POINTWISE EQUIDISTRIBUTION WITH AN ERROR RATE AND WITH RESPECT TO UNBOUNDED FUNCTIONS

DMITRY KLEINBOCK, RONGGANG SHI, AND BARAK WEISS

Abstract. Consider $G = \mathrm{SL}_d(\mathbb{R})$ and $\Gamma = \mathrm{SL}_d(\mathbb{Z})$. It was recently shown by the second-named author [21] that for some diagonal subgroups $\{g_t\} \subset G$ and unipotent subgroups $U \subset G$, g_t-trajectories of almost all points on all U-orbits on G/Γ are equidistributed with respect to continuous compactly supported functions φ on G/Γ. In this paper we strengthen this result in two directions: by exhibiting an error rate of equidistribution when φ is smooth and compactly supported, and by proving equidistribution with respect to certain unbounded functions, namely Siegel transforms of Riemann integrable functions on $\mathbb{R}^d$. For the first part we use a method based on effective double equidistribution of g_t-translates of U-orbits, which generalizes the main result of [13]. The second part is based on Schmidt's results on counting of lattice points. Number-theoretic consequences involving spiraling of lattice approximations, extending recent work of Athreya, Ghosh and Tseng [1], are derived using the equidistribution result.

1. Introduction

Fix an integer $d \geq 2$, let $G = \mathrm{SL}_d(\mathbb{R})$ and $\Gamma = \mathrm{SL}_d(\mathbb{Z})$, and denote by X the homogeneous space G/Γ, which can be identified with the space of unimodular lattices in $\mathbb{R}^d$. The group G acts on X by left translations preserving the probability measure μ induced by Haar measure. This action has been intensively studied due to its intrinsic interest as a dynamical system and for its number-theoretic applications. See [14, Chapter 5] for a survey of this topic.

Let $D = \{g_t\}$ be a one-parameter subgroup of G, and let φ be a real-valued function on X. We say that $\Lambda \in X$ is (D^+, φ)-*generic* if

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \varphi(g_t\Lambda)\, \mathrm{d}t = \int_X \varphi \, \mathrm{d}\mu\,, \tag{1.1}$$

2000 *Mathematics Subject Classification.* Primary 28A33; Secondary 37C85, 22E40.

Key words and phrases. homogeneous dynamics, equidistribution, ergodic theorem, Diophantine approximation.

and, for a collection $\mathcal{S}$ of functions, that Λ is $(D^+,\mathcal{S})$-*generic* if it is (D^+,φ)-generic for every $\varphi \in \mathcal{S}$. Let $C_c(X)$ denote the space of continuous compactly supported functions on X. It is a well-known consequence of Moore's ergodicity theorem and Birkhoff's ergodic theorem that for any unbounded subgroup D of G, μ-almost every $\Lambda \in X$ is $\big(D^+, C_c(X)\big)$-generic. Also, using effective mixing estimates for the D-action on X one can conclude that whenever φ belongs to the space $C_c^\infty(X)$ of smooth compactly supported functions on X, the convergence in (1.1) takes place with a certain rate for μ-a.e. $\Lambda \in X$. More precisely, see [10, Theorem 16] or [9, Theorem 4(vii)], for any $\varepsilon > 0$ and μ-a.e. $\Lambda \in X$ one has

$$\frac{1}{T}\int_0^T \varphi\big(g_t\Lambda\big)\,\mathrm{d}t = \int_X \varphi\,\mathrm{d}\mu + o(T^{-1/2}\log^{\frac{3}{2}+\varepsilon} T). \tag{1.2}$$

(The notation $f(T) = o\big(g(T)\big)$ means $\lim_{T\to\infty}\frac{|f(T)|}{g(T)} = 0$.)

Now let U^+ be the *unstable horospherical subgroup of G relative to D^+*, defined by

$$U^+ := \{g \in G : g_{-t}gg_t \to e \text{ as } t \to \infty\}. \tag{1.3}$$

Assume in addition that D is diagonalizable. Then, using a local decomposition of G as the product of its unstable, neutral and stable horospherical subgroups with respect to D^+, see e.g. [12, §1.3], it is easy to conclude that for any $\Lambda \in X$ the element $g\Lambda$ is $\big(D^+, C_c(X)\big)$-generic for Haar-a.e. $g \in U^+$.

Recently in [21], as a corollary of a more general result, the second-named author obtained a similar conclusion for some proper subgroups of U^+ – namely, for the class of so-called D^+-*expanding* subgroups, introduced in [15] (see also [20] for some related ideas). More precisely, the following is a special case of [21, Theorem 1.2]: if D is a diagonalizable one-parameter subgroup of G and U is a connected D^+-expanding abelian subgroup of U^+, then

$$\forall\,\Lambda \in X,\ g\Lambda \text{ is } \big(D^+, C_c(X)\big)\text{-generic for Haar-a.e. } g \in U. \tag{1.4}$$

In the present paper we consider a specific family of pairs (D,U), where $D \subset G$ is one-parameter and $U \subset G$ is D^+-expanding, which are important for number-theoretic applications. Namely, take $m,n \in \mathbb{N}$ with $m+n = d$, denote by M the space of $m\times n$ real matrices, and consider

$$U := u(M) \text{ where } u : M \to G \text{ is given by } u(\vartheta) := \begin{pmatrix} 1_m & \vartheta \\ 0 & 1_n \end{pmatrix} \tag{1.5}$$

(here and hereafter 1_k stands for the identity matrix of order k). Also fix two 'weight vectors'

$$\mathbf{a} = (a_1, \dots, a_m) \in \mathbb{R}^m_{>0} \quad \text{and} \quad \mathbf{b} = (b_1, \dots, b_n) \in \mathbb{R}^n_{>0}$$
$$\text{such that } \sum_{i=1}^m a_i = \sum_{i=1}^n b_i = 1.$$

We will refer to the case

$$\mathbf{a} = \mathbf{m} = (1/m, \dots, 1/m), \ \mathbf{b} = \mathbf{n} = (1/n, \dots, 1/n)$$

as the case of *equal weights.* Then take

$$D = \{g_t\}, \quad \text{where } g_t = \operatorname{diag}(e^{a_1 t}, \dots, e^{a_m t}, e^{-b_1 t}, \dots, e^{-b_n t}). \tag{1.6}$$

It is easy to see that for any choice of weight vectors $\mathbf{a}, \mathbf{b}$, the group U as in (1.5) is contained in the unstable horospherical subgroup relative to D^+ (and coincides with it in the case of equal weights). It was shown in [15] that U is D^+-expanding for any D as in (1.6), which, in view of [21], implies (1.4). For the rest of the paper we will fix arbitrary weight vectors $\mathbf{a}, \mathbf{b}$ and choose D as in (1.6).

Our first main result gives an analogue of (1.2) for almost all points on U-orbits:

Theorem 1.1. *Let $\Lambda \in X$, $\varphi \in C_c^\infty(X)$ and $\varepsilon > 0$ be given. Then for almost every $\vartheta \in M$*

$$\frac{1}{T} \int_0^T \varphi\big(g_t u(\vartheta)\Lambda\big)\, \mathrm{d}t = \int_X \varphi \, \mathrm{d}\mu + o(T^{-1/2} \log^{\frac{3}{2}+\varepsilon} T). \tag{1.7}$$

Here and hereafter 'almost every ϑ' means almost every with respect to Lebesgue measure on $M \cong \mathbb{R}^{mn}$; this corresponds to taking a Haar measure on U.

Our proof of Theorem 1.1 is completely independent of [21], and, in view of the density of $C_c^\infty(X)$ in $C_c(X)$, provides an alternative demonstration (1.4) for the case (1.5)–(1.6). It is based on the following 'effective double equidistribution' of g_t-translates of U, generalizing the 'effective equidistribution' result of [13]:

Theorem 1.2. *There exists $\delta > 0$ such that the following holds. Given $f \in C_c^\infty(M)$ and $\varphi, \psi \in C_c^\infty(X)$ and a compact subset L of X, there exists $C > 0$ such that for any $\Delta, \Lambda \in L$ and $t, w \geq 0$ one has*

$$\left| I_{\Delta,\Lambda,f,\varphi,\psi}(t,w) - \int_M f \, \mathrm{d}\vartheta \int_X \varphi \, \mathrm{d}\mu \int_X \psi \, \mathrm{d}\mu \right| \leq C e^{-\delta \min(t,w,|w-t|)}, \tag{1.8}$$

where

$$I_{\Delta,\Lambda,f,\varphi,\psi}(t,w) := \int_M f(\vartheta)\varphi\big(g_t u(\vartheta)\Delta\big)\psi\big(g_w u(\vartheta)\Lambda\big)\,\mathrm{d}\vartheta. \tag{1.9}$$

Under the same assumptions one can also prove 'effective k-fold equidistribution' of g_t-translates of U for all $k \in \mathbb{N}$. We hope to return to this topic elsewhere. See also [5] and [16] for related results on 'effective k-mixing'. To derive pointwise equidistribution of g_t-trajectories of points on U-orbits with an error rate from effective double equidistribution of g_t-translates of U we use a method borrowed from Schmidt's work [18] and originally due to Cassels [6]. It is also similar to the argument used in [9] to derive the rate of convergence in Birkhoff's Theorem from the rate of decay of correlations. As a byproduct of the method, we show in §3 (Remark 3.6) how the estimate (1.2) for μ-a.e. Λ can be derived from the exponential mixing of the D-action on X.

The second main theme of the present paper is establishing (D^+,φ)-genericity with respect to some unbounded functions φ. It follows from Birkhoff's Theorem that for any $\varphi \in L^1(X,\mu)$, μ-almost every $\Lambda \in X$ is (D^+,φ)-generic; however a passage from μ-a.e. $\Lambda \in X$ to Haar-almost all points on U^+-orbits requires some regularity of φ, such as continuity and Lipschitz property away from a compact subset. When U^+ is replaced by its proper subgroup, the situation becomes more complicated. In particular, our method of proof of Theorem 1.1, as well as the argument in [21], do not work for unbounded functions. Yet we are able to prove a partial result for the setup (1.5)–(1.6) and explore its number-theoretic consequences.

In order to control the rate of growth at infinity of unbounded functions on X, following [8], let us introduce a function measuring 'penetration into the thin part of X'. For a lattice $\Lambda \in X$, given a subgroup $\Lambda' \subset \Lambda$, let $d(\Lambda')$ denote the covolume of Λ' in $\mathrm{span}_{\mathbb{R}}(\Lambda')$ (measured with respect to the standard Euclidean structure on $\mathbb{R}^d$). We denote

$$\alpha(\Lambda) = \max\left\{d(\Lambda')^{-1} : \Lambda' \text{ a subgroup of } \Lambda\right\}.$$

It is well-known that this maximum is attained, and defines a proper map $X \to [1,\infty)$.

Now let us denote by $C_\alpha(X)$ the space of functions φ on X satisfying the following two properties:

(C_α-1) φ is continuous except on a set of μ-measure zero;

(C_α-2) the growth of φ is majorized by α, namely there exists $C > 0$ such that for all $\Lambda \in X$, we have

$$|\varphi(\Lambda)| \le C\alpha(\Lambda)$$

(in particular, φ is bounded on compact sets).

One can show that the space $C_\alpha(X)$ contains certain unbounded functions which often arise in number-theoretic applications. Recall that $f : \mathbb{R}^d \to \mathbb{R}$ is said to be *Riemann integrable* if f is bounded with bounded support and continuous except on a set of Lebesgue measure zero. For such f, we define a function $\widehat{f}$ on X by

$$\widehat{f}(\Lambda) := \sum_{\mathbf{v} \in \Lambda \smallsetminus \{\mathbf{0}\}} f(\mathbf{v}). \tag{1.10}$$

The *Siegel integral formula*, established in [22], asserts that for any f as above, $\int_X \widehat{f}\, \mathrm{d}\mu = \int_{\mathbb{R}^d} f$. We show in Lemma 5.1 that for any Riemann integrable f, the function $\widehat{f}$ satisfies conditions (C_α-1) and (C_α-2). In fact, up to constants, when f is non-negative and nonzero on a set of positive measure, the order of growth of $\widehat{f}$ is precisely the same as that of α. As a consequence, in view of the Siegel integral formula, $C_\alpha(X)$ is contained in $L^1(X, \mu)$.

For our second main result, we denote by Λ_0 the standard lattice $\mathbb{Z}^d \subset \mathbb{R}^d$.

Theorem 1.3. *Let U and D be as in* (1.5)–(1.6). *Then for almost every $\vartheta \in M$, $u(\vartheta)\Lambda_0$ is $\big(D^+, C_\alpha(X)\big)$-generic.*

Since $\big(D^+, C_c(X)\big)$-genericity has already been proved, the main additional point in the proof of Theorem 1.3 is to obtain upper bounds for Birkhoff averages of restrictions of non-negative $\varphi \in C_\alpha(X)$ to complements of large compact subsets of X. To this end we employ a lattice point counting result of Schmidt [18]. Lattice point counting was used for a similar purpose in [17], and the connection between Schmidt's result and the action of D^+ was already noted in [3] in the equal weights case. Note that in our result we assume that the lattices are of the form $u(\vartheta)\Lambda_0$; we expect a similar result to be true if Λ_0 is replaced by any other lattice in X. However our proof does not yield this more general statement.

Let us now explain the number-theoretic implications of Theorem 1.3. In Diophantine approximation one interprets $\vartheta \in M$ as a system of m linear forms in n variables and studies how close to integers are the values $\vartheta\mathbf{q}$ of those forms at integer vectors $\mathbf{q}$. During recent years there have been many developments in *Diophantine approximation with weights*, which allows to treat individual components of $\mathbf{q}$ and $\vartheta\mathbf{q}$ differently. This is done by choosing weight vectors $\mathbf{a}, \mathbf{b}$ and considering 'weighted quasi-norms' introduced in [11]:

$$\|\mathbf{x}\|_{\mathbf{a}} = \max_{1\le i\le m} |x_i|^{\frac{1}{a_i}} \quad \left(\text{resp. } \|\mathbf{y}\|_{\mathbf{b}} = \max_{j\le i\le n} |y_j|^{\frac{1}{b_j}}\right).$$

It is a consequence of the general version of the Khintchine-Groshev Theorem proved by Schmidt [18] that for a.e. $\vartheta \in M$ and any $c > 0$ there are infinitely many solutions $(\mathbf{p},\mathbf{q}) \in \mathbb{Z}^m \times \mathbb{Z}^n$ to the inequality

$$\|\vartheta\mathbf{q} - \mathbf{p}\|_{\mathbf{a}} < \frac{c}{\|\mathbf{q}\|_{\mathbf{b}}}. \tag{1.11}$$

More precisely, the number of integer solutions of (1.11) with $\mathbf{q}$ satisfying

$$1 \le \|\mathbf{q}\|_{\mathbf{b}} < e^T \tag{1.12}$$

has the same asymptotic growth as $2^d cT$ (here and hereafter we say that $f(T)$ and $g(T)$ *have the same asymptotic growth*, denoted $f(T) \sim g(T)$, if $\lim\limits_{T\to\infty} f(T)/g(T) = 1$). We remark that the set of nonzero integer solutions of (1.11)–(1.12) is in one-to-one correspondence with the intersection of the lattice

$$u(\vartheta)\Lambda_0 = \left\{ \begin{pmatrix} \vartheta\mathbf{q} - \mathbf{p} \\ \mathbf{q} \end{pmatrix} : \mathbf{p} \in \mathbb{Z}^m,\ \mathbf{q} \in \mathbb{Z}^n \right\}$$

with the set

$$E_{T,c} := \left\{ (\mathbf{x},\mathbf{y}) \in \mathbb{R}^m \times \mathbb{R}^n : \|\mathbf{x}\|_{\mathbf{a}} < \frac{c}{\|\mathbf{y}\|_{\mathbf{b}}},\quad 1 \le \|\mathbf{y}\|_{\mathbf{b}} < e^T \right\},$$

and an elementary computation shows that the volume of this set is equal to $2^d cT$.

A finer question concerning directions of vectors $\vartheta\mathbf{q} - \mathbf{p}$ and $\mathbf{q}$ was addressed in a recent paper [1] by Athreya, Ghosh and Tseng. In the equal weights case they considered integer solutions of (1.11)–(1.12) in addition satisfying

$$\pi(\vartheta\mathbf{q} - \mathbf{p}) \in A \text{ and } \pi(\mathbf{q}) \in B, \tag{1.13}$$

where π stands for the radial projection from $\mathbb{R}^m$ and $\mathbb{R}^n$ to the unit spheres $\mathbb{S}^{m-1}$ and $\mathbb{S}^{n-1}$ respectively, and $A \subset \mathbb{S}^{m-1}$, $B \subset \mathbb{S}^{n-1}$ are two measurable subsets. When the boundaries of A and B are of measure zero, they showed that for almost every $\vartheta \in M$ the number of solutions of (1.11)–(1.13) (with $\mathbf{a} = \mathbf{m}$ and $\mathbf{b} = \mathbf{n}$) has the same asymptotic growth as the volume of the set

$$\left\{ (\mathbf{x},\mathbf{y}) \in \mathbb{R}^m \times \mathbb{R}^n : \begin{array}{c} \|\mathbf{x}\|^m < \dfrac{c}{\|\mathbf{y}\|^n},\quad 1 \le \|\mathbf{y}\|^n < e^T \\ \pi(\mathbf{x}) \in A,\quad \pi(\mathbf{y}) \in B \end{array} \right\}$$

(here $\|\cdot\|$ stands for the supremum norm; note that [1] deals only with the case $m = 1$ and uses Euclidean norm instead of the supremum norm, but the argument can be applied to an arbitrary m and arbitrary norm).

In the case of arbitrary weight vectors it no longer makes sense to project radially, since these projections accumulate near directions corresponding to smallest (resp., largest) weights. We remedy this as follows. Let $F_{\mathbf{a}t}$ be the $\mathbf{a}$-weighted flow on $\mathbb{R}^m$ defined by

$$F_{\mathbf{a}t}(\mathbf{x}) := \left(e^{a_1 t}x_1, \ldots, e^{a_m t}x_m\right)$$

(note that in the case of equal weights, these are just homotheties of $\mathbb{R}^m$). Similarly we define the transformation $F_{\mathbf{b}t}$ of $\mathbb{R}^n$. Then for nonzero $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^n$ let

(1.14)

$$\pi_{\mathbf{a}}(\mathbf{x}) := \{F_{\mathbf{a}t}(\mathbf{x}) : t \in \mathbb{R}\} \cap \mathbb{S}^{m-1}, \ \pi_{\mathbf{b}}(\mathbf{y}) := \{F_{\mathbf{b}t}(\mathbf{y}) : t \in \mathbb{R}\} \cap \mathbb{S}^{n-1}$$

(these intersection point are clearly unique), replace (1.13) with

$$\pi_{\mathbf{a}}(\vartheta\mathbf{q} - \mathbf{p}) \in A \text{ and } \pi_{\mathbf{b}}(\mathbf{q}) \in B, \tag{1.15}$$

and define

$$E_{T,c}(A,B) := \left\{ (\mathbf{x},\mathbf{y}) \in \mathbb{R}^m \times \mathbb{R}^n : \begin{array}{c} \|\mathbf{x}\|_{\mathbf{a}} < \dfrac{c}{\|\mathbf{y}\|_{\mathbf{b}}}, \quad 1 \le \|\mathbf{y}\|_{\mathbf{b}} < e^T \\ \pi_{\mathbf{a}}(\mathbf{x}) \in A, \quad \pi_{\mathbf{b}}(\mathbf{y}) \in B \end{array} \right\}.$$

Using Theorem 1.3, we obtain:

Theorem 1.4. *Let $c > 0$ and measurable subsets $A \subset \mathbb{S}^{m-1}$, $B \subset \mathbb{S}^{n-1}$ with boundaries of measure zero be given. Then for a.e. $\vartheta \in M$, as $T \to \infty$, the number of integer solutions to* (1.11), (1.12) *and* (1.15) *has the same asymptotic growth as the volume of $E_{T,c}(A,B)$.*

Note that the above counting result does not follow from the techniques of [18, 19]. See also [2] for some related results. The reduction of Theorem 1.4 to Theorem 1.3 is based on the observation that the number of integer solutions to (1.11), (1.12) and (1.15) is equal to $\sharp\big(u(\vartheta)\Lambda_0 \cap E_{T,c}(A,B)\big)$, that is, to $\widehat{f}$, where f is a certain Riemann integrable function on $\mathbb{R}^d$.

The structure of the paper is as follows: in §2 we prove Theorem 1.2, and use it in §3 to prove Theorem 1.1. §4 contains a discussion of Schmidt's asymptotic formula (Theorem 4.1) and some auxiliary results, which are needed in the final section of the paper, where Theorems 1.3 and 1.4 are proved.

Acknowledgements. The support of grants NSFC (11201388), NSFC (11271278), ERC starter grant DLGAPS 279893, NSF DMS-1101320 and BSF 2010428 is gratefully acknowledged. We would like to thank Manfred Einsiedler for a discussion of effective double equidistribution and MSRI for its hospitality during Spring 2015.

2. Effective double equidistribution

In this section we prove Theorem 1.2. We first recall several facts from the paper [13]. Its main result, i.e. effective equidistribution of g_t-translates of U, can be stated as follows:

Proposition 2.1 ([13], Theorem 1.3)**.** *There exists $\delta_1 > 0$ such that for any $f \in C_c^\infty(M)$, $\varphi \in C_c^\infty(X)$, and for any compact $L \subset X$ there exists $C_1 = C_1(f, \varphi, L)$ such that for any $\Lambda \in L$ and $t \geq 0$,*

$$\left| \int_M f(\vartheta)\varphi\big(g_t u(\vartheta)\Lambda\big)\,\mathrm{d}\vartheta - \int_M f\,\mathrm{d}\vartheta \int_X \varphi\,\mathrm{d}\mu \right| \leq C_1 e^{-\delta_1 t}. \tag{2.1}$$

Let

$$b = \min\left\{ \frac{a_i}{n}, \frac{b_j}{m} : 1 \leq i \leq m,\ 1 \leq j \leq n \right\},$$

and let

$$g'_t = \begin{pmatrix} e^{nbt}1_m & 0 \\ 0 & e^{-mbt}1_n \end{pmatrix}, \quad g''_t = g_t g'_{-t}. \tag{2.2}$$

Note that g'_t (after reparametrization) corresponds to the case of equal weights; in particular, U is the unstable horospherical subgroup relative to $\{g'_t : t \geq 0\}$. On the other hand g''_t is of the form

$$\mathrm{diag}\left(e^{a'_1 t}, \ldots, e^{a'_m t}, e^{-b'_1 t}, \ldots, e^{-b'_n t}\right)$$

where a'_i and b'_j are nonegative.

Let B_r^G be the ball of radius r centered at the identity element of G with respect to the metric induced by some right invariant Riemannian metric. For $\Lambda \in X$ we let $\pi_\Lambda : G \to X$ be the map $g \mapsto g\Lambda$, and define

$$\mathrm{Inj}_\varepsilon := \left\{\Lambda \in X : \pi_\Lambda|_{B_\varepsilon^G} \text{ is injective}\right\}.$$

Also let B_r denote the ball of radius r centered at zero with respect to the Euclidean norm on $M \cong \mathbb{R}^{mn}$. The following quantitative nondivergence result is a combination of [13, Corollary 3.4] and [13, Proposition 3.5]:

Proposition 2.2. *Let $L \subset X$ be compact and let $r > 0$. There exists $t_0 = t_0(L, r) > 0$ and $C_3 > 0$ such that for every $0 < \varepsilon < 1$, $\Lambda \in L$, $s \geq 0$ and $t \geq t_0$ one has*

$$|\{\vartheta \in B_r : g''_s g_t u(\vartheta)\Lambda \notin \mathrm{Inj}_\varepsilon\}| \leq C_3 \varepsilon^{1/d^4} |B_r|.$$

Here and hereafter for a measurable subset S of a Euclidean space we use $|S|$ to denote the Lebesgue measure of S. For $h \in C_c^\infty(M)$ and a nonnegative integer k we define the k-th Sobolev norm of h to be

$$\|h\|_k = \max_{|\beta| \le k} \left\| \partial^\beta h \right\|_{L^2(M)}$$

where $|\beta|$ is the order of the multi-index β.

We will also need the following effective equidistribution estimate for the g'_t-action:

Proposition 2.3 ([13], Theorem 2.3)**.** *There exist $\delta_2, r_0 > 0$ and $k \in \mathbb{N}$ with the following property: for any $\psi \in C_c^\infty(X)$ there exists $C_2 > 0$ such that for any $0 < r < r_0$, any $h \in C_c^\infty(M)$ with* $\operatorname{supp}(h) \subset B_r$*, any $\Lambda \in \mathrm{Inj}_{2r}$ and any $t \ge 0$ one has*

$$\left| \int_M h(\vartheta) \psi\big(g'_t u(\vartheta)\Lambda\big) \, \mathrm{d}\vartheta - \int_M h \int_X \psi \right| \le C_2 \left(r^{-k} \|h\|_k e^{-\delta_2 t} + r \int_M |h| \right).$$

Note that the statement is not precisely the one which appears in [13] but can be deduced from it by taking k large enough.

Proof of Theorem 1.2. Recall that we are given a compact subset L of X, $\varphi, \psi \in C_c^\infty(X)$ and $f \in C_c^\infty(M)$. In this proof, the notation $y \ll x$ will mean that $y \le Cx$ where C is a constant independent of x, but which could depend on f, φ, ψ or L.

Let $\delta_0 = \min(\delta_1, \delta_2)$ where δ_1 and δ_2 are given in Propositions 2.1 and 2.3 respectively. We also fix $k \in \mathbb{N}$ so that Proposition 2.3 holds. Without loss of generality we assume that $k \ge mn$. Now let

$$a := \min\{a_i + b_j : 1 \le i \le m, 1 \le j \le n\}$$

and

$$\delta := \min\left(\frac{\delta_0}{2(1+3k)}, \frac{a}{2} \right). \tag{2.3}$$

Let Δ, Λ, t and w be as in the statement of Theorem 1.2. Put

$$r := e^{-\delta|w-t|} \quad \text{and } s := \frac{t+w}{2}.$$

We can assume with no loss of generality that $w \ge t \ge t_0$, where t_0 is as in Proposition 2.2, and moreover that $r < r_0$, where r_0 is as in Proposition 2.3. According to [13, Lemma 2.2] there exists $h \in C_c^\infty(M)$ such that $\operatorname{supp}(h) \subset B_r$, $h \ge 0$, $\int_M h(\vartheta) \, \mathrm{d}\vartheta = 1$ and $\|h\|_k \ll r^{-2k}$. Given $\zeta \in M$, we define $\zeta_1, \zeta_2 \in M$ by the formulae

$$u(\zeta_1) := g''_{t-s} g_{-s} u(\zeta) g_s g''_{s-t}, \quad u(\zeta_2) := g''_{t-s} g_{t-s} u(\zeta) g_{s-t} g''_{s-t},$$

which imply

$$g_t u(\zeta_1) = u(\zeta_2) g_t \quad \text{and} \quad g_w u(\zeta_1) = g'_{s-t} u(\zeta) g_s g''_{s-t}. \tag{2.4}$$

Let $I_{\Delta,\Lambda,f,\varphi,\psi}(t,w)$ be as in (1.9). Then, using Fubini's theorem and, for each ζ, making a change of variables $\vartheta \mapsto \vartheta + \zeta_1$, we find in view of (2.4) that

(2.5)

$$\begin{aligned} I_{\Delta,\Lambda,f,\varphi,\psi}(t,w) &= \int_M f(\vartheta)\varphi\big(g_t u(\vartheta)\Delta\big)\psi\big(g_w u(\vartheta)\Lambda\big)\,\mathrm{d}\vartheta \int_M h(\zeta)\,\mathrm{d}\zeta \\ =&\int_M \int_M f(\vartheta+\zeta_1)\varphi\big(u(\zeta_2)g_t u(\vartheta)\Delta\big)\psi\big(g'_{s-t}u(\zeta)g''_{s-t}g_s u(\vartheta)\Lambda\big)h(\zeta)\,\mathrm{d}\vartheta\,\mathrm{d}\zeta. \end{aligned}$$

Note that for all $\zeta \in \mathrm{supp}(h)$ we have

$$\|\zeta_1\| \ll e^{-sa}\|\zeta\| \ll e^{-\frac{w-t}{2}a} \overset{(2.3)}{\ll} e^{-\delta(w-t)},$$

and similarly $\|\zeta_2\| \ll e^{-\delta(w-t)}$. Denote

$$\Psi(\vartheta) := \int_M \psi\left(g'_{s-t}u(\zeta)g''_{s-t}g_s u(\vartheta)\Lambda\right) h(\zeta)\,\mathrm{d}\zeta.$$

Then, by approximating the function $f(\vartheta+\zeta_1)\varphi\big(u(\zeta_2)g_t u(\vartheta)\Delta\big)$ by $f(\vartheta)\varphi\big(g_t u(\vartheta)\Delta\big)$ in (2.5) we get

$$\left| I_{\Delta,\Lambda,f,\varphi,\psi}(t,w) - \int_M \Psi(\vartheta) f(\vartheta)\varphi\big(g_t u(\vartheta)\Delta\big)\,\mathrm{d}\vartheta \right| \ll e^{-\delta(w-t)}. \tag{2.6}$$

Let $r_1 > 0$ be such that $\mathrm{supp}\, f \subset B_{r_1}$, let $\varepsilon = e^{-\delta(w-t)d^4}$, and denote

$$E := \{\vartheta \in B_{r_1} : g''_{s-t} g_s u(\vartheta)\Lambda \in \mathrm{Inj}_\varepsilon\}.$$

It follows from Proposition 2.2 that

$$|B_{r_1} \smallsetminus E| \ll e^{-\delta(w-t)}|B_{r_1}|,$$

and hence

$$\left| \int_{B_{r_1}\smallsetminus E} \Psi(\vartheta) f(\vartheta)\varphi\big(g_t u(\vartheta)\Delta\big)\,\mathrm{d}\vartheta \right| \ll e^{-\delta(w-t)}. \tag{2.7}$$

By Proposition 2.3 for $\vartheta \in E$ (with $g''_{s-t}g_s u(\vartheta)\Lambda$ in place of Λ) one has

$$\left| \Psi(\vartheta) - \int_X \psi \right| \ll r + r^{-k}\|h\|_k e^{-\delta_2(s-t)} \ll e^{-\delta(w-t)} + e^{-(\delta_0/2-3k\delta)(w-t)},$$

which, in view of (2.3) and (2.7), implies

$$\begin{aligned}\left|\int_M \Psi(\vartheta) f(\vartheta)\varphi\big(g_t u(\vartheta)\Delta\big)\,\mathrm{d}\vartheta - \int_M f(\vartheta)\varphi\big(g_t u(\vartheta)\Delta\big)\,\mathrm{d}\vartheta \int_X \psi\right| \\ \ll e^{-\delta(w-t)} + e^{-\frac{\delta_0(w-t)}{2-3k\delta}} \ll e^{-\delta(w-t)}.\end{aligned} \tag{2.8}$$

On the other hand from Proposition 2.1 one gets

$$\left|\int_M f(\vartheta)\varphi\big(g_t u(\vartheta)\Delta\big)\,\mathrm{d}\vartheta - \int_M f \int_X \varphi\right| \ll e^{-\delta_0 t}. \tag{2.9}$$

Combining (2.6), (2.8) and (2.9), one arrives at (1.8). □

3. Pointwise equidistribution with an error rate

In this section we prove Theorem 1.1. The method works in a general framework as follows:

Theorem 3.1. *Let (Y,ν) be a probability space, and let $F : Y\times\mathbb{R}_+ \to \mathbb{R}$ be a bounded measurable function. Suppose there exist $\delta > 0$ and $C > 0$ such that for any $w \geq t \geq 0$,*

$$\left|\int_Y F(x,t)F(x,w)\,\mathrm{d}\nu(x)\right| \leq Ce^{-\delta\min(t,w-t)}. \tag{3.1}$$

Then given $\varepsilon > 0$ we have

$$\frac{1}{T}\int_0^T F(y,t)\,\mathrm{d}t = o(T^{-1/2}\log^{\frac{3}{2}+\varepsilon} T) \tag{3.2}$$

for ν-almost every $y \in Y$.

We begin with some lemmas. In the statements below the notation and assumptions are as in Theorem 3.1.

Lemma 3.2. *Let $[b,c]$ be a closed interval in $[0,\infty)$. Then*

$$\int_Y \left(\int_b^c F(x,t)\,\mathrm{d}t\right)^2 \mathrm{d}\nu(x) \leq 4C\delta^{-1}(c-b).$$

Proof. The left hand side can be written as

$$\begin{aligned}
&\int_Y \left(\int_b^c F(x,t)\,\mathrm{d}t\right)^2 \mathrm{d}\nu(x) \\
=&\int_b^c \int_b^c \left[\int_Y F(x,t)F(x,w)\,\mathrm{d}\nu(x)\right] \mathrm{d}w\,\mathrm{d}t && \text{by Fubini} \\
\le& 2\int_b^c \int_t^c \left|\int_Y F(x,t)F(x,w)\,\mathrm{d}\nu(x)\right| \mathrm{d}w\,\mathrm{d}t \\
\le& 2C\int_b^c \left[\int_t^c (e^{-\delta(w-t)} + e^{-\delta t})\,\mathrm{d}w\right] \mathrm{d}t && \text{by (3.1)} \\
\le& 2C\int_b^c \left(\delta^{-1} + e^{-\delta t}(c-t)\right) \mathrm{d}t \\
\le& 2C\int_b^c \left(\delta^{-1} + e^{-\delta t}(c-b)\right) \mathrm{d}t \ \le\ 4C\delta^{-1}(c-b),
\end{aligned}$$

and the proof is finished. □

For a positive integer s we let L_s be the set of intervals of the form $[2^i j, 2^i(j+1)]$ where i, j are nonnegative integers and $2^i(j+1) < 2^s$.

Lemma 3.3. *One has*

$$\sum_{[b,c]\in L_s} \int_Y \left(\int_b^c F(x,t)\,\mathrm{d}t\right)^2 \mathrm{d}\nu(x) \le 4C\delta^{-1}s2^s. \tag{3.3}$$

Proof. We estimate the left hand side as

$$\begin{aligned}
&\sum_{[b,c]\in L_s} \int_Y \left(\int_b^c F(x,t)\,\mathrm{d}t\right)^2 \mathrm{d}\nu(x) \\
\le& \sum_{i=0}^{s-1} \sum_{j=0}^{2^{s-i}-1} \int_Y \left(\int_{2^i j}^{2^i(j+1)} F(x,t)\,\mathrm{d}t\right)^2 \mathrm{d}\nu(x) \\
\le& \sum_{i=0}^{s-1} \sum_{j=0}^{2^{s-i}-1} 4C\delta^{-1}2^i, && \text{by Lemma 3.2}
\end{aligned}$$

which is clearly bounded from above by $4C\delta^{-1}s2^s$. □

Lemma 3.4. *Let k, s be positive integers with $k < 2^s$. Then the interval $[0, k]$ can be covered by at most s intervals in L_s.*

Proof. These intervals can be easily constructed using binary expansion of k. See also [18, Lemma 1]. □

Lemma 3.5. *For every $\varepsilon > 0$, there exists a sequence of measurable subsets $\{Y_s\}_{s\in\mathbb{N}}$ of Y such that:*

(i) $\nu(Y_s) \leq 4C\delta^{-1}s^{-(1+2\varepsilon)}$.

(ii) *For every positive integer k with $k < 2^s$ and every $y \notin Y_s$ one has*

$$\left|\int_0^k F(y,t)\,\mathrm{d}t\right| \leq 2^{s/2}s^{\frac{3}{2}+\varepsilon}. \tag{3.4}$$

Proof. Let

$$Y_s = \left\{ y \in Y : \sum_{I\in L_s} \left(\int_I F(y,t)\,\mathrm{d}t\right)^2 > 2^s s^{2+2\varepsilon} \right\}.$$

Assertion (i) follows from Lemma 3.3 and Markov's Inequality. By Lemma 3.4 there exists a subset $L(k)$ of L_s with cardinality at most s such that $[0,k] = \bigcup_{I\in L(k)} I$. For $k < 2^s$ and $y \notin Y_s$ we estimate

$$\begin{aligned} &\left(\int_0^k F(y,t)\,\mathrm{d}t\right)^2 \\ &\leq \left(\sum_{I\in L(k)} \int_I F(y,t)\,\mathrm{d}t\right)^2 \\ &\leq s \sum_{I\in L(k)} \left(\int_I F(y,t)\,\mathrm{d}t\right)^2 && \text{by Cauchy's inequality} \\ &\leq s \sum_{I\in L_s} \left(\int_I F(y,t)\,\mathrm{d}t\right)^2 \leq\ 2^s s^{3+2\varepsilon} && \text{since } y \notin Y_s. \end{aligned}$$

Now (3.4) follows by taking square roots. □

Proof of Theorem 3.1. We fix $\varepsilon > 0$ and choose a sequence of measurable subsets $\{Y_s\}_{s\in\mathbb{N}}$ as in Lemma 3.5. Note that

$$\sum_{s=1}^{\infty} \nu(Y_s) \leq \sum_{s=1}^{\infty} 4C\delta^{-1}s^{-(1+2\varepsilon)} < \infty.$$

The Borel-Cantelli lemma implies that there exists a measurable subset $Y(\varepsilon)$ of Y with full measure such that for every $y \in Y(\varepsilon)$ there exists $s_y \in \mathbb{N}$ such that $y \notin Y_s$ whenever $s \geq s_y$.

We will show that for every $y \in Y(\varepsilon)$ one has

$$\frac{1}{T}\left|\int_0^T F(y,t)\,\mathrm{d}t\right| \ll T^{-1/2}\log^{\frac{3}{2}+\varepsilon}(T) \tag{3.5}$$

provided T is large enough, where the implicit constant depends only on F. Given $T > 2$, let $k = \lfloor T \rfloor$ and $s = 1 + \lfloor \log T \rfloor$, so that $2^{s-1} \leq k \leq T < k+1 \leq 2^s$. Suppose $T \geq 2^{s_y - 1}$, then $s \geq s_y$ and hence $y \notin Y_s$. Therefore we have

$$\begin{aligned} \left| \int_0^T F(y,t)\,\mathrm{d}t \right| &\leq \|F\|_\infty + \left| \int_0^k F(y,t)\,\mathrm{d}t \right| \\ &\leq \|F\|_\infty + 2^{s/2} s^{\frac{3}{2}+\varepsilon} && \text{by (3.4)} \\ &\leq \|F\|_\infty + (2T)^{1/2} \log^{\frac{3}{2}+\varepsilon}(2T). \end{aligned}$$

and (3.5) follows. This clearly implies (3.2) for $y \in \cap_{k\in\mathbb{N}} Y(1/k)$. $\square$

Proof of Theorem 1.1. Recall that we are given $\Lambda \in X$, $\varphi \in C_c^\infty(X)$ and $\varepsilon > 0$. Take $f \in C_c^\infty(M)$ with $f \geq 0$ and $\int_M f\,\mathrm{d}\vartheta = 1$. Let ν be the probability measure on X defined by

$$\int_X \psi\,\mathrm{d}\nu = \int_M f(\vartheta)\psi\big(u(\vartheta)\Lambda\big)\,\mathrm{d}\vartheta$$

for every $\psi \in C_c(X)$. Denote $\alpha = \int_X \varphi\,\mathrm{d}\mu$ and for $t, w \geq 0$ write

$$\begin{aligned} \int_X \big(\varphi(g_t\Lambda) - \alpha\big)\big(\varphi(g_w\Lambda) - \alpha\big)\,\mathrm{d}\nu = {} & I_{\Lambda,\Lambda,f,\varphi,\varphi}(t,w) - \int_M f\,\mathrm{d}\vartheta \left(\int_X \varphi\,\mathrm{d}\mu\right)^2 \\ & - \alpha\left(\int_X f(\vartheta)\varphi\big(g_t u(\vartheta)\Lambda\big)\,\mathrm{d}\vartheta - \int_M f\,\mathrm{d}\vartheta \int_X \varphi\,\mathrm{d}\mu\right) \\ & - \alpha\left(\int_X f(\vartheta)\varphi\big(g_w u(\vartheta)\Lambda\big)\,\mathrm{d}\vartheta - \int_M f\,\mathrm{d}\vartheta \int_X \varphi\,\mathrm{d}\mu\right). \end{aligned}$$

Applying Theorem 1.2 and Proposition 2.1, we conclude that there exist $C > 0$ and $\delta > 0$ such that the estimate

$$\left| \int_X \big(\varphi(g_t\Lambda) - \alpha\big)\big(\varphi(g_w\Lambda) - \alpha\big)\,\mathrm{d}\nu \right| \leq Ce^{-\delta \min(t, w-t)}$$

holds for any $w \geq t \geq 0$. Then we can apply Theorem 3.1 with $F(x,t) = \varphi(g_t x) - \alpha$ and obtain (1.7) for almost every $\vartheta \in S_f$, where $S_f := \{\vartheta \in M : f(\vartheta) > 0\}$. Since countably many sets of the form S_f exhaust M, we reach the desired conclusion. $\square$

Remark 3.6. Arguing similarly with ν replaced by μ and using exponential mixing of the g_t-action, see e.g. [13, Theorem 1.1], instead of Theorem 1.2, one can easily obtain (1.2) for almost every $\Lambda \in X$.

4. Lattice points counting

In this section we recall a result of Schmidt [18] concerning a counting problem arising from Diophantine approximation, and relate it to the D-action on X. From this we will deduce some estimates which will be used in the proof of Theorem 1.3.

Let $\mathbf{a}, \mathbf{b}$ be weight vectors, let $\pi_{\mathbf{a}}, \pi_{\mathbf{b}}$ be as in (1.14), and let c, A, B be as in Theorem 1.4. For an individual vector $\mathbf{v} = (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^d$ we write $g_t\mathbf{v} = (\mathbf{x}_t, \mathbf{y}_t)$, and then we have

$$\|\mathbf{y}_t\|_{\mathbf{b}} = e^{-t}\|\mathbf{y}\|_{\mathbf{b}}$$

and

$$\|\mathbf{x}_t\|_{\mathbf{a}} \cdot \|\mathbf{y}_t\|_{\mathbf{b}} = \|\mathbf{x}\|_{\mathbf{a}} \cdot \|\mathbf{y}\|_{\mathbf{b}}, \ \pi_{\mathbf{a}}(\mathbf{x}_t) = \pi_{\mathbf{a}}(\mathbf{x}), \ \pi_{\mathbf{b}}(\mathbf{y}_t) = \pi_{\mathbf{b}}(\mathbf{y})$$

for all t.

For $r > 0$, define

$$f_{A,B,r,c} := \mathbb{1}_{E_{r,c}(A,B)} \tag{4.1}$$

to be the characteristic function of $E_{r,c}(A, B)$. It follows that

$$f_{A,B,r,c}(g_t\mathbf{v}) = \begin{cases} 1 & \begin{array}{c} \text{if } e^t \leq \|\mathbf{y}\|_{\mathbf{b}} < e^{t+r}, \|\mathbf{x}\|_{\mathbf{a}} \cdot \|\mathbf{y}\|_{\mathbf{b}} < c \\ \text{and } \big(\pi_{\mathbf{a}}(\mathbf{x}), \pi_{\mathbf{b}}(\mathbf{y})\big) \in A \times B \end{array} \\ 0 & \text{otherwise.} \end{cases}$$

Therefore

$$\mathbf{v} \in E_{T,c}(A, B) \smallsetminus E_{r,c}(A, B) \implies |\{t \in [0, T] : g_t\mathbf{v} \in E_{r,c}(A, B)\}| = r$$

and

$$g_t\mathbf{v} \in E_{r,c}(A, B) \text{ for some } t \in [0, T] \implies \mathbf{v} \in E_{r+T,c}(A, B).$$

Using (1.10), and changing the order of summation and integration, it follows that for any $\Lambda \in X$ and any $T > r$ we have

$$\begin{aligned} \sharp\Big(\Lambda \cap \big(E_{T,c}(A, B) \smallsetminus E_{r,c}(A, B)\big)\Big) &\leq \frac{1}{r}\int_0^T \widehat{f}_{A,B,r,c}(g_t\Lambda)\,\mathrm{d}t \\ &\leq \sharp\big(E_{r+T,c}(A, B) \cap \Lambda\big). \end{aligned} \tag{4.2}$$

Let $\mathbb{S}_+^{m-1} := \{\mathbf{x} \in \mathbb{S}^{m-1} : x_i \geq 0\}$ and $\mathbb{S}_+^{n-1} := \{\mathbf{y} \in \mathbb{S}^{n-1} : y_j \geq 0\}$. Also for brevity let us denote by Λ_ϑ the lattice $u(\vartheta)\Lambda_0$. Our main tool for proving Theorem 1.3 is the following result of Schmidt.

Theorem 4.1 ([18], Theorems 1 and 2)**.** *For almost every $\vartheta \in M$,*

$$\sharp\left(E_{T,c}(\mathbb{S}_+^{m-1}, \mathbb{S}_+^{n-1}) \cap \Lambda_\vartheta\right) \sim cT \qquad \textit{as } T \to \infty. \tag{4.3}$$

The case $\mathbf{b} = \mathbf{n}$ of this result follows from [18, Thm. 1 and 2] (setting $h = e^T$, $\psi_i(q) = \left(\frac{c}{q}\right)^{a_i}$ in Schmidt's notation), and in fact Schmidt also obtains an error estimate $o\left(T^{1/2}\log^{\tau+\varepsilon}(T)\right)$ where $\tau = 2$ if $n = 1$ and $\tau = 1.5$ otherwise. The case $\mathbf{b} \neq \mathbf{n}$ does not appear in [18], but the argument given there works for arbitrary $\mathbf{b}$.

Corollary 4.2. *Let $r, c > 0$. Then for almost every $\vartheta \in M$,*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \widehat{f}_{\mathbb{S}^{m-1},\mathbb{S}^{n-1},r,c}(g_t\Lambda_\vartheta)\,\mathrm{d}t = |E_{r,c}|. \tag{4.4}$$

Proof. Let $N_{T,c}(\vartheta)$ be the number of solutions $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^m \times \mathbb{Z}^n_{\geq 0}$ of the system

$$\begin{cases} 0 \leq (\vartheta\mathbf{q})_i - p_i < c^{a_i}\|\mathbf{q}\|_{\mathbf{b}}^{-a_i} & (i = 1, \ldots, m) \\ 1 \leq \|\mathbf{q}\|_{\mathbf{b}} < e^T. \end{cases}$$

It follows directly from the definitions that

$$N_{T,c}(\vartheta) = \sharp\left(E_{T,c}(\mathbb{S}^{m-1}_+, \mathbb{S}^{n-1}_+) \cap \Lambda_\vartheta\right). \tag{4.5}$$

Let $\{\mathbf{e}_i : 1 \leq i \leq m\}$ be the standard basis of $\mathbb{R}^m$. For $I \subset \{1, \ldots, m\}$ we let $\zeta_I : \mathbb{R}^m \to \mathbb{R}^m$ be the linear transformation defined by

$$\zeta_I\mathbf{e}_i = \begin{cases} -\mathbf{e}_i & i \in I \\ \mathbf{e}_i & i \notin I. \end{cases}$$

Analogously we define $\eta_J : \mathbb{R}^n \to \mathbb{R}^n$ for $J \subset \{1, \ldots, n\}$. It follows from Theorem 4.1 that for almost every $\vartheta \in M$

$$N_{T,c}(\zeta_I\vartheta\eta_J) \sim cT \qquad \text{as } T \to \infty. \tag{4.6}$$

On the other hand, it is easy to see that $N_{T,c}(\zeta_I\vartheta\eta_J)$ is the number of solutions $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^d$ of the system

$$\begin{aligned} & c^{a_i}\|\mathbf{q}\|_{\mathbf{b}}^{-a_i} < (\vartheta\mathbf{q})_i - p_i \leq 0 && (i \in I) \\ & 0 \leq (\vartheta\mathbf{q})_i - p_i < c^{a_i}\|\mathbf{q}\|_{\mathbf{b}}^{-a_i} && (i \notin I) \\ & 1 \leq \|\mathbf{q}\|_{\mathbf{b}} < e^T && \\ & q_j \leq 0 && (j \in J) \\ & q_j \geq 0 && (j \notin J) \end{aligned}$$

Now let $\widetilde{N}_{T,c}(\vartheta)$ be the number of solutions $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^d$ of the system (1.11)–(1.12). Then for almost every $\vartheta \in M$ (i.e. for those ϑ for which (4.6) holds for all I and $(\vartheta\mathbf{q})_i - p_i$ is not equal to zero for any $\mathbf{p}, \mathbf{q}$ with $\mathbf{q} \neq 0$ and any i), one has

$$\widetilde{N}_{T,c}(\vartheta) \sim 2^d cT \qquad \text{as } T \to \infty. \tag{4.7}$$

As was mentioned in the introduction,

$$\widetilde{N}_{T,c}(\vartheta) = \sharp\,(E_{T,c} \cap \Lambda_\vartheta) \quad \text{and} \quad |E_{T,c}| = 2^d cT. \tag{4.8}$$

It follows from (4.8) and (4.2) that (4.4) holds for those ϑ which satisfy (4.7). □

Corollary 4.3. *For $r, c > 0$ we let*

$$F_{r,c} := \left\{(\mathbf{x},\mathbf{y}) \in \mathbb{R}^d : \|\mathbf{x}\|_{\mathbf{a}}\|\mathbf{y}\|_{\mathbf{b}} < c,\ 1 \le \|\mathbf{x}\|_{\mathbf{a}} < e^r\right\}. \tag{4.9}$$

Then for almost every $\vartheta \in M$

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \widehat{\mathbb{1}}_{F_{r,c}}(g_t\Lambda_\vartheta)\,\mathrm{d}t = |F_{r,c}|.$$

Proof. Let $\widetilde{N}_{T,c}$ be as in Corollary 4.2. For every $T > |\log c|$ and $\vartheta \in M$, similarly to (4.2) one has

$$\frac{1}{r}\int_0^T \widehat{\mathbb{1}}_{F_{r,c}}(g_t\Lambda_\vartheta)\,dt \le \widetilde{N}_{T+\log c}(\vartheta) + \sharp(\widetilde{F} \cap \Lambda_\vartheta)$$

where

$$\widetilde{F} := \{(\mathbf{x},\mathbf{y}) \in \mathbb{R}^d : \|\mathbf{x}\|_{\mathbf{a}} < e^r, \|\mathbf{y}\|_{\mathbf{b}} < c\}.$$

For every ϑ, $\sharp(\widetilde{F} \cap \Lambda_\vartheta)$ is a number independent of T and $|F_{r,c}| = |E_{r,c}|$. So for $\vartheta \in M$ satisfying (4.7) one has

$$\limsup_{T\to\infty} \frac{1}{T}\int_0^T \widehat{\mathbb{1}}_{F_{r,c}}(g_t\Lambda_\vartheta)\,\mathrm{d}t \le |F_{r,c}|.$$

On the other hand by [21, Corollary 1.3], there is a conull subset of $\vartheta \in M$ for which Λ_ϑ is $\big(D^+, C_c(X)\big)$-generic. For any $\varepsilon > 0$, since $\mathbb{1}_{F_{r,c}}$ is a Riemann integrable non-negative function, there is $\varphi_1 \in C_c(\mathbb{R}^d)$ such that

$$\mathbb{1}_{F_{r,c}}(\mathbf{v}) \ge \varphi_1(\mathbf{v}) \text{ for all } \mathbf{v} \text{ and } \int_{\mathbb{R}^d} \varphi_1 > |F_{r,c}| - \varepsilon.$$

Now $\widehat{\varphi}_1$ is a continuous non-negative integrable function on X, and therefore there exists $\varphi_2 \in C_c(X)$ such that

$$\varphi_2(\Lambda) \le \widehat{\varphi}_1(\Lambda) \le \widehat{\mathbb{1}}_{F_{r,c}}(\Lambda) \text{ for all } \Lambda \tag{4.10}$$

and

$$\int_X \varphi_2\,\mathrm{d}\mu \ge \int_X \widehat{\varphi}_1\,\mathrm{d}\mu - \varepsilon = \int_{\mathbb{R}^d} \varphi_1 - \varepsilon \ge |F_{r,c}| - 2\varepsilon. \tag{4.11}$$

Since $\varepsilon > 0$ was arbitrary, (4.10) and (4.11) imply that for almost every $\vartheta \in M$

$$\liminf_{T\to\infty} \frac{1}{T}\int_0^T \widehat{\mathbb{1}}_{F_{r,c}}(g_t\Lambda_\vartheta)\,\mathrm{d}t \ge |F_{r,c}|.$$

This completes the proof. □

5. Pointwise equidistribution with respect to unbounded functions

In this section we prove Theorems 1.3 and 1.4. We first show that for Riemann integrable f, the function $\widehat{f}$ as in (1.10) belongs to the class $C_\alpha(X)$.

Lemma 5.1. *For any d and all sufficiently large r there are constants c_1, c_2 such that if $\mathbb{1}_{B_r}$ is the characteristic function of the open ball of radius r centered at origin, then for all $\Lambda \in X$,*

$$c_1\alpha(\Lambda) \leq \widehat{\mathbb{1}}_{B_r}(\Lambda) \leq c_2\alpha(\Lambda). \tag{5.1}$$

In particular, for any Riemann integrable $f : \mathbb{R}^d \to \mathbb{R}$, $\widehat{f} \in C_\alpha(X)$.

Proof. For any discrete subgroup $\Delta \subset \mathbb{R}^d$, the i-th Minkowski successive minimum of Δ with respect to B_r is defined to be

$$\lambda_i(\Delta) := \inf\left\{t > 0 : \dim\left(\operatorname{span}(tB_r \cap \Delta)\right) \geq i\right\}.$$

Let r_0 be large enough (depending on d) so that for any $r \geq r_0$ and any $\Lambda \in X$, $\lambda_1(\Lambda) < 1$. The notation $x \asymp y$ will mean that x and y are functions of discrete subgroups of $\mathbb{R}^d$ and there are positive constants C_1, C_2, depending on d and r, such that $C_1 \leq x/y \leq C_2$. By Minkowski's second theorem, see e.g. [7, §VIII.2],

$$\lambda_1(\Delta) \cdots \lambda_\ell(\Delta) \asymp d(\Delta),$$

where $d(\Delta)$ is the covolume of Δ in $\operatorname{span}(\Delta)$ and ℓ is the rank of Δ. Now for $\Lambda \in X$ define Δ to be the subgroup of Λ generated by $\Lambda \cap \overline{B_r}$. Then

$$\alpha(\Lambda) \asymp \left(\lambda_1(\Lambda) \cdots \lambda_j(\Lambda)\right)^{-1} = \left(\lambda_1(\Delta) \cdots \lambda_j(\Delta)\right)^{-1},$$

where j is the index for which $\lambda_j(\Lambda) \leq 1 < \lambda_{j+1}(\Lambda)$.

It follows from [4, Prop. 2.1 and Cor. 2.1] (applied to $K = B_r$, $d = j$ and $\mathbb{L} = \Delta$) that

$$\sharp(\Delta \cap B_r) \asymp \left(\lambda_1(\Delta) \cdots \lambda_j(\Delta)\right)^{-1}.$$

Since $\widehat{\mathbb{1}}_{B_r}(\Lambda) = \sharp(\Lambda \cap B_r) - 1 = \sharp(\Delta \cap B_r) - 1$, (5.1) follows.

For the second assertion, note that for any Riemann integrable $f : \mathbb{R}^d \to \mathbb{R}$ there are positive r and C so that $|f| \leq C \cdot \mathbb{1}_{B_r}$ and hence condition (C_α-2) follows from (5.1). To prove (C_α-1), let S be the set of discontinuities of f in $\mathbb{R}^d \smallsetminus \{0\}$, so that $|S| = 0$. From (1.10) it follows that the set S' of discontinuities of $\widehat{f}$ is contained in

$$S'' := \{\Lambda : \Lambda \cap S \neq \varnothing\}.$$

For each $\mathbf{v} \in \mathbb{Z}^d \smallsetminus \{0\}$, the set of $g \in G$ such that $g\mathbf{v} \in S$ has Haar measure zero in G, and hence S'' is a countable union of sets of μ-measure zero. In particular $\mu(S') = 0$. $\square$

We now derive some general properties from equidistribution of measures.

Lemma 5.2. *Let $\psi \in C_\alpha(X)$ and let $\{\mu_i\}$ be a sequence of probability measures on X such that $\mu_i \to \mu$, with respect to the weak-$*$ topology. Then for any non-negative $\varphi \in C_c(X)$ one has*

$$\lim_{i\to\infty} \int_X \varphi\psi \,\mathrm{d}\mu_i = \int_X \varphi\psi \,\mathrm{d}\mu. \tag{5.2}$$

Proof. Using assumption (C_α-1) we see that the function $\varphi\psi$ is bounded, compactly supported and continuous except on a set of measure zero. By using a partition of unity, without loss of generality one can assume that φ is supported on a coordinate chart. Applying Lebesgue's criterion for Riemann integrability to $\varphi\psi$, one can write $\int_X \varphi\psi \,\mathrm{d}\mu$ as the limit of upper and lower Riemann sums. From this it easily follows that for any $\varepsilon > 0$ there exist $h_1, h_2 \in C_c(X)$ such that $h_1 \le \varphi\psi \le h_2$ and

$$\int_X (h_2 - h_1) \,\mathrm{d}\mu \le \varepsilon. \tag{5.3}$$

Thus we have

$$\limsup_{i\to\infty} \int_X \varphi\psi \,\mathrm{d}\mu_i \le \int_X h_2 \,\mathrm{d}\mu \tag{5.4}$$

$$\liminf_{i\to\infty} \int_X \varphi\psi \,\mathrm{d}\mu_i \ge \int_X h_1 \,\mathrm{d}\mu \tag{5.5}$$

$$\int_X h_1 \,\mathrm{d}\mu \le \int_X \varphi\psi \,\mathrm{d}\mu \le \int_X h_2 \,\mathrm{d}\mu. \tag{5.6}$$

Therefore by taking $\varepsilon \to 0$, (5.2) follows from (5.4)–(5.6) and (5.3). $\square$

Corollary 5.3. *Let the notation be as in Lemma 5.2 and assume that*

$$\lim_{i\to\infty} \int_X \psi \,\mathrm{d}\mu_i = \int_X \psi \,\mathrm{d}\mu. \tag{5.7}$$

Then for any $\varepsilon > 0$ there exists $i_0 > 0$ and $\varphi \in C_c(X)$ with $0 \le \varphi \le 1$ such that

$$\left| \int_X (1 - \varphi)\psi \,\mathrm{d}\mu_i \right| < \varepsilon \tag{5.8}$$

for any $i \ge i_0$.

Proof. Since $\psi \in L^1(X, \mu)$, there exists a compactly supported continuous function $\varphi : X \to [0, 1]$ such that

$$\left| \int_X (1 - \varphi)\psi \, \mathrm{d}\mu \right| < \frac{\varepsilon}{3}. \tag{5.9}$$

By Lemma 5.2 and (5.7), there exists $i_0 > 0$ such that for $i \geq i_0$

$$\left| \int_X \varphi\psi \, \mathrm{d}\mu_i - \int_X \varphi\psi \, \mathrm{d}\mu \right| < \frac{\varepsilon}{3} \tag{5.10}$$

$$\left| \int_X \psi \, \mathrm{d}\mu_i - \int_X \psi \, \mathrm{d}\mu \right| < \frac{\varepsilon}{3}. \tag{5.11}$$

It follows from (5.9), (5.10) and (5.11) that (5.8) holds if $i \geq i_0$. □

Corollary 5.4. *Let the notation be as in Lemma 5.2. Assume that there exists a non-negative Riemann integrable function f_0 on $\mathbb{R}^d$ such that $|\psi| \leq \widehat{f_0}$ and*

$$\lim_{i \to \infty} \int_X \widehat{f_0} \, \mathrm{d}\mu_i = \int_X \widehat{f_0} \, \mathrm{d}\mu. \tag{5.12}$$

Then

$$\lim_{i \to \infty} \int_X \psi \, \mathrm{d}\mu_i = \int_X \psi \, \mathrm{d}\mu.$$

Proof. By Lemma 5.2, Corollary 5.3 and (5.12), for any $\varepsilon > 0$ there exists $i_0 > 0$ and a continuous compactly supported function $\varphi : X \to [0, 1]$ such that for $i \geq i_0$ one has

$$\left| \int_X \varphi\psi \, \mathrm{d}\mu_i - \int_X \varphi\psi \, \mathrm{d}\mu \right| < \frac{\varepsilon}{3},$$

$$\int_X (1 - \varphi)\widehat{f_0} \, \mathrm{d}\mu_i < \frac{\varepsilon}{3},$$

$$\int_X (1 - \varphi)\widehat{f_0} \, \mathrm{d}\mu < \frac{\varepsilon}{3}.$$

Using the assumption $|\psi| \leq \widehat{f_0}$ and these inequalities, one finds

$$\left| \int_X \psi \, \mathrm{d}\mu_i - \int_X \psi \, \mathrm{d}\mu \right| < \varepsilon$$

for $i \geq i_0$. This completes the proof. □

Now, given $\psi \in C_\alpha(X)$, we are going to apply the results of §4 to construct a function f_0 satisfying the assumption of Corollary 5.4.

Lemma 5.5. *For $r > d$ let A_r be the annular region defined by*

$$A_r := \{\mathbf{v} \in \mathbb{R}^d : d < \|\mathbf{v}\| < r\}.$$

Then for almost every $\vartheta \in M$,

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \widehat{\mathbb{1}}_{A_r}\big(g_t \Lambda_\vartheta\big)\, \mathrm{d}t = |A_r|. \tag{5.13}$$

Proof. For $r', c > 0$ let $F_{r',c}$ be as in (4.9). Since $r > d$, there exist $c, r' > 1$ such that $A_r \subset E_{r',c} \cup F_{r',c}$. It follows that

$$\mathbb{1}_{A_r} \le \mathbb{1}_{E_{r',c}} + \mathbb{1}_{F_{r',c}}.$$

It follows from Corollaries 4.2 and 4.3 that (5.12) holds for

$$f_0 = \mathbb{1}_{E_{r',c}} + \mathbb{1}_{F_{r',c}}.$$

Hence the conclusion of the lemma follows from Corollary 5.4. □

Lemma 5.6. *Let B_r be the open Euclidean ball of radius $r > 0$ centered at the origin in $\mathbb{R}^d$. Then*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \widehat{\mathbb{1}}_{B_r}(g_t \Lambda_\vartheta)\, \mathrm{d}t = |B_r|$$

for almost every $\vartheta \in M$.

Proof. We first observe that any annular region of width d centered at the origin contains a lattice point of any unimodular lattice in $\mathbb{R}^d$. Also for any unimodular lattice Λ in $\mathbb{R}^d$ and any $\mathbf{v} \in \Lambda$ one has

$$\sharp(B_r \cap \Lambda) = \sharp\big((B_r + \mathbf{v}) \cap \Lambda\big)\,.$$

It follows from these two observations that for any lattice $\Lambda \in X$, the number of lattice points in B_r is at least the number of lattice points in $A_{r'}$, where $r' = 2d + 2r$. That is, $\widehat{\mathbb{1}}_{B_r} \le \widehat{\mathbb{1}}_{A_{r'}}$. The conclusion now follows from Lemma 5.5 and Corollary 5.4. □

Proof of Theorem 1.3. It follows from Lemma 5.6 and the Siegel integral formula that there is a Borel subset M' of M of full measure such that for any $\vartheta \in M'$ and $r \in \mathbb{N}$ one has

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \widehat{\mathbb{1}}_{B_r}(g_t \Lambda_\vartheta)\, \mathrm{d}t = \int_X \widehat{\mathbb{1}}_{B_r}\, \mathrm{d}\mu.$$

Moreover we assume that Λ_ϑ is $\big(D^+, C_c(X)\big)$-generic for every $\vartheta \in M'$. In view of (5.1) for any $\psi \in C_\alpha(X)$ there is $c > 0$ and $r \in \mathbb{N}$ such that

$$|\psi(\Lambda)| \le c\widehat{\mathbb{1}}_{B_r}(\Lambda)$$

for every $\Lambda \in X$. It follows from Corollary 5.4 that Λ_ϑ is (D^+, ψ)-generic for every $\vartheta \in M'$, and the proof is complete. □

Proof of Theorem 1.4. Our argument is similar to the one used in [3]. The assumption on the boundaries of A and B in $\mathbb{S}^{m-1}$ and $\mathbb{S}^{n-1}$ respectively implies that the boundary of $E_{r,c}(A,B)$ in $\mathbb{R}^d$ has zero Lebesgue measure. With the notation $f_{A,B,r,c}$ as in (4.1), it follows from (4.2) and Theorem 1.3 that for almost every $\vartheta \in M$

$$\begin{aligned}\sharp\big(E_{T,c}(A,B)\cap\Lambda_\vartheta\big)\cdot r &\sim \int_0^T \widehat{f}_{A,B,r,c}(g_t\Lambda_\vartheta)\,\mathrm{d}t \\ &\sim |E_{r,c}(A,B)|\cdot T \quad \text{as } T\to\infty.\end{aligned} \tag{5.14}$$

In order to conclude the proof, it suffices to show that

$$|E_{r,c}(A,B)|\cdot T = |E_{T,c}(A,B)|\cdot r. \tag{5.15}$$

Indeed, when $T = kr$ for some $k\in\mathbb{N}$ we have

$$E_{T,c}(A,B) = \bigcup_{j=0}^{k-1} g_{-jr}\big(E_{r,c}(A,B)\big)$$

(a disjoint union), and (5.15) follows from the fact that the g_t-action preserves Lebesgue measure. From this one deduces (5.15) when $T/r\in\mathbb{Q}$, and finally one gets (5.15) for arbitrary $T,r>0$ by continuity. $\square$

References

[1] J.S. Athreya, A. Ghosh, and J. Tseng, *Spiraling of approximations and spherical averages of Siegel transforms,* preprint, (2013) http://arxiv.org/abs/1305.0296

[2] ———, *Spherical averages of Siegel transforms for higher rank diagonal actions and applications*, preprint, (2014) http://arxiv.org/abs/1407.3573

[3] J.S. Athreya, A. Parrish and J. Tseng, *Ergodic theory and Diophantine approximation for linear forms and translation surfaces*, preprint (2014), http://arxiv.org/abs/1401.4148

[4] U. Betke, M. Henk and J. M. Wills, *Successive-minima type inequalities*, Disc. Comp. Geom. **9** (1993), 165–175.

[5] M. Björklund, M. Einsiedler and A. Gorodnik, in preparation.

[6] J.W.S. Cassels, *Some metrical theorems in diophantine approximation III*, Proc. Camb. Phil. Soc. **46** (1950), 219–225.

[7] ———, *An introduction to the geometry of numbers.* Corrected reprint of the 1971 edition. Classics in Mathematics, Springer-Verlag, Berlin, 1997.

[8] A. Eskin, G. A. Margulis, and S. Mozes, *Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture*, Ann. of Math. **147** (1998), 93–141.

[9] V.P. Gaposhkin, *On the Dependence of the Convergence Rate in the Strong Law of Large Numbers for Stationary Processes on the Rate of Decay of the Correlation Function,* Theory of Probability & Its Applications, **26** (1982), no. 4, 706–720.

[10] A.G. Kachurovskii, *The rate of convergence in ergodic theorems,* Russian Math. Surveys, **51** (1996), no. 4, 653–703.

[11] D. Y. Kleinbock, *Flows on homogeneous spaces and Diophantine properties of matrices,* Duke Math. J. **95** (1998), 107–124.

[12] D. Y. Kleinbock and G. A. Margulis, *Bounded orbits of nonquasiunipotent flows on homogeneous spaces*, Sinai's Moscow Seminar on Dynamical Systems, 141-172, Amer. Math. Soc. Transl. Ser. 2, vol. 171, Amer. Math. Soc., Providence, RI, 1996.

[13] ———, *On effective equidistribution of expanding translates of certain orbits in the space of lattices*, Number theory, analysis and geometry, 385–396, Springer, New York, 2012.

[14] D. Y. Kleinbock, N. A. Shah and A. Starkov, *Dynamics of subgroup actions on homogeneous spaces of Lie groups and applications to number theory*, in **Handbook on Dynamical Systems, Volume 1A**, Elsevier Science, North Holland (2002).

[15] D. Y. Kleinbock and B. Weiss, *Dirichlet's theorem on diophantine approximation and homogeneous flows*, J. Mod. Dyn. **4** (2008), 43–62.

[16] I. Konstantoulas, *Effective multiple mixing in semidirect product actions*, preprint (2013), `http://arxiv.org/abs/1310.6974v2`.

[17] J. Marklof and A. Strömbergsson, *Free path lengths in quasicrystals*, Comm. Math. Physics **330** (2014), 723–755.

[18] W. M. Schmidt, *A metrical theorem in diophantine approximation*, Canad. J. Math. **12** (1960), 619–631.

[19] ———, *A metrical theorem in geometry of numbers*, Trans. Amer. Math. Soc. **95** (1960), 516–529.

[20] N. Shah and B. Weiss, *On actions of epimorphic subgroups on homogeneous spaces*, Ergodic Theory Dynam. Systems **20** (2000), no. 2, 567–592.

[21] R. Shi, *Pointwise equidistribution for one parameter diagonalizable group action on homogeneous space*, preprint (2014), `http://arxiv.org/abs/1405.2067`.

[22] C. L. Siegel, *A mean value theorem in geometry of numbers*, Ann. Math. **46** (1945), 340–347.

Department of Mathematics, Brandeis University, Waltham MA 02454, USA

E-mail address: `kleinboc@brandeis.edu`

School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel, and School of Mathematical Sciences, Xiamen University, Xiamen 361005, PR China

E-mail address: `ronggang@xmu.edu.cn`

School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel

E-mail address: `barakw@post.tau.ac.il`