Advances in Mathematics 327 (2018) 410–424

Contents lists available at ScienceDirect

Advances in Mathematics

www.elsevier.com/locate/aim

New bounds in reduction theory of indefinite ternary integral quadratic forms

Han Li [a,*], Gregory A. Margulis [b]

[a] *Department of Mathematics and Computer Science, Wesleyan University, Middletown, CT 06459, USA*
[b] *Department of Mathematics, Yale University, New Haven, CT 06511, USA*

A R T I C L E   I N F O

Article history:
Received 30 October 2016
Accepted 24 March 2017
Available online 4 July 2017

Dedicated to D. Kazhdan on the occasion of his 70th birthday

MSC:
primary 11E12
secondary 37A25

Keywords:
Integral quadratic forms
Reduction theory
Homogeneous dynamics

A B S T R A C T

Using dynamics on homogeneous spaces we obtain some new and improved estimates for reduction of indefinite ternary integral quadratic forms.

1. Introduction and statement of results

1.1. Historical perspective

An important problem in reduction theory of integral quadratic forms is to decide to what extent one can simplify a given integral quadratic form by taking an equivalent

* Corresponding author.
E-mail addresses: hli03@wesleyan.edu (H. Li), margulis@math.yale.edu (G.A. Margulis).

http://dx.doi.org/10.1016/j.aim.2017.06.033

form. In this context an explicit question is *given an integral quadratic form, how small can the coefficients of an equivalent form be?* Let Q be a quadratic form and $A = (a_{ij})$ the symmetric matrix of Q. Sometimes we will use $Q = (a_{ij})$ with abuse of notation. Recall that the quadratic form Q is called integral if the entries A are integers. To measure the complexity of Q we will consider $\det A$ the *determinant* of Q and

$$\mathrm{ht}(Q) := \max(|a_{ij}|)$$

the *height* of Q. Two integral quadratic forms are said to be equivalent if their symmetric matrices A and B satisfy $A = \gamma^t B\gamma$ for some $\gamma \in \mathrm{GL}_n(\mathbb{Z})$.

An early achievement in reduction theory, due to Lagrange, is that any non-degenerate binary integral quadratic form Q is equivalent to a form $ax^2+bxy+cy^2$ with $|b| \leq |a| \leq |c|$. This implies

$$\mathrm{ht}(ax^2 + bxy + cy^2) = |c| \leq \frac{4}{3} \cdot |ac - \frac{b^2}{4}| = 4|\det Q|/3. \tag{1}$$

In other words, any binary quadratic form is equivalent to a form whose coefficients are all small. Since *equivalent integral quadratic forms have the same determinant* we have, by the finiteness of the choice of a, b and c, for any integer $D > 0$ there are only finitely many equivalence classes of binary quadratic forms with the determinant equal to D. Historically, an important application of reduction theory was to prove the finiteness for the number of equivalence classes of integral quadratic forms of a given determinant, in any number of variables.

1.2. Statement of results on quadratic forms

The goal of the present paper is to prove some new and improved estimates for reduction of indefinite ternary integral quadratic forms, using an approach based on the theory of homogeneous flows. For the dynamical setting we let, throughout the paper, G stand for $\mathrm{SL}_3(\mathbb{R})$ and Γ for $\mathrm{SL}_3(\mathbb{Z})$. We will fix the indefinite ternary integral quadratic form

$$Q_0(\boldsymbol{w}) = 2w_1w_3 - w_2^2 \qquad (\boldsymbol{w} \in \mathbb{R}^3),$$

and denote by $H = \mathrm{SO}(Q_0)^\circ_{\mathbb{R}}$ the identity component of the real points of the special orthogonal group of Q_0. We shall also fix a right invariant Riemannian metric on G which induces a G-invariant probability measure on G/Γ. The metric restricts to the closed subgroup H of G and gives rise a Haar measure m_H on H. For any indefinite ternary real quadratic form Q, there exists an element $g \in G$ such that

$$Q(\cdot) = (\det Q)^{\frac{1}{3}} \cdot Q_0(g\cdot). \tag{2}$$

Recall that if Q is integral, then the subgroup $H \cap g\Gamma g^{-1}$ is a lattice in H. The covolume of this lattice has been used by people to measure the complexity of an integral quadratic form, besides height and determinant.

Definition 1.1. Let Q be an indefinite ternary integral quadratic form. The volume of Q, denoted by $\mathrm{vol}(Q)$, is by definition the co-volume of $H \cap g\Gamma g^{-1}$ in H with respect to the Haar measure m_H.

The reader may readily verify that (1) the value $\mathrm{vol}(Q)$ does not depend on the choice of $g \in G$; (2) $\mathrm{vol}(Q) = \mathrm{vol}(kQ)$ for any integer k; and (3) equivalent integral quadratic forms have the same volume. We are now in a position of stating our main results.

Theorem 1.2. *There is a constant $C > 0$ such that for any indefinite ternary integral quadratic form Q_1 and $\xi \in (-1, 1)$, there exists a quadratic form $Q = (a_{ij})$ equivalent to Q_1 and satisfying*

$$\left| \frac{a_{11}}{(\det Q_1)^{\frac{1}{3}}} - \xi \right| < C \cdot \mathrm{vol}(Q_1)^{-\frac{1}{80}}, \tag{3}$$

$$\mathrm{ht}(Q) < C \cdot |\det Q_1|^{\frac{1}{3}}. \tag{4}$$

We note that the estimate in (4) is optimal, up to a multiplicative constant C, because plainly one has $\mathrm{ht}(Q) \gg |\det Q|^{\frac{1}{3}} = |\det Q_1|^{\frac{1}{3}}$. Notice also that such estimate is not valid for definite forms, as the height of any quadratic form equivalent to $x_1^2 + x_2^2 + Dx_3^2$, where $D > 0$, must be at least D.

It is also well known (see for instance [7, 2.6] and [2, Sect. 17.3]) that the determinant of any primitive form is bounded above by a polynomial of $\mathrm{vol}(Q)$. In view of Theorem 1.2, it might be of interest to give an explicit exponent for such estimate. Hence we have the following result.

Proposition 1.3. *There exists a constant $C > 0$, such that for every primitive indefinite ternary integral quadratic form Q,*

$$|\det Q| < C \cdot \mathrm{vol}(Q)^{39}. \tag{5}$$

It is necessary to assume the quadratic form to be primitive in the above proposition because $\mathrm{vol}(Q) = \mathrm{vol}(kQ)$ but $\det(Q) = k^3 \det(Q)$, for any integer k.

The following corollary is an easy consequence of Theorem 1.2 and Proposition 1.3.

Corollary 1.4. *There is a constant $C > 0$ such that for any primitive indefinite ternary integral quadratic form Q_1, there exists a quadratic form $Q = (a_{ij})$, which is equivalent to Q_1, satisfying*

$$0 < |a_{11}| < C \cdot |\det Q_1|^{\frac{1}{3} - \frac{1}{80 \cdot 39}}, \qquad \mathrm{ht}(Q) < C \cdot |\det Q_1|^{\frac{1}{3}}.$$

Another direct consequence of Theorem 1.2 is the existence of many equivalent forms of small height.

Corollary 1.5. *There exist constants C, $c > 0$ such that for every indefinite integral quadratic form Q_1,*

$$\#\left\{Q \textit{ equivalent to } Q_1 : \mathrm{ht}(Q) < C \cdot |\det Q_1|^{\frac{1}{3}}\right\} > c \cdot \mathrm{vol}(Q_1)^{\frac{1}{80}}.$$

The asymptotics of $\#\{Q \text{ equivalent to } Q_1 : \mathrm{ht}(Q) < T\}$ $(T \to \infty)$ can be analyzed by the lattice point counting theory. The readers are referred to a nice survey of Oh [10] for more information on this topic. However, Corollary 1.5 is mainly concerned with the number of forms whose height are relatively small.

Finally, a result of Siegel [12] asserts that one can effectively determine the equivalence of two given integral quadratic forms (see also a recent paper of the authors [6] for polynomially effective results). Hence for any given indefinite ternary integral quadratic form Q_1, one can in principle effectively find all quadratic forms Q satisfying Theorem 1.2, Corollary 1.4 and Corollary 1.5.

1.3. Outline of the paper

Section 2 introduces the preliminaries for later sections. In Section 3 we will state the main dynamical result Theorem 3.2 and use it to prove Theorem 1.2 combining a result in the geometry of numbers Lemma 3.1. The proof of Theorem 3.2 will be given in Section 6. It involves the study of the transversal behavior of closed H-orbits in G/Γ which will be carried out in Section 4 and the spectral theory of automorphic representations which will be discussed in Section 5. Section 7 is devoted to the proof of Proposition 1.3.

2. Preliminaries

2.1. Notations and basic facts

Let $G = \mathrm{SL}_3(\mathbb{R})$, $\Gamma = \mathrm{SL}_3(\mathbb{Z})$, $X = G/\Gamma$ and $H = \mathrm{SO}(Q_0)^\circ_{\mathbb{R}}$ be as in Section 1.2. For any element $g \in G$ we will denote by $[g]$ the element $g\Gamma \in X$ which can be naturally identified with the unimodular lattice $g\mathbb{Z}^3$ in $\mathbb{R}^3$.

In the sequel $\|\cdot\|$ will be a fixed matrix norm on G. For any $r > 0$ we denote by

$$B_G(r) := \{g \in G : \|g - 1_G\| < r\}, \qquad B_H(r) := H \cap B_G(r).$$

They are identity neighborhoods of G and H, respectively. We will make frequent use of the matrices

$$D(\lambda) = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda^{-1} \end{pmatrix} \in H, \qquad V(s) = \begin{pmatrix} 1 & 0 & s \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in G.$$

Let $\mathfrak{g} = \mathrm{Lie}(G)$ and $\mathfrak{h} = \mathrm{Lie}(H)$ be the Lie algebras. Under the adjoint action of H, the vector space $\mathfrak{g}$ is decomposed into irreducible H-modules $\mathfrak{g} = \mathfrak{h} \dot{+} \mathfrak{h}'$. Let E_{ij} be the matrix whose ijth entry is 1 and 0 otherwise. We shall fix a Euclidean norm $\|\cdot\|$ on $\mathfrak{g}$ such that it gives rise to a G-invariant probability measure on X, and that E_{13}, E_{31}, $E_{12} - E_{23}$, $E_{21} - E_{32}$, $E_{11} - 2E_{22} + E_{33}$ form an orthonormal basis of $\mathfrak{h}'$. Thus the vector space $\mathfrak{h}'$ has an orthogonal decomposition

$$\mathfrak{h}' = \mathfrak{h}'_0 + \mathfrak{h}'_1, \quad \text{where } \mathfrak{h}'_0 = \mathbb{R}E_{13}. \tag{6}$$

We are interested in this decomposition because for the adjoint of action of $D(\lambda)$ with $\lambda > 1$, $\mathfrak{h}'_0$ is the subspace of $\mathfrak{h}'$ with the maximal expansion. We will write $r = r_0 + r_1$ for any $r \in \mathfrak{h}'$ according to (6).

The group $K = \exp(\mathbb{R}(E_{12} + E_{23} - E_{21} - E_{32}))$ is a maximal compact subgroup of H. It will be useful to keep in mind the following fact.

$$\begin{gathered}\text{There does not exist a nonzero linear subspace of } \mathfrak{h}'_1 \\ \text{that is invariant under the adjoint action of } K.\end{gathered} \tag{7}$$

2.2. Invariant measures

Let $H.x$ be a closed orbit in X with $x = [g]$. Then $H \cap g\Gamma g^{-1}$ is a lattice in H. The orbit $H.x$ supports a unique H-invariant probability measure which will be denoted as μ_x. We will also consider the measure vol_H on $H.x$ induced by the Haar measure m_H (see Section 1.2). The two H-invariant measures are related in such a way that for any Borel subset $Y \subset H.x$

$$\mathrm{vol}_H(H.x) \cdot \mu_x(Y) = \mathrm{vol}_H(Y).$$

2.3. Compact subsets of X

Let α_1 be the function on X defined as $\alpha_1([g]) = \max\{\|w\|^{-1} : w \in g\mathbb{Z}^3 \backslash \{\mathbf{0}\}\}$. By Mahler's criterion $\mathfrak{S}(R) = \{x \in X : \alpha_1(x) \leq R\}$ is a compact subset in X for every $R > 0$. By [2, Lemma 3.2] we can fix a large number $R_0 > 0$ such that for every closed orbit $H.x$ in X

$$\mu_x(H.x \cap \mathfrak{S}(R_0)) > 1 - 10^{-11}. \tag{8}$$

For later use (in Proposition 4.4), we will consider an enlarged K-invariant compact subset of $\mathfrak{S}(R_0)$

$$X_{\mathrm{cpt}} = \{k.x : k \in K, x \in \mathfrak{S}(R_0)\} \subseteq X.$$

For technical reason we shall fix some other constants related to X_{cpt} whose existence can be easily seen by compactness argument: a positive constant $R_1 > 0$ such that

$$X_{\text{cpt}} \subseteq B_G(R_1)[1_G]; \tag{9}$$

and positive constants ϵ_0, $R_0' > 0$ such that $B_G(\epsilon_0)X_{\text{cpt}} \subset \mathfrak{S}(R_0')$ and the map $g \mapsto g.x$ is injective on $B_G(\epsilon_0)$ for every $x \in \mathfrak{S}(R_0')$.

2.4. Effective constants and their dependencies

In the sequel, the notation $c_1, c_2, \cdots$ will stand for positive constants which are in principle effectively computable and may depend on G, H, X, and the choice of R_0, ϵ_0, R_1, R_0' in Section 2.3; we also allow c_j depends on the choice of $c_1, \cdots, c_{j-1}$. We will use $A \ll B$ to represent the inequalities $A < \eta B$ for some sufficiently large and effectively computable multiplicative constant $\eta > 0$, which might depend on the choice of R_0, ϵ_0, R_1, R_0', c_1, c_2, $\cdots$.

3. The main dynamical result and proof of Theorem 1.2

3.1. A result in geometry of numbers

Lemma 3.1. *Let R_1 be as defined in Section 2.3 and $g \in B_G(R_1)$. Then there exists a primitive vector $\boldsymbol{w} \in g\mathbb{Z}^3$ (that is, $\boldsymbol{w} = g\boldsymbol{v}$ for some $\boldsymbol{v} \in \mathbb{Z}^3$ with the g.c.d. of the coordinates of $\boldsymbol{v}$ equal to 1) satisfying*

$$\|\boldsymbol{w}\| \ll 1, \quad Q_0(V(1)\boldsymbol{w}) < -1, \quad Q_0(V(4)\boldsymbol{w}) > 1.$$

Proof. Let $\mathcal{V} = \left\{\boldsymbol{v} \in \mathbb{R}^3 : Q_0(V(1)\boldsymbol{v}) < -1,\ Q_0(V(4)\boldsymbol{v}) > 1\right\}$. First we notice that there exists $t_1 > 0$ such that for any $t > t_1$,

$$B_G(R_1)B_{\mathbb{R}^3}((-t,t,t), t/32) \subseteq B_{\mathbb{R}^3}((-t,t,t), t/16) \subseteq \mathcal{V},$$

where $B_{\mathbb{R}^3}(\boldsymbol{v}, r) = \{x \in \mathbb{R}^3 : \|x - \boldsymbol{v}\| \leq r\}$. Let $\mathcal{D} = B_{\mathbb{R}^3}((-1,1,1), 1/32)$. It is well known that

$$\#\{\boldsymbol{v} \in \mathbb{Z}^3 \cap t\mathcal{D} :\ \boldsymbol{v} \text{ primitive}\} \sim \frac{\text{vol}(t\mathcal{D})}{\zeta(3)}.$$

Therefore, there exists $t_2 > 0$ such that for any $t > t_2$ the ball $t\mathcal{D} = B_{\mathbb{R}^3}((-t,t,t), t/32) \subseteq \mathbb{R}^3$ contains a primitive lattice point in $\mathbb{Z}^3$. Let us fix a t_0 which is larger than both t_1 and t_2. We see from the above that there exists a primitive lattice point $\boldsymbol{v_0} \in \mathbb{Z}^3$ satisfying

$$B_G(R_1)\boldsymbol{v_0} \subseteq B_G(R_1)B_{\mathbb{R}^3}((-t_0,t_0,t_0), t_0/32) \subseteq B_{\mathbb{R}^3}((-t_0,t_0,t_0), t_0/16) \subseteq \mathcal{V}.$$

Hence the primitive vector $\boldsymbol{w} = g\boldsymbol{v_0}$ satisfies the lemma. $\square$

3.2. The main dynamical theorem

We are now in a position of stating our main dynamical result.

Theorem 3.2. *Let $H.x$ be a closed orbit in X with $\mathrm{vol}_H(H.x) = M$. Then, there exists $y \in X_{\mathrm{cpt}}$ such that for any $t \in (1, 4)$ there exists $u \in G$ satisfying*

$$uV(t).y \in H.x, \qquad \|u - 1_G\| \ll M^{-\frac{1}{80}}.$$

Roughly speaking, the theorem says that any closed H-orbit has some concentration in the compact part of X which can be seen in the direction of $\{V(t) : 1 < t < 4\}$ that is transversal to the direction of H. The proof of Theorem 3.2 will be postponed to Section 6. We now deduce Theorem 1.2 from Theorem 3.2.

3.3. Proof of Theorem 1.2

Proof. As Q_1 is integral, there exists $g_1 \in G$ such that $Q_1 = (\det Q_1)^{\frac{1}{3}} Q_0(g_1 \cdot)$ and the orbit $H.[g_1]$ is closed. By Theorem 3.2, there exists $[g] \in X_{\mathrm{cpt}}$ such that for any $t \in (1, 4)$ there exists $u \in G$ satisfying

$$uV(t).[g] \in H.[g_1], \qquad \|u - 1_G\| \ll M^{-\frac{1}{80}} \tag{10}$$

where $M = \mathrm{vol}_H(H.[g_1])$. Moreover, by (9), we can choose $g \in G$ such that $g \in B_G(R_1)$.

By Lemma 3.1, there exists a primitive vector $\boldsymbol{w} \in g\mathbb{Z}^3$ satisfying

$$\|\boldsymbol{w}\| \ll 1, \qquad Q_0(V(1)\boldsymbol{w}) < -1, \qquad Q_0(V(4)\boldsymbol{w}) > 1.$$

Let $\xi \in (-1, 1)$ be given. Since $Q_0(V(s)\boldsymbol{w}) = Q_0(\boldsymbol{w}) + 2sw_3^2$, there exists $t_\xi \in (1, 4)$ such that $Q_0(V(t_\xi)\boldsymbol{w}) = \xi$. For $t = t_\xi$, let u_ξ be the element u satisfying (10). Consider the quadratic form

$$Q_2(\cdot) = (\det Q_1)^{\frac{1}{3}} Q_0(u_\xi V(t_\xi) g \cdot).$$

First, as $u_\xi V(t_\xi).[g] \in H.[g_1]$, the quadratic form Q_2 is integral and equivalent to Q_1. Second, letting $\boldsymbol{v} = g^{-1}\boldsymbol{w} \in \mathbb{Z}^3$ and since $\|\boldsymbol{w}\| \ll 1$ and $\|V(t_\xi)\| \ll 1$, we have

$$\left| \frac{Q_2(\boldsymbol{v})}{(\det Q_1)^{\frac{1}{3}}} - \xi \right| = \left| Q_0(u_\xi V(t_\xi).\boldsymbol{w}) - Q_0(V(t_\xi).\boldsymbol{w}) \right| \ll \|u_\xi - 1_G\| \ll M^{-\frac{1}{80}}.$$

Finally, as $\|u_\xi V(t_\xi) g\| \ll 1$, the height of the quadratic form Q_2 satisfies

$$\mathrm{ht}(Q_2) \ll |\det Q_1|^{\frac{1}{3}}.$$

We proceed to find a form Q satisfying our theorem. Notice that the vector $\boldsymbol{v} \in \mathbb{Z}^3$ is primitive and

$$\|\boldsymbol{v}\| \ll \|g^{-1}\| \cdot \|\boldsymbol{w}\| \ll 1.$$

Let $\boldsymbol{e} = (1, 0, 0)^t$. From the above we conclude that there exists $\gamma \in \Gamma$ such that $\gamma.\boldsymbol{e} = \boldsymbol{v}$ and $\|\gamma\| \ll 1$. We shall prove that $Q = Q_2 \circ \gamma$ satisfies Theorem 1.2. Obviously Q is equivalent to Q_1. Since $\|\gamma\| \ll 1$, we have $\mathrm{ht}(Q_2 \circ \gamma) \ll (\det Q_1)^{\frac{1}{3}}$ and thus Q satisfies the height bound (4) for the theorem. Finally to check (3), with $Q = (a_{ij})$ we have

$$\left|\frac{a_{11}}{(\det Q_1)^{\frac{1}{3}}} - \xi\right| = \left|\frac{Q(\boldsymbol{e})}{(\det Q_1)^{\frac{1}{3}}} - \xi\right| = \left|\frac{Q_2 \circ \gamma(\boldsymbol{e})}{(\det Q_1)^{\frac{1}{3}}} - \xi\right| = \left|\frac{Q_2(\boldsymbol{v})}{(\det Q_1)^{\frac{1}{3}}} - \xi\right| \ll M^{-\frac{1}{80}}. \qquad \square$$

4. Recurrence properties of closed *H*-orbits

4.1. The drift of transversally nearby points on closed H-orbits

We now state some basic facts concerning the drift of nearby points on closed H-orbits under the translates of $D(\lambda)$ (see [8,2,9]). Suppose $H.x_1, H.x_2$ are (not necessarily distinct) closed orbits in X, and assume that $\exp(r).x_1 = x_2$ for some $r \in \mathfrak{h}'$. Then

$$D(\lambda).x_2 = D(\lambda)\exp(r).x_1 = \exp(\mathrm{Ad}(D(\lambda))r)D(\lambda).x_1. \tag{11}$$

That is, the drift of x_1 and x_2 is given by $\exp(\mathrm{Ad}(D(\lambda))r)$. Writing $r = r_0 + r_1$ as (6), we have

$$\mathrm{Ad}(D(\lambda))r = \lambda^2 r_0 + \mathrm{Ad}(D(\lambda))r_1, \quad \|\mathrm{Ad}(D(\lambda))r_1\| \ll \lambda\|r_1\|. \tag{12}$$

Lemma 4.1. *The following two statements are valid.*

1. *Let m_K be the Haar measure on K as in Sect. 1.2. There exists $c_1 > 0$ such that for any $r \in \mathfrak{h}' \backslash \{\mathbf{0}\}$*

$$m_K\left(\left\{k \in K : \frac{\|(\mathrm{Ad}(k)r)_0\|}{\|\mathrm{Ad}(k)r\|} > c_1\right\}\right) > 0.99 m_K(K). \tag{13}$$

2. *Let $r \in \mathfrak{h}' \backslash \{\mathbf{0}\}$ be such that $\|r_0\| > c_1\|r\|$. Then for any $\lambda > 0$ with $\lambda^2\|r_0\| < 4$ we have*

$$\|\exp(\mathrm{Ad}(D(\lambda)r)) - \exp(\lambda^2 r_0)\| \ll \lambda^{-1} \tag{14}$$

Remark. See also [9, Proposition 3.2, 3.3].

Proof. For (13) we only need to work with $\|r\| = 1$. By (7), for every unit vector $r \in \mathfrak{h}'\backslash\{\mathbf{0}\}$ we can find $c = c(r)$ for which (13) holds. Since the unit sphere in Euclidean space is compact, we can choose a constant $c_1 > 0$ which satisfies (1). The assertion (2) follows directly from (12). □

4.2. Transversal injectivity radius for points on closed H-orbits

Definition 4.2. Let ϵ_0 be as defined in Section 2.3, and $B_{\mathfrak{h}'}(t) := \{r \in \mathfrak{h}' : \|r\| < t\}$. We fix $\epsilon_1 > 0$ so that $\exp(B_{\mathfrak{h}'}(\epsilon_1)) \subset B_G(\epsilon_0/2)$ and $\exp(B_{\mathfrak{h}'}(\epsilon_1)) \cap H = \{e\}$. Let $H.[g]$ be a closed orbit in X. By definition the transversal injectivity radius of $x \in H.[g]$ is

$$\sigma(x) = \sup\{0 < \epsilon < \epsilon_1 : \exp(B_{\mathfrak{h}'}(\epsilon))x \cap H.[g] = \{x\}\}.$$

Roughly speaking, transversal injectivity radius measures how close an H-orbit can approach itself at a given point from transversal direction.

The following lemma concerns polynomially quantitative isolation of closed H-orbits.

Lemma 4.3. *[2, Lemma 10.1] There are constants $c_2, c_3 > 0$, so that for any closed orbit $H.[g] \subseteq X$ and $x \in H.[g] \cap X_{\text{cpt}}$, we have $\sigma(x) > c_2 M^{-c_3}$.*

Our next result provides for a large fraction of points on a closed H-orbit an upper bound on their transversal injectitivity radius.

Proposition 4.4. *There exists $c_4 > 0$ such that, for any closed orbit $H.x$ in X, the subset*

$$\begin{aligned}\mathfrak{B} = \{y \in H.x \cap X_{\text{cpt}} : \exists\, r \in \mathfrak{h}' \text{ such that } \exp(r)y \in H.x,\\ \|r\| < \min(c_1^{-1}\|r_0\|, c_4 M^{-0.2})\}\end{aligned}$$

satisfies $\mu_x(\mathfrak{B}) > 0.9$. Here the decomposition $r = r_0 + r_1$ is as (6).

Remark. The fact that $\mathfrak{B}$ is nonempty for large enough constant c_4 is proved in [2, Proposition 14.2] (see also [9, Proposition 3.3]). Here we show in our setting that $\mathfrak{B}$ consumes a large portion.

Proof. (**Step I**) We begin by fixing some notation. Let ϵ_1 be as in Definition 4.2, $\mathfrak{A} = H.x \cap X_{\text{cpt}}$ and $\mathfrak{A}_\delta = \{y \in \mathfrak{A} : \sigma(y) < \delta\}$ where $0 < \delta < \epsilon_1$. We first show that

$$\text{vol}_H(\mathfrak{A} - \mathfrak{A}_\delta) \ll \delta^{-5}.$$

Notice that there is a constant $\kappa > 0$ such that $m(\exp(B_{\mathfrak{h}'}(\delta))\Omega) \geq \kappa\delta^5 \,\text{vol}(\Omega)$ for any $0 < \delta < \epsilon_1$ and Borel set $\Omega \subset B_H(\epsilon_0/2)$. The compact set $\mathfrak{A} - \mathfrak{A}_\delta$ can be covered by finitely many sets of the form $B_H(\delta).x_i$ ($x_i \in \mathfrak{A} - \mathfrak{A}_\delta$) with multiplicity 4 since $\dim H = 3$.

Because the map $\varphi : B_{\mathfrak{h}'}(\delta/2) \times (\mathfrak{A} - \mathfrak{A}_\delta) \to X, \varphi(r,x) = \exp(r).x$ is injective (recall this assumption for X_{cpt} from Sect. 2.3), it follows that

$$
\begin{aligned}
4 &> \sum_i m_G\Big(\{\exp(r).y : r \in B_{\mathfrak{h}'}(\delta/2), y \in (\mathfrak{A} - \mathfrak{A}_\delta) \cap (B_H(\delta)x_i)\} \Big) \\
&> \kappa\delta^5 \operatorname{vol}_H(\mathfrak{A} - \mathfrak{A}_\delta). \qquad (15)
\end{aligned}
$$

(**Step II**) Let $M = \operatorname{vol}_H(H.x)$. Recall from Sect. 2.3 that $\operatorname{vol}_H(\mathfrak{A}) > M(1 - 10^{-11})$. Let $c > 0$ be such that $\kappa c^5 > 400$. By (15), $\operatorname{vol}_H(\mathfrak{A} - \mathfrak{A}_{cM^{-\frac{1}{5}}}) < 0.01M$. So if the volume $M = \operatorname{vol}_H(H.x)$ of the orbit $H.x$ is so large that $cM^{-\frac{1}{5}} < \epsilon_1$, then

$$
\operatorname{vol}_H(\mathfrak{A}_{cM^{-\frac{1}{5}}}) > 0.98M. \qquad (16)
$$

(**Step III**) We are ready to prove our proposition. Let

$$
l_0 = \min\{l : \|\operatorname{Ad}(k)r\| \le l\|r\|, \forall\, k \in K, r \in \mathfrak{h}'\},
$$

and set the constant $c_4 = c\max(l_0, 1)$. We shall now prove that for any closed orbit $H.x$ such that $c_4M^{-0.2} < \epsilon_1$, the constant c_4 satisfies the proposition. As there are only finitely many closed H-orbits for which $c_4M^{-0.2} \ge \epsilon_1$, by enlarge c_4 if necessary, the proposition will be proved for any closed orbit.

Recall that X_{cpt} is K-invariant, and notice that $k\exp(r)k^{-1}k.y \in \mathfrak{A}$ whenever $\exp(r).y \in \mathfrak{A}$. By (1) of Lemma 4.1, for any $y \in \mathfrak{A}_{cM^{-\frac{1}{5}}}$ we have

$$
m_K\big(\{k \in K : kx \in \mathfrak{B}\}\big) > 0.99m_K(K).
$$

By Fubini and (16)

$$
\frac{\operatorname{vol}_H(\mathfrak{B})}{M} = \int\limits_{H.x} \Big(\frac{1}{m_K(K)} \int\limits_K \chi_{\mathfrak{B}}(kz)dk\Big) d\mu_x(z) > 0.99 \times 0.98 > 0.9. \qquad \square
$$

4.3. An effective result on recurrence of closed H-orbits

Proposition 4.5. *There are constants c_5, n_0 with the following properties. Let $H.x$ be a closed orbit in X with $\operatorname{vol}_H(H.x) = M$ and suppose $n > n_0$. Then there exist $A \subseteq H.x$ and $\lambda_i \gg M^{\frac{1}{10}}$, such that $\mu_x(A) \gg (n\log M)^{-1}$, and that for every $y \in B_H(\frac{c_5}{n})D(\lambda_i)B_H(\frac{c_5}{n})A$, at least one of the following holds*

$$
d(V(4^{\frac{i}{n}}).y, H.x) \ll (n^{-1} + M^{-\frac{1}{10}}), \text{ for any } i = 1, \cdots, n; \qquad (17)
$$

$$
d(V(-4^{\frac{i}{n}}).y, H.x) \ll (n^{-1} + M^{-\frac{1}{10}}), \text{ for any } i = 1, \cdots, n. \qquad (18)
$$

Proof. Lemma 4.3 and Proposition 4.4 assert that there exists $\mathfrak{B} \subseteq H.x$ such that $\operatorname{vol}_H(\mathfrak{B}) > 0.9M$, and for any $y \in \mathfrak{B}$ there exists $r = r_0 + r_1 \in \mathfrak{h}'$ (as in (6)) satisfying

(a1) $\exp(r).y \in H.x$, $\|r_0\| > c_1\|r\|$,
(b1) $c_1c_2M^{-c_3} < \|r_0\| < c_4M^{-\frac{1}{5}}$.

For every $r_0 \in \mathfrak{h}_0$ we write $r_0 = \mathfrak{p}(r_0)E_{13}$. Hence, without loss of generality, we assume that there exists $\mathfrak{B}' \subseteq \mathfrak{B}$ so that $\mathrm{vol}_H(\mathfrak{B}') > 0.45M$, and for every $y \in \mathfrak{B}'$ there exists $r \in \mathfrak{h}'$ satisfying (a) and

(b2) $c_1c_2M^{-c_3} < \mathfrak{p}(r_0) < c_4M^{-\frac{1}{5}}$.

By the Pigeonhole principle (for the values of $\mathfrak{p}(r_0)$), there exist $A \subset H.x$ and $0 < a < c_4M^{-\frac{1}{5}}$ such that $\mu_x(A) \gg (n \log M)^{-1}$, and for every $y \in A$ there is an $r \in \mathfrak{h}'$ satisfying (a1) and

(b3) $a4^{-\frac{1}{2n}} < \mathfrak{p}(r_0) < a$.

Notice that if $\exp(r)x \in H.x$ and $h \in H$, then $\exp(\mathrm{Ad}(h)r)hx \in H.x$. Hence there exists $c_5 > 0$ such that for every $y \in B_H(\frac{c_5}{n})A$, there exists $r \in \mathfrak{h}'$ with

(a2) $\exp(r).y \in H.x$, $\|r_0\| > \frac{1}{2}c_1\|r\|$,
(b4) $a(1 - \frac{1}{5n})4^{-\frac{1}{2n}} < \mathfrak{p}(r_0) < (1 + \frac{1}{5n})a$.

Let n_0 be the smallest natural number with $\frac{5n+1}{5n-1} < 4^{\frac{1}{2n}}$. We have, for any $n > n_0$, (b4) can be replaced by: $a'4^{-\frac{1}{n}} < \mathfrak{p}(r_0) < a'$, where $a' = (1 + \frac{1}{5n})a$.

Let λ_i be such that $\lambda_i^2 a' = 4^{\frac{i}{n}}$ $(1 \le i \le n)$. Clearly $\lambda_i > \frac{1}{2}c_4M^{\frac{1}{10}}$. If $y \in B_H(\frac{c_5}{n})D(\lambda_i)B_H(\frac{c_5}{n})A$, then

$$V(4^{\frac{i}{n}}).y \in V(4^{\frac{i}{n}})B_H(\frac{c_5}{n})V(-4^{\frac{i}{n}})V(4^{\frac{i}{n}})D(\lambda_i)B_H(\frac{c_5}{n})A,$$

$$\subseteq B_G(\frac{16c_5}{n})V(4^{\frac{i}{n}})D(\lambda_i)B_H(\frac{c_5}{n})A$$

For every $z \in B_H(\frac{c_5}{n})A$, let $r = r_0 + r_1 \in \mathfrak{h}'$ which satisfies condition (a2) and (b4). We have $D(\lambda_i)\exp(r).z \in H.x$. Notice that $4^{\frac{i}{n-1}} < \mathfrak{p}(\lambda_i^2 r_0) < 4^{\frac{i}{n}}$. It follows from (2) of Lemma 4.1 that

$$d(V(4^{\frac{i}{n}})D(\lambda_i).z, H.x) \ll n^{-1} + M^{-\frac{1}{10}}.$$

Hence (17) holds. Regrading the cases for which (b2) is $c_1c_2M^{-c_3} < -\mathfrak{p}(r_0) < c_4M^{-\frac{1}{5}}$, the same proof will lead to conclusion (18) instead of (17). □

5. Uniform spectral gap for closed H-orbits

Let H, K, X be as before. We say a unitary representation $(\pi, \mathcal{H})$ of H is strongly L^p if for a dense subset $\mathcal{V} \subset \mathcal{H}$, the matrix coefficient $< \pi(\cdot)u, v > \in L^p(H, m_H)$ for every $u, v \in \mathcal{V}$. We say that $(\pi, \mathcal{H})$ is strongly $L^{p+\epsilon}$ if it is strongly L^q for every $q > p$.

Let $H.x$ be a closed H-orbit in X. The H action on $H.x$ is ergodic with respect to μ_x. Let $L^2_0(H.x)$ be the orthogonal complement of the constant functions in $L^2(H.x)$. By a result of Kim–Sarnak ([5]) combined with Jacquet–Langlands correspondence ([4]), the unitary representation of H on $L^2_0(H.x)$ is strongly $L^{\frac{64}{25}+\epsilon}$ for *every* closed orbit $H.x$. Let Ω be an element in the Lie algebra of K of norm one, and consider for any $f \in C^\infty(H.x)$ the Sobolev norm along K:

$$S(f) := \|\Omega(f)\|_2 + \|f\|_2, \quad \text{where } \Omega(f)(z) := \lim_{t \to 0} \frac{f(\exp(t\Omega)z) - f(z)}{t} \quad \forall\, z \in H.x.$$

The next lemma, whose proof can be found in [11, Theorem 2.1] or [2, 6.2.1], shows that the correlations of H-actions on its closed orbits have exponential decay whose rates are related to Sobolev norm.

Lemma 5.1. *For every closed orbit $H.x$ in X, smooth functions $\psi, \phi \in C^\infty(H.x)$ and $\lambda > 1$, we have*

$$| < D(\lambda)\psi, \phi > -\mu_x(\psi)\mu_x(\phi)| \ll \lambda^{-\frac{5}{13}} S(\psi) S(\phi).$$

Smooth approximations of indicator functions are essential in our approach. To establish this we shall now fix for every $0 < \epsilon < 1$ a function $\theta_\epsilon \in C^\infty(H)$ such that

1. $\mathrm{supp}(\theta_\epsilon) \subset B_H(\epsilon)$;
2. $\theta_\epsilon \geq 0$, and $\int \theta_\epsilon = 1$;
3. $\|\Omega(\theta_\epsilon)\|_{L^1} \ll \epsilon^{-1}$.

Lemma 5.2. *Let $H.x$ be a closed orbit in X and $f \in L^2(H.x)$. Define*

$$\theta_\epsilon \star f(z) := \int_H f(h^{-1}z)\theta_\epsilon(h)\, dm_H(h), \quad \forall\, z \in H.x.$$

Then $\theta_\epsilon \star f \in C^\infty(H.x) \cap L^2(H.x)$, and

$$S(\theta_\epsilon \star f) \ll (\|\Omega(\theta_\epsilon)\|_{L^1} + \|\theta_\epsilon\|_{L^1}) \cdot \|f\|_2 \ll \epsilon^{-1} \cdot \|f\|_2.$$

Proof. For any element $Y \in \mathfrak{h}$ we have $Y(\theta_\epsilon \star f) = Y(\theta_\epsilon) \star f$. Hence $\theta_\epsilon \star f$ is smooth, and is square integrability and the norm estimate of $\theta_\epsilon \star f$ follow from the Minkowski's inequality. □

Proposition 5.3. *Let $H.x$ be a closed orbit in X, and $A \subseteq H.x$ be a subset with $\mu_x(A) > 0$. Then for every $\epsilon > 0$ and $\lambda > 1$ the subset $A' = H.x \backslash \big(B_H(\epsilon)D(\lambda)B_H(\epsilon)A\big) \subseteq H.x$ satisfies*

$$\mu_x(A') \ll \lambda^{-\frac{10}{13}}\mu_x(A)^{-1}\epsilon^{-4}.$$

Remark: The result says that $D(\lambda)$ translates of $B_H(\epsilon)A$ become equidistributed in $H.x$ in the sense that the measure of $B_H(\epsilon)D(\lambda)B_H(\epsilon)A$, the ϵ-thickening of the translated image $D(\lambda)B_H(\epsilon)A$, is close to 1.

Proof. Let $\psi = \theta_\epsilon \star \chi_A$ and $\phi = \theta_\epsilon \star \chi_{A'}$ with χ_A and $\chi_{A'}$ are the indicator functions. Notice that $< D(\lambda)\psi, \phi >= 0$ because $D(\lambda)\text{supp}(\psi) \cap \text{supp}(\phi) = \emptyset$. By Lemma 5.1 and 5.2 we have

$$\mu_x(A)\mu_x(A') \ll \lambda^{-\frac{5}{13}}\mu_x(A)^{1/2}\epsilon^{-1}\mu_x(A')^{1/2}\epsilon^{-1},$$

which directly implies our proposition. □

6. Proof of Theorem 3.2

Proof. It suffices to show that there exists an absolute constant $M_0 > 0$, such that Theorem 3.2 holds for any closed orbit $H.x$ with $M = \text{vol}_H(H.x) > M_0$.

We first deal with the case of (17). Applying Proposition 4.5 with $n = [M^{\frac{1}{80}}]$, there exists $A \subseteq H.x$ and $\lambda_i \gg M^{\frac{1}{10}}$, such that $\mu_x(A) \gg (n \log M)^{-1}$ and for every $z \in B_H(\frac{c_5}{n})D(\lambda_i)B_H(\frac{c_5}{n})A$

$$d(V(4^{\frac{i}{n}}).z, H.x) \ll M^{-\frac{1}{80}} + M^{-\frac{1}{10}} \ll M^{-\frac{1}{80}} \quad (i = 1, \cdots, n).$$

By Proposition 5.3, for every $i = 1, \cdots, n$

$$\mu_x\Big(H.x\backslash\Big(B_H(\frac{c_5}{n})D(\lambda_i)B_H(\frac{c_5}{n})A\Big)\Big) \ll M^{-\frac{1}{78}}.$$

Hence

$$\mu_x\Big(H.x\backslash\Big(\bigcap_{i=1}^{n} B_H(\frac{c_5}{n})D(\lambda_i)B_H(\frac{c_5}{n})A\Big)\Big) \ll M^{-\frac{1}{78}+\frac{1}{80}}.$$

Comparing this with (8), we get that there exists $M_0 > 0$ such that if $\text{vol}_H(H.x) = M > M_0$, then

$$\Big(\bigcap_{i=1}^{n} B_H(\frac{c_5}{n})D(\lambda_i)B_H(\frac{c_5}{n})A\Big) \bigcap X_{\text{cpt}} \neq \emptyset. \tag{19}$$

Because $|4^{\frac{i+1}{n}} - 4^{\frac{i}{n}}| \ll 1/n \ll M^{-\frac{1}{80}}$, any element y in the intersection of (19) satisfies Theorem 3.2.

To finish the proof of the theorem, it remains to deal with the case of (18). The same argument as above produces a point $y_1 \in X_{\mathrm{cpt}}$ such that for any $t \in (-4, -1)$ there exists $u = u(t) \in G$ satisfying

$$uV(t).y_1 \in H.x, \qquad \|u - 1_G\| \ll M^{-\frac{1}{80}}.$$

Therefore $y = V(-5).y_1$ satisfies the theorem. □

7. Proof of Proposition 1.3

Proof. Recall from (8) that for every closed H-orbit in X, a large proportion lies in X_{cpt}. Let Q be a primitive indefinite ternary integral quadratic form. Then there exists $g \in G$ such that $Q(\cdot) = (\det Q)^{\frac{1}{3}} \cdot Q_0(g\cdot)$ and $\|g\| \ll 1$. For later use let us record that $g^{-1}Hg$ is the identity component of $\mathrm{SO}_Q(\mathbb{R})$; and $g^{-1}Hg \cap \Gamma$ is a lattice of $g^{-1}Hg$.

As X_{cpt} is compact there exists $\delta > 0$ such that, for any $x \in X_{\mathrm{cpt}}$ the map $B_H(\delta) \to X$ given by $h \mapsto h.x$ is injective. Then there exists, by Proposition 5.3, $|\lambda| \ll \mathrm{vol}(Q)^{\frac{13}{5}}$ such that

$$D(\lambda)B_H(\delta)y \cap B_H(\delta)y \neq \emptyset$$

for any point $y \in X_{\mathrm{cpt}}$ with $H.y$ closed in X. Hence, by Anosov's closing lemma (see [3]) there exist $|\lambda_1| \ll \mathrm{vol}(Q)^{\frac{13}{5}}$ and $z \in B_H(\delta)g$, such that $D(\lambda_1).[z] = [z]$. We have thus found a hyperbolic element $\gamma = z^{-1}D(\lambda_1)z$ such that

$$\gamma \in g^{-1}Hg \cap \Gamma, \qquad \|\gamma\| \ll \mathrm{vol}(Q)^{\frac{13}{5}}.$$

Notice that the only connected, two-dimensional, algebraic subgroup of SO_Q is the Borel subgroup. Using the same argument with g replaced by elements in $B_H(1)g$, we see that there exist two hyperbolic elements $\gamma_1, \gamma_2 \in g^{-1}Hg \cap \Gamma$ generating a Zariski dense subgroup of SO_Q and satisfying

$$\|\gamma_1\| \ll \mathrm{vol}(Q)^{\frac{13}{5}}, \qquad \|\gamma_2\| \ll \mathrm{vol}(Q)^{\frac{13}{5}}.$$

On the other hand, the matrix equations

$$X^t = X, \qquad \gamma_1^t X \gamma_1 = X, \qquad \gamma_2^t X \gamma_2 = X \tag{20}$$

have integral solutions kA, where k is any integer and A the symmetric matrix of Q. So the rank of (20), as a system of linear Diophantine equations, is at most 5. By [1] it has a small solution with $\|X\| \ll \mathrm{vol}(Q)^{13}$. Because Q is primitive and γ_1, γ_2 generate a Zariski dense subgroup of SO_Q, we have $\|A\| \ll \mathrm{vol}(Q)^{13}$. Therefore $|\det Q| = |\det A| \ll \mathrm{vol}(Q)^{39}$. □

Acknowledgments

We would like to thank Zeev Rudnick and Peter Sarnak for sharing their insight in the theory of quadratic forms. H. Li would like to thank Wai-Kiu Chan and Han Wu for helpful discussions. H. Li was supported in part by NSF Grant #1700109 and a grant from the Simons Foundation (#426090, Han Li). G.A. Margulis was supported in part by NSF Grant #1265695.

References

[1] I. Borosh, M. Flahive, D. Rubin, B. Treybig, A sharp bound for solutions of linear Diophantine equations, Proc. Amer. Math. Soc. 105 (4) (1989) 844–846.

[2] M. Einsiedler, G.A. Margulis, A. Venkatesh, Effective equidistribution for closed orbits of semisimple groups on homogeneous spaces, Invent. Math. 177 (1) (2009) 137–212.

[3] B. Hasselblatt, A. Katok, Introduction to the Modern Theory of Dynamical Systems, Encyclopedia Math. Appl., vol. 54, Cambridge University Press, Cambridge, 1995.

[4] H. Jacquet, R. Langlands, Automorphic Forms on GL(2), Lecture Notes in Math., vol. 114, Springer-Verlag, Berlin, New York, 1970.

[5] H.H. Kim, P. Sarnak, Refined estimates towards the Ramanujan and Selberg conjectures, Appendix to H. Kim, Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2, J. Amer. Math. Soc. 16 (1) (2003) 139–183.

[6] H. Li, G.A. Margulis, Effective estimates on integral quadratic forms: Masserís conjecture, generators of orthogonal groups, and bounds in reduction theory, Geom. Funct. Anal. 26 (3) (2016) 874–908.

[7] E. Lindenstrauss, G.A. Margulis, Effective estimates on indefinite ternary forms, Israel J. Math. 203 (1) (2014) 445–499.

[8] G.A. Margulis, Indefinite quadratic forms and unipotent flows on homogeneous spaces, in: Proceeding of "Semester on Dynamical Systems and Ergodic Theory", Warsaw, 1986, in: Banach Center Publ., vol. 23, PWN, Warsaw, 1989, pp. 399–409.

[9] A. Mohammadi, A special case of effective equidistribution with explicit constants, Ergodic Theory Dynam. Systems 32 (1) (2012) 237–247.

[10] H. Oh, Orbital counting via mixing and unipotent flows, in: Homogeneous Flows, Moduli Spaces and Arithmetic, in: Clay Math. Proc., vol. 10, 2010, pp. 339–375.

[11] Y. Shalom, Rigidity, unitary representations of semisimple groups, and fundamental groups of manifolds with rank one transformation group, Ann. of Math. (2) 152 (1) (2000) 113–182.

[12] C.L. Siegel, Zur Theorie der quadratischen Formen, Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II (1972) 21–46.