



ELSEVIER

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com



Number theory/Dynamical systems

A variational principle in the parametric geometry of numbers, with applications to metric Diophantine approximation



Un principe variationnel en géométrie paramétrique des nombres, illustré par des applications à la théorie métrique de l'approximation diophantienne

Tushar Das^a, Lior Fishman^b, David Simmons^c, Mariusz Urbański^b

^a University of Wisconsin – La Crosse, Department of Mathematics & Statistics, 1725 State Street, La Crosse, WI 54601, USA

^b University of North Texas, Department of Mathematics, 1155 Union Circle #311430, Denton, TX 76203-5017, USA

^c University of York, Department of Mathematics, Heslington, York YO10 5DD, UK

ARTICLE INFO

Article history:

Received 10 May 2017

Accepted after revision 11 July 2017

Available online 21 July 2017

Presented by the Editorial Board

ABSTRACT

We establish a new connection between metric Diophantine approximation and the parametric geometry of numbers by proving a variational principle facilitating the computation of the Hausdorff and packing dimensions of many sets of interest in Diophantine approximation. In particular, we show that the Hausdorff and packing dimensions of the set of singular $m \times n$ matrices are both equal to $mn(1 - \frac{1}{m+n})$, thus proving a conjecture of Kadyrov, Kleinbock, Lindenstrauss, and Margulis as well as answering a question of Bugeaud, Cheung, and Chevallier. Other applications include computing the dimensions of the sets of points witnessing conjectures of Starkov and Schmidt.

© 2017 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

RÉSUMÉ

Nous établissons un nouveau lien entre la théorie métrique de l'approximation diophantienne et la géométrie paramétrique des nombres, en démontrant un principe variationnel permettant le calcul des dimensions de Hausdorff et d'entassement de nombreux ensembles d'intérêt en approximation diophantienne. Comme cas particulier, nous démontrons que les dimensions de Hausdorff et d'entassement de l'ensemble des matrices singulières de dimensions $m \times n$ sont toutes deux égales à $mn(1 - \frac{1}{m+n})$, démontrant ainsi une conjecture de Kadyrov, Kleinbock, Lindenstrauss et Margulis, et répondant par là même à une question soulevée par Bugeaud, Cheung et Chevallier. D'autres exemples d'applica-

E-mail addresses: tdas@uwlax.edu (T. Das), lior.fishman@unt.edu (L. Fishman), David.Simmons@york.ac.uk (D. Simmons), urbanski@unt.edu (M. Urbański).

URLs: <https://sites.google.com/a/uwlax.edu/tdas/> (T. Das), <https://sites.google.com/site/davidsimmonsmath/> (D. Simmons), <http://www.urbanskimath.com/> (M. Urbański).

<http://dx.doi.org/10.1016/j.crma.2017.07.007>

1631-073X/© 2017 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

tion incluent le calcul des dimensions des ensembles de points satisfaisant des conjectures énoncées par Starkov et Schmidt.

© 2017 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

1. Main results

The notion of singularity (in the sense of Diophantine approximation) was introduced by Khintchine, first in 1937 in the setting of simultaneous approximation [11], and later in 1948 in the more general setting of matrix approximation [12].¹ Since then this notion has been studied within Diophantine approximation and allied fields, see Moshchevitin's 2010 survey [13]. An $m \times n$ matrix A is called *singular* if for all $\epsilon > 0$, there exists Q_ϵ such that for all $Q \geq Q_\epsilon$, there exist integer vectors $\mathbf{p} \in \mathbb{Z}^m$ and $\mathbf{q} \in \mathbb{Z}^n$ such that

$$\|A\mathbf{q} + \mathbf{p}\| \leq \epsilon Q^{-n/m} \quad \text{and} \quad 0 < \|\mathbf{q}\| \leq Q.$$

Here $\|\cdot\|$ denotes an arbitrary norm on \mathbb{R}^m or \mathbb{R}^n . We denote the set of singular $m \times n$ matrices by $\text{Sing}(m, n)$. For 1×1 matrices (i.e. numbers), being singular is equivalent to being rational, and in general any matrix A which satisfies an equation of the form $A\mathbf{q} = \mathbf{p}$, with \mathbf{p}, \mathbf{q} integral and \mathbf{q} nonzero, is singular. However, Khintchine proved that there exist singular 2×1 matrices whose entries are linearly independent over \mathbb{Q} [10, Satz II], and his argument generalizes to the setting of $m \times n$ matrices for all $(m, n) \neq (1, 1)$. The name *singular* derives from the fact that $\text{Sing}(m, n)$ is a Lebesgue nullset for all m, n , see e.g. [11, p. 431] or [2, Chapter 5, §7]. Note that singularity is a strengthening of the property of *Dirichlet improvability* introduced by Davenport and Schmidt [6].

In contrast to the measure zero result mentioned above, the computation of the Hausdorff dimension of $\text{Sing}(m, n)$ has been a challenge that so far only met with partial progress. The first breakthrough was made in 2011 by Cheung [3], who proved that the Hausdorff dimension of $\text{Sing}(2, 1)$ is $4/3$; this was extended in 2016 by Cheung and Chevallier [4], who proved that the Hausdorff dimension of $\text{Sing}(m, 1)$ is $m^2/(m+1)$ for all $m \geq 2$; while most recently Kadyrov, Kleinbock, Lindenstrauss, and Margulis [8] proved that the Hausdorff dimension of $\text{Sing}(m, n)$ is at most $\delta_{m,n} := mn(1 - \frac{1}{m+n})$, and went on to conjecture that their upper bound is sharp for all $(m, n) \neq (1, 1)$ (see also [1, Problem 1]).

In this paper, we announce a proof that their conjecture is correct. We will also show that the packing dimension of $\text{Sing}(m, n)$ is the same as its Hausdorff dimension, thus answering a question of Bugeaud, Cheung, and Chevallier [1, Problem 7]. To summarize:

Theorem 1.1. *For all $(m, n) \neq (1, 1)$, we have*

$$\dim_H(\text{Sing}(m, n)) = \dim_P(\text{Sing}(m, n)) = \delta_{m,n} \stackrel{\text{def}}{=} mn \left(1 - \frac{1}{m+n}\right),$$

where $\dim_H(S)$ and $\dim_P(S)$ denote the Hausdorff and packing dimensions of a set S , respectively.

1.1. Dani correspondence

The set of singular matrices is linked to homogeneous dynamics via the *Dani correspondence principle*. For each $t \in \mathbb{R}$ and for each matrix A , let

$$g_t \stackrel{\text{def}}{=} \begin{bmatrix} e^{t/m} I_m & \\ & e^{-t/n} I_n \end{bmatrix}, \quad u_A \stackrel{\text{def}}{=} \begin{bmatrix} I_m & A \\ & I_n \end{bmatrix},$$

where I_k denotes the k -dimensional identity matrix. Finally, let $d = m + n$, and for each $j = 1, \dots, d$, let $\lambda_j(\Lambda)$ denote the j th successive minimum of a lattice $\Lambda \subset \mathbb{R}^d$ (with respect to some fixed norm on \mathbb{R}^d), i.e. the infimum of λ such that the set $\{\mathbf{r} \in \Lambda : \|\mathbf{r}\| \leq \lambda\}$ contains j linearly independent vectors. Then the Dani correspondence principle is a dictionary between the Diophantine properties of a matrix A on the one hand, and the dynamical properties of the orbit $(g_t u_A \mathbb{Z}^d)_{t \geq 0}$ on the other. A particular example is the following result:

Theorem 1.2 ([5, Theorem 2.14]). *An $m \times n$ matrix A is singular if and only if the trajectory $(g_t u_A \mathbb{Z}^d)_{t \geq 0}$ is divergent in the space of unimodular lattices in \mathbb{R}^d , or equivalently if*

$$\lim_{t \rightarrow \infty} \lambda_1(g_t u_A \mathbb{Z}^d) = 0.$$

¹ Although Khintchine's 1926 paper [10] includes a proof of the existence of 2×1 and 1×2 matrices possessing a certain property which implies that they are singular, it does not include a definition of singularity nor discuss any property equivalent to singularity.

It is natural to ask about the set of matrices such that the above limit occurs at a prescribed rate, such as the set of matrices such that $-\log \lambda_1(g_t u_A \mathbb{Z}^d)$ grows linearly with respect to t . This question is closely linked with the concept of uniform exponents of irrationality. The *uniform exponent of irrationality* of an $m \times n$ matrix A , denoted $\widehat{\omega}(A)$, is the supremum of ω such that for all Q sufficiently large, there exist integer vectors $\mathbf{p} \in \mathbb{Z}^m$ and $\mathbf{q} \in \mathbb{Z}^n$ such that

$$\|A\mathbf{q} + \mathbf{p}\| \leq Q^{-\omega} \text{ and } 0 < \|\mathbf{q}\| \leq Q.$$

By Dirichlet’s theorem, every $m \times n$ matrix A satisfies $\widehat{\omega}(A) \geq \frac{n}{m}$. Moreover, it is immediate from the definitions that any matrix A satisfying $\widehat{\omega}(A) > \frac{n}{m}$ is singular. We call a matrix *very singular* if it satisfies the inequality $\widehat{\omega}(A) > \frac{n}{m}$, in analogy with the set of *very well approximable* matrices, which satisfy a similar inequality for the regular (non-uniform) exponent of irrationality (see (1.2)). We denote the set of very singular $m \times n$ matrices by $\text{VSing}(m, n)$. The relationship between uniform exponents of irrationality and very singular matrices on the one hand, and homogeneous dynamics on the other, is given as follows:

Theorem 1.3. *A matrix A is very singular if and only $\widehat{\tau}(A) > 0$, where*

$$\widehat{\tau}(A) \stackrel{\text{def}}{=} \liminf_{t \rightarrow \infty} \frac{-1}{t} \log \lambda_1(g_t u_A \mathbb{Z}^d).$$

Moreover, the quantities $\tau = \widehat{\tau}(A)$ and $\omega = \widehat{\omega}(A)$ are related by the formula

$$\tau = \frac{1}{n} \frac{\omega - \frac{n}{m}}{\omega + 1}. \tag{1.1}$$

This theorem is a straightforward example of the Dani correspondence principle and is probably well-known, but we have not been able to find a reference.

1.2. Dimensions of very singular matrices

Perhaps unsurprisingly, the set of very singular matrices has the same dimension properties as the set of singular matrices.

Theorem 1.4. *For all $(m, n) \neq (1, 1)$, we have*

$$\dim_H(\text{VSing}(m, n)) = \dim_P(\text{VSing}(m, n)) = \delta_{m,n}.$$

One can also ask for more precise results regarding the function $\widehat{\omega}$. Specifically, for each $\omega > \frac{n}{m}$ we can consider the levelset²

$$\text{Sing}_{m,n}(\omega) \stackrel{\text{def}}{=} \{A : \widehat{\omega}(A) = \omega\} = \{A : \widehat{\tau}(A) = \tau\},$$

where τ is given by (1.1). It would be desirable to give precise formulas for the Hausdorff and packing dimensions of $\text{Sing}_{m,n}(\omega)$ in terms of ω, m , and n , see e.g. [1, Problem 2]. However, this appears quite challenging at the present juncture, though we have made significant progress towards this question which we will describe in the next section. Thus, instead of precise formulas we will give asymptotic formulas of two types: estimates valid when ω is small and estimates valid when ω is large. Note that while the minimum value of $\widehat{\omega}$ is always $\frac{n}{m}$ (corresponding to $\widehat{\tau} = 0$), the maximum value depends on whether or not n is at least 2. If $n \geq 2$, then the maximum value of $\widehat{\omega}$ is ∞ (corresponding to $\widehat{\tau} = \frac{1}{n}$), while if $n = 1$, then the maximum value of $\widehat{\omega}$ (excluding rational points) is 1 (corresponding to $\widehat{\tau} = \frac{m-1}{2m}$). Consequently, we have two different asymptotic estimates of the dimensions of $\text{Sing}_{m,n}(\omega)$ when ω is large corresponding to these two cases.

Recall that $\Theta(x)$ denotes any number such that $x/C \leq \Theta(x) \leq Cx$ for some uniform constant C . In all of the formulas below, τ is given by (1.1).

Theorem 1.5. *Suppose that $(m, n) \neq (1, 1)$. Then for all $\omega > \frac{n}{m}$ sufficiently close to $\frac{n}{m}$, we have*

$$\begin{aligned} \dim_H(\text{Sing}_{m,n}(\omega)) &= \delta_{m,n} - \Theta\left(\sqrt{\omega - \frac{n}{m}}\right) & \dim_P(\text{Sing}_{m,n}(\omega)) &= \delta_{m,n} - \Theta\left(\omega - \frac{n}{m}\right) \\ &= \delta_{m,n} - \Theta(\sqrt{\tau}) & &= \delta_{m,n} - \Theta(\tau) \end{aligned}$$

² For results considering the superlevelset, see Theorem 2.9.

unless $m = n = 2$, in which case

$$\begin{aligned} \dim_H(\text{Sing}_{m,n}(\omega)) &= \delta_{m,n} - \Theta\left(\omega - \frac{n}{m}\right) & \dim_P(\text{Sing}_{m,n}(\omega)) &= \delta_{m,n} - \Theta\left(\omega - \frac{n}{m}\right) \\ &= \delta_{m,n} - \Theta(\tau) & &= \delta_{m,n} - \Theta(\tau). \end{aligned}$$

Theorem 1.6. Suppose that $n \geq 2$. Then for all $\omega < \infty$ sufficiently large, we have

$$\begin{aligned} \dim_H(\text{Sing}_{m,n}(\omega)) &= mn - 2m + \Theta\left(\frac{1}{\omega}\right) & \dim_P(\text{Sing}_{m,n}(\omega)) &= mn - m. \\ &= mn - 2m + \Theta\left(\frac{1}{n} - \tau\right) & & \end{aligned}$$

Theorem 1.7. Suppose that $n = 1$ and $m \geq 2$. Then for all $\omega < 1$ sufficiently close to 1, we have

$$\begin{aligned} \dim_H(\text{Sing}_{m,n}(\omega)) &= \Theta(1 - \omega) & \dim_P(\text{Sing}_{m,n}(\omega)) &= 1. \\ &= \Theta\left(\frac{m-1}{2m} - \tau\right) & & \end{aligned}$$

Remark 1.8. Call a matrix A *trivially singular* if there exists $j = 1, \dots, d-1$ such that

$$\log \lambda_{j+1}(g_t u_A \mathbb{Z}^d) - \log \lambda_j(g_t u_A \mathbb{Z}^d) \rightarrow \infty \text{ as } t \rightarrow \infty.$$

Then all of the above formulas remain true if $\text{Sing}_{m,n}(\omega)$ is replaced by the set

$$\text{Sing}_{m,n}^*(\omega) = \{A \in \text{Sing}_{m,n}(\omega) : A \text{ is not trivially singular}\}.$$

Moreover, for $n \geq 2$ we have

$$\dim_H(\text{Sing}_{m,n}^*(\infty)) = mn - 2m \quad \dim_P(\text{Sing}_{m,n}^*(\infty)) = mn - m$$

and for $n = 1, m \geq 2$ we have

$$\dim_H(\text{Sing}_{m,n}^*(1)) = 0 \quad \dim_P(\text{Sing}_{m,n}^*(1)) = 1.$$

Note that the class of trivially singular matrices is smaller than the class of matrices with degenerate trajectories in the sense of [5, Definition 2.8], but larger than the class considered in [1, p. 2] consisting of matrices A such that the group $A\mathbb{Z}^n + \mathbb{Z}^m$ does not have full rank. A $d \times 1$ or $1 \times d$ matrix is trivially singular if and only if it is contained in a rational hyperplane of \mathbb{R}^d .

1.3. 1×2 and 2×1 matrices

Although we cannot give precise formulas for the Hausdorff and packing dimensions of $\text{Sing}_{m,n}(\omega)$ for all pairs (m, n) , the special cases $(m, n) = (1, 2)$ and $(m, n) = (2, 1)$ are easier to handle.

Theorem 1.9. For all $\omega \in (1/2, 1)$ we have

$$\begin{aligned} \dim_H(\text{Sing}_{1,2}(\omega)) &= \begin{cases} \frac{4}{3} - \frac{4}{3}\sqrt{\tau - 6\tau^3 + 4\tau^4} - 2\tau + \frac{8}{3}\tau^2 & \text{if } \tau \leq \tau_0 \stackrel{\text{def}}{=} \frac{3\sqrt{2}-2}{14} \\ \frac{1-2\tau}{1+\tau} & \text{if } \tau \geq \tau_0 \end{cases} \\ \dim_P(\text{Sing}_{1,2}(\omega)) &= \begin{cases} \frac{4-8\tau}{3} & \text{if } \tau \leq \tau_1 \stackrel{\text{def}}{=} \frac{1}{8} \\ 1 & \text{if } \tau \geq \tau_1 \end{cases} \end{aligned}$$

(cf. Fig. 1).

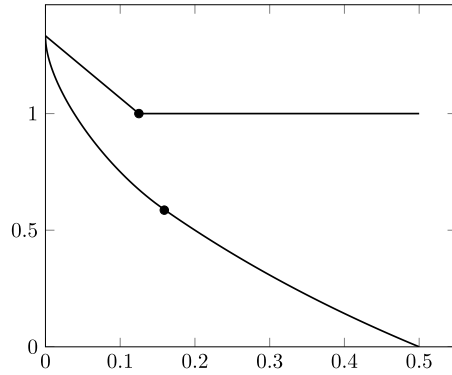


Fig. 1. The functions $f_0(\tau) = \dim_H(\text{Sing}_{1,2}(\omega))$ and $f_1(\tau) = \dim_P(\text{Sing}_{1,2}(\omega))$. The function f_0 is real-analytic on the intervals $[0, \tau_0]$ and $[\tau_0, 1/2]$, where $\tau_0 = \frac{3\sqrt{2}-2}{14} \sim 0.1602$, while f_1 is linear on the intervals $[0, 1/8]$ and $[1/8, 1/2]$.

Remark. By Jarník’s identity [7], for all $\omega \in [2, \infty]$ we have

$$\text{Sing}_{1,2}(\omega) = \text{Sing}_{2,1}(\omega')$$

where $\omega' = 1 - \frac{1}{\omega}$. Thus by applying an appropriate substitution to the above formulas, it is possible to get explicit formulas for $\dim_H(\text{Sing}_{2,1}(\omega'))$ and $\dim_P(\text{Sing}_{2,1}(\omega'))$, either in terms of ω' or in terms of $\tau' = \frac{\omega' - \frac{1}{2}}{\omega' + 1} = \frac{\tau}{1 + 2\tau}$. However, the resulting formulas are not very elegant so we omit them.

Remark. The transition point $\tau_0 = \frac{3\sqrt{2}-2}{14}$ in the above formula for Hausdorff dimension corresponds to $\omega_0 = 2 + \sqrt{2}$, $\omega'_0 = \frac{\sqrt{2}}{2}$, $\tau'_0 = \frac{4-3\sqrt{2}}{2}$, and $\dim_H(\text{Sing}_{1,2}(\omega_0)) = 2 - \sqrt{2}$. The transition point $\tau_1 = \frac{1}{8}$ for packing dimension corresponds to $\omega_1 = 3$, $\omega'_1 = \frac{2}{3}$, $\tau'_1 = \frac{1}{10}$, and $\dim_P(\text{Sing}_{1,2}(\omega_1)) = 1$.

Remark. Theorem 1.9 implies that $\dim_H(\text{Sing}_{1,2}(\omega)) < \dim_P(\text{Sing}_{1,2}(\omega))$ for all $\omega \in (1/2, 1)$. This answers the first part of [1, Problem 7] in the affirmative.

Remark. There has been a lot of partial progress towards the Hausdorff dimension part of Theorem 1.9. In particular, the \geq direction follows from [1, Corollary 2 and Theorem 3]. For $\tau \geq \tau_0$ the upper bound follows from [1, Corollary 2] and for $\tau < \tau_0$, a non-optimal upper bound is given in [1, Theorem 1]. We refer to [1] for a detailed history of the prior results.

1.4. Singularity on average

A different way of quantifying the notion of singularity is the notion of *singularity on average* introduced in [8]. Given a matrix A , we define the *proportion of time spent in the cusp* to be the number

$$\mathcal{P}(A) = \lim_{\epsilon \rightarrow 0} \liminf_{T \rightarrow \infty} \frac{1}{T} \lambda(\{t \in [0, T] : \lambda_1(g_t u_A \mathbb{Z}^d) \leq \epsilon\}) \in [0, 1],$$

where λ denotes Lebesgue measure. The matrix A is said to be *singular on average* if $\mathcal{P}(A) = 1$. Clearly, every singular matrix is singular on average.

Theorem 1.10. For all $p \in [0, 1]$, we have

$$\dim_H(\{A : \mathcal{P}(A) = p\}) = \dim_P(\{A : \mathcal{P}(A) = p\}) = p\delta_{m,n} + (1 - p)mn.$$

In particular, the dimension of the set of matrices singular on average is $\delta_{m,n}$.

Note that the fact that the Hausdorff dimension of the set of matrices singular on average is $\leq \delta_{m,n}$ was proven in [8], while the fact that this number is $\geq \delta_{m,n}$ follows from Theorem 1.1.

1.5. Starkov’s conjecture

In [22, p. 213], Starkov asked whether there exists a singular vector (i.e. $m \times 1$ singular matrix) which is not very well approximable. Here, we recall that a matrix A is called *very well approximable* if for some $\omega > \frac{n}{m}$, there exist infinitely many pairs $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^m \times \mathbb{Z}^n$ such that

$$\|A\mathbf{q} + \mathbf{p}\| \leq \|\mathbf{q}\|^{-\omega}, \quad (1.2)$$

or equivalently in terms of the Dani correspondence principle, a matrix A is very well approximable if $\limsup_{t \rightarrow \infty} -\frac{1}{t} \log \lambda_1(g_t u_A \mathbb{Z}^d) > 0$. This question was answered affirmatively by Cheung [3, Theorem 1.4] in the case $m = 2$. In fact, Cheung showed that if ψ is any function such that $q^{1/2} \psi(q) \rightarrow 0$ as $q \rightarrow \infty$, then there exists a 2×1 singular vector which is not ψ -approximable. Here, a matrix A is called ψ -approximable if there exist infinitely many pairs $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^m \times \mathbb{Z}^n$ such that $\mathbf{q} \neq \mathbf{0}$ and

$$\|A\mathbf{q} + \mathbf{p}\| \leq \psi(\|\mathbf{q}\|).$$

The following theorem improves on Cheung's result both by generalizing it to the case of arbitrary m, n , and also by computing the dimension of the set of matrices with the given property:

Theorem 1.11. *If ψ is any function such that $q^{n/m} \psi(q) \rightarrow 0$ as $q \rightarrow \infty$, then the set of $m \times n$ singular matrices that are not ψ -approximable has Hausdorff dimension $\delta_{m,n}$. Equivalently, if ϕ is any function such that $\phi(t) \rightarrow \infty$ as $t \rightarrow \infty$, then the set of $m \times n$ singular matrices A such that $-\log \lambda_1(g_t u_A \mathbb{Z}^d) \leq \phi(t)$ for all t sufficiently large has Hausdorff dimension $\delta_{m,n}$. The same is true for the packing dimension.*

Note that this theorem is optimal in the sense that if $\psi(q) \geq cq^{-n/m}$ for some constant c , then it is easy to check that every singular $m \times n$ matrix is ψ -approximable.

1.6. Schmidt's conjecture

In [19, p. 273], Schmidt conjectured that for all $2 \leq k \leq m$, there exists an $m \times 1$ matrix A such that

$$\lambda_{k-1}(g_t u_A \mathbb{Z}^d) \rightarrow 0 \text{ and } \lambda_{k+1}(g_t u_A \mathbb{Z}^d) \rightarrow \infty \text{ as } t \rightarrow \infty. \quad (1.3)$$

(Note that any matrix satisfying (1.3) is singular by Theorem 1.2.) This conjecture was proven by Moshchevitin [14], who constructed an $m \times 1$ matrix A satisfying (1.3) and not contained in any rational hyperplane of \mathbb{R}^m (see also [9,17]). We will improve Moshchevitin's result by computing the Hausdorff and packing dimensions of the set of matrices witnessing this conjecture:

Theorem 1.12. *For all $(m, n) \neq (1, 1)$ and for all $2 \leq k \leq m + n - 1$, the Hausdorff and packing dimensions of the set of matrices A that satisfy (1.3) are both equal to*

$$\max(f_{m,n}(k), f_{m,n}(k-1))$$

where

$$f_{m,n}(k) = mn - \frac{kmn}{m+n} \left(1 - \frac{k}{m+n}\right) - \left\{ \frac{km}{m+n} \right\} \left\{ \frac{kn}{m+n} \right\}. \quad (1.4)$$

Here $\{x\}$ denotes the fractional part of a real number x . The same formulas are valid for the set of matrices A that satisfy (1.3) and are not trivially singular.

Remark 1.13. The function $f_{m,n}$ satisfies $f_{m,n}(m+n-k) = f_{m,n}(k)$ and $f_{m,n}(1) = f_{m,n}(m+n-1) = \delta_{m,n}$. Moreover, for all $1 \leq k \leq m+n-1$ we have $f_{m,n}(k) \leq \delta_{m,n}$. It follows that when $k=2$ or $m+n-1$, the Hausdorff and packing dimensions of the set of matrices A that satisfy (1.3) are both equal to $\delta_{m,n}$.

Remark 1.14. When $m=1$ or $n=1$, the fractional parts appearing in (1.4) can be computed explicitly, leading to the formula

$$f_{m,n}(k) = mn - \frac{k(m+n-k)}{m+n}.$$

However, this formula is not valid when $m, n \geq 2$.

2. The variational principle

All the theorems in the previous section (with the exception of Theorems 1.2 and 1.3) are consequences of a single *variational principle* in the parametric geometry of numbers. This variational principle is a quantitative analogue of theorems due to Schmidt and Summerer [21, §2] and Roy [15, Theorem 1.3]. However, we will state their results in language somewhat different from the language used in their papers, due to the fact that the fundamental object we consider is the one-parameter family of unimodular lattices $(g_t u_A \mathbb{Z}^d)_{t \geq 0}$ used by the Dani correspondence principle, rather than a one-parameter family of (non-unimodular) convex bodies as is done in [21,15]. We leave it to the reader to verify that the

theorems we attribute below to [21] and [15] are indeed faithful translations of their results to our setting. We note that these papers, unlike ours, do not consider the case of matrices.

The fundamental question of our version of the parametric geometry of numbers will be as follows: given a matrix A , what does the function $\mathbf{h} = \mathbf{h}_A = (h_1, \dots, h_d) : [0, \infty) \rightarrow \mathbb{R}^d$ defined by the formula

$$h_i(t) \stackrel{\text{def}}{=} \log \lambda_i(g_t u_A \mathbb{Z}^d) \tag{2.1}$$

look like? The function \mathbf{h}_A will be called the *successive minima function* of the matrix A . The Dani correspondence principle shows that many interesting Diophantine questions about the matrix A are equivalent to questions about its successive minima function.

The main restriction on the successive minima function comes from an application of Minkowski’s second theorem on successive minima to certain subgroups of $g_t u_A \mathbb{Z}^d$. Specifically, fix $j = 1, \dots, d - 1$ and let I be an interval such that $h_j(t) < h_{j+1}(t)$ for all $t \in I$. For each $t \in I$, let $V_t \subset \mathbb{R}^d$ be the linear span of the set

$$\{\mathbf{r} \in \mathbb{Z}^d : \|g_t u_A \mathbf{r}\| \leq \lambda_j(g_t u_A \mathbb{Z}^d)\}.$$

Then a continuity argument shows that the map $t \mapsto V_t$ is constant on I , see [20, Lemma 2.1] for the case of simultaneous approximation. Write $V_t = V$. By Minkowski’s second theorem, we have

$$\sum_{i \leq j} h_i(t) \asymp_+ F_{j,I}(t) \stackrel{\text{def}}{=} \log \|g_t u_A (V \cap \mathbb{Z}^d)\|,$$

where $\|\Gamma\|$ denotes the covolume of a discrete group $\Gamma \subset \mathbb{R}^d$, and $A \asymp_+ B$ means that there exists a constant C such that $|B - A| \leq C$. Now an argument based on the exterior product formula for covolume and the definition of g_t shows that $F_{j,I} \asymp_+ G_{j,I}$ for some convex, piecewise linear function $G_{j,I}$ whose slopes are in the set

$$Z(j) \stackrel{\text{def}}{=} \left\{ \frac{L_+}{m} - \frac{L_-}{n} : L_{\pm} \in [0, d_{\pm}] \mathbb{Z}, L_+ + L_- = j \right\}, \tag{2.2}$$

where for convenience we write $d_+ = m$, $d_- = n$, and $[a, b]_{\mathbb{Z}} = [a, b] \cap \mathbb{Z}$. This suggests that \mathbf{h} can be approximated by a piecewise linear function \mathbf{f} such that whenever $f_j < f_{j+1}$ on an interval I , the function $F_j = \sum_{i \leq j} f_i$ is convex on I with slopes in $Z(j)$. Moreover, it is obvious that $h_1 \leq \dots \leq h_d$, and the formula for g_t implies that for all i , we have $-\frac{1}{n} \leq h'_i \leq \frac{1}{m}$ wherever h_i is differentiable. We therefore make the following definition:

Definition 2.1. An $m \times n$ *template* is a piecewise linear map $\mathbf{f} : [t_0, \infty) \rightarrow \mathbb{R}^d$ with the following properties:

- (I) $f_1 \leq \dots \leq f_d$.
- (II) $-\frac{1}{n} \leq f'_i \leq \frac{1}{m}$ for all i .
- (III) For all $j = 1, \dots, d$ and for every interval I such that $f_j < f_{j+1}$ on I , the function $F_j := \sum_{i \leq j} f_i$ is convex on I with slopes in $Z(j)$. Here we use the convention that $f_0 = -\infty$ and $f_{d+1} = +\infty$.

When $m = 1$, templates are a slight generalization of reparameterized versions of the *rigid systems* of [15]. We denote the space of $m \times n$ templates by $\mathcal{T}_{m,n}$.

The fundamental relation between templates and successive minima functions is given as follows:

Theorem 2.2.

- (i) For every $m \times n$ matrix A , there exists an $m \times n$ template \mathbf{f} such that $\mathbf{h}_A \asymp_+ \mathbf{f}$.
- (ii) For every $m \times n$ template \mathbf{f} , there exists an $m \times n$ matrix A such that $\mathbf{h}_A \asymp_+ \mathbf{f}$.

In the case $m = 1$, part (i) of Theorem 2.2 is due to Schmidt and Summerer [21, §2] and part (ii) is due to Roy ([15, Theorem 1.3] and [16, Corollary 4.7]).

Theorem 2.2(ii) asserts that for every template \mathbf{f} , the set

$$\mathcal{M}(\mathbf{f}) \stackrel{\text{def}}{=} \{A : \mathbf{h}_A \asymp_+ \mathbf{f}\}$$

is nonempty. It is natural to ask how big this set is in terms of Hausdorff and packing dimension. Moreover, given a collection of templates \mathcal{F} , we can ask the same question about the set

$$\mathcal{M}(\mathcal{F}) = \bigcup_{\mathbf{f} \in \mathcal{F}} \mathcal{M}(\mathbf{f}).$$

It turns out to be easier to answer the second question than the first, assuming that the collection of templates \mathcal{F} is closed under finite perturbations. Here, \mathcal{F} is said to be *closed under finite perturbations* if whenever $\mathbf{g} \asymp_+ \mathbf{f} \in \mathcal{F}$, we have $\mathbf{g} \in \mathcal{F}$.

Theorem 2.3 (Variational principle, version 1). Let \mathcal{F} be a collection of templates closed under finite perturbations. Then

$$\dim_H(\mathcal{M}(\mathcal{F})) = \sup_{\mathbf{f} \in \mathcal{F}} \underline{\delta}(\mathbf{f}), \quad \dim_P(\mathcal{M}(\mathcal{F})) = \sup_{\mathbf{f} \in \mathcal{F}} \bar{\delta}(\mathbf{f}) \tag{2.3}$$

where the functions $\underline{\delta}, \bar{\delta} : \mathcal{T}_{m,n} \rightarrow [0, mn]$ are as in Definition 2.5 below.

Corollary 2.4. With \mathcal{F} as above, we have

$$\dim_H(\mathcal{M}(\mathcal{F})) = \sup_{\mathbf{f} \in \mathcal{F}} \dim_H(\mathcal{M}(\mathbf{f})), \quad \dim_P(\mathcal{M}(\mathcal{F})) = \sup_{\mathbf{f} \in \mathcal{F}} \dim_P(\mathcal{M}(\mathbf{f})). \tag{2.4}$$

However, note that Theorem 2.3 does not imply that $\dim_H(\mathcal{M}(\mathbf{f})) = \underline{\delta}(\mathbf{f})$ for an individual template \mathbf{f} , since the family $\{\mathbf{f}\}$ is not closed under finite perturbations. And indeed, since the function $\underline{\delta}$ is sensitive to finite perturbations, the formula $\dim_H(\mathcal{M}(\mathbf{f})) = \underline{\delta}(\mathbf{f})$ cannot hold for all $\mathbf{f} \in \mathcal{T}_{m,n}$.

Definition 2.5. We define the lower and upper average contraction rate of a template \mathbf{f} as follows. Let I be an open interval on which \mathbf{f} is linear. For each $q = 1, \dots, d$ such that $f_q < f_{q+1}$ on I , let $L_{\pm} = L_{\pm}(\mathbf{f}, I, q) \in [0, d_{\pm}]_{\mathbb{Z}}$ be chosen to satisfy $L_+ + L_- = q$ and

$$F'_q = \sum_{i=1}^q f'_i = \frac{L_+}{m} - \frac{L_-}{n} \text{ on } I, \tag{2.5}$$

as guaranteed by (III) of Definition 2.1. An interval of equality for \mathbf{f} on I is an interval $(p, q]_{\mathbb{Z}}$, where $0 \leq p < q \leq d$ satisfy

$$f_p < f_{p+1} = \dots = f_q < f_{q+1} \text{ on } I. \tag{2.6}$$

Note that the collection of intervals of equality forms a partition of $[1, d]_{\mathbb{Z}}$. If $(p, q]_{\mathbb{Z}}$ is an interval of equality for \mathbf{f} on I , then we let $M_{\pm}(p, q) = M_{\pm}(\mathbf{f}, I, p, q)$, where

$$M_{\pm}(\mathbf{f}, I, p, q) = L_{\pm}(\mathbf{f}, I, q) - L_{\pm}(\mathbf{f}, I, p), \tag{2.7}$$

and we let

$$S_+(\mathbf{f}, I) = \bigcup_{(p,q]_{\mathbb{Z}}} (p, p + M_+(p, q)]_{\mathbb{Z}} \tag{2.8}$$

$$S_-(\mathbf{f}, I) = \bigcup_{(p,q]_{\mathbb{Z}}} (p + M_+(p, q), q]_{\mathbb{Z}} \tag{2.9}$$

where the unions are taken over all intervals of equality for \mathbf{f} on I . Note that $M_{\pm}(p, q) \geq 0$ by (II) of Definition 2.1, and further that S_+ and S_- are disjoint and satisfy $S_+ \cup S_- = [1, d]_{\mathbb{Z}}$, and that $\#(S_+) = m$ and $\#(S_-) = n$. Next, let

$$\delta(\mathbf{f}, I) = \#\{(i_+, i_-) \in S_+ \times S_- : i_+ < i_-\} \in [0, mn]_{\mathbb{Z}}, \tag{2.10}$$

and note that

$$mn - \delta(\mathbf{f}, I) = \#\{(i_+, i_-) \in S_+ \times S_- : i_+ > i_-\}. \tag{2.11}$$

The lower and upper average contraction rates of \mathbf{f} are the numbers

$$\underline{\delta}(\mathbf{f}) \stackrel{\text{def}}{=} \liminf_{T \rightarrow \infty} \Delta(\mathbf{f}, T), \quad \bar{\delta}(\mathbf{f}) \stackrel{\text{def}}{=} \limsup_{T \rightarrow \infty} \Delta(\mathbf{f}, T), \tag{2.12}$$

where

$$\Delta(\mathbf{f}, T) \stackrel{\text{def}}{=} \frac{1}{T} \int_0^T \delta(\mathbf{f}, t) dt.$$

Here we abuse notation by writing $\delta(\mathbf{f}, t) = \delta(\mathbf{f}, I)$ for all $t \in I$.

Definition 2.5 can be understood intuitively in terms of a simple version of one-dimensional physics with sticky collisions and conservation of momentum. Suppose that we observe particles P_1, \dots, P_d travelling along trajectories f_1, \dots, f_d during a time interval I along which \mathbf{f} is linear, and we want to infer the velocities of these particles before they collided, based on the following background information: before the collision m of the particles were travelling upwards at a speed of $\frac{1}{m}$, and n of the particles were travelling downwards at a speed of $\frac{1}{n}$. When particles collide (that is, when the velocities of the particles of lower index are more upwards than the velocities of the particles of higher index at the same location), they

join forces to move as a unit, and their new velocity is determined by conservation of momentum. However, we can still think of the group as being composed of a certain number of “upwards” particles and a certain number of “downwards” particles.

The equations (2.8) and (2.9) can be understood as suggesting a particular solution to this problem of inference: assume that within each group, all of the upwards-travelling particles started out below all of the downwards-travelling particles. This is not the only possible solution, but it is the nicest one for certain purposes. Specifically, we can imagine a force of “gravity” attempting to bring all of the particles together, which acts between any two particles by imposing a fixed energy cost if the two particles are travelling away from each other.³ The total energy cost is then the codimension $mn - \delta(\mathbf{f}, I)$ defined by (2.11). The equations (2.8) and (2.9) can then be thought of as giving the solution that minimizes this cost.

The idea of codimension as an energy cost is also useful for computing the suprema (2.3) in certain circumstances, since it suggests principles like the conservation of energy. However, one needs to be careful since the stickiness of collisions means that some naive formulations of conservation of energy are violated.

In most cases of interest, the collection \mathcal{F} in Theorem 2.3 is defined by some Diophantine condition. In this case, generally rather than $\mathcal{M}(\mathcal{F})$ the set we are really interested in is the set of all matrices whose corresponding successive minima functions satisfy the same Diophantine condition. Although these two sets are *a priori* different, Theorem 2.2(i) implies that they are the same and thus Theorem 2.3 is equivalent modulo Theorem 2.2(i) to the following:

Theorem 2.6 (Variational principle, version 2). *Let S be a collection of functions from $[0, \infty)$ to \mathbb{R}^d which is closed under finite perturbations, and let*

$$\mathcal{M}(S) = \{A : \mathbf{h}_A \in S\}.$$

Then

$$\dim_H(\mathcal{M}(S)) = \sup_{\mathbf{f} \in S \cap \mathcal{T}_{m,n}} \underline{\delta}(\mathbf{f}), \quad \dim_P(\mathcal{M}(S)) = \sup_{\mathbf{f} \in S \cap \mathcal{T}_{m,n}} \overline{\delta}(\mathbf{f}). \tag{2.13}$$

Proof of equivalence. Theorem 2.6 implies Theorem 2.3 since we can take $S = \{\mathbf{g} : \mathbf{g} \asymp_+ \mathbf{f} \in \mathcal{F}\}$. Conversely, Theorem 2.3 implies Theorem 2.6 modulo Theorem 2.2(i) since we can take $\mathcal{F} = S \cap \mathcal{T}_{m,n}$. \square

Theorem 2.6 can be thought of as a quantitative analogue of Theorem 2.2, as shown by the following equivalent formulation:

Theorem 2.7 (Variational principle, version 3).

- (i) *Let S be a set of $m \times n$ matrices of Hausdorff (resp. packing) dimension $> \delta$. Then there exists a matrix $A \in S$ and a template $\mathbf{f} \asymp_+ \mathbf{h}_A$ whose lower (resp. upper) average contraction rate is $> \delta$.*
- (ii) *Let \mathbf{f} be a template whose lower (resp. upper) average contraction rate is $> \delta$. Then there exists a set S of $m \times n$ matrices of Hausdorff (resp. packing) dimension $> \delta$, such that $\mathbf{h}_A \asymp_+ \mathbf{f}$ for all $A \in S$.*

Proof of equivalence. Part (i) is equivalent to the \leq direction of (2.13), and part (ii) to the \geq direction. For the first equivalence, for the forwards direction take $S = \{A : \mathbf{h}_A \in S\}$, and for the backwards direction take $S = \{\mathbf{g} : \mathbf{g} \asymp_+ \mathbf{h}_A, A \in S\}$. For the second equivalence, for the backwards direction take $S = \mathcal{M}(\mathbf{f})$ and $S = \{\mathbf{g} : \mathbf{g} \asymp_+ \mathbf{f}\}$. \square

It is worth stating the special case of Theorem 2.6 that occurs when the collection S is defined by the Diophantine conditions defining $\text{Sing}_{m,n}(\omega)$ and $\text{Sing}_{m,n}^*(\omega)$ for some $\omega \geq \frac{n}{m}$. Thus, we define the *uniform dynamical exponent* of a map $\mathbf{f} : [0, \infty) \rightarrow \mathbb{R}^d$ to be the number

$$\widehat{\tau}(\mathbf{f}) = \liminf_{t \rightarrow \infty} \frac{-1}{t} f_1(t).$$

Similarly, \mathbf{f} is said to be *trivially singular* if $f_{j+1}(t) - f_j(t) \rightarrow \infty$ as $t \rightarrow \infty$ for some $j = 1, \dots, d - 1$. Letting $S = \{\mathbf{f} : \widehat{\tau}(\mathbf{f}) = \tau\}$ or $S = \{\mathbf{f} : \widehat{\tau}(\mathbf{f}) = \tau, \mathbf{f} \text{ not trivially singular}\}$ in Theorem 2.6 yields the following result:

Theorem 2.8 (Special case of variational principle). *For all $\omega \geq \frac{n}{m}$, we have*

$$\begin{aligned} \dim_H(\text{Sing}_{m,n}(\omega)) &= \sup\{\underline{\delta}(\mathbf{f}) : \mathbf{f} \in \mathcal{T}_{m,n}, \widehat{\tau}(\mathbf{f}) = \tau\} \\ \dim_P(\text{Sing}_{m,n}(\omega)) &= \sup\{\overline{\delta}(\mathbf{f}) : \mathbf{f} \in \mathcal{T}_{m,n}, \widehat{\tau}(\mathbf{f}) = \tau\} \end{aligned}$$

³ This is of course unlike real gravity, which imposes an energy cost variable with respect to distance.

$$\dim_H(\text{Sing}_{m,n}^*(\omega)) = \sup\{\underline{\delta}(\mathbf{f}) : \mathbf{f} \in \mathcal{T}_{m,n}, \widehat{\tau}(\mathbf{f}) = \tau, \mathbf{f} \text{ not trivially singular}\}$$

$$\dim_P(\text{Sing}_{m,n}^*(\omega)) = \sup\{\overline{\delta}(\mathbf{f}) : \mathbf{f} \in \mathcal{T}_{m,n}, \widehat{\tau}(\mathbf{f}) = \tau, \mathbf{f} \text{ not trivially singular}\},$$

where τ is as in (1.1).

Theorem 2.6 can also be used to compute the dimensions of the set

$$\widetilde{\text{Sing}}_{m,n}^*(\omega) = \{A : \widehat{\omega}(A) \geq \omega, A \text{ not trivially singular}\} = \bigcup_{\omega' \geq \omega} \text{Sing}_{m,n}^*(\omega').$$

Theorem 2.9 (Special case of variational principle). For all $\omega \geq \frac{n}{m}$, we have

$$\dim_H(\widetilde{\text{Sing}}_{m,n}^*(\omega)) = \sup_{\omega' \geq \omega} \dim_H(\text{Sing}_{m,n}^*(\omega'))$$

$$\dim_P(\widetilde{\text{Sing}}_{m,n}^*(\omega)) = \sup_{\omega' \geq \omega} \dim_P(\text{Sing}_{m,n}^*(\omega')).$$

(Theorem 2.9 is also true with the stars removed, but in that case it is not as interesting because $\dim_H(\text{Sing}_{m,n}(\infty))$ is “too large”, whereas $\dim_H(\text{Sing}_{m,n}^*(\infty))$ is the “correct” size according to Remark 1.8.)

It is natural to expect that the map $\omega \mapsto \dim_H(\text{Sing}_{m,n}^*(\omega))$ is monotonically decreasing, in which case Theorem 2.9 would imply that

$$\dim_H(\widetilde{\text{Sing}}_{m,n}^*(\omega)) = \dim_H(\text{Sing}_{m,n}^*(\omega)).$$

Conjecture 2.10. The functions

$$\omega \mapsto \dim_H(\text{Sing}_{m,n}^*(\omega)), \quad \omega \mapsto \dim_P(\text{Sing}_{m,n}^*(\omega))$$

are decreasing and continuous, and furthermore are computable in the sense of [23].

The main difficulty in proving this conjecture is the rigidity of templates – one would like to show that every template can be perturbed into a new template whose uniform dynamical exponent is either slightly larger, or much smaller, than that of the original template, and whose average contraction rates are not too much smaller than those of the original template. However, it is not at all clear how one would perform such a perturbation except in a few special cases.

3. A characterization of Hausdorff and packing dimensions using games

The proof of the variational principle is based on a new variant of Schmidt’s game which is in principle capable of computing the Hausdorff and packing dimensions of any set. In Schmidt’s game [18], players take turns choosing a descending sequence of balls and compete to determine whether or not the intersection point of these balls is in a certain target set. The key feature of our new variant is that instead of requiring the rate at which the players’ moves contribute information to the game to be constant, the new variant allows the rate of information transfer to be variable, with the first player, Alice, getting to choose the rate of information transfer. However, Alice is penalized if she exerts too much control over the game over long periods of time without giving her opponent Bob a chance to exert control over the game.

Definition 3.1. Given $0 < \beta < 1$, Alice and Bob play the δ -dimensional Hausdorff (resp. packing) β -game as follows:

- the turn order is alternating, with Alice playing first. Thus, Bob’s k th turn occurs after Alice’s k th turn and before Alice’s $(k + 1)$ st turn;
- Alice begins by choosing a starting radius $\rho_0 > 0$;
- on the k th turn, Alice chooses a nonempty $3\rho_k$ -separated set $A_k \subset \mathbb{R}^d$, and Bob responds by choosing a ball $B_k = B(\mathbf{x}_k, \rho_k)$, where $\mathbf{x}_k \in A_k$ and $\rho_k = \beta^k \rho_0$;
- on the first (0th) turn, Alice’s choice A_0 can be any finite set, but on subsequent turns she must choose it to satisfy

$$A_{k+1} \subset B(\mathbf{x}_k, (1 - \beta)\rho_k). \tag{3.1}$$

Note that this condition guarantees that

$$B_0 \supset B_1 \supset B_2 \supset \dots$$

After infinitely many turns have passed, the point

$$\mathbf{x}_\infty = \lim_{k \rightarrow \infty} \mathbf{x}_k \in \bigcap_{k=0}^{\infty} B_k$$

is computed (note that the right-hand side is always a singleton). It is called the *outcome* of the game. Also, we let $\mathcal{A} = (A_k)_{k \in \mathbb{N}}$, and we compute the numbers

$$\underline{\delta}(\mathcal{A}) \stackrel{\text{def}}{=} \liminf_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^k \frac{\log \#(A_i)}{-\log(\beta)} \quad (3.2)$$

and

$$\overline{\delta}(\mathcal{A}) \stackrel{\text{def}}{=} \limsup_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^k \frac{\log \#(A_i)}{-\log(\beta)} \quad (3.3)$$

which are called Alice's *lower* and *upper scores*, respectively. Alice's goal will be to ensure that the outcome is in a certain set S , called the *target set*, and simultaneously to maximize her score while doing so.

To be precise, a set $S \subset \mathbb{R}^d$ is said to be *δ -dimensionally Hausdorff (resp. packing) β -winning* if Alice has a strategy to simultaneously ensure that the outcome \mathbf{x}_∞ is in S , and that her lower (resp. upper) score is at least δ . It is said to be *δ -dimensionally Hausdorff (resp. packing) winning* if it is δ -dimensionally Hausdorff (resp. packing) β -winning for all sufficiently small $\beta > 0$. The reader may contrast the definition of our game with that of Cheung's *self-similar coverings* in [3, Section 3].

The following result is one of the key ingredients in the proof of the variational principle:

Theorem 3.2. *The Hausdorff (resp. packing) dimension of a Borel set S is the supremum of δ such that S is δ -dimensionally Hausdorff (resp. packing) winning.*

Acknowledgements

This research began when the authors met at the American Institute of Mathematics via their SQuaRE program. We thank them for their hospitality. The first-named author was supported in part by a 2017–2018 Faculty Research Grant from the University of Wisconsin–La Crosse. The second-named author was supported in part by the Simons Foundation grant #245708. The third-named author was supported by the EPSRC Programme Grant EP/J018260/1. The fourth-named author was supported in part by the National Science Foundation grant DMS-1361677. We thank Nicolas Chevallier, Elon Lindenstrauss, Damien Roy, and Johannes Schleischitz for helpful comments.

References

- [1] Y. Bugeaud, Y. Cheung, N. Chevallier, Hausdorff dimension and uniform exponents in dimension two, <https://arxiv.org/abs/1610.06374>, 2016.
- [2] J.W.S. Cassels, *An Introduction to Diophantine Approximation*, Cambridge Tracts in Mathematics and Mathematical Physics, vol. 45, Cambridge University Press, New York, 1957.
- [3] Y. Cheung, Hausdorff dimension of the set of singular pairs, *Ann. of Math. (2)* 173 (1) (2011) 127–167.
- [4] Y. Cheung, N. Chevallier, Hausdorff dimension of singular vectors, *Duke Math. J.* 165 (12) (2016) 2273–2329, MR 3544282.
- [5] S.G. Dani, Divergent trajectories of flows on homogeneous spaces and Diophantine approximation, *J. Reine Angew. Math.* 359 (1985) 55–89.
- [6] H. Davenport, W.M. Schmidt, Dirichlet's theorem on Diophantine approximation. II, *Acta Arith.* 16 (1969/1970) 413–424, MR 0279040.
- [7] V. Jarník, Zum Khintchineschen "Übertragungssatz", *Trav. Inst. Math. Tbilissi* 3 (1938) 193–212 (in German).
- [8] S. Kadyrov, D. Kleinbock, E. Lindenstrauss, G. Margulis, Singular systems of linear forms and non-escape of mass in the space of lattices, <http://arxiv.org/abs/1407.5310>, 2014.
- [9] A. Keita, On a conjecture of Schmidt for the parametric geometry of numbers, *Mosc. J. Comb. Number Theory* 6 (2–3) (2016) 166–176.
- [10] A. Khinchin, Über eine Klasse linearer diophantischer Approximationen, *Rend. Circ. Mat. Palermo* 50 (1926) 170–195 (in German).
- [11] A. Khinchin, Über singuläre Zahlensysteme, *Compos. Math.* 4 (1937) 424–431, MR 1556985.
- [12] A. Khinchin, Regular systems of linear equations and a general problem of Čebyšev, *Izv. Akad. Nauk SSSR, Ser. Mat.* 12 (1948) 249–258, MR 0025513.
- [13] N. Moshchevitin, Khintchine's singular Diophantine systems and their applications, *Russ. Math. Surv.* 65 (3) (2010) 433–511.
- [14] N. Moshchevitin, Proof of W.M. Schmidt's conjecture concerning successive minima of a lattice, *J. Lond. Math. Soc. (2)* 86 (1) (2012) 129–151, MR 2959298.
- [15] D. Roy, On Schmidt and Summerer parametric geometry of numbers, *Ann. of Math. (2)* 182 (2) (2015) 739–786, MR 3418530.
- [16] D. Roy, Spectrum of the exponents of best rational approximation, *Math. Z.* 283 (1–2) (2016) 143–155, MR 3489062.
- [17] J. Schleischitz, Diophantine approximation and special Liouville numbers, *Commun. Math.* 21 (1) (2013) 39–76, MR 3067121.
- [18] W.M. Schmidt, On badly approximable numbers and certain games, *Trans. Amer. Math. Soc.* 123 (1966) 27–50.
- [19] W.M. Schmidt, Open problems in Diophantine approximation, in: *Diophantine Approximations and Transcendental Numbers*, Luminy, 1982, in: *Prog. Math.*, vol. 31, Birkhäuser Boston, Boston, MA, 1983, pp. 271–287, MR 702204.
- [20] W.M. Schmidt, L. Summerer, Parametric geometry of numbers and applications, *Acta Arith.* 140 (1) (2009) 67–91, MR 2557854.
- [21] W.M. Schmidt, L. Summerer, Diophantine approximation and parametric geometry of numbers, *Monatshefte Math.* 169 (1) (2013) 51–104, MR 3016519.

- [22] A. Starkov, *Dynamical Systems on Homogeneous Spaces*, Translations of Mathematical Monographs, vol. 190, American Mathematical Society, Providence, RI, 2000, translated from the 1999 Russian original by the author. MR 1746847.
- [23] K. Weihrauch, *Computable Analysis. An Introduction*, Texts in Theoretical Computer Science. An EATCS Series, Springer-Verlag, Berlin, 2000, MR 1795407.