\mathbb{R})$, let $\Gamma = \mathrm{SL}_3(\mathbb{Z})$, and let U be a generic one parameter unipotent⁶ subgroup of G. In this context, the paper [20], proves that \overline{Ux} is homogeneous for all $x \in X$.

Theorem 5.1 has also been generalized to the S-arithmetic context, i.e., product of real and p-adic groups, independently by Margulis and Tomanov [57], and Ratner [72].

Let us recall the basic strategy in the proof of Raghunathan's conjectures, see also the discussion after Theorem 4.1. The starting point, á la Margulis and Ratner, is a set of "generic points" (a dynamical notion) for our unipotent group U. The heart of the matter then is to carefully investigates divergence of the U-orbits of two nearby generic points; slow (polynomial like) nature of this divergence implies that nearby points diverge in directions that are stable under the action of U. I.e., the divergence is in the direction of the normalizer of U – this is in sharp contrast to hyperbolic dynamics where points typically diverge along the unstable directions for the flow. The goal is to conclude that unless some explicit algebraic obstructions exist, the closure of a U-orbit contains an orbit of a subgroup $V \supseteq U$.

As was mentioned above, the slow divergence of unipotent orbits is a major player in the analysis. This actually is not the only place where polynomial like behavior of unipotent actions is used in the proofs. Indeed in passing from measure classification to topological rigidity (and more generally equidistribution theorem) non-divergence of unipotent orbits [55, 16, 17], plays an essential role, see also §5.4.

Generic sets play a crucial role in the above study, and the existing notions, e.g. minimal sets in the topological approach or a generic set for Birkhoff's ergodic theorem in Ratner's argument, are rather non-effective. Providing an effective notion of a generic set which is also compatible with the nice algebraic properties of unipotent flows is the first step towards an effective account of the above outline. With that in place one then may try to carry out the above analysis in an effective fashion. The caveat though is that the estimates one gets from such arguments are usually rather poor, i.e., rather than obtaining a negative power of complexity one typically gets a negative power of an iterated logarithmic function of the complexity; see

⁶A one parameter unipotent subgroup of $SL_d(\mathbb{R})$ is called generic if it is contained in only one Borel subgroup of $SL_d(\mathbb{R})$

the discussions in §4 for an instance where this argument is carried out successfully and actually with a polynomial rate.

5.2. Effective versions of the Oppenheim conjecture. The resolution of the Oppenheim conjecture by Margulis [56], has played a crucial role in the developments of the field.

Let us recall the setup. Let Q be a non-degenerate, indefinite quadratic form in $d \geq 3$ variables on \mathbb{R}^d . The Oppenheim and Davenport conjecture stated that $\overline{Q(\mathbb{Z}^d)} = \mathbb{R}$ if and only if Q is not a multiple of a form with integral coefficients. Quantitative (or equidistribution) versions were also obtained [32, 33, 62, 59]; see also [74, 9, 2] where effective results for generic forms (in different parameter spaces) are obtained.

On an effective level one might ask the following question. Given $\epsilon > 0$ what is the smallest $0 \neq v \in \mathbb{Z}^d$ so that $|Q(v)| \leq \epsilon$. Analytic methods, which were used prior to Margulis' work to resolve special cases of the Oppenheim conjecture, yield such estimates. Marguli's proof, however, is based on dynamical ideas and does not provide information on the size of such solution.

The paper [43] proves a polynomial estimate for $n \geq 5$ and under explicit Diophantine conditions on Q – this paper combines analytic methods together with some ideas related to systems of inequalities which were developed in [32]. In [74] and [9] analytic methods are used to obtain polynomial estimates for almost every form in certain families of forms in dimensions 3 and 4.

In general, however, the best known results in dimension 3 are due to Lindenstrauss and Margulis as we now discuss.

5.3. **Theorem** ([54]). There exist absolute constants $A \ge 1$ and $\alpha > 0$ so that the following holds.

Let Q be an indefinite, ternary quadratic form with $\det Q = 1$ and $\epsilon > 0$. Then for any $T \geq T_0(\epsilon) ||Q||^A$ at least one of the following holds.

(1) For any $\xi \in [-(\log T)^{\alpha}, (\log T)^{\alpha}]$ there is a primitive integer vector $v \in \mathbb{Z}^3$ with $0 < ||v|| < T^A$ satisfying

$$|Q(v) - \xi| \ll (\log T)^{-\alpha}.$$

(2) There is an integral quadratic form Q' with $|\det Q'| < T^{\epsilon}$ so that

$$||Q - \lambda Q'|| \ll ||Q||T^{-1}$$

where
$$\lambda = |\det Q'|^{-1/3}$$
.

The implied multiplicative constants are absolute constants.

The above theorem provides a dichotomy: unless there is an explicit Diophantine (algebraic) obstruction, a density result holds; in this sense the result is similar to Theorem 2.2 and Theorem 3.1. Note, however, that the quality of the effective rate one obtains here is $(\log T)^{-\alpha}$ – ideally one would like to have a result where $(\log T)^{-\alpha}$ is replaced by $T^{-\alpha}$, however, such an improvement seems to be out of the reach of the current technology.

We now highlight some of the main features of the proof of Theorem 5.3. An important ingredient in the proof is an explicit Diophantine condition [54, §4]; this is used in place of the notion of minimal sets which was used in [56, 19, 20]. We will discuss a related Diophantine condition in the next section; one important feature of the notion used in [54] is that it is inherited for most points along a unipotent orbit, see also Theorem 5.8.

The proof in [54] then proceed by making effective and improving on several techniques from [56, 19, 20]. If one follows this scheme of the proof, the quality of the estimates in Theorem 5.3(1) would be $(\log \log T)^{-\alpha}$. Instead [54] uses a combinatorial lemma about rational functions to *increase* the density of points, see [54, §9] – this lemma, which is of independent interest, is responsible for the better error rate in Theorem 5.3(1).

5.4. Effective avoidance principles for unipotent orbits. Let \mathbf{G} be a connected \mathbb{Q} -group and put $G = \mathbf{G}(\mathbb{R})$. We assume Γ is an arithmetic lattice in G. More specifically, we assume fixed an embedding $\iota : \mathbf{G} \to \mathrm{SL}_d$, defined over \mathbb{Q} so that $\iota(\Gamma) \subset \mathrm{SL}_d(\mathbb{Z})$. Using ι we identify \mathbf{G} with $\iota(\mathbf{G}) \subset \mathrm{SL}_d$ and hence $G \subset \mathrm{SL}_d(\mathbb{R})$. Let $U = \{u(t) : t \in \mathbb{R}\} \subset G$ be a one parameter unipotent subgroup of G and put $X = G/\Gamma$.

Define the following family of subgroups

$$\mathcal{H} = \{ \mathbf{H} \subset \mathbf{G} : \mathbf{H} \text{ is a connected } \mathbb{Q}\text{-subgroup and } \mathbf{R}(\mathbf{H}) = \mathbf{R}_u(\mathbf{H}) \}$$

where $R(\mathbf{H})$ (resp. $R_u(\mathbf{H})$) denotes the solvable (resp. unipotent) radical of \mathbf{H} . Alternatively, $\mathbf{H} \in \mathcal{H}$ if and only if \mathbf{H} is a connected \mathbb{Q} -subgroup and $\mathbf{H}(\mathbb{C})$ is generated by unipotent subgroups. We will always assume that $\mathbf{G} \in \mathcal{H}$.

For any $\mathbf{H} \in \mathcal{H}$ we will write $H = \mathbf{H}(\mathbb{R})$; examples of such groups are $H = \mathrm{SL}_d(\mathbb{R})$, $\mathrm{SL}_d(\mathbb{R}) \ltimes \mathbb{R}^d$ (with the standard action), and $\mathrm{SO}_d(\mathbb{R})$. By a theorem of Borel and Harish-Chandra, $H \cap \Gamma$ is a lattice in H for any $\mathbf{H} \in \mathcal{H}$. Define $\mathrm{N}_G(U, H) = \{g \in G : Ug \subset gH\}$. Put

$$\mathcal{S}(U) = \bigg(\bigcup_{\substack{H \in \mathcal{H} \\ H \neq G}} \mathcal{N}_G(U, H)\bigg) / \Gamma \text{ and } \mathcal{G}(U) = X \setminus \mathcal{S}(U)$$

Following Dani and Margulis [22], points in $\mathcal{S}(U)$ are called *singular* with respect to U, and points in $\mathcal{G}(U)$ are called *generic* with respect to U. This notion of a generic point is a priori different from measure theoretically generic points for the action of U on X with respect to vol_X ; however, any measure theoretically generic point is generic in this new sense as well.

By Theorem 5.1(2), for every $x \in \mathcal{G}(U)$ we have $\overline{Ux} = X$.

In [22], Dani and Margulis established strong avoidance properties for unipotent orbits, see also [77]. These properties, which are often referred to as *linearization* of unipotent flows in the literature, go hand in hand with Ratner's theorems in many applications, see e.g. [65, 37].

In this section we will state a polynomially effective version of the results and techniques in [22]. These effective results as well as their S-arithmetic generalizations are proved in [52].

The main effective theorem, Theorem 5.8, requires some further preparation. Let us first begin with the following theorems which are corollaries of Theorem 5.8.

5.5. **Theorem** ([52]). There exists a compact subset $\mathcal{K} \subset \mathcal{G}(U)$ with the following property. Let $x \in \mathcal{G}(U)$, then $\overline{Ux} \cap \mathcal{K} \neq \emptyset$.

Theorem 5.5 is a special case of the following.

5.6. **Theorem** ([52]). There exists some D > 1 depending on d, and some E > 1 depending on G, d and Γ so that the following holds.

For every $0 < \eta < 1/2$ there is a compact subset $\mathcal{K}_{\eta} \subset \mathcal{G}(U)$ with the following property. Let $\{x_m\}$ be a bounded sequence of points in X, and let $T_m \to \infty$ be a sequence of real numbers. For each m let $I_m \subset [-T_m, T_m]$ be a measurable set whose measure is $> E\eta^{1/D}(2T_m)$. Then one of the following holds.

- (1) $\overline{\bigcup_{m}\{u(t)x_m:t\in I_m\}}\cap \mathcal{K}_{\eta}\neq\emptyset$, or
- (2) there exists a finite collection $H_1, \ldots, H_r \in \mathcal{H}$ and for each $1 \leq i \leq r$ there is a compact subset $C_i \subset N(U, H_i)$, so that all the limit points of $\{x_m\}$ lie in $\bigcup_{i=1}^r C_i \Gamma / \Gamma$.

Theorem 5.6 is yet another reminiscent of the sort of dichotomy that we have seen in previous sections: unless an explicit algebraic obstruction exists, the pieces of the U-orbits intersect the generic set; see also Theorem 5.8 where this dichotomy is more apparent.

The polynomial dependence, $E\eta^{1/D}$, in Theorem 5.6 is a consequence of the fact that Theorem 5.8 is *polynomially* effective. We note however that even for Theorem 5.5 it is not clear how it would follow from the statements in [22].

We now fix the required notation to state Theorem 5.8. Let $\|\cdot\|$ denote the maximums norm on $\mathfrak{sl}_N(\mathbb{R})$ with respect to the standard basis; this induces a family of norms, $\|\cdot\|$ on $\bigwedge^m \mathfrak{sl}_N(\mathbb{R})$ for $m=1,2,\ldots$ Let furthermore $\mathfrak{g}=\mathrm{Lie}(G)$ and put $\mathfrak{g}(\mathbb{Z}):=\mathfrak{g}\cap\mathfrak{sl}_N(\mathbb{Z})$.

For any $\eta > 0$, set

$$X_{\eta} = \big\{g\Gamma \in X: \min_{0 \neq v \in \mathfrak{g}(\mathbb{Z})} \|\operatorname{Ad}(g)v\| \geq \eta\big\}.$$

By Mahler's compactness criterion, X_{η} is compact for any $\eta > 0$.

Let $\mathbf{H} \in \mathcal{H}$ be a proper subgroup and put

$$\rho_H := \bigwedge^{\dim \mathbf{H}} \operatorname{Ad} \quad \text{ and } \quad V_H := \bigwedge^{\dim \mathbf{H}} \mathfrak{g}.$$

The representation ρ_H is defined over \mathbb{Q} . Let $\mathbf{v_H}$ be a primitive integral vector in $\wedge^{\dim \mathbf{H}} \mathfrak{g}$ corresponding to the Lie algebra of \mathbf{H} , i.e., we fix a \mathbb{Z} -basis for $\operatorname{Lie}(H) \cap \mathfrak{sl}_N(\mathbb{Z})$ and let $\mathbf{v_H}$ be the corresponding wedge product.

We define the height of $\mathbf{H} \in \mathcal{H}$ by

Given $\mathbf{H} \in \mathcal{H}$ and $n \in \mathbb{N}$, define

$$\operatorname{ht}(\mathbf{H}, n) := e^n \operatorname{ht}(\mathbf{H}).$$

Given a finite collection $\mathcal{F} = \{(\mathbf{H}, n)\} \subset \mathcal{H} \times \mathbb{N}$, define

$$ht(\mathcal{F}) = \max\{ht(\mathbf{H}, n) : (\mathbf{H}, n) \in \mathcal{F}\}.$$

Using the element $\mathbf{v_H} \in \bigwedge^{\dim \mathbf{H}} \mathfrak{g}$ we define the orbit map

$$\eta_H(g) := \rho_H(g)\mathbf{v}_H$$
 for every $g \in G$.

Given a nonzero vector $w \in \bigwedge^r \mathfrak{g}$, for some $0 < r \le \dim \mathbf{G}$, we define $\overline{w} := \frac{w}{\|w\|}$.

Let $z \in \mathfrak{g}$ with ||z|| = 1 be so that $u(t) = \exp(tz)$. Let $\mathbf{H} \in \mathcal{H}$; the definition of $N_G(U, H)$ then implies that

$$N_G(U, H) = \{ g \in G : z \land \eta_H(g) = 0 \}.$$

We will need the following definition of an effective notion of a generic point from [52].

5.7. **Definition.** Let $\varepsilon : \mathcal{H} \times \mathbb{N} \to (0,1)$ be a function so that $\varepsilon(\mathbf{H}, \cdot)$ is decreasing and $\varepsilon(\cdot, n)$ is decreasing in $ht(\mathbf{H})$.

A point $g\Gamma$ is said to be ε -Diophantine for the action of U if the following holds. For every nontrivial $\mathbf{H} \in \mathcal{H}$, with $\mathbf{H} \neq \mathbf{G}$, and every $n \in \mathbb{N}$ we have

(5.2) for every
$$\gamma \in \Gamma$$
 with $\|\eta_H(g\gamma)\| < e^n$ that $\|z \wedge \overline{\eta_H(g\gamma)}\| \ge \varepsilon(\mathbf{H}, n)$.

Given a finite collection $\mathcal{F} \subset \mathcal{H} \times \mathbb{N}$, we say that $g\Gamma$ is $(\varepsilon; \mathcal{F})$ -Diophantine if (5.2) holds for all $(\mathbf{H}, n) \in \mathcal{F}$.

This is a condition on the pair $(U, g\Gamma)$. We note that the definition of singular points $\mathcal{S}(U)$ using the varieties $N(U, H) = \{g \in G : g^{-1}Ug \subset H\}$ for various subgroups H is defined using polynomial equations. As such, its behavior may change dramatically under small perturbations. Definition 5.7 behaves in that respect much better.

Moreover, one checks easily that $\#\{\mathbf{H} \in \mathcal{H} : \operatorname{ht}(\mathbf{H}) \leq T\} \ll T^{O(1)}$; now for a given pair (\mathbf{H}, n) , the condition in (5.2) is given using continuous functions. This implies that any $x \in \mathcal{G}(U)$ is ε -Diophantine for some ε as above.

Normal subgroups of **G** are fixed points for the adjoint action of G, and hence for U. Thus, we need to control the *distance* from them separately. For any T > 0, define

$$\sigma(T) = \min \Big\{ \| z \wedge \overline{\mathbf{v}_{\mathbf{H}}} \| : \begin{array}{c} \mathbf{H} \in \mathcal{H}, \mathbf{H} \triangleleft \mathbf{G}, \\ \mathrm{ht}(\mathbf{H}) \leq T, \{1\} \neq \mathbf{H} \neq \mathbf{G} \end{array} \Big\}.$$

For every $(\mathbf{H}, n) \in \mathcal{H} \times \mathbb{N}$ and any C > 0 set

(5.3)
$$\ell_C(\mathbf{H}, n) := \min\{\operatorname{ht}(\mathbf{H}, n)^{-C}, \sigma(\operatorname{ht}(\mathbf{H}, n)^{C})\}.$$

Let $\| \|$ be a norm on $\mathrm{SL}_d(\mathbb{R})$ fixed once and for all. For every $g \in \mathrm{SL}_d(\mathbb{R})$, in particular, for any $g \in G$ let

$$|g| = \max\{||g||, ||g^{-1}||\}.$$

As we discussed after Theorem 5.3, an important property one anticipates from generic points is that *genericity* is inherited by many points along the orbit. The following theorem guarantees this for the notion defined in Definition 5.7.

5.8. **Theorem** ([52]). There are constants C and D, depending only on d, and a constant E depending on d, G, and Γ so that the following holds. Let $\mathcal{F} \subset \mathcal{H} \times \mathbb{N}$ be a finite subset. For any $g \in G$, $k \geq 1$, and $0 < \eta < 1/2$ at least one of the following holds.

(1)

$$\left| \left\{ t \in [-1,1] : \begin{array}{c} u(e^k t)g\Gamma \not\in X_{\eta} \ or \\ u(e^k t)g \ is \ not \ (\eta^D \ell_C; \mathcal{F}) \text{-}Diophantine} \end{array} \right\} \right| < E\eta^{1/D}.$$

(2) There exist a nontrivial proper subgroup $\mathbf{H}_0 \in \mathcal{H}$ and some $n_0 \in \mathbb{N}$ with

$$\operatorname{ht}(\mathbf{H}_0, n_0) \le E \max\{\operatorname{ht}(\mathcal{F}), |g|\eta^{-1}\}^D,$$

so that the following hold.

(a) For all $t \in [-1, 1]$ we have

$$\|\eta_{H_0}(u(e^kt)g)\| \le Ee^{n_0}.$$

(b) For every $t \in [-1, 1]$ we have

$$||z \wedge \overline{\eta_{H_0}(u(e^{k-1}t)g)}|| \leq Ee^{-k/D} \max\{\operatorname{ht}(\mathcal{F}), |g|\eta^{-1}\}^D.$$

As was alluded to before the effective notion of a generic point, Definition 5.7 above, is one of the main innovations in [52]. In addition to this, the proof of Theorem 5.8 also takes advantage of the role played by the subgroup $\mathbf{L} = \{g \in \mathbf{G} : g\mathbf{v_H} = \mathbf{v_H}\}$ to control the *speed* of unipotent orbits – the *distance* between U and subgroup $\mathbf{L}(\mathbb{R})$ controls the speed of $t \mapsto u_t\mathbf{v_H}$. Note that \mathbf{L} is a \mathbb{Q} -subgroup of \mathbf{G} whose height is controlled by $\mathrm{ht}(\mathbf{H})^{O(1)}$ – it is defined as the stabilizer of the vector $\mathbf{v_H}$. However, \mathbf{L} may not belong to the class \mathcal{H} . Actually it turns out, one may use the fact that U is a unipotent group to replace \mathbf{L} by a subgroup $\mathbf{M} \subset \mathbf{L}$ in \mathcal{H} which already controls the aforementioned speed.

The general strategy of the proof of Theorem 5.8, however, is again based on polynomial like behavior of unipotent orbits; and it relies on effectivizing the approach in [22].

References

- 1. M. Aka, M. Einsiedler, H. Li, and A. Mohammadi, On effective equidistribution for quotients of $SL(d, \mathbb{R})$, Preprint, arXiv:1607.08464.
- J. Athreya and G. Margulis, Values of random polynomials at integer points, Preprint, arXiv:1802.00792.
- L. Auslander, L. Green, and F. Hahn, Flows on homogeneous spaces, With the assistance of L. Markus and W. Massey, and an appendix by L. Greenberg. Annals of Mathematics Studies, No. 53, Princeton University Press, Princeton, N.J., 1963. MR 0167569
- 4. Yves Benoist and Hee Oh, Effective equidistribution of S-integral points on symmetric varieties, Ann. Inst. Fourier (Grenoble) 62 (2012), no. 5, 1889–1942. MR 3025156
- 5. Yves Benoist and Jean-François Quint, Mesures stationnaires et fermés invariants des espaces homogènes, Ann. of Math. (2) 174 (2011), no. 2, 1111–1162. MR 2831114
- 6. ______, Stationary measures and invariant subsets of homogeneous spaces (II), J. Amer. Math. Soc. **26** (2013), no. 3, 659–734. MR 3037785
- 7. _____, Stationary measures and invariant subsets of homogeneous spaces (III), Ann. of Math. (2) 178 (2013), no. 3, 1017–1059. MR 3092475
- Armand Borel and Gopal Prasad, Finiteness theorems for discrete subgroups of bounded covolume in semi-simple groups, Inst. Hautes Études Sci. Publ. Math. (1989), no. 69, 119–171. MR 1019963
- Jean Bourgain, A quantitative Oppenheim theorem for generic diagonal quadratic forms, Israel J. Math. 215 (2016), no. 1, 503–512. MR 3551907
- Jean Bourgain, Alex Furman, Elon Lindenstrauss, and Shahar Mozes, Stationary measures and equidistribution for orbits of nonabelian semigroups on the torus, J. Amer. Math. Soc. 24 (2011), no. 1, 231–280. MR 2726604
- M. Burger and P. Sarnak, Ramanujan duals. II, Invent. Math. 106 (1991), no. 1, 1–11.
 MR 1123369
- Marc Burger, Horocycle flow on geometrically finite surfaces, Duke Math. J. 61 (1990), no. 3, 779–803. MR 1084459
- 13. Laurent Clozel, Démonstration de la conjecture τ , Invent. Math. **151** (2003), no. 2, 297–328. MR 1953260
- S. G. Dani, Invariant measures of horospherical flows on noncompact homogeneous spaces, Invent. Math. 47 (1978), no. 2, 101–138. MR 0578655
- 15. _____, Invariant measures and minimal sets of horospherical flows, Invent. Math. **64** (1981), no. 2, 357–385. MR 629475
- 16. _____, On orbits of unipotent flows on homogeneous spaces, Ergodic Theory Dynam. Systems 4 (1984), no. 1, 25–34. MR 758891
- 17. _____, On orbits of unipotent flows on homogeneous spaces. II, Ergodic Theory Dynam. Systems 6 (1986), no. 2, 167–182. MR 857195
- 18. _____, Orbits of horospherical flows, Duke Math. J. ${\bf 53}$ (1986), no. 1, 177–188. MR 835804
- S. G. Dani and G. A. Margulis, Values of quadratic forms at primitive integral points, Invent. Math. 98 (1989), no. 2, 405–424. MR 1016271
- 20. _____, Orbit closures of generic unipotent flows on homogeneous spaces of SL(3, **R**), Math. Ann. **286** (1990), no. 1-3, 101–128. MR 1032925
- 21. ______, Asymptotic behaviour of trajectories of unipotent flows on homogeneous spaces, Proc. Indian Acad. Sci. Math. Sci. 101 (1991), no. 1, 1–17. MR 1101994
- 22. _____, Asymptotic behaviour of trajectories of unipotent flows on homogeneous spaces, Proc. Indian Acad. Sci. Math. Sci. 101 (1991), no. 1, 1–17. MR 1101994
- 23. W. Duke, Z. Rudnick, and P. Sarnak, Density of integer points on affine homogeneous varieties, Duke Math. J. **71** (1993), no. 1, 143–179. MR 1230289

- 24. M. Einsiedler, A. Margulis, G. Mohammadi, and A. Venkatesh, *Effective equidistribution and property tau*, Preprint, arXiv:1503.05884.
- M. Einsiedler, G. Margulis, and A. Venkatesh, Effective equidistribution for closed orbits of semisimple groups on homogeneous spaces, Invent. Math. 177 (2009), no. 1, 137–212. MR 2507639
- M. Einsiedler, R. Rühr, and Wirth P., Distribution of shapes of orthogonal lattices, Preprint, arXiv:1609.03070.
- 27. M. Einsiedler and P. Wirth, Effective equidistribution of closed hyperbolic surfaces on congruence quotients of hyperbolic spaces, in this volume.
- 28. Manfred Einsiedler, Ratner's theorem on SL(2, ℝ)-invariant measures, Jahresber. Deutsch. Math.-Verein. 108 (2006), no. 3, 143–164. MR 2265534
- 29. Manfred Einsiedler, Anatole Katok, and Elon Lindenstrauss, *Invariant measures and the set of exceptions to Littlewood's conjecture*, Ann. of Math. (2) **164** (2006), no. 2, 513–560.
- Manfred Einsiedler and Elon Lindenstrauss, On measures invariant under tori on quotients of semisimple groups, Ann. of Math. (2) 181 (2015), no. 3, 993–1031.
 MR 3296819
- 31. Jordan S. Ellenberg and Akshay Venkatesh, Local-global principles for representations of quadratic forms, Invent. Math. 171 (2008), no. 2, 257–279. MR 2367020
- 32. Alex Eskin, Gregory Margulis, and Shahar Mozes, On a quantitative version of the Oppenheim conjecture, Electron. Res. Announc. Amer. Math. Soc. 1 (1995), no. 3, 124–130. MR 1369644
- 33. ______, Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture, Ann. of Math. (2) 147 (1998), no. 1, 93–141. MR 1609447
- 34. Alex Eskin and Curt McMullen, Mixing, counting, and equidistribution in Lie groups, Duke Math. J. 71 (1993), no. 1, 181–209. MR 1230290
- 35. Alex Eskin and Maryam Mirzakhani, *Invariant and stationary measures for the* SL(2, ℝ) *action on moduli space*, Publ. Math. Inst. Hautes Études Sci. **127** (2018), 95–324. MR 3814652
- 36. Alex Eskin, Maryam Mirzakhani, and Amir Mohammadi, Isolation, equidistribution, and orbit closures for the SL(2, ℝ) action on moduli space, Ann. of Math. (2) 182 (2015), no. 2, 673–721. MR 3418528
- Alex Eskin, Shahar Mozes, and Nimish Shah, Unipotent flows and counting lattice points on homogeneous varieties, Ann. of Math. (2) 143 (1996), no. 2, 253–299. MR 1381987
- 38. Alex Eskin and Hee Oh, Representations of integers by an invariant polynomial and unipotent flows, Duke Math. J. 135 (2006), no. 3, 481–506. MR 2272974
- 39. Livio Flaminio and Giovanni Forni, Equidistribution of nilflows and applications to theta sums, Ergodic Theory Dynam. Systems 26 (2006), no. 2, 409–433. MR 2218767
- Harry Furstenberg, The unique ergodicity of the horocycle flow, (1973), 95–115. Lecture Notes in Math., Vol. 318. MR 0393339
- 41. Alex Gorodnik, François Maucourant, and Hee Oh, Manin's and Peyre's conjectures on rational points and adelic mixing, Ann. Sci. Éc. Norm. Supér. (4) **41** (2008), no. 3, 383–435. MR 2482443
- 42. Ben Green and Terence Tao, The quantitative behaviour of polynomial orbits on nilmanifolds, Ann. of Math. (2) 175 (2012), no. 2, 465–540. MR 2877065
- 43. F. Götze and G. Margulis, Distribution of values of quadratic forms at integral points, Preprint, arXiv:1004.5123.
- 44. Gustav A. Hedlund, Fuchsian groups and transitive horocycles, Duke Math. J. 2 (1936), no. 3, 530–542. MR 1545946
- 45. H. Jacquet and R. P. Langlands, *Automorphic forms on GL*(2), Lecture Notes in Mathematics, Vol. 114, Springer-Verlag, Berlin-New York, 1970. MR 0401654

- 46. D. A. Každan, On the connection of the dual space of a group with the structure of its closed subgroups, Funkcional. Anal. i Priložen. 1 (1967), 71–74. MR 0209390
- 47. D. Y. Kleinbock and G. A. Margulis, *Bounded orbits of nonquasiunipotent flows on homogeneous spaces*, SinaUi's Moscow Seminar on Dynamical Systems, Amer. Math. Soc. Transl. Ser. 2, vol. 171, Amer. Math. Soc., Providence, RI, 1996, pp. 141–172.
- Bounded orbits of nonquasiunipotent flows on homogeneous spaces, SinaUı's Moscow Seminar on Dynamical Systems, Amer. Math. Soc. Transl. Ser. 2, vol. 171, Amer. Math. Soc., Providence, RI, 1996, pp. 141–172. MR 1359098
- 49. _____, Flows on homogeneous spaces and Diophantine approximation on manifolds, Ann. of Math. (2) 148 (1998), no. 1, 339–360. MR 1652916
- 50. ______, On effective equidistribution of expanding translates of certain orbits in the space of lattices, Number theory, analysis and geometry, Springer, New York, 2012, pp. 385–396. MR 2867926
- 51. A. Leibman, Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold, Ergodic Theory Dynam. Systems **25** (2005), no. 1, 201–213. MR 2122919
- 52. E. Lindenstrauss, G. Margulis, A. Mohammadi, and N. Shah, Effective avoidance principle for unipotent orbits, In preparation.
- 53. Elon Lindenstrauss, Invariant measures and arithmetic quantum unique ergodicity, Ann. of Math. (2) 163 (2006), no. 1, 165–219. MR 2195133 (2007b:11072)
- 54. Elon Lindenstrauss and Gregory Margulis, Effective estimates on indefinite ternary forms, Israel J. Math. 203 (2014), no. 1, 445–499. MR 3273448
- G. A. Margulis, The action of unipotent groups in a lattice space, Mat. Sb. (N.S.)
 86(128) (1971), 552–556. MR 0291352
- 56. G. A Margulis, Indefinite quadratic forms and unipotent flows on homogeneous spaces, Dynamical systems and ergodic theory (Warsaw, 1986) 23 (1989), 399–409.
- 57. Gregori Aleksandrovitch Margulis and Georges Metodiev Tomanov, *Invariant measures for actions of unipotent groups over local fields on homogeneous spaces*, Invent. Math. **116** (1994), no. 1-3, 347–392.
- 58. Gregory Margulis, *Problems and conjectures in rigidity theory*, Mathematics: frontiers and perspectives, Amer. Math. Soc., Providence, RI, 2000, pp. 161–174.
- Gregory Margulis and Amir Mohammadi, Quantitative version of the Oppenheim conjecture for inhomogeneous quadratic forms, Duke Math. J. 158 (2011), no. 1, 121–160. MR 2794370
- 60. Grigorii Margulis, Dynamical and ergodic properties of subgroup actions on homogeneous spaces with applications to number theory, ICM-90, Mathematical Society of Japan, Tokyo; distributed outside Asia by the American Mathematical Society, Providence, RI, 1990, A plenary address presented at the International Congress of Mathematicians held in Kyoto, August 1990. MR 1127160
- 61. Grigoriy A. Margulis, On some aspects of the theory of Anosov systems, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2004, With a survey by Richard Sharp: Periodic orbits of hyperbolic flows, Translated from the Russian by Valentina Vladimirovna Szulikowska. MR 2035655
- 62. Jens Marklof, Pair correlation densities of inhomogeneous quadratic forms, Ann. of Math. (2) 158 (2003), no. 2, 419–471. MR 2018926
- 63. Taylor McAdam, Almost-primes in horospherical flows on the space of lattices, Preprint, arXiv:1802.08764.
- A. Mohammadi, A special case of effective equidistribution with explicit constants, Ergodic Theory Dynam. Systems 32 (2012), no. 1, 237–247. MR 2873169
- 65. Shahar Mozes and Nimish Shah, On the space of ergodic invariant measures of unipotent flows, Ergodic Theory Dynam. Systems 15 (1995), no. 1, 149–159. MR 1314973

- 66. Hee Oh, Uniform pointwise bounds for matrix coefficients of unitary representations and applications to Kazhdan constants, Duke Math. J. 113 (2002), no. 1, 133–192. MR 1905394
- William Parry, Dynamical systems on nilmanifolds, Bull. London Math. Soc. 2 (1970), 37–40. MR 0267558
- 68. Gopal Prasad, Volumes of S-arithmetic quotients of semi-simple groups, Inst. Hautes Études Sci. Publ. Math. (1989), no. 69, 91–117, With an appendix by Moshe Jarden and the author. MR 1019962
- Marina Ratner, On measure rigidity of unipotent subgroups of semisimple groups, Acta Math. 165 (1990), no. 3-4, 229-309.
- 70. _____, On Raghunathan's measure conjecture, Ann. of Math. (2) **134** (1991), no. 3, 545–607.
- 71. ______, Raghunathan's topological conjecture and distributions of unipotent flows, Duke Math. J. **63** (1991), no. 1, 235–280.
- 72. _____, Raghunathan's conjectures for cartesian products of real and p-adic lie groups, Duke Math. J. 77 (1995), no. 2, 275–382.
- 73. Peter Sarnak, Asymptotic behavior of periodic orbits of the horocycle flow and Eisenstein series, Comm. Pure Appl. Math. 34 (1981), no. 6, 719–739. MR 634284
- 74. ______, Values at integers of binary quadratic forms, Harmonic analysis and number theory (Montreal, PQ, 1996), CMS Conf. Proc., vol. 21, Amer. Math. Soc., Providence, RI, 1997, pp. 181–203. MR 1472786
- Peter Sarnak and Adrián Ubis, The horocycle flow at prime times, J. Math. Pures Appl. (9) 103 (2015), no. 2, 575–618. MR 3298371
- Atle Selberg, On the estimation of Fourier coefficients of modular forms, Proc. Sympos. Pure Math., Vol. VIII, Amer. Math. Soc., Providence, R.I., 1965, pp. 1–15. MR 0182610
- 77. Nimish A. Shah, Uniformly distributed orbits of certain flows on homogeneous spaces, Math. Ann. 289 (1991), no. 2, 315–334. MR 1092178
- Andreas Strömbergsson, On the uniform equidistribution of long closed horocycles, Duke Math. J. 123 (2004), no. 3, 507–547. MR 2068968
- 79. _____, An effective Ratner equidistribution result for $SL(2, \mathbb{R}) \ltimes \mathbb{R}^2$, Duke Math. J. **164** (2015), no. 5, 843–902. MR 3332893
- William A. Veech, Minimality of horospherical flows, Israel J. Math. 21 (1975), no. 2-3, 233–239, Conference on Ergodic Theory and Topological Dynamics (Kibbutz Lavi, 1974). MR 0385009
- 81. Akshay Venkatesh, Sparse equidistribution problems, period bounds and subconvexity, Ann. of Math. (2) 172 (2010), no. 2, 989–1094. MR 2680486
- 82. D. Zagier, Eisenstein series and the Riemann zeta function, Automorphic forms, representation theory and arithmetic (Bombay, 1979), Tata Inst. Fund. Res. Studies in Math., vol. 10, Tata Inst. Fundamental Res., Bombay, 1981, pp. 275–301. MR 633666

M.E.: Departement Mathematik, ETH Zürich, Rämistrasse 101, 8092, Zürich, Switzerland

E-mail address: manfred.einsiedler@math.ethz.ch

 $\rm A.M.:~Department$ of Mathematics, The University of California, San Diego, CA 92093, USA

E-mail address: ammohammadi@ucsd.edu