

Diss. ETH No. 22,060

EXCEPTIONAL ORBITS IN HOMOGENEOUS
SPACES

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES OF ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

BEVERLY LYTLE

M. Sc. Math. Ohio State University

born the 6th of August, 1984

citizen of the United States of America

Accepted on the recommendation of:

PROF. MANFRED EINSIEDLER, examiner

DR. ANISH GHOSH, co-examiner

2014

Diss. ETH No. 22,060

EXCEPTIONAL ORBITS IN HOMOGENEOUS
SPACES

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES OF ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

BEVERLY LYTLE

M. Sc. Math. Ohio State University

born the 6th of August, 1984

citizen of the United States of America

Accepted on the recommendation of:

PROF. MANFRED EINSIEDLER, examiner

DR. ANISH GHOSH, co-examiner

2014

ABSTRACT. We begin with an introduction to and a brief history of the interactions of homogeneous dynamics and number theory. Three self-contained, yet related articles follow. In the first, we show that the set of points on C^1 curves which are badly approximable by rationals in a number field form a winning set in the sense of W. Schmidt. As a consequence, we obtain a number field version of Schmidt's conjecture in Diophantine approximation. Within the second article, we show that for a fixed vector, the set of inhomogeneous linear forms which are badly approximable by the rationals of a fixed number field form a winning set (in the sense of Schmidt), even when restricted to a fractal. In addition, we show an analogous result for a fixed matrix rather than a fixed vector. Now, let S and T be hyperbolic endomorphisms of \mathbb{T}^d with the property that the span of the maximal subspace contracted by S along with the maximal subspace contracted by T is \mathbb{R}^d . In the third article, we show that the Hausdorff dimension of the intersection of the set of points with equidistributing orbits under S with the set of points with nondense orbit under T is full. In the case that S and T are quasihyperbolic automorphisms, we prove that the Hausdorff dimension of the intersection is again full when we assume that \mathbb{R}^d is spanned by the maximal subspaces contracted by S and T along with the central eigenspaces of S and T .

ZUSAMMENFASSUNG. Zunächst geben wir eine Einführung und einen kurzen historischen Überblick der Interaktionen zwischen homogener Dynamik und Zahlentheorie. Danach folgen drei in sich geschlossene Artikel, die miteinander in Verbindung stehen. Im ersten zeigen wir, dass auf C^1 -Kurven die Menge der Punkte, die schlecht durch rationale Zahlen in einem (gegebenen) Zahlkörper approximierbar sind, eine “winning” Menge im Sinne von W. Schmidt bildet. Daraus erhalten wir eine Variante der Schmidtschen Vermutung für Zahlkörper, einer Vermutung in der Diophantischen Approximation. Im zweiten Artikel zeigen wir, dass für einen festen Vektor die Menge inhomogener Linearformen, die schlecht durch rationale Zahlen eines festen Zahlkörpers approximierbar sind, eine “winning” Menge (im Sinne von Schmidt) bilden, auch wenn sie auf ein Fraktal eingeschränkt werden. Weiters zeigen wir ein analoges Resultat, wenn man den festen Vektor mit einer festen Matrix ersetzt. Seien nun S und T hyperbolische Endomorphismen von \mathbb{T}^d mit der Eigenschaft, dass die lineare Hülle der Vereinigung des kontrahierenden Hauptraumes von S und des kontrahierenden Hauptraumes von T ganz \mathbb{R}^d ist. Im dritten Artikel zeigen wir, dass die Schnittmenge der Menge der Punkte mit gleichverteilenden Bahnen unter S mit der Menge der Punkte mit nicht-dichten Bahnen unter T volle Hausdorff-Dimension hat. Im Fall quasi-hyperbolischer invertierbarer S und T zeigen wir, dass die Schnittmenge volle Hausdorff-Dimension hat, unter der Annahme, dass \mathbb{R}^d durch die Vereinigung der kontrahierenden Haupträume von S und T und den zentralen Eigenräumen von S und T aufgespannt wird.

Contents

Chapter 1. Introduction	3
1. Diophantine approximation	3
2. Homogeneous dynamics and applications	10
3. A detailed example	15
4. Exceptional orbits for multiple transformations	18
Chapter 2. Badly approximable vectors, C^1 curves and number fields	21
1. Introduction	21
2. Mahler's Compactness Criterion	25
3. Dani's Correspondence	28
4. A special case	30
5. Proof in the general case	34
Chapter 3. Badly approximable affine forms, fractals and number fields	37
1. Introduction	37
2. Statement of results	39
3. Background	41
4. Proof of fixed matrix case	45
5. Proof of homogeneous fixed vector case	50
6. Proof of inhomogeneous fixed vector case	52
Chapter 4. Simultaneous dense and nondense orbits for noncommuting toral endomorphisms	55
1. Introduction	55
2. Preliminaries	57
3. Hyperbolic toral endomorphisms	60
4. Quasihyperbolic toral endomorphisms	65
Bibliography	73

CHAPTER 1

Introduction

In this dissertation, dynamical methods are used to prove results in diophantine approximation. This introduction serves to familiarize the reader, first, with a brief history of diophantine approximation, including various recent results. The second section will be an overview of dynamics, in particular, of homogeneous dynamics. The third section will provide an explicit example to show how dynamical methods can be employed to achieve interesting results in diophantine approximation. The final section will concern homogeneous dynamics on compact spaces and exceptional orbits therein. There are three chapters following the introduction containing original content, each chapter with a self-contained article in preparation for submission for publication in scholarly journals outside of this thesis.

1. Diophantine approximation

The area of mathematical study known as diophantine approximation takes its name from the Alexandrian mathematician Diophantus. In his *Arithmetica*, he studied particular and approximate solutions to equations with integer coefficients. Modern treatment of diophantine approximation is generally considered to have begun with Dirichlet's Theorem (1842) on approximation of real numbers by rationals [20].

THEOREM 1.1 (Dirichlet's Theorem). *Let α be a real number and N a positive integer. Then there exist integers p and q with $1 \leq q \leq N$ and*

$$|q\alpha - p| \leq \frac{1}{N}.$$

Hence, there are infinitely many pairs p and q satisfying

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{q^2}.$$

The proof of this theorem is an early instance of the use of the pigeonhole principle.

PROOF. Partition $[0, 1)$ into the $N + 1$ subintervals $[\frac{i}{N+1}, \frac{i+1}{N+1})$ for $0 \leq i \leq N$. Consider the collection $q\alpha - [q\alpha]$ for $1 \leq q \leq N$ where

$\lfloor x \rfloor$ denotes the integer part of x . If one of these numbers lies in one of the intervals $[0, \frac{1}{N+1})$ or $[\frac{N}{N+1}, 1)$, then we have found the desired approximation. Suppose this is not the case. Then the N elements $q\alpha - \lfloor q\alpha \rfloor$ for $1 \leq q \leq N$ lie in the $N-1$ sets $[\frac{i}{N+1}, \frac{i+1}{N+1})$ for $1 \leq i \leq N-1$. By the pigeonhole principle, one of these sets must contain two points, say $q_1\alpha - \lfloor q_1\alpha \rfloor$ and $q_2\alpha - \lfloor q_2\alpha \rfloor$. Then

$$|(q_1\alpha - \lfloor q_1\alpha \rfloor) - (q_2\alpha - \lfloor q_2\alpha \rfloor)| = |(q_1 - q_2)\alpha - (\lfloor q_1\alpha \rfloor - \lfloor q_2\alpha \rfloor)| \leq \frac{1}{N+1},$$

which completes the proof. \square

While it is clear that every number can be approximated by rationals, Dirichlet's Theorem is a first step toward quantifying the rate of approximation. A natural question to follow is asking whether this rate is optimal, that is, can the exponent 2 be replaced with something larger? Real numbers α with infinitely many integral solutions to $|\alpha - \frac{p}{q}| < \frac{1}{q^\tau}$ for $\tau > 2$ are called very well approximable. Liouville started investigating such properties of algebraic numbers. This line of thought was pursued by Thue, Siegel, and Roth culminating in what is now known as Roth's Theorem (for which he won a Fields Medal in 1958) [58].

THEOREM 1.2 (Roth's Theorem). *Let α be an irrational algebraic integer and let $\varepsilon > 0$. Then the inequality*

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{q^{2+\varepsilon}}$$

has only finitely many coprime integer solutions p and q . Hence there exists $c(\alpha) > 0$ such that

$$\left| \alpha - \frac{p}{q} \right| > \frac{c}{q^{2+\varepsilon}}$$

for all integers p and $q \neq 0$.

We need not restrict ourselves to approximations with quality measured by the function q^{-2} . Let $\psi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be a non-increasing function and define

$$W(\psi) =$$

$$\left\{ \alpha \in \mathbb{R} \mid \left| \alpha - \frac{p}{q} \right| < \psi(q) \text{ for infinitely many integers } p \text{ and } q \neq 0 \right\}.$$

This theorem on approximation relative to ψ is due to Khintchine [39]:

THEOREM 1.3 (Khintchine's Theorem). *Let $\psi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be a non-increasing function. Then for the Lebesgue measure m*

$$m(W(\psi)) = \begin{cases} 0 & \text{if } \sum_q q\psi(q) < \infty \\ 1 & \text{if } \sum_q q\psi(q) = \infty \end{cases}.$$

Jarnik proved a similar statement, but this time involving s -dimensional Hausdorff measure \mathcal{H}^s [38].

THEOREM 1.4 (Jarnik's Theorem). *Let $s \in (0, 1)$ and let $\psi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be a non-increasing function. Then*

$$\mathcal{H}^s(W(\psi)) = \begin{cases} 0 & \text{if } \sum_q q\psi(q)^s < \infty \\ 1 & \text{if } \sum_q q\psi(q)^s = \infty \end{cases}.$$

Another direction one might take is to focus not the exponent, but rather the constant 1 appearing in Dirichlet's theorem. Take for example Hurwitz' Theorem [37].

THEOREM 1.5 (Hurwitz' Theorem). *For any real number α there exists infinitely many pairs p and q satisfying*

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{\sqrt{5}q^2}.$$

Furthermore, the constant $1/\sqrt{5}$ is the smallest possible for such a statement to hold for all real numbers α .

All of these theorems tell us how close certain points are to the rationals, but what about points which are far away? To wit, a number α is called badly approximable if there exists a constant $c(\alpha) > 0$ such that for every pair of integers p and $q \neq 0$,

$$\left| \alpha - \frac{p}{q} \right| > \frac{c}{q^2}.$$

There is a tight connection between diophantine properties of numbers and the theory of continued fractions. Recall that the continued fraction expansion for a real number α is given by

$$\alpha = [a_0 : a_1, a_2, \dots] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots}}},$$

and that if α is irrational, this expansion is infinite and unique. It can be shown that a number α is badly approximable if and only if the terms a_i appearing in the continued fraction expansion are bounded from above. (We note also that the partial quotients of the continued

fraction contain the best approximations of α by rationals.) Immediately, we have that there are uncountably many badly approximable numbers, and that every quadratic irrational is badly approximable (since quadratic irrationals are precisely the numbers with periodic continued fraction expansions). It can also be deduced that the set of badly approximable numbers has Lebesgue measure 0. (For more on this topic, see, for example, [62].)

Jarnik went further and showed, in contrast, that the set of badly approximable numbers has Hausdorff dimension 1. Schmidt reproved this result by showing something stronger, that the set of badly approximable numbers is winning. To define winning, we must first introduce Schmidt games.

Let X be a complete metric space. Let α and β lie in $(0, 1)$ and let $S \subset X$. Two players, Alice and Bob, compete in the (α, β) game. Bob begins by choosing $B_0 = B(x_0, \rho) \subset X$, the closed ball of radius ρ around x_0 . In the n th round of the game, Alice chooses a center point y_n so that $A_n := B(y_n, \rho\alpha(\alpha\beta)^{n-1})$ is contained in B_{n-1} . Then Bob chooses x_n so that $B_n := B(x_n, \rho(\alpha\beta)^n)$ is contained in A_n . Alice and Bob continue alternating choosing center points ad nauseam. At the “end” of the game, by the assumptions on α and β , there is one point x_∞ left in $\bigcap_n B_n$. If $x_\infty \in S$, Alice has won the game. Otherwise, Bob wins. If there exists α so that Alice finds a strategy to win every (α, β) game for any $\beta > 0$, then the set S is called α -winning. Often we omit α and simply refer to S as winning.

Schmidt introduced this game in [59], and proved various properties of winning sets. For example, if $X = \mathbb{R}^d$, then the Hausdorff dimension of any winning set in \mathbb{R}^d is d . Moreover, given a countable collection of sets S_i with S_i α_i -winning and $\inf_i \alpha_i = \alpha_0 > 0$, then $\bigcap_i S_i$ is α_0 -winning. This demonstrates that the property of winning is much stronger than that of having full Hausdorff dimension. Schmidt went on to prove that the set of badly approximable numbers is a winning set in \mathbb{R} . This statement will be re-proven in Section 3, but using dynamical structures, in order to give the basic ideas for the proofs which appear in later chapters.

So far, the discussion has been limited to the real line. Despite the utility of continued fractions in the one dimensional setting, this is not a natural obstruction. Indeed, one may study the approximation theory of vectors, or more generally linear forms. Below is the analogue of Dirichlet’s Theorem in higher dimension.

THEOREM 1.6. *Let $L = (l_{i,j})$ be a collection of n linear forms in m variables. Then there are infinitely many $p \in \mathbb{Z}^n$ and $q \in \mathbb{Z}^m$ with*

$q \neq 0$ satisfying

$$\|L(q) - p\|^n \leq \frac{1}{\|q\|^m},$$

where $\|\cdot\|$ denotes the maximum norm.

From this one can easily define a badly approximable linear form. There are results measuring the size of the set of badly approximable forms, and also results similar to those of Hurwitz. Further, Groshev proved a generalization of Khintchine's Theorem to the set of linear forms [35].

Before moving to versions of metric diophantine approximation in more exotic spaces than \mathbb{R}^n , let us also mention several results regarding "mixed" approximation. These types of approximations involve simultaneous approximation of a vector, perhaps involving weights in each coordinate. Originally a conjecture by Schmidt, Badziahin, Pollington and Velani recently proved the following result [5]:

THEOREM 1.7. *For any $i, j \geq 0$ with $i + j = 1$, let $\text{Bad}(i, j)$ denote the set of points $(x, y) \in \mathbb{R}^2$ for which there exists a constant c such that $\max\{|qx - p_1|^{1/i}, |qy - p_2|^{1/j}\} > \frac{c}{q}$ for all $q, p_1, p_2 \in \mathbb{N}$. Then the intersection of any finite collection $\text{Bad}(i_k, j_k)$ has full Hausdorff dimension.*

The proof, while involved, uses elementary technology. This is not true of the most recent development toward the related Littlewood Conjecture:

CONJECTURE 1.8 (Littlewood's Conjecture). *For all pairs $(x, y) \in \mathbb{R}^2$*

$$\liminf_{q \rightarrow \infty} q \|qx\|_{\mathbb{Z}} \|qy\|_{\mathbb{Z}} = 0,$$

where $\|\cdot\|_{\mathbb{Z}}$ denotes distance to nearest integer.

Had one been able to find a counterexample to the conjecture of Schmidt, that is, that there exist two pairs (i_1, j_1) and (i_2, j_2) with $i_1 + j_1 = i_2 + j_2 = 1$ and $\text{Bad}(i_1, j_1) \cap \text{Bad}(i_2, j_2) = \emptyset$, this would have provided a proof of the Littlewood Conjecture. By the work of Badziahin et al, this is not the case. Einsiedler, Katok, and Lindenstrauss have made the following contribution toward the Littlewood conjecture, stating that the set of exceptions to the Conjecture is small [22].

THEOREM 1.9. *The set containing all pairs $(x, y) \in \mathbb{R}^2$ such that*

$$\liminf_{q \rightarrow \infty} q \|qx\|_{\mathbb{Z}} \|qy\|_{\mathbb{Z}} > 0$$

has Hausdorff dimension 0.

The approach to the proof of this theorem begins with a conjecture of Margulis on a classification of measures which are invariant under a particular flow within some dynamical system. More on this will be discussed in the later sections.

There are three other directions in which to take diophantine approximation which are of concern within this dissertation: approximation on manifolds, approximation on fractals and approximation by integral elements in other fields or division algebras.

First, let us review some work on diophantine approximation on manifolds. Mahler conjectured in the 1930s that almost every point on the Veronese curve $\{(t, t^2, \dots, t^n) \mid t \in \mathbb{R}\}$ is not very well approximable. This was proven by Sprindzhuk in 1968 [64]. Later, he conjectured the following: Suppose $M = \{(f_1(x), \dots, f_n(x)) \mid x \in U\}$ is a manifold where U is a domain in \mathbb{R}^d and the functions f_i are real analytic which together with 1 are linearly independent over \mathbb{R} . Then almost every point of M is not very well approximable. This conjecture was proven by Kleinbock and Margulis by refining results on nondivergence of unipotent flows in the homogeneous space of lattices [43]. This paper inspired a number of subsequent articles on diophantine approximation on manifolds, in particular in proving Khintchine-type theorems for manifolds. Much like in the theorem of Kleinbock and Margulis, these theorems require non-degeneracy on the manifold, for example, ensuring that the manifold does not lie in a rational subspace. See, also, [33].

Since 2004, there have been a number of papers asking what can be said about diophantine approximation on fractals. Questions in this vein usually take the form of calculating the Hausdorff dimension of the intersection of the set of badly approximable objects with a sufficiently nice fractal, usually one which is the support of a so-called absolutely friendly measure. Examples of sufficiently nice fractals are the Cantor set, the Koch curve and the Sierpinski gasket. Kleinbock, Lindenstrauss and Weiss showed that almost every point on the support of an absolutely friendly measure is not very well approximable [42]. Fishman proved that the set of badly approximable linear forms, when intersect with a nice fractal, is a winning set within said fractal, and he proved further that winning sets within nice fractals have full Hausdorff dimension [30, 29].

The next type of problem in diophantine approximation deals with approximation not by integers in \mathbb{Q} , but by integral elements in number or p -adic fields. These types of questions replace the denominator of a rational as the measure of the quality of approximation with a height function H on the algebraic elements. For number fields and algebraic

integers, the height can be defined in one of several ways. One common definition of the height of the algebraic number β is the maximum of the absolute values of the coefficients appearing in the minimal polynomial of β . Another, the one that is used in the following chapters, is the maximum of the absolute values of the Galois conjugates of β . There are in general two types of questions for approximation over number fields. The first tries to approximate a fixed number by algebraic numbers of a fixed degree, and the other fixes a field, finding estimates for a fixed number or vector by integers or elements within the number field. One example of a theorem of many from the book [62] of this type is given here.

THEOREM 1.10. *Let k be a real algebraic number field. Then there exists a constant c_k such that for every real α not in k there are infinitely many $\beta \in k$ satisfying*

$$|\alpha - \beta| < c_k \max\{1, |\alpha|^2\} H(\beta)^{-2}.$$

The work contained in this thesis uses estimates by the integers of a fixed number field k , but the approximation is simultaneous by an integer and its conjugates. Suppose, for the moment, that k is totally real and has degree d over \mathbb{Q} . (This assumption is made now only for ease of explication, however, the results stated below hold for general number fields). Denote by $\sigma_1, \dots, \sigma_d$ the embeddings of k into \mathbb{R} , and by \mathcal{O}_k the ring of integers of k . We say a vector $\mathbf{x} \in \mathbb{R}^d$ is k -badly approximable if there exists a constant $c > 0$ such that for all $p, q \in \mathcal{O}_k$ with $q \neq 0$, we have

$$\max_{1 \leq i \leq d} \{|\sigma_i(q)|\} \max_{1 \leq i \leq d} \{|\sigma_i(q)x_i + \sigma_i(p)|\} > c.$$

Similarly, a d -tuple $(\langle A_i, b_i \rangle)$ of m inhomogeneous linear forms in n variables is k -badly approximable if there exists a constant $c > 0$ such that for all $p \in \mathcal{O}_k^m$ and $q \in \mathcal{O}_k^n$ with $q \neq 0$, we have

$$\max_{1 \leq i \leq d} \{\|\sigma_i(q)\|\}^n \max_{1 \leq i \leq d} \{\|A_i \sigma_i(q) + b_i + \sigma_i(p)\|\}^m > c.$$

Estimates of this type for Dirichlet-like theorems for k -badly approximable vectors have been considered in [13, 36]. In this thesis several results are proven about the Hausdorff dimension of the set of k -badly approximable vectors when restricted to a smooth curve (Chapter 3) and of the set of k -badly approximable inhomogeneous linear forms intersection with a nice fractal (Chapter 2). Specifically, in joint work with Einsiedler and Ghosh, we prove:

THEOREM 1.11. *Let $\phi = (\phi_i) : [0, 1] \rightarrow \mathbb{R}^d$ be a continuously differentiable map. We assume that $\phi'_i(x) \neq 0$ for all but finitely many*

$x \in [0, 1]$ and for at least $d/2$ of the indices i . Then the set of $x \in [0, 1]$ such that $\phi(x)$ is k -badly approximable is winning, and hence has Hausdorff dimension 1.

We also prove a collection of results in the vein of the following:

THEOREM 1.12. *Let $F \subset (\mathbb{R}^m)^d$ be a “nice” fractal. Fix a d -tuple of $n \times m$ real matrices (A_i) . Then the set of vectors $(b_i) \in F$ so that $(\langle A_i, b_i \rangle)$ is k -badly approximable is winning in F .*

The proofs of these results and the results of Chapter 4, use the structures developed in dynamical systems on homogenous spaces. We give a brief introduction to those ideas in the next section.

2. Homogeneous dynamics and applications

In this section, we introduce the basics of dynamical systems on homogenous spaces and some the connections between homogenous dynamics and number theory. To begin, let X be a measure space with the σ -algebra \mathcal{A} and the probability measure μ . Let $T : X \rightarrow X$ be measure-preserving, that is, for every $A \in \mathcal{A}$, $\mu(T^{-1}(A)) = \mu(A)$. Then T induces a transformation U_T on the Hilbert space $L^2(\mu)$ of square μ -integrable functions on X given by $U_T(f) = f \circ T$. This is known as the Koopman operator of T . Since T is measure-preserving, $\langle U_T f, U_T g \rangle = \langle f, g \rangle$ for any $f, g \in L^2(\mu)$. Thus, U_T is a unitary operator. The transformation T is said to be ergodic with respect to μ (or μ is ergodic relative to T) if the only T -invariant functions in $L^2(\mu)$ are the constant functions. Ergodic theory is the study of how time averages relate to space averages. Take for example one of the first ergodic theorems [26].

THEOREM 2.1 (Pointwise ergodic theorem). *Let (X, \mathcal{A}, μ, T) be a measure preserving probability space. Then for any $f \in L^1(\mu)$,*

$$\frac{1}{N} \sum_{n=0}^{N-1} f(T^n x)$$

converges almost everywhere to a T -invariant function. If T is in addition ergodic, then the convergence is to the constant function taking the value $\int f d\mu$.

Denote by δ_y , the point mass at y . Then the above theorem may be interpreted as saying for almost every x the weak* limit of $\frac{1}{N} \sum_{n=0}^{N-1} \delta_{T^n x}$ is μ , or that this sequence of measures equidistributes with respect to μ . A central question in homogeneous dynamics is determining the limiting behavior of various invariant measures.

As it has been described so far, we have an action of the discrete group \mathbb{Z} on the space X . In homogenous dynamics, we look at a continuous analogue. The typical set-up is the following: Let G be a “nice” Lie group and let Γ be a discrete subgroup. The group G admits a Haar measure m_G which descends to the quotient $X = G/\Gamma$. This measure on the quotient will be denoted m_X . This is the unique G -invariant measure on X . We assume further that G/Γ has finite volume, that is, that Γ is a lattice in G . Thus, we may assume that m_X has been normalized to be a probability measure. Suppose that $H = \{u_t \mid t \in \mathbb{R}\}$ is a subgroup of G containing precisely a one parameter unipotent flow. Now we have a continuous action of H on X . Then analogous to the pointwise ergodic theorem one may ask about the limits

$$\frac{1}{T} \int_0^T f(u_t x) dt \text{ as } T \rightarrow \infty.$$

For such unipotent actions, this was completely answered by Ratner in the following theorem [57].

THEOREM 2.2 (Ratner’s Equidistribution Theorem). *Let G be a Lie group, and Γ a lattice in G . Let $H = \{u_t\}$ be a one parameter unipotent flow in G . Then for each $x \in X = G/\Gamma$ there exists a closed connected subgroup L of G such that $\{u_t\} \subset L$, $H.x \subset \overline{H.x} = L.x$, and the unique L invariant measure, m_L , on $L.x$ satisfies*

$$\frac{1}{T} \int_0^T f(u_t x) dt \rightarrow \int_{L.x} f dm_L \text{ as } T \rightarrow \infty$$

for all $f \in C_c(X)$.

This powerful theorem uses a complete classification of all u_t invariant measures on X , which is also due to Ratner. It is also a key ingredient for a proof of the Oppenheim Conjecture.

THEOREM 2.3. *Let Q be a real, indefinite, nondegenerate quadratic form in $n \geq 3$ variables. If Q is not a scalar multiple of a form with integer coefficients, then $Q(\mathbb{Z}^n)$ is dense in \mathbb{R} .*

We remark that the first proof of the Oppenheim Conjecture in complete generality was given by Margulis in 1986. Ragunathan originally conjectured what are now called Ratner’s Theorems and showed that the Oppenheim Conjecture would follow from the measure classification given above. Ratner published her proofs around 1990. We now provide a sketch of the proof of this theorem using Ratner’s Theorem. For a thorough exposition on this topic, see [52].

SKETCH OF PROOF. Let Q be a real, indefinite, nondegenerate quadratic form in $n \geq 3$ variables. We will assume that $n = 3$ (although the general result can be reduced to this case). Let $G = \mathrm{SL}_3(\mathbb{R})$ and $\Gamma = \mathrm{SL}_3(\mathbb{Z})$. Since Q is indefinite, it must be of signature $(2, 1)$ or $(1, 2)$, that is, after some change of variables and multiplication by a scalar Q takes the form $Q_0(x_1, x_2, x_3) = x_1^2 + x_2^2 - x_3^2$. Suppose that $g \in \mathrm{SL}_3(\mathbb{R})$ and $\lambda \in \mathbb{R}$ are such that $Q = \lambda Q_0 \circ g$. Recall $\mathrm{SO}(Q) = \{h \in \mathrm{SL}_3(\mathbb{R}) \mid Q(hv) = Q(v) \text{ for all } v \in \mathbb{R}^3\}$. Then we have that $H := \mathrm{SO}(Q_0)^\circ \simeq \mathrm{PSL}_2(\mathbb{R})$ and $\mathrm{SO}(Q)^\circ = gHg^{-1}$. (Here $^\circ$ denotes the connected component containing the identity.) Since $\mathrm{SL}_2(\mathbb{R})$ is generated by unipotent elements, we may apply Ratner's Theorem to find a closed, connected subgroup $L < G$ such that $H < L$, $\overline{H.g} = L.g$, and there exists a unique L invariant probability measure on $L.g$. Analyzing the structure of the Lie algebra of G reveals that the only closed connected subgroups of G containing H are precisely H and G themselves.

First suppose that $L = G$. Then $H.g\Gamma$ is dense in G . Since \mathbb{Z}^3 is invariant under Γ and by construction of g , H and Q_0 ,

$$Q(\mathbb{Z}^3) = Q_0(H.g\Gamma\mathbb{Z}^3)$$

which is dense in $Q_0(G\mathbb{Z}^3) = Q_0(\mathbb{R}^3 \setminus \{0\})$. Thus, $Q(\mathbb{Z}^3)$ is dense in \mathbb{R} .

In the case that $L = H$, it turns out Q is a scalar multiple of a form with integer coefficients. To show this, one appeals to the theory of algebraic groups and the Borel Density Theorem. \square

As mentioned in the previous section, another partial measure classification devised by Einsiedler, Katok and Lindenstrauss gives the strongest existing result toward the Littlewood Conjecture. Instead of classifying measure invariant under unipotent flows, their theorem classifies measures invariant under diagonal flows provided there is an assumption made on the entropy of the measure. Before we state the theorem, let us take a moment to define entropy. This is an invariant of a dynamical system which is a measure of the "expansiveness" of a measure preserving system. As in the beginning of the section, let (X, \mathcal{A}, μ, T) be a measure preserving system. Let \mathcal{P} be a finite measurable partition of X and define

$$H_\mu(\mathcal{P}) = - \sum_{P \in \mathcal{P}} \mu(P) \log \mu(P).$$

For two finite measurable partitions \mathcal{P} and \mathcal{Q} , we define their common refinement to be $\mathcal{P} \vee \mathcal{Q} = \{P \cap Q \mid P \in \mathcal{P}, Q \in \mathcal{Q}\}$. Then for partitions \mathcal{P} with $H_\mu(\mathcal{P}) < \infty$ we define the metric entropy of T with respect to

μ and \mathcal{P} to be

$$h_\mu(T, \mathcal{P}) = \lim_{N \rightarrow \infty} \frac{1}{N} H_\mu \left(\bigvee_{i=0}^{N-1} T^{-i} \mathcal{P} \right).$$

It can be shown that this limit always exists. Finally, we define the metric entropy of T with respect to μ to be

$$h_\mu(T) = \sup_{\substack{\mathcal{P} \\ H_\mu(\mathcal{P}) < \infty}} h_\mu(T, \mathcal{P}).$$

For more details on entropy, see for example [66].

The theorem of Einsiedler, Katok and Lindenstrauss is the following:

THEOREM 2.4. *Let $X = \mathrm{SL}_d(\mathbb{R})/\mathrm{SL}_d(\mathbb{Z})$ for some $d \geq 3$. Let A be the space of diagonal matrices in $\mathrm{SL}_d(\mathbb{R})$. Let μ be an A invariant and ergodic measure on X . If there is a one parameter subgroup of A acting on X with positive entropy, then there is a closed connected intermediate subgroup L between A and G such that μ is the L invariant measure on a single closed L orbit in X .*

While we will not give a full explanation of how Theorem 1.9 follows from this result, we will provide some indication of the relationship between diophantine approximation and properties of orbits under diagonal flows in homogeneous spaces.

First, we will construct the space X_d of unimodular lattices of rank d . Let Λ be a unimodular lattice in \mathbb{R}^d . This means that Λ is a discrete subgroup of \mathbb{R}^d with a \mathbb{Z} basis consisting of d vectors v_1, v_2, \dots, v_d (linearly independent over \mathbb{R}) so that the parallelepiped P given by $P = \{\sum t_i v_i \mid t_i \in [0, 1]\}$ has volume 1. Thus $\Lambda = g\mathbb{Z}^d$, where the columns of g are the vectors v_i . That the parallelepiped P has volume 1 tells us that $\det g = 1$. This proves that the action of $\mathrm{SL}_d(\mathbb{R})$ on the set of rank d unimodular lattices is transitive. We choose as a base point the standard lattice \mathbb{Z}^d . The stabilizer of this lattice is $\mathrm{SL}_d(\mathbb{Z})$. Therefore $X_d \simeq \mathrm{SL}_d(\mathbb{R})/\mathrm{SL}_d(\mathbb{Z})$.

Now $\mathrm{SL}_d(\mathbb{Z})$ is a discrete subgroup of $\mathrm{SL}_d(\mathbb{R})$. It can be seen by construction near-fundamental domains of $\mathrm{SL}_d(\mathbb{R})/\mathrm{SL}_d(\mathbb{Z})$ known as Siegel sets that X_d has finite volume. Thus $\mathrm{SL}_d(\mathbb{Z})$ is a lattice in $\mathrm{SL}_d(\mathbb{R})$. However, this space is not compact. Consider the sequence of unimodular lattices with \mathbb{Z} basis given by $\{\frac{1}{n}e_1, ne_2, e_3, \dots, e_d\}$ where the e_i are the standard basis vectors of \mathbb{R}^d . As n tends to infinity, the first basis vector tends to 0 so that the limiting object is degenerate and is no longer a rank d lattice. This is the premise behind Mahler's Compactness Criterion:

THEOREM 2.5 (Mahler's Compactness Criterion). *A subset $A \subset X_d$ is precompact (that is, has compact closure) if and only if there exists a constant c such that for all $\Lambda \in A$, for all $v \in \Lambda \setminus \{0\}$, we have $\|v\| > c$.*

Here $\|\cdot\|$ can be any norm on \mathbb{R}^d , but generally we refer to the maximum norm.

Let us fix a vector $x = (x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1}$, and define the matrix

$$u(x) = \begin{pmatrix} 1 & & & & & \\ x_1 & 1 & & & & \\ x_2 & 0 & 1 & & & \\ \vdots & \vdots & \ddots & \ddots & & \\ x_d & 0 & \cdots & 0 & 1 & \end{pmatrix}.$$

Set $\Lambda_x = u(x)\mathbb{Z}^d$. Then a typical element of Λ_x is

$$u(x) \begin{pmatrix} q \\ p_1 \\ \vdots \\ p_{d-1} \end{pmatrix} = \begin{pmatrix} q \\ qx_1 + p_1 \\ \vdots \\ qx_{d-1} + p_{d-1} \end{pmatrix},$$

and so contains precisely all of the integer approximations of the vector x . Dani observed the following theorem, now referred to as the Dani Correspondence [15].

THEOREM 2.6 (Dani's Correspondence). *A vector $x \in \mathbb{R}^{d-1}$ is badly approximable if and only if the trajectory $g_t\Lambda_x$ for $t > 0$ is bounded in X_d , where g_t is the diagonal matrix with entries $(e^{-(d-1)t}, e^t, \dots, e^t)$.*

This theorem, in conjunction with Mahler's compactness criterion, is the link between statements in diophantine approximation and dynamics on homogeneous spaces. In fact finer information can be derived with a thorough understanding of this link between dynamics and diophantine approximation. There is a computable relationship between the constant appearing in the definition of badly approximable and the constant appearing in Mahler's compactness criterion. Other geometric characteristics of orbits correspond to other types of diophantine behavior. For example, an orbit $g_t\Lambda_x$ avoiding a fixed sufficiently large compact set for all large t relates to whether x satisfies the property of being Dirichlet improvable (see [63]). Indeed, the work of Kleinbock and Margulis toward proving the Sprindzhuk Conjecture is based in estimating from above the proportion of time in a fixed interval that an orbit under a unipotent flow may spend outside of a sufficiently large compact set. This is known as quantitative nondivergence.

However, Moore's ergodicity theorem, which states that the action of g_t is ergodic, tells us that almost every orbit is spread evenly throughout X_d . Thus, the set of badly approximable vectors, the set of very well approximable vectors and the set of Dirichlet improvable all have measure zero. Thus, the basic underlying principle is exceptional orbits correspond to exceptional diophantine properties.

The results contained in Chapters 2 and 3 of this thesis mimic the construction of the space of lattices to construct the space of \mathcal{O}_k -modules, and a corresponding homogenous space. The link of the Dani Correspondence between diophantine properties over a number field k and geometric properties of orbits within the homogeneous space is reestablished. Finally, the results described at the end of Section 1 are proven, using techniques that are similar to those developed in the next section.

3. A detailed example

This section contains a warm-up example of using dynamics combined with Schmidt's games to prove a now standard theorem in diophantine approximation. The approach used below has the same basic structure as the arguments given Chapters 2 and 3. The presentation here is of a simple situation, to give a scaffolding for the more ornamented versions later.

THEOREM 3.1. *The set Bad of badly approximable numbers is winning, whence has Hausdorff dimension 1.*

PROOF. To each real number x , we associate the matrix

$$u(x) = \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix}$$

and the lattice

$$\Lambda_x = u(x)\mathbb{Z}^2 = \left\{ \begin{pmatrix} qx + p \\ q \end{pmatrix} \mid p, q \in \mathbb{Z} \right\}.$$

This may be viewed as both as an element in the space X_2 of rank 2 unimodular lattices, and as the element $u(x) \mathrm{SL}_2(\mathbb{Z})$ on the modular curve $\mathrm{SL}_2(\mathbb{R}) \backslash \mathrm{SL}_2(\mathbb{Z})$. Define

$$g_t = \begin{pmatrix} e^t & \\ & e^{-t} \end{pmatrix}.$$

Let us fix, throughout this paragraph, a real number x . To determine whether x is badly approximable, we consider the quantity $|q||qx + p|$ as q and p range over the integers. When this quantity is uniformly bounded below, x is badly approximable. The vectors in

the lattice Λ_x contain precisely all integer approximations of x . For a generic point $v = (qx + p, q)^\top \in \Lambda_x$ consider the effect of g_t on v . The first coordinate, $qx + p$, which may possibly be quite small, is dilated by a factor of e^t . The second coordinate, q , is contracted by a proportionate amount. If e^t is roughly of size $|q|$, then the second coordinate is of size roughly 1, and checking if $|q||qx + p| = |e^{-t}q||e^t(qx + p)|$ is small amounts to checking if the vector g_tv is near the origin. (This is the heart of the Dani Correspondence.) We may think of t as a time parameter, and as the time parameter progresses, checking for short vectors in the lattice $g_t\Lambda_x$ will show if there are good rational approximations of x with denominator of size roughly e^t .

In order to demonstrate that the set is winning, we must give a winning strategy for Alice. Before we begin playing the game, let us describe the connection between the progression of time in the flow g_t on $\mathrm{SL}_2(\mathbb{R}) \setminus \mathrm{SL}_2(\mathbb{Z})$ and the progression of the rounds of the game. Between round n and round $n + 1$, the playing field is reduced from a ball B_n centered at the point x_n of radius $\rho(\alpha\beta)^n$ to the ball B_{n+1} centered at x_{n+1} of radius $\rho(\alpha\beta)^{n+1}$. Alice's intermediate choice A_{n+1} between B_n and B_{n+1} will be informed by considering the rational approximations $\frac{p}{q}$ of the center point x_n with denominator of size roughly $(\rho(\alpha\beta)^n)^{-1}$, and trying to move the playing field so that all such $\frac{p}{q}$ are sufficiently (uniformly in n) distance to x for $x \in A_{n+1}$ remaining in the game. Notice that there is essentially one pair (p, q) with q roughly of size $(\rho(\alpha\beta)^n)^{-1}$ and p/q lying in the ball of radius $\rho(\alpha\beta)^n$ around x_n . To more easily make this choice, Alice will look at the dynamical picture. Define the sequence t_n so that $e^{2t_n} = (\rho(\alpha\beta)^n)^{-1}$. The points x in B_n , along with the normalization allowing Alice to focus on the potentially good approximations with denominator of size e^{t_n} , correspond to the lattices

$$\begin{aligned} & \{g_{t_n}\Lambda_x \mid x \in B_n\} \\ &= \left\{ \begin{pmatrix} e^{t_n} & \\ & e^{-t_n} \end{pmatrix} \begin{pmatrix} 1 & x_n + x \\ & 1 \end{pmatrix} \mathbb{Z}^2 \mid x \in [-\rho(\alpha\beta)^n, \rho(\alpha\beta)^n] \right\} \\ &= \left\{ \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \begin{pmatrix} e^{t_n} & \\ & e^{-t_n} \end{pmatrix} \begin{pmatrix} 1 & x_n \\ & 1 \end{pmatrix} \mathbb{Z}^2 \mid x \in [-1, 1] \right\}. \end{aligned}$$

Using this normalization of g_{t_n} , Alice reduces a relatively complicated choice to choosing a ball A'_{n+1} of radius α in $[-1, 1]$ so that no vector in any of the lattices

$$\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} g_{t_n}\Lambda_{x_n}$$

for $x \in A'_{n+1}$ is too small. Indeed she should be able to make this choice independent of n .

Suppose, then, Alice and Bob have arrived at round n of the (α, β) game, where $\beta > 0$ and $\frac{1}{2} > \alpha > 0$. Bob has just selected the ball $B_n = (x_n, \rho(\alpha\beta)^n)$. We will now describe the strategy of Alice. If in the lattice $\Lambda_n = g_{t_n}\Lambda_{x_n}$ there are no nonzero vectors of norm less than $\frac{1}{2}$, Alice chooses the ball A_{n+1} of radius α at random in B_n . Suppose, on the other hand, that there is a nonzero vector $v = \begin{pmatrix} p \\ q \end{pmatrix}$ in the unimodular lattice Λ_n , with the maximum norm $\|v\| < \frac{1}{2}$. Notice that if there is another vector $w \in \Lambda_n$ with $\|w\| < \frac{1}{2}$ then w must be linearly dependent on v by the unimodularity of Λ_n . Thus, Alice does not concern herself with a particular short vector, but a line containing the short vectors. Moreover, we may take v to be the shortest nonzero vector of Λ_n , and we have that $\|g_{-t}v\| \geq \frac{1}{2}$ for all $t > 0$. Consider the neighboring points of v in the lattices $\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \Lambda$:

$$\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} qx + p \\ q \end{pmatrix}.$$

For $x \in [-1, 1]$, this collection of vectors describes a line segment which intersects the unit ball. Depending on which quadrant v lies in, that is, on whether $\frac{p}{q}$ is positive or negative, Alice may choose as A_{n+1} the rightmost, resp. leftmost, subinterval of $[-1, 1]$ of length 2α . This maximizes the norm of the vectors $\begin{pmatrix} qx + p \\ q \end{pmatrix}$ for $x \in A_{n+1}$ remaining in the game, however it does not directly give a uniform lower bound on the size of these vectors. What is guaranteed though is that the angle of a vector of the form $\begin{pmatrix} qx + p \\ q \end{pmatrix}$ to the contracting direction (relative to g_t) is (uniformly) bounded below by a constant depending only on α . Indeed, we have that the ratio $|\frac{qx+p}{q}| > 1 - 2\alpha$. So while this vector $\begin{pmatrix} qx + p \\ q \end{pmatrix}$ may initially continue to shrink under application of the flow g_t , it may not do so indefinitely. For $t > t_0 = -\frac{1}{2} \log(1 - 2\alpha)$ and $x \in A'_{n+1}$, $\|g_t \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} v\|$ is growing. Thus, for any t and any $x \in A'_{n+1}$, $\|g_t \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} v\|$ will never be smaller than $\xi = \frac{1}{2}\rho(\alpha\beta)^{n_0+1}$ where n_0 is the least number such that $\rho(\alpha\beta)^{n_0} < 1 - 2\alpha$.

Let us verify that this strategy does result in the point x_∞ remaining at the end of the game being badly approximable. Recall, that x_∞ is badly approximable is equivalent to the orbit $g_t\Lambda_{x_\infty}$, $t > 0$, having all nonzero vectors bounded away from the origin, which is equivalent to the collection $g_{t_n}\Lambda_{x_\infty}$, $n \in \mathbb{N}$, having all nonzero vectors bounded away from the origin. Suppose there exists n and $v \in g_{t_n}\Lambda_{x_\infty}$ with $\|v\| < \frac{1}{2}$. Then since $x_\infty \in A_{n+1}$, we know that Alice has shifted the playing field so that v will be growing in a bounded amount of time and g_tv will never have norm shorter than ξ . Thus, all nonzero vectors of $g_{t_n}\Lambda_{x_\infty}$ are bounded away from the origin. \square

4. Exceptional orbits for multiple transformations

The previous sections have focused on orbits having exceptional properties under only one flow or transformation. However, it is also interesting to consider the behavior of a single point under two different transformations. Let T and S be ergodic endomorphisms of the torus \mathbb{T}^d . This means that T and S can be written as integer matrices with no eigenvalues that are roots of unity. Let $x \in \mathbb{T}^d$. The torus is compact, so naturally the sets $\{T^n x\}_{n \in \mathbb{N}}$ and $\{S^n x\}_{n \in \mathbb{N}}$ are bounded. So within this set up, every orbit is bounded and, by ergodicity, almost every orbit equidistributes throughout the torus. Thus, the exceptional orbits are those which do not equidistribute, or even stronger, those which are not dense. Denote, then, by $D(S)$ the set of points with dense orbit under S and by $ND(T)$ the set of points with orbit which is not dense under T . How big is the set $ND(T) \cap D(S)$? By the assumption of ergodicity, we know that $ND(T)$, hence also $ND(T) \cap D(S)$, has measure zero. However, Broderick, Fishman and Kleinbock showed that $N(D)$ is 1/2-winning and so has Hausdorff dimension 1 [11]. So there is hope that the Hausdorff dimension of $ND(T) \cap D(S)$ is positive. With extra assumptions on T and S , this has been shown to be true.

Before we give the first example of such a result, we must first give some definitions for those extra assumptions. A toral endomorphism is hyperbolic if it has no eigenvalue of modulus 1. Suppose now that S and T are commuting automorphisms. Then, together they generate a \mathbb{Z}^2 -action α on \mathbb{T}^d . A \mathbb{Z}^2 action β on \mathbb{T}^k is an algebraic factor of α if there is a surjective homomorphism $h : \mathbb{T}^d \rightarrow \mathbb{T}^k$ such that $\beta \circ h = h \circ \alpha$. In addition, β is a rank-one factor if $\beta(\mathbb{Z}^2)$ has a finite index subgroup generated by a single map. In [9], Bergelson, Einsiedler and Tseng establish the following theorem:

THEOREM 4.1. *Let S and T be commuting hyperbolic automorphisms of the torus \mathbb{T}^d for $d \geq 2$ such that S and T generate an algebraic \mathbb{Z}^2 -action without rank-one factors. Then $ND(T) \cap D(S)$ has full Hausdorff dimension.*

As with the proof on the size of the set of exceptions to the Littlewood Conjecture, their proof relies on a classification of ergodic measures for \mathbb{Z}^d -actions on the torus and makes use of entropy.

In joint work with Maier, we consider the case that S and T are not necessarily commuting. Our assumptions on the endomorphisms ask the eigenspaces are, as opposed to the commuting case, not aligned. In Chapter 4, we present a proof of:

THEOREM 4.2. *Let S and T be hyperbolic endomorphisms of \mathbb{T}^d . Let \mathfrak{s}_- be the maximal subspace contracted by S and let \mathfrak{t}_- be the maximal subspace contracted by T . Assume that both \mathfrak{s}_- and \mathfrak{t}_- are nontrivial, and that, together, they span \mathbb{R}^d . Then the Hausdorff dimension of $ND(T) \cap Eq(S)$ is d .*

Here, $Eq(S)$ is the set of points which equidistribute under S . Denote by $NEq(T)$ the set of points which do not equidistribute under T . We also provide a proof for the following theorem.

THEOREM 4.3. *Let S and T be ergodic toral endomorphisms. Let \mathfrak{s}_- be the subspace contracted by S and let \mathfrak{t}_- be the subspace contracted by T . Let \mathfrak{s}'_0 (\mathfrak{t}'_0 , resp.) denote the sum of the eigenspaces of S (of T , resp.) of eigenvalues with modulus 1. Assume that $\mathfrak{s}_- \oplus \mathfrak{s}'_0$ and $\mathfrak{t}_- \oplus \mathfrak{t}'_0$ are nontrivial and that, together, they span \mathbb{R}^d . Then the Hausdorff dimension of the set $NEq(T) \cap Eq(S)$ is d .*

The proofs of both of these theorems follow essentially the same strategy. We show that given a point $x \in \mathbb{T}^d$, translating x by an element y in the contracting direction for one of the endomorphisms does not change the asymptotic behavior of the element. That is, either both x and $x + y$ equidistribute under S or both do not. In the case that T is hyperbolic we can show that either both x and $x + y$ have dense orbit or both do not. We use these transverse partitions $x + \mathfrak{s}_-$ for $x \in Eq(S)$ or $x \in ND(T)$ and $x + \mathfrak{t}_-$ for x in $Eq(S)$ or $x \in ND(T)$ along with the known features of $Eq(S)$ and $ND(T)$ and the Marstrand Slicing Theorem to come to the desired conclusion.

As an end remark, we reiterate that this thesis is cumulative. The chapters to follow are self-contained and maybe read independently. As such, the notation used in one paper may not match that of the next. We apologize to the reader for the inconvenience.

CHAPTER 2

Badly approximable vectors, C^1 curves and number fields

MANFRED EINSIEDLER, ANISH GHOSH, AND BEVERLY LYTLE

ABSTRACT. We show that the set of points on C^1 curves which are badly approximable by rationals in a number field form a winning set in the sense of W. Schmidt. As a consequence, we obtain a number field version of Schmidt's conjecture in Diophantine approximation.

1. Introduction

Recall that a real number x is *badly approximable* if there exists $c > 0$ such that

$$|qx - p| > \frac{c}{q} \quad (1)$$

for all $q \in \mathbb{N}$ and $p \in \mathbb{Z}$. It is well known that badly approximable vectors have zero Lebesgue measure and full Hausdorff dimension (Jarnik [38] for $n = 1$ and Schmidt [59, 62] for arbitrary n). In fact, Schmidt showed that they are *winning* for a certain game, a stronger and more versatile property than having full Hausdorff dimension.

1.1. Schmidt's Game. In [59], Schmidt introduced the following game. Two players, say Player A and Player B, start with a complete metric space X , a subset $W \subseteq X$, and two parameters $0 < \alpha, \beta < 1$. Player A begins by choosing an arbitrary ball $A_0 = B(x_0, \rho)$. The Player B then chooses a ball $B_0 = B(y_1, \alpha\rho)$ contained in A_0 . Player A makes his next move by choosing a ball $A_1 \subset B_0$ of radius $\alpha\beta\rho$. The n th step of the game consists of first the Player A choosing a ball $A_n = B(x_n, (\alpha\beta)^n\rho) \subset B_{n-1}$ and Player B following by choosing the next ball $B_n = B(y_n, \alpha(\alpha\beta)^n\rho) \subset A_n$. As the radii of the balls are shrinking to zero and X is complete, at the end of the infinite game, Player A and Player B are left with a single point $\{x_\infty\} = \bigcap_n A_n$. We say that Player B has won this (α, β) game if $x_\infty \in W$. The set W is called (α, β) -winning if Player B can find a winning strategy, α -winning if it is (α, β) -winning for all $0 < \beta < 1$ and winning if it is α -winning

for some $\alpha > 0$. Schmidt games have the following properties (cf. [59], [18]):

- (1) A winning subset of \mathbb{R}^n is *thick*, i.e. the intersection of a winning set with every open set in \mathbb{R}^n has Hausdorff dimension n .
- (2) A countable intersection of α -winning sets is α -winning.
- (3) Winning sets are preserved by bi-Lipschitz homeomorphisms of \mathbb{R}^n .
- (4) The set of badly approximable vectors is (α, β) -winning whenever $2\alpha < 1 + \alpha\beta$; in particular, it is α -winning for any $0 < \alpha \leq 1/2$.

1.2. Diophantine approximation in number fields. Let K be a number field of degree d with r real and s complex embeddings. Denote by S the set of Galois embeddings σ , where for the complex embeddings, one chooses one of the pair σ and $\bar{\sigma}$. Let \mathcal{O}_K be the ring of integers of K . Denote by $K_S := \mathbb{R}^r \times \mathbb{C}^s \cong \mathbb{R}^d$. We denote by τ the twisted diagonal embedding of K into K_S by

$$\tau(x) = (\sigma(x))_{\sigma \in S},$$

where we identify each coordinate of K_S with an element of S . The notation will be extended to vectors and matrices and τ will be omitted in the notation when it causes no confusion.

It is natural to ask if analogues of the traditional theorems in Diophantine approximation hold in the setting of number fields. More precisely, we wish to approximate elements in K_S using ratios of elements in \mathcal{O}_K . Analogues of Dirichlet's theorem in this setting have been established by several authors (cf. [61], [13], [56], [36]) using appropriate adaptations of the geometry of numbers. Moreover, [13] and [36] also show the existence of badly approximable vectors¹ in this setting.

Say that a vector $\mathbf{x} = (x^\sigma)_{\sigma \in S} \in K_S$ is K -badly approximable ($\mathbf{x} \in \text{Bad}(K)$) if there exists $c > 0$ such that for all $p, q \in \mathcal{O}_K$ with $q \neq 0$

$$\max_{\sigma \in S} \{|\sigma(p) + x^\sigma \sigma(q)|\} \max_{\sigma \in S} \{|\sigma(q)|\} > c. \quad (2)$$

¹We note, however, that our notion of badly approximable vectors differs slightly from the notion considered elsewhere as we do not square the absolute value at the complex places in our definition.

Here and in the rest of the paper, $||$ will be used to denote both real and complex absolute values depending on context.

S.G. Dani [15] showed that a real number is badly approximable if and only if a lattice associated to the number has a bounded trajectory in the space $\mathrm{SL}_2(\mathbb{Z}) \backslash \mathrm{SL}_2(\mathbb{R})$ under the action of a certain subsemigroup. A version of the Dani correspondence (§3) states that a vector \mathbf{x} is K -badly approximable if and only if the trajectory

$$\left\{ \mathrm{SL}_2(\mathcal{O}_K) \left(\left(\begin{array}{cc} 1 & x^\sigma \\ & 1 \end{array} \right) \right)_{\sigma \in S} g_t \mid t \geq 0 \right\}$$

for the flow

$$g_t := \left\{ \left(\left(\begin{array}{cc} e^{-t} & \\ & e^t \end{array} \right) \right)_{\sigma \in S} : t \geq 0 \right\} \quad (3)$$

is bounded in the quotient space $\mathrm{SL}_2(\mathcal{O}_K) \backslash \mathrm{SL}_2(K_S)$. In conjunction with the Moore ergodicity theorem, we can conclude that K -badly approximable vectors have zero Lebesgue measure. Nevertheless, they constitute a winning set for Schmidt's game and therefore have full Hausdorff dimension.

1.3. Main Results. We show that the set of badly approximable vectors are winning even when the game is restricted to a curve. We need slightly separate conditions on the curve in question in different cases. We recall that $r + s$ is the number of simple factors of $\mathrm{SL}_2(K_S)$.

THEOREM 1.1. *Let $\phi = (\phi_\sigma)_{\sigma \in S} : [0, 1] \rightarrow K_S$ be a continuously differentiable map. We assume that $\phi'_\sigma(x) \neq 0$ for all but finitely many $x \in [0, 1]$ and for all σ in a subset $S' \subset S$ (possibly depending on x) with*

$$|\{\sigma \in S' : \sigma \text{ is real}\}| + 2|\{\sigma \in S' : \sigma \text{ is complex}\}| > \lfloor \frac{d}{2} \rfloor.$$

Define

$$\Phi(x) := \left(\left(\begin{array}{cc} 1 & \phi_\sigma(x) \\ & 1 \end{array} \right) \right)_{\sigma \in S}, \quad (4)$$

and let g_t as defined in (3) act on $\mathrm{SL}_2(\mathcal{O}_K) \backslash \mathrm{SL}_2(K_S)$ by right multiplication. Let $\Lambda = \mathrm{SL}_2(\mathcal{O}_K)g \in \mathrm{SL}_2(\mathcal{O}_K) \backslash \mathrm{SL}_2(K_S)$. Then the set

$$\{x \in [0, 1] \mid \Lambda\Phi(x) \text{ has bounded trajectory under the flow } g_t\} \quad (5)$$

is winning in the sense of Schmidt, and hence has Hausdorff dimension 1.

We note that the condition on the curve in the theorem above is not of a technical nature. If Φ simply parametrizes a line segment that is parallel a coordinate axis (which are special directions as they

correspond to the simple factors of the ambient Lie group), then there may not be any points with bounded trajectory on the line segment if e.g. $d = r = 3$, see §4.2.

Theorem 1.1 coupled with Dani's correspondence gives us:

COROLLARY 1.2. *Let $\phi : [0, 1] \rightarrow K_S$ be as in Theorem 1.1. Then the set*

$$\{x \in [0, 1] \mid \phi(x) \in \text{Bad}(K)\}$$

is winning in the sense of Schmidt, and hence has Hausdorff dimension 1.

Now we let K be a real quadratic extension of \mathbb{Q} where $d = r = 2$. In this case, we can choose different directions in the two-dimensional Cartan subgroup.

THEOREM 1.3. *Let K be a real quadratic field. Let $\phi = (\phi_\sigma)_{\sigma \in S} : [0, 1] \rightarrow K_S$ be a continuously differentiable map such that $\phi'_\sigma(x) \neq 0$ for all but finitely many $x \in [0, 1]$ and every $\sigma \in S$. Let $\mathbf{r} \in \mathbb{R}^2$ be a real vector with $r_\sigma \geq 0$ for $\sigma \in S$ and $\sum_\sigma r_\sigma = 1$. For $t \geq 0$, let*

$$g(\mathbf{r})_t := \left(\left(\begin{array}{cc} e^{-r_\sigma t} & \\ & e^{r_\sigma t} \end{array} \right) \right)_{\sigma \in S} \quad (6)$$

act on $\text{SL}_2(\mathcal{O}_K) \backslash \text{SL}_2(K_S)$ by right multiplication. Let $\Lambda = \text{SL}_2(\mathcal{O}_K)g \in \text{SL}_2(\mathcal{O}_K) \backslash \text{SL}_2(K_S)$. Then the set

$$\{x \in [0, 1] \mid \Lambda\Phi(x) \text{ has bounded trajectory under the flow } g(\mathbf{r})_t\} \quad (7)$$

is winning in the sense of Schmidt, and hence has Hausdorff dimension 1.

The α in the winning statement above does not depend on \mathbf{r} . Taking intersections over rational vectors \mathbf{r} and using the fact that countable intersections of α -winning sets are α -winning therefore shows us that

COROLLARY 1.4. *With notation as in Theorem 1.3,*

$$\bigcap_{\mathbf{r} \in \mathbb{Q}_+^2} \{x \in [0, 1] \mid \Lambda\Phi(x) \text{ has bounded trajectory under the flow } g(\mathbf{r})_t\} \quad (8)$$

is winning.

Remarks:

- (1) Corollary 1.4 provides an analogue of Schmidt's conjecture for real quadratic fields, a theorem of Badziahin-Pollington-Velani [5] in the real case. See also the works [1, 2] of Jinpeng An for stronger results in this vein.

- (2) In contrast with Corollary 1.4, we note that it follows from results in [22] (see also [23]) that bounded orbits for the full Cartan subgroup have *zero* Hausdorff dimension.
- (3) Using Theorem 1.1 and the Marstrand Slicing Theorem, we see that $\text{Bad}(K)$ has full Hausdorff dimension.
- (4) As far as we are aware, the only other result regarding abundance of badly approximable vectors in the context of number fields is for certain quadratic extensions whose rings of integers have unique factorization [27].
- (5) Theorems 1.1, 1.3 and Corollary 1.2 are among the very few existing results which show that badly approximable points on curves are winning. In a recent work, V. Beresnevich [7] show that badly approximable vectors on “nondegenerate” manifolds have full Hausdorff dimension. See also the works [47, 48, 49] for results regarding badly approximable vectors on certain classes of fractals.
- (6) We refer to the r_i 's which appear in $g(\mathbf{r})_t$ as weights. Thus Theorem 1.1 deals with equal weights and Theorem 1.3 with unequal weights. The equal weights version of our results is closely related to a result from [12] (see Proposition 4.9). We note, however, that in the context of this paper certain special directions (e.g. line segments parallel to a coordinate axis corresponding to one of three real factors) may fail to have any badly approximable points on them while line segments in general directions are covered by Theorem 1.1.
- (7) In a related, earlier result [3], it is shown that points on C^1 curves in rank 1 locally symmetric spaces which have bounded orbits under the geodesic flow are winning. The result of this paper may be viewed as a generalization of this result to certain quotients of higher rank groups with \mathbb{Q} -rank 1.

Acknowledgements. AG thanks ETH Zürich for hospitality during visits.

2. Mahler's Compactness Criterion

In this section, we state and prove Mahler's compactness criterion for S -adic homogeneous spaces. Theorem 2.2, which is the main result in this section is almost certainly well known to experts. For instance, see [44, 45]. We provide a proof for completeness. The original statement of Mahler's compactness criterion is as follows:

THEOREM 2.1. *A subset $A \subset \mathrm{SL}_n(\mathbb{Z}) \backslash \mathrm{SL}_n(\mathbb{R})$ is relatively compact if and only if there exists $\varepsilon > 0$ such that for all $\mathrm{SL}_n(\mathbb{Z})g \in A$ and for all $v \in \mathbb{Z}^n g$, $\|v\| > \varepsilon$.*

Here we take $\|v\|$ to be the maximum norm. We wish to rephrase this theorem for the space $\mathrm{SL}_2(\mathcal{O}_K) \backslash \mathrm{SL}_2(K_S)$. To do this we will use restriction of scalars to map $\mathrm{SL}_2(\mathcal{O}_K) \backslash \mathrm{SL}_2(K_S)$ in $\mathrm{SL}_{2d}(\mathbb{Z}) \backslash \mathrm{SL}_{2d}(\mathbb{R})$. More concretely, we define the embedding as follows:

Since the degree of K over \mathbb{Q} is d , we may view K as a d dimensional vector space over \mathbb{Q} . We choose a basis $\{a_1, \dots, a_d\}$ of K over \mathbb{Q} so that the \mathbb{Z} -span of these elements is \mathcal{O}_K . Left multiplication by an element of K is a \mathbb{Q} -linear transformation on K . We have an algebraic embedding $\iota : K \rightarrow \mathrm{Mat}_{d,d}(\mathbb{Q})$ with respect to this chosen basis. We then have an induced map

$$\begin{aligned} \iota : \mathrm{SL}_2(K) &\rightarrow \mathrm{SL}_{2d}(\mathbb{R}) \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} &\mapsto \begin{pmatrix} \iota(a) & \iota(b) \\ \iota(c) & \iota(d) \end{pmatrix} \end{aligned}$$

Thus, we have defined the algebraic subgroup

$$\mathrm{Res}_{K/\mathbb{Q}} \mathrm{SL}_2 = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \mid A, B, C, D \in \mathrm{Im}(\iota), AD - BC = \mathrm{Id} \right\},$$

such that $\iota(\mathrm{SL}_2(K)) = (\mathrm{Res}_{K/\mathbb{Q}} \mathrm{SL}_2)(\mathbb{Q}) \subseteq \mathrm{SL}_{2d}(\mathbb{Q})$.

To give another idea of the structure of this space, consider the basis of K given by $\{1, \xi, \xi^2, \dots, \xi^{d-1}\}$, where ξ is a primitive element of K , that is, $K = \mathbb{Q}(\xi)$. (Note that the transformation from this basis to the previously mentioned one is rational.) Then as a \mathbb{Q} -linear transformation of K , left multiplication by ξ is represented by the companion matrix

$$T_\xi = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -c_0 & -c_1 & -c_2 & \cdots & -c_{d-1} \end{pmatrix},$$

where the $p_{\min}(x) = x^d + c_{d-1}x^{d-1} + \cdots + c_1x + c_0$ is the minimal polynomial of ξ over \mathbb{Q} . It is well known that this is conjugated by the Vandermonde matrix V_ξ (associated to the various Galois embeddings of ξ) to the diagonal form

$$V_\xi^{-1} T_\xi V_\xi = \mathrm{diag}(\sigma(\xi))_{\sigma \in S}$$

Since the elements of $\mathrm{Im}(\iota)$ commute, they are simultaneously diagonalizable. Moreover for any $\zeta \in K$ there exists $p_\zeta(x) \in \mathbb{Q}[x]$ of degree

less than $d - 1$ with $\zeta = p_\zeta(\xi)$, and so the transformation of left multiplication by ζ is given by $T_\zeta = p_\zeta(T_\xi)$ and is conjugate to

$$\text{diag}(p_\zeta(\sigma(\xi)))_{\sigma \in S} = \text{diag}(\sigma(\zeta))_{\sigma \in S}$$

(since the maps σ are ring homomorphisms). After a simple change of bases, we have an embedding $\text{SL}_2(K) \rightarrow \text{SL}_{2d}(\mathbb{R})$ given by

$$\begin{pmatrix} \zeta_1 & \zeta_2 \\ \zeta_3 & \zeta_4 \end{pmatrix} \mapsto \text{diag} \left(\begin{pmatrix} \sigma(\zeta_1) & \sigma(\zeta_2) \\ \sigma(\zeta_3) & \sigma(\zeta_4) \end{pmatrix} \right)_{\sigma \in S}.$$

This is precisely $\iota(\text{SL}_2(K)) \subset \text{SL}_2(K_S)$, where the latter is sitting in $\text{SL}_{2d}(\mathbb{R})$ in block diagonal form and $\iota(\text{SL}_2(K))$ forms the \mathbb{Q} -points of the variety defined above after conjugation with the appropriate change of basis matrix.

As with the identification of $\text{SL}_2(\mathbb{Z}) \backslash \text{SL}_2(\mathbb{R})$ with the space of co-volume 1 lattices in \mathbb{R}^2 , we have an identification of $\text{SL}_2(\mathcal{O}_K) \backslash \text{SL}_2(K_S)$ with the set, denoted X , of discrete (as subsets of K_S^2) rank 2 \mathcal{O}_K -modules with the property that for each $\Lambda \in X$ there exists a basis $\{v, w\}$ of Λ so that for each σ , $\sigma(v)$ and $\sigma(w)$ form the sides of a parallelepiped of area 1 (in \mathbb{R}^2 or \mathbb{C}^2 , appropriately). Now that we have a proper embedding $\text{SL}_2(\mathcal{O}_K) \backslash \text{SL}_2(K_S) \rightarrow \text{SL}_{2d}(\mathbb{Z}) \backslash \text{SL}_{2d}(\mathbb{R})$, we use Mahler's compactness criterion on the second space to derive the statement:

THEOREM 2.2. *A subset $A \subset X$ is relatively compact if and only if there exists $\varepsilon > 0$ such that for all $\Lambda \in A$ and for all vectors $v \in \Lambda = \tau(\mathcal{O}_K^2)g$, $\|v\| > \varepsilon$.*

For a vector v in K_S^2 , denote by v^σ the projection of v onto the factor associated with the embedding σ . We define a height function $H : K_S^2 \rightarrow \mathbb{R}$ by

$$H(v) := \prod_{\sigma} \|v^\sigma\|_{\sigma} = \prod_{\sigma \text{ real}} \|v^\sigma\| \prod_{\sigma \text{ complex}} \|v^\sigma\|^2,$$

where we write $\|\cdot\|_{\sigma}$ for the norm respectively the square of the norm depending on whether the place σ is real or complex. It will be useful to think of this height function as a measure of depth into the cusp. We wish to say that a set A is relatively compact if and only if the height function is uniformly bounded below by a positive constant over all \mathcal{O}_K -modules in A . To prove this statement, we first need some properties of the function H .

LEMMA 2.3. *Let $\Lambda = \tau(\mathcal{O}_K)g \in X$ and $v \in \Lambda \setminus \{0\}$. Then*

- (1) $H(v) \neq 0$
- (2) for $\xi \in \mathcal{O}_K^\times$, $H(\xi v) = H(v)$.

PROOF. For the first property, suppose $H(v) = 0$. Then $\|v^\sigma\| = \max\{|v_1^\sigma|, |v_2^\sigma|\} = 0$ for some σ . Since $v = \tau(a, b)g$ for $a, b \in \mathcal{O}_K$, and since g is invertible, we have $(\sigma(a), \sigma(b)) = (0, 0)$. Thus, $(a, b) = (0, 0)$ and hence $v = 0$. The other property follows from the product formula $\prod_{\sigma \text{ real}} |\sigma(\xi)| \prod_{\sigma \text{ complex}} |\sigma(\xi)|^2 = 1$ for units $\xi \in \mathcal{O}^*$. \square

The following lemma is essentially taken from the preprint [44] of Kleinbock and Tomanov.

LEMMA 2.4. *There exists a constant C such that if $v \in K_S^2$ with $H(v) \neq 0$ then there exists a unit ξ with*

$$C^{-1} H(v)^{\frac{1}{d}} \leq \|\xi v\| \leq C H(v)^{\frac{1}{d}}.$$

PROOF. Let

$$Z = \left\{ \mathbf{x} = (x^\sigma)_{\sigma \in S} \in \mathbb{R}_{>0}^{r+s} \mid \prod_{\sigma \text{ real}} x^\sigma \prod_{\sigma \text{ complex}} (x^\sigma)^2 = 1 \right\}.$$

The morphism $\xi \mapsto (|\sigma(\xi)|)_{\sigma \in S}$ sends \mathcal{O}_K^* to a subgroup of the multiplicative group Z . By the proof of the Dirichlet Unit Theorem, this is a cocompact lattice. Thus, there exists a constant C so that for any $(x^\sigma) \in Z$ there exists $\xi \in \mathcal{O}_K^*$ with

$$C^{-1} \leq |\sigma(\xi)| x^\sigma \leq C.$$

Let $v \in K_S^2$ with $H(v) \neq 0$. Then the vector $\left(\frac{\|v^\sigma\|}{H(v)^{1/d}} \right)$ is in Z . Applying the previous lemma, we have the claim. \square

PROPOSITION 2.5. *A subset $A \subset X$ is relatively compact if and only if there exists $\delta > 0$ such that for all $\Lambda \in A$ and for all nonzero vectors $v \in \Lambda = \tau(\mathcal{O}_K^2)g$, $H(v) > \delta$.*

PROOF. The first implication follows from the previous lemma along with Theorem 2.2. The reverse implication is immediate from continuity of H . \square

3. Dani's Correspondence

We prove a version of Dani's correspondence for number fields. As in the introduction we will consider here the notion of badly approximable vectors with equal weights, which as we will now show corresponds to the dynamics of the flow g_t .

PROPOSITION 3.1. *A vector $\mathbf{x} \in K_S$ is K -badly approximable, that is, there exists $c > 0$ such that for all $p, q \in \mathcal{O}_K$ with $q \neq 0$*

$$\max_{\sigma \in S} \{|\sigma(q)x^\sigma + \sigma(p)|\} \max_{\sigma \in S} \{|\sigma(q)|\} > c, \quad (9)$$

if and only if the trajectory

$$\left\{ \mathrm{SL}_2(\mathcal{O}_K) \left(\left(\begin{array}{cc} 1 & x^\sigma \\ & 1 \end{array} \right)_{\sigma \in S} g_t \mid t \geq 0 \right\} \quad (10)$$

in the quotient space $\mathrm{SL}_2(\mathcal{O}_K) \backslash \mathrm{SL}_2(K_S)$ is bounded.

PROOF. By Mahler's compactness criterion, we know that boundedness of the trajectory is equivalent to the existence of positive c' such that

$$\max_{\sigma \in S} \left\| \tau(q, p) \left(\left(\begin{array}{cc} 1 & x^\sigma \\ & 1 \end{array} \right) \left(\begin{array}{cc} e^{-t} & \\ & e^t \end{array} \right) \right)_{\sigma \in S} \right\| > c'$$

or equivalently

$$\max_{\sigma \in S} \|(e^{-t}\sigma(q), e^t(\sigma(q)x^\sigma + \sigma(p)))\| > c'$$

for all $t \geq 0$ and all nonzero pairs $(q, p) \in \mathcal{O}_K^2$.

Assume first that the orbit is unbounded. Then there exists for every $c' > 0$ some $t \geq 0$ and some nonzero vector $(q, p) \in \mathcal{O}_K^2$ with

$$\max_{\sigma \in S} \|(e^{-t}\sigma(q), e^t(\sigma(q)x^\sigma + \sigma(p)))\| < c'.$$

Note that since $t \geq 0$, that $q = 0$ would contradict this inequality (at least for small enough c' since $p \in \mathcal{O}_K$ cannot be small at all places $\sigma \in S$). Hence $q \neq 0$. By splitting the above inequality into two inequalities for the first and second coordinates of the vectors involved and taking the product we obtain

$$\max_{\sigma \in S} \{|\sigma(q)x^\sigma + \sigma(p)|\} \max_{\sigma \in S} \{|\sigma(q)|\} < (c')^2.$$

As c' was arbitrary we see that the vector $(x^\sigma)_{\sigma \in S}$ is not badly approximable.

Assume now that $\mathbf{x} = (x^\sigma)_{\sigma \in S}$ is not badly approximable. Then we have by definition that for every $c > 0$ there exists $p, q \in \mathcal{O}_K$ with $q \neq 0$ such that

$$\max_{\sigma \in S} \{|\sigma(p) + x^\sigma \sigma(q)|\} \max_{\sigma \in S} \{|\sigma(q)|\} < c.$$

We choose $t = -\frac{1}{2} \log c + \log \max_{\sigma \in S} \{|\sigma(q)|\}$, which will be positive if only c is sufficiently small (as the second summand is bounded from below for $q \in \mathcal{O}_K \setminus \{0\}$). Note that this gives $\max_{\sigma \in S} e^{-t} |\sigma(q)| = c^{\frac{1}{2}}$. Dividing our assumed inequality by the latter equality we also get $\max_{\sigma \in S} e^t |\sigma(p) + x^\sigma \sigma(q)| < c^{\frac{1}{2}}$. As c was arbitrary, this shows that the orbit is not bounded. \square

Notice that it is sufficient to have that the trajectory is bounded for a discrete sequence of times t_n where the consecutive differences are uniformly bounded.

4. A special case

In this section, we give proofs of Theorems 1.1 and 1.3 in a simplified linear case. In the next section, we show that a modification of this simplified argument suffices for the general C^1 case as well.

4.1. A special case of Theorem 1.1. We fix notation as in Theorem 1.1. Fix $0 < \alpha < \frac{1}{2}$ and let $0 < \beta < 1$. Notice that we choose α is independent of β as required by the definition of winning. Fix the base point $\text{SL}_2(\mathcal{O}_K)g \in \text{SL}_2(\mathcal{O}_K) \setminus \text{SL}_2(K_S)$ and denote by $\Lambda = \tau(\mathcal{O}_K^2)g$ the associated discrete \mathcal{O}_K -module viewed as a subset of \mathbb{R}^{2d} . In the simplified case we suppose that the function Φ is such that $\phi_\sigma(x) = a_\sigma x$ with $a_\sigma \in \mathbb{R}$ and $a_\sigma \neq 0$ for sufficiently many $\sigma \in S$ as required in the theorem, to be precise at least one half of the factors should satisfy $a_\sigma \neq 0$ with the complex places counting double. In this setting we will describe the strategy of Player B and prove that it is indeed winning.

The ultimate goal of Player B is to have the point x_∞ which remains at the end of the game satisfying that the set $\{v \in \Lambda\Phi(x_\infty)g_t \setminus \{0\} \mid t \geq 0\}$ is bounded away from zero. As remarked before, it is sufficient to have that $\{v \in \Lambda\Phi(x_\infty)g_{t_n} \setminus \{0\} \mid n \in \mathbb{N}\}$ is bounded away from zero where t_n is a positive sequence tending to infinity with bounded gaps. During each round of the game, Player B will want to monitor the short vectors that will appear in the module $\Lambda\Phi(x_n)g_{t_n}$ and play in such a way that these vectors are expanding under g_t after a finite amount of time (which needs to be independent of the short vector and the length of time the game has already been played). The initial step of the game by Player A may force a vector to be short for a long time, and the initial step of Player B will only make sure that the perturbed vector grows at some point in the future. However, for the later steps of the game it is important that we give a uniform lower bound on how short the perturbed vectors can become.

After the initial steps of the game, the n th round of the game plays out as follows: Player A has chosen a subinterval $A_n = B(x_n, \rho(\alpha\beta)^n) \subseteq B_{n-1}$. This corresponds to the collection of \mathcal{O}_K -modules

$$\{\Lambda\Phi(x_n + x)g_{t_n} \mid x \in B(0, \rho(\alpha\beta)^n)\}$$

where $t_n = \frac{1}{2} \log \frac{1}{\rho(\alpha\beta)^n}$. Player B focuses on the modules associated to the midpoint of A_n , namely $\Lambda\Phi(x_n)g_{t_n}$. If this module contains no short vectors, i.e. no nonzero $v \in \Lambda\Phi(x_n)g_{t_n}$ with $H(v) < 1$, then Player B chooses the new ball B_n heedlessly as allowed by the rules

of the game. Suppose there does exist a nonzero $v \in \Lambda\Phi(x_n)g_{t_n}$ with $H(v) < 1$. The strategy Player B employs in choosing the ball B makes use of the following phenomena²:

LEMMA 4.1. *Let $v = (v_1, v_2)$ and $w = (w_1, w_2)$ be two nonzero vectors of an \mathcal{O}_K -module $\Delta = \tau(\mathcal{O}_K^2)h \in X$. Suppose $H(v)H(w) < 1$. Then $Kv = Kw$.*

PROOF. We may write $v = (a_1, a_2)h$ and $w = (b_1, b_2)h$ with $a_1, a_2, b_1,$ and b_2 elements of \mathcal{O}_K . Recall that Δ admits a basis as an \mathcal{O}_K -module whose projection to each factor corresponding to some σ determines a parallelogram with area 1. Consider, then, the parallelograms formed by the projections of v and w to each factor. On one hand, the product of their areas (resp. the areas squared for the complex places) is given by

$$\begin{aligned} \prod_{\sigma} \left| \det \begin{pmatrix} v^{\sigma} \\ w^{\sigma} \end{pmatrix} \right|_{\sigma} &= \prod_{\sigma} \left| \det \begin{pmatrix} v_1^{\sigma} & v_2^{\sigma} \\ w_1^{\sigma} & w_2^{\sigma} \end{pmatrix} \right|_{\sigma} \\ &\leq \prod_{\sigma} \max\{|v_1^{\sigma}|_{\sigma}, |v_2^{\sigma}|_{\sigma}\} \max\{|w_1^{\sigma}|_{\sigma}, |w_2^{\sigma}|_{\sigma}\} \\ &= H(v)H(w) < 1. \end{aligned}$$

On the other hand, using the fact that $h \in \mathrm{SL}_2(K_S)$ and so $\det(h^{\sigma}) = 1$ for all σ , we have

$$\begin{aligned} \prod_{\sigma} \left| \det \begin{pmatrix} v^{\sigma} \\ w^{\sigma} \end{pmatrix} \right|_{\sigma} &= \prod_{\sigma} \left| \det \begin{pmatrix} a_1^{\sigma} & a_2^{\sigma} \\ b_1^{\sigma} & b_2^{\sigma} \end{pmatrix} \right|_{\sigma} \\ &= \prod_{\sigma} |\sigma(a_1b_2 - b_1a_2)|_{\sigma} \\ &= |N_{K|\mathbb{Q}}(a_1b_2 - b_1a_2)|, \end{aligned}$$

which is an integer as $a_1b_2 - b_1a_2 \in \mathcal{O}_K$. Therefore,

$$\prod_{\sigma} \det \begin{pmatrix} v^{\sigma} \\ w^{\sigma} \end{pmatrix} = 0,$$

implying that a and b are K -multiples of each other. As multiplication by elements of K on K_S^2 commutes with h , we see that also v and w are K -linearly dependent. \square

Thus in any round of the game Player B need only worry about a single K -span of short vectors. Define $\Phi_n(x) = g_{t_n}^{-1}\Phi(x)g_{t_n}$ so that $\Lambda\Phi(x_n + x)g_{t_n} = \Lambda\Phi(x_n)g_{t_n}\Phi_n(x)$ and the vectors in $\Lambda\Phi(x_n + x)g_{t_n}$ corresponding to the short vector $v \in \Lambda\Phi(x_n)g_{t_n}$ are $v\Phi_n(x)$. Player B

²This is where we use that our quotient has \mathbb{Q} -rank one.

wishes to choose B_n so that for $x \in B_n$ the neighbors $v\Phi_n(x)$ are all eventually expanding (or at least not further contracting) under g_{t_n} , where the notion of “eventually” depends only on α , β and Φ , but not on v , n or the game play up to this point. If the height of $v\Phi_n(x)$ is growing (in a uniform way for all $x \in B_n$) under g_{t_m} for $m \geq n$, then the same holds for all other vectors in $Kv\Phi_n(x)$. Moreover, by the lemma we know that for the period of time that the height of our vector is < 1 there cannot be any other vector outside the K -span with height < 1 . Also note that $\Phi_n(x)$ is uniformly bounded for $|x| < \rho(\alpha\beta)^n$, which implies that all vectors $w \in \Lambda\Phi(x_n)g_{t_n}\Phi_n(x)$ of sufficiently small norm are by the lemma in $Kv\Phi_n(x)$ and so controlled by the move of Player B.

Recall that under an application of g_t , v_1^σ is contracted by e^{-t} and v_2^σ is expanded by e^t . Consider the ratio between the expanding direction in the factor corresponding to some $\sigma \in S$ of $v\Phi_n(x)$ with $a_\sigma \neq 0$ and the norm of v^σ :

$$\frac{|v_1^\sigma a_\sigma e^{2t_n} x + v_2^\sigma|}{\|v^\sigma\|} = \left| a_\sigma (\rho(\alpha\beta)^n)^{-1} x \frac{v_1^\sigma}{\|v^\sigma\|} + \frac{v_2^\sigma}{\|v^\sigma\|} \right|,$$

where a_σ denotes the slope of Φ at σ . Also recall that x is restricted, at this stage in the game, to $x \in (-\rho(\alpha\beta)^n, \rho(\alpha\beta)^n)$. This shows that it is possible to choose x such that this ratio is bounded below by a constant $\varepsilon_\sigma > 0$ (depending on a_σ and α only). In fact, depending on the signs of v_1^σ and v_2^σ we can choose a subinterval $B \subset (-\rho(\alpha\beta)^n, \rho(\alpha\beta)^n)$ of radius α such that the ratio is uniformly bounded from below by α . We may choose $B = (\rho(\alpha\beta)^n(1 - 2\alpha), \rho(\alpha\beta)^n)$ if the real part of $\frac{v_2^\sigma}{v_1^\sigma}$ is nonnegative, or choose $B = (-\rho(\alpha\beta)^n, -\rho(\alpha\beta)^n(1 - 2\alpha))$ if the real part of $\frac{v_2^\sigma}{v_1^\sigma}$ is negative.

If we make this choice then it follows that the expanding component of $(v\Phi_n(x))^\sigma$ is of norm at least $\varepsilon_\sigma \|v^\sigma\|$ for all x in the new ball and this remains true in the future even if the vector initially is contracted. Moreover, if we know that the coordinate in the expanding direction is of norm at least $\varepsilon_\sigma \|v^\sigma\|$, then after time $t = |\log \varepsilon_\sigma|$, the expanding direction of the factor $(v\Phi_n(x)g_t)^\sigma$ will be at least as big as $\|v^\sigma\|$ for all $x \in B$.

Take $\varepsilon = \min_\sigma \varepsilon_\sigma$. If half of the factors with nonzero a_σ agree in the choice of B , then B is chosen to be the rightmost (resp., leftmost) subinterval of $(-\rho(\alpha\beta)^n, \rho(\alpha\beta)^n)$. After this Player A makes his move. Now we again look at all components and repeat the vote among those factors that voted differently the first time. After a uniformly bounded number of repetitions of this voting procedure, say m

repetitions (during which the game moves on) we have ensured that the norm of the expanding component v_1^σ is at least $\varepsilon e^{-2tm} \|v^\sigma\|$ for all places σ with $a_\sigma \neq 0$. Let $k = n + m$ and let B_k be the ball that is chosen by Player B at the last step, then it follows that the height of $v\Phi_k(x)g_t$ is uniformly bounded away from zero for all $x \in B_k$ and all $t \geq t_k$. Indeed, for all places $\sigma \in S$ with $a_\sigma \neq 0$ we ensured that the expanding direction is significant in size, and even if all the remaining directions are contracted by g_t the height will be bounded away from 0 (depending on ε and $H(v)$).

If during the above procedure (or later in the game) a new vector becomes of height less than one, its height is bounded away from zero by a constant depending on m . We then repeat the procedure with the new vector.

Notice that the strategy for Player B has not depended on the short vector v but on the direction of v . Moreover, this is indeed a winning strategy for Player B. The resulting point $x_\infty \in \bigcap A_n$ satisfies: For all $v \in \Lambda\Phi(x_\infty)$, either $H(vg_{t_n}) \geq 1$ for all n , or in the first round n with $H(vg_{t_n}) < 1$, we have that $H(vg_{t_{n+m}})$ is increasing for positive n , whence $H(vg_t)$ is bounded away from zero for all $t \geq t_n$. Thus, no vector in $\{v \in \Lambda\Phi(x_\infty)g_{t_n} \setminus \{0\} \mid n \in \mathbb{N}\}$ is ever shorter than a constant depending on $\alpha, \beta, \varepsilon$ and m , except for perhaps the short vectors which appear in $\Lambda\Phi(x_\infty)$ initially.

4.2. A counterexample. Note that we are crucially using the fact that $a_\sigma \neq 0$ for at least half the σ 's (counting complex places double). Indeed, suppose $d = r = 3$ and in two of the factors, we have $a_{\sigma_1} = a_{\sigma_2} = 0$. In this case, it may happen that the vector v considered above satisfies that v^{σ_1} and v^{σ_2} are contracted eigenvectors. However, in that case Player B will always lose — no matter of the choices in the game the height of the vector corresponding to v will go to zero. As we will see below, this behavior becomes even more significant in the unequal weights case.

4.3. Proof of Theorem 1.3 in a special case. The proof of Theorem 1.3, i.e. the weighted result for quadratic extensions in the linear case follows along the same general lines as above. We explain the strategy at the n -th stage following the notation in Theorem 1.3. In particular, we are now acting by $g(\mathbf{r})_t$. As before, Player A has chosen a subinterval $A_n = B(x_n, \rho(\alpha\beta)^n) \subseteq B_{n-1}$. This corresponds to the collection of \mathcal{O}_K -modules

$$\{\Lambda\Phi(x_n + x)g(\mathbf{r})_{t_n} \mid x \in B(0, \rho(\alpha\beta)^n)\}$$

where $t_n = \frac{1}{2r} \log \frac{1}{\rho(\alpha\beta)^n}$ and $r = \max_{\sigma \in S} r_\sigma$.

We assume that there exists nonzero $v \in \Lambda\Phi(x_n)g(\mathbf{r})_{t_n}$ with $H(v) < 1$. Let $\sigma_0 \in S$ be a place with $r_{\sigma_0} = r$. By Lemma 4.1, we again have to worry about at most one direction. Define $\Phi_n(x) = g(\mathbf{r})_{t_n}^{-1}\Phi(x)g(\mathbf{r})_{t_n}$ so that $\Lambda\Phi(x_n+x)g(\mathbf{r})_{t_n} = \Lambda\Phi(x_n)g(\mathbf{r})_{t_n}\Phi_n(x)$, denote by v the short vector and consider $v\Phi_n(x)$. Under an application of $g(\mathbf{r})_t$, $v_1^{\sigma_0}$ is contracted by e^{-rt} and $v_2^{\sigma_0}$ is expanded by e^{rt} . In the component corresponding to σ_0 consider the ratio between the expanding coordinate $v\Phi_n(x)$ and the norm of the vector:

$$\left| \frac{v_1^{\sigma_0} a_{\sigma_0} e^{2t_n r} x + v_2^{\sigma_0}}{\|v^{\sigma_0}\|} \right| = \left| a_{\sigma_0} (\rho(\alpha\beta)^n)^{-1} x \frac{v_1^{\sigma_0}}{\|v^{\sigma_0}\|} + \frac{v_2^{\sigma_0}}{\|v^{\sigma_0}\|} \right|$$

This ratio can be guaranteed to be larger than $|a_{\sigma_0}|(1 - 2\alpha)$ for all x in a subinterval $B \subset (-\rho(\alpha\beta)^n, \rho(\alpha\beta)^n)$ of radius α , choosing $B = (\rho(\alpha\beta)^n(1 - 2\alpha), \rho(\alpha\beta)^n)$ if $\frac{v_2^{\sigma_0}}{v_1^{\sigma_0}}$ is nonnegative, or choosing $B = (-\rho(\alpha\beta)^n, -\rho(\alpha\beta)^n(1 - 2\alpha))$ if $\frac{v_2^{\sigma_0}}{v_1^{\sigma_0}}$ is negative. (If $v_1^{\sigma_0} = 0$ then the vector v^{σ_0} is an expanding eigenvector already anyway and the ratio is one.)

We now argue as before. As the modified vector $v\Phi_n(x)$ for x in the new subinterval B has a significant expanding component in the place σ_0 and since this place is the one with the faster dynamics, it follows that the height of $v\Phi_n(x)g(\mathbf{r})_t$ will be ≥ 1 for $x \in B$ and for $t \geq t_0 = t_0(a_{\sigma_0}, \alpha, \mathbf{r})$. As before, since $\Phi_n(x)$ is bounded no vector in $\Lambda\Phi(x_n+x)g(\mathbf{r})_{t_n} = \Lambda\Phi(x_n)g(\mathbf{r})_{t_n}\Phi_n(x)$ will have height much smaller than v . Once more we obtain a winning strategy for Player B.

In this case, once more the lower bound on derivatives which amounts to $a_\sigma \neq 0$ for *both* σ is crucial. If not, we may consider a similar example to the one in §4.2, i.e. suppose $a_{\sigma'} = 0$ for one factor. If now $\sigma' = \sigma_0$ for some weight \mathbf{r} , then the above strategy fails. Moreover, in that case it may happen that our lattice contains a vector $v = (v^{\sigma_0}, v^{\sigma_1})$ with v^{σ_0} being contracted by $e^{r_{\sigma_0}t}$ under the dynamics of $g(\mathbf{r})_t$. Even if v^{σ_1} is now expanded by $e^{r_{\sigma_1}t}$, then height of the vector $vg(\mathbf{r})_t$ will still go to zero and this remains true for every possible outcome of the game. Hence in this case there cannot exist a winning strategy.

5. Proof in the general case

We now discuss the proof of Theorem 1.1 in the general case of ϕ in C^1 . Recall that we have assumed that we have a subset $S' \subset S$ with $|\{\sigma \in S' \mid \sigma \text{ is real}\}| + 2|\{\sigma \in S' \mid \sigma \text{ is complex}\}| > \lfloor \frac{d}{2} \rfloor$ and for each $\sigma \in S'$, only finitely many points x have $\phi'_\sigma(x) = 0$. Here we use linear approximations to ϕ , and the strategy explained above for linear functions. At the beginning of the game, Player B acts by moving the

playing field away from all points x with $\phi'_\sigma(x) = 0$ for any $\sigma \in S'$. Since this is assumed to be a finite set, this takes only finitely many rounds of the game. Thus, the game arrives at round N in the situation that for all $\sigma \in S'$, $\phi'_\sigma(x) \neq 0$ for all $x \in \overline{B_N}$. Since ϕ'_σ is uniformly continuous on $\overline{B_N}$, there exists for each $\sigma \in S'$, $m_\sigma, M_\sigma \in \mathbb{R}$ such that $0 \notin [m_\sigma, M_\sigma]$ while for all $x \in B_N$, $\phi'_\sigma(x) \in [m_\sigma, M_\sigma]$. Player B will use these bounds to produce the piecewise linear approximation to ϕ .

Suppose that in round $n > N$ there is a nonzero vector $v \in \Lambda\Phi(x_n)$ with $H(vg_{t_n}) < 1$. (Again we choose $t_n = \frac{1}{2} \log \rho(\alpha\beta)^{-n}$.) Define

$$\hat{\Phi}_n(x) = g_{t_n}^{-1} \left(\begin{pmatrix} 1 & \phi_\sigma(x_n + x) - \phi_\sigma(x_n) \\ 0 & 1 \end{pmatrix} \right)_\sigma g_{t_n}.$$

Then as before Player B would like to choose B_n so that the ‘‘angle’’ between $v\hat{\Phi}_n(x)$ and the contracting direction is significant. Thus for $v\hat{\Phi}_n(x) \in \Lambda\Phi(x_n + x)g_{t_n}$ and $\sigma \in S'$, Player B considers the ratio

$$\begin{aligned} & \left| \frac{v_1^\sigma (\phi_\sigma(x_n + x) - \phi_\sigma(x_n)) e^{2t_n} + v_2^\sigma}{\|v^\sigma\|} \right| \\ &= \left| (\phi_\sigma(x_n + x) - \phi_\sigma(x_n)) (\rho(\alpha\beta)^n)^{-1} \frac{v_1^\sigma}{\|v^\sigma\|} + \frac{v_2^\sigma}{\|v^\sigma\|} \right|, \end{aligned}$$

and wishes to bound this quantity from below uniformly over x in the yet to be determined B_n . Since ϕ is monotone, let us first suppose that ϕ is increasing, that is, we assume $0 < m_\sigma \leq M_\sigma$. Then on $[0, \rho(\alpha\beta)^n]$ we have that $\phi_\sigma(x_n + x) - \phi_\sigma(x_n) \geq m_\sigma x$ and on $(-\rho(\alpha\beta)^n, 0)$, $\phi_\sigma(x_n + x) - \phi_\sigma(x_n) \geq M_\sigma x$. Using these linear approximations of ϕ , Player B uses the same strategy as before. If the real part of $\frac{v_2^\sigma}{v_1^\sigma}$ is nonnegative, then for x in $((1 - 2\alpha)\rho(\alpha\beta)^n, \rho(\alpha\beta)^n)$,

$$\left| (\phi_\sigma(x_n + x) - \phi_\sigma(x_n)) (\rho(\alpha\beta)^n)^{-1} \frac{v_1^\sigma}{\|v^\sigma\|} + \frac{v_2^\sigma}{\|v^\sigma\|} \right| \geq m_\sigma (1 - 2\alpha).$$

Similarly, if the real part of $\frac{v_2^\sigma}{v_1^\sigma}$ is negative, then for x in $(-\rho(\alpha\beta)^n, -(1 - 2\alpha)\rho(\alpha\beta)^n)$,

$$\left| (\phi_\sigma(x_n + x) - \phi_\sigma(x_n)) (\rho(\alpha\beta)^n)^{-1} \frac{v_1^\sigma}{\|v^\sigma\|} + \frac{v_2^\sigma}{\|v^\sigma\|} \right| \geq M_\sigma (1 - 2\alpha).$$

Player B uses a similar analysis in the case that ϕ_σ is decreasing. So for fixed $\sigma \in S'$, Player B can choose B_n so that for all $x \in B_n$ the ratio between the coordinates in the factor corresponding to σ of any neighbor $v\hat{\Phi}_n(x)$ is greater than $\inf_{y \in B_N} |\phi'_\sigma(y)| (1 - 2\alpha)$.

As before, Player B should execute this strategy over several, say m , rounds of the game to ensure that $H(v\hat{\Phi}_n(x)g_t)$ is increasing for all

$t > t_m + |\log \varepsilon|$, where $\varepsilon = \min_{\sigma \in S'} \inf_{y \in \overline{B_N}} |\phi'_\sigma(y)|(1 - 2\alpha)$. The same reasoning as in the special case shows that this strategy is winning. The same strategy works for the more general case of Theorem 1.3.

CHAPTER 3

Badly approximable affine forms, fractals and number fields

BEVERLY LYTLE

ABSTRACT. We show that for a fixed vector, the set of inhomogeneous linear forms which are badly approximable by the rationals of a fixed number field form a winning set (in the sense of Schmidt), even when restricted to a fractal. In addition, we show an analogous result for a fixed matrix rather than a fixed vector.

1. Introduction

In diophantine approximation, one studies quantitatively the density of the rational points in a given space. For example, a classical theorem of Dirichlet states that no point of \mathbb{R} is too far away from some point of \mathbb{Q} . Specifically, for any $x \in \mathbb{R}$, there exist infinitely many $q \in \mathbb{Z}$ such that $|x - \frac{p}{q}| < q^{-2}$ for some $p \in \mathbb{Z}$. A robust area of study has developed from this statement investigating generalizations of this statement to other spaces and to pushing the bounds on the rate of approximation (for example, see [1, 14, 25, 41, 51, 62]). This article studies the inhomogeneous linear systems which are badly approximable, not by rationals, but rather by integers in an algebraic number field.

We begin by recalling the definitions of rationally badly approximable inhomogeneous forms. For $n, m \in \mathbb{N}$, denote by $M_{n,m}(\mathbb{R})$ the set of $m \times n$ real matrices and by $\tilde{M}_{n,m}(\mathbb{R}) = M_{n,m}(\mathbb{R}) \times \mathbb{R}^m$, the set of inhomogeneous linear systems. Consider the subset

$\text{Bad}(m, n) = \{\langle A, \mathbf{b} \rangle \in \tilde{M}_{n,m}(\mathbb{R}) \mid \text{there exists } c > 0 \text{ such that}$

$$\|A\mathbf{q} - \mathbf{b} + \mathbf{p}\| \geq \frac{c}{\|q\|^{n/m}} \text{ for all } \mathbf{q} \in \mathbb{Z}^n \setminus \{0\} \text{ and } \mathbf{p} \in \mathbb{Z}^m\},$$

where $\|\cdot\|$ denotes the supremum norm, and the related sets $\text{Bad}^{\mathbf{b}}(m, n) = \{A \in M_{n,m}(\mathbb{R}) \mid \langle A, \mathbf{b} \rangle \in \text{Bad}(m, n)\}$ for fixed $\mathbf{b} \in \mathbb{R}^m$ and $\text{Bad}_A(m, n) = \{\mathbf{b} \in \mathbb{R}^m \mid \langle A, \mathbf{b} \rangle \in \text{Bad}(m, n)\}$ for fixed $A \in M_{n,m}(\mathbb{R})$. In [25], Einsiedler and Tseng showed that while these sets

have zero Lebesgue measure, they are still large in the sense that they have full Hausdorff dimension, even when restricted to sufficiently nice fractals within $M_{n,m}(\mathbb{R})$ and \mathbb{R}^m , respectively. In fact, they prove the stronger result that the sets are winning in the sense of Schmidt. Winning sets are not only of full Hausdorff dimension within these nice fractals, but they also satisfy a countable intersection property.

The aim of this article is to prove a generalization of this result, where the approximations of the affine forms will be by the ring of integers \mathcal{O}_k of an algebraic number field k rather than the integral points \mathbb{Z} of \mathbb{Q} . Several authors have developed analogues of the standard theorems of diophantine approximation, e.g. Dirichlet's theorem and the existence of k -badly approximable numbers, in [13, 36, 56].

Let S be the set of all embeddings σ of k into \mathbb{R} or \mathbb{C} where we choose one of σ and $\bar{\sigma}$ for each complex embedding (and where we will identify \mathbb{C} with \mathbb{R}^2 wherever appropriate). Let d be the degree of k over \mathbb{Q} with $d = r + 2s$ where r is the number of real embeddings of k and s the number of complex embeddings. Denote by $k_S = k \otimes \mathbb{R} \simeq \mathbb{R}^r \oplus \mathbb{C}^s$. Note then that $\tilde{M}_{n,m}(k_S)$ is the set of d -tuples of pairs $\langle A_\sigma, \mathbf{b}_\sigma \rangle$ with, depending on σ , each pair in $\tilde{M}_{n,m}(\mathbb{R})$ or $\tilde{M}_{n,m}(\mathbb{C})$, and similarly for $M_{n,m}(k_S)$. We define

$$k\text{-Bad}(m, n) = \{(\langle A_\sigma, \mathbf{b}_\sigma \rangle) \in \tilde{M}_{n,m}(k_S) \mid \exists c > 0 \text{ such that} \\ \max_S \{\|\mathbf{q}^\sigma\|^{\frac{n}{m}}\} \max_S \{\|A_\sigma \mathbf{q}^\sigma - \mathbf{b}_\sigma + \mathbf{p}^\sigma\|\} > c \forall \mathbf{q} \in \mathcal{O}_k^n \setminus \{0\}, \mathbf{p} \in \mathcal{O}_k^m\}$$

where for a vector \mathbf{x} with entries in \mathcal{O}_k and $\sigma \in S$, \mathbf{x}^σ denotes the coordinatewise image of \mathbf{x} under σ . As above we define for fixed $(\mathbf{b}_\sigma) \in (k_S)^m$,

$$k\text{-Bad}^{(\mathbf{b}_\sigma)}(m, n) = \{(A_\sigma) \in M_{n,m}(k_S) \mid (\langle A_\sigma, \mathbf{b}_\sigma \rangle) \in k\text{-Bad}(m, n)\},$$

and again, for fixed $(A_\sigma) \in M_{n,m}(k_S)$, set

$$k\text{-Bad}_{(A_\sigma)}(m, n) = \{(\mathbf{b}_\sigma) \in k_S^m \mid (\langle A_\sigma, \mathbf{b}_\sigma \rangle) \in k\text{-Bad}(m, n)\}.$$

These two types of sets are the primary concern here. Using methods similar to [25] it will be shown that they are winning and hence of full Hausdorff dimension even when restricted to certain fractals.

We also mention the set of infinitely k -badly approximable affine systems. For fixed $(A_\sigma) \in M_{n,m}(k_S)$, we define this set as

$$k\text{-Bad}_{(A_\sigma)}^\infty(m, n) = \\ \{\mathbf{b}_\sigma \in k_S^m \mid \liminf_{\mathbf{q} \in \mathcal{O}_k^n \setminus \{0\}, \mathbf{p} \in \mathcal{O}_k^m} \max_S \{\|\mathbf{q}^\sigma\|\}^n \max_S \{\|A_\sigma \mathbf{q}^\sigma - \mathbf{b}_\sigma + \mathbf{p}^\sigma\|\}^m = \infty\}.$$

Notice that $k\text{-Bad}_{(A_\sigma)}^\infty(m, n) \subset k\text{-Bad}_{(A_\sigma)}(m, n)$. Such a set can only be non-empty for very particular (A_σ) . Indeed the matrices must be

singular in the sense of diophantine approximation, that is, for all $\varepsilon > 0$ and large enough N there exist solutions $\mathbf{q} \in \mathcal{O}_k^n$ and $\mathbf{p} \in \mathcal{O}_k^m$ to

$$\max_S \{\|A_\sigma \mathbf{q}^\sigma + \mathbf{p}^\sigma\|\} \leq \frac{\varepsilon}{N^{n/m}} \quad \text{and} \quad 0 < \max_S \{\|\mathbf{q}^\sigma\|\} < N.$$

A statement about this type of set will follow as a corollary to the main theorem.

The paper is structured as follows. In Section 2, we give the statements of the results proven in this paper. We introduce, in Section 3, the background material including Schmidt's game, properties of the measures involved and the construction of the space of special affine \mathcal{O}_k -modules and its properties. In the fourth section, we give a proof of the first and second of our theorems on k -badly approximable inhomogeneous forms where the matrix is fixed. Section 5 contains a proof of the third theorem on k -badly approximable vectors in the homogeneous setting and Section 6 builds on this proof to give a proof relating to k -badly approximable vectors in the inhomogeneous setting.

2. Statement of results

The results of this paper are listed in this section. All of the properties absolutely friendly, absolutely decaying, and δ -fitting of measures are defined in Section 3. For the moment, one may think of measures with support given by a "nice" fractal, for example Lebesgue measure or the standard measure on the Cantor set. Throughout, \dim refers to Hausdorff dimension. We begin with the statements of the size of the k -badly approximable inhomogeneous forms where the matrix is fixed.

THEOREM 2.1. *Let $F \subset k_S^m$ be the support of an absolutely decaying μ . Then for any $(A_\sigma) \in M_{nm}(k_S)$,*

$$k\text{-Bad}_{(A_\sigma)}(m, n) \cap F$$

is a winning set on F .

If our attention is restricted to singular matrices (as defined in the previous section), we are able to modify the proof of theorem 2.1 and extend it to a proof of the following theorem.

THEOREM 2.2. *Let $F \subset k_S^m$ be the support of an absolutely decaying μ . Then for any singular $(A_\sigma) \in M_{nm}(k_S)$,*

$$k\text{-Bad}_{(A_\sigma)}^\infty(m, n) \cap F$$

is a winning set on F .

Now we turn our attention to the set of k -badly approximable inhomogeneous linear forms for a fixed vector. Unfortunately, these statements are restricted to low dimensions. First we begin with the homogeneous case.

THEOREM 2.3. *Let $F \subset M_{1,1}(k_S)$ be the support of an absolutely decaying measure μ . Then,*

$$k\text{-Bad}^0(1, 1) \cap F$$

is a winning set on F .

Since Lebesgue measure is absolutely friendly, we immediately have the corollary:

COROLLARY 2.4. *The set $k\text{-Bad}^0(1, 1)$ is winning on k_S .*

This is a strengthening of the conclusion in [21] that $k\text{-Bad}^0(1, 1)$ (or in the notation used in [21], $\text{Bad}(K)$) is of full Hausdorff dimension. We are able to use the proof of Theorem 2.3 to obtain the stronger inhomogeneous statement.

THEOREM 2.5. *Let $F \subset M_{1,1}(k_S)$ be the support of an absolutely decaying measure μ . Then for any $(\mathbf{b}_\sigma) \in k_S$,*

$$k\text{-Bad}^{(\mathbf{b}_\sigma)}(1, 1) \cap F$$

is a winning set on F .

COROLLARY 2.6. *Let $F \subset M_{1,1}(k_S)$ be the support of an absolutely decaying measure μ . Then for any countable sequence $(\mathbf{b}_{\sigma,i}) \in k_S$,*

$$\left(\bigcap_i k\text{-Bad}^{(\mathbf{b}_{\sigma,i})}(1, 1)\right) \cap F$$

is a winning set on F . If in addition μ is $\dim(F)$ -fitting and absolutely decaying, then

$$\left(\bigcap_i k\text{-Bad}^{(\mathbf{b}_{\sigma,i})}(1, 1)\right) \cap F$$

has Hausdorff dimension equal to that of F .

This statement is analogous to Kleinbock's conjecture in [41] which was proven in [25]. It follows from the properties of Schmidt's games in [47] and from [30], and the fact that the winning parameter for Theorem 2.5 is independent of the fixed vector.

3. Background

3.1. Schmidt games and winning sets. In [59], W. Schmidt introduced the following game now bearing his name. Let M be a complete metric space, and for any point $x \in M$, denote by $B(x, r)$ the closed ball of radius r . This game is played by two opponents, say Alice and Bob, with each alternately choosing a closed metric ball of a prescribed radius contained within the previously chosen ball. The rules of the game are determined by two parameters $0 < \alpha < 1$ and $0 < \beta < 1$. This (α, β) -game begins with Bob choosing a point x_0 and a radius ρ determining $B_0 = B(x_0, \rho)$. The n th round of the game begins with Alice choosing a point y_n such that $A_n = B(y_n, \alpha(\alpha\beta)^{n-1}\rho)$ is contained in $B_{n-1} = B(x_{n-1}, (\alpha\beta)^{n-1}\rho)$. Then Bob chooses a point x_n so that $B_n = B(x_n, (\alpha\beta)^n\rho)$ is contained in A_n . This produces a nested sequence of nonempty closed balls of shrinking radius. We say a set $W \subset M$ is (α, β) -winning if Alice can always find a strategy so that the point remaining in $\bigcap B_n$ lies in W . We call the set W winning if there exists α so that W is (α, β) -winning for all β .

In [51], McMullen introduced a variant game and the notion of strong winning. The game is played with the same rules as Schmidt's original game with one change. The radii of the balls chosen by Alice and Bob are not fixed by the game. Instead, in round n , Alice chooses a ball $A_n \subset B_{n-1}$ with radius of size at least as big as α times the radius of B_{n-1} . Similarly, Bob chooses a ball $B_n \subset A_n$ with radius of size at least as big as β times the radius of A_n . The definitions of (α, β) -strong winning and α -strong winning follow analogously. It is clear that strong winning implies winning.

The following properties of winning sets (hence also of strong winning sets) are relevant here [30, 42, 59]:

- Countable intersections of α -winning sets are again α -winning.
- For a set F which is the support of an absolutely friendly and $\dim(F)$ -fitting measure, winning subsets of F have full Hausdorff dimension in F .

Complete definitions concerning the measures will be given in the next subsection.

3.2. Fractals. For an affine hyperplane H of \mathbb{R}^n and for $\varepsilon > 0$, denote by $H^{(\varepsilon)}$ the ε -thickening of H . Let μ be a locally finite Borel measure on \mathbb{R}^n , and let $B(x, r)$ be the closed euclidean ball about x of radius r . We say that μ is absolutely decaying if there exist positive constants C_0 , η and r_0 such that for any affine hyperplane H , $\varepsilon > 0$,

$x \in \text{supp}(\mu)$ and $r < r_0$,

$$\mu(B(x, r) \cap H^{(\varepsilon)}) \leq C_0 \left(\frac{\varepsilon}{r}\right)^\eta \mu(B(x, r)).$$

We call μ Federer (or doubling) if there exist positive constants C_1 and r_1 such that for any $x \in \text{supp}(\mu)$ and $r < r_1$

$$\mu\left(B\left(x, \frac{1}{2}r\right)\right) > C_1\mu(B(x, r)).$$

Finally, μ is absolutely friendly if it is both absolutely decaying and Federer.

Again, we take M to be a metric space. For $x \in M$ and real numbers $r > 0$ and $0 < \beta < 1$, let $N_M(\beta, x, r)$ be the maximal number of disjoint balls of radius βr contained in $B(x, r)$. For $\delta > 0$, we call μ δ -fitting if there are positive constants $r_2 \leq 1$ and C_2 such that for every $0 < r \leq r_2$, $0 < \beta < 1$, and $x \in \text{supp}(\mu)$,

$$N_{\text{supp}(\mu)}(\beta, x, r) \geq C_2\beta^{-\delta}.$$

Examples of sets in \mathbb{R}^n which are the support of an absolutely friendly and fitting measure include the Cantor set, the Koch curve and the attractor of an irreducible family of contracting similarity maps of \mathbb{R}^k satisfying the open set condition [30, 42].

3.3. Affine \mathcal{O}_k -modules. In this section we will construct for the algebraic number field k the relevant homogeneous space identifiable with the set of affine \mathcal{O}_k -modules of rank $n + m$ satisfying a special property. This construction proceeds analogously to the construction of the space of affine unimodular lattices in \mathbb{R}^n . First recall that the twisted embedding of \mathcal{O}_k into \mathbb{R}^d (where d is the degree of k over \mathbb{Q}) given by

$$q \in \mathcal{O}_k \mapsto (q^\sigma)_{\sigma \in S}$$

in fact defines a lattice in \mathbb{R}^d of covolume $2^{-s}|\Delta_k|^{1/2}$, where s is the number of complex embeddings of k and Δ_k is the discriminant of k . Naturally, this embedding extends to higher dimensions $\mathcal{O}_k^{n+m} \hookrightarrow (\mathbb{R}^d)^{n+m}$. Furthermore, with restriction of scalars, we have $\text{SL}_{n+m}(k_S) = \text{SL}_{n+m}(\mathbb{R})^r \times \text{SL}_{n+m}(\mathbb{C})^s$. Then $\text{SL}_{n+m}(k_S)$ acts naturally on $\mathcal{O}_k^{n+m} \subset (\mathbb{R}^d)^{n+m}$, with stabilizer $\text{SL}_{n+m}(\mathcal{O}_k)$. The collection $\text{SL}_{n+m}(k_S)\mathcal{O}_k^{n+m}$ consists of all discrete (as a subset of $(\mathbb{R}^d)^{n+m}$) \mathcal{O}_k -modules Λ of rank $n + m$ with the following property: There exists an \mathcal{O}_k -basis v_1, v_2, \dots, v_{n+m} of Λ such that for each $\sigma \in S$, $\{\sigma(v_i)\}$ form the edges of a parallelotope of covolume 1 within \mathbb{R}^{n+m} (or $\mathbb{R}^{2(n+m)}$ as appropriate).

Since we will be dealing with affine modules, it makes sense to view these objects as subsets of $(\mathbb{R}^{n+m+1})^d$ sitting at level set 1, that is, we identify $(\mathbb{R}^{n+m})^d$ with $(\mathbb{R}^{n+m} \times \{1\})^d$. The reference module is (the twisted embedding of) $\mathcal{O}_k^{n+m} \times \{1\}$. From now on, we will use the shorthand \mathcal{O}_k^{n+m} and $(\mathbb{R}^{n+m})^d$ for the identified subsets of $\mathcal{O}_k^{n+m} \times \{1\}$ and $(\mathbb{R}^{n+m} \times \{1\})^d$.

Let $(\langle A_\sigma, \mathbf{b}_\sigma \rangle)$ be a d -tuple of m inhomogeneous linear forms in n variables over \mathbb{R} or \mathbb{C} as appropriate. We wish to encode the diophantine information coming from the field k of this system in an affine \mathcal{O}_k -module. For $\sigma \in S$, define the matrices

$$L_{A_\sigma}(b_\sigma) = \begin{pmatrix} I_m & A_\sigma & -\mathbf{b}_\sigma \\ & I_n & \\ & & 1 \end{pmatrix}$$

where I_m and I_n denote the $m \times m$ and $n \times n$ identity matrices, respectively. Then we have the associated discrete affine \mathcal{O}_k -module

$$(L_{A_\sigma}(\mathbf{b}_\sigma)) \mathcal{O}_k^{n+m},$$

with general element

$$(L_{A_\sigma}(\mathbf{b}_\sigma)) \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \\ 1 \end{pmatrix} = \left(\begin{pmatrix} A_\sigma \mathbf{q}^\sigma - \mathbf{b}_\sigma + \mathbf{p}^\sigma \\ \mathbf{q}^\sigma \\ 1 \end{pmatrix} \right).$$

These are precisely the quantities we are interested in analyzing.

Finally, we define for $t \in \mathbb{R}$ the matrices

$$g_t = \begin{pmatrix} e^{t/m} I_m & & \\ & e^{-t/n} I_n & \\ & & 1 \end{pmatrix}.$$

For $t \in \mathbb{R}$, the d -tuple $\mathbf{g}_t = (g_t, \dots, g_t)$ acts on the space $(\mathbb{R}^{m+n+1})^d$ leaving our copy of $(\mathbb{R}^{m+n})^d$ invariant. Indeed, it acts on the space of affine \mathcal{O}_k -modules. Notice also that the matrices \mathbf{g}_t expand the subspace $(\mathbb{R}^m \times \{0\}^n)^d$ and contracts the subspace $(\{0\}^m \times \mathbb{R}^n)^d$.

These objects are useful to us so that we can make use of the following version of a theorem of Dani [15, 21]:

THEOREM 3.1. *The system $(\langle A_\sigma, \mathbf{b}_\sigma \rangle)$ is in k -Bad(n, m) if and only if all of the nonzero points of the affine lattices $\mathbf{g}_t(L_{A_\sigma}(\mathbf{b}_\sigma))\mathcal{O}_k^{n+m}$ for $t > 0$ are uniformly bounded away from the origin of $(\mathbb{R}^{n+m})^d$.*

Notice that the statement remains true if instead of considering the entire trajectory under $\{\mathbf{g}_t \mid t > 0\}$, one looks only at the lattices corresponding to a discrete time sample with the gaps between any two consecutive sample points uniformly bounded above.

To compress notation, we will refer to a general full rank \mathcal{O}_k -module in $(\mathbb{R}^{n+m})^d$ as Λ and its affine translates as $\Lambda + \mathbf{c}$. We say a subspace $V \subset (\mathbb{R}^{n+m})^d$ is Λ -rational if $\Lambda \cap V$ spans V .

Let P be an $d(n+m) - 1$ -dimensional parallelotope, and denote by $|P|$ its $d(n+m) - 1$ -dimensional volume. For H , a Λ -rational hyperplane, we will also write $|H|$ for the volume of the parallelotope spanned by a basis for H in $H \cap \Lambda$. We may abuse this notation for smaller dimensional objects. Define

$$\xi_0 = \frac{2}{\pi} \Gamma\left(\frac{d(m+n)}{2} + 1\right)^{\frac{1}{d(n+m)}} (2^{-s} |\Delta_k|^{1/2})^{\frac{(n+m)d-1}{(n+m)d}},$$

where s is the number of complex places of k . We call a hyperplane H small if $|H| \leq \xi_0$ and big otherwise. We make this choice of ξ_0 in order to streamline some sections of the proofs, however the choice is essentially arbitrary.

We will be applying the flow \mathbf{g}_t to various Λ -rational hyperplanes H throughout this article, and will be interested in the deformation of $|\mathbf{g}_t H|$ as $t \rightarrow \infty$. It will not be explicitly written, but the measurement $|\mathbf{g}_t H|$ is with respect to the module $\mathbf{g}_t \Lambda$. We observe that for any hyperplane H , the volume $|\mathbf{g}_t H|$ tends to zero or infinity as $t \rightarrow \infty$. To this this, we consider the exterior product $\bigwedge^{d(n+m)-1} \mathbb{R}^{d(n+m)}$ with the associated action $\bigwedge^{d(n+m)-1} \mathbf{g}_t$. The space can be thought of as the space of volume elements of hyperplanes contained in $\mathbb{R}^{d(n+m)}$ acted on by \mathbf{g}_t . Let $\{e_{i,j}\}$ for $1 \leq i \leq d$ and $1 \leq j \leq n+m$ be a basis of eigenvectors of $(\mathbb{R}^{n+m})^d$ for the action of \mathbf{g}_t so that the contracted subspace $(\{0\}^m \times \mathbb{R}^n)^d$ is spanned by $\{e_{i,j}\}$ for $1 \leq i \leq d$ and $m+1 \leq j \leq n+m$. This gives a basis $\{e^{i,j}\}$ of $\bigwedge^{d(n+m)-1} \mathbb{R}^{d(n+m)}$ where $e^{i,j} = e_{1,1} \wedge \cdots \wedge \hat{e}_{i,j} \wedge \cdots \wedge e_{d,n+m}$ is the wedge product of all basis elements of $\mathbb{R}^{d(n+m)}$ except $e_{i,j}$. Then as an element of the exterior product H may be represented as $\sum_{i=1}^d \sum_{j=1}^{n+m} \alpha_{i,j} e^{i,j}$ with $|H|$ equal to the norm of this vector. Then

$$\begin{aligned} |\mathbf{g}_t H| &= \left\| \bigwedge \mathbf{g}_t \sum_{i=1}^d \sum_{j=1}^{n+m} \alpha_{i,j} e^{i,j} \right\| \\ &= \left\| \sum_{i=1}^d \sum_{j=1}^m e^{-t/m} \alpha_{i,j} e^{i,j} + \sum_{i=1}^d \sum_{j=m+1}^{m+n} e^{t/n} \alpha_{i,j} e^{i,j} \right\|. \end{aligned}$$

Hence if $\alpha_{i,j} = 0$ for all $1 \leq i \leq d$ and $m+1 \leq j \leq n+m$, that is, if H contains the contracted subspace, then $|\mathbf{g}_t H| \rightarrow 0$ as $t \rightarrow \infty$. If just one $\alpha_{i,j} \neq 0$ for i, j as before, then $|\mathbf{g}_t H| \rightarrow \infty$ as $t \rightarrow \infty$.

Another geometric lemma which will be of use concerns the distance between nearest cosets of rational hyperplanes.

LEMMA 3.2. *Let Λ be a full rank \mathcal{O}_k -module of $(\mathbb{R}^{n+m})^d$, and suppose H is a Λ -rational hyperplane. Then the distance between any two nearest cosets $H + \mathbf{v}$ and $H + \mathbf{w}$ with $\mathbf{v}, \mathbf{w} \in \Lambda$ is equal to $1/|H|$.*

We remark that this Lemma encodes the ‘‘co-rank’’ 1 phenomenon of this setting: For any set of $d(n+m)$ linearly independent vectors in Λ , at least one of these vectors must have large norm. This is the central idea of the proof to follow.

One final observation about small hyperplanes will be of use, which is that they exist and that there are only finitely many of them. As well-known (see e.g. [21]), the full rank \mathcal{O}_k -modules are indeed lattices in $\mathbb{R}^{(n+m)d}$, that is, discrete subgroups with finite covolume, and it is convenient to think of them as such. Given a lattice $\Lambda \subset \mathbb{R}^{(n+m)d}$ of covolume $V = 2^{-s}|\Delta_k|^{1/2}$, consider the image of Λ inside the exterior product $\bigwedge^{d(n+m)-1} \mathbb{R}^{d(n+m)}$, that is, the set $\{v_1 \wedge \cdots \wedge v_{d(n+m)-1} \mid v_i \in \Lambda\}$. This is, again, a lattice, but with covolume $V^{d(n+m)-1}$. (This can be seen by viewing $\Lambda = g\mathbb{Z}^{(n+m)d}$ for some matrix g with determinant V and noting that the volume of $\bigwedge^{d(n+m)-1} \Lambda = \bigwedge^{d(n+m)-1} g\mathbb{Z}^{(n+m)d}$ is $\det(\bigwedge^{d(n+m)-1} g) = V^{d(n+m)-1}$). By Minkowski’s convex body theorem, there must be vectors $w \in \bigwedge^{d(n+m)-1} \Lambda$ with

$$\|w\| \leq \frac{2}{\pi} \Gamma\left(\frac{(n+m)d}{2} + 1\right)^{\frac{1}{(n+m)d}} V^{\frac{(n+m)d-1}{(n+m)d}} = \xi_0.$$

Furthermore, since $\bigwedge^{d(n+m)-1} \Lambda$ is a lattice, there may be only finitely many such vectors. However, there is no upper bound on the number of such vectors.

4. Proof of fixed matrix case

Fix $(A_\sigma) \in M_{n,m}(k_S)$. Let μ be an absolutely decaying, Federer measure on k_S^m with associated constants C_0, C_1, η , and $\tilde{r} = \min\{r_0, r_1\}$ as given in the definitions in Section 3.2, and set $F = \text{supp}(\mu)$. We wish to show that the set of $(\mathbf{b}_\sigma) \in F$ for which $(\langle A_\sigma, \mathbf{b}_\sigma \rangle)$ is k -badly approximable is winning. We will describe the winning strategy that Alice will employ as Alice and Bob play their game on the metric space F . Let $0 < \beta < 1$, and choose $\alpha < (4(2\xi_0 C_0)^{1/\eta})^{-1}$. Define $T_i = m \log(\rho(\alpha\beta)^i)$, where ρ designates the radius of the initial ball chosen by Bob.

The aim of Alice is to force the point $\mathbf{x}_{\sigma,\infty} \in \bigcap_n B_n$ remaining at the end of the game to satisfy that $(\langle A_\sigma, \mathbf{x}_{\sigma,\infty} \rangle) \in k$ -Bad. During each

round of the game, Alice must choose a sub-ball $A_n \subset B_{n-1}$ with center point $\mathbf{y}_{\sigma,n}$ and radius $\rho(\alpha\beta)^{n-1}\alpha$. She will consider the affine lattice associated to the center point of the ball $B_{n-1} = B(\mathbf{x}_{\sigma,n-1}, \rho(\alpha\beta)^{n-1})$ given to her by Bob. Within this affine lattice she will look at all of the vectors of a particular height and determine if any of them are good approximants to the system $(L_{A_\sigma}(\mathbf{x}_{\sigma,n-1}))$. A good approximant in \mathcal{O}_k^{n+m} for the affine system $(L_{A_\sigma}(\mathbf{x}_{\sigma,n-1}))$ of height e^t corresponds to a short vector in the affine lattice $\mathbf{g}_t(L_{A_\sigma}(\mathbf{x}_{\sigma,n-1}))\mathcal{O}_k^{n+m}$, that is, the application of the matrix \mathbf{g}_t acts as a normalization between the norms of the sub-vectors $\|A_\sigma\mathbf{q}^\sigma - \mathbf{x}_{\sigma,n-1} + \mathbf{p}^\sigma\|$ and $\|\mathbf{q}^\sigma\|$. Alice's search for the good approximants she wishes to avoid is reduced to checking whether certain affine lattices have vectors inside a fixed ball about the origin in $(\mathbb{R}^{n+m})^d$. If in this round of the game she does find a short vector, she will choose $\mathbf{y}_{\sigma,n}$ in a position so that the good approximant for $(L_{A_\sigma}(\mathbf{x}_{\sigma,n-1}))$ is not such a good approximant for $(L_{A_\sigma}(\mathbf{y}_{\sigma,n}))$. However, she must be very careful in this choice, as she may introduce new good approximants of height e^t for $(L_{A_\sigma}(\mathbf{y}_{\sigma,n}))$. In particular, she may be put in the situation that there is a $(L_{A_\sigma}(0))\mathcal{O}_k^{n+m}$ -rational hyperplane H in $(\mathbb{R}^{n+m})^d$ such that the points of $\mathbf{g}_t(L_{A_\sigma}(\mathbf{x}_{\sigma,n-1}))\mathcal{O}_k^{n+m}$ are very crowded in \mathbf{g}_tH , and Alice must make her move away from this hyperplane containing many short and eventually short vectors. This is where Alice will make use of the condition that the measure μ is absolutely decaying.

One other convenience of using the normalizing matrices \mathbf{g}_t is that they transfer the playing field $B_n = B(\mathbf{x}_{\sigma,n}, \rho(\alpha\beta)^n)$ to the ball of radius 1 in the following manner. A general point under consideration in the beginning of the n th round of the game is of the form $\mathbf{x}_{\sigma,n} + \mathbf{b}_\sigma$, where $\mathbf{b}_\sigma \in B(0, \rho(\alpha\beta)^n)$. By conjugating matrices, we see

$$\begin{aligned} \mathbf{g}_t(L_{A_\sigma}(\mathbf{x}_{\sigma,n} + \mathbf{b}_\sigma)) &= \mathbf{g}_t(L_0(\mathbf{b}_\sigma))\mathbf{g}_{-t}\mathbf{g}_t(L_{A_\sigma}(\mathbf{x}_{n,\sigma})) \\ &= (L_0(e^{2t/m}\mathbf{b}_\sigma))\mathbf{g}_t(L_{A_\sigma}(\mathbf{x}_{n,\sigma})) \end{aligned}$$

The choice of $t = T_i = m \log(\rho(\alpha\beta)^i)$ and with a change of variables, we see that Alice can think of choosing her next move by choosing a ball of radius α within $B(0, 1)$. Of course, she also needs to be mindful that her center point is still in F .

Now we begin with the description of the formal strategy that Alice will use to win the game. For clarity of exposition, we will break the strategy down into first a special, easier to handle case, and then the general case.

4.1. Case 1: There exists a $(L_{A_\sigma}(0))\mathcal{O}_k^{n+m}$ -rational hyperplane whose covolume tends to zero. Bob begins the game by

choosing the ball $B_0 = B(\mathbf{x}_{\sigma,0}, \rho)$ with $\mathbf{x}_{\sigma,0} \in F$. Denote by H the $(L_{A_\sigma}(0))\mathcal{O}_k^{n+m}$ -rational hyperplane with $|\mathbf{g}_t H| \rightarrow 0$ as $t \rightarrow \infty$. As discussed in Section 3, this means that H contains the contracted subspace $(\{0\}^m \times \mathbb{R}^n)^d$ and intersects the expanded subspace $(\mathbb{R}^m \times \{0\}^n)^d$ in a hyperplane. For a winning strategy, Alice may play with no special goal until the hyperplane H is small. Let i_0 be the first time that $\mathbf{g}_{T_{i_0}} H$ has covolume less than $\frac{1}{3}$ and also the radius $\rho(\alpha\beta)^{i_0}$ of the ball B_{i_0} satisfies $\rho(\alpha\beta)^{i_0}(1-\alpha) < \tilde{r}$. Then by Lemma 3.2, we know that any two parallel cosets $\mathbf{g}_{T_{i_0}} H + \mathbf{v}$ and $\mathbf{g}_{T_{i_0}} H + \mathbf{w}$ for $\mathbf{v}, \mathbf{w} \in \mathbf{g}_{T_{i_0}}(L_{A_\sigma}(0))\mathcal{O}_k^{n+m}$ must be at least distance 3 apart. In this case, Alice simply wishes to ensure that the point $\mathbf{x}_{\sigma,\infty}$ constructed by the game is such that $(L_{A_\sigma}(\mathbf{x}_{\sigma,\infty}))\mathcal{O}_k^{n+m} + H$ does not contain the origin. Indeed consider the trajectory of such an affine system. As the hyperplane shrinks, its cosets move farther and farther apart, allowing no vector or sequence of vectors to become arbitrarily small.

Now suppose we have also allowed the game to run until there is a short vector, that is, we have arrived at round i_1 and for some $v \in \mathbf{g}_{T_{i_1}}(L_{A_\sigma}(0))\mathcal{O}_k^{n+m}$ the coset $v + \mathbf{g}_{T_{i_1}} H$ intersects the ball $(\rho(\alpha\beta)^{i_1})^{-1}B(\mathbf{x}_{\sigma,i_1}, \rho(\alpha\beta)^{i_1})$. By assumption, there can be only one such coset. Define $\mathcal{L} = \mathbf{g}_{T_{i_0}}^{-1}v + H \cap (\mathbb{R}^m \times \{0\}^n)^d$. Set $\rho_1 = \rho(\alpha\beta)^{i_1}$. Alice now uses the assumption on the size of α and the fact that μ is absolutely decaying with respect to the ball $B(\mathbf{x}_{\sigma,i_1}, \rho_1(1-\alpha))$ and the ε -thickening of the hyperplane \mathcal{L} , where $\varepsilon = 2\alpha\rho_1$, to find a point

$$\mathbf{y}_{\sigma,i_1+1} \in F \cap B(\mathbf{x}_{\sigma,i_1}, \rho_1(1-\alpha)) \setminus \mathcal{L}^{(\varepsilon)}.$$

(Notice that $\rho_1(1-\alpha) \geq \frac{1}{2}\rho_1$.) Alice chooses this point to be the center of the ball A_{i_1+1} , which is allowed by the rules of game since $A_{i_1+1} \subset B_{i_1}$.

We claim that after this step, Alice may play at random and still win the game. The point $\mathbf{x}_{\sigma,\infty}$ remaining at the end of the game satisfies that $\mathbf{x}_{\sigma,\infty} \notin \mathcal{L}^{(\alpha\rho_1)}$. Thus, the parallelotope spanned by $\left(\left(\begin{smallmatrix} \mathbf{x}_{\sigma,\infty} \\ 0 \end{smallmatrix}\right)\right)$ and a basis of $H \cap (L_{A_\sigma}(0))\mathcal{O}_k^{n+m}$ has positive $d(n+m)$ -dimensional volume. Since \mathbf{g}_t has determinant 1 and does not alter the volumes of such parallelotopes and since the covolume $|\mathbf{g}_t H|$ goes to zero, the distance of $\mathbf{g}_t \left(\left(\begin{smallmatrix} \mathbf{x}_{\sigma,\infty} \\ 0 \end{smallmatrix}\right)\right)$ to $\mathbf{g}_t H$ must go to infinity. The same is true for any parallelotope spanned by $\left(\left(\begin{smallmatrix} \mathbf{x}_{\sigma,\infty} \\ 0 \end{smallmatrix}\right)\right)$ and any translation of a basis of $H \cap (L_{A_\sigma}(0))\mathcal{O}_k^{n+m}$ by an element of \mathcal{O}_k^{n+m} . Therefore, the

trajectory $\mathbf{g}_t(L_{A_\sigma})(\mathbf{x}_{\sigma,\infty})\mathcal{O}_k$ cannot contain arbitrarily small vectors as $t \rightarrow \infty$, concluding the argument in this case.

4.2. Case 2: No $(L_{A_\sigma}(0))\mathcal{O}_k^{n+m}$ -rational hyperplane remains small. Let $B_0 = B(\mathbf{x}_{\sigma,0}, \rho)$ be, again, the initial ball chosen by Bob. In this more complicated case, Alice will again pay attention to the appearance of small hyperplanes, and make her moves so that the cosets of these small hyperplanes are far from the origin. However, this time there is not simply one special hyperplane to consider, and hence not simply one special move to ensure a winning strategy. Alice will have to be diligent throughout.

Small $\mathbf{g}_t(L_{A_\sigma}(0))$ -rational hyperplanes must exist, and in this case we know that no hyperplane is perpetually small. Thus, some large hyperplane must eventually become small. Let J be the smallest integer satisfying both that $\rho(\alpha\beta)^J(1-\alpha) < \tilde{r}$ and also that some $(L_{A_\sigma}(0))$ -rational hyperplane H satisfies that $\mathbf{g}_{T_{J-1}}H$ is big and $\mathbf{g}_{T_J}H$ is small. If there are two such hyperplanes, choose H to be the hyperplane which is small for the longest period of time. Alice, ignoring any small hyperplanes which appear initially in the game, plays the first J rounds of the game with no purpose in mind except to pass the time.

The game arrives at round J with Bob having chosen $B_J = B(\mathbf{x}_{\sigma,J}, \rho_J)$ with $\rho_J = \rho(\alpha\beta)^J$. Alice must choose a point $\mathbf{y}_{\sigma,J+1} \in B(\mathbf{x}_{\sigma,J}, \rho_J(1-\alpha))$, which is also in the fractal F , while trying to avoid letting any coset of the short hyperplane becoming too close to the origin.

To do this, she will make use of the following geometric observation. The hyperplane $\mathbf{g}_{T_{J-1}}H$ is shrinking. Then as noted in Section 3, $\mathbf{g}_{T_{J-1}}H$ cannot contain the entire expanding eigenspace $(\mathbb{R}^m \times \{0\}^n)^d$. Moreover we claim that the hyperplane $\mathbf{g}_{T_{J-1}}H$ must be at a significant angle to this expanding eigenspace in the sense that: there exists $\delta > 0$ such that for any vector $v \in (\mathbb{R}^m \times \{0\}^n)^d$ which is at distance D from $\mathbf{g}_{T_{J-1}}H \cap (\mathbb{R}^m \times \{0\}^n)^d$ forms, along with a $(L_{A_\sigma}(0))\mathcal{O}_k^{n+m}$ -rational basis of \mathbf{g}_{T_J} , a parallelepiped of volume no less than $\delta D|\mathbf{g}_{T_J}H|$.

To prove this, consider the exterior algebra $\bigwedge((\mathbb{R}^m \times \mathbb{R}^n)^d)$. Then $|\mathbf{g}_{T_J}H|$ corresponds to the norm of the vector $v_1 \wedge \cdots \wedge v_{d(n+m)-1}$ where $v_1, \dots, v_{d(n+m)-1}$ forms a basis of $\mathbf{g}_{T_J}H \cap (L_{A_\sigma}(0))\mathcal{O}_k^{n+m}$. Let $T = \log(\alpha\beta)$, and notice that $\mathbf{g}_T\mathbf{g}_{T_J} = \mathbf{g}_{T_{J+1}}$. The action of $\bigwedge^{d(n+m)-1}\mathbf{g}_T$ on the degree $d(n+m) - 1$ elements of $\bigwedge((\mathbb{R}^m \times \mathbb{R}^n)^d)$ has eigenvalues $e^{-T/(dm)}$ of multiplicity dm and $e^{T/(dn)}$ of multiplicity dn . Thus we may decompose $v_1 \wedge \cdots \wedge v_{d(n+m)-1}$ into the sum of two eigenvectors, w_+ of eigenvalue $e^{T/(dn)}$ and w_- of eigenvalue $e^{-T/(dm)}$. We are interested in bounding from below the quantity $|(w_+ + w_-) \wedge v|$. First, notice that v

is in the subspace determined by w_+ , and so $|(w_+ + w_-) \wedge v| = |w_- \wedge v|$. The vector w_- corresponds to a hyperplane \tilde{H} in $(\mathbb{R}^m \times \mathbb{R}^n)^d$. Notice that $\tilde{H} \cap (\mathbb{R}^m \times \{0\}^n)^d = \mathbf{g}_{T_{J-1}}H \cap (\mathbb{R}^m \times \{0\}^n)^d$, which can be seen by considering the normal vectors attached to $\mathbf{g}_{T_{J-1}}H \cap (\mathbb{R}^m \times \{0\}^n)^d$, $\tilde{H} \cap (\mathbb{R}^m \times \{0\}^n)^d$ and the hyperplane in $(\mathbb{R}^m \times \{0\}^n)^d$ corresponding to w_+ . Thus, v is also at distance D to \tilde{H} , and so $|v \wedge w_-| = D|w_-|$. Let us estimate $|w_-|$. Since $\mathbf{g}_{T_{J-1}}H$ is shrinking, it must be that the size of the component w_- is not too small in comparison to the size of w_+ , that is, there exists some $\delta' > 0$ depending on the norm and the eigenvalues such that $|w_-| \geq \delta'|w_+|$. Since $|w_-| + |w_+| > \xi_0$ and $|\mathbf{g}_{T_{J-1}}H| \geq \xi_0 > e^{-T/(dm)}|\mathbf{g}_{T_{J-1}}H|$, we see that $|w_-| > \delta|\mathbf{g}_{T_{J-1}}H|$, where δ is defined in terms of δ' and the eigenvalues. Combing all of these facts, we may compute $|v_1 \wedge \cdots \wedge v_{d(n+m)-1} \wedge v| = |w_- \wedge v| = \delta D|\mathbf{g}_{T_{J-1}}H|$, yielding the claim.

The hyperplane $\mathbf{g}_{T_J}H$ has covolume no more than ξ_0 . By a previous observation, this means that the distance between cosets $v + \mathbf{g}_{T_J}H$, for $v \in \mathbf{g}_{T_J}(L_{A_\sigma}(\mathbf{x}_{\sigma,J}))\mathcal{O}_k^{n+m}$, is at least ξ_0^{-1} . Therefore, there are at most $2\xi_0$ of such cosets of $\mathbf{g}_{T_J}H$ which intersect the unit ball. Denote by \mathcal{L} the hyperplane $\mathbf{g}_{T_J}H \cap (\mathbb{R}^m \times \{0\}^n)^d$. Alice wants to choose the center of the new ball at sufficient distance from these cosets. Using the equation

$$\mathbf{g}_{T_J}(L_{A_\sigma}(\mathbf{x}_{\sigma,J} + e^{-2T_J/m}\mathbf{b})) = (L_0(\mathbf{b}))\mathbf{g}_{T_J}(L_{A_\sigma}(\mathbf{x}_{\sigma,J}))$$

we may view the current field of play as the unit ball in $(\mathbb{R}^m \times \{0\}^n)^d$. More specifically, Alice wants to choose the ball $B(\mathbf{c}, \alpha) \subset B(0, 1)$ so that after shifting by $\mathbf{b} \in B(\mathbf{c}, \alpha)$ the distance between any of the cosets $(L_0(\mathbf{b}))\mathbf{g}_{T_J}(L_{A_\sigma}(\mathbf{x}_{\sigma,J})) + \mathbf{g}_{T_J}H$ to the origin within $(\mathbb{R}^m \times \{0\}^n)^d$ is at least α .

Moreover, this center point must be in the fractal F . Here Alice uses that the measure is absolutely decaying. Choosing $\varepsilon = 2\alpha\rho_J$ and $r = \rho_J(1 - \alpha)$ and applying the definition of absolutely decaying $2\xi_0$ times, Alice finds a point $\mathbf{y}_{\sigma,J+1}$ in $F \cap B(\mathbf{x}_{\sigma,J}, \rho_J(1 - \alpha))$ and outside of the ε neighborhoods of the $2\xi_0$ relevant translates of \mathcal{L} . The assumption that $\alpha < (4(2\xi_0 C_0)^{1/\eta})^{-1}$ makes this possible. Her move this round is to choose the ball $B(\mathbf{y}_{\sigma,J+1}, \rho_J\alpha)$.

As before, consider the volume of the $d(n + m)$ -dimensional parallelotope with edges given by any translate of a basis of $\mathbf{g}_{T_J}H$ which is $\mathbf{g}_{T_J}(L_{A_\sigma}(0))\mathcal{O}_k^{n+m}$ -rational by a vector $v \in (L_0(\mathbf{b}))\mathbf{g}_{T_J}(L_{A_\sigma}(\mathbf{x}_{\sigma,J}))$ along with the vector \mathbf{b} , where \mathbf{b} is any element of the ball $B(\mathbf{c}, \alpha)$ just picked by Alice. Playing in this way guarantees that the volume of this object is at least $|\mathbf{g}_{T_J}H|\delta\alpha \geq \xi_0 e^{-T/(dm)}\delta\alpha$. Notice that this quantity does

not depend on J . Transformation under \mathbf{g}_t does not change the volume of this parallelopete and so the vectors in the sets $\mathbf{g}_{T_{J+i}}(L_{A_\sigma}(\mathbf{x}_{\sigma,J} + e^{-T_J/(dm)}\mathbf{b}))\mathcal{O}_k^{n+m}$ cannot have norm less than $\xi_0 e^{-T/(dm)}\alpha\delta|\mathbf{g}_{T_J}H|^{-1} \geq e^{-T/(dm)}\alpha\delta$ whenever $i \geq 0$ is such that $|\mathbf{g}_{T_{J+i}}H| \leq \xi_0$. Thus Alice has protected against the appearance of short vectors for the length of time the hyperplane H is small.

If at some future point $J' > J$, another big hyperplane H' becomes small, Alice repeats this procedure, again preventing short vectors from appearing during a certain time interval. Doing this infinitely often constructs the vector $\mathbf{x}_{\sigma,\infty}$ which forms with (A_σ) a k -badly approximable inhomogeneous linear system.

As a final remark, we would like to note that the basic strategy used above can also be used to deduce strong winning.

4.3. Proof of Theorem 2.2. Suppose now that (A_σ) is singular. We claim that the strategy employed by Alice as described above results in an affine linear system which is infinitely badly approximable, that is, $\mathbf{x}_{\sigma,\infty} \in k\text{-Bad}_{(A_\sigma)}^\infty(m, n)$. As was shown above, no vector of $\mathbf{g}_t(L_{A_\sigma}(\mathbf{x}_{\sigma,\infty}))\mathcal{O}_k^{n+m}$ can have norm less than $\xi_0 e^{-T/(dm)}\alpha\delta|\mathbf{g}_tH|^{-1}$. Since (A_σ) is singular and the map sending a lattice to its dual is continuous, we may apply Mahler's compactness criterion to the dual of $(L_{A_\sigma}(0))\mathcal{O}_k^{n+m}$ to conclude that $\min_H |\mathbf{g}_tH| \rightarrow 0$ as $t \rightarrow \infty$. Thus the shortest vector of $\mathbf{g}_t(L_{A_\sigma}(\mathbf{x}_{\sigma,\infty}))\mathcal{O}_k^{n+m}$ has norm tending to infinity, and the claim is proven.

5. Proof of homogeneous fixed vector case

In this section we combine the basic ideas of the previous section along with those contained within [21] to give the winning strategy that for the set $k\text{-Bad}^0(1, 1) \cap F$ where $F \subseteq k_S$ is the support of an absolutely decaying measure μ . Here we slightly change notation so as to (hopefully) avoid clutter and confusion. The set $k\text{-Bad}^0(1, 1)$ is a subset of $\text{Mat}_{1,1}(k_S) \simeq \mathbb{R}^d$. Instead of referring to the d -tuples of matrices (A_σ) , it will be simpler to discuss d -dimensional vectors \mathbf{a} , indexed by S . Center points of balls chosen in the game by Alice and Bob will be denoted by \mathbf{y}_n and \mathbf{x}_n , resp., while elements of the \mathcal{O}_k -modules will be denoted by \mathbf{v} . To prevent the over-proliferation of parentheses we will also use the notation

$$L_{\mathbf{a}}(\mathbf{b}) = \left(\left(\begin{pmatrix} 1 & a_\sigma & b_\sigma \\ & 1 & 0 \\ & & 1 \end{pmatrix} \right) \right).$$

Fix $0 < \beta < 1$. Let C_0 , η and r_0 be the constants associated to the decaying property of μ . Let $0 < \alpha < (2(C_0)^{1/\eta})^{-1}$. Define $T_i = \frac{1}{2} \log(\rho(\alpha\beta)^i)$ where again ρ designates the radius of the initial ball $B_0 = B(\mathbf{x}_0, \rho)$ chosen by Bob.

For the first few rounds of the game, Alice makes her choices at random. Her goal is only to make the radius small enough that the properties of μ are applicable and a short vector appears in the game. More precisely, the implementation of Alice's strategy begins at round n where $\rho(\alpha\beta)^n < r_0$ and there exists $\mathbf{v} \in \mathfrak{g}_{T_n} L_{\mathbf{x}_n}(0) \mathcal{O}_k^2$ with $\|\mathbf{v}\| < C^{-1}$ where C is the constant appearing in Lemma 2.4 of [21]. By Lemma 4.1 in combination with Lemma 2.4 of [21], we know that any other short vector must be in the k span of \mathbf{v} . Alice must chose a new center point \mathbf{y}_{n+1} in the ball $B_n = B(\mathbf{x}_n, \rho(\alpha\beta)^n(1 - \alpha))$ and also within the set F . This will correspond to the collection of \mathcal{O}_k -modules $L_{\mathbf{d}}(0) \mathfrak{g}_{T_n} L_{\mathbf{x}_n}(0) \mathcal{O}_k^2$ where \mathbf{d} ranges over a ball of radius α in $B(0, 1)$. She wishes to make this choice in such a manner as to ensure that the contribution of the expanding component of $L_{\mathbf{d}}(0)\mathbf{v}$ is significant in comparison to the component which is contracted under the application of \mathfrak{g}_t . Recall,

$$L_{\mathbf{d}}(0)\mathbf{v} = \left(\begin{pmatrix} v_1^\sigma + d_\sigma v_2^\sigma \\ v_2^\sigma \end{pmatrix} \right),$$

with \mathfrak{g}_t expanding the first components by e^t and contracting the seconds by e^{-t} . So there is only one point \mathbf{d}_0 at this step in the game that Alice must avoid. Let \mathcal{L} be a hyperplane in \mathbb{R}^d containing this point $\rho(\alpha\beta)^n \mathbf{d}_0 + \mathbf{x}_n$ which is to be avoided (where now we using the fact that $L_{\mathbf{d}}(0) \mathfrak{g}_{T_n} L_{\mathbf{x}_n}(0) \mathcal{O}_k^2 = \mathfrak{g}_{T_n} L_{\rho(\alpha\beta)^n \mathbf{d} + \mathbf{x}_n}(0) \mathcal{O}_k^2$). Define $\varepsilon = \alpha \rho(\alpha\beta)^n$. Using the definition of absolutely decaying along with the assumption that $\alpha < (2(C_0)^{1/\eta})^{-1}$ we have that

$$\begin{aligned} & \mu(B(\mathbf{x}_n, \rho(\alpha\beta)^n(1 - \alpha)) \cap \mathcal{L}^{(\alpha \rho(\alpha\beta)^n)}) \\ & \leq C_0 \left(\frac{\alpha \rho(\alpha\beta)^n}{\rho(\alpha\beta)^n(1 - \alpha)} \right)^\eta \mu(B(\mathbf{x}_n, \rho(\alpha\beta)^n(1 - \alpha))) \\ & < \mu(B(\mathbf{x}_n, \rho(\alpha\beta)^n(1 - \alpha))). \end{aligned}$$

(Observe that $\rho(\alpha\beta)^n(1 - \alpha) \geq \frac{1}{2} \rho(\alpha\beta)^n$.) Thus, Alice is able to choose her center point

$$\mathbf{y}_{n+1} \in F \cap B(\mathbf{x}_n, \rho(\alpha\beta)^n(1 - \alpha)) \setminus \mathcal{L}^{(\varepsilon)}.$$

Thus for \mathbf{d} with $\rho(\alpha\beta)^n \mathbf{d} \in A_{n+1} = B(\mathbf{y}_{n+1}, \rho(\alpha\beta)^n \alpha)$ we have that the size of the components of $L_{\mathbf{d}}(0)\mathbf{v}$ which expand under \mathfrak{g}_t , namely $((v_1^\sigma + d_\sigma v_2^\sigma))$, is at least $\alpha \|\mathbf{v}\|$. Since this bound is independent of n , we know that for t large enough (with "large enough" also being

independent of n) the vectors (in the k -span of) $L_{\mathbf{d}}(0)\mathbf{v}$ will be growing under \mathbf{g}_t . Alice's strategy is to repeat this procedure each time she encounters a sufficiently small vector $\mathbf{v} \in \mathbf{g}_{T_i}L_{\mathbf{x}_i}(0)\mathcal{O}_k^2$.

We claim that this strategy is indeed winning, that is, the point \mathbf{x}_∞ remaining at the end of the game satisfies that the corresponding \mathcal{O}_k -module, $L_{\mathbf{x}_\infty}(0)\mathcal{O}_k^2$, has bounded trajectory under \mathbf{g}_t . Let c be a constant such that for all $\mathbf{d} \in B(0,1)$ and all $\mathbf{v} \in \mathbb{R}^d$ we have $\|L_{\mathbf{d}}(0)\mathbf{v}\| < c\|\mathbf{v}\|$. Suppose there is a vector $\mathbf{v} \in \mathbf{g}_{T_n}L_{\mathbf{x}_\infty}(0)\mathcal{O}_k^2$ with $\|\mathbf{v}\| < (cC)^{-1}$. Then for some $\mathbf{d} \in B(0,1)$,

$$L_{\mathbf{d}}(0)\mathbf{v} \in L_{\mathbf{d}}(0)\mathbf{g}_{T_n}L_{\mathbf{x}_\infty}(0)\mathcal{O}_k^2 = \mathbf{g}_{T_n}L_{\mathbf{x}_n}(0)\mathcal{O}_k^2,$$

and $\|L_{\mathbf{d}}(0)\mathbf{v}\| < C^{-1}$. Thus, $L_{\mathbf{d}}(0)\mathbf{v}$ must be in the k -span of vectors against which Alice has protected herself. Thus, after a bounded length of time, the image of this vector under \mathbf{g}_t must be growing in norm. Therefore, there is some $\delta > 0$ so that all vectors contained in any of the \mathcal{O}_k -modules $\mathbf{g}_{T_n}L_{\mathbf{x}_\infty}(0)\mathcal{O}_k^2$ must have norm greater than δ . Thus, the trajectory $\mathbf{g}_tL_{\mathbf{x}_\infty}(0)\mathcal{O}_k^2$ is bounded and \mathbf{x}_∞ is k -badly approximable.

6. Proof of inhomogeneous fixed vector case

From the last section we have that the set of k -badly approximable vectors is winning. We will modify the winning strategy from that situation to arrive at the stronger conclusion that the k -badly approximable vectors in the inhomogeneous setting form a winning set as well. Moreover, we will have that this is still true when restricted to our fractal F .

Fix $\mathbf{b} \in k_S$, where we assume that $\mathbf{b} \notin \mathcal{O}_k$ (as this is the case covered in the previous section). Let μ be an absolutely decaying measure on k_S with support F . We will show that $k\text{-Bad}^{\mathbf{b}\sigma}(1,1) \cap F$ is winning. We will be using the notational conventions established in the previous section.

Let $\alpha_0 > 0$ be such that $k\text{-Bad}^0(1,1)$ is (α_0, β_0) -winning for every $\beta_0 > 0$. Set $\alpha = (4(2C_0)^{1/\eta})^{-1}\alpha_0$ and for every $\beta > 0$, define $\beta_0 = \beta(4(2C_0)^{1/\eta})^{-1}$ so that $\alpha_0\beta_0 = \alpha\beta$. Again, define T_n so that $\mathbf{g}_{T_{n+1}}L_{\mathbf{d}}(0) = L_{(\alpha_0\rho_n)^{-1}(0)}\mathbf{d}\mathbf{g}_{T_{n+1}}$ for any $\mathbf{d} \in \mathbb{R}^d$. In essence, two games will be played concurrently. In round n of the game, Alice will be presented with a ball $B_n = B(\mathbf{x}_n, \rho_n)$. She will use the winning strategy for $k\text{-Bad}^0(1,1)$ described in the previous section to first choose $A'_{n+1} = B(\mathbf{y}'_{n+1}, \alpha_0\rho_n) \subset B_n$ with $\mathbf{y}'_{n+1} \in F$ and then with an extra step she will choose $A_{n+1} = B(\mathbf{y}_{n+1}, \alpha\rho_n) \subset A'_{n+1}$ again with $\mathbf{y}_{n+1} \in F$.

The strategy of Alice is implemented in each round of the game as follows. In round n , Bob has chosen B_n and Alice has made the

first half of her move by choosing $A'_{n+1} = B(\mathbf{y}'_{n+1}, \alpha_0 \rho_n)$. To finish the move, Alice considers the shortest vector \mathbf{v} appearing in the affine \mathcal{O}_k -module $\mathbf{g}_{T_{n+1}} L_{\mathbf{y}'_{n+1}}(\mathbf{b}_\sigma) \mathcal{O}_k^2$. As before, she wishes to choose a subball B of $B(\mathbf{y}'_{n+1}, \alpha_0 \rho_n)$ in a way that ensures that the size of the components of $L_{(\alpha_0 \rho_n)^{-1} \mathbf{d}}(0) \mathbf{v}$ which are expanded under the application of \mathbf{g}_t are significant in comparison to the norm of $L_{(\alpha_0 \rho_n)^{-1} \mathbf{d}}(0) \mathbf{v}$ for any $\mathbf{d} \in B$. Denote by $(\mathbf{v})_e$ the projection of \mathbf{v} onto the subspace expanded by \mathbf{g}_t and similarly, denote by $(\mathbf{v})_c$ the projection of \mathbf{v} onto the subspace contracted by \mathbf{g}_t . Specifically,

$$\left(\begin{pmatrix} v_1^\sigma \\ v_2^\sigma \end{pmatrix} \right)_e = ((v_1^\sigma)) \text{ and } \left(\begin{pmatrix} v_1^\sigma \\ v_2^\sigma \end{pmatrix} \right)_c = ((v_2^\sigma)).$$

There exists a single point $\mathbf{d}_0 \in \mathbb{R}^d$ such that $0 = (L_{\mathbf{d}_0}(0) \mathbf{v})_e = (v_1^\sigma + d_{\sigma,0} v_2^\sigma)$. Moreover, if \mathbf{d} is at distance ε to \mathbf{d}_0 , then $\|(L_{\mathbf{d}}(0) \mathbf{v})_e\| \geq \varepsilon \|(\mathbf{v})_c\|$. Alice chooses a hyperplane \mathcal{L} in \mathbb{R}^d containing \mathbf{d}_0 and applies the definition of absolutely decaying to the ball $B(\mathbf{y}'_{n+1}, (\alpha_0 - \alpha) \rho_n)$ and with $\varepsilon = 2(4(2C_0)^{1/\eta})^{-1} \alpha_0 \rho_n$ to find a new center point $\mathbf{y}_{n+1} \in F \setminus \mathcal{L}^{(\varepsilon)}$ satisfying $A_{n+1} = B(\mathbf{y}_{n+1}, \alpha \rho_n) \subset B(\mathbf{y}'_{n+1}, \alpha_0 \rho_n)$. Thus for any $\mathbf{d} \in A_{n+1}$, we have that

$$\|(L_{\mathbf{d}}(0) \mathbf{v})_e\| \geq (4(2C_0)^{1/\eta})^{-1} \|(\mathbf{v})_c\|.$$

We claim that this two-step strategy is winning for Alice. By the first step along with the assumptions on α, α_0, β , and β_0 , we have that point \mathbf{x}_∞ determined by the course of the game must be an element of k -Bad $^0(1, 1)$. Hence, there exists $\varepsilon > 0$ such that for all $t \geq 0$, the \mathcal{O}_k -module $\mathbf{g}_t L_{\mathbf{x}_\infty}(0) \mathcal{O}_k^2$ contains no nonzero vector of norm less than ε . We also have the affine \mathcal{O}_k -module $L_{\mathbf{x}_\infty}(\mathbf{b}_\sigma) \mathcal{O}_k^2$ is discrete as a subset of k_S^2 . So there must exist $\delta > 0$ so that no vector of $L_{\mathbf{x}_\infty}(\mathbf{b}_\sigma) \mathcal{O}_k^2$ has norm less than δ . (Recall we have assumed that $\mathbf{b}_\sigma \notin \mathcal{O}_k$.) Finally, we observe that for any $\mathbf{d} \in B(0, 1)$ and for any $\mathbf{v} \in k_S^2$ we have $\|(d_\sigma v_2^\sigma)\| \leq \|(v_2^\sigma)\|$. This implies that $\|L_{\mathbf{d}}(0) \mathbf{v}\| \leq 2\|\mathbf{v}\|$ for appropriate \mathbf{d} . We claim that for any n , $\mathbf{g}_{T_n} L_{\mathbf{x}_\infty}(\mathbf{b}_\sigma) \mathcal{O}_k^2$ does not intersection the ball $B(0, r)$ where

$$r = \frac{\min(\varepsilon, \delta)}{4} (\alpha\beta)^{1/2}.$$

This gives that $\mathbf{x}_\infty \in k$ -Bad $^{(b)}(1, 1)$.

Clearly this claim holds for $n = 0$ by the assumption on δ . Suppose that $\mathbf{v}'' \in \mathbf{g}_{T_n} L_{\mathbf{x}_\infty}(\mathbf{b}) \mathcal{O}_k^2 \cap B(0, r)$ and that n is minimal with respect to this property. The discrepancy between the \mathcal{O}_k -module at the end of the game, $L_{\mathbf{x}_\infty}(\mathbf{b}) \mathcal{O}_k^2$, and the \mathcal{O}_k -module from round n of the game, $L_{\mathbf{x}_n}(\mathbf{b}) \mathcal{O}_k^2$, is measured by an application of the matrix $L_{\mathbf{d}}(0)$ for some $\mathbf{d} \in A'_{n+1} = B(\mathbf{y}'_{n+1}, \alpha_0 \rho_n)$. Thus, $\mathbf{g}_{T_n} L_{\mathbf{x}_\infty}(\mathbf{b}) \mathcal{O}_k^2 =$

$L_{\mathbf{d}''}(0)\mathbf{g}_{T_n}L_{\mathbf{x}_n}(\mathbf{b})\mathcal{O}_k^2$ for some $\mathbf{d}'' \in B(0, 1)$, and so \mathbf{v}'' is the preimage of some $\mathbf{v}' \in \mathbf{g}_{T_n}L_{\mathbf{x}_n}(\mathbf{b})\mathcal{O}_k^2$ under $L_{\mathbf{d}}(0)$ with $\|\mathbf{v}'\| \leq 2\|\mathbf{v}''\| \leq 2r$. This vector \mathbf{v}' corresponds to a vector \mathbf{v} coming from the \mathcal{O}_k -module from one round earlier in the game, namely $\mathbf{g}_{T_{n-1}}L_{\mathbf{x}_{n-1}}(\mathbf{b})\mathcal{O}_k^2$, in the sense that there exists $\mathbf{d}' \in B(0, 1)$ with

$$\mathbf{g}_{T_n}L_{\mathbf{x}_n}(\mathbf{b})\mathcal{O}_k^2 = \mathbf{g}_T L_{\mathbf{d}'}(0)\mathbf{g}_{T_{n-1}}L_{\mathbf{x}_{n-1}}(\mathbf{b})\mathcal{O}_k^2$$

where T is chosen such that $L_{\mathbf{x}}(0)\mathbf{g}_T = \mathbf{g}_T L_{(\alpha\beta)^{-1}\mathbf{x}}(0)$ and \mathbf{v} is the preimage of \mathbf{v}' under $\mathbf{g}_T L_{\mathbf{d}'}(0)$. Therefore, we are able to conclude

$$\|\mathbf{v}\| \leq 2 \cdot 2(\alpha\beta)^{-1/2}r \leq \frac{\varepsilon}{2}.$$

We have assumed, however, that $\mathbf{g}_{T_n}L_{\mathbf{x}_\infty}(0)\mathcal{O}_k^2$ does not contain any element of norm less than ε . So if there were another vector $\mathbf{w} \in \mathbf{g}_{T_{n-1}}L_{\mathbf{x}_{n-1}}(\mathbf{b})\mathcal{O}_k^2$ with norm $\|\mathbf{w}\| \leq \|\mathbf{v}\|$, this would result in a contradiction. Thus, \mathbf{v} must be the shortest vector of $\mathbf{g}_{T_{n-1}}L_{\mathbf{x}_{n-1}}(\mathbf{b})\mathcal{O}_k^2$ against which Alice has protected herself. By the minimality assumption on n , we also may conclude that $\|\mathbf{v}\| \geq \frac{1}{2}r$.

We claim that these facts imply that the size of $(L_{\mathbf{d}'}(0)\mathbf{v})_e$ is at least $(4(2C_0)^{1/\eta})^{-1}r$. First let us assume that $\|(\mathbf{v})_c\| \geq \frac{1}{2}\|(\mathbf{v})_e\|$. Then by the strategy employed by Alice, we have that

$$\|(L_{\mathbf{d}'}(0)\mathbf{v})_e\| \geq (4(2C_0)^{1/\eta})^{-1}\|(\mathbf{v})_c\| \geq (4(2C_0)^{1/\eta})^{-1}r.$$

If, instead, we assume that $\frac{1}{2}\|(\mathbf{v})_e\| > \|(\mathbf{v})_c\|$, then we see

$$\begin{aligned} \|(L_{\mathbf{d}'}(0)\mathbf{v})_e\| &= \|(v_1^\sigma + d'_\sigma v_2^\sigma)\| \geq \|(v_1^\sigma)\| - \|(d'_\sigma v_2^\sigma)\| \geq \|(v_1^\sigma)\| - \|(v_2^\sigma)\| \\ &\geq \frac{1}{2}\|(v_1^\sigma)\| \geq \frac{1}{4}r \geq (4(2C_0)^{1/\eta})^{-1}r, \end{aligned}$$

whence our claim holds.

Observe that applying \mathbf{g}_T results in the expanding components being multiplied by a factor of $(\alpha\beta)^{-1/2}$. Thus \mathbf{v}' has norm greater than $(4(2C_0)^{1/\eta})^{-1}r(\alpha\beta)^{-1/2}$. But we also already know that $\|\mathbf{v}'\| \leq 2r$, which yields the inequality $(4(2C_0)^{1/\eta})^{-1}r(\alpha\beta)^{-1/2} \leq 2r$. However, if we take β to be sufficiently small, this results in a contradiction. Thus, no such vector \mathbf{v}'' exists, and so the strategy is winning.

CHAPTER 4

Simultaneous dense and nondense orbits for noncommuting toral endomorphisms

BEVERLY LYTLE AND ALEX MAIER

ABSTRACT. Let S and T be hyperbolic endomorphisms of \mathbb{T}^d with the property that the span of the maximal subspace contracted by S along with the maximal subspace contracted by T is \mathbb{R}^d . We show that the Hausdorff dimension of the intersection of the set of points with equidistributing orbits under S with the set of points with nondense orbit under T is full. In the case that S and T are quasihyperbolic automorphisms, we prove that the Hausdorff dimension of the intersection is again full when we assume that \mathbb{R}^d is spanned by the maximal subspaces contracted by S and T along with the central eigenspaces of S and T .

1. Introduction

Questions about the size of the sets with points satisfying certain conditions on their orbits has received much attention in recent years. Most notably, there are many results in the area of diophantine approximation (for example, [9, 11, 18, 17, 22, 43] and further [28, 40, 47, 48]). Many of diophantine properties can be expressed in terms of the behavior of an orbit of a point in a homogeneous space. For example, the notion of bad approximability can be interpreted in terms of bounded orbits, while very well approximable objects correspond to divergent orbits. However, these types of results generally refer to the behavior of a point under one transformation or flow. In this work we attempt to address the simultaneous behavior of a point under two different transformations.

Let us begin with introducing our setting. Let $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ be the d dimensional torus, and let S be an ergodic endomorphism of \mathbb{T}^d . Then the transformation S can be realized as an integer matrix with no eigenvalue a root of unity. In this case S is called quasihyperbolic, and in the case that S has no eigenvalue of modulus 1, S is called hyperbolic. It is well known that, since S is ergodic, the set of points,

denoted $Eq(S)$, whose orbits equidistribute under S is a full measure set of \mathbb{T}^d . Consequently, the set of points, denoted $ND(S)$, which have nondense orbit under S has measure zero as it is contained in the complement, $NEq(S)$, of $Eq(S)$. While the set $ND(S)$ is small in a measure theoretic sense, it does have full Hausdorff dimension [11]. Indeed, it satisfies a stronger property known as winning. (Definitions of these terms are given in Section 2.) All this is to say that in the compact case, much is known about the behavior of points under a single transformation.

Let us then introduce T , a second ergodic endomorphism of \mathbb{T}^d . Clearly, as the intersection of full measure sets, $Eq(S) \cap Eq(T)$ is a full measure set, hence of full Hausdorff dimension. Further, it is a property of winning sets, that $ND(S) \cap ND(T)$ is again winning, and therefore of full Hausdorff dimension. Now what can be said of $Eq(S) \cap ND(T)$? In [9], Bergelson, Einsiedler and Tseng showed that for two commuting hyperbolic toral endomorphisms S and T which generate an algebraic \mathbb{Z}^2 action without rank 1 factors, $\dim(ND(T) \cap D(S))$ is greater than the dimension of the unstable manifold determined by T . (Here $D(S)$ is the set of points of \mathbb{T}^d which have dense orbit, a slightly larger set than $Eq(S)$, and \dim refers to Hausdorff dimension.) Their work relies on the measure rigidity theorems of Einsiedler and Lindenstrauss [24] and that the generalized eigenspaces of S and T are aligned (since the maps commute). In this article, we wish to determine the size of the set $Eq(S) \cap ND(T)$ in the situation that the generalized eigenspaces of S and T are not aligned in the following sense.

THEOREM 1.1. *Let $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ be the d dimensional torus. Let S and T be quasihyperbolic automorphisms of \mathbb{T}^d . Denote by \mathfrak{s}_- (resp. \mathfrak{t}_-) the maximal subspace contracted by S (resp. by T) and denote by \mathfrak{s}'_0 (resp. \mathfrak{t}'_0) the sum of the eigenspaces of S (resp. of T) of eigenvalue of modulus 1. Assume that \mathbb{R}^d is spanned by $\mathfrak{s}_- \oplus \mathfrak{s}'_0$ and $\mathfrak{t}_- \oplus \mathfrak{t}'_0$. Then*

$$\dim(Eq(S) \cap ND(T)) = d.$$

We have two versions of the proofs of Theorem 1.1 and two corollaries each following from one version of the proof and the properties of winning sets. We explicate the corollaries following from the proof of Theorem 1.1.

COROLLARY 1.2. *Let S be a quasihyperbolic endomorphism of \mathbb{T}^d . Let $\{T_k\}_k$ be a countable collection of quasihyperbolic automorphisms*

such that each, when paired with S , satisfy the conditions of Theorem 1.1 on the generalized eigenspaces of S and T . Then

$$\dim(Eq(S) \cap \bigcap_k ND(T_k)) = d.$$

COROLLARY 1.3. *Let T be a quasihyperbolic automorphism of \mathbb{T}^d . Let $\{S_k\}_k$ be a countable collection of quasihyperbolic endomorphisms such that each, when paired with T , satisfy the conditions of Theorem 1.1 on the generalized eigenspaces of T and S . Then*

$$\dim\left(\bigcap_k Eq(S_k) \cap ND(T)\right) = d.$$

We remark that with the methods we use it is not possible to say anything about sets of the form $Eq(S_1) \cap Eq(S_2) \cap ND(T_1) \cap ND(T_2)$.

We remark further that if S and T are non-commuting ergodic automorphisms of the 2-torus, then the hypotheses of the theorem are automatically satisfied. Indeed, if S and T share one eigenvector, then, by considering the underlying splitting field for the characteristic polynomial of T (and also of S), one sees that the other eigenspace of T and hence also of S must be the conjugate. Thus, if S and T share one eigenspace, then they share both, whence they commute. (See [65].)

The outline of the paper is as follows. In Section 2, we give some preliminaries on the basics of winning sets and ergodic toral endomorphisms. The third section is devoted to the proof of the theorem in the case that S and T are hyperbolic, exposing some structure of the sets $ND(T)$ and $Eq(S)$, taking advantage of that structure to prove a lemma about winning sets, and a recollection and application of Marstrand's slicing theorem. Section 4 contains the proofs of Theorem 1.1 which in structure is the same as the proof contained in Section 3 but to complete the details, requires some spectral theory, the notion of joinings of dynamical systems and topological entropy.

2. Preliminaries

Let $\mathbb{T}^d = \mathbb{R}^d \backslash \mathbb{Z}^d$ be the d -dimensional torus equipped with a Haar measure m normalized to be a probability measure. We also choose $\|\cdot\|$ to be the maximum norm on \mathbb{R}^d . Let S be an ergodic toral endomorphism, that is, S can be given as a nonsingular $d \times d$ matrix with integer entries and each eigenvalue of this matrix is not a root of unity. This property is often named quasihyperbolicity. It will be convenient to think of S as both a transformation on \mathbb{R}^d and on \mathbb{T}^d . Hopefully, on which space S is acting will be clear from context.

We define the following sets related to the \mathbb{Z} -action of S on \mathbb{T}^d :

$$\begin{aligned} D(S) &= \{x \in \mathbb{T}^d : \overline{\{S^n x\}_{n \in \mathbb{N}_0}} = \mathbb{T}^d\}, \\ Eq(S) &= \{x \in \mathbb{T}^d : \frac{1}{N} \sum_{i=0}^{N-1} f(S^i x) \xrightarrow{N \rightarrow \infty} \int f dm \ \forall f \in C(\mathbb{T}^d)\}, \\ ND(S) &= \{x \in \mathbb{T}^d : \overline{\{S^n x\}_{n \in \mathbb{N}_0}} \subsetneq \mathbb{T}^d\}, \text{ and} \\ NEq(S) &= \mathbb{T}^d \setminus Eq(S). \end{aligned}$$

Certainly, we have $ND(S) = \mathbb{T}^d \setminus D(S)$ and $Eq(S) \subset D(S)$. Applying the Pointwise Ergodic Theorem of Birkhoff, we have the following fact:

PROPOSITION 2.1. *With respect to m , almost every point $x \in \mathbb{T}^d$ equidistributes under S . Thus, $1 = m(Eq(S)) = m(D(S))$.*

It follows then that $NEq(S)$ and $ND(S) \subset NEq(S)$ are sets of measure zero. However, $ND(S)$ is “large” in another sense. To understand this notion of “large”, we recall the definitions of Hausdorff dimension and winning.

For a subset $A \subset \mathbb{R}^d$, $s \geq 0$ and $\delta > 0$, define the function

$$H_{s,\delta}(A) = \inf \left\{ \sum_{i=1}^{\infty} (\text{diam } B_i)^s : A \subset \bigcup_i B_i \text{ and } \text{diam } B_i \leq \delta \right\},$$

where $\text{diam } B$ refers to the supremum of the distances between any two points within B . As the parameter δ decreases, the value $H_{s,\delta}(A)$ increases. Thus, we define the s -dimensional Hausdorff measure of A to be the limit

$$H_s(A) = \lim_{\delta \rightarrow 0} H_{s,\delta}(A).$$

It is a fact that as a function of S , $H_s(A)$ is a jump function, taking the value ∞ for $s \in [0, \gamma)$ and taking the value 0 for $s \in (\gamma, \infty)$ for some $\gamma \in \mathbb{R}$. (Strictly speaking the value γ could be 0, however we will not be concerned with such sets in this article.) This value γ is the Hausdorff dimension of A , that is, $\dim(A) = \gamma$. (For further details on Hausdorff dimension we refer to Folland [31].)

There are a number of sets with known Hausdorff dimension. For example, the Hausdorff dimension of \mathbb{R}^d is d and the Hausdorff dimension of the ternary Cantor set is $\log 2 / \log 3$. While the definitions of Lebesgue measure and Hausdorff dimension both rely on the standard euclidean distance, the relationship between the two quantities is not straightforward. As we will see, there are sets of measure zero (namely $ND(S)$) which also have full Hausdorff dimension.

In general, Hausdorff dimension can be difficult to compute. We will make use of tools developed by W. Schmidt. In [59], he introduced a game and the definition of winning along with a few properties of winning sets. The game is played on (X, dist) , a complete metric space. Denote by $B(x, r)$ the closed metric ball around a point x of radius r . The setup of the two player game is given by two parameters $0 < \alpha, \beta < 1$, a set $S \subset X$, and the choice of one of the players, let's call him Bob, of a ball $B_0 = B(x_0, \rho)$. The first round begins with the other player, called Alice, choosing a center point of a ball y_1 such that $A_1 = B(y_1, \rho\alpha) \subset B_0$. Bob chooses the next center point of a ball x_1 such that $B_1 = B(x_1, \rho\alpha\beta) \subset A_1$. This procedure is iterated with the n th round of the game beginning with Alice choosing a point y_n with $A_n = B(y_n, \rho\alpha(\alpha\beta)^{n-1}) \subset B_{n-1}$, and continuing with Bob choosing a point x_n satisfying $B_n = B(x_n, \rho(\alpha\beta)^n) \subset A_n$. At the end of the game, there remains one point $x_\infty \in \bigcap B_n$. If $x_\infty \in S$, then Alice wins. If Alice can always find a winning strategy independent of the moves of Bob, the set S is (α, β) -winning. If there exists α such that S is (α, β) -winning for all $\beta > 0$, then S is an α -winning set, which may be shortened to S is a winning set.

Winning sets have a number of useful properties for computing Hausdorff dimension. Schmidt showed in [59] that winning sets within $X = \mathbb{R}^d$ have Hausdorff dimension d (although more general statements exist [29, 30]). Moreover he showed in [59] that for a countable collection $\{S_i\}$ of α_i -winning sets with $\inf \alpha_i = \alpha_0 > 0$, the intersection $\bigcap_i S_i$ is α_0 -winning.

The following theorem of Broderick, Fishman and Kleinbock [11] will be useful.

THEOREM 2.2. *For any $M \in \text{GL}_d(\mathbb{R}) \cap \text{Mat}_{d,d}(\mathbb{Z})$ and for any $y \in \mathbb{T}^d$, the set $E(M, y)$ is $1/2$ -winning in \mathbb{T}^d .*

Here $E(M, y) = \{x \in \mathbb{T}^d : y \notin \overline{\{M^n x\}}\}$. Clearly for any $y \in \mathbb{T}^d$, $E(S, y)$ is a subset of $ND(S)$. Thus we have that:

COROLLARY 2.3. *For any ergodic toral endomorphism S , the set $ND(S)$ is $1/2$ -winning, and hence $\dim(ND(S)) = d$.*

To finish we establish some notation regarding the eigenspaces of S . Viewing S as a transformation of \mathbb{R}^d (which maps \mathbb{Z}^d into itself), we have the following decomposition

$$\mathbb{R}^d = \mathfrak{s}_- \oplus \mathfrak{s}_0 \oplus \mathfrak{s}_+,$$

where \mathfrak{s}_- is the subspace corresponding to the eigenvalues of S of modulus less than 1, \mathfrak{s}_+ is the subspace corresponding to the eigenvalues of

S of modulus greater than 1, and \mathfrak{s}_0 is the subspace corresponding to all other eigenvalues of S . For a second ergodic toral endomorphism T , we will use similar notation for the decomposition

$$\mathbb{R}^d = \mathfrak{t}_- \oplus \mathfrak{t}_0 \oplus \mathfrak{t}_+.$$

When speaking about a measure we always mean the Lebesgue measure. Of which dimension should be clear from context, since we are often talking about subspaces.

3. Hyperbolic toral endomorphisms

In this section, we specialize to the case of hyperbolic (and of course ergodic) toral endomorphisms, and give a proof of the main theorem in this case. Specifically, we prove:

THEOREM 3.1. *Let $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ be the d -dimensional torus. Let S and T be hyperbolic endomorphisms of \mathbb{T}^d and let \mathfrak{s}_- be the maximal subspace contracted by S and \mathfrak{t}_- the maximal subspace contracted by T . Assume that both \mathfrak{s}_- and \mathfrak{t}_- are nontrivial and that they span \mathbb{R}^d . Then*

$$\dim(Eq(S) \cap ND(T)) = d.$$

We give two proofs of this theorem. By properties of winning sets, we will have two immediately corollaries, one from each version of the proof of the theorem.

COROLLARY 3.2. *Let S be a hyperbolic endomorphism of \mathbb{T}^d . Let $\{T_k\}_k$ be a countable collection of hyperbolic endomorphisms such that each, when paired with S , satisfy the conditions on the generalized eigenspaces of Theorem 3.1 on S and T . Then*

$$\dim(Eq(S) \cap \bigcap_k ND(T_k)) = d.$$

COROLLARY 3.3. *Let T be a hyperbolic endomorphism of \mathbb{T}^d . Let $\{S_k\}_k$ be a countable collection of hyperbolic endomorphisms such that each, when paired with T , satisfy the conditions on the generalized eigenspaces of Theorem 3.1 on T and S . Then*

$$\dim(\bigcap_k Eq(S_k) \cap ND(T)) = d.$$

Let S and T be endomorphisms of \mathbb{T}^d such that no eigenvalue of S or of T is of modulus 1 or 0. This means that $\mathfrak{s}_0 = \mathfrak{t}_0 = 0$ and we have

$$\mathbb{R}^d = \mathfrak{s}_- \oplus \mathfrak{s}_+ = \mathfrak{t}_- \oplus \mathfrak{t}_+.$$

We also assume that each of the subspaces \mathfrak{s}_- and \mathfrak{t}_- are nontrivial, and that \mathbb{R}^d is spanned by these subspaces. We choose nontrivial subspaces $\mathfrak{s} \subset \mathfrak{s}_-$ and $\mathfrak{t} \subset \mathfrak{t}_-$ such that $\mathbb{R}^d = \mathfrak{s} \oplus \mathfrak{t}$. Define $\pi_{\mathfrak{s}} : \mathbb{R}^d \rightarrow \mathfrak{s}$ to be the projection parallel to \mathfrak{t} and similarly define $\pi_{\mathfrak{t}} : \mathbb{R}^d \rightarrow \mathfrak{t}$ to be the projection parallel to \mathfrak{s} . In this section, there is no harm in identifying \mathfrak{s} and \mathfrak{t} and their elements with their images in \mathbb{T}^d .

The following lemma is a key observation for the proof of this special case.

LEMMA 3.4. *Let $x \in \mathbb{T}^d$. Then for any $y \in \mathfrak{s}$, the orbits of the points x and $x + y$ under S are both dense in \mathbb{T}^d or both not dense in \mathbb{T}^d . Furthermore, if one of the orbits equidistributes, then both do. An analogous statement holds for $y \in \mathfrak{t}$ and T .*

PROOF. Since y is contracted by S we have that

$$\|S^n(x) - S^n(x + y)\| = \|S^n(y)\| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Thus the tails of the orbits of x and of $x + y$ are nearing identical. Suppose $\{S^n x\}_n$ is not dense in \mathbb{T}^d . Then there exists $z \in \mathbb{T}^d$ and $r > 0$ such that $\{S^n x\}_n \cap B(z, r) = \emptyset$. Since the orbits of x and $x + y$ have the same asymptotic behavior, there exists N such that for all $n > 0$, we have $\|S^{N+n}(x) - S^{N+n}(x + y)\| < \frac{r}{2}$. Therefore, $\{S^{N+n}(x + y)\}_n \cap B(z, \frac{r}{2}) = \emptyset$. Since $\{x + y, S(x + y), \dots, S^{N-1}(x + y)\}$ is a finite set, there exists some point $z' \in B(z, \frac{r}{2})$ and some $r' < r/2$ such that $\{S^n(x + y)\}_n \cap B(z', r') = \emptyset$. Thus both $\{S^n x\}_n$ and $\{S^n(x + y)\}_n$ are not dense in \mathbb{T}^d .

On the other hand, assume $\{S^n x\}_n$ is dense, and let $z \in \mathbb{T}^d$. Then there exists a subsequence $S^{n_k} x$ with z as the limit point. Then since the tails of $\{S^n x\}_n$ and $\{S^n(x + y)\}_n$ are approaching identical, the subsequence $S^{n_k}(x + y)$ also tends to z . Since this is true for any point z , $\{S^n(x + y)\}_n$ is also dense.

Now suppose that $\{S^n x\}$ equidistributes in \mathbb{T}^d . Let $f \in C(\mathbb{T}^d)$. Then we have

$$\lim_{N \rightarrow \infty} \left| \frac{1}{N} \sum_{n=0}^{N-1} f(S^n x) - \int_{\mathbb{T}^d} f dm \right| = 0.$$

Since \mathbb{T}^d is compact, f is uniformly continuous. Let $\varepsilon > 0$ and $\delta > 0$ be the constant coming from the definition of uniform continuity for f . Let M be such that $\|S^n(x) - S^n(x + y)\| < \varepsilon$ for all $n \geq M$. Then for

$N > M$ we have

$$\begin{aligned}
& \left| \frac{1}{N} \sum_{n=0}^{N-1} f(S^n(x+y)) - \int f \, dm \right| \\
& \leq \left| \frac{1}{N} \sum_{n=0}^{N-1} f(S^n(x+y)) - \frac{1}{N} \sum_{n=0}^{N-1} f(S^n x) \right| \\
& \quad + \left| \frac{1}{N} \sum_{n=0}^{N-1} f(S^n x) - \int f \, dm \right| \\
& \leq \left| \frac{1}{N} \sum_{n=0}^{M-1} (f(S^n(x+y)) - f(S^n x)) \right| + \frac{1}{N} \sum_{n=M}^{N-1} \varepsilon \\
& \quad + \left| \frac{1}{N} \sum_{n=0}^{N-1} f(S^n x) - \int f \, dm \right|.
\end{aligned}$$

Since M is a fixed finite value, as N tends to infinity, each term of the last expression tends to 0. \square

This means that for any $x \in \mathbb{T}^d$, every point of the leaf $x + \mathfrak{s}$ has the same behavior asymptotically, either nondense or dense (and perhaps equidistributing). Therefore $Eq(S)$ can be written as the disjoint union of sets of the form $x + \mathfrak{s}$ where x ranges over $\mathfrak{t} \cap Eq(S) = \pi_{\mathfrak{t}}(Eq(S))$. Similarly, we have a disintegration of $D(S)$ into the sets $x + \mathfrak{s}$ for $x \in \mathfrak{t} \cap D(S) = \pi_{\mathfrak{t}}(D(S))$, and a disintegration of $ND(T)$ into the sets $x + \mathfrak{t}$ for $x \in \mathfrak{s} \cap ND(T) = \pi_{\mathfrak{s}}(ND(T))$. This type of foliated structure leads to the next lemma.

LEMMA 3.5. *For almost every $x \in \mathfrak{t}$, we have that x is in $Eq(S)$. In other words, $Eq(S) \cap \mathfrak{t}$ is a full measure set within \mathfrak{t} (with respect to Lebesgue measure on \mathfrak{t}). An analogous statement holds for $D(S)$.*

PROOF. As shown in the previous section, $Eq(S)$ is a full measure set, and each line parallel to \mathfrak{s} is either completely contained in $Eq(S)$ or disjoint from it. Moreover every such line is transverse to \mathfrak{t} . More formally, this is an application of the theorem of Fubini: Using the equalities

$$1 = m(Eq(S)) = \int_{x \in \mathfrak{t}} \int_{x+\mathfrak{s}} \chi_{Eq(S)} = \int_{x \in \mathfrak{t} \cap Eq(S)} 1,$$

we can conclude that $\mathfrak{t} \cap Eq(S)$ has full measure in \mathfrak{t} .

The same argument holds for $D(S)$. \square

We would also like to say something about the size of the set $\mathfrak{s} \cap ND(T)$. Since $ND(T)$ is a measure zero set, the same arguments do not apply. We will instead have to appeal to the fact that it is a winning set and use the following lemma.

LEMMA 3.6. *Let (X, dist) be a complete metric space with a subset $U \subset X$ admitting a product structure $U = V \times W$ (meaning dist restricted to U is the maximum of the distance on V and the distance on W). Denote by π_V the projection from U to V . Suppose that $A \subset V$ is such that $\pi_V^{-1}(A)$ is an α -winning set in U . Then A is an α -winning set in V .*

PROOF. Since $\pi_V^{-1}(A)$ is (α, β) -winning for all $\beta > 0$, Alice will simply use the strategy for winning the (α, β) -game on U to win the (α, β) -game on V by projecting her moves using π_V to V . Explicitly, suppose in round n of an (α, β) -game on V Bob has chosen the ball $B(x_n, \rho_n)$. Alice must choose a subball $A_{n+1} = B(y_{n+1}, \alpha\rho_n) \subset B(x_n, \rho_n)$. She does this by considering $\pi_V^{-1}(B(x_n, \rho_n))$, which is a metric ball in U , employing the known winning strategy there to find $A'_{n+1} \subset \pi_V^{-1}(B(x_n, \rho_n))$, and setting $A_{n+1} = \pi_V(A'_{n+1})$. Since the end point of the (α, β) -game on U is in $\pi_V^{-1}(A)$, the end point of the corresponding (α, β) -game on V is in A . \square

This lemma leads to the following analogue of Lemma 3.5 for $ND(T)$ and \mathfrak{s} .

LEMMA 3.7. *The set $ND(T) \cap \mathfrak{s} = \pi_{\mathfrak{s}}(ND(T))$ is a $1/2$ -winning set in \mathfrak{s} .*

PROOF. This is an application of Lemma 3.6 with the result of Theorem 2.2. \square

For the next step, we compute the Hausdorff dimension of $Eq(S) \cap ND(T)$. For this we use Kleinbock and Margulis' version of the Marstrand slicing theorem [43]:

THEOREM 3.8. *Let M_1 and M_2 be Riemannian manifolds, $A \subset M_1$, $B \subset M_1 \times M_2$. Denote by B_a the intersection of B with $\{a\} \times M_2$ and assume that B_a is nonempty for all $a \in A$. Then*

$$\dim(B) \geq \dim(A) + \inf_{a \in A} \dim(B_a).$$

Now we have all ingredients and are ready to prove Theorem 3.1:

PROOF OF THEOREM 3.1. For this theorem we give two proofs. Both are needed to prove each of the corollaries.

Version 1: Define $A = \mathfrak{t} \cap Eq(S)$. This is a full measure set in \mathfrak{t} (with respect to Lebesgue measure on \mathfrak{t}), and therefore has full Hausdorff dimension, that is, $\dim(A) = \dim(\mathfrak{t})$. For every $a \in A$, define $B_a = (a + \mathfrak{s}) \cap ND(T)$. Notice then that $B = \bigcup_{a \in A} B_a$ is equal to $Eq(S) \cap ND(T)$. Due to the foliated structure of $ND(T)$, we have that $a + \pi_{\mathfrak{s}}(ND(T)) = B_a$. Since the projection (and translate) of a 1/2-winning set is still 1/2-winning by Lemma 3.6, B_a is a winning subset of $a + \mathfrak{s}$ and so $\dim(B_a) = \dim(a + \mathfrak{s}) = \dim(\mathfrak{s})$. Now we apply the above theorem to conclude

$$\dim(Eq(S) \cap ND(T)) = \dim(B) \geq \dim(\mathfrak{t}) + \dim(\mathfrak{s}) = n.$$

Since the left side of the inequality is bounded above by n , we have equality and our theorem is proven.

Version 2: This time we reverse the roles of \mathfrak{s} and \mathfrak{t} , and we define $A = \mathfrak{s} \cap ND(T)$. This we have shown to be a 1/2-winning set in the subspace \mathfrak{s} and therefore $\dim(A) = \dim(\mathfrak{s})$. For every $a \in A$ define $B_a = (a + \mathfrak{t}) \cap Eq(S)$. Due to the foliated structure of $Eq(S)$, we have that $a + \pi_{\mathfrak{t}}(Eq(S)) = B_a$. By Fubini's theorem, we know that B_a has full measure within $a + \mathfrak{t}$, and so $\dim(B_a) = \dim(a + \mathfrak{t}) = \dim(\mathfrak{t})$. Again we apply Theorem 3.8 to conclude

$$\dim(Eq(S) \cap ND(T)) = \dim(B) \geq \dim(\mathfrak{t}) + \dim(\mathfrak{s}) = n,$$

where $B = \bigcup_{a \in A} B_a$ as before. Thus, the theorem is proven again. \square

PROOF OF COROLLARY 3.2. We extend the first version of the proof of Theorem 3.1 to a proof of the corollary. Since we have more than two endomorphisms, we must be more flexible in the choice of the spaces \mathfrak{s} and \mathfrak{t} . We define $\mathfrak{s} = \mathfrak{s}_-$. Note that $\mathfrak{s} \neq \mathbb{R}^d$ because S maps \mathbb{Z}^d to itself and cannot contract in all directions simultaneously. Now we choose an arbitrary subspace \mathfrak{t} such that $\mathfrak{s} \oplus \mathfrak{t} = \mathbb{R}^d$. With this choice of \mathfrak{t} we then take the same definition of the set A , that is $A = \mathfrak{t} \cap Eq(S)$. For each T_k we choose a subspace $\mathfrak{t}^k \subseteq \mathfrak{t}$ such that $\mathfrak{s} \oplus \mathfrak{t}^k = \mathbb{R}^d$, and for every $a \in A$, set $B_a^k = (a + \mathfrak{s}) \cap ND(T_k)$. Using the subspaces \mathfrak{t}_k , we see as before that each the B_a^k are 1/2-winning. Hence, $\bigcap_k B_a^k$ is 1/2-winning as well. We proceed in the proof by applying the Marstrand slicing theorem to the sets A and $\bigcap_k B_a^k$. \square

PROOF OF COROLLARY 3.3. For the second corollary we extend the second version of the proof of Theorem 4.4 in an analogous way: We define $\mathfrak{t} = \mathfrak{t}_-$. Then we choose an arbitrary subspace \mathfrak{s} such that we have $\mathfrak{s} \oplus \mathfrak{t} = \mathbb{R}^d$ and we take the same definition of the set A , this time with a different choice of \mathfrak{s} . We replace the sets B_a with $\bigcap_k B_a^k$ where $B_a^k = (a + \mathfrak{t}) \cap Eq(S_k)$. Using the analogously constructed

subspaces \mathfrak{s}^k each of the sets B_a^k is of full measure, whence $\bigcap_k B_a^k$ is of full measure and of full Hausdorff dimension. The rest of the proof follows as above. \square

4. Quasihyperbolic toral endomorphisms

In this section we develop the ideas introduced in the previous section to extend the result to certain quasihyperbolic endomorphisms of the torus. Let S and T be endomorphisms of \mathbb{T}^d such that no eigenvalue of either is zero or a root of unity. Now we have the decomposition

$$\mathbb{R}^d = \mathfrak{s}_- \oplus \mathfrak{s}_0 \oplus \mathfrak{s}_+ = \mathfrak{t}_- \oplus \mathfrak{t}_0 \oplus \mathfrak{t}_+$$

where \mathfrak{s}_- is the subspace corresponding to the eigenvalues of S of modulus less than 1, \mathfrak{s}_0 is the subspace corresponding to eigenvalues of S of modulus equal to 1, and \mathfrak{s}_+ to those of modulus greater than 1, and similarly for $\mathfrak{t}_-, \mathfrak{t}_0, \mathfrak{t}_+$ and T . Again we assume that neither of \mathfrak{s}_- nor \mathfrak{t}_- are trivial. In the previous section we assumed that \mathfrak{s}_- and \mathfrak{t}_- spanned \mathbb{R}^d , but in this section our assumption is more complicated. Namely, we require that the subspaces $\mathfrak{s}_- \oplus \mathfrak{s}'_0$ and $\mathfrak{t}_- \oplus \mathfrak{t}'_0$ span \mathbb{R}^d where \mathfrak{s}'_0 and \mathfrak{t}'_0 are the sums of eigenspaces of eigenvalues of modulus 1 for S and T , respectively. The nature of this restriction will become apparent later in the section. Choose nontrivial subspaces $\mathfrak{s} \subset \mathfrak{s}_- \oplus \mathfrak{s}'_0$ and $\mathfrak{t} \subset \mathfrak{t}_- \oplus \mathfrak{t}'_0$ such that

$$\mathbb{R}^d = \mathfrak{s} \oplus \mathfrak{t}. \tag{11}$$

As in the previous section, we want to make use of the Marstrand slicing theorem. In order to do this, we must demonstrate that sets of the form $x + \mathfrak{s}$ are contained entirely in or are disjoint from $Eq(S)$, and similarly for $x + \mathfrak{t}$ and $ND(T)$. For this we need an analogue of Lemma 3.1 which states that for any $x \in \mathbb{T}^d$ and $y \in \mathfrak{s}$, the point $x + y$ has the same asymptotic behavior as x (that is, either have nondense or equidistributed orbit) under S . Let us first focus on the property of equidistribution. Consider the action of S on the space \mathfrak{s}_0 . As explained by Lind [50] it can be realized as the action of the block matrix

$$R = \begin{pmatrix} J(R_1, n_1) & & \\ & \ddots & \\ & & J(R_m, n_m) \end{pmatrix}$$

where each $J(R_i, n_i)$ is a Jordan block consisting of n_i copies of the 2×2 rotation matrix R_i ,

$$J(R_i, n_i) = \begin{pmatrix} R_i & I & & \\ & \ddots & \ddots & \\ & & \ddots & I \\ & & & R_i \end{pmatrix}.$$

Observe that when such a Jordan block is applied to a vector $y = (y_1, y_2, \dots, y_{n_i})^\top \in \mathbb{R}^{2n_i}$,

$$J(R_i, n_i)y = \begin{pmatrix} R_i y_1 + y_2 \\ R_i y_2 + y_3 \\ \vdots \\ R_i y_{n-1} + y_n \\ R_i y_n \end{pmatrix},$$

there is drift or shearing in all but the last coordinates. We are unable to control this drift and so cannot easily deduce the analogue of Lemma 3.1 for the full subspace $\mathfrak{s}_- \oplus \mathfrak{s}_0$. This is why we instead insist on working with the subspace $\mathfrak{s} = \mathfrak{s}_- \oplus \mathfrak{s}'_0$. On \mathfrak{s}'_0 , the action of S is realized by the matrix

$$R = \begin{pmatrix} R_1 & & \\ & \ddots & \\ & & R_m \end{pmatrix}.$$

(When $\mathfrak{s}'_0 = \mathfrak{s}_0$ S is said to be *central spin*, in which case the restriction of S to \mathfrak{s}_0 is always of the latter form.)

Let us look in closer detail at the actions of R and S on $\mathfrak{s}'_0 \simeq \mathbb{R}^{2m}$ and \mathbb{T}^d , respectively. In particular, let us determine the spectral types of R and S . Recall that for a unitary action U on a separable Hilbert space \mathcal{H} there is a decomposition $\mathcal{H} = \bigoplus_i Z(x_i)$ of \mathcal{H} into cyclic subspaces along with a sequence of measures $\sigma_{x_1} \gg \sigma_{x_2} \gg \dots$ on the unit circle such that Fourier transforms of these measures are defined by $\widehat{\sigma}_{x_i}(n) = \langle U^n x_i, x_i \rangle$. These measures are called the spectral measures of x_i with respect to U , and the measure class of σ_{x_1} is called the spectral type of U . (For further details on spectral types we refer to Glasner [34].)

Now R is not an integer matrix and so does not act on \mathbb{T}^d directly. However R does act by rotation on each factor of $\mathbb{R}^2 \oplus \dots \oplus \mathbb{R}^2 = \mathbb{R}^{2m}$ and can be thought of as irrational rotation by a vector α on \mathbb{T}^m . This gives rise to a unitary action U_R on $L_0^2(\mathbb{T}^m)$ with $U_R(f) = f \circ R$. It is well known that the nontrivial characters of \mathbb{T}^m , namely $f_{\mathbf{j}}(x) = e^{2\pi i \mathbf{j} \cdot \mathbf{x}}$ for $\mathbf{j} \in \mathbb{Z}^m$ with $\mathbf{j} \neq 0$, span a dense subspace of $L_0^2(\mathbb{T}^m)$. Now any

character $f_{\mathbf{j}}$ is an eigenfunction of U_R with eigenvalue $e^{2\pi i \mathbf{j} \cdot \alpha}$. Thus R has purely discrete spectrum, and the spectral type of R is an atomic measure.

On the other hand, S has continuous spectral type. Indeed, consider the spectral measure $\sigma_{\mathbf{j}}$ associated to any nontrivial character $f_{\mathbf{j}}$ for $\mathbf{j} \in \mathbb{Z}^d$ of \mathbb{T}^d and the action U_S on $L_0^2(\mathbb{T}^d)$ induced by S . We may compute the Fourier transform of $\sigma_{\mathbf{j}}$ as

$$\begin{aligned} \widehat{\sigma}_{\mathbf{j}}(n) &= \langle U_S^n f_{\mathbf{j}}, f_{\mathbf{j}} \rangle \\ &= \int_{\mathbb{T}^d} e^{2\pi i \mathbf{j} \cdot S^n x} e^{-2\pi i \mathbf{j} \cdot x} \\ &= \int_{\mathbb{T}^d} e^{2\pi i \mathbf{j} \cdot (S^n - \text{Id})x}, \end{aligned}$$

where Id is the $d \times d$ identity matrix. For $n = 0$, we see that $\widehat{\sigma}_{\mathbf{j}}(n) = 1$. Now take $n > 0$. Observe that $S^n - \text{Id}$, by assumption, has trivial kernel. Thus, $\mathbf{j} \cdot (S^n - \text{Id})x = j_1 n_1 x_1 + \cdots + j_d n_d x_d$ for some integers n_i with $n_1 x_1 + \cdots + n_d x_d$ nonzero for all nonzero x . Clearly, $j_1 n_1 x_1 + \cdots + j_d n_d x_d$ is identically zero only when $\mathbf{j} = 0$, a situation we have excluded. Therefore, for $n > 0$, $\widehat{\sigma}_{\mathbf{j}}(n) = 1$ if $n = 0$ and $\widehat{\sigma}_{\mathbf{j}}(n) = 0$ otherwise. Hence $\sigma_{\mathbf{j}}$ is the Lebesgue measure, and S has continuous spectral type. These distinct spectral types of S and R will be an important ingredient in the next lemma. As a result of our discussion above we know that the spectral types of S and R are mutually singular.

We remark here that the action associated to the restriction of S to the sum \mathfrak{s}_0 of the generalized eigenspaces with eigenvalues of modulus 1 has mixed spectrum, yet still has Lebesgue spectral type (in the case that S is not central spin), and so is not, in this way, distinguishable from the full action of S on \mathbb{T}^d .

We will use one more tool to prove the next lemma and that is the structure of joinings of measure preserving systems. Let X and Y be measure spaces with probability measures μ and ν , respectively. Suppose A is a measure preserving and ergodic transformation of X and that B is a measure preserving and ergodic transformation of Y . Then we may consider the product space $X \times Y$ with the product action $A \times B$. Denote by $\pi_X : X \times Y \rightarrow X$ and $\pi_Y : X \times Y \rightarrow Y$ the canonical projections. A joining of X and Y is an $A \times B$ -invariant probability measure λ on $X \times Y$ with the property that $\pi_X^* \lambda = \mu$ and $\pi_Y^* \lambda = \nu$. The product measure $\mu \times \nu$ is clearly a joining of X and Y , however there may be others. If $\mu \times \nu$ is the only joining of X and Y , then X and Y are called disjoint. The following lemma follows almost immediately from the definitions.

LEMMA 4.1. *Given two measure preserving and ergodic probability spaces (X, A, μ) and (Y, B, ν) which are disjoint and two points $x \in X$ and $y \in Y$ which equidistribute under A and B , respectively, then (x, y) equidistributes in $X \times Y$ under $A \times B$ with respect to the product measure $\mu \times \nu$.*

PROOF. Define a probability measure λ_N on $X \times Y$ as the average of point masses:

$$\lambda_N = \frac{1}{N} \sum_{n=0}^{N-1} \delta_{(A^n x, B^n y)}.$$

Let λ be a weak* limit of the sequence λ_N . Then λ is $A \times B$ -invariant from construction. Let $f \in C_c(X)$, and define $F \in C_c(X \times Y)$ to be $F(x, y) = f(x)$. Then

$$\begin{aligned} \int f d\pi_X^* \lambda &= \int F d\lambda = \lim_N \int F d\lambda_N \\ &= \lim_N \frac{1}{N} \sum_{n=0}^{N-1} F(A^n x, B^n y) = \lim_N \frac{1}{N} \sum_{n=0}^{N-1} f(A^n x) = \int f d\mu, \end{aligned}$$

and so $\pi_X^* \lambda = \mu$. Similarly, $\pi_Y^* \lambda = \nu$. Therefore λ is a joining of X and Y , and $\lambda = \mu \times \nu$, whence (x, y) equidistributes under $A \times B$. \square

The relationship between spectral types and joinings that will be of use to us is the following proposition. For completeness we include a proof [34]:

PROPOSITION 4.2. *For two probability spaces X and Y with ergodic actions A and B , respectively, if the spectral types of the corresponding L_0^2 -actions are mutually singular, then X and Y are disjoint.*

PROOF. Let λ be a joining of X and Y . There are natural inclusions of $L^2(\mu)$ and $L^2(\nu)$ in $L^2(\lambda)$ as indicated in the previous proof. Let $f \in L^2(\mu)$ and $g \in L^2(\nu)$. Let $Z(f_0)$ be the cyclic subspace generated by $f_0 = f - \int f$ and let $P : L^2(\lambda) \rightarrow Z(f_0)$ be the natural orthogonal projection. The spectral measure σ_{Pg} is absolutely continuous with respect to both the spectral type σ_X of X and the spectral type σ_Y of Y . Since these are assumed to be mutually singular, σ_{Pg} is the zero measure. Therefore, $Pg = 0$ and

$$\int f_0 d\lambda \int g d\lambda = \int f_0 g d\lambda = \int P f_0 g d\lambda = \int f_0 P g d\lambda = 0.$$

Hence, $\int f g d\lambda = \int f d\mu \int g d\nu$. Thus, $\lambda = \mu \times \nu$ and X and Y are disjoint. \square

We are now ready to prove the analogue of the equidistribution statement of Lemma 3.4.

LEMMA 4.3. *Let $x \in \mathbb{T}^d$ and let $y \in \mathfrak{s}$. Then x equidistributes under S if and only if $x + y$ equidistributes under S . The analogous statement holds for the action of T and $y \in \mathfrak{t}$.*

PROOF. Let $x \in \mathbb{T}^d$ and $y \in \mathfrak{s}$ with $y \neq 0$ (otherwise the statement is a tautology). Suppose that x equidistributes in \mathbb{T}^d . We wish to show that $x + y$ equidistributes as well. The result for y in the contracting direction of S was already shown in Section 3, so without loss of generality we will assume that y is in the central eigenspace of S and that S acts on y as the irrational rotation $R : \mathbb{T}^m \rightarrow \mathbb{T}^m$ described above. Now the orbit of y under R is dense in a copy of \mathbb{T}^m in \mathfrak{s} . In fact, as R is an irrational rotation and so is uniquely ergodic, y equidistributes in \mathbb{T}^m under R . In the last few paragraphs, it has been established that the action of S on \mathbb{T}^d and the action of R on \mathbb{T}^m are disjoint. Since we may identify the action of S on $x + y$ with the action of $S \times R$ on (x, y) , we see that by Lemma 4.1, $x + y$ equidistributes under S . \square

At this point we are in a position to supply the proof of the following weaker version of Theorem 1.1.

THEOREM 4.4. *Let $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$ be the d dimensional torus. Let S and T be quasihyperbolic endomorphisms of \mathbb{T}^d . Denote by \mathfrak{s}_- (resp. \mathfrak{t}_-) the subspace contracted by S (resp. by T) and denote by \mathfrak{s}'_0 (resp. \mathfrak{t}'_0) the sum of the eigenspaces of S (resp. of T) of eigenvalue of modulus 1. Assume that \mathbb{R}^d is spanned by $\mathfrak{s}_- \oplus \mathfrak{s}'_0$ and $\mathfrak{t}_- \oplus \mathfrak{t}'_0$ and that both of these subspaces are nontrivial. Then*

$$\dim(Eq(S) \cap NEq(T)) = d.$$

PROOF OF THEOREM 4.4. Lemma 4.3 tells us that any set of the form $x + \mathfrak{s}$ is either contained in or disjoint from $Eq(S)$. Similarly, any set of the form $x + \mathfrak{t}$ is either contained in or disjoint from $NEq(T)$. Thus we may write

$$Eq(S) = \bigcup_{x \in \mathfrak{t} \cap Eq(S)} x + \mathfrak{s} \text{ and } NEq(T) = \bigcup_{x \in \mathfrak{s} \cap NEq(T)} x + \mathfrak{t}.$$

The same proofs as given in Lemmas 3.2 and 3.4 show that $\mathfrak{t} \cap Eq(S)$ has full measure in \mathfrak{t} and $\mathfrak{s} \cap NEq(T)$ is winning in \mathfrak{s} (where we note that $NEq(T)$ is winning because it contains the winning set $ND(T)$). Furthermore, both versions of the proof of Theorem 3.1 may be applied in the current situation to finish the proof. \square

To prove Theorem 1.1, we must work a bit harder to show the analogue of Lemma 3.4 for points with nondense orbit under the transformation T . To prove that for any $x \in ND(T)$ and $y \in \mathfrak{t}$, the point $x + y$ also has nondense orbit, we must introduce and use topological entropy. (See [67] for complete details.)

Let X be a compact topological space and let A be a continuous self-map of X . We recall the definition of the topological entropy of A . For two open covers \mathcal{U} and \mathcal{V} of X define their common refinement (or join) $\mathcal{U} \vee \mathcal{V}$ to be the open cover consisting of sets of the form $U \cap V$ for $U \in \mathcal{U}$ and $V \in \mathcal{V}$. For an open cover \mathcal{U} of X , let $N(\mathcal{U})$ be the number of elements in a minimal subcover of X . We define $H(\mathcal{U}) = \log N(\mathcal{U})$. The topological entropy of A relative to \mathcal{U} is

$$h(A, \mathcal{U}) = \lim_{n \rightarrow \infty} \frac{1}{n} H\left(\bigvee_{i=0}^{n-1} A^{-i}\mathcal{U}\right).$$

This limit always exists by a lemma of Fekete. Finally, define the topological entropy of A to be

$$h(A) = \sup_{\mathcal{U}} h(A, \mathcal{U}),$$

where the supremum ranges over all open covers of X .

LEMMA 4.5. *For a closed subset $Y \subseteq X$ with $AY \subseteq Y$,*

$$h(A|_Y) \leq h(A).$$

PROOF. Let \mathcal{U} be an open cover of Y . By definition of the subspace topology, each element $U \in \mathcal{U}$ is of the form $U = \tilde{U} \cap Y$ for some open set \tilde{U} of X . Set $\tilde{\mathcal{U}}$ to be the collection consisting of these sets \tilde{U} along with any other open set completing $\tilde{\mathcal{U}}$ to a cover of X . Thus, $N(\mathcal{U}) \leq N(\tilde{\mathcal{U}})$. Since $AY \subseteq Y$, $A|_Y^{-1}(\mathcal{U})$ is again an open cover of Y , as is $A^{-1}(\tilde{\mathcal{U}})$ of X . Moreover, elements of $A|_Y^{-1}(\mathcal{U})$ are of the form

$$A|_Y^{-1}(U) = A^{-1}(U) \cap Y = A^{-1}(\tilde{U} \cap Y) \cap Y = A^{-1}(\tilde{U}) \cap Y.$$

Thus $N(\mathcal{U} \vee A|_Y^{-1}(\mathcal{U})) \leq N(\tilde{\mathcal{U}} \vee A^{-1}(\tilde{\mathcal{U}}))$. Inductively, we arrive at the inequality $N(\bigvee_{i=0}^{n-1} A|_Y^{-i}(\mathcal{U})) \leq N(\bigvee_{i=0}^{n-1} A^{-i}(\tilde{\mathcal{U}}))$. By convexity of logarithm and continuity of limits, we see $h(A|_Y, \mathcal{U}) \leq h(A, \tilde{\mathcal{U}})$. Since each open cover of Y induces an open cover of X , we have that $h(A|_Y) \leq h(A)$. \square

In the specific case of ergodic toral automorphisms, we have a more specific version of Lemma 4.5. The next lemma states that the density of an orbit can be measured by entropy. It is in this lemma, where a

result of Berg is applied, that it is essential that T is an automorphism. Denote by $O(x)$ the closure of the orbit $\{T^n x\}_{n \in \mathbb{N}}$ of x under T .

LEMMA 4.6. *For $x \in \mathbb{T}^d$, $O(x) \subsetneq \mathbb{T}^d$ if and only if $h(T|_{O(x)}) < h(T)$.*

PROOF. By Lemma 4.5, for any $x \in \mathbb{T}^d$, we have $h(T|_{O(x)}) \leq h(T)$. If $x \in D(T)$, then $O(x) = \mathbb{T}^d$ and $h(T|_{O(x)}) = h(T)$. Suppose, then, that $x \in ND(T)$. We want to show $h(T|_{O(x)}) < h(T)$; assume otherwise. By the variational principle, there exists a sequence of probability measures μ_n supported on $O(x)$ with $h_{\mu_n}(T|_{O(x)})$ increasing to $h(T|_{O(x)}) = h(T)$. (Here $h_\nu(T)$ denotes the metric entropy of T with respect to a probability measure ν , and the variational principle states that $h(T) = \sup_\nu h_\nu(T)$. See [67] for the definition of metric entropy and details of the variational principle.) Let μ be a weak* limit of the sequence μ_n . Since $h_\nu(T)$ is an upper semi-continuous function of ν (see Theorem 4.1 of [53]), $h_\mu(T) = h(T|_{O(x)}) = h(T)$. By [8], the unique measure of maximal metric entropy for the ergodic toral automorphism T is the Lebesgue measure on \mathbb{T}^d , whence $\mu = m$. However, μ is supported on $O(x) \neq \mathbb{T}^d$. Thus we arrive at a contradiction, and so $h(T|_{O(x)}) < h(T)$. \square

Now we will use these lemmas to prove the analogue of Lemma 3.4 for nondense orbits.

LEMMA 4.7. *Let $x \in \mathbb{T}^d$ and let $y \in \mathfrak{t}$. Then x has nondense orbit under T if and only if $x + y$ has nondense orbit under T .*

PROOF. Assume $x \in ND(T)$, so that $O(x) \subsetneq \mathbb{T}^d$ and by the previous lemma $h(T|_{O(x)}) < h(T)$. We will show that $h(T|_{O(x+y)}) \leq h(T|_{O(x)})$. By the previous lemma and symmetry, we will have proven the desired result.

The result for y in the contracting direction of T was already shown in Section 3, so without loss of generality we will assume that y is a nonzero element in the central eigenspace of T and that T acts on y by the irrational rotation $R : \mathbb{T}^m \rightarrow \mathbb{T}^m$ described above. The system $(O(x), T|_{O(x)})$ is a factor of the system $(O(x) \times \mathbb{T}^m, T|_{O(x)} \times R)$ in the natural way. It is a basic fact of topological entropy that $h(T|_{O(x)} \times R) = h(T|_{O(x)}) + h(R)$. Moreover, since R is a rotation $h(R) = 0$. (See [67].) Hence $h(T|_{O(x)} \times R) = h(T|_{O(x)})$. Now, $O(x + y)$ is in a natural manner included in $O(x) \times \mathbb{T}^d$ as a closed $T \times R$ invariant set. Applying Lemma 4.4, we have that $h(T|_{O(x+y)}) \leq h(T|_{O(x)})$, finishing the proof. \square

All of the pieces are in place to prove Theorem 1.1.

PROOF OF THEOREM 1.1. Lemma 4.3 tells us that any set of the form $x + \mathfrak{s}$ is either contained in or disjoint from $Eq(S)$. Similarly, Lemma 4.6 states any set of the form $x + \mathfrak{t}$ is either contained in or disjoint from $ND(T)$. Thus we may write

$$Eq(S) = \bigcup_{x \in \mathfrak{t} \cap Eq(S)} x + \mathfrak{s} \text{ and } ND(T) = \bigcup_{x \in \mathfrak{s} \cap ND(T)} x + \mathfrak{t}.$$

The same proofs as given in Lemmas 3.5 and 3.7 show that $\mathfrak{t} \cap Eq(S)$ has full measure in \mathfrak{t} and $\mathfrak{s} \cap ND(T)$ is winning in \mathfrak{s} . Furthermore, both versions of the proof of Theorem 3.1 may be applied in the current situation to finish the proof. \square

The proofs of Corollaries 1.2 and 1.3 follow exactly the same line of reasoning given in Section 3 for Corollaries 3.2 and 3.3.

Bibliography

- [1] Jinpeng An, *Badziahin-Pollington-Velani's theorem and Schmidt's game*, Bull. London. Math. Soc. 45 (2013), no. 4, 721–733.
- [2] Jinpeng An, *Two dimensional badly approximable vectors and Schmidt's game*, preprint.
- [3] C. S. Aravinda, *Bounded geodesics and Hausdorff dimension*, Math. Proc. Cambridge Philos. Soc. 116 (1994), no. 3, 505–511.
- [4] C. S. Aravinda and E. Leuzinger, *Bounded geodesics in rank-1 locally symmetric spaces*, Ergodic Theory Dynam. Systems 15 (1995), no. 5, 813–820.
- [5] D. Badziahin, A. Pollington and S. Velani, *On a problem in simultaneous Diophantine approximation: Schmidt's conjecture*, Ann. of Math. (2) 174 (2011), no. 3, 1837–1883.
- [6] V. Beresnevich *Rational points near manifolds and metric Diophantine approximation* Annals of Math. 187–235 Vol. 175 (2012), no. 1
- [7] V. Beresnevich, *Badly approximable points on manifolds*, preprint.
- [8] K. Berg, *Convolution of invariant measures, maximal entropy*, Math. Systems Theory 3 1969 146–150. 28.75.
- [9] V. Bergelson, M. Einsiedler, J. Tseng, *Simultaneous dense and nondense orbits for commuting maps*, Preprint, arXiv:1309.4823v1.
- [10] V. Bernik, D. Kleinbock and G. Margulis, *Khintchine-type theorems on manifolds: the convergence case for standard and multiplicative versions*. Internat. Math. Res. Notices 2001, no. 9, 453–486.
- [11] R. Broderick, L. Fishman, D. Kleinbock, *Schmidt's game, fractals, and orbits of toral endomorphisms*, Ergodic Theory Dynam. Systems 31 (2011), no. 4, 1095–1107.
- [12] R. Broderick, L. Fishman, D. Kleinbock, A. Reich, B Weiss, *The set of badly approximable vectors is strongly C^1 incompressible*, Math. Proc. Cambridge Philos. Soc. 153 (2012), no. 2, 319–339.
- [13] E. Burger, *Homogeneous Diophantine approximation in S -integers*, Pacific J. Math. 152 (1992), no. 2, 211–253.
- [14] J. Cassels, *An introduction to diophantine approximation*, Cambridge Tracts Math. 45, Math. Phys., Cambridge Univ. Press, Cambridge, UK, 1957.
- [15] S. G. Dani, *Divergent trajectories of flows on homogeneous spaces and Diophantine approximation*. J. Reine Angew. Math. 359 (1985), 55–89.
- [16] ———, *Bounded orbits of flows on homogeneous spaces*, Comment. Math. Helv. 61 (1986), 636–660.
- [17] ———, *On orbits of endomorphisms of tori and the Schmidt game*, Ergodic Theory Dunam. Systems 8 (1988), 523–529.

- [18] ———, *On badly approximable numbers, Schmidt games and bounded orbits of flows*. In *Number theory and dynamical systems*, (York, 1987), London Math. Soc. Lecture Note Ser., 134, 69–86. Cambridge Univ. Press, 1989.
- [19] S. G. Dani and H. Shah, *Badly approximable numbers and vectors in Cantor-like sets*, Proc. Amer. Math. Soc. 140 (2012), no. 8, 2575–2587.
- [20] L. Dirichlet, *Verallgemeinerung eines satzes aus der lehre von der kettenbrüchen nebst einiger anwendungen auf die theorie der zahlen*, S.-B. Preuss., (1842) 93–95.
- [21] M. Einsiedler, A. Ghosh and B. Lytle, *Badly approximable vectors, C^1 curves and number fields*, Preprint: arXiv:1401.0992.
- [22] M. Einsiedler, A. Katok and E. Lindenstrauss, *Invariant measures and the set of exceptions to Littlewood’s conjecture*, Ann. Math. 164 (2006), 513–560.
- [23] M. Einsiedler and D. Kleinbock, *Measure rigidity and p -adic Littlewood-type problems*, Compositio Math. 143 (2007), 689–702.
- [24] M. Einsiedler and E. Lindenstrauss, *Rigidity properties of Zd -actions on tori and solenoids*, Electron. Res. Announc. Amer. Math. Soc. 9 (2003), 99–110.
- [25] M. Einsiedler and J. Tseng, *Badly approximable systems of affine forms, fractals, and Schmidt games*, J. Reine Angew. Math 660 (2011), 83–97.
- [26] M. Einsiedler and T. Ward, *Ergodic theory - with a view towards number theory*, Graduate Texts in Math., Vol. 259, Springer, 2011.
- [27] R. Esdahl-Schou and S. Kristensen, *On badly approximable complex numbers*, Glasg. Math. J. 52(2) (2010), 349–355.
- [28] D. Färm, *Simultaneously non-dense orbits under different expanding maps*, Dyn. Syst. 25 (2010), 531–545.
- [29] L. Fishman, *Schmidt’s games, badly approximable matrices, and fractals*, J. Number Th. 129(2009), 2133–2153
- [30] ———, *Schmidt’s game on fractals*, Israel J. Math. 171 (2009), 77–92.
- [31] G. Folland, *Real analysis : modern techniques and their applications*, Pure and applied mathematics, Wiley, New York, 1984.
- [32] A. Ghosh, *A Khintchine-type theorem for hyperplanes*. J. London Math. Soc. (2) 72 (2005), no. 2, 293–304.
- [33] A. Ghosh, A. Gorodnik and A. Nevo, *Diophantine approximation and automorphic spectrum*. Int. Math. Res. Not. IMRN 2013, no. 21, 5002–5058. 11J13
- [34] E. Glasner, *Ergodic theory via joinings*, Mathematical Surveys and Monographs, vol. 101, 2003.
- [35] A. Groshev, *A theorem on a system of linear forms*, Dokl. Akad. Nauk SSSR 19 (1938), 151–152, (in Russian).
- [36] T. Hattori, *Some Diophantine approximation inequalities and products of hyperbolic spaces*, J. Math. Soc. Japan 59 (2007), no. 1, 239–264.
- [37] A. Hurwitz, *Ueber die angenäherte Darstellung der Irrationalzahlen durch rationale Brüche*. (German) Math. Ann. 39 (1891), no. 2, 279–284.
- [38] V. Jarnik, *Diophantischen Approximationen und Hausdorffsches Mass*, Mat. Sb. 36 (1929), 371–382.
- [39] A. Khintchine, *Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der Diophantischen Approximationen*. (German) Math. Ann. 92 (1924), no. 1-2, 115–125.

- [40] D. Kleinbock, *Nondense orbits of flows on homogeneous spaces*, Ergodic Theory Dynam. Systems 18, (1998), 373–396.
- [41] D. Kleinbock, *Badly approximable systems of affine forms*, J. Number Theory 79 (1999), 83–102.
- [42] D. Kleinbock, E. Lindenstrauss and B. Weiss, *On fractal measures and diophantine approximation*, Selecta Math. (N.S.) 10 (2004), no. 4, 479–523.
- [43] D. Kleinbock and G. Margulis, *Bounded orbits of nonquasiunipotent flows on homogeneous spaces*, Amer. Math. Soc. Translations 171 (1996), 141–172.
- [44] D. Kleinbock and G. Tomanov, *Flows on S -arithmetic homogeneous spaces and applications to metric Diophantine approximation*, Max Planck Institute Preprint 2003–65 (2003).
- [45] D. Kleinbock, and G. Tomanov, *Flows on S -arithmetic homogenous spaces and application to metric Diophantine approximation*, Comm. Math. Helv. 82 (2007) 519–581.
- [46] D. Kleinbock and B. Weiss, *Badly approximable vectors on fractals*, Israel Journal of Mathematics 149 (2005), 137–170.
- [47] ———, *Modified Schmidt games and Diophantine approximation with weights*, Adv. Math. 223 (2010), no. 4, 1276–1298.
- [48] ———, *Modified Schmidt games and a conjecture of Margulis*, preprint arXiv:1001.5017v2.
- [49] S. Kristensen, R. Thorn and S. Velani, *Diophantine approximation and badly approximable sets*, Adv. Math. 203 (2006), 132–169.
- [50] D. Lind, *Dynamical properties of quasihyperbolic toral automorphisms*, Ergodic Theory Dynamical Systems 2 (1982), no. 1, 49–68.
- [51] C. McMullen, *Winning sets, quasiconformal maps and Diophantine approximation*, Geometric and Functional Analysis (2010), Volume 20, Issue 3, 726–740.
- [52] D. Morris, *Ratner’s theorems on unipotent flows*. Chicago Lectures in Math. Series, Chicago Uni. Press, 2005
- [53] S. Newhouse, *Continuity properties of entropy*, Ann. of Math. (2) 129 (1989), no. 2, 215–235.
- [54] A. Pollington and S. Velani, *On simultaneously badly approximable numbers*, J. London Math. Soc. (2) 66 (2002), no. 1, 29–40.
- [55] A. Pollington and S. Velani, *Metric Diophantine approximation and ‘absolutely friendly’ measures*, Selecta Mathematica 11 (2005), 297–307.
- [56] R. Quême, *On Diophantine approximation by algebraic numbers of a given number field: a new generalization of Dirichlet approximation theorem*, Journées Arithmétiques, 1989 (Luminy, 1989). Astérisque No. 198-200 (1991), 273–283 (1992).
- [57] M. Ratner, *Distribution rigidity for unipotent actions on homogeneous spaces*. Bull. Amer. Math. Soc. (N.S.) 24 (1991), no. 2, 321–325.
- [58] K. Roth, *Rational approximations to algebraic numbers*. Mathematika 2 (1955), 1?20; corrigendum, 168.
- [59] W. Schmidt, *On badly approximable numbers and certain games*, Trans. Amer. Math. Soc. 123 (1966), 178–199.
- [60] ———, *Badly approximable systems of linear forms*, J. Number Theory 1(1969), 139–154.

- [61] ———, *Simultaneous approximation to algebraic numbers by elements of a number field*, *Monatsh. Math.* 79 (1975), 55–66.
- [62] ———, *Diophantine approximation*, *Lecture Notes in Mathematics*, vol. 785, Springer-Verlag, Berlin, 1980.
- [63] N. Shah, *Expanding translates of curves and Dirichlet-Minkowski theorem on linear forms*. *J. Amer. Math. Soc.* 23 (2010), no. 2, 563–589.
- [64] V. Sprindzhuk, *More on Mahler's conjecture*, *Dokl. Akad. Nauk SSSR* 155 (1964), 54–56
- [65] J. Tseng, *Simultaneous dense and nondense orbits for toral automorphisms*, Preprint, arXiv:1406.1970
- [66] P. Walters, *An introduction to ergodic theory*, *Graduate Texts in Math.*, Springer, 2000.
- [67] B. Weiss, *Almost no points on a Cantor set are very well approximable*, *Proc. R. Soc. Lond. A* (2001) 457 949–952.