Received: 1 March 2024 | Revised: 10 April 2024 | Accepted: 16 April 2024

DOI: 10.1112/mtk.12262

RESEARCH ARTICLE

Mathematika

Badly approximable grids and k-divergent lattices

Nikolay Moshchevitin[1,2] | **Anurag Rao**[3] | **Uri Shapira**[4]

[1]Mathematics Department, Technische Universität Wien, Wien, Austria

[2]Faculty of Computer Science, HSE University, Moscow Center of Fundamental and Applied Mathematics, Moscow, Russia

[3]Mathematics Department, BICMR, Peking University, Beijing, China

[4]Mathematics Department, Technion – Israel Institute for Technology, Haifa, Israel

Correspondence
Uri Shapira, Mathematics Department, Technion – Israel Institute for Technology, Haifa, Israel.
Email: ushapira@technion.ac.il

Funding information
European Research Council, Grant/Award Number: 754475; Austrian Science Fund, Grant/Award Number: I 5554; RSF, Grant/Award Number: 22-41-05001

Abstract

Let $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ be a matrix. In this paper, we investigate the set $\mathrm{Bad}_A \subset \mathbb{T}^m$ of badly approximable targets for A, where $\mathbb{T}^m$ is the m-torus. It is well known that Bad_A is a winning set for Schmidt's game and hence is a dense subset of full Hausdorff dimension. We investigate the relationship between the measure of Bad_A and Diophantine properties of A. On the one hand, we give the first examples of a nonsingular A such that Bad_A has full measure with respect to some nontrivial algebraic measure on the torus. For this, we use transference theorems due to Jarnik and Khintchine, and the parametric geometry of numbers in the sense of Roy. On the other hand, we give a novel Diophantine condition on A that slightly strengthens nonsingularity, and show that under the assumption that A satisfies this condition, Bad_A is a null-set with respect to any nontrivial algebraic measure on the torus. For this, we use naive homogeneous dynamics, harmonic analysis, and a novel concept that we refer to as mixing convergence of measures.

MSC 2020
11-XX, 37-XX (primary)

1 | INTRODUCTION

1.1 | The origins of this work

This paper stems from an effort to understand the state of affairs regarding the validity of several statements claimed to be true in [26, sections 2.2, 3.2], whose proof relied on a careless observation made by the third author, which is false, as was pointed out to us by David Simmons.

In this paper, we revisit this discussion and prove some of the statements claimed in [26] under slightly stronger assumptions. We also provide constructions of examples showing that some of the statements made in [26] are false and maybe more importantly, we present some open problems.

In order to present the discussion in an organic manner and keep the reader motivated, we open with a self-contained introduction to the subject. In Section 2.8, we will elaborate further regarding what is exactly the mistake in [26] and the current state of affairs regarding the questionable statements there.

1.2 | Inhomogeneous Diophantine approximation

One of the main themes in the theory of inhomogeneous Diophantine approximations is to analyze, for a matrix

$$A = \begin{pmatrix} \theta^1 & \cdots & \theta^n \end{pmatrix} \in \mathrm{Mat}_{m\times n}(\mathbb{R}),$$

the rate at which the group generated by the columns of A in $\mathbb{R}^m$ approximates a given target vector $\eta \in \mathbb{R}^m$ modulo the integers. More precisely, let $\|\cdot\|$ denote choices of norms on $\mathbb{R}^n$ and $\mathbb{R}^m$ and let $\langle\cdot\rangle$ denote the induced distance on the m-torus $\mathbb{T}^m := \mathbb{R}^m/\mathbb{Z}^m$. This theory tries to understand, for a given $\eta \in \mathbb{T}^m$ the rate at which the sequence

$$\min\{\langle Aq - \eta\rangle : q \in \mathbb{Z}^n, \|q\| \leqslant Q\}$$

approaches zero (if at all) as $Q \to \infty$. One way to do this is to fix a monotonely increasing function $\psi : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ and investigate

$$\liminf_{q\in\mathbb{Z}^n, \|q\|\to\infty} \psi(\|q\|)\langle Aq - \eta\rangle. \tag{1.1}$$

The most natural and widely investigated function ψ is $\psi(t) = t^{n/m}$. A heuristic reason for why this is a natural choice for ψ is that under the constraint $\|q\| \leqslant Q$ on the coefficient vector $q \in \mathbb{Z}^n$, we have roughly Q^n points in $\{Aq \mod \mathbb{Z} : \|q\| \leqslant Q\} \subset \mathbb{T}^m$ and since the m-torus is m-dimensional, if we split it to Q^n boxes of the same size, their side length is $Q^{-n/m}$. This is why the rescaling by $Q^{n/m}$ leads to an interesting discussion.

The most basic question one might ask about (1.1) is whether it is positive or not. Indeed a classical subset of the torus, which is widely investigated, is the set of *badly approximable targets* *for A*:

$$\mathrm{Bad}_A := \left\{ \eta \in \mathbb{T}^m : \liminf_{q\in\mathbb{Z}^n, \|q\|\to\infty} \|q\|^{n/m} \langle Aq - \eta \rangle > 0 \right\}. \tag{1.2}$$

In this paper, we investigate the structure of Bad_A and its relation to Diophantine conditions on A. The following classical result says that Bad_A is never empty.

Theorem 1.1 (Theorem X, Chapter IV [4]). *For any $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, there exists an $\eta \in \mathbb{R}^m$ for which*

$$\inf_{q\in\mathbb{Z}^n\setminus\{0\}} \|q\|^{n/m} \langle Aq - \eta \rangle > 0. \tag{1.3}$$

This result was amplified significantly as follows:

Theorem 1.2 (Theorem 1.4 in [8] or Theorem 1 in [16]). *For any $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, the set $\mathrm{Bad}_A \subset \mathbb{T}^m$ is a winning set for Schmidt's game. As a consequence, it is a dense subset of full Hausdorff dimension.*

For the notion of winning and Schmidt games, see the seminal paper [23]. See also [2] where the fact that Bad_A has full Hausdorff dimension was first proved. Theorem 1.2 says that Bad_A is large from the dimension point of view. In this paper, we are interested in investigating its size from the measure theoretical point of view. The basic question that we investigate is the following:

Question 1.3. Given a probability measure μ on $\mathbb{T}^m$, what can be said about $\mu(\mathrm{Bad}_A)$ and how does this relate to the Diophantine properties of A.

To examplify a situation where Question 1.3 has an easy answer, we note the following. Consider the case $n = 1$ in which case $A = \theta$ is a single column vector (or rather a point in $\mathbb{T}^m$). Assume its coordinates satisfy that $1, \theta_1, \ldots, \theta_m$ are linearly dependent over $\mathbb{Q}$. This is equivalent to the fact that the cyclic subgroup generated by θ in $\mathbb{T}^m$ is not dense. In fact, in such a case, $\overline{\{q\theta : q \in \mathbb{Z}\}}$ is a union of finitely many cosets of a lower dimensional subtorus. In particular, any η outside this lower dimensional submanifold belongs to Bad_θ. In particular, we get

$$\mu(\mathrm{Bad}_\theta) = 1$$

for various measures μ. For example, for μ being the Lebesgue measure on $\mathbb{T}^m$ as well as many *algebraic measures* according to the following definition.

Definition 1.4. A probability measure μ on $\mathbb{T}^m$ is said to be algebraic if there is a subspace $U < \mathbb{R}^m$ whose image in $\mathbb{T}^m$ is closed (i.e., $U \cap \mathbb{Z}^m$ is a lattice in U), and μ is the U-invariant probability measure supported on a single U-orbit in $\mathbb{T}^m$. We exclude the possibility of $U = \{0\}$ by saying that μ is nontrivial.

The Diophantine condition saying that $1, \theta_1, \ldots, \theta_m$ are linearly dependent over $\mathbb{Q}$ is very strong. A relaxation of it is the famous notion of *singularity*:

Definition 1.5 (Singularity). We say that $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ is singular if, for every $\varepsilon > 0$, for all large enough Q,

$$Q^{n/m} \min\{\langle Aq\rangle : q \in \mathbb{Z}^n \setminus \{0\}, \|q\| \leqslant Q\} < \varepsilon. \tag{1.4}$$

Assuming nonsingularity of A, we have the following result from [26] that answers Question 1.3 for the Lebesgue measure on $\mathbb{T}^m$.

Theorem 1.6 [26]. *Let $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ be nonsingular and let $\lambda_{\mathbb{T}^m}$ denote the Lebesgue probability measure on $\mathbb{T}^m$. Then, $\lambda_{\mathbb{T}^m}(\mathrm{Bad}_A) = 0$.*

As pointed out in Section 1.1, there is a mistake in the paper [26]. Theorem 1.6 is a special case of [26, Theorem 2.3], which in the generality stated there turns out to be false as we will see below. But, the argument presented there, for the case of the measure $\lambda_{\mathbb{T}^m}$, is robust enough to carry through. We will provide a full proof of Theorem 1.6 reproducing the argument of [26] in Section 4. This proof will also serve as an introduction to the proof of one of our main results Theorem 2.11 in Section 6. We note that since [26], several alternative proofs for Theorem 1.6 have appeared. See [13, Corollary 1.4], [18, Theorem 1], and [1, Theorem 1.3] and also [12, Theorem 1.2] for the 1-dimensional case. See also Remark 1.21 and Theorem 8.4 for a slightly stronger result than Theorem 1.6, which we prove in Section 8.

Since David Simmons spotted the gap in [26, sections 2.2, 3.2], the third author tried to rule regarding the validity/falsity of the results there, most notably regarding [26, Theorem 2.3], which implies that if A is nonsingular, then $\mu(\mathrm{Bad}_A) = 0$ for every nontrivial algebraic measure μ on $\mathbb{T}^m$. One of the main results we present in this paper is the following construction, which shows that this statement may fail drastically. In it we choose $n = 1$ and take $A = \theta$ to be a vector.

Theorem 1.7. *Let $m \in \mathbb{N}$ with $m > 2$. There exists $\theta \in \mathbb{R}^m$, which is nonsingular and an $\eta \in \mathbb{R}^m$ such that, for every $(t_3, \dots, t_m) \in \mathbb{R}^{m-2}$, we have*

$$\inf_{q\in\mathbb{Z}\setminus\{0\}} |q|^{1/m}\langle q\theta - (\eta + t_3\mathbf{e}_3 + \cdots + t_m\mathbf{e}_m)\rangle > 0.$$

In particular, Bad_θ contains a coset of a subtorus of codimension 2 and as a consequence, there are nontrivial algebraic measures μ satisfying $\mu(\mathrm{Bad}_\theta) = 1$.

This theorem is proved as Theorem 7.7. Here, $\mathbf{e}_3, \dots, \mathbf{e}_m$ denote last $m-2$ standard basis vectors for $\mathbb{R}^m$.

Remark 1.8. We note here that it is very likely that Theorem 1.7 could be generalized to $n > 1$. This would probably involve using the new theory of parametric geometry of numbers as in [7]. We were content with using this theory as presented in [20], which seems to give only the $n = 1$ case.

Remark 1.9. We note here that we were unable to provide a construction of nonsingular $\theta \in \mathbb{R}^m$ such that $\mu(\mathrm{Bad}_\theta) > 0$ for an algebraic measure corresponding to a codimension 1 subtorus. In particular, the case $m = 2$ remains open. That is, is there a nonsingular vector $\theta \in \mathbb{R}^2$ and an algebraic measure μ on $\mathbb{T}^2$ of dimension 1 such that $\mu(\mathrm{Bad}_\theta) > 0$?

We proceed to state versions of our main result, which could be viewed as an amendment of [26, Theorem 2.3]. We introduce a novel Diophantine condition on A that strengthens nonsingularity and ensures that $\mu(\mathrm{Bad}_A) = 0$ for any nontrivial algebraic measure μ on $\mathbb{T}^m$. This Diophantine condition is a bit elaborate to state and is dynamical in nature.

Let $d := n + m$. Let X denote the space of unimodular (i.e., covolume 1) lattices in $\mathbb{R}^d$. As usual X is identified with the quotient $\mathrm{SL}_d(\mathbb{R})/\mathrm{SL}_d(\mathbb{Z})$ via $g\,\mathrm{SL}_d(\mathbb{Z}) \leftrightarrow g\mathbb{Z}^d$. This identification gives rise to a topology on X (the quotient topology), as well as to a natural action of $\mathrm{SL}_d(\mathbb{R})$ on X. Given $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, we consider the lattice

$$x_A := \begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix} \mathbb{Z}^d. \tag{1.5}$$

Of particular interest to our discussion is the following one-parameter subgroup:

$$h_t := \begin{bmatrix} e^{nt} I_m & 0 \\ 0 & e^{-mt} I_n \end{bmatrix} \in \mathrm{SL}_d(\mathbb{R}), \quad t \in \mathbb{R}. \tag{1.6}$$

The so-called *Dani correspondence* (see [5]) says that many of the Diophantine properties of A could be read from the topological and statistical properties of the orbit $\{h_t x_A : t \geqslant 0\}$. In particular, we have the following famous characterization of singularity.

Theorem 1.10 [5, Theorem 2.14]. *A matrix $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ is singular if and only if the orbit $\{h_t x_A : t \geqslant 0\}$ is divergent.*

Here the orbit is said to be divergent if the map $t \mapsto h_t x_A$ is a *proper* map from $\mathbb{R}_{\geqslant 0}$ to X (i.e., preimages of compact sets are compact). Our novel Diophantine condition on A is a relaxation of singularity that is stated in terms of a topological property of the orbit $\{h_t x_A : t \geqslant 0\}$.

Definition 1.11 (Asymptotic accumulation points). Let $x \in X$. We define the set of asymptotic accumulation points of x as

$$\partial(x) := \left\{ y \in X : \text{there is an unbounded sequence } (t_k)_{k\in\mathbb{N}} \subset \mathbb{R}_{\geqslant 0} \text{ with } \lim_{k\to\infty} h_{t_k} x = y \right\}.$$

Definition 1.12 (Accumulation sequences and k-divergence). Let $x \in X$. A sequence

$$I = (x_0 = x, x_1, \dots, x_k) \subset X^{k+1} \ \text{ with } \ x_{i+1} \in \partial(x_i) \text{ for } i = 0, \dots k-1$$

is called an accumulation sequence of length $k + 1$ for x. If

$$k = \sup\{\mathrm{length}(I) - 1 : I \text{ is an accumulation sequence for } x\},$$

we say x is k-divergent. A matrix A is said to be k-divergent if the lattice x_A is.

Remark 1.13. The set $\partial(x)$ and as a consequence the definition of k-divergent lattices depend on the semigroup $\{h_t : t \geqslant 0\}$. Versions of these notions could be investigated for other groups and semigroups. We choose to work with this particular semigroup because of the relation to Diophantine approximation given by the Dani correspondence as reflected in Theorem 1.10.

Note that by Theorem 1.10, saying that A is singular is the same as saying it is 0-divergent. In the following results, we assume A is not k-divergent for small values of k. This gives enough dynamical richness for the orbit $\{h_t x_A : t \geqslant 0\}$ for our arguments to carry through.

Theorem 1.14. *Assume $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ is not k-divergent for any $0 \leqslant k \leqslant d-2$ (in other words, the lattice x_A has an accumulation sequence of length d), and that $\gcd(m,n) = 1$. Then, for any nontrivial algebraic measure μ on $\mathbb{T}^m$, we have*

$$\mu(\mathrm{Bad}_A) = 0.$$

Remark 1.15. The assumption $\gcd(n,m) = 1$ appearing in some of our results is curious. We are not sure how much of it is an artifact of the method of proof. See Lemma 6.2 and Corollary 6.3 where this assumption enters the discussion. This assumption is not entirely redundant though, as the case $n = m$ is generally false. See Example 2.12.

Theorem 1.14 could be restated more explicitly as follows. Since any ℓ-dimensional subtorus of $\mathbb{T}^m$ can be presented as a product of a 1-dimensional subtorus and an $\ell - 1$-dimensional one, by Fubini's theorem, we see that the general statement of Theorem 1.14 follows from the corresponding statement for 1-dimensional subtori. Since 1-dimensional subtori of $\mathbb{T}^m$ are in 1-2 correspondence to primitive integral vectors (up to sign), the following is a restatement of Theorem 1.14.

Theorem 1.16. *Assume $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ is not k-divergent for any $0 \leqslant k \leqslant d-2$ (in other words, the lattice x_A has an accumulation sequence of length d), and $\gcd(n,m) = 1$. Then, for any $p_0 \in \mathbb{Z}^m$ and $\eta \in \mathbb{R}^m$, we have for Lebesgue almost every $t \in \mathbb{R}$*

$$\liminf_{q\in\mathbb{Z}^n, \|q\|\to\infty} \|q\|^{n/m} \langle Aq - (tp_0 + \eta)\rangle = 0.$$

Theorems 1.14 and 1.16 give the answer to Question 1.3 for the case when μ is algebraic. Our techniques give the following stronger result, which deals with the 1-dimensional Lebesgue measure sitting on an immersed line in the direction of $Aq_0 + p_0$ for any given $q_0 \in \mathbb{Z}^n, p_0 \in \mathbb{Z}^m$. These immersed lines are closed if $q_0 = 0$.

Theorem 1.17. *Assume $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ is not k-divergent for any $0 \leqslant k \leqslant d-2$ (in other words, the lattice x_A has an accumulation sequence of length d), and that $\gcd(n,m) = 1$. Then, for any $p_0 \in \mathbb{Z}^m, q_0 \in \mathbb{Z}^n$, and $\eta \in \mathbb{R}^m$, we have for Lebesgue almost every $t \in \mathbb{R}$*

$$\liminf_{q\in\mathbb{Z}^n, \|q\|\to\infty} \|q\|^{n/m} \langle Aq - (t(Aq_0 + p_0) + \eta)\rangle = 0.$$

Theorem 1.17 reduces to Theorem 1.16 if we choose $q_0 = 0$. Theorem 1.17 follows from the more general Theorem 2.11 below as we show after its formulation in Section 2.6.

Remark 1.18. Note that when $m = 1$ and $n \geqslant 2$, Theorems 1.16, 1.14 and 1.17 are weaker than Theorem 1.6 because there is only one nontrivial algebraic measure on $\mathbb{T}^1 = \mathbb{T}^m$. So in this case, there is no need to assume x_A is not k-divergent for $0 \leqslant k \leqslant d-2$. It is enough to assume it is not k-divergent for $k = 0$.

Remark 1.19. It is nontrivial to construct A, which is k-divergent. Using the emerging theory of *parametric geometry of numbers*, k-divergent lattices are constructed in [14] and in fact, the Hausdorff dimension of the set of k-divergent lattices is calculated there. Curiously enough, the vector $\theta \in \mathbb{R}^m$ constructed in Theorem 1.7 must be k-divergent for some $k = 1, \dots, d-2$.

Remark 1.20. This remark pertains to a possible variant on how to define Bad_A. In Equation (1.2), we took a liminf as $q \in \mathbb{Z}^n$, $\|q\| \to \infty$. It is sometimes desirable to restrict attention to coefficient vectors q, which belong to some subset. The most natural situation in which this arises is when $n = 1$, that is, $q \in \mathbb{Z}$ but we are interested in $q \to \infty$ rather than $|q| \to \infty$. In many discussions in Diophantine approximations, it is trivial to connect two such questions but in our inhomogeneous setting, it **does not** seem to be true that for a given vector $\theta \in \mathbb{R}^m$ and a target $\eta \in \mathbb{R}^m$

$$\liminf_{q \to \infty} q^{1/m} \langle q\theta - \eta \rangle = 0 \Longleftrightarrow \liminf_{|q| \to \infty} |q|^{1/m} \langle q\theta - \eta \rangle = 0.$$

If one defines a variant Bad_A^+ of Bad_A using the restriction that $\|q\| \to \infty$ and all the coordinates of q are positive, then we expect that the techniques of this paper should be strong enough to prove versions of Theorems 1.6 and 1.14. Such statements are stronger because $\mathrm{Bad}_A \subset \mathrm{Bad}_A^+$. Here is an outline of how one might do this: One defines a version of the value set $V_F(y)$ (see Definition 2.4) that restricts attention to the values F takes on the grid points in the cone in $\mathbb{R}^d$ corresponding to the last n-coordinates being positive. One then proves a version of Lemma 2.9, where instead of the nondegeneracy degree defined in Definition 2.8, one considers only grid points in the cone. The rest of the argument then follows the same path as in this paper. We note that there are results in the literature regarding homogeneous Diophantine approximations with sign constraints on the coefficients. See [17, 19, 24, 27, 28].

Remark 1.21. As we saw in Theorem 1.6, if A is nonsingular, then $\lambda(\mathrm{Bad}_A) = 0$. A natural question that comes to mind is if the opposite statement is true, namely, is it true that if A is singular, then $\lambda(\mathrm{Bad}_A) > 0$. As we shall now explain, this is not the case. In fact, we can give an explicit Diophantine condition on A, which gives rise to a class of matrices strictly containing the nonsingular matrices (and in particular, this class contains some singular matrices), such that for A in this class, $\lambda(\mathrm{Bad}_A) = 0$. In particular, this shows that there are singular matrices A with $\lambda(\mathrm{Bad}_A) = 0$. The definition of the Diophantine class of matrices is stated in terms of the sequence of best approximations and is tightly related and motivated by the discussions in the papers [18] and [13]. We note that in [13], another complementary Diophantine class is defined, which is a subclass of the singular matrices, and it is proved there that for such matrices A, one has $\lambda(\mathrm{Bad}_A) = 1$.

In order to define our new Diophantine condition properly, one needs a bit of notation and terminology and hence we postpone the exact formulation of the result to Section 8. See Definition 8.1 and Theorem 8.4.

1.3 | A few open problems

Nontrivial coset intersection

In Theorem 1.7, the codimension 2 coset is completely contained in Bad_A. Example 2.12 also shows a similar phenomenon. The most basic question one might state here is the following: Are there examples of nonsingular matrices A and 1-dimensional algebraic measures μ on $\mathbb{T}^m$ such that $\mu(\mathrm{Bad}_A) > 0$ and the support of μ not fully contained in Bad_A.

Codimension 1 algebraic measures

Can one construct a nonsingular matrix A such that $\mu(\mathrm{Bad}_A) > 0$ for an algebraic measure μ of dimension $d - 1$?

The coprimality condition

It is not clear to us at the moment if the condition $\gcd(m, n) = 1$ in Theorem 1.14 can be relaxed. Example 2.12 suggests that the case $n = m$ is false.

k-divergence

Our definition of k-divergence for a matrix A is dynamical but we expect that similarly to singularity, there should be a purely Diophantine characterization of this property, although we were not able to pin down such a condition. Even in the case $n = m = 1$, where A is a number, and 0-divergence is characterized by rationality, we do not have a clean characterization of the notion of k-divergence (for general k) using continued fractions.

A question that emerges from Theorem 1.7 is the following: In Theorem 1.7, we construct a nonsingular vector that violates the conclusion of Theorem 1.14 and hence it must be k-divergent for some $1 \leqslant k \leqslant d - 2$. For which values of k can this be achieved? A possible reformulation of this question might be: For a given $1 \leqslant k \leqslant d - 2$, is it possible to construct k-divergent matrices A and algebraic measures μ on $\mathbb{T}^m$ such that $\mu(\mathrm{Bad}_A) > 0$?

Beyond algebraic measures

In the context of Theorems 1.6 and 1.14, it is natural to ask what happens for other natural classes of measures. Given a natural class $\mathcal{A}$ of measures on $\mathbb{T}^m$, can one pin down exact Diophantine conditions on A, which ensure that for all $\mu \in \mathcal{A}$, $\mu(\mathrm{Bad}_A) = 0$? Examples of $\mathcal{A}$ can be, smooth measures on submanifolds, Hausdorff measures on certain fractals, or measures whose Fourier coefficients satisfy certain conditions.

Zero-one law

Is it always the case that $\mu(\mathrm{Bad}_A) \in \{0, 1\}$ for algebraic measures on $\mathbb{T}^m$. For what classes of measures does such a zero-one law hold? The following simple observation gives an affirmative answer in the case that μ is an algebraic measure supported on a subgroup (rather than a coset):

Proposition 1.22. *Let μ be an algebraic measure supported on a subgroup of $\mathbb{T}^m$. Then, for any $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, we have $\mu(\mathrm{Bad}_A) \in \{0, 1\}$.*

Proof. Let $T : \mathbb{T}^m \to \mathbb{T}^m$ denote the homomorphism $\eta \mapsto T(\eta) := 2\eta$. Then, as μ is the Haar measure on a subgroup, μ is T-invariant and ergodic (see [9, Corollary 2.20]). We will therefore finish the proof once we show $T^{-1}(\mathrm{Bad}_A) \subset \mathrm{Bad}_A$. This is equivalent to showing

$$\eta \notin \mathrm{Bad}_A \Rightarrow 2\eta \notin \mathrm{Bad}_A .$$

Indeed, if $\eta \notin \mathrm{Bad}_A$, then by definition

$$\liminf_{q\in\mathbb{Z}^n, \|q\|\to\infty} \|q\|^{n/m}\langle Aq - 2\eta\rangle = 0$$

and so necessarily

$$\begin{aligned}\liminf_{q\in\mathbb{Z}^n, \|q\|\to\infty} \|q\|^{n/m}\langle Aq - 2\eta\rangle &\leqslant \liminf_{q\in\mathbb{Z}^n, \|q\|\to\infty} \|2q\|^{n/m}\langle A2q - 2\eta\rangle \\ &\leqslant 2^{n/m+1} \liminf_{q\in\mathbb{Z}^n, \|q\|\to\infty} \|q\|^{n/m}\langle Aq - \eta\rangle = 0\end{aligned}$$

and so $2\eta \notin \mathrm{Bad}_A$. □

2 | FROM INHOMOGENEOUS APPROXIMATIONS TO VALUE SETS OF GRIDS

In this section, we introduce most of the notation and terminology used in the paper, state the main Theorem 2.11, and prove some basic results like the Inheritance Lemma 2.6.

2.1 | Euclidean space

Throughout the paper, we fix integers $m, n \geqslant 1$ and set

$$d := m + n.$$

Vectors in $\mathbb{R}^d$ will be denoted u, v, w, and so forth. The decomposition of a vector in $\mathbb{R}^d$ to its first m coordinates and last n coordinates will be important from time to time. Hence, when we write a vector as $\binom{v}{w}$, the reader should interpret $v \in \mathbb{R}^m, w \in \mathbb{R}^n$. Given two subsets $A, B \subset \mathbb{R}^d$, we denote $A + B = \{v + w : v \in A, w \in B\}$.

2.2 | The space of lattices

We let X denote the space of unimodular lattices in $\mathbb{R}^d$. Although it can be confusing, we denote lattices by small letters x, x_1, x_2, and so forth. It will be important to us to think of X as a topological space, and thus be able to discuss, for example, a converging sequence $(x_l)_{l\in\mathbb{N}}$ in X. At the same time, we wish to keep in mind that a lattice $x \in X$ is also viewed simply as a subset of $\mathbb{R}^d$.

A simple way to understand the topology on X is to note that the group $G := \mathrm{SL}_d(\mathbb{R})$ acts transitively on X, where the action $(g, x) \mapsto gx$ is induced by the linear action of G on $\mathbb{R}^d$. That is, given a lattice $x \in X$ and $g \in G$, the lattice gx is simply the linear image of x under the linear map induced by multiplication by g on column vectors.

It is easy to see that this action is transitive and that the stabilizer group of the standard lattice $\mathbb{Z}^d$ is $\Gamma := \mathrm{SL}_d(\mathbb{Z})$. We therefore obtain bijection $g\Gamma \leftrightarrow g\mathbb{Z}^d$ between the formal coset space G/Γ and the space X. An immediate advantage of this is that the quotient G/Γ is naturally equipped with a topology—the quotient topology with respect to the natural projection map $G \to G/\Gamma$. In fact, since $\Gamma < G$ is discrete, this map is a covering map. In particular, we have the following simple result, which allows us to explicitly understand convergence in X and will be used without reference throughout.

Lemma 2.1. *A sequence $(x_l)_{l\in\mathbb{N}}$ in X converges to a lattice $x \in X$ if and only if there exists a sequence $\varepsilon_l \in G$, $\varepsilon_l \to e$ (e being the identity element), such that for all l, $\varepsilon_l x_l = x$.*

2.3 | The space of grids

Most of our discussion will take place in a topological space $X \subset Y$. Again, we first introduce Y as an abstract set and then explain how to equip it with a topology.

Definition 2.2. A unimodular grid in $\mathbb{R}^d$ is a subset of the form

$$x + v := \{u + v : u \in x\},$$

where $x \in X$ and $v \in \mathbb{R}^d$.

The space of all unimodular grids will be denoted by Y. Note that in the representation of a grid $y \in Y$ as $y = x + v$, the lattice x is uniquely determined (it is obtained by translating y so that it would contain 0) but the translation vector v is only well-defined modulo x. Note also that $X \subset Y$—any lattice is a grid and a grid y is a lattice if and only if $0 \in y$.

In order to induce a topology on Y, we follow the same line of thought as before. Let

$$G' := \mathrm{ASL}_d(\mathbb{R}) := \left\{ \begin{pmatrix} g & v \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}_{d+1}(\mathbb{R}) : g \in \mathrm{SL}_d(\mathbb{R}), v \in \mathbb{R}^d \right\}$$

and consider the action of G' on $\mathbb{R}^d$ given by

$$\begin{pmatrix} g & v \\ 0 & 1 \end{pmatrix} \cdot u = gu + v.$$

Thus, the matrix $\left(\begin{smallmatrix} g & v \\ 0 & 1 \end{smallmatrix}\right) \in G'$ acts on $\mathbb{R}^d$ as an affine map obtained by applying the linear map g followed by the translation by v.

The action of G' on $\mathbb{R}^d$ induces an action on subsets of it and as a consequence G' acts on Y. It is easy to see that this action is transitive and that the stabilizer group of the standard grid $\mathbb{Z}^d$ is $\Gamma' = \mathrm{ASL}_d(\mathbb{Z})$ (i.e., the matrices in G' whose entries are integral). We therefore identify Y with G'/Γ' via $\left(\begin{smallmatrix} g & v \\ 0 & 1 \end{smallmatrix}\right)\Gamma' \leftrightarrow g\mathbb{Z}^d + v$ and equip Y with the quotient topology with respect to the natural quotient map $G' \to G'/\Gamma'$. The following lemma is left to be verified by the reader.

Lemma 2.3. *Let $y, y_l \in Y, l \in \mathbb{N}$ be grids and assume $y_l \to_{l\to\infty} y$. Then, for any vector $v \in y$, there exist vectors $u_l \in y_l$ such that $u_l \to v$.*

We note that although the transitive action of G' on Y is used to define the topology, this action will not play any other role in this paper. But, the group G acts on subsets of $\mathbb{R}^d$ and thus acts on Y in a way that extends its action on X. The G-action on Y will be paramount to our discussion.

2.4 | The projection from grids to lattices and its fibers

Let $\pi : Y \to X$ denote the map that sends a grid y to the unique lattice x such that $y = x + v$ for some $v \in \mathbb{R}^d$. We leave it to the enthusiastic reader to check that π is a continuous proper map, where proper means that preimages of compact sets are compact.

Given a lattice $x \in X$, the fiber $\pi^{-1}(x) = \left\{x + v : v \in \mathbb{R}^d\right\}$ is nothing but the collection of all cosets of the lattice x in $\mathbb{R}^d$. As such it is **equal** to the torus $\mathbb{R}^d/x$. The space Y can be thus thought of as the union of all possible tori (of volume 1) of $\mathbb{R}^d$. We alternate between writing $\pi^{-1}(x)$ and $\mathbb{R}^d/x$ throughout but wish to stress here that we can (and will) alternate between thinking of a tori $\mathbb{R}^d/x$ as

- the compact abelian group $\mathbb{R}^d/x$,
- a closed subset of Y obtained as a fiber $\pi^{-1}(x)$,
- a collection of grids (subsets of $\mathbb{R}^d$) $\left\{x + v : v \in \mathbb{R}^d\right\}$.

Finally note that π intertwines the G-actions on X, Y. That is, the following diagram commutes: For any $g \in G$,

$$\begin{tikzcd} Y \arrow[r, "g"] \arrow[d, "\pi"'] & Y \arrow[d, "\pi"] \\ X \arrow[r, "g"] & X \end{tikzcd}$$

This captures the fact that if we are given an element $g \in G$ and two lattices $x_1, x_2 \in X$ such that $gx_1 = x_2$, then when we act with g on Y, g maps the fiber $\pi^{-1}(x_1)$ onto the fiber $\pi^{-1}(x_2)$. Indeed the map $x_1 + v \mapsto g(x_1 + v)$ is an isomorphism between the compact abelian groups $\mathbb{R}^d/x_1, \mathbb{R}^d/x_2$.

2.5 | Subtori, Haar measures, and algebraic measures

As noted above, for a lattice $x \in X$, the fiber $\pi^{-1}(x) = \mathbb{R}^d/x$ is a d-dimensional compact abelian group we refer to as a torus. We now discuss its closed connected subgroups, their Haar probability measures, and their translates.

Let $x \in X$ be a lattice. A linear subspace $U < \mathbb{R}^d$ is called *x-rational* if $U \cap x$ contains a basis of U. It is well known that there is a 1-1 correspondence between x-rational subspaces and closed connected subgroups of $\mathbb{R}^d/x$. Let $U < \mathbb{R}^d$ be a x-rational subspace. For any grid $y \in \mathbb{R}^d/x$, the set $y + U \subset \mathbb{R}^d/x$ is a coset of the closed subgroup $x + U < \mathbb{R}^d/x$ and it supports a unique U-invariant probability measure. We refer to such measures as *algebraic measures* on $\mathbb{R}^d/x$. When the lattice x (the trivial element of the torus) belongs to the support of the measure, we refer to the measure as a *Haar measure* on a subtorus.

2.6 | Value sets and inheritance

Let $F : \mathbb{R}^d = \mathbb{R}^m \oplus \mathbb{R}^n \to \mathbb{R}$ be the map

$$F\left(\begin{pmatrix} v \\ w \end{pmatrix}\right) = \|v\|^m \cdot \|w\|^n.$$

Recalling the one-parameter subgroup $\{h_t : t \in \mathbb{R}\} < G$ defined in (1.6), we note that F is h_t-invariant. That is, for any $t \in \mathbb{R}$ and any $u \in \mathbb{R}^d$, $F(h_t u) = F(u)$. Geometrically, when h_t acts linearly on $\mathbb{R}^d$, it acts on the level sets of F. Note that although h_t does not act transitively on the level sets, it does act cocompactly on each level set of the form $F^{-1}(s)$ for $s > 0$. This will not be used explicitly in the paper but has conceptual importance toward the claim that understanding the values F takes on a subset of $\mathbb{R}^d$ can be attacked by analyzing how this set changes under the action of h_t. This is the idea behind our results.

Definition 2.4. Given a grid $y \in Y$, we define the *value set* of y to be

$$V_F(y) = \{F(u) : u \in y\}.$$

The fundamental questions in *geometry of numbers*, which guide us are as follows: What can be said about the value sets $V_F(y)$? Is it dense or discrete? Is 0 an accumulation point or not? Does its closure contain a ray? For the sake of the discussion in this paper, we make the following definitions.

Definition 2.5. A grid $y \in Y$ is a *dense values grid* or y is DV_F if

$$\overline{V_F(y)} = F(\mathbb{R}^d) = [0, \infty).$$

The following lemma is a fundamental tool in our discussion. It serves as the entry point of dynamics to our discussion.

Lemma 2.6 (Inheritance lemma). *Let $y_1, y_2 \in Y$ be grids and assume the orbit-closure $\overline{\{h_t y_1 : t \in \mathbb{R}\}}$ in Y contains y_2. Then, $\overline{V_F(y_2)} \subset \overline{V_F(y_1)}$.*

Proof. Let $v \in y_2$. We show that $F(v)$ can be approximated by elements from $V_F(y_1)$. Let $t_l, l \in \mathbb{N}$ be a sequence of real numbers such that $h_{t_l} y_1 \to y_2$. By Lemma 2.3, there are vectors $u_l \in h_{t_l} y_1$ such that $u_l \to v$. The vector $w_l := h_{t_l}^{-1} u_l$ belongs to the grid y_1 and by the h_t-invariance of F that we have together with the continuity of F, we have

$$F(w_l) = F(h_{t_l} w_l) = F(u_l) \to_{l\to\infty} F(v).$$

This shows $F(v) \in \overline{V_F(y_1)}$ and finishes the proof. $\square$

The way in which the Inheritance Lemma 2.6 will enter our proofs and help us establish that certain grids are DV_F is the following.

Proposition 2.7. *Let $y \in Y$ be a grid. If there exists a lattice $x \in X$ such that the orbit-closure $\overline{\{h_t y : t \in \mathbb{R}\}}$ in Y contains a full coset of a $d-1$-dimensional subtorus in the fiber $\pi^{-1}(x) = \mathbb{R}^d/x$, then y is DV_F.*

For the proof, we will need the following definition and lemma.

Definition 2.8. The nondegeneracy degree of F is the minimal dimension ℓ such that for any grid y and any ℓ-dimensional subspace $U < \mathbb{R}^d$, one has $F(y+U) = [0,\infty)$.

Lemma 2.9. *The nondegeneracy degree of F is $d-1$.*

Proof. Let

$$y = \mathbb{Z}^d + 1/2 \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Let $U < \mathbb{R}^d$ be the codimension 2 subspace given by $U = \{\left(\begin{smallmatrix} v \\ w \end{smallmatrix}\right) : v_1 = 0, w_n = 0\}$. Note that the first and last coordinates of the vectors in $y+U$ are at least $1/2$ in absolute value. Therefore, their F-value is $\geqslant 2^{-n} \cdot 2^{-m} = 2^{-d}$. This shows that the nondegeneracy degree is at least $d-1$.

On the other hand, let $U < \mathbb{R}^d$ be a $d-1$-dimensional space, let y be a grid. We will show $F(y+U) = [0,\infty)$. Let $\left(\begin{smallmatrix} v^1 \\ w^1 \end{smallmatrix}\right), \dots, \left(\begin{smallmatrix} v^{d-1} \\ w^{d-1} \end{smallmatrix}\right)$ be a basis for U. Let

$$r_1 = \operatorname{rank}\left(v^1 \cdots v^{d-1} \right),\ r_2 = \operatorname{rank}\left(w^1 \cdots w^{d-1} \right).$$

Since the rank of the matrix whose columns are the basis of U is $d-1 = m+n-1$, it is impossible to have both $r_1 < m$ and $r_2 < n$. Assume for concreteness that $r_1 = m$. Choose a vector $\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \in y$ such that $F(\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right)) > 0$ (this is always possible because grids always contain points in $\mathbb{R}^d_{>0}$ for example). We can solve the system of equations

$$-p = a_1 v^1 + \cdots + a_{d-1} v^{d-1}$$

because we assume $r_1 = m$. Therefore, the affine subspace $\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) + U$ contains a vector of the form $\left(\begin{smallmatrix} 0 \\ w \end{smallmatrix}\right)$.

Assume for the moment that the norms defining F are the Euclidean norms. The restricted function $F|_{\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right)+U}$ is the square root of a polynomial in $d-1$ variables that attains the value 0 because of the above argument and also attains the value $F((\frac{p}{q})) > 0$. Hence it is a nonconstant polynomial and so we conclude that $F((\frac{p}{q}) + U) = [0, \infty)$, which finishes the proof.

We leave it to the reader to deduce the general case, where the norms used to define F are not assumed to be the Euclidean norms, from the Euclidean one (use the equivalence of norms and the mean-value theorem). □

Proof of Proposition 2.7. A subtorus of dimension $d-1$ in $\pi^{-1}(x) = \mathbb{R}^d/x$ is a subgroup of the form $x + U$ where $U < \mathbb{R}^d$ is a $d-1$-dimensional subspace whose image in $\mathbb{R}^d/x$ is closed. A coset of such a subtorus is a set of the form $x + v + U$ for such a subspace U and a vector $v \in \mathbb{R}^d$. If the closure $\overline{\{h_t y \,:\, t \in \mathbb{R}\}}$ contains all the grids in such a set, then by the Inheritance Lemma 2.6 we have

$$\overline{V_F(y)} \supset \bigcup_{u \in U} V_F(x + v + u) = [0, \infty),$$

where the last equality follows from Lemma 2.9. □

In our discussion, we will fix a lattice x and a probability measure μ on the torus $\mathbb{R}^d/x$ and try to say something about $\mu\left(\left\{y \in \pi^{-1}(x) \,:\, y \text{ is } DV_F\right\}\right)$.

Definition 2.10. Given $x \in X$ and a probability measure μ on $\mathbb{R}^d/x$, we say that x is μ-almost surely DV_F if μ-almost any y is DV_F.

The following is one of the main results of the paper. The reader should recall Definition 1.12.

Theorem 2.11. *Assume* $\gcd(m, n) = 1$. *If $x \in X$ has an accumulation sequence of length d, in other words x is not k-divergent for any* $0 \leqslant k \leqslant d-2$, *then x is μ-almost surely grid DV_F with respect to any algebraic measure μ on* $\pi^{-1}(x)$. *In fact, μ-almost any* $y \in \pi^{-1}(x)$ *satisfies that the orbit closure* $\overline{\{h_t y \,:\, t \geqslant 0\}} \subset Y$ *contains a coset of a* $d-1$*-dimensional subtorus of* $\mathbb{R}^d/x_{d-1}$.

We now deduce Theorem 1.17 from Theorem 2.11 and explain how to link the discussion about DV_F grids to the discussion about the set of badly approximable targets Bad_A.

Proof of Theorem 1.17 assuming Theorem 2.11. Let A, p_0, q_0, η be as in the statement of Theorem 1.17. Consider the lattice x_A defined in (1.5) and the 1-dimensional algebraic measure μ on $\mathbb{R}^d/x_A$ supported on the 1-dimensional subtorus given by the x_A-rational subspace

$$U := \left\{ \begin{pmatrix} t(Aq_0 + p_0) \\ tq_0 \end{pmatrix} \,:\, t \in \mathbb{R} \right\}$$

and translation $\begin{pmatrix} -\eta \\ 0 \end{pmatrix}$. Theorem 2.11 says that μ-almost any grid is DV_F. In other words, for Lebesgue almost any $t \in \mathbb{R}$, $F\left(x_A - \begin{pmatrix} t(Aq_0 + p_0) + \eta \\ tq_0 \end{pmatrix}\right)$ is dense in $[0, \infty)$. In particular, for

Lebesgue almost any t, F attains arbitrarily small positive values on the grid

$$x_A - \begin{pmatrix} t(Aq_0 + p_0) + \eta \\ tq_0 \end{pmatrix} = \left\{ \begin{pmatrix} p + Aq - t(Aq_0 + p_0) - \eta \\ q - tq_0 \end{pmatrix} : q \in \mathbb{Z}^n, p \in \mathbb{Z}^m \right\}.$$

For a vector $\begin{pmatrix} p + Aq - t(Aq_0 + p_0) - \eta \\ q - tq_0 \end{pmatrix}$ in the above set, we have

$$F\left(\begin{pmatrix} p + Aq - t(Aq_0 + p_0) - \eta \\ q - tq_0 \end{pmatrix}\right) = \|q - tq_0\|^n \|p + Aq - t(Aq_0 + p_0) - \eta\|^m.$$

This quantity cannot become positive and arbitrarily small using finitely many q's and so we get that there exists a sequence $q_i \in \mathbb{Z}^n$ and $p_i \in \mathbb{Z}^m$ such that $\|q_i\| \to \infty$ and $\|q_i\|^n \|p_i + Aq_i - t(Aq_0 + p_0) - \eta\|^m$ is a sequence of positive numbers going to zero. In particular,

$$\liminf_{q \in \mathbb{Z}^n, \|q\| \to \infty} \|q\|^{n/m} \langle Aq - (-t(Aq_0 + p_0) + \eta) \rangle = 0$$

as desired. □

2.7 | An intriguing example

We end this section with the following example, which shows that the assumption $\gcd(n, m) = 1$ is not entirely an artifact of the method of proof (see Lemma 6.2 and Corollary 6.3, for the point where this assumption enters our discussion). It shows that the statement of Theorem 2.11 can fail drastically when $n = m \geqslant 2$.

Example 2.12. Take any 2-lattice in the plane $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mathbb{Z}^2$ and find a translate $\left(\begin{smallmatrix} s_1 \\ s_2 \end{smallmatrix}\right)$ satisfying

$$\inf_{q_1, q_2 \in \mathbb{Z}} |aq_1 + bq_2 + s_1| \cdot |cq_1 + dq_2 + s_2| := \varepsilon > 0. \tag{2.1}$$

The existence of such s_1, s_2 is tightly related to Theorems 1.1 and 1.2. A proof for this existence can be found in [6, Theorem 1].

Now choose $n = m$ for any positive integer m and take a matrix $g \in \mathrm{SL}_d(\mathbb{R})$ whose 1st and $m + 1$th rows are

$$(a, 0, \dots, 0, b, 0, \dots, 0)$$

$$(c, 0, \dots, 0, d, 0, \dots, 0)$$

where the b and the d are the $m + 1$-coordinates.

Let us denote by $\mathbf{p} : \mathbb{R}^d \to \mathbb{R}^2$ the projection on the 1st and $m + 1$th coordinates. For such a choice of g, for any grid of the form $y = g\mathbb{Z}^d + u$, where

$$u = (s_1, *, \dots, *, s_2, *, \dots, *)^{\mathrm{tr}}, \tag{2.2}$$

we have

$$\mathbf{p}(y) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\mathbb{Z}^2 + \begin{pmatrix} s_1 \\ s_2 \end{pmatrix}.$$

Since we have the lower bound, for $\binom{v}{w} \in \mathbb{R}^d$,

$$F(\tbinom{v}{w}) = \|v\|^m \|w\|^m \geqslant |v_1|^m |w_1|^m,$$

we deduce from (2.1) that

$$\inf V_F(y) \geqslant \varepsilon^m > 0$$

and in particular y is not DV_F. Note that this is where it is important that $n = m$. By the general shape (2.2) of the translate u giving the grid y, this shows that the set of grids of $g\mathbb{Z}^d$, which are not DV_F, contains a coset of a codimension 2 subtorus.

We note that it is not hard to make the choices in the above general construction in a way that the h_t-orbit of $g\mathbb{Z}^d$ is **bounded** and in particular, this lattice is not k-divergent for any k. This shows that Theorem 2.11 fails drastically when $n = m$. It also serves as a counterexample to [26, Theorem 2.3] for the regime $n = m \geqslant 2$. The counterexample we present in Theorem 1.7 is more sophisticated and deals with the regime $n = 1$ and $m > 2$.

A possible choice of the parameters is given by choosing $n = m = 2$, $d = 4$,

$$g = \alpha \begin{pmatrix} 1 & 0 & \sqrt{2} & 0 \\ 0 & 1 & 0 & \sqrt{2} \\ 1 & 0 & -\sqrt{2} & 0 \\ 0 & 1 & 0 & -\sqrt{2} \end{pmatrix},$$

where α is chosen so that $\det g = 1$. We shall ignore the multiplicative factor α below. To see that the orbit $h_t g\mathbb{Z}^4$ is bounded, we use Mahler's criterion, which says that a set of lattices in X has compact closure if and only if there is a uniform lower bound on the lengths of nonzero vectors in lattices in the set. Here, the general form of a nonzero vector in a lattice in $\{h_t g\mathbb{Z}^4 : t \in \mathbb{R}\}$ is of the form

$$\begin{pmatrix} e^t(q_1 + q_3\sqrt{2}) \\ e^t(q_2 + q_4\sqrt{2}) \\ e^{-t}(q_1 - q_3\sqrt{2}) \\ e^{-t}(q_2 - q_4\sqrt{2}) \end{pmatrix},$$

where $(q_1, q_2, q_3, q_4) \neq (0, 0, 0, 0)$. Assuming for concreteness that $(q_1, q_3) \neq (0, 0)$, we get that the length of the above 4-dimensional vector is bounded below by the length of the 2-dimensional vector

$$\begin{pmatrix} e^t(q_1 + q_3\sqrt{2}) \\ e^{-t}(q_1 - q_3\sqrt{2}) \end{pmatrix}.$$

Now this is a vector whose product of coordinates equals $q_1^2 - 2q_3^2$, which is a nonzero integer and so it is impossible that both of the coordinates are smaller than 1 in absolute value. This shows that the length of the 4-dimensional vector above is at least 1 as well.

2.8 | The mistake in [26]

The main result in [26] is Theorem 2.2 there. This theorem is solid and has interesting applications beyond the discussion of this paper, namely, for functions other than F. The mistake in that paper is in sections 2.2 and 3.2 where the main theorem was applied to the function F (denoted there by $P_{n,m}$) under the false statement that F has nondegeneracy degree 1 according to Definition 2.8 of this paper. We now comment on the validity/falsity of each of the statements made in [26, sections 2.2, 3.2].

- [26, Theorem 2.3] is not true in the generality stated. It is true for the Haar measure on $\mathbb{R}^d/x$ as we prove in Corollary 4.3 in this paper. On the other hand, we saw in Example 2.12 and in Theorem 1.7 that [26, Theorem 2.3] is not true for some choices of n, m and algebraic measures of codimension 2. We do not know at this point if the statement holds for codimension 1 algebraic measures. In particular, the simplest case that remains open is the following:

  Question 2.13. In the notation of this paper, let $d = 3, m = 2, n = 1$. Let $x \in X$ be a lattice with a nondivergent h_t-orbit. Is it true that for any nontrivial algebraic measure μ on $\mathbb{R}^3/x$, we have that x is μ-almost surely DV_F? Is it true that for μ-almost any grid y, $0 \in \overline{V_F(y)}$? Specializing to vectors, given a nonsingular column vector $\theta \in \mathbb{R}^2$, is it true that Bad_θ is a null set with respect to any algebraic measure on $\mathbb{T}^2$?

- [26, Theorem 3.4] is true as stated. This follows from Corollary 4.3 and Theorem 1.6 of this paper.
- [26, Theorem 3.7] is false in the generality stated. Theorem 1.7 shows that it is false for some algebraic measures.
- [26, Corollary 3.8] is under question. We do not yet have a counterexample. The statement does hold if one replaces the nonsingularity assumption by the stronger assumption that the vector is not k-divergent for $0 \leqslant k \leqslant d-1$. This is a special case of Theorem 1.17.

3 | GENERALITIES ABOUT MEASURES

3.1 | Borel measures and weak* convergence

In this section, we present the elementary theory we will use regarding regular Borel measures on topological spaces. Although the discussion is widely known, we include it here for the sake of completeness, which some readers might appreciate.

Throughout this section, we assume all topological spaces to be locally compact second countable Hausdorff spaces. This puts us in a setting in which every Borel measure, which gives finite measure to compact sets, is automatically regular. Such measures are called Radon measures. Furthermore, we will only discuss finite measures so the condition one finiteness on compacta will be automatic.

In later sections, when we apply the theory presented here, we will only be concerned with the spaces Y, X and their subspaces. Nevertheless, we try to keep the discussion here abstract for clarity.

Definition 3.1 (Topology on regular complex Borel measures). Let Z be locally compact second countable Hausdorff space. Consider the Banach space $C_0(Z, \mathbb{C})$ of complex continuous functions vanishing at infinity with the supremum norm. Its dual, by [21, Theorem 6.19], is isometric to $\mathcal{M}(Z)$, the vector space of complex Radon measures on Z endowed with the norm induced by total variation (cf. [21, Chapter 6]). The bilinear duality pairing is given by integration:

$$(\phi, \mu) \mapsto \int_Z \phi d\mu.$$

We endow $\mathcal{M}(Z)$ with the weak* topology. That is, the smallest topology making all the functionals

$$\mu \mapsto \int_Z \phi d\mu \,(\text{where } \phi \in C_0(Z, \mathbb{C}))$$

continuous. We denote by $\mathcal{M}_1(Z)$ the set of measures with norm $\leqslant 1$, the set of real positive measures by $\mathcal{M}_+(Z)$, and the set of real positive measures of norm 1 by $\mathcal{P}(Z)$. This last set is the set of Borel probabilities on Z. We have a partial order on $\mathcal{M}(Z)$ given by

$$(\mu \geqslant \nu) \iff \big(\mu - \nu \in \mathcal{M}_+(Z)\big).$$

Given $\mu \in \mathcal{P}(Z)$, we define the *support* of μ, $\operatorname{supp} \mu$ to be the complement of the set

$$\bigcup \{V \subset Z \,:\, V \text{ is open and } \mu(V) = 0\}.$$

Lemma 3.2. *Let Z be a second countable locally compact Hausdorff space and let $\mu, \nu \in \mathcal{M}(Z)$. If $D \subset C_0(Z, \mathbb{C})$ spans a dense subspace of functions, then*

$$(\mu = \nu) \iff \left(\int_Z \phi d\mu = \int_Z \phi d\nu \text{ for all } \phi \in D\right).$$

Proof. The implication $\Longrightarrow$ is trivial. The other implication follows from the fact that μ, ν are continuous linear functionals on $C_0(Z, \mathbb{C})$. ☐

An example to keep in mind for the applicability of the above lemma is when $Z = \mathbb{T}^m$ and D is the set of characters $D = \left\{e^{2\pi i\langle x,p\rangle} \,:\, p \in \mathbb{Z}^m\right\}$.

Lemma 3.3. *Let Z be a locally compact second countable Hausdorff space. In the weak* topology, the unit ball $\mathcal{M}_1(Z)$ is compact. Furthermore, when Z is compact, $\mathcal{P}(Z)$ is compact.*

Proof. The first statement is the Banach–Alaoglu theorem. See [22, Theorem 3.15]. For the second statement, note that

$$\mathcal{M}_+(Z) = \cap_{\phi \geqslant 0, \phi \in C_0(Z,\mathbb{C})} \left\{\mu \,:\, \int_Z \phi d\mu \geqslant 0\right\}$$

is weak* closed and when Z is compact, $\mathbf{1}_Z$ (the constant function 1) is an element of $C_0(Z, \mathbb{C})$ and so the subset $\mathcal{P}(Z) = \left\{\mu \in \mathcal{M}_+(Z) : \int_Z \mathbf{1}_Z d\mu = 1\right\}$ is also closed. Therefore, $\mathcal{P}(Z) \subset \mathcal{M}_1(Z)$ is a closed subset and hence compact (note that the weak* topology is Hausdorff). □

Lemma 3.4. *Let Z be a locally compact second countable Hausdorff space. Then, the unit ball $\mathcal{M}_1(Z)$ with the weak* topology is metrizable.*

Proof. Under our assumptions, $C_0(Z, \mathbb{C})$ contains a countable dense set. Call it $(\phi_i)_{i\in\mathbb{N}} \subset C_0(Z, \mathbb{C})$. We leave it to the reader to check that the topology on $\mathcal{M}_1(Z)$ induced by the metric

$$\operatorname{dist}(\mu, \nu) := \sum_{i=1}^{\infty} \frac{1}{2^i \|\phi_i\|} \left| \int \phi_i d(\mu - \nu) \right|$$

coincides with topology on $\mathcal{M}_1(Z)$ induced by the weak* topology on $\mathcal{M}(Z)$. Alternatively, apply Lemmas 3.3 and 3.2 to use [22, (c) of p. 63]. □

Definition 3.5 (Pushforwards of measures). Let Z and Z' be locally compact second countable Hausdorff spaces. If $f : Z \to Z'$ is a Borel measurable map between them and $\mu \in \mathcal{M}(Z)$, we have the pushforward of μ by f in $\mathcal{M}(Z')$ defined by the formula:

$$(f_*\mu)(E) := \mu(f^{-1}(E)) \text{ for every Borel } E \subset Z'.$$

It is straightforward to check that for any bounded measurable function $\phi : Z' \to \mathbb{C}$, we have the equality

$$\int_{Z'} \phi d(f_*\mu) = \int_Z \phi \circ f d\mu,$$

and that this formula characterizes $f_*\mu$.

Lemma 3.6. *Let Z, Z', Z'' be locally compact second countable Hausdorff spaces. Let $\mu \in \mathcal{M}(Z)$, and say we have measurable maps $f : Z \to Z'$ and $g : Z' \to Z''$. Then,*

$$(g \circ f)_*(\mu) = g_*(f_*\mu) \textit{ in } \mathcal{M}(Z'').$$

Proof. This follows at once from Definition 3.5. □

Lemma 3.7. *Let Z, Z' be locally compact second countable Hausdorff spaces and let $f : Z \to Z'$ be continuous and proper. Then,*

$$f_* : \mathcal{M}(Z) \to \mathcal{M}(Z')$$

is continuous with respect to the weak topologies.*

Proof. It is clear that f_* is a linear map. It is thus enough to show that preimages of basic open neighborhoods of 0 are open. A basic weak* open neighborhood of $0 \in \mathcal{M}(Z')$ is given by, for $\epsilon > 0, \phi \in C_0(Z', \mathbb{C})$, $\left\{\nu : |\int \phi d\nu| < \epsilon\right\}$. It follows from the definition of f_* that for

$\phi \in C_0(Z, \mathbb{C}), \epsilon > 0$,

$$(f_*)^{-1}\left\{\nu \in \mathcal{M}(Z') : \left|\int \phi d\nu\right| < \epsilon\right\} = \left\{\mu \in \mathcal{M}(Z) : \left|\int \phi \circ f d\mu\right| < \epsilon\right\}.$$

Since f is continuous and proper, $\phi \circ f \in C_0(Z, \mathbb{C})$ and so the right-hand side is a basic open neighborhood of 0 as well. □

The following lemma is used to justify a step in a particular argument. The reader should skip it and return to it when referred to.

Lemma 3.8. *Let Z, Z', Z'' be locally compact second countable Hausdorff spaces. Let $(f_l)_{l\in\mathbb{N}}$ and $(g_l)_{l\in\mathbb{N}}$ be sequences of continuous proper functions from $Z \to Z'$ and $Z' \to Z''$, respectively. Let $\mu \in \mathcal{M}(Z)$, $\nu \in \mathcal{M}(Z')$, and $\eta \in \mathcal{M}(Z'')$ be measures and assume we have the weak* convergence*

$$\lim_{n\to\infty}(f_l)_*\mu = \nu \ \text{ and } \ \lim_{l\to\infty}(g_l)_*\nu = \eta.$$

Then, there are subsequences $(p_l)_{l\in\mathbb{N}}$ and $(q_l)_{l\in\mathbb{N}}$ of $\mathbb{N}$ such that we have the weak convergence*

$$\lim_{l\to\infty}(g_{p_l} \circ f_{q_l})_*\mu = \eta.$$

Proof. Since the maps on measures $(f_l)_*$ and $(g_l)_*$ have operator norm less than or equal to 1, we might as well assume μ, ν, and η are in the unit balls of their respective measure spaces.

By Lemma 3.4, the weak* topologies on the unit balls $\mathcal{M}_1(Z)$, $\mathcal{M}_1(Z')$, and $\mathcal{M}_1(Z'')$ are metrizable. Let $l \in \mathbb{N}$ and consider the metric ball

$$B_{\mathcal{M}_1(Z'')}(\eta, l^{-1}) \subset \mathcal{M}_1(Z'')$$

with center η and radius l^{-1} in this metric. By convergence of $((g_l)_*\nu)_{l\in\mathbb{N}}$, there exists some $p_l \in \mathbb{N}$ with

$$(g_{p_l})_*\nu \in B_{\mathcal{M}_1(Z'')}(\eta, l^{-1}).$$

By Lemma 3.7, there is some $\delta > 0$ with

$$\nu \in B_{\mathcal{M}_1(Z')}(\nu, \delta) \subset (g_{p_l})_*^{-1}\left(B_{\mathcal{M}_1(Z'')}(\eta, l^{-1})\right). \tag{3.1}$$

By the convergence of $((f_l)_*\mu)_{l\in\mathbb{N}}$, we see that there exists $q_l \in \mathbb{N}$ with

$$(f_{q_l})_*\mu \in B_{\mathcal{M}_1(Z')}(\nu, \delta). \tag{3.2}$$

Combining Equations (3.1) and (3.2), we see that

$$(g_{p_l})_* \circ (f_{q_l})_*\mu \in B_{\mathcal{M}_1(Z'')}(\eta, l^{-1}).$$

Varying $l \in \mathbb{N}$ gives us our sequence. □

The following lemma is used to justify a step in a particular argument. The reader should skip it and return to it when referred to.

Lemma 3.9. *Say Z and Z' are locally compact second countable metric spaces with distance functions* dist_Z *and* $\operatorname{dist}_{Z'}$*, respectively. Let $\mu \in \mathcal{M}(Z)$. Suppose we have sequences of continuous maps $(f_l)_{l\in\mathbb{N}}$, $(g_l)_{l\in\mathbb{N}}$ from $Z \to Z'$ with the property:*

$$\lim_{l\to\infty} \sup_{z\in Z} \operatorname{dist}_{Z'}(f_l(z), g_l(z)) = 0.$$

If there exists a weak limit*

$$\lim_{l\to\infty} (f_l)_*\mu = \nu \ \textit{ in } \mathcal{M}(Z'),$$

then we also have the same weak limit*

$$\lim_{l\to\infty} (g_l)_*\mu = \nu.$$

Proof. We can assume $\mu \neq 0$. Let $\phi \in C_0(Z', \mathbb{C})$ and let $\varepsilon > 0$. Since ϕ is uniformly continuous, there is a $\delta > 0$ such that

$$\left(\operatorname{dist}_{Z'}(z_1', z_2') < \delta\right) \Rightarrow \left(|\phi(z_1') - \phi(z_2')| < \frac{\varepsilon}{2}(\|\mu\|)^{-1}\right).$$

Choose $l_1 \in \mathbb{N}$ large enough so that, for all $l > l_1$,

$$\sup_{z\in Z'} d_{Z'}(f_l(z), g_l(z)) < \delta.$$

Choose $l_2 \in \mathbb{N}$ large enough so that, for all $l > l_2$,

$$\left| \int_{Z'} \phi d(f_l)_*\mu - \int_{Z'} \phi d\nu \right| < \frac{\varepsilon}{2}.$$

For $n > \max\{l_1, l_2\}$, we compute

$$\begin{aligned} \left| \int_{Z'} \phi d(g_l)_*\mu - \int_{Z'} \phi d\nu \right| &= \left| \int_{Z} \phi \circ g_l d\mu - \int_{Z'} \phi d\nu \right| \\ &= \left| \int_{Z} (\phi\circ g_l - \phi\circ f_l)) d\mu + \int_{Z'} \phi d(f_l)_*\mu - \int_{Z'} \phi d\nu \right| \\ &\leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2}. \end{aligned}$$

□

3.2 | Mixing convergence

A key concept that plays a role in the proof of Theorem 2.11 is the notion of mixing convergence of measures, which was introduced in [26, Definition 4.1]. We review this concept for completeness.

Definition 3.10 (Mixing convergence). Let Z and Z' be locally compact second countable Hausdorff spaces. Let $\mu \in \mathcal{P}(Z)$ and let $\nu \in \mathcal{P}(Z')$. Let $(f_l)_{l\in\mathbb{N}}$ be a sequence of measurable maps from $Z \to Z'$. We say $(f_l)_{l\in\mathbb{N}}$ mixes μ to ν if, for every absolutely continuous probability measure $\eta \ll \mu$ in $\mathcal{P}(Z)$, we have

$$\lim_{l\to\infty}(f_l)_*\eta = \nu. \tag{3.3}$$

Note that by taking $\eta = \mu$, we see that mixing convergence is stronger than weak* convergence. We will use the following reformulation of mixing convergence.

Proposition 3.11. *In the notation of Definition 3.10, $(f_l)_*$ mixes μ to ν if and only if*

$$\lim_{l\to\infty}\int_Z \phi(f_l(z))\psi(z)d\mu(z) = \left(\int_Z \psi d\mu\right)\left(\int_{Z'}\phi d\nu\right) \tag{3.4}$$

for all $\phi \in C_0(Z',\mathbb{C}), \psi \in L^1(Z,\mu)$.

Further, the condition of (3.4) holds if it does on subsets of $L^1(Z,\mu)$ and $C(Z',\mathbb{C})$ spanning dense subspaces.

Proof. Consider the bilinear map $\langle\cdot,\cdot\rangle_l : C_0(Z',\mathbb{C})\times L^1(Z,\mu)\to\mathbb{C}$ given by

$$\langle\phi,\psi\rangle_l = \int_Z \phi(f_l(z))\psi(z)d\mu(z).$$

By the Radon–Nikodym theorem, the condition $\eta \ll \mu, \eta \in \mathcal{P}(Z)$, in Definition 3.10 is equivalent to the fact that there exists a density $\psi \in L^1(Z,\mu)$ such that $\eta = \psi d\mu$ and $\psi \geqslant 0$, $\int_Z \psi d\mu = 1$. Interpreting the definition of weak* convergence, we see that Definition 3.10 is equivalent to requiring that Equation (3.4) holds for any $\phi \in C_0(Z',\mathbb{C})$ and any $\psi \in L^1(Z,\mu)$ with $\psi \geqslant 0$ and $\int_Z \psi d\mu = 1$. Since both sides of (3.4) are linear in ψ and $L^1(Z,\mu)$ is spanned by positive functions of integral 1, this equation must hold for all $\psi \in L^1(Z,\mu)$ as claimed in the proposition.

For the last sentence in the proposition, we note the inequality

$$|\langle\phi,\psi\rangle_l| \leqslant \|\phi\|_\infty \cdot \|\psi\|_{L^1(Z,\mu)}. \tag{3.5}$$

If we know (3.4) for collections $D_1 \subset C_0(Z',\mathbb{C}), D_2 \subset L^1(Z,\mu)$, then the bilinearity of $\langle\cdot,\cdot\rangle_l$ and the inequality (3.5) allows to propagate (3.4) to the closures of the spans of D_1, D_2 in the corresponding Banach space. ☐

The following proposition is the source of the *almost any* part in our main Theorem 2.11 and its derivatives.

Proposition 3.12. *Let Z and Z' be locally compact second countable Hausdorff spaces. Let $\mu \in \mathcal{M}(Z)$ and let $\nu \in \mathcal{M}(Z')$. Let $(f_l)_{l\in\mathbb{N}}$ be a sequence of continuous maps from $Z \to Z'$, which mixes μ to ν. Then, for μ almost every $z \in Z$, we have*

$$\overline{\{f_l(z) : l \in \mathbb{N}\}} \supset \operatorname{supp}\nu.$$

Proof. As $\operatorname{supp}\nu \subset Z'$, it has a countable dense set D. It is enough to show that for any $z' \in D$, μ-almost any z satisfies that $\{f_l(z)\}$ visits any open neighborhood of z. Since there is a countable base for the topology of Z', it is enough to fix a neighborhood U of z' and show that μ-almost any z satisfies that for some l, $f_l(z) \in U$. By way of contradiction, there exists $z' \in D$ and an open neighborhood U of z' such that the set

$$S = \{z \in Z : \forall l, f_l(z) \notin U\}$$

has positive μ-measure.

Consider the nonzero indicator function $\mathbb{1}_S \in L^1(\mu)$ and the probability measure $\eta := \mu(S)^{-1}\mathbb{1}_S d\mu$, which is absolutely continuous with respect to μ. By definition of mixing convergence, we have $(f_l)_*\eta \to \nu$. In particular, if we choose a nonzero function $\phi \in C_c(Z', \mathbb{C})$ with compact support contained in U and such that $\phi \geqslant 0$ and $\phi(z') > 0$, then the above weak* convergence implies

$$\lim_{l\to\infty} \mu(S)^{-1} \int_Z \phi(f_l(z))\mathbb{1}_S(z)d\mu(z) = \int_{Z'} \phi d\nu.$$

However, the left-hand side is zero for all l by the definition of S, U, and ϕ, while the integral on the right-hand side is positive since $z' \in \operatorname{supp}\nu$ and ϕ was chosen nonnegative and positive at z'. We arrive at a contradiction and conclude the proof. □

4 | PROOF OF THEOREM 1.6

In this section, we reproduce (as promised in the introduction) the proof of Theorem 1.6, which appeared implicitly in [26]. The proof relies on the notion of mixing convergence of measures and elementary theory of characters. Although Theorem 1.6 is not essential for the proof of our main result Theorem 2.11, its proof is a simplified version of the one for Theorem 2.11 and prepares the reader for the discussion in Section 6.

Definition 4.1 (Tori and characters). Let $d \in \mathbb{N}$. If $x \subset \mathbb{R}^d$ is a lattice, a character on $\mathbb{R}^d/x$ is a continuous group homomorphism

$$\chi : \mathbb{R}^d/x \to \mathbb{C} \setminus \{0\}.$$

Linear combinations of characters are dense in $C(\mathbb{R}^d/x, \mathbb{C})$. See, for example, [21, 4.24]. If $\mu \in \mathcal{M}(\mathbb{R}^d/x)$, Lemma 3.2 and the linearity of integrals show that μ is uniquely determined in $\mathcal{M}(\mathbb{R}^d/x)$ by the values

$$\left\{ \int_{\mathbb{R}^d/x} \chi d\mu \in \mathbb{C} : \chi \text{ is a character} \right\}. \tag{4.1}$$

When μ is the Haar probability on $\mathbb{R}^d/x$, the set of values in (4.1) is 1 if χ is the identity character and 0 otherwise. On $\mathbb{R}^d$, we use the standard inner product $(\cdot|\cdot)$ and, for the lattice $x \subset \mathbb{R}^d$, we define the dual lattice

$$x^* := \{v \in \mathbb{R}^d \ : \ \text{for all } a \in x \text{ we have } (v|a) \in \mathbb{Z}\}.$$

The set of characters of $\mathbb{R}^d/x$ can then be identified with x^* where, for $v \in x^*$ and $w \in \mathbb{R}^d/x$, we have

$$\chi_v(w + x) := \exp(2i\pi(v|w)).$$

The lattice $\mathbb{Z}^d$ is self-dual.

Recall the notation h_t in (1.6) and Definition 1.11. The following proposition shows how nondivergence relates to mixing convergence. Its corollary together with the the Inheritance Lemma 2.6 relate the discussion to the notion of almost sure DV_F (see Definition 2.5).

Proposition 4.2. *Let $x_0, x_1 \in X$ be two lattices with $x_1 \in \partial(x_0)$. Accordingly, let $(t_l)_{l\in\mathbb{N}}$ be a divergent sequence of positive reals and let $(\varepsilon_l)_{l\in\mathbb{N}} \subset G$ be a sequence converging to the identity such that*

$$\varepsilon_l h_{t_l} x_0 = x_1 \text{ for all } l \in \mathbb{N}. \tag{4.2}$$

Then, $\varepsilon_l h_{t_l}$ mixes the Haar measure $\lambda_{\mathbb{R}^d/x_0}$ to the Haar measure $\lambda_{\mathbb{R}^d/x_1}$.

Proof. Note, the condition (4.2) ensures that the sequence of matrices $(\varepsilon_l h_{t_l})_{l\in\mathbb{N}}$ induces group morphisms

$$\mathbb{R}^d/x_0 \to \mathbb{R}^d/x_1.$$

We use the same notation for these induced maps. We let μ be the Haar probability on $\mathbb{R}^d/x_0$, ν be the Haar probability on $\mathbb{R}^d/x_1$, and claim that the sequence $(\varepsilon_l h_{t_l})_{l\in\mathbb{N}}$ mixes μ to ν. In order to show this, we must check, in light of Proposition 3.11 and the comments in Definition 4.1, that for every $b \in x_1^*$ and $a \in x_0^*$ and corresponding characters χ_a and χ_b on $\mathbb{R}^d/x_0$ and $\mathbb{R}^d/x_1$, respectively, we have

$$\lim_{l\to\infty} \int_{\mathbb{R}^d/x_0} \chi_b(\varepsilon_l h_{t_l}(z))\chi_a(z)d\mu(z) \overset{?}{=} \left(\int_{\mathbb{R}^d/x_0} \chi_a d\mu\right)\left(\int_{\mathbb{R}^d/x_1} \chi_b d\nu\right). \tag{4.3}$$

For the right side, we have

$$\left(\int_{\mathbb{R}^d/x_0} \chi_a d\mu\right)\left(\int_{\mathbb{R}^d/x_1} \chi_b d\nu\right) = \begin{cases} 1 & \text{if } a = 0 \text{ and } b = 0 \\ 0 & \text{if } a \neq 0 \text{ or } b \neq 0 \end{cases}. \tag{4.4}$$

For the left side, we rewrite it as

$$\lim_{l\to\infty} \int_{\mathbb{R}^d/x_0} \exp(2i\pi((\varepsilon_l h_{t_l})^T b + a)|z)d\mu(z), \tag{4.5}$$

where we have taken the transpose of the matrix $(\varepsilon_l h_{t_l})$. When $b = 0$, Equations (4.4) and (4.5) clearly coincide. We now check equality when $b \neq 0$.

Fix $0 \neq b \in x_1^*$ and $a \in x_0^*$. The right-hand side of (4.3) equals 0. The sequence on left-hand side of this equation, $\int_{\mathbb{R}^d/x_0} \chi_{(\varepsilon_l h_{t_l})^T b+a}(z)d\mu(z)$, is a sequence of integrals of characters on $\mathbb{R}^d/x_0$. As such, it is a sequence of 1's and 0's according to whether the character is trivial or not. We will therefore finish once we establish that the vectors $(\varepsilon_l h_{t_l})^T b + a \in x_0^*$ are nonzero for all large enough l. We have

$$(\varepsilon_l h_{t_l})^T b + a = 0 \iff -h_{t_l}^{-1} a = \varepsilon_l^T b.$$

Write $a = \binom{v}{w}$ with $v \in \mathbb{R}^m, w \in \mathbb{R}^n$. By (1.6), we see that $h_{t_l}^{-1} a = \begin{pmatrix} e^{-nt_l} v \\ e^{mt_l} w \end{pmatrix}$ and since $t_l \to \infty$, this sequence of vectors either diverges or converges to zero. In any case, it is impossible that it will coincide with the sequence $\varepsilon_l^T b \to b \neq 0$ for infinitely many l's (where here we have used that ε_l converges to the identity and the assumption that $b \neq 0$). □

Applying Proposition 3.12, we obtain the following corollary.

Corollary 4.3. *Let $x_0 \in X$ be a lattice in $\mathbb{R}^d$ and assume $\{h_t x_0 : t \geqslant 0\}$ is nondivergent. Then, x_0 is $\lambda_{\mathbb{R}^d/x_0}$-almost surely DV_F.*

Proof. By the nondivergence assumption, we deduce the existence of a lattice $x_1 \in \partial(x_0)$. As a consequence, we get the existence of a divergent sequence of reals t_l and a sequence $\varepsilon_l \to e$ in G satisfying $\varepsilon_l h_{t_l} x_0 = x_1$. By Proposition 4.2, we conclude that $\varepsilon_l h_{t_l}$ mixes $\lambda_{\mathbb{R}^d/x_0}$ to $\lambda_{\mathbb{R}^d/x_1}$. An application of Proposition 3.12 shows that, for $\lambda_{\mathbb{R}^d/x_0}$-almost every grid $y \in \pi^{-1}(x_0) = \mathbb{R}^d/x_0$, we have $\overline{\left\{\varepsilon_l h_{t_l} y : l \in \mathbb{N}\right\}} \supset \pi^{-1}(x_1)$. Since $\varepsilon_l \to e$, we deduce that $\overline{\left\{h_{t_l} y : l \in \mathbb{N}\right\}} \supset \pi^{-1}(x_1)$. An application of the Inheritance Lemma 2.6 shows that for $\lambda_{\mathbb{R}^d/x_0}$-almost any grid y,

$$\overline{V_F(y)} \supset \bigcup_{v \in \mathbb{R}^d} V_F(x_1 + v) = F(\mathbb{R}^d) = [0, \infty).$$

This means exactly that x_0 is $\lambda_{\mathbb{R}^d/x_0}$-almost surely grid DV. □

We end this section with the following:

Proof of Theorem 1.6. Let A be as in the statement and let x_A be as in (1.5). By Theorem 1.10, the nonsingularity of A is equivalent to the fact that $\partial(x_A) \neq \emptyset$ and so Corollary 4.3 applies and gives us that x_A is $\lambda_{\mathbb{R}^d/x_A}$-almost surely DV_F. This can be rephrased as saying that for Lebesgue almost any $\binom{\eta}{w} \in \mathbb{R}^d$ the grid $x_A - \binom{\eta}{w}$ is DV_F. For each such $\binom{\eta}{w}$, we therefore have

$$\begin{aligned} F\left(x_A - \begin{pmatrix} \eta \\ w \end{pmatrix}\right) &= F(\left\{ \begin{pmatrix} p + Aq - \eta \\ q - w \end{pmatrix} : p \in \mathbb{Z}^m, q \in \mathbb{Z}^n \right\} \\ &= \{\|q - w\|^n \|p + Aq - \eta\|^m : p \in \mathbb{Z}^m, q \in \mathbb{Z}^n\} \end{aligned}$$

contains arbitrarily small positive values. Because using finitely many q's cannot produce infinitely many arbitrarily small positive values, we deduce the existence of a sequence $q_i \in \mathbb{Z}^n$ with $\|q_i\| \to \infty$ and $p_i \in \mathbb{Z}^m$ such that $\|q_i - w\|^n \|p_i + Aq_i - \eta\|^m$ is a sequence of positive numbers approaching zero. In particular,

$$\liminf_{q \in \mathbb{Z}^n \|q\| \to \infty} \|q\|^{n/m} \langle Aq - \eta \rangle = 0$$

and so $\eta \notin \mathrm{Bad}_A$. This shows $\lambda_{\mathbb{T}^m}(\mathrm{Bad}_A) = 0$ as desired. □

5 | CONVERGENCE OF PROBABILITIES ON TORI AND THE COSET LEMMA

In the previous section, we proved Corollary 4.3, which is reminiscent of Theorem 2.11. In it one assumes x_0 is not divergent and deduces that x_0 is almost surely grid DV_F with respect to the algebraic probability measure $\lambda_{\mathbb{R}^d/x_0}$. A key input in the proof was that pushforwards of $\lambda_{\mathbb{R}^d/x_0}$ converge mixingly (Proposition 4.2). Since Theorem 2.11 deals with any nontrivial algebraic probability measure on $\mathbb{R}^d/x_0$, we will need to further analyze how such probabilities converge under pushforwards.

Theorem 5.1. *Let $x \in X$ be a lattice and $\mathbb{R}^d/x$ be the corresponding torus. Let μ_l be a sequence of Haar probability measures on it corresponding to the x-rational subspaces V_l. Assume, $\mu_l \to \nu$ in the weak* topology. Then:*

(1) *the subspace $V_\infty := \bigcap_j \operatorname{span}\left(\bigcup_{l>j} V_l\right)$ is x-rational;*
(2) *V_∞ is the smallest subspace containing all but finitely many of the V_l's;*
(3) *ν is the Haar probability measure corresponding to the x-rational subspace V_∞;*
(4) *$\dim \nu \geqslant \limsup_l \dim \mu_l$;*
(5) *$\dim \nu = \limsup_l \dim \mu_l$ if and only if there are infinitely many l's with $V_l = V_\infty$.*

Proof. (1) Each space $U_j := \operatorname{span}\left(\bigcup_{l>j} V_l\right)$ is x-rational as it is spanned by x-rational spaces. In turn, V_∞ is x-rational as the intersection of such spaces.

(2) We have that U_j is a descending sequence of subspaces and hence it stabilizes. Let j_0 be such that $U_{j_0} = V_\infty$. It follows from the definition of U_{j_0} that V_∞ contains all but finitely many of the V_j's. Moreover, if V is a subspace containing all but finitely many of the V_j's, then by definition $U_j < V$ for some j and so $V_\infty < V$.

(3) Because the characters $\{\chi_a : a \in x^*\}$ span a dense subspace of $C_0(\mathbb{R}^d/x, \mathbb{C})$, Lemma 3.2 implies that any probability measure μ on $\mathbb{R}^d/x$ is characterized by its Fourier coefficients

$$\left\{\hat{\mu}(a) := \int \chi_a d\mu \,:\, a \in x^*\right\}.$$

Since characters on compact abelian groups integrate to 0 or 1 with respect to the Haar measure according to the triviality of the character, the subset of $\mathcal{P}(\mathbb{R}^d/x)$ consisting of Haar measures on subtori is characterized as follows: μ is the Haar measure corresponding to the x-rational subspace

$V < \mathbb{R}^d$ if and only if

$$\begin{aligned}&\text{for any } a \in x^*,\ \hat{\mu}(a) \in \{0,1\},\\ &\text{and } \hat{\mu}(a) = 1 \iff a \in V^\perp \cap x^*.\end{aligned} \tag{5.1}$$

From the definition of V_∞ and elementary properties of $\perp$, we have

$$V_\infty^\perp = \bigcup_j \bigcap_{l>j} V_l^\perp. \tag{5.2}$$

We show that ν is the Haar measure corresponding to the x-rational subspace V_∞ using the characterization (5.1) of Haar measures. Since characters are continuous and ν is a weak* limit of Haar measures, we have that for any $a \in x^*$, $\hat{\nu}(a) \in \{0,1\}$. For each $a \in x^*$, we are left to check that $\hat{\nu}(a) = 1$ if and only if $a \in V_\infty^\perp$.

We know by the weak* convergence assumption that $\hat{\mu}_l(a)$ converges and thus stabilizes. Thus, $\hat{\nu}(a) = 1$ if and only if for all large enough l, we have $\hat{\mu}_l(a) = 1$. Because μ_l is the Haar measure corresponding to V_l, this is equivalent to saying that for all large enough l, $a \in V_l^\perp$, which by (5.2) means that $a \in V_\infty^\perp$.

(4) As noted before, $V_\infty = \operatorname{span}\left(\cup_{l>j_0} V_l\right)$ for some j_0 and so the inequality $\dim \nu = \dim V_\infty \geqslant \limsup_l \dim V_l = \limsup \dim \mu_l$ is immediate.

(5) From the equality $V_\infty = \operatorname{span}\left(\cup_{l>j_0} V_l\right)$, it is clear that $\dim V_\infty = \limsup_l \dim V_l$ if and only if $V_\infty = V_l$ for infinitely many l's. □

The following corollary deals with limits of algebraic probability measures on a torus.

Corollary 5.2. *Let $x \in X$ be a lattice and let $\mu_l \in \mathcal{P}(\mathbb{R}^d/x)$ be a sequence of algebraic measures corresponding to the x rational subspaces V_l and translations w_l. If μ_l converges to ν, then ν is algebraic,* $\dim \nu \geqslant \limsup_l \dim \mu_l$, *and if there is equality, for infinitely many l's V_l is a fixed subspace of dimension* $\dim \nu$.

Proof. It is enough to prove the statement for a subsequence. Denote by μ_l^0 the Haar probability measures corresponding to V_l so that $\mu_l = (\ell_{w_l})_* \mu_l^0$. Here, $\ell_\cdot$ denotes translation. By taking a subsequence, we may assume that the $\mu_l^0 \to \nu^0$ for some $\nu^0 \in \mathcal{P}(\mathbb{R}^d/x)$ and furthermore that $w_l \to w$ in $\mathbb{R}^d/x$. By Theorem 5.1, ν^0 is a Haar probability measure. It follows from Lemma 3.9 that

$$\nu = \lim_l \mu_l = \lim_l (\ell_{w_l})_* \mu_l^0 = (\ell_w)_* \nu^0,$$

which means that ν is algebraic. Moreover, Theorem 5.1 also gives

$$\dim \nu = \dim \nu^0 \geqslant \limsup_l \dim V_l$$

with equality possible only if infinitely many of the V_l's are equal to the x-rational subspace corresponding to ν^0. □

The next lemma captures the following example as a general phenomenon.

Example 5.3. Consider the sequence of maps $(\gamma_l)_{l\in\mathbb{N}} : \mathbb{T}^1 \to \mathbb{T}^2$ induced by the sequence of matrix maps

$$\begin{bmatrix} l \\ 1 \end{bmatrix} : \mathbb{R} \to \mathbb{R}^2.$$

Let $S \subset \mathbb{T}^2$ be the 1-dimensional closed connected subgroup corresponding to $\mathbb{R}\mathbf{e}_1$, the span of the first basis vector in $\mathbb{R}^2$. Then, for $t \in [0,1]$, which is irrational, the closure of the set

$$\left\{ \begin{bmatrix} tl \\ t \end{bmatrix} + \mathbb{Z}^2 \in \mathbb{T}^2 : l \in \mathbb{N} \right\}$$

contains a coset of S.

Lemma 5.4 (The Coset Lemma). *Let $q \in \mathbb{N}$ and let $(\gamma_l)_{l\in\mathbb{N}}$ be a sequence of continuous group morphisms from $\mathbb{T}^1 \to \mathbb{T}^q$. Let $\mu \in \mathcal{P}(\mathbb{T}^1)$ and $\nu \in \mathcal{P}(\mathbb{T}^q)$ denote the full Haar probabilities. Assume we have the weak* convergence*

$$\lim_{l\to\infty} (\gamma_l)_*\mu = \nu. \tag{5.3}$$

Then, there is a codimension 1 closed connected subgroup $S \subset \mathbb{T}^q$ such that, for μ almost every $z \in \mathbb{T}^1$,

$$\overline{\{\gamma_l(z) \in \mathbb{T}^q : l \in \mathbb{N}\}} \supset S + z' \textit{ for some } z' \in \mathbb{T}^q \textit{ depending on } z.$$

Proof. The lemma is trivially true if $q = 1$ since we may take the identity subgroup. We assume $q > 1$. Without loss of generality, assume that each $\gamma_l \in \mathrm{Mat}_{q\times 1}(\mathbb{Z})$ and acts in $\mathbb{T}^1$ by multiplication on column vectors (a 1×1 column vector). We use γ_l^T to denote the transpose matrix in $\mathrm{Mat}_{1\times q}(\mathbb{Z})$. If $(\gamma_l)_{l\in\mathbb{N}}$ mixes μ to ν, we are done by Proposition 3.12. In particular, we get something much stronger than the desired conclusion. That is, for μ almost every z, the set $\{\gamma_l(z) \in \mathbb{T}^q : l \in \mathbb{N}\}$ is dense. We proceed, assuming that $(\gamma_l)_{l\in\mathbb{N}}$ fails to mix μ to ν.

Note that it is enough to prove the statement for a subsequence of γ_l. Moreover, if we replace γ_l by $\gamma\gamma_l$ for a fixed $\gamma \in \mathrm{SL}_q(\mathbb{Z})$, then it is enough to prove the statement for the new sequence $\gamma\gamma_l$. Thus, along the proof, we will take subsequences and replace γ_l by appropriate linear images of γ_l and by abuse of notation, allow ourselves to continue and denote the new sequence by γ_l.

By the bilinearity of formula (3.4), the failure of mixing means that there are integer vectors $a \in \mathbb{Z}$ and $b \in \mathbb{Z}^q$ such that, if we denote by χ_a and χ_b the respective characters on $\mathbb{T}^1$ and $\mathbb{T}^q$, we have

$$\lim_{l\to\infty} \int_{\mathbb{T}^1} \chi_b(\gamma_l(z))\chi_a(z)d\mu(z) \neq \left(\int_{\mathbb{T}^1} \chi_a d\mu\right)\left(\int_{\mathbb{T}^q} \chi_b d\nu\right).$$

In other words, we have

$$\lim_{l\to\infty} \int_{\mathbb{T}^1} \exp(2i\pi(a + \gamma_l^T(b)|z))d\mu(z) \neq \left(\int_{\mathbb{T}^1} \chi_a d\mu\right)\left(\int_{\mathbb{T}^q} \chi_b d\nu\right).$$

Thus, b is nonzero. As the right-hand side is 0 in this case, the left-hand side must equal 1 for infinitely many l's, so we may assume after passing to a subsequence that

$$\gamma_l^T(b) = -a \text{ for all } l \in \mathbb{N}. \tag{5.4}$$

Without loss of generality, we can assume b is primitive. Let $\gamma \in \mathrm{SL}_q(\mathbb{Z})$ be such that

$$\gamma^T(\mathbf{e}_q) = b, \tag{5.5}$$

where $\mathbf{e}_q \in \mathbb{Z}^q$ is the qth vector in the standard basis. On replacing γ_l by $\gamma\gamma_l$, we see that we may assume without loss of generality $b = \mathbf{e}_q$. By (5.4), we deduce that

$$\gamma_l = \begin{bmatrix} \gamma_l' \\ -a \end{bmatrix} \tag{5.6}$$

for all l, where $\gamma_l' \in \mathrm{Mat}_{(q-1)\times 1}$.

Now we view $\gamma_l' : \mathbb{T}^1 \to \mathbb{T}^{q-1}$ as a sequence of homomorphisms and let ν' denote the Haar probability measure on $\mathbb{T}^{q-1}$. Note that the weak* convergence $(\gamma_l')_*\mu \to \nu'$ follows from the assumed convergence $(\gamma_l)_*\mu \to \nu$ by composition with the projection $\mathbb{T}^q \to \mathbb{T}^{q-1}$ that forgets the last coordinate. This relies on Lemma 3.7.

Claim 5.4.1. The sequence γ_l' must mix μ to ν'.

Proof. If not, by the same logic as above, there must be some $\gamma' \in \mathrm{SL}_{q-1}(\mathbb{Z})$ and $a' \in \mathbb{Z}$ such that, on replacing γ_l' by $\gamma'\gamma_l'$ and on passing to a subsequence, we have

$$\gamma_l' = \begin{bmatrix} \gamma_l'' \\ -a' \end{bmatrix}.$$

This means that after an appropriate replacement of γ_l, we have

$$\gamma_l = \begin{bmatrix} \gamma_l'' \\ -a' \\ -a \end{bmatrix}. \tag{5.7}$$

Let $\delta \in \mathrm{SL}_2(\mathbb{Z})$ be such that $\delta\left(\begin{smallmatrix} a' \\ a \end{smallmatrix}\right) = \left(\begin{smallmatrix} * \\ 0 \end{smallmatrix}\right)$ and let $\gamma \in \mathrm{SL}_q(\mathbb{Z})$ be equal to the identity matrix with the bottom right 2×2 block replaced by δ. Then, by the choice of δ and (5.7), we have that the last coordinate of $\gamma\gamma_l$ is 0. Replacing γ_l by $\gamma\gamma_l$, we see that the image of the homomorphism $\gamma_l : \mathbb{T}^1 \to \mathbb{T}^q$ is contained in the subtorus defined by requiring the last coordinate to vanish. This contradicts the weak* convergence $(\gamma_l)_*\mu \to \nu$ and so finishes the proof of the claim. □

We apply Proposition 3.12 to the sequence of homomorphisms $\gamma_l' : \mathbb{T}^1 \to \mathbb{T}^{q-1}$ and conclude that for μ-almost any $y \in \mathbb{T}^1$, the sequence of images $\left\{\gamma_l' y \in \mathbb{T}^{q-1}\right\}$ is dense. For each such y, we have by (5.6) that

$$\{\gamma_l y : l \in \mathbb{N}\} = \left\{ \begin{pmatrix} \gamma_l' y \\ -ay \end{pmatrix} \in \mathbb{T}^q : l \in \mathbb{N} \right\}$$

and so $\overline{\{\gamma_l y : l \in \mathbb{N}\}}$ contains a coset (depending on y) of the subtorus obtained by requiring the last coordinate to vanish. □

6 | PROOF OF THEOREM 2.11

Here is a sketch of the proof of Theorem 2.11 that should motivate the technicalities to come. The proof uses as an input an accumulation sequence $(x_0, x_1, \dots, x_{d-1})$. We start with a 1-dimensional algebraic measure μ on the torus $\mathbb{R}^d/x_0$ and track what happens to the measure μ when we push it using sequences h_{t_l} witnessing that $x_1 \in \partial(x_0), x_2 \in \partial(x_1)$, and so forth. At each step, we may assume the measures converge and we move from an algebraic measure on $\mathbb{R}^d/x_i$ to an algebraic measure on $\mathbb{R}^d/x_{i+1}$. The requirement regarding the length of the accumulation sequence is there to ensure that at the last step we must see the Haar probability measure on $\mathbb{R}^d/x_{d-1}$. The proof ends by invoking the Coset Lemma 5.4 and using the Inheritance Lemma 2.6 together with the fact that the nondegeneracy degree of F is $d-1$, which is Lemma 2.9.

In order for this strategy to work, we must know that the dimension of the algebraic measure we obtain in the limit at each stage jumps by at least 1. This is the goal of the next section and the reason for our assumption $\gcd(m,n) = 1$.

6.1 | Jump in dimension

Consider the Euclidean space $\mathbb{R}^d$ with the standard inner product. Fix a dimension $1 \leqslant k \leqslant d$ and consider the kth exterior power $\wedge^k\mathbb{R}^d$. On this vector space, we induce an inner product by declaring the standard basis vectors $\mathbf{e}_I$ orthogonal. Here, I is a multi-index of length k, that is, a sequence $1 \leqslant i_1 < \cdots < i_k \leqslant d$ and $\mathbf{e}_I := \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_k}$. We denote by $\mathcal{I}_k$ the set of all multi-indices of length k.

The following lemma is left as an exercise. It explains how the exterior power relates to our discussion: We deal with algebraic measures supported on a subtorus of $\mathbb{R}^d/x$ where $x \in X$ is a lattice. Such a subtorus corresponds to an x-rational subspace V and to a vector in the appropriate exterior power obtained by wedging a $\mathbb{Z}$-basis of $V \cap x$.

Lemma 6.1. *Let $\mathbf{u} = u_1 \wedge \cdots \wedge u_k \in \wedge^k\mathbb{R}^d$. Then, $\|u\|$ is the k-dimensional volume of the parallelepiped spanned by $u_1, \dots, u_k$ in $\mathbb{R}^d$. In particular, if $V < \mathbb{R}^d$ is a k-dimensional subspace and $v_1, \dots, v_k, w_1, \dots, w_k$ are $\mathbb{Z}$-bases of a lattice in V, then $v_1 \wedge \cdots \wedge v_k = \pm w_1 \wedge \cdots \wedge w_k$.*

Each linear map $g : \mathbb{R}^d \to \mathbb{R}^d$ induces uniquely a linear map

$$\wedge g : \wedge^k\mathbb{R}^d \to \wedge^k\mathbb{R}^d$$

characterized by the way it acts on pure tensors: For all $v_1, \dots, v_k \in \mathbb{R}^d$, $\wedge g(v_1 \wedge \cdots \wedge v_k) = (gv_1) \wedge \cdots \wedge (gv_k)$. We consider the semigroup of matrices $(h_t)_{t\in\mathbb{R}_{\geqslant 0}} \in G$ as in (1.6) and consider their linear action on $\mathbb{R}^d$ by multiplication on column vectors. These induce the family of linear maps

$$(\wedge h_t)_{t\in\mathbb{R}_{\geqslant 0}} : \wedge^k\mathbb{R}^d \to \wedge^k\mathbb{R}^d,$$

which are simultaneously diagonalizable. A basis of orthogonal eigenvectors is given by the standard basis vectors $\{\mathbf{e}_I : I \in \mathcal{I}_k\}$ as we will show below. We split the indices $[d] := \{1, \dots, d\}$ into $J_+ := \{1, \dots, m\}$ and $J_- := [d] \setminus J_+$. With this notation it is easy to see that for any multi-index $I \in \mathcal{I}_k$,

$$(\wedge h_t)\mathbf{e}_I = \exp\big((|I \cap J_+|n - |I \cap J_-|m)t\big) \cdot \mathbf{e}_I. \tag{6.1}$$

We denote

$$\alpha_I(t) := \exp\big((|I \cap J_+|n - |I \cap J_-|m)t\big) \tag{6.2}$$

and note that if the exponent $(|I \cap J_+|n - |I \cap J_-|m)$ is not zero, then $\alpha_I(t) \to \infty$ or $\alpha_I(t) \to 0$ as $t \to \infty$.

Lemma 6.2. *Assume* $\gcd(n, m) = 1$. *Then, for any* $1 \leqslant k < d$, *there is no* $I \in \mathcal{I}_k$ *such that* $\alpha_I(t) \equiv 1$. *Moreover, for any* $u \in \wedge^k\mathbb{R}^d$, *we have that* $\|(\wedge h_t)u\|$ *either diverges or converges to 0 as* $t \to \infty$.

Proof. Fix k as in the statement and let $I \in \mathcal{I}_k$. By Equation (6.2), if $\alpha_I(t) \equiv 1$, then there are $1 \leqslant a \leqslant b \leqslant k$ such that $a + b = k$ and $an - mb = 0$. By the gcd assumption, we deduce that $m|a$ and $n|b$. But this gives that $d = m + n \leqslant a + b = k < d$, which is absurd.

Finally, if $u \in \wedge^k\mathbb{R}^d$, we may write $u = \sum_{I\in\mathcal{I}_k} a_I\mathbf{e}_I$. By (6.1), we have

$$(\wedge h_t)u = \sum_{I\in\mathcal{I}_k} a_I(\wedge h_t)\mathbf{e}_I = \sum_{I\in\mathcal{I}_k} \alpha_I(t)a_I\mathbf{e}_I.$$

By the first part, for each I, $\lim_{t\to\infty} \alpha_I(t) \in \{0, \infty\}$, and hence we deduce that $\|(\wedge h_t)u\| \to \infty$ unless $a_I = 0$ for every I for which $\alpha_I(t) \to \infty$, in which case, $\|(\wedge h_t)u\| \to 0$. ☐

Corollary 6.3. *Assume* $\gcd(m, n) = 1$. *Suppose* $x_0, x_1 \in X$ *are lattices such that* $x_1 \in \partial(x_0)$. *Let* $t_l \to \infty$ *and* $G \ni \varepsilon_l \to e$ *be such that* $\varepsilon_l h_{t_l} x_0 = x_1$. *Then, if* $U < \mathbb{R}^d$ *is a* k*-dimensional* x_0*-rational subspace with* $1 \leqslant k < d$, *it is not possible that the sequence of subspaces* $\varepsilon_l h_{t_l} U$ *stabilizes.*

Proof. Suppose on the contrary that the x_1-rational subspace $\varepsilon_l h_{t_l} U$ is independent of l and denote it by V. Note that $\varepsilon_l h_{t_l}(U \cap x_0) = V \cap x_1$. Choose a basis $u_1, \dots, u_k$ of $U \cap x_0$ and denote $\mathbf{u} := u_1 \wedge \cdots \wedge u_k \in \wedge^k\mathbb{R}^d$. We have that for any l,

$$(\wedge\varepsilon_l)(\wedge h_{t_l})(\mathbf{u}) = \wedge(\varepsilon_l h_{t_l})(\mathbf{u}) = (\varepsilon_l h_{t_l} u_1) \wedge \cdots \wedge (\varepsilon_l h_{t_l} u_k).$$

The right-hand side is a pure tensor obtained by wedging a basis of the lattice $x_1 \cap V$ in V and hence, by Lemma 6.1 does not depend on l (up to a sign maybe). In particular, the norm of this vector is bounded above and below (away from 0). On the other hand, the left-hand side goes to infinity or to zero by Lemma 6.2 because the operators $\wedge \varepsilon_l$ converge to the identity map on $\wedge^k \mathbb{R}^d$. This gives the desired contradiction. □

Theorem 6.4 (Jump in dimension). *Assume* $\gcd(m, n) = 1$. *Let* $x_0, x_1 \in X$ *with* $x_1 \in \partial(x_0)$. *Let* μ *be an algebraic measure on* $\mathbb{R}^d/x_0$ *of dimension* $1 \leqslant k < d$. *Then, there is a divergent sequence* $(t_l)_{l\in\mathbb{N}} \subset \mathbb{R}_{\geqslant 0}$ *with*

$$\lim_{l\to\infty} (h_{t_l})_* \mu = \nu \ \textit{ in the weak* topology of } \mathcal{M}(Y),$$

where ν *is an algebraic measure on* $\mathbb{R}^d/x_1$ *and moreover,*

$$\dim \nu \geqslant k + 1. \tag{6.3}$$

Proof. Choose a divergent sequence $(t_l)_{l\in\mathbb{N}} \subset \mathbb{R}_{\geqslant 0}$ witnessing $x_1 \in \partial(x_0)$. Let $(\varepsilon_l)_{l\in\mathbb{N}} \subset G$ be a sequence converging to the identity with

$$\varepsilon_l h_{t_l} x_0 = x_1 \ \text{ for all } \ l \in \mathbb{N}.$$

Acting on Y rather than on X, the above equation says that $\varepsilon_l h_{t_l}$ maps $\pi^{-1}(x_0) = \mathbb{R}^d/x_0$ to $\pi^{-1}(x_1) = \mathbb{R}^d/x_1$ and is in fact a group homomorphism between these two tori. By Lemma 3.3, we can assume that we have a weak* limit

$$\lim_{l\to\infty} (h_{t_l})_* \mu = \lim_{l\to\infty} (\varepsilon_l h_{t_l})_* \mu = \nu \in \mathcal{P}(\pi^{-1}(x_1)), \tag{6.4}$$

where the equality on the left follows by applying Lemma 3.9 and noting that $(\varepsilon_l)_{l\in\mathbb{N}}$ converges to the identity in G. An application of Corollary 5.2 shows that ν must be algebraic.

It remains to establish (6.3). Assume that the algebraic measure μ on $\mathbb{R}^d/x_0$ corresponds to the x_0-rational subspace $U \subset \mathbb{R}^d$. Then, the measures $(\varepsilon_l h_{t_l})_* \mu \in \mathcal{P}(\mathbb{R}^d/x_1)$ are algebraic measures that correspond to the x_1-rational subspaces $V_l := \varepsilon_l h_{t_l} U$ and are thus of a fixed dimension k. Applying Corollary 5.2, we deduce that $\dim \nu \geqslant k$ and that if $\dim \nu = k$, then along a subsequence, $\varepsilon_l h_{t_l} U$ is a fixed x_1-rational subspace. This contradicts Corollary 6.3 because we assume $\gcd(m, n) = 1$ and $1 \leqslant k < d$. □

The following corollary bootstraps Theorem 6.4 for a long accumulation sequence. This is used in order to reach the dimension of the nondegeneracy degree of F as appears in Lemma 2.9.

Corollary 6.5. *Assume* $\gcd(m, n) = 1$. *Let* $x_0 \in X$ *be a lattice with an accumulation sequence* $(x_0, \ldots, x_r) \in X^{r+1}$. *Let* μ *be an algebraic measure on* $\mathbb{R}^d/x_0$ *of dimension* $1 \leqslant k < d$. *Then, there is a divergent sequence* $(t_l)_{l\in\mathbb{N}} \subset \mathbb{R}_{\geqslant 0}$ *with*

$$\lim_{l\to\infty} (h_{t_l})_* \mu = \nu \ \textit{ in the weak* topology of } \mathcal{M}(Y),$$

where ν *is an algebraic measure on* $\mathbb{R}^d/x_r$ *and moreover,*

$$\dim \nu \geqslant \min\{k + r, d\}. \tag{6.5}$$

Proof. We prove this by induction on the length of the accumulation sequence r. When $r = 1$, this is Theorem 6.4. Assume the statement holds for accumulation sequences of length $r - 1 \geqslant 1$ and let $x_0, \dots, x_r, \mu$ be as in the statement. By the induction hypothesis, there exists a sequence $t_l \to \infty$ such that $(h_{t_l})_*\mu \to \eta$, where η is an algebraic measure in $\mathcal{P}(\mathbb{R}^d/x_{r-1})$ of dimension $\geqslant \min\{k + r - 1, d\}$. Applying Theorem 6.4 to x_{r-1}, x_r we deduce the existence of a sequence $s_l \to \infty$ such that $(h_{s_l})_*\eta \to \nu$ where $\nu \in \mathcal{P}(\mathbb{R}^d/x_r)$ is an algebraic measure with $\dim \nu \geqslant \min\{k + r, d\}$. An application of Lemma 3.8 gives the existence of sequences p_l, q_l of natural numbers such that

$$h_{s_{p_l}+t_{q_l}} = (h_{s_{p_l}})_*(h_{t_{q_l}}) * \mu \to \nu,$$

which finishes the proof. □

We now come to the following proof:

Proof of Theorem 2.11. Let $x \in X$ be as in the statement and

$$(x = x_0, x_1, \dots, x_{d-1}) \in X^d$$

be an accumulation sequence of length d for x. Let μ on $\pi^{-1}(x)$ be an algebraic measure. Similarly to the argument showing the equivalence between the formulations of Theorems 1.14 and 1.16, we may assume that μ is 1-dimensional. Applying Corollary 6.5, we see that there is a sequence of times $(t_l)_{l\in\mathbb{N}} \subset \mathbb{R}_{\geqslant 0}$ and an algebraic measure ν on $\pi^{-1}(x_{d-1})$ such that $\lim_{l\to\infty}(h_{t_l})_*\mu = \nu$. Moreover, because of the inequality (6.5), we must have $\dim \nu = d$ so that in fact, ν is the Haar probability measure on $\mathbb{R}^d/x_{d-1}$. Let $(\varepsilon_l)_{l\in\mathbb{N}} \subset G$ be a sequence converging to the identity with

$$\varepsilon_l h_{t_l} x = x_{d-1} \text{ for all } l \in \mathbb{N}.$$

Acting on Y rather on X, the above equation means that acting with $\varepsilon_l h_{t_l}$ maps the fiber $\pi^{-1}(x_0) = \mathbb{R}^d/x_0$ to $\pi^{-1}(x_{d-1}) = \mathbb{R}^d/x_{d-1}$. Applying the Coset Lemma 5.4, we see that for μ almost every $y \in Y$, the closure $\overline{\left\{\varepsilon_l h_{t_l} y \in Y \,:\, l \in \mathbb{N}\right\}}$ contains a coset of a subtorus of $\mathbb{R}^d/x_{d-1}$ of dimension $d - 1$. Since $(\varepsilon_l)_{l\in\mathbb{N}}$ converges to the identity, the same is true for the set $\overline{\left\{h_{t_l} y \in Y \,:\, l \in \mathbb{N}\right\}}$ where y lies in a set of full μ measure. Proposition 2.7 implies that each such grid y is DV_F and we are done. □

7 | DUAL APPROXIMATION, TRANSFERENCE THEOREMS, AND A COUNTEREXAMPLE

In order to prove Theorem 1.7, we switch to the dual setting of approximation of linear forms to make use of some powerful transference theorems due to Jarnik and Khintchine (see Theorem 7.5). We need some notation.

Definition 7.1 (Irrationality measure function). Let $\theta = (\theta_1, \theta_2, \dots, \theta_m) \in \mathbb{R}^m$. We define $\Psi : [1, \infty) \to [0, \infty)$ by

$$\Psi(t) = \min\{\langle a_1\theta_1 + \dots + a_m\theta_m\rangle \,:\, a_i \in \mathbb{Z} \text{ and } 0 < |a_i| \leqslant t \text{ for } i = 1, \dots m\}.$$

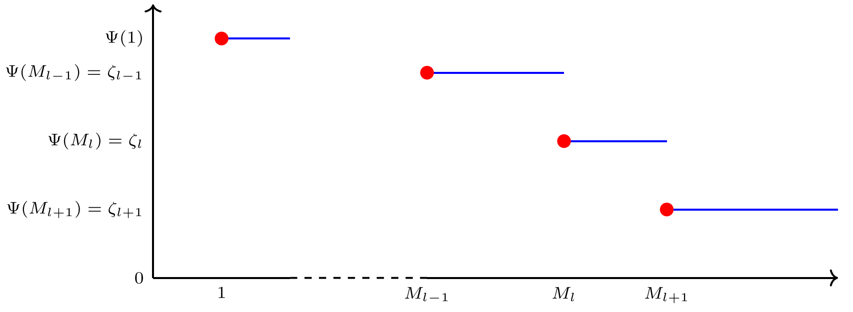

FIGURE 1 Here is a graph of Ψ to remember the notation.

This function is piecewise constant and nonincreasing. Its points of discontinuity occur at integer values. We denote these integers by the increasing sequence $(M_l)_{l\in\mathbb{N}} \subset \mathbb{N}$ and write the images as (see Figure 1)

$$\zeta_l := \Psi(M_l) \text{ for all } l \in \mathbb{N}.$$

We have the following basic estimate.

Lemma 7.2. *For every $\theta \in \mathbb{R}^m$ and associated irrationality measure function Ψ, we have*

$$\Psi(t) \leqslant t^{-m}.$$

Proof. This follows from Minkowski's convex body theorem [25, II, Theorem 2B]. ☐

A key fact we use is the following.

Proposition 7.3. *Let m be an integer with $m > 2$. There exists $\theta \in \mathbb{R}^2$ with irrationality measure function Ψ satisfying, for some positive constants C_1 and C_2, the inequalities*

$$\Psi(t)t^m \leqslant C_1 \textit{ for all } t \geqslant 1,$$

and

$$C_2 \leqslant \Psi(t)t^m \textit{ for an unbounded set of } t \geqslant 1.$$

Remark 7.4. Since the proof of this uses the parametric geometry of numbers as in the paper [20] and requires additional notation, its proof is deferred to the Appendix.

The conditions above are related to the Diophantine properties of the (column) vector $\theta \in \mathbb{R}^2$ by the following transference theorem.

Theorem 7.5 [11, Theorem 7]. *Suppose $\theta \in \mathbb{R}^2$ with associated irrationality measure function Ψ. Suppose further, we have an integer $m > 2$ and a constant $C_1 > 0$ with*

$$\Psi(t)t^m \leqslant C_1 \textit{ for all } t \geqslant 1.$$

Then, there exists $\eta = (\eta_1, \eta_2) \in \mathbb{R}^2$ *with*

$$\inf_{q \in \mathbb{Z}\setminus\{0\}} |q|^{1/m} \langle q\theta - \eta \rangle > 0.$$

Next, we see that the second property in 7.3 is useful in extending the components of the (column) vector $\theta \in \mathbb{R}^2$ to obtain a nonsingular one in higher dimensions.

Theorem 7.6. *Assume* $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$ *with irrationality measure function* Ψ. *Suppose* m *is an integer with* $m > 2$, *and we have a constant* $C_2 > 0$ *with*

$$C_2 \leqslant \Psi(t)t^m \text{ for an unbounded set of } t \geqslant 1. \tag{7.1}$$

Then, for Lebesgue almost every $(\theta_3, \dots, \theta_m) \in \mathbb{R}^{m-2}$, *the vector* $(\theta_1, \theta_2, \theta_3, \dots, \theta_m) \in \mathbb{R}^m$ *is nonsingular.*

Proof. First, recall the sequence of integers $(M_l)_{l \in \mathbb{N}}$ associated to the irrationality measure function Ψ. The condition (7.1) means there is an infinite subset $I \subset \mathbb{N}$ such that, for each $l \in I$, we have

$$C_2 \leqslant \Psi(M_l) M_{l+1}^m. \tag{7.2}$$

Next, we need some auxiliary claims:

Claim 7.6.1. Let $0 < \delta < 1/2$. Given a nonzero integer vector $(a_3, \dots, a_m) \in \mathbb{Z}^{m-2}$ and $M \in \mathbb{N}$, let $S(M, \delta, a_3, \dots, a_m)$ denote the set of integer vectors $(a_0, a_1, a_2) \in \mathbb{Z}^3$ for which

$$\max_{i=1,2} |a_i| < M$$

and for which there exists $(\theta_3, \dots, \theta_m) \in [0, 1]^{m-2}$ with

$$|a_0 + a_1\theta_1 + a_2\theta_2 + a_3\theta_3 + \cdots + a_m\theta_m| \leqslant \delta.$$

Then,

$$\#S(M, \delta, a_3, \dots, a_m) \leqslant 18M^2(|a_3| + \dots |a_m| + 1).$$

Proof of Claim 7.6.1. Say $(a_0, a_1, a_2) \in S(M, \delta, a_3, \dots, a_m)$. We see that a_0 must certainly satisfy

$$-(a_1\theta_1 + a_2\theta_2) - (|a_3| + \cdots + |a_m|) - \delta \leqslant a_0 \leqslant -(a_1\theta_1 + a_2\theta_2) + (|a_3| + \cdots + |a_m|) + \delta.$$

Whence, for fixed (a_1, a_2), there are at most

$$2(|a_3| + \dots |a_m| + 1)$$

options for a_0. Moreover, there are $2M + 1 \leqslant 3M$ options for each of a_1 and a_2. This gives us the required bound. $\square$

For an integer vector $(a_0, a_1, \dots, a_m) \in \mathbb{Z}^{m+1}$ with $(a_3, \dots, a_m)$ nonzero, $M \in \mathbb{N}$ and $0 < \varepsilon < 1/2$, we define the open subset $B_{M,\varepsilon}(a_0, \dots, a_m) \subset [0,1]^{m-2}$ as:

$$\left\{ (\theta_3, \dots, \theta_m) \in [0,1]^{m-2} \, : \, |a_0 + a_1\theta_1 + a_2\theta_2 + a_3\theta_3 + \cdots + a_m\theta_m| < \frac{\varepsilon}{M^m} \right\}.$$

Further, we define $B_{M,\varepsilon}$ as the union

$$\bigcup \left\{ B_{M,\varepsilon}(a_0, \dots, a_m) \, : \, (a_0, \dots, a_m) \in \mathbb{Z}^{m+1}, \; (a_3, \dots, a_m) \neq 0, \; \max_{i=1,\dots,m} |a_i| < M \right\}.$$

We use λ to denote the Lebesgue measure on $\mathbb{R}^{m-2}$. Viewing the inequality occurring in the definition of $B_{M,\varepsilon}(a_0, \dots, a_m)$ in terms of an inner product with a unit vector multiple of $(a_3, \dots, a_m)$, we see that

$$\lambda(B_{M,\varepsilon}(a_0, \dots, a_m)) \leqslant (m-2)^{(m-3)/2} \frac{2\varepsilon}{M^m \sqrt{a_3^2 + \cdots + a_m^2}}. \tag{7.3}$$

Claim 7.6.2. For any $M \in \mathbb{N}$ and ε with $0 < \varepsilon < 1/2$, we have

$$\lambda(B_{M,\varepsilon}) < K\varepsilon,$$

where $K > 0$ is a constant depending only on m.

Proof of Claim 7.6.2. We have the union

$$B_{M,\varepsilon} = \bigcup_{\substack{(a_3,\dots,a_m) \in \mathbb{Z}^{m-2} \setminus \{0\}, \\ |a_3|,\dots,|a_m| < M}} \bigcup \left\{ B_{M,\varepsilon}(a_0, \dots, a_m) \, : \, (a_0, a_1, a_2) \in S(M, \varepsilon/M^m, a_3, \dots, a_m) \right\}.$$

Applying the estimates of Claim 7.6.1 and (7.3), we get

$$\lambda(B_{M,\varepsilon}) \leqslant \sum_{\substack{(a_3,\dots,a_m) \in \mathbb{Z}^{m-2} \setminus \{0\}, \\ |a_3|,\dots,|a_m| < M}} 18M^2(|a_3| + \dots |a_m| + 1) \cdot \frac{2\varepsilon(m-2)^{(m-3)/2}}{M^m \sqrt{a_3^2 + \cdots + a_m^2}}.$$

This gives us the claim. $\square$

Consider the sets, for a fixed $0 < \varepsilon < 1/2$ and each $l \in I$,

$$W_l(\varepsilon) := [0,1]^{m-2} \setminus B_{M_l,\varepsilon}.$$

Let $W(\varepsilon)$ be the limsup set

$$W(\varepsilon) := \bigcap_{p \in \mathbb{N}} \bigcup_{\substack{l > p \\ l \in I}} W_l(\varepsilon)$$

and let

$$W := \bigcup_{0<\varepsilon<1/2} W(\varepsilon).$$

For each $l \in I$ and $0 < \varepsilon < 1/2$, we have from Claim 7.6.2

$$\lambda(W_l(\varepsilon)) \geqslant 1 - K\varepsilon.$$

Thus, $W(\varepsilon)$ is a decreasing intersection of sets each of which has measure greater than or equal to $(1 - K\varepsilon)$. This gives

$$\lambda(W(\varepsilon)) \geqslant 1 - K\varepsilon \text{ for all } 0 < \varepsilon < 1/2.$$

In particular, we have $\lambda(W) = 1$. Now let $(\theta_3, \ldots, \theta_m) \in W$ and consider the vector $(\theta_1, \ldots, \theta_m) \in \mathbb{R}^m$. There is some $0 < \varepsilon < 1/2$ for which $(\theta_3, \ldots, \theta_m) \in W(\varepsilon)$. In particular, we have an infinite number of $l \in I$ such that $(\theta_3, \ldots, \theta_m) \notin B_{M_l,\varepsilon}$.

Fix such an $l > 1$. It follows from the definitions that, for each $(a_0, \ldots, a_m) \in \mathbb{Z}^{m+1}$ with $(a_3, \ldots, a_m) \neq 0$ and $\max_{i=1,\ldots,m} |a_i| < M_l$, we have

$$|a_0 + a_1\theta_1 + a_2\theta_2 + a_3\theta_3 + \cdots + a_m\theta_m| \geqslant \frac{\varepsilon}{M_l^m}.$$

Further, if we have $(a_0, \ldots, a_m) \in \mathbb{Z}^{m+1}$ with $\max_{i=1,\ldots m} |a_i| < M_l$ and $(a_3, \ldots, a_m) = 0$, then

$$\begin{aligned}|a_0 + a_1\theta_1 + a_2\theta_2 + a_3\theta_3 + \cdots + a_m\theta_m| &= |a_0 + a_1\theta_1 + a_2\theta_2| \\ &\geqslant \Psi(M_{l-1}) \\ &\geqslant \frac{C_2}{M_l^m},\end{aligned}$$

where we use the hypothesis (7.2) in the last inequality. If we put $t = M_l - \frac{1}{2}$ and if we let Φ denote the irrationality measure function of $(\theta_1, \ldots, \theta_m) \in \mathbb{R}^m$, we get

$$\Phi(t)t^m \geqslant \frac{1}{2^m} \min\{\varepsilon, C_3\} > 0.$$

Allowing l to vary, we see that $(\theta_1, \ldots, \theta_m) \in \mathbb{R}^m$ is nonsingular. □

We finally come to our counterexample.

Theorem 7.7. *Let m be an integer with $m > 2$. Let $(\theta_1, \theta_2) \in \mathbb{R}^2$ be as in Proposition 7.3. Then, there exists a $(\theta_3, \ldots, \theta_m) \in \mathbb{R}^{m-2}$ and a $(\eta_1, \eta_2) \in \mathbb{R}^2$ such that*

(a) *the augmented vector*

$$\theta = (\theta_1, \ldots, \theta_m) \in \mathbb{R}^m$$

is nonsingular;

(b) *for any* $(\eta_3, \dots, \eta_m) \in \mathbb{R}^{m-2}$*, the augmented vector* $\eta = (\eta_1, \dots, \eta_m) \in \mathbb{R}^m$ *satisfies*

$$\inf_{q \in \mathbb{Z}\setminus\{0\}} |q|^{1/m} \langle q\theta - \eta \rangle > 0.$$

In particular, we have a nonsingular $\theta \in \mathbb{R}^m$ *with* Bad_θ *containing the translate of an* $m-2$*-dimensional subtorus in* $\mathbb{T}^m$*.*

Proof. We apply Theorem 7.6 to obtain part (a). We apply Theorem 7.5 to obtain $(\eta_1, \eta_2) \in \mathbb{R}^2$. We then compute

$$\inf_{q \in \mathbb{Z}\setminus\{0\}} |q|^{1/m} \max_{i=1,\dots,m} \langle q\theta_i - \eta_i \rangle \geqslant \inf_{q \in \mathbb{Z}\setminus\{0\}} |q|^{1/m} \max_{i=1,2} \langle q\theta_i - \eta_i \rangle,$$

$$> 0$$

where we used the conclusion of Theorem 7.5 in the last inequality. □

8 | A STRENGTHENING OF THEOREM 1.6

In this section, we wish to define a Diophantine class of matrices in $C \subset \mathrm{Mat}_{m\times n}(\mathbb{R})$ such that C strictly contains the class of nonsingular matrices and such that for any $A \in C$, $\lambda(\mathrm{Bad}_A) = 0$, where λ is the m-dimensional Lebesgue measure on $\mathbb{T}^m$. We begin by introducing the necessary terminology and notation for the definition of C.

In this part of the paper, we assume that $||\cdot||$ is the sup-norm, that is for a vector $z = (z_1, \dots, z_s) \in \mathbb{R}^s$, we have $||z|| = \max_{1\leqslant i\leqslant s} |z_i|$.

Recall that in our notation $d = m + n$. Suppose that the columns $\theta^1, \dots, \theta^n$ of our $m \times n$ matrix

$$A = (\theta^1 \cdots \theta^n)$$

are linearly independent over $\mathbb{Q}$ and consider the irrationality measure function

$$\Psi_A(t) = \min_{q \in \mathbb{Z}^n : 0 < ||q|| \leqslant t} \langle Aq \rangle,$$

which is a piecewise constant function decreasing to zero. Let

$$M_1 < M_2 < \dots < M_l < M_{l+1} < \dots$$

be the infinite sequence of all the points where $\Psi_A(t)$ is not continuous and

$$\zeta_l = \Psi_A(M_l), \quad \zeta_l > \zeta_{l+1}$$

be the corresponding sequence of values of the function. In particular,

$$\zeta_l = \min_{q \in \mathbb{Z}^n : 0 < ||q|| < M_{l+1}} \langle Aq \rangle.$$

The main properties of the function $\Psi_A(t)$ are discussed, for example, in [15]. From Minkowski convex body theorem, it follows that

$$\Delta_l := M_{l+1}^n \zeta_l^m \leqslant 1 \tag{8.1}$$

and A is singular if and only if $\Delta_l \to 0$ as $l \to \infty$.

It is well known (see, e.g., [3]) that the values of M_l grow exponentially, that is,

$$M_{l+3^d+1} \geqslant 2M_l \tag{8.2}$$

for all l.

We now have enough notation and terminology in order to define the Diophantine class of matrices for which our result holds.

Definition 8.1. Let $C \subset \mathrm{Mat}_{m\times n}(\mathbb{R})$ be the collection of matrices A such that there exists an increasing sequence l_k such that

(1) $\sum_{k=1}^{\infty} \Delta_{l_k}^{d-1} = \infty$;

(2) $H_k := \sup_{k_1 \geqslant k+1} \left(\frac{\zeta_{l_{k_1}}}{\Delta_{l_{k_1}}}\right)^m \cdot \left(\frac{M_{l_k+1}}{\Delta_{l_k}}\right)^n \to 0, \quad k \to \infty.$

Proposition 8.2. *If A is a nonsingular matrix, then A is in C.*

Proof. For nonsingular A, there exists a sequence l_k such that

$$\inf_k \Delta_{l_k} > 0. \tag{8.3}$$

So item (1) of Definition 8.1 holds for any subsequence of l_k. It also follows from (8.3) and monotonicity of ζ_l that we can choose a subsequence of l_k to satisfy

$$H_k \ll \zeta_{l_{k+1}}^m M_{l_k+1}^n = \left(\frac{M_{l_k+1}}{M_{l_{k+1}+1}}\right)^n \zeta_{l_{k+1}}^m M_{l_{k+1}+1}^n \leqslant \left(\frac{M_{l_k+1}}{M_{l_{k+1}+1}}\right)^n,$$

where we have used (8.1) in the last inequality. We can now take a subsequence of l_k and use the exponential growth of the sequence M_l to arrange that item (2) will hold. □

The following proposition shows that the class C strictly contains the class of nonsingular matrices.

Proposition 8.3. *There exists a singular matrix $A \in C$.*

Proof. It is well known, and may be easily deduced using the theory of templates, that for any function $\rho(l)$ that satisfies $\lim_{l\to\infty} \rho(l) = 0$, there exists a matrix A with $\Delta_l \asymp \rho(l)$. For our construction below, we choose A such that

$$\Delta_l \asymp l^{-\frac{1}{2(d-1)}}.$$

The fact that A is singular follows since $\Delta_l \to 0$. We now show that the lower bound $\Delta_l \gg l^{-\frac{1}{2(d-1)}}$ implies $A \in \mathcal{C}$.

Take $l_k = k^2$. Then, because of the lower bound, we have that $\sum_k \Delta_{l_k}^{d-1}$ diverges, so item (1) is satisfied. Moreover, the lower bound for Δ_l implies that for any $k_1 \geqslant k$, we have the inequality

$$\frac{1}{\Delta_{l_{k_1}}^n \Delta_{l_k}^m} \ll k_1^{\frac{d}{d-1}},$$

meanwhile from (8.1) and (8.2), it follows that

$$\zeta_{l_{k_1}}^m M_{l_k+1}^n \leqslant \frac{M_{k^2+1}^n}{M_{k_1^2+1}^n} \ll 2^{-\frac{n(k_1^2-k^2)}{3^d+1}}.$$

So

$$\begin{aligned} H_k &\leqslant \sup_{k_1 \geqslant k+1} k_1^{\frac{d}{d-1}} \cdot 2^{-\frac{n(k_1^2-k^2)}{3^d+1}} \\ &= \sup_{j \geqslant 1} (k+j)^{\frac{d}{d-1}} \cdot 2^{-\frac{n(2kj+j^2)}{3^d+1}} \\ &\ll k^{\frac{d}{d-1}} \cdot 2^{-\frac{2nk}{3^d+1}} \to 0, \text{ as } k \to \infty, \end{aligned}$$

and so item (2) is also valid. □

We now arrive to the main result of this section, which strengthens Theorem 1.6 and also shows that the fact that Bad_A is a null-set with respect to the m-dimensional Lebesgue measure does not characterize nonsingularity.

The proof of the following result is motivated by [13, Theorem 1.7] and is based on an application of Minkowski successive minima theory and modifies the proof from [18].

Theorem 8.4. *For any $A \in \mathcal{C}$, one has $\lambda_{\mathbb{T}^m}(\mathrm{Bad}_A) = 0$.*

Proof. Let A be any matrix and let Δ_l, ζ_l be as above. For $\eta = (\eta_1, \dots, \eta_m) \in \mathbb{R}^m$, let us denote

$$B_l(\eta) = \left[\eta_1, \eta_1 + \frac{2d\zeta_l}{\Delta_l}\right] \times \cdots \times \left[\eta_m, \eta_m + \frac{2d\zeta_l}{\Delta_l}\right] \subset \mathbb{R}^m. \tag{8.4}$$

We will need the following:

Lemma 8.5. *Let $R_l = \frac{dM_{l+1}}{\Delta_l}$. Then, for any η, the box $B_l(\eta)$ contains a point of the form $Aq - p$ with*

$$||q|| \leqslant R_l, \quad q \in \mathbb{Z}^n, \quad p \in \mathbb{Z}^m. \tag{8.5}$$

Proof. Fix l and consider parallelepiped

$$\Pi_l = \{z = (q, p) \in \mathbb{R}^d \ : \ ||q|| < M_{l+1}, ||Aq - p|| < \zeta_l\}.$$

Let $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_d$ be the successive minima for the parallelepiped Π_l with respect to the lattice $\mathbb{Z}^d$. As the volume of Π_l is equal to $2^d \Delta_l$, by Minkowski theorem, we know that

$$\lambda_1 \cdots \lambda_d \leqslant \Delta_l^{-1}.$$

But Π_l does not contain nontrivial integer points and so $\lambda_1 \geqslant 1$. We see that $\lambda_d \leqslant \Delta_l^{-1}$ and the closure of $\lambda_d \cdot \Pi_l$ contains d independent integer points. So, the closed parallelepiped

$$\widehat{\Pi}_l = \text{closure of } d\Delta_l^{-1}\Pi_l \supset d\lambda_d \Pi_l$$

contains a fundamental domain of the lattice $\mathbb{Z}^d$. This means that any shift

$$\widehat{\Pi}_l + w, \quad w \in \mathbb{R}^d \tag{8.6}$$

contains an integer point. But any parallelepiped

$$\left\{z = (q, p) \in \mathbb{R}^d \ : \quad ||q|| \leqslant R_l, \ \ Aq - p \in B_l(\eta)\right\}$$

is of the form (8.6), which proves the Lemma. □

We continue with the proof of Theorem 8.4 by following the argument from [18] and choosing the parameters there more carefully.

Assume now that the matrix A belongs to C and let H_k be as in Definition 8.1. Fix a positive $\varepsilon < \frac{1}{2}$ and define a sequence of positive reals $\phi_k < 1$ such that

$$\lim_{k \to \infty} \phi_k = 0, \tag{8.7}$$

$$\phi_k \geqslant d^d \left(\frac{3}{\varepsilon}\right)^m H_k, \tag{8.8}$$

and

$$\sum_{k=1}^{\infty} \phi_k \Delta_{l_k}^{d-1} = \infty. \tag{8.9}$$

This is possible because by item (2) of Definition 8.1 the right-hand side of (8.8) tends to zero, and as we have the divergence in item (1), ϕ_k can tend to zero slow enough to satisfy divergence condition (8.9). Define

$$W_k = \left\lceil \frac{\Delta_{l_k}}{2d\zeta_{l_k}} \right\rceil^m .$$

Note that as $\frac{M_{l_k+1}}{\Delta_{l_k}} \to \infty$ as $k \to \infty$, it follows from item (2) of Definition 8.1 that

$$\frac{\zeta_{l_k}}{\Delta_{l_k}} \to 0, \quad k \to \infty, \tag{8.10}$$

and so using (8.10), we see that $W_k \to +\infty$, as $k \to \infty$. We cover the cube $[0,1)^m$ by W_k boxes $B_{l_k}(\eta_{i_1,\dots,i_n})$ of the form (8.4) with

$$\eta_{i_1,\dots,i_n} = \left(\frac{i_1 d\zeta_{l_k}}{\Delta_{l_k}}, \dots, \frac{i_n d\zeta_{l_k}}{\Delta_{l_k}}\right), \quad 0 \leqslant i_1, \dots, i_n < \left[\frac{\Delta_{l_k}}{2d\zeta_{l_k}}\right], \quad i_1, \dots, i_n \in \mathbb{Z},$$

which have disjoint interiors.

We note that at least $W_k' = \left(\left[\frac{\Delta_{l_k}}{2d\zeta_{l_k}}\right] - 1\right)^m \sim W_k$ of these boxes are contained in the cube $[0,1)^m$. Moreover, Lemma 8.5 shows that in each of these boxes there is a point of the form $Aq - p$ satisfying (8.5) with $l = l_k$.

Put

$$\delta_k = \frac{\phi_k^{\frac{1}{m}}}{R_{l_k}^{\frac{n}{m}}} = \phi_k^{\frac{1}{m}} \cdot \left(\frac{\Delta_{l_k}}{dM_{l_k+1}}\right)^{\frac{n}{m}}.$$

It is clear that $\delta_k \leqslant \phi_k^{\frac{1}{m}} \cdot \frac{\zeta_{l_k}}{\Delta_{l_k}}$ because from (8.1) we deduce $\Delta_{l_k}^d \leqslant \Delta_{l_k} = M_{l_k+1}^n \zeta_{l_k}^m$, and this gives $\left(\frac{\Delta_{l_k}}{dM_{l_k+1}}\right)^{\frac{n}{m}} < \frac{\zeta_{l_k}}{\Delta_{l_k}}$.

As in the proof from [18], we take

$$W_k'' = \left[\frac{\left[\frac{\Delta_{l_k}}{2d\zeta_{l_k}}\right] - 1}{3}\right]^m \sim \frac{W_k'}{3^m} \sim \frac{W_k}{3^m}$$

boxes of the form

$$I_i^{[k]} = \left[\xi_{1,i}^{[k]} - \delta_k, \xi_{1,i}^{[k]} + \delta_k\right] \times \dots \times \left[\xi_{m,i}^{[k]} - \delta_k, \xi_{m,i}^{[k]} + \delta_k\right], \quad 1 \leqslant i \leqslant W_k'' \tag{8.11}$$

with the centers at certain points

$$\xi_i^{[k]} = (\xi_{1,i}^{[k]}, \dots, \xi_{m,i}^{[k]}) = Aq_i^{[k]} - p_i^{[k]}, \quad q_i^{[k]} \in \mathbb{Z}^n, \ p_i^{[k]} \in \mathbb{Z}^n, \quad |q_i^{[k]}| \leqslant R_{l_k}, \quad 1 \leqslant i \leqslant W_k'',$$

which belong to boxes

$$B_{l_k}(\eta_{i_1,\dots,i_n}) \quad \text{with} \quad i_1 \equiv \dots \equiv i_n \equiv 1 \pmod 3. \tag{8.12}$$

In each of such boxes $B_{l_k}(\eta_{i_1,\dots,i_n})$, we take just one point ξ_i. Then,

$$I_i^{[k]} \cap I_{i'}^{[k]} = \emptyset, \quad i \neq i'. \tag{8.13}$$

Consider two values $k_1 > k$. Recall that $I_i^{[k]}$ is an m-dimensional box with edge $2\delta_k$. The centers of the boxes $I_i^{[k_1]}$ by the construction are distributed in $I_i^{[k]}$ uniformly and so

$$\sharp\{i_1\ :\ 1 \leqslant i_1 \leqslant W''_{k_1},\ \ I_{i_1}^{[k_1]} \cap I_i^{[k]} \neq \emptyset\} \leqslant \left(2\delta_k \Big/ \left(6 \cdot \frac{\zeta_{l_{k_1}}}{\Delta_{l_{k_1}}}\right) + 1\right)^m. \tag{8.14}$$

We note that by condition (8.8) we have

$$1 \leqslant \varepsilon \cdot 2\delta_k \Big/ \left(6 \cdot \frac{d\zeta_{l_{k_1}}}{\Delta_{l_{k_1}}}\right) \tag{8.15}$$

because as $k_1 \geqslant k+1$, we have

$$\delta_k = \phi_k^{\frac{1}{m}} \cdot \left(\frac{\Delta_{l_k}}{dM_{l_k+1}}\right)^{\frac{n}{m}} \geqslant d^{\frac{d}{m}}\, \frac{3}{\varepsilon}\, \frac{\zeta_{l_{k_1}}}{\Delta_{l_{k_1}}}.$$

Recall that $\lambda(I_i^{[k]}) = (2\delta_k)^m$ and $W''_{k_1} \sim \frac{1}{3^m} \cdot \left(\frac{\Delta_{l_{k_1}}}{2d\zeta_{l_{k_1}}}\right)^m$. So for k large enough, we deduce from (8.14) and (8.15) an upper bound

$$\sharp\{i_1\ :\ 1 \leqslant i_1 \leqslant W''_{k_1},\ \ I_{i_1}^{[k_1]} \cap I_i^{[k]} \neq \emptyset\} \leqslant \lambda(I_i^{[k]}) W''_{k_1}(1 + 2^m\varepsilon). \tag{8.16}$$

Now we consider the union

$$E_k = \bigcup_{i=1}^{W''_k} I_i^{[k]}.$$

By (8.13), we get

$$\lambda(E_k) = \sum_i \lambda(I_i^{[k]}) \asymp \delta_k^m W_k \asymp \phi_k \Delta_{l_k}^{d-1}$$

and so by (8.9), we see that

$$\sum_{k=1}^{\infty} \lambda(E_k) = \infty. \tag{8.17}$$

Then from (8.16) for $k_1 > k$, we have

$$\lambda(E_k \cap E_{k_1}) \leqslant (1 + 2^m\varepsilon)\lambda(E_k)\lambda(E_{k_1}). \tag{8.18}$$

It follows from [10, Theorem 18.10] that for the set

$$E = \{\eta\ :\ \exists \text{ infinitely many } k \text{ such that } \ \eta \in E_k\},$$

we have

$$\lambda(E) \geqslant \limsup_{t\to\infty} \frac{\left(\sum_{k=1}^t \lambda(E_k)\right)^2}{\sum_{k,k_1=1}^t \lambda(E_k \cap E_{k_1})} \geqslant \frac{1}{1 + 2^m\varepsilon},$$

by (8.17) and (8.18). As ε is arbitrary, we see that $\lambda(E) = 1$.

Finally, if $\eta \in E$, then there exist infinitely many k such that for each k, there exists $q \in \mathbb{Z}^n$ with $||q|| \leqslant R_{l_k}$ and $p \in \mathbb{Z}^n$ such that $||Aq - p - \eta|| \leqslant \delta_k$. In particular, from the definition of R_{l_k} for these (p, q), we get

$$||Aq - p - \eta|| \cdot ||q||^{\frac{n}{m}} \leqslant \delta_k R_{l_k}^{\frac{n}{m}} = \phi_k^{\frac{1}{m}}.$$

We see that condition (8.7) leads to the inclusion $\mathrm{Bad}_A \cap [0,1)^m \subset [0,1)^m \setminus E$, which finishes the proof. □

APPENDIX A: ROY'S PARAMETRIC GEOMETRY OF NUMBERS

As it was mentioned before, the existence results from Propositions 7.3 and 8.3 as well as the main result from the paper [14] can be deduced by means of existence results from parametric geometry of numbers obtained in [20] and [7]. In particular, here in the Appendix, we show how one can get Proposition 7.3.

Definition A.1 (Log-minima functions). Let $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$. For $Q \in \mathbb{R}$ with $Q \geqslant 1$, we define the convex body

$$\mathcal{C}(Q) := \left\{ v = (v_1, v_2, v_3) \in \mathbb{R}^3 \,:\, \|v\| \leqslant 1,\ |v_1\theta_1 + v_2\theta_2 + v_3| \leqslant Q^{-1} \right\}.$$

For $j = 1, 2, 3$, we define the successive minima functions $\lambda_j \,:\, [1, \infty) \to \mathbb{R}$:

$$\lambda_j(Q) := \min\left\{ r \in \mathbb{R} \,:\, r\mathcal{C}(Q) \text{ contains } j \text{ independent vectors of } \mathbb{Z}^3 \right\}$$

and the log-minima functions $L_j \,:\, [0, \infty) \to \mathbb{R}$:

$$L_j(q) = \log(\lambda_j(e^q)).$$

The relation between the first log-minima function and the irrationality measure function for a vector $\theta \in \mathbb{R}^2$ is given by the following proposition.

Proposition A.2. *Let $\theta \in \mathbb{R}^2$. Let L_1 be the first associated log-minima function and let Ψ be the associated irrationality measure function. If, for some constants $0 < A < 1$ and $0 < B$, the equation*

$$L_1(q) > Aq - B \tag{A.1}$$

is satisfied for an unbounded set of $q \in \mathbb{R}_{\geqslant 0}$, then the equation

$$\Psi(t) \geqslant e^{\frac{-B - \ln(1+\|\theta\|)}{A}} t^{1-\frac{1}{A}}$$

is satisfied for an unbounded set of $t \in \mathbb{R}_{\geqslant 0}$.

Proof. If, for some $q > 0$, (A.1) holds, then from the definitions we have

$$\left\{ v \in \mathbb{R}^3 \,:\, \|v\| \leqslant e^{Aq-B},\ |v_1\theta_1 + v_2\theta_2 + v_3| \leqslant e^{Aq-B}e^{-q} \right\} \cap \mathbb{Z}^3 = \{0\}. \tag{A.2}$$

Now let $C = \ln(1 + \|\theta\|)$. Suppose, by way of contradiction,

$$\Psi(e^{Aq-B-C}) < e^{Aq-B-C}e^{-q}.$$

This gives the existence of $a = (a_1, a_2) \in \mathbb{Z}^2$ with $0 < \|a\| \leqslant e^{Aq-B-C}$ and some $b \in \mathbb{Z}$ for which

$$|a_1\theta_1 + a_2\theta_2 + b| < e^{Aq-B-C}e^{-q}. \tag{A.3}$$

Thus,

$$\begin{aligned} |b| &= |b + (a_1\theta_1 + a_2\theta_2) - (a_1\theta_1 + a_2\theta_2)| \\ &< e^{Aq-B-C}e^{-q} + \|\theta\|e^{Aq-B-C} \\ &\leqslant e^{Aq-B}e^{-C}(1 + \|\theta\|) = e^{Aq-B}. \end{aligned} \tag{A.4}$$

Since $C \geqslant 0$, we have

$$\max\{|a_1|, |a_1|, |b|\} \leqslant e^{Aq-B} \;\text{ and }\; |a_1\theta_1 + a_2\theta_2 + b| \leqslant e^{Aq-B}e^{-q},$$

which contradicts (A.2). Thus, we must have

$$\Psi(e^{Aq-B-C}) \geqslant e^{Aq-B-C}e^{-q}.$$

If we substitute $t = e^{Aq-B-C}$, we get

$$\Psi(t) \geqslant e^{\frac{-B-C}{A}}t^{1-\frac{1}{A}},$$

which is the required result. □

Proposition A.3. *Let $\theta \in \mathbb{R}^2$, let L_1 be the associated first log-minima function, and let Ψ be the associated irrationality measure function. Suppose, for some constants $0 < A < 1$ and $0 < B$, we have*

$$L_1(q) \leqslant Aq + B \textit{ for all sufficiently large } q \in \mathbb{R}_{\geqslant 0}. \tag{A.5}$$

Then, the irrationality measure function satisfies

$$\Psi(t) \leqslant e^{B/A}t^{1-\frac{1}{A}} \textit{ for all } t \in \mathbb{R}_{\geqslant 0} \textit{ sufficiently large.}$$

Proof. This is again a change of variables. Equation (A.5) implies that, for all large q,

$$e^{Aq+B}C(e^q) \cap \mathbb{Z}^3 \neq \{0\}.$$

Unwinding the definitions, we obtain

$$\Psi(e^{Aq+B}) \leqslant e^{Aq+B}e^{-q} \text{ for all sufficiently large } q.$$

Putting $t = e^{Aq+B}$, we get the required conclusion. □

The log-minima functions (L_1, L_2, L_3) are modeled by the following functions.

Definition A.4 [20], Definition 4.1 of n-systems. Let $I \subset [0, \infty)$ be a subinterval. A 3-system on I is a continuous piecewise linear map $P = (P_1, P_2, P_3) : I \to \mathbb{R}^3$ such that

(S1) for each $q \in I$, we have

$$0 \leqslant P_1(q) \leqslant P_2(q) \leqslant P_3(q) \text{ and } P_1(q) + P_2(q) + P_3(q) = q;$$

(S2) if H is a nonempty open subset on which P is differentiable, then there is an integer r with $1 \leqslant r \leqslant 3$ such that P_r has slope 1 on H while the other components of P are constant;

(S3) if q is an interior point of I where P is not differentiable and if the integers r, s for which $P_r'(q^-) = P_s'(q^+) = 1$ also satisfy $r < s$, then we have

$$P_r(q) = P_{r+1}(q) = \cdots = P_s(q).$$

Remark A.5. Piecewise linear means the points of I where P is not differentiable has discrete closure in $\mathbb{R}$ and P is linear on each connected component of the complement (in I) of the nondifferentiable points. The notation $P_r'(q^-)$ and $P_s'(q^+)$ denote the left and right derivatives, respectively (assuming q is not in the boundary of I).

We have the following fundamental theorem.

Theorem A.6 [20] Theorem 4.2. *For each 3-system P on an interval $[q_0, \infty)$, there is a $\theta \in \mathbb{R}^2$ such that, considering the log-minima function $L = (L_1, L_2, L_3) : [q_0, \infty) \to \mathbb{R}^3$ associated to θ, we have*

$$L - P \text{ is bounded on } [q_0, \infty).$$

We are finally ready for the following:

Proof of Proposition 7.3. Let $Q \in \mathbb{R}$ satisfy $Q > 2$. Consider the following graphs of three piecewise linear functions on the interval $[1, Q-1]$ (see Figure A.1).

Here is the precise definition of the function $P : [1, Q-1] \to \mathbb{R}^3$:

$$(P_1(q), P_2(q), P_3(q)) := \begin{cases} \left(\frac{1}{Q+1}, \frac{1}{Q+1} + q - 1, \frac{Q-1}{Q+1}\right) & ; 1 \leqslant q \leqslant \frac{2Q-1}{Q+1} \\ \left(\frac{1}{Q+1}, \frac{Q-1}{Q+1}, \frac{Q-1}{Q+1} + q - \frac{2Q-1}{Q+1}\right) & ; \frac{2Q-1}{Q+1} \leqslant q \leqslant \frac{Q^2-Q+1}{Q+1} \\ \left(\frac{1}{Q+1} + q - \frac{Q^2-Q+1}{Q+1}, \frac{Q-1}{Q+1}, \frac{(Q-1)^2}{Q+1}\right) & ; \frac{Q^2-Q+1}{Q+1} \leqslant q \leqslant Q-1 \end{cases}. \tag{A.6}$$

We leave it to the reader to check that P satisfies Definition A.4 on $[1, Q-1]$. We now extend this function to a 3-system on all of $[1, \infty)$; We let $I_0 := [1, Q-1]$. For each $l \in \mathbb{N}$, we inductively define the interval I_l as the closed interval of length $(Q-1)^l \times \text{length}(I_0)$ that is contiguous and to the right of I_{l-1}. We let f_l be the linear, orientation-preserving bijection from I_l to I_0, and define

$$\widetilde{P}(q) := (Q-1)^l P(f_l(q)) \text{ for } q \in I_l. \tag{A.7}$$

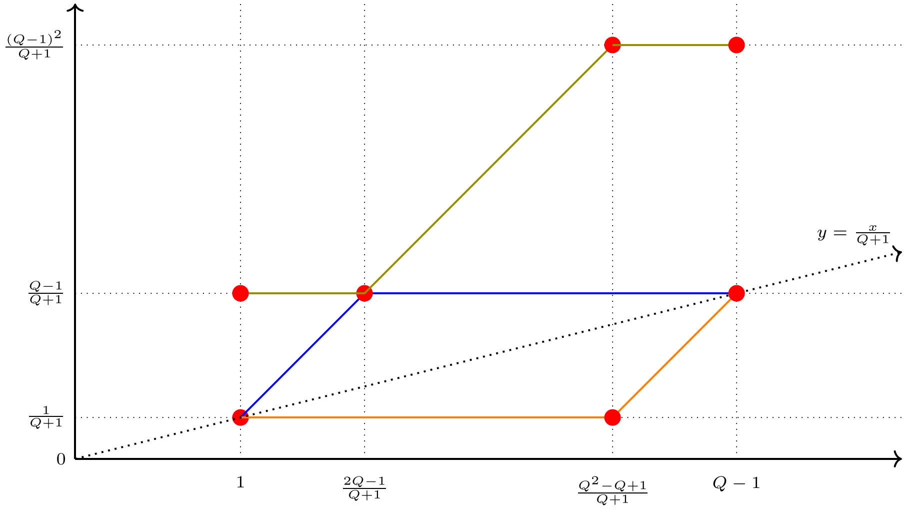

FIGURE A.1 The orange, blue and olive graphs are called P_1, P_2, P_3 respectively.

Note that $\widetilde{P}$ is well defined: For if we have $q \in I_{l-1} \cap I_l$,

$$\begin{aligned}
(Q-1)^l P(f_l(q)) &= (Q-1)^l P(1) \\
&= (Q-1)^l \left(\frac{1}{Q+1}, \frac{1}{Q+1}, \frac{Q-1}{Q+1} \right) \\
&= (Q-1)^{l-1} \left(\frac{Q-1}{Q+1}, \frac{Q-1}{Q+1}, \frac{(Q-1)^2}{Q+1} \right) \\
&= (Q-1)^{l-1} P(Q-1) \\
&= (Q-1)^{l-1} P(f_{l-1}(q)).
\end{aligned}$$

We have the following straightforward claim, which we write out completely. The reader might prefer to do it themselves.

Claim A.6.1. The function $\widetilde{P} = (\widetilde{P}_i)_{i=1,2,3} : [1, \infty) \to \mathbb{R}$ is a 3-system with

(a) $\widetilde{P}_1(q) \leqslant \frac{q}{Q+1}$ for all $q \geqslant 1$;
(b) if q is the right endpoint of I_n, then $\widetilde{P}_1(q) = \frac{q}{Q+1}$.

Proof. $\widetilde{P}$ is continuous and piecewise linear since it is so on each interval I_l. Moreover, the maps $f_l : I_l \to I_0$ have slopes $(Q-1)^{-l}$, which cancels with the scaling factor $(Q-1)^l$ of formula (A.7). Thus, Definition A.4(S1, S2) become clear. We leave it to the reader to check Definition A.4(S3) at the end points of I_l for each $l \in \mathbb{N}$. Thus, $\widetilde{P}$ is a 3-system.

For parts (a) and (b), we first prove that for each $l \in \mathbb{N} \cup \{0\}$, we have

$$I_l = \left[(Q-1)^l, (Q-1)^{l+1} \right].$$

This is true for I_0 by definition. We let $l \geqslant 1$ and write $I_l = [a_n, b_n]$. We compute, using induction, that

$$a_l = a_{l-1} + (Q-1)^{l-1}\,\mathrm{length}(I_0) = (Q-1)^{l-1} + (Q-1)^{l-1}(Q-2) = (Q-1)^l.$$

Similarly, for b_n, we have

$$b_l = b_{l-1} + (Q-1)^l(Q-2) = (Q-1)^l + (Q-1)^l(Q-2) = (Q-1)^{l+1}.$$

Part (a) for $q \in I_0$ is clear from formula (A.6). For $l \in \mathbb{N}$ and $q \in I_l = [a_l, b_l]$, we can then compute

$$\begin{aligned}\widetilde{P}_1(q) &= (Q-1)^l P_1(f_l(q)) \\ &\leqslant (Q-1)^l \frac{f_l(q)}{Q+1}.\end{aligned}$$

Now $q \mapsto (Q-1)^l f_l(q)$ is a linear function of slope 1 passing through the point

$$(a_l, (Q-1)^l) = ((Q-1)^l, (Q-1)^l).$$

Thus, $(Q-1)^l f_l(q) = q$ on the domain of f_l and so we have $\widetilde{P}_1(q) \leqslant q(Q+1)^{-1}$.

For part (b), we have

$$\widetilde{P}_1(b_l) = (Q-1)^l P_1(f_l(b_l)) = (Q-1)^l \frac{Q-1}{Q+1} = \frac{b_l}{Q+1}.$$

□

We now apply Theorem A.6 to the 3-system $\widetilde{P}$ and obtain a $\theta \in \mathbb{R}^2$ with a constant $B > 0$ such that

$$|L_1(q) - \widetilde{P}_1(q)| \leqslant B \text{ for all } q \in [1, \infty).$$

Here L_1 is the first log-minima function associated to θ as in Definition A.1. Claim A.6.1 then shows

$$L_1(q) \leqslant \frac{q}{Q+1} + B \text{ for all } q \in [1, \infty)$$

and that

$$-B - 1 + \frac{q}{Q+1} < L_1(q) \text{ for an unbounded set of } q \in [1, \infty).$$

Applying Propositions A.2 and A.3, we see the existence of constants C_1, C_2 for which

$$\Psi(t) \leqslant C_1 t^{-Q} \text{ for all sufficiently large } t \geqslant 1$$

and

$$C_2 t^{-Q} \leqslant \Psi(t) \text{ for an unbounded set of } t \geqslant 1.$$

Here, Ψ is the irrationality measure function associated to θ. □

ACKNOWLEDGMENTS

This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program, Grant Agreement Number 754475. N. Moshchevitin was also supported by the Austrian Science Fund FWF, Project I 5554 "Diophantine approximation, arithmetic sequences & analytic number theory" (joint project with RSF project 22-41-05001). An essential part of this work was done during the first and second author's stay at Israel Institute of Technology (Technion). They thank the people from Technion for the extremely friendly atmosphere and wonderful opportunities for work. We thank the referee for a careful report, which helped us improve the writing.

JOURNAL INFORMATION

Mathematika is owned by University College London and published by the London Mathematical Society. All surplus income from the publication of *Mathematika* is returned to mathematicians and mathematics research via the Society's research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

REFERENCES

1. V. Beresnevich, S. Datta, A. Ghosh, and B. Ward, *Rectangular shrinking targets for $\mathbb{Z}^m$ actions on tori: well and badly approximable systems*, 2023, https://arxiv.org/abs/2307.10122
2. Y. Bugeaud, S. Harrap, S. Kristensen, and S. Velani, *On shrinking targets for $\mathbb{Z}^m$ actions on tori*, Mathematika **56** (2010), no. 2, 193–202.
3. Y. Bugeaud and M. Laurent, *On exponents of homogeneous and inhomogeneous Diophantine approximation*, Mosc. Math. J. **5** (2005), no. 4, 747–766.
4. J. W. S. Cassels, *An introduction to Diophantine approximation*, Cambridge Tracts in Mathematics and Physics, vol. 35, Cambridge University Press, Cambridge, 1957.
5. S. G. Dani, *Divergent trajectories of flows on homogeneous spaces and Diophantine approximation*,J. Reine Angew. Math. **359**, 55–89, 1985.
6. H. Davenport, *Indefinite binary quadratic forms, and Euclid's algorithm in real quadratic fields*, Proc. Lond. Math. Soc. (2) **53**, 65–82, 1951.
7. T. Das, L. Fishman, D. Simmons, and M. Urbański, *A variational principle in the parametric geometry of numbers*, Adv. Math. **437**, 109435, 2024.
8. M. Einsiedler and J. Tseng, *Badly approximable systems of affine forms, fractals, and Schmidt games*, J. Reine Angew. Math. **660**, 83–97, 2011.
9. M. Einsiedler and T. Ward, *Ergodic theory with a view towards number theory*, Graduate Texts in Mathematics, vol. 259, Springer, London, 2011.
10. A. Gut, *Probability: a graduate course*, Springer Texts in Statistics, Springer, New York, 2014.
11. V. Jarník, *Sur les approximations diophantiques linéaires non homogénes*, Bull. internat. de l'Académie des sciences de Bohême, 1946, 145–160. https://dml.cz/handle/10338.dmlcz/500769?show=full
12. D. H. Kim, *The shrinking target property of irrational rotations*, Nonlinearity **20** (2007), 1637–1644.
13. T. Kim, *On a kurzweil type theorem via ubiquity*, Acta Arithmetica **213** (2024), 181–191.
14. G. Lachman, A. Rao, U. Shapira, and Y. Yifrach, *k-divergent lattices*, 2023, https://arxiv.org/abs/2307.09054
15. N. G. Moshchevitin, *Singular Diophantine systems of A. Ya. Khinchin and their application*, Uspekhi Mat. Nauk **65** (2010), no. 3, 43–126.
16. N. Moshchevitin, *A note on badly approximable affine forms and winning sets*, Mosc. Math. J. **11** (2011), no. 1, 129–137.
17. N. G. Moshchevitin, *Positive integers: counterexample to W. M. Schmidt's conjecture*, Mosc. J. Comb. Number Theory **2** (2012), no. 2, 63–84.
18. N. Moshchevitin, *A note on well distributed sequences*, Unif. Distrib. Theory **18** (2023), 141–146.
19. D. Roy, *Diophantine approximation with sign constraints*, Monatsh. Math. **173** (2014), no. 3, 417–432.
20. D. Roy, *Spectrum of the exponents of best rational approximation*, Math Z. **283** (2016), no. 1, 143–155.

21. W. Rudin, *Real and complex analysis*, 3rd ed., McGraw-Hill Book Co., New York, 1987.
22. W. Rudin, *Functional analysis*, 2nd ed., McGraw-Hill Book Co., New York, 1991.
23. W. M. Schmidt, *badly approximable numbers and certain games*, Trans. Amer. Math. Soc. **123** (1966), 178–199.
24. W. M. Schmidt, *Two questions in Diophantine approximation*, Monatsh. Math. **82** (1976), no. 3, 237–245.
25. W. M. Schmidt, *Diophantine approximation*, Springer lecture notes in mathematics, vol. 785, Springer-Verlag, New York, 1980.
26. U. Shapira, *Grids with dense values*, Comment. Math. Helv. **88** (2013), no. 2, 485–506.
27. P. Thurnheer, *Zur diophantischen approximation von zwei reellen Zahlen*, Acta Arith. **44** (1984), no. 3, 201–206.
28. P. Thurnheer, *On Dirichlet's theorem concerning Diophantine approximation*, Acta Arith. **54** (1990), no. 3, 241–250.