# DIRICHLET'S THEOREM ON DIOPHANTINE APPROXIMATION AND HOMOGENEOUS FLOWS

DMITRY KLEINBOCK AND BARAK WEISS

ABSTRACT. Given an $m \times n$ real matrix $Y$, an unbounded set $\mathcal{T}$ of parameters $\mathbf{t} = (t_1, \ldots, t_{m+n}) \in \mathbb{R}_+^{m+n}$ with $\sum_{i=1}^m t_i = \sum_{j=1}^n t_{m+j}$ and $0 < \varepsilon \le 1$, we say that Dirichlet's Theorem can be $\varepsilon$-improved for $Y$ along $\mathcal{T}$ if for every sufficiently large $\mathbf{t} \in \mathcal{T}$ there are nonzero $\mathbf{q} \in \mathbb{Z}^n$ and $\mathbf{p} \in \mathbb{Z}^m$ such that

$$\begin{cases} |Y_i\mathbf{q} - p_i| < \varepsilon e^{-t_i}, & i = 1, \ldots, m \\ |q_j| < \varepsilon e^{t_{m+j}}, & j = 1, \ldots, n \end{cases}$$

(here $Y_1, \ldots, Y_m$ are rows of $Y$). We show that for any $\varepsilon < 1$ and any $\mathcal{T}$ 'drifting away from walls', see (1.8), Dirichlet's Theorem cannot be $\varepsilon$-improved along $\mathcal{T}$ for Lebesgue almost every $Y$. In the case $m = 1$ we also show that for a large class of measures $\mu$ (introduced in [KLW]) there is $\varepsilon_0 > 0$ such that for any unbounded $\mathcal{T}$, any $\varepsilon < \varepsilon_0$, and for $\mu$-almost every $Y$, Dirichlet's Theorem cannot be $\varepsilon$-improved along $\mathcal{T}$. These measures include natural measures on sufficiently regular smooth manifolds and fractals.

Our results extend those of several authors beginning with the work of Davenport and Schmidt done in late 1960s. The proofs rely on a translation of the problem into a dynamical one regarding the action of a diagonal semigroup on the space $\mathrm{SL}_{m+n}(\mathbb{R})/\mathrm{SL}_{m+n}(\mathbb{Z})$.

## 1. INTRODUCTION

Let $m, n$ be positive integers, and denote by $M_{m,n}$ the space of $m \times n$ matrices with real entries. Dirichlet's Theorem (hereafter abbreviated by 'DT') on simultaneous diophantine approximation states that for any $Y \in M_{m,n}$ (viewed as a system of $m$ linear forms in $n$ variables) and for any $t > 0$ there exist $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{Z}^n \smallsetminus \{0\}$ and $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{Z}^m$ satisfying the following system of inequalities:

$$\|Y\mathbf{q} - \mathbf{p}\| < e^{-t/m} \quad \text{and} \quad \|\mathbf{q}\| \le e^{t/n}. \tag{1.1}$$

Here and hereafter, unless otherwise specified, $\|\cdot\|$ stands for the norm on $\mathbb{R}^k$ given by $\|\mathbf{x}\| = \max_{1 \le i \le k} |x_i|$. See [Sc2] for a discussion of

_Date_: September 2007.

two ways of proving this theorem, due to Dirichlet and Minkowski respectively.

Given $Y$ as above and positive $\varepsilon < 1$, we will say that DT *can be $\varepsilon$-improved* for $Y$, and write $Y \in \mathrm{DI}_\varepsilon(m, n)$, or $Y \in \mathrm{DI}_\varepsilon$ when the dimensionality is clear from the context, if for every sufficiently large $t$ one can find $\mathbf{q} \in \mathbb{Z}^n \smallsetminus \{0\}$ and $\mathbf{p} \in \mathbb{Z}^m$ with

$$\|Y\mathbf{q} - \mathbf{p}\| < \varepsilon e^{-t/m} \quad \text{and} \quad \|\mathbf{q}\| < \varepsilon e^{t/n} \,, \tag{1.2}$$

that is, satisfy (1.1) with the right hand side terms multiplied by $\varepsilon$ (for convenience we will also replace $\leq$ in the second inequality by $<$). Also note that $Y$ is called *singular* if $Y \in \mathrm{DI}_\varepsilon$ for any $\varepsilon > 0$ (in other words, DT can be 'infinitely improved' for $Y$).

The two papers [DS1, DS2] by H. Davenport and W. Schmidt give a few basic results concerning the properties defined above. For example, the following is proved there:

**Theorem 1.1** ([DS2]). *For any[1] $m, n \in \mathbb{N}$ and any $\varepsilon < 1$, the sets $\mathrm{DI}_\varepsilon(m, n)$ have Lebesgue measure zero.*

In other words, $\lambda$-generic systems of linear forms do not allow any improvement to DT ($\lambda$ will denote Lebesgue measure throughout the paper).

Another question raised by Davenport and Schmidt concerns the possibility of improving DT for matrices with some functional relationship between entries. Specifically they considered row matrices

$$\mathbf{f}(x) = \begin{pmatrix} x & x^2 \end{pmatrix} \in M_{1,2} \,, \tag{1.3}$$

and proved

**Theorem 1.2** ([DS1]). *For any $\varepsilon < 4^{-1/3}$, the set of $x \in \mathbb{R}$ for which $\mathbf{f}(x) \in \mathrm{DI}_\varepsilon(1, 2)$ has zero Lebesgue measure.*

In other words, generic matrices of the form (1.3) do not allow a sufficiently drastic improvement to DT. This result was subsequently extended by R. Baker, Y. Bugeaud, and others. Namely, for some other smooth submanifolds of $\mathbb{R}^n$ they exhibited constants $\varepsilon_0$ such that almost no points on these submanifolds (viewed as row or column matrices) are in $\mathrm{DI}_\varepsilon$ for $\varepsilon < \varepsilon_0$. We will discuss the history in more detail in §4.

In the present paper we significantly generalize Theorems 1.1 and 1.2 by using a homogeneous dynamics approach and following a theme

---

[1] Even though the results of [DS2] are stated for matrices with one row or one column, i.e. for a vector or a single linear form, the proofs can be generalized to the setting of systems of linear forms.

developed in the paper [KW2] which dealt with infinite improvement to DT, that is, with singular systems. Namely, we study the subject of improvement of the multiplicative version of DT. In what follows we fix $m, n \in \mathbb{N}$ and let $k = m+n$. Let us denote by $\mathfrak{a}^+$ the set of $k$-tuples $\mathbf{t} = (t_1, \ldots, t_k) \in \mathbb{R}^k$ such that

$$t_1, \ldots, t_k > 0 \quad \text{and} \quad \sum_{i=1}^m t_i = \sum_{j=1}^n t_{m+j} \, . \tag{1.4}$$

It is not hard to see that both Dirichlet's and Minkowski's proofs of DT easily yield the following statement:

**Theorem 1.3.** *For any system of $m$ linear forms $Y_1, \ldots, Y_m$ (rows of $Y \in M_{m,n}$) in $n$ variables and for any $\mathbf{t} \in \mathfrak{a}^+$ there exist solutions $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{Z}^n \smallsetminus \{0\}$ and $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{Z}^m$ of*

$$\begin{cases} |Y_i \mathbf{q} - p_i| < e^{-t_i} \, , & i = 1, \ldots, m \\ |q_j| \leq e^{t_{m+j}} \, , & j = 1, \ldots, n \, . \end{cases} \tag{1.5}$$

Now, given an unbounded subset $\mathcal{T}$ of $\mathfrak{a}^+$ and positive $\varepsilon < 1$, say that *DT can be $\varepsilon$-improved for $Y$ along $\mathcal{T}$*, or $Y \in \mathrm{DI}_\varepsilon(\mathcal{T})$, if there is $t_0$ such that for every $\mathbf{t} = (t_1, \ldots, t_k) \in \mathcal{T}$ with $\|\mathbf{t}\| > t_0$, the inequalities

$$\begin{cases} |Y_i \mathbf{q} - p_i| < \varepsilon e^{-t_i} \, , & i = 1, \ldots, m \\ |q_j| < \varepsilon e^{t_{m+j}} \, , & j = 1, \ldots, n \, . \end{cases} \tag{1.6}$$

i.e., (1.5) with the right hand side terms multiplied by $\varepsilon$, have nontrivial integer solutions. Clearly $\mathrm{DI}_\varepsilon = \mathrm{DI}_\varepsilon(\mathcal{R})$ where

$$\mathcal{R} \stackrel{\text{def}}{=} \left\{ \left( \tfrac{t}{m}, \ldots, \tfrac{t}{m}, \tfrac{t}{n}, \ldots, \tfrac{t}{n} \right) : t > 0 \right\} \tag{1.7}$$

is the 'central ray' in $\mathfrak{a}^+$. Also say that $Y$ is *singular along $\mathcal{T}$* if it belongs to $\mathrm{DI}_\varepsilon(\mathcal{T})$ for every positive $\varepsilon$. The latter definition was introduced in [KW2] for $\mathcal{T}$ contained in an arbitrary ray in $\mathfrak{a}^+$ emanating from the origin (the setup of *diophantine approximation with weights*) in the special case $n = 1$.

To state our first result we need one more definition. Denote

$$\lfloor \mathbf{t} \rfloor \stackrel{\text{def}}{=} \min_{i=1,\ldots,k} t_i \, ,$$

and say that $\mathcal{T} \subset \mathfrak{a}^+$ *drifts away from walls* if

$$\forall \, s > 0 \quad \exists \, \mathbf{t} \in \mathcal{T} \text{ with } \lfloor \mathbf{t} \rfloor > s \, . \tag{1.8}$$

In other words, the distance from $\mathbf{t} \in \mathcal{T}$ to the boundary of $\mathfrak{a}^+$ is unbounded (hence, in partucular, $\mathcal{T}$ itself is unbounded).

Using Lebesgue's Density Theorem and an elementary argument which can be found e.g. in [Ca, Chapter V, §7] and dates back to Khintchine, one can show that for any $m, n$ and $\mathcal{T} \subset \mathfrak{a}^+$ drifting away from walls, $\mathrm{DI}_\varepsilon(\mathcal{T})$ has Lebesgue measure zero as long as $\varepsilon < 1/2$. However proving the same for any $\varepsilon < 1$ is more difficult, even in the case $\mathcal{T} = \mathcal{R}$ considered by Davenport and Schmidt. In [DS2] they derived this fact from the density of generic trajectories of certain one-parameter subgroups on the space $G/\Gamma$, where

$$G = \mathrm{SL}_k(\mathbb{R}) \text{ and } \Gamma = \mathrm{SL}_k(\mathbb{Z}). \qquad (1.9)$$

In §2.2 we show how the fact that $\mathrm{DI}_\varepsilon(\mathcal{T})$ has Lebesgue measure zero for any $\varepsilon < 1$ and any unbounded $\mathcal{T} \subset \mathcal{R}$ follows from mixing of the $G$-action on $G/\Gamma$. We also explain why mixing is not enough to obtain such a result for an arbitrary $\mathcal{T} \subset \mathfrak{a}^+$ drifting away from walls, even contained in a single 'non-central' ray, and prove the following multiplicative analogue of Theorem 1.1:

**Theorem 1.4.** *For any $\mathcal{T} \subset \mathfrak{a}^+$ drifting away from walls and any $\varepsilon < 1$, the set $\mathrm{DI}_\varepsilon(\mathcal{T})$ has Lebesgue measure zero.*

In other words, given any $\mathcal{T}$ as above, DT in its generalized form (Theorem 1.3) cannot be improved for generic systems of linear forms. Theorem 1.4 is derived from the equidistribution of translates of certain measures on $G/\Gamma$ (Theorem 2.2). Our proof of this theorem, described in §2, is a modification of arguments used in analyzing epimorphic groups [We, SW]. It relies on S. G. Dani's classification of measures invariant under horospherical subgroups, and the 'linearization method' developed by Dani, G. A. Margulis, M. Ratner, J. Smillie, N. Shah and others. See also [KM3] for an alternative approach suggested to the authors by Margulis.

Next, building on the approach of [KLW, KW2], we generalize Theorem 1.2, namely, consider measures other than Lebesgue. Here we will restrict ourselves to measures $\mu$ on $\mathbb{R}^n \cong M_{1,n}$ and study diophantine properties of $\mu$-almost all $\mathbf{y} \in \mathbb{R}^n$ interpreted as row vectors (linear forms); that is, we put $m = 1$ and $k = n + 1$. The dual case of simultaneous approximation ($n = 1$, which was the setup of [KLW] and [KW2]) can be treated along the same lines.

All measures on Euclidean spaces will be assumed to be Radon (locally finite regular Borel). Generalizing the setup of Theorem 1.2, we will consider measures on $\mathbb{R}^n$ of the form $\mathbf{f}_*\nu$, where $\nu$ is a measure on $\mathbb{R}^d$ and $\mathbf{f}$ a map from $\mathbb{R}^d$ to $\mathbb{R}^n$. Our assumptions on $\mathbf{f}$ and $\nu$ rely on definitions of:

- measures $\nu$ which are *D-Federer* on open subsets $U$ of $\mathbb{R}^d$, and
- functions $f : U \to \mathbb{R}$ which are $(C, \alpha)$-*good on $U$ w.r.t.* $\nu$

(here $D, C, \alpha$ are positive constants). We postpone the precise definitions, introduced in [KM2, KLW], until §3.2.

Now given a measure $\nu$ on $\mathbb{R}^d$, an open $U \subset \mathbb{R}^d$ with $\nu(U) > 0$ and a map $\mathbf{f} : \mathbb{R}^d \to \mathbb{R}^n$, say that a pair $(\mathbf{f}, \nu)$ is

- $(C, \alpha)$-*good on $U$* if any linear combination of $1, f_1, \ldots, f_n$ is $(C, \alpha)$-good on $U$ with respect to $\nu$;
- *nonplanar on $U$* if for any ball $B \subset U$ centered in $\operatorname{supp} \nu$, the restrictions of $1, f_1, \ldots, f_n$ to $B \cap \operatorname{supp} \nu$ are linearly independent over $\mathbb{R}$; in other words, if $\mathbf{f}(B \cap \operatorname{supp} \nu)$ is not contained in any proper affine subspace of $\mathbb{R}^n$;
- $(C, \alpha)$-*good* (resp., *nonplanar*) if for $\nu$-a.e. $\mathbf{x}$ there exists a neighborhood $U$ of $\mathbf{x}$ such that $\nu$ is $(C, \alpha)$-good (resp., nonplanar) on $U$.

Similarly, we will say that $\nu$ is *D-Federer* if for $\nu$-a.e. $\mathbf{x} \in \mathbb{R}^d$ there exists a neighborhood $U$ of $\mathbf{x}$ such that $\nu$ is $D$-Federer on $U$.

In §3 we prove

**Theorem 1.5.** *For any $d, n \in \mathbb{N}$ and $C, \alpha, D > 0$ there exists $\varepsilon_0 = \varepsilon_0(d, n, C, \alpha, D)$ with the following property. Let $\nu$ be a measure on $\mathbb{R}^d$ and $\mathbf{f}$ a continuous map from open $U \subset \mathbb{R}^d$ with $\nu(U) > 0$ to $\mathbb{R}^n$. Assume that $\nu$ is $D$-Federer, and $(\mathbf{f}, \nu)$ is $(C, \alpha)$-good and nonplanar. Then for any $\varepsilon < \varepsilon_0$*

$$\mathbf{f}_* \nu \big( \mathrm{DI}_\varepsilon(\mathcal{T}) \big) = 0 \ \text{ for any unbounded } \mathcal{T} \subset \mathfrak{a}^+. \qquad (1.10)$$

One can easily check, see §4.1, that $(\mathbf{f}, \lambda)$ as in (1.3) satisfies the assumptions of the above theorem. Moreover, Theorem 1.5 generalizes all known results on improvement of DT on manifolds, and is applicable to many other examples of interesting measures, see §4.3 for more examples. The proof is based on so-called 'quantitative nondivergence' technique, a generalization of non-divergence of unipotent flows on homogeneous spaces first established by Margulis in early 1970s.

Note that $\varepsilon_0$ in the above theorem can be explicitly computed in terms of $d, n, C, \alpha, D$, although the value that can be obtained seems to be far from optimal. In particular they are not as good as the values of $\varepsilon_0$ obtained in [DS1, DRV, Bu]. On the other hand the setup of Theorem 1.5 is much more general than that of the aforementioned papers, as illustrated by examples considered in §4.

We remark that the above theorem does not require to impose an additional condition (1.8) on $\mathcal{T}$, unlike Theorem 1.4 which is in general false if (1.8) is not assumed. Counterexamples are discussed in §4.4.

## 2. Theorem 1.4 and equidistribution

2.1. **The correspondence.** Let $G$ and $\Gamma$ be as in (1.9), and denote by $\pi : G \to G/\Gamma$ the quotient map. $G$ acts on $G/\Gamma$ by left translations via the rule $g\pi(h) = \pi(gh)$, $g, h \in G$. Define

$$\tau(Y) \stackrel{\mathrm{def}}{=} \begin{pmatrix} I_m & Y \\ 0 & I_n \end{pmatrix}, \quad \bar\tau \stackrel{\mathrm{def}}{=} \pi \circ \tau,$$

where $I_\ell$ stands for the $\ell \times \ell$ identity matrix. Since $\Gamma$ is the stabilizer of $\mathbb{Z}^k$ under the action of $G$ on the set of lattices in $\mathbb{R}^k$, $G/\Gamma$ can be identified with $G\mathbb{Z}^k$, that is, with the set of all unimodular lattices in $\mathbb{R}^k$. To highlight the relevance of the objects defined above to the diophantine problems considered in the introduction, note that

$$\bar\tau(Y) = \left\{ \begin{pmatrix} Y\mathbf{q} - \mathbf{p} \\ \mathbf{q} \end{pmatrix} : \mathbf{p} \in \mathbb{Z}^m, \ \mathbf{q} \in \mathbb{Z}^n \right\}. \tag{2.1}$$

Now for $\varepsilon > 0$ let

$$K_\varepsilon \stackrel{\mathrm{def}}{=} \pi\big(\big\{ g \in G : \|g\mathbf{v}\| \geq \varepsilon \quad \forall\, \mathbf{v} \in \mathbb{Z}^k \smallsetminus \{0\} \big\}\big), \tag{2.2}$$

i.e., $K_\varepsilon$ is the collection of all unimodular lattices in $\mathbb{R}^k$ which contain no nonzero vector of norm smaller than $\varepsilon$. By Mahler's compactness criterion (see e.g. [Ra, Chapter 10]), each $K_\varepsilon$ is compact, and for each compact $K \subset G/\Gamma$ there is $\varepsilon > 0$ such that $K \subset K_\varepsilon$. Note also that $K_\varepsilon$ is empty if $\varepsilon > 1$ by Minkowski's Lemma, see [Sc2], and has nonempty interior if $\varepsilon < 1$.

Now to any $\mathbf{t} \in \mathfrak{a}^+$ let us associate the diagonal matrix

$$g_\mathbf{t} \stackrel{\mathrm{def}}{=} \mathrm{diag}(e^{t_1}, \ldots, e^{t_m}, e^{-t_{m+1}}, \ldots, e^{-t_k}) \in G.$$

Then, using (2.1), it is straightforward to see that the system (1.6) has a nonzero integer solution if and only if $g_\mathbf{t}\bar\tau(Y) \notin K_\varepsilon$. We therefore arrive at

**Proposition 2.1.** *For $Y \in M_{m,n}$, $0 < \varepsilon < 1$ and unbounded $\mathcal{T} \subset \mathfrak{a}^+$, one has $Y \in \mathrm{DI}_\varepsilon(\mathcal{T})$ if and only if $g_\mathbf{t}\bar\tau(Y)$ is outside of $K_\varepsilon$ for all $\mathbf{t} \in \mathcal{T}$ with large enough norm. Equivalently,*

$$\mathrm{DI}_\varepsilon(\mathcal{T}) = \bigcup_{t>0} \bigcap_{\mathbf{t} \in \mathcal{T}, \|\mathbf{t}\| > t} \{Y : g_\mathbf{t}\bar\tau(Y) \notin K_\varepsilon\}. \tag{2.3}$$

In particular, $Y$ is singular along $\mathcal{T}$ if and only if the trajectory $\{g_{\mathbf{t}}\bar{\tau}(Y) : \mathbf{t} \in \mathcal{T}\}$ is divergent (i.e. eventually leaves $K_\varepsilon$ for any $\varepsilon > 0$). The latter observation was first made in [Da1, Proposition 2.12] for $\mathcal{T} = \mathcal{R}$, and then in [Kl1, Theorem 7.4] for an arbitrary ray in $\mathfrak{a}^+$.

## 2.2. **Lebesgue measure and uniform distribution.**

We recall that one of the goals of this paper is to show that, whenever $\varepsilon < 1$ and $\mathcal{T} \subset \mathfrak{a}^+$ drifts away from walls, $\lambda$-a.e. $Y \in M_{m,n}$ does not belong to $\mathrm{DI}_\varepsilon(\mathcal{T})$ (Theorem 1.4). The latter, in view of the above proposition, amounts to showing that the set $\{\mathbf{t} \in \mathcal{T} : g_{\mathbf{t}}\bar{\tau}(Y) \in K_\varepsilon\}$ is unbounded for $\lambda$-a.e. $Y \in M_{m,n}$. In this section we will prove a stronger statement. Here and hereafter 'vol' stands for the $G$-invariant probability measure on $G/\Gamma$.

**Theorem 2.2.** *Let a continuous compactly supported function $\varphi$ on $G/\Gamma$ and bounded $B \subset M_{m,n}$ with positive Lebesgue measure be given. Then for any compact subset $L$ of $G/\Gamma$ and any $\delta > 0$ there exists $T > 0$ such that*

$$\left| \frac{1}{\lambda(B)} \int_B \varphi\big(g_{\mathbf{t}}\tau(Y)z\big) \, d\lambda(Y) - \int_{G/\Gamma} \varphi \, d\mathrm{vol} \right| < \delta \qquad (2.4)$$

*for all $z \in L$ and $\mathbf{t} \in \mathfrak{a}^+$, $\lfloor \mathbf{t} \rfloor \geq T$.*

Let us denote by $\lambda_{z,B}$ the pushforward of $\frac{1}{\lambda(B)}\lambda|_B$ to $G/\Gamma$ by the map $Y \mapsto \tau(Y)z$. That is, for a Borel subset $A$ of $G/\Gamma$, let

$$\lambda_{z,B}(A) \overset{\mathrm{def}}{=} \frac{\lambda\big(\{Y \in B : \tau(Y)z \in A\}\big)}{\lambda(B)}.$$

The above theorem asserts that $g_{\mathbf{t}}$-translates of $\lambda_{z,B}$ weak-$*$ converge to vol as $\lfloor \mathbf{t} \rfloor \to \infty$, and the convergence is uniform in $z$ when the latter is restricted to a compact subset of $G/\Gamma$. By taking $z$ to be the standard lattice $\mathbb{Z}^k \in G/\Gamma$ and approximating $K_\varepsilon$ (which has nonempty interior and boundary of measure zero for any $\varepsilon < 1$) by continuous functions on $G/\Gamma$, from Theorem 2.2 one obtains that

$$\lambda\big(\{Y \in B : g_{\mathbf{t}}\bar{\tau}(Y) \in K_\varepsilon\}\big) \to \lambda(B)\mathrm{vol}(K_\varepsilon) \quad \text{as } \lfloor \mathbf{t} \rfloor \to \infty.$$

This immediately rules out the existence of $\mathcal{T} \subset \mathfrak{a}^+$ drifting away from walls and $B \subset M_{m,n}$ of positive Lebesgue measure such that the set $\{\mathbf{t} \in \mathcal{T} : g_{\mathbf{t}}\bar{\tau}(Y) \in K_\varepsilon\}$ is bounded for any $Y \in B$. Thus Theorem 1.4 follows from Theorem 2.2.

Note that the conclusion of Theorem 2.2 is not new for $\mathcal{T}$ being a subset of the 'central ray' $\mathcal{R}$ defined in (1.7). Indeed, one can immediately see that the group

$$H \stackrel{\text{def}}{=} \tau(M_{m,n}) \tag{2.5}$$

is *expanding horospherical* (see e.g. [KSS, Chapter 1] for the definition) with respect to $g_{\mathbf{t}}$ for any $\mathbf{t} \in \mathcal{R}$. In this case Theorem 2.2 coincides with [KM1, Proposition 2.2.1], where it was deduced from mixing of the $G$-action on $G/\Gamma$, using an argument dating back to the Ph.D. Thesis of Margulis [Ma2]. On the other hand, $H$ is strictly contained in the expanding horospherical subgroup relative to $g_{\mathbf{t}}$ for any $\mathbf{t} \in \mathfrak{a}^+ \smallsetminus \mathcal{R}$, thus the aforementioned argument does not prove Theorem 2.2. In the remainder part of this section we show how to bypass this difficulty. See also [KM3] for an alternative proof.

2.3. **Expanding vectors along cones.** In this section we will discuss an important representation-theoretic property of the pair $(\mathfrak{a}^+, H)$.

**Lemma 2.3.** *Let $\rho : G \to \mathrm{GL}(V)$ be a representation (of algebraic groups) on a finite-dimensional normed vector space $V$ without nonzero fixed vectors, and let*

$$V^H = \{v \in V : \rho(h)v = v \ \forall h \in H\}.$$

*Then there are positive $c, c_0$ such that for any $v \in V^H$ and $\mathbf{t} \in \mathfrak{a}^+$ one has*

$$\|\rho(g_{\mathbf{t}})v\| \geq c_0 e^{c\lfloor \mathbf{t} \rfloor} \|v\|.$$

*Proof.* Let us denote by $A$ the group of positive diagonal matrices in $G$, let $\mathfrak{a}$ be its Lie algebra, and let $\mathfrak{h}$ be the Lie algebra of $H$. Since $A$ normalizes $H$, $V^H$ is a $\rho(A)$-invariant subspace, and we may write

$$V^H = \bigoplus_{\chi \in \Psi} V_\chi,$$

where $\Psi$ is a finite set of weights (linear functionals on $\mathfrak{a}$) and

$$V_\chi = \{v \in V : \rho(\exp Y)v = e^{\chi(Y)}v \ \forall Y \in \mathfrak{a}\}$$

is nonzero for any $\chi \in \Psi$. There is no loss of generality in assuming that $\|\cdot\|$ is the sup-norm with respect to a basis of $\rho(A)$-eigenvectors; thus it suffices to show that for any $\chi \in \Psi$, $\inf_{\mathbf{t} \in \mathfrak{a}^+} \chi(\mathbf{t})/\lfloor \mathbf{t} \rfloor$ is positive.
    Let

$$\mathcal{I} = \{1, \ldots, m\} \times \{m+1, \ldots, k\},$$

and for each $(i, j) \in \mathcal{I}$ let $G_0 = G_0(i, j)$ be the Lie subgroup of $G$ whose Lie algebra $\mathfrak{g}_0$ is generated by $E_{ij}, E_{ji}, F_{ij}$ (here $E_{rs}$ stands for the $k \times k$ matrix with 1 in position $(r, s)$ and 0 elsewhere, and $F_{rs}$

stands for $E_{rr} - E_{ss}$). Then $G_0$ is a copy of $\mathrm{SL}_2(\mathbb{R})$ contained in $G$, such that $\mathfrak{h} \cap \mathfrak{g}_0 = \mathrm{span}\, E_{ij}$ and $\mathfrak{a} \cap \mathfrak{g}_0 = \mathrm{span}\, F_{ij}$. Since all vectors in $V_\chi$ are fixed by $\rho\big(\tau(E_{ij})\big)$, by the representation theory of $\mathfrak{sl}_2$ (see e.g. [Se]) we have $\chi(F_{ij}) \geq 0$, and

$$\chi(F_{ij}) = 0 \iff \rho(g)v = v \;\forall\, g \in G_0,\; v \in V_\chi\,.$$

Since $V$ contains no nonzero vectors fixed by $\rho(G)$, and since the group generated by $\{G_0(i,j) : (i,j) \in \mathcal{I}\}$ is equal to $G$, there is at least one $(i_0, j_0)$ for which $\chi(F_{i_0 j_0}) > 0$. Since any $Y \in \overline{\mathfrak{a}^+}$ can be written as a linear combination of $\{F_{ij} : (i,j) \in \mathcal{I}\}$ with non-negative coefficients, and since $\mathbf{t}_0 = \mathbf{t} - \lfloor \mathbf{t} \rfloor (F_{i_0 j_0}) \in \overline{\mathfrak{a}^+}$, we have:

$$\chi(\mathbf{t}) = \chi(\mathbf{t}_0) + \chi(\lfloor \mathbf{t} \rfloor F_{i_0 j_0}) \geq \lfloor \mathbf{t} \rfloor \chi(F_{i_0 j_0})\,,$$

finishing the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

To put this result in context, recall that a subgroup $L_1$ of an algebraic group $L_2$ is said to be *epimorphic* in $L_2$ if for any representation $\rho : L_2 \to \mathrm{GL}(V)$, any vector fixed by $\rho(L_1)$ is also fixed by $\rho(L_2)$. For example, in our present notation, $AH$ is epimorphic in $G$. The 'cone lemma' ([We, Lemma 1]) shows that if $TU$ is epimorphic in $L$, where $U$ is unipotent, $T$ is diagonalizable and normalizes $U$, and $L$ is generated by unipotents, then for any $\rho : L \to \mathrm{GL}(V)$ without nonzero fixed vectors there is a nonempty open cone $T^+$ in $T$ such that for any nonzero $v \in V^U$, $\rho(a)v \to \infty$ as $a \to \infty$ in $T^+$. The proof in [We] is non-constructive. The true meaning of Lemma 2.3 is a precise determination of a cone which works for all representations $\rho$ in the case $T = A,\, U = H$.

The next proposition is a consequence of Lemma 2.3.

**Proposition 2.4.** *Let $V$, $\rho$, $c$ be as in Lemma 2.3, and let $B$ be a neighborhood of $0$ in $M_{m,n}$. Then there exists $b > 0$ such that for any $v \in V$ and $\mathbf{t} \in \mathfrak{a}^+$ one has*

$$\sup_{Y \in B} \big\| \rho\big(g_{\mathbf{t}} \tau(Y)\big) v \big\| \geq b e^{c \lfloor \mathbf{t} \rfloor} \| v \|\,. \tag{2.6}$$

*Proof.* Denote by $p$ the $\rho(A)$-equivariant projection $V \to V^H$. By [Sh1, Lemma 5.1] there exists $c_1 > 0$ (dependent on $\rho$, $B$ and the choice of the norm) such that for any $v \in V$,

$$\sup_{Y \in B} \big\| p \circ \rho\big(\tau(Y)\big) v \big\| \geq c_1 \| v \|\,. \tag{2.7}$$

Also choose $c_2 > 0$ such that $\|v\| \geq c_2\|p(v)\|$ for all $v \in V$. Then for any $Y \in B$, $v \in V$ and $\mathbf{t} \in \mathfrak{a}^+$ one can write

$$\big\|\rho\big(g_{\mathbf{t}}\tau(Y)\big)v\big\| \geq c_2\big\|p \circ \rho\big(g_{\mathbf{t}}\tau(Y)\big)v\big\| = c_2\big\|\rho(g_{\mathbf{t}}) \circ p \circ \rho\big(\tau(Y)\big)v\big\|$$
$$\underset{\text{by Lemma 2.3}}{\geq} c_0c_2e^{c\lfloor\mathbf{t}\rfloor}\big\|p \circ \rho\big(\tau(Y)\big)v\big\|,$$

hence

$$\sup_{Y \in B}\big\|\rho\big(g_{\mathbf{t}}\tau(Y)\big)v\big\| \geq c_0c_2e^{c\lfloor\mathbf{t}\rfloor}\sup_{Y \in B}\big\|p \circ \rho\big(\tau(Y)\big)v\big\| \underset{\text{by (2.7)}}{\geq} c_0c_1c_2e^{c\lfloor\mathbf{t}\rfloor}\|v\|.$$

$\square$

2.4. **Recurrence to compact sets.** In order to establish the equidistribution of $g_{\mathbf{t}}$-translates of $\lambda_{z,B}$ as $\lfloor\mathbf{t}\rfloor \to \infty$, one needs to at least show the existence of one limit point (which is not guaranteed apriori since $G/\Gamma$ is not compact). In other words, there must exist a compact subset $K$ of $G/\Gamma$ such that $\lambda_{z,B}\big(g_{\mathbf{t}}^{-1}(K)\big)$ is big enough whenever $\lfloor\mathbf{t}\rfloor$ is large. We show in this section how to construct such a compact set using Proposition 2.4 and a theorem of Dani and Margulis.

Denote by $\mathfrak{g}$ the Lie algebra of $G$, let $V \overset{\text{def}}{=} \bigoplus_{j=1}^{\dim(G)-1}\bigwedge^j\mathfrak{g}$, and let $\rho : G \to \mathrm{GL}(V)$ be the representation obtained by acting on $V$ via the adjoint representation and its exterior powers. Note that $V$ has no nonzero $G$-fixed vectors since $G$ is simple. For any proper connected Lie subgroup $W$ of $G$ we will denote by $\mathbf{p}_W$ an associated vector in $V$.

We have the following result of Dani and Margulis, see [Sh1, Thm. 2.2] for a more general statement:

**Proposition 2.5.** *Let $G$, $\Gamma$ and $\pi : G \to G/\Gamma$ be as above. Then there exist finitely many closed subgroups $W_1, \ldots, W_\ell$ of $G$ such that $\pi(W_i)$ is compact and $\rho(\Gamma)\mathbf{p}_{W_i}$ is discrete for each $i \in \{1, \ldots, \ell\}$, and the following holds: for any positive $\alpha, \varepsilon$ there is a compact $K \subset G/\Gamma$ such that for any $g \in G$, $\mathbf{t} \in \mathfrak{a}$ and bounded convex open $B \subset M_{m,n}$, one of the following is satisfied:*

    (1) *There is $\gamma \in \Gamma$ and $i \in \{1, \ldots, \ell\}$ such that*

$$\sup_{Y \in B}\big\|\rho\big(g_{\mathbf{t}}\tau(Y)g\gamma\big)\mathbf{p}_{W_i}\big\| < \alpha.$$

    (2) $(g_{\mathbf{t}})_*\lambda_{\pi(g),B}(K) \geq 1 - \varepsilon.$

**Corollary 2.6.** *For any compact subset $L$ of $G/\Gamma$ and any $\varepsilon > 0$ there exists a compact $K \subset G/\Gamma$ with the following property: for any bounded convex open neighborhood $B$ of $0$ in $M_{m,n}$ there exists $T > 0$ such that $(g_{\mathbf{t}})_*\lambda_{z,B}(K) \geq 1 - \varepsilon$ whenever $z \in L$ and $\lfloor\mathbf{t}\rfloor \geq T$.*

*Proof.* Let $W_1, \ldots, W_\ell$ be as in Proposition 2.5, and for any compact $L \subset G/\Gamma$ consider

$$\delta(L) \stackrel{\text{def}}{=} \inf_{\pi(g) \in L, \ \gamma \in \Gamma, \ i=1,\ldots,\ell} \|\rho(g\gamma)\mathbf{p}_{W_i}\|.$$

It is positive since $L$ is compact and $\rho(\Gamma)\mathbf{p}_{W_i}$ is discrete. Proposition 2.4 then implies that for any neighborhood $B$ of 0 in $M_{m,n}$ there exist constants $b, c$ such that for any $g \in \pi^{-1}(L)$, $\gamma \in \Gamma$, $i = 1, \ldots, \ell$ and $\mathbf{t} \in \mathfrak{a}^+$, one has

$$\sup_{Y \in B} \|\rho\big(g_\mathbf{t}\tau(Y)\big)\rho(g\gamma)\mathbf{p}_{W_i}\| \geq be^{c\lfloor \mathbf{t}\rfloor}\delta(L). \qquad (2.8)$$

Now take an arbitrary $\varepsilon > 0$ and $\alpha = 1$, and choose $K$ according to Proposition 2.5. Then it follows from (2.8) that for any $B$ there exists $T$ such that whenever $\lfloor \mathbf{t}\rfloor \geq T$ and $\pi(g) \in L$, the second alternative of Proposition 2.5 must hold. $\qquad\square$

### 2.5. The linearization method.

As remarked in the introduction, our proof of Theorem 2.2 relies on the work of many mathematicians. Although we do not require Ratner's results on the classification of measures invariant under unipotent flows (the earlier results of Dani on horospherical subgroups are sufficient for us), we do use the linearization method developed by many authors following Ratner's work. These results are described in detail in [KSS]. Since our argument will be very close to arguments in [Sh1, SW] we will rely on the notation and results as stated in [SW], where additional references to the literature may be found.

Let $\mathcal{H}$ be the set of all closed connected subgroups $W$ of $G$ such that $W \cap \Gamma$ is a lattice in $W$, and the subgroup of $W$ generated by its one-parameter unipotent subgroups acts ergodically on $W/(W \cap \Gamma)$. This is a countable collection. For any $W \in \mathcal{H}$, we define

$$N(W, H) = \{g \in G : Hg \subset gW\}$$

$$S(W, H) = \bigcup \big\{N(W', H) : W' \in \mathcal{H}, W' \subset W, \dim W' < \dim W\big\}$$

$$N^*(W, H) = N(W, H) \smallsetminus S(W, H).$$

Recall that the subgroup $H$ of $G$ is horospherical. Dani [Da2] classified all the measures on $G/\Gamma$ invariant under the $H$-action. The following is a consequence of Dani's classification and ergodic decomposition:

**Proposition 2.7.** *Let $\mu$ be a finite $H$-invariant measure on $G/\Gamma$ which is not equal to* vol. *Then $\mu\big(\pi(N(W, H))\big) > 0$ for some $W \in \mathcal{H}$ which is a proper subgroup of $G$.*

For $W \in \mathcal{H}$ let $V_W$ be the span of $\rho(N(W, H))\mathbf{p}_W$ in $V$ and let $N_G^1(W) = \{g \in G : \rho(g)\mathbf{p}_W = \mathbf{p}_W\}$. Then the orbit $\rho(\Gamma)\mathbf{p}_W$ is discrete in $V$.

The next proposition uses the representation $\rho$ defined in §2.4 to detect orbits which stay close to $\pi\big(N(W, H)\big)$ for some $W \in \mathcal{H}$. The idea has a long history and is used in a similar context by Dani and Margulis in [DM].

**Proposition 2.8.** [Sh1, Theorem 4.1] *Given $W \in \mathcal{H}$ and $\varepsilon > 0$, for any compact $C \subset \pi(N^*(W, H))$ there exists a compact $D \subset V_W$ with the following property: for any neighborhood $D'$ of $D$ in $V_W$ there exists a neighborhood $C'$ of $C$ in $G/\Gamma$ such that for any $g \in G$, any $\mathbf{t} \in \mathfrak{a}^+$ and any bounded convex open $B \subset M_{m,n}$, one of the following holds:*

(1) *There is $\gamma \in \Gamma$ such that $\rho(g_\mathbf{t}\tau(B)g\gamma)\mathbf{p}_W \subset D'$.*
(2) $(g_\mathbf{t})_*\lambda_{\pi(g),B}(C') < \varepsilon$.

2.6. **Proof of Theorem 2.2.** Since finite linear combinations of indicator functions of balls are dense in $L^1(M_{m,n}, \lambda)$, there is no loss of generality in assuming that $B$ is an open ball in $M_{m,n}$. Take a sequence of points $z_n \in L$ and a sequence $\mathbf{t}_n \in \mathfrak{a}^+$ drifting away from walls. It follows from Corollary 2.6 that the sequence of translated measures $(g_{\mathbf{t}_n})_*\lambda_{z_n,B}$ is weak-$*$ precompact, that is, along a subsequence we have $g_{\mathbf{t}_n}\lambda_{z_n,B}$ to $\mu$, where $\mu$ is a Borel probability measure on $G/\Gamma$. Our goal is thus to show that $\mu = \text{vol}$.

The hypothesis about drifting away from walls implies that for any $h \in H$ we have $g_{\mathbf{t}_n}^{-1}hg_{\mathbf{t}_n} \to e$, where $e$ is the identity element in $H$. A simple computation (see [SW, Claim 3.2]) shows that $\mu$ is $H$-invariant. By Proposition 2.7, if $\mu \neq \text{vol}$, there is a proper subgroup $W \in \mathcal{H}$ such that $\mu\big(\pi(N(W, H))\big) > 0$. Making $W$ smaller if necessary we can assume that $\mu\big(\pi(N^*(W, H))\big) > 0$. Let $C \subset \pi\big(N^*(W, H)\big)$ be compact with $\mu(C) > 0$, and put $\varepsilon \overset{\text{def}}{=} \mu(C)/2$. Let $\widetilde{L} \subset G$ be a compact subset such that $\pi(\widetilde{L}) = L$, and let $g_n \in \pi^{-1}(z_n) \cap \widetilde{L}$. Applying Proposition 2.8, we find that there is a compact $D \subset V_W$ such that the following holds. For each $n$, let $D_{n+1} \subset D_n$ be a compact neighborhood of $D$ in $V_W$ such that $\bigcap_n D_n = D$. Then there is an open neighborhood $C_n$ of $C$ in $G/\Gamma$ such that one of the following holds:

(1) there is $v_n \in \rho(g_n\Gamma)\mathbf{p}_W$ such that $\rho(g_{\mathbf{t}_n}\tau(B))v_n \subset D_n$;
(2) $(g_{\mathbf{t}_n})_*\lambda_{z_n,B}(C_n) < \varepsilon$.

Since $g_{\mathbf{t}_n}\lambda_{z_n,B} \to \mu$, and the sets $C_n$ are neighborhoods of $C$, we find that $\lambda_{z_n,B}(g_{\mathbf{t}_n}^{-1}C_n) > \mu(C)/2 = \varepsilon$ for all sufficiently large $n$, so condition

(2) above does not hold. Therefore

$$\rho\big(g_{\mathbf{t}_n}\tau(B)\big)v_n \subset D_n \subset D_1\,,$$

a bounded subset of $V_W$. On the other hand, since $\widetilde{L}$ is compact and $\rho(\Gamma)\mathbf{p}_W$ is discrete, we have

$$\inf_n \|v_n\| > 0,$$

hence, by (2.6), $\sup_{Y\in B} \|\rho\big(g_{\mathbf{t}_n}\tau(Y)\big)v_n\| \to \infty$, a contradiction.          $\square$

## 3. THEOREM 1.5 AND QUANTITATIVE NONDIVERGENCE

3.1. **A sufficient condition.** The second goal of this paper is to show that sets $\mathrm{DI}_\varepsilon(\mathcal{T})$ are null with respect to certain measures $\mu$ on $M_{m,n}$ other than Lebesgue. We will use Proposition 2.1 to formulate a condition sufficient for having $\mu\big(\mathrm{DI}_\varepsilon(\mathcal{T})\big) = 0$ for fixed $\varepsilon > 0$ and all unbounded $\mathcal{T} \subset \mathfrak{a}^+$. Similarly to the setup of Theorem 1.5, we will consider measures $\mu$ of the form $F_*\nu$, where $\nu$ is a measure on $\mathbb{R}^d$ and $F$ a map from $\mathbb{R}^d$ to $M_{m,n}$.

**Proposition 3.1.** *Let a measure $\nu$ on $\mathbb{R}^d$, an open subset $U$ of $\mathbb{R}^d$, a map $F : U \to M_{m,n}$, and $0 < \varepsilon, c < 1$ be given. Suppose that for any ball $B \subset U$ there exists $s > 0$ such that the inequality*

$$\nu \left(\big\{\mathbf{x} \in B : g_{\mathbf{t}}\bar{\tau}\big(F(\mathbf{x})\big) \notin K_\varepsilon\big\}\right) \le c\nu(B) \qquad (3.1)$$

*holds for any $\mathbf{t} \in \mathfrak{a}^+$ with $\|\mathbf{t}\| \ge s$. Then $F_*\nu\big(\mathrm{DI}_\varepsilon(\mathcal{T})\big) = 0$ for any unbounded $\mathcal{T} \subset \mathfrak{a}^+$.*

*Proof.* Since $\mathcal{T}$ is unbounded, it follows from the assumption of the proposition that for any ball $B \subset U$ and any positive $t$ one has

$$\nu \bigg(\bigcap_{\mathbf{t}\in\mathcal{T},\,\|\mathbf{t}\|\ge t} \big\{\mathbf{x} \in B : g_{\mathbf{t}}\bar{\tau}\big(F(\mathbf{x})\big) \notin K_\varepsilon\big\}\bigg) \le c\nu(B)\,.$$

Therefore, by (2.3), $\nu\big(\{\mathbf{x} \in B : F(\mathbf{x}) \in \mathrm{DI}_\varepsilon(\mathcal{T})\}\big) \le c\nu(B)$. In view of a density theorem for Radon measures on Euclidean spaces [Mat, Corollary 2.14], this forces $F^{-1}\big(\mathrm{DI}_\varepsilon(\mathcal{T})\big)$ to have $\nu$-measure zero.          $\square$

3.2. **A quantitative nondivergence estimate.** The proof of Theorem 2.2 given in §2 relies on Corollary 2.6, which is a quantitative nondivergence estimate for translates of unipotent trajectories. Estimates of this kind have their origin in the proof by Margulis [Ma1] that orbits of unipotent flows do not diverge, see [KSS] for a historical account. During the last decade, starting from the paper [KM2], these techniques were transformed into a powerful method yielding measure estimates as in (3.1) for a certain broad class of measures $\nu$ and maps $F$.

To introduce these we need to elaborate on the definitions mentioned in the introduction.

If $B = B(\mathbf{x}, r)$ is a ball in $\mathbb{R}^d$ and $c > 0$, $cB$ will denote the ball $B(\mathbf{x}, cr)$. A measure $\nu$ on $\mathbb{R}^d$ is said to be *D-Federer* on an open $U \subset \mathbb{R}^d$ if for all balls $B$ centered at $\operatorname{supp}\nu$ with $3B \subset U$ one has $\nu(3B)/\nu(B) \leq D$.

If $\nu$ is a measure on $\mathbb{R}^d$, $B$ a subset of $\mathbb{R}^d$ with $\nu(B) > 0$, and $f$ a real-valued measurable function on $B$, we let

$$\|f\|_{\nu,B} \overset{\text{def}}{=} \sup_{\mathbf{x} \in B \cap \operatorname{supp}\nu} |f(\mathbf{x})|.$$

Given $C, \alpha > 0$, open $U \subset \mathbb{R}^d$ and a measure $\nu$ on $\mathbb{R}^d$, say that $f : U \to \mathbb{R}$ is *$(C, \alpha)$-good on $U$ with respect to $\nu$* if for any ball $B \subset U$ centered in $\operatorname{supp}\nu$ and any $\varepsilon > 0$ one has

$$\nu\big(\{y \in B : |f(y)| < \varepsilon\}\big) \leq C \left(\frac{\varepsilon}{\|f\|_{\nu,B}}\right)^{\alpha} \nu(B). \qquad (3.2)$$

We need to introduce some more notation in order to state a theorem from [KLW]. Let

$$\mathcal{W} \overset{\text{def}}{=} \text{ the set of proper nonzero rational subspaces of } \mathbb{R}^k.$$

From here until the end of this section, we let $\|\cdot\|$ stand for the Euclidean norm on $\mathbb{R}^k$, induced by the standard inner product $\langle\cdot,\cdot\rangle$, which we extend from $\mathbb{R}^k$ to its exterior algebra. For $V \in \mathcal{W}$ and $g \in G$, let

$$\ell_V(g) \overset{\text{def}}{=} \|g(\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_j)\|,$$

where $\{\mathbf{v}_1, \ldots, \mathbf{v}_j\}$ is a generating set for $\mathbb{Z}^k \cap V$; note that $\ell_V(g)$ does not depend on the choice of $\{\mathbf{v}_i\}$.

**Theorem 3.2** ([KLW], Theorem 4.3). *Given $d, k \in \mathbb{N}$ and positive constants $C, D, \alpha$, there exists $C_1 = C_1(d, k, C, \alpha, D) > 0$ with the following property. Suppose a measure $\nu$ on $\mathbb{R}^d$ is $D$-Federer on a ball $\widetilde{B}$ centered at $\operatorname{supp}\nu$, $0 < \rho \leq 1$, and $h$ is a continuous map $\widetilde{B} \to G$ such that for each $V \in \mathcal{W}$,*

*(i) the function $\ell_V \circ h$ is $(C, \alpha)$-good on $\widetilde{B}$ with respect to $\nu$,*

*and*

*(ii) $\|\ell_V \circ h\|_{\nu,B} \geq \rho$, where $B = 3^{-(k-1)}\widetilde{B}$.*

*Then for any $0 < \varepsilon \leq \rho$,*

$$\nu\big(\{\mathbf{x} \in B : \pi\big(h(\mathbf{x})\big) \notin K_\varepsilon\}\big) \leq C_1 \left(\frac{\varepsilon}{\rho}\right)^{\alpha} \nu(B).$$

3.3. **Checking (i) and (ii).** At this point we restrict ourselves to the setup of Theorem 1.5, that is consider measures on $\mathbb{R}^n \cong M_{1,n}$ of the form $\mathbf{f}_*\nu$, where $\nu$ is a measure on $\mathbb{R}^d$ and $\mathbf{f} = (f_1, \ldots, f_n)$ is a map from an open $U \subset \mathbb{R}^d$ with $\nu(U) > 0$ to $\mathbb{R}^n$. In order to combine Proposition 3.1 with Theorem 3.2, one needs to work with functions $\ell_V \circ h_{\mathbf{t}}$ for each $V \in \mathcal{W}$, where

$$h_{\mathbf{t}} \stackrel{\text{def}}{=} g_{\mathbf{t}} \circ \tau \circ \mathbf{f} \,, \tag{3.3}$$

and find conditions sufficient for the validity of (i) and (ii) of Theorem 3.2 for large enough $\mathbf{t} \in \mathcal{T}$.

The explicit computation that is reproduced below first appeared in [KM2]. Let $\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis of $\mathbb{R}^{n+1}$, and for

$$I = \{i_1, \ldots, i_j\} \subset \{0, \ldots, n\}, \quad i_1 < i_2 < \cdots < i_j \,, \tag{3.4}$$

let $\mathbf{e}_I \stackrel{\text{def}}{=} \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_j}$; then $\{\mathbf{e}_I \mid \#I = j\}$ is an orthonormal basis of $\bigwedge^j(\mathbb{R}^{n+1})$. Similarly, it will be convenient to put $\mathbf{t} = (t_0, t_1, \ldots, t_n) \in \mathfrak{a}^+$ where

$$\mathfrak{a}^+ = \left\{ (t_0, t_1, \ldots, t_n) \in \mathbb{R}^{n+1} : t_i > 0, \ t_0 = \sum_{i=1}^{n} t_i \right\} . \tag{3.5}$$

Then one immediately sees that for any $I$ as in (3.4),

$$\mathbf{e}_I \text{ is an eigenvector for } g_{\mathbf{t}} \text{ with eigenvalue } e^{t_I} \,, \tag{3.6}$$

where

$$t_I \stackrel{\text{def}}{=} \begin{cases} t_0 - \sum_{i \in I \smallsetminus \{0\}} t_i & \text{if } 0 \in I \\ -\sum_{i \in I} t_i & \text{otherwise.} \end{cases}$$

We remark that in view of (3.5), $e^{t_I}$ is not less than 1 for any $\mathbf{t} \in \mathfrak{a}^+$ and $0 \in I$. Moreover, let $\ell$ be such that $t_\ell = \max_{i=1,\ldots,n} t_i$. Then

$$t_I \geq t_\ell \geq \|\mathbf{t}\|/n \quad \forall I \text{ containing 0 and not containing } \ell \,. \tag{3.7}$$

Since the action of $\tau(\mathbf{y})$, where $\mathbf{y} \in \mathbb{R}^n$, leaves $\mathbf{e}_0$ invariant and sends $\mathbf{e}_i$, $i > 0$, to $\mathbf{e}_i + y_i \mathbf{e}_0$, one can write[2]

$$\tau(\mathbf{y})\mathbf{e}_I = \begin{cases} \mathbf{e}_I & \text{if } 0 \in I \\ \mathbf{e}_I + \sum_{i \in I} \pm y_i \, \mathbf{e}_{I \cup \{0\} \smallsetminus \{i\}} & \text{otherwise.} \end{cases} \tag{3.8}$$

Now take $V \in \mathcal{W}$, choose a generating set $\{\mathbf{v}_1, \ldots, \mathbf{v}_j\}$ for $\mathbb{Z}^k \cap V$, and expand $\mathbf{w} \stackrel{\text{def}}{=} \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_j$ with respect to the above basis by writing

---

[2]The choice of $+$ or $-$ in (3.8) depends on the parity of the number of elements of $I$ less than $i$ and is not important for our purposes. See however [Kl2] for a more precise computation.

$\mathbf{w} = \sum_{I \subset \{0,\dots,n\}, \#I=j} w_I \mathbf{e}_I \in \bigwedge^j(\mathbb{Z}^k) \smallsetminus \{0\}$. Then one has

$$\tau(\mathbf{y})\mathbf{w} = \sum_{0 \notin I} w_I \mathbf{e}_I + \sum_{0 \in I} \left( w_I + \sum_{i \notin I} \pm w_{I \cup \{i\} \smallsetminus \{0\}} y_i \right) \mathbf{e}_I. \qquad (3.9)$$

Here is an immediate implication of the above formula:

**Lemma 3.3.** *For any $\mathbf{w}$ and $\mathbf{t}$, the map $\mathbf{y} \mapsto g_{\mathbf{t}}\tau(\mathbf{y})\mathbf{w}$ is affine; in other words, for any $I \subset \{0,\dots,n\}$ the projection of $g_{\mathbf{t}}\tau(\mathbf{y})\mathbf{w}$ onto $\mathbf{e}_I$ has the form*

$$\langle g_{\mathbf{t}}\tau(\mathbf{y})\mathbf{w}, \mathbf{e}_I \rangle = c_0 + \sum_{i=1}^{n} c_i y_i \qquad (3.10)$$

*for some $c_0, c_1, \dots, c_n \in \mathbb{R}$.*

It is also clear from (3.9) that $\tau(\cdot)\mathbf{w} \equiv \mathbf{w}$ if the subspace $V$ represented by $\mathbf{w}$ contains $\mathbf{e}_0$ (in other words, if $w_I = 0$ whenever $0 \notin I$). In this case for any $I \ni 0$ all the coefficients $c_i$ in (3.10) with $i \geq 1$ are equal to 0, and $|c_0| = e^{t_I}|w_I| \geq 1$ as long as $w_I \neq 0$.

Let us now consider the complementary case.

**Lemma 3.4.** *Suppose that $\mathbf{w} \in \bigwedge^j(\mathbb{Z}^k)$, $1 \leq j \leq n$, is such that $w_J \neq 0$ for some $J \subset \{1,\dots,n\}$. Then for any $\mathbf{t} \in \mathfrak{a}^+$ there exists $I$ such that the absolute value of one of the coefficients $c_i$ in (3.10) is at least $e^{\|\mathbf{t}\|/n}$.*

*Proof.* Let $\ell$ be such that $t_\ell = \max_{i=1,\dots,n} t_i$. If $J$ as above contains $\ell$, take $I = J \cup \{0\} \smallsetminus \{\ell\}$. Then $J = I \cup \{\ell\} \smallsetminus \{0\}$, hence, by (3.9) and (3.6), the coefficient $c_\ell$ in the expression (3.10) for this $I$ is equal to $\pm e^{t_I} w_J$. The claim then follows from (3.7), since $|w_J| \geq 1$ and $I$ contains 0 and does not contain $\ell$.

If $J$ does not contain $\ell$, the argument is similar: choose any $i \in J$ and take $I = J \cup \{0\} \smallsetminus \{i\}$. Then $J = I \cup \{\imath\} \smallsetminus \{0\}$, hence the coefficient $c_i$ in (3.10) for this $I$ is equal to $\pm e^{t_I} w_J$. As before, $I$ contains 0 and does not contain $\ell$, so (3.7) applies again. $\square$

3.4. **Proof of Theorem 1.5.** The next theorem generalizes [KM2, Theorem 5.4]:

**Theorem 3.5.** *For any $d, n \in \mathbb{N}$ and any $C, \alpha, D > 0$ there exists $C_2 = C_2(d, n, C, \alpha, D)$ with the following property. Suppose a measure $\nu$ on $\mathbb{R}^d$ is $D$-Federer on a ball $\widetilde{B}$ centered at $\operatorname{supp}\nu$ and $\mathbf{f} : \tilde{B} \to \mathbb{R}^n$ is continuous. Assume that:*

    (1) *any linear combination of $1, f_1, \dots, f_n$ is $(C, \alpha)$-good on $\tilde{B}$ with respect to $\nu$;*

(2) *the restrictions of* $1, f_1, \ldots, f_n$ *to* $B \cap \operatorname{supp} \nu$, *where* $B = 3^{-n} \widetilde{B}$,
    *are linearly independent over* $\mathbb{R}$.

*Then there exists* $s > 0$ *such that for any* $\mathbf{t} \in \mathfrak{a}^+$ *with* $\|vt\| \geq s$ *and
any* $\varepsilon < 1$, *one has*

$$\nu\big(\{\mathbf{x} \in B : g_{\mathbf{t}} \bar{\tau}\big(\mathbf{f}(\mathbf{x})\big) \notin K_\varepsilon\}\big) \leq C_2 \varepsilon^\alpha \nu(B).$$

*Proof.* We will apply Theorem 3.2 with $h = h_{\mathbf{t}}$ as in (3.3). Take $V \in \mathcal{W}$
and, as before, represent it by $\mathbf{w} = \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_j$, where $\{\mathbf{v}_1, \ldots, \mathbf{v}_j\}$ is a
generating set for $\mathbb{Z}^k \cap V$. From Lemma 3.3 and assumption (1) above
it follows that for any $\mathbf{t}$, each coordinate of $h_{\mathbf{t}}(\cdot)\mathbf{w}$ is $(C, \alpha)$-good on $\tilde{B}$
with respect to $\nu$. Hence the same, with $C$ replaced by $\binom{n+1}{j}^{\alpha/2} C$, can
be said about $\|h_{\mathbf{t}}(\cdot)\mathbf{w}\| = \ell_V \circ h_{\mathbf{t}}$, see [KLW, Lemma 4.1]. This verifies
condition (i) of Theorem 3.2.

Now observe that assumption (2) implies the existence of $\delta > 0$
(depending on $B$) such that $\|c_0 + \sum_{i=1}^n c_i f_i\|_{\nu, B} \geq \delta$ for any $c_0, c_1, \ldots, c_n$
with $\max |c_i| \geq 1$. Using Lemma 3.4 and the remark preceding it,
we conclude that either $\|\ell_V \circ h_{\mathbf{t}}\|_{\nu, B} \geq 1$ (in the case $\mathbf{e}_0 \in V$) or
$\|\ell_V \circ h_{\mathbf{t}}\|_{\nu, B} \geq \delta e^{\|\mathbf{t}\|/n}$ (in the complementary case). So condition (ii) of
Theorem 3.2 holds with $\rho = 1$ whenever $\|\mathbf{t}\|$ is at least $s \overset{\text{def}}{=} -n \log \delta$. $\square$

We can now proceed with the

*Proof of Theorem 1.5.* It suffices to show that for $\nu$-a.e. $\mathbf{x}$ there exists
a ball $B$ centered at $\mathbf{x}$ such that

$$\nu\big(\{\mathbf{x} \in B : \mathbf{f}(\mathbf{x}) \in \mathrm{DI}_\varepsilon(\mathcal{T})\}\big) = 0. \tag{3.11}$$

Since $(\mathbf{f}, \nu)$ is $(C, \alpha)$-good and nonplanar, for $\nu$-a.e. $\mathbf{x}$ one can choose
$B$ centered at $\mathbf{x}$ such that $(\mathbf{f}, \nu)$ is nonplanar on $B$ and $(C, \alpha)$-good on
$\tilde{B} = 3^n B \subset U$, which implies that conditions (1) and (2) of Theorem
3.5 are satisfied. Then, combining Theorem 3.5 with Proposition 3.1,
one concludes that (3.11) holds whenever $C_2 \varepsilon^\alpha < 1$. $\square$

We remark that the constant $C_1$ from Theorem 3.2, and hence $C_2$
from Theorem 3.5 and $\varepsilon_0$ from Theorem 1.5, can be explicitly extimated
in terms of the input data of those theorems, see [KM2, BKM, KT, Kl2].
However we chose not to bother the reader with explicit computations,
the reason being that in the special cases previously considered in the
literature our method produces much weaker estimates. More on that
in the next section.

## 4. Examples and applications

4.1. **Polynomial maps.** A model example of functions which are $(C, \alpha)$-good with respect to Lebesgue measure is given by polynomials: it is shown in [KM2, Lemma 3.2] that any polynomial of degree $\ell$ is $(C, \alpha)$-good on $\mathbb{R}$ with respect to $\lambda$, where $\alpha = \frac{1}{\ell}$ and $C$ depends only on $\ell$. The same can be said about polynomials in $d \geq 1$ variables, with $\alpha = \frac{1}{d\ell}$ and $C$ depending on $\ell$ and $d$. Obviously $\lambda$ is $3^d$-Federer on $\mathbb{R}^d$. Thus, as a corollary of Theorem 1.5, we obtain the existence of $\varepsilon_1 = \varepsilon_1(n, d, \ell)$ such that whenever $\varepsilon < \varepsilon_1$, (1.10) holds for $\nu = \lambda$ and any polynomial map $\mathbf{f} = (f_1, \ldots, f_n)$ of degree $\ell$ in $d$ variables such that $1, f_1, \ldots, f_n$ are linearly independent over $\mathbb{R}$. This in particular applies to $\mathbf{f}(x) = (x, \ldots, x^n)$, a generalization of the setup of Theorem 1.2 considered by Baker in [Ba1] for $n = 3$ and then by Bugeaud for an arbitrary $n$. Note that it is proved in [Bu] that $(x, \ldots, x^n)$ is almost surely not in $\mathrm{DI}_\varepsilon$ for $\varepsilon < 1/8$. Our method, in comparison, shows that $(x, \ldots, x^n)$ is almost surely not in $\mathrm{DI}_\varepsilon(\mathcal{T})$ for any $\mathcal{T}$ drifting away from walls and $\varepsilon < 1/n^n(n + 1)^2 2^{n^2+n}$.

Improving these results to any $\varepsilon < 1$ is a natural and challenging problem. Recently [Sh2] the following was obtained:

**Theorem 4.1** (N.A. Shah). *Let* $\mathbf{f} : \mathbb{R} \to M_{1,2} \cong \mathbb{R}^2$ *be a nonplanar polynomial map. For an interval* $B \subset \mathbb{R}$ *and for* $z \in \mathrm{SL}_3(\mathbb{R})/\mathrm{SL}_3(\mathbb{Z})$, *let* $\nu_{z,B}$ *be the natural probability measure on* $\tau\big(\mathbf{f}(B)\big)z$, *that is the push-forward of* $\frac{1}{\lambda(B)}\lambda|_B$ *via the map* $x \mapsto \tau \circ \mathbf{f}(x)z$. *Then for any unbounded sequence* $\mathbf{t}_n \in \mathcal{R}$, *the sequence of translated measures* $(g_{\mathbf{t}_n})_* \nu_{z,B}$ *weak-\* converges to* vol.

The proof follows a similar strategy as our proof of Theorem 2.2, but is considerably more difficult. Repeating the argument of §2.2, one obtains:

**Corollary 4.2.** *For any* $\varepsilon < 1$ *and any nonplanar polynomial curve* $\mathbf{f} : \mathbb{R} \to M_{1,2}$, $\mathbf{f}(x)$ *does not belong to* $\mathrm{DI}_\varepsilon$ *for* $\lambda$-*a.e.* $x \in \mathbb{R}$.

Note also that Bugeaud's result mentioned above can be rephrased in terms of small values of integer polynomials at almost all real $x$. Theorem 1.5 produces a similar result for polynomials of $d$ variables. Namely, for some $c = c(d, n)$ and for $\lambda$-almost every $\mathbf{x} \in \mathbb{R}^d$ there are infinitely many integers $N$ for which there are no polynomials $P \in \mathbb{Z}[X_1, \ldots, X_d]$ of degree at most $n$ and height less than $N$ with $|P(\mathbf{x})| < cN^{-m}$, where $m$ is the dimension of the space of non-constant polynomials in $d$ variables of degree at most $n$.

4.2. **Nondegenerate maps.** Here is another situation in which $(C, \alpha)$-good functions arise. The following lemma is a strengthening of [KM2, Lemma 3.3]:

**Lemma 4.3.** *For any $d \in \mathbb{N}$ there exists $C_d > 0$ with the following property. Let $B$ be a cube in $\mathbb{R}^d$ (product of intervals of the same length), and let $f \in C^\ell(B)$, $\ell \in \mathbb{N}$, be such that for some positive constants $a_1, \ldots, a_d$ and $A_1, \ldots, A_d$ one has*

$$a_i \leq |\partial_i^\ell f(\mathbf{x})| \leq A_i \quad \forall\, \mathbf{x} \in B, \ i = 1, \ldots, d. \tag{4.1}$$

*Then*

$$\lambda\big(\{\mathbf{x} \in B : |f(\mathbf{x})| < \varepsilon\}\big) \leq \ell C_d \max_i \left(\frac{A_i}{a_i}\right)^{1/\ell} \left(\frac{\varepsilon}{\|f\|_{\lambda, B}}\right)^{1/d\ell} \lambda(B).$$

We remark that [KM2, Lemma 3.3] instead of (4.1) assumed

$$a \leq |\partial_i^\ell f(\mathbf{x})| \quad \forall\, \mathbf{x} \in B, \ i = 1, \ldots, d$$

and

$$|\partial^\beta f(\mathbf{x})| \leq A \quad \forall\, \mathbf{x} \in B \ \forall \ \text{multiindex } \beta \text{ with } |\beta| \leq \ell,$$

and produced the same conclusion as the above lemma, with $A_i = A$ and $a_i = a$ for all $i$.

*Sketch of Proof.* The case $d = 1$ can be proved by a verbatim repetition of the argument from [KM2] – it is easy to verify that a bound on just the top derivative is enough for the proof. The general case then follows using [KT, Corollary 2.3]. ☐

Recall that a map $\mathbf{f}$ from $U \subset \mathbb{R}^d$ to $\mathbb{R}^n$ is called $\ell$-*nondegenerate* at $\mathbf{x} \in U$ if partial derivatives of $\mathbf{f}$ at $\mathbf{x}$ up to order $\ell$ span $\mathbb{R}^n$, and $\ell$-*nondegenerate* if it is $\ell$-nondegenerate at $\lambda$-a.e. $\mathbf{x} \in U$. Arguing as in the proof of [KM2, Proposition 3.4], from the above lemma one deduces

**Proposition 4.4.** *For any $d, \ell \in \mathbb{N}$ there exists $C_{d,\ell} > 0$ with the following property. Let $n \in \mathbb{N}$ and let $\mathbf{f} = (f_1, \ldots, f_n) : U \to \mathbb{R}^n$ be $\ell$-nondegenerate at $\mathbf{x} \in U \subset \mathbb{R}^d$. Then for any $C > C_{d,\ell}$ there exists a neighborhood $V \subset U$ of $\mathbf{x}$ such that any linear combination of $1, f_1, \ldots, f_n$ is $(C, 1/d\ell)$-good on $V$.*

This implies that if $\mathbf{f} : U \to \mathbb{R}^n$ is $\ell$-nondegenerate, then the pair $(\mathbf{f}, \lambda)$ is $(C, 1/d\ell)$-good for any $C > C_{d,\ell}$. Also, nonplanarity is clearly an immediate consequence of nondegeneracy. Thus, by Theorem 1.5, there exists $\varepsilon_2 = \varepsilon_2(n, d, \ell)$ such that whenever $\varepsilon < \varepsilon_2$, (1.10) holds for $\nu = \lambda$ and any $\ell$-nondegenerate $\mathbf{f} : U \to \mathbb{R}^n$, $U \subset \mathbb{R}^d$. This was previously established in the case $d = 1$, $l = n = 2$ in [Ba2], with an additional assumption that $\mathbf{f}$ be $C^3$ rather than $C^2$. Also, M. Dodson,

B. Rynne, and J. Vickers considered $C^3$ submanifolds of $\mathbb{R}^n$ with 'two-dimensional definite curvature almost everywhere', a condition which implies 2-nondegeneracy (and requires the dimension of the manifold to be at least 2). It is proved in [DRV] that almost every point on such a manifold is not in $\mathrm{DI}_\varepsilon$ for $\varepsilon < 2^{-\frac{n}{n+1}}$. We also remark that the result of [Sh2] extends to nondegenerate curves in $\mathbb{R}^2$.

4.3. **Friendly measures.** The class of friendly measures was introduced in [KLW], the word 'friendly' being an approximate abbreviation of 'Federer, nonplanar and decaying'. Using the terminology of the present paper, we can define this class as follows: a measure $\mu$ on $\mathbb{R}^n$ is *friendly* if for $\mu$-a.e. $\mathbf{x} \in \mathbb{R}^n$ there exist a neighborhood $U$ of $\mathbf{x}$ and $D, C, \alpha > 0$ such that $\mu$ is $D$-Federer on $U$, and $(\mathrm{Id}, \mu)$ is both $(C, \alpha)$-good and nonplanar[3] on $U$. In order to apply Theorem 1.5 we would like to use somewhat more uniform version: given $C, \alpha, D > 0$, define $\mu$ to be $(D, C, \alpha)$-*friendly* if for $\mu$-a.e. $\mathbf{x} \in \mathbb{R}^n$ there exists a neighborhood $U$ of $\mathbf{x}$ such that $\mu$ is $D$-Federer on $U$ and $(\mathrm{Id}, \mu)$ is both $(C, \alpha)$-good and nonplanar on $U$. In view of Theorem 1.5, almost all points with respect to those measures are not in $\mathrm{DI}_\varepsilon(\mathcal{T})$ for any $\mathcal{T}$ drifting away from walls and small enough $\varepsilon$, where $\varepsilon$ depends only on $C, D, \alpha$.

As discussed in the previous subsections, smooth measures on nondegenerate submanifolds of $\mathbb{R}^n$ satisfy the above properties. Furthermore, the class of friendly measures is rather large; many examples are described in [KLW, KW1, Ur1, Ur2, SU]. A notable class of examples is given by limit measures of finite irreducible systems of contracting similarities [KLW, §8] (or, more generally, self-conformal contractions, [Ur1]) of $\mathbb{R}^n$ with the open set condition. These measures were shown to be $(D, C, \alpha)$-friendly for some $D, C, \alpha$, thus satisfy the conclusions of Theorem 1.5.

4.4. **Improving DT along non-drifting $\mathcal{T}$.** Comparing Theorem 1.4 with Theorem 1.5, one sees that the former has more restrictive assumptions, namely $\mathcal{T}$ has to drift away from walls as opposed to just be unbounded. This is not an accident: the drift condition is in fact necessary for the main technical tools of the proof, that is, equidistribution results of §2.

To see this, for simplicity let us restrict ourselves to the case $m = 2$, $n = 1$; the argument for the general case is similar. Suppose that $\mathcal{T} = \{\mathbf{t}^{(\ell)} = (t_1^{(\ell)}, t_2^{(\ell)}, t_3^{(\ell)}) : \ell \in \mathbb{N}\} \subset \mathfrak{a}^+$ is unbounded but does not satisfy (1.8); that is, either $\{t_1^{(\ell)}\}$ or $\{t_2^{(\ell)}\}$ is bounded. Without loss

---

[3] We remark that the nonplanarity of $(\mathrm{Id}, \mu)$ as defined in this paper is a condition weaker than the nonplanarity of $\mu$ as defined in [KLW].

of generality, and passing to a subsequence, we can assume that $t_1^{(\ell)}$ is convergent as $\ell \to \infty$; that is, for any $\ell$ we can write $\mathbf{t}^{(\ell)} = \mathbf{s}^{(\ell)} + \mathbf{u}^{(\ell)}$ where

$$\mathbf{s}^{(\ell)} = (0, s^{(\ell)}, s^{(\ell)}), \quad s^{(\ell)} \to \infty \text{ as } \ell \to \infty, \tag{4.2}$$

and $g_{\mathbf{u}^{(\ell)}} \to g_0 \in \mathrm{SL}_3(\mathbb{R})$ as $\ell \to \infty$. Note that for any $\ell$ and any $Y \in M_{2,1}$, $g_{s^{(\ell)}}\tau(Y)$ belongs to the group

$$H \overset{\mathrm{def}}{=} \left\{ \begin{pmatrix} 1 & \mathbf{x} \\ 0 & g \end{pmatrix} : g \in \mathrm{SL}_2(\mathbb{R}), \ \mathbf{x} \in \mathbb{R}^2 \right\}, \tag{4.3}$$

a semi-direct product of $\mathrm{SL}_2(\mathbb{R})$ and $\mathbb{R}^2$. Thus for any $Y$, the trajectory $\{g_{\mathbf{t}}\tau(Y)\mathbb{Z}^3 : \mathbf{t} \in \mathcal{T}\}$ must approach $g_0 H \mathbb{Z}^3$, which is a proper submanifold inside the space of lattices in $\mathbb{R}^3$ (in fact, $H\mathbb{Z}^3$ is the set of lattices in $\mathbb{R}^3$ containing $\mathbf{e}_1 = (1,0,0)$ as a primitive vector); therefore the translates $g_{\mathbf{t}}\tau(B)\mathbb{Z}^3$, where $B$ is any subset of $M_{2,1}$, do not become equidistributed. We conclude that it is not possible to prove the analogue of Theorem 2.2 with $\lfloor\mathbf{t}\rfloor$ replaced by $\|\mathbf{t}\|$.

Note that this apriori does not rule out proving Theorem 1.4 with a relaxed assumption on $\mathcal{T}$: recall that our goal was to make almost every orbit return to a specific set $K_\varepsilon$, not just any nonempty open set. And indeed, it is easy to show, using Theorem 1.4, induction and Fubini's Theorem, that the set $\mathrm{DI}_\varepsilon(\mathcal{T})$ is $\lambda$-null whenever $\varepsilon < 1$ and $\mathcal{T}$ is such that for every $i$, $\{t_i : \mathbf{t} \in \mathcal{T}\}$ is either unbounded or converges to 0.

However, in general Theorem 1.4 is false if one just assumes that $\mathcal{T}$ is unbounded. Here is a simple counterexample, also in the case $m = 2$, $n = 1$. Fix $0 < \varepsilon < 1$ and write $\mathbf{t}^{(\ell)} = \mathbf{s}^{(\ell)} + \mathbf{u}$, where $\mathbf{s}^{(\ell)}$ is as in (4.2) and $\mathbf{u} = (u, u, 0)$. Then for any $\ell$ and any $Y$, $g_{\mathbf{s}^{(\ell)}}\tau(Y)\mathbb{Z}^3$ belongs to $H\mathbb{Z}^3$ where $H$ is as in (4.3), that is, contains $\mathbf{e}_1$ as a primitive vector. Therefore for any $\mathbf{t} \in \mathcal{T}$ and any $\mathbf{y}$, $g_{\mathbf{t}}\tau(Y)\mathbb{Z}^3$ contains $e^u \mathbf{e}_1$ as a primitive vector. Now suppose that $1/\varepsilon^2 < e^u < 2\varepsilon$. Let $B_\varepsilon \subset \mathbb{R}^3$ be given by $\{|x_1| < e^u, |x_2| < \varepsilon, |x_3| < \varepsilon\}$. It is a convex centrally symmetric domain of volume greater than 8, hence, by Minkowski's Lemma, it must contain a nonzero vector $\mathbf{v} \in g_{\mathbf{t}}\tau(Y)\mathbb{Z}^3$. However, since $e^u < 2\varepsilon$, the sup-norm distance of $\mathbf{v}$ to either $e^u \mathbf{e}_1$ or $-e^u \mathbf{e}_1$ is less than $\varepsilon$, and it is positive since $\pm e^u \mathbf{e}_1 \notin B_\varepsilon$. This proves that $g_{\mathbf{t}}\tau(Y)\mathbb{Z}^3$ is always disjoint from $K_\varepsilon$; thus, under those assumptions on $\mathcal{T}$ and $\varepsilon$, the set $\mathrm{DI}_\varepsilon(\mathcal{T})$ is equal to $M_{2,1}$.

Similar counterexamples exist in any dimension. Still, it seems plausible that Theorem 1.4 will remain valid if (1.8) is replaced by an assumption that $\liminf_{\ell\to\infty} \lfloor\mathbf{t}^{(\ell)}\rfloor$ is large enough. The proof of this

requires equidisiribution results for more general homogeneous spaces (specifically, spaces similar to $H\mathbb{Z}^3$ in the above example).

4.5. **Weighted badly approximable systems.** We conclude the paper with another application of Theorem 2.2. Let $g_t$ be a one-parameter subgroup of $G$. Suppose a subgroup $H$ of $G$ normalized by $g_t$ is such that (a) the conjugation by $g_t$, $t > 0$, restricted to $H$ is an expanding automorphism of $H$, and (b) $g_t$-translates of the leaves $Hx$, $x \in G/\Gamma$, become equidistributed as $t \to \infty$, with the convergence uniform as $x$ ranges over compact subsets of $G/\Gamma$. These conditions were shown in [KM1] to imply that for any $x \in G/\Gamma$, the set

$$\big\{ h \in H : \text{ the trajectory } \{g_t hx : t > 0\} \text{ is bounded} \big\}$$

is thick (that is, has full Hausdorff dimension at every point). When $H$ is as in (2.5) and $\{g_t : t > 0\}$ is any one-parameter subsemigroup of $G$ contained in $\exp(\mathfrak{a}^+)$, both (a) and (b) are satisfied, the latter being a consequence of Theorem 2.2.

Let us now take $g_t$ of the form

$$g_t = \text{diag}(e^{r_1 t}, \ldots, e^{r_m t}, e^{-s_1 t}, \ldots, e^{-s_n t}),$$

where

$$r_i, s_j > 0 \quad \text{and} \quad \sum_{i=1}^m r_i = 1 = \sum_{j=1}^n s_j. \tag{4.4}$$

Then it is known [Kl1] that the trajectory $\{g_t \bar\tau(Y) : t > 0\}$ is bounded in $G/\Gamma$ if and only if $Y$ is $(\mathbf{r}, \mathbf{s})$-*badly approximable*, which by definition means

$$\inf_{\mathbf{p} \in \mathbb{Z}^m, \, \mathbf{q} \in \mathbb{Z}^n \smallsetminus \{0\}} \max_i |Y_i \mathbf{q} - p_i|^{1/r_i} \cdot \max_j |q_j|^{1/s_j} > 0.$$

The components of vectors $\mathbf{r}, \mathbf{s}$ should be thought of as weights assigned to linear forms $Y_i$ and integers $q_j$. Thus one can obtain a weighted generalization of W.M. Schmidt's theorem [Sc1] on the thickness of the set of badly approximable systems of linear forms:

**Corollary 4.5.** *For any choice of $\mathbf{r}, \mathbf{s}$ as in (4.4), the set of $(\mathbf{r}, \mathbf{s})$-badly approximable $Y \in M_{m,n}$ is thick.*

This was previously established by A. Pollington and S. Velani [PV] in the case $n = 1$.

## REFERENCES

[Ba1]    R. C. Baker, *Metric diophantine approximation on manifolds*, J. Lond. Math. Soc. (2) **14** (1976), 43–48.

[Ba2]    _____, *Dirichlet's theorem on diophantine approximation*, Math. Proc. Cambridge Phil. Soc. **83** (1978), 37–59.

[BKM]    V. Bernik, D. Kleinbock and G. A. Margulis, *Khintchine-type theorems on manifolds: convergence case for standard and multiplicative versions*, Internat. Math. Res. Notices **2001**, no. 9, 453–486.

[Bu]     Y. Bugeaud, *Approximation by algebraic integers and Hausdorff dimension,* J. London Math. Soc. (2) **65** (2002), no. 3, 547–559.

[Ca]     J. W. S. Cassels, **An introduction to Diophantine approximation**, Cambridge Tracts **45**, Cambridge University Press, New York, 1957.

[Da1]    _____, *Invariant measures and minimial sets of horospherical flows*, Inv. Math. **64** (1981) 357–385.

[Da2]    S. G. Dani, *Divergent trajectories of flows on homogeneous spaces and diophantine approximation*, J. Reine Angew. Math. **359** (1985), 55–89.

[DM]     S. G. Dani and G. A. Margulis, *Limit distributions of orbits of unipotent flows and values of quadratic forms*, I.M. Gelfand Seminar, 91–137, Adv. Soviet Math., **16**, Part 1, Amer. Math. Soc., Providence, RI, 1993.

[DS1]    H. Davenport and W. M. Schmidt, *Dirichlet's theorem on diophantine approximation*, in: Symposia Mathematica, Vol. IV (INDAM, Rome, 1968/69), pp. 113–132, 1970.

[DS2]    _____, *Dirichlet's theorem on diophantine approximation. II*, Acta Arith. **16** (1969/1970) 413–424.

[DRV]    M. Dodson, B. Rynne, and J. Vickers, *Dirichlet's theorem and Diophantine approximation on manifolds*, J. Number Theory **36** (1990), no. 1, 85–88.

[Kl1]    D. Kleinbock, *Flows on homogeneous spaces and Diophantine properties of matrices*, Duke Math. J. **95** (1998), no. 1, 107–124.

[Kl2]    _____, *An extension of quantitative nondivergence and applications to Diophantine exponents*, Trans. Amer. Math. Soc. (to appear).

[KLW]    D. Kleinbock, E. Lindenstrauss and B. Weiss, *On fractal measures and diophantine approximation*, Selecta Math. **10** (2004), 479–523.

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

[KM2]    _____, *Flows on homogeneous spaces and Diophantine approximation on manifolds,* Ann. Math. **148** (1998), 339–360.

[KM3]    _____, *On equidistribution of expanding translates of certain orbits in the space of lattices*, Preprint (2007), `arXiv:math/0702433`.

[KSS]    D. 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, pp. 813–930.

[KT]     D. Kleinbock and G. Tomanov, *Flows on S-arithmetic homogeneous spaces and applications to metric Diophantine approximation*, Comment. Math. Helv. **82** (2007), 519–581.

[KW1]    D. Kleinbock and B. Weiss, *Badly approximable vectors on fractals*, Israel J. Math. **149** (2005), 137–170.

[KW2]    _____, *Friendly measures, homogeneous flows and singular vectors*, in: **Algebraic and Topological Dynamics**, Contemp. Math. **211**, Amer. Math. Soc., Providence, RI, 2005, pp. 281–292.

[Ma1]   G.A. Margulis, *On the action of unipotent groups in the space of lattices*, in: **Lie groups and their representations**, Proc. Summer School, Bolyai, János Math. Soc., Budapest, 1971), Halsted, New York, 1975, pp. 365–370.

[Ma2]   G.A. Margulis, **On some aspects of the theory of Anosov systems**, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2004.

[Mat]   P. Mattila, **Geometry of sets and measures in Euclidean space. Fractals and rectifiability**, Cambridge Stud. Adv. Math. **44** Cambridge University Press, Cambridge, 1995.

[Mo]    D. Morris, **Ratner's theorems on unipotent flows**, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 2005.

[PV]    A. Pollington and S. Velani, *On simultaneously badly approximable numbers,* J. London Math. Soc. (2) **66** (2002), no. 1, 29–40.

[Ra]    M.S. Raghunathan, **Discrete Subgroups of Lie Groups**, Springer, Berlin, 1972.

[Sc1]   W.M. Schmidt, *Badly approximable systems of linear forms*, J. Number Theory **1** (1969), 139–154.

[Sc2]   ———, **Diophantine approximation**, Lecture Notes in Mathematics, vol. 785, Springer-Verlag, Berlin, 1980.

[Se]    J.-P. Serre, **Complex semisimple Lie algebras**, Springer monographs in mathematics, Berlin, 2001 (reprint of the 1987 edition).

[Sh1]   N. A. Shah, *Limit distributions of expanding translates of certain orbits on homogeneous spaces*, Proc. Ind. Acad. Sci. (Math. Sci.) **106** (1996), 105–125.

[Sh2]   ———, personal communication.

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

[SU]    B. Stratmann and M. Urbanski, *Diophantine extremality of the Patterson measure*, Math. Proc. Cambridge Phil. Soc. **140** (2006), 297–304.

[Ur1]   M. Urbanski, *Diophantine approximation of self-conformal measures*, J. Number Theory **110** (2005), no. 2, 219–235.

[Ur2]   ———, *Diophantine approximation for conformal measure of one-dimensional iterated function systems*, Compos. Math. **141** (2005), no. 4, 869–886.

[We]    B. Weiss, *Finite dimensional representations and subgroup actions on homogeneous spaces*, Israel J. Math. **106** (1998), 189–207.

BRANDEIS UNIVERSITY, WALTHAM MA 02454-9110 `kleinboc@brandeis.edu`

BEN GURION UNIVERSITY, BE'ER SHEVA, ISRAEL 84105 `barakw@math.bgu.ac.il`