RIGIDITY OF SOME ABELIAN-BY-CYCLIC SOLVABLE GROUP ACTIONS ON $\mathbb{T}^N$

AMIE WILKINSON, JINXIN XUE

ABSTRACT. In this paper, we study the rigidity properties of a class of affine solvable group actions on $\mathbb{T}^N$. In these actions, elliptic and hyperbolic dynamics coexist. We first show, using the KAM method, that any small and sufficiently smooth perturbation of an affine action can be conjugated smoothly to an affine action, provided certain Diophantine conditions are met. In dimension two, under natural dynamical hypotheses, we get a complete classification of such actions; namely, any such group action by C^r diffeomorphims can be conjugated to the affine action by $C^{r-\varepsilon}$ conjugacy. Next, we show that in any dimension, C^1 small perturbations can be conjugated to an affine action via C^2 conjugacy. The method is a generalization of the Herman theory for circle diffeomorphisms to higher dimensions in the presence of a foliation structure provided by the hyperbolic dynamics.

CONTENTS

Date: May 22, 2017.

1. Introduction

This paper is motivated by an attempt to understand the action on $\mathbb{T}^2$ generated by the diffeomorphisms

$$g_0(x,y) = (2x+y, x+y), \quad g_1(x,y) = (x+\rho, y), \quad g_2(x,y) = (x, y+\rho).$$

The map g_0 is a hyperbolic linear automorphism, and g_1, g_2 are translations. They satisfy the group relations

$$g_0 g_1 = g_1^2 g_2 g_0, \quad g_0 g_2 = g_1 g_2 g_0, \quad g_1 g_2 = g_2 g_1,$$

and no other relations if ρ is irrational. Our aim is to classify all diffeomorphisms g_0, g_1, g_2 satisfying these relations.

To place the problem in a more general context, in this paper, we establish rigidity properties of certain solvable group actions on the torus $\mathbb{T}^N = \mathbb{R}^N/\mathbb{Z}^N$, for $N > 1$.

The solvable groups Γ considered here are the finitely presented, torsion-free, *Abelian-by-cyclic* (ABC) groups, which admit a short exact sequence

$$0 \hookrightarrow \mathbb{Z}^d \to \Gamma \to \mathbb{Z} \to 0.$$

All such groups are of the form $\Gamma = \Gamma_B$, where $B = (b_{ij})$ is an integer valued, $d \times d$ matrix with $\det(B) \neq 0$, and

(1.1)

$$\Gamma_B = \mathbb{Z} \ltimes \mathbb{Z}^d = \left\langle g_0, g_1, \ldots, g_d \mid g_0 g_i = \left(\prod_{j=1}^{d} g_j^{b_{ji}} \right) g_0, \quad [g_i, g_j] = 1, \quad i, j = 1, 2, \ldots, d \right\rangle.$$

ABC groups have been studied intensively in geometric group theory, as it is an open problem to classify finitely generated solvable groups up to quasi-isometry. The classification problem for ABC groups has been solved in [FM1] (the non-polycyclic case, $|\det B| > 1$) and [EFW1, EFW2] (the polycyclic case, $|\det B| = 1$), where the authors also revealed close connections between the geometry of these groups and dynamics [FM2, EF]. Here we consider actions of polycyclic ABC groups.

Recall that a C^r action α of a finitely generated group Γ with generators $g_1, \ldots, g_k$ on a closed manifold M is a homomorphism $\alpha : \Gamma \to \mathrm{Diff}^r(M)$. The action is determined completely by $\alpha(g_1), \ldots, \alpha(g_k)$. The polycylic ABC groups admit natural affine actions on tori, as follows.

Up to rearranging the standard basis for $\mathbb{R}^d$, every matrix $B \in \mathrm{SL}(d; \mathbb{Z})$ can be written in the form

$$B = \begin{pmatrix} \bar{A} & 0 \\ 0 & I_{d-N} \end{pmatrix},$$

for some $N \leq d$, where $\bar{A} \in \mathrm{SL}(N; \mathbb{Z})$, and I_{d-N} is the $(d - N) \times (d - N)$ identity matrix, chosen to be maximal. To avoid obviously degenerate actions, we restrict our attention to the cases where $d = KN$, for some $K \geq 1$. Then

$$\begin{aligned} \Gamma_B = \Gamma_{\bar{A},K} := \langle g_0, g_{i,k},\ i = 1, \ldots, N,\ k = 1, \ldots, K \mid \ [g_{i,k}, g_{j,\ell}] = 1, \\ g_0 g_{i,k} = \left(\prod_{j=1}^{N} g_{j,k}^{a_{ji}} \right) g_0, \quad i, j = 1, \ldots, N,\ k, \ell = 1, \ldots, K \right\rangle. \end{aligned} \tag{1.2}$$

Note that $\Gamma_{\bar{A}} = \Gamma_{\bar{A},1}$.

In the affine actions of $\Gamma_{\bar{A},K}$ we consider, the element g_0 acts on $\mathbb{T}^N$ by the automorphism $x \mapsto \bar{A}x \pmod{\mathbb{Z}^N}$ induced by $\bar{A}$, and the elements $g_{i,k}, i = 1, \ldots, N,\ k = 1, \ldots, K$ act as translations $x \mapsto x + \rho_{i,j} \pmod{\mathbb{Z}^N}$, where $\rho_{i,j} \in \mathbb{R}^N$. The group relations in $\Gamma_{\bar{A},K}$ restrict the possible values of $\rho_{i,j}$; we describe precisely these restrictions in the next subsection. We will see that for a typical $\bar{A}$, the affine actions define a finite dimensional space of distinct (i.e. nonconjugate) actions on the torus.

The main rigidity results of this paper can be grouped into two classes: local and global. We obtain local rigidity results for the actions of $\Gamma_{\bar{A}} = \Gamma_{\bar{A},1}$, which implies local rigidity for $\Gamma_{\bar{A}} = \Gamma_{\bar{A},K}$, $K \geq 1$. To each action α on $\mathbb{T}^N$ sufficiently C^r close to an affine action, we define an $N \times N$ *rotation matrix* $\boldsymbol{\rho}(\alpha)$. Under suitable hypotheses on $\bar{A}$, if the columns of this matrix satisfy a simultaneous Diophantine condition, then α is smoothly conjugate to the affine action with rotation matrix $\boldsymbol{\rho}(\alpha)$. The fact that the action α is a smooth perturbation of an affine action is crucial.

More generally, for each affine action α_K of $\Gamma_{\bar{A},K}$ there are K rotation matrices $\boldsymbol{\rho}_1(\alpha_K), \dots, \boldsymbol{\rho}_K(\alpha_K)$. In the section on global rigidity, we consider actions α of $\Gamma_{\bar{A},K}$ for which $\bar{A}$ acts as an Anosov diffeomorphism, but the rotation matrices $\boldsymbol{\rho}_i(\alpha)$ are not *a priori* well-defined. Under relatively weak additional assumptions on the action, we obtain that the collection of $\boldsymbol{\rho}_i(\alpha_K)$ can be defined and and forms a complete invariant of the action, up to *topological* conjugacy. We then establish conditions under which this topological conjugacy is smooth. In particuar, if K is sufficiently large (depending on the spectrum of $\bar{A}$ and the Anosov element $\alpha(g_0)$), then for almost every set of rotation matrices $\boldsymbol{\rho}_1(\alpha), \dots, \boldsymbol{\rho}_K(\alpha)$, the conjugacy is smooth.

Before stating these results, we describe precisely the space of affine actions of $\Gamma_{\bar{A},K}$ we consider.

1.1. The affine actions of $\Gamma_{\bar{A},K}$.

The following proposition can be verified directly using the group relation 3.1.

Proposition 1.1. *Let $\bar{A} \in \mathrm{SL}(N, \mathbb{Z})$, and suppose that $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \dots, \boldsymbol{\rho}_K$ are real-valued, $N \times N$ matrices such that each $\boldsymbol{\rho} = \boldsymbol{\rho}_i$ satisfies:*

$$\bar{A}\boldsymbol{\rho} = \boldsymbol{\rho}\bar{A} \mod \mathbb{Z}^{N\times N}. \tag{1.3}$$

Denote by $\rho_{i,j}$ the j-th column of $\boldsymbol{\rho}_i$. Then the affine maps

$$\bar{\alpha}(g_0)(x) := \bar{A}x \ (\mathrm{mod}\, \mathbb{Z}^N), \quad \text{and} \quad \bar{\alpha}(g_{i,j})(x) := x + \rho_{i,j} \ (\mathrm{mod}\, \mathbb{Z}^N)$$

define an action $\bar{\alpha} = \bar{\alpha}_K(\bar{A}, \boldsymbol{\rho}) : \ \Gamma_{\bar{A},K} \to \mathrm{SL}(N, \mathbb{Z}) \ltimes \mathbb{R}^N$ on $\mathbb{T}^N$.

Conversely, if $\alpha : \ \Gamma_{\bar{A},K} \to \mathrm{SL}(N, \mathbb{Z}) \ltimes \mathbb{R}^N$ is an action on $\mathbb{T}^N$ with

$$\alpha(g_0)(x) = \bar{A}x \ (\mathrm{mod}\, \mathbb{Z}^N), \quad \text{and} \quad \alpha(g_i)(x) = x + \beta_{i,j} \ (\mathrm{mod}\, \mathbb{Z}^N),$$

for some vectors $\beta_{i,j} \in \mathbb{R}^N$, then for each $i = 1, \dots, K$, the matrix $\boldsymbol{\rho}_i$ whose columns are formed by the $\beta_{i,j}$ satisfies (1.3).

We further focus on the case $K = 1$. For $\bar{A} \in \mathrm{SL}(N, \mathbb{Z})$ and $\boldsymbol{\rho} \in M_N(\mathbb{T})$, where $M_N(\mathbb{T})$ denotes $N \times N$ matrices with entries in $\mathbb{T}$, we denote by $\bar{\alpha}(\bar{A}, \boldsymbol{\rho}) = \bar{\alpha}_1(\bar{A}, \boldsymbol{\rho})$ the action $\bar{\alpha}$ on $\mathbb{T}^N$ defined in Proposition 1.1. Let

$$\mathrm{Aff}(\Gamma_{\bar{A}}) := \{\bar{\alpha}(\bar{A}, \boldsymbol{\rho}) : \boldsymbol{\rho} \text{ satisfies } (1.3)\}.$$

Proposition 1.2. *The action $\alpha(\bar{A}, \boldsymbol{\rho}) \in \mathrm{Aff}(\Gamma_{\bar{A}})$ is faithful if and only if $\bar{A}$ is not of finite order, and the column vectors $\rho_1, \ldots, \rho_N$ of $\boldsymbol{\rho}$ are linearly independent over $\mathbb{Z}$; that is, if there exists $(p_1, \ldots, p_N) \in \mathbb{Z}^N$ with $\sum_{i=1}^N p_i \rho_i = 0 \mod \mathbb{Z}^N$, then $p_1 = \cdots = p_N = 0$.*

For $\bar{A} \in \mathrm{SL}(N, \mathbb{Z})$ not of finite order, we thus define the set of faithful affine actions by

$$\mathrm{Aff}_\star(\Gamma_{\bar{A}}) := \{\bar{\alpha}(\bar{A}, \boldsymbol{\rho}) \in \mathrm{Aff}(\Gamma_{\bar{A}}) : \bar{\alpha}(\bar{A}, \boldsymbol{\rho}) \text{ is faithful}\}.$$

Note that $\mathrm{Aff}_\star(\Gamma_{\bar{A}})$ is dense in $\mathrm{Aff}(\Gamma_{\bar{A}})$, as is its complement $\mathrm{Aff}(\Gamma_{\bar{A}}) \backslash \mathrm{Aff}_\star(\Gamma_{\bar{A}})$, consisting of the non-faithful actions. Since faithful actions can be approximated arbitrarily well by ones that are not faithful, there is no form of local rigidity in $\mathrm{Aff}(\Gamma_{\bar{A}})$; what is more, local rigidity even fails in $\mathrm{Aff}_\star(\Gamma_{\bar{A}})$.

Proposition 1.3. *Given $\bar{A} \in \mathrm{SL}(N, \mathbb{Z})$, two actions $\bar{\alpha}_1, \bar{\alpha}_2 \in \mathrm{Aff}_\star(\Gamma_{\bar{A}})$ are conjugate by a homeomorphism homotopic to identity if and only if $\bar{\alpha}_1 = \bar{\alpha}_2$.*

Thus actions in $\mathrm{Aff}_\star(\Gamma_{\bar{A}})$ are not locally rigid even among affine actions. Returning to our original example, let

$$\bar{A} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \tag{1.4}$$

and

$$\Gamma_{\bar{A}} = \langle g_0, g_1, g_2 \mid g_0 g_1 g_0^{-1} = g_1^2 g_2, \quad g_0 g_2 g_0^{-1} = g_1 g_2, \quad [g_1, g_2] = 1 \rangle.$$

Fix $a_0, a_1 \in \mathbb{T}^1$, and let $\boldsymbol{\rho}(a_0, a_1) := \begin{pmatrix} a_0 + 2a_1 & a_1 \\ a_1 & a_0 + a_1 \end{pmatrix}$. Then $\{\bar{\alpha}(\bar{A}, \boldsymbol{\rho}(a_0, a_1)) : a_0, a_1 \in \mathbb{T}^1\}$ defines a 2-parameter family of non-conjugate actions on $\mathbb{T}^2$; in the next subsection we explain that these are all such affine actions.

Thus the matrix $\boldsymbol{\rho}$ is a complete invariant of the faithful affine representations $\bar{\alpha}(\bar{A}, \boldsymbol{\rho})$ of $\Gamma_{\bar{A}}$. The columns of $\boldsymbol{\rho}$ are rotation vectors of the corresponding translations. We will show that these rotation vectors, and hence the invariant $\boldsymbol{\rho}$, extend continuously to a neighborhood of the affine representations in such a way that $\boldsymbol{\rho}$ gives a complete invariant under smooth conjugacy, under the hypotheses that the columns of $\boldsymbol{\rho}$ satisfy a simultaneous Diophantine condition.

Further properties of the affine representations are discussed in Section C, which also contains the proofs of the results in this section.

1.2. **Local rigidity of $\Gamma_{\bar{A},K}$ actions.** An action $\alpha : \ \Gamma \to \mathrm{Diff}^r(M)$ is called $C^{r,k,\ell}$ *locally rigid* if any sufficiently C^k small C^r perturbation $\tilde{\alpha}$ is C^ℓ conjugate to α, i.e., there exists a diffeomorphism h of M, C^ℓ close to identity, that conjugates $\tilde{\alpha}$ to α: $h \circ \alpha(g) = \alpha(g) \circ h$ for all $g \in \Gamma$. The paper of Fisher [Fi] contains background and an excellent overview of the local rigidity problem for general group actions.

Local rigidity results for solvable group actions are relatively rare. In [DK], Damjanovic and Katok proved $C^{\infty,1,\infty}$ local rigidity for $\mathbb{Z}^k$ ($k \geq 2$) (Abelian) higher rank partially hyperbolic actions by toral automorphisms, by introducing a new KAM iterative scheme. In [HSW] and [W], the authors proved local rigidity for higher rank ergodic nilpotent actions by toral automorphisms on $\mathbb{T}^N$, for any even $N \geq 6$. Burslem and Wilkinson in [BW] studied the solvable Baumslag-Solitar groups

$$BS(1,n) = \langle a, b \mid aba^{-1} = b^n; n \geq 2 \rangle),$$

acting on $\mathbb{T}^1$ and obtained a classification of such actions and a global rigidity result in the analytic setting. Asaoka in [A1, A2] studied the local rigidity of the action on $\mathbb{T}^N$ or $\mathbb{S}^N$ of non-polycyclic Abelian-by-cyclic groups, where the cyclic factor is uniformly expanding.

Unless assumptions are made on the action (or the manifold), solvable group actions are typically not locally rigid but can enjoy a form of partial local rigidity: that is, local rigidity subject to constraints that certain invariants be preserved. The simplest example occurs in dimension 1, where the rotation number of a single C^2 circle diffeomorphism supplies a complete topological invariant, provided that it is irrational, and a complete smooth invariant, provided it satisfies a Diophantine condition. This result extends to actions of higher rank abelian groups on $\mathbb{T}^1$, under a simultaneous Diophantine assumption on the rotation numbers of the generators of the action. In fact, these results are not just local in nature but apply to all diffeomorphisms of the circle.

For higher dimensional tori, even local rigidity results of this type are scarce, one problem being the lack of invariants analogous to the rotation number. One result in this direction is by Damjanovic and Fayad [DF], who proved local rigidity of ergodic affine $\mathbb{Z}^k$ actions on the torus that have a rank-one factor in their linear part, under certain Diophantine conditions.

As described in the previous section, the affine actions of $\Gamma_{\bar{A}}$ we consider here are not locally rigid.

Definition 1.1. *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*

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

We denote $(v_1, \dots, v_m) \in \mathrm{SDC}(C, \tau)$.

For example, the matrix $\rho \mathrm{Id}_N$ is simultaneously Diophantine if ρ is a Diophantine number. It is known that simultaneous Diophantine vectors in $\mathrm{SDC}(C, \tau)$ form a full measure set in $\mathbb{T}^{N \times m}$, if we take the union with respect to all $C > 0$ for fixed $\tau > N-1$.

Definition 1.2. *Given a homeomorphism $f : \mathbb{T}^N \to \mathbb{T}^N$ preserving a probability measure μ, the vector*

$$\rho_\mu(f) := \int_{\mathbb{T}^N} (\tilde{f}(x) - x)\, d\mu, \ \bmod \mathbb{Z}^N, \tag{1.5}$$

where $\tilde{f} :\ \mathbb{R}^N \to \mathbb{R}^N$ is any lift of f, is independent of the choice of the lift $\tilde{f}$. We call $\rho_\mu(f)$ the rotation vector *of f with respect to μ.*

The main local rigidity result is:

Theorem 1.4. *For any $\bar{A} \in \mathrm{SL}(N, \mathbb{Z})$, any $C, \tau > 0$, there exist $\varepsilon > 0$ and $\ell \in \mathbb{N}$, such that for any $\boldsymbol{\rho} \in \mathrm{SDC}(C, \tau)$ satisfying (1.3) the following holds. Let $\alpha : \Gamma_{\bar{A}} \to \mathrm{Diff}^\infty(\mathbb{T}^N)$ be any representation satisfying*

1. *$\alpha(g_0)$ is homotopic to $\bar{\alpha}(\bar{A}, \boldsymbol{\rho})(g_0)$;*
2. *$\max_{1\le i\le N} \|\alpha(g_i) - \bar{\alpha}(\bar{A}, \boldsymbol{\rho})(g_i)\|_{C^\ell} < \varepsilon$;*
3. *there exist $\alpha(g_i)$-invariant probability measures μ_i, $i = 1, \dots, N$, such that the matrix formed by the rotation vectors $(\rho_{\mu_1}(\alpha(g_1)), \dots, \rho_{\mu_N}(\alpha(g_N)))$ is equal to $\boldsymbol{\rho}$.*

Then there exists a C^∞ diffeomorphism h that is C^1 close to identity such that $h \circ \alpha = \bar{\alpha} \circ h$. Moreover, we have that $\mu_i = h_* m$ for all $i = 1, 2, \dots, N$, where m is Haar on $\mathbb{T}^N$, and $h_* m$ is uniquely ergodic under the action α.

We will prove in Appendix B that the simultaneously Diophantine condition is actually satisfied by a big class of matrices $\boldsymbol{\rho}$ and $\bar{A}$ satisfying (1.3). One special case is when $\bar{A} \in \mathrm{SL}(N, \mathbb{Z})$ has simple spectrum, in which case any matrix $\boldsymbol{\rho} \in M_N(\mathbb{R})$ commuting with $\bar{A}$ has the form $\boldsymbol{\rho} = \sum_{i=1}^N a_i \bar{A}^{i-1}$, $a_i \in \mathbb{R}$, $i = 1, \dots, N$. The columns of the matrix $\boldsymbol{\rho}$ is simultaneously Diophantine if the nonvanishing a_i's form a Diophantine vector.

Remark 1.1. *We remark that the faithfulness (guaranteed by the Diophantine condition) of the action is necessary for smooth conjugacy. For instance, consider $\rho = 1/2$ in (1.4) and $\alpha(g_i) = \bar{\alpha}(g_i)$, $i = 1, 2$, and for any $\varepsilon > 0$,*

$$\alpha(g_0) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \varepsilon \begin{bmatrix} \sin(4\pi x) \\ \sin(4\pi x) \end{bmatrix}.$$

One can verify that this gives rise to a $\Gamma_{\bar{A}}$ action. We will see in Theorem 1.5 that for sufficiently small ε, there exists a bi-Hölder conjugacy h satisfying $h \circ \alpha = \bar{\alpha} \circ h$. However, the conjugacy h is not C^1. Indeed, 0 is a fixed point for both $\alpha(g_0)$ and $\bar{A}$. The derivative $D(\alpha(g_0))(0) = \bar{A} + 4\pi\varepsilon \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ has determinant 1 but different trace than $\bar{A}$, so it is not conjugate to $\bar{A}$.

It is natural to ask the following question:

Question: *Suppose the action is faithful and close to an algebraic action, is it always possible to smoothly conjugate the action to an algebraic one?*

1.3. Global rigidity.

1.3.1. *Topological conjugacy for $N = 2$.* The proof of the above local rigidity theorem is an application of the KAM techniques for $\mathbb{Z}^N$ actions initiated by Moser [M] in the context of $\mathbb{Z}^N$ action by circle diffeomorphisms. The KAM technique is essentially perturbative. It is natural to ask if our solvable group action has certain rigidity in the nonperturbative sense, i.e. whether it is globally rigid. A class of actions of a group, not necessarily close to a algebraic actions, is called *globally rigid* if any action from this class is conjugate to an algebraic one. There is a nonperturbative global rigidity theory for circle maps known as *Herman-Yoccoz theory.* For Abelian group actions by circle diffeomorphisms, the global version of Moser's theorem was proved by Fayad and Khanin [FK]. These global rigidity results rely on the Denjoy theorem stating that a C^2 circle diffeomorphism with irrational rotation number is topologically conjugate to the irrational rotation by the rotation number.

In the higher dimensional case, there is no corresponding Herman-Yoccoz theory for diffeomorphisms of $\mathbb{T}^N$ isotopic to rotations. The reason is that rotation vectors are not well-defined in general. Even when rotation vectors are uniquely defined, they are not complete invariants for conjugacy analogous to rotation numbers for circle maps. In particular, the obvious analogue of the topological conjugacy given by the Denjoy theorem is missing for diffeomorphisms of $\mathbb{T}^N,\ N > 1$.

On the other hand, by a theorem of Franks (Theorem 3.1 below), Anosov diffeomorphisms on $\mathbb{T}^N$ are topologically conjugate to a toral automorphisms. A $C^r,\ r \geq 1$ diffeomorphism $f:\ M \to M$ is called *Anosov* if for each $x \in M$, there exists a splitting of the tangent space $T_xM = E^s(x) \oplus E^u(x)$ and constant C and $0 < \lambda < 1$ such that for every $x \in M$, we have

- $DfE^s(x) = E^s(f(x))$ and $DfE^u(x) = E^u(f(x))$,
- $\|Df^n v\| \leq C\lambda^n \|v\|$ for $v \in E^s(x) \setminus \{0\}$ and $n \geq 0$, and $\|Df^n v\| \leq C\lambda^{-n} \|v\|$ for $v \in E^u(x) \setminus \{0\}$ and $n \leq 0$.

As the starting point of a global rigidity result of our $\Gamma_{\bar{A}}$ action, we assume $\alpha(g_0)$ acts by an Anosov diffeomorphism homotopic to $\bar{A}$. With the topological conjugacy at hand, the next question is to show the topological conjugacy given by Franks's theorem also linearizes the Abelian subgroup action. The new problem that arises is that for toral diffeomorphisms homotopic to identity the rotation vector is in general not well-defined, and it only makes sense to talk about the rotation set. When there is more than one vector in the rotation set, the diffeomorphism cannot be conjugate to a translation.

In the circle map case, the existence and uniqueness of rotation number relies crucially on the fact that the graph in $\mathbb{R}^2$ of every lifted orbit stays within distance 1 of a straight line, and the rotation number is simply the slope of the line. This fact is also important in the study of Euler class and bounded cohomology for groups acting on circle [Gh]. We impose a similar, but much weaker, analogous condition that we call sublinear oscillation (see Definition 1.3) on the Abelian subgroup action to guarantee the existence of a common conjugacy.

Theorem 1.5. *Suppose $N = 2$. Let $\alpha : \Gamma_{\bar{A}} \to \mathrm{Diff}^r(\mathbb{T}^2)$, $r > 1$, be a representation satisfying*

1. *$\alpha(g_0)$ is Anosov and homotopic to $\bar{A}$;*
2. *the sub-action generated by $\alpha(g_1), \ldots, \alpha(g_N)$ has sub-linear oscillation (see Definition 1.3 below).*

Then there exist $\boldsymbol{\rho}$ satisfying (1.3) and a unique bi-Hölder homeomorphism $h : \mathbb{T}^2 \to \mathbb{T}^2$ homotopic to the identity satisfying

$$h \circ \alpha = \bar{\alpha}(\bar{A}, \boldsymbol{\rho}) \circ h.$$

Remark 1.2. *In this theorem faithfulness of the action is not necessary, since we do not need the rotation vectors of $\bar{\alpha}(g_i)$ to be irrational.*

To motivate the definition of sublinear oscillation, let us first recall the following definition. We say a diffeomorphism $f : \mathbb{T}^N \to \mathbb{T}^N$ is of *bounded deviation* if there exists $\rho \in \mathbb{T}^N$ and a constant $C > 0$, such that

$$\|\tilde{f}^n(x) - x - n\rho\|_{C^0} \le C, \quad \forall\, n \in \mathbb{Z}.$$

Being of bounded deviation implies that each orbit of $\tilde{f}$ stays within bounded distance of the line $\mathbb{R}\rho$. The concept of bounded deviation was first introduced by Morse, who called it of class A, in the case of geodesic flows on surfaces of genus greater than 1 [Mo]. It was later shown by Hedlund that globally minimizing geodesics for arbitrary smooth metric on $\mathbb{T}^2$ are also of bounded deviation [He]. A generalization to Gromov hyperbolic spaces can be found in [BBI]. In the one-dimensional case, all circle maps are of bounded deviation, from which follows immediately the existence of the rotation number. Being of bounded deviation does not guarantee the existence of a conjugacy of the map f to a rigid translation. In the one dimensional case, a circle map with irrational rotation number is only known to be semi-conjugate to a rotation. Denjoy's counter-example shows that the semi-conjugacy cannot be improved to a conjugacy without further assumptions. In the two dimensional case, it is known [Ja] that for a conservative pseudo-rotation of bounded oscillation, the rotation vector being totally irrational is equivalent to the existence of a semi-conjugacy to the rigid translation.

Examples of diffeomorphisms on $\mathbb{T}^2$ of bounded deviation can be found in [MS], which are higher dimensional generalizations of Denjoy's examples on $\mathbb{T}^1$.

Let $T \in \mathrm{Diff}_0(\mathbb{T}^N)$ and let $\tilde{T}:\ \mathbb{R}^N \to \mathbb{R}^N$ be a lift of T. Denote by π_i the projection to the i-th component of a vector in $\mathbb{R}^N$. Define the *oscillation* $\mathrm{Osc}(\tilde{T})$ of $\tilde{T}$ by:

$$\mathrm{Osc}(\tilde{T}) := \max_{x,i}\{\pi_i(\tilde{T}(x) - x)\} - \min_{x,i}\{\pi_i(\tilde{T}(x) - x)\}.$$

It is easy to see that Osc is independent of the choice of the lift. We define $\mathrm{Osc}(T) = \mathrm{Osc}(\tilde{T})$.

Definition 1.3. (1) *For given $c \in [0, 1)$, we say that the Abelian group action $\beta:\ \mathbb{Z}^N \to \mathrm{Diff}^r_0(\mathbb{T}^N)$ is* of c-slow oscillation *if*

$$\limsup_{\|\mathbf{p}\|\to\infty} \frac{\mathrm{Osc}(\beta(\mathbf{p}))}{\|\mathbf{p}\|^c} < \infty.$$

(2) *We say the action β has* sublinear oscillation *if it is of c-slow oscillation for some $c < 1$.*

It is easy to see that bounded deviation implies c-slow oscillation with $c = 0$. Sublinear oscillation occurs in first passage percolation (see Section 4.2 of [ADH]) where paths minimizing a cost defined for random walks on $\mathbb{Z}^2$ have c-slow oscillation with a power law $c \le 3/4$ and conjecturally $c = 2/3$.

1.3.2. *Smooth conjugacy for $N = 2$.* The conjugacy h in Theorem 1.5 is only known to be Hölder. It is natural to ask if we could improve the regularity. In hyperbolic dynamics, there is a periodic data rigidity theory for Anosov diffeomorphisms, which implies that in the two-dimensional case that if the regularity of h is known to be C^1, then h is in fact as smooth as the Anosov diffeomorphism $\alpha(g_0)$ (see Theorem 5.7 below and Section 7.2).

So the problem is now to find sufficient conditions for our action to ensure that the conjugacy h is C^1. The invariant foliation structure given by the Anosov diffeomorphism enables us to generalize the Herman-Yoccoz theory for circle maps to the higher dimensional setting. Employing an idea of Kra [K], we need to increase the rank of the Abelian component in $\Gamma_{\bar{A}}$.

To obtain higher regularity of the conjugacy, we consider a slightly different class of ABC groups. For $K \ge 1$, let

$$\begin{aligned}\Gamma_{\bar{A},K} = \Big\langle g_0, g_{i,k},\ i = 1, \dots, N,\ k = 1, \dots, K \ |\ &[g_{i,k}, g_{j,\ell}] = 1,\\ g_0 g_{i,k} = \left(\prod_{j=1}^N g_{j,k}^{a_{ji}}\right) g_0,\quad i, j = 1, \dots, N,\ &k, \ell = 1, \dots, K \Big\rangle.\end{aligned} \tag{1.6}$$

Note that $\Gamma_{\bar{A},K} = \mathbb{Z} \ltimes (\mathbb{Z}^N)^K$ is ABC and $\Gamma_{\bar{A},1} = \Gamma_{\bar{A}}$.

Theorem 1.6. *Given an Anosov diffeomorphism $A: \mathbb{T}^2 \to \mathbb{T}^2$ homotopic to $\bar{A} \in \mathrm{SL}(2,\mathbb{Z})$, there is a C^1 open set $\mathcal{O}$ of Anosov diffeomorphisms containing A, and a number K_0 such that for any integer $K \geq K_0$, there exists a full measure set $\mathcal{R}_{2,K} \subset (\mathbb{T}^2)^K$ such that the following holds.*

Let $\alpha: \Gamma_{\bar{A},K} \to \mathrm{Diff}^r(\mathbb{T}^2)$ be a representation satisfying:

1. *$\alpha(g_0) \in \mathcal{O}$,*
2. *the sub-action generated by $\alpha(g_{1,1}), \alpha(g_{2,1})$ has sublinear oscillation,*
3. *for some $i: \{1,\dots,K\} \to \{1,2\}$, the rotation vectors $(\rho_{i(1),1},\dots,\rho_{i(K),K})$ lie in $\mathcal{R}_{2,K}$, where $\rho_{j,k}$ is the rotation vector of $\alpha(g_{j,k})$ with respect to an invariant probability measure $\mu_{j,k}$, $j=1,2$ and $k=1,\dots,K$.*

Then there is a unique $C^{r-\varepsilon}$ conjugacy h conjugating the action α to an affine action for ε arbitrarily small.

1.3.3. *The higher dimensional case.* We next turn to the higher dimensional case $N > 2$.

First we have the following result on the topological conjugacy parallel to Theorem 1.5.

Theorem 1.7. *Suppose $N > 2$. Given a hyperbolic $\bar{A} \in \mathrm{SL}(N,\mathbb{Z})$, there exists $0 \leq c < 1$ such that the following holds. Let $\alpha: \Gamma_{\bar{A}} \to \mathrm{Diff}^r(\mathbb{T}^N)$, $r > 1$, be a representation satisfying*

1. *$\alpha(g_0)$ is Anosov and homotopic to $\bar{A}$,*
2. *the sub-action generated by $\alpha(g_1),\dots,\alpha(g_N)$ has c-slow oscillation.*

Then there exist $\boldsymbol{\rho}$ satisfying (1.3) and a unique bi-Hölder homeomorphism $h: \mathbb{T}^N \to \mathbb{T}^N$ homotopic to the identity with

$$h \circ \alpha = \bar{\alpha}(\bar{A}, \boldsymbol{\rho}) \circ h.$$

The constant c in Theorem 1.7 can be made explicit as follows. Suppose $\bar{A}$ has eigenvalues $\lambda_1^u,\dots,\lambda_k^u$ and $\lambda_1^s,\dots,\lambda_\ell^s$, $k \geq 1$, $\ell \geq 1$, $k+\ell = N$, (complex eigenvalues and repeated eigenvalues are allowed), ordered as follows

$$|\lambda_\ell^s| \leq \ldots \leq |\lambda_1^s| < 1 < |\lambda_1^u| \leq \ldots \leq |\lambda_k^u|. \tag{1.7}$$

Then c can be chosen to be any number satisfying

$$0 \leq c < \min\left\{\frac{\ln|\lambda_1^u|}{\ln|\lambda_k^u|}, \frac{\ln|\lambda_1^s|}{\ln|\lambda_\ell^s|}\right\}. \tag{1.8}$$

We next study how to improve the regularity of the conjugacy h.

Theorem 1.8. *Given a hyperbolic $\bar{A} \in \mathrm{SL}(N,\mathbb{Z})$, $N > 2$, with simple real spectrum, there exist a C^1 neighborhood $\mathcal{O}$ of $\bar{A}$, a number $0 \le c < 1$ and a number K_0, such that for any integer $K > K_0$, there exists a full measure set $\mathcal{R}_{N,K} \subset (\mathbb{T}^N)^K$ such that the following holds.*

Let $\alpha: \Gamma_{\bar{A},K} \to \mathrm{Diff}^r(\mathbb{T}^N)$ be a representation satisfying

1. *$\alpha(g_0) \in \mathcal{O}$;*
2. *the sub-action generated by $\alpha(g_{1,1}), \dots, \alpha(g_{N,1})$ has c-slow oscillation;*
3. *for some $i: \{1,\dots,K\} \to \{1,\dots,N\}$, the rotation vectors $(\rho_{i(1),1}, \dots, \rho_{i(K),K})$ lie in $\mathcal{R}_{N,K}$, where $\rho_{j,k}$ is the rotation vector of $\alpha(g_{j,k})$ with respect to an invariant probability measure $\mu_{j,k}$, $j = 1,\dots,N$ and $k = 1,\dots,K$.*

Then there is a unique $C^{1,\nu}$ conjugacy h conjugating α to an affine action for some $\nu > 0$.

In dimension 3, the regularity of the conjugacy can be improved applying the work of Gogolev in [G2] (Theorem 5.8 below).

Corollary 1.9. *Under the same assumptions as Theorem 1.8, suppose in addition that $N = 3$ and $r > 3$. Then the conjugacy $h \in C^{r-3-\varepsilon}$, for arbitrarily small ε. Moreover, there exists a $\kappa \in \mathbb{Z}$ such that if $r \notin (\kappa, \kappa+3)$, then $h \in C^{r-\varepsilon}$.*

Further relaxation of the assumptions of Theorem 1.8 and Corollary 1.9 is possible, which will be discussed in Section 6.3. In particular, in many cases the condition on the C^1 closeness of A to $\bar{A}$ can be relaxed.

For $N > 2$ case, the elliptic dynamics techniques from the two-dimensional case carry over completely. However, there are two obstructions coming from the hyperbolic dynamics. On the one hand when restricted to the invariant foliation of $\bar{A}$, it turns out that a conjugacy between two Anosov diffeomorphisms sends (un)stable leaves to (un)stable leaves. On the other hand the affine foliations parallel to the eigenspaces of $\bar{A}$ might not be sent to A invariant foliations with smooth leaves. Adding to the difficulty is the fact that the regularity of the weakest stable and unstable distributions are bad (only Hölder in general). These obstructions prevent people from developing a theory of periodic data rigidity as strong as the two-dimensional setting. The general result [G1, GKS] for the $N > 2$ case of periodic data rigidity of Anosov diffeomorphisms is to get that the conjugacy h is in C^{1+} (i.e. Dh and Dh^{-1} are Hölder) if A is C^1 close to $\bar{A}$ and they have the same periodic data (see Section 7.2).

The paper is organized as follows. We prove the local rigidity Theorem 1.4 in Section 2. All the remaining sections are devoted to the proof of the global rigidity. In Section 3, we prove that there is a common conjugacy (Theorem 1.5 and 1.7). In Section 4, we prepare techniques from elliptic dynamics and hyperbolic dynamics. In Section 5, we state and prove the main propositions needed for the proof of Theorem 1.6 and 1.8. In Section 6, we prove the main Theorems 1.6 and 1.8. In Section 7, we discuss our result from the viewpoints of elliptic dynamics, hyperbolic dynamics and amenable group actions. Finally, in the Appendix A, we give the proof of the number theoretic result Theorem 5.3. In Section C, we prove the results about affine action stated in Section 1.1.

2. Local rigidity: proofs

In this section, we prove Theorem 1.4. We sketch the idea of the proof. Given representation $\alpha:\ \Gamma_{\bar{A}} \to \mathrm{Diff}^\infty(\mathbb{T}^N)$ with $(\rho_{\mu_1}(\alpha(g_1)), \dots, \rho_{\mu_N}(\alpha(g_N))) = \boldsymbol{\rho} \in \mathrm{SDC}(C, \tau)$, where μ_i is a invariant probability measure of $\alpha(g_i)$, we can proceed as in [M] using the KAM method to show that the Abelian subgroup action can be smoothly conjugated to rigid translations. Using the group relation, we can further show that this conjugacy also conjugates the diffeomorphism $\alpha(g_0)$ to a linear one.

We have the following lemma which is proved by the standard KAM iteration procedure.

Lemma 2.1 (KAM for Abelian group actions)**.** *Given $C, \tau > 0$, there exist constants ℓ and ε_0 such that the following holds.*

Let $T_1, \dots, T_m \in \mathrm{Diff}_0^\infty(\mathbb{T}^N)$ be commuting diffeomorphisms with $m > 1$. Suppose the rotation vectors $\rho_{\mu_i}(T_i)$, $i = 1, \dots, m$, satisfy the simultaneous Diophantine condition with constants C, τ, for some T_i-invariant probability measure μ_i, and

$$\max_{1 \le i \le m} \|T_i - \mathrm{id} - \rho_{\mu_i}(T_i)\|_{C^\ell} < \varepsilon_0.$$

Then there exists a C^∞ diffeomorphism h that is C^1 close to the identity such that

$$h \circ T_i(x) = h(x) + \rho_{\mu_i}(T_i), \quad x \in \mathbb{T}^N, \quad i = 1, \dots, m.$$

Moreover, the invariant measures coincide: $\mu_i = h_* m$ where m is the Haar measure on $\mathbb{T}^N$.

The proof of this lemma is essentially the same as Moser [M]. A proof was sketched by F. Rodriguez-Hertz in the case of $m = 2, N = 2$ (see Theorem 6.5 of [R1]). There is no difficulty to adapt the proof to the case $N > 2, m \geq 2$, so we skip it.

Now we are ready to prove the local rigidity.

Proof of Theorem 1.4, local rigidity. Denote by $T_i = \alpha(g_i)$ and $\bar{T}_i = \bar{\alpha}(\bar{A}, \boldsymbol{\rho})(g_i) : x \mapsto x + \rho_{\mu_i}(T_i)$, $i = 1, \dots, N$. We next apply Lemma 2.1 to show that such h exists that simultaneously conjugates $h \circ T_i = \bar{T}_i \circ h$, $i = 1, \dots, N$, using the commutativity of T_j and the simultaneous Diophantine condition. We then compose this h^{-1} to the right and h to the left on both the LHS and the RHS of the group relation $AT_i = (\prod_{j=1}^N T_j^{a_{ji}})A$ to get

$$hAh^{-1}\bar{T}_i = (\prod_{j=1}^N \bar{T}_j^{a_{ji}})hAh^{-1},$$

in other words

$$hAh^{-1}(x+\rho_j) = hAh^{-1}(x) + \sum_{j=1}^N a_{ji}\rho_j, \text{ mod } \mathbb{Z}^N, \quad i = 1, \dots, N.$$

We introduce the function $B(x) = hAh^{-1}(x) - \bar{A}x$, then using (1.3) the above equation gives $B(x+\rho_i) = B(x), \quad i = 1, \dots, N$. Due to the ergodicity of the affine Abelian subgroup action following from the Diophantine property of the vectors $\rho_1, \dots, \rho_N$, we get that $B(x) = hAh^{-1}(x) - \bar{A}x = B$ is constant. To kill this constant, we introduce a translation $t(x) = x + (\mathrm{id} - \bar{A})^{-1}B$. It can be easily verified that t conjugates $\bar{A}x + B$ and $\bar{A}x$, i.e. $\bar{A}t(x) + B = t(\bar{A}x)$. Composing the above h with t, we get the conjugacy in the statement of the theorem. □

3. The existence of the common conjugacy

In this section, we prove Theorem 1.5 and 1.7. We need the following result of Franks [Fr].

Theorem 3.1. *If $A : \mathbb{T}^N \to \mathbb{T}^N$ is an Anosov diffeomorphism, then A is topologically conjugate to a hyperbolic toral automorphism induced by $A_* :\ H_1(\mathbb{T}^N, \mathbb{Z}) \to H_1(\mathbb{T}^N, \mathbb{Z})$.*

This result has been generalized to the infranilmanifold case by Manning. It is also known ([HK] Theorem 19.1.2) that the conjugacy h is bi-Hölder, i.e. both h and h^{-1} are Hölder.

Proof of Theorem 1.7 and Theorem 1.5. Notice that Theorem 1.5 follows from Theorem 1.7 since the right hand side of the inequality (1.8) is 1 in the two dimensional case. So here we only prove Theorem 1.7.

Suppose we are given an action $\alpha :\ \Gamma_{\bar{A}} \to \mathrm{Diff}^r(\mathbb{T}^N)$ with $\alpha(g_0)(x) = A(x)$ Anosov and homotopic to $\bar{A}$, and $\alpha(g_i)(x) = T_i(x)$, $i = 1, \dots, N$, $x \in \mathbb{T}^N$, with c-slow oscillation where c satisfies (1.8). By Theorem 3.1, we have a conjugacy h such that $hAh^{-1} = \bar{A}$. We next denote $hT_ih^{-1} = R_i,\ i = 1, \dots, N$. We will show that $R_i(x) = x + \rho_i$ where ρ_i is the rotation vector of T_i. We lift h to $\tilde{h} :\ \mathbb{R}^N \to \mathbb{R}^N$ and decompose

$$\tilde{h}(x) = x + g(x), \quad \tilde{h}^{-1}(x) = x + g_-(x), \quad \tilde{T}^{\mathbf{p}}(x) = x + \Delta T^{\mathbf{p}}(x)$$

where $g(x)$, $g_-(x)$ and $\Delta T^{\mathbf{p}}(x)$ are $\mathbb{Z}^N$-periodic. For $\mathbf{p} \in \mathbb{Z}^N$, we have

$$\begin{aligned}\tilde{R}^{\mathbf{p}}(x) &= \tilde{h}\tilde{T}^{\mathbf{p}}\tilde{h}^{-1}(x)\\ &= \tilde{T}^{\mathbf{p}}\tilde{h}^{-1}(x) + g(\tilde{T}^{\mathbf{p}}\tilde{h}^{-1}(x))\\ &= \tilde{h}^{-1}(x) + \Delta T^{\mathbf{p}}(\tilde{h}^{-1}(x)) + g(\tilde{T}^{\mathbf{p}}\tilde{h}^{-1}(x))\\ &= x + \Delta T^{\mathbf{p}}(\tilde{h}^{-1}(x)) + g_-(x) + g(\tilde{T}^{\mathbf{p}}\tilde{h}^{-1}(x)).\end{aligned}$$

Since both g_- and g are uniformly bounded, we get that if $\{T^{\mathbf{p}}\}$ has c-slow oscillation, then so does $\{R^{\mathbf{p}}\}$. From the group relation, we get for $\mathbf{p} \in \mathbb{Z}^N$ and $n \in \mathbb{Z}$,

$$\bar{A}^n\tilde{R}^{\mathbf{p}}(x) = \tilde{R}^{(\bar{A}^t)^n\mathbf{p}}\bar{A}^n(x)+Q_{\mathbf{p},n}, \quad \bar{A}^n(\tilde{R}^{\mathbf{p}}(x)-x) = (\tilde{R}^{(\bar{A}^t)^n\mathbf{p}}-\mathrm{id})\bar{A}^n(x)+Q_{\mathbf{p},n}, \tag{3.1}$$

where $Q_{\mathbf{p},n}$ is an integer vector in $\mathbb{Z}^N$ depending on $\mathbf{p}, n$ and the choice of the lifts.

For each $\tilde{R}^{\mathbf{p}}$, $\mathbf{p} \in \mathbb{Z}^N$, we take the Fourier expansion $\tilde{R}^{\mathbf{p}}(x)-x = \sum_{k\in\mathbb{Z}^N} \hat{R}_k(\mathbf{p})e^{2\pi i\langle k,x\rangle}$, where the coefficient for $k \neq 0$ is

$$\hat{R}_k(\mathbf{p}) = \int_{\mathbb{T}^N} (\tilde{R}^{\mathbf{p}}(x) - x)e^{-2\pi i\langle k,x\rangle}\, dx.$$

The condition that $\{R^{\mathbf{p}}\}$ has c-slow oscillation implies that there exist C, P such that when $\|\mathbf{p}\| \geq P$, we have $\|\hat{R}_k(\mathbf{p})\| \leq C\|\mathbf{p}\|^c$, uniformly for all $k \neq 0$. From the above equation (3.1) we get for all $k \in \mathbb{Z}^N \setminus \{0\}$

$$\hat{R}_k(\mathbf{p}) = \bar{A}^{-n}\hat{R}_{(\bar{A}^t)^{-n}k}((\bar{A}^t)^n\mathbf{p}).$$

We next consider the splitting of $\mathbb{R}^N$ into $\bar{W}^u \oplus \bar{W}^s$, the direct sum decomposition into unstable and stable eigenspaces of $\bar{A}$. We project the above equation into $\bar{W}^u$. We get the estimate for $\|(\bar{A}^t)^n\mathbf{p}\| \geq P$,

$$\|\mathrm{Proj}_{\bar{W}^u}\hat{R}_k(\mathbf{p})\| \leq \frac{1}{|\lambda_1^u|^n}\|\hat{R}_{(\bar{A}^t)^{-n}k}((\bar{A}^t)^n\mathbf{p})\| \leq C\frac{\|(\bar{A}^t)^n\mathbf{p}\|^c}{|\lambda_1^u|^n} \leq C\|\mathbf{p}\|^c\left(\frac{|\lambda_k^u|^c}{|\lambda_1^u|}\right)^n \to 0$$

as $n \to \infty$, if $c < \frac{\ln|\lambda_1^u|}{\ln|\lambda_k^u|}$. Similarly, letting $n \to -\infty$ and projecting to the $\bar{W}^s$ in the above argument, we get that the projection of $\hat{R}_k(\mathbf{p})$ to $\bar{W}^s$ is also 0. Therefore $\hat{R}_k(\mathbf{p}) = 0$ for all $k \neq 0$. This implies that each $R^{\mathbf{p}}(x) - x$, $\mathbf{p} \in \mathbb{Z}^N$, is a constant. Since a conjugacy does not change the rotation vector, we have $R_i(x) = x + \rho_i$, where ρ_i the rotation vector of T_i, $i = 1, \ldots, N$. Next we have $R^{\mathbf{p}}(x) = x + \boldsymbol{\rho}\mathbf{p}$, $\mathbf{p} \in \mathbb{Z}^N$. This completes the proof. □

4. Preliminaries: elliptic and hyperbolic dynamics

In this section, we explain and develop techniques from elliptic dynamics and hyperbolic dynamics that we will use to prove our main results. We first introduce the framework of Herman-Yoccoz-Katznelson-Ornstein for obtaining regularity of the conjugacy of circle maps and generalize it to Abelian group actions on $\mathbb{T}^N$. Next, we state facts about Anosov diffeomorphisms, including the invariant foliation structure and the regularity properties.

4.1. **Elliptic dynamics: the framework of Herman-Yoccoz-Katznelson-Ornstein.** In this section, we generalize to Abelian group actions the framework of Herman-Yoccoz theory for circle maps after Katznelson-Ornstein.

Definition 4.1. (1) *Denote by* $\mathcal{H}^k$ *the group of* C^k *diffeomorphisms on* $\mathbb{T}^N$, $k \in \mathbb{N}$.

(2) *Let* $\mathcal{F}$ *be a continuous foliation of* $\mathbb{T}^N$ *by one-dimensional uniformly* C^1 *leaves* $\mathcal{F}(x),\ x \in \mathbb{T}^N$. *Let* $\mathcal{H}^k_{\mathcal{F}}$ $(< \mathcal{H}^k)$ *be the group of diffeomorphisms in* $\mathcal{H}^k$ *preserving the foliation, i.e.* $f\mathcal{F}(x) = \mathcal{F}(f(x)),\ \forall\ f \in \mathcal{H}^k_{\mathcal{F}}$ *and* $\forall\ x \in \mathbb{T}^N$. *Denote by* $v(x)$ *a unit tangent vector field to the leaves* $\mathcal{F}(x)$, $x \in \mathbb{T}^N$.

We introduce the C^k *norm* $\|\cdot\|_{C^k(\mathcal{F})}$ *along the foliation:*

$$\|\varphi\|_{C^k(\mathcal{F})} := \sum_{i=0}^{k} \sup_x \|(D_{v(x)})^i \varphi(x)\|,$$

for $\varphi \in C^k(\mathbb{T}^N, \mathbb{R}^n)$ *with* $k \in \mathbb{N}, n \geq 1$.

4.1.1. *Generalization of the framework of Herman after Katznelson-Ornstein.* The following statement about circle maps was known to Herman [H]:

Suppose $f = h^{-1}(h(x) + \rho)$, *where* $h:\ \mathbb{T}^1 \to \mathbb{T}^1$ *is a homeomorphism and* $\rho \notin \mathbb{Q}$ *is the rotation number. Then* $h \in \mathcal{H}^k$ *if and only if the iterates* $\{f^j\}_{j\in\mathbb{Z}}$ *are uniformly bounded in* $\mathcal{H}^k$.

Here we generalize this statement to our Abelian subgroup actions after Katznelson-Ornstein [KO].

Definition 4.2. *A collection of* m *vectors* $\rho_1, \dots, \rho_m \in \mathbb{T}^N$ *is said to* rationally generate $\mathbb{T}^N$, *if* $\{\sum_{i=1}^m p_i\rho_i,\ p_i \in \mathbb{Z},\ i = 1, \dots, m\}$ *is dense on* $\mathbb{T}^N$.

Proposition 4.1. *Suppose* $T_i \in \mathcal{H}^k,\ i = 1, \dots, m,\ k \geq 0$ *commute. Suppose also that there exists a homeomorphism* $h:\ \mathbb{T}^N \to \mathbb{T}^N$ *such that* $T_i = h^{-1}\bar{T}_i h$, *where* $\bar{T}_i(x) = x + \rho_i, \mathrm{mod}\ \mathbb{Z}^N,\ i = 1, 2, \dots, m$ *and* $\rho_1, \dots, \rho_m$ *rationally generate* $\mathbb{T}^N$. *Then the following equality holds for all* x*:*

$$\tilde{h}(x) = \mathrm{const.} + \lim_{n\to\infty} \frac{1}{(2n+1)^N} \sum_{\|\mathbf{p}\|_{\ell^\infty} \leq n} \left(\tilde{T}^{\mathbf{p}}(x) - \boldsymbol{\rho}\mathbf{p}\right). \tag{4.1}$$

where $\mathbf{p} = (p_1, \dots, p_m) \in \mathbb{Z}^m$, $\boldsymbol{\rho} = (\rho_1, \dots, \rho_m) \in \mathbb{T}^{N\times m}$ *and* $\tilde{T}^{\mathbf{p}}$ *is the lift of* $T^{\mathbf{p}}$ *satisfying* $\tilde{T}^{\mathbf{p}}(0) = \tilde{h}^{-1}(\tilde{h}(0) + \boldsymbol{\rho}\mathbf{p})$, *and* $\tilde{h}$ *is a fixed lift of* h.

Proof of Proposition 4.1. From $T^{\mathbf{p}} = h^{-1}\bar{T}^{\mathbf{p}}h$, we get $T^{\mathbf{p}} = h^{-1}(h(x) + \boldsymbol{\rho}\mathbf{p})$. We next lift $T^{\mathbf{p}}$ to $\tilde{T}^{\mathbf{p}}$ satisfying $\tilde{T}^{\mathbf{p}}(0) = \tilde{h}^{-1}(\tilde{h}(0) + \boldsymbol{\rho}\mathbf{p})$ and obtain

$$\tilde{T}^{\mathbf{p}}(x) - \boldsymbol{\rho}\mathbf{p} - \tilde{h}(x) = (\tilde{h}^{-1} - \mathrm{id}) \circ (\tilde{h}(x) + \boldsymbol{\rho}\mathbf{p}).$$

Averaging over all $\mathbf{p} \in \mathbb{Z}^m$ with $\|\mathbf{p}\|_{\ell^\infty} \leq n$, and letting $n \to \infty$, we get

$$\tilde{h}(x) = -\int_{\mathbb{T}^N} (\tilde{h}^{-1}(x) - x)dx + \lim_{n\to\infty} \frac{1}{(2n+1)^N} \sum_{\|\mathbf{p}\|_{\ell^\infty} \leq n} \left(\tilde{T}^{\mathbf{p}}(x) - \boldsymbol{\rho}\mathbf{p}\right),$$

where to get the integral, we use the fact that the affine action of $\mathbb{Z}^m$ via the rigid translations $\bar{T}_i,\ i = 1, \dots, m$ is uniquely ergodic and we use a version of Birkhoff ergodic theorem for Abelian group actions [OW, L]. □

Corollary 4.2. (1) *Let the Abelian group $\mathcal{A} = \{T^{\mathbf{p}} :\ \mathbf{p} \in \mathbb{Z}^m\}$ $(< \mathcal{H}^k)$ and the conjugacy h be as in Proposition 4.1.*

(2) *Consider a foliation $\bar{\mathcal{F}}$ of $\mathbb{T}^N$ by straight lines $\bar{\mathcal{F}}(x),\ x \in \mathbb{T}^N$ with fixed slope. Let $\mathcal{F}$ be the foliation of $\mathbb{T}^N$ whose leaves are $\mathcal{F}(x) = h^{-1}(\bar{\mathcal{F}}(h(x)))$ through the point $x \in \mathbb{T}^N$.*

(3) *Assume the leaves $\mathcal{F}(x)$ of the foliation $\mathcal{F}$ are uniformly C^1. Note that this implies that $\mathcal{A} < \mathcal{H}^k_{\mathcal{F}}$.*

(4) *Denote by $v(x)$ (resp. $e(x)$) a unit tangent vector field to the leaves $\mathcal{F}(x)$ (resp. $\bar{\mathcal{F}}(x)$), $x \in \mathbb{T}^N$.*

If we have that the set $\{\tilde{T}^{\mathbf{p}}(x) - \tilde{T}^{\mathbf{p}}(0) - x :\ \mathbf{p} \in \mathbb{Z}^m\} \subset C^k(\mathbb{T}^N, \mathbb{R}^N)$ is precompact in the $\|\cdot\|_{C^k(\mathcal{F})}$ norm, then

$$(D_v)^k h \in C^0,$$

hence h is uniformly C^k along the leaves of $\mathcal{F}$. Moreover, in the case of $k = 1$, we also have that h^{-1} is uniformly C^1 along the leaves of $\bar{\mathcal{F}}$.

The proof of Corollary 4.2 is given in Section 4.1.2.

Given a continuous increasing function $\psi(x) :\ \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ with $\psi(0) = 0$, we say that a function $f : (X, d) \to (X', d')$ between two metric spaces has *modulus of continuity* ψ if there exists a constant $C > 0$ such that

$$d'(f(x), f(y)) \leq C\psi(Cd(x, y)),\ \forall\ x, y \in X.$$

Proposition 4.3. *Let the Abelian group $\mathcal{A}$, the conjugacy h, the foliations $\mathcal{F}$, $\bar{\mathcal{F}}$ and the vector fields v, e be as in Corollary 4.2. Assume that $\{\tilde{T}^{\mathbf{p}}(x) - \tilde{T}^{\mathbf{p}}(0) - x :\ \mathbf{p} \in \mathbb{Z}^m\} \subset C^1(\mathbb{T}^N, \mathbb{R}^N)$ is uniformly bounded in the $\|\cdot\|_{C^1(\mathcal{F})}$ norm, and that the mapping*

$$\boldsymbol{\rho}\mathbf{p} \pmod{\mathbb{Z}^N} \mapsto \tilde{T}^{\mathbf{p}}(x) - \tilde{T}^{\mathbf{p}}(0) - x$$

from $\{\boldsymbol{\rho}\mathbf{p},\ \mathrm{mod}\ \mathbb{Z}^N\} \subset \mathbb{T}^N$ (with the metric induced by $\mathbb{T}^N$) into $C^1(\mathcal{F})$ (with the $\|\cdot\|_{C^1(\mathcal{F})}$ metric) has modulus of continuity ψ. Then both $D_v h$ and $D_e h^{-1}$ have modulus of continuity ψ.

The proof of Proposition 4.3 is given in Section 4.1.3.

4.1.2. *Proof of Corollary 4.2.* We only prove the case of $k=1$. Higher derivatives are similar. We denote the n-th Birkhoff average on the right hand side of of (4.1) by $\tilde{S}_n$, so (4.1) can be rephrased as $\tilde{h}(x)=\lim_{n\to\infty}\tilde{S}_n$ up to an additive constant. Since $\mathcal{A}$ is assumed to be pre-compact in $\mathcal{H}^1_{\mathcal{F}}$ and the pointwise convergence is given by (4.1), we get that $D_v\tilde{S}_n$ is precompact in C^0 by Theorem 5.35 of [AB], which states that the convex hull of compact sets is compact in a completely metrizable locally convex space.

We choose a subsequence $\tilde{S}_{n_j}$ of Birkhoff averages with $\tilde{S}_{n_j}(x)\to\tilde{h}(x)$ and $D_v\tilde{S}_{n_j}(x)\to\tilde{V}(x),\ \forall\ x\in\mathbb{T}^N$ as $j\to\infty$. We consider a leaf $\mathcal{F}(x)$ passing through a given point $x\in\mathbb{T}^N$ and another point $y\in\mathcal{F}(x)$. Such a leaf $\mathcal{F}(x)$ can be given a chart $\Gamma:\ [-1,1]\to\mathcal{F}(x)$ with $\Gamma(t)=x,\ \Gamma(-1)=y$ and $\Gamma'(t)=v(x)$ and $\Gamma\in C^r$. As above, after taking limit we have

$$h(x)-h(y)=h(\Gamma(t))-h(\Gamma(-1))=\int_{-1}^{t}\tilde{V}(\Gamma(t))\Gamma'(t)\,dt,$$

where the integral is taken along the leaf $\mathcal{F}(x)$ from y to x. This gives $D_vh(x)=\tilde{V}(\Gamma(t))\Gamma'(t)=\tilde{V}(x)v(x)$.

We have proved that $D_vh(x)$ is continuous. To show that D_eh^{-1} is also continuous, by the implicit function theorem, it is enough to show that $\|D_vh(x)\|$ is bounded away from zero. If $\|D_vh(x)\|<\varepsilon$ for some given small ε at some point x_0, then the same inequality holds in a small neighborhood $B(x_0)$ of x_0. By the ergodicity of $T^{\mathbf{p}}$, there exist finitely many $\mathbf{p}_i,\ i=1,\dots,n$ such that $\cup_{i=1}^n T^{\mathbf{p}_i}(B(x_0))=\mathbb{T}^N$. Therefore there exists a constant C independent of $\mathbf{p}$ such that $\|D_vh(x)\|<C\varepsilon$ for all $x\in\mathbb{T}^N$ using the equation $D_{v(T^{\mathbf{p}}(x))}h(T^{\mathbf{p}}(x))\|D_{v(x)}T^{\mathbf{p}}(x)\|=D_{v(x)}h(x)$. Consider a leaf $\mathcal{F}(x)$ with two endpoints x and x'. We lift the leaf to the universal cover and consider its image under $\tilde{h}$, i.e. the line segment between $\tilde{h}(x)$ and $\tilde{h}(x')$. Since $\tilde{h}(x)=x+g(x)$ where $g(x)$ is $\mathbb{Z}^N$ periodic, choosing x and x' far apart we can make $\|\tilde{h}(x')-\tilde{h}(x)\|\geq 1$. On the other hand, since $\mathcal{F}(x)$ is a C^1 leaf, the length of $\mathcal{F}(x)$ between x and x' is bounded by a number $C_{x,x'}$. We consider the integral $1\leq\|\tilde{h}(x)-\tilde{h}(x')\|=\int_{\mathcal{F}(x)}\|D_vh(x)\|\,dx\leq\varepsilon C_{x,x'}$. For fixed x,x', the ε cannot be arbitrarily small. This completes the proof. $\square$

4.1.3. *Proof of Proposition 4.3.* By the assumption, $\{\tilde{T}^{\mathbf{p}}(x)-\tilde{T}^{\mathbf{p}}(0)-x,\ \mathbf{p}\in\mathbb{Z}^m\}$ is pre-compact in $C^1(\mathcal{F})$ since the map from $\boldsymbol{\rho}\mathbf{p}\in\mathbb{T}^N$ to $C^1(\mathcal{F})$ is continuous and $\mathbb{T}^N$ is compact. Therefore, by Corollary 4.2, we get $D_vh, D_eh^{-1}\in C^0$. Differentiating the expression $T^{\mathbf{p}}(x)=h^{-1}(h(x)+\boldsymbol{\rho}\mathbf{p})$ along the leaf $\mathcal{F}(x)$, we get that

$$D_vT^{\mathbf{p}}(x)-v=\left(D_eh^{-1}(h(x)+\boldsymbol{\rho}\mathbf{p})-D_eh^{-1}(h(x))\right)\|D_vh(x)\|.$$

Since the LHS satisfies the modulus of continuity ψ by assumption, i.e.

$$\|D_vT^{\mathbf{p}}(x)-v\|_{C^0}\leq\psi(\|\boldsymbol{\rho}\mathbf{p}\|),$$

for $\mathbf{p}$ such that $\|\boldsymbol{\rho}\mathbf{p}\|$ small, we get that the RHS satisfies

$$\| \left(D_e h^{-1}(h(x)+\boldsymbol{\rho}\mathbf{p}) - D_e h^{-1}(h(x))\right) \|D_v h(x)\| \|_{C^0} \leq \psi(\|\boldsymbol{\rho}\mathbf{p}\|).$$

This implies

$$\| \left(D_e h^{-1}(h(x)+\boldsymbol{\rho}\mathbf{p}) - D_e h^{-1}(h(x))\right) \|_{C^0} \leq \psi(\|\boldsymbol{\rho}\mathbf{p}\|)(\min_x \|(D_{v(x)}h(x))\|)^{-1},$$

hence $D_e h^{-1}$ has modulus of continuity ψ. To get the same modulus of continuity for $D_v h$, we use the fact that $D_v h$ is a vector parallel to e and $\|D_v h\| = \|D_e h^{-1}\|^{-1}$. □

4.2. **Hyperbolic dynamics: invariant foliations of Anosov diffeomorphisms.** In this section, we state results from hyperbolic dynamics. We next state only the theorems and propositions for the unstable manifolds W^u and unstable distribution E^u. The stable analogues also hold.

Definition 4.3. *We say that $A: \mathbb{T}^N \to \mathbb{T}^N$ is a C^r, $r > 1$, Anosov diffeomorphism with* simple Mather spectrum *if there exists a DA-invariant splitting of the tangent space*

$$T_x\mathbb{T}^N = E^s_\ell(x) \oplus \ldots \oplus E^s_1(x) \oplus E^u_1(x) \oplus \ldots \oplus E^u_k(x), \quad k+\ell = N,\ k,\ell \geq 1$$

and numbers

$$\underline{\mu}^s_\ell \leq \bar{\mu}^s_\ell < \ldots < \underline{\mu}^s_1 \leq \bar{\mu}^s_1 < 1 < \underline{\mu}^u_1 \leq \bar{\mu}^u_1 < \ldots < \underline{\mu}^s_k \leq \bar{\mu}^s_k$$

such that for some constant $C > 1$,

$$\frac{1}{C}(\underline{\mu}^{u,s}_i)^n \leq \frac{\|DA^n v\|}{\|v\|} \leq C(\bar{\mu}^{u,s}_i)^n, \quad \forall v \in E^{u,s}_i \setminus \{0\},$$

where $i = 1, \ldots, \ell$ for s and $i = 1, \ldots, k$ for u.

The next result is classical (see [HPS]).

Proposition 4.4. *For any C^r, $r > 1$ Anosov diffeomorphism $A: \mathbb{T}^N \to \mathbb{T}^N$ with simple Mather spectrum, the strong invariant distribution $E^u_{i\leq} := E^u_i \oplus \ldots \oplus E^u_k$ is uniquely integrable into a foliation $W^u_{i\leq}$ of $\mathbb{T}^N$ whose leaf $W^u_{i\leq}(x)$ passing through x is $C^r, x \in \mathbb{T}^N$. This gives rise to a flag of strong unstable foliations*

$$W^u_k(x) \subset W^u_{(k-1)\leq}(x) \subset \ldots \subset W^u_{2\leq}(x) \subset W^u_{1\leq}(x) := W^u(x), \quad x \in \mathbb{T}^N,$$

where each of the inclusions is proper and $W^u_{i\leq}$ sub-foliates $W^u_{(i-1)\leq}$ with C^r leaves for $i = 2, \ldots, k$.

It is known that simple Mather spectrum is an open property in the C^1 topology. In particular, if $\bar{A}$ is a toral automorphism with simple real spectrum, then an Anosov diffeormophism that is C^1 close to $\bar{A}$ has simple Mather spectrum.

Proposition 4.5 (Hölder regularity of the invariant distribution, Theorem 19.1.6 of [HK])**.** *Each distribution $E^u_i(x)$ is Hölder in the base point x.*

The Holder exponents of the distributions depend only on the expansion and contraction rates $\bar{\mu}_i^{u,s}$ and $\underline{\mu}_i^{u,s}$.

We next denote the weak unstable bundles for A by $E^u_{\leq i} := E^u_1(x) \oplus E^u_2(x) \oplus \cdots \oplus E^u_i(x)$, and that of $\bar{A}$ by $\bar{E}^u_{\leq i}(x) := \bar{E}^u_1(x) \oplus \bar{E}^u_2(x) \oplus \cdots \oplus \bar{E}^u_i(x)$, $i = 1, \ldots, k$, $x \in \mathbb{T}^N$. Denote the unstable manifolds of A by $W^u(x)$ and that of $\bar{A}$ by $\bar{W}^u(x)$.

Proposition 4.6 (Lemma 6.1-6.3 of [G1])**.** *Consider A a C^r Anosov diffeomorphism that is C^1 close to a linear toral automorphism $\bar{A}$ with simple real spectrum, and the bi-Hölder conjugacy h given by Theorem 3.1 with $h \circ A = \bar{A} \circ h$. Then*

(1) *The conjugacy h preserves the unstable manifold, $h(W^u(x)) = \bar{W}^u(h(x))$, $x \in \mathbb{T}^N$.*
(2) *Each weak unstable distribution $E^u_{\leq i}$ is uniquely integrable into a foliation $W^u_{\leq i}$ of $\mathbb{T}^N$, whose leaf $W^u_{\leq i}(x)$ passing through $x \in \mathbb{T}^N$ is C^{1+}.*
(3) *Each distribution $E^u_{i,j} := E^u_{\geq i} \cap E^u_{\leq j}$, $i \leq j$, is uniquely integrable into C^{1+} leaves.*
(4) *The weak unstable foliation is preserved by the conjugacy, i.e. $h(W^u_{\leq i}(x)) = \bar{W}^u_{\leq i}(h(x))$, $i = 1, \ldots, k$, $\forall\ x \in \mathbb{T}^N$.*

We remark that the item (1) does not need the C^1 closeness of A to $\bar{A}$, and it holds under the same assumption as Theorem 3.1.

This proposition gives us also a weak flag of foliations

$$W^u_1 \subset W^u_{\leq 2} \subset \ldots \subset W^u_{\leq k-1} \subset W^u_{\leq k} := W^u,$$

where each of the inclusions is proper and $W^u_{\leq (i-1)}$ sub-foliates $W^u_{\leq i}$ with C^{1+} leaves for $i = 2, \ldots, k$. This flag is preserved by the conjugacy h.

When the weak distributions E^u_i are known to be uniquely integrable, we have the following proposition by the standard graph transform technique.

Proposition 4.7. (1) *Each weak unstable leaf $W^u_{\leq i}$ in item (2) of Proposition 4.6 is again subfoliated by W^u_i each of whose leaves are C^r.*
(2) *The weakest unstable leaf $W^u_1(x)$ is C^{1+}, and its tangent distribution $E^u_1(x)$ is Hölder.*

5. The elliptic regularity with the help of the invariant distribution

5.1. **The elliptic dynamics and the invariant distribution.** In this section, we show how to get regularity of the conjugacy using the elliptic dynamics combined with the invariant distribution provided by the hyperbolic dynamics.

Lemma 5.1. *Let $\mathcal{F}$ be a foliation of $\mathbb{T}^N$ by C^1 leaves $\mathcal{F}(x)$, $x \in \mathbb{T}^N$. Suppose a homeomorphism h conjugates the Abelian group $\{T^{\mathbf{p}},\ \mathbf{p} \in \mathbb{Z}^m\}$ to translations $\{\bar{T}^{\mathbf{p}}(x) = x + \boldsymbol{\rho}\mathbf{p},\ \mathbf{p} \in \mathbb{Z}^m\}$ and send the foliation $\mathcal{F}$ into a foliation $\bar{\mathcal{F}}$ by straight lines $\bar{\mathcal{F}}(x)$, $x \in \mathbb{T}^N$ with fixed slope. Denote by $E(x)$ the one-dimensional distribution that is tangent to $\mathcal{F}(x)$. Then the distribution $E(x)$ is invariant under the $DT^{\mathbf{p}}$. Namely*

$$DT^{\mathbf{p}}(x)E(x) = E(T^{\mathbf{p}}(x)),\ \mathbf{p} \in \mathbb{Z}^m.$$

Proof. The straight line foliation $\bar{\mathcal{F}}$ is invariant under translations, so after the conjugation the foliation $\mathcal{F}$ is also invariant under $\{T^{\mathbf{p}},\ \mathbf{p} \in \mathbb{Z}^m\}$. The lemma follows directly by differentiating the equation $T^{\mathbf{p}}\mathcal{F}(x) = \mathcal{F}(T^{\mathbf{p}}(x))$ along the leaves. □

Lemma 5.2. *Suppose the conjugacy h in the previous Lemma 5.1 is bi-Hölder. Then for each $\mathbf{p} \in \mathbb{Z}^m$, all the Lyapunov exponents of $T^{\mathbf{p}}$ with respect to the invariant measure $\mu = h_* m$ are zero, where m is the Lebesgue measure supported on the orbit closure of $x + t\boldsymbol{\rho}\mathbf{p},\ t \in \mathbb{R}$ mod $\mathbb{Z}^N$.*

Proof. Suppose not, by Pesin theory, non vanishing Lyapunov exponent implies the existence of local (un)stable manifolds of a.e. $x \in \mathrm{supp}(\mu)$. Pick two points y, y' on a same stable manifold (for unstable manifold, we will use backward iterate) of a point x, then we have $\|T^{n\mathbf{p}}(y) - T^{n\mathbf{p}}(y')\|$ converges to zero exponentially with a rate being the Lyapunov exponent. On the other hand, using the conjugacy we have

$$\|T^{n\mathbf{p}}(y) - T^{n\mathbf{p}}(y')\| = \|h^{-1}(n\boldsymbol{\rho}\mathbf{p} + h(y)) - h^{-1}(n\boldsymbol{\rho}\mathbf{p} + h(y'))\| \geq \text{const.}\|h(y) - h(y')\|^{\eta}.$$

where η is the Hölder exponent of h^{-1}, since h is bi-Hölder. We get a contradiction. □

5.2. **A quantitative Kronecker theorem.** We need the following number theoretic result, whose proof is postponed to the Appendix.

Theorem 5.3. *Let $N, K \in \mathbb{N}$ be given. Then there exists a full measure set $\mathcal{O}$ in the set $\mathcal{M}_{N\times K}(\mathbb{T})$ of matrices of $N \times K$ such that for all $M \in \mathcal{O}$, the following holds.*

For any small $\epsilon > 0$, there exists a constant C such that for any $y \in \mathbb{T}^N$ and any $n \in \mathbb{N}$ there exist $q \in \mathbb{Z}^N$, $p \in \mathbb{Z}^K$ satisfying $\|p\| < n$, such that the following inequality holds

$$\|Mp - q - y\| \leq Cn^{-\frac{K}{N}-\epsilon}.$$

This theorem is a quantitative version of the classical Kronecker's approximation theorem. When $K = 1$, this is the classical Dirichlet's simultaneous Diophantine approximation theorem where we can set $\epsilon = 0$. The $N = 1$ case was proved in [K].

This theorem inspires the following definition.

Definition 5.1. *Suppose the m vectors $\rho_1,\dots,\rho_m \in \mathbb{T}^N$ rationally generate $\mathbb{T}^N$, and consider the set of finite linear combinations*

$$S(\rho_1,\dots,\rho_m) := \left\{\sum_{i=1}^m p_i\rho_i \bmod \mathbb{Z}^N \mid p_i \in \mathbb{Z},\quad i=1,\dots,m\right\}. \tag{5.1}$$

For each element $\gamma \in S$, we denote by $\|\gamma\|_w$ the word length $\|\gamma\|_w := \|\mathbf{p}\|_{\ell_1}$, where $\mathbf{p} = (p_1,\dots,p_m) \in \mathbb{Z}^m$ and by $\|\gamma\|$ the closest Euclidean distance of γ (mod $\mathbb{Z}^N$) to zero.

We say that S has dimension d *if there exists a constant c such that for any $x \in \mathbb{T}^N$ and any $\ell > 0$ there exists a point $\gamma \in S$ satisfying*

$$\|\gamma\|_w \le \ell, \quad \|\gamma - x\| \le c\ell^{-d}.$$

Theorem 5.3 implies that for almost every choice of vector tuple $\rho_1,\dots,\rho_m \in \mathbb{T}^N$, the set S formed by linear combinations as above has dimension $m/N - \epsilon$ for all $\epsilon > 0$ small.

5.3. **Organization of the proofs of Theorem 1.6 and Theorem 1.8.** The proofs of Theorem 1.6 and Theorem 1.8 are based on the following propositions.

The first proposition chooses the K_0 in Theorem 1.6 and Theorem 1.8.

Proposition 5.4. *Given $\eta \in (0,1)$ and $d > 2/\eta^2$, there exists K_0 such that the following holds: for all $K > K_0$, there exists a full measure set $\mathcal{R}_{N,K} \subset (\mathbb{T}^N)^K$ such that the set S generated by any tuple of vectors $(\rho_1,\dots,\rho_K)$ lying in $\mathcal{R}_{N,K}$ is dense on $\mathbb{T}^N$ and has dimension d.*

Proof of Proposition 5.4. To satisfy the inequality $2/d < \eta^2$, we choose $K_0 > 2N/\eta^2$. Applying Theorem 5.3, we get a full measure set in $(\mathbb{T}^N)^K$, $K > K_0$, each point of which generates a set S of dimension $d = K/N - \epsilon$ satisfying $2/d < \eta^2$, where ϵ is arbitrarily small. Next, removing further a zero measure set to guarantee that the vectors rationally generate $\mathbb{T}^N$, we get the full measure set $\mathcal{R}_{N,K}$ as claimed. $\square$

The next proposition gives the choice of η in Proposition 5.4, and will give the C^{1+} regularity of h along the one-dimensional leaves of a foliation applying Corollary 4.2 and Proposition 4.3.

Proposition 5.5. *Suppose*

1. *the Abelian group $\mathcal{A}(< \mathcal{H}^r)$ is generated by*
$$\{T_{i,j} \mid \ i=1,\dots,N,\ j=1,\dots,K,\ T_{i,j}T_{i',j'} = T_{i',j'}T_{i,j}\};$$
2. *there is an η-bi-Hölder conjugacy h such that $T_{i,j}(x) = h^{-1}(h(x)+\rho_{i,j})$;*

(3) *there is a $\{T_{i,j}\}$-invariant foliation $\mathcal{F}$ into one-dimensional C^1 leaves $\mathcal{F}(x)$ with tangential distributions $E(x)$, $x \in \mathbb{T}^N$ that is η-Hölder in x. Denote by $v(x)$ a unit vector field tangent to $\mathcal{F}(x)$, $x \in \mathbb{T}^N$;*

(4) *the set S generated by the rotation vectors $\rho_{i,j}$ has dimension d, with $2/d < \eta^2$.*

For any $\gamma \in S$, we denote

$$T_\gamma := \prod_{j=1}^{K} \prod_{i=1}^{N} T_{i,j}^{q_{i,j}}$$

where $q_{i,j} \in \mathbb{Z}$ are the coefficients in the linear combination of γ, i.e. $\gamma = \sum_j \sum_i q_{i,j}\rho_{i,j}$.

Then for all $\gamma \in S$ with $\|\gamma\|$ small enough, we have

$$|\|DT_\gamma(x)v(x)\|_{C^0} - 1| \leq \text{const.}\|\gamma\|^\nu$$

where $\nu \leq \eta^2 - 2/d$.

This proposition will be proved in Section 5.4. We next cite the following well-known theorem of Journé.

Theorem 5.6 ([J])**.** *Suppose $\mathcal{F}^1$, $\mathcal{F}^2$ are two transverse uniformly $C^{n,\nu}$ foliations of a manifold M. Suppose that a continuous function $u : M \to \mathbb{R}$ is uniformly $C^{n,\nu}$ when restricted to each local leaf $\mathcal{F}^1_\varepsilon(x), \mathcal{F}^2_\varepsilon(x)$, $x \in M$. Then u is $C^{n,\nu}$ on M.*

In the two dimensinonal case, we apply Proposition 5.5 and Proposition 4.3 to get that h is C^{1+} along the stable and unstable foliations of the Anosov diffeomorphism A. Applying Theorem 5.6, we get that h is C^{1+} on $\mathbb{T}^2$.

An application of the next result completes the proof of Theorem 1.6. More details of the proof of Theorem 1.6 will be given in Section 6.1.

Theorem 5.7 ([LMM, Ll])**.** *Suppose f and g are two C^r, $r > 1$, Anosov diffeomorphims $\mathbb{T}^2$ that are topologically conjugated by h, i.e. $f \circ h = h \circ g$. Suppose the periodic data of f and g coincide, namely, $Df^q(h(x))$ is conjugate to $Dg^q(x)$ at every q-periodic point x of g for all $q \in \mathbb{Z}$. Then $h \in C^{r-\varepsilon}$ for ε arbitrarily small.*

The proof of Theorem 1.8 in the $N > 2$ case follows from the same general strategy. However, there is some more work needed to show that the conjugacy h sends the one-dimensional leaves $W_i^{u,s}$ to the straight lines parallel to the eigenvectors of $\bar{A}$. We will give the proof of the C^{1+} regularity of h in Section 6.2.

In dimension three, we get improved regularity (Corollary 1.9) applying the following result of Gogolev in [G2].

Theorem 5.8 (Addendum 1.2 of [G2])**.** *Suppose $\bar{A} \in \mathrm{SL}(3, \mathbb{Z})$ has simple real spectrum and $A :\ \mathbb{T}^3 \to \mathbb{T}^3$ is C^r, $r > 3$ that is C^1 close to $\bar{A}$. Suppose also that $\bar{A}$ and A have* *the same periodic data, then there exists* $h:\ \mathbb{T}^3\to\mathbb{T}^3$ *in* $C^{r-3-\varepsilon}$ *with* $h\circ A=\bar{A}\circ h$. *Furthermore there exists* $\kappa\in\mathbb{Z}$, *such that if* $r\notin(\kappa,\kappa+3)$, *then* $h\in C^{r-\varepsilon}$, *where* ε *is arbitrarily small.*

5.4. **Elliptic regularity in the presence of invariant distributions.** In this section, we prove Proposition 5.5.

Proof of Proposition 5.5. Let T_γ and $v(x)$ be as in the statement and $\mu=h_*m$ be as in Lemma 5.2. Note that by assumption, the set S generated by the rotation vectors has dimension $d>0$ and by Definition 5.1, the measure m is Lebesgue on $\mathbb{T}^N$, so μ is supported on $\mathbb{T}^N$. We get from Lemma 5.2 that

$$\lim_k \frac{1}{k}\log\|D(T_\gamma)^k(x)v(x)\| = \int \log\|DT_\gamma(x)v(x)\|\,d\mu = 0. \tag{5.2}$$

This shows that $\log\|DT_\gamma(x)v(x)\|$ vanishes at some point on $\mathbb{T}^N$.

To simplify notation, we reindex the $T_{i,j}$ appearing in T_γ by $T_1,\dots,T_{\|\gamma\|_w}$, and write $T_\gamma=\prod_{i(\gamma)=1}^{\|\gamma\|_w}T_{i(\gamma)}$. (Due to the commutativity of the $T_{i,j}$'s, the ordering of the (i,j) appearing in $i(\gamma)$ does not matter). We also write $T_{i(\gamma)}x=x_{i(\gamma)}$.

We next introduce $\log\|DT_{i(\gamma)+1}(x)v(x)\|:=\ell_{i(\gamma)}(x)$ which is η-Hölder in x. By the invariance of the distribution E, we get

$$\begin{aligned}\log\|DT_\gamma(x)v(x)\| = \log\left\|D\prod_{i(\gamma)=1}^{\|\gamma\|_w}T_{i(\gamma)}(x)v(x)\right\| &= \sum_{i(\gamma)=1}^{\|\gamma\|_w}\log\|DT_{i(\gamma)+1}(x_{i(\gamma)})v(x_{i(\gamma)})\|\\ &= \sum_{i(\gamma)=1}^{\|\gamma\|_w}\ell_{i(\gamma)}(x_{i(\gamma)}).\end{aligned}$$

To prove the lemma, it is enough to consider in a neighborhood of $\gamma=0$. We consider a dyadic decomposition of a small neighborhood of 0 by

$$D_m=\{\gamma\in S\mid c2^{-d(m+1)/2}<\|\gamma\|\le c2^{-dm/2}\},$$

where c is the constant in Definition 5.1. Next, for D_m, we introduce a $c2^{-dm}$-net by defining

$$S_m:=\{\gamma\in D_m\mid \|\gamma\|_w\le 2^m\}.$$

The remaining proof is split into two steps. In the first step, we prove the following

Claim 1: *For any* $\gamma\in S_m$, *we have*

$$|\log\|DT_\gamma(x)v(x)\||\le \text{const.}\|\gamma\|^{\eta^2-2/d},\ \forall\, x\in\mathbb{T}^N.$$

Proof of Claim 1. First, by (5.2), for any given $\gamma \in S_m$, there exists $y \in \mathbb{T}^N$ such that $\log \|DT_\gamma(y)v(y)\| = 0$. Next, it follows from the definition of the dimension of the set S that there exists $\delta \in S$ with

$$\|\delta\|_w \le \|\gamma\|_w, \quad \|\delta + \bar{y} - \bar{x}\| \le c\|\gamma\|_w^{-d},\ \bar{x} = h(x),\ \bar{y} = h(y).$$

We denote $y_\delta = T_\delta y$ and $y_\gamma = T_\gamma y$.

Since h is bi-Hölder, we have for all $i(\gamma) = 0, 1, \ldots, \|\gamma\|_w - 1$ and $i(\delta) = 0, 1, \ldots, \|\delta\|_w - 1$, the following estimate

$$\|x_{i(\gamma)} - (y_\delta)_{i(\gamma)}\| \le \text{const.}\|\gamma\|_w^{-d\eta}, \quad \|y_{i(\delta)} - (y_\gamma)_{i(\delta)}\| \le \text{const.}\|\gamma\|^\eta.$$

Next we estimate $\log \|DT_\gamma(x)v(x)\|$ as follows

$$\begin{aligned}
&|\log \|DT_\gamma(x)v(x)\||\\
&= |\log \|DT_\gamma(x)v(x)\| - \log \|D(T_\gamma T_\delta)(y)v(y)\| + \log \|D(T_\delta T_\gamma)(y)v(y)\||\\
&= |\log \|DT_\gamma(x)v(x)\| - \log \|DT_\gamma(T_\delta y)v(T_\delta y)\|\\
&\ - \log \|DT_\delta(y)v(y)\| + \log \|DT_\delta(T_\gamma y)v(T_\gamma y)\| + \log \|DT_\gamma(y)v(y)\||\\
&= |\log \|DT_\gamma(x)v(x)\| - \log \|DT_\gamma(T_\delta y)v(y_\delta)\|\\
&\ - \log \|DT_\delta(y)v(y)\| + \log \|DT_\delta(T_\gamma y)v(y_\gamma)\||\\
&= \Big| \sum_{i(\gamma)=1}^{\|\gamma\|_w - 1} (\ell_{i(\gamma)}(x_{i(\gamma)}) - \ell_{i(\gamma)}((y_\delta)_{i(\gamma)})) + \sum_{i(\delta)=1}^{\|\delta\|_w - 1} (\ell_{i(\delta)}(y_{i(\delta)}) - \ell_{i(\delta)}((y_\gamma)_{i(\delta)}))\Big| \qquad (5.3)\\
&\le \text{const.}(\|\gamma\|_w \cdot \|\gamma\|_w^{-\eta^2 d} + \|\delta\|_w \|\gamma\|^{\eta^2})\\
&\le \text{const.}(2^{m(1-\eta^2 d)} + 2^m \|\gamma\|^{\eta^2})\\
&\le \text{const.}(\|\gamma\|^{2(\eta^2 - 1/d)} + \|\gamma\|^{\eta^2 - 2/d})\\
&\le \text{const.}\|\gamma\|^{\eta^2 - 2/d}.
\end{aligned}$$

□

In the second step, we prove the following.

Claim 2: *Suppose for any $\gamma_m \in S_m$, we have $\| \log \|DT_{\gamma_m}(x)v(x)\|\|_{C^0} \le \text{const.}\|\gamma_m\|^\nu$, $\nu > 0$. Then for any $\gamma \in S$, we have $\| \log \|DT_\gamma(x)v(x)\|\|_{C^0} \le \text{const.}\|\gamma\|^\nu$*

Proof of Claim 2. By the definition of D_m and S_m and Definition 5.1, we get that each annulus D_m in the dyadic decomposition is covered by at least $O(2^{dNm/2})$ balls of radius $c2^{-dm}$ centered at points in S_m.

We claim that for any $\gamma \in S$ with small norm $\|\gamma\|$, there exists a finite number $\kappa(\gamma)$ and $\{\gamma_{m_k},\ k = 1, 2, \ldots, \kappa(\gamma)\}$ satisfying $\gamma_{m_k} \in D_{m_k}$, and $m_{k+1} \ge 2m_k$ and $\gamma = \sum_{k=1}^{\kappa(\gamma)} \gamma_{m_k}$.

The algorithm is as follows. First find m such that $\gamma \in D_m$. Denote this m by m_1 and find $\gamma_{m_1} \in S_{m_1}$ that is closest to γ. The closest distance is bounded by $c2^{-dm_1}$. Next consider the vector $\gamma-\gamma_{m_1}$ and repeat the above procedure to it in place of γ. We see that $\gamma-\gamma_{m_1} \in D_{m_2}$ for some $m_2 \geq 2m_1$. This procedure terminates after finitely many steps since $\gamma \in S$ is a finite integer linear combination of the rotation vectors $\rho_i,\ i=1,\dots,m$.

We next denote $x_{m_i}=\prod_{j=i}^{\kappa(\gamma)} T_{\gamma_{m_j}}(x)$ and $x_{m_{\kappa(\gamma)+1}}=x$. So we get

$$\begin{aligned}|\log\|DT_\gamma(x)v(x)\|| &= |\log\|D\prod_i T_{\gamma_{m_i}}(x)v(x)\||\\ &\leq \sum_i |\log\|DT_{\gamma_{m_i}}(x_{m_{i+1}})v(x_{m_{i+1}})\||\\ &\leq \text{const.}\sum_{i=1}^{\kappa(\gamma)}\|\gamma_{m_i}\|^\nu.\end{aligned}$$

By the construction of D_m and S_m, we have that $\frac{1}{2}\|\gamma_{m_1}\| \leq \|\gamma\| \leq 2\|\gamma_{m_1}\|$ and $\|\gamma_{m_k}\|$ decays exponentially with uniform exponential rate for all $\gamma \in S$. This gives that $|\log\|D_vT_\gamma\|| \leq \text{const.}\|\gamma\|^\nu$ for every $\gamma \in S$ close to zero. □

This completes the proof of Proposition 5.5. □

6. Proof of the Theorems

In this section, we prove Theorem 1.6 and Theorem 1.8.

6.1. **Proof of Theorem 1.6.**

Proof of Theorem 1.6. First, we explain how to choose K_0 and the open set $\mathcal{O}$ in the statement of Theorem 1.6. We choose $\mathcal{O}$ to be a C^1 neighborhood of $\bar{A}$ in the set of Anosov diffeomorphisms with simple spectrum.

By Proposition 5.4, in order to determine K_0 it is enough to determine η. Given $\bar{A}$ and an Anosov diffeomorphsm $A:\mathbb{T}^2\to\mathbb{T}^2$ homotopic to $\bar{A}$. By Theorem 3.1, there is a bi-Hölder conjugacy h such that $h\circ A=\bar{A}\circ h$. The Hölder regularity of the conjugacy h depends on both the spectrum of $\bar{A}$ and the Mather spectrum of A ([HK] Theorem 19.1.2), and the Hölder regularity of the invariant distribution $E_i(x)$ of the Anosov diffeomorphism A depends on the Mather spectrum of A. The η is chosen to be the smaller one of the two Hölder exponents.

For given $K>K_0$, we get a full measure set $\mathcal{R}_{2,K}$ in $\mathbb{T}^{N\times K}$ by applying Proposition 5.4. For given function $i:\ \{1,\dots,K\}\to\{1,2\}$, if the rotation vectors of

$(\rho_{i(1),1},\dots,\rho_{i(K),K})$ lie in $\mathcal{R}_{2,K}$, then the set S generated by the set of all rotation vectors $\{\rho_{i,j},\ i=1,2,\ j=1,\dots,K\}$ has dimension $d\in(K/2,K)$. For $K>K_0$, we have $2/d<\eta^2$ by Proposition 5.4. Moreover S is dense on $\mathbb{T}^N$.

We consider the action $\alpha:\ \Gamma_{\bar{A},K}\to \mathrm{Diff}^r(\mathbb{T}^2)$ with $\alpha(g_0)=A:\ \mathbb{T}^2\to\mathbb{T}^2$ Anosov, and $\alpha(g_{i,k}),\ i=1,2,\ k=1,\dots,K$ generating an Abelian subgroup action $(\mathbb{Z}^2)^K\to \mathrm{Diff}^r(\mathbb{T}^2)$. By assumption, the subgroup action generated by $\alpha(g_{i,1}),\ i=1,2$, satisfies sublinear oscillation. So we can apply Theorem 1.5 to the $\Gamma_{\bar{A}}$ action generated by $\alpha(g_0)$ and $\alpha(g_{i,1}),\ i=1,2$, to get a bi-Hölder conjugacy h linearizing the $\Gamma_{\bar{A}}$ action.

We show that the conjugacy h given in Theorem 1.5 also linearizes the whole $\Gamma_{\bar{A},K}$ action α. Indeed, for any diffeomorphism f that commutes with $\alpha(g_{1,1}),\alpha(g_{2,1})$, we have

$$hfh^{-1}(x+\rho_{i,1})=hfh^{-1}(x)+\rho_{i,1},\quad i=1,2$$

after applying the conjugacy h. Since the rotation vectors $\rho_{i,1},\ i=1,2$, rationally generate $\mathbb{T}^N$, by taking Fourier expansion, we get that hfh^{-1} is a rigid rotation by a constant vector that is the rotation vector of f. Therefore, we get that h conjugates the whole action α to an affine action by rigid translations.

We next apply Proposition 5.5, Corollary 4.2 and Proposition 4.3 to get that the conjugacy h is C^{1+} along each stable and unstable leaves of the Anosov diffeomorphism A. By Theorem 5.6, we get that h is C^{1+} on $\mathbb{T}^2$ and finally by Theorem 5.7, we get that h is $C^{r-\varepsilon}$ for any ε small enough. □

6.2. **Proof of Theorem 1.8, the N dimensional case.** The main difficulty in generalizing the above argument to the N-dimensional case is that it is in general unknown if the one dimensional distributions E^u_i (or E^s_i) that are invariant under DA are also invariant under DT_γ. It is only known that the weakest stable and unstable distributions E^u_1 and E^s_1 are invariant under DT_γ by Proposition 4.6 (4) and Lemma 5.1.

We cite the following Lemma from [GKS].

Lemma 6.1 (Proposition 2.4 of [GKS])**.** *Let A, $\bar{A}$ and h be as in Proposition 4.6. Suppose h is C^{1+} along $W^u_{\le i}$ and $h(W^u_j(x))=\bar{W}^u_j(h(x))$, $1\le j\le i$, then*

$$h(W^u_{i+1}(x))=\bar{W}^u_{i+1}(h(x)),\ x\in\mathbb{T}^N.$$

Using this lemma, we now prove that $h\in C^{1+}$ in the general case $N>2$.

Proof of Theorem 1.8. The proof follows the strategy of the proof of Theorem 1.6 with small modifications dealing with the high dimensionality.

We first choose K_0 and the open set $\mathcal{O}$ of Anosov diffeomorphisms. Since $\bar{A}$ is assumed to have simple spectrum, it has a C^1 small neighborhood in which the Anosov diffeomorphisms have simple Mather spectrum. We choose such a neighborhood and denote it by $\mathcal{O}$. We will choose K_0 to satisfy $2/d < \eta^2$ using Proposition 5.4, where d is the dimension of the set S generated by the rotation vectors $\rho_{i,j}$ and η is the smallest one among the Hölder exponent of the conjugacy h and the Hölder exponents of all the distributions $E_i^{u,s}$ tangent to the one-dimensional stable and unstable leaves $W_i^{u,s}$, for all the Anosov diffeomorphisms in $\mathcal{O}$.

By Proposition 5.4, we get the full measure set $\mathcal{R}_{N,K}$ in $\mathbb{T}^{N\times K}$. We next get a bi-Hölder conjugacy h which linearizes the whole action $\alpha:\ \Gamma_{\bar{A},K}\to \mathrm{Diff}^r(\mathbb{T}^N)$ applying Theorem 1.7 and the argument in the proof of Theorem 1.6.

It remains to improve the regularity of h to C^{1+}. To start, we first have that the weakest leaves are preserved $h(W_1^u(x)) = \bar{W}_1^u(h(x))$ by Proposition 4.6 (4). Then we apply Lemma 5.1 to get that the weakest distribution E_1^u is invariant under the Abelian group action generated by $\alpha(g_{i,k}),\ i=1,\dots,N,\ k=1,\dots,K$. Next, we apply Proposition 5.5, Corollary 4.2 and Proposition 4.3 to conclude that h is C^{1+} along the weakest leaves $W_1^u(x)$. Therefore the assumption of the Lemma 6.1 is satisfied with $i=1$, and we conclude that the second weakest leaves are preserved $h(W_2^u(x)) = \bar{W}_2^u(h(x))$. We next apply Lemma 5.1, Proposition 5.5, Corollary 4.2 and Proposition 4.3 to conclude that h is C^{1+} along W_2^u. By Journé's theorem 5.6, we get that h is C^{1+} along the leaves $W^u_{\leq 2}$.

Performing induction in i in Lemma 6.1, we conclude h is C^{1+} along each unstable leaves $W^u(x),\ x\in\mathbb{T}^N$. Similarly, we prove that h is also C^{1+} along each stable leaves $W^s(x),\ x\in\mathbb{T}^N$. Hence by Journé's theorem 5.6, we have that $h\in C^{1+}$. □

6.3. **Alternative assumptions.** In this section, we discuss possible alternative assumptions for Theorem 1.8. Our technique developed in Section 4 relies on the existence of foliations by one dimensional leaves that is invariant under the Abelian group action. To prove the theorems in this paper, the foliation is provided by the Anosov diffeomorphism. The foliations being invariant under the Abelian group action follows from the common conjugacy h. In other words we need that the leaves (straight lines) of the invariant foliation of the toral automorphism $\bar{A}$ are mapped to the leaves of the invariant foliations of A by the conjugacy h^{-1} (Proposition 4.6). This is true when $N=2$ or in higher dimensions when we assume that A is C^1 close to $\bar{A}$. There are also circumstances Proposition 4.6 can be proved without the C^1 smallness assuption. We mention here mainly two cases.

In [G1], the author assumed simple Mather spectrum of the Anosov diffeomorphism A, in each of the connected component of the Mather spectrum, there lies exactly one eigenvalue of $\bar{A}$. Moreover it is assumed that the invariant distribution $E_i^{u,s}$ always form an angle that is less than $\pi/2$ with the corresponding distributions $\bar{E}_i^{u,s}$ in the linear case $\bar{A}$. This assumption guarantees certain quasi-isometry of the stable and unstable leaves, which is sufficient to prove Proposition 4.6.

In [FPS], the authors proved dynamical coherence and leaf conjugacy assuming that A is isotopic to $\bar{A}$ along a path of Anosov diffeomorphisms with simple Mather spectrum (see [FPS] for the terminologies and more details).

7. Perspectives

The action we consider here has both elliptic and hyperbolic dynamics. In the local and global rigidity theorems, the roles played by the hyperbolic and elliptic dynamics are different. In the local rigidity theorem, the elliptic dynamics provides the conjugacy and its smoothness while the hyperbolic dynamics put restrictions on the joint Diophantine condition. In the global rigidity theorem, the hyperbolic dynamics provides a non smooth conjugacy while the elliptic dynamics improves the regularity with the help of the invariant distribution provided by the hyperbolic dynamics. In a broad sense, the hyperbolic systems enjoy good geometric and topological structures such as the invariant foliations, while the elliptic systems enjoy good regularities. Our result is a case in which two different types of dynamics collaborate to achieve better results than one can expect from either of them. We discuss the outreach of our result from both the elliptic and hyperbolic viewpoint in this section.

7.1. **Towards a higher dimensional global KAM theory.** One dimensional orientation preserving circle maps were first studied by Poincaré, who introduced the important invariant of the rotation number. Later Denjoy showed topological conjugacy if the map $f : \mathbb{T}^1 \to \mathbb{T}^1$ has irrational rotation number and Df is of bounded variation. After the birth of the Kolmogorov-Arnold-Moser (KAM) theory, which asserts the persistence of invariant tori with Diophantine rotation vectors under small perturbations in nearly integrable Hamiltonian systems, Arnold considered the circle map as a simple model for the KAM method. The KAM method is essentially perturbative. The nonperturbative global theory for circle maps was initiated by Herman [H]. This theory was later developed by Katznelson-Ornstein, Khanin, Sinai, Teplinsky, Yoccoz [Y, KO, KS, KT, SK] *etc.* and achieved its current state of the art. Namely, for $f : \mathbb{T}^1 \to \mathbb{T}^1$, a C^r orientation preserving diffeomorphism with rotation number ρ such that $(1, \rho) \in \mathrm{DC}(2, C, \tau)$, and $r > \tau + 1$, then the conjugacy h is $C^{r-\tau-\varepsilon}$ for every small $\varepsilon > 0$. Moser [M] also studied the problem of local rigidity of Abelian groups actions on $\mathbb{T}^1$ by circle maps having rotation numbers satisfying certain simultaneous Diophantine conditions (see B.2), which are in general weaker than the Diophantine condition. The corresponding global rigidity problem was solved by Fayad and Khanin [FK] in the C^∞ and the analytic category. In the very low regularity setting, Kra [K] showed that a global theory does not exist by constructing commuting diffeomorphisms in $C^{2-\varepsilon}$, for any given set of rotation numbers, that are not conjugate to rotations.

We may ask the possible higher dimensional generalization of such a global KAM theory. The main difficulty is the lack of a circle order in the higher dimensional case. Since the Denjoy theorem does not hold when dimension $N > 1$, i.e. there is no a priori given topological conjugacy, we may put it as an assumption. We learned from K. Khanin the following open problem.

Problem: *Suppose a C^r, $r > N+1$ torus diffeomorphism $f : \mathbb{T}^N \to \mathbb{T}^N$ is topologically conjugate to a translation $\bar{T} : \mathbb{T}^N \to \mathbb{T}^N$ with $\bar{T}(x) = x + \rho \mod \mathbb{Z}^N$. Assume that ρ satisfies*

$$|\langle(1,\rho),k\rangle| \geq \frac{C}{|k|^\tau}, \quad \forall\, k \in \mathbb{Z}^{N+1} \setminus \{0\}.$$

with $N < \tau < r-1$. Is it true that the conjugacy is in $C^{r-\tau-\varepsilon}$ for any small $\varepsilon > 0$?

Parallel to Moser and Fayad-Khanin's theorems, it is also natural to ask a similar problem on the global rigidity of Abelian group actions on $\mathbb{T}^N$. Our result is a special case where the problem is partially answered. The new idea that we introduce to this problem is that the dimension can be reduced with the help of the invariant distributions.

7.2. **The periodic data rigidity of Anosov diffeomorphisms.** It has been known since Anosov that the Anosov diffeomorphisms are structurally stable. Consider for instance $\bar{A} \in \mathrm{SL}(N,\mathbb{Z})$ hyperbolic, and A a $C^r, r \geq 1$, small perturbation of $\bar{A}$. Then there exists a homeomorphism h, such that $h \circ A = \bar{A} \circ h$. The conjugacy h is known to be Hölder but is in general not C^1. If h is C^1, then we have $Dh(x)DA^q(x) = \bar{A}^q Dh(x)$ at every q-periodic point x of A. We say that A has the same periodic data as $\bar{A}$ if $DA^q(x)$ is similar to $\bar{A}^q$ at every q-periodic point x of A. It is a natural question if the periodic data is the only obstruction for h to be C^1 or of higher regularity. The question is answered affirmatively in the two dimensional case by Theorem 5.7 of de la Llave *et al.*

For $N \geq 3$, it was shown by Gogolev [G1] and Gogolev-Kalinin-Sadovskaya [GKS] that $h \in C^{1+}$ if f and g are C^1 close and have the same periodic data under some further assumptions. In the case $N = 3$, the regularity of h was first shown to be C^{1+} in [GG] and later improved to C^{r-2} in [G2], for C^1 small perturbations of linear Anosov toral automorphisms. In $N = 4$, de la Llave [Ll] constructed an example in which the periodic data of f and g coincide but the conjugacy is only Hölder. The rigidity problem of our solvable group action shares many things in common with the periodic data rigidity problem for the Anosov diffeomorphisms. In particular, the elliptic dynamics effectively guarantees the coincidence of the periodic data.

7.3. **Rigidity of solvable group actions.** Much work has been devoted to study the rigidity of higher rank Abelian group action by Anosov diffeomorphisms on torus or nil-manifolds, which was in turn motivated by Zimmer's conjecture that the standard action of $\mathrm{SL}(n,\mathbb{Z})$ on $\mathbb{T}^N$, $N > 2$, is locally rigid. See [FKS, R2, RW], etc and the references therein.

Though we work with different type of group acting by different type of diffeomorphisms using different techniques, we share some common general ideas with the previous work. The rigidity results of [FKS], and [RW] are essentially due to the coexistence of (almost) isometric (by elements close to the walls of a Weyl chamber) and hyperbolic behavior in the actions. The exponential mixing property of Anosov diffeomorphisms played a crucial role in their arguments. See also [GS]. In our case, the elliptic dynamics acts isometrically. However, it is not mixing. Effectively, introducing more translations ($K\ \mathbb{Z}^N$ factors in the action) makes the elliptic dynamics look "more mixing" (Theorem 5.3).

Appendix A. Proof of Theorem 5.3

We[1] only need to consider matrices $M = (m_{ij}),\ m_{i,j} \in \mathbb{T},\ i = 1, \dots, N$ and $j = 1, \dots, K$. So $\mathcal{M}_{N\times K}(\mathbb{T})$ is identified with $\mathbb{T}^{N\times K}$ endowed with Lebesgue measure.

Fix a smooth function $\psi \in C^\infty(\mathbb{R})$ with $\mathrm{supp}\psi \subset (-1,1)$, $\psi \geq 0$ and $\int \psi = 1$. Let ε be fixed. We next introduce $r_n = n^\varepsilon n^{-K/N}$, $n \in \mathbb{N}$, and put $\Psi_n(x) = r_n^{-N} \prod_{i=1}^N \psi(x_i/r_n)$ and consider the periodic function $\Phi_{n,y}(x) = \sum_{q\in\mathbb{Z}^N} \Psi_n(x-q-y)$ for each $y \in \mathbb{T}^N$. Then we claim that

Given $\varepsilon > 0$, there exists a $\delta > 0$ such that for each $n \in \mathbb{N}$, there exists a set $\mathcal{U}_n \subset \mathcal{M}_{N\times K}(\mathbb{T})$ with $Leb(\mathcal{U}_n) < n^{-\delta}$, and for each $M \notin \mathcal{U}_n$ and any $y \in \mathbb{T}^N$, there exists $p \in \mathbb{Z}^K$ with $\|p\| \leq n$ and $\Phi_{n,y}(Mp) \neq 0$.

Assuming the claim, considering $n = 2^\ell, \ell \in \mathbb{N}$ and using Borel-Cantelli, we get $Leb(\limsup_n \mathcal{U}_n) = 0$. This means that the probability for M lying in infinitely many $\mathcal{U}_n$ is zero. This completes the proof of the theorem.

It remains to prove the claim. Decompose $\Phi_{n,y}(x)$ into Fourier series $\Phi_{n,y}(x) = \sum_{k\in\mathbb{Z}^N} c_k(n,y)e^{2\pi i\langle k,x\rangle}$. Notice that for each x, there is only one $q \in \mathbb{Z}^N$ such that $x - q - y \in (-1,1)^N$. It follows that $\|c_k(n,y)\| \leq 1$ for all $k \in \mathbb{Z}^N$ and $c_0(n,y) = 1$ is independent of n, y. Moreover, for each $\ell > 0$, there exists C_ℓ (depending only on ψ) such that $\|c_k(n,y)\| \leq C_\ell r_n^{-N-\ell}/\|k\|^\ell$ due to the C^∞ smoothness of ψ. Next for any matrix $M \in \mathcal{M}_{N\times K}(\mathbb{T})$ write

$$S_n(M,k) := \sum_{\|p\|_\infty \leq n} e^{2\pi i\langle k,Mp\rangle}, \quad \Lambda_{n,y}(M) := \sum_{\|p\|_\infty \leq n} \Phi_{n,y}(Mp) = \sum_{k\in\mathbb{Z}^N} c_k(n,y)S_n(M,k).$$

[1]The proof was communicated to us by the user Fedja on MathOverflow, http://mathoverflow.net/questions/227817/a-quantitative-kronecker-theorem.

We get $S_n(M,0)=n^K$ and for $\beta>0$ to be determined later

$$\left|\sum_{\|k\|\ge r_n^{-\beta}} c_k(n,y)S_n(M,k)\right|\le n^K C_\ell r_n^{\beta\ell-N-\ell}.$$

It remains to investigate the sum

$$\Gamma_{n,r_n}(M):=\sum_{0<\|k\|_\infty\le r_n^{-\beta}}|S_n(M,k)|=\sum_{0<\|k\|_\infty\le r_n^{-\beta}}\left|\sum_{\|p\|_\infty\le n} e^{2\pi i\langle z,p\rangle}\right|,\quad z=M^tk.$$

We use the fact that $|\sum_{\|p\|_\infty\le n} e^{2\pi i\langle z,p\rangle}|\le C\prod_{j=1}^K(\min\{n,\|z_j\|^{-1}\})$, where $\|z_j\|$ is the distance from z_j to the nearest integer. Consider a map $F_k:\ \mathcal{M}_{N\times K}(\mathbb{T})\to\mathbb{T}^K$ via $F_k(M)=M^tk$, mod $\mathbb{Z}^K$, then F_k pushes forward the Lebesgue measure on $\mathcal{M}_{N\times K}(\mathbb{T})$ to a Lebesgue measure on $\mathbb{T}^K$. We immediately get that

$$\int_{\mathcal{M}_{N\times K}(\mathbb{T})}|S_n(M,k)|\,d\mathrm{Leb}\le\int_{\mathbb{T}^K}C\prod_{j=1}^K(\min\{n,\|z_j\|^{-1}\})\,dz\le C\log^K n,$$

so there exists a set $\mathcal{U}_n\subset\mathcal{M}_{N\times K}(\mathbb{T})$ with $\mathrm{Leb}(\mathcal{U}_n)\le n^{-\delta}$ such that we have

$$\Gamma_{n,r_n}(M)\le Cn^\delta r_n^{-\beta N}\log^K n,\quad \forall\ M\in\mathcal{M}_{N\times K}(\mathbb{T})\setminus\mathcal{U}_n.$$

Note that this set $\mathcal{U}_n$ is independent of y since $\Gamma_{n,r_n}(M)$ is. Now we get

$$|\Lambda_{n,y}(M)|\ge n^K-n^KC_\ell r_n^{\beta\ell-N-\ell}-Cn^\delta r_n^{-\beta N}\log^K n,\quad\forall\ M\in\mathcal{M}_{N\times K}(\mathbb{T})\setminus\mathcal{U}_n,\ \forall y\in\mathbb{T}^N.$$

We choose $r_n=n^\varepsilon n^{-K/N}$, and $\beta(>1)$ and $\delta(>0)$ sufficiently close to 1 and 0 respectively to satisfy the inequality $(\beta-1)K+2\delta<\beta N\varepsilon$ for given ε, and choose ℓ large enough to satisfy $(\beta-1)\ell>N$. Hence $|\Lambda_{n,y}(M)|\ge\frac{1}{2}n^K$. This completes the proof of the claim hence the theorem. □

Appendix B. The simultaneous Diophantine condition

In this section, we discuss the assumption in Theorem 1.4 on $\bar{A}$ and $\boldsymbol{\rho}$. We want to show how to get the assumption (1.3) and the Diophantine assumption $\boldsymbol{\rho}$ satisfied simultanously.

Given $\bar{A}$, we solve equation (1.3) for $\boldsymbol{\rho}$. Lifting (1.3) to $\mathbb{R}^N$, we get the following equation

$$\boldsymbol{\rho}\bar{A}=\bar{A}\boldsymbol{\rho}+\mathbf{P},\quad \mathbf{P}\in\mathbb{Z}^{N\times N}. \tag{B.1}$$

As usual, we first set $\mathbf{P}=0$ and consider the homogeneous equation. We recall a fact and definition from linear algebra:

Lemma B.1 (Corollary 4.1.4 of [HJ])**.** *Let $A \in M_N(\mathbb{R})$ where $M_N(\mathbb{R})$ is the set of $N \times N$ matrices with entries in $\mathbb{R}$. The set of matrices in $M_N(\mathbb{R})$ that commute with A is a subspace of $M_N(\mathbb{R})$ with dimension at least N. The dimension is equal to N if and only if M is non-derogatory, i.e. each eigenvalue of A has geometric multiplicity exactly 1. Thus if A is nonderogary, the centralizer $Z(A)$ of A is*

$$Z(\bar{A}) = \operatorname{span}_{\mathbb{R}}\{\mathrm{id}, A, \dots, A^{N-1}\}$$

If $\bar{A}$ is non-derogatory, then for any $\boldsymbol{\rho}$ satisfying $\boldsymbol{\rho}\bar{A} = \bar{A}\boldsymbol{\rho}$, we can thus write each $\boldsymbol{\rho} \in Z(\bar{A})$ as a linear combination $\boldsymbol{\rho} = \sum_{i=1}^{N} a_i A^{i-1}$, where $a = (a_1, \dots, a_N) \in \mathbb{R}^N$.

Lemma B.2. *Let $\boldsymbol{\rho} = \sum a_i \bar{A}^{i-1}$ for some $a = (a_1, \dots, a_N)$ and $\bar{A}^{i-1} \in \mathrm{SL}(N, \mathbb{Z})$. Suppose the nonvanishing entries of a form a vector $a' \in \mathbb{R}^k$, $1 \le k \le N$ satisfying the Diophantine condition: there exist $C, \tau > 0$ such that*

$$|\langle a', m\rangle| \ge \frac{C}{|m|^{\tau}}, \quad \forall\, m \in \mathbb{Z}^k \setminus \{0\}.$$

Then the columns of $\boldsymbol{\rho}$, denoted by $\rho_1, \dots, \rho_N$, satisfy the simultaneous Diophantine condition for some $C' > 0$, i.e.

$$\max_{1\le j\le N}\{|\langle n, \rho_j\rangle|\} \ge \frac{C'}{\|n\|^{\tau}}, \quad \forall\, n \in \mathbb{Z}^N \setminus \{0\}. \tag{B.2}$$

Proof. Denote by v_j^i is the j-th column of $\bar{A}^{i-1}$, $i, j = 1, \dots, N$, and by ρ_j the j-th column of $\boldsymbol{\rho}$. Hence we have $\rho_j = \sum_i a_i v_j^i$.

Assume the vector formed by the non vanishing entries of a satisfies the Diophantine condition and denote by $\mathcal{I}$ the set of indices of the non vanishing entries of the vector a, then we have for each j

$$\begin{aligned} |\langle n, \rho_j\rangle| = \left|\sum_{i=1}^{N} a_i \langle n, v_j^i\rangle\right| &\ge \frac{C}{(\sum_{i\in\mathcal{I}} |\langle n, v_j^i\rangle|)^{\tau}} \\ &= \frac{C}{\|n\|^{\tau}(\sum_{i\in\mathcal{I}}\left|\langle \frac{n}{\|n\|}, v_j^i\rangle\right|)^{\tau}} \\ &\ge \frac{C}{\|n\|^{\tau}(\sum_{i\in\mathcal{I}}\|v_j^i\|)^{\tau}} \end{aligned} \tag{B.3}$$

if $\langle n, v_j^i\rangle \ne 0$ for some $i \in \mathcal{I}$.

To show that the simultaneous Diophantine condition holds for $\rho_1, \dots, \rho_N$, it remains to show that for each $u \in \mathbb{S}^{N-1}$, there exist $i \in \mathcal{I}$, $j \in \{1, 2, \dots, N\}$, such that $\langle u, v_j^i\rangle \ne 0$. This follows from the non-degeneracy of $\bar{A}$. We fix any $i \in \mathcal{I}$, then v_j^i, $j = 1, 2, \dots, N$, form the matrix $\bar{A}^{i-1}$ which is non-degenerate. Hence the vectors v_j^i, $j = 1, 2, \dots, N$, are linearly independent. The compactness of $\mathbb{S}^{N-1}$ implies that there does not exists $u \in \mathbb{S}^{N-1}$ that is simultaneously orthogonal to all of $v_j^i,\ j = 1, 2, \ldots, N$. $\square$

Next, in order to solve the inhomogeneous equation $\boldsymbol{\rho}\bar{A} = \bar{A}\boldsymbol{\rho} + \mathbf{P}$, it is enough produce a particular solution for given $\mathbf{P} \in \mathbb{Z}^{N\times N}$ in addition to the general solutions to the homogeneous equation. Note that the (B.1) might not be solvable for some $\mathbf{P}$. We have the following result.

Theorem B.3 (Theorem 4.2.22 of [HJ])**.** *Given matrices $A, B, C \in M_N(\mathbb{R})$. Then there exists some $X \in M_N(\mathbb{R})$ solving the equation $AX - XB = C$ if and only if the matrices $\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$ and $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ are similar.*

So to solve (B.1), the necessary and sufficient condition is the similarity of the matrices $\begin{bmatrix} \bar{A} & \mathbf{P} \\ 0 & \bar{A} \end{bmatrix}$ and $\begin{bmatrix} \bar{A} & 0 \\ 0 & \bar{A} \end{bmatrix}$. Given a particular solution $\boldsymbol{\rho}_*(\mathbf{P})$ of (B.1). If $\bar{A}$ is nonderogatory, then the general solution of (B.1) can be written as

$$\boldsymbol{\rho} = \sum a_i \bar{A}^{i-1} + \boldsymbol{\rho}_*$$

for some $a = (a_1, \ldots, a_N)$, if $\boldsymbol{\rho}_*$ happens to be rational, then $\boldsymbol{\rho}$ is simultaneously Diophantine, if the nonvanishing entries of a form a Diophantine vector.

Appendix C. Affine action and vanishing Lyapunov exponents

In this appendix, we prove the results in Section 1.1. Proposition 1.1 and 1.2 are straightforward from the group relation. We prove first Proposition 1.3.

Proof of Proposition 1.3. Suppose we have two affine actions $\bar{\alpha} = \bar{\alpha}(\bar{A}, \boldsymbol{\rho})$ and $\bar{\alpha}' = \bar{\alpha}(\bar{A}, \boldsymbol{\rho}')$ conjugate by a homeomorphism h of the form $h(x) = x + \tilde{h}(x),\ x \in \mathbb{T}^N$ where $\tilde{h}$ is $\mathbb{Z}^N$-periodic. We want to show that $\boldsymbol{\rho} = \boldsymbol{\rho}'$. Denote by ρ_j and ρ_j' the j-th column of $\boldsymbol{\rho}$ and $\boldsymbol{\rho}'$ respectively. We have

$$h(x + \rho_j) = h(x) + \rho_j', \quad j = 1, \ldots, N.$$

This is equivalent to

$$\rho_j + \tilde{h}(x + \rho_j) = \tilde{h}(x) + \rho_j', \quad j = 1, \ldots, N.$$

Integrating over $\mathbb{T}^N$, we get that $\int_{\mathbb{T}^N} \tilde{h}(x + \rho_j)\, dx = \int_{\mathbb{T}^N} \tilde{h}(x)\, dx$, hence $\rho_j = \rho_j'$. $\square$

Proposition C.1. *Suppose $N = 2$ and $\bar{A}$ has real simple spectrum. Given an action $\alpha : \Gamma_{\bar{A}} \to \mathrm{Diff}^r(\mathbb{T}^2),\ r > 1$, let μ be an ergodic measure of the Abelian subaction $\beta : \mathbb{Z}^2 \to \mathrm{Diff}^r(\mathbb{T}^2)$ generated by $\alpha(g_1)$ and $\alpha(g_2)$. Then with respect to μ, all the Lyapunov exponents of the Abelian subaction β is zero.*

We first cite the following result on the Weyl chamber of the $\mathbb{Z}^N$ actions.

Theorem C.2 (Proposition 2.1 of [FKS])**.** *Suppose μ is an ergodic measure for the action $\beta: \mathbb{Z}^N \to \mathrm{Diff}^r(M)$, $r > 1$. Then there are finitely many linear functionals $\chi: \mathbb{Z}^N \to \mathbb{R}$, a set $\mathcal{P}$ of full measure and a β-invariant measurable splitting of the tangent bundle $T_xM = \oplus E_\chi(x)$, $x \in \mathcal{P}$ such that for all $a \in \mathbb{Z}^N$ and $v \in E_\chi$, the Lyapunov exponent of v is*

$$\lim_{n\to\pm\infty} n^{-1} \log \|D\beta(a^n)(v)\| = \chi(a).$$

Proof of Proposition C.1. Consider an action $\alpha: \Gamma_{\bar{A}} \to \mathrm{Diff}^r(\mathbb{T}^2)$ with $\alpha(g_0)(x) = A(x)$ and $\alpha(g_i)(x) = T_i$, $i = 1, 2$, $x \in \mathbb{T}^2$. Denote by β the Abelian subaction generated by T_1 and T_2 and suppose μ is an ergodic invariant measure of β.

We apply the group relation $A^n T^{\mathbf{p}} A^{-n} = T^{(\bar{A}^t)^n \mathbf{p}}$, $n \in \mathbb{Z}$, $\mathbf{p} \in \mathbb{Z}^2$. For given χ and $v \in E_\chi(x)$ given by Theorem C.2 for the subaction β, we consider

$$\frac{1}{\|\bar{A}^n\|} \log \|D(A^n T^{\mathbf{p}} A^{-n})v\| = \frac{1}{\|\bar{A}^n\|} \log \|D(T^{(\bar{A}^t)^n \mathbf{p}})v\|.$$

Since $\|\bar{A}^n\|$ grows exponentially, the LHS goes to zero as $n \to \pm\infty$. We denote by u_i an eigenvector of the eigenvalue λ_i of $\bar{A}^t$ and decompose $\mathbf{p} = \sum_i p_i u_i$. Since u_i has irrational slope and $\mathbf{p} \in \mathbb{Z}^2 \setminus \{0\}$, the coefficients $p_i \neq 0$, $i = 1, 2$, and we get $(\bar{A}^t)^n \mathbf{p} = \sum \lambda_i^n p_i u_i$. It is not hard to get the following from the RHS using Theorem C.2

$$p_i \chi(u_1) = 0, \text{ for } n \to \infty, \quad \text{and } p_2 \chi(u_2) = 0, \text{ for } n \to -\infty.$$

Since u_1 and u_2 are linearly independent, the only linear functional that is simultaneously orthogonal to both u_1 and u_2 is 0, so $\chi = 0$. $\square$

To see the naturality of the above affine representation, we consider the following space of special affine representations

$$\mathrm{aff}(\Gamma_{\bar{A}}) := \{\alpha: \Gamma_{\bar{A}} \to \mathrm{Diff}^r(\mathbb{T}^N) \mid \alpha(g_0)(x) = \bar{A}x,\ \alpha(g_i)(x) = B_i x + \rho_i,\ i = 1, \dots, N,$$
$$\text{where } B_i \in \mathrm{SL}(N, \mathbb{Z}), \text{ and } \rho_i \in \mathbb{T}^N,\ i = 1, \dots, N, \text{ linearly independent over } \mathbb{Z}\}.$$

Proposition C.3. *Suppose $N \geq 2$ and $\bar{A} \in \mathrm{SL}(N, \mathbb{Z})$ has no eigenvalue 1. For an affine action α in $\mathrm{aff}(\Gamma_{\bar{A}})$, all the eigenvalues of B_i, $i = 1, \dots, N$, has modulus 1.*

Proposition C.4. *In dimension $N = 2$, suppose $\bar{A}$ has simple real eigenvalues. Then any affine representation α in $\mathrm{aff}(\Gamma_{\bar{A}})$ has the following form*

$$\alpha(g_0) = \bar{A}, \quad \alpha(g_i): x \mapsto x + \rho_i,\ i = 1, 2,$$

where ρ_1, ρ_2 satisfy (1.3).

Proof of Proposition C.3. Suppose $\alpha \in \mathrm{aff}(\Gamma_{\bar{A}})$, then we have $\alpha(g_0)x = \bar{A}x$ and $\alpha(g_i)x = B_i x + \rho_i,\ i = 1, \ldots, N$, where B_i is a matrix in $\mathrm{SL}(N, \mathbb{Z})$ and ρ_i is a vector in $\mathbb{T}^N$. Differentiating the group relation we get

$$\bar{A}B^{\mathbf{p}} = B^{\bar{A}^t\mathbf{p}}\bar{A}, \quad B_iB_j = B_jB_i.$$

Since $\Gamma_{\bar{A}}$ is a solvable group, there exists a nondegenerate matrix $C \in \mathrm{GL}(N, \mathbb{R})$ such that $C^{-1}B_iC$ and $C^{-1}\bar{A}C$ are simultaneously of block upper triangular form in $\mathrm{GL}(N, \mathbb{R})$ (see [Mr]), where each block on the diagonal is a scalar multiple of an orthogonal matrix. Denote by $\{b_{i,j},\ j = 1, \ldots, k\}$ and $\{a_j,\ j = 1, \ldots, k\}$ the diagonal block of $C^{-1}B_iC$ and $C^{-1}\bar{A}C$ respectively for some $1 \le k \le N$ and it is also known that for the each j, the matrices a_j and $b_{i,j}$ are of the same size $n_j \times n_j$ with $\sum_{j=1}^k n_j = N$. We also denote by $b_j^{\mathbf{p}} = \prod_i b_{i,j}^{p_i}$ and by $\lambda(a_j)$ and $\lambda(b_j^{\mathbf{p}}) = \prod_i \lambda(b_{i,j})^{p_i}$ the modulus of the scalar part of a_j and $b_j^{\mathbf{p}}$ respectively.

The group relation

$$(C^{-1}\bar{A}C)(C^{-1}BC)^{\mathbf{p}} = (C^{-1}BC)^{\bar{A}^t\mathbf{p}}(C^{-1}\bar{A}C)$$

gives the following on the diagonal

$$a_jb_j^{\mathbf{p}} = b_j^{\bar{A}^t\mathbf{p}}a_j, \quad \text{and} \quad \lambda(a_j)\lambda(b_j^{\mathbf{p}}) = \lambda(b_j^{\bar{A}^t\mathbf{p}})\lambda(a_j).$$

After cancelling $\lambda(a_j)$, taking log and taking $\mathbf{p} = (1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)$, we get

$$\mathsf{b} = \bar{A}^t\mathsf{b}, \quad \text{i.e.} \quad (\bar{A}^t - \mathrm{Id})\mathsf{b} = 0,$$

where b is the matrix $(\log\lambda(b_{i,j}))$. Since $\bar{A}^t$ does not have eigenvalue 1, the only solution is $\mathsf{b} = 0$ so all the $\lambda(b_{i,j})$ is 1. This completes the proof. $\square$

We next focus on the $N = 2$ case and show that each B_i is in fact identity.

Proof of Proposition C.4. We consider only the case $N = 2$. From the commutativity of $B_i \in \mathrm{SL}(2, \mathbb{Z}),\ i = 1, 2$, in the proof of Lemma C.3, we get that B_1 and B_2 are simultaneously put into Jordan normal form. We know the eigenvalues have to be of modulus 1 by Lemma C.3. Next, the eigenvalues cannot be imaginary, since otherwise by the same result of [Mr] used in Lemma C.3, the blocks in the upper-triangular form have to be 2×2 (in fact, not upper-triangular), but the matrix $\bar{A}$ with simple real spectrum cannot be conjugate to a scalar multiple of an orthogonal matrix. So we get that the Jordan normal form of B_i can only be $\pm\mathrm{Id}$ or $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}$. Our goal is to show that $B_1 = B_2 = \mathrm{Id}$. First it is easy to check that the map $x \mapsto -x + \rho_1$ does not commute with $x \mapsto \pm x + \rho_2$ under the assumption that ρ_1, ρ_2 are linearly independent over $\mathbb{Z}$, so neither B_1 nor B_2 can be $-\mathrm{Id}$. Next, we consider only the case when both B_1 and B_2 are conjugate to $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. All the other cases are either trivial or similar.

Suppose the matrix $C \in \mathrm{GL}(2,\mathbb{C})$ is such that $C^{-1}B_1C = C^{-1}B_2C = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $C^{-1}\bar{A}^nC = \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix}$. We next choose $\mathbf{p}_k = (p_{k,1}, p_{k,2})$ such that $\frac{\mathbf{p}_k}{\|\mathbf{p}_k\|}$ converges to an eigenvector of the larger modulus eigenvalue λ of $\bar{A}^t$. Since $\frac{\mathbf{p}_k}{\|\mathbf{p}_k\|}$ tends to an irrational direction as $k \to \infty$, we have $(p_{k,1} + p_{k,2})/\|\mathbf{p}_k\|$ is bounded away from zero.

Then from the group relation

$$(C^{-1}\bar{A}^nC)(C^{-1}BC)^{\mathbf{p}} = (C^{-1}BC)^{(\bar{A}^t)^n\mathbf{p}}(C^{-1}\bar{A}^nC)$$

we get

$$\begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix}\begin{bmatrix} 0 & p_{k,1}+p_{k,2} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \lambda^n(p_{k,1}+p_{k,2}) + \lambda^n o_k(\|\mathbf{p}_k\|) \\ 0 & 0 \end{bmatrix}\begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix}.$$

We get that

$$\begin{bmatrix} 0 & (p_{k,1}+p_{k,2})a_n \\ 0 & (p_{k,1}+p_{k,2})c_n \end{bmatrix} = (\lambda^n(p_{k,1}+p_{k,2}) + \lambda^n o_k(\|\mathbf{p}_k\|))\begin{bmatrix} c_n & d_n \\ 0 & 0 \end{bmatrix}.$$

We get immediately $c_n = 0$. The first possibility is that $c_n = 0$, $a_n = \lambda^n$ and $d_n = \lambda^{-n}$ and $(p_{k,1}+p_{k,2})\lambda^n = (p_{k,1}+p_{k,2})+o_k(\|\mathbf{p}_k\|)$, which is impossible since $(p_{k,1}+p_{k,2})/\|\mathbf{p}_k\|$ is bounded away from zero. Similarly the possibility $c_n = 0$, $a_n = \lambda^{-n}$ and $d_n = \lambda^n$ is also excluded. The contradiction excludes the possibility that both B_1 and B_2 are conjugate to $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. All other possibilities can be analyzed and excluded similarly and finally we conclude that $B_1 = B_2 = \mathrm{Id}$. □

Acknowledgment

A.W. is supported by NSF grant DMS-1316534. J. X. is supported by NSF grant DMS-1500897. We would like to thank K. Khanin and Pengfei Zhang for the helpful discussions.

References

[A1] Asaoka, Masayuki. *Rigidity of certain solvable actions on the sphere.* Geometry & Topology 16.3 (2012): 1835-1857.

[A2] Asaoka, Masayuki. *Rigidity of certain solvable actions on the torus.* (2014).

[AB] Aliprantis, Charalambos D., and Kim Border. *Infinite dimensional analysis: a hitchhiker's guide.* Springer Science & Business Media, 2006.

[ADH] Auffinger, Antonio, Michael Damron, and Jack Hanson. *50 years of first passage percolation.* arXiv preprint arXiv:1511.03262 (2015).

[BBI] Burago, Dmitri, Yuri Burago, and Sergei Ivanov. *A course in metric geometry.* Vol. 33. Providence: American Mathematical Society, 2001.

[BW] Burslem, Lizzie, and Amie Wilkinson. *Global rigidity of solvable group actions on* $\mathbb{S}^1$. Geometry & Topology 8.2 (2004): 877-924.

[DF] Damjanovic, Danijela, and Bassam Fayad. *KAM rigidity for partially hyperbolic affine $\mathbb{Z}^k$ actions on the torus with a rank one factor.* arXiv preprint arXiv:1308.3546 (2013).

[DK] Damjanović, Danijela, and Anatole Katok. *Local rigidity of partially hyperbolic actions I. KAM method and $\mathbb{Z}^k$ actions on the torus.* Annals of mathematics (2010): 1805-1858.

[EF] Eskin, Alex, and David Fisher. *Quasi-isometric rigidity of solvable groups.* Proceedings of the International Congress of Mathematicians. Vol. 3. 2010.

[EFW1] Eskin, Alex, David Fisher, and Kevin Whyte. *Coarse differentiation of quasi-isometries I: Spaces not quasi-isometric to Cayley graphs.* Annals of Mathematics 176.1 (2012): 221-260.

[EFW2] Alex Eskin, David Fisher, Kevin Whyte, *Coarse differentiation of quasi-isometries II: Rigidity for Sol and Lamplighter groups,* Annals of Mathematics, Volume 177, Issue 3, (2013): 869-910

[FK] Fayad, Bassam, and Kostantin Khanin. *Smooth linearization of commuting circle diffeomorphisms.* Annals of Mathematics (2009): 961-980.

[Fr] Franks, John. *Anosov diffeomorphisms on tori.* Transactions of the American Mathematical Society 145 (1969): 117-124.

[Fi] Fisher, David. *Local rigidity of group actions: past, present, future.* arXiv preprint math/0507455 (2005).

[FKS] Fisher, David, Boris Kalinin, and Ralf Spatzier. *Global rigidity of higher rank Anosov actions on tori and nilmanifolds.* JAMS 26.1 (2013): 167-198.

[FM1] Farb, Benson, and Lee Mosher. *On the asymptotic geometry of abelian-by-cyclic groups.* Acta Mathematica 184.2 (2000): 145-202.

[FM2] Farb, Benson, and Lee Mosher. *Problems on the geometry of finitely generated solvable groups.* Crystallographic groups and their generalizations (Kortrijk, 1999), 121-134, Contemp. Math 262.

[FPS] Fisher, Todd, Rafael Potrie, and Martin Sambarino. *Dynamical coherence of partially hyperbolic diffeomorphisms of tori isotopic to Anosov.* Mathematische Zeitschrift 278.1-2 (2014): 149-168.

[G1] Gogolev, Andrey. *Smooth conjugacy of Anosov diffeomorphisms on higher dimensional tori.* Journal of Modern Dynamics, 2, no. 4, 645-700 (2008).

[G2] Gogolev, Andrey. *Bootstrap for local rigidity of Anosov automorphisms on the 3-torus.* arXiv preprint arXiv:1407.7771 (2014).

[Gh] Ghys, Étienne. *Sur les groupes engendrés par des difféomorphismes proches de l'identité.* Boletim da Sociedade Brasileira de Matemática-Bulletin/Brazilian Mathematical Society 24.2 (1993): 137-178.

[GG] Gogolev, Andrey, and Misha Guysinsky.*C^1 Differentiable Conjugacy of Anosov Diffeomorphisms on Three Dimensional Torus.* DCDS-A, 22, no. 1/2, 183-200 (2008).

[GKS] Gogolev, Andrey, Boris Kalinin, and Victoria Sadovskaya. *Local rigidity for Anosov automorphisms,* (appendix by Rafael de la Llave) Mathematical Research Letters, 18, no. 5, 843-858 (2011)

[GS] Gorodnik, Alexander, and Ralf Spatzier. *Mixing properties of commuting nilmanifold automorphisms.* Acta Mathematica 215.1 (2015): 127-159.

[H] Herman, Michael Robert. *Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations.* Publications Mathématiques de l'Institut des Hautes Études Scientifiques 49.1 (1979): 5-233.

[He] Hedlund, Gustav A. *Geodesics on a two-dimensional Riemannian manifold with periodic coeffcients.* Annals of Mathematics (1932): 719-739.

[HK] Katok, Anatole, and Boris Hasselblatt. *Introduction to the modern theory of dynamical systems.* Vol. 54. Cambridge university press, 1997.

[HJ] Horn, Roger A., and Charles R. Johnson. *Topics in matrix analysis.* 1991. Cambridge University Presss, Cambridge 37: 39.

[HP] Hammerlindl, Andy, and Rafael Potrie. *Classification of partially hyperbolic diffeomorphisms in 3-manifolds with solvable fundamental group.* Journal of Topology 8.3 (2015): 842-870.

[HPS] Hirsch, M. W., C. C. Pugh, and M. Shub. *Invariant manifolds,* volume 583 of Lecture notes in mathematics. (1977).

[HSW] Hu, Huyi, Enhui Shi, and Zhenqi Jenny Wang. *Some ergodic and rigidity properties of discrete Heisenberg group actions.* arXiv preprint arXiv:1405.1120 (2014).

[J] Journé, Jean-Lin. *A regularity lemma for functions of several variables.* Revista Matematica Iberoamericana 4.2 (1988): 187-194.

[Ja] Jäger, T. *Linearization of conservative toral homeomorphisms.* Inventiones mathematicae 176.3 (2009): 601-616.

[K] Kra, Bryna. *The conjugating map for commutative groups of circle diffeomorphisms.* Israel Journal of Mathematics 93.1 (1996): 303-316.

[Kl] Kleppner, Adam. *Measurable homomorphisms of locally compact groups.* Proceedings of the American Mathematical Society 106.2 (1989): 391-395.

[KO] Katznelson, Yitzhak, and Donald Ornstein. *The differentiability of the conjugation of certain diffeomorphisms of the circle.* Ergodic Theory and Dynamical Systems 9.04 (1989): 643-680.

[KS] Khanin, K. M., and Ya G. Sinai. *A new proof of M. Hermans theorem.* Communications in Mathematical Physics 112.1 (1987): 89-101.

[KT] Khanin, Konstantin, and Alexey Teplinsky. *Hermans theory revisited.* Inventiones mathematicae 178.2 (2009): 333-344.

[L] Lindenstrauss, Elon. *Pointwise theorems for amenable groups.* Inventiones mathematicae 146.2 (2001): 259-295.

[Ll] de la Llave, Rafael. *Smooth conjugacy and SRB measures for uniformly and non-uniformly hyperbolic systems.* Communications in mathematical physics 150.2 (1992): 289-320.

[LMM] de la Llave, Rafael, Marco, J. M., and R. Moriyon. *Invariants for smooth conjugacy of hyperbolic dynamical systems I-V.* Communications in Mathematical Physics 109, 112, 116 (1987, 1988).

[M] Moser, Jürgen. *On commuting circle mappings and simultaneous Diophantine approximations.* Mathematische Zeitschrift 205.1 (1990): 105-121.

[Mc] McCarthy, Anne. *Rigidity of trivial actions of Abelian-by-cyclic groups.* Proceedings of the American Mathematical Society 138.4 (2010): 1395-1403.

[Mo] Morse, Harold Marston. *A fundamental class of geodesics on any closed surface of genus greater than one.* Transactions of the American Mathematical Society 26.1 (1924): 25-60.

[Mr] Moore, Calvin C. *Amenable subgroups of semi-simple groups and proximal flows.* Israel Journal of Mathematics 34.1-2 (1979): 121-138.

[MS] McSwiggen, Patrick D. *Diffeomorphisms of the torus with wandering domains.* Proceedings of the American Mathematical Society 117.4 (1993): 1175-1186.

[N] Nevo, A. *Pointwise ergodic theorems for actions of groups.* Handbook of Dynamical Systems, vol. IB, Eds. B. Hasselblatt and A. Katok, 2006. 871-982.

[OW] Ornstein, Donald S., and Benjamin Weiss. *Entropy and isomorphism theorems for actions of amenable groups.* Journal d'Analyse Mathématique 48.1 (1987): 1-141.

[R1] Hertz, Federico Rodriguez. *Stable ergodicity of certain linear automorphisms of the torus.* Annals of mathematics (2005): 65-107.

[R2] Hertz, F. Rodriguez, *Global rigidity of certain abelian actions by toral automorphisms,* Journal of Modern Dynamics. 10/2006; 1(3).

[RW] Hertz, Federico Rodriguez, Zhiren Wang, *Global rigidity of higher rank abelian Anosov algebraic actions,* Inventiones mathematicae, October 2014, Volume 198, Issue 1, pp 165-209

[S] Shub, Michael. Global stability of dynamical systems. Springer Science & Business Media, 2013.

[SK] Sinai, Yakov, and Konstantin Khanin. *Smoothness of conjugacies of diffeomorphisms of the circle with rotations.* Russian Mathematical Surveys 44.1 (1989): 69-99.

[W] Wang, Zhenqi Jenny. *Local rigidity of higher rank non-abelian action on tours.* arXiv preprint arXiv:1405.7737 (2014).

[Y] Yoccoz, Jean-Christophe. *Analytic linearization of circle diffeomorphisms.* Dynamical systems and small divisors. Springer Berlin Heidelberg, 2002. 125-173.

Department of mathematics, the University of Chicago, Chicago, IL, US, 60637

E-mail address: wilkinso@math.uchicago.edu

Department of mathematics, the University of Chicago, Chicago, IL, US, 60637

E-mail address: jxue@math.uchicago.edu