Journal of Number Theory **79**, 83–102 (1999)
Article ID jnth.1999.2419, available online at http://www.idealibrary.com on IDEAL®

Badly Approximable Systems of Affine Forms[1]

Dmitry Kleinbock

Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903
E-mail: kleinboc@math.rutgers.edu

Communicated by M. Waldschmidt

Received May 21, 1998

We prove an inhomogeneous analogue of W. M. Schmidt's theorem on the Hausdorff dimension of the set of badly approximable systems of linear forms. The proof is based on ideas and methods from the theory of dynamical systems, in par-

1. INTRODUCTION

1.1. For $m, n \in \mathbb{N}$, we will denote by $M_{m,n}(\mathbb{R})$ the space or real matrices with m rows and n columns. Vectors will be denoted by lowercase boldface letters, such as $\mathbf{x} = (x_1, ..., x_k)^T$. A 0 will mean zero vector in any dimension, as well as zero matrix of any size. The norm $\|\cdot\|$ on $\mathbb{R}^k$ will be always given by $\|\mathbf{x}\| = \max_{1 \leqslant i \leqslant k} |x_i|$.

All distances (diameters of sets) in various metric spaces will be denoted by "dist" ("diam"), and $B(x, r)$ will stand for the open ball of radius r centered at x. To avoid confusion, we will sometimes put subscripts indicating the underlying metric space. If the metric space is a group and e is its identity element, we will write $B(r)$ instead of $B(e, r)$. The Hausdorff dimension of a subset Y of a metric space X will be denoted by $\dim(Y)$, and we will say that Y is *thick* (in X) if for any nonempty open subset W of X, $\dim(W \cap Y) = \dim(W)$ (i.e., Y has full Hausdorff dimension at any point of X).

A system of m linear forms in n variables given by $A \in M_{m,n}(\mathbb{R})$ is called *badly approximable* if there exists a constant $c > 0$ such that for every $\mathbf{p} \in \mathbb{Z}^m$ and all but finitely many $\mathbf{q} \in \mathbb{Z}^n$ the product $\|A\mathbf{q} + \mathbf{p}\|^m \|\mathbf{q}\|^n$ is greater than c; equivalently, if

$$c_A \stackrel{\text{def}}{=} \liminf_{\mathbf{p} \in \mathbb{Z}^m,\, \mathbf{q} \in \mathbb{Z}^n,\, \mathbf{q} \to \infty} \|A\mathbf{q} + \mathbf{p}\|^m \|\mathbf{q}\|^n > 0.$$

[1] Supported in part by NSF Grants DMS-9304580 and DMS-9704489.

0022-314X/99 $30.00

Denote by $\mathscr{BA}_{m,n}$ the set of badly approximable $A \in M_{m,n}(\mathbb{R})$. It has been known since the 1920s that $\mathscr{BA}_{m,n}$ is infinite (Perron, 1921) and of zero Lebesgue measure in $M_{m,n}(\mathbb{R})$ (Khintchine, 1926), and that $\mathscr{BA}_{1,1}$ is thick in $\mathbb{R}$ (Jarnik, 1929; the latter result was obtained using continued fractions). In 1969 W. M. Schmidt [S3] used the technique of (α, β)-games to show that the set $\mathscr{BA}_{m,n}$ is thick in $M_{m,n}(\mathbb{R})$.

1.2. The subject of the present paper is an inhomogeneous analogue of the above notion. By an *affine form* we will mean a linear form plus a real number. A system of m affine forms in n variables will be then given by a pair $\langle A, \mathbf{b}\rangle$, where $A \in M_{m,n}(\mathbb{R})$ and $\mathbf{b} \in \mathbb{R}^m$. We will denote by $\tilde{M}_{m,n}(\mathbb{R})$ the direct product of $M_{m,n}(\mathbb{R})$ and $\mathbb{R}^m$. Now say that a system of affine forms given by $\langle A, \mathbf{b}\rangle \in \tilde{M}_{m,n}(\mathbb{R})$ is *badly approximable* if

$$\tilde{c}_{A,\mathbf{b}} \stackrel{\text{def}}{=} \liminf_{\mathbf{p} \in \mathbb{Z}^m,\, \mathbf{q} \in \mathbb{Z}^n,\, \mathbf{q} \to \infty} \|A\mathbf{q} + \mathbf{b} + \mathbf{p}\|^m \, \|\mathbf{q}\|^n > 0, \tag{1.1}$$

and *well approximable* otherwise. We will denote by $\widehat{\mathscr{BA}}_{m,n}$ the set of badly approximable $\langle A, \mathbf{b}\rangle \in \tilde{M}_{m,n}(\mathbb{R})$.

Before going further, let us consider several trivial examples of badly approximable systems of affine forms.

1.3. Example. For comparison let us start with the homogeneous case. Suppose that $A\mathbf{q}_0 \in \mathbb{Z}^m$ for some $\mathbf{q}_0 \in \mathbb{Z}^n \setminus \{0\}$. Then clearly there exist infinitely many $\mathbf{q} \in \mathbb{Z}^n$ (integral multiples of $\mathbf{q}_0$) for which $A\mathbf{q} \in \mathbb{Z}^m$, hence such A is well approximable. On the other hand, the assumption

$$A\mathbf{q}_0 + \mathbf{b} + \mathbf{p}_0 = 0 \tag{1.2}$$

does not in general guarantee the existence of any other $\mathbf{q} \in \mathbb{Z}^n$ with $A\mathbf{q} + \mathbf{b} \in \mathbb{Z}^m$, and, in view of the definition above, just one integral solution is not enough for $\langle A, \mathbf{b}\rangle$ to be well approximable. We will say that $\langle A, \mathbf{b}\rangle \in \tilde{M}_{m,n}(\mathbb{R})$ is *rational* if (1.2) holds for some $\mathbf{p}_0 \in \mathbb{Z}^m$ and $\mathbf{q}_0 \in \mathbb{Z}^n$, and *irrational* otherwise.

Because of the aforementioned difference of the homogeneous and inhomogeneous cases, rational systems of forms will have to be treated separately. In fact, as mentioned in [C, Chap. III, Sect. 1], (1.2) allows one to reduce the study of a rational system $\langle A, \mathbf{b}\rangle \in \tilde{M}_{m,n}(\mathbb{R})$ to that of A. Indeed, for all $\mathbf{q} \neq \mathbf{q}_0$ one can write

$$\begin{aligned}\|A\mathbf{q} + \mathbf{b} + \mathbf{p}\|^m \, \|\mathbf{q}\|^n &= \|A(\mathbf{q} - \mathbf{q}_0) + \mathbf{p} - \mathbf{p}_0\|^m \, \|\mathbf{q}\|^n \\ &= \|A(\mathbf{q} - \mathbf{q}_0) + \mathbf{p} - \mathbf{p}_0\|^m \, \|\mathbf{q} - \mathbf{q}_0\|^n \frac{\|\mathbf{q}\|^n}{\|\mathbf{q} - \mathbf{q}_0\|^n},\end{aligned}$$

which shows that for rational $\langle A, \mathbf{b}\rangle \in \tilde{M}_{m,n}(\mathbb{R})$ one has $\tilde{c}_{A,\mathbf{b}} = c_A$; in particular, $\langle A, \mathbf{b}\rangle$ is badly approximable iff A is.

1.4. EXAMPLE. Another class of examples is given by

KRONECKER'S THEOREM (see [C, Chap. III, Theorem IV]). *For $\langle A, \mathbf{b}\rangle \in \tilde{M}_{m,n}(\mathbb{R})$, the following are equivalent*:

(i) *there exists $\varepsilon > 0$ such that for any $\mathbf{p} \in \mathbb{Z}^m$ and $\mathbf{q} \in \mathbb{Z}^n$ one has $\|A\mathbf{q} + \mathbf{b} + \mathbf{p}\| \geqslant \varepsilon$*;

(ii) *there exists $\mathbf{u} \in \mathbb{Z}^m$ such that $A^T\mathbf{u} \in \mathbb{Z}^n$ but $\mathbf{b}^T\mathbf{u}$ is not an integer.*

The above equivalence is straightforward in the $m = n = 1$ case; if $a = k/l \in \mathbb{Q}$, then $|aq + b + p| \geqslant \mathrm{dist}(b, (1/l)\,\mathbb{Z})$, and $a \notin \mathbb{Q}$ implies that $\{(aq + b) \bmod 1\}$ is dense in $[0, 1]$. In general, it is easy to construct numerous examples of systems $\langle A, \mathbf{b}\rangle$ satisfying (ii), and for such systems one clearly has $\tilde{c}_{A,\mathbf{b}} = +\infty$ in view of (i). Here one notices another difference from the homogenous case: in view of Dirichlet's Theorem, one has $c_A < 1$ for any $A \in M_{m,n}(\mathbb{R})$.

1.5. It follows from the inhomogeneous version of the Khintchine–Groshev Theorem (see [C, Chap. VIII, Theorem II]) that the set $\widehat{\mathscr{BA}}_{m,n}$ has Lebesgue measure zero. (See Sect. 5 for a stronger statement and other extensions.) A natural problem to consider is to measure the magnitude of this set in terms of the Hausdorff dimension. One can easily see that the systems of forms $\langle A, \mathbf{b}\rangle$ which are badly approximable by virtue of the two previous examples (that is, either are rational with $A \in \mathscr{BA}_{m,n}$ or satisfy the assumption (ii) of Kronecker's Theorem) belong to a countable union of proper submanifolds of $\tilde{M}_{m,n}(\mathbb{R})$ and, consequently, form a set of positive Hausdorff codimension. Nevertheless, the following is true and constitutes the main result of the paper:

THEOREM. *The set $\widehat{\mathscr{BA}}_{m,n}$ is thick in $\tilde{M}_{m,n}(\mathbb{R})$.*

This theorem will be proved using results and methods of the paper [KM1]. More precisely, we will derive Theorem 1.5 from Theorem 1.6 below. Before stating the latter, let us introduce some notation and terminology from the theory of Lie groups and homogeneous spaces.

1.6. Let G be a connected Lie group, $\mathfrak{g}$ its Lie algebra. Any $X \in \mathfrak{g}$ gives rise to a one-parameter semigroup $F = \{\exp(tX) \mid t \geqslant 0\}$, where exp stand for the exponential map from $\mathfrak{g}$ to G. We will be interested in the left action of F on homogeneous spaces $\Omega \stackrel{\mathrm{def}}{=} G/\Gamma$, where Γ is a discrete subgroup of G.

Many properties of the above action can be understood by looking at the adjoint action of X on $\mathfrak{g}$. For $\lambda \in \mathbb{C}$, we denote by $\mathfrak{g}_\lambda(X)$ the *generalized eigenspace* of ad X corresponding to λ, i.e., the subspace of the complexification $\mathfrak{g}_{\mathbb{C}}$ of $\mathfrak{g}$ defined by

$$\mathfrak{g}_\lambda(X) = \{ Y \in \mathfrak{g}_{\mathbb{C}} \mid (\text{ad } X - \lambda I)^j \, Y = 0 \text{ for some } j \in \mathbb{N} \}.$$

We sill say that X is *semisimple* if $\mathfrak{g}_{\mathbb{C}}$ is spanned by eigenvectors of ad X. Further, we will define the X- (or F-) *expanding horospherical subgroup* of G s follows: $H = \exp \mathfrak{h}$, where $\mathfrak{h}$ is the subalgebra of $\mathfrak{g}$ with complexification $\mathfrak{h}_{\mathbb{C}} = \bigoplus_{\operatorname{Re} \lambda > 0} \mathfrak{g}_\lambda(X)$.

Say that a discrete subgroup Γ of G is a *lattice* if the quotient space $\Omega = G/\Gamma$ has finite volume with respect to a G-invariant measure. Note that Ω may or may not be compact. Any group admitting a lattice is unimodular; we will choose a Haar measure μ on G and the corresponding Haar measure $\bar{\mu}$ on Ω so that $\bar{\mu}(\Omega) = 1$. The F-action on Ω is said to be *mixing* if

$$\lim_{t \to +\infty} \bar{\mu}(\exp(tX) \, W \cap K) = \bar{\mu}(W) \, \bar{\mu}(K) \tag{1.3}$$

for any two measurable subsets K, W of Ω.

The last piece of notation comes from the papers [K2, K3]. Consider the one-point compactification $\Omega^* \stackrel{\text{def}}{=} \Omega \cup \{\infty\}$ of Ω, topologized so that the complements to all compact sets constitute the basis of neighborhoods of ∞. We will use the notation $Z^* \stackrel{\text{def}}{=} Z \cup \{\infty\}$ for any subset Z of Ω. Now for a subset W of Ω^* and a subset F of G define $E(F, W)$ to be the set of points of Ω with F-orbits escaping W, that is,

$$E(F, W) \stackrel{\text{def}}{=} \{ x \in \Omega \mid \overline{Fx} \cap W = \varnothing \},$$

with the closure taken in the topology of Ω^*. In particular, if Z is a subset of Ω, $E(F, Z^*)$ stands for the set of $x \in \Omega$ such that orbits Fx are bounded and stay away from Z.

We are now ready to state

THEOREM. *Let G be a real Lie group, Γ a lattice in G, X a semisimple element of the Lie algebra of G, H the X-expanding horospherical subgroup of G such that the action of $F = \{\exp(tX) \mid t \geqslant 0\}$ on $\Omega = G/\Gamma$ is mixing. Then for any closed F-invariant null subset Z of Ω and any $x \in \Omega$, the set $\{ h \in H \mid hx \in E(F, Z^*) \}$ is thick in H.*

1.7. The reduction of Theorem 1.5 to the above theorem is described in Section 4 and is based on a version of S. G. Dani's (see [D1, K3, Sect. 2.5]) correspondence between badly approximable systems of linear forms and certain orbits of lattices in the Euclidean space $\mathbb{R}^{m+n}$. More precisely, let $G_0 = SL_{m+n}(\mathbb{R})$ and let G be the group $\mathrm{Aff}(\mathbb{R}^{m+n})$ of measure-preserving affine transformations of $\mathbb{R}^{m+n}$, i.e. the semidirect product of G_0 and $\mathbb{R}^{m+n}$. Further, put $\Gamma_0 = SL_{m+n}(\mathbb{Z})$ and let $\Gamma = \mathrm{Aff}(\mathbb{Z}^{m+n}) \stackrel{\text{def}}{=} \Gamma_0 \ltimes \mathbb{Z}^{m+n}$ be the subgroup of G leaving the standard lattice $\mathbb{Z}^{m+n}$ invariant. It is easy to check that Γ is a non-cocompact lattice in G, and that $\Omega \stackrel{\text{def}}{=} G/\Gamma$ can be identified with the space of *free unimodular lattices* in $\mathbb{R}^{m+n}$, i.e.,

$$\Omega \cong \{\Lambda + \mathbf{w} \mid \Lambda \text{ a lattice in } \mathbb{R}^{m+n} \text{ of covolume } 1, \mathbf{w} \in \mathbb{R}^{m+n}\}.$$

One can show that the quotient topology on Ω coincides with the natural topology on the space of lattices, so that two lattices are close to each other if their generating elements are. We will also consider $\Omega_0 \stackrel{\text{def}}{=} G_0/\Gamma_0$ and identify it with the subset of Ω consisting of "true" lattices, i.e., those containing $0 \in \mathbb{R}^{m+n}$.

We will write elements of G in the form $\langle L, \mathbf{w}\rangle$, where $L \in G_0$ and $\mathbf{w} \in \mathbb{R}^{m+n}$, so that $\langle L, \mathbf{w}\rangle$ sends $\mathbf{x} \in \mathbb{R}^{m+n}$ to $L\mathbf{x} + \mathbf{w}$. If $\mathbf{w} = 0$, we will simply write L instead of $\langle L, 0\rangle$; the same convention will apply to elements of the Lie algebra $\mathfrak{g}$ of G. We will fix an element X of $\mathfrak{sl}_{m+n}(\mathbb{R}) \subset \mathfrak{g}$ of the form

$$X = \mathrm{diag}\Bigg(\underbrace{\frac{1}{m}, ..., \frac{1}{m}}_{m \text{ times}}, \underbrace{-\frac{1}{n}, ..., -\frac{1}{n}}_{n \text{ times}}\Bigg),$$

and let $F = \{\exp(tX) \mid t \geqslant 0\}$.

Recall that in the standard version of Dani's correspondence, to a system of linear forms given by $A \in M_{m,n}(\mathbb{R})$ one associates a lattice $L_A \mathbb{Z}^{m+n}$ in $\mathbb{R}^{m+n}$, where L_A stands for $\left(\begin{smallmatrix} I_m & A \\ 0 & I_n \end{smallmatrix}\right)$. It is proved in [D1] that $A \in \mathscr{BA}_{m,n}$ if and only if the trajectory $FL_A\mathbb{Z}^{m+n} \subset \Omega_0$ is bounded (see Subsection 4.3 for more details). From this and the aforementioned 1969 result of Schmidt, Dani deduced in [D1] that the set of lattices in Ω_0 with bounded F-trajectories is thick.

It was suggested by Dani and then conjectured by G. A. Margulis [Ma, Conjecture (A)] that the abundance of bounded orbits is a general feature of nonquasiunipotent (see Subsection 2.3) flows on homogeneous spaces of Lie groups. This conjecture was settled in [KM1], thus giving an alternative (dynamical) proof of Dani's result, and hence of the thickness of the set $\mathscr{BA}_{m,n}$.

Our goal in this paper is to play the same game in the inhomogeneous setting. Given $\langle A, \mathbf{b}\rangle \in \widetilde{M}_{m,n}(\mathbb{R})$ one can consider a free lattice in $\mathbb{R}^{m+n}$ of the form $\tilde{L}_{A,\mathbf{b}}\mathbb{Z}^{m+n}$, where $\tilde{L}_{A,\mathbf{b}}$ is the element of G given by $\langle L_A, (\mathbf{b}, 0)^T\rangle$. In Section 4 we will prove the following

THEOREM. *Let G, Γ, Ω, Ω_0, and F be as above. Assume that $\tilde{L}_{A,\mathbf{b}}\mathbb{Z}^{m+n}$ belongs to $E(F, (\Omega_0)^*)$. Then a system of affine forms given by $\langle A, \mathbf{b}\rangle$ is badly approximable.*

To see that this theorem provides a link from Theorem 1.6 to Theorem 1.5, it remains to observe that the F-action on Ω is mixing (all the necessary facts related to mixing of actions on homogeneous spaces are collected in Section 2), and that $\{\tilde{L}_{A,\mathbf{b}} \mid \langle A, \mathbf{b}\rangle \in \tilde{M}_{m,n}(\mathbb{R})\}$ is the F-expanding horospherical subgroup of G. In fact, Theorem 1.7 is obtained as a corollary from a necessary and sufficient condition for an irrational system of affine forms to be badly approximable, an inhomogeneous analogue of Dani's correspondence [D1, Theorem 2.20] and the man result of Section 4 of the present paper.

For the sake of making this paper self-contained, in Section 3 we present the complete proof of Theorem 1.6, which is basically a simplified version of the argument from [KM1]. The last section of the paper is devoted to several concluding remarks and open questions.

2. MIXING AND THE EXPANDING HOROSPHERICAL SUBGROUP

2.1. Throughout the next two sections, we let G be a connected Lie group, $\mathfrak{g}$ its Lie algebra, X an element of $\mathfrak{g}$, $g_t = \exp(tX)$, $F = \{g_g \mid t \geqslant 0\}$, Γ a lattice in G and $\Omega = G/\Gamma$. For $x \in \Omega$, denote by π_x the quotient map $G \to \Omega$, $g \to gx$. The following restatement of the definition of mixing of the F-action on Ω is straightforward:

PROPOSITION (cf. [KM1, Theorem 2.1.2]). *The action of F on Ω is mixing iff or any compact $Q \subset \Omega$, any measurable $K \subset \Omega$ and any measurable $U \subset G$ such that π_z is injective on U for all $x \in Q$, one has*

$$\forall \varepsilon > 0\ \exists T > 0 \text{ such that}$$

$$|\bar{\mu}(g_t Ux \cap K) - \mu(U)\,\bar{\mu}(K)| \leqslant \varepsilon \text{ for all } t \geqslant T \text{ and } x \in Q. \tag{2.1}$$

Proof. To get (1.3) from (2.1), take any $x \in \Omega$, put $Q = \{x\}$ and take U to be any-to-one π_x-preimage of W. For the converse, one considers the family of sets $W_x = Ux$ and observes that the difference $\bar{\mu}(g_t W_x \cap K) - \bar{\mu}(W_x)\,\bar{\mu}(K)$ goes to zero uniformly when x belongs to a compact subset of Ω. ∎

2.2. If G is a connected semisimple Lie group without compact factors and Γ an irreducible lattice in G, one has C. Moore's criterion [Mo] for mixing of one-parameter subgroups of G: the F-action on Ω is mixing iff F is not relatively compact in G. Since the group $\mathrm{Aff}(\mathbb{R}^{m+n})$ is not semisimple, we need a reduction to the semisimple case based on the work [BM] of Brezin and Moore. Following [Ma], say that a homogeneous space G/Δ is a *quotient* of Ω if Δ is a closed subgroup of G containing Γ. If Δ contains a closed normal subgroup L of G such that G/L is semisimple (resp. Euclidean[2]) then the quotient G/Δ is called *semisimple* (resp. *Euclidean*). It is easy to show that the maximal semisimple (resp. Euclidean) quotient of Ω exists (by the latter we mean the quotient G/Δ of Ω such that any other semisimple (resp. Euclidean) quotient of Ω is a quotient of G/Δ).

The following proposition is a combination of Theorems 6 and 9 from [Ma]:

PROPOSITION. *Suppose that*

(i) *there are no nontrivial Euclidean quotients of* Ω, *and*

(ii) F *acts ergodically on the maximal semisimple quotient of* Ω.

Then the action of F *on* Ω *is mixing.*

2.3. Choose a Euclidean structure on $\mathfrak{g} = \mathrm{Lie}(G)$, inducing a right-invariant Riemannian metric on G and a corresponding Riemannian metric on Ω. We will fix a positive σ_0 such that

$$\begin{gathered}\text{the restriction of } \exp : \mathfrak{g} \to G \text{ to } B_{\mathfrak{g}}(4\sigma_0) \text{ is one-to-one} \\ \text{and distorts distances by at most a factor of 2.}\end{gathered} \tag{2.2}$$

Denote by $\tilde{\mathfrak{h}}$ the subalgebra of $\mathfrak{g}$ with complexification $\tilde{\mathfrak{h}}_{\mathbb{C}} = \bigoplus_{\mathrm{Re}\,\lambda > 0} \mathfrak{g}_\lambda(X)$, and put $\tilde{H} = \exp \tilde{\mathfrak{h}}$. Clearly $\mathfrak{g} = \mathfrak{h} \oplus \tilde{\mathfrak{h}}$, which implies that the multiplication map $\tilde{H} \times H \to G$ is one-to-one in a neighborhood of the identity in $\tilde{H} \times H$.

We now assume that the X-expanding horospherical subgroup H of G is nontrivial (in the terminology of [Ma], F is not *quasiunipotent*). Denote by χ the trace of $\mathrm{ad}\, X|_{\mathfrak{h}}$ and by λ the real part of an eigenvalue of $\mathrm{ad}\, X|_{\mathfrak{h}}$ with the smallest real part. Denote also by Φ_t the inner automorphism $g \to g_t g g_{-t}$ of G. Since $\mathrm{Ad}\, g_t = e^{\mathrm{ad}\, tX}$ is the differential of Φ_t at the identity, the Jacobian of $\Phi_t|_H$ is equal to $e^{\chi t}$, and local metric properties of Φ_t are determined by eigenvalues of $\mathrm{ad}\, tX$. In particular, the following is true:

[2] A connected solvable Lie group is called *Euclidean* if it is locally isomorphic to an extension of a vector group by a compact Abelian Lie group.

Lemma. *For all $t>0$ one has*

(a) $\operatorname{dist}(\Phi_t^{-1}(g), \Phi_t^{-1}(h)) \leqslant 4e^{-\lambda t} \operatorname{dist}(g, h)$ *for all* $g, h \in B_H(\sigma_0)$;

(b) *if X is semisimple*, $\operatorname{dist}(\Phi_t(g), \Phi_t(h)) \leqslant 4 \cdot \operatorname{dist}(g, h)$ *for all* $g, h \in B_{\tilde{H}}(\sigma_0)$.

In other words, Φ_t acts as an expanding map of H and as a non-expanding map of $\tilde{H}$.

2.4. We now turn to a crucial application of mixing of F-action on Ω. Choose a Haar measure ν on H. Roughly speaking, our goal is to replace a subset U of G in the formula (2.1) with a ν-measurable subset V of H.

Proposition. *Let V be a bounded measurable subset of H, K a bounded measurable subset of Ω with $\bar{\mu}(\partial K)=0$, Q a compact subset of Ω. Assume that X is semisimple and that the F-action on Ω is mixing. Then for any $\varepsilon>0$ there exists $T_1 = T_1(V, K, Q, \varepsilon)>0$ such that*

$$t \geqslant T_1 \Rightarrow \forall x \in Q \qquad \nu(\{h \in V \mid g_t hx \in K\}) \geqslant \nu(V)\,\bar{\mu}(K) - \varepsilon. \tag{2.3}$$

Proof. Since V is bounded, Q is compact, and Γ is discrete, V can be decomposed as a disjoint union of subsets V_j of H with π_x injective on some neighborhood of V_j for all $x \in Q$ and for each j. Hence one can without loss of generality assume that the maps π_x are injective on some neighborhood U' of V for all $x \in Q$. Similarly, one can safely assume that $V \subset B_H(\sigma_0)$ and $\nu(V) \leqslant 1$.

Choose a subset K' of K such that $\bar{\mu}(K') \geqslant \bar{\mu}(K) - \varepsilon/2$ and the distance σ_1 between K' and ∂K is positive. Then choose $\sigma \leqslant \min(\sigma_0, \sigma_1/4)$. After that, pick a neighborhood $\tilde{V} \subset B_{\tilde{H}}(\sigma)$ of the identity in $\tilde{H}$ such that $U \stackrel{\text{def}}{=} \tilde{V}V$ is contained in U'.

Given $x \in \Omega$ and $t>0$, denote by V' the set $\{h \in V \mid g_t hx \in K\}$ that we need to estimate the measure of.

Claim. *The set $Ux \cap g_t^{-1}K'$ is contained in $\tilde{V}V'x$.*

Proof. For any $h \in H$ and $\tilde{h} \in \tilde{V}$, $g_t \tilde{h} hx = \Phi_t(\tilde{h})\, g_t hx \in B(g_t hx, 4\sigma) \subset B(g_t hx, \sigma_1)$ by Lemma 2.3(b) and the choice of $\tilde{V}$. Therefore $g_t hx$ belongs to k whenever $g_t \tilde{h} hx \in K'$. ∎

Now, using Proposition 2.1, find $T_1>0$ such that $|\bar{\mu}(g_t Ux \cap K') - \mu(U)\,\bar{\mu}(K')| \leqslant \varepsilon/2$ for all $t \geqslant T_1$ and $x \in Q$. In order to pass from U to V, choose a left Haar measure $\tilde{\nu}$ on $\tilde{H}$ such that μ is the product of ν and $\tilde{\nu}$ (cf. [Bou, Chap. VII, Sect. 9, Proposition 13]). Then for all $t \geqslant T_1$ and $x \in Q$ one can write

$$\begin{aligned}
\tilde{\nu}(\tilde{V})\,\nu(V') &= \mu(\tilde{V}V') = \bar{\mu}(\tilde{V}V'x) \\
&\underset{\text{(Claim)}}{\geqslant} \bar{\mu}(Ux \cap g_t^{-1}K') = \bar{\mu}(g_t Ux \cap K') \\
&\underset{(\text{if } t \geqslant T_1)}{\geqslant} \mu(U)\,\bar{\mu}(K') - \varepsilon/2 \geqslant \tilde{\nu}(\tilde{V})\,\nu(V)(\bar{\mu}(K) - \varepsilon/2) - \varepsilon/2 \\
&\geqslant \tilde{\nu}(\tilde{V})(\nu(V)\,\bar{\mu}(K) - \varepsilon),
\end{aligned}$$

which immediately implies (2.3). ∎

Note that similarly one can estimate $\nu(V')$ from above, see [KM1, Proposition 2.2.1] for a more general statement.

It will be convenient to denote the Φ_t-image of V' by $V(x, K, t)$, i.e., to let

$$V(x, K, t) \stackrel{\text{def}}{=} \Phi_t(\{h \in V \mid g_t hx \in K\}) = \{h \in \Phi_t(V) \mid hg_t x \in K\}.$$

Roughly speaking, Proposition 2.4 says that the relative measure of $V(x, K, t)$ in $\Phi_t(V)$ is big when t is large enough: indeed, (2.3) can be rewritten in the form

$$t \geqslant T_1 \Rightarrow \forall x \in Q \qquad \nu(V(x, K, t)) \geqslant e^{\chi t}(\nu(V)\,\bar{\mu}(K) - \varepsilon). \tag{2.4}$$

3. PROOF OF THEOREM 1.6

3.1. Following [KM1], says that an open subset V of H is a *tessellation domain* for the right action of H on itself relative to a countable subset Λ of H if

(i) $\nu(\partial V) = 0$,

(ii) $V\gamma_1 \cap V\gamma_2 = \varnothing$ for different $\gamma_1, \gamma_2 \in \Lambda$, and

(iii) $X = \bigcup_{\gamma \in \Lambda} \bar{V}\gamma$.

The pair (V, Λ) will be called a *tessellation* of H. Note that it follows easily from (ii) and (iii) that for any measurable subset A of H one has

$$\begin{aligned}
\frac{\nu(A)}{\nu(V)} &\leqslant \#\{\gamma \in \Lambda \mid V\gamma \cap A \neq \varnothing\} \\
&\leqslant \frac{\nu(\{h \in H \mid \operatorname{dist}(h, A) \leqslant \operatorname{diam}(V)\})}{\nu(V)}.
\end{aligned} \tag{3.1}$$

Let k stand for the dimension of H. We will use a one-parameter family of tessellations of H defined as follows: if $\{X_1, ..., X_k\}$ is a fixed orthonormal *strong Malcev basis* of $\mathfrak{h}$ (see [CG; KM1, Sect. 3.3] for the definition), we let $I = \{\sum_{j=1}^k x_j X_j \mid |x_j| < 1/2\}$ be the unit cube in $\mathfrak{h}$, and then take $V_r = \exp((r/\sqrt{k})\, I)$. It was proved in [KM1] that V_r is a tessellation domain of H; let Λ_r be a corresponding set of translations. It is clear from (2.2) that V_r is contained in $B(r)$ provided $r \leqslant \sigma_0$, where σ_0 is as in Subsection 2.3.

The main ingredient of the proof of Theorem 1.6 is given by the following procedure: we look at the expansion $\Phi_t(V_r)$ of the set V_r by the automorphism Φ_t, and then consider the translates $V_r\gamma$ which lie entirely inside $\Phi_t(V_r)$. It was shown in [KM1] (see also [K2, Proposition 2.6]) that when t is large enough, the measure of the union of all such translates is approximately equal to the measure of $\Phi_t(V_r)$; in other words, boundary effects are negligible. More precisely, the following is what will be needed for the proof of the main theorem:

PROPOSITION. *For any $r \leqslant \sigma_0$ and any $\varepsilon > 0$ there exists $T_2 = T_2(r, \varepsilon) > 0$ such that*

$$t \geqslant T_2 \Rightarrow \#\{\gamma \in \Lambda_r \mid V_r\gamma \cap \partial(\Phi_t(V_r)) \neq \varnothing\} \leqslant \varepsilon e^{\varkappa t}.$$

Proof. One can write

$$\#\{\gamma \in \Lambda_r \mid V_r\gamma \cap \partial(\Phi_t(V_r)) \neq \varnothing\} = \#\{\gamma \in \Lambda_r \mid \Phi_t^{-1}(V_r\gamma) \cap \partial V_r \neq \varnothing\}.$$

Observe that $(\Phi_t^{-1}(V_r), \Phi_t^{-t}(\Lambda_r))$ is also a tessellation of H, and, in view of Lemma 2.3(a), the diameter of $\Phi_t^{-1}(V_r)$ is at most $8re^{-\lambda t}$. Therefore, by (3.1), the number in the right hand side is not greater than the ratio of the measure of the $8re^{-\lambda t}$ neighborhood of ∂V_r (which, in view of condition (i) above, tends to zero as $t \mapsto \infty$) and $\nu(\Phi_t^{-1}(V_r)) = e^{-\varkappa t}\nu(V_r)$. This shows that $\lim_{t\to\infty} e^{-\varkappa t} \#\{\gamma \in \Lambda_r \mid V_r\gamma \cap \partial(\Phi_t(V_r)) \neq \varnothing\} = 0$, hence the proposition. ∎

3.2. Suppose a subset K of Ω, a point $x \in \Omega$, $t > 0$ and positive $r \leqslant \sigma_0$ are given. Consider a tessellation (V_r, Λ_r) of H, and recall that we defined $V_r(x, K, t)$ as the set of all elements h in $\Phi_t(V_r)$ for which $hg_t x$ belongs to K. Our goal now is to approximate this set by the union of translates of V_r. More precisely, let us denote by $\Lambda_r(x, K, t)$ the set of translations $\gamma \in \Lambda_r$ such that $V_r\gamma$ lies entirely inside $V_r(x, K, t)$; in other words, if $V_r\gamma \subset \Phi_t(V_r)$ and $V_r\gamma g_t x = g_t \Phi_t^{-1}(V_r\gamma)\, x$ is contained in K. Then the union

$$\bigcup_{\gamma \in \Lambda_r(x, K, t)} V_r\gamma = \bigcup_{\gamma \in \Lambda,\ V_r\gamma \subset V_r(x, K, t)} V_r\gamma$$

can be thought of as a "tessellation approximation" to $V_r(x, K, t)$. We can therefore think of the theorem below as of a "tessellation approximation" to Proposition 2.4.

THEOREM. *Let K be a subset of Ω with $\bar{\mu}(\partial K)=0$, Q a compact subset of Ω. Then for any $\varepsilon>0$ there exists $r_0=r_0(K, \varepsilon)\in(0, \sigma_0)$ such that for any positive $r\leqslant r_0$ one can find $T_0=T_0(K, Q, \varepsilon, r)>0$ with the property*

$$t\geqslant T_0 \Rightarrow \forall x\in Q \qquad \# \Lambda_r(x, K, t)\geqslant e^{\varkappa t}(\bar{\mu}(K)-\varepsilon). \tag{3.2}$$

Proof. If $\bar{\mu}(K)=0$, there is nothing to prove. Otherwise, pick a compact subset K' of K with $\bar{\mu}(\partial K')=0$, which satisfies $\bar{\mu}(K')\geqslant\bar{\mu}(K)-\varepsilon/3$ and lies at a positive distance from the complement of K. Take $r_0\leqslant\sigma_0$ such that $V_{r_0}V_{r_0}^{-1}K'\subset K$ (hence $V_rV_r^{-1}K'\subset K$ for any positive $r\leqslant r_0$). Then for any $t>0$ and $x\in\Omega$ one has

$$\begin{aligned} V_r\gamma g_t x\subset K &\Leftarrow v_r\gamma g_t x\subset V_rV_r^{-1}K' \Leftarrow \gamma g_t x\in V_r^{-1}K' \\ &\Leftarrow V_r\gamma g_t x\cap K'\neq\varnothing. \end{aligned}$$

Therefore

$$\begin{aligned} \#\Lambda_r(x, K, t) &= \#\{\gamma\in\Lambda_r \mid V_r\gamma\subset\Phi_t(V_r)\ \&\ V_r\gamma g_t x\subset K\} \\ &\geqslant \#\{\gamma\in\Lambda_r \mid V_r\gamma\subset\Phi_t(V_r)\ \&\ V\gamma g_t x\cap K'\neq\varnothing\} \\ &\geqslant \#\{\gamma\in\Lambda_r \mid V_r\gamma\cap V_r(x, K', t)\neq\varnothing\} \\ &\quad -\#\{\gamma\in\Lambda_r \mid V_r\gamma\cap\partial(\Phi_t(V_r))\neq\varnothing\}. \end{aligned}$$

Now take

$$T_0=\max\left(T_1\left(V_r, K', Q, \frac{\varepsilon\nu(V)}{3}\right)\text{ from Proposition 2.4,}\right.$$
$$\left. T_2\left(r, \frac{\varepsilon}{3}\right)\text{ from Proposition 3.1}\right).$$

Then the number of $\gamma\in\Lambda_r$ for which $V_r\gamma$ has nonempty intersection with $V_r(x, K', t)$ is, in view of (2.4) and (3.1), for all $x\in Q$ and $t\geqslant T_0$ not less than

$$\frac{e^{\varkappa t}(\nu(V_r)\,\bar{\mu}(K')-\varepsilon\nu(V_r)/3)}{\nu(V_r)}=e^{\varkappa t}(\bar{\mu}(K')-\varepsilon/3)\geqslant e^{\varkappa t}(\bar{\mu}(K)-2\varepsilon/3).$$

On the other hand, the number of translates $V_r\gamma$ nontrivially intersecting with $\partial(\Phi_t(V_r))$ is, by Proposition 3.1, not greater than $\varepsilon e^{\varkappa t}/3$, and (3.2) follows. ∎

3.3. We now describe a construction of a class of sets for which there is a natural lower estimate for the Hausdorff dimension. Let X be a Riemannian manifold, ν a Borel measure on X, A_0 a compact subset of X. Say that a countable collection $\mathscr{A}$ of compact subsets of A_0 of positive measure ν is *tree-like* relative to ν if $\mathscr{A}$ is the union of finite nonempty subcollections A_j, $j=0, 1, \ldots$, such that $\mathscr{A}_0=\{A_0\}$ and the following two conditions are satisfied:

(TL1) $\forall j \in \mathbb{N}\ \forall A, B \in \mathscr{A}_j$ either $A=B$ or $\nu(A \cap B)=0$;

(TL2) $\forall j \in \mathbb{N}\ \forall B \in \mathscr{A}_j\ \exists A \in \mathscr{A}_{j-1}$ such that $B \subset A$.

Say also that $\mathscr{A}$ is *strongly tree-like* if it is tree-like and in addition

(STL) $d_j(\mathscr{A}) \stackrel{\text{def}}{=} \sup_{A \in \mathscr{A}_j} \operatorname{diam}(A) \to 0$ as $j \to \infty$.

Let $\mathscr{A}$ be a tree-like collection of sets. For each $j=0, 1, \ldots,$ let $\mathbf{A}_j=\bigcup_{A \in \mathscr{A}_j} A$. These are nonempty compact sets, and from (TL2) it follows that $\mathbf{A}_j \subset \mathbf{A}_{j-1}$ for any $j \in \mathbb{N}$. Therefore one can define the (nonempty) *limit set* of $\mathscr{A}$ to be

$$\mathbf{A}_\infty=\bigcap_{j=0}^{\infty} \mathbf{A}_j.$$

Further, for any subset B of A_0 with $\nu(B)>0$ and any $j \in \mathbb{N}$, define the *jth stage density* $\delta_j(B, \mathscr{A})$ of B in $\mathscr{A}$ by

$$\delta_j(B, \mathscr{A})=\frac{\nu(\mathbf{A}_j \cap B)}{\nu(B)},$$

and the *jth stage density* $\delta_j(\mathscr{A})$ of $\mathscr{A}$ by $\delta_j(\mathscr{A})=\inf_{B \in \mathscr{A}_{j-1}} \delta_j(B, \mathscr{A})$.

The following estimate, based on an application of Frostman's Lemma, is essentially proved in [Mc, U]:

Lemma. *Assume that there exists $k>0$ such that*

$$\liminf_{r \to 0} \frac{\log \nu(B(x, r))}{\log r} \geqslant k \tag{3.3}$$

for any $x \in A_0$. Then for any strongly tree-like (relative to ν) collection $\mathscr{A}$ of subsets of A_0

$$\dim(\mathbf{A}_\infty) \geqslant k-\limsup_{j \to \infty} \frac{\sum_{i=1}^{j} \log \delta_i(\mathscr{A})}{\log d_j(\mathscr{A})}.$$

3.4. Now everything is ready for the

Proof of Theorem 1.6. Let $x \in \Omega$ and a nonempty open subset V of H be given. We need to prove that the Hausdorff dimension of the set $\{h \in V \mid hx \in E(F, Z^*)\}$ is equal to $k = \dim(H)$. Replacing x by hx for some $h \in V$ we can assume that V is a neighborhood of identity in H.

Pick a compact set $K \subset \Omega \backslash Z$ with $\bar{\mu}(\partial K) = 0$, and choose arbitrary $\varepsilon > 0$, $\varepsilon < \mu(K)$. Then, using Theorem 3.2, find $r \leqslant r_0(K, \varepsilon)$ such that the corresponding tessellation domain $\bar{V}_r$ is contained in V, and then take $t \geqslant \max(T_0(K, K \cup \{x\}, \varepsilon, r), 1/\varepsilon)$. We claim that

$$\dim(\{h \in \overline{V_r} \mid hx \in E(F, Z^*)\}) \geqslant k - \frac{\log(1/(\bar{\mu}(K) - \varepsilon))}{\lambda t - \log 4}. \tag{3.4}$$

Since $\overline{V_r} \subset V$, ε is arbitrary small and t is greater than $1/\varepsilon$, it follows from (3.4) that $\dim(\{h \in V \mid hx \in E(F, Z^*)\})$ is equal to k.

To demonstrate (3.4), for all $y \in K \cup \{x\}$ let us define strongly tree-like (relative to the Haar measure ν on H) collections $\mathscr{A}(y)$ inductively as follows. We first let $\mathscr{A}_0(y) = \{\overline{V_r}\}$ for all y, then define

$$\mathscr{A}_1(y) = \{\Phi_{-t}(\overline{V_r}\gamma) \mid \gamma \in \Lambda_r(y, K, t)\}. \tag{3.5}$$

More generally, if $\mathscr{A}_i(y)$ is defined for all $y \in K \cup \{x\}$ and $i < j$, we let

$$\mathscr{A}_j(y) = \{\Phi_{-t}(A\gamma) \mid \gamma \in \Lambda_r(y, K, t),\ A \in \mathscr{A}_{j-1}(\gamma g_t y)\}. \tag{3.6}$$

By definition, $\gamma \in \Lambda_r(y, K, t)$ implies that $\gamma g_t y \in K$; therefore $\mathscr{A}_{j-1}(\gamma g_t y)$ in (3.6) is defined and the inductive procedure goes through. The properties (TL1) and (TL2) follow readily from the construction and V_r being a tessellation domain. Also, by Lemma 2.3(a), the diameter of $\Phi_{-t}(A\gamma)$ is not greater than $4e^{-\lambda t}\operatorname{diam}(A)$, which implies that $d_j(\mathscr{A}(y))$ is for all $j \in \mathbb{N}$ and $y \in K$ not greater than $2r \cdot (4e^{-\lambda t})^j$, and therefore (STL) is satisfied.

Let us now show by induction that the jth stage density $\delta_j(\mathscr{A}(y))$ of $\mathscr{A}(y)$ is for all $y \in K \cup \{x\}$ and $j \in \mathbb{N}$ bounded from below by $\bar{\mu}(K) - \varepsilon$. Indeed, by definition

$$\begin{aligned}\delta_1(\overline{V_r}, \mathscr{A}(y)) = \frac{\nu(\mathbf{A}_1(y))}{\nu(\bar{V}_r)} &\underset{(\text{by } (3.5))}{=} \frac{\nu\left(\bigcup_{\gamma \in \Lambda_r(y, K, t)} \Phi_t^{-1}(V_r\gamma)\right)}{\nu(V_r)} \\ &= e^{-\chi t} \# \Lambda_r(y, K, t) \underset{(\text{by } (3.2))}{\geqslant} \bar{\mu}(K) - \varepsilon.\end{aligned}$$

On the other hand, if $j \geqslant 2$ and $B \in \mathscr{A}_{j-1}(y)$ is of the form $\Phi_t^{-1}(A\gamma)$ for $A \in \mathscr{A}_{j-2}(\gamma g_t y)$, the formula (3.6) gives

$$\begin{aligned}\delta_j(B, \mathscr{A}(y)) &= \frac{\nu(B \cap \mathbf{A}_j(y))}{\nu(B)} = \frac{\nu(\Phi_t^{-1}(B \cap \mathbf{A}_j(y)))}{\nu(\Phi_t^{-1}(B))} \\ &= \frac{\nu(A\gamma \cap \mathbf{A}_{j-1}(\gamma g_t y)\,\gamma)}{\nu(A\gamma)} = \frac{\nu(A \cap \mathbf{A}_{j-1}(\gamma g_t y))}{\nu(A)} \\ &= \delta_{j-1}(A, \mathscr{A}(\gamma g_t y)),\end{aligned}$$

and induction applies. Finally, the measure ν clearly satisfies (3.3) with $k = \dim(H)$, and an application of Lemma 3.3 yields that for all $y \in K \cup \{x\}$ one has

$$\dim(\mathbf{A}_\infty(y)) \geqslant k - \limsup_{j \to \infty} \frac{j \log(\bar{\mu}(K) - \varepsilon)}{\log(2r \cdot (4e^{-\lambda t})^j)},$$

which is exactly the right hand side of (3.4).

To finish the proof it remains to show that $\mathbf{A}_\infty(x)\, x$ is a subset of $E(F, Z^*)$. Indeed, from (3.5) and the definition of $\Lambda_r(y, K, t)$ it follows that $g_t \mathbf{A}_1(y)\, y \subset K$ for all $y \in K \cup \{x\}$. Using (3.6) one can then inductively prove that $g_{jt} \mathbf{A}_j(y)\, y \subset K$ for all $y \in K \cup \{x\}$ and $j \in \mathbb{N}$. This implies that

$$g_{jt} \mathbf{A}_\infty(x) \subset K \qquad \text{for all} \quad j \in \mathbb{N}. \tag{3.7}$$

It remains to define the set $C = \bigcup_{s=-t}^{0} g_s K$, which is compact and disjoint from Z due to the F-invariance of the latter. From (3.7) it easily follows that for any $h \in \mathbf{A}_\infty(x)$, the orbit Fhx is contained in C, and therefore is bounded and disjoint from Z. ∎

4. DIOPHANTINE APPROXIMATION AND ORBITS OF LATTICES

4.1. We return to the notation introduced in Section 1, i.e., put $G = SL_{m+n}(\mathbb{R}) \ltimes \mathbb{R}^{m+n}$, $\Gamma = SL_{m+n}(\mathbb{Z}) \ltimes \mathbb{Z}^{m+n}$, $G_0 = SL_{m+n}(\mathbb{R})$, $\Gamma_0 = SL_{m+n}(\mathbb{Z})$, $\Omega = G/\Gamma$, the space of free lattices in $\mathbb{R}^{m+n}$, $L_A = \left(\begin{smallmatrix} I_m & A \\ 0 & I_n \end{smallmatrix}\right)$, $\tilde{L}_{A,\mathbf{b}} = \langle L_A, (\mathbf{b}, 0)^T \rangle$,

$$X = \operatorname{diag}\Big(\underbrace{\frac{1}{m}, \dots, \frac{1}{m}}_{m \text{ times}}, \underbrace{-\frac{1}{n}, \dots, -\frac{1}{n}}_{n \text{ times}}\Big),$$

$g_t = \exp(tX)$ and $F = \{g_t \mid t \geqslant 0\}$.

As is mentioned in the introduction to [KM1], $\{L_A \mid A \in M_{m,n}(\mathbb{R})\}$ is the F-expanding horospherical subgroup of G_0. Similarly, one has

LEMMA. *$\{\tilde{L}_{A,\mathbf{b}} \mid \langle A, \mathbf{b} \rangle \in \tilde{M}_{m,n}(\mathbb{R})\}$ is the F-expanding horospherical subgroup of G.*

Proof. It is a straightforward computation that ad X sends $\langle \begin{pmatrix} B & A \\ C & D \end{pmatrix}, \begin{pmatrix} \mathbf{b} \\ \mathbf{c} \end{pmatrix} \rangle \in \mathfrak{g}$ (here $A \in M_{m,n}(\mathbb{R})$, $B \in M_{m,m}(\mathbb{R})$, $C \in M_{n,m}(\mathbb{R})$, $D \in M_{n,n}(\mathbb{R})$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} \in \mathbb{R}^n$) to the element

$$\left\langle \begin{pmatrix} 0 & \left(\frac{1}{m} + \frac{1}{n}\right) A \\ -\left(\frac{1}{m} + \frac{1}{n}\right) C & 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{m}\mathbf{b} \\ -\frac{1}{n}\mathbf{c} \end{pmatrix} \right\rangle. \quad \blacksquare$$

4.2. We also need

LEMMA. *The action of F on Ω is mixing.*

Proof. Since $\mathbb{R}^{m+n}$ is the only nontrivial closed normal subgroup of G, the homogeneous space Ω has no nontrivial Euclidean quotients and its maximal semisimple quotient is equal to G/Δ, where $\Delta = \Gamma_0 \ltimes \mathbb{R}^{m+n}$. Denote by p the quotient map $G \to G_0 \cong G/\mathbb{R}^{m+n}$. Then G/Δ, as a G-space, is p-equivariantly isomorphic to G_0/Γ_0, and clearly $p(F)$ is not relatively compact in G_0. It follows from Moore's theorem that the F-action on G/Δ is mixing, therefore, by Proposition 2.2, so is the F-action on Ω. $\blacksquare$

4.3. We are now going to connect Diophantine properties of $\langle A, \mathbf{b} \rangle \in \tilde{M}_{m,n}(\mathbb{R})$ with orbit properties of $\tilde{L}_{A,\mathbf{b}}\mathbb{Z}^{m+n}$. For comparison, let us first state Dani's correspondence [D1, Theorem 2.20] for homogeneous approximation.

THEOREM. *$A \in M_{m,n}(\mathbb{R})$ is badly approximable iff there exists $\varepsilon > 0$ such that $\|g_t L_A \mathbf{v}\| \geqslant \varepsilon$ for all $t \geqslant 0$ and $\mathbf{v} \in \mathbb{Z}^{m+n} \setminus \{0\}$.*

In view of Mahler's compactness criterion, the latter assertion is equivalent to the orbit $FL_A\mathbb{Z}^{m+n}$ being bounded (in other words, to $L_A\mathbb{Z}^{m+n}$ being an element of $E(F, \{\infty\})$). Therefore, as is mentioned in [K3], one can use the result of [KM1] to get an alternative proof of Schmidt's theorem on thickness of the set of badly approximable systems of linear forms. In order to move to affine forms, we need an inhomogeneous analogue of the above criterion:

4.4. THEOREM. *$\langle A, \mathbf{b}\rangle \in \widetilde{M}_{m,n}(\mathbb{R})$ is irrational and badly approximable iff*

$$\exists \varepsilon > 0 \text{ such that } \|g_t \widetilde{L}_{A,\mathbf{b}}\mathbf{v}\| \geqslant \varepsilon \text{ for all } t \geqslant 0 \text{ and } \mathbf{v} \in \mathbb{Z}^{m+n}. \tag{4.1}$$

Proof. We essentially follow the argument of [K1, Proof of Proposition 5.2(a)]. Write $\mathbf{v} = (\mathbf{p}, \mathbf{q})^T$, where $\mathbf{p} \in \mathbb{Z}^m$ and $\mathbf{q} \in \mathbb{Z}^n$. Then

$$g_t \widetilde{L}_{A,\mathbf{b}}\mathbf{v} = (e^{t/m}(A\mathbf{q} + \mathbf{b} + \mathbf{p}), e^{-t/n}\mathbf{q})^T.$$

This shows that (4.1) does not hold iff there exist sequences $t_j \geqslant 0$, $\mathbf{p}_j \in \mathbb{Z}^m$ and $\mathbf{q}_j \in \mathbb{Z}^n$ such that

$$e^{t_j/m}(A\mathbf{q}_j + \mathbf{b} + \mathbf{p}_j) \to 0 \tag{4.2a}$$

and

$$e^{-t_j/n}\mathbf{q}_j \to 0 \tag{4.2b}$$

as $j \to \infty$. On the other hand, $\langle A, \mathbf{b}\rangle$ is well approximable iff there exist sequences $\mathbf{p}_j \in \mathbb{Z}^m$ and $\mathbf{q}_j \in \mathbb{Z}^n$, $j \in \mathbb{N}$, such that $\mathbf{q}_j \to \infty$ and

$$\|A\mathbf{q}_j + \mathbf{b} + \mathbf{p}_j\|^m \, \|\mathbf{q}_j\|^n \to 0 \qquad \text{as} \quad j \to \infty. \tag{4.3}$$

We need to prove that (4.1) doe snot hold if and only if $\langle A, \mathbf{b}\rangle$ is either rational or well approximable. If $\langle A, \mathbf{b}\rangle$ is rational, one can take $\mathbf{p}_j = \mathbf{p}_0$ and $\mathbf{q}_j = \mathbf{q}_0$, with $\mathbf{p}_0$ and $\mathbf{q}_0$ as in (1.2), and arbitrary $t_j \to \infty$; then the left hand side of (4.2a) is zero, and (4.2b) is satisfied as well. On the other hand, for irrational and well approximable $\langle A, \mathbf{b}\rangle$ one can define $e^{t_j} \stackrel{\text{def}}{=} \sqrt{\|\mathbf{q}_j\|^n / \|A\mathbf{q}_j + \mathbf{b} + \mathbf{p}_j\|^m}$ and check that (4.2a), (4.2b) holds.

Conversely, multiplying the norm of the left hand side of (4.2a) risen to he mth power and the norm of the left hand side of (4.2b) risen to the nth power, one immediately sees that (4.3) follows from (4.2a), (4.2b). It remains to observe that either $\langle A, \mathbf{b}\rangle$ is rational or $A\mathbf{q}_j + \mathbf{b} + \mathbf{p}_j$ is never zero, therefore (4.2a) forces the sequences $\mathbf{q}_j$ to tend to infinity. ▮

4.5. Recall that we denoted by Ω_0 the set of "true" (containing the zero vector) lattices in $\mathbb{R}^{m+n}$. It is now easy to complete the

Proof of Theorem 1.7. Take a well approximable $\langle A, \mathbf{b}\rangle \in \widetilde{M}_{m,n}(\mathbb{R})$ such that $\widetilde{L}_{A,\mathbf{b}}\mathbb{Z}^{m+n}$ belongs to $E(F, (\Omega_0)^*)$. In view of the above criterion,

$$\exists \text{ a sequence } \Lambda_j \in F\widetilde{L}_{A,\mathbf{b}}\mathbb{Z}^{m+n} \text{ and vectors } \mathbf{v}_j \in \Lambda_j \text{ with } \mathbf{v}_j \to 0. \tag{4.4}$$

Since $F\widetilde{L}_{A,\mathbf{b}}\mathbb{Z}^{m+n}$ is relatively compact, one can without loss of generality assume that there exists $\Lambda \in \Omega$ with $\Lambda_j \to \Lambda$ in the topology of Ω. Clearly the presence of arbitrarily small vectors in the lattices Λ_j forces Λ to contain 0, i.e., belong to Ω_0, which is a contradiction. ∎

4.6. *Remark.* Note that the converse to Theorem 1.7 is not true: by virtue of Theorem 4.4, any rational $\langle A, \mathbf{b}\rangle \in \tilde{M}_{m,n}(\mathbb{R})$ satisfies (4.4), hence $\tilde{L}_{A,\mathbf{b}}\mathbb{Z}^{m+n}$ is not in $E(F, (\Omega_0)^*)$. Restriction to the irrational case gives a partial converse: indeed, the above proof basically shows that the existence of a limit point $\Lambda \in \Omega_0$ of the orbit $F\tilde{L}_{A,\mathbf{b}}\mathbb{Z}^{m+n}$ violates (4.1); hence $\tilde{L}_{A,\mathbf{b}}\mathbb{Z}^{m+n}$ belongs to $E(F, \Omega_0)$ whenever $\langle A, \mathbf{b}\rangle$ is irrational and badly approximable. But the orbit $F\tilde{L}_{A,\mathbf{b}}\mathbb{Z}^{m+n}$ does not have to be bounded, as can be shown using the explicit construction given by Kronecker's Theorem (see Example 1.4). Perhaps the simplest possible example is the irrational badly approximable form $\langle A, \mathbf{b}\rangle = \langle 0, 1/2\rangle$ (here $m = n = 1$): it is easy to see that the orbit

$$\mathrm{diag}(e^t, e^{-t})\, \tilde{L}_{A,\mathbf{b}}\mathbb{Z}^2 = \{(e^t(p+1/2), e^{-t}q)^T \mid p, q \in \mathbb{Z}\}$$

has no limit points in the space of free lattices in $\mathbb{R}^2$.

4.7. We conclude this section with the

Proof of Theorem 1.5. Observe that $\Omega_0 = G_0\mathbb{Z}^{m+n}$ is the orbit of a proper subgroup of G containing F, which makes it null and F-invariant subset of Ω. The fact that Ω_0 is closed is also straightforward. From Theorem 1.6 and Lemmas 4.1 and 4.2 it follows that the set $\{\langle A, \mathbf{b}\rangle \in \tilde{M}_{m,n}(\mathbb{R}) \mid \tilde{L}_{A,\mathbf{b}}\mathbb{Z}^{m+n} \in E(F, (\Omega_0)^*)\}$ is thick in $\tilde{M}_{m,n}(\mathbb{R})$. In view of Theorem 1.7, systems of forms which belong to the latter set are badly approximable, hence the thickness of the set $\widehat{\mathscr{BA}}_{m,n}$. ∎

5. CONCLUDING REMARKS AND OPEN QUESTIONS

It is worthwhile to look at the main result of this paper in the context of other results in inhomogeneous Diophantine approximation. In what follows, $\psi: \mathbb{N} \mapsto (0, \infty)$ will be a non-increasing function, and we will say, following [KM3], that a system of affine forms given by $\langle A, \mathbf{b}\rangle \in \tilde{M}_{m,n}(\mathbb{R})$ is *ψ-approximable*[3] if there exists infinitely many $\mathbf{q} \in \mathbb{Z}^n$ such that

$$\|A\mathbf{q} + \mathbf{b} + \mathbf{p}\|^m \leqslant \psi(\|\mathbf{q}\|^n) \tag{5.1}$$

[3] We are grateful to M. M. Dodson for a permission to modify his terminology introduced in [Do1]. In our opinion, the use of (5.1) instead of traditional $\|A\mathbf{q} + \mathbf{b} + \mathbf{p}\| \leqslant \psi(\|\mathbf{q}\|)$ makes the structure, including the connection with homogeneous flows, more transparent. See [KM2, KM3] for justification.

for some $\mathbf{p} \in \mathbb{Z}^m$. Denote by $\mathscr{W}_{m,n}(\psi)$ the set of ψ-approximable systems $\langle A, \mathbf{b} \rangle \in \tilde{M}_{m,n}(\mathbb{R})$.

The main result in the present paper is *doubly metric* in its nature; that is, the object of study is the set of *all* pairs $\langle A, \mathbf{b} \rangle$. On the contrary, in *singly metric* inhomogeneous problems one is interested in the set of pairs $\langle A, \mathbf{b} \rangle$ where $\mathbf{b}$ is fixed. For example, the doubly metric inhomogeneous Khintchine–Groshev Theorem [C, Chap. VII, Theorem II] says that

$$\text{the set } \mathscr{W}_{m,n}(\psi) \text{ has } \begin{cases} \text{full measure} & \text{if } \sum_{k=1}^{\infty} \psi(k) = \infty \\ \text{zero measure} & \text{if } \sum_{k=1}^{\infty} \varphi(k) < \infty. \end{cases}$$

The singly metric strengthening is due to Schmidt. For $\mathbf{b} \in \mathbb{R}^m$, denote by $\mathscr{W}_{m,n}(\psi, \mathbf{b})$ the set of matrices $A \in M_{m,n}(\mathbb{R})$ such that $\langle A, \mathbf{b} \rangle \in \mathscr{W}_{m,n}(\psi)$. It follows from the main result of [S1] that given any $\mathbf{b} \in \mathbb{R}^m$,

$$\text{the set } \mathscr{W}_{m,n}(\psi, \mathbf{b}) \text{ has } \begin{cases} \text{full measure} & \text{if } \sum_{k=1}^{\infty} \psi(k) = \infty \\ \text{zero measure} & \text{if } \sum_{k=1}^{\infty} \varphi(k) < \infty. \end{cases}$$

Denote $\psi_0(x) = 1/x$. A quick comparison of (5.1) with (1.1) shows that $\langle A, \mathbf{b} \rangle$ is badly approximable iff it is not $\varepsilon\psi_0$-approximable for some $\varepsilon > 0$; in other words, $\widehat{\mathscr{BA}}_{m,n} = \tilde{M}_{m,n}(\mathbb{R}) \setminus \bigcup_{\varepsilon > 0} \mathscr{W}_{m,n}(\varepsilon\psi_0)$. Therefore it follows from the above result that for any $\mathbf{b} \in \mathbb{R}^m$, the set

$$\mathscr{BA}_{m,n}(\mathbf{b}) \stackrel{\text{def}}{=} \{ A \in M_{m,n}(\mathbb{R}) \mid \langle A, \mathbf{b} \rangle \in \widehat{\mathscr{BA}}_{m,n} \}$$

has zero measure.

Another class of related problems involves a connection between the rate of decay of ψ and the Hausdorff dimension of $\mathscr{W}_{m,n}(\psi)$. The corresponding result in homogeneous approximation is called the Jarnik–Besicovitch Theorem, see [Do1]. The doubly metric inhomogeneous version was done by M. M. Dodson [Do2] and H. Dickinson [Di], and recently J. Levesley [L] obtained a singly metric strengthening. More precisely he proved that for any $\mathbf{b} \in \mathbb{R}^m$,

$$\dim(\mathscr{W}_{m,n}(\psi, \mathbf{b})) = \begin{cases} mn\left(1 - \dfrac{\lambda - 1}{m + n\lambda}\right) & \text{if } \lambda > 1 \\ mn & \text{if } \lambda \leqslant 1, \end{cases}$$

where $\lambda = \liminf_{k \to \infty} (\log(1/\psi(k))/\log k)$ is the *lower order* of the function $1/\psi$.

In view of the aforementioned results, one can ask whether it is possible to prove that $\mathscr{BA}_{m,n}(\mathbf{b})$ is thick in $M_{m,n}(\mathbb{R})$ for every $\mathbf{b} \in \mathbb{R}^m$. (It is a consequence of Theorem 1.7 and slicing properties of the Hausdorff dimension that vectors $\mathbf{b} \in \mathbb{R}^m$ such that $\mathscr{BA}_{m,n}(\mathbf{b})$ is thick form a thick subset of $\mathbb{R}^m$.)

Note also that in the paper [S3], Schmidt proved that $\mathscr{B}\mathscr{A}_{m,n}$ is a *winning* (a property stronger than thickness, cf. [S2, D2]) subset of $M_{m,n}(\mathbb{R})$. It is not clear to the author whether Schmidt's methods can be modified to allow treatment of inhomogeneous problems. It seems natural to conjecture that $\widehat{\mathscr{B}\mathscr{A}}_{m,n}$ is a wining subset of $\tilde{M}_{m,n}(\mathbb{R})$, and, moreover, that $\mathscr{B}\mathscr{A}_{m,n}(\mathbf{b})$ is a winning subset of $M_{m,n}(\mathbb{R})$ for every $\mathbf{b} \in \mathbb{R}^m$. This seems to be an interesting and challenging problem in metric number theory.

ACKNOWLEDGMENTS

The author is grateful to Professor G. A. Margulis for his interest in this problem, and to the referee for useful comments and suggestions.

REFERENCES

[Bou] N. Bourbaki, "Éléments de mathématique. Livre VI. Intégration," Chaps. 7, 8, Hermann, Paris, 1963.

[BM] J. Brezin and C. C. Moore, Flows on homogeneous spaces: A new look, *Amer. J. Math.* **103** (1981), 571–613.

[C] J. W. S. Cassels, "An Introduction to Diophantine Approximation," Cambridge Tracts in Math., Vol. 45, Cambridge Univ. Press, Cambridge, 1957.

[CG] L. G. Corwin and F. P. Greenleaf, "Representations of Nilpotent Lie Groups and Their Applications, Part I," Cambridge Stud. Adv. Math., Vol. 18, Cambridge Univ. Press, Cambridge, 1990.

[D1] S. G. Dani, Divergent trajectories of flows on homogeneous spaces and Diophantine approximation, *J. Reine Angew. Math.* **359** (1985), 55–89.

[D2] S. G. Dani, On badly approximable numbers, Schmidt games and bounded orbits of flows, *in* "Proceedings of the Number Theory and Dynamical Systems Conference, York, 1987," pp. 69–86, Cambridge Univ. Press, Cambridge/New York, 1990.

[Di] H. Dickinson, A remark on the Jarnik–Besicovitch theorem, *Glasgow Math. J.* **39** (1997), 233–236.

[Do1] M. M. Dodson, Geometric and probabilistic ideas in metric Diophantine approximation, *Russian Math. Surveys* **48** (1993), 73–102.

[Do2] M. M. Dodson, A note on metric inhomogeneous Diophantine approximation, *J. Austral. Math. Soc.* **62** (1997), 175–185.

[K1] D. Kleinbock, Flows on homogeneous spaces and Diophantine properties of matrices, *Duke Math. J.* **95** (1998), 107–124.

[K2] D. Kleinbock, Nondense orbits of flows on homogeneous spaces, *Ergodic Theory Dynam. Systems* **18** (1998), 373–396.

[K3] D. Kleinbock, "Nondense Orbits of Nonquasiunipotent Flows and Applications to Diophantine Approximation," Ph.D. thesis, Yale University, 1996.

[KM1] D. Kleinbock and G. A. Margulis, Bounded orbits of nonquasiunipotent flows on homogeneous spaces, *Amer. Math. Soc. Transl.* **171** (1996), 141–172.

[KM2] D. Kleinbock and G. A. Margulis, Flows on homogeneous spaces and Diophantine approximation on manifolds, *Ann. Math.* **148** (1998), 339–360.

[KM3] D. Kleinbock and G. A. Margulis, Logarithm laws for flows on homogeneous spaces, *Invent. Math.*, in press.

[L] J. Levesley, A general inhomogeneous Jarnik–Besicovitch theorem, *J. Number Theory* **71** (1998), 65–80.

[Ma] G. A. Margulis, Dynamical and ergodic properties of subgroup actions on homogeneous spaces with applications to number theory, *in* "Proceedings of the International Congress of Mathematicians, Kyoto, 1990," Vol. I, pp. 193–215, Math. Soc. Japan, Tokyo, 1991.

[Mc] C. McMullen, Area and Hausdorff dimension of Julia sets of entire functions, *Trans. Amer. Math. Soc.* **300** (1987), 329–342.

[Mo] C. C. Moore, Ergodicity of flows on homogeneous spaces, *Amer. J. Math.* **88** (1966), 154–178.

[S1] W. M. Schmidt, Metrical theorems on fractional parts of sequences, *Trans. Amer. Math. Soc.* **110** (1964), 493–518.

[S2] W. M. Schmidt, On badly approximable numbers and certain games, *Trans. Amer. Math. Soc.* **123** (1966), 178–199.

[S3] W. M. Schmidt, Badly approximable systems of linear forms, *J. Number Theory* **1** (1969), 139–154.

[U] M. Urbanski, The Hausdorff dimension of the set of points with nondense orbit under a hyperbolic dynamical system, *Nonlinearity* **2** (1991), 385–397.