UNIVERSITY of York

DIOPHANTINE TRANSFERENCE INEQUALITIES: WEIGHTED, INHOMOGENEOUS, AND INTERMEDIATE EXPONENTS

SAM CHOW, ANISH GHOSH, LIFAN GUAN, ANTOINE MARNAT, AND DAVID SIMMONS

ABSTRACT. We extend the Khintchine transference inequalities, as well as a homogeneous–inhomogeneous transference inequality for lattices, due to Bugeaud and Laurent, to a weighted setting. We also provide applications to inhomogeneous Diophantine approximation on manifolds and to weighted badly approximable vectors. Finally, we interpret and prove a conjecture of Beresnevich–Velani (2010) about inhomogeneous intermediate exponents.

1. INTRODUCTION

Dirichlet's approximation theorem [12] is a foundational result in Diophantine approximation, and follows straightforwardly from the pigeonhole principle.

Theorem 1.1 (Dirichlet's approximation theorem). *If $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m) \in \mathbb{R}^m$ and $N \in \mathbb{N}$ then there exists $q \leqslant N$ such that*

$$\|q\theta_i\|_{\mathbb{R}/\mathbb{Z}} \leqslant N^{-1/m} \qquad (1 \leqslant i \leqslant m).$$

In general it is sharp, but for some $\boldsymbol{\theta}$ there are closer rational approximations. This leads to the notion of *exponents of Diophantine approximation*, as introduced by Khintchine [25] and Jarník [21]. In this article, we concern ourselves with the very general setting of weighted Diophantine exponents, uniform or otherwise.

Definition 1.2. Let $m, n \in \mathbb{N}$ denote dimensions, and let $\mathbf{s} = (s_1, \ldots, s_m) \in \mathbb{R}_{>0}^m$ and $\mathbf{r} = (r_1, \ldots, r_n) \in \mathbb{R}_{>0}^n$ be weights, that is

$$\sum_{i=1}^{m} s_m = \sum_{j=1}^{n} r_n = 1.$$

Write

$$\|(x_1, \ldots, x_m)\|_{\mathbf{s}} = \max_{1 \leqslant i \leqslant m} |x_i|^{\frac{1}{s_i}}, \qquad \|(y_1, \ldots, y_n)\|_{\mathbf{r}} = \max_{1 \leqslant j \leqslant n} |y_j|^{\frac{1}{r_j}}.$$

Let $A \in M_{m \times n}(\mathbb{R})$ and $\boldsymbol{\theta} \in \mathbb{R}^m$. The *inhomogeneous weighted exponent*, denoted $\omega_{\mathbf{s},\mathbf{r}}(A, \boldsymbol{\theta})$, is the supremum of the real numbers ω for which, for some

---

1

arbitrarily large real numbers T, the inequalities

$$\|\mathbf{q}\|_{\mathbf{r}} < T, \qquad \|A\mathbf{q} - \mathbf{p} - \boldsymbol{\theta}\|_{\mathbf{s}} < T^{-\omega} \qquad (1.1)$$

have a solution $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^m \times (\mathbb{Z}^n \setminus \{\mathbf{0}\})$. The *uniform inhomogeneous weighted exponent*, denoted $\hat{\omega}_{\mathbf{s},\mathbf{r}}(A, \boldsymbol{\theta})$, is the supremum of the real numbers $\hat{\omega}$ for which, for **all** sufficiently large real numbers T, the inequalities

$$\|\mathbf{q}\|_{\mathbf{r}} < T, \qquad \|A\mathbf{q} - \mathbf{p} - \boldsymbol{\theta}\|_{\mathbf{s}} < T^{-\hat{\omega}}$$

have a solution $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^m \times (\mathbb{Z}^n \setminus \{\mathbf{0}\})$. Moreover, define the homogeneous exponents

$$\omega_{\mathbf{s},\mathbf{r}}(A) = \omega_{\mathbf{s},\mathbf{r}}(A, \mathbf{0}), \qquad \hat{\omega}_{\mathbf{s},\mathbf{r}}(A) = \hat{\omega}_{\mathbf{s},\mathbf{r}}(A, \mathbf{0}).$$

Finally, in the unweighted case $\mathbf{s} = (1/m, \ldots, 1/m)$ and $\mathbf{r} = (1/n, \ldots, 1/n)$, write $\omega(A, \boldsymbol{\theta})$ and $\hat{\omega}(A, \boldsymbol{\theta})$ for $\omega_{\mathbf{s},\mathbf{r}}(A, \boldsymbol{\theta})$ and $\hat{\omega}_{\mathbf{s},\mathbf{r}}(A, \boldsymbol{\theta})$, respectively.

Remark 1.3. Diophantine exponents are allowed to equal $+\infty$. We have normalised in such a way that Dirichlet's approximation theorem delivers the lower bound 1 for the exponent, as in [4]; it is more common to normalise so that Dirichlet's approximation theorem delivers the lower bound n/m. This value is said to be *critical*, as it is attained by almost all matrices A, especially in the context of §7.

Exponents of multiplicative Diophantine approximation can be similarly defined.

Definition 1.4. Let $A \in M_{m \times n}(\mathbb{R})$ and $\boldsymbol{\theta} \in \mathbb{R}^m$. The *inhomogeneous multiplicative exponent*, written $\omega^{\times}(A, \boldsymbol{\theta})$, is the supremum of the real numbers ω for which, for some arbitrarily large real numbers T, the inequalities

$$\Pi_+(\mathbf{q}) < T, \qquad \Pi(A\mathbf{q} - \mathbf{p} - \boldsymbol{\theta}) < T^{-\omega}$$

have a solution $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^m \times (\mathbb{Z}^n \setminus \{\mathbf{0}\})$, where

$$\Pi_+(\mathbf{q}) = \prod_{j=1}^n \max\{1, |q_j|\}, \qquad \Pi(\mathbf{y}) = \prod_{i=1}^m |y_i|$$

with $\mathbf{y} = (y_1, \ldots, y_m)$. The *uniform inhomogeneous multiplicative exponent*, written $\hat{\omega}^{\times}(A, \boldsymbol{\theta})$, is the supremum of the real numbers $\hat{\omega}$ for which, for **all** sufficiently large real numbers T, the inequalities

$$\Pi_+(\mathbf{q}) < T, \qquad \Pi(A\mathbf{q} - \mathbf{p} - \boldsymbol{\theta}) < T^{-\hat{\omega}}$$

have a solution $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^m \times (\mathbb{Z}^n \setminus \{\mathbf{0}\})$. The homogeneous multiplicative exponents are

$$\omega^{\times}(A) := \omega^{\times}(A, \mathbf{0}), \qquad \hat{\omega}^{\times}(A) := \hat{\omega}^{\times}(A, \mathbf{0}).$$

Remark 1.5. It follows directly from the definitions that for any weights $(\mathbf{s}, \mathbf{r})$ and any $(A, \boldsymbol{\theta})$ we have

$$\omega^{\times}(A, \boldsymbol{\theta}) \geqslant \omega_{\mathbf{s},\mathbf{r}}(A, \boldsymbol{\theta}), \qquad \hat{\omega}^{\times}(A, \boldsymbol{\theta}) \geqslant \hat{\omega}_{\mathbf{s},\mathbf{r}}(A, \boldsymbol{\theta}),$$

and in particular

$$\omega^{\times}(A) \geqslant \omega_{\mathbf{s},\mathbf{r}}(A), \qquad \hat{\omega}^{\times}(A) \geqslant \hat{\omega}_{\mathbf{s},\mathbf{r}}(A).$$

It is well-known that, for $A \in M_{m \times n}(\mathbb{R})$, the Diophantine exponent $\omega(A)$ and that of its transpose $\omega(^tA)$ are related by *transference inequalities*. This was first observed by Khintchine [25] in the $n = 1$ (simultaneous approximation) or $m = 1$ (dual approximation) cases. The following generalisation is due to Dyson [13, Theorem 4], see also [19, Chapter 6, §45, Theorem 8] and [44, Chapter IV, §5].

Theorem 1.6. *Let $A \in M_{m \times n}(\mathbb{R})$. Write $\omega = \omega(A)$ and $^t\omega = \omega(^tA)$. Then*

$$^t\omega \geqslant \frac{n\omega + m - 1}{(n-1)\omega + m} \quad and \quad \omega \geqslant \frac{m \, ^t\omega + n - 1}{(m-1)^t\omega + n}.$$

In particular, we have $\omega = 1$ if and only if $^t\omega = 1$. The inequalities also hold for the uniform exponents with $\omega = \hat{\omega}(A)$ and $^t\omega = \hat{\omega}(^tA)$.

We extend Dyson's transference inequalities to the weighted case. Set

$$\rho_{\mathbf{s}} = \max_{1 \leqslant i \leqslant m} s_i, \quad \delta_{\mathbf{s}} = \min_{1 \leqslant i \leqslant m} s_i, \quad and \quad \rho_{\mathbf{r}} = \max_{1 \leqslant j \leqslant n} r_j, \quad \delta_{\mathbf{r}} = \min_{1 \leqslant j \leqslant n} r_j.$$

Theorem 1.7. *Let $A \in M_{m \times n}(\mathbb{R})$. Write $\omega = \omega_{\mathbf{s},\mathbf{r}}(A)$ and $^t\omega = \omega_{\mathbf{r},\mathbf{s}}(^tA)$. Then*

$$^t\omega \geqslant \frac{(m+n-1)\rho_{\mathbf{s}}\rho_{\mathbf{r}}(\delta_{\mathbf{r}} + \delta_{\mathbf{s}}\omega) + \rho_{\mathbf{s}}\delta_{\mathbf{r}}\delta_{\mathbf{s}}(\omega - 1)}{(m+n-1)\rho_{\mathbf{s}}\rho_{\mathbf{r}}(\delta_{\mathbf{r}} + \delta_{\mathbf{s}}\omega) - \rho_{\mathbf{r}}\delta_{\mathbf{r}}\delta_{\mathbf{s}}(\omega - 1)}$$

and

$$\omega \geqslant \frac{(m+n-1)\rho_{\mathbf{s}}\rho_{\mathbf{r}}(\delta_{\mathbf{s}} + \delta_{\mathbf{r}}{}^t\omega) + \rho_{\mathbf{r}}\delta_{\mathbf{r}}\delta_{\mathbf{s}}(^t\omega - 1)}{(m+n-1)\rho_{\mathbf{s}}\rho_{\mathbf{r}}(\delta_{\mathbf{s}} + \delta_{\mathbf{r}}{}^t\omega) - \rho_{\mathbf{s}}\delta_{\mathbf{r}}\delta_{\mathbf{s}}(^t\omega - 1)}.$$

In particular, $\omega = 1$ if and only if $^t\omega = 1$. The inequalities also hold for the uniform exponents with $\omega = \hat{\omega}_{\mathbf{s},\mathbf{r}}(A)$ and $^t\omega = \hat{\omega}_{\mathbf{r},\mathbf{s}}(^tA)$.

Our next result concerns the following elegant result due to Bugeaud and Laurent [8].

Theorem 1.8 (Bugeaud–Laurent)**.** *For $A \in M_{m \times n}(\mathbb{R})$ and $\boldsymbol{\theta} \in \mathbb{R}^n$ we have*

$$\omega(^tA, \boldsymbol{\theta}) \geqslant \hat{\omega}(A)^{-1}, \qquad \hat{\omega}(^tA, \boldsymbol{\theta}) \geqslant \omega(A)^{-1},$$

with equality for Lebesgue-almost all $\boldsymbol{\theta}$.

The proof in [8] fails to deliver a weighted version. Using a slightly different proof, we are able to establish the following more general assertion.

Theorem 1.9. *If $A \in M_{m \times n}(\mathbb{R})$ and $\boldsymbol{\theta} \in \mathbb{R}^n$ then*

$$\omega_{\mathbf{r},\mathbf{s}}(^tA, \boldsymbol{\theta}) \geqslant \hat{\omega}_{\mathbf{s},\mathbf{r}}(A)^{-1}, \qquad \hat{\omega}_{\mathbf{r},\mathbf{s}}(^tA, \boldsymbol{\theta}) \geqslant \omega_{\mathbf{s},\mathbf{r}}(A)^{-1}, \tag{1.2}$$

with equality for Lebesgue-almost all $\boldsymbol{\theta}$.

The inequalities (1.2) still hold for the natural multiplicative analogue, in view of Remark 1.5, as was already noted in [18]. However, if $\omega^\times(A) > \omega_{\mathbf{s},\mathbf{r}}(A)$ then equality is never attained.

Remark 1.10. The homogeneous case $\boldsymbol{\theta} = \mathbf{0}$ can be quickly seen as follows. There is a weighted form of Dirichlet's approximation theorem (whose proof is essentially the same, using Minkowski's first convex body theorem) which implies that $\omega_{\mathbf{s},\mathbf{r}}(A) \geqslant \hat{\omega}_{\mathbf{s},\mathbf{r}}(A) \geqslant 1$. Hence

$$\omega_{\mathbf{r},\mathbf{s}}({}^{t}A) \geqslant 1 \geqslant \hat{\omega}_{\mathbf{s},\mathbf{r}}(A)^{-1}.$$

In the course of the proof of Theorem 1.9, we extend the theory of best approximations to a weighted setting in §4. As a by-product, we obtain results on the dimension of a certain set of inhomogeneous shifts for weighted ε-badly approximable matrices, namely a weighted generalisation of [7, Theorem 1.5].

Theorem 1.11. *Let* $A \in M_{m \times n}(\mathbb{R})$ *be a matrix for which the group* $G := {}^{t}A\mathbb{Z}^{m} + \mathbb{Z}^{n}$ *has rank*

$$\mathrm{rank}_{\mathbb{Z}}(G) = m + n.$$

Let $(\mathbf{q}_{k})_{k \geqslant 1}$ *be a sequence of weighted best approximations associated to* ${}^{t}A$, *and suppose* $\lim_{k \to \infty} \|\mathbf{q}_{k}\|_{\mathbf{r}}^{1/k} = \infty$. *For* $\varepsilon > 0$, *define*

$$\mathrm{Bad}_{\mathbf{r},\mathbf{s}}^{\varepsilon}({}^{t}A) = \{\boldsymbol{\theta} \in \mathbb{R}^{n} : \liminf_{(\mathbf{p},\mathbf{q}) \in (\mathbb{Z}^{m} \setminus \{\mathbf{0}\}) \times \mathbb{Z}^{n}} \|\mathbf{p}\|_{\mathbf{s}} \|{}^{t}A\mathbf{p} - \mathbf{q} - \boldsymbol{\theta}\|_{\mathbf{r}} \geqslant \varepsilon\}.$$

Then there exists $\varepsilon = \varepsilon(A) > 0$ *such that*

$$\dim_{H} \mathrm{Bad}_{\mathbf{r},\mathbf{s}}^{\varepsilon}({}^{t}A) = n.$$

This is often referred to as *twisted* diophantine approximation: the inhomogeneous shift is metric. The analogous problem for weighted badly approximable matrices has hitherto been investigated in [20, 6]; therein the object of study is $\mathrm{Bad}_{\mathbf{r},\mathbf{s}}({}^{t}A) := \cup_{\varepsilon > 0} \mathrm{Bad}_{\mathbf{r},\mathbf{s}}^{\varepsilon}({}^{t}A)$. Our conclusion is stronger than the assertion that $\dim_{H} \mathrm{Bad}_{\mathbf{r},\mathbf{s}}({}^{t}A) = n$.

Remark 1.12. As noted in [7] and prior works, if the maximal rank condition is not met then ${}^{t}A\mathbf{x} = \mathbf{0}$ possesses infinitely many solutions $\mathbf{x} \in \mathbb{Z}^{m}$, and the theory of best approximations breaks down. This excluded case is less interesting: the conclusion is still valid, since any $\boldsymbol{\theta} \in \mathrm{Bad}_{\mathbf{r},\mathbf{s}}^{\varepsilon}({}^{t}A)$ would need to lie close to a discrete family of parallel hyperplanes in $\mathbb{R}^{n}$. The rank condition will be discussed further in §5, in a similar context.

When $m = 1$ or $n = 1$, the matrix A may be interpreted as a vector. In these special cases, the exponents defined above are the classical exponents of simultaneous and dual approximation; as discussed, these are related by Khintchine's transference inequality [25]. In [34], Laurent introduced intermediate exponents, refining the above quantities.

In a nutshell, the dth intermediate exponent quantifies a vector's proximity to d-dimensional rational subspaces. We can write down a d-dimensional linear subvariety $\mathcal{L}$ of $\mathbb{P}_{\mathbb{R}}^{n}$ using homogeneous coordinates. These span a $(d+1)$-dimensional subspace V of $\mathbb{R}^{n+1}$, and $\mathcal{L} = \mathbb{P}(V)$ is *rational* if V has a rational basis, that is, a basis

$$\{\mathbf{x}_{0}, \ldots, \mathbf{x}_{d}\} \subset \mathbb{Q}^{n+1}.$$

To define the height $H(\mathcal{L})$ of a d-dimensional rational linear subvariety, Schmidt [43] began by using the Plücker embedding

$$\mathrm{Gr}(d, \mathbb{P}^n_{\mathbb{R}}) \hookrightarrow \mathbb{P}^{\binom{n+1}{d+1}-1}_{\mathbb{R}}$$

to obtain Grassmannian coordinates for $\mathcal{L}$. Explicitly, this yields

$$\mathcal{L} \mapsto \mathbf{X} := \mathbf{x}_0 \wedge \cdots \wedge \mathbf{x}_d \in \mathbb{P}^{\binom{n+1}{d+1}-1}_{\mathbb{R}}$$

and, since the basis is rational, we in fact have $\mathbf{X} \in \mathbb{P}^{\binom{n+1}{d+1}-1}_{\mathbb{Q}}$. The *height $H(\mathcal{L})$* of $\mathcal{L}$ is the Weil height $|\mathbf{X}|$ of $\mathbf{X}$ (one rescales the projective coordinates to obtain a primitive integer vector, then evaluates the supremum norm). Schmidt did not work projectively; in this aspect we follow Laurent [34].

The distance generalises the notion of *projective distance* between two points [1, 40, 42]. Recall that there is a unique defined inner product on $\Lambda^t(\mathbb{R}^{n+1})$ such that for any two multivectors

$$\mathbf{u} = \mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_t \text{ and } \mathbf{v} = \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_t$$

we have

$$\langle \mathbf{u}, \mathbf{v} \rangle = \det(\langle \mathbf{u}_i, \mathbf{v}_j \rangle)^t_{i,j=1}, \tag{1.3}$$

see [1, §3]. Then the *Euclidean norm* $|\mathbf{u}|$ of a multivector

$$\mathbf{u} = \mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_t$$

is given by

$$|\mathbf{u}|^2 = |\det(\langle \mathbf{u}_i, \mathbf{u}_j \rangle)^t_{i,j=1}|.$$

For $P, Q \in \mathbb{P}^n_{\mathbb{R}}$, the *projective distance* between P and Q is

$$\mathrm{d}(P, Q) = \frac{|\mathbf{x} \wedge \mathbf{y}|}{|\mathbf{x}| \cdot |\mathbf{y}|},$$

where $\mathbf{x}$ and $\mathbf{y}$ are homogeneous coordinates for P and Q respectively. If $P \in \mathbb{P}^n_{\mathbb{R}}$ then

$$\mathrm{d}(P, \mathcal{L}) := \min_{Q \in \mathcal{L}} \mathrm{d}(P, Q) \tag{1.4}$$

is the least projective distance between P and a point of $\mathcal{L}$.

Definition 1.13. Let d be an integer in the range $0 \leqslant d \leqslant n - 1$, and let $\boldsymbol{\alpha} \in \mathbb{R}^n$. Define the *$d$th ordinary exponent $\omega_d(\boldsymbol{\alpha})$* (resp. the *$d$th uniform exponent $\hat{\omega}_d(\boldsymbol{\alpha})$*) as the supremum of the real numbers ω for which there exist d-dimensional rational linear subspaces $\mathcal{L} \subset \mathbb{R}^n$ such that

$$H(\mathcal{L})^{\binom{n}{d}} \leqslant T, \qquad (T \cdot \mathrm{d}([1 : \boldsymbol{\alpha}], \mathcal{L}))^{\binom{n}{d+1}} \leqslant T^{-\omega}$$

for some arbitrarily large real numbers T (resp. for every sufficiently large real number T).

Here we have chosen the normalisation with powers $\binom{n}{d}$ and $\binom{n}{d+1}$ so that the exponent is generically 1. The cases $d = 0$ and $d = n - 1$ correspond to the simultaneous and dual cases, respectively, see [34].

In [4] an associated inhomogeneous Diophantine exponent $\omega_d(\boldsymbol{\alpha}, \boldsymbol{\theta})$ is posited but not defined, where

$$\boldsymbol{\alpha} \in \mathbb{R}^n, \qquad d \in \{0, 1, \ldots, n-1\}, \qquad \boldsymbol{\theta} \in \mathbb{R}^{n-d}.$$

Then, the following transference inequality is conjectured [4, Conjecture 3].

Conjecture 1.14 (Beresnevich–Velani)**.** *Let $\boldsymbol{\alpha} \in \mathbb{R}^n$ and $d \in \{0, \ldots, n-1\}$. Then for all $\boldsymbol{\theta} \in \mathbb{R}^{n-d}$ we have*

$$\omega_d(\boldsymbol{\alpha}, \boldsymbol{\theta}) \geqslant \frac{1}{\hat{\omega}_{n-1-d}(\boldsymbol{\alpha})}.$$

In [34, §2], Laurent introduced an equivalent definition of $\mathrm{dist}(\boldsymbol{\alpha}, \mathcal{L})$. We will use this to formally define the inhomogeneous intermediate exponents $\omega_d(\boldsymbol{\alpha}, \boldsymbol{\theta})$ in §8, and establish the resulting interpretation of Conjecture 1.14. This applies to a shift $\boldsymbol{\theta} \in \mathbb{R}^{\binom{n}{d+1}}$.

Theorem 1.15. *Let $\boldsymbol{\alpha} \in \mathbb{R}^n$ and $d \in \{0, \ldots, n-1\}$. Then for $\boldsymbol{\theta} \in \mathbb{R}^{\binom{n}{d+1}}$ we have*

$$\omega_d(\boldsymbol{\alpha}, \boldsymbol{\theta}) \geqslant \frac{1}{\hat{\omega}_{n-1-d}(\boldsymbol{\alpha})},$$

with equality for Lebesgue-almost all $\boldsymbol{\theta}$.

After working with the definitions, we will see that this follows directly from Theorem 1.8.

Organisation. In §2, we use the geometry of numbers to establish a property for general approximating functions, which would imply (1.2). Then, in §3, we prove Theorem 1.7. In §4, we extend the theory of best approximations to a weighted setting. This allows us to finish the proof of Theorem 1.9 in §5. Theorem 1.11 is established in §6. In §7, we discuss applications to the theory of inhomogeneous Diophantine approximation on manifolds. Finally, in §8, we define $\omega_d(\mathbf{x}, \boldsymbol{\theta})$ and prove Theorem 1.15.

2. THE GEOMETRY OF NUMBERS

For the proof of Theorem 1.9, the following lemma is pivotal, and we anticipate that it will find uses in other contexts. Evertse has recently pointed out in a survey article [14] that this lemma is implicit in Mahler's work [37] from the late 1930s (in German). For completeness, we supply the details below.

Lemma 2.1. *Let $d \in \mathbb{N}$, and let $C = C_d = d!(3/2)^{\frac{d-1}{2}}d$. Let Λ be a full lattice in $\mathbb{R}^d$, and let $\mathcal{R} \subseteq \mathbb{R}^d$ be a symmetric, convex body such that $\mathcal{R} \cap \Lambda = \{\mathbf{0}\}$. Then for all $\boldsymbol{\gamma} \in \mathbb{R}^d$ we have*

$$(C\mathcal{R}^* + \boldsymbol{\gamma}) \cap \Lambda^* \neq \emptyset,$$

and moreover

$$C\mathcal{R}^* \cap (\Lambda^* \setminus \{\mathbf{0}\}) \neq \emptyset.$$

Here and throughout, we use an asterisk to denote a dual/polar convex set or lattice [10].

Proof. By assumption, the first successive minimum satisfies $\mu_1(\mathcal{R}) > 1$, and so [10, Ch. VIII, Theorem VI] gives

$$\mu_d(\mathcal{R}^*) < d!. \tag{2.1}$$

Now the inequality at the end of the proof of the First Finiteness Theorem in [45, Lecture X, §6] implies that $\frac{C}{d}\mathcal{R}^*$ contains a basis $\mathbf{b}_1^*, \ldots, \mathbf{b}_d^*$ for Λ^*. Hence $C\mathcal{R}^*$ contains $d\mathbf{b}_1^*, \ldots, d\mathbf{b}_d^*$, as well as the origin, and therefore also contains the convex hull of these vectors, which in turn contains the fundamental parallelepiped

$$\{w_1\mathbf{b}_1^* + \cdots + w_d\mathbf{b}_d^* : \mathbf{w} \in [0, 1]^d\}$$

for Λ^*. Therefore any translate of $C\mathcal{R}^*$ intersects Λ^*. The second assertion follows directly from (2.1). $\qquad\square$

Remark 2.2. The reader may consult [23, 24] for sharper and more general results in this direction.

Lemma 2.1 enables us to tackle transference inequalities for more general approximation functions.

Definition 2.3. Let $\psi : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ be a strictly decreasing function. A pair $(A, \boldsymbol{\theta}) \in M_{m \times n}(\mathbb{R}) \times \mathbb{R}^n$ is $(\psi, \mathbf{s}, \mathbf{r})$-*approximable* (resp. *uniformly* $(\psi, \mathbf{s}, \mathbf{r})$-*approximable*) if for some arbitrarily large (resp. all sufficiently large) real numbers T the inequalities

$$\|\mathbf{q}\|_{\mathbf{r}} < T, \qquad \|A\mathbf{q} - \mathbf{p} - \boldsymbol{\theta}\|_{\mathbf{s}} < \psi(T) \tag{2.2}$$

have a solution $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^m \times (\mathbb{Z}^n \setminus \{\mathbf{0}\})$. The matrix $A \in M_{m \times n}(\mathbb{R})$ is $(\psi, \mathbf{s}, \mathbf{r})$-*approximable* (resp. uniformly $(\psi, \mathbf{s}, \mathbf{r})$-approximable) if this holds for $(A, \mathbf{0})$. In the special case $\psi(T) = T^{-\omega}$, we also write (uniformly) $(\omega, \mathbf{s}, \mathbf{r})$-approximable to mean (uniformly) $(\psi, \mathbf{s}, \mathbf{r})$-approximable.

Lemma 2.4. *Let $A \in M_{m \times n}(\mathbb{R})$ and $\boldsymbol{\theta} \in \mathbb{R}^n$. Let $\psi, \phi : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ be strictly decreasing functions with*

$$\lim_{T \to \infty} \psi(T) = \lim_{T \to \infty} \phi(T) = 0, \tag{2.3}$$

and suppose that if C and T are sufficiently large then

$$\phi(C\psi(T)^{-1}) > CT^{-1}. \tag{2.4}$$

If A is not $(\psi, \mathbf{s}, \mathbf{r})$-approximable, then $(^tA, \boldsymbol{\theta})$ is uniformly $(\phi, \mathbf{r}, \mathbf{s})$-approximable for all $\boldsymbol{\theta} \in \mathbb{R}^n$. If A is not uniformly $(\psi, \mathbf{s}, \mathbf{r})$-approximable, then $(^tA, \boldsymbol{\theta})$ is $(\phi, \mathbf{r}, \mathbf{s})$-approximable for all $\boldsymbol{\theta} \in \mathbb{R}^n$.

Proof. For $T > 0$, write

$$g_{\mathbf{r}}(T) = \mathrm{diag}(T^{r_1}, \ldots, T^{r_n}) \in \mathrm{GL}_n(\mathbb{R}),$$

and define $g_{\mathbf{s}}(T) \in \mathrm{GL}_m(\mathbb{R})$ similarly. For $Q, T > 0$ and $A \in M_{m \times n}(\mathbb{R})$, we define a full lattice in $\mathbb{R}^{m+n}$ as follows,

$$\Lambda(Q, T, A) = \begin{pmatrix} g_{\mathbf{s}}(Q^{-1}) & \\ & g_{\mathbf{r}}(T^{-1}) \end{pmatrix} \begin{pmatrix} \mathbf{I}_m & A \\ & \mathbf{I}_n \end{pmatrix} \mathbb{Z}^{m+n}.$$

Observe that (2.2) has a non-zero integer solution with $\boldsymbol{\theta} = \mathbf{0}$ if and only if

$$\Lambda(\psi(T), T, A) \cap \mathcal{B} \neq \{\mathbf{0}\},$$

where $\mathcal{B} = [-1, 1]^{m+n}$. The dual region is

$$\mathcal{B}^* = \left\{ (x_1, \ldots, x_{m+n}) \in \mathbb{R}^{m+n} : \sum_{i=1}^{m+n} |x_i| \leqslant 1 \right\},$$

and in particular $\mathcal{B}^* \subset \mathcal{B}$.

We now set about proving the first assertion. The proof of the second statement is similar, and omitted. Let T be a large positive real number, and let $\boldsymbol{\theta} \in \mathbb{R}^n$. If A is not $(\psi, \mathbf{s}, \mathbf{r})$-approximable then (2.2) has no solution with $\boldsymbol{\theta} = \mathbf{0}$. In light of the discussion above, we have

$$\Lambda(\psi(T), T, A) \cap \mathcal{B} = \{\mathbf{0}\},$$

and now Lemma 2.1 yields

$$\Lambda(\psi(T), T, A)^* \cap (C\mathcal{B} + \boldsymbol{\gamma}) \neq \emptyset \qquad (\boldsymbol{\gamma} \in \mathbb{R}^{m+n}). \tag{2.5}$$

A standard calculation gives

$$\Lambda(\psi(T), T, A)^* = \begin{pmatrix} g_{\mathbf{s}}(\psi(T)) & \\ & g_{\mathbf{r}}(T) \end{pmatrix} \begin{pmatrix} \mathbf{I}_m & \\ -^tA & \mathbf{I}_n \end{pmatrix} \mathbb{Z}^{m+n}.$$

Applying (2.5) with $\boldsymbol{\gamma} = \boldsymbol{\gamma}_{T,\boldsymbol{\theta}} := (\mathbf{0}, g_{\mathbf{r}}(T)\boldsymbol{\theta}) \in \mathbb{R}^{m+n}$, we find that the inequalities

$$\|\mathbf{p}\|_{\mathbf{s}} < C_1\psi(T)^{-1}, \qquad \|^tA\mathbf{p} - \mathbf{q} - \boldsymbol{\theta}\|_{\mathbf{r}} < C_1T^{-1} \tag{2.6}$$

have a solution $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^{m+n}$. Here $C_1 = C^{\frac{1}{\delta_{\mathbf{s}}}}$.

We claim that $\mathbf{p}$ can be chosen to be non-zero. There are two cases to consider. When $\boldsymbol{\theta} \notin \mathbb{Z}^n$, we know that $\|\mathbf{q} - \boldsymbol{\theta}\|_{\mathbf{r}}$ is bounded away from 0, so as T is large the second inequality of (2.6) cannot be satisfied when $\mathbf{p} = \mathbf{0}$. On

the other hand, when $\boldsymbol{\theta} \in \mathbb{Z}^n$, we may freely suppose that $\boldsymbol{\theta} = \mathbf{0}$, whereupon the second part of Lemma 2.1 allows us to take $(\mathbf{p}, \mathbf{q}) \neq (\mathbf{0}, \mathbf{0})$, and then the largeness of T forces $\mathbf{p} \neq \mathbf{0}$. Using (2.3) and (2.4), we finally conclude that if T_1 is large then the inequalities

$$\|\mathbf{p}\|_{\mathbf{s}} < T_1, \qquad \|{}^t A \mathbf{p} - \mathbf{q} - \boldsymbol{\theta}\|_{\mathbf{r}} < \phi(T_1)$$

have a solution $(\mathbf{p}, \mathbf{q}) \in (\mathbb{Z}^m \setminus \{\mathbf{0}\}) \times \mathbb{Z}^n$. Therefore $({}^t A, \boldsymbol{\theta})$ is uniformly $(\phi, \mathbf{r}, \mathbf{s})$-approximable. $\square$

3. Dyson's theorem with weights

The purpose of this section is to prove Theorem 1.7. As in the unweighted case, this essentially follows from Minkowski's second convex body theorem.

By symmetry, it suffices to prove the first assertion of the theorem. For $t_1, t_2 \in \mathbb{R}$, set

$$L(t_1, t_2, A) = \Lambda(e^{t_1}, e^{t_2}, A), \qquad h(t_1, t_2) = \operatorname{diag}(g_{\mathbf{s}}(e^{t_1}), g_{\mathbf{r}}(e^{t_2})).$$

Let $v < \omega$. By definition, for some arbitrarily large t, we have

$$L(-vt, t, A) \cap \mathcal{B} \neq \{\mathbf{0}\},$$

where $\mathcal{B} = [-1, 1]^{m+n}$, or equivalently

$$h\left(\delta_{\mathbf{r}}(\delta_{\mathbf{s}} + \delta_{\mathbf{r}})^{-1}(1 - v)t, \delta_{\mathbf{s}}(\delta_{\mathbf{s}} + \delta_{\mathbf{r}})^{-1}(1 - v)t\right)\left(L(-vt, t, A) \cap \mathcal{B}\right) \neq \{\mathbf{0}\}.$$

It then follows that

$$L(-t_1, t_1, A) \cap h\left(\delta_{\mathbf{r}}(\delta_{\mathbf{s}} + \delta_{\mathbf{r}})^{-1}(1 - v)t, \delta_{\mathbf{s}}(\delta_{\mathbf{s}} + \delta_{\mathbf{r}})^{-1}(1 - v)t\right)\mathcal{B} \neq \{\mathbf{0}\},$$

where $t_1 = (\delta_{\mathbf{s}} + \delta_{\mathbf{r}})^{-1}(\delta_{\mathbf{r}} + \delta_{\mathbf{s}} v)t$. Hence

$$\mu_1(L(-t_1, t_1, A), \mathcal{B}) \leqslant e^{-t_0},$$

where

$$t_0 = (\delta_{\mathbf{s}} + \delta_{\mathbf{r}})^{-1}\delta_{\mathbf{s}}\delta_{\mathbf{r}}(v - 1)t = (\delta_{\mathbf{s}} v + \delta_{\mathbf{r}})^{-1}\delta_{\mathbf{s}}\delta_{\mathbf{r}}(v - 1)t_1.$$

By [10, Ch. VIII, Theorem VI], we now have

$$\mu_{m+n}(L(-t_1, t_1, A)^*, \mathcal{B}) \geqslant e^{t_0}.$$

Minkowski's second convex body theorem [9, Appendix B, Theorem V] now gives

$$\mu_1(L(-t_1, t_1, A)^*, \mathcal{B}) \leqslant Ce^{-(m+n-1)^{-1}t_0},$$

for some constant $C > 0$ depending only on $m + n$, and so

$$\begin{pmatrix} g_{\mathbf{s}}(e^{-t_1}) & \\ & g_{\mathbf{r}}(e^{t_1}) \end{pmatrix}\begin{pmatrix} \mathbf{I}_m & \\ -{}^t A & \mathbf{I}_n \end{pmatrix}\mathbb{Z}^{m+n} \cap e^{-(m+n-1)^{-1}t_0 + \log C}\mathcal{B} \neq \{\mathbf{0}\}.$$

Therefore

$$L(-t', t'', {}^t A) \cap \mathcal{B} \neq \{\mathbf{0}\},$$

where

$$t' = t_1 - \rho_{\mathbf{s}}^{-1}((m+n-1)^{-1}t_0 - \log C) \text{ and } t'' = t_1 + \rho_{\mathbf{r}}^{-1}((m+n-1)^{-1}t_0 - \log C).$$

Since C is constant, we get

$$^t\omega \geqslant \frac{1 + \rho_{\mathbf{r}}^{-1} t_1^{-1} (m+n-1)^{-1} t_0}{1 - \rho_{\mathbf{s}}^{-1} t_1^{-1} (m+n-1)^{-1} t_0}.$$

As $v < \omega$ is arbitrary, a direct calculation completes the proof of Theorem 1.7.

If we consider uniform exponents, the proof is the very same by choosing **any** sufficiently large t, instead of **some** arbitrarily large t.

Remark 3.1. We do not know whether Theorem 1.7 is optimal for every choice of weights. The special case in which the weights are uniform, namely Theorem 1.6, is however known to be optimal: Jarník [22] established this in quite some generality, for example if

$$1 \leqslant m \leqslant n, \qquad 1 \leqslant \omega \leqslant \infty$$

then the first inequality in Theorem 1.6 is sharp. All of this is discussed more broadly after the proof of [19, Chapter 6, §45, Theorem 8]; the reader should be wary of the difference in normalisation therein.

4. Best approximations

When working in higher dimensions, the theory of *best approximations* often acts as a proxy for the theory of continued fractions. Best approximations were introduced by Voronoi [49] as minimal points in lattices, and Rogers [41] was the first to define them in the context of exponents of Diophantine approximation. We require a weighted version of the best approximations employed in [8, §3]. The properties presented therein generalise cleanly to a broad setting, which includes the weighted case. We supply full details for completeness, closely following [8].

Definition 4.1. Let Λ be a lattice in a real vector space. Let $N, L : \Lambda \to [0, \infty)$ be functions such that

(i) L attains its minimum on sets of the form

$$\{\mathbf{X} \in \Lambda \setminus \{\mathbf{0}\} : N(\mathbf{X}) \leqslant B\} \text{ for any } B \in \mathbb{R}_{\geqslant 0},$$

(ii)
$$L(\mathbf{X}) \neq \mathbf{0} \qquad (\mathbf{0} \neq \mathbf{X} \in \Lambda), \tag{4.1}$$

(iii) and
$$\inf_{\mathbf{X} \in \Lambda \setminus \{\mathbf{0}\}} L(\mathbf{X}) = 0. \tag{4.2}$$

A *sequence of (N, L)-best approximations* is $(\mathbf{X}_i)_{i=1}^\infty \in \Lambda^{\mathbb{N}}$ such that

(i) the sequence $N(\mathbf{X}_1), N(\mathbf{X}_2), \ldots$ is strictly increasing;
(ii) the sequence $L(\mathbf{X}_1), L(\mathbf{X}_2), \ldots$ is strictly decreasing;
(iii) for every $\mathbf{X} \in \Lambda \setminus \{\mathbf{0}\}$,

$$\text{if } N(\mathbf{X}) < N(\mathbf{X}_{i+1}) \text{ then } L(\mathbf{X}) \geqslant L(\mathbf{X}_i). \tag{4.3}$$

Informally, the function N (usually a height) measures the size of a point in Λ. The function L will, in practice, depend on some point in E that we wish to approximate; it measures the quantitative strength of the best approximations. The condition $\inf_{\mathbf{X} \in \Lambda \setminus \{\mathbf{0}\}} L(\mathbf{X}) = 0$ ensures that good approximations exist at all.

In the context of weighted exponents, given $A \in M_{m \times n}(\mathbb{R})$ we shall consider

$$\Lambda = \mathbb{Z}^m, \qquad N(\mathbf{X}) = \|\mathbf{X}\|_{\mathbf{s}}, \qquad L(\mathbf{X}) = \inf_{\mathbf{p} \in \mathbb{Z}^n} \|{}^t A \mathbf{X} - \mathbf{p}\|_{\mathbf{r}}, \qquad (4.4)$$

and call $\mathbf{X}_1, \mathbf{X}_2, \dots$ *weighted best approximations* for A. We will require an additional hypothesis on A to ensure that

$$\{{}^t A \mathbf{X} : \mathbf{X} \in \mathbb{Z}^m \setminus \{\mathbf{0}\}, \mathbf{p} \in \mathbb{Z}^n\}$$

does not contain the origin, but contains points arbitrarily close to it.

Observe that the requirement (i) on N and L is met if N possesses the *Northcott property*

$$\#\{\mathbf{X} \in \Lambda : N(\mathbf{X}) \leqslant B\} < \infty \text{ for any } B \in \mathbb{R}_{\geqslant 0}.$$

For example, one may consider best approximation vectors and exponents of best approximation for Λ being the ring of integers of a Northcott field.

For $i \in \mathbb{N}$ we write $Y_i = N(\mathbf{X}_i)$ and $M_i = L(\mathbf{X}_i)$. The sequences $(Y_i)_{i=1}^\infty$ and $(M_i)_{i=1}^\infty$ are respectively strictly increasing and strictly decreasing, and furthermore

$$\lim_{i \to \infty} Y_i = \infty, \qquad \lim_{i \to \infty} M_i = 0.$$

We will see from the construction below that there can be several distinct sequences of (N, L)-best approximations. Notwithstanding, the sequences $(Y_i)_{i>0}$ and $(M_i)_{i>0}$ are uniquely determined.

We now demonstrate, by construction, the existence of a sequence of (N, L)-best approximations. When $i = 0$, let M_0 be the minimum of L on the set of $\mathbf{X} \in \Lambda \setminus \{\mathbf{0}\}$ such that $N(\mathbf{X}) \leqslant 1$, and choose $\mathbf{X}_0$ to be a point were this minimum in reached. Neither $\mathbf{X}_0$ nor $Y_0 := N(\mathbf{X}_0)$ is uniquely determined. Next, suppose that $\mathbf{X}_1, \dots, \mathbf{X}_k$ have already been chosen, that (4.3) holds for $i \leqslant k - 1$, and that

$$\text{if } \mathbf{X} \in \Lambda \setminus \{\mathbf{0}\} \text{ and } N(\mathbf{X}) < N(\mathbf{X}_k) \text{ then } L(\mathbf{X}) \geqslant L(\mathbf{X}_k).$$

Let $Y > N(\mathbf{X}_k)$ be minimal such that

$$\min\{L(\mathbf{X}) : \mathbf{X} \in \Lambda, \ N(\mathbf{X}) \leqslant Y\} < M_k.$$

This minimum is well-defined, since $\inf_{\mathbf{X} \in \Lambda} L(\mathbf{X}) = 0$ and L is strictly positive. Putting $Y_{k+1} = Y$, by the definition of Y there exists a (non-unique) point $\mathbf{X}_{k+1}$ such that $Y_{k+1} = N(\mathbf{X}_{k+1})$ and $L(\mathbf{X}_{k+1}) = M_{k+1} < M_k$ satisfy (4.3). From our recursive definition, it is clear that the sequence $\mathbf{X}_1, \mathbf{X}_2, \dots$ so constructed is a sequence of (N, L)-best approximations.

For a sequence of (N, L)-best approximations, define the exponents

$$\omega_{N,L} := \limsup_{i \to \infty} \frac{\log(M_i)}{\log(Y_i)}, \qquad \hat{\omega}_{N,L} := \liminf_{i \to \infty} \frac{\log(M_i)}{\log(Y_{i+1})}.$$

The quantities are well-defined, and compatible with our previous definitions.

Lemma 4.2. *In the setting (4.4), we have*

$$\omega_{\mathbf{r},\mathbf{s}}(^tA) = \omega_{N,L}, \qquad \hat{\omega}_{\mathbf{r},\mathbf{s}}(^tA) = \hat{\omega}_{N,L}.$$

Proof. From the definitions, if $M_i \leqslant Y_i^{-\omega}$ for infinitely many indices i, then $\omega \leqslant \omega_{\mathbf{r},\mathbf{s}}(^tA)$. This shows that $\omega_{\mathbf{r},\mathbf{s}}(^tA) \geqslant \omega_{N,L}$.

Conversely, for any $\omega < \omega_{\mathbf{r},\mathbf{s}}(^tA)$, we can choose $T > 0$ arbitrarily large such that

$$\|^tA\mathbf{p} - \mathbf{q}\|_{\mathbf{r}} < T^{-\omega}, \qquad \|\mathbf{p}\|_{\mathbf{s}} < T$$

for some $(\mathbf{p}, \mathbf{q}) \in (\mathbb{Z}^m \setminus \{\mathbf{0}\}) \times \mathbb{Z}^n$. Let k be the index such that $Y_k \leqslant T < Y_{k+1}$. From the definition of weighted best approximations, we have

$$M_k \leqslant \|^tA\mathbf{p} - \mathbf{q}\|_{\mathbf{r}} < T^{-\omega} \leqslant Y_k^{-\omega}.$$

We conclude that $\omega_{N,L} \geqslant \omega_{\mathbf{r},\mathbf{s}}(^tA)$.

We have proved the first assertion, and the second follows by a similar argument. $\qquad\square$

Next we show, under a further assumption on L and N, that the sequence $(Y_i)_{i=1}^{\infty}$ exhibits geometric growth, generalising the fact that continued fraction denominators enjoy this property. This is well-understood in the context of unweighted best rational approximations, see [8, Lemma 1] and [32, 33].

Lemma 4.3. *Suppose that L and N satisfy the slack triangle inequalities*

$$N(\lambda(\mathbf{a} + \mathbf{b})) \leqslant \lambda^{\delta}(N(\mathbf{a}) + N(\mathbf{b})) \qquad (\lambda \leqslant 1)$$

$$L(U^{-1}(\mathbf{a} + \mathbf{b})) \leqslant \frac{1}{2}(L(\mathbf{a}) + L(\mathbf{b})) \tag{4.5}$$

for some $\delta > 0$ and all $U \geqslant U_0(\delta)$. Then there exist $c > 0$ and $\gamma > 1$ such that

$$Y_i \geqslant c\gamma^i \qquad (i \in \mathbb{N}).$$

Remark 4.4. Note that the inequalities (4.5) hold in the setting (4.4) of weighted exponents, with $\delta = \min\{\delta_{\mathbf{s}}, \delta_{\mathbf{r}}\} = \min\{s_i, r_j : 1 \leqslant i \leqslant m, \ 1 \leqslant j \leqslant n\}$.

Proof. Let us take $U \in \mathbb{N}$ constant, but large enough to ensure that

$$U^{\delta} > 3. \tag{4.6}$$

Consider any $V := 2U^n$ consecutive vectors $\mathbf{X}_{i+1}, \ldots, \mathbf{X}_{i+V}$. By the pigeonhole principle, there exist indices j, k with $1 \leqslant j < k \leqslant V$ such that $\mathbf{X}_{i+j} - \mathbf{X}_{i+k} \in U\mathbb{Z}^n$. By the slack triangle inequality, the vector

$$\mathbf{X} = U^{-1}(\mathbf{X}_{i+k} - \mathbf{X}_{i+j})$$

satisfies

$$N(\mathbf{X}) = N(U^{-1}(\mathbf{X}_{i+k} - \mathbf{X}_{i+j})) \leqslant U^{-\delta}(Y_{i+j} + Y_{i+k}) \leqslant \frac{1}{3}(Y_{i+j} + Y_{i+k})$$

and

$$L(\mathbf{X}) = L(U^{-1}(\mathbf{X}_{i+k} - \mathbf{X}_{i+j})) \leqslant \frac{1}{3}(M_{i+j} + M_{i+k}) < M_{i+j}.$$

By the definition of (N, L)-best approximations, we must have

$$Y_{i+j+1} \leqslant N(\mathbf{X}) \leqslant \frac{1}{3}(Y_{i+j} + Y_{i+k}),$$

and consequently

$$Y_{i+V} \geqslant Y_{i+k} \geqslant 3Y_{i+j+1} - Y_{i+j} \geqslant 2Y_{i+j+1} \geqslant 2Y_i \qquad (i \in \mathbb{N}).$$

With $\gamma = 2^{1/V}$, the inequality $Y_i \gg \gamma^i$ now follows by induction. $\qquad \square$

5. The Bugeaud–Laurent theorem with weights

Our objective in this section is to establish Theorem 1.9. First and foremost, we use Lemma 2.4 to deduce (1.2). By definition, if $\hat{\omega} > \hat{\omega}_{\mathbf{s},\mathbf{r}}(A)$ then A is not uniformly $(\hat{\omega}, \mathbf{s}, \mathbf{r})$-approximable. Observe that the functions $\phi(T) = T^{-\omega}$ and $\psi(T) = T^{-\hat{\omega}}$ satisfy the condition (2.4) whenever $\omega\hat{\omega} > 1$. Therefore $({}^tA, \boldsymbol{\theta})$ is $(\omega, \mathbf{r}, \mathbf{s})$-approximable for all $\boldsymbol{\theta} \in \mathbb{R}^n$, by Lemma 2.4. Since $\omega, \hat{\omega}$ are arbitrary real numbers for which $\omega_{\mathbf{r},\mathbf{s}}({}^tA, \boldsymbol{\theta}) \geqslant \omega > \hat{\omega}^{-1}$, this confirms the second inequality in (1.2). The proof of the first inequality is similar.

Next, we show that for any fixed A, equality holds in the first inequality of (1.2) for almost all $\boldsymbol{\theta} \in \mathbb{R}^n$; the analogous statement for the second inequality will follow by similar reasoning. We will apply the theory of weighted best approximations in the setting (4.4), when $A \in M_{m \times n}(\mathbb{R})$ has the property that $G := {}^tA\mathbb{Z}^m + \mathbb{Z}^n \leqslant \mathbb{R}^m$ has maximal rank $m + n$ as a group. This implies (4.1) and, by Kronecker's theorem [9, Chapter III, Theorem IV], also ensures that the condition (4.2) is met.

When G does not have maximal rank, the observation made by Bugeaud and Laurent [8] still applies in our weighted framework: the exponents $\omega_{\mathbf{r},\mathbf{s}}({}^tA, \boldsymbol{\theta})$ and $\hat{\omega}_{\mathbf{r},\mathbf{s}}({}^tA, \boldsymbol{\theta})$ vanish unless $\boldsymbol{\theta}$ lies in a discrete family of parallel hyperplanes in $\mathbb{R}^n$. The upshot is that, in this case, equality in (1.2) certainly holds almost surely.

We proceed on the assumption that G has maximal rank. It suffices to prove that if $0 < \varepsilon < \hat{\omega}_{\mathbf{s},\mathbf{r}}(A)$ then the set of $\boldsymbol{\theta} \in [0, 1]^n$ satisfying

$$\omega_{\mathbf{r},\mathbf{s}}({}^tA, \boldsymbol{\theta}) > \frac{1 + \varepsilon}{\hat{\omega}_{\mathbf{s},\mathbf{r}}(A) - \varepsilon} + \varepsilon \tag{5.1}$$

has Lebesgue measure 0. Denote by $\langle \cdot, \cdot \rangle$ the inner product on $\mathbb{R}^n$, that is, write

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n.$$

We compute that

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leqslant \sum_{i=1}^{n} |x_i y_i| \leqslant \sum_{i=1}^{n} \|\mathbf{x}\|_{\mathbf{r}}^{r_i} \|\mathbf{y}\|_{\mathbf{r}}^{r_i} \leqslant n \max\{\|\mathbf{x}\|_{\mathbf{r}}^{\delta}\|\mathbf{y}\|_{\mathbf{r}}^{\delta}, \|\mathbf{x}\|_{\mathbf{r}}\|\mathbf{y}\|_{\mathbf{r}}\},$$

where δ is as in Remark 4.4.

Fix a positive real number $\varepsilon < \hat{\omega}_{\mathbf{s},\mathbf{r}}(A)$, as well as a sequence $\mathbf{q}_1, \mathbf{q}_2, \ldots$ of weighted best approximations for ^tA, and for $k \in \mathbb{N}$ write $Y_k = \|\mathbf{q}_k\|_{\mathbf{r}}$ as before. We claim that for $\eta = \frac{1}{2}\delta\varepsilon > 0$, the set of $\boldsymbol{\theta} \in [0,1]^n$ satisfying (5.1) is contained in $\limsup_{k \to \infty} S_k$, where

$$S_k = \{\mathbf{y} \in [0,1]^n : \operatorname{dist}(\langle \mathbf{y}, \mathbf{q}_k \rangle, \mathbb{Z}) < Y_k^{-\eta}\} \qquad (n \in \mathbb{N}).$$

It is easy to check that the Lebesgue measure of S_k does not exceed $2nY_k^{-\eta}$. Thus, by Lemma 4.3 and the first Borel–Cantelli lemma, the theorem will follow from our claim. It remains to confirm the claim.

Suppose $\boldsymbol{\theta} \in [0,1]^n$ satisfies (5.1), and put

$$\hat{\omega} = \hat{\omega}_{\mathbf{s},\mathbf{r}}(A) - \varepsilon, \qquad \omega = \frac{1+\varepsilon}{\hat{\omega}} + \varepsilon.$$

By (5.1) we have $\omega < \omega_{\mathbf{r},\mathbf{s}}(^tA, \boldsymbol{\theta})$, so we may select an arbitrarily large positive real number T such that the inequalities

$$\|\mathbf{p}\|_{\mathbf{s}} < T, \qquad \|^tA\mathbf{p} - \mathbf{q} - \boldsymbol{\theta}\|_{\mathbf{s}} < T^{-\omega}$$

have a solution $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^m \times (\mathbb{Z}^n \setminus \{\mathbf{0}\})$. Let k be the unique index for which $Y_k \leqslant T^{\omega-\varepsilon} < Y_{k+1}$, so that Lemma 4.2 gives

$$M_k < Y_{k+1}^{-\hat{\omega}} < T^{-\hat{\omega}(\omega-\epsilon)} = T^{-1-\varepsilon}.$$

Since A is the Hermitian adjoint of ^tA, we obtain

$$\begin{aligned}
\langle \mathbf{q}_k, \boldsymbol{\theta} \rangle &= \langle \mathbf{q}_k, {}^tA\mathbf{p} \rangle + \langle \mathbf{q}_k, \mathbf{q} \rangle - \langle \mathbf{q}_k, {}^tA\mathbf{p} - \mathbf{q} - \boldsymbol{\theta} \rangle \\
&\equiv \langle A\mathbf{q}_k - \mathbf{p}_k, \mathbf{p} \rangle - \langle \mathbf{q}_k, {}^tA\mathbf{p} - \mathbf{q} - \boldsymbol{\theta} \rangle \qquad \mod 1, \qquad (5.2)
\end{aligned}$$

where $\mathbf{p}_k$ is the integer vector nearest to $A\mathbf{q}_k$. Moreover, as T is large we have

$$|\langle A\mathbf{q}_k - \mathbf{p}_k, \mathbf{p} \rangle| \leqslant m \cdot \max\{M_k^\delta \|\mathbf{p}\|_{\mathbf{s}}^\delta, M_k\|\mathbf{p}\|_{\mathbf{s}}\} \leqslant mT^{-\delta\varepsilon} < \frac{1}{2}T^{-\eta}$$

and

$$|\langle \mathbf{q}_k, {}^tA\mathbf{p} - \mathbf{q} - \boldsymbol{\theta} \rangle| \leqslant n \cdot \max\{Y_k^\delta T^{-\omega\delta}, Y_k T^{-\omega}\} \leqslant nT^{-\delta\varepsilon} < \frac{1}{2}T^{-\eta}.$$

The triangle inequality now reveals that $\boldsymbol{\theta} \in \limsup_{k \to \infty} S_k$, establishing our claim and hence the theorem.

6. Weighted inhomogeneous Bad

In this section, we establish Theorem 1.11, closely following [7]. We begin by recalling the setup. For $\varepsilon > 0$, set

$$\operatorname{Bad}_{\mathbf{r},\mathbf{s}}^\varepsilon(^tA) = \{\boldsymbol{\theta} \in \mathbb{R}^n : \liminf_{(\mathbf{p},\mathbf{q}) \in (\mathbb{Z}^m \setminus \{\mathbf{0}\}) \times \mathbb{Z}^n} \|\mathbf{p}\|_{\mathbf{s}} \|^tA\mathbf{p} - \mathbf{q} - \boldsymbol{\theta}\|_{\mathbf{r}} \geqslant \varepsilon\},$$

and

$$\operatorname{Bad}_{\mathbf{r},\mathbf{s}}(^tA) = \bigcup_{\varepsilon > 0} \operatorname{Bad}_{\mathbf{r},\mathbf{s}}^\varepsilon(^tA).$$

It is known that $\mathrm{Bad}_{\mathbf{r},\mathbf{s}}({}^t A)$ has full Hausdorff dimension, see [31], however the task of determining the Hausdorff dimension of $\mathrm{Bad}_{\mathbf{r},\mathbf{s}}^{\varepsilon}({}^t A)$ is much more delicate. The case of vectors in $\mathbb{R}^n$ has been recently studied in [35], where it is proved that for $\varepsilon > 0$ and for an explicit class of vectors $\mathbf{v}$ termed *heavy*, we have

$$\dim(\mathrm{Bad}_{\mathbf{r},\mathbf{s}}^{\varepsilon}(\mathbf{v})) < n. \tag{6.1}$$

Heavy vectors form a set of full Lebesgue measure.

Subsequently, the unweighted sets $\mathrm{Bad}^{\varepsilon}(A)$ were investigated in [7], where necessary and sufficient conditions were obtained in dimension 1, so that $\dim(\mathrm{Bad}^{\varepsilon}(v)) = 1$ for some $\varepsilon > 0$. These conditions were expressed in terms of the continued fraction expansion of v, and were shown to be equivalent to v being "singular on average". In higher dimensions, [7, Theorem 1.5] states sufficient conditions, in terms of best approximation vectors for a matrix A, to ensure that $\dim(\mathrm{Bad}^{\varepsilon}(A)) = n$ for some $\varepsilon > 0$. Theorem 1.11 is a weighted extension of the latter result.

We begin by proving a weighted version of [7, Theorem 5.1].

Theorem 6.1. *For any $\alpha \in (0, 1/2)$, there exists $R = R(\alpha) > 1$ with the following property. Let $(\mathbf{y}_k)_{k \geqslant 1}$ be a sequence in $\mathbb{R}^n \backslash \{\mathbf{0}\}$ such that $\|\mathbf{y}_{k+1}\|_{\mathbf{r}}/\|\mathbf{y}_k\|_{\mathbf{r}} \geqslant R$ for all $k \geqslant 1$ and $\lim_{k \to \infty} \|\mathbf{y}_k\|_{\mathbf{r}}^{1/k} = \infty$. Then the set*

$$S_{\alpha} := \{\boldsymbol{\theta} \in [0, 1]^n : \mathrm{dist}(\langle \mathbf{y}_k, \boldsymbol{\theta} \rangle, \mathbb{Z}) \geqslant \alpha \text{ for all } k \geqslant 1\}$$

has Hausdorff dimension n.

Proof. Let $\alpha \in (0, 1/2)$ be fixed. Choose R large enough that

$$c := 1 - 2\alpha - 3nR^{-\delta_{\mathbf{r}}} > 0.$$

We proceed to verify the theorem for such R. By the mass distribution principle [15, Chapter 4], it suffices to demonstrate the existence of a measure μ, supported on S_{α}, with the following property: if $\varepsilon > 0$ then there exist $C(\varepsilon), r_0(\varepsilon) > 0$ such that for any Euclidean ball B of radius $r \in (0, r_0(\varepsilon))$ we have

$$\mu(B) \leqslant C(\varepsilon) r^{n-\varepsilon}. \tag{6.2}$$

The measure μ will be constructed in a standard way. Write

$$Y_k = \|\mathbf{y}_k\|_{\mathbf{r}}, \qquad Z_{k,\alpha} = \{\boldsymbol{\theta} \in [0, 1]^n : \mathrm{dist}(\langle \mathbf{y}_k, \boldsymbol{\theta} \rangle, \mathbb{Z}) < \alpha\}.$$

The first step is to construct two sequences $(\mathcal{I}_k)_{k \geqslant 1}$ and $(\mathcal{J}_k)_{k \geqslant 1}$ of collections of subsets of $I_0 := [0, 1]^n$. Set $\mathcal{I}_0 = \{I_0\}$. The sequences $(\mathcal{I}_k)_{k \geqslant 1}$ and $(\mathcal{J}_k)_{k \geqslant 1}$ will be defined recursively, so that $\mathcal{J}_k \subset \mathcal{I}_k$, and so that $\mathcal{I}_k$ comprises a collection of translates of $\Pi(Y_k)$, where

$$\Pi(Y_k) = [0, Y_k^{-r_1}] \times \cdots \times [0, Y_k^{-r_n}].$$

It is easily seen that there exists a collection $\mathcal{P}_{k+1}$ of translates of $\Pi(Y_{k+1})$ satisfying:

    (1) Elements from $\mathcal{P}_{k+1}$ are subsets of $\Pi(Y_k)$ and have mutually disjoint interiors;

(2)
$$\Delta_{k+1} := \#\mathcal{P}_{k+1} = \prod_{1 \leqslant i \leqslant n} \lfloor Y_k^{-r_i} Y_{k+1}^{r_i} \rfloor.$$

We record the following lower bound on Δ_{k+1} for later use:

$$\Delta_{k+1} \geqslant \prod_{1 \leqslant i \leqslant n} (Y_k^{-r_i} Y_{k+1}^{r_i} - 1) \geqslant Y_k^{-1} Y_{k+1} (1 - \sum_{1 \leqslant i \leqslant n} Y_k^{r_i} Y_{k+1}^{-r_i})$$

$$\geqslant (1 - nR^{-\delta_{\mathbf{r}}}) Y_k^{-1} Y_{k+1}.$$

Now assume that $\mathcal{I}_k$ and $\mathcal{J}_k$ have been defined. For any $I = \boldsymbol{\theta} + \Pi(Y_k) \in \mathcal{I}_k$, set

$$\mathcal{I}(I) = \{\boldsymbol{\theta} + I_1 : I_1 \in \mathcal{P}_{k+1}\} \text{ and } \mathcal{J}(I) = \{I_2 \in \mathcal{I}(I) : I_2 \cap Z_{k,\alpha} = \emptyset\}.$$

Then choose

$$\mathcal{I}_{k+1} = \bigcup_{I \in \mathcal{I}_k} \mathcal{I}(I) \text{ and } \mathcal{J}_{k+1} = \bigcup_{I \in \mathcal{J}_k} \mathcal{J}(I).$$

For any $I \in \mathcal{I}_k$, we have $\#\mathcal{I}(I) = \Delta_{k+1}$. Next, we estimate $\#\mathcal{J}(I)$. Observe that if $I_3 \in \mathcal{I}(I)$, with $I_3 \cap Z_{k,\alpha} \neq \emptyset$, then $I_3 \subset Z_{k,\beta}$, where

$$\beta = \alpha + \sum_{1 \leqslant i \leqslant n} Y_k^{r_i} Y_{k+1}^{-r_i} \leqslant \alpha + nR^{-\delta_{\mathbf{r}}}.$$

Hence

$$\#\mathcal{J}(I) \geqslant \Delta_{k+1} - \frac{\lambda(Z_{k,\beta} \cap I)}{\lambda(I_3)} \geqslant \Delta_{k+1} - 2\beta Y_k^{-1} Y_{k+1} \geqslant cY_k^{-1} Y_{k+1},$$

where λ denotes Lebesgue measure on $\mathbb{R}^n$. Consequently, we have the lower bound

$$\#\mathcal{J}_k \geqslant c^{k-1} Y_1^{-1} Y_k.$$

We are ready to specify the measure μ, as promised. Define

$$\mu = \lim_{k \to \infty} \mu_k,$$

where

$$\mu_k = (\#\mathcal{J}_k)^{-1} Y_k \sum_{I \in \mathcal{J}_k} \lambda|_I.$$

It follows from the construction that μ is a probability measure supported on S_α. Let B be a ball of radius $r \in (0, r_0(\varepsilon))$. Choose k such that

$$Y_{k+1}^{-\delta_{\mathbf{r}}} < r \leqslant Y_k^{-\delta_{\mathbf{r}}}.$$

Then B can be covered by at most $4^n r^n Y_{k+1}$ many elements from $\mathcal{I}_{k+1}$. Thus

$$\mu_{k+1}(B) \leqslant 4^n r^n Y_{k+1} (\#\mathcal{J}_{k+1})^{-1} \leqslant 4^n r^n c^k Y_1 \leqslant C_1 r^{n - \frac{k \log c}{\delta_{\mathbf{r}} \log Y_k}},$$

for some $C_1 > 0$. By our assumption that $Y_k^{1/k} \to \infty$, this implies (6.2), which completes the proof. $\qquad\square$

We are equipped to prove Theorem 1.11. Fix $\alpha \in (0, 1/2)$, and let $R = R(\alpha)$ be as in Theorem 6.1. For $k \in \mathbb{N}$, write

$$Y_k = \|\mathbf{q}_k\|_{\mathbf{r}}, \qquad M_k = \inf_{\mathbf{p} \in \mathbb{Z}^m} \|A\mathbf{q}_k - \mathbf{p}\|_{\mathbf{s}} = \|A\mathbf{q}_k - \mathbf{p}_k\|_{\mathbf{s}},$$

as before. Since $\lim_{k \to \infty} Y_k^{1/k} = \infty$, the proof of [7, Theorem 2.2] reveals that there exists a function $\phi : \mathbb{N} \to \mathbb{N}$ for which

$$Y_{\phi(k+1)} \geqslant R Y_{\phi(k)} \quad \text{and} \quad Y_{\phi(k)+1} \geqslant R^{-1} Y_{\phi(k+1)}.$$

By Theorem 6.1, it suffices to show that the set

$$S := \{\boldsymbol{\theta} \in [0, 1]^n : \operatorname{dist}(\langle \mathbf{q}_{\phi(k)}, \boldsymbol{\theta} \rangle, \mathbb{Z}) \geqslant \alpha \text{ for all } k \geqslant 1\}$$

is a subset of $\operatorname{Bad}_{\mathbf{r}, \mathbf{s}}^{\varepsilon}({}^t A)$, for some $\varepsilon > 0$.

Let $\boldsymbol{\theta} \in S$. For any $(\mathbf{p}, \mathbf{q}) \in (\mathbb{Z}^m \setminus \{\mathbf{0}\}) \times \mathbb{Z}^n$, let k be the unique index for which

$$Y_{\phi(k)} \leqslant \varepsilon_1^{-1} \|\mathbf{p}\|_{\mathbf{s}} < Y_{\phi(k+1)},$$

where $\varepsilon_1 = R^{-1}(\alpha/2m)^{1/\delta}$. With δ as in Remark 4.4, we have

$$|\langle A\mathbf{q}_{\phi(k)} - \mathbf{p}_{\phi(k)}, \mathbf{p} \rangle| \leqslant m \cdot \max\{M_{\phi(k)}^{\delta}\|\mathbf{p}\|_{\mathbf{s}}^{\delta}, M_{\phi(k)}\|\mathbf{p}\|_{\mathbf{s}}\}$$

$$\leqslant m \cdot \max\{Y_{\phi(k)+1}^{-\delta}\|\mathbf{p}\|_{\mathbf{s}}^{\delta}, Y_{\phi(k)+1}^{-1}\|\mathbf{p}\|_{\mathbf{s}}\} \leqslant m(R\varepsilon_1)^{\delta} \leqslant \frac{\alpha}{2}.$$

Now the calculation (5.2) yields, for any $(\mathbf{p}, \mathbf{q}) \in (\mathbb{Z}^m \setminus \{\mathbf{0}\}) \times \mathbb{Z}^n$, the inequality

$$|\langle \mathbf{q}_{\phi(k)}, {}^t A\mathbf{p} - \mathbf{q} - \boldsymbol{\theta} \rangle| \geqslant \frac{\alpha}{2}.$$

Hence

$$Y_{\phi(k)}\|{}^t A\mathbf{p} - \mathbf{q} - \boldsymbol{\theta}\|_{\mathbf{r}} \geqslant \min\left\{\frac{\alpha}{2n}, \left(\frac{\alpha}{2n}\right)^{1/\delta}\right\} = \left(\frac{\alpha}{2n}\right)^{1/\delta},$$

which implies

$$\|\mathbf{p}\|_{\mathbf{s}} \cdot \|{}^t A\mathbf{p} - \mathbf{q} - \boldsymbol{\theta}\|_{\mathbf{r}} \geqslant \varepsilon_1 \left(\frac{\alpha}{2n}\right)^{1/\delta} = \frac{1}{R}\left(\frac{\alpha^2}{4mn}\right)^{1/\delta}.$$

Therefore $\boldsymbol{\theta} \in \operatorname{Bad}_{\mathbf{r}, \mathbf{s}}^{\varepsilon}({}^t A)$, for $\varepsilon = \frac{1}{R}\left(\frac{\alpha^2}{4mn}\right)^{1/\delta}$, completing the proof of Theorem 1.11.

7. Applications to inhomogeneous Diophantine approximation on manifolds

Transference principles such as Theorem 1.9 play a crucial role in inhomogeneous Diophantine approximation on manifolds by providing *lower bounds* for inhomogeneous Diophantine exponents. In this section, we illustrate this principle in a number of examples. The corresponding upper bound in each case is found using a (suitable adaptation of) the transference principle of Beresnevich and Velani. We begin with

Proposition 7.1. *Assume that $\omega_{\mathbf{s},\mathbf{r}}(A) = 1$. Then for every $\boldsymbol{\theta} \in \mathbb{R}^m$,*

$$\omega_{\mathbf{s},\mathbf{r}}(A, \boldsymbol{\theta}) \geqslant 1,$$

with equality for Lebesgue-almost all $\boldsymbol{\theta}$.

Proof. We follow the argument in [4]. Let $A \in M_{m \times n}(\mathbb{R})$ be as above. By the weighted version of Dyson's transference principle (Theorem 1.7), we have $\omega_{\mathbf{r},\mathbf{s}}(^t A) = 1$. Now applying the weighted form of Dirichlet's approximation theorem and the trivial inequality $\omega_{\mathbf{r},\mathbf{s}}(^t A) \geqslant \hat{\omega}_{\mathbf{r},\mathbf{s}}(^t A)$ yields

$$1 = \omega_{\mathbf{r},\mathbf{s}}(^t A) \geqslant \hat{\omega}_{\mathbf{r},\mathbf{s}}(^t A) \geqslant 1.$$

These inequalities must be equalities, so in particular $\hat{\omega}_{\mathbf{r},\mathbf{s}}(^t A) = 1$. By Theorem 1.9, we finally have $\omega_{\mathbf{s},\mathbf{r}}(A, \boldsymbol{\theta}) \geqslant 1$, with equality for almost all $\boldsymbol{\theta}$. $\qquad\square$

As a consequence we have:

Corollary 7.2. *Let μ be a measure on $M_{m \times n}(\mathbb{R})$. Assume that $\omega_{\mathbf{s},\mathbf{r}}(A) = 1$ for μ-almost all A. Then for every $\boldsymbol{\theta} \in \mathbb{R}^m$ and μ-almost all A, we have*

$$\omega_{\mathbf{s},\mathbf{r}}(A, \boldsymbol{\theta}) \geqslant 1.$$

We now discuss the above corollary in the context of some interesting measures. Dirichlet's theorem implies that $\omega(A) \geqslant 1$ for every $A \in M_{m \times n}(\mathbb{R})$. A matrix $A \in M_{m \times n}(\mathbb{R})$ is said to be *very well approximable* if $\omega(A) > 1$. Metric Diophantine approximation on manifolds is concerned with the question of whether typical Diophantine properties in $M_{m \times n}(\mathbb{R})$, i.e. those which are generic for Lebesgue measure, are inherited by proper submanifolds (or, more generally, supports of suitable measures).

In 1932, Mahler [36] conjectured that for almost every $x \in \mathbb{R}$ the vector

$$(x, x^2, \ldots, x^n) \tag{7.1}$$

is not very well approximable. A measure μ is called *extremal* if μ-almost every A is not very well approximable. A manifold is *extremal* if a measure in its natural volume class—for example, the pushforward of Lebesgue measure by a map parametrising the manifold—is extremal. We can similarly define (weighted) *inhomogeneously very well approximable* matrices (see [3, §§6.1–6.2]) and (inhomogeneously) *very well multiplicatively approximable* matrices, as well as the corresponding notions of extremality.

Mahler's conjecture was resolved by Sprindžuk [46, 47], who in turn formulated a more general conjecture [48] that was proved by Kleinbock and Margulis [29]. They showed that almost every point on a smooth, "nondegenerate" submanifold of $\mathbb{R}^n$ is not very well (multiplicatively) approximable. Subsequently, there have been numerous advances in the subject. The inhomogeneous version of Sprindžuk's conjectures were established by Beresnevich and Velani [4] using their *transference principle*. They proved, for instance, the following result ([4, Theorem 1]).

Theorem 7.3. *Let μ be a measure on $M_{m \times n}(\mathbb{R})$. If μ is contracting almost everywhere, then for every $\boldsymbol{\theta} \in \mathbb{R}^m$, we have that*

$$\omega(A, \boldsymbol{\theta}) = 1$$

for μ-almost every $A \in M_{m \times n}(\mathbb{R})$.

We refer the interested reader to the papers above for the definitions of the terms "nondenegenerate" and "contracting". Since there is no inhomogeneous version of Dirichlet's theorem, the notion of extremality is more delicate and both upper and lower bounds for the exponent are required. The proof of the Theorem above accordingly has two steps. The lower bound for the exponent proceeds using the transference results of Bugeaud and Laurent, combined with Dyson's transference principle. The upper bound is proved using a homogeneous–inhomogeneous transference principle introduced for the purpose, by Beresnevich and Velani.

Subsequently, Beresnevich, Kleinbock and Margulis [3] considered the more general problem of Diophantine approximation on manifolds in the space of matrices. We recall their notation. Let X be a Euclidean space. Given $x \in X$ and $r > 0$, let $B(x, r)$ denote the open ball of radius r centred at x. If $V = B(x, r)$ and $c > 0$, let cV stand for $B(x, cr)$. Let μ be a Radon measure on X. Given $V \subset X$ such that $\mu(V) > 0$ and a function $f : V \to \mathbb{R}$, let

$$\|f\|_{\mu, V} := \sup_{x \in V \cap \text{supp } \mu} |f(x)|.$$

A Radon measure μ will be called *D-Federer* on U, where $D > 0$ and U is an open subset of X, if $\mu(3V) < D\mu(V)$ for any ball $V \subset U$ centred in the support of μ. The measure μ is called *Federer* if for μ-almost every point $x \in X$ there is a neighbourhood U of x and $D > 0$ such that μ is D-Federer on U.

Given $C, \alpha > 0$ and an open subset $U \subset X$, we say that $f : U \to \mathbb{R}$ is called (C, α)-*good* on U with respect to the measure μ if for any ball $V \subset U$ centred in supp μ and any $\varepsilon > 0$ one has

$$\mu(\{x \in V \ : \ |f(x)| < \varepsilon\}) \leqslant C \left(\frac{\varepsilon}{\|f\|_{\mu, V}} \right)^{\alpha} \mu(V).$$

Given $f = (f_1, \ldots, f_N) : U \to \mathbb{R}^N$, we say that the pair (f, μ) is *good* if for μ-almost every $x \in U$ there is a neighbourhood $V \subset U$ of x and $C, \alpha > 0$ such that any linear combination of $1, f_1, \ldots, f_N$ over $\mathbb{R}$ is (C, α)-good on V. The pair (f, μ) is called *non-planar* if for any ball $V \subset U$ centred in supp μ the set $f(V \cap \text{supp } \mu)$ is not contained in any affine hyperplane of $\mathbb{R}^N$. Non-planarity is a generalisation of the nondegeneracy property of smooth manifolds mentioned above. The theorem below deals with Diophantine properties of submanifolds in the space of matrices, and therefore requires a slightly more general notion. Informally, we say that the pair (F, μ) is weakly non-planar if $F(\text{supp } \mu)$ does not locally lie entirely inside a certain polynomial hypersurface. The precise definition may be found in [3, §2].

The following is [3, Theorem 6.5].

Theorem 7.4. *Let U be an open subset of $\mathbb{R}^d$, μ a Federer measure on U, and $F : U \to M_{m \times n}(\mathbb{R})$ a continuous map such that (F, μ) is*

    (1) good, and
    (2) weakly non-planar.

Then $F_*\mu$ is inhomogeneously $(\mathbf{s}, \mathbf{r})$-extremal for any $(m + n)$-tuple $(\mathbf{s}, \mathbf{r})$ of weights.

The conclusion of the theorem above asserts that $\omega_{\mathbf{s},\mathbf{r}}(A, \boldsymbol{\theta}) \leqslant 1$ for μ-almost every A in $F(U)$. We may apply Corollary 7.2 using the homogeneous ($\boldsymbol{\theta} = \mathbf{0}$) case of Theorem 7.4, along with the weighted version of Dirichlet's theorem, to conclude thus.

Theorem 7.5. *Let U, μ, F be as in Theorem 7.4. Then for μ-almost every $A \in F(U)$ and every $\boldsymbol{\theta}$, we have*

$$\omega_{\mathbf{s},\mathbf{r}}(A, \boldsymbol{\theta}) = 1.$$

Whilst the scope of the result is very general, we note that it is new even for the Veronese curve (7.1) in both the simultaneous and the dual approximation contexts.

We now turn our attention to measures which do not satisfy the weak non-planarity condition. Natural examples are provided by affine subspaces, and the reader may consult [17] for a recent survey. We specialise to the case of simultaneous approximation. The results below also hold in the dual setting: see [2] for examples of inhomogeneous dual Diophantine approximation on affine subspaces.

Theorem 7.6. *Let L be an $\mathbf{r}$-extremal affine subspace of $\mathbb{R}^n$ and let M be a smooth nondegenerate submanifold of L. Then, for almost every $A \in M$ and every $\boldsymbol{\theta} \in \mathbb{R}^n$,*

$$\omega_{\mathbf{r}}(A, \boldsymbol{\theta}) = 1.$$

Proof. The unweighted analogue of this result was proved in [4]. First we note that L is $\mathbf{r}$-extremal if and only if M is. This is more or less proved by Kleinbock in [26]. More precisely, the unweighted (Theorem 1.2) and multiplicative (Theorem 1.4) variants of the statement are known. We indicate the minor changes required to prove the weighted version. The argument in [26] is based on the dynamical approach developed in [29]. To a row vector $\mathbf{y} \in \mathbb{R}^n$, one can associate a unimodular lattice

$$u_{\mathbf{y}} \mathbb{Z}^{n+1} := \begin{pmatrix} 1 & y \\ 0 & I_n \end{pmatrix}.$$

The Diophantine properties of $\mathbf{y}$ may be viewed in terms of the F-orbit of $u_{\mathbf{y}} \mathbb{Z}^{n+1}$ on the moduli space of lattices $\mathrm{SL}_{n+1}(\mathbb{R})/\mathrm{SL}_{n+1}(\mathbb{Z})$, where

$$F = \{g_t \mid t \geqslant 0\} \text{ and } g_t = \mathrm{diag}(e^t, e^{-t/n}, \ldots, e^{-t/n}).$$

One then invokes a quantitative nondivergence result [29, Theorem 5.2]. In order to address the weighted case, we replace F by $F_{\mathbf{r}}$, where for $\mathbf{r} = (r_1, \ldots, r_n)$, we set

$$F_{\mathbf{r}} := \{g(\mathbf{r})_t \mid t \geqslant 0\} \text{ and } g(\mathbf{r})_t = \operatorname{diag}(e^t, e^{-r_1 t}, \ldots, e^{-r_n t}).$$

It remains to apply the arguments in [26, §3], *mutatis mutandis*, to deliver our weighted analogue of [loc. cit., Theorem 4.2].

Assume therefore that M is an $\mathbf{r}$-extremal, smooth, nondegenerate submanifold of an affine subspace L. Then for almost every $A \in M$, we have that $\omega_{\mathbf{r}}(A) = 1$. We may therefore apply Corollary 7.2 to conclude that for almost every $A \in M$ and every $\boldsymbol{\theta}$ we have $\omega_{\mathbf{r}}(A, \boldsymbol{\theta}) \geqslant 1$.

To prove the lower bound, we apply the transference principle developed by Beresnevich and Velani [4]; see also [5] where a simplified exposition of the simultaneous transference principle is given, and additionally [3] where the weighted theory is developed. The proof proceeds via a more or less verbatim repetition of [4], so we omit it. $\square$

The results of Kleinbock [27], alluded to above, apply to more general situations than extremality. In particular, it has been shown that arbitrary Diophantine exponents (not just the critical exponent) of affine subspaces are inherited by smooth nondegenerate submanifolds. These results also have counterparts in inhomogeneous approximation. All of this was investigated in the non-weighted and multiplicative cases in [18] by two of the authors of the present article, where it was shown that the Beresnevich–Velani transference principle can be adapted to the non-extremal setting to get an upper bound for inhomogeneous exponents. The corresponding lower bound was then obtained using the results of Bugeaud and Laurent along with more transference results of German [16]. We can provide a weighted analogue in the non-extremal case, using a similar adaptation of the Beresnevich–Velani transference principle for the upper bound. The weighted analogues of the Bugeaud–Laurent and Dyson theorems, namely theorems 1.9 and 1.7, together with the trivial relation $\omega_{\mathbf{r},\mathbf{s}} \geqslant \hat{\omega}_{\mathbf{r},\mathbf{s}}$, provide the lower bound as follows.

Theorem 7.7. *Let U, μ, F be as in Theorem 7.4. Suppose that for μ-almost every $A \in F(U)$ we have $\omega_{\mathbf{s},\mathbf{r}}(A) =: \omega$. Then for μ-almost every $A \in F(U)$ and every $\boldsymbol{\theta}$, we have*

$$\omega \geqslant \omega_{\mathbf{s},\mathbf{r}}(A, \boldsymbol{\theta}) \geqslant \frac{(m + n - 1)\rho_{\mathbf{s}}\rho_{\mathbf{r}}(\delta_{\mathbf{r}} + \delta_{\mathbf{s}}\omega) - \rho_{\mathbf{s}}\delta_{\mathbf{r}}\delta_{\mathbf{s}}(\omega - 1)}{(m + n - 1)\rho_{\mathbf{s}}\rho_{\mathbf{r}}(\delta_{\mathbf{r}} + \delta_{\mathbf{s}}\omega) + \rho_{\mathbf{r}}\delta_{\mathbf{r}}\delta_{\mathbf{s}}(\omega - 1)}.$$

Note that in [18] the trivial relation $\omega \geqslant \hat{\omega}$ is also used, and the results can be refined by using the recent optimal lower bound $\omega \geqslant G(\hat{\omega})$ for vectors obtained in [39]. The corresponding optimal lower bounds for matrices, or in the weighted case, remain open.

8. Inhomogeneous intermediate exponents

In this final section, we prove Theorem 1.15, first defining $\omega_d(\boldsymbol{\alpha}, \boldsymbol{\theta})$. As a prelude, we briefly discuss Laurent's "algebraic" definition [34, §2] in the homogeneous setting $\boldsymbol{\theta} = \boldsymbol{0}$, with the same notations as in the introduction. Laurent made the following observations.

(1) Recall from the introduction that to a rational subspace $\mathcal{L}$ we may associate Grassmannian coordinates $\mathbf{X}$. Let $\boldsymbol{\alpha} \in \mathbb{R}^n$. Let $\boldsymbol{\alpha}' = (1, \boldsymbol{\alpha}) \in \mathbb{R}^{n+1}$ and $\mathbf{X}' \in \mathbb{R}^{\binom{n+1}{d+1}}$ be homogeneous coordinates for $[1 : \boldsymbol{\alpha}] \in \mathbb{P}^n_{\mathbb{R}}$ and $\mathbf{X} \in \mathbb{P}^{\binom{n+1}{d+1}-1}_{\mathbb{R}}$ respectively. As shown in [34, Lemma 1], we have
$$d([1 : \boldsymbol{\alpha}], \mathcal{L}) = \frac{|\boldsymbol{\alpha}' \wedge \mathbf{X}'|}{|\boldsymbol{\alpha}'| \cdot |\mathbf{X}'|}.$$
This circumvents the need for the extremal definition (1.4).
(2) The Plücker embedding is known to establish a bijection from $\mathrm{Gr}(d, \mathbb{P}^n_{\mathbb{R}})$ to the set of non-zero decomposable multivectors in the exterior algebra $\Lambda^{d+1}(\mathbb{R}^{n+1})$, up to homothety.
(3) When optimising over $\mathcal{L} \in \mathrm{Gr}(d, \mathbb{P}^n_{\mathbb{R}})$, or equivalently over multivectors in $\Lambda^{d+1}(\mathbb{R}^{n+1})$, one may drop the assumption of decomposability—see the remark following [34, Definition 4].

We thus arrive at the following variant of [34, Definition 4], which is equivalent to Definition 1.13 (we have normalised in accordance with Remark 1.3).

Definition 8.1. Let $n \in \mathbb{N}$, let $d \in \{0, 1, \dots, n-1\}$, and let $\boldsymbol{\alpha}' = (1, \boldsymbol{\alpha})$ with $\boldsymbol{\alpha} \in \mathbb{R}^n$. The *dth ordinary exponent* $\omega_d(\boldsymbol{\alpha})$ (resp. the *dth uniform exponent* $\hat{\omega}_d(\boldsymbol{\alpha})$) is the supremum of $\omega \in \mathbb{R}$ such that there exists $\mathbf{X} \in \Lambda^{d+1}(\mathbb{Z}^{n+1})$ for which
$$|\mathbf{X}|^{\binom{n}{d}} \leqslant T, \qquad |\boldsymbol{\alpha}' \wedge \mathbf{X}|^{\binom{n}{d+1}} \leqslant T^{-\omega}$$
for some arbitrarily large real numbers T (resp. for every sufficiently large real number T).

Before we define inhomogeneous intermediate exponents, we give another equivalent definition of intermediate exponents.

Definition 8.2. Let $n \in \mathbb{N}$, let $d \in \{0, 1, \dots, n-1\}$ and let $\boldsymbol{\alpha} \in \mathbb{R}^n$. The *dth ordinary exponent* $\omega_d(\boldsymbol{\alpha})$ (resp. the *dth uniform exponent* $\hat{\omega}_d(\boldsymbol{\alpha})$) is the supremum of $\omega \in \mathbb{R}$ such that there exists $\mathbf{Z} \in \Lambda^d(\mathbb{Z}^n) \setminus \{\mathbf{0}\}$ and $\mathbf{Y} \in \Lambda^{d+1}(\mathbb{Z}^n)$ for which
$$|\mathbf{Z}|^{\binom{n}{d}} \leqslant T, \qquad |\boldsymbol{\alpha} \wedge \mathbf{Z} + \mathbf{Y}|^{\binom{n}{d+1}} \leqslant T^{-\omega}$$
for some arbitrarily large real numbers T (resp. for every sufficiently large real number T).

Proof of equivalence of these two definitions. We set $\mathbf{e}_0 = (1, 0, \dots, 0) \in \mathbb{Z}^{n+1}$, and identify $\{\mathbf{x} = (x_0, \dots, x_n) \in \mathbb{Z}^{n+1} : x_0 = 0\}$ and $\{\mathbf{x} = (x_0, \dots, x_n) \in \mathbb{R}^{n+1} : x_0 = 0\}$ with $\mathbb{Z}^n$ and $\mathbb{R}^n$, respectively. We have
$$\Lambda^{d+1}(\mathbb{Z}^{n+1}) = \mathbf{e}_0 \wedge (\Lambda^d(\mathbb{Z}^n)) \oplus \Lambda^{d+1}(\mathbb{Z}^n),$$

so we can uniquely decompose any $\mathbf{X} \in \Lambda^{d+1}(\mathbb{Z}^{n+1})$ as $\mathbf{e}_0 \wedge \mathbf{Z} - \mathbf{Y}$, with $\mathbf{Z} \in \Lambda^d(\mathbb{Z}^n)$ and $\mathbf{Y} \in \Lambda^{d+1}(\mathbb{Z}^n)$. This gives

$$\boldsymbol{\alpha}' \wedge \mathbf{X} = (\mathbf{e}_0 + \boldsymbol{\alpha}) \wedge (\mathbf{e}_0 \wedge \mathbf{Z} - \mathbf{Y}) = -\mathbf{e}_0 \wedge \mathbf{Y} + \boldsymbol{\alpha} \wedge \mathbf{e}_0 \wedge \mathbf{Z} - \boldsymbol{\alpha} \wedge \mathbf{Y}$$
$$= -(\mathbf{e}_0 + \boldsymbol{\alpha}) \wedge (\boldsymbol{\alpha} \wedge \mathbf{Z} + \mathbf{Y}).$$

From the definition (1.3) of the inner product on $\Lambda^{d+1}(\mathbb{R}^n)$, we see that $\mathbf{e}_0 \wedge (\boldsymbol{\alpha} \wedge \mathbf{Z} + \mathbf{Y})$ is orthogonal to $\boldsymbol{\alpha} \wedge (\boldsymbol{\alpha} \wedge \mathbf{Z} + \mathbf{Y})$. Thus it follows that

$$|\boldsymbol{\alpha} \wedge \mathbf{Z} + \mathbf{Y}| \leqslant |\boldsymbol{\alpha}' \wedge \mathbf{X}| \leqslant |\boldsymbol{\alpha}'||\boldsymbol{\alpha} \wedge \mathbf{Z} + \mathbf{Y}|.$$

Note that $\mathbf{e}_0 \wedge \mathbf{Z}$ is orthogonal to $\mathbf{Y}$, hence

$$\max\{|\mathbf{Z}|, |\mathbf{Y}|\} \leqslant |\mathbf{X}| \leqslant 2\max\{|\mathbf{Z}|, |\mathbf{Y}|\}.$$

The equivalence of Definition 8.1 and Definition 8.2 follows from the inequalities above. $\qquad\square$

We proceed to define inhomogeneous intermediate exponents. By lexicographically ordering a basis for $\Lambda^{d+1}(\mathbb{R}^{n+1})$, we obtain a canonical isomorphism

$$\Lambda^{d+1}(\mathbb{R}^{n+1}) \simeq \mathbb{R}^{\binom{n+1}{d+1}},$$

which we use to identify the two spaces.

Definition 8.3. Let $n \in \mathbb{N}$, let $d \in \{0, 1, \ldots, n-1\}$ and let $\boldsymbol{\alpha} \in \mathbb{R}^n$. Let $\boldsymbol{\theta} \in \mathbb{R}^{\binom{n}{d+1}}$. The *$d$th ordinary exponent* $\omega_d(\boldsymbol{\alpha}, \boldsymbol{\theta})$ (resp. the *dth uniform exponent* $\hat{\omega}_d(\boldsymbol{\alpha}, \boldsymbol{\theta})$) is the supremum of $\omega \in \mathbb{R}$ such that there exists $\mathbf{Z} \in \Lambda^d(\mathbb{Z}^n) \setminus \{\mathbf{0}\}$ and $\mathbf{Y} \in \Lambda^{d+1}(\mathbb{Z}^n)$ for which

$$|\mathbf{Z}|^{\binom{n}{d}} \leqslant T, \qquad |\boldsymbol{\alpha} \wedge \mathbf{Z} + \mathbf{Y} + \boldsymbol{\theta}|^{\binom{n}{d+1}} \leqslant T^{-\omega} \tag{8.1}$$

for some arbitrarily large real numbers T (resp. for every sufficiently large real number T).

Our definition matches the usual definitions for the $d = 0$ (simultaneous) and $d = n - 1$ (dual) cases. The lemma below formalises this, and generalises the discussion following [34, Definition 2].

Lemma 8.4. *If we interpret $\boldsymbol{\theta}$ as an element of $\mathbb{R}^n$ in the former case and $\boldsymbol{\theta}$ as an element of $\mathbb{R}$ in the latter, then*

$$\omega_0(\boldsymbol{\alpha}, \boldsymbol{\theta}) = \omega(\boldsymbol{\alpha}, \boldsymbol{\theta}), \qquad \hat{\omega}_0(\boldsymbol{\alpha}, \boldsymbol{\theta}) = \hat{\omega}(\boldsymbol{\alpha}, \boldsymbol{\theta})$$

and

$$\omega_{n-1}(\boldsymbol{\alpha}, \boldsymbol{\theta}) = \omega({}^t\boldsymbol{\alpha}, \boldsymbol{\theta}), \qquad \hat{\omega}_{n-1}(\boldsymbol{\alpha}, \boldsymbol{\theta}) = \hat{\omega}({}^t\boldsymbol{\alpha}, \boldsymbol{\theta})$$

Proof. By comparing Definition 8.3 with Definition 1.2, it is clear that to prove the lemma, it suffices to compare (8.1) with (1.1). When $d = 0$, $\mathbf{Z} \in \mathbb{Z} \setminus \{\mathbf{0}\}$ and $\mathbf{Y} \in \mathbb{Z}^n$. Writing $\mathbf{Z} = z$ and $\mathbf{Y} = \mathbf{y}$, we have $\boldsymbol{\alpha} \wedge \mathbf{Z} = z\boldsymbol{\alpha}$. Now (8.1) becomes

$$|z| \leqslant T \qquad |z\boldsymbol{\alpha} + \mathbf{y} + \boldsymbol{\theta}|^n \leqslant T^{-\omega},$$

which coincides with (1.1).

When $d = n - 1$, we have $\mathbf{Z} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$ and $\mathbf{Y} \in \mathbb{Z}$. Writing $\mathbf{Z} = \mathbf{z}$, $\mathbf{Y} = y$ and $\boldsymbol{\theta} = \theta$, we have $\boldsymbol{\alpha} \wedge \mathbf{Z} = \langle \boldsymbol{\alpha}, \mathbf{z} \rangle$. Now (8.1) becomes

$$|\mathbf{z}|^n \leqslant T \qquad |\langle \boldsymbol{\alpha}, \mathbf{z} \rangle + y + \theta| \leqslant T^{-\omega},$$

which coincides with (1.1). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We are ready to prove Theorem 1.15. The wedge product with $\boldsymbol{\alpha}$ defines a linear map from $\Lambda^d(\mathbb{R}^n)$ to $\Lambda^{d+1}(\mathbb{R}^n)$. Hence, in view of Definition 8.3, Theorem 1.15 follows from Theorem 1.8; we expound on this below.

Instead of using matrices, we may define our Diophantine exponents for linear maps and their transpose linear maps. Taking the wedge product with $\boldsymbol{\alpha} \in \mathbb{R}^n$ defines linear maps

$$\mathbb{R}^N \simeq \Lambda^d(\mathbb{R}^n) \xrightarrow{f = \boldsymbol{\alpha} \wedge \cdot} \Lambda^{d+1}(\mathbb{R}^n) \simeq \mathbb{R}^M$$

and

$$\mathbb{R}^M \simeq \Lambda^{n-d-1}(\mathbb{R}^n) \xrightarrow{g = \boldsymbol{\alpha} \wedge \cdot} \Lambda^{n-d}(\mathbb{R}^n) \simeq \mathbb{R}^N.$$

These are transposes of one another, up to sign, for if $\boldsymbol{\beta} \in \Lambda^{n-d-1}(\mathbb{R}^n) \simeq \Lambda^{d+1}(\mathbb{R}^n)^\vee$ and $\boldsymbol{\gamma} \in \Lambda^d(\mathbb{R}^n)$ then

$$|\boldsymbol{\beta} \wedge f(\boldsymbol{\gamma})| = |\boldsymbol{\beta} \wedge \boldsymbol{\alpha} \wedge \boldsymbol{\gamma}| = |g(\boldsymbol{\beta}) \wedge \boldsymbol{\gamma}|.$$

Theorem 1.8 therefore reveals that

$$\omega_d(\boldsymbol{\alpha}, \boldsymbol{\theta}) \geqslant \frac{1}{\hat{\omega}_{n-1-d}(\boldsymbol{\alpha})},$$

with equality for almost all $\boldsymbol{\theta}$.

References

[1] V. Beresnevich, *Rational points near manifolds and metric Diophantine approximation* Ann. of Math. (2) **175** (2012), 187–235.

[2] V. Beresnevich, A. Ganguly, A. Ghosh and S. Velani, *Inhomogeneous Dual Diophantine Approximation on Affine Subspaces*, Int. Math. Res. Not., doi:10.1093/imrn/rny124, 32 pp.

[3] V. Beresnevich, D. Kleinbock and G. Margulis, *Non-planarity and metric Diophantine approximation for systems of linear forms*, J. Théor. Nombres Bordeaux **27** (2015), 1–31.

[4] V. Beresnevich and S. Velani, *An inhomogeneous transference principle and Diophantine approximation*, Proc. Lond. Math. Soc. (3) **101** (2010), 821–851.

[5] V. Beresnevich and S. Velani, *Simultaneous inhomogeneous Diophantine approximations on manifolds*, Fundam. Prikl. Mat. **16** (2010), 3–17.

[6] P. Bengoechea and N. Moshchevitin, *Badly approximable points in twisted Diophantine approximation and Hausdorff dimension*, Acta Arith. **177** (2017), 301–314.

[7] Y. Bugeaud, D. Kim, S. Lim and M. Rams, *Hausdorff dimension in inhomogeneous Diophantine approximation*, arXiv:1805.10436.

[8] Y. Bugeaud and M. Laurent, *On exponents of homogeneous and inhomogeneous Diophantine approximation*, Mosc. Math. J. **5** (2005), 747–766.

[9] J. W. S. Cassels, *An introduction to Diophantine Approximation*, Cambridge Tracts in Math. and Math. Phys., vol. 99, Cambridge University Press, 1957.

[10] J. W. S. Cassels, *An introduction to the Geometry of Numbers*, Springer Verlag, 1997.

[11] N. Chevallier, *Best simultaneous Diophantine approximations and multidimensional continued fraction expansions*, Mosc. J. Comb. Number Theory **3** (2013), 3–56.

[12] J. P. G. Lejeune Dirichlet, *Verallgemeinerung eines Satzes aus der Lehre ven Kettenbrüchen nebst einigen Anwendungen auf die Theorie die Zahlen*, S.-B. Preus Akad. Wiss. (1842), 93–95.

[13] F. J. Dyson, *On simultaneous Diophantine approximations*, Proc. Lond. Math. Soc. (2) **49** (1947), 409–420.

[14] J.-H. Evertse, *Mahler's work on the geometry of numbers*, arXiv:1806.00356.

[15] K. Falconer, *Fractal geometry. Mathematical foundations and applications*, third edition, Wiley & Sons, Chichester, 2014.

[16] O. N. German, *On Diophantine exponents and Khintchine's transference principle*, Mosc. J. Comb. Number Theory **2** (2012), 22–51.

[17] A. Ghosh, *Diophantine approximation on subspaces of $\mathbb{R}^n$ and dynamics on homogeneous spaces*, Handbook of Group Actions (Vol. IV) ALM 41, Ch. 9, pp. 509–527. Editors Lizhen Ji, Athanase Papadopoulos, Shing-Tung Yau.

[18] A. Ghosh and A. Marnat, *On Diophantine transference principles*, Math. Proc. Camb. Phil. Soc., to appear.

[19] P. M. Gruber and C. G. Lekkerkerker, *Geometry of numbers*, North-Holland Mathematical Library, vol. 37, North-Holland Publishing Co., Amsterdam, 1937.

[20] S. Harrap and N. Moshchevitin, *A note on weighted badly approximable linear forms*, Glasg. Math. J. **59** (2017), 349–357.

[21] V. Jarník, *Zum khintchineschen "Übertragungssatz"*, Trav. Inst. Math. Tbilissi **3** (1938), 193–212.

[22] V. Jarník, *Eine Bemerkung zum Übertragimgssatz*, Büalgar. Akad. Nauk Izv. Mat. Inst. **3** (1959), 169–175.

[23] R. Kannan and L. Lovász, *Covering minima and lattice-point-free convex bodies*, Ann. of Math. **2** (1998) **128**, 507–602.

[24] W. Banaszczyk, *Inequalities for convex bodies and polar reciprocal lattices in $\mathbb{R}^n$. II. Application of K-convexity*, Discrete Comput. Geom. **16** (1996), 305–311.

[25] A. Ya. Khintchine, *Über eine Klasse linearer diophantischer Approximationen*, Rendiconti Circ. Mat. Palermo **50** (1926), 170–195.

[26] D. Kleinbock, *Extremal subspaces and their submanifolds*, Geom. Funct. Anal. **13** (2003), 437–466.

[27] D. Kleinbock, *An extension of quantitative nondivergence and applications to Diophantine exponents*, Trans. Amer. Math. Soc. **360** (2008), 6497–6523.

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

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

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

[31] D. Kleinbock and B. Weiss, *Modified Schmidt games and Diophantine approximation with weights*, Adv. Math. **223** (2010), 1276–1298.

[32] J. C. Lagarias, *Best simultaneous Diophantine approximations. I. Growth rates of best approximation denominators*, Trans. Amer. Math. Soc. **272** (1982), 545–554.

[33] J. C. Lagarias, *Best Diophantine approximations to a set of linear forms*, J. Austral. Math. Soc. Ser. A **34** (1983), 114–122.

[34] M. Laurent, *On transfer inequalities in Diophantine approximation.* Analytic number theory, 306–314, Cambridge Univ. Press, Cambridge, 2009.

[35] S. Lim, N. de Saxće and U. Shapira, *Dimension bound for badly approximable grids*, Int. Math. Res. Not., to appear.

[36] K. Mahler, *Über das Mass der Menge aller S-Zahlen*, Math. Ann. **106** (1932), 131–139.

[37] K. Mahler, *Ein Übertragungsprinzip für konvexe Körper*, Math. Časopis **68** (1939), 93–102.

[38] K. Mahler, *On compound convex bodies. I.*, Proc. London Math. Soc. (3) **5** (1955), 358–379.

[39] A. Marnat and N. G. Moshchevitin, *An optimal bound for the ratio between ordinary and uniform exponents of Diophantine approximation*, arXiv:1802.03081.

[40] P. Philippon, *Critères pour l'indépendance algébrique*, Inst. Hautes Études Sci. Publ. Math. **64** (1986), 5–52.

[41] C. A. Rogers, *The signature of the errors of some Diophantine approximations*, Proc. Lond. Math. Soc. **52** (1951) , 186–190.

[42] D. Roy, *Diophantine approximation in small degree*, Number Theory, CRM Proc. Lecture Notes, **36** (2004), 269–285.

[43] W. M. Schmidt, *On heights of algebraic subspaces and Diophantine approximations*, Ann. of Math. (2) **85** (1967), 430–472.

[44] W. M. Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics, vol. 785, Springer, Berlin, 1980.

[45] C. L. Siegel, *Lectures on the Geometry of Numbers*, Springer–Verlag, Berlin–Heidelberg–New York 1989.

[46] V. Sprindžuk, *More on Mahler's conjecture*, Dokl. Akad. Nauk SSSR **155** (1964), 54–56 (Russian); English transl. in Soviet Math. Dokl. **5** (1964), 361–363.

[47] V. Sprindžuk, *Mahler's problem in metric number theory*, Translations of Mathematical Monographs, vol. 25, Amer. Math. Soc., Providence, RI, 1969.

[48] V. Sprindžuk, *Achievements and problems in Diophantine approximation theory*, Russian Math. Surveys **35** (1980), 1–80.

[49] G. F. Voronoi, *On one generalization of continued fractions algorithm*, USSR Ac. Sci., v.1 (1952), 197–391.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF YORK, HESLINGTON, YORK, YO10 5DD, UNITED KINGDOM

E-mail address: `Sam.Chow@maths.ox.ac.uk`

SCHOOL OF MATHEMATICS, TATA INSTITUTE OF FUNDAMENTAL RESEARCH, MUMBAI, INDIA 400005

E-mail address: `ghosh@math.tifr.res.in`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF YORK, HESLINGTON, YORK, YO10 5DD, UNITED KINGDOM

E-mail address: `lifan.guan@york.ac.uk`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF YORK, HESLINGTON, YORK, YO10 5DD, UNITED KINGDOM

E-mail address: `antoine.marnat@york.ac.uk`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF YORK, HESLINGTON, YORK, YO10 5DD, UNITED KINGDOM

E-mail address: `david.simmons@york.ac.uk`