THE PENNSYLVANIA STATE UNIVERSITY
The Graduate School

TWO PROBLEMS IN DYNAMICS ON TORI

A Dissertation in
Mathematics
by
Qiao Liu

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2021

The dissertation of Qiao Liu was reviewed and approved by the following:

Zhiren Wang
Associate Professor of Mathematics
Dissertation Advisor
Chair of Committee

Federico Juan Rodriguez Hertz
Anatole Katok Chair Professor of Mathematics

Boris Kalinin
Professor of Mathematics

Martin Bojowald
Professor of Physics

Carina Curto
Professor of Mathematics
co-Associate Head for Graduate Studies

Abstract

In this dissertation, two problems about dynamics by toral automorphisms are investigated.

In chapter 1 and 2, local rigidity property of affine $\mathbb{Z} \ltimes_\lambda \mathbb{R}$-action on tori generated by an irreducible toral automorphism and a linear flow along an eigenspace is studied. Such an action exhibits a weak version of local rigidity, i.e., any smooth perturbation close enough to an affine action is smoothly conjugate to the affine action up to a constant time change.

In chapter 3 and 4, exceptional points whose orbits under an ergodic toral automorphism A are sequentially ϵ-concentrating on atoms are defined and studied, More precisely, we prove the Hausdorff dimension of the set

$$Z_{\epsilon,A} = \left\{ x \in X \;\middle|\; \begin{array}{l} \exists \{n_k\}_k \subset \mathbb{N} \text{ such that } \mu_x = \lim_{k\to\infty} \delta_x^{n_k} \text{ exists} \\ \text{and } \mu_x \text{ has at least } \epsilon\text{-portion supported on atoms.} \end{array} \right\}$$

is at least $\kappa_A \epsilon$ where κ_A only depends on A.

Table of Contents

Acknowledgments

I was born in a small beautiful village in Hunan, China. Never expect to be on this journey along which I was helped by so many people.

I cannot begin to express my greatest thanks to my advisors, the late Professor Anatole Katok and Professor Zhiren Wang, for their patient guidance, continuous encouragement, generous support. This dissertation couldn't get finished without their help. It is my honor to study with such experts. I learned from them not only the mathematical knowledge but also the enthusiasm and devotion to research and education. Professor Katok supported me up to his unfortunate passing. I will always cherish the hours studying and discussing with such great mathematician. He will forever live in my heart. Professor Wang helped me a lot both academically and personally. It is beyond words to express my appreciation.

I would also like to thank my other committee members Federico Rodriguez Hertz, Boris Kalinin and Martin Bojowald who are more than generous with their expertise and precious time.

I also want to extend my thanks to PSU dynamics and geometry group for providing such great research atmosphere. Different level courses, various seminars, the annual conference provide the best nourishment I can imagine, not to mention the coffee and cream puff, the best combo you could have. I feel so fortunate to be a student, colleague and friend with you. This is a lifelong memory for me.

A special thanks to Prof. Kainan Xiang, for encouraging me through my whole graduate study, and Prof. Marcelo Viana for mentoring me during my study in IMPA.

Life is bigger than math. The friends I met along this journey, including but not limited to, Shuangjian, Fang, Tiecheng, Allan, Diego, Renata, Changguang, Xiaofei, Yajie, Daren, Simo, Zelin, Honglei, I deeply appreciate your friendship.

Last but not least, I am forever indebted to my family.

This thesis work has been partially supported by NSF grants DMS-1501095 and DMS-1753042. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

Part I

Local rigidity of certain solvable group action on tori

Chapter 1
Introduction to KAM method in dynamics on tori

1.1 Introduction to local rigidity

In this section, we introduce a few definitions of rigidity of group actions. Generally speaking, an action of group Γ, usually a group generated by finite elements, is *rigid* if deformation of this action is trivial in some sense, that is, they are all conjugate. There are several way to formulate this notion.

Definition 1.1.1. *Let G be a topological group and $\rho : \Gamma \to G$ be a representation of Γ. The representation ρ is called local rigid if there exist an open neighborhood U of ρ such that ρ' is conjugate to ρ for any $\rho' \in U$ and the conjugacy is an element in G close to the identity element.*

When G has differential structure, $G = \mathrm{Diff}^\infty(M)$ or occasionally $\mathrm{Diff}^k(M)$, we can develop the more refined notion of local rigidity. Here we think of a representation $\rho : \Gamma \to \mathrm{Diff}^k(M)$ as C^k-action $\rho : \Gamma \times M \to M$. The following is definition of local rigidity cited directly from Fisher's paper [Fis05] which is probably the most exhaustive one to my best knowledge.

Definition 1.1.2. *Let Γ be a group and $\rho : \Gamma \to \mathrm{Diff}^k(M)$ be a homomorphism where k is either a positive integer or ∞. We say that ρ is $C^{k,l,i,j,m}$ locally rigid if any $\rho' : \Gamma \to \mathrm{Diff}^l(M)$ which is close to ρ in the C^i topology is conjugate to ρ by a C^j diffeomorphism h which is C^m small. Here l, i, j, m are all either non-negative integers or ∞ and the only a priori constraints are $i \leq min(k, l)$ and $m \leq j$.*

This definition shows many possible rigidity can be considered under various differential topologies of the group action itself, its perturbation and their closeness. For simplicity, we will avoid using this cumbersome notation when at all possible. There is a classical dynamical notion called structural stability which is equivalent to $C^{1,1,1,0,0}$ local rigidity, more precisely, stated as below:

Definition 1.1.3 (Structural stability)**.** *A group action $\rho : \Gamma \to \mathrm{Diff}^1(M)$ is structurally stable if there is a neighborhood U of ρ in $\mathrm{Hom}(\Gamma, \mathrm{Diff}^1(M))$ consisting of group action C^0 conjugates to ρ.*

From a dynamical point of view structural stability is important since it allows one to control dynamical properties of an open set of actions in $\mathrm{Diff}^1(M)$. Local differentiable rigidity can be viewed as a strengthening of this property in that it shows that an open set of actions is exhausted by smooth conjugates of a single action.

Definition 1.1.4 (Local differentiable rigidity, classical sense)**.** *A group action $\rho : \Gamma \to \mathrm{Diff}^\infty(M)$ has local differentiable rigidity if every group action in $\mathrm{Hom}(\Gamma, \mathrm{Diff}^\infty(M))$ sufficiently close to ρ in the C^i topology on $\mathrm{Diff}^\infty(M)$ is conjugate to ρ by a C^∞ diffeomorphism.*

1.2 Stability in hyperbolic dynamics

Definition 1.2.1. *Consider a diffeomorphism $f : M \to M$ on a Riemannian manifold M, f is called hyperbolic or Anosov if there exist numbers $0 < \lambda < 1 < \mu$, $C > 0$ and subspace $E^s(x)$ and $E^u(x)$ for all x such that*

(1) $E^i(x)$ form an invariant continuous splitting of the tangent space, i.e., $T_xM = E^s(x) \oplus E^u(x)$ and $d_xfE^i(x) = E^i(f(x))$ for $i = s, u$.

(2) if $n \in \mathbb{N}$ then $||d_xf^n|E^s(x)|| \le C\lambda^n$ and $||d_xf^{-n}|E^u(x)|| \le C\mu^{-n}$ $E^s(x)$, $E^u(x)$ are called stable and unstable bundle.

Example 1.2.2. *Let M be the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. Then A has eigenvalues $\lambda_1 = \frac{3+\sqrt{5}}{2}, \lambda_2 = \frac{3-\sqrt{5}}{2}$ where $0 < \lambda_2 < 1 < \lambda_1$. The eigenspaces rise as a natural invariant splitting into stable and unstable bundle.*

It deserves to mention all known concrete examples of Ansov diffeomorphism is algebraic, in the broader sense that they are continuously conjugate to automorphisms of tori, or more generally nilmanifolds and infranilmanifolds..

1.2.1 Structural stability

Theorem 1.2.3 (Structural Stability)**.** *If $f : M \to M$ is Anosov, so is any $g : M \to M$ sufficiently close to f in the C^1-topology. Moreover, there is a homotopically trivial Hölder homeomorphism h that is C^0 close to id_M such that $f = h^{-1} \circ g \circ h$.*

In general the conjugacy is only Bi-Hölder. For instance, consider the following perturbation of A in example 1.2.2, $\tilde{A} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \epsilon \begin{pmatrix} \sin 2\pi x \\ \sin 2\pi y \end{pmatrix}$. We will see in Theorem 1.2.3 that for sufficient small ϵ, there exists a bi-Hölder conjugacy h satisfying $A \circ h = h \circ \tilde{A}$. However, this conjugacy h can't be C^1. Indeed, 0 is a fixed point for both A and $\tilde{A}$. Taking the derivative at 0, $D\tilde{A} = \begin{pmatrix} 2 + 2\pi\epsilon & 1 \\ 1 + 2\pi\epsilon & 1 \end{pmatrix}$ has determinant 1 but different trace from that of A for $\epsilon \neq 0$, so $D\tilde{A}$ is not conjugate to A.

So the question is: under what condition does the conjugacy have higher regularity?

One possible direction is to study sufficiency of some necessary conditions carried with smooth conjugacy, like same periodic data. De la Llave, Marco and Moriyón [dlLMM86] gave a positive answer to this question in dimension 2. In higher dimension, there is a counterexample constructed by de la Llave [dlL92] showing that one has to add extra assumption such as irreducibility of the linear map, real simple spectrum, etc to guarantee sufficiency. Along this direction, there are lots of contribution by de la Llave, Gogolev, Kalinin, Sadovskaya and many others. Having the same Lyapunov exponents is another necessary condition. Saghin and Yang [SY18] obtained smoothness of the conjugacy for a volume preserving perturbation of an irreducible $A \in \mathrm{GL}(n, \mathbb{Z})$ assuming they have the same simple Lyapunov exponents. See also the recent work [GKS18], where the assumption on the simplicity of Lyapunov exponents is relaxed.

Another possible way to study regularity of the conjugacy in Theorem 1.2.3 is to consider higher rank abelian group actions. Since a single Anosov diffeomorphism induces a representation of group $\mathbb{Z}$ which can be thought as a group action by $\mathbb{Z}$. Instead of studying group action by $\mathbb{Z}$, it is more accessible to consider $\mathbb{Z}^k$ action or more general one like $\mathbb{Z}^k \times \mathbb{R}^k$.

1.2.2 Local rigidity of partially hyperbolic $\mathbb{Z}^k$ actions on the torus

In Damjanovic and Katok's paper [DK10], they proved the following.

Theorem 1.2.4. *Let $\alpha : \mathbb{Z}^k \times \mathbb{T}^N \to \mathbb{T}^N$ be an ergodic higher rank action of $\mathbb{Z}^k (k \geq 2)$ by automorphism of the N-dimensional torus. Then there exists a constant $l = l(\alpha, N) \in \mathbb{N}$*

such that α is $C^{\infty,l,\infty}$ local rigid.

Actually the automorphism they consider is *partial hyperbolic* which is much more general than *hyperbolic*. Unlike the hyperbolic diffeomorphism whose Lyapunov exponents are separated by the unit circle which giving us the splitting into the stable and unstable directions, the partial hyperbolic diffeomorphism has a triple splitting than includes a central direction with weaker contraction and expansion.

Definition 1.2.5. *Consider a manifold M and a diffeomorphism $f : M \mapsto M$, f is said to be partially hyperbolic if there exist numbers $0 < \lambda < \lambda_1 < 1 < \lambda_2 < \mu$, $C > 0$ and subspace $E^s(x)$, $E^c(x)$ and $E^u(x)$ for all x such that*

(1) $E^i(x)$ form an invariant splitting of the tangent space, i.e., $T_xM = E^s(x) \oplus E^c(x) \oplus E^u(x)$ and $d_xfE^i(x) = E^i(f(x))$ for $i = s, c, u$.

(2) if $n \in \mathbb{N}$ then $||d_xf^n|E^s(x)|| \leq C\lambda^n, C^{-1}\lambda_1^n < ||d_xf^n|E^c(x)|| < C\lambda_2^n$ and $C^{-1}\mu^n \leq ||d_xf^n|E^u(x)||$

$E^s(x)$, $E^c(x)$ and $E^u(x)$ are called stable, center and unstable bundle respectively.

Note that partial hyperbolic diffeomorphism is in general not structurally stable.

1.3 Historical background on rigidity in low dimensional dynamics

1.3.1 Rotation number for circle homeomorphism

Notice that there is a natural projection map $\pi : \mathbb{R} \to \mathbb{T}$ given by $x \to x+\mathbb{Z}$. This provides a lifting of an orientation preserving homeomorphism $f : \mathbb{T} \to \mathbb{T}$ to a homeomorphism $\bar{f} : \mathbb{R} \to \mathbb{R}$ with the property $\pi \circ \bar{f} = f \circ \pi$. Such a lift $\bar{f}$ is unique up to an additive integer constant, i.e., $\bar{f}(x + k) = f(x) + k$.

Proposition 1.3.1. *Let $f : \mathbb{T} \to \mathbb{T}$ be an orientation-preserving homeomorphism and $\bar{f} : \mathbb{R} \to \mathbb{R}$ a lifting of f. Then*

$$\rho(\bar{f}) := \lim_{n\to\infty} \frac{\bar{f}^n(x) - x}{n}$$

exists for all $x \in \mathbb{R}$. $\rho(\bar{f})$ is independent of x and well defined up to an integer, that is, if $\bar{f}_1, \bar{f}_2$ are liftings of f then $\rho(\bar{f}_1) - \rho(\bar{f}_2) = \bar{f}_1 - \bar{f}_2 \in \mathbb{Z}$.

This proposition justifies the following terminology:

Definition 1.3.2 (rotation number)**.** $\rho(f) := \pi(\rho(\bar{f}))$ *is called the rotation number of* f*.*

Remark 1.3.3. *If f has a periodic point then $\rho(f)$ is rational. The rotation number is invariant under orientation-preserving topological conjugacies.*

It is natural to ask if this definition can be generalized to higher dimensional tori. Unfortunately the same strategy doesn't work in higher dimension. Using the same notation, $\bar{f}$ is a lifting of $f : \mathbb{T}^d \to \mathbb{T}^d$, $\lim_{n\to\infty} \frac{\bar{f}^n(x)-x}{n}$ may not even exist for $x \in \mathbb{R}^d$. The following is one way of generating rotation number by using invariant measures. Define $\phi(x) = \bar{f}(y) - y$ where $y \in \pi^{-1}(x)$. Since $\bar{f}(x+k) = \bar{f}(x) + k$, $\phi(x)$ is well-defined. Denote the space of all f-invariant probabilistic measures on $\mathbb{T}^d$ by $\mathcal{M}(f)$ and its subspace consisting of ergodic measures by $\mathcal{M}_E(f)$. If $\mu \in \mathcal{M}_E(f)$ then, by the Birkhoff ergodic theorem,

$$\frac{1}{n}\sum_{i=0}^{n-1} \phi \circ f^i(x) \to \int \phi d\mu \ \mu\text{-almost everywhere.} \tag{1.1}$$

On the other hand, if $\pi(y) = x$ then

$$\frac{1}{n}\sum_{i=0}^{n-1} \phi \circ f^i(x) = \frac{1}{n}\sum_{i=0}^{n-1}(\bar{f}^{i+1}(y) - \bar{f}^i(y)) = \frac{\bar{f}^n(y) - y}{n}. \tag{1.2}$$

By 1.1 and 1.2, for μ-almost every $x \in \mathbb{T}^d$ and every $y \in \pi^{-1}(x)$, the sequence $\frac{\bar{f}^n(y)-y}{n}$ converges to $\int \phi d\mu$. With ergodic decomposition theorem, this can easily generalize to any f-invariant measures. Therefore we can define a rotation vector of f with respect to a invariant measure as below:

Definition 1.3.4 (rotation vector)**.** *Given a homeomorphism $f : \mathbb{T}^d \to \mathbb{T}^d$ homotopic to identity and preserving a probability measure μ, the vector*

$$\rho_\mu(f) = \int_{\mathbb{T}^d} (\bar{f}(x) - x) d\mu, \mod \mathbb{Z}^d \tag{1.3}$$

where $\bar{f}$ is any lifting of f, is independent of the choice of lifting $\bar{f}$. We call $\rho_\mu(f)$ the rotation vector of f with respect to μ.

Definition 1.3.5 (rotation set)**.** *The set $\rho_{meas} := \{\rho_\mu(f) : \mu \, is \, f\text{-invariant measure}\}$ is called rotation set of f.*

Remark 1.3.6. *There is no difficult to define rotation number, rotation vector and rotation set similarly for flow on tori. We will omit the details here. There are other ways to define these concept, a good resource would be the paper [MZ89] by Misiurewicz and Ziemian.*

1.3.2 Diophantine Condition

Definition 1.3.7. *A vector $v \in \mathbb{R}^d$ is said to satisfy a Diophantine condition of type (C, τ), if $|\langle k, v\rangle| \geq \frac{C}{|k|^\tau}$ for all $k \in \mathbb{Z}^d \backslash \{0\}$.*

We denote by $DC(C, \tau)$ the set of such v . Another more general notion is called simultaneously Diophantine.

Definition 1.3.8. *A collection of vectors $v_1, \dots, v_m \in \mathbb{R}^N$ is simultaneously Diophantine if there exist $\tau > 0$ and $C > 0$ such that*

$$\max_{1\leq i\leq m} |\langle n, v_i\rangle| \geq \frac{C}{\|n\|^\tau}, \forall n \in \mathbb{Z}^N \setminus \{0\}. \tag{1.4}$$

We denote by $SDC(C, \tau)$ the set of $(v_1, \dots, v_m)$ satisfying (1.4)

1.3.3 Local rigidity of translation and linear flow on $\mathbb{T}^N$

Theorem 1.3.9 (Kolmogorov, Arnold and Moser)**.** *Any smooth diffeomorphism $f \in \mathrm{Diff}^\infty_+(\mathbb{T})$ with $\rho(f)$ satisfying Diophantine condition and f sufficiently close to $R_{\rho(f)}$ is smoothly conjugated to $R_{\rho(f)}$.*

The main tool used is what we call KAM theory now. The name comes from the initials of A. N. Kolmogorov, V. I. Arnold and J. Moser who initiated the theory. In chapter 2, the proof of Theorem1.3.9 will be presented as a glimpse of KAM theory. The Diophantine condition can't be dropped due to their crucial role in solving small divisor problems. The closeness to rotation hypothesis can be dropped [Yoc02] which is referred to as the global rigidity of circle diffeomorphism.

In higher dimension, assuming all orbits have the same asymptotic behavior, i.e., the rotation vector $\lim_{n\to\infty} \frac{\bar{f}^n(x)-x}{n}$ exists and are the same for all x, Dias [Dia08] showed, assuming analytic regularity, such flows have similar properties, at least for typical rotation vectors. His result is stated below:

Theorem 1.3.10 (Dias)**.** *Let $v \in \mathcal{Y} \subset \mathbb{R}^d$. If a C^ω flow on $\mathbb{T}^d$ has an unique rotation vector v and it is C^ω close enough to linear, then it is C^ω conjugate to R^t_v.*

Here $\mathcal{Y} \subset \mathbb{R}^d$ is a set that contains all Diophantine vectors and has full Lebesgue measure. The definition of $\mathcal{Y}$ comes from a multidimensional continued fractions algorithm which is well paired with renormalization method used in the proof. Although the algorithm to find smooth conjugacy is uses renormalization method, it still follows the

KAM strategy. Since the rotation vector is invariant under C^0-conjugacies, this result yields an immediate consequence:

Corollary 1.3.11. *Let $v \in \mathcal{Y} \subset \mathbb{R}^d$. If a C^ω flow on $\mathbb{T}^d$ is C^0-conjugate to the translation flow R_v^t along v and is C^ω-close to linear, then the conjugacy is in fact C^ω.*

Actually the assumption that all orbits have the same asymptotic behavior can be relaxed. Karaliolios [Kar18] investigated a family of perturbations of Diophantine translation on tori that does not destroy all orbits with rotation vectors and obtained the following result.

Theorem 1.3.12 (Karaliolios)**.** *Let $R_v \in \mathbb{T}^d$ be a Diophantine rotation and $f \in \mathrm{Diff}^\infty(\mathbb{T}^d)$be a small enough perturbation. If v is in the convex hull of the rotation set of f, then the diffeomorphism f is smoothly conjugate to the translation by R_v.*

Remark 1.3.13. *v being in the convex hull of the rotation set means $v = \rho_\mu(f)$ for some f-invariant measure μ.*

1.3.4 Local rigidity of some abelian-by-cyclic solvable group action on $\mathbb{T}^N$

In paper [WX20], Winkinson and Xue studied rigidity of a solvable, abelian-by-cyclic, group acting as affine transformations of $\mathbb{T}^N$. Both local and global rigidity are obtained under certain assumptions. For the purpose of our paper, we only mention the local rigidity part. The abelian-by-cyclic group itself is defined as

$$\Gamma_{\bar{B},K} := \left\{ \begin{array}{l} g_0, g_{i,k}, i = 1, \ldots, N, k = 1, \ldots, K \left| \, [g_{i,k}, g_{j,l}] = 0, \right. \\ g_0 g_{i,k} = (\Pi_{j=1}^N g_{j,k}^{\bar{b}_{ji}}) g_0, \; i,j = 1, \ldots, N, k, l = 1, \ldots, K \end{array} \right\}$$

They studied the affine action of $\Gamma_{\bar{B},K}$ where the element g_0 acts on $\mathbb{T}^N$ by toral automorphism $A \in SL(N, \mathbb{Z})$, and element $g_{i,k}$ acts as translation $x \to x + \rho_{i,k}$, where $\rho_{i,k} \in \mathbb{R}^N$.

Local rigidity of affine $\Gamma_{\bar{B},1}$-action on $\mathbb{T}^N$, is the following:

Theorem 1.3.14 (Wilkinson-Xue)**.** *For any $\bar{A}, \bar{B} \in SL(N, \mathbb{Z})$ and any $C, \tau > 0$, there exist $\epsilon > 0$ and $l \in \mathbb{N}$ such that for any $\rho = (\rho_{1,1}, \ldots, \rho_{i,1}, \ldots, \rho_{N,1}) \in SDC(C, \tau)$ satisfying faithfulness the following holds. Let $\alpha : \Gamma_{\bar{B},1} \to \mathrm{Diff}^\infty(\mathbb{T}^N)$ be any representation satisfying*

1. $\alpha(g_0)$ *is homotopic to* $\bar{\alpha}(\bar{A}, \rho) = \bar{A}$*;*

2. $\max_{1 \le i \le N} \|\alpha(g_i) - \bar{\alpha}(\bar{A}, \rho)\|_{C^l} < \epsilon$;

3. *there exist* $\alpha(g_i)$*-invariant probability measures* $\mu_i, i = 1, \ldots, N$ *such that the matrix formed by the rotation vectors* $(\rho_{\mu_1}(\alpha(g_1)), \ldots, \rho_{\mu_N}(\alpha(g_N)))$ *is equal to* ρ*.*

Then there exists a C^∞ *diffeomorphism* h *that is* C^1 *close to identity such that* $h \circ \alpha = \bar{\alpha} \circ h$*. Moreover, the measure* $\mu = h_*^{-1}Leb$*, where Leb is Haar measure on* $\mathbb{T}^N$*, is the unique* α*-invariant measure and thus satisfies* $\rho_\mu(g_i) = \rho_{\mu_i}(g_i) = \rho_i, i = 1, \ldots, N$*.*

In the Theorem 1.3.14, assumption 3 says the perturbation of translation preserves rotation vectors. This is one essential requirement for their proof. They apply KAM method to the abelian group action $\mathbb{Z}^N$ induced by $\langle g_1, \ldots, g_N \rangle$ to get the conjugacy. It turns out the conjugacy obtained solely through information of $\mathbb{Z}^N$ action also conjugates $\alpha(g_0)$ to $\bar{A}$ up to a constant vector. This constant vector can easily be killed by composing conjugacy with a translation. Therefore, Theorem 1.3.14 is reduced to the following proposition.

Proposition 1.3.15 (Wilkinson-Xue)**.** *Given* $C, \tau > 0$*, there exist* $l \ge 1$ *and* $\epsilon_0 > 0$ *such that the following holds. Let* $T_1, \ldots, T_m \in Diff_0^\infty(\mathbb{T}^N)$ *be commuting diffeomorphism with* $m > 1$*. Suppose there exists* T_k*-invariant measures* μ_k *such that the rotation vectors* $\rho_{\mu_k}, k = 1, \ldots, m$*, satisfying the simultaneous Diophantine condition with constant* C, τ*, and* $\max \|T_k - id - \rho_{\mu_k}(T_k)\|_{C^l} \le \epsilon$*. Then there exists a* C^∞ *diffeomprhism* h *that is* C^1 *close to the identity such that* $h \circ T_k(x) = x + \rho_{\mu_k}(T_k), x \in \mathbb{T}^N, k = 1, \ldots, m$*. Moreover, the measure* $\mu = h_*^{-1}Leb$*, where Leb is Haar measure on* $\mathbb{T}^N$*, is the unique* α*-invariant measure and thus satisfies* $\rho_\mu(T_k) = \rho_{\mu_k}(T_k), k = 1, \ldots, m$*.*

1.4 KAM: a simple example demonstration

In this section, as a demonstration of the KAM method, we present the proof of the following classical theorem:

Theorem 1.4.1. *Given* $C, \tau > 0$*, there exits* $l \ge 1$ *and* ϵ *such that the following holds. Let* F *be an orientation preserving smooth circle diffeomorphism, i.e.,* $F \in Diff_+^\infty(\mathbb{T})$*. Suppose rotation number* $\alpha = \rho(F)$ *satisfies Diophantine condition* $\alpha \in DC(C, \tau)$ *and* $\|F - R_\alpha\|_{C^l} \le \epsilon$*. Then there exists a* C^∞ *diffeomorphism* H *that is* C^1 *close to identity such that* $H \circ F = H \circ R_\alpha$*.*

Since some of the lemmas in the proof will be used in the proof of our main result in Chapter 2 as well, we will include them later in Chapter 3 and quote them from there when proving Theorem 1.4.1.

1.4.1 The cohomological equation

Rewrite $F = R_\alpha + f$ and $H = Id + h$, $H \circ F = R_\alpha \circ H$ can be transformed into

$$h \circ (R_\alpha + f) - h = -f. \tag{1.5}$$

To find the conjugacy is equivalent to solve h in (1.5). Instead of solving (1.5) directly which is hard as it is a nonlinear differential equation, consider its linearization.

$$h \circ (R_\alpha) - h = -f \tag{1.6}$$

Define an operator $L : h \to h \circ (R_\alpha) - h$. So the question is whether this operator can be inverted, or in which space this can be done. If $h \in C^0(\mathbb{T}^1)$ with Fourier series expansion $h(x) = \sum_{n\in\mathbb{Z}} \hat{h}(n) \exp(2\pi i n x)$ then $L(\hat{h}(0)) = \hat{f}(0) = 0$. Passing to Fourier coefficients at nonzero frequency, $\hat{h}(n) \exp(2\pi i\alpha) - \hat{h}(n) = -\hat{f}(n)$. Obviously L is only invertible in a proper space. Given a function f with average $\hat{f}(0) = 0$, $\hat{h}(n) = \frac{\hat{f}(n)}{1-\exp(2\pi i n\alpha)}$ for every $n \neq 0$. This can be done because as $\alpha \in DC(C, \tau), |n\alpha| \geq \frac{C}{n^\tau}$ is uniformly bounded from zero. Moreover $\frac{1}{1-\exp(2\pi i n\alpha)} \leq |n|^\tau$. We have $|\hat{h}(n)| \leq |\hat{f}(n)||n|^\tau$. As the smoothness of a function on $\mathbb{T}$ depends on the speed of decay of its Fourier coefficient. $\hat{h}(n)$ decays slower than $\hat{f}(n)$. The above can be stated as the following lemma. We will use notation $\|\cdot\|_r$ for $\|\cdot\|_{C^r}$ and $|\cdot|_r$ for $|f|_r = \sup_n |\hat{f}(n)||n|^r$. The relation between the two norms are $\|f\|_r \leq |f|_{r+2}, |f|_r \leq \|f\|_r$.

Lemma 1.4.2. *Let $\alpha \in DC(C, \tau)$, let $f \in C^\infty(\mathbb{T}, \mathbb{R})$ with average $f(0) = 0$. Then there exists a solution h to the equation 1.6 such that $\|h\|_r \leq C_r \|f\|_{r+\sigma}$ for all $r \geq 0$ and $\sigma \geq 2 + \tau$.*

Proof. Use the construction above, $n \neq 0, \hat{h}(n) = \frac{\hat{f}(n)}{1-\exp(2\pi i n\alpha)}$.

$$\|h\|_r \leq |h|_{r+2} \leq \sup_n |\hat{f}(n)||n|^{r+2+\tau} = |f|_{r+2+\tau} \leq \|f\|_{r+2+\tau}. \tag{1.7}$$

□

1.4.2 Fourier cutoff

As the Lemma 1.4.2 shows we have to use higher order derivatives of f to control low order derivatives of h. There is a certain loss of regularity for the solution. In other words, L^{-1} sends C^r maps to $C^{r-\sigma}$ maps. It is unsustainable if we want to repeat the process. To overcome this obstruction, We take a Fourier cutoff of f. Define a smoothing operator $S_N : C^\infty(\mathbb{T}, \mathbb{R}) \to C^\infty(\mathbb{T}, \mathbb{R})$

$$S_N(f) = \sum_{|n|<N} \hat{f}(n) \exp(2\pi i n x) \tag{1.8}$$

$$\dot{S}_N(f) = \sum_{|n|\geq N} \hat{f}(n) \exp(2\pi i n x) \tag{1.9}$$

More general way of cutting off Fourier coefficients for higher dimensional toral diffeomorphism can be found in Lemma 2.3.2. We adapt the result here.

$$|S_N f|_{a+b} \leq N^b |f|_a$$
$$||S_N f||_{a+b} \leq N^{b+2} ||f||_a$$

and

$$|\dot{S}_N f|_{a-b} \leq N^{-b} |f|_a$$
$$||\dot{S}_N f||_{a-b} \leq N^{-b+2} ||f||_a$$

Instead of solving (1.6), we will sacrifice some accuracy of h in exchange of higher regularity. Solve the following equation:

$$h \circ (R_\alpha) - h = -S_N f \tag{1.10}$$

Lemma 1.4.3. *If $\alpha \in DC(C, \tau)$, and $f \in C^\infty(\mathbb{T}, \mathbb{R})$ with average $\hat{f}(0) = 0$, then there exists a solution h to the equation 1.10 such that $\|h\|_r \leq C_r N^{2+\sigma} \|f\|_r$ for all $r \geq 0$ and $\sigma \geq 2 + \tau$.*

Proof. Following Lemma 1.4.2, $\|h\|_r \leq C_r \|S_N f\|_{r+\sigma} \leq C_r N^{2+\sigma} \|f\|_r$. □

1.4.3 Inductive lemma

With all preparation above, we can state the following lemma:

Lemma 1.4.4. *Suppose F is in* $\mathrm{Diff}^\infty(\mathbb{T})$ *with* $\rho(F) = \alpha \in DC(C, \tau)$. *Assuming* $F = R_\alpha + f$ *and* $\|f\|_1 < 1$, *if there exists a* $h \in C^\infty(\mathbb{T})$ *such that:*

1. $\|h\|_r \le C_r N^{2+\sigma} \|f\|_r$
2. $\|h\|_1 \le \frac{1}{2}$*and* $H = id + h, H^{-1}$ *exists.*

Then for the map $\tilde{f} = H \circ F \circ H^{-1} - R_\alpha$ *we have:*

(i) $\|\tilde{f}\|_0 \le CN^{2+\sigma}\|f\|_0\|f\|_1 + C_l N^{-l}\|f\|_l$ *for all* $l \ge 0$

(ii) $\|\tilde{f}\|_l \le C_l N^{2+\sigma}(\|f\|_l + 1)$ *for all* $l \ge 0$.

Proof. Let $F^1 = H \circ F \circ H^{-1}, \tilde{f} = F^1 - R_\alpha$. Expanding the expression $F^1 \circ H = H \circ F$, we get:

$$\begin{aligned}(R_\alpha + \tilde{f}) \circ (id + h) &= (id + h) \circ (R_\alpha + f)\\ R_\alpha + h + \tilde{f} \circ (id + h) &= R_\alpha + f + h \circ (R_\alpha + f)\\ \tilde{f} \circ (id + h) &= f + h \circ (R_\alpha + f) - h\\ &= \hat{f}(0) + S_N(f) + \dot{S}_N(f) + h \circ (R_\alpha + f) - h \circ R_\alpha + h \circ R_\alpha - h\\ &= (S_N(f) + h \circ R_\alpha - h) + \dot{S}_N(f) + \hat{f}(0) + h \circ (R_\alpha + f) - h \circ R_\alpha\\ &= \dot{S}_N(f) + \hat{f}(0) + (h \circ (R_\alpha + f) - h \circ R_\alpha)\end{aligned} \tag{1.11}$$

To estimate C^0 norm of $\tilde{f}$, it suffices to estimate $\|\tilde{f}(id + h)\|_0$. Obviously we already have $\|\dot{S}_N(f)\|_0 \le N^{-l+2}\|f\|_l$. and $\|(h \circ (R_\alpha + f) - h \circ R_\alpha)\|_0 \le \|h\|_1\|f\|_0 \le C_1 N^{2+\sigma}\|f\|_1\|f\|_0$. The only trouble is $\hat{f}(0)$. It turns out the constant $\hat{f}(0)$ can be absorbed by the left side of (1.11). As we assume $\rho(F) = \alpha$, from this it follows that $\rho(F^1) = \alpha$ since the rotation number is topological conjugacy invariant. Now since $F^1 = R_\alpha + \tilde{f}$ and its rotation number is α, it follows that $\tilde{f}$ must have a zero, and so does $\tilde{f} \circ (id + h)$. Using a basic property of continuous map g with a zero, that is, $\|g\|_0 \le 2\|g - C\|_0$ for any constant C, we have $\|\tilde{f} \circ (id + h)\|_0 \le 2\|\tilde{f} \circ (id + h) - f(0)\|$. Let us take a few lines to show this property. Since g is continuous and has a zero,

$$\begin{aligned}\|g\|_0 &\le \max_{g(x)\ge 0} f(x) + \max_{g(x)<0}(-f(x)) \le \max_{g(x)\ge 0}(g(x) - C) + \max_{g(x)<0}(-(g(x) - C))\\ &\le 2\max_x |g(x) - C| = 2\|g - C\|_0.\end{aligned}$$

Based on all this discussion above, we have $\|\tilde{f}\|_0 \le CN^{2+\sigma}\|f\|_0\|f\|_1 + C_l N^{-l}\|f\|_l$.

Now it comes to estimate the C^l norm of $\tilde{f}^l$.

$$\begin{aligned}\tilde{f} &= H \circ F \circ H^{-1} - R_\alpha \\ &= (id + h) \circ F \circ H^{-1} - R_\alpha \\ &= F \circ H^{-1} - R_\alpha + h \circ F \circ H^{-1} \\ &= R_\alpha \circ H^{-1} - R_\alpha + f \circ H^{-1} + h \circ F \circ H^{-1} \\ &= H^{-1} - id + f \circ H^{-1} + h \circ F \circ H^{-1}.\end{aligned} \tag{1.12}$$

Applying Lemma 2.2.5, $\|\tilde{f}\|_l \leq C(\|h\|_l + \|f\|_l + 1) \leq C_l N^{2+\sigma}(\|f\|_l + 1)$. □

1.4.4 Iteration process

Now we will construct a sequence of perturbations F^k and a sequence of conjugacies H^k which converge. Starting with the initial perturbation $F^0 = F$ and $H^0 = id$. For $k \geq 0$, given $F^k = R_\alpha + f^k$, applying Lemma 1.4.4 with proper Fourier cutoff N_k, to f^k to obtain h^k and define $H^k = id + h^k$. Assuming invertibility of H^k is assured, define $F^{k+1} = H^k \circ F^k \circ (H^k)^{-1}$ and repeat this process. We shall fix the following parameters: speed of convergence $b = \frac{3}{2}$, choice of Fourier cutoff $N_{k+1} = N_k^b$.

Lemma 1.4.5. *Assuming the C^0-error at the n-th step of iteration satisfies $\|f^k\|_0 \ll N_k^{-a}$ and the C^l-error satisfies $\|f^k\|_l \ll N_k^a$ where $a = 3(\sigma + 2)$ and $l = 6(\sigma + 2)$, then $\|h_{k+1}|_1 \ll N_k^{-3}$, $\|f^{k+1}\|_0 \ll N_{k+1}^{-a}$ and $\|f^{k+1}\|_l \ll N_{k+1}^a$.*

Proof.

$$\begin{aligned}\|h_{k+1}\|_1 &\leq C_1 N_k^{2+\sigma} \|f^k\|_1 \\ &\ll N_k^{2+\sigma} \|f^k\|_0^{1-\frac{1}{l}} \|f^k\|_l^{\frac{1}{l}} \\ &\ll N_k^{2+\sigma-a(1-\frac{1}{l})+\frac{a}{l}} \\ &= N_k^{2+\sigma-3(2+\sigma)+1} = N_k^{-2\sigma-3} < N_k^{-3}\end{aligned} \tag{1.13}$$

h_{k+1} is invertible, therefore f^{k+1} is well defined. Applying the inductive lemma 1.4.4,

$$\|f^{k+1}\|_l \leq C N_k^{2+\sigma}(\|f^k\|_l + 1) \ll N_k^{2+\sigma+a} \leq N_k^{\frac{3}{2}a} = N_{k+1}^a \tag{1.14}$$

and

$$
\begin{aligned}
\|f^{k+1}\|_0 &\le CN_k^{2+\sigma}\|f^k\|_0\|f^k\|_1 + C_l N_k^{-l}\|f^k\|_l \\
&\ll N_k^{2+\sigma}\|f^k\|_0\|f^k\|_0^{1-\frac{1}{l}}\|f^k\|_l^{\frac{1}{l}} + N_k^{-l}N_k^a \\
&\ll N_k^{2+\sigma-a(2-\frac{1}{l})+\frac{a}{l}} + N_k^{-l+a} \\
&= N_k^{-5\sigma-10+\frac{1}{2}} + N_k^{-3(2+\sigma)} \\
&\ll N_k^{-\frac{9}{2}\sigma-9} = N_k^{-\frac{3}{2}a} = N_{k+1}^{-a}
\end{aligned}
\tag{1.15}
$$

□

To initialize this iteration process, it suffices to check $\|f^0\|_0 \ll N_0^{-a}$ and $\|f^0\|_l \ll N_0^a$. Since we can always require perturbation F to be small enough, it is not difficult to find ϵ and N_0 satisfying assumption in Lemma 1.4.5. Through this iteration process, we obtain a sequence of perturbation F^k satisfying $\|F^k - R_\alpha\|_0 \ll N_k^{-a} = (N_0^{(\frac{3}{2})^n})^{-a}$ which decays to 0. Convergence of $H_k \circ \cdots \circ H_i \circ \cdots \circ H_0$ follows from the fact $\|h_{k+1}|_1 \ll N_k^{-3}$ decays exponentially fast, its limit is the C^1 conjugacy between F and R_α. To improve the regularity of the conjugacy from C^1 to C^∞, standard argument can be found in Chapter 2 where a similar local rigidity problem is studied.

Chapter 2
Local rigidity of $\mathbb{Z} \ltimes_\lambda \mathbb{R}$ group actions on tori

2.1 Introduction

Consider the following solvable group,

$$\mathbb{Z} \ltimes_\lambda \mathbb{R} = \{(n,t) : (n_1,t_1)(n_2,t_2) = (n_1+n_2, \lambda^{-n_2}t_1 + t_2), n_i \in \mathbb{Z}, t_i \in \mathbb{R}, i = 1,2\}, \quad (2.1)$$

α is an affine $\mathbb{Z} \ltimes_\lambda \mathbb{R}$-action on torus $\mathbb{T}^d$. More precisely α is a group morphism $\alpha : \mathbb{Z} \ltimes_\lambda \mathbb{R} \to \mathrm{Diff}^\infty(\mathbb{T}^d)$ equipped with $\alpha(n,0).x = A^n x$, $\alpha(0,t).x = x + vt$ where A is an ergodic toral automorphism and λ, v are respectively a real eigenvalue of A and a corresponding eigenvector. We call α the affine $\mathbb{Z} \ltimes_\lambda \mathbb{R}$-action generated by the pair (A, v) where v represents the constant vector field that generates the linear flow. Note that if A is irreducible, the eigenspace V corresponding to the eigenvalue λ is 1-dimensional.

Let β be a smooth $\mathbb{Z} \ltimes_\lambda \mathbb{R}$-action. We say β is generated by the pair $(\tilde{A}, \tilde{v})$, where $\tilde{A}$ is a diffeomorphism homotopic to A and $\tilde{v}$ is a smooth vector field, if $\beta(1,0) = \tilde{A}$ and $\beta(0,t) = \tilde{\Phi}^t$ is the flow along $\tilde{v}$. The distance between α and β in C^r topology can be taken to be

$$d_r(\beta, \alpha) = max\{||\tilde{A} - A||_r, ||\tilde{v} - v||_r\}.$$

Here $||\cdot||_r$ represents the general C^r norm $||\cdot||_{C^r}$. We obtained a weak local rigidity of this group action in the following theorem.

Theorem 2.1.1. *Let α be the affine $\mathbb{Z} \ltimes_\lambda \mathbb{R}$-action on torus $\mathbb{T}^d$ generated by pair (A, v), where A is an irreducible toral automorphism and v is a non-zero eigenvector whose corresponding eigenvalue is $\lambda \in \mathbb{R}$. Then there exists $r = r(d)$ and $\delta = \delta(d, \alpha) > 0$ such* *that if β is a C^∞-smooth $\mathbb{Z} \ltimes_\lambda \mathbb{R}$-action satisfying $d_r(\alpha, \beta) < \delta$, then there is a vector v^* proportional to v such that β is C^∞-conjugate to the affine action generated by (A, v^*). Moreover, one may choose $r = 42(d+1)$.*

We prove our theorem using a KAM mechanism, details can be found in chapter 3. In classical KAM method studying local rigidity of translation or linear flow on torus, certain conditions on the preservation of rotation vectors are needed to absorb the zeroth Fourier coefficients at each step. In our proof, the Fourier coefficient of zero frequency, or at least its component in the direct sum of all the eigenspaces except the one containing v, can be transformed by the automorphism A and absorbed thereafter. Wikinson and Xue studied local rigidity of ABC-group actions on tori in theorem1.3.14 assuming rotation vectors are preserved as well as a simultaneous Diophantine condition on them. Our result studies a smaller group action with less group relations, and does not assume the existence of rotation vectors.

2.2 Preliminaries

Let P_V be the projection into V in the decomposition

$$\mathbb{R}^d = V \oplus V^\perp, \tag{2.2}$$

where $V^\perp$ denotes the direct sum of all eigenspaces of A other than V. In the rest of this thesis, $x \ll_z y$ denotes $x \leq Cy$, where C depends only on z.

Lemma 2.2.1. [Kat71, Lemma 3] *If A is an irreducible $d \times d$ matrix of integer coefficients, and v is a non-zero eigenvector of A from a 1-dimensional eigenspace, then v is Diophantine of type (C, τ) where $\tau = d-1$.*

To carry out the KAM scheme, we will set $\tau = d - 1$ and work under the following assumption:

Assumption 2.2.2. *The eigenvalue v and the perturbation error $w = \tilde{v} - v$ satisfy:*

1. *$|v| \in [\frac{1}{2}, 2]$;*

2. *$\hat{w}(0) \in V^\perp$, where $V^\perp$ is as in the decomposition (2.2).*

Remark 2.2.3. *By Lemma 2.2.1, part (1) of Assumption 2.2.2 implies that v is (C, τ)-diophantine for some constant $C = C(d, A)$.*

We will use $|\cdot|_r$ for $|f|_r = \sup_{n\in\mathbb{Z}} |\hat{f}(n)| \cdot |n|^r$ where $\hat{f}(n)$ is the n-th Fourier coefficient. $|\cdot|_r$ is a semi-norm. The following lemma is standard and describes the relation between $||\cdot||_r$ and $|\cdot|_r$. (See e.g. [DlL01].)

Lemma 2.2.4. *Let $f \in C^\infty(\mathbb{T}^d, \mathbb{R}^d)$ with $\hat{f}(0) = 0$, then $|f|_r \le ||f||_r$, $||f||_r \ll_{r,d} |f|_{r+d+1}$.*

When F, G are diffeomorphisms on $\mathbb{T}^d$, the C^r norm of their composition can be controlled linearly provided that $\|F\|_1$ and $\|G\|_1$ are bounded:

Lemma 2.2.5. [Ham82, Lemma 2.3.4] *Suppose $F, G \in C^r(\mathbb{T}^d, \mathbb{T}^d)$ and $\|F\|_1, \|G\|_1 \le M$, then*

$$\|F \circ G\|_r \ll_{r,d,M} 1 + ||F||_r + ||G||_r.$$

Also the following interpolation inequalities will be used a lot in our proof. Suppose $f \in C^\infty(\mathbb{T}^d)$, then

$$||f||_r \ll_{s,r_1,r_2,d} ||f||_{r_1}^{1-s} ||f||_{r_2}^{s}, \tag{2.3}$$

where $r = (1-s)r_1 + sr_2$ and $0 \le s \le 1$.

2.3 Preparatory Steps

The perturbation β may not be conjugate to α. Finding the conjugacy can be transformed into solving the equation $\alpha \circ H = H \circ \beta$, or equivalently to finding H that satisfies the following two equations.

$$A \circ H = H \circ \tilde{A}$$
$$\Phi^t \circ H = H \circ \tilde{\Phi}^t$$

Here Φ^t and $\tilde{\Phi}^t$ are respectively the smooth flows generated by v and $\tilde{v}$. Differentiating with respect to t in the second equation above, we see

$$v \circ H = DH \cdot \tilde{v} \tag{2.4}$$

There is no direct way to solve this nonlinear equation (2.4). In the classical KAM theory of studying linear flow on torus, instead of solving it directly, one linearizes the problem first and then resolves the linearized equation, using the solution to the linearized equation to construct a new approximate solution to the non-linear problem which is much closer to being an actual solution compared with previous one. By iterating this process, a convergent sequence of solutions will be obtained and its limit is the solution to original nonlinear equation. we adapt the KAM method in our proof.

Let $H = \mathrm{Id} + h$ and $\tilde{v} = v + w$, linearizing equation (2.4) gives

$$-w = Dh \cdot v \tag{2.5}$$

2.3.1 Construction of conjugacy

Equation (2.5), in terms of Fourier coefficients, can be rewritten as

$$-\hat{w}(n) = \hat{h}(n) 2\pi i n \cdot v$$

This equation gives the solution

$$\hat{h}(n) = \frac{-\hat{w}(n)}{2\pi i n \cdot v}, \ \forall n \in \mathbb{Z}^d \setminus \{0\} \tag{2.6}$$

To get $\hat{h}(0)$, let $f = \tilde{A} - A$, one can linearize $A \circ H = H \circ \tilde{A}$ to get

$$A \circ h = f + h \circ A$$

Rewriting it in terms of the Fourier coefficient at 0 frequency, we have

$$A\hat{h}(0) = \hat{f}(0) + \hat{h}((A^{-1})^{\mathrm{T}}(0))$$

Since A is an ergodic toral automorphism, there exists a unique solution

$$\hat{h}(0) = (A - \mathrm{Id})^{-1} \hat{f}(0) \tag{2.7}$$

With all the $\hat{h}(n)$'s solved above in (2.6) and (2.7), define

$$h(x) = \sum_{n \in \mathbb{Z}^d} \hat{h}(n) \exp(2\pi i n \cdot x) \tag{2.8}$$

Since β is a $\mathbb{Z} \ltimes_\lambda \mathbb{R}$-action on torus, it satisfies $\tilde{A} \circ \tilde{\Phi}^t = \tilde{\Phi}^{\lambda t} \circ \tilde{A}$. Differentiating in t, we get

$$(A + Df) \cdot (v + w) = \lambda(v + w) \circ (A + f)$$

or

$$A \cdot v + A \cdot w + Df \cdot v + Df \cdot w = \lambda v + \lambda w \circ (A + f)$$

Applying $A \cdot v = \lambda v$ and subtracting both sides by $\lambda w \circ A$,

$$A \cdot w + Df \cdot v - \lambda w \circ A = \lambda w \circ (A + f) - \lambda w \circ A - Df \cdot w \tag{2.9}$$

For notational simplicity, let E denote the right side of (2.9), i.e.,

$$E \coloneqq \lambda w \circ (A + f) - \lambda w \circ A - Df \cdot w$$

Its C^0 norm can be estimated by

$$||E||_0 \leq |\lambda| ||w||_1 ||f||_0 + ||f||_1 ||w||_0 \tag{2.10}$$

Writing equation (2.9) in Fourier coefficients, we get

$$A\hat{w}(n) + \hat{f}(n) 2\pi i n \cdot v - \lambda \hat{w}((A^{-1})^{\mathrm{T}}(n)) = \widehat{E}(n) \tag{2.11}$$

When $n = 0$, $A\hat{w}(0) - \lambda \hat{w}(0) = \hat{E}(0)$. This forces $\hat{E}(0)$ to lie in $V^{\perp}$. By Assumption 2.2.2,

$$\hat{w}(0) = (A - \lambda \cdot \mathrm{Id})|_{V^{\perp}}^{-1} \hat{E}(0). \tag{2.12}$$

When $n \neq 0$, $|2\pi i n \cdot v| > 0$ because of the Diophantine property of v. Dividing both sides of (2.11) by $2\pi i n \cdot v$ yields, with definition of h in (2.8),

$$-A(\hat{h}(n)) + \hat{f}(n) + \hat{h}((A^{-1})^{\mathrm{T}}(n)) = \frac{\hat{E}(n)}{2\pi i n \cdot v} \tag{2.13}$$

When $n = 0$, from the construction of $\hat{h}(0)$, it satisfies

$$A\hat{h}(0) = \hat{f}(0) + \hat{h}((A^{-1})^{\mathrm{T}}(0)). \tag{2.14}$$

Combining these, one has the following equality:

$$-A \circ h + f + h \circ A = E^* \tag{2.15}$$

where

$$E^* = \sum_{n \in \mathbb{Z} \setminus \{0\}} \frac{\hat{E}(n)}{2\pi i n \cdot v} \exp(2\pi i n \cdot x).$$

The following norm estimates for $\hat{E}(0)$ and E^* are straightforward from Remark 2.2.3

and Lemma 2.2.4.

$$||E^*||_r \ll_{r,d} |E^*|_{r+d+1} \ll |E|_{r+d+1+\tau} \ll_{r,d} ||E||_{r+d+1+\tau} \tag{2.16}$$

After all the preparation above, we state the following lemma.

Lemma 2.3.1. *Let α be an affine $\mathbb{Z} \ltimes_\lambda \mathbb{R}$-action on torus $\mathbb{T}^d$ generated by pair (A, v) and β be a $C^{r+d+1+\tau}$ $\mathbb{Z} \ltimes_\lambda \mathbb{R}$-action on $\mathbb{T}^d$ generated by pair $(A + f, v + w)$. Under assumption 2.2.2, there exists a solution h to equation* (2.15) *and*

$$-w + \hat{w}(0) = Dh \cdot v \tag{2.17}$$

Moreover

$$||h||_r \ll ||f||_0 + ||w - \hat{w}(0)||_{r+d+1+\tau} \leq ||f||_0 + ||w||_{r+d+1+\tau}. \tag{2.18}$$

Proof. Let h be given by (2.6) and (2.7). Then h satisfies (2.15) and (2.17) naturally. To estimate the norm of h, notice for $n \neq 0$, $|\hat{h}(n)| = |\frac{-\hat{w}(n)}{2\pi i n \cdot v}| \ll_v |\hat{w}(n)| \cdot |n|^\tau$ because $v \in DC(C, \tau)$. Combining with Lemma 2.2.4, we have

$$||h - \hat{h}(0)||_r \ll ||w - \hat{w}(0)||_{r+d+1+\tau}.$$

Moreover,

$$|\hat{h}(0)| = |(A - \mathrm{Id})^{-1}\hat{f}(0)| \ll_A ||f||_0.$$

This proves the lemma. □

2.3.2 Smoothing operators pair

It is clear from the above Lemma that there is certain loss of smoothness. In other words, we can only estimate the C^r norm of h by the $C^{r+d+1+\tau}$ norm of w.

To overcome this obstacle, the standard method in KAM is to introduce a family of smoothing operators $S_t : C^r(T^d, \mathbb{R}^d) \mapsto C^\infty(T^d, \mathbb{R}^d)$. Instead of solving $-w + \hat{w}(0) = Dh \cdot v$, one can solve the following equation:

$$-S_t w = Dh \cdot v$$

Here we use a discrete version of smoothing operators.

Lemma 2.3.2. *Suppose $\Lambda \subset \mathbb{Z}^d$ satisfies $\{n \in \mathbb{Z}^d : 0 < |n| \leq M\} \subseteq \Lambda \subseteq \{n \in \mathbb{Z}^d : 0 < |n| \leq N\}$ where $M < N$. For a sufficiently regular function $f(x) = \sum_{n \in \mathbb{Z}^d} \hat{f}(n) \exp(2\pi i n \cdot$*

x), Let

$$S_\Lambda f(x) = \sum_{n\in\Lambda} \hat{f}(n)\exp(2\pi i n\cdot x)$$

$$\dot{S}_\Lambda f(x) = \sum_{n\in\mathbb{Z}^d\setminus(\Lambda\cup\{0\})} \hat{f}(n)\exp(2\pi i n\cdot x).$$

Then,

$$|S_\Lambda f|_{a+b} \le N^b |f|_a$$
$$||S_\Lambda f||_{a+b} \ll_{a,b,d} N^{b+d+1}||f||_a$$

and

$$|\dot{S}_\Lambda f|_{a-b} \le M^{-b}|f|_a$$
$$||\dot{S}_\Lambda f||_{a-b} \ll_{a,b,d} M^{-b+d+1}||f||_a$$

for $a \ge b \ge 0$*.*

Proof. Since $S_\Lambda f$ and $\dot{S}_\Lambda f$ both satisfy assumption in Lemma 2.2.4, it suffices to estimate the $|\cdot|$ norms.

$$|S_\Lambda f|_{a+b} \le \sup_{0<|n|\le N} |\hat{f}(n)|\cdot|n|^{a+b}$$
$$\le \sup_{0<|n|,|A^{-1}n|\le N} |\hat{f}(n)|\cdot|n|^a N^b \le N^b|f|_a.$$

Similarly,

$$|\dot{S}_\Lambda f|_{a-b} = \sup_{n\notin\Lambda, n\neq 0} |\hat{f}(n)|\cdot|n|^{a-b} \le \sup_{|n|>M} |\hat{f}(n)|\cdot|n|^{a-b}$$
$$\le \sup_{|n|>M} ||\hat{f}(n)|\cdot|n|^a M^{-b} \le M^{-b}|f|_a.$$

□

We now define a group of operators using Lemma 2.3.2 as follows:

- $S_N = S_\Lambda$ and $\dot{S}_N = \dot{S}_\Lambda$, when

$$\Lambda = \{n\in\mathbb{Z}^d : 0 < |n| \le N\};$$

- $T_N^{\#} = S_\Lambda$ and $\dot{T}_N^{\#} = \dot{S}_\Lambda$, when

$$\Lambda = \{n \in \mathbb{Z}^d : 0 < |n| \le N, 0 < |A^{\mathrm{T}} n| \le N\};$$

- $T_N = S_\Lambda$ and $\dot{T}_N = \dot{S}_\Lambda$, when

$$\Lambda = \{n \in \mathbb{Z}^d : 0 < |n| \le N, 0 < |(A^{\mathrm{T}})^{-1} n| \le N\}.$$

Then they all satisfy the condition of Lemma 2.3.2 with $M = \|A\|^{-1} N$.

The same proof of Lemma 2.3.1 implies there exists h which solves the equation

$$-S_N w = Dh \cdot v. \tag{2.19}$$

However, (2.15) no longer holds. Instead, with the new operators constructed above, , we apply T_N to both sides of (2.11) to obtain

$$-A \circ T_N h + T_N f + (T_N^{\#} h) \circ A = T_N E^* \tag{2.20}$$

Instead of (2.18), we have by Lemma 2.3.2

$$||h||_r \ll_{r,d,A} \|f\|_0 + ||S_N w||_{r+d+1+\tau} \ll_{r,d} \|f\|_0 + N^{2d+2+\tau} ||w||_r. \tag{2.21}$$

We also have

$$||h||_{r+1} \ll_{r,d,A} \|f\|_0 + ||S_N w||_{r+d+2+\tau} \ll_{r,d} \|f\|_0 + N^{2d+3+\tau} ||w||_r. \tag{2.22}$$

All the above is summarized in the following lemma:

Lemma 2.3.3. *Let α be an affine $\mathbb{Z} \ltimes_\lambda \mathbb{R}$-action on torus $\mathbb{T}^d$ generated by pair (A, v) and β be a C^r $\mathbb{Z} \ltimes_\lambda \mathbb{R}$-action on $\mathbb{T}^d$ generated by pair $(A+f, v+w)$. Under Assumption 2.2.2, there exists a solution h to equation* (2.19), (2.20) *and the norm estimates* (2.21),(2.22).

2.3.3 Inductive lemma

Let α be the affine $\mathbb{Z} \ltimes_\lambda \mathbb{R}$-action generated by (A, v) and β be a C^r smooth perturbation generated by $(\tilde{A}, \tilde{v}) = (A + f, v + w)$. Denote $\epsilon_k = \|f\|_k$ and $\eta_k = \|w\|_k$.

If the perturbation β is close enough to α in C^r with $r > d + 1$, i.e. C^r norms of f and w are small enough, with proper chosen N, h's C^r norm would be small by Lemma

2.3.3 and invertibility of $\mathrm{Id}+h$ is guaranteed by inverse function theorem. We now state and prove the inductive lemma.

Lemma 2.3.4. *Under Assumption 2.2.2, there is a positive constant δ that depends only on d and A, suppose*

$$\epsilon_1 < 1, \tag{2.23}$$

and for some $N \in \mathbb{N}$,

$$\epsilon_0 + N^{2d+2+\tau}\eta_1 < \delta. \tag{2.24}$$

Then there exists $H \in \mathrm{Diff}^\infty(\mathbb{T}^d)$ such that $\beta' = H \circ \beta \circ H^{-1}$ is generated by $(\tilde{A}', \tilde{v}')$ and the following hold true for the $\mathbb{Z} \ltimes_\lambda \mathbb{R}$-action β' thus defined.

$$\begin{aligned}
||\tilde{A}' - A||_0 \ll_{r,d,A} & \, N^{-r+d+1}\epsilon_r + N^{-r+3d+3+\tau}\eta_r + N^{2d+2+\tau}\eta_0^{1-\frac{1}{r}}\eta_r^{\frac{1}{r}}\epsilon_0 \\
& + N^{d+1+\tau}\eta_0\epsilon_0^{1-\frac{1}{r}}\epsilon_r^{\frac{1}{r}} \\
||\tilde{A}' - A||_r \ll_{r,d,A} & \, 1 + \epsilon_r + N^{2d+2+\tau}\eta_r \\
||\tilde{v}' - v||_0 \ll_{r,d,A} & \, N^{-r+d+1}\eta_r + N^{2d+2+\tau}\eta_0^{2-\frac{1}{r}}\eta_r^{\frac{1}{r}} + \eta_0^{1-\frac{1}{r}}\eta_r^{\frac{1}{r}}\epsilon_0 + \eta_0\epsilon_0^{1-\frac{1}{r}}\epsilon_r^{\frac{1}{r}} \\
||\tilde{v}' - v||_r \ll_{r,d,A} & \, 1 + N^{2d+3+\tau}\eta_r.
\end{aligned}$$

Proof. Let h be as in Lemma 2.3.3. With a properly chosen constant δ, the condition (2.24) implies $||h||_1 < \frac{1}{2}$. In this case, $H = \mathrm{Id} + h$ is invertible, and H^{-1} is also C^r with $\|H^{-1} - \mathrm{Id}\|_r \ll_{r,d} \|h\|_r$ by inverse function theorem.

$$\begin{aligned}
& H \circ \tilde{A} \circ H^{-1} - A \\
= & H \circ (A + f) \circ H^{-1} - A \circ H \circ H^{-1} \\
= & ((\mathrm{Id} + h) \circ (A + f) - A \circ (\mathrm{Id} + h)) \circ H^{-1} \\
= & (f - A \circ h + h \circ A + h \circ \tilde{A} - h \circ A) \circ H^{-1} \\
= & (T_N f + \dot{T}_N f + \hat{f}(0) - A \circ T_N h - A \circ \dot{T}_N h - A \circ \hat{h}(0) \\
& + T_N^{\#} h \circ A + \dot{T}_N^{\#} h \circ A + \hat{h}(0) + h \circ \tilde{A} - h \circ A) \circ H^{-1} \\
= & (\dot{T}_N f) \circ H^{-1} - (A \circ \dot{T}_N h) \circ H^{-1} + (\dot{T}_N^{\#} h \circ A) \circ H^{-1} + \\
& (h \circ \tilde{A} - h \circ A) \circ H^{-1} + (T_N f - A \circ T_N h + T_N^{\#} h \circ A) \circ H^{-1}
\end{aligned}$$

where $\hat{f}(0) - A \circ \hat{h}(0) + \hat{h}(0) = 0$ with definition of $\hat{h}(0)$ in (2.7). Enumerate parentheses in the last expression above from the left to the right by $\Omega_1, \cdots, \Omega_5$. Their C^0 norm are bounded as below, using Lemma 2.3.2 and Lemma 2.3.3.

$$||\Omega_1||_0 = ||\dot{T}_N f||_0 \ll_{r,d} N^{-r+d+1}||f||_r$$

$$\begin{aligned}||\Omega_2||_0 &\leq ||A||_1||\dot{T}_N h||_0 = ||A||_1||\dot{T}_N(h - \hat{h}(0))||_0 \\ &\ll_{r,d,A} N^{-r+d+1}||h - \hat{h}(0)||_r \ll_{r,d} lN^{-r+3d+3+\tau}||w||_r\end{aligned}$$

$$\begin{aligned}||\Omega_3||_0 &\leq ||\dot{T}_N^{\#} h||_0 = ||\dot{T}_N^{\#}(h - \hat{h}(0))||_0 \\ &\ll_{r,d} N^{-r+d+1}||h - \hat{h}(0)||_r \ll_{r,d} N^{-r+3d+3+\tau}||w||_r\end{aligned}$$

$$\begin{aligned}||\Omega_4||_0 &\leq ||Dh||_0||f||_0 = ||D(h - \hat{h}(0))||_0||f||_0 \\ &\leq ||h - \hat{h}(0)||_1||f||_0 \ll_{r,d,A} N^{2d+2+\tau}||w||_1||f||_0\end{aligned}$$

Here, as $\dot{T}_N h$ and $\dot{S}_N h$ and Dh don't depend on zeroth Fourier frequency of h, we could replace h with $h - \hat{h}(0)$. Finally,

$$\begin{aligned}||\Omega_5||_0 &= ||T_N E^*||_0 \ll_d |T_N E^*|_{d+1} \ll_{d,A} N^{d+1+\tau}|E|_0 \leq N^{d+1+\tau}||E||_0 \\ &\ll N^{d+1+\tau}(||w||_1||f||_0 + ||f||_1||w||_0)\end{aligned}$$

by Lemma 2.2.4 and inequality (2.10).

Combining above inequalities and applying interpolation inequalities, we have the following norm error estimates

$$\begin{aligned}||H \circ \tilde{A} \circ H^{-1} - A||_0 \ll & N^{-r+d+1}||f||_r + N^{-r+3d+3+\tau}||w||_r \\ & + (N^{2d+2+\tau} + N^{d+1+\tau})||w||_0^{(1-\frac{1}{r})}||w||_r^{\frac{1}{r}}||f||_0 \\ & + N^{d+1+\tau}||w||_0||f||_0^{(1-\frac{1}{r})}||f||_r^{\frac{1}{r}}\end{aligned}$$

Now we estimate the C^r norm error of $H \circ \tilde{A} \circ H^{-1} - A$. By (2.23), (2.24), $\|f\|_0, \|h\|_1 <$

1. Moreover, $\|H^{-1}-\mathrm{Id}\|_r \ll_{r,d} \|h\|_r$. Hence by Lemma 2.2.5,

$$\begin{aligned}
&||H\circ\tilde{A}\circ H^{-1}-A||_r\\
\leq&||h\circ\tilde{A}\circ H^{-1}||_r+||f\circ H^{-1}||_r+||A\circ H^{-1}-A||_r\\
\ll_{r,d,A}&1+\|f|_r+\|h\|_r\\
\ll_{r,d,A}&1+\|f\|_r+N^{2d+2+\tau}\|w\|_r
\end{aligned}$$

In this norm estimate, error of the automorphism part depends on that of the flow part since Fourier coefficients of h at nonzero frequencies are solved from the flow part.

It remains to bound the error in the flow part, for which we have following decomposition:

$$\begin{aligned}
&(DH\cdot\tilde{v})\circ H^{-1}-v\\
=&(\mathrm{Id}+Dh)\cdot(v+w)\circ H^{-1}-v\\
=&(w+Dh\cdot v+Dh\cdot w)\circ H^{-1}\\
=&(\hat{w}(0)+\dot{S}_Nw+S_Nw+Dh\cdot v+Dh\cdot w)\circ H^{-1}\\
=&\hat{w}(0)\circ H^{-1}+\dot{S}_Nw\circ H^{-1}+(Dh\cdot w)\circ H^{-1}
\end{aligned}$$

Each term's C^0 norm are bounded as follows:

$$||\hat{w}(0)\circ H^{-1}||_0=||(A-\lambda\mathrm{Id})^{-1}\hat{E}(0)||_0\ll_{d,A}|\hat{E}(0)|\leq||E||_0$$

because of the assumption $\hat{w}(0)\in V^\perp$ from Assumption 2.2.2.

$$||\dot{S}_Nw\circ H^{-1}||_0=||\dot{S}_Nw||_0\ll_{r,d}N^{-r+d+1}||w||_r$$

$$\begin{aligned}
||(Dh\cdot w)\circ H^{-1}||_0\leq&||Dh||_0||w||_0\leq||h-\hat{h}(0)||_1||w||_0\\
\ll_{d,A}&N^{2d+2+\tau}||w||_1||w||_0
\end{aligned}$$

Combining them and applying interpolation inequalities, it can be improved into the following.

$$\begin{aligned}
||DH\cdot\tilde{v}\circ H^{-1}-v||_0\ll&N^{-r+d+1}||w||_r+N^{2d+2+\tau}||w||_0^{(2-\frac{1}{r})}||w||_r^{\frac{1}{r}}\\
&+||w||_0^{(1-\frac{1}{r})}||w||_r^{\frac{1}{r}}||f||_0+||w||_0||f||_0^{(1-\frac{1}{r})}||f||_r^{\frac{1}{r}},
\end{aligned}$$

Finally, we estimate the C^r norm error using Lemma 2.2.5, (2.21), (2.22) and the fact that $\|w\|_1, \|h\|_1 < 1$:

$$
\begin{aligned}
&||DH \cdot \tilde{v} \circ H^{-1} - v||_r \\
\ll_{r,d,A} & 1 + ||DH||_r + ||\tilde{v}||_r + |v| + \|H^{-1}\|_r \\
\ll_{r,d,A} & 1 + ||h||_{r+1} + \|w\|_r \\
\ll_{r,d,A} & 1 + N^{2d+3+\tau}||w||_r + ||w||_r \\
\ll_{r,d,A} & 1 + N^{2d+3+\tau}||w||_r
\end{aligned}
$$

By now we have proved the inductive lemma. □

2.4 Proof of Theorem 2.1.1

We now prove Theorem 2.1.1. Let $r \in \mathbb{N}$ be a parameter that will be determined later. For simplicity, we will omit the subscripts $_{r,d,A}$ and write $\ll$ for $\ll_{r,d,A}$ from now on.

In our proof, we shall not linearize $(\tilde{A}, \tilde{v})$ around (A, v). This is because, as we don't assume preservation of rotation vector, the flow part is not guaranteed to be conjugate to the original linear flow, and one has to allow a linear time change. To overcome this, in every step of the KAM scheme, one will linearize around a different affine action, which is obtained by a linear time change. This will not affect the condition (1) from Assumption 2.2.2, and the linear time change will be chosen so that condition (2) from Assumption 2.2.2 remains true.

To begin the process, we set up $\tilde{A}_0 = \tilde{A}$, $f_0 = f$, $\tilde{v}_0 = \tilde{v}$, $v_0 = v$, $\beta_0 = \beta$ and $H_0 = \text{Id}$. It will be assumed that this configuration satisfies Assumption 2.2.2 as well as the inequality (2.24).

In the n-th step, given a smooth action β_n generated by $\tilde{A}_n = A + f_n$ and $\tilde{v}_n = v_n + w_n$. Suppose v_n and w_n satisfies Assumption 2.2.2, and the inequality (2.24) with a properly chosen large integer N_n, then one can apply Lemma 2.3.4 to obtain a new smooth action β_{n+1}, generated by $\tilde{A}_{n+1} = A_n + f'_n$ and $\tilde{v}_{n+1} = v_n + w'_n$, and a conjugacy $H_{n+1} = \text{Id} + h_{n+1}$ between β_{n+1} and β_n. To insure Assumption 2.2.2 in the following step, we define

$$
f_{n+1} = f'_n, v_{n+1} = v_n + P_V(\widehat{w'_n}(0)), w_{n+1} = w'_n - P_V(\widehat{w'_n}(0)). \tag{2.25}
$$

Then $\tilde{A}_{n+1} = A_n + f_{n+1}$, $\tilde{v}_{n+1} = v_{n+1} + w_{n+1}$. Moreover, v_{n+1} is a vector from V, and w_{n+1} satisfies condition (2) in Assumption 2.2.2.

It should be remarked that, because w_{n+1} and w'_n differ by the constant vector $P_V(\widehat{w'_n}(0))$, which satisfies $|P_V(\widehat{w'_n}(0))| \ll \|w'_n\|_0$,

$$\|w_{n+1}\|_l \leq \|w'_n\|_l, \forall l \geq 0. \tag{2.26}$$

In the rest of the proof, we will write $\epsilon_{n,l} = ||f_n||_l$ and $\eta_{n,l} = ||w_n||_l$ for all $l \geq 0$ once f_n and w_n are defined.

In order to be able to iterate the inductive lemma, it remains to verify: condition (1) in Assumption 2.2.2 for v_{n+1}, namely that

$$|v_{n+1}| \in [\frac{1}{2}, 2]; \tag{2.27}$$

and the inequalities (2.23), (2.24) for f_{n+1}, w_{n+1} and a properly chosen integer N_{n+1}, namely that

$$\epsilon_{n+1,1} < 1. \tag{2.28}$$

$$\epsilon_{n+1,0} + N^{2d+3+\tau}\eta_{n+1,1} < \delta. \tag{2.29}$$

The sequence $\{N_n\}$ is determined as follows: let $N_0 = N_0(r, d, A)$ is a sufficiently large integer which in particular makes (2.24) hold for $\epsilon_{0,0}$ and $\eta_{0,1}$. We then fix some $\sigma \in (0, 1)$, say $\sigma = \frac{1}{2}$, and define N_n inductively by $N_{n+1} = N_n^{1+\sigma}$. With these choices, (2.27), (2.28), and (2.29) will be guaranteed by the following lemmas.

Lemma 2.4.1. *For all $r \geq 0$ and $k > \frac{2d+3+\tau}{\sigma}$, if $\epsilon_{n,r} \ll_r N_n^k$ and $\eta_{n,r} \ll_r N_n^k$, then $\epsilon_{n+1,r} \ll N_{n+1}^k$ and $\eta_{n+1,r} \ll N_{n+1}^k$.*

Proof. From Lemma 2.3.4, one deduce

$$\epsilon_{r,n+1} \ll \|f_n\|_r \ll 1 + N_n^k + N_n^{2d+2+\tau+k} \ll N_n^{2d+2+\tau+k} \ll (N_n^{1+\sigma})^k \ll N_{n+1}^k.$$

Using (2.26), we also get

$$\eta_{r,n+1} \ll \|w'_n\|_r \ll 1 + N_n^{2d+3+\tau+k} \ll N_n^{2d+3+\tau+k} \ll (N_n^{1+\sigma})^k \ll N_{n+1}^k.$$

This completes the proof of the lemma. □

We then prove that $\epsilon_{n+1,0}$ and $\eta_{n+1,0}$ are small.

Lemma 2.4.2. *Given σ and k as in Lemma 2.4.1. If*

$$r > \max\Big(\frac{2}{1-\sigma}, (1+\sigma)\Big(3d+3+\tau+2k+\frac{2(2d+2+\tau)}{1-\sigma}\Big)\Big),$$

then there exists a constant $y > 0$ such that:

If, in addition to the conditions in Lemma 2.4.1, $\epsilon_{n,0} \ll N_n^{-y}$ and $\eta_{n,0} \ll N_n^{-y}$, then $\epsilon_{n+1,0} \ll N_{n+1}^{-y}$ and $\eta_{n+1,0} \ll N_{n+1}^{-y}$.

Proof. By Lemma 2.3.4,

$$\begin{aligned}\epsilon_{n+1,0} \ll & N_n^{-r+d+1}\epsilon_{r,n} + N_n^{-r+3d+3+\tau}\eta_{r,n} + N_n^{2d+2+\tau}\eta_{0,n}^{1-\frac{1}{r}}\eta_{r,n}^{\frac{1}{r}}\epsilon_{0,n} \\ & + N_n^{d+1+\tau}\eta_{0,n}\epsilon_{0,n}^{1-\frac{1}{r}}\epsilon_{r,n}^{\frac{1}{r}} \\ \eta_{n+1,0} \ll & \|w_n'\|_0 \\ \ll & N_n^{-r+d+1}\eta_{r,n} + N^{2d+2+\tau}\eta_{0,n}^{2-\frac{1}{r}}\eta_{r,n}^{\frac{1}{r}} + \eta_{0,n}^{1-\frac{1}{r}}\eta_{r,n}^{\frac{1}{r}}\epsilon_{0,n} + \eta_{0,n}\epsilon_{0,n}^{1-\frac{1}{r}}\epsilon_{r,n}^{\frac{1}{r}}\end{aligned}$$

By our assumptions,

$$\begin{aligned}\epsilon_{0,n+1} \ll & N_n^{-r+d+1+k} + N_n^{-r+3d+3+\tau+k} + N_n^{2d+2+\tau+\frac{k}{r}}\eta_{0,n}^{1-\frac{1}{r}}\epsilon_{0,n} \\ & + N_n^{d+1+\tau+\frac{k}{r}}\eta_{0,n}\epsilon_{0,n}^{1-\frac{1}{r}} \\ \ll & N_n^{-r+3d+3+\tau+k} + N_n^{2d+2+\tau+\frac{k}{r}-(2-\frac{1}{r})y} + N_n^{d+1+\tau+\frac{k}{r}-(2-\frac{1}{r})y} \\ \eta_{0,n+1} \ll & N_n^{-r+d+1+k} + N^{2d+2+\tau+\frac{k}{r}}\eta_{0,n}^{2-\frac{1}{r}} + N_n^{\frac{k}{r}}\eta_{0,n}^{1-\frac{1}{r}}\epsilon_{0,n} \\ \ll & N_n^{-r+d+1+k} + N_n^{2d+2+\tau+\frac{k}{r}-(2-\frac{1}{r})y} + N_n^{\frac{k}{r}-(2-\frac{1}{r})y}\end{aligned}$$

To get $\epsilon_{0,n+1} \leq N_{n+1}^{-y}$ and $\eta_{0,n+1} \ll N_{n+1}^{-y}$, following inequalities must hold.

$$-r+3d+3+\tau+k \leq -(1+\sigma)y \tag{2.30}$$

$$2d+2+\tau+\frac{k}{r}-(2-\frac{1}{r})y \leq -(1+\sigma)y \tag{2.31}$$

$$d+1+\tau+\frac{k}{r}-(2-\frac{1}{r})y \leq -(1+\sigma)y \tag{2.32}$$

$$-r+d+1+k \leq -(1+\sigma)y \tag{2.33}$$

$$\frac{k}{r}-(2-\frac{1}{r})y \leq -(1+\sigma)y \tag{2.34}$$

Obviously, (2.30) ⇒ (2.33) and (2.31) ⇒ (2.32) ⇒ (2.34). We just need to solve the inequalities (2.30) and (2.31), which can be reduced to

$$\frac{2d+2+\tau+\frac{k}{r}}{1-\sigma-\frac{1}{r}} \leq y \leq \frac{r-3d-3-\tau-k}{1+\sigma} \tag{2.35}$$

For our choice of r, $\frac{r-3d-3-\tau-k}{1+\sigma} > \frac{2d+2+\tau+\frac{k}{r}}{1-\sigma-\frac{1}{r}} > 0$ always hold. Therefore such a parameter y exist. □

Corollary 2.4.3. *If N_0 is chosen to be sufficiently large, in the settings of Lemma 2.4.1 and Lemma 2.4.2, the inequalities* (2.28) *and* (2.29) *are both true.*

Proof. By Lemma 2.4.1, Lemma 2.4.2, interpolation inequalities, and (2.31),

$$\epsilon_{n+1,1} \ll \epsilon_{n+1,0}^{1-\frac{1}{r}} \epsilon_{n+1,r}^{\frac{1}{r}} \ll N_{n+1}^{-(1-\frac{1}{r})y+\frac{k}{r}} \leq N_{n+1}^{(-(2d+2+\tau)-\sigma)y}; \tag{2.36}$$

$$\eta_{n+1,1} \ll N_{n+1}^{-\sigma y}. \tag{2.37}$$

As $N_{n+1} > 1$, (2.23) follows immediately. For (2.24), notice

$$\begin{aligned} &\epsilon_{n+1,0} + N_{n+1}^{2d+2+\tau}\eta_{n+1,1} \\ \ll &N_{n+1}^{-y} + N_{n+1}^{2d+2+\tau+\frac{k}{r}-(1-\frac{1}{r})y} \ll N_{n+1}^{-y} + N_{n+1}^{-\sigma y}. \end{aligned}$$

As $\delta = \delta(r, d, A)$ is given and $N_{n+1} > N_0$, this implies (2.24) when N_0 is chosen to be sufficiently large. □

Remark 2.4.4. *Since $\tau = d - 1$, with $\sigma = \frac{1}{2}$ one may choose $k = 6d + 6$ in Lemma 2.4.1. It is then easy to verify that $r = 42(d + 1)$ is sufficient for Lemma 2.4.2.*

It remains to establish (2.27).

Corollary 2.4.5. *When N_0 is chosen to be sufficiently large, $|v_0| = 1$ and the conditions in Lemma 2.4.1 and Lemma 2.4.2 hold in the m-th step for every m from 0 to n, then the inequality* (2.27) *is true.*

Proof. Using the assumptions, we know for every m between 0 and n,

$$|v_{m+1} - v_m| = |P_V \widehat{w'_m(0)}| \ll \|w'_m(0)\|_0.$$

Remark that the proofs of Lemma 2.4.1 and Lemma 2.4.2 also applies to $\|w'_m\|_r$ and

$\|w_m'(0)\|_0$ instead of $\epsilon_{m,r}$ and $\epsilon_{m,0}$, and thus $\|w_m'(0)\| \ll N_{m+1}^{-y}$. It follows that

$$|v_{n+1}| - 1 = |v_{n+1}| - |v_0| \ll \sum_{m=0}^{n} N_{m+1}^{-y} = \sum_{m=0}^{n} \left(N_0^{((1+\sigma)^{m+1})}\right)^{-y} \ll_{\sigma,y} N_0^{-y}.$$

Once σ and y are given, one may choose a sufficiently large N_0, which now depends only on r, d and A, such that $|v_{n+1}| - 1 \le \frac{1}{2}$. □

What we have proved so far can now be summarized into the following:

Theorem 2.4.6. *There exists k, y, N_0 and δ_0, which depend only on d and A, such that, with $r = 42(d+1)$ and $\sigma = \frac{1}{2}$, if $|v_0| = 1$, $\|f_0\|_r < \delta_0$, $\|w_0\|_r < \delta_0$, then the inductive construction above can be iterated for all $n \ge 0$, and Lemma 2.4.1 and Lemma 2.4.2 can be applied in every step.*

Furthermore, the conjugacy $\tilde{H}_n = H_n \circ H_{n-1} \circ \cdots \circ H_0$ converges in C^1 topology to a diffeomorphism $H : \mathbb{T}^d \to \mathbb{T}^d$, which satisfies

$$H \circ \tilde{A} \circ H^{-1} = A, \quad DH \cdot \tilde{v} \circ H^{-1} = v^*,$$

where v^* is a vector proportional to v and $|v^*| \in [\frac{1}{2}, 2]$.

Proof. We may choose δ_0 to reflect the initial conditions $\epsilon_{0,0} \ll N_0^{-y}$, $\eta_{0,0} \ll N_0^{-y}$, $\epsilon_{0,r} \ll N_0^k$ and $\eta_{0,r} \ll N_0^k$. Then the validity of the inductive construction, and the applicability of the lemmas in all steps, follow from Corollaries 2.4.3 and 2.4.5 .

The C^1 convergence of $\tilde{H}_n$ to a diffeomorphism H is deduced from the bound

$$\begin{aligned}\|H_n - \mathrm{Id}\|_1 =& \|h\|_1 \ll \epsilon_{n+1,0} + N_{n+1}^{2d+2+\tau}\eta_{n+1,1} \\ &\ll N_{n+1}^{-y} + N_{n+1}^{2d+2+\tau+\frac{k}{r}-(1-\frac{1}{r})y} \ll N_{n+1}^{-y} + N_{n+1}^{-\sigma y}\end{aligned}$$

and the fact that $N_n = N_0^{((1+\sigma)^m)}$ is fast growing. As long as N_0 is large enough, $\|H - \mathrm{Id}\|_1 < 1$ and thus H is invertible.

Finally, by (2.36) and (2.37), as $n \to \infty$,

$$\|\tilde{H}_n \circ \tilde{A} \circ \tilde{H}_n^{-1} - A\|_1 = \epsilon_{n+1,1} \to 0.$$

Hence $H \circ \tilde{A} \circ H^{-1} = A$.

On the other hand, it follows from the proof of Corollary 2.4.5 that $\{v_n\}$ is a Cauchy sequence, thus $v^* = \lim_{n\to\infty} v_n \in V$ exists and satisfies $|v^*| \in [\frac{1}{2}, 2]$. Since

$$\|D\tilde{H}_n \cdot \tilde{v} \circ \tilde{H}_n^{-1} - v_{n+1}\|_0 = \eta_{n+1,0} \to 0,$$

we obtain $DH \cdot v \circ H^{-1} = v^*$. □

The last step in establishing Theorem 2.1.1 is to upgrade the regularity of H from C^1 to C^∞.

Proof of Theorem 2.1.1. By passing to a linear time change of the flow part, one may assume without loss of generality that $v_0 = v$ has length 1. Hence Theorem 2.4.6 applies.

The proof of the C^∞ convergence of the sequence $\{\tilde{H}_n\}$ from the proof of Theorem 2.4.6 to H is standard. Indeed, by Lemma 2.3.4, for all $p > 0$, there exists $U_p > 0$ that depends on p, d, A, such that

$$1 + \eta_{n+1,p} \le U_p(1 + N_n^{2d+3+\tau}\eta_{n,p}).$$

In consequence, there exists $n_0 = n_0(p, d, A)$, such that for all $n \ge n_0$,

$$\begin{aligned}\eta_{n,p} \le& U_p^n\Big(\prod_{i=0}^{n-1} N_i^{2d+3+\tau}\Big)(1+\eta_{0,p})\\ =& U_p^n N_0^{(2d+3+\tau)\sum_{i=0}^{n-1}(1+\sigma)^i}(1+\eta_{0,p})\\ =& U_p^n N_0^{(2d+3+\tau)\frac{(1+\sigma)^n-1}{\sigma}}(1+\eta_{0,p})\\ \le& N_0^{(2d+3+\tau)\frac{(1+\sigma)^n}{2\sigma}}(1+\eta_{0,p})\\ =& N_n^{\frac{2d+3+\tau}{2\sigma}}(1+\eta_{0,p})\end{aligned} \tag{2.38}$$

Choose $\mu \in (0, 1)$ that is sufficiently close to 0, so that

$$y_\mu := -\mu\frac{2d+3+\tau}{2\sigma} + (1-\mu)y > 0.$$

Then by interpolation between (2.38) and Lemma 2.4.2, for every $q \le \mu p$ and sufficiently large n,

$$\begin{aligned}\eta_{n,q} \ll_{p,q}& \eta_{n,0}^{1-\frac{q}{p}}\eta_{n,p}^{\frac{q}{p}}\\ \ll_{p,q,d,A}& N_n^{-(1-\frac{q}{p})\cdot y} N_n^{\frac{q}{p}\cdot\frac{2d+3+\tau}{2\sigma}}(1+\eta_{0,p})^{\frac{q}{p}}\\ \le& N_n^{-y_\mu}(1+\eta_{0,p})^\mu.\end{aligned} \tag{2.39}$$

Combining (2.39) with (2.21) and Lemma 2.2.4, we obtain that for $m < q-(3d+3+\tau)$,

$$\begin{aligned}\|H_n - \mathrm{Id}\|_m =&\|h_n\|_m \ll_{q,d,A} \epsilon_{n,0} + N_n^{3d+3+\tau}\eta_{n,m} \\ &\ll_{q,d,A}\epsilon_{n,0} + \eta_{n,q} \ll_{p,q,d,A} N_n^{-y} + N_n^{-y_\mu}(1+\eta_{0,p})^\mu.\end{aligned} \tag{2.40}$$

Since $N_n = N_0^{(1+\sigma)^n}$ is fast growing as $n \to \infty$, and $y > 0$, $y_\mu > 0$, the bound above leads to the convergence of $\tilde{H}^n$ to H in C^m. By letting $p \to \infty$, q and m can be arbitrarily large. Therefore the convergence holds in C^∞. □

Part II

Hausdorff dimension of ϵ-concentrating orbits on tori

Chapter 3
Introduction to ϵ-concentrating orbits on tori

3.1 Introduction

Given a dynamics system $f : X \to X$, to capture the asymptotic behavior of any orbits, it is convenient to set up some observing window $U \subset X$. We are interested in statistically describing the frequency an orbit hits U.

Here X is a topological space and f can be either an endomorphism or a homeomorphism. But we will mainly work with the forward orbits of points $x \in X$. We say that a point has dense orbit if its orbit closure equals X and nondense orbit otherwise. We will denote $ND(f)$ the set of points with nondense orbit. The study of nondense orbits is important for the interaction between ergodic theory and Diophantine approximation. Actually $ND(f)$ can be decomposed as $ND(f) = \cup_{y \in X} E(f, y)$ where

$$E(f, y) = \{x \in X : y \notin \overline{\{f^n(x) : n \in \mathbb{N} \text{ or } \mathbb{Z}\}}\}. \tag{3.1}$$

Hausdorff dimension of $E(f, y)$ has been investigated in various scenarios. It usually has full Hausdorff dimension if the system have some hyperbolic or expanding behavior. To name a few, Dani [Dan88] showed $E(f, y)$ has full Hausdorff dimension when $X = \mathbb{T}^n$, $f \in \mathrm{GL}(n, \mathbb{Z})$, f is semisimple and $y \in \mathbb{Q}^n/\mathbb{Z}^n$. It was later extended to arbitrary $f \in \mathrm{GL}(n, \mathbb{Z})$ and $y \in \mathbb{R}^n/\mathbb{Z}^n$ by Broderick, Fishman and Kleinbock [BFK11].

Similar idea was adapted to more general homogeneous spaces. Let $X = G/\Gamma$ where G is a Lie group and Γ is a lattice in G. In [Kad15], Kadyrov used exponential mixing results for diagonalizable flow in compact homogeneous space to show the Hausdorff dimension of set of points that lie on trajectories missing a particular open ball $U = B(x, r)$ of radius r is at most $\dim X + C\frac{r^{\dim X}}{\log r}$. Kleinbock and Mirzadeh [KM19] extended the work of Kadyrov [Kad15] to noncompact homogeneous space with U being a subset of X whose complement is compact and obtained similar result.

When the orbit (or trajectory) appears in the observing window U statistically, or equivalently saying the orbit leave the complement set U^c statistically, various problems has been formulated and well understood. One of them is the study of divergent trajectories in homogeneous dynamics. Let $X = G/\Gamma$ where G is a Lie group and Γ is a lattice in G, D be the direction of one parameter action by $\{\exp(tD)\}_{t\geq 0}$. A trajectory is divergent if it eventually leaves every compact set without returning. It was showed divergent trajectories of one parameter flows

$$\{\exp(tD)\}_{t\geq 0}, D = diag(m, \cdots, m, -n, \cdots, -n) \tag{3.2}$$

in $X = SL_d(\mathbb{R})/SL_d(\mathbb{Z})$ where $m + n = d$ has exact Hausdorff codimension $\frac{mn}{m+n}$ in works by Cheung and Chevallier [CC+16] and by Das, Fishman, Simmons and Urbanski [DFSU17, DFSU19]. Same flow was studied by Kadyrov [Kad12] and Kadyrov-Kleinbock-Lindenstrauss-Margulis [KKLM17]. Instead of considering divergent trajectories, they defined a more general notion called escaping on average. A point $x \in X$ is *escaping on average* (with respect to the semigroup $\exp(tD)$) if $\lim_{N\to\infty} \frac{1}{N}|\{l \in \{1, \ldots, N\} : \exp(lD)x \in Q\}| = 0$ for any compact set Q in X. Observe this notion is independent of the parametrization of the orbit, in other words x is *escaping on average* if and only if $\lim_{N\to\infty} \frac{1}{N}|\{l \in \{1, \ldots, N\} : \exp(alD)x \in Q\}| = 0$ for any $a > 0$. This definition means the empirical probability measure along the trajectories converges to 0 in the weak-* topology. In other words, even though the trajectory is allowed to return to a given compact set, on average it only spends a zero portion of time inside the compact set. In the same setting, [KKLM17] obtains an upper bound of Hausdorff codimension of points escaping on average $\frac{mn}{m+n}$.

In a recent work of Hertz and Wang [HW21], they played with trajectories which are called sequentially ϵ-escaping on average, more precisely, the following

Definition 3.1.1. *Given G/Γ where G, Γ are connected simple Lie subgroup and its lattice, $D \in \mathfrak{g}$ and $\epsilon \in (0, 1]$, a point $x \in G$ is sequentially ϵ-escaping with respect to the one parameter subgroup $\{\exp(tD)\}$, if there exists a sequence $T_k \to \infty$ and a weak-* limit μ of the sequence of probability measures*

$$\mu_{T^k} := \frac{1}{T^k}\int_0^{T^k} \delta_{\exp(tD)\cdot x}dt, \tag{3.3}$$

such that $\mu(G/\Gamma) \leq 1 - \epsilon$.

Compared with escaping on average, this definition loosen sampling times and allow a flexibility for an ϵ-portion of time which makes it capable of capturing more wild trajectories. They achieved a positive lower bound of the form $c\epsilon$ to the Hausdorff codimension of the set of sequentially ϵ-escaping on average.

In this paper, we consider a similar problem but in a compact scenario. Let X be a compact metric space and $f : X \to X$ a continuous map, $P(X)$ denotes space of Borel probability measure on X with weak-* topology . Given any point $x \in X$,

$$\delta_x^n = \frac{1}{n} \sum_{i=0}^{n-1} \delta_{f^i(x)},\ n \geq 1, \tag{3.4}$$

called the empirical measure where δ_y is Dirac measure at $y \in X$, define

$$Z_{\epsilon,f} = \left\{ x \in M \;\middle|\; \begin{array}{l} \exists \{n_k\}_k \subset \mathbb{N} \text{ such that } \mu_x = \lim_{k\to\infty} \delta_x^{n_k}, \\ \mu_x \ \text{ has at least } \epsilon \text{ mass supported on atoms} \end{array} \right\}$$

We are concerned about the Hausdorff dimension of $Z_{\epsilon,f}$.

This problem is motivated by the works mentioned above [KKLM17, HW21]. The ambient space G/Γ they work with is noncompact. Instead of escaping trajectories, we will handle trajectories spending a prescribed amount of time in an open neighborhood of a given set. Since we are working on compact space, it is possible to describe this concentrating behavior directly. If we image one point $x_0 \in X$ as ∞ like noncompact homogeneous space, following the same terminology as sequentially ϵ-escaping on average, x_0 has to be an atom and orbit that spends at least ϵ-portion amount of time concentrating to x_0 should be considered. Moreover, we shall expect similar result with [HW21]. It turns out this is what happens. We proved the following theorem:

Theorem 3.1.2. *If A is a toral automorphism on $\mathbb{T}^d$, then $\exists\ \kappa = \kappa(A)$ such that $\dim_H Z_{\epsilon,A} \leq d - \kappa\epsilon$. Moreover, if A has at least one eigenvalue with absolute value other than 1, κ can precisely be $\kappa = \kappa(d, A) = \frac{d}{2(d+1)^2} \cdot \frac{\log \rho}{\log \rho_{\max}}$ where ρ_{max} is the most expanding rate and ρ is the least expanding rate back and forth.*

Below is the precise definition of $\rho_{\max}$ and ρ. Suppose A has at least one eigenvalue that is not root of unity, then $\mathbb{R}^d$ can decompose as $E^u \oplus E^{cs}$ where both E^u and E^{cs} are A-invariant subspaces, and the absolute values of eigenvalues of A restricted to them are respectively > 1 and ≤ 1. Let Σ be the set of eigenvalues of A. Fix $\rho > 1$ by

$$\rho = \min\left(\{\left|\frac{1}{\sigma}\right| : \sigma \in \Sigma, |\sigma| < 1\} \cup \{|\sigma| : \sigma \in \Sigma, |\sigma| > 1\}\right) \tag{3.5}$$

and

$$\rho_{\max} = \max\{|\sigma| : \sigma \in \Sigma, |\sigma| > 1\} \tag{3.6}$$

Note that ρ and $\rho_{\max}$ only depend on A.

3.2 Hausdorff measure and Hausdorff dimension

Let (X, d) be a metric space. If U is a nonempty subset of X, $\operatorname{diam}(U)$ denotes the diameter of U which is defined as $\operatorname{diam}(U) = \sup\{d(x,y) : x, y \in U\}$. Let $E \subset X$ be any subset of a metric space, and $\epsilon > 0$ a real number. We define the Hausdorff outer measure

$$\mathcal{H}^{\delta}_{\epsilon}(E) = \inf_{U}\{\sum_{i=1}^{\infty} \operatorname{diam}(U_i)^{\delta} : E \subset \cup_{i=1}^{\infty} U_i \text{ and } \operatorname{diam}(U_i) \leq \epsilon\} \tag{3.7}$$

The s-dimensional Hausdorff measure of E is defined as $\mathcal{H}^{\delta}(E) = \lim_{\epsilon \to 0} \mathcal{H}^{\delta}_{\epsilon}(E)$. Finally, the Hausdorff dimension of E is defined as

$$\dim_H(E) = \inf\{\delta : \mathcal{H}^{\delta}(E) = 0\} \tag{3.8}$$

Proposition 3.2.1 (Elementary Properties of Hausdorff Dimension)**.** *Hausdorff dimension has the following properties:*

1. *if* $X \subset Y$, *then* $\dim_H(X) \leq \dim_H(Y)$*;*
2. *if* X_i *is a countable collection of sets with* $\dim_H(X_i) \leq d$, *then* $\dim_H(\cup_i X_i) \leq d$*;*
3. *if* $f : X \to f(X)$ *is a Lipschitz map, then* $\dim_H(f(X)) \leq \dim_H(X)$*.*

Note that the Hausdorff dimension is invariant under bi-Lipschitz transformation. Generally, Hausdorff dimensions are difficult to calculate. The box dimension, defined by

$$\dim_B(E) = \limsup_{\epsilon \to 0} \frac{\log N(\epsilon)}{\log(1/\epsilon)}, \tag{3.9}$$

gives an upper bound. Here $N(\epsilon)$ denotes the smallest number of ϵ-ball needed to cover E.

3.3 Properties of $Z_{\epsilon,f}$

Given a measure space $(X, \mathcal{A}, \mu)$, a set $A \subset X$ in $\mathcal{A}$ is called an atom if $\mu(A) > 0$ and for any measurable subset $E \subset A$ with $\mu(E) < \mu(A)$ the set E has measure zero. A singleton $\{a\} \in \mathcal{B}$ is a point mass if $\mu(\{a\}) > 0$. The following Lemma 3.3.1 shows atoms are point masses and disjoint to each other up to a measure zero set. By countable additivity, there are countable many point masses for every Borel probability supported on a separable space.

Lemma 3.3.1. *Let (X, d) be a separable metric space, $\mathcal{A}$ its Borel σ-algebra, and μ a σ-finite measure on $\mathcal{A}$. Then each atom of μ is the union of a point mass and a null set.*

Proof. Let C be a countable dense subset of X. For each integer $k \in \mathbb{N}$, we have $\cup_{x \in C} B(x, 1/k) = X$ where $B(x, 1/k)$ is the open ball centered at x with radius $1/k$. Thus $\cup_{x \in C}(A \cap B(x, 1/k)) = A$. By countable additivity, there exists $x_k \in C$ such that $\mu(A \cap B(x_k, 1/k)) > 0$. Since A is an atom. $\mu(A \setminus B(x_k, 1/k)) = 0$. Let $S = \cap_k B(x_k, 1/k)$. Since for each k, S is contained in a ball of radius $1/k$, S contains at most one point. On the other hand, by De Morgan's law and countable additivity,

$$\mu(A \setminus S) = \mu(\cup_k A \setminus B(x_k, 1/k)) = 0$$

Since $\mu(S \cap A) = \mu(A) > 0$, $A \cap S$ is not empty, so $A \cap S$ is a singleton. Hence $A \cap S$ is a point mass and $A \setminus S$ is a null set. □

Proposition 3.3.2. *If f is a bi-Lipschitz homeomorphism, then $\dim_H Z_{f,\epsilon} = \dim_H Z_{f^l,\epsilon}$, $\forall\, l \in \mathbb{N}^+$.*

Proof. Let $x \in Z_{\epsilon,f}$, that is, $\exists\ \{n_k\}$ and weak-* limit $\mu_x = \lim_{k\to\infty} \frac{1}{n_k} \sum_{i=0}^{n_k-1} \delta_{f^i(x)}$ such that μ_x has at least ϵ mass supported on atoms. For each n_k, rewrite $n_k = m_k + j_k$ where m_k is the largest integer smaller than n_k that has l as a factor and $0 \le j_k \le l-1$.

$$\frac{1}{n_k}\sum_{i=0}^{n_k-1} \delta_{f^i(x)} - \frac{1}{n_k}\sum_{i=0}^{m_k-1} \delta_{f^i(x)} = \frac{1}{n_k}\sum_{i=m_k}^{n_k-1} \delta_{f^i(x)} \le \frac{l-1}{n_k} \tag{3.10}$$

The limit of $\frac{1}{n_k}\sum_{i=m_k}^{n_k-1} \delta_{f^i(x)}$ exists since $\frac{l-1}{n_k}$ converges to zero since $n_k \to \infty$. Therefore $\frac{1}{m_k}\sum_{i=0}^{m_k-1} \delta_{f^i(x)} = \frac{n_k}{m_k}\frac{1}{n_k} \cdot \sum_{i=0}^{m_k-1} \delta_{f^i(x)}$ converges to μ_x. Without lost of generality, we can assume each n_k in the subsequence $\{n_k\}$ has l as a factor. Rearrange the Birkhoff sum,

$$\begin{aligned} \frac{1}{n_k}\sum_{i=0}^{n_k-1}\delta_{f^i(x)} &= \frac{1}{\frac{n_k}{l}\cdot l}\sum_{j=0}^{l-1}\sum_{i=0}^{\frac{n_k}{l}-1}\delta_{f^{l\cdot i+j}(x)} \\ &= \frac{1}{l}\sum_{j=0}^{l-1}\frac{1}{\frac{n_k}{l}}\sum_{i=0}^{\frac{n_k}{l}-1}\delta_{f^{l\cdot i}(f^j(x))} \end{aligned}$$

For notation simplicity, let $\{\mu^k_{x,j}\}$ denote sequences $\delta^{\frac{n_k}{l}}_{f^j(x)} = \frac{1}{\frac{n_k}{l}}\sum_i \delta_{f^{l\cdot i}(f^j(x))}$ for each $0 \le j < l$, and $\{\mu^k_x\}$ denote $\delta^{n_k}_x$. Let A be the union of all atoms of μ_x, $\mu_x(A) \ge \epsilon$. With Lemma 3.3.1, A can be assumed as a union of countable point masses. We have $\mu^k_x = \frac{1}{l}\sum_{j=0}^{l-1}\mu^k_{x,j}$. $\mu_x(A) = \lim \mu^k_x(A) \le \frac{1}{l}\sum_{j=0}^{j=l-1}\limsup_k \mu^k_{x,j}(A)$. Obviously there exists at least one j such that $\limsup_k \mu^k_{x,j}(A) \ge \mu_x(A) \ge \epsilon$. Passing to a subsequence of $\{\mu^k_{x,j}\}$ if necessary as $\{\mu^k_{x,j}\}$ may not converge, we can find a convergent subsequence of $\{\mu^k_{x,j}\}$ with a limit μ^j_x that has at least ϵ mass supported on A. Also A is a union of countable many points, which implies μ^j_x has at least ϵ mass supported on its own atoms. Therefore $f^j(x) \in Z_{f^l,\epsilon}$ and $x \in f^{-j}Z_{f^l,\epsilon}$ for some $0 \le j < l$. Equivalently, $Z_{f,\epsilon} \subset \cup_j f^{-j}Z_{f^l,\epsilon}$. Based on properties of Hausdorff dimension, $\dim_H Z_{f,\epsilon} \le \max_j \dim_H f^{-j}Z_{f^l,\epsilon} = \dim_H Z_{f^l,\epsilon}$.
The next step is to prove $\dim_H Z_{f,\epsilon} \ge \dim_H Z_{f^l,\epsilon}$. It suffices to show $Z_{f^l,\epsilon} \subset Z_{f,\epsilon}$. Consider $x \in Z_{f^l,\epsilon}$, suppose $\exists\{n_k\} \subset \mathbb{N}$ such that $\lim \frac{1}{nk}\sum_{i=0}^{n_k}\delta^{l\cdot i}_f x = \mu_x$. It is easy to verify $\lim \frac{1}{n_k}\sum_{i=0}^{n_k}\delta_{f^{l\cdot i}(f^j(x))} = f^j\mu_x$ where $f^j\mu_x$ is the pullback measure of μ under f^j. Obviously $f^j(x) \in Z_{f^l,\epsilon}$ for any $j \in \mathbb{Z}$. Moreover, we can pick the convergent subsequence simultaneously for them. Therefore

$$\frac{1}{l}\sum_{j=0}^{l-1}\frac{1}{n_k}\sum_{i=0}^{n_k-1}\delta_{f^{l\cdot i}(f^j(x))} = \frac{1}{n_k\cdot l}\sum_{i=0}^{n_k\cdot l-1}\delta_{f^i(x)} \tag{3.11}$$

The left side of equation converges and has limit $\frac{1}{l}\sum_{j=0}^{l-1}f^j\mu_x$. Since f^js are bijection and atoms are singletons, $f^j\mu_x$ also has at least ϵ mass supported on atoms as μ_x. So does their average. This implies the right side of (3.11) has a limit with at least ϵ mass supported on atoms. $x \in Z_{f,\epsilon}$ is proved. □

Proposition 3.3.3. *Let f be a continuous map on the compact metric space X and g be a continuous map on the compact metric space Y. Suppose $g : Y \to Y$ is factor of $f : X \to X$, that is, there exists a continuous map $h : X \to Y$ satisfying $h \circ f = g \circ h$. Then we have $Z_{\epsilon,f} \subset h^{-1}(Z_{\epsilon,g})$.*

Proof. Given a point $x \in Z_{\epsilon,f}$, it suffices to show $h(x) \in Z_{\epsilon,g}$. According proposition 3.3.1, all atoms coincide with a point mass up to a measure zero set. Without losing generality, we can assume atoms are point masses. Suppose x generates μ along its empirical measure subsequence at $\{n_k\}$and p is a point mass of μ, then $h(p)$ is a point mass of $h_*\mu$ which is the pushforward of μ under h and $h_*\mu(\{h(p)\}) \geq \mu(\{p\})$. Moreover $\mu = \lim_{k\to\infty} \delta_x^{n_k}$ implies $h_*\mu = \lim_{k\to\infty} h_*\delta_x^{n_k} = \lim_{k\to\infty} \delta_{h(x)}^{n_k}$. It follows that $h(x) \in Z_{\epsilon,g}$. ☐

3.4 Reduction to simple cases

Corollary 3.4.1. *To prove Theorem 3.1.2, we can assume A has no roots of unity other than 1 among its eigenvalues.*

Proof. Applying Proposition 3.3.2, $\dim_H Z_{\epsilon,A} = \dim_H Z_{\epsilon,A^l}$ for all $l \in \mathbb{N}^+$. We can always find a $l \in \mathbb{N}^+$ such that A^l has no roots of unity other than 1 among its eigenvalues. ☐

In the case that A is an unipotent toral automorphism, there is no expansiveness in the system. We simply set $\kappa_A = 0$ in Theorem 3.1.2 and obtain the trivial result. Therefore, from now on we can assume A has no roots of unity.

Corollary 3.4.2. *To prove Theorem 3.1.2, we can assume A is irreducible over $\mathbb{Q}$.*

Proof. Here we use induction method to prove this corollary. When $d = 1$, the only toral automorphism is identity. The result is trivial. When $d = 2$, A is either irreducible over $\mathbb{Q}$, or reducible and unipotent. If A is reducible and unipotent, $\kappa_A = 0$ always work as $\dim_H Z_{\epsilon,A} \leq d$. Therefore, to prove theorem 3.1.2 suffices to prove the case A is irreducible. Suppose the corollary holds for $d = n-1, n \geq 3$. Now we verify the case $d = n$. If A is reducible over $\mathbb{Q}$, then there exists a proper A-invariant rational subspace $V \subset \mathbb{T}^d$. It induces a quotient map $\bar{A} : \mathbb{T}^d/V \to \mathbb{T}^d/V$. Obviously $\bar{A}$ is a factor of A, Moreover the projection triple $h : \mathbb{T}^d \to \mathbb{T}^d/V$ is a vector bundle and $\mathbb{T}^d/V$ is a torus. According to Proposition 3.3.3, $\dim_H Z_{\epsilon,A} \leq \dim_H h^{-1}(Z_{\epsilon,\bar{A}}) \leq \dim V + \dim_H Z_{\epsilon,\bar{A}} \leq \dim V + \dim(\mathbb{T}^d/V) - \kappa_{\bar{A}}\epsilon \leq d - \kappa_{\bar{A}}\epsilon$ where $\kappa_{\bar{A}}$ depends on $\bar{A}$. Since $\bar{A}$ is a factor of A, the set of eigenvalues of $\bar{A}$ is a subset of the set of eigenvalue of A. Moreover, $\dim \mathbb{T}^d/V < d$. By the construction of κ, we have $\kappa_A \leq \kappa_{\bar{A}}$ which implies $\dim_H Z_{\epsilon,A} \leq d - \kappa_A\epsilon$. To finish the proof of theorem 3.1.2 is left to prove the case A is irreducible.

☐

Proof of Theorem 3.1.2. According Corollary 3.4.1 and Corollary 3.4.2, it suffices to prove Theorem 3.1.2 when $A \in \mathrm{GL}(d, \mathbb{Z})$ is irreducible matrix over $\mathbb{Q}$ without roots of unity among its eigenvalues. In this case, A cannot have all of its eigenvalues in the unit circle. This is because the eigenvalues are Galois conjugate algebraic units, and it is known that the only algebraic units with all Galois conjugates in the circle are the roots of unity. Based on Proposition 3.3.1, all atoms of μ_x are point masses up to a zero measure set. Since μ_x is A-invariant measure, for any point mass $\{y\}$ of μ_x, $\{A(y)\}$ is also a point mass of μ_x. Moreover, $\mu_x(\{y\}) = \mu_x(\{A(y)\})$. According to countable additivity, orbit $\{A^i(y), i = 1, \ldots, n, \ldots\}$ of y has to been a finite set. Or equivalently y is a periodic point of f. Therefore atoms of all such μ_x are periodic points up to a zero measure set. Periodic points of an irreducible toral automorphism with no roots of unity among its eigenvalues have to been rational. Let F denote all periodic points of A, then $F \subset \mathbb{Q}^d/\mathbb{Z}^d$. Let $\{F_h\}_h$ be an increasing family of finite subsets of $\mathbb{Q}^d/\mathbb{Z}^d \subseteq \mathbb{T}^d$ such that $\bigcup_{h=1}^{\infty} F_h = \mathbb{Q}^d/\mathbb{Z}^d$. Then the set $Z_{\epsilon,A}$ can be written as

$$Z_{\epsilon,A} = \bigcup_{h=1}^{\infty} Z_{\epsilon,A,F_h}.$$

Applying properties of Hausdorff dimension, $\dim_H(Z_{\epsilon,A}) \leq \dim_H(Z_{\epsilon,A,F_h})$. With Theorem 4.1.3, which will be stated and proved in Chapter 4, $\dim_H(Z_{\epsilon,A}) \leq \kappa_A \epsilon$. $\square$

Chapter 4 is devoted to obtaining Theorem 4.1.3.

Chapter 4
Hausdorff dimension estimate of ϵ-concentrating orbits

4.1 Preliminaries

Recall ρ and $\rho_{\max}$ which are defined in Chapter 3. $\mathbb{R}^d$ decomposes as $E^u \oplus E^{cs}$ where both E^u and E^{cs} are A-invariant subspaces, and the absolute values of eigenvalues of A restricted to them are respectively > 1 and ≤ 1. Let Σ be the set of eigenvalues of A. Fix $\rho > 1$ by

$$\rho = \min\left(\{\left|\frac{1}{\sigma}\right| : \sigma \in \Sigma, |\sigma| < 1\} \cup \{|\sigma| : \sigma \in \Sigma, |\sigma| > 1\}\right) \tag{4.1}$$

and

$$\rho_{\max} = \max\{|\sigma| : \sigma \in \Sigma, |\sigma| > 1\} \tag{4.2}$$

We also chose

$$\lambda \in (0, \rho^{\frac{1}{d+1}}). \tag{4.3}$$

Note that the choices of λ and ρ can be made dependent only on A.

The following lemma is obvious.

Lemma 4.1.1. *Equip $\mathbb{R}^d$ with the metric induced by a positively definite quadratic form Q on $\mathbb{R}^d$. The choice of Q does not affect the Hausdorff dimension of a subset.*

This is because any two such metrics are bounded by each other up to a finite multiplicative constant.

Notation 4.1.2. *Since A is assumed to be irreducible, and in particular diagonalizable, we can fix a positive definite quadratic form Q_A with respect to which the eigenspaces of*

A are mutually orthogonal. Without further notice, $|\cdot|$ will denote the norm induced by Q_A and the distance we work with will be by default the one induced by this norm. By the lemma above, this will not affect the calculation of Hausdorff dimensions. Write $\mathrm{m}_{\mathbb{T}^d}$ for the volume form on $\mathbb{T}^d$. We can always renormalize Q_A by a constant so that $\mathrm{m}_{\mathbb{T}^d}(\mathbb{T}^d) = 1$.

We will denote by $B_r(x)$ the ball of radius r around a point x and $B_r(X) = \bigcup_{x\in X} B_r(x)$ the neighborhood of radius r around a subset X (both with respect to the distance from Notation 4.1.2).

The set $Z_{\epsilon,A,F}$ is defined as:

$$\begin{aligned} & Z_{\epsilon,A,F} \\ =& \{x \in \mathbb{T}^d : \exists N_k \to \infty \text{ and a weak-* limit } \mu \text{ of } \frac{1}{N_k}\sum_{n=0}^{N_k-1} \delta_{A^n x} \text{ s.t. } \mu(F) \geq \epsilon\} \\ =& \{x \in \mathbb{T}^d : \exists N_k \to \infty \text{ and a weak-* limit } \mu \text{ of } \frac{1}{N_k}\sum_{n=1}^{N_k} \delta_{A^n x} \text{ s.t. } \mu(F) \geq \epsilon\}. \end{aligned} \tag{4.4}$$

Then the set $Z_{\epsilon,A}$ can be written as

$$Z_{\epsilon,A} = \bigcup_{h=1}^{\infty} Z_{\epsilon,A,F_h}, \tag{4.5}$$

where $\{F_h\}_h$ is an increasing family of finite subsets of $\mathbb{Q}^d/\mathbb{Z}^d \subseteq \mathbb{T}^d$ such that $\bigcup_{h=1}^{\infty} F_h = \mathbb{Q}^d/\mathbb{Z}^d$.

$Z_{\epsilon,A,F}$ is further related to

$$Z_{\epsilon,A,F,r} = \{x \in \mathbb{T}^d : \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_{B_r(F)}(A^n x) \geq \epsilon\} \tag{4.6}$$

by

$$Z_{\epsilon,A,F} = \bigcap_{r>0} Z_{\epsilon,A,F,r}. \tag{4.7}$$

And $Z_{\epsilon,A,F,r}$ is in turn related to

$$Z_{\epsilon,A,F,r,N} = \{x \in \mathbb{T}^d : \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_{B_r(F)}(A^n x) \geq \epsilon\} \tag{4.8}$$

by

$$Z_{\epsilon,A,F,r} = \bigcap_{\epsilon'\in(0,\epsilon)} \bigcap_{M\geq 1} \bigcup_{N\geq M} Z_{\epsilon',A,F,r,N}. \tag{4.9}$$

The goal is to bound $\dim_H Z_{\epsilon,A,F_h}$ uniformly for all h by the following theorem.

Theorem 4.1.3. *Suppse $A \in \mathrm{GL}(d,\mathbb{Z})$ has at least one eigenvalue that is not a root of unity. Then for all $\epsilon > 0$ and finite set F,*

$$\dim_H Z_{\epsilon,A,F} \leq d - \kappa\epsilon,$$

where the constant $\kappa = \kappa(d,A) = \frac{d}{2(d+1)^2} \cdot \frac{\log \rho}{\log \rho_{\max}}$ is determined by the constants ρ, $\rho_{\max}$ defined above.

Remark 4.1.4. *In the special case where $d = 2$, ρ and $\rho_{\max}$ are equal, and $\kappa = \frac{1}{9}$ is independent of A.*

Below, we will suppose the parameter $r = r(d,A,F,\epsilon)$ is sufficiently small and leave it to be determined later. In particular, the following hypothesis will be in place:

$$r \leq \frac{1}{10} \min\Big(1, \min_{\substack{x,x' \in F \\ x \neq x'}} d(x,x')\Big). \tag{4.10}$$

In the decomposition $\mathbb{R}^d = E^u \oplus E^{cs}$, let m_{E^u} and $\mathrm{m}_{E^{cs}}$ be respectively the volume on E^u and E^{cs} arising from the Euclidean metric in Notation 4.1.2. Then $E^u \perp E^{cs}$. Denote the ball of radius r inside the E^u leaf around a point $x \in \mathbb{T}^d$ by $B^u_r(x)$. By abusing notation, we also let $B^u_r(0)$ be the ball of radius r around 0 inside E^u. Note that if $r < \frac{1}{2}$, this ball is injectively embedded into $\mathbb{T}^d$. We shall denote indifferently by m_{E^u} the uniform measure induced by m_{E^u} on $B^u_r(x)$. The neighborhood $B^{cs}_r(x)$ and the uniform measure $\mathrm{m}_{E^{cs}}$ on it are defined similarly. Then

$$d\mathrm{m}_{\mathbb{R}^d}(v^u \oplus v^{cs}) = d\mathrm{m}_{E^u}(v^u) \cdot d\mathrm{m}_{E^{cs}}(v^{cs}). \tag{4.11}$$

Further, due to orthogonality between eigenspaces, we have

$$|Av| \leq \rho_{\max}|v|, \forall v \in \mathbb{R}^d, \tag{4.12}$$

$$|Av| \geq \rho|v|,\ |A^{-1}v| \leq \rho^{-1}|v|, \forall v \in E^u. \tag{4.13}$$

$$|Av| \leq |v|, \forall v \in E^{cs}. \tag{4.14}$$

Our proof is based on the following exponential mixing property due to Lind:

Proposition 4.1.5. [Lin82] *Given $\gamma \in (0,1]$ and $f, g \in C^\gamma(\mathbb{T}^d)$, for sufficiently large $n \in \mathbb{N}$ (depending on d, A, γ),*

$$\left|\int_{\mathbb{T}^d} f(A^n(x))g(x)\mathrm{d}x - \int_{\mathbb{T}^d} f(x)\mathrm{d}x \cdot \int_{\mathbb{T}^d} g(x)\mathrm{d}x\right| \ll_{d,A,\lambda} \lambda^{-\gamma n}\|f\|_{C^\gamma}\|g\|_{C^\gamma}.$$

This is [Lin82, Theorem 6]. We note that though the range for γ was $(0,1)$ in [Lin82], the proof therein works for $\gamma = 1$ as well. Our methods below will work with any value in $(0,1]$ for γ. But we will stick to the choice $\gamma = 1$ for simplicity and best outcome.

We shall fix a smooth bump function $f : \mathbb{T}^d \to [0,1]$ supported on $B_{3r}(F)$, such that: $f|_{B_{2r}(F)} = 1$ and $\|f\|_{C^1} \ll \frac{1}{r}$. Indeed, it suffices to take a radial bump function in the ball of radius $3r$ around each point from F and sum them up. By (4.10) these balls are mutually disjoint. Note that

$$|F|r^d \ll_d \mathrm{m}_{\mathbb{T}^d}(B_{2r}(F)) \le \int_{\mathbb{T}^d} f(x)\mathrm{d}x \le \mathrm{m}_{\mathbb{T}^d}(B_{3r}(F)) \ll_d |F|r^d \tag{4.15}$$

In addition, for any $x_0 \in \mathbb{T}^d$, let g_{x_0} be a $[0,1]$-valued smooth bump function supported on $B_{3r}(x_0)$, such that: $g_{x_0}|_{B_{2r}(x_0)} = 1$ and $\|g_{x_0}\|_{C^1} \ll \frac{1}{r}$. Similar to (4.15),

$$r^d \ll_d \mathrm{m}_{\mathbb{T}^d}(B_{2r}(x_0)) \le \int_{\mathbb{T}^d} g_{x_0}(x)\mathrm{d}x \le \mathrm{m}_{\mathbb{T}^d}(B_{3r}(x_0)) \ll_d r^d \tag{4.16}$$

Applying Proposition 4.1.5 to f_x and g yields

$$\begin{aligned}&\left|\int_{\mathbb{T}^d} f(A^n(x))g_{x_0}(x)\mathrm{d}x - \int_{\mathbb{T}^d} f(x)\mathrm{d}x \cdot \int_{\mathbb{T}^d} g_{x_0}(x)\mathrm{d}x\right| \\ \ll_{d,A,\lambda}&\lambda^{-n}\|f\|_{C^1}\|g_{x_0}\|_{C^1} \ll_d \lambda^{-n}\cdot r^{-1}\cdot r^{-1} = \lambda^{-n}r^{-2}.\end{aligned}$$

Combining with (4.15), (4.16), we know that

$$\int_{\mathbb{T}^d} f(A^n(x))g_{x_0}(x)\mathrm{d}x \ll_{d,A,\lambda} \quad |F|r^d\cdot r^d + \lambda^{-n}r^{-2} \ll_{d,A,\lambda,F} \quad r^{2d} + \lambda^{-n}r^{-2}.$$

Corollary 4.1.6. *If $q = q(d, A, \lambda, r) \in \mathbb{N}$ satisfies*

$$\rho^{-q} \le \frac{1}{3} \text{ and } \lambda^{-q} \le r^{2d+2}; \tag{4.17}$$

and for all $n \ge q$ and $x_0 \in \mathbb{T}^d$,

$$\mathrm{m}_{\mathbb{T}^d}(\{x \in B_{2r}(x_0) : A^n x \in B_{2r}(F)\}) \ll_{d,A,\lambda,F} r^{2d}.$$

Proof. For $n \geq q$,
$$\int_{\mathbb{T}^d} f(A^n(x)) g_{x_0}(x) \mathrm{d}x \ll_{d,A,\lambda,F} r^{2d}.$$
It then suffices to notice that, because $\mathbf{1}_{B_r(F)} \leq f$ and $\mathbf{1}_{B_{2r}(x_0)} \leq g_{x_0}$,

$$\begin{aligned}
&\mathrm{m}_{\mathbb{T}^d}(\{x \in B_{2r}(x_0) : A^n x \in B_{2r}(F)\}) \\
&= \int_{\mathbb{T}^d} \mathbf{1}_{B_{2r}(F)}(A^n(x)) \mathbf{1}_{B_{2r}(x_0)}(x) \mathrm{d}x \\
&\leq \int_{\mathbb{T}^d} f(A^n(x)) g_{x_0}(x) \mathrm{d}x.
\end{aligned}$$

□

Corollary 4.1.7. *For all $n \geq q$ and $x_0 \in \mathbb{T}^d$,*

$$\mathrm{m}_{E^u}(\{v \in B_r^u(0) : A^n(x_0 + v) \in B_r(F)\}) \ll_{d,A,\lambda,F} r^{d + \dim E^u}.$$

Proof. Suppose $x \in B_r^u(x_0)$ and $A^n(x) \in B_r(F)$. We claim that for all $v \in E^{cs}$ such that $|v| \leq r$, $x + v \in B_{2r}(x_0)$ and $A^n(x + v) \in B_{2r}(F)$. The first assertion is obvious because $|v| \leq r$. For the second assert, since by (4.14) $|A^n v| \leq |v| \leq r$, the assumption that $A^n(x) \in B_r(F)$ implies $A^n(x + v) = A^n x + A^n v \in B_{2r}(F)$.

In other words, the map $(x, y) \to x + y$ sends $\{x \in B_r^u(x_0) : A^n x \in B_r(F)\} \times B_r^{cs}(0)$ to $\{x \in B_{2r}(x_0) : A^n x \in B_{2r}(F)\}$. Furthermore, as $r + r \leq 2r < 1$, this map is injective. Hence, by (4.11), we know that

$$\begin{aligned}
&\mathrm{m}_{E^u}(\{v \in B_r^u(0) : A^n(x_0 + v) \in B_r(F)\}) \cdot r^{\dim E^{cs}} \\
&= \mathrm{m}_{E^u}(\{x \in B_r^u(x_0) : A^n x \in B_r(F)\}) \cdot r^{\dim E^{cs}} \\
&\ll_d \mathrm{m}_{E^u}(\{x \in B_r^u(x_0) : A^n x \in B_r(F)\}) \cdot \mathrm{m}_{E^{cs}}(B_r^{cs}(0)) \\
&\leq \mathrm{m}_{\mathbb{T}^d}(\{x \in B_{2r}(x_0) : A^n x \in B_{2r}(F)\}) \\
&\ll_{d,A,\lambda,F} r^{2d}.
\end{aligned}$$

The corollary follows as $d = \dim E^u + \dim E^{cs}$. □

4.2 Coding of the exceptional set

To make use of Corollary 4.1.7, we argue as in [HW21] and define a sequence of probability measures μ_n on E^u as follows. First fix

$$\nu_1 = \frac{\mathrm{m}_{E^u}|_{B^u_r(0)}}{\mathrm{m}_{E^u}(B^u_r(0))}$$

which is the uniform probability measure on the ball of radius r around 0 in E^u.

For all $n \in \mathbb{N}$, the pushforward $\nu_n := (A^{-(n-1)q})_* \nu_1$ is a probability measure supported on $A^{-(n-1)q} B^u_r(0)$. By the choice of q in Corollary 4.1.6 and the definition of ρ,

$$\mathrm{supp}\nu_n \subseteq B^u_{\rho^{-(n-1)q}r}(0) \subseteq B^u_{3^{-(n-1)}r}(0). \tag{4.18}$$

We then define the convolution

$$\mu_n = \nu_1 * \cdots * \nu_n.$$

Lemma 4.2.1. *For all $n \in \mathbb{N}$, μ_n is a probability measure supported on $B^u_{\frac{3}{2}r}(0)$, and $\mu_n|_{B^u_{\frac{1}{2}r}(0)} = \nu_1|_{B^u_{\frac{1}{2}r}(0)}$.*

Proof. This is basically a special case of [HW21, Lemma 7.1], we provide a proof for completeness.

First, by (4.18), the support of μ_n lies in the ball around 0 of radius $\sum_{k=1}^{\infty} 3^{-(k-1)} r = \frac{3}{2} r$.

On the other hand, for all $v_k \in \mathrm{supp}\nu_k$, $k = 1, \cdots, n$,

$$|\sum_{k=2}^{n} v_k| \leq \sum_{k=2}^{n} 3^{-(k-1)} r = \frac{1}{2} r.$$

Therefore, for all subset $D \subseteq B^u_{\frac{1}{2}r}(0)$,

$$\begin{aligned}
\mu_n(D) =& \int_{\prod_{k=2}^n \mathrm{supp}\nu_k} \left(\int_{\mathrm{supp}\nu_1} \mathbf{1}_D(v_1 + \sum_{k=2}^{n} v_k) \mathrm{d}\nu_1(v_1) \right) \left(\prod_{k=2}^{n} \mathrm{d}\nu_k(v_k) \right) \\
=& \int_{\prod_{k=2}^n \mathrm{supp}\nu_k} \nu_1 \Big(D - \sum_{k=2}^{n} v_k \Big) \Big(\prod_{k=2}^{n} \mathrm{d}\nu_k(v_k) \Big) \\
=& \int_{\prod_{k=2}^n \mathrm{supp}\nu_k} \nu_1(D) \Big(\prod_{k=2}^{n} \mathrm{d}\nu_k(v_k) \Big) \\
=& \nu_1(D).
\end{aligned}$$

Here we used the fact that when $D \subseteq B^u_{\frac{1}{2}r}(0)$ and $v \in B^u_{\frac{1}{2}r}(0)$, $D - v \subseteq B^u_r(0)$ and $\nu_1(D-v) = \nu_1(D)$ by the translation invariance of $\nu_1 = \frac{\mathrm{m}_{E^u}|_{B^u_r(0)}}{\mathrm{m}_{E^u}(B^u_r(0))}$. Therefore μ_n coincides with ν_1 on $B^u_{\frac{1}{2}r}(0)$. □

Given $N \in \mathbb{N}$, instead of $x \in B^u_r(x_0)$ in Corollary 4.1.7, we will consider a random tuple $\mathbf{y} = (y_1, y_2, \cdots, y_N) \in (B^u_r(0))^N$. For $n = 0, \cdots, N$ and $x_0 \in \mathbb{T}^d$, write $\iota_{n,x_0}(\mathbf{y}) = x_0 + \sum_{k=1}^n A^{-(k-1)q} y_k$. (The summation is meant to be empty when $n = 0$.)

Write $\mathbf{y}|_n = (y_1, y_2, \cdots, y_n)$ for the truncation of $\mathbf{y}$ of length n. Then for all $n \leq m \leq N$, $\iota_{n,x_0}(\mathbf{y}) = \iota_{n,x_0}(\mathbf{y}|_m)$.

Lemma 4.2.2. *For $\mathbf{y} \in (B^u_r(0))^N$, $x_0 \in \mathbb{T}^d$ and $0 \leq n \leq m \leq N$, if $A^{nq}\iota_{m,x_0}(\mathbf{y}) \in B_r(F)$, then $A^{nq}\iota_{n,x_0}(\mathbf{y}) \in B_{2r}(F)$.*

Proof. Observe that $\iota_{m,x_0}(\mathbf{y}) = \iota_{n,x_0}(\mathbf{y}) + \sum_{k=n+1}^m A^{-(k-1)q} y_k$. Moreover,

$$\left| A^{nq} \sum_{k=n+1}^{m} A^{-(k-1)q} y_k \right| = \left| \sum_{k=0}^{m-n-1} A^{-kq} y_{n+1+k} \right| \leq \frac{1}{2} r.$$

Here the fact $\left| \sum_{k=0}^{m-n-1} A^{-kq} y_{n+1+k} \right| \leq \frac{1}{2} r$ is obtained as in the proof of Lemma 4.2.1.

Therefore $A^{nq}\iota_{n,x_0}(\mathbf{y}) = A^{nq}\iota_{m,x_0}(\mathbf{y}) - A^{nq} \sum_{k=n+1}^m A^{-(k-1)q} y_k$ is within distance $\frac{1}{2}r$ from $A^{nq}\iota_{m,x_0}(\mathbf{y}) \in B_r(F)$, and is thus in $B_{2r}(F)$. □

For $N \in \mathbb{N}$ and $x \in \mathbb{T}^d$, define

$$J^q_{A,F,r,x,N} = \{1 \leq n \leq N : A^{nq} x \in B_r(F)\}.$$

For any subset $J \subseteq \{1, \cdots, N\}$ and $x_0 \in \mathbb{T}^d$, define

$$Z^q_{A,F,r,N,J} := \{x : J^q_{A,F,r,x,N} = J\}.$$

Suppose for a tuple $\mathbf{y} \in (B^u_r(0))^N$, $\iota_{N,x_0}(\mathbf{y}) \in Z^q_{A,F,r,N,J}$. Then for every $n \in J \backslash \{0\}$, $A^{nq}\iota_{N,x_0}(y) \in B_r(F)$ and, by the lemma above, $A^{nq}\iota_{n,x_0}(\mathbf{y}) \in B_{2r}(F)$. It follows that

$$A^{nq}\iota_{n-1,x_0}(\mathbf{y}) + A^q y_n = A^{nq}(\iota_{n-1,x_0}(\mathbf{y}) + A^{-(n-1)q} y_n) \in B_{2r}(F),$$

or equivalently

$$\begin{aligned} y_n \in &\{v \in B^u_r(0) : A^q(A^{(n-1)q}\iota_{n-1,x_0}(\mathbf{y}) + v) \in B_{2r}(F)\} \\ = &\{v \in B^u_r(0) : A^q(A^{(n-1)q}\iota_{n-1,x_0}(\mathbf{y}|_{n-1}) + v) \in B_{2r}(F)\}. \end{aligned} \tag{4.19}$$

We denote the set in the last line by $\Lambda_{A,F,r,n,x_0}(\mathbf{y}|_{n-1})$.

By Corollary 4.1.7 and the choice of ν_1,

$$\nu_1\Big(\Lambda_{A,F,r,n,x_0}(\mathbf{y}|_{n-1})\Big) \ll_{d,A,\lambda,F} \frac{r^{d+\dim E^u}}{\mathrm{m}_{E^u}(B_r^u(0))} \ll_d \frac{r^d}{\mathrm{m}_{E^u}(B_r^u(0))} r^{\dim E^u} \ll_d r^d. \tag{4.20}$$

Write $C = C(d, A, \lambda, F)$ for the implied constant in (4.20).

Corresponding to $Z^q_{A,F,r,N,J}$, define a new set

$$\widetilde{Z}^q_{A,F,r,x_0,N,J} = \{\mathbf{y} \in (B_r^u(0))^N : \iota_{N,x_0}(\mathbf{y}) \in Z^q_{A,F,r,N,J}\}$$

Lemma 4.2.3. *For all $N \in \mathbb{N}$, $J \subset \{1, \cdots, N\}$ and $x_0 \in \mathbb{T}^d$,*

$$\nu_1^N(\widetilde{Z}^q_{A,F,r,x_0,N,J}) \leq (Cr^d)^{|J|}.$$

Proof. First define for all $n \leq N$ the projection set

$$(\widetilde{Z}^q_{A,F,r,x_0,N,J})|_n = \{\mathbf{y}' \in (B_r^u(0))^n : \exists \mathbf{y}'' \in (B_r^u(0))^{N-n} \text{ s.t. } (\mathbf{y}', \mathbf{y}'') \in \widetilde{Z}^q_{A,F,r,x_0,N,J}\}.$$

It then suffices to prove the inductive claim

$$\nu_1^n\Big((\widetilde{Z}^q_{A,F,r,x_0,N,J})|_n\Big) \leq (Cr^d)^{|J\cap\{1,\cdots,n\}|}, \forall 0 \leq n \leq N. \tag{4.21}$$

Indeed, the statement of the lemma is the case of $n = N$.

When $n = 0$, both sides of (4.21) are equal to 1 if we define ν_1^0 as the probability measure on the set of a singleton. This constitutes the base step of the induction.

Suppose (4.21) is known for $n-1$. We distinguish between two cases. First, if $n \notin J$, then because $(\widetilde{Z}^q_{A,F,r,x_0,N,J})|_{n-1}$ is the projection of $(\widetilde{Z}^q_{A,F,r,x_0,N,J})|_n$ to the first $n-1$ coordinates,

$$\begin{aligned}\nu_1^n\Big((\widetilde{Z}^q_{A,F,r,x_0,N,J})|_n\Big) \leq& \nu_1^{n-1}\Big((\widetilde{Z}^q_{A,F,r,x_0,N,J})|_{n-1}\Big)\\ \leq&(Cr^d)^{|J\cap\{1,\cdots,n-1\}|}\\ =&(Cr^d)^{|J\cap\{1,\cdots,n\}|}.\end{aligned}$$

If $n \in J$ instead, then by (4.19) and (4.20),

$$\begin{aligned}
&\nu_1^n\Big((\widetilde{Z}^q_{A,F,r,x_0,N,J})|_n\Big)\\
=&\int_{(\widetilde{Z}^q_{A,F,r,x_0,N,J})|_{n-1}} \nu_1\Big(\Lambda_{A,F,r,n,x_0}(\mathbf{y}|_{n-1})\Big)\mathrm{d}\nu_1^{n-1}(\mathbf{y}|_{n-1})\\
\leq& Cr^d \cdot (Cr^d)^{|J\cap\{1,\cdots,n-1\}|}\\
=&(Cr^d)^{|J\cap\{1,\cdots,n\}|}.
\end{aligned}$$

The lemma is established. □

Corollary 4.2.4. *For all $N \in \mathbb{N}$, $J \subset \{1, \cdots, N\}$ and $x_0 \in \mathbb{T}^d$,*

$$\frac{\mathrm{m}_{E^u}(B^u_{\frac{1}{2}r}(x_0) \cap Z^q_{A,F,r,N,J})}{\mathrm{m}_{E^u}(B^u_{\frac{1}{2}r}(0))} \leq (Cr^d)^{|J|}.$$

Proof. Remark that by construction,

$$\mathrm{d}(\iota_{N,x_0})_*\nu_1^N(x_0+v) = \mathrm{d}\mu_N(v). \tag{4.22}$$

From this and Lemma 4.2.3 we derive:

$$\mu_N(\{v \in B^u_{\frac{3}{2}r}(0) : x_0 + v \in Z^q_{A,F,r,N,J}\}) \leq (Cr^d)^{|J|},$$

where the radius $\frac{3}{2}$ comes from the bound on suppμ_N from Lemma 4.2.1. By the last assertion in Lemma 4.2.1, we see that

$$\begin{aligned}
&\nu_1(\{v \in B^u_{\frac{1}{2}r}(0) : x_0 + v \in Z^q_{A,F,r,N,J}\})\\
\leq&\mu_N(\{v \in B^u_{\frac{3}{2}r}(0) : x_0 + v \in Z^q_{A,F,r,N,J}\}) \leq (Cr^d)^{|J|},
\end{aligned}$$

which is equivalent to the statement of the corollary by the construction of ν_1. □

We can now switch back, from working within an E^u-leaf, to measurements inside $\mathbb{T}^d$.

Proposition 4.2.5. *There exists $C = C(d, A, \lambda, F)$ such that, for all $r > 0$ satisfying* (4.10) *and $q \in \mathbb{N}$ satisfying* (4.17)*, for all $N \in \mathbb{N}$, $J \subset \{1, \cdots, N\}$,*

$$\mathrm{m}_{\mathbb{T}^d}(Z^q_{A,F,r,N,J}) \ll_{d,A} (Cr^d)^{|J|}.$$

Proof. We first show that: for all $x_0 \in \mathbb{T}^d$,

$$\frac{\mathrm{m}_{\mathbb{T}^d}(B_{\frac{1}{2}r}(x_0) \cap Z^q_{A,F,r,N,J})}{\mathrm{m}_{\mathbb{T}^d}(B_{\frac{1}{2}r}(0))} \ll_{d,A} (Cr^d)^{|J|}. \tag{4.23}$$

Cover $B_{\frac{1}{2}r}(x_0)$ by the union $P = \bigcup_{x' \in B^{cs}_{\frac{1}{2}r}(x_0)} B^u_{\frac{1}{2}r}(x')$ where $B^{cs}_{\frac{1}{2}r}(x_0)$ is the ball of radius r around x_0 in the leaf along E^{cs}, which is the orthogonal complement of E^u. We have

$$\begin{aligned}
\mathrm{m}_{\mathbb{T}^d}(B_{\frac{1}{2}r}(x_0) \cap Z^q_{A,F,r,N,J}) \ll_d & \mathrm{m}_{\mathbb{T}^d}(P \cap Z^q_{A,F,r,N,J}) \\
\leq & \mathrm{m}_{(E^u)^{cs}}(B^{cs}_{\frac{1}{2}r}(0)) \cdot \mathrm{m}_{E^u}(B^u_{\frac{1}{2}r}(0)) \cdot (Cr^d)^{|J|} \\
\ll_d & \mathrm{m}_{\mathbb{T}^d}(B_{\frac{1}{2}r}(0)) \cdot (Cr^d)^{|J|}.
\end{aligned}$$

This establishes (4.23). To deduce the proposition, note that we can cover $\mathbb{T}^d$ with $O_d(r^{-d})$ balls $B_{\frac{1}{2}r}(x_i)$ of radius $\frac{1}{2}$. Therefore

$$\begin{aligned}
\mathrm{m}_{\mathbb{T}^d}(Z^q_{A,F,r,N,J}) \leq & \sum_i \mathrm{m}_{\mathbb{T}^d}(B_{\frac{1}{2}r}(x_i) \cap Z^q_{A,F,r,N,J}) \\
\ll_d & r^{-d} \mathrm{m}_{\mathbb{T}^d}(B_{\frac{1}{2}r}(x_0)) \cdot (Cr^d)^{|J|} \\
\ll_d & (Cr^d)^{|J|}.
\end{aligned}$$

□

4.3 Hausdorff dimension estimate

Define

$$Z^q_{\epsilon,A,F,r,N} := \{x \in \mathbb{T}^d, |J^q_{A,F,r,x,N}| \geq \epsilon N\},$$

and note $Z_{\epsilon,A,F,r,N} = Z^1_{\epsilon,A,F,r,N}$.

Corollary 4.3.1. *There exists $C = C(d, A, \lambda, F)$ such that, for all $r > 0$ satisfying* (4.10) *and q satisfying* (4.17)*, for all $N \in \mathbb{N}$, $\epsilon \in (0, 1]$, we have*

$$\mathrm{m}_{\mathbb{T}^d}(Z^q_{\epsilon,A,F,r,N}) \ll_d 2^N (Cr^d)^{\epsilon N}.$$

Proof. Clearly $\mathrm{m}_{\mathbb{T}^d}(Z^q_{\epsilon,A,F,r,N}) = \sum_{\substack{J \subseteq \{1,\cdots,N\} \\ |J| \geq \epsilon N}} \mathrm{m}_{\mathbb{T}^d}(Z^q_{A,F,r,N,J})$. We bound the number of such subset J's trivially by 2^N. □

Proposition 4.3.2. *There exists $C = C(d, A, \lambda, F)$ such that, for all $r > 0$ satisfying (4.10) and $q \in \mathbb{N}$ satisfying (4.17), for all integers $N \geq q$ and $\epsilon \in (0, 1]$,*

$$\mathrm{m}_{\mathbb{T}^d}(Z_{\epsilon,A,F,r,N}) \ll_d q 2^{\frac{N}{q}} (Cr^d)^{\frac{\epsilon N}{q}+\epsilon}.$$

Proof. Remark that for a point $x \in \mathbb{T}^d$ and a remainder $b = 1, \cdots, q$, the congruence class within $J_{A,F,r,x,N}$ at remainder b modulo q is given by

$$J_{A,F,r,x,N} \cap (q\mathbb{Z} + b) = \{qn - q + b : n \in J^q_{A,F,r,A^{-q+b}x,\lfloor \frac{N-b+q}{q} \rfloor}\}.$$

If $|J_{A,F,r,x,N}| \geq \epsilon N$, then for at least one $b \in \{1, \cdots, q\}$,

$$|J^q_{A,F,r,A^{-q+b}x,\lfloor \frac{N-b+q}{q} \rfloor}| \geq \epsilon \lfloor \frac{N-b+q}{q} \rfloor.$$

It follows that

$$Z_{\epsilon,A,F,r,N} \subseteq \bigcup_{b=1}^{q} A^{q-b} Z^q_{A,F,r,\lfloor \frac{N-b+q}{q} \rfloor,\epsilon}. \tag{4.24}$$

As A^{q-b} preserves the volume $\mathrm{m}_{\mathbb{T}^d}$, (4.24) and Corollary 4.3.1 together implies

$$\begin{aligned} \mathrm{m}_{\mathbb{T}^d}(Z_{\epsilon,A,F,r,N}) &\ll_d q 2^{\lfloor \frac{N-1+q}{q} \rfloor} (Cr^d)^{\epsilon \lfloor \frac{N-1+q}{q} \rfloor} \\ &\ll_d q 2^{\frac{N}{q}} (Cr^d)^{\frac{\epsilon N}{q}+\epsilon} \end{aligned} \tag{4.25}$$

□

From this volume upper bound, we deduce an upper bound for the covering of $Z_{\epsilon,A,F,r,N}$

Corollary 4.3.3. *In the setting of Proposition 4.3.2, $Z_{\epsilon,A,F,\frac{1}{2}r,N}$ can be covered by $I_{\epsilon,A,F,\frac{1}{2}r,N} \ll_d \rho_{\max}^{dN} r^{-d} \cdot q 2^{\frac{N}{q}} (Cr^d)^{\frac{\epsilon N}{q}+\epsilon}$ balls of radius $\frac{1}{4}\rho_{\max}^{-N} r$.*

Proof. One can always find a finite covering $\{B_{\frac{1}{4}\rho_{\max}^{-N} r}(y_i)\}$ of $\mathbb{T}^d$ by no more than $O_{d,A}\Big((\rho_{\max}^{-N} r)^{-d}\Big)$ balls of radius $\frac{1}{4}\rho_{\max}^{-N} r$, such that each point of $\mathbb{T}^d$ is covered by at most $O_d(1)$ such balls.

Hence it follows from Proposition 4.3.2 that $B_{\frac{1}{4}\rho_{\max}^{-N} r}(y_i) \subseteq Z_{\epsilon,A,F,r,N}$ holds for no more than

$$O_d(\rho_{\max}^{-N} r)^{-d} \mathrm{m}_{\mathbb{T}^d}(Z_{\epsilon,A,F,r,N}) \ll_{d,A} \rho_{\max}^{dN} r^{-d} \cdot q 2^{\frac{N}{q}} (Cr^d)^{\frac{\epsilon N}{q}+\epsilon} \tag{4.26}$$

such balls.

We claim that: if for some y_i, $B_{\frac{1}{4}\rho_{\max}^{-N}r}(y_i)\cap Z_{\epsilon,A,F,\frac{1}{2}r,N}\neq\emptyset$, then $B_{\frac{1}{4}\rho_{\max}^{-N}r}(y_i)\subseteq Z_{\epsilon,A,F,r,N}$. In fact, suppose $x\in B_{\frac{1}{4}\rho_{\max}^{-N}r}(y_i)\cap Z_{\epsilon,A,F,\frac{1}{2}r,N}$. For all $x'\in B_{\frac{1}{4}\rho_{\max}^{-N}r}(y_i)$, $d(x,x')\leq\frac{1}{2}\rho_{\max}^{-N}r$. By (4.12), $d(A^nx,A^nx')\leq\frac{1}{2}r$ for all $0\leq n\leq N$. Therefore $A^nx'\in B_r(F)$ whenever $A^nx\in B_{\frac{1}{2}r}(F)$, or in other words $J_{A,F,\frac{1}{2}r,x,N}\subseteq J_{A,F,\frac{1}{2}r,x',N}$. Because $x\in Z_{\epsilon,A,F,\frac{1}{2}r,N}$, $|J_{A,F,r,x',N}|\geq|J_{A,F,\frac{1}{2}r,x,N}|\geq\epsilon N$ and thus $x'\in Z_{\epsilon,A,F,r,N}$. This proves the claim, and leads to the first statement of the corollary with the bound (4.26). □

Proposition 4.3.4. *In the setting of Proposition 4.3.2, given $\epsilon>0$, if in addition r and q satisfy*

$$(Cr^d)^\epsilon<\frac{1}{2}\ , \tag{4.27}$$

then the Hausdorff dimension of the set $\bigcap_{M\geq1}\bigcup_{N\geq M}Z_{\epsilon,A,F,r,N}$ is bounded by

$$d+\frac{\log(2^{\frac{1}{q}}(Cr^d)^{\frac{\epsilon}{q}})}{\log\rho_{\max}}.$$

Proof. By Corollary 4.3.3, the union $\bigcup_{N\geq M}Z_{\epsilon,A,F,r,N}$, and therefore the intersection set $\bigcap_{M\geq1}\bigcup_{N\geq M}Z_{\epsilon,A,F,r,N}$ itself, can be covered by a countable union of balls

$$\{B_{N,i}:N\geq M,1\leq i\leq I_{\epsilon,A,F,\frac{1}{2}r,N}\}$$

such that $B_{N,i}$ has radius $\frac{1}{4}\rho_{\max}^{-N}r$.

Since for $s\geq0$,

$$\begin{aligned}\sum_{N=M}^{\infty}\sum_{i=1}^{I_{\epsilon,A,F,\frac{1}{2}r,N}}(\frac{1}{4}\rho_{\max}^{-N}r)^s\lesssim_{d,A}&\sum_{N=M}^{\infty}\rho_{\max}^{dN}r^{-d}\cdot q2^{\frac{N}{q}}(Cr^d)^{\frac{\epsilon N}{q}+\epsilon}\rho_{\max}^{-sN}r^s\\ \leq&\frac{1}{2}qr^{s-d}\sum_{N=M}^{\infty}\left(2^{\frac{1}{q}}(Cr^d)^{\frac{\epsilon}{q}}\rho_{\max}^{(d-s)}\right)^N.\end{aligned}\tag{4.28}$$

converges if and only if

$$2^{\frac{1}{q}}(Cr^d)^{\frac{\epsilon}{q}}\rho_{\max}^{(d-s)}<1. \tag{4.29}$$

Furthermore, as long as (4.29) holds, both the sum (4.28) and the maximal radius $\frac{1}{4}\rho_{\max}^{-M}r$ of the covering tends to 0 as $M\to\infty$. In other words, the Hausdorff measure $\mathcal{H}^s(Z_{\epsilon,A,F,r,N})$ is 0. Because (4.29) holds when s is greater than the critical value $d+\frac{\log(2^{\frac{1}{q}}(Cr^d)^{\frac{\epsilon}{q}})}{\log\rho_{\max}}$, the proposition follows. □

We next drop the quantifier N.

Proposition 4.3.5. *In the setting of Proposition 4.3.2, if r, q and ϵ satisfy the conditions* (4.10), (4.17) *and* (4.27), *then*

$$\dim_H Z_{\epsilon,A,F,r} \le d + \frac{\log(2^{\frac{1}{q}}(Cr^d)^{\frac{\epsilon}{q}})}{\log \rho_{\max}}.$$

Proof. By taking limit, it suffices to show that for any $\epsilon' < \epsilon$ that is sufficiently close to ϵ so that (4.27) continues to be true for ϵ', $\dim_H Z_{\epsilon,A,F,r} \le d + \frac{\log(2^{\frac{1}{q}}(Cr^d)^{\frac{\epsilon'}{q}})}{\log \rho_{\max}}$.

This later fact follows from (4.9), which says $Z_{\epsilon,A,F,r} \subseteq \bigcap_{M\ge 1}\bigcup_{N\ge M} Z_{\epsilon',A,F,r,N}$, and Proposition 4.3.4. □

We are now ready to establish Theorem 4.1.3.

Proof of Theorem 4.1.3. First, let q rely on r by choosing the minimal q satisfying (4.17), i.e.

$$q = q(d, A, r) = \max\left(\left\lceil \frac{-(2d+2)\log r}{\log \lambda} \right\rceil, \frac{\log 3}{\log \rho}\right). \tag{4.30}$$

While keeping ϵ, F and λ fixed, let r tend to 0. By doing this, the conditions (4.10) and (4.27) are guaranteed. Moreover, in this case $-\log r \to \infty$ and $q = \left\lceil \frac{-(2d+2)\log r}{\log \lambda} \right\rceil$. So we have

$$\begin{aligned}
&\liminf_{r\to 0} \dim_H Z_{\epsilon,A,F,r}\\
\le&\lim_{r\to 0}\left(d + \frac{\log(2^{\frac{1}{q}}(Cr^d)^{\frac{\epsilon}{q}})}{\log \rho_{\max}}\right)\\
=&\lim_{r\to 0}\left(d + \frac{\log 2 + \epsilon\log C + \epsilon d \log r}{\left\lceil \frac{-(2d+2)\log r}{\log \lambda}\right\rceil \log \rho_{\max}}\right)\\
=&\lim_{r\to 0}\left(d + \frac{\epsilon d \log r}{\frac{-(2d+2)\log r}{\log \lambda}\cdot \log \rho_{\max}}\right)\\
=&d - \kappa_\lambda \epsilon,
\end{aligned}$$

where $\kappa_\lambda = \frac{d}{2(d+1)} \cdot \frac{\log \lambda}{\log \rho_{\max}}$.

By the relation (4.7),

$$\dim_H Z_{\epsilon,A,F} \le \liminf_{r\to 0} \dim_H Z_{\epsilon,A,F,r} \le d - \kappa_\lambda \epsilon.$$

Finally, notice that the construction of $Z_{\epsilon,A,F}$ does not depend on the parameter λ, which can take any value in $(1, \rho^{\frac{1}{d+1}})$. By letting λ tend to $\rho^{\frac{1}{d+1}}$, κ_λ converges to $\kappa = \frac{d}{2(d+1)^2} \cdot \frac{\log \rho}{\log \rho_{\max}}$. In consequence, $\dim_H Z_{\epsilon,A,F} \le d - \kappa\epsilon$. □

Bibliography

[BFK11] R. Broderick, L. Fishman, and D. Kleinbock, *Schmidt's game, fractals, and orbits of toral endomorphisms*, Ergodic Theory and Dynamical Systems **31** (2011), no. 4, 1095–1107.

[CC$^+$16] Y. Cheung, N. Chevallier, et al., *Hausdorff dimension of singular vectors*, Duke Mathematical Journal **165** (2016), no. 12, 2273–2329.

[Dan88] S. Dani, *On orbits of endomorphisms of tori and the schmidt game*, Ergodic Theory and Dynamical Systems **8** (1988), no. 4, 523–529.

[DFSU17] T. Das, L. Fishman, D. Simmons, and M. Urbański, *A variational principle in the parametric geometry of numbers, with applications to metric diophantine approximation*, Comptes Rendus Mathematique **355** (2017), no. 8, 835–846.

[DFSU19] T. Das, L. Fishman, D. Simmons, and M. Urbański, *A variational principle in the parametric geometry of numbers*, arXiv preprint arXiv:1901.06602 (2019).

[Dia08] J. L. Dias, *Local conjugacy classes for analytic torus flows*, Journal of Differential Equations **245** (2008), no. 2, 468–489.

[DK10] D. Damjanović and A. Katok, *Local rigidity of partially hyperbolic actions I. KAM method and $\mathbb{Z}^k$ actions on the torus*, Annals of mathematics **172** (2010), no. 3, 1805–1858.

[DlL01] R. De la Llave, *A tutorial on kam theory*, Proceedings of symposia in pure mathematics, 2001, pp. 175–296.

[dlL92] R. de la Llave, *Smooth conjugacy and srb measures for uniformly and non-uniformly hyperbolic systems*, Communications in mathematical physics **150** (1992), no. 2, 289–320.

[dlLMM86] R. de la Llave, J. M. Marco, and R. Moriyón, *Canonical perturbation theory of anosov systems and regularity results for the livsic cohomology equation*, Annals of Mathematics **123** (1986), no. 3, 537–611.

[Fis05] D. Fisher, *Local rigidity of group actions: past, present, future*, arXiv preprint math (2005).

[GKS18] A. Gogolev, B. Kalinin, and V. Sadovskaya, *Local rigidity of lyapunov spectrum for toral automorphisms*, arXiv preprint arXiv:1808.06249 (2018).

[Ham82] R. S Hamilton, *The inverse function theorem of nash and moser*, American Mathematical Society **7** (1982), no. 1, 65–122.

[HW21] F. R. Hertz and Z. Wang, *On ϵ-escaping trajectories in homogeneous spaces*, Discrete & Continuous Dynamical Systems-A **41** (2021), no. 1, 329.

[Kad12] S. Kadyrov, *Entropy and escape of mass for hilbert modular spaces*, Journal of Lie Theory **22** (2012), no. 3, 701–722.

[Kad15] S. Kadyrov, *Exceptional sets in homogeneous spaces and hausdorff dimension*, Dynamical Systems **30** (2015), no. 2, 149–157.

[Kar18] N. Karaliolios, *Local rigidity of diophantine translations in higher-dimensional tori*, Regular and Chaotic Dynamics **23** (2018), no. 1, 12–25.

[Kat71] Y. Katzenlson, *Ergodic automorphisms oft n are bernoulli shifts*, Israel Journal of Mathematics **10** (1971), no. 2, 186–195.

[KKLM17] S. Kadyrov, D. Kleinbock, E. Lindenstrauss, and G. A. Margulis, *Singular systems of linear forms and non-escape of mass in the space of lattices*, Journal d'Analyse Mathématique **133** (2017), no. 1, 253–277.

[KM19] D. Kleinbock and S. Mirzadeh, *Dimension estimates for the set of points with non-dense orbit in homogeneous spaces*, Mathematische Zeitschrift (2019), 1–29.

[Lin82] D. A Lind, *Dynamical properties of quasihyperbolic toral automorphisms*, Ergodic Theory and Dynamical Systems **2** (1982), no. 1, 49–68.

[MZ89] M. Misiurewicz and K. Ziemian, *Rotation sets for maps of tori*, Journal of the London Mathematical Society **2** (1989), no. 3, 490–506.

[SY18] R. Saghin and J. Yang, *Lyapunov exponents and rigidity of anosov automorphisms and skew products*, arXiv preprint arXiv:1802.08266 (2018).

[WX20] A. Wilkinson and J. Xue, *Rigidity of some abelian-by-cyclic solvable group actions on* $\mathbb{T}^n$, Communications in Mathematical Physics **376** (2020), no. 2, 1223–1259.

[Yoc02] J.-C. Yoccoz, *Analytic linearization of circle diffeomorphisms*, Dynamical systems and small divisors, 2002, pp. 125–173.

Vita

Qiao Liu

Education

2014-2021 Ph.D in Mathematics -Pennsylvania State University, University Park, USA
Advisors: Anatole Katok and Zhiren Wang

2012-2014 M.Sc in Mathematics - IMPA, Rio de Janeiro, Brazil

2008-2012 B.Sc in Mathematics - Nankai University, Tianjin, China

Publications

1. Local rigidity of certain solvable group actions on tori, *Discrete & Continuous Dynamical Systems - A*, 41(2020), no. 2, 553