WEIGHTED UNIFORM DIOPHANTINE APPROXIMATION OF SYSTEMS OF LINEAR FORMS

DMITRY KLEINBOCK AND ANURAG RAO

Abstract. Following the development of weighted asymptotic approximation properties of matrices, we introduce the analogous uniform approximation properties (that is, study the improvability of Dirichlet's Theorem). An added feature is the use of general norms, rather than the supremum norm, to quantify the approximation. In terms of homogeneous dynamics, the approximation properties of an $m \times n$ matrix are governed by a trajectory in $\mathrm{SL}_{m+n}(\mathbb{R})/\mathrm{SL}_{m+n}(\mathbb{Z})$ avoiding a compact subset of the space of lattices called the critical locus defined with respect to the corresponding norm. The trajectory is formed by the action of a one-parameter diagonal subgroup corresponding to the weights. We first state a very precise form of Dirichlet's theorem and prove it for some norms. Secondly we show, for these same norms, that the set of Dirichlet-improvable matrices has full Hausdorff dimension. Though the techniques used vary greatly depending on the chosen norm, we expect these results to hold in general.

1. Introduction

Let m and n be positive integers and let $d = m + n$. We will denote by $M_{m,n}$ the space of $m \times n$ real matrices, and by $\|\cdot\|_\infty$ the supremum norm on $\mathbb{R}^m$, $\mathbb{R}^n$ and $\mathbb{R}^d$. The classical theorem of Dirichlet, see e.g. [C1, §I.1.5], asserts that for any $A \in M_{m,n}$ and $t > 1$ there exists $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^m \times (\mathbb{Z}^n \smallsetminus \{\mathbf{0}\})$ satisfying

$$\|A\mathbf{q} - \mathbf{p}\|_\infty^m \le 1/t \quad \text{and} \quad \|\mathbf{q}\|_\infty^n < t. \tag{1.1}$$

Here A is viewed as a system of m linear forms $A_1, \dots, A_m$ (rows of A) in n variables, and the goal is to approximate the values of these forms at integer points by integers. A natural question to ask is whether one can improve (1.1) by replacing $1/t$ with a smaller function, that is, consider the following system of inequalities:

$$\|A\mathbf{q} - \mathbf{p}\|_\infty^m < \psi(t) \quad \text{and} \quad \|\mathbf{q}\|_\infty^n < t, \tag{1.2}$$

where ψ is a positive function such that $\psi(t)$ is strictly less than $\psi_1(t) := 1/t$ for all large enough t. One says that A is *ψ-Dirichlet* (see [KWa1, KWa2, KSY]) if the system (1.2) has solutions in $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^m \times (\mathbb{Z}^n \smallsetminus \{\mathbf{0}\})$ for all sufficiently large t. We will denote the set of ψ-Dirichlet matrices by $D_\infty(\psi)$. (The use of the subscript ∞ in (1.3) and in other occurrences below refers to the use of the supremum norm in (1.2).)

The above set-up is usually referred to as *uniform approximation*, as opposed to *asymptotic approximation* dealing with the system (1.2) being solvable for an unbounded set of t. Note that from Dirichlet's Theorem it trivially follows that $D_\infty(c\psi_1) = M_{m,n}$ if $c > 1$, and with a little more work, caused by the difference between '$<$' in (1.1) and '$\le$' in (1.2), one can show that $D_\infty(\psi_1) = M_{m,n}$ as well, see Theorem 1.1 below for a more general statement.

Date: February, 2022.

2010 *Mathematics Subject Classification.* 11J13; 11J83, 11H06, 37A17.

D.K. was supported by NSF grant DMS-1900560.

The problem of improving Dirichlet's theorem was initiated by Davenport and Schmidt [DS] who showed that the set

$$\mathbf{DI}_\infty := \bigcup_{0<c<1} D_\infty(c\psi_1) \tag{1.3}$$

of *Dirichlet improvable* matrices is of Lebesgue measure zero, while having full Hausdorff dimension mn. Furthermore, Davenport and Schmidt showed that **DI** contains the set **BA** of *badly approximable* matrices

$$\mathbf{BA} := \left\{ A \in M_{m,n} : \inf_{\mathbf{p}\in\mathbb{Z}^m,\, \mathbf{q}\in\mathbb{Z}^n\smallsetminus\{\mathbf{0}\}} \|A\mathbf{q}-\mathbf{p}\|_\infty^m \|\mathbf{q}\|_\infty^n > 0 \right\},$$

which was known to be *thick*, that is, have full Hausdorff dimension at any point of $M_{m,n}$ [Sc2].

In this paper we will generalize the above set-up in several different ways. It is known that many results in Diophantine approximation extend to *approximation with weights*, an approach allowing to treat forms A_i and components of $\mathbf{q}$ differently. Namely, given a tuple of positive weights

$$\boldsymbol{\omega} = (\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{R}_+^m \times \mathbb{R}_+^n \text{ with } \sum_{i=1}^m \alpha_i = \sum_{i=1}^n \beta_i = 1, \tag{1.4}$$

one introduces *quasi-norms* associated with $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ respectively:

$$\|\mathbf{x}\|_{\boldsymbol{\alpha}} := \max_i |x_i|^{1/\alpha_i} \quad \text{and} \quad \|\mathbf{y}\|_{\boldsymbol{\beta}} := \max_j |y_j|^{1/\beta_j}.$$

Then, for ψ as above, one says that $A \in M_{m,n}$ is $(\psi, \boldsymbol{\omega})$-*Dirichlet*, denoted by $A \in D_{\infty,\boldsymbol{\omega}}(\psi)$, if the system of inequalities

$$\|A\mathbf{q}-\mathbf{p}\|_{\boldsymbol{\alpha}} < \psi(t) \quad \text{and} \quad \|\mathbf{q}\|_{\boldsymbol{\beta}} < t$$

in $(\mathbf{p},\mathbf{q}) \in \mathbb{Z}^m \times (\mathbb{Z}^n \smallsetminus \{\mathbf{0}\})$ for all sufficiently large t. In other words, we are considering the solvability of the system

$$\begin{cases} |A_i \cdot \mathbf{q} - p_i| < \psi(t)^{\alpha_i}, & i = 1,\dots,m; \\ |q_j| < t^{\beta_j}, & j = 1,\dots,n. \end{cases} \tag{1.5}$$

Clearly the unweighted case corresponds to the choice

$$\boldsymbol{\alpha} = (1/m,\dots,1/m) \text{ and } \boldsymbol{\beta} = (1/n,\dots,1/n).$$

A lot of what can be proved for unweighted approximation easily extends to the weighted case. A weighted analogue of Dirichlet's theorem, which is a straightforward consequence of Minkowski's Convex Body Theorem [C1, §III.2.2], implies that $D_{\boldsymbol{\omega}}(c\psi_1) = M_{m,n}$ if $c > 1$. And with a little more work one can prove a stronger result:

Theorem 1.1. *For any choice of weights $\boldsymbol{\omega}$, we have $D_{\infty,\boldsymbol{\omega}}(\psi_1) = M_{m,n}$.*

As for the set

$$\mathbf{DI}_{\infty,\boldsymbol{\omega}} := \bigcup_{0<c<1} D_{\infty,\boldsymbol{\omega}}(c\psi_1),$$

the fact that it has Lebesgue measure zero was established by the first named author and Weiss using the correspondence between Diophantine approximation and dynamics, see [KWe1, Theorem 1.4]. In this paper we prove

Theorem 1.2. *For any choice of weights $\boldsymbol{\omega}$, the set $\mathbf{DI}_{\infty,\boldsymbol{\omega}}$ contains the set $\mathbf{BA}_{\boldsymbol{\omega}}$ of $\boldsymbol{\omega}$-badly approximable matrices, defined by*

$$\mathbf{BA}_{\boldsymbol{\omega}} := \left\{ A \in M_{m,n} : \inf_{\mathbf{p}\in\mathbb{Z}^m,\, \mathbf{q}\in\mathbb{Z}^n\smallsetminus\{\mathbf{0}\}} \|A\mathbf{q}-\mathbf{p}\|_{\boldsymbol{\alpha}}\|\mathbf{q}\|_{\boldsymbol{\beta}} > 0 \right\}. \tag{1.6}$$

Note that the latter set is thick, as shown in [KWe1, §4.5], see also [PV] and [KWe2]. It should also be noted that in [Su, Theorem 4.6], Suëss proved the above result in the case when $m = 1$. Our proof here is different and is written in the language of dynamics on the space of lattices.

We remark that the problem of determining conditions on ψ under which the set $D_{\infty,\boldsymbol{\omega}}(\psi)$ has zero/full measure is rather tricky. A complete solution for the case $m = n = 1$ is given in [KWa1], and a recent paper [KSY] by the first named author, Strombergsson and Yu deals with the general case, including arbitrary weights, and provides a partial result.

In order to generalize the set-up further, let us restate the definition of $(\psi,\boldsymbol{\omega})$-Dirichlet matrices in a geometric language. Let X_d denote the space of unimodular lattices in $\mathbb{R}^d$, identified with $\mathrm{SL}_d(\mathbb{R})/\mathrm{SL}_d(\mathbb{Z})$ via $g \mapsto g\mathbb{Z}^d$. Given $A \in M_{m,n}$, we define

$$u_A := \begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix}, \quad \Lambda_A := u_A\mathbb{Z}^d.$$

Then it is easy to see that $A \in D_\infty(\psi)$ if and only if

$$\Lambda_A \cap \begin{bmatrix} \psi(t)^{1/m} I_m & 0 \\ 0 & t^{1/n} I_n \end{bmatrix} B_\infty(1) \neq \{\mathbf{0}\} \tag{1.7}$$

for all sufficiently large t (here $B_\infty(1)$ is the unit open ball centered at zero with respect to the norm $\|\cdot\|_\infty$) And for a weighted version it will be convenient to use the following notation for a number raised to a vector power: if $c > 0$ and $\mathbf{x} \in \mathbb{R}^k$, define

$$c^{\mathbf{x}} := \mathrm{diag}(c^{x_1}, \dots, c^{x_k}).$$

Then, similarly to (1.7), one can state that $A \in D_{\infty,\boldsymbol{\omega}}(\psi)$ if and only if

$$\Lambda_A \cap \begin{bmatrix} \psi(t)^{\boldsymbol{\alpha}} & 0 \\ 0 & t^{\boldsymbol{\beta}} \end{bmatrix} B_\infty(1) \neq \{\mathbf{0}\} \tag{1.8}$$

for all sufficiently large t.

At this point one might wonder: what will change if in the above definition the supremum norm $\|\cdot\|_\infty$ is replaced by some other norm ν? and indeed this type of questions have appeared in the literature, first for the case $m = n = 1$ [AD], and then for arbitrary m, n in the unweighted case [KR1]. We will now use (1.8) to state a general weighted definition. In order to do that, for an arbitrary norm ν on $\mathbb{R}^d$ let us define the *critical radius* of ν as follows:

$$r_\nu := \sup\big\{ r : \Lambda \cap B_\nu(r) = \{\mathbf{0}\} \text{ for some } \Lambda \in X_d \big\}.$$

Here $B_\nu(r) := \{\mathbf{x} \in \mathbb{R}^d : \nu(\mathbf{x}) < r\}$; clearly $r_\infty = 1$. (Throughout the paper we will use the notation p when ν is the ℓ^p norm, in particular when $p = \infty$.)

Now let us define the most general sets of ψ-Dirichlet matrices.

Definition 1.3. *Given a function $\psi : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ and a tuple of weights $\boldsymbol{\omega} = (\boldsymbol{\alpha}, \boldsymbol{\beta})$ as in (1.4), we say that $A \in M_{m,n}$ is $(\psi, \nu, \boldsymbol{\omega})$-Dirichlet if*

$$\Lambda_A \cap \begin{bmatrix} \psi(t)^{\boldsymbol{\alpha}} & 0 \\ 0 & t^{\boldsymbol{\beta}} \end{bmatrix} B_\nu(r_\nu) \neq \{\mathbf{0}\}$$

for all sufficiently large t.

For brevity, we write the set of $(\psi, \nu, \boldsymbol{\omega})$-Dirichlet matrices as $D_{\nu,\boldsymbol{\omega}}(\psi)$. Note that the above property in general cannot be written in a way similar to (1.5), with separate conditions involving the linear forms A_i and the variables q_j. For example, in the case where $m = n = 1$, ν is the Euclidean norm on $\mathbb{R}^2$ and $\boldsymbol{\omega} = (1, 1)$ is the only possible choice for the weights, it is easy to see that $r_2 = \left(\frac{4}{3}\right)^{1/4}$. The corresponding condition for a real number α to be $(\psi, \nu, \boldsymbol{\omega})$-Dirichlet is that the inequality

$$\left(\frac{\alpha q - p}{\psi(t)}\right)^2 + \left(\frac{q}{t}\right)^2 < \frac{2}{\sqrt{3}}$$

has a solution in $(p, q) \in \mathbb{Z} \times \mathbb{N}$ for all sufficiently large t.

It immediately follows from the definition of r_ν that $D_{\nu,\boldsymbol{\omega}}(c\psi_1) = M_{m,n}$ for any $c > 1$. Also one can define

$$\mathbf{DI}_{\nu,\boldsymbol{\omega}} := \bigcup_{0<c<1} D_{\nu,\boldsymbol{\omega}}(c\psi_1),$$

the set of *weighted Dirichlet-improvable matrices with respect to* ν, and use the same dynamical argument as in [KWe1, Theorem 1.4] to prove

Theorem 1.4. *For any choice of a norm ν on $\mathbb{R}^d$ and a weight vector $\boldsymbol{\omega}$, the set $\mathbf{DI}_{\nu,\boldsymbol{\omega}}$ has Lebesgue measure zero.*

We are thus left with the following two problems:

1. Find norms ν and weight vectors $\boldsymbol{\omega}$ such that
$$D_{\nu,\boldsymbol{\omega}}(\psi_1)^c = \varnothing. \tag{1.9}$$
2. Find norms ν and weight vectors $\boldsymbol{\omega}$ such that
$$\mathbf{DI}_{\nu,\boldsymbol{\omega}} \text{ is thick.} \tag{1.10}$$

Both problems will be addressed in this paper for some specific choices of norms ν, using a dynamical restatement of the property of being $(\psi, \nu, \boldsymbol{\omega})$-Dirichlet. The choice of norms in the theorems below arise from what is known or can be proved regarding the densest lattice-packings of their unit balls. This will be made abundantly clear in the proofs.

With regards to problem 1 above, we have, like Theorem 1.1, a precise form of Dirichlet theorem in the following additional cases.

Theorem 1.5. *We have that $D_{\nu,\boldsymbol{\omega}}(\psi_1) = M_{m,n}$*

(a) *when $m = n = 1$ and ν is any ℓ^p norm on $\mathbb{R}^2$;*
(b) *when $m = 2$, $n = 1$, $\boldsymbol{\omega}$ is arbitrary, and ν on $\mathbb{R}^3$ is of the form*
$$(x, y, z) \mapsto \max\left\{\eta(x, y), |z|\right\} \text{ for some norm } \eta \text{ on } \mathbb{R}^2. \tag{1.11}$$

For Problem 2, when $m = n = 1$ and with only one possible choice of weights, the thickness result was established in [KR1, Theorem 1.3]. For the unweighted case of the Euclidean norm in arbitrary dimensions it was established in [KR1, Theorem 3.7]. The result for the weighted supremum norm in arbitrary dimension follows from Theorem 1.2. Presently we prove

Theorem 1.6. *The set $\mathbf{DI}_{\nu,\boldsymbol{\omega}}$ is thick*

(a) *for any $m, n, \boldsymbol{\omega}$, and when ν is the Euclidean norm on $\mathbb{R}^d$;*
(b) *when $m = 2$, $n = 1$, $\boldsymbol{\omega}$ is arbitrary, and ν on $\mathbb{R}^3$ is of the form (1.11).*

Theorems 1.5 and 1.6 can be proved for certain other norms as well. See Proposition 3.2 and Corollary 5.4 below for general results applicable to other norms.

One might also ask whether or not the inclusion

$$\mathbf{BA}_{\boldsymbol{\omega}} \subset \mathbf{DI}_{\nu,\boldsymbol{\omega}} \tag{1.12}$$

holds for some norms ν other than $\|\cdot\|_\infty$. In Proposition 4.1 we give a condition sufficient for (1.12), which in particular is valid for norms of the form (1.11) as in Theorem 1.6(b). However in general (1.12) is false: in fact for any $A \in \mathbf{BA}_{\boldsymbol{\omega}}$ one can find a norm ν such that $A \notin \mathbf{DI}_{\nu,\boldsymbol{\omega}}$. Moreover, the same holds for any $A \in M_{m,n}$ except for the case when A is $\boldsymbol{\omega}$-singular, or $A \in \mathbf{Sing}_{\boldsymbol{\omega}}$. The latter set is defined as

$$\mathbf{Sing}_{\boldsymbol{\omega}} := \bigcap_{0<c<1} D_{\nu,\boldsymbol{\omega}}(c\psi_1).$$

(It is easy to see that the choice of the norm does not make a difference in this definition.) We prove

Theorem 1.7. *For any weight vector $\boldsymbol{\omega}$,*

$$\mathbf{Sing}_{\boldsymbol{\omega}} = \bigcap_{\nu \text{ a norm on } \mathbb{R}^d} \mathbf{DI}_{\nu,\boldsymbol{\omega}}$$

In fact, for any fixed norm ν on $\mathbb{R}^d$, we have

$$\mathbf{Sing}_{\boldsymbol{\omega}} = \bigcap_{g\in \mathrm{SL}_d(\mathbb{R})} \mathbf{DI}_{\nu\circ g,\boldsymbol{\omega}}.$$

This characterization of singular systems of linear forms is new even in the unweighted case.

The structure of the paper is as follows; in the next section we give a dynamical interpretation of Dirichlet-improvability. In particular, the relation to the *critical locus* of a norm is clarified. An effective equidistribution result on the space of lattices then yields the coarse form of Dirichlet's theorem as in Theorem 1.4. Theorems 1.1, 1.2, 1.5, 1.6(b) and 1.7 are proved in the next two sections by using the geometry of numbers to identify certain divergent subsets in the space of lattices. Part (a) of Theorem 1.6 is proved in §5 using results of the first-named author along with An and Guan.

Acknowledgements. The authors are grateful to Nikolay Moshchevitin for helpful discussions, and to the anonymous referee for several useful comments.

2. Dirichlet improvable matrices form a null set

As before, X_d denotes the space of unimodular lattices in $\mathbb{R}^d$, and ν stands for a norm on $\mathbb{R}^d$. For any $r > 0$ define

$$\mathcal{K}_\nu(r) := \big\{\Lambda \in X_d : \Lambda \cap B_\nu(r) = \{\mathbf{0}\}\big\}.$$

These sets are compact in view of Mahler's Compactness Criterion, and empty for $r > r_\nu$, whereas for $0 < r < r_\nu$, these give a system of neighborhoods of the non-empty compact *critical locus* $\mathcal{L}_\nu := \mathcal{K}_\nu(r_\nu)$. Up to scaling, $\mathcal{L}_\nu$ gives the set of lattices witnessing the densest lattice-packings of the unit ball of ν. Further, given a weight vector as in (1.4), we have the following one-parameter subgroup of $\mathrm{SL}_d(\mathbb{R})$:

$$a_s = \begin{bmatrix} (e^s)^{\boldsymbol{\alpha}} & 0 \\ 0 & (e^{-s})^{\boldsymbol{\beta}} \end{bmatrix}. \tag{2.1}$$

Proposition 2.1. *An $m \times n$ matrix A belongs to $\mathbf{DI}_{\nu,\boldsymbol{\omega}}$ if and only if there is some $0 < r < r_\nu$ and $s_0 > 0$ such that*

$$\{a_s\Lambda_A : s > s_0\} \cap \mathcal{K}_\nu(r) = \varnothing.$$

Proof. Say $A \in \mathbf{DI}_{\nu,\boldsymbol{\omega}}$, so that there is some $0 < c < 1$ with $A \in D_{\nu,\boldsymbol{\omega}}(c\psi_1)$. The defining intersection condition for $D_{\nu,\boldsymbol{\omega}}(c\psi_1)$ can be changed to

$$a_s\Lambda_A \cap a_s \begin{bmatrix} \big(c\psi_1(t)\big)^{\boldsymbol{\alpha}} & 0 \\ 0 & t^{\boldsymbol{\beta}} \end{bmatrix} B_\nu(r_\nu) \neq \{\mathbf{0}\}$$

for all sufficiently large t. Putting

$$s = \frac{1}{2}\ln\frac{t^2}{c}, \tag{2.2}$$

the condition becomes

$$a_s\Lambda_A \cap \begin{bmatrix} (\sqrt{c})^{\boldsymbol{\alpha}} & 0 \\ 0 & (\sqrt{c})^{\boldsymbol{\beta}} \end{bmatrix} B_\nu(r_\nu) \neq \{\mathbf{0}\}$$

for all sufficiently large s. Let

$$r = r_\nu \cdot \max\left\{c^{\alpha_1/2}, \dots, c^{\alpha_m/2}, c^{\beta_1/2}, \dots, c^{\beta_n/2}\right\}.$$

Since $c < 1$, r is less than r_ν. Thus, we have that

$$a_s\Lambda_A \notin \mathcal{K}_\nu(r) \tag{2.3}$$

for all sufficiently large s.

Conversely, say we have a matrix A for which there is an $0 < r < r_\nu$ such that (2.3) holds for all sufficiently large s. Thus

$$a_s\Lambda_A \cap B_\nu(r) \neq \{\mathbf{0}\} \tag{2.4}$$

for all sufficiently large s. Condition (2.4) can be rewritten as

$$\Lambda_A \cap \begin{bmatrix} (e^{-s})^{\boldsymbol{\alpha}} & 0 \\ 0 & (e^{s})^{\boldsymbol{\beta}} \end{bmatrix} B_\nu(r) \neq \{\mathbf{0}\}. \tag{2.5}$$

So, if we define

$$c := \left(\frac{r}{r_\nu}\right)^{\frac{2}{\gamma}} \text{ with } \gamma := \max\{\beta_j\},$$

and define $t > 0$ by the equation (2.2), we see that

$$e^{-s} = \frac{\sqrt{c}}{t} = \psi_1(t)\left(\frac{r}{r_\nu}\right)^{1/\gamma} \quad \text{and} \quad e^{s} = \frac{t}{\sqrt{c}} = t\left(\frac{r_\nu}{r}\right)^{1/\gamma}.$$

From this we see that

$$\frac{r}{r_\nu}e^{-s\alpha_i} = \left(\frac{r}{r_\nu}\right)^{1+\alpha_i/\gamma}\psi_1(t)^{\alpha_i} \quad \text{and} \quad \frac{r}{r_\nu}e^{s\beta_j} = \left(\frac{r}{r_\nu}\right)^{1-\frac{\beta_j}{\gamma}} t^{\beta_j}.$$

By choice of γ, we see that $\left(\frac{r}{r_\nu}\right)^{1-\frac{\beta_j}{\gamma}} \leq 1$. Defining $c_1 := \left(\frac{r}{r_\nu}\right)^{\frac{1}{\alpha_i}+\frac{1}{\gamma}}$, which is less than 1, condition (2.5) then implies

$$\Lambda_A \cap \begin{bmatrix} \big(c_1\psi_{n/m}(t)\big)^{\boldsymbol{\alpha}} & 0 \\ 0 & (t)^{\boldsymbol{\beta}} \end{bmatrix} B_\nu(r_\nu) \neq \{\mathbf{0}\}.$$

From this we can see that $A \in \mathbf{DI}_{\nu,\boldsymbol{\omega}}$. □

Propositions of the above sort first appeared in [D] and now go by the name 'Dani's correspondence'.

Corollary 2.2. *We have the equivalence*

$$A \notin D_{\nu,\boldsymbol{\omega}}(\psi_1) \iff a_s\Lambda_A \in \mathcal{L}_\nu \text{ for an unbounded set of positive times } s.$$

Proof. It suffices to go through the above proof putting $c = 1$ and $r = r_\nu$ in the forward and backward directions of the equivalence respectively. □

In order to prove Theorem 1.4 we need the following equidistribution theorem of Kleinbock–Weiss [KWe1, Theorem 2.2], see also [KM, Theorem 1.3] for an effective version. The argument appears in [KSY] in case of ν being the supremum norm and applies with little changes to the general case.

Theorem 2.3. *Let $f \in C_c(X_d)$, $B \subset M_{m,n}$ be bounded with positive Lebesgue measure, and $\delta > 0$ be given. Then there exists an $s_0 > 0$ such that for all $s > s_0$,*

$$\left| \frac{1}{\lambda(B)} \int_B f\left(a_s \Lambda_A\right) \, d\lambda(A) - \int_{X_d} f(x) \, d\mu \right| < \delta.$$

Here, the integrals are taken with respect to the Lebesgue measure λ on $M_{m,n}$ and the Haar probability measure μ on X_d. □

Proof of Theorem 1.4. We have $c < 1$. Let r be associated to c as in Proposition 2.1. We aim to show that for almost every $A \in M_{m,n}$, there is an unbounded positive sequence (s_k) such that

$$a_{s_k} \Lambda_A \in \mathcal{K}_\nu(r). \tag{2.6}$$

This and Proposition 2.1 then show that almost every $A \notin D_{\nu,\boldsymbol{\omega}}(c\psi_1)$. For $i \in \mathbb{N}$, if the set

$$B_i := \bigcap_{s>i} \{A \in M_{m,n} : a_s \Lambda_A \notin \mathcal{K}_\nu(r)\}$$

has positive Lebesgue measure, choose $B \subset B_i$ compact with positive measure as well. Take a non-negative $f \in C_c(X_d)$ which is supported on $\mathcal{K}_\nu(r)$, and choose $\delta = \frac{1}{2} \int_{X_d} f \, d\mu$. Applying Theorem 2.3 with $s > i$, we get a contradiction. Thus each B_i has measure zero and thus so does their union. Hence we have shown that Lebesgue almost every $A \in M_{m,n}$ has an unbounded positive sequence (s_k) for which (2.6) holds. □

3. Dirichlet's theorem via divergence

For the rest of the paper we fix a weight vector as in (1.4) and the one-parameter subgroup $\{a_s\}$ of $\mathrm{SL}_d(\mathbb{R})$ as in (2.1). We now address Problem 1 regarding Dirichlet's theorem in the form (1.9). First, a general condition implying the result.

Proposition 3.1. *Say ν is a norm in $\mathbb{R}^d$ with $\mathcal{L}_\nu = \bigcup \mathcal{Z}_i$ a finite union of compact subsets such that each $\mathcal{Z}_i$ has either one of the following properties.*

(i) *For every $\Lambda \in \mathcal{Z}_i$ and compact $\mathcal{K} \subset X_d$, there is a t_0 such that for all $s > t_0$, $a_s \Lambda \notin \mathcal{K}$. That is, every $\Lambda \in \mathcal{Z}_i$ is forward divergent.*

(ii) *For every $\Lambda \in \mathcal{Z}_i$ and compact $\mathcal{K} \subset X_d$, there is a t_0 such that for all $s < t_0$, $a_s \Lambda \notin \mathcal{K}$. That is, every $\Lambda \in \mathcal{Z}_i$ is backward divergent.*

Then $D_{\nu,\boldsymbol{\omega}}(\psi_1) = M_{m,n}$.

Proof. For the sake of contradiction, say that $A \notin D_{\nu,\boldsymbol{\omega}}(\psi_1)$. By Corollary 2.2, there is an unbounded positive sequence (s_k) such that for each k, $a_{s_k} \Lambda_A \in \mathcal{L}_\nu$. By the above finiteness hypothesis we might as well assume $\mathcal{L}_\nu$ itself has one of the properties (i) or (ii). Observe that compactness implies that there is a uniform t_0 in the above conditions which works for every $\Lambda \in \mathcal{L}_\nu$. We now separate into two cases.

(i) We can find t_0 such that for all $s > t_0$,

$$a_s \mathcal{L}_\nu \cap \mathcal{L}_\nu = \varnothing. \tag{3.1}$$

This contradicts the fact that for every k, $a_{s_k} \Lambda_A (= a_{s_k - s_1} a_{s_1} \Lambda_A)$ belongs to $\mathcal{L}_\nu$.

(ii) Find t_0 such that for all $s < t_0$, (3.1) holds. This contradicts the fact that for every k, $a_{s_1}\Lambda_A (= a_{s_1-s_k}a_{s_k}\Lambda_A) \in \mathcal{L}_\nu$.

Thus $D_{\nu,\boldsymbol{\omega}}(\psi_1) = M_{m,n}$. □

Proof of Theorem 1.1. Let B denote the set of upper triangular unipotent $d \times d$ matrices. It is a well-known theorem of Hajós [H] that the set $\mathcal{L}_\infty$ is exactly the union

$$\bigcup \{wBw\,\mathrm{SL}_d(\mathbb{Z}) : w \text{ is a permutation matrix}\}. \tag{3.2}$$

From this we get that for every permutation matrix w, there is some fixed standard basis vector $\mathbf{e}_i$ which belongs to every $\Lambda \in wBw\,\mathrm{SL}_d(\mathbb{Z})$. From the description of a_s in (2.1), we see that, according to whether $m < i$ or $i \le m$, $\mathbf{e}_i$ is contracted by a_s either for $s > 0$ or $s < 0$. Thus, for each permutation matrix w, we are in one of the two situations of Proposition 3.1. □

Proof of Theorem 1.5(b). [KR2, Proposition 5.1] asserts that whenever ν is a cylindrical norm on $\mathbb{R}^3$ as in (1.11), the critical locus in X_3 is contained in the union of

$$\mathcal{Z}_1 := \left\{ \begin{bmatrix} * & * & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix} \mathbb{Z}^3 \right\} \text{ and } \mathcal{Z}_2 := \left\{ \begin{bmatrix} * & * & * \\ * & * & * \\ 0 & * & * \end{bmatrix} \mathbb{Z}^3 \right\}. \tag{3.3}$$

Moreover, since we have $m = 2$ and $n = 1$ by hypothesis,

$$a_s = \begin{bmatrix} e^{s\alpha_1} & 0 & 0 \\ 0 & e^{s\alpha_2} & 0 \\ 0 & 0 & e^{-s} \end{bmatrix}. \tag{3.4}$$

Thus, if $\Lambda \in \mathcal{Z}_1$, it contains a vector contracted by a_s for $s > 0$. And if $\Lambda \in \mathcal{Z}_2$, it contains a vector contracted by a_s for $s < 0$. Applying Proposition 3.1, we are done. □

We also have the following simple but useful result:

Proposition 3.2. *Let ν be a norm on $\mathbb{R}^d$ such that the critical locus $\mathcal{L}_\nu$ is finite. Then $D_{\nu,\omega}(\psi_1) = M_{m,n}$.*

Proof. Again, by Proposition 2.1, any $A \notin D_{\nu,\omega}$ would give rise to a periodic orbit $\{a_s\Lambda_A\}$. On the other hand, Λ_A is backward divergent under the flow a_s. □

Proof of Theorem 1.5(a). This has already been proved for $p = 2$ in [KR1, Theorem 1.4], and for $p = \infty$ in Theorem 1.1. For the other cases, the work [GGM] shows that $\mathcal{L}_p$ is finite. Thus we are done by applying Proposition 3.2. □

Remark 3.3. *Other examples of norms which are known to have finite critical locus are norms in $\mathbb{R}^2$ induced by hexagons, as well as the ℓ^1 norm in $\mathbb{R}^3$. For the former fact see* [C2, §V.8.4, Lemma 13] *and for the latter see* [M] *or the discussion in the pages prior to* [GL, Equation (4), page 346].

4. Thickness results via divergence

Some similar observations about divergence in the space of lattices lead us to solutions of Problem 2 as well. Recall the set $\mathbf{BA}_{\boldsymbol{\omega}}$ of $\boldsymbol{\omega}$-badly approximable matrices defined in (1.6). It is well known (see [K, Theorem 2.5]) that

$$A \in \mathbf{BA}_{\boldsymbol{\omega}} \iff \{a_s\Lambda : s > 0\} \text{ is bounded in } X_d.$$

We now give a general proposition giving sufficient conditions (on the norm ν) which ensure that $\mathbf{BA}_{\boldsymbol{\omega}}$ is a subset of $\mathbf{DI}_\nu$.

Proposition 4.1. *If ν is a norm on $\mathbb{R}^d$ such that every $\Lambda \in \mathcal{L}_\nu$ has the property that*

$$\{a_s\Lambda : s \in \mathbb{R}\} \text{ is unbounded in } X_d,$$

then $\mathbf{BA}_{\boldsymbol{\omega}}$ is contained in $\mathbf{DI}_{\nu,\boldsymbol{\omega}}$.

Remark 4.2. *To be precise, being unbounded means that for each compact $\mathcal{K} \subset X_d$, there is some $s \in \mathbb{R}$ such that $a_s\Lambda \notin \mathcal{K}$.*

Proof. We again use the characterization in Proposition 2.1. Say ν is a norm with the property as above. Say $A \in \mathbf{BA}_{\boldsymbol{\omega}}$. Say further, contrary to the theorem, that there is an unbounded positive sequence (s_k) and a lattice $\Lambda \in \mathcal{L}_\nu$ such that $a_{s_k}\Lambda_A \to \Lambda$. Let $\mathcal{K} \subset X_d$ be a compact set such that

$$\{a_s\Lambda_A : s > 0\} \subset \mathcal{K}.$$

We consider two cases.

(i) $\{a_s\Lambda : s > 0\}$ is unbounded. This implies that there is a positive time t for which $a_t\Lambda \notin \mathcal{K}$. Let $\mathcal{V}$ be a neighborhood of Λ such that

$$a_t\mathcal{V} \subset X_d \smallsetminus \mathcal{K}. \tag{4.1}$$

Thus, for large enough k, we have $a_{t+s_k}\Lambda_A \notin \mathcal{K}$, a contradiction.

(ii) For the second case, we assume that $\{a_s\Lambda : s < 0\}$ is unbounded. This means we have a negative t for which $a_t\Lambda \notin \mathcal{K}$. Let $\mathcal{V}$ again be a neighborhood such that (4.1) holds. We have that for large k, $a_{t+s_k}\Lambda_A \notin \mathcal{K}$. On observing that $t + s_k$ is positive for large k, we have a contradiction.

Thus, any A in $\mathbf{BA}_{\boldsymbol{\omega}}$ must belong to $\mathbf{DI}_{\nu,\boldsymbol{\omega}}$. □

Proof of Theorem 1.2. As was observed before, it follows from the expression (3.2) for the critical locus $\mathcal{L}_\infty$ that every $\Lambda \in \mathcal{L}_\infty$ contains one of the basis vectors $\mathbf{e}_i$. So, according to whether $m < i$ or $i \le m$, $\mathbf{e}_i$ is contracted by a_s either for $s > 0$ or $s < 0$. This, of course, implies that $\{a_s\Lambda\}$ is unbounded and we can apply Proposition 4.1. □

Proof of Theorem 1.6(b). Again, from (3.3) and (3.4) describing the critical locus and the flow respectively, we see that each $\Lambda \in \mathcal{L}_\nu$ is either forward or backward divergent (hence also unbounded) with respect to a_s. Thus Proposition 4.1 applies. □

Perhaps now is a good time to observe that the conclusion of Proposition 4.1 does not always hold. More precisely, for any $A \notin \mathbf{Sing}_{\boldsymbol{\omega}}$ there exists a norm ν on $\mathbb{R}^d$ such that A does not belong to $\mathbf{DI}_{\boldsymbol{\omega}}$.

Proof of Theorem 1.7. It is well known (see [K, Theorem 7.4], or [D, Proposition 2.12] for a version with equal weights) that $A \in \mathbf{Sing}_{\boldsymbol{\omega}}$ if and only if Λ_A is forward divergent under a_s. And by divergence, any such element must avoid any given critical locus after a certain time. Thus $\mathbf{Sing}_{\boldsymbol{\omega}}$ is contained in each of the intersections in the theorem.

To complete the proof, it now suffices to show that, for a fixed norm ν,

$$\bigcap_{g \in \mathrm{SL}_d(\mathbb{R})} \mathbf{DI}_{\nu\circ g,\boldsymbol{\omega}} \subset \mathbf{Sing}_{\boldsymbol{\omega}}.$$

Take $A \in M_{m,n}$ that is Dirichlet-improvable for all norms of the form $\nu \circ g$. In order to show that A is singular, it suffices to show that for every $\Lambda \in X_d$, there is a neighborhood $\mathcal{V}$ of Λ and some time s_0 such that the orbit $\{a_s\Lambda_A : s > s_0\}$ avoids $\mathcal{V}$.

Fix $\Lambda \in X_d$ and pick some $g \in \mathrm{SL}_d(\mathbb{R})$ such that $g\Lambda \in \mathcal{L}_\nu$. Since $g^{-1}\mathcal{L}_\nu = \mathcal{L}_{\nu\circ g}$, we see that $\Lambda \in \mathcal{L}_{\nu\circ g}$. By Dirichlet-improvability of A with respect to $\nu \circ g$, we see that there is an $r < r_{\nu\circ g}$ and some s_0 such that

$$a_s\Lambda_A \notin \mathcal{K}_{\nu\circ g}(r) \text{ for all } s > s_0.$$

As observed before, $\mathcal{K}_{\nu\circ g}(r)$ for $r < r_{\nu\circ g}$ is an open neighborhood of $\mathcal{L}_{\nu\circ g}$, and so we are done. □

5. Thickness results via transversality

In order to prove the thickness result for the Euclidean norm, we use a result of the first-named author with An and Guan [AGK]. They give a very general condition on the critical locus $\mathcal{L}_\nu$ which guarantees that the set of $A \in M_{m,n}$ such that the trajectory $\{a_s hx : s > 0\}$ eventually stays away from $\mathcal{L}_\nu$ is winning in the sense of Schmidt. More precisely, the results in [AGK] deal with a modified version of Schmidt's winning property called *hyperplane absolute winning* (HAW). For the definition of the HAW property, see [BFKRW, §2] or [AGK, §2.1]. HAW implies winning in the sense of Schmidt [Sc1], and this in turn implies thickness. Furthermore, the class of HAW sets, like those which are winning, is closed under countable intersections.

To state the aforementioned condition we need some notation. Let G denote $\mathrm{SL}_d(\mathbb{R})$, and let $\mathfrak{g}$ denote its Lie algebra $\mathfrak{sl}_d(\mathbb{R})$. Let $H \subset G$ denote the subgroup

$$H = \{u_A : A \in M_{m,n}\},$$

and let $\mathfrak{h}$ denote its Lie algebra. Fixing weights $\boldsymbol{\omega} = (\boldsymbol{\alpha}, \boldsymbol{\beta})$, let $F \subset G$ denote the subgroup

$$F = \{a_s : s \in \mathbb{R}\}$$

where a_s is as in (2.1). Let $D \in \mathfrak{g}$ denote the the diagonal element

$$D = \begin{bmatrix} (1)^{\boldsymbol{\alpha}} & 0 \\ 0 & -(1)^{\boldsymbol{\beta}} \end{bmatrix} = \mathrm{diag}(\alpha_1, \dots, \alpha_m, -\beta_1, \dots, -\beta_n)$$

so that

$$a_s = \exp(sD).$$

The adjoint action $\mathrm{ad}(D) : \mathfrak{g} \to \mathfrak{g}$ is diagonable: If we let $E^{i,j}$ denote the $d \times d$ matrix with 1 in the (i,j)-entry and 0 everywhere else we see that, when $i \neq j$,

$$\mathrm{ad}(D)E^{i,j} = (D_{ii} - D_{jj})\, E^{i,j},$$

and that

$$\mathrm{ad}(D)\left(E^{i,i} - E^{j,j}\right) = 0.$$

Here D_{ij} denotes the (i,j)-entry of D. Thus, if we let λ run over the eigenvalues of $\mathrm{ad}(D)$, we have an eigenspace decomposition

$$\mathfrak{g} = \bigoplus \mathfrak{g}_\lambda.$$

Let ρ denote the largest eigenvalue, and let $q : \mathfrak{g} \to \mathfrak{g}$ be the projection with image and kernel

$$\bigoplus_{\lambda=\rho} \mathfrak{g}_\lambda \text{ and } \bigoplus_{\lambda<\rho} \mathfrak{g}_\lambda$$

respectively. Let $\mathfrak{h}^{max}$ denote the image $q(\mathfrak{h})$, and let H^{max} denote the connected subgroup of G generated by $\mathfrak{h}^{max}$.

Remark 5.1. *Note that, from the definition of D, the eigenvectors with maximal eigenvalues must occur as matrices $E^{i,j}$ with $i \leq m$ and $j \geq n$. Thus $\mathfrak{h}^{max}$ (which is a subset of $\mathfrak{h}$) is never the zero subspace.*

We also use the notation $T_x(M)$ to denote the tangent space of a submanifold M of X_d at a point $x \in X_d$.

Definition 5.2. *A compact submanifold $Z \subset X_d$ is said to be (F, H^{max})-transversal if for all $z \in Z$,*

(i) $T_z(Fz) \not\subset T_z(Z)$;
(ii) $T_z(H^{max}z) \not\subset T_z(Z) \oplus T_z(Fz)$.

We can finally state the relevant result from [AGK, Theorem 2.8].

Theorem 5.3. *Keeping with the notation above, if $Z \subset X_d$ is an (F, H^{max})-transversal compact submanifold, then for any $x \in X_d$,*

$$\left\{h \in H : \overline{\{a_s hx : s > 0\}} \cap Z = \varnothing\right\}$$

is HAW in H.

Clearly zero-dimensional submanifolds are (F, H^{max})-transversal. And since countable intersections of winning sets are winning, on applying the above theorem to the case where $x \in X_d$ is the standard lattice, we have

Corollary 5.4. *If ν is a norm on $\mathbb{R}^d$ such that $\mathcal{L}_\nu$ is finite, then $\mathbf{DI}_{\nu,\omega}$ is thick.* □

We can also apply Theorem 5.3 to get

Proof of Theorem 1.6(a). Recall from Proposition 2.1 that $A \in \mathbf{DI}_{2,\boldsymbol{\omega}}$ if and only if there is some $r < r_\nu$ such that

$$a_s \Lambda_A \notin \mathcal{K}_2(r)$$

for all sufficiently large s. Here we are considering a neighborhood of the compact set $\mathcal{L}_2 \subset X_d$ which is a finite union of $\mathrm{SO}(d)$-orbits (see [KR1, Theorem 3.7]). The Lie algebra $\mathfrak{so}(d)$ consists of skew-symmetric matrices and it then becomes straightforward to check that each $\mathrm{SO}(d)$-orbit is an (F, H^{max})-transversal submanifold. Indeed, after identifying with $\mathfrak{g}$, we see that $T_z(Fz) = \mathrm{span}_{\mathbb{R}}\{D\}$, while $T_z(H^{max}z)$ includes nonzero upper triangular matrices, so that it is not contained in $\mathfrak{so}(d) \oplus \mathrm{span}_{\mathbb{R}}\{D\}$. Thus, Theorem 5.3 shows that $\mathbf{DI}_{2,\boldsymbol{\omega}}$ contains a finite intersection of winning sets, and thus is itself thick. □

References

[AD] N. Andersen and W. Duke, *On a theorem of Davenport and Schmidt*, Acta Arith. **198** (2021), no. 1, 37–75.

[AGK] J. An, L. Guan and D. Kleinbock, *Nondense orbits on homogeneous spaces and applications to geometry and number theory*, Ergodic Theory Dynam. Systems **42** (2022), no. 4, 1327–1372.

[BFKRW] R. Broderick, L. Fishman, D. Kleinbock, A. Reich and B. Weiss, *The set of badly approximable vectors is strongly $\mathcal{C}^1$ incompressible*, Math. Proc. Cambridge Philos. Soc. **153** (2012), no. 2, 319–339.

[C1] J. W. S. Cassels, *An Introduction to Diophantine Approximation*, Cambridge Tracts in Mathematics and Physics, Cambridge University Press, London, 1957.

[C2] ———, *An introduction to the geometry of numbers*, Die Grundlehren der mathematischen Wissenschaften, Bd. 99 Springer-Verlag, 1959, viii+344.

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

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

[H] G. Hajós, *Über einfache und mehrfache Bedeckung des n-dimensionalen Raumes mit einem Würfelgitter*, Math. Z. **47** (1941), 427–467.

[GGM] N. M. Glazunov, A. S. Golovanov and A. V. Malyshev, *Proof of the Minkowski conjecture on the critical determinant of the region $|x|^p + |y|^p < 1$*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov (LOMI) **151** (1986), 40–53.

[GL] P. Gruber and C. Lekkerkerker, *Geometry of numbers*, North-Holland Mathematical Library, **37**, North-Holland Publishing Co., Amsterdam, 1987.

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

[KM] D. Kleinbock and G. A. Margulis, *On effective equidistribution of expanding translates of certain orbits in the space of lattices*, in: Number Theory, Analysis and Geometry, Springer, New York, 2012, pp. 385–396.

[KR1] D. Kleinbock and A. Rao, *A zero-one law for uniform Diophantine approximation in Euclidean norm*, Internat. Math. Res. Notices **2022**, no. 8, 5617–5657.

[KR2] ———, *Abundance of Dirichlet-improvable pairs with respect to arbitrary norms*, Mosc. J. Comb. Number Theory **11** (2022), no. 1, 97–114.

[KSY] D. Kleinbock, A. Strömbergsson, and S. Yu, *A measure estimate in geometry of numbers and improvements to Dirichlet's theorem*, Proc. London Math. Soc., to appear.

[KWa1] D. Kleinbock and N. Wadleigh, *A zero-one law for improvements to Dirichlet's Theorem*, Proc. Amer. Math. Soc. **146** (2018), no. 5, 1833–1844.

[KWa2] ———, *An inhomogeneous Dirichlet theorem via shrinking targets*, Compos. Math. **155** (2019), no. 7, 1402–1423.

[KWe1] D. Kleinbock and B. Weiss, *Dirichlet's theorem on Diophantine approximation and homogeneous flows*, J. Mod. Dyn. **2** (2008), 43–62.

[KWe2] ———, *Modified Schmidt games and Diophantine approximation with weights*, Adv. Math. **223** (2010), 1276–1298.

[M] H. Minkowski, *Dichteste gitterförmige Lagerung kongruenter Körper*, Nachr. K. Ges. Wiss. Göttingen (1904), 311-355. (Reprinted in Gesammelte Abhandlungen II, 3-42)

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

[Sc1] W. M. Schmidt, *On badly approximable numbers and certain games*, Trans. Amer. Math. Soc. **123** (1966), 178–199.

[Sc2] ———, *Badly approximable systems of linear forms*, J. Number Theory **1** (1969), 139–154.

[Su] F. Süess, *Simultaneous Diophantine approximation on affine subspaces and Dirichlet improvability*, arXiv preprint `https://arxiv.org/abs/1711.08288` (2017).

Brandeis University, Waltham MA 02454-9110 `kleinboc@brandeis.edu`

Centre for Excellence in Basic Sciences, Mumbai 400098 `anurag.rao@cbs.ac.in`