*Ergod. Th. & Dynam. Sys.* (2012), **32**, 237–247 doi:10.1017/S0143385710000799
© *Cambridge University Press, 2011*

# A special case of effective equidistribution with explicit constants

A. MOHAMMADI

*Mathematics Department, University of Chicago, Chicago, IL, USA*
*(e-mail: amirmo@math.uchicago.edu)*

(*Received* 18 *July* 2010 *and accepted in revised form* 9 *October* 2010)

*Abstract*. An effective equidistribution with explicit constants for the isometry group of rational forms with signature (2, 1) is proved. As an application we get an effective discreteness of the Markov spectrum.

## 1. *Introduction*

Let $B(v) = 2v_1v_3 - v_2^2$ for any $v \in \mathbb{R}^3$ and let $H = \mathrm{SO}(B)$. We denote $X = \mathrm{SL}_3(\mathbb{R})/\mathrm{SL}_3(\mathbb{Z})$. Let $m$ be the $\mathrm{SL}_3(\mathbb{R})$-invariant Haar measure on $X$ normalized so that $m(X) = 1$. Let $\mathcal{A} \neq X$ be a closed subset of $X$ which is $H$-invariant. It was proved by Dani and Margulis [**DM89**] that $\mathcal{A}$ is a union of finitely many closed $H$-orbits. Indeed [**DM89**] is a generalization of the ideas and results of Margulis' work in the proof of the Oppenheim conjecture [**Mar86, Mar90a**] and it is a special case of Ratner's equidistribution theorem [**MT94, R90, R91, R92, R95**] for the proof of the measure classification and equidistribution theorems.

Recently Einsiedler *et al* in a landmark paper [**EMV09**] proved a polynomially effective equidistribution theorem for closed orbits of semisimple groups. Their result is quite general and the proof is very involved. Our result in this paper gives the explicit exponent in a very special case of the main theorem in [**EMV09**]. Since we are dealing with a very special case, the proof simplifies quite considerably. In this special case the calculation of the exponent becomes much simpler and this is what is carried out here. Let us also mention that the main result of [**EMV09**] and our result here are indeed effective versions of a special case of the results obtained by Mozes and Shah [**MS95**] regarding the weak*-limits of measures invariant and ergodic with respect to a unipotent flow.

As we mentioned, we borrow extensively from the results and techniques in [**EMV09**]. There is however one main technical difference between this paper and [**EMV09**]. In this paper we use translates of a small piece of a unipotent orbit in $H$ by a particular torus in $H$ in order to get an 'effective ergodic' theorem. In [**EMV09**] however the authors utilize such an ergodic theorem for the action of unipotent subgroups. See also [**E06**] where the action of unipotent groups is used to get measure classification for the action of semisimple groups. The main reason for this variant in the method is to get a 'bigger'

https://doi.org/10.1017/S0143385710000799 Published online by Cambridge University Press

exponent. The idea of using the torus in this way goes back to Margulis, where in an unpublished manuscript he proved a topological version of the main result of this paper.

Let $x_i \in X$, $1 \le i \le N$, be such that $Hx_i$ is closed for all $1 \le i \le N$. Let

$$\mathcal{A} = \bigcup_{1 \le i \le N} Hx_i \quad \text{and let } \operatorname{vol}(\mathcal{A}) = \sum_i \operatorname{vol}(Hx_i). \tag{1}$$

We let $\mu_i$ denote the $H$-invariant probability measure on $Hx_i$. We prove the following theorem.

THEOREM 1.1. *Let the notation be as above. In particular, let $\mathcal{A}$ be a union of finitely many closed $H$-orbits. There exist absolute constants $c, M_0 > 0$ such that if $\operatorname{vol}(\mathcal{A}) = M > M_0$ then there exists $1 \le i_0 \le N$ such that for any $f \in C_c^\infty(X)$ we have*

$$\left| \int_X f(x)\, d\mu_{i_0}(x) - \int_X f(x)\, dm(x) \right| < c\mathcal{S}_{20}(f) \operatorname{vol}(\mathcal{A})^{-\delta}, \tag{2}$$

*where $\mathcal{S}_{20}$ is a certain Sobolev norm; see §2 for the definition. In particular $\delta = 10^{-7}$ satisfies (2).*

An application of this theorem to effective discreteness of the Markov spectrum is given at the end of this paper; see Corollary 5.1.

## 2. *Notation and preliminaries*

Throughout the paper we let $G = \mathrm{SL}_3(\mathbb{R})$ and $\Gamma = \mathrm{SL}_3(\mathbb{Z})$. We let

$$a_t = \begin{pmatrix} e^t & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-t} \end{pmatrix}, \quad u_s = \begin{pmatrix} 1 & s & s^2/2 \\ 0 & 1 & s \\ 0 & 0 & 1 \end{pmatrix}. \tag{3}$$

The subgroups $\{a_t : t \in \mathbb{R}\}$ and $\{u_s : s \in \mathbb{R}\}$ are one-parameter subgroups of $H = \mathrm{SO}(B)$. Let $\mathfrak{g} = \mathfrak{sl}_3(\mathbb{R})$ (respectively $\mathfrak{h}$) denote the Lie algebra of $G$ (respectively $H$). We let $\|\cdot\|$ be a fixed Euclidean norm on $\mathfrak{g}$ and fix an orthonormal basis for $\mathfrak{g}$ with respect to this norm. We may write $\mathfrak{g} = \mathfrak{r} \oplus \mathfrak{h}$ where $\mathfrak{r}$ is the $\mathrm{Ad}(H)$-invariant complement to $\mathfrak{h}$. Further we let $\mathfrak{r}_0$ denote the centralizer of $\{u_s\}$ in $\mathfrak{r}$ and let $\mathfrak{r}_1$ denote the orthogonal complement of $\mathfrak{r}_0$ in $\mathfrak{r}$ with respect to $\|\cdot\|$. For any $r \in \mathfrak{r}$ let $r_1$ (respectively $r_0$) denote the $\mathfrak{r}_1$ (respectively $\mathfrak{r}_0$) component of $r$ and define $Z_r = r_0/\|r_0\|$ with $Z_r = 0$ if $r_0 = 0$. Throughout exp will denote the exponential map from $\mathfrak{g}$ into $G$.

The Euclidean norm $\|\cdot\|$ on $\mathfrak{g}$ defines a right-invariant Riemannian metric on $G$ which in turn gives a metric on $X = G/\Gamma$. The Riemannian metric on $G$ gives a volume form on subgroups of $G$; this will be denoted by vol in the following. We will use Greek letter $\mu$ and $\mu_i$ to denote the $H$-invariant probability Haar measure on closed $H$-orbits.

Let $\Gamma = \mathrm{SL}_3(\mathbb{Z})$ and $X = G/\Gamma$. This is the space of unimodular lattices in $\mathbb{R}^3$. For any $x \in X$, we let $g_x\mathbb{Z}^3$ denote the corresponding lattice in $\mathbb{R}^3$; and we define $\alpha_1(x) = \max_{0 \ne v \in g_x\mathbb{Z}^3} 1/|v|$. By Mahler's compactness criteria the set

$$\mathfrak{S}(R) = \{x \in X : \alpha_1(x) \le R\} \tag{4}$$

is compact. We will use the Greek letter $\mu$ and also $\mu_i$ for $H$-invariant probability measures on closed $H$-orbits in $X$ and will save the letter $m$ for the $G$-invariant probability measure

https://doi.org/10.1017/S0143385710000799 Published online by Cambridge University Press

on $X$. Recall that $H$ is a maximal subgroup of $G$. In particular $H$ is not contained in any proper parabolic subgroup of $G$. Using this and quantitative non-divergence results proved by Kleinbock and Margulis [**KM98**] (see also [**EMV09**, Appendix B]), we have the following lemma.

LEMMA 2.1. *With the notation as above:*

(i) *there exists $R_1 > 1$ such that any $H$-orbit in $X$ intersects $\mathfrak{S}(R_1)$; and*

(ii) *there exists a constant $c$ depending on $R_1$ such that for any $H$-invariant measure $\mu$ on $X$ we have $\mu(\{x \notin \mathfrak{S}(R)\}) \leq cR^{-1/2}$.*

Let $R_0$ be such that $\mu(X \backslash \mathfrak{S}(R_0)) < 10^{-11}$. We define $X_{\text{cpct}} = \mathfrak{S}(R_0)$. This will be our favorite fixed compact set throughout this paper.

We will define $\|\cdot\|$ on $G$ by letting $\|g\|$ be the maximum of the absolute value of matrix coefficients of $g$ and $g^{-1}$. This should not cause any confusion with $\|\cdot\|$ on $\mathfrak{g}$ in the following.

The letter $f$ will denote an element of $C_c^\infty(X)$ throughout the following. If $f \in C_c^\infty(X)$ then $\|f\|_2$ (respectively $\|f\|_\infty$) denotes the $L^2$ norm (respectively $L^\infty$ norm) of $f$; and $G$ acts on these spaces by $g \cdot f(x) = f(g^{-1}x)$. For the sake of simplicity in notation if $\sigma$ is a measure on $X$ we will denote $\sigma(f) = \int_X f \, d\sigma$. For any $g \in G$ we let $\sigma^g$ be the measure defined by $\sigma^g(f) = \sigma(g^{-1}f)$, for any $f \in C_c^\infty(X)$.

We now recall the definition of a certain family of Sobolev norms which were also used in [**EMV09**]. For any integer $d \geq 0$ we let $\mathcal{S}_d$ be the Sobolev norm on $X$ defined by

$$\mathcal{S}_d(f)^2 = \sum_{\mathcal{D}} \|\alpha_1(\cdot)^d \mathcal{D} f\|_2^2, \tag{5}$$

where the sum is taken over all monomials of degree at most $d$ in the fixed basis for $\mathfrak{g}$. We will need the following three properties of $\mathcal{S}_d$. Let $d \geq 8$, then for any $g \in G$ and $f \in C_c^\infty(X)$ we have:

(i) $\mathcal{S}_d(gf) \leq c(d)\|g\|^{2d}\mathcal{S}_d(f)$;

(ii) $\|f\|_\infty \leq c'(d)\mathcal{S}_d(f)$; and

(iii) $\|gf - f\|_\infty \leq d(e, g)\mathcal{S}_d(f)$.

For a discussion of the Sobolev norm and the proofs of the above properties we refer to [**EMV09**, §5].

Let $\mathcal{S}_d$ be as above and let $\varepsilon > 0$. We say that the measure $\sigma$ is $\varepsilon$-almost invariant under $g \in G$ if $|\sigma(gf) - \sigma(f)| < \varepsilon \mathcal{S}_d(f)$. We say that $\sigma$ is $\varepsilon$-almost invariant under $G$ if it is $\varepsilon$-almost invariant under all $g \in G$ with $\|g\| \leq 2$.

In the proof of Proposition 3.1 we will need to use the notion of relative trace; see [**EMV09**, §5.2] and [**BR02**, Appendix 2] for a more extensive study of these ideas. Let $V$ be a complex vector space and $A$ and $B$ be two non-negative Hermitian forms on $V$. Further assume that $B$ is positive definite. If $V$ is finite dimensional, define the relative trace by $\text{Tr}(A|B) = \text{Tr}(B^{-1}A)$. It follows from the definition that we have the following formula for the relative trace:

$$\text{Tr}(A|B) = \sum_i \frac{A(e_i)}{B(e_i)}, \tag{6}$$

https://doi.org/10.1017/S0143385710000799 Published online by Cambridge University Press

where $\{e_i\}$ is any orthonormal basis for $V$ with respect to $B$. If $V$ is infinite dimensional, define

$$\mathrm{Tr}(A|B) = \sup_{W \subset V} \mathrm{Tr}(A_W|B_W), \tag{7}$$

where the supremum is taken over all finite dimensional subspaces and $A_W$, $B_W$ denote restriction of $A$ and $B$ to $W$. Note that $\mathcal{S}_d$ defines a Hermitian norm in $C_c^\infty(X)$. We have that for every $d$ there exists $d' > d$ such that

$$\mathrm{Tr}(\mathcal{S}_d^2|\mathcal{S}_{d'}^2) < \infty. \tag{8}$$

That is to say there exists an orthonormal basis $\{e_i\}$ for the completion of $C_c^\infty(X)$ with respect to $\mathcal{S}_{d'}$ such that $\sum_i \mathcal{S}_d(e_i)^2 < \infty$.

## 3. *Producing effective extra invariant*

In this section we will use an effective 'ergodic' theorem (see Proposition 3.1) and show that we can effectively produce elements outside $H$ which will move one closed orbit $Hx_i$ in $\mathcal{A}$ 'close' to another closed orbit $Hx_j$ in $\mathcal{A}$; see Proposition 3.2 below.

*Generic points.* Let $t \geq 1$. For any $x \in X$ define

$$D_t^f(x) = e^{7t/24} \int_0^{e^{-7t/24}} f(a_t u_s x)\, ds - \int_X f\, d\mu. \tag{9}$$

Let $t \in \mathbb{R}$ be a positive number. A point $x \in X$ is called *t-generic* with respect to the Sobolev norm $\mathcal{S}$ if for any integer $n \geq t$ we have

$$|D_n^f(x)| \leq e^{-n/8}\mathcal{S}(f). \tag{10}$$

As is clear from the definition above, generic points provide us with an effective ergodic theorem. The following proposition shows that 'most' points are generic.

PROPOSITION 3.1. *There exists $d_0$ depending on $G$ and $H$ only such that for all $d \geq d_0$ the $\mu$-measure of points which are not $t$-generic with respect to $\mathcal{S}_d$ is $O(e^{-t/24})$. Furthermore $d = 20$ satisfies this property.*

*Proof.* Let $f$ be a fixed function in $C_c^\infty(X)$. Let $\langle \cdot, \cdot \rangle$ denote the inner product on $L^2(X, \mu)$. Note that $\dim H = 3$. As a result of an explicit estimate for the decay of matrix coefficients [**KS03**] we have

$$\left| \langle u_s f, f \rangle - \left( \int_X f\, d\mu \right)^2 \right| \leq (1 + |s|)^{-0.7} \mathcal{S}_3(f)^2. \tag{11}$$

Let $A = \{(s_1, s_2) \in I \times I : |s_1 - s_2| < e^{-7t/12}\}$ and $B = I \times I \backslash A$, where $I = [0, e^{-7t/24}]$. Squaring $D_t^f$ we have

$$\begin{aligned}
\int_X |D_t^f(x)|^2\, d\mu &= e^{7t/12} \int_I \int_I \langle a_t u_{s_1 - s_2} a_{-t} f, f \rangle\, ds_1\, ds_2 - \left( \int_X f\, d\mu \right)^2 \\
&= e^{7t/12} \left( \int_A \langle a_t u_{s_1 - s_2} a_{-t} f, f \rangle\, ds_1\, ds_2 + \int_B \langle a_t u_{s_1 - s_2} a_{-t} f, f \rangle\, ds_1\, ds_2 \right) \\
&\quad - \left( \int_X f\, d\mu \right)^2 \leq 3e^{-7t/24} \mathcal{S}_3(f)^2.
\end{aligned}$$

https://doi.org/10.1017/S0143385710000799 Published online by Cambridge University Press

Hence we have

$$\mu(\{x \in X : |D_t^f(x)| > r\}) \leq 3r^{-2}e^{-7t/24}\mathcal{S}_3(f)^2. \tag{12}$$

The rest of the proof is *mutatis mutandis* that of [**EMV09**, proof of Proposition 9.2]. Let us recall this proof. First choose $d > d' > 3$ such that $\mathrm{Tr}(\mathcal{S}_{d'}^2|\mathcal{S}_d^2)$ and $\mathrm{Tr}(\mathcal{S}_3^2|\mathcal{S}_{d'}^2)$ are both finite; see §2 and references therein. Now choose an orthonormal basis $\{e_i : i \geq 1\}$ for the completion of $C_c^\infty(X)$ with respect to $\mathcal{S}_d$ which is also orthogonal with respect to $\mathcal{S}_{d'}$; note that this choice can be made thanks to the spectral theorem.

Let $c = (\sum_i \mathcal{S}_{d'}(e_i)^2)^{-1/2}$ and define

$$E = \bigcup_{n \geq t, i \geq 1} \{x \in X : |D_n^{e_i}(x)| \geq 3ce^{-n/8}\mathcal{S}_{d'}(e_i)\}. \tag{13}$$

Using (12) and the fact that $\mathrm{Tr}(\mathcal{S}_3^2|\mathcal{S}_{d'}^2)$ is finite, we have

$$\mu(E) \leq \frac{\mathrm{Tr}(\mathcal{S}_3^2|\mathcal{S}_{d'}^2)}{c^2} e^{-n/24}. \tag{14}$$

Now using the linearity of $D_n$ and the Cauchy–Schwarz inequality we have that if $f = \sum_i f_i e_i \in C_c^\infty(X)$ then for any $n \geq t$ and any $x \notin E$ we have $|D_n^f(x)| \leq e^{-n/8}\mathcal{S}_d(f)$. This combined with (14) finishes the proof of the proposition. To see the last conclusion note that $\dim(G) = 8$. Hence it follows from the properties of the Sobolev norm [**EMV09**, §5] that $d' = 3 + 8$ and $d = 11 + 8$ satisfy the properties we need in the above argument. □

*Producing extra invariants.* Let $\mu_1$ and $\mu_2$ be two $H$-invariant probability measures on $X$. We use Proposition 3.1 to show that if the supports of these measures come 'very close' to each other then we may effectively produce elements in the direction of $\mathfrak{r}_0$ which move $\mu_1$ 'close' to $\mu_2$. This idea is by no means new and has been used by several people—to give two important examples we refer to Margulis's proof of the Oppenheim conjecture and Ratner's proof of the measure rigidity conjecture. Here using generic points we will give an effective version of this phenomena. This effective form is one of the main ingredients in [**EMV09**] as well. The following is the precise statement.

PROPOSITION 3.2. *Let $\mathcal{S}_d$ be as in Proposition 3.1 and let $\iota > 0$ be a constant. Suppose that $\mu_1$ and $\mu_2$ are two $H$-invariant measures and that $y_1, y_2 \in X$ are such that $y_2 = \exp(r)y_1$, where $r \in \mathfrak{r}$ for some $\|r\| < 1$ is implicitly assumed to be small enough and that $\|r_0\| \geq \iota\|r\|$. Moreover assume that the $y_i$ are $\log(1/\sqrt{\|r\|})$-generic for $\mu_i$ with respect to $\mathcal{S}_d$ for $i = 1, 2$. Then for any $\tau_1 \geq 1$ there exists a constant $c = c(d, \tau_1, \iota)$ such that*

$$|\mu_1^{\exp(\tau Z_r)}(f) - \mu_2(f)| \leq c\|r\|^{-1/336}\mathcal{S}_d(f) \quad \text{for all } 0 \leq \tau \leq \tau_1, \tag{15}$$

*where $Z_r$ is defined as in §2.*

The proof of this proposition is essentially based on the following fact: if $r \in \mathfrak{g}$ is 'generic enough' then $\mathrm{Ad}(a_t)r$ will be 'close' to $\mathfrak{r}_0$. This indeed is a consequence of the fact that $\mathfrak{r}_0$ is the direction of maximum expansion for the conjugation action of $a_t$ for $t > 0$. The genericity condition we impose is $\|r_0\| \geq \iota\|r\|$. Note that the adjoint action of a 'small piece' of $u_s$ leaves $r$ 'almost' fixed. We now flow by $a_t u_s$ from $y_2 = \exp(r)y_1$. The above discussion implies that this orbit is 'close' to a translate of the orbit of $y_1$ under this flow, by an element in $\exp(\mathfrak{r}_0)$. The proposition then follows from the definition of a generic point and some continuity arguments.

https://doi.org/10.1017/S0143385710000799 Published online by Cambridge University Press

*Proof.* Let us start with the following simple calculation: if $r = r_0 + r_1 \in \mathfrak{r}$ then

$$\mathrm{Ad}(u_s)(r) = r + sp(r, s) \quad \text{where } p(r, s) \in \mathfrak{r} \text{ and } \|p(r, s)\| \le c(s)\|r\|$$

and $c(s)$ is a constant depending on $s$ and the norm $\|\cdot\|$. In particular if $|s| \le 1$, then $c(s)$ is an absolute constant and we will omit it from the notation from now. Hence for any positive integer $n > 0$ and any $0 \le s \le e^{-7n/24}$ we have

$$\|(\mathrm{Ad}(u_s)(r)) - r\| \le e^{-7n/24}\|r\|.$$

Now let $r$ be as in the statement of the proposition. For any positive integer $n > 0$ we have $\|\mathrm{Ad}(a_n)r_1\| \le e^n\|r\|$. Let $n > 0$ be an integer so that $1 \le e^{2n}\|r_0\| \le e^4$. Our assumption, $\|r_0\| \ge \iota\|r\|$, implies that $e^n\|r\| \le c(\iota)e^{-n}$. We thus have

$$\|\mathrm{Ad}(a_n u_s)r - e^{2n}r_0\| \le e^n\|r\| \le c(\iota)e^{-n}. \tag{16}$$

Using (16) and the property (iii) of the Sobolev norm $\mathcal{S}_d$ given above, we have

$$\|f(a_n u_s \exp(r)y_1) - f(\exp(e^{2n}r_0)a_n u_s y_1)\|_\infty \le c(\iota)e^{-n}\mathcal{S}_d(f). \tag{17}$$

Recall now that the $y_i$ are $\log(1/\sqrt{\|r\|})$-generic for $\mu_i$. This together with our choice of the integer $n$ gives

$$\left|e^{7n/24}\int_0^{e^{-7n/24}} f(a_n u_s y_i)\,ds - \int_X f\,d\mu_i\right| \le e^{-n/8}\mathcal{S}_d(f) \quad \text{for } i = 1, 2. \tag{18}$$

We now combine (17) and (18) and get

$$\left|e^{7n/24}\int_0^{e^{-7n/24}} f(a_n u_s \exp(r)y_1)\,ds - \mu_1^{\exp(e^{2n}r_0)}(f)\right| \le 2c(\iota)e^{-n/8}\mathcal{S}_d(f). \tag{19}$$

The fact that $y_2 = \exp(r)y_1$, (18) and (19) now imply

$$|\mu_1^{\exp(e^{2n}r_0)}(f) - \mu_2(f)| \le ce^{-n/8}\mathcal{S}_d(f) \tag{20}$$

for some constant $c$ depending on $\iota$, which proves (15) for some $\tau_0 \ge 1$. To prove the proposition for all values of $0 \le \tau \le \tau_1$, note that

$$|\mu_1^{a_t \exp(\tau_0 Z_r)a_{-t}}(f) - \mu_2(f)| = |\mu_1^{a_t \exp(\tau_0 Z_r)}(f) - \mu_2^{a_t}(f)| \le ce^{-n/8}\mathcal{S}_d(a_t \cdot f). \tag{21}$$

Recall that we may let $d = 20$ in the above calculations, hence property (i) of $\mathcal{S}_d$ gives $\mathcal{S}_d(a_t \cdot f) \le e^{40t}\mathcal{S}_d(f)$. This and (21) give the result for $e^{-n/168} \le \tau \le \tau_1$. The case $0 \le \tau < e^{-n/168}$ follows using a continuity argument; see property (iii) of the Sobolev norm. □

*Finding generic points close to each other.* Let the notation be as before. In particular $\mathcal{A}$ is a subset of $X$ which consists of a finite number of closed $H$-orbits and let $M = \mathrm{vol}(\mathcal{A})$ be as in (1). In order to be able to apply Proposition 3.2 we need to produce close-by generic points. The following proposition provides us with such points. This proposition is essentially a result of the Dirichlet pigeonhole principle.

https://doi.org/10.1017/S0143385710000799 Published online by Cambridge University Press

PROPOSITION 3.3. *There exist absolute constants $c, \iota > 0$ and a Sobolev norm $\mathcal{S}_d$ with the following property: there are $Hx_i \in \mathcal{A}$ and points $y_i \in Hx_i$ for $i = 1, 2$, so that $y_2 = \exp(r)y_1$, where $r \in \mathfrak{r}$ with $\|r\| < cM^{-1/5}$ and $\|r_0\| \geq \iota\|r\|$. Moreover $y_i$ is $\log(1/\sqrt{\|r\|})$-generic for $\mu_i$ with respect to $\mathcal{S}_d$, where $\mu_i$ is the $H$-invariant probability measure on the closed orbit $Hx_i$ for $i = 1, 2$.*

*Proof.* This is virtually a special case of [**EMV09**, Proposition 14.2]. We reproduce the proof here. We first work with a single closed $H$-orbit. So let $Hx$ be a closed $H$ orbit and $\mu$ be the $H$-invariant probability measure on $Hx$. Let $t$ be a large number so that $e^{-t/24} < 10^{-11}$ and let $E'_{Hx}$ be the set of $t$-generic points obtained by Proposition 3.1 and $E_{Hx} = E'_{Hx} \cap X_{\text{cpct}}$. For any $\delta > 0$ let $\mathfrak{r}_\delta$ (respectively $\mathfrak{h}_\delta$) denote the ball of radius $\delta$ in $\mathfrak{r}$ (respectively $\mathfrak{h}$) around the origin with respect to the norm $\|\cdot\|$. Let $\delta_0$ be small enough such that for any $\delta < \delta_0$ and any $z \in X_{\text{cpct}}$ the natural map from $\mathfrak{r}_\delta \times \mathfrak{h}_\delta$ to $X$ given by $\pi_z(r, h) = \exp(r)\exp(h)z$ is a diffeomorphism. Indeed we are also assuming that the restriction of the exponential map to $\mathfrak{r}_\delta \times \mathfrak{h}_\delta$ into $G$ is a diffeomorphism. We let $\Omega = \exp(\mathfrak{h}_{\delta_0/2})$. Define the following function on $X$:

$$\phi(z) = \frac{1}{\text{vol}(\Omega)} \int_\Omega \chi_E(hz)\, d\,\text{vol}(h). \tag{22}$$

We have $\int_X \phi(z)\, d\mu = \mu(E_{Hx}) > 1 - (2/10^{11})$. Hence the set

$$F_{Hx} = \{z \in E_{Hx} : \phi(z) > 0.99\}$$

has measure at least 0.9. Define

$$\mathcal{F} = \mathcal{F}(\mathcal{A}) = \bigcup_{Hx \in \mathcal{A}} F_{Hx}. \tag{23}$$

Note that $\text{vol}(\mathcal{F}) \geq 0.9M$. Now let $\delta < \delta_0$ be given; this will be determined in terms of $M$ later. We let $B_\delta z = \pi_z(\mathfrak{r}_\delta \times \mathfrak{h}_\delta)$, i.e. the image of $\pi_z$. We may cover $\mathcal{F}$ by finitely many sets of the form $B_\delta z$ with finite multiplicity $c''$ independent of $\delta$. In this covering the center $z$ of each $B_\delta z$ is assumed to be in $\mathcal{F}$. Now using the Dirichlet pigeonhole principle we have that there is an absolute constant $c'$ depending on $c''$ such that if we let $\delta = c'M^{-1/5}$ and if $M$ is large enough such that $\delta < \delta_0$ then the following holds: there are $Hx_1$ and $Hx_2$ in $\mathcal{A}$ and $y'_i \in F_{Hx_i}$ for $i = 1, 2$ such that $y'_i \in B_\delta z$ for some $z$ and $y_1 \neq hy_2$ for any $h \in \Omega$. We now want to perturb the $y'_i$ within $F_{Hx_i}$ to guarantee that they satisfy the conclusion of the lemma. First let us remark on the following simple statement: there is a constant $\iota > 0$ such that

$$\text{vol}\left\{h \in \Omega : \frac{\|(\text{Ad}(h)r)_0\|}{\|r\|} \leq \iota\right\} < \text{vol}(\Omega)/2. \tag{24}$$

This indeed is a result of a compactness argument and the fact that the Ad-representation of $H$ on $\mathfrak{r}$ is irreducible. Now if we apply the implicit function theorem and use the fact that $\phi(y'_i) > 0.99$, we can find $h_i \in \Omega$ such that $h_i y'_i \in F_{Hx_i}$ and $h_2 y_2 = \exp(r)h_1 y'_1$, where $\|r\| \leq cM^{-1/5}$, for some absolute constant depending on $c'$; moreover we have $\|r_0\| > \iota\|r\|$. So $y_i = h_i y'_i$ for $i = 1, 2$ satisfy the conclusion of the proposition. □

https://doi.org/10.1017/S0143385710000799 Published online by Cambridge University Press

## 4. *Proof of Theorem 1.1*

We will use the results from the last section here and prove the main theorem. We first need the following statement.

*G-almost invariant measures are close to the Haar measure.* This follows from the fact that the action of $G$ on $X$ has spectral gap. This has indeed been shown in [**EMV09**, Proposition 15.1]. Before stating the proposition let us recall here that the conclusion of Proposition 3.2 holds for $d = 20$. This will be used when we iterate the $\varepsilon$-almost invariance inequality. We have the following proposition.

PROPOSITION 4.1. *Let $\sigma$ be a probability measure on $X$ which is $\varepsilon$-almost invariant under $G$ with respect to $\mathcal{S}_d$. Then there is some constant $C$ such that*

$$|\sigma(f) - m(f)| \le C\varepsilon^{0.0002}\mathcal{S}_d(f), \tag{25}$$

*where $f$ is in $C_c^\infty(X)$ as before.*

*Proof.* We will reproduce the proof given in [**EMV09**]. We begin by fixing some notation. For any real number $r \ge 4$ we let

$$A_r = \left\{ a = \operatorname{diag}(a_1, a_2, a_3) : 1 \le \frac{a_1}{a_2} \le r,\ 1 \le \frac{a_2}{a_3} \le r \right\}.$$

It is an exercise in linear algebra to see that for any $a, b \in A_r$ with $d(a, e) > r/2$ and $d(b, e) > r/2$ the measure of $\{k \in K : d(akb, e) < r/2\}$ is at most half the measure of $K$. Let $\chi$ be the characteristic function of $KA_6K$ normalized to have mean one. It follows from this fact and [**Oh02**] (see in particular the main theorem and the discussion in §8 therein) that for any $f \in C_c^\infty(X)$ with $m(f) = 0$ we have $\|\chi \star f\|_2 \le (1/2 + \delta/2)\|f\|_2$, where $\star$ denotes the convolution and $\delta \le (11.09 + \log 36)/6 < 1/2$. Hence we get

$$\|\chi \star f\|_2 \le \tfrac{3}{4}\|f\|_2. \tag{26}$$

It suffices to show (25) for function $f$ with $m(f) = 0$. We have $|\sigma(\chi \star f) - \sigma(f)| < c(d)\varepsilon\mathcal{S}_d(f)$ as $\sigma$ is $\varepsilon$-almost invariant under $G$. Iterating this, see [**EMV09**, §8.2], we have

$$|\sigma(\chi^n \star f) - \sigma(f)| < c(d)6^{40n}\varepsilon\mathcal{S}_d(f). \tag{27}$$

Note that we have the trivial bound $|\chi^n \star f(x)| \le \|f\|_\infty \le \mathcal{S}_d(f)$. Note also that the fibers of the map $x \mapsto gx$ are at most of size $c\alpha_1(x)^8$ for a constant depending on $r$. Now using the Cauchy–Schwarz inequality we have

$$|\chi \star f(x)| \le \alpha_1(x)^8\|f\|_2. \tag{28}$$

Hence we get the more interesting bound $|\chi^n \star f(x)| \le \alpha_1(x)^8(3^n/4^n)\|f\|_2$. If we combine these inequalities now, we obtain

$$\sigma(f) \le C(d)\mathcal{S}_d(f)\left(6^{40n}\varepsilon + \int_X \min(1, \alpha_1(x)^8)\frac{3^n}{4^n}\,d\sigma\right). \tag{29}$$

Divide $X$ into $\mathfrak{S}(R)$ and its complement. Using (29) and Lemma 2.1 we have

$$\sigma(f) \le C(d)\mathcal{S}_d(f)\left(6^{40n}\varepsilon + R^8\frac{3^n}{4^n} + R^{-1/2}\right). \tag{30}$$

https://doi.org/10.1017/S0143385710000799 Published online by Cambridge University Press

Choosing $R$ and $n$ such that the terms in parentheses are of the same size we get $|\sigma(f)| < \varepsilon^{\kappa_2}\mathcal{S}_d(f)$ and

$$\kappa_2 = \frac{1/(17\log(3/4))}{1/(17\log(3/4)) - 40\log 6}.$$

Hence $\kappa_2 \geq 0.0002$. □

*Proof of Theorem 1.1.* Let $\mathcal{A}$ be as in the statement of Theorem 1.1 and let $M = \mathrm{vol}(\mathcal{A})$. We apply Proposition 3.3. Hence we find $Hx_i \in \mathcal{A}$ and points $y_i \in Hx_i$ for $i = 1, 2$, so that $y_2 = \exp(r)y_1$, where $r \in \mathfrak{r}$ with $\|r\| < cM^{-1/5}$ and $\|r_0\| \geq \iota\|r\|$. Moreover $y_i$ is $\log(1/\sqrt{\|r\|})$-generic for $\mu_i$ with respect to $\mathcal{S}_d$, where $\mu_i$ is the $H$-invariant probability measure on the closed orbit $Hx_i$ for $i = 1, 2$. We now apply Proposition 3.2 to these $\mu_i$ and $y_i$ and get

$$|\mu_1^{\exp(\tau Z_r)}(f) - \mu_2(f)| \leq c(d, \tau_0, \iota)M^{-1/1680}\mathcal{S}_d(f) \quad \text{for all } 0 \leq \tau \leq \tau_0. \tag{31}$$

From this we have the following claim.

CLAIM. *There is an absolute constant $c$ such that for any $g \in G$ with $d(1, g) \leq 2$ we have*

$$|\mu_1^g(f) - \mu_1(f)| \leq M^{-1/1680}\mathcal{S}_d(f). \tag{32}$$

*In other words $\mu_1$ is $M^{-1/1680}$-almost invariant under $G$ with respect to $\mathcal{S}_d$.*

*Proof of the claim.* This indeed is a corollary of (31) above and the fact that $H$ is a maximal subgroup of $G$. Let $H_1 = \exp(-Z_r)H\exp(Z_r)$ and let $g \in H_1$. Now using (31), the $H$-invariance of $\mu_2$ and the properties of the Sobolev norm we have

$$|\mu_1^g(f) - \mu_1(f)| \leq c(g, d)M^{-1/1680}\mathcal{S}_d(f) \tag{33}$$

and $c(g, d)$ is uniform on compact sets. Let $\mathcal{Y}_\iota = Y_1 \times \cdots \times Y_\iota$, where each $Y_i$ is either $H$ or $H_1$. As $G$ is generated by $H$ and $H_1$ it follows from standard facts about algebraic groups that there is $\iota$ such that the product map from $\mathcal{Y}_\iota$ to $G$ is dominant. This plus (33) and the $H$-invariance of $\mu_1$ finish the proof.

Thanks to this claim we may now apply Proposition 4.1 to the measure $\mu_1$. Recall that $d$ may be taken to be 20. We have that there is an absolute constant $c$ such that

$$|\mu_1(f) - m(f)| \leq cM^{-\delta}\mathcal{S}_d(f), \tag{34}$$

where we may take $\delta = \frac{1}{84} \times 10^{-5}$.

## 5. *Effective discreteness of Markov spectrum*

Our main theorem has an application in giving effective estimates for the discreteness of the Markov spectrum. Let us first recall the definition and some known facts; see [**Mar91**] for a more detailed discussion.

Let $\mathcal{Q}_n$ denote the set of all non-degenerate indefinite quadratic forms in $n$ variables. For any $Q \in \mathcal{Q}_n$ let $m(Q) = \inf\{|Q(v)| : v \in \mathbb{Z}^n \setminus \{0\}\}$ and let $d(Q)$ denote the discriminant of $Q$. Let $\mu(Q) = m(Q)^n/d(Q)$. It is clear that $\mu(Q) = \mu(Q')$ if $Q$ and $Q'$ are equivalent. Let $M_n = \mu(\mathcal{Q}_n)$. This is called the $n$-dimensional Markov spectrum. It follows from Mahler's compactness criteria that $M_n$ is bounded and closed.

https://doi.org/10.1017/S0143385710000799 Published online by Cambridge University Press

In 1880, Markov [**Mark**] described $M_2 \cap (4/9, \infty)$ and the corresponding quadratic forms. It follows from this description that $M_2 \cap (4/9, \infty)$ is a discrete subset of $(4/9, 4/5]$ and, for any $a > 4/9$, $(a, \infty) \cap M_2$ is a finite set. On the other hand, $M_2 \cap (0, 4/9]$ is uncountable and has a quite complicated topological structure.

It follows from Meyer's theorem and Margulis' proof of the Oppenheim conjecture [**Mar86**] that $M_n = \{0\}$ if $n \geq 5$. It follows from [**CaS55, Mar86**] that for $n = 3, 4$ and any $\varepsilon > 0$ we have that $M_n \cap (\varepsilon, \infty)$ is a finite set.

We will give an effective statement here. Let us mention before proceeding to the relevant statement that we have been informed that, using techniques similar to the work of Cogdell, Piateski-Shapiro and Sarnak, reported on in [**Co01**], one can give an alternative quantitative treatment of the discreteness of the Markov spectrum. It is quite likely that the bound one gets using this method could be better than the bound we obtain here.

Let $V$ and $\varepsilon > 0$ be given positive numbers, with the understanding that $\varepsilon$ is small and $V$ is large. Let

$$\mathcal{A}(V, \varepsilon) = \{Hx : Hx \text{ is closed, } V \leq \mathrm{vol}(Hx) \leq 2V \text{ and } |B(g_x\mathbb{Z}^3)|^3 > \varepsilon\}. \tag{35}$$

Note that $\mathcal{A}(V, \varepsilon)$ is a finite set. We let $N(V, \varepsilon) = |\mathcal{A}(V, \varepsilon)|$ be the cardinality of this set, and let $M = N(V, \varepsilon)V$. Indeed $M \leq \mathrm{vol}(\mathcal{A}(V, \varepsilon)) \leq 2M$.

COROLLARY 5.1. *With the notation as above we have*

$$\mathrm{vol}(\mathcal{A}(V, \varepsilon)) \leq C\varepsilon^{-\eta},$$

*where $C$ is an absolute constant. Furthermore we can let $\eta = 784 \times 10^5$.*

*Proof.* Let $y = \mathbb{Z}^3 \in X$. Let $f_y$ be a smooth bump function supported in the $\varepsilon^{1/3}$ neighborhood of $y$ such that $0 \leq f_y \leq 1$ and $f_y = 1$ in the ball of radius $\varepsilon^{1/3}/2$. We may and will choose such $f_y$ so that $\mathcal{S}_{20}(f_y) \leq \varepsilon^{-20/3}$ and $\varepsilon^{8/3}/\iota \leq m(f) \leq \iota\varepsilon^{8/3}$ for some absolute constant $\iota > 0$. Now apply Theorem 1.1 to the function $f_y$. We have that, if $M > M_0$, then there exists some $Hx_0 \in \mathcal{A}(V, \varepsilon)$ such that

$$|\mu_0(f) - m(f)| < c\mathcal{S}_{20}(f_y)M^{-\delta}, \tag{36}$$

where $\mu_0$ is the $H$-invariant probability measure on $Hx_0$ and $c$ is a universal constant. The inequality (36) together with the bounds on $\mathcal{S}_{20}(f_y)$ and $m(f_y)$ imply that if $M > \varepsilon^{-\eta}$, with the given $\eta$, then $\mu_0(f_y) > 0$. In particular we have that $Hx_0$ intersects the ball of radius $\varepsilon^{1/3}$ about $y$, i.e. $Hx_0 \notin \mathcal{A}(V, \varepsilon)$, which is a contradiction. □

It is worth mentioning that Corollary 5.1 and a geometric series argument imply that $\#(M_3 \cap (\varepsilon, \infty)) \leq C\varepsilon^{-\eta}$, where $C$ is a (computable) absolute constant and $\eta = 784 \times 10^5$.

*Acknowledgements.* We would like to thank Professor G. A. Margulis for his interest in this work and many enlightening conversations. We are in debt to M. Einsiedler for reading the first draft of this paper and for many helpful comments. We would like to thank E. Lindenstrauss for helpful conversations on the Markov spectrum. We also thank the anonymous referee for many advantageous comments.

https://doi.org/10.1017/S0143385710000799 Published online by Cambridge University Press

## References


**[BR02]** J. Bernstein and A. Reznikov. Sobolev norms of automorphic functionals. *Int. Math. Res. Not.* (40) (2002), 2155–2174.

**[CaS55]** J. W. Cassels and H. P. Swinnerton-Dyer. On the product of three homogeneous linear forms and indefinite ternary quadratic forms. *Philos. Trans. R. Soc. Lond.* **248**(940) (1955), 73–96.

**[Co01]** J. W. Cogdell. On sums of three squares. *Les XXIIémes Journées Arithmetiques (Lille, 2001). J. Théor. Nombres Bordeaux* **15**(1) (2003), 33–44.

**[DM89]** S. G. Dani and G. A. Margulis. Values of quadratic forms at primitive integral points. *Invent. Math.* **98**(2) (1989), 405–424.

**[E06]** M. Einsiedler. Ratner's theorem on $\mathrm{SL}_2(\mathbb{R})$-invariant measures. *Jahresber. Deutsch. Math.-Verein.* **108**(3) (2006), 143–164.

**[EMV09]** M. Einsiedler, G. A. Margulis and A. Venkatesh. Effective equidistribution for closed orbits of semisimple groups on homogeneous spaces. *Invent. Math.* **177**(1) (2009), 137–212.

**[KS03]** H. H. Kim and P. Sarnak. Refined estimates towards the Ramanujan and Selberg conjectures, Appendix to H. Kim, Functoriality for the exterior square of $\mathrm{GL}_4$ and the symmetric fourth of $\mathrm{GL}_2$. *J. Amer. Math. Soc.* **16**(1) (2003), 139–183.

**[KM98]** D. Kleinbock and G. A. Margulis. Flows on homogeneous spaces and Diophantine approximation on manifolds. *Ann. of Math.* (2) **148** (1998), 339–360.

**[Mar86]** G. A. Margulis. Indefinite quadratic forms and unipotent flows on homogeneous spaces. *Proc. Semester on Dynamical Systems and Ergodic Theory (Warsaw 1986) (Banach Center Publications, 23)*. PWN, Warsaw, 1989, pp. 399–409.

**[Mar90a]** G. A. Margulis. Orbits of group actions and values of quadrtic forms at integral points. *Festschrift in Honour of I. I. Piatetski-Shapiro (Israel Mathematical Conference Proceedings, 3)*. Weizmann Science Press, Israel, Jerusalem, 1990, pp. 127–151.

**[Mar91]** G. A. Margulis. Dynamical and ergodic properties of subgroup actions on homogeneous spaces with applications to number theory. *Proc. Int. Congress of Mathematicians (Kyoto, 1990),* Vol. I, II. Mathematical Society of Japan, Tokyo, 1991, pp. 193–215.

**[MT94]** G. A. Margulis and G. Tomanov. Invariant measures for actions of unipotent groups over local fields on homogeneous spaces. *Invent. Math.* **116**(1–3) (1994), 347–392.

**[Mark]** A. A. Markov. On binary quadratic forms of positive determinant. SPb, 1880. *Usp. Mat. Nauk* (3) **5** (1948), 7–51.

**[MS95]** S. Mozes and N. Shah. On the space of ergodic invariant measures of unipotent flows. *Ergod. Th. & Dynam. Sys.* **15**(1) (1995), 149–159.

**[Oh02]** H. Oh. Uniform pointwise bounds for matrix coefficients of unitary representations and applications to Kazhdan constants. *Duke Math. J.* **113** (2002), 133–192.

**[R90]** M. Ratner. On measure rigidity of unipotent subgroups of semi-simple groups. *Acta Math.* **165** (1990), 229–309.

**[R91]** M. Rather. Raghunathan topological conjecture and distributions of unipotent flows. *Duke Math. J.* **63** (1991), 235–280.

**[R92]** M. Ratner. On Raghunathan's measure conjecture. *Ann. of Math.* (2) **134** (1992), 545–607.

**[R95]** M. Ratner. Raghunathan's conjectures for Cartesian products of real and $p$-adic Lie groups. *Duke Math. J.* **77**(2) (1995), 275–382.


https://doi.org/10.1017/S0143385710000799 Published online by Cambridge University Press