GLOBAL RIGIDITY OF SOME ABELIAN-BY-CYCLIC GROUP ACTIONS ON $\mathbb{T}^2$

SEBASTIAN HURTADO AND JINXIN XUE

ABSTRACT. For groups of diffeomorphisms of $\mathbb{T}^2$ containing an Anosov diffeomorphism, we give a complete classification for polycyclic Abelian-by-Cyclic group actions on $\mathbb{T}^2$ up to both topological conjugacy and smooth conjugacy under mild assumptions. Along the way, we also prove a Tits alternative type theorem for some groups of diffeomorphisms of $\mathbb{T}^2$.

1. INTRODUCTION

In this article, we study the action of certain abelian-by-cyclic groups (ABC, for simplicity) on $\mathbb{T}^2$. Our groups are given as follows. Let $B \in \text{SL}_n(\mathbb{Z})$ be an invertible matrix with integer coefficients and determinant one, and let $\Gamma_B = \mathbb{Z} \ltimes_B \mathbb{Z}^n$ be the semi-direct product, where $\mathbb{Z}$ acts on $\mathbb{Z}^n$ via the linear standard action of $B$ on $\mathbb{Z}^n$.

1.1. The topological classification result. We consider the case where $n = 2$, that is $B \in \text{SL}_2(\mathbb{Z})$ and $\Gamma_B = \mathbb{Z} \ltimes_B \mathbb{Z}^2$. We define an affine action of $\Gamma_B$ on $\mathbb{T}^2$ to be a group homomorphism $\Phi : \Gamma_B \rightarrow \text{Aff}(\mathbb{T}^2)$, where $\text{Aff}(\mathbb{T}^2)$ is the group of orientation preserving affine transformations on $\mathbb{T}^2$, and

$$\Phi(B) = A \in \text{SL}_2(\mathbb{Z}), \quad \Phi(e_i) : x \mapsto x + \rho_i \mod \mathbb{Z}^2, \quad i = 1, 2,$$

(1)

where $e_1 = (1, 0), e_2 = (0, 1) \in \mathbb{Z}^2$ and $\rho_i \in \mathbb{R}^2, \quad i = 1, 2$.

In this case, we let $\rho = (\rho_1, \rho_2)$ to be the matrix having $\rho_1$ and $\rho_2$ as columns and call it the rotation matrix of $\Phi$. The group relation $\Phi(B)\Phi(v) = \Phi(Bv)\Phi(B), \forall v \in \mathbb{Z}^2$, is equivalent to

$$A\rho = \rho B, \quad \text{mod } \mathbb{Z}^{2 \times 2},$$

(2)

in other words, for the equation to define an action, there must exists a matrix $C \in \mathbb{Z}^{2 \times 2}$ such that $A\rho = \rho B + C$. As discussed in Section 5.1 (Corollary 5.1), these actions can only be faithful in the case where $|\text{tr} A| = |\text{tr} B|$. Affine actions will be classified in Section 4.2.

One of our main results is the following global rigidity result for actions of $\Gamma_B$ on $\mathbb{T}^2$ by diffeomorphisms. We denote by $\text{Diff}^r(\mathbb{T}^2)$ the group of $C^r$ diffeomorphisms on $\mathbb{T}^2$ and by $\text{Homeo}(\mathbb{T}^2)$ the group of homeomorphisms on $\mathbb{T}^2$.

**Theorem 1.1.** Suppose that $\Phi : \Gamma_B \rightarrow \text{Diff}^r(\mathbb{T}^2), \quad r > 2$, is such that:

1. $B \in \text{SL}_2(\mathbb{Z})$ is an Anosov linear map (i.e. $B$ has its eigenvalues with norm different than one);
2. $\Phi(B)$ is an Anosov diffeomorphism of $\mathbb{T}^2$ homotopic to $A \in \text{SL}_2(\mathbb{Z})$.

Then $\Phi$ is topologically conjugate to an affine action of $\Gamma_B$ up to finite index. More concretely, there exist a finite index subgroup $\Gamma' \subset \Gamma_B$ and $h \in \text{Homeo}(\mathbb{T}^2)$ such that $h\Phi(\gamma)h^{-1}$, for all $\gamma \in \Gamma'$, coincides with an affine action of the form (1.1).
Remark 1.2. When $|\text{tr}A| \geq |\text{tr}B|$, the differentiability $r > 2$ can be relaxed to $r \geq 1$. The assumption $r > 2$ is only needed in the proof of the $|\text{tr}A| < |\text{tr}B|$ case where we apply the Herman-Yoccoz theory for circle maps.

This result gives a complete classification of actions of ABC groups containing an Anosov element up to topological conjugacy and up to finite covers. It strengthens Theorem 1.5 of [39] by removing the slow oscillation condition therein.

Remark 1.3. We do not know how to remove the assumption of the existence of an Anosov element in Theorem 1.1 since the proof relies strongly on the conjugacy linearizing an Anosov diffeomorphism provided by Franks’ theorem, but it seems plausible the result still holds without this assumption in the smooth category. In the topological category, it is possible to construct faithful actions of $\Gamma_B$ by homeomorphisms in any surface (even by homeomorphisms homotopic to the identity) as the group $\Gamma_B$ is left-orderable if $B$ has infinite order and therefore the group $\Gamma_B$ admits an action on $\text{Homeo}(\mathbb{R})$.

1.2. Smooth conjugacy. We next show how to improve the regularity of the conjugacy $h$ in Theorem 1.1. The main strategy is to use the ergodic effect of the $\mathbb{Z}^2$ part of the action to promote an SRB measure to a measure of maximal entropy.

In the next theorem, we get complete classification of faithful volume-preserving actions up to smooth conjugacy.

Theorem 1.4. Let $\Phi : \Gamma_B \to \text{Diff}^r(T^2)$, $r \geq 2$, be a faithful action by volume-preserving diffeomorphisms satisfying:

1) $B \in \text{SL}_2(\mathbb{Z})$ is an Anosov linear map;
2) $\Phi(B)$ is an Anosov diffeomorphism.

Then $h$ and $h^{-1}$ in Theorem 1.1 are $C^{r-\varepsilon}$ for all $\varepsilon > 0$.

We remark that the volume-preservation assumption in the above Theorem 1.4 might be unnecessary and we believe our methods might lead to a proof. Also, the volume-preservation assumption can be replaced by the assumption of preserving a smooth measure.

In the proof of Theorem 1.4 we can reduce to the case where $\text{tr}A = \text{tr}B$, since the rotation matrix $\rho$ has to be rational (hence is nonfaithful) if $\text{tr}A \neq \text{tr}B$, as we will see in the classification of affine actions given in Proposition 4.1 in Section 4.2. We will use the following notation.

Notation: For a diagonalizable matrix $D \in \text{SL}_2(\mathbb{Z})$, we let $u_D$ to be a choice of an eigenvector associated to the larger in norm eigenvalue $\lambda_D$. We define $D^{-t}$ be the transpose of $D^{-1}$. For vectors $v, w \in \mathbb{R}^2$, we let $v \otimes w$ denote the matrix product of $v$ treated as a column vector and $w$ treated as a row vector. More formally, this matrix represents the endomorphism of $\mathbb{R}^2$ defined by the tensor product of the element $v$ of $\mathbb{R}^2$ and the element $w^t$ of the dual $(\mathbb{R}^2)^*$ (defined by the standard inner product).

We will see in Proposition 4.1 that a rotation matrix $\rho$ satisfying the equation $A\rho = \rho B + C$ for some $C \in \mathbb{Z}^{2 \times 2}$ and $\text{tr}A = \text{tr}B$, $|\text{tr}A| > 2$, has the form

$$\rho = \rho_C + c_1 u_A \otimes u_B + c_2 u_{A^{-1}} \otimes u_{B^{-1}}, \quad (c_1, c_2) \in \mathbb{R}^2,$$

where $\rho_C \in \mathbb{Q}^{2 \times 2}$ is a particular solution to $A\rho = \rho B + C$. 


An affine action of the form (1.1) is faithful if and only if \(A\) is of infinite order and \(\{pp, \ p \in \mathbb{Z}^2\}\) is dense on \(\mathbb{T}^2\) (c.f. Proposition 1.2 of [39]). Therefore in terms of the representation (3), the rotation matrices of faithful affine actions correspond to a full measure set of \((c_1, c_2) \in \mathbb{R}^2\). In general the full measure set cannot be \(\mathbb{R}^2\). For example, we consider \(A = B\) then \(\rho = \rho_C + a_1 \text{id} + a_2 A\), \((a_1, a_2) \in \mathbb{R}^2\) and the affine action is not faithful if we take \(a_1, a_2 \in \mathbb{Q}\). Given \(\rho\), it can be checked directly if the affine action is faithful or not.

Next, replacing the above full measure set by another implicit full measure set, we obtain the following theorem without assuming the volume preservation assumption.

**Theorem 1.5.** There exists a full measure subset \(C\) of \(\mathbb{R}^2\) such that if \(\Phi : \Gamma_B \to \text{Diff}^r(\mathbb{T}^2), \ r \geq 2\), is an action satisfying:

1. \(B \in \text{SL}_2(\mathbb{Z})\) is an Anosov linear map;
2. \(\Phi(B)\) is an Anosov diffeomorphism;
3. the rotation matrix (3) has \((c_1, c_2) \in C\).

Then \(h\) and \(h^{-1}\) in Theorem 1.1 are \(C^{r-\varepsilon}\) for all \(\varepsilon > 0\).

For special rotation vectors translating along the stable or unstable leaves of \(A\), we can combine the Herman-Yoccoz theory for circle maps and the Ledrappier-Young entropy formula to obtain the following theorem.

**Theorem 1.6.** Suppose that \(\Phi : \Gamma_B \to \text{Diff}^r(\mathbb{T}^2), \ r > 2\), is such that:

1. \(B \in \text{SL}_2(\mathbb{Z})\) is an Anosov linear map;
2. \(\Phi(B)\) is a volume-preserving Anosov diffeomorphism of \(\mathbb{T}^2\) homotopic to \(A \in \text{SL}_2(\mathbb{Z})\) with \(\text{tr} A = \text{tr} B\);
3. The matrix \(\rho = (\rho_1, \rho_2)\) in (1.1) has the form \(\rho = cu_A \otimes u_B^{t}\) or \(\rho = cu_A^{-1} \otimes u_B^{-t}\), \(c \in \mathbb{R} \setminus \{0\}\).

Then \(h\) and \(h^{-1}\) in Theorem 1.1 are \(C^{r-\varepsilon}\) for all \(\varepsilon > 0\).

The proofs of Theorem 1.4 and 1.5 rely crucially on the two dimensionality. However, the proof of Theorem 1.6 admits a partial higher dimensional generalization as follows.

Let \(B \in \text{SL}_n(\mathbb{Z})\) \(n \geq 2\) be an Anosov linear map with simple real spectrum. We extend the group \(\Gamma_B\) to a group \(\Gamma_{B,n} = \mathbb{Z} \times_B (\mathbb{Z}^n)^n\), where each \(\mathbb{Z}^n\) factor form a copy of \(\Gamma_B = \mathbb{Z} \times_B \mathbb{Z}^n\) with \(\mathbb{Z}\) and the \(\mathbb{Z}^n\) factors commute. An affine action \(\Phi\) of \(\Gamma_{B,n}\) on torus \(\mathbb{T}^n\) is given by \(\Phi(B) = A \in \text{SL}_n(\mathbb{Z})\) and the \(i\)-th \(\mathbb{Z}^n\) factor acts by translation with rotation matrix \(\rho_i \in \mathbb{R}^{n \times n}\) such that \(A \rho_i = \rho_i B\) mod \(\mathbb{Z}^{n \times n}\) holds for all \(i = 1, \ldots, n\).

**Theorem 1.7.** Let \(\Phi : \Gamma_{B,n} \to \text{Diff}^r(\mathbb{T}^n), \ r > n\), be an action satisfying:

1. \(B \in \text{SL}_n(\mathbb{Z})\) is an Anosov linear map with simple real spectrum \(\lambda_1, \ldots, \lambda_n\), and \(A \in \text{SL}_n(\mathbb{Z})\) has the same spectrum;
2. \(\Phi(B)\) is volume-preserving Anosov \(C^1\) close to \(A\);
3. There is a homeomorphism \(h\) conjugating the action to an affine action satisfying \(\rho_i = c_i u_{A,i} \otimes u_{B^{t},i}\), where \(c_i \neq 0\) and \(u_{A,i}\) (resp. \(u_{B^{t},i}\)) is the normalized eigenvector of \(A\) (resp. \(B^{t}\)) corresponding to eigenvalue \(\lambda_i, i = 1, \ldots, n\). Moreover, the Anosov part of the affine action is \(A\).
Then \( h \) and \( h^{-1} \) are \( C^{1,\nu} \) for some \( \nu > 0 \).

We have the following remarks improving the result slightly.

**Remark 1.8.** (1) The \( c_i \) corresponding to one of the largest in norm or smallest in norm eigenvalue is allowed to be zero by the volume preserving condition.

(2) In the case of \( n = 3 \), when \( \Phi(B) \) is sufficiently smooth (\( C^r \) for \( r \) large enough), the conjugacy \( h \) can be \( C^{r-\epsilon} \) for all \( \epsilon > 0 \) applying [18].

In recent years, there has been an emerging interest in understanding actions of non-abelian groups on manifolds by homeomorphisms or diffeomorphisms. Among the simplest interesting examples are solvable groups. For example, for the Bamslag-Solitar group

\[
\text{BS}(p, q) = \langle a, b \mid ab^p a^{-1} = b^q \rangle,
\]

there are results by Farb and Franks [12], by Burslem and Wilkinson [8], who classified analytic BS(1, \( n \)) action on \( S^1 \), which was later improved in [19, 7]. And there are results about Baumslag Solitar group actions on surfaces by [3, 4] and [20, 21].

Actions of ABC groups have been studied by many authors. In [1, 2], Asaoka studied the local rigidity of some non polycyclic (\( |\det B| > 1 \)) ABC group actions on \( \mathbb{T}^n \) and \( S^n \) generalizing [8]. In [39], Wilkinson and the second author studied the local rigidity for polycyclic (\( |\det B| = 1 \)) ABC group actions on \( \mathbb{T}^n \) using a KAM method and also obtained a number of semi-global rigidity results by combining the Herman-Yoccoz theory for circle maps and the geometry of the invariant foliations. In [33], McCarthy proved certain \( C^1 \) local rigidity of the trivial action of ABC groups on compact manifolds, which is generalized to actions by semi products in [9] by Bonatti-Kim-Koberda-Triestino. In [7], Bonatti-Monteverde-Navas-Rivas obtained a number of rigidity results for the \( C^1 \) action of ABC groups on \( \mathbb{T}^n \) using a KAM method and also obtained a number of semi-global rigidity results by combining the Herman-Yoccoz theory for circle maps and the geometry of the invariant foliations. We also note that the local rigidity result of [38] was proved in the Lie group setting which can be considered as a continuous analogue of our discrete ABC group action. In [29], it is proved a local rigidity result for a solvable group \( (\mathbb{Z} \ltimes \lambda \mathbb{R}) \) acting on \( \mathbb{T}^n \).

1.3. Free subgroups of diffeomorphism groups. One novel discovery in the proof of Theorem 1.1 is a version of Tits alternative for subgroups of the diffeomorphism group \( \text{Diff}^1(\mathbb{T}^2) \) containing an Anosov element.

The Tits alternative is a well-known theorem stating that a linear group (i.e. a subgroup of the group of invertible matrices \( GL_n(\mathbb{R}) \)) either contains a free subgroup on two generators or it is virtually solvable. This result has far-reaching applications in various fields in mathematics such as group theory, geometry and dynamics, etc. Non-linear transformation groups such as the groups of homeomorphisms or diffeomorphisms of a closed connected manifold \( M \) (denoted by \( \text{Homeo}(M) \) and \( \text{Diff}(M) \) respectively) are conjectured to behave in many ways similar to linear groups (See [16]) and therefore one can ask whether the Tits Alternative still holds for such groups.

Unfortunately, the Tits alternative does not hold for these groups, since \( \text{Homeo}(S^1) \) and \( \text{Diff}(S^1) \) contain subgroups isomorphic to Thompson’s group \( F \), which is known to not contain free subgroups and neither be solvable (c.f. [35, 23]). Nonetheless, up to what extent such an alternative fails for \( \text{Homeo}(M) \) and \( \text{Diff}(M) \) is not clear. As an example, a well known theorem of Margulis (Conjectured by Ghys) for the group of homeomorphisms of the circle \( S^1 \) states the following:
Theorem 1.9. (Margulis) A group $\Gamma \subseteq \text{Homeo}(S^1)$ either contains a free subgroup on two generators or preserves a probability measure on $S^1$.

In the case when the dimension of $M$ is greater than one, very few similar results are known due to the complexity of elements in $\text{Homeo}(M)$, but there have been some recent results in this direction. For example, Franks-Handel [13] have shown that groups of real-analytic area-preserving diffeomorphisms of $S^2$ which contain certain entropy zero diffeomorphisms satisfy the Tits alternative.

Our discovery is a Tits alternative for $\text{Diff}^1(M)$ in the presence of an Anosov element. To fix the notations, for an Anosov diffeomorphism $f$ on $M$, we denote by $W^s_f$ and $W^u_f$ its stable and unstable foliations respectively.

Theorem 1.10. Let $\Gamma$ be a subgroup of $\text{Diff}^1(T^2)$ containing an Anosov element $f \in \text{Diff}^1(T^2)$, then there is a dichotomy:

1. either there is an index 2 subgroup of $\Gamma$ which preserves either $W^s_f$ or $W^u_f$;
2. or $\Gamma$ contains a free subgroup $F_2$.

The above result relies crucially on the two dimensionality. We next state two results as partial generalizations of Theorem 1.10 to higher dimensions. We denote further by $E^s_f$ and $E^u_f$ the stable and unstable distributions of an Anosov diffeomorphism $f$ respectively. The next result is the main technical ingredient in the proof of Theorem 1.10.

Theorem 1.11. Let $\Gamma$ be a subgroup of $\text{Diff}^1(M)$ satisfying the following:

1. There exists $f \in \Gamma$ which is a transitive Anosov diffeomorphism such that $\dim W^s_f = \dim W^u_f$.
2. There exists $h \in \Gamma$ and $x \in M$ such that each of the subspaces $D_x h(E^s_f(x))$ and $D_x h(E^u_f(x))$ intersect transversely both $E^s_f(h(x))$ and $E^u_f(h(x))$.

Then, there exists an integer $N$ such that $\{f^N, hf^N h^{-1}\}$ generate a free group $F_2$.

Relaxing assumption (2) in Theorem 1.11 we have the following result.

Theorem 1.12. Let $\Gamma$ be a subgroup in $\text{Diff}^1(M)$ satisfying

1. there is a transitive Anosov element $f \in \Gamma$ with stable and unstable foliations $W^s_f$ and $W^u_f$ respectively and $\dim W^s_f \leq \dim W^u_f$;
2. there exist an element $h \in \Gamma$ and a point $x \in M$, such that $D_x h E^u_f(x)$ intersects transversally $E^s_f(h(x))$.

Then there exists an integer $N$ such that $\{f^N, hf^N h^{-1}\}$ generate a nontrivial free semi-group.

The paper is organized as follows: In Section 2, we first give an outline and then proceed to prove Theorem 1.11 and Theorem 1.12. The main observation in the proof is a ping-pong argument. In Section 3, we prove Theorem 1.10. The next four sections are devoted to the proof of Theorem 1.1. The proof is divided into three cases depending on the comparison of $|\lambda_A|$ and $|\lambda_B|$ respectively. In Section 6 we prove Theorem 1.1 in the case of $|\lambda_A| > |\lambda_B|$. In Section 7, we prove Theorem 1.1 in the case of $|\lambda_A| = |\lambda_B|$. In Section 8 we prove Theorem 1.1 in the case of $|\lambda_A| < |\lambda_B|$. In Section 9, we prove Theorem 1.6 and 1.7. In Section 10, we prove Theorem 1.4. In Section 11, we prove Theorem 1.5.
2. Ping-Pong: Proofs of Theorem 1.11 and Theorem 1.12

In this section, we prove Theorem 1.11 and Theorem 1.12. We first give an outline of the proof in Section 2.1. Afterwards, we give the formal proof. We will start by recalling some notations and results from hyperbolic dynamics in Section 2.2. In Section 2.3, we give the ping-pong argument. In the remaining two subsections, we prove Theorem 1.11 and Theorem 1.12.

2.1. Outline of the proof of Theorem 1.11. We now proceed to give an outline of the proof of Theorem 1.11.

If we consider the two Anosov diffeomorphism $f_1 := f$ and $f_2 := hfh^{-1}$, both the stable foliations and unstable foliations of $f_1$ and $f_2$ are transverse in some open neighborhood of by assumption 2) in Theorem 1.11. The existence of a free group follows from a ping-pong argument using the following elementary observation: Suppose $B$ is a small open ball in $M$, if $N > 0$ is large enough, the ball $f_1^N(B)$ looks similar to a small neighborhood of a long piece of a leaf of the unstable foliation $W^u_{f_1}$, as the foliation $W^u_{f_1}$ is transverse to both foliations $W^u_{f_2}$ and $W^u_{f_2}$, by the inclination lemma 2.3 the set $f_2^N f_1^N(B)$ (or $f_2^{-N} f_1^N(B)$) must contain an open set that looks like an open neighborhood of a leaf of $W^u_{f_2}$ (or $W^u_{f_2}$) near every point in $M$ and so we conclude that the words $f_2^N f_1^N$ and $f_2^{-N} f_1^N$ are non-trivial elements in $Γ$, see Figure 1. In general, we will show that if $N$ is large enough, for any non-trivial word $w$ in the elements $f_1^N$ and $f_2^N$ the set $w(B)$ will contain a subset that looks like the stable or unstable foliations of $f_1$ or $f_2$ and this will guarantee that $w(B) ≠ B$ and so the group generated by $\{f_1^N, f_2^N\}$ must be necessarily free.

2.2. Preparations from hyperbolic dynamics. We recall some definitions and well known facts about hyperbolic dynamics and prepare notations for later use. Let $p$ be a hyperbolic fixed point of a diffeomorphism $f$ of $M^m$.

Definition 2.1. (1) By Theorem 6.2.3 in [26], there exists a neighborhood $O$ of $p$ and a coordinate chart $Φ : O → R^m$ such that $Φ(W^s_p(p) ∩ O) ⊂ R^k ⊕ \{0\}$ and $Φ(W^u_p(p) ∩ O) ⊂ \{0\} ⊕ R^{m-k}$ where $m = \dim(M)$ and $k = \dim(W^u_p(p))$, this construction is known as “adapted coordinates” for the dynamics of $f$ near $p$.

(2) In those coordinates there are obvious projections from $R^m$ into $R^k$ or $R^{m-k}$ that we denote by $π_1, π_2$ respectively.

(3) In the decomposition $R^k ⊕ R^{m-k}$, we say that $R^k$ is horizontal and $R^{m-k}$ is vertical.

Definition 2.2. (1) Let $D_k ⊂ R^k$ be the $k$-dimensional open unit disk, a $C^1$ embedded $k$-disk $D$ in $M$ is defined as a $C^1$ embedding $f : D_k → M$.

(2) Consider the $k$-Grassmanian $Gr_k(M)$ of $M$, which is a fiber bundle over $M$ whose fiber over point $p$ of $M$ consist of all the $k$-dimensional subspaces of the tangent space $T_p(M)$, we choose an arbitrary metric on $Gr_k(M)$ that we denote by $d$. A $C^1$ map $f : D_k → M$ induces a natural map $G_k(f) : D_k → Gr_k(M)$ via $G_k(f)(p) = (f(p), D_pf(T_p D_k))$.

(3) We say that two disks $f_1 : D_k → M$ and $f_2 : D_k → M$ are $C^1$ $ε$-close if for every $p ∈ D_k$ we have $d(G_k(f_1(p)), G_k(f_2(p))) < ε$.

Our main technical tool is the following theorem known as the Inclination Lemma or $λ$-Lemma.
Lemma 2.3 (Inclination Lemma, Lemma 6.2.23 [26]). Considering $C^1$ adapted coordinates on a neighborhood $O$ of a hyperbolic fixed point $p$ of a diffeomorphism $f : M \to M$. Given $\varepsilon, K, \nu > 0$, there exists $N$ such that if $D$ is a $C^1$ embedded $k$-disk containing $q \in W^s_f \cap O$ with all tangent spaces in horizontal $K$-cones and such that $\pi_2(D)$ contains a $\nu$-ball around $0 \in \mathbb{R}^k$, then there exists $D_1 \subset D$ embedded disk such that for any $n \geq N$, the set $\pi_1(f^n(D_1))$ contains a unit ball in $\mathbb{R}^k$ and $T_zf^n(D_1)$ is contained in a horizontal $\varepsilon$-cone for every $z \in f^n(D_1)$.

Remark 2.4. See ([26], Sec. 6.2) for the definition of cones. In [26] the theorem is stated slightly different but the equivalence is easy. See also Lemma 5.7.1. in [10].

We have the following corollary of the inclination lemma.

Corollary 2.5. Suppose that $D_1$ is a $C^1$ embedded $k$-disk intersecting $W^s_f(p)$ transversely and $D_2 \subset W^u_f(p)$ is an embedded $k$-disk, then there exists $\varepsilon_0 > 0$ (just depending on $D_1$) such for any $0 < \varepsilon < \varepsilon_0$ there exists $N > 0$ such that if $D$ is a $k$-disk which is $C^1$ $\varepsilon_0$-close to $D_1$, then $D$ intersects $W^s_f(p)$, and for all $n \geq N$, $f^n(D)$ contains a $C^1$ embedded disk $D'$ which is $C^1 \varepsilon$-close to $D_2$.

For a fixed Riemannian metric on $M$, $x \in M$, $\delta > 0$, let $W^s_f(p, \delta)$ and $W^u_f(p, \delta)$ be the local stable and unstable manifolds at distance $\delta$, which are the subsets of $M$ consists of points whose forward orbit (or backward orbit) stay at distance less than $\delta$ to the orbit of $p$. If $p$ is a hyperbolic periodic point and $\delta$ is small enough, then $W^s_f(p, \delta)$ and $W^u_f(p, \delta)$ are embedded manifolds.

2.3. The Ping-Pong argument. The main idea in the proof of Theorem 1.11 is enclosed in the following proposition:

Proposition 2.6 (Ping Pong argument). Suppose $f_1, f_2 \in \text{Diff}(M)$ have hyperbolic fixed points $p_1, p_2 \in M$ and $\delta > 0$ is such that the local stable and unstable disks $W^s_{f_1}(p_1, \delta)$, $W^u_{f_1}(p_1, \delta)$ are smooth and both disks intersect transversely and non-trivially both $W^s_{f_2}(p_2, \delta)$ and $W^u_{f_2}(p_2, \delta)$, then there exists $N > 0$ such that for all $n \geq N$ the group generated by $\{f_1^n, f_2^n\}$ is free.

Proof. Let $O_1, O_2$ be adapted coordinates for $f_1$ and $f_2$ with corresponding system of coordinates $\phi_i : O_i \to \mathbb{R}^n$, we can assume that $O_1$ and $O_2$ are disjoint sets and for simplicity assume that $\phi_i(O_i)$ contains the ball $B_4^k(0) \subset \mathbb{R}^n$ of radius 4 for each $i = 1, 2$. Let

$$C_i := \phi_i^{-1}(\{2\} \times B_1^k(0) \cup B_1^{n-k}(0) \times \{2\})$$

and let $C = C_1 \cup C_2$. Therefore $C$ is the union of four “small” disks, each one parallel to one of the stable or unstable manifolds of $f_1$ at $p_1$ or $f_2$ at $p_2$. See figure 1.

As an application of Corollary 2.5, there exists $\varepsilon_1 > 0$, such that for any $0 < \varepsilon < \varepsilon_1$, there is $N_1 > 0$ such that if a $C^1$ smooth disk $D$ is $\varepsilon_1$-close to $W^s_{f_1}(p_1, \delta)$ (or $W^u_{f_1}(p_1, \delta)$), then $D$ intersects transversely $W^s_{f_2}(p_2, \delta)$ and $W^u_{f_2}(p_2, \delta)$ and if $n \geq N_1$, then $f_2^n(D)$ contains a disk which is $C^1$ $\varepsilon$-close to $W^u_{f_2}(p_2, \delta)$ and $f_2^{-n}(D)$ contains a disk which is $C^1$ $\varepsilon$-close to $W^s_{f_2}(p_2, \delta)$.

The same argument implies there exists $\varepsilon_2 > 0$, such that for any $0 < \varepsilon < \varepsilon_2$ there is $N_2 > 0$ so that if a $C^1$ smooth disk $D$ is $\varepsilon_2$-close to $W^s_{f_2}(p_2, \delta)$ (or $W^u_{f_2}(p_2, \delta)$), then $D$ intersects transversely
both $W_s^u(p_1, \delta)$ and $W_u^s(p_1, \delta)$ and if $n \geq N_2$, then $f^n_1(D)$ contains a disk which is $C^1$-close to $W_{f_1}^u(p_1, \delta)$ and also $f^{-n}_1(D)$ contains a disk which is $C^1$-close to $W_{f_1}^s(p_1, \delta)$.

Also, from the inclination Lemma, for any $\varepsilon > 0$ there exists $N_3 > 0$ so that if $n \geq N_3$, then for every $i = 1, 2$, $f^n_i(C)$ contains an embedded disk $\varepsilon$-close to $W_{f_i}^u(p, \delta)$ and that $f^{-n}_i(C)$ contains an embedded disk $\varepsilon$-close to $W_{f_i}^s(p, \delta)$.

Let $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$ and let $N = \max\{N_1, N_2, N_3\}$. By definition of $\varepsilon_1, \varepsilon_2, N_1, N_2, N_3$ one obtains by an easy induction in the word length the following:

**Claim 2.7.** If $n \geq N$, for any non trivial word $w$ in $\{f^n_1, f^n_2, f^{-n}_1, f^{-n}_2\}$ we have:

1. If $w$ starts by $f^n_1$ ($f^n_2$ respectively), the set $w(C)$ contains an embedded disk $D$ that is $C^1$-$\varepsilon$-close to $W_{f_1}^u(p_1, \delta)$ ($W_{f_2}^u(p_2, \delta)$ respectively).

2. If $w$ starts by $f^{-n}_1$ ($f^{-n}_2$ respectively), the set $w(C)$ contains an embedded disk $D$ that is $C^1$-$\varepsilon$-close to $W_{f_1}^s(p_1, \delta)$ ($W_{f_2}^s(p_2, \delta)$ respectively).

As a consequence of Claim 2.7, for any non-trivial word $w(C)$ contains an embedded disk $C^1$-$\varepsilon$-close to one among $W_{f_1}^u(p_i)$ or $W_{f_2}^s(p_i)$ for some $i$ and in particular choosing $\varepsilon$ small enough, we can guarantee that $w(C) \neq C$. In that case, the group generated by $\{f^n_1, f^n_2\}$ is free.

\[\square\]

2.4. **Proof of Theorem 1.11.** We now finish the proof of Theorem 1.11:

**Proof of Theorem 1.11.** As the condition of $D_xh(E^u_f(x))$ and $D_xh(E^s_f(x))$ intersecting transversely both $E^u_f(h(x))$ and $E^s_f(h(x))$ is an open condition on $x \in M$ and periodic points of $f$ are dense (consequence of the Anosov Closing lemma and $f$ being transitive) we can assume that $x$ is a periodic point of $f$. 


By choosing \( \delta \) small enough we can assume that \( W_f^s(x, \delta) \) and \( W_f^u(x, \delta) \) are embedded disjoint disks such that each of the disks \( h(W_f^s(x, \delta)) \) and \( h(W_f^u(x, \delta)) \) are both transverse to \( W_f^s(h(x)) \) and \( W_f^u(h(x)) \).

Let \( p_2 := h(x) \), as the periodic points of \( f \) are dense, we can assume that there exists a periodic point \( p_1 \) near the point \( p_2 \) such that the local stable and unstable manifolds \( W_f^s(p_1, \delta) \) and \( W_f^u(p_1, \delta) \) are embedded manifolds that intersect transversely \( h(W_f^s(x, \delta)) \) and \( h(W_f^u(x, \delta)) \) (we are using that \( W^s \) and \( W^u \) vary continuously). By taking \( p_1 \) close enough to \( h(x) \), we can assume that the intersection of any of the disks \( h(W_f^s(x, \delta)) \) and \( h(W_f^u(x, \delta)) \) with any of the disks \( W_f^s(p_1, \delta) \) (or \( W_f^u(p_1, \delta) \)) is a transverse intersection at exactly one point. See figure 1.

Let \( f_1 := f^N \) and \( f_2 := h f^N h^{-1} \) for some \( N > 0 \). Observe that both \( f_1 \) and \( f_2 \) are Anosov diffeomorphisms, by choosing \( N \) multiple of both the periods of \( p_1 \) and \( p_2 \), we can assume that the points \( p_1, p_2 \) are hyperbolic fixed points of \( f_1^N \) and \( f_2^N \). We can therefore apply proposition 2.6 and we are done.

\[ \square \]

2.5. Proof of Theorem 1.12. Similarly, we finish the proof of Theorem 1.12.

**Proof of Theorem 1.12.** Similar to the proof of Theorem 1.11, by the assumption, we can find \( N_1 \) and \( p_1, p_2 \) such that the following hold:

1. \( f_1 = f^{N_1} \), \( f_2 = h f^{N_1} h^{-1} \) have \( p_1 \) and \( p_2 \) as hyperbolic fixed points respectively;
2. \( W_f^s(p_1) \) intersects \( W_f^u(p_2) \) transversally.

We can now repeat the proof of Proposition 2.6 for words generated by \( \{f_1^{N_2}, f_2^{N_2}\} \) for some \( N_2 \) large to show that \( \{f_1^{N_2}, f_2^{N_2}\} \) generates a semigroup. Indeed, we pick a disk \( D_i \) of dimension \( \dim W_f^s \) near \( p_i \) transverse to \( W_f^s(p_i) \) and disjoint from \( W_f^u(p_i) \), \( i = 1, 2 \). After the application of any word starting with \( f_1^{N_2} \), the image of \( D_1 \) would be very close to either of \( W_f^u(p_i) \), so is transverse to both \( W_f^{u}(p_{2-i}) \) and \( W_f^{u}(p_{2-i}) \), \( i = 1, 2 \). In particular, the image can never by \( D_1 \). Similarly for a word starting with \( f_2^{N_2} \), we look at the image of \( D_2 \) instead. This completes the proof. \[ \square \]

3. **Proof of Theorem 1.10**

In this section we give a proof of Theorem 1.10. Observe that by Theorem 1.11 to obtain a free subgroup of \( \Gamma \), we only need to show the existence of an element \( h \in \Gamma \) and a point \( x \in \mathbb{T}^2 \) such that the subspaces \( D_x h(E_f^s(x)) \) and \( D_x h(E_f^u(x)) \) intersect transversely both \( E_f^s(h(x)) \) and \( E_f^u(h(x)) \). This will be shown in Proposition 3.2.

**Proposition 3.1.** Suppose \( f \) is an Anosov diffeomorphism in a subgroup \( \Gamma \subset \text{Diff}^1(\mathbb{T}^2) \) and suppose there exist \( h_1, h_2 \in \Gamma \) and \( x_1, x_2 \in \mathbb{T}^2 \) such that:

1. \( D_{x_1} h_1(E_f^s(x_1)) \) intersect transversely both \( E_f^s(h_1(x_1)) \) and \( E_f^u(h_1(x_1)) \).
2. \( D_{x_2} h_2(E_f^u(x_2)) \) intersect transversely both \( E_f^s(h_2(x_2)) \) and \( E_f^u(h_2(x_2)) \).

Then, there exists \( h \in \Gamma \) and \( x \in \mathbb{T}^2 \) such \( D_x h(E_f^s(x)) \) and \( D_x h(E_f^u(x)) \) intersect transversely both \( E_f^s(h(x)) \) and \( E_f^u(h(x)) \).
Proof. If \( D_{x_1} h_1(E_f^s(x_1)) \) intersect transversely both \( E_f^u(h_1(x_1)) \) and \( E_f^s(h_1(x_1)) \) then we can take \( h = h_1 \) and \( x = x_1 \) and we are done. Otherwise we have either of the following two cases holds:

(a) \( D_{x_1} h_1(E_f^u(x_1)) = E_f^s(h_1(x_1)) \);
(b) \( D_{x_1} h_1(E_f^s(x_1)) = E_f^u(h_1(x_1)) \).

Suppose the former is true, i.e. we have \( D_{x_1} h_1(E_f^u(x_1)) = E_f^s(h_1(x_1)) \).

We may further assume that \( D_{x'} h_1(E_f^u(x')) = E_f^s(h_1(x')) \) for all \( x' \) sufficiently close to \( x_1 \). The assumption (1) still holds with \( x' \) in place of \( x_1 \) by the continuity. We also have \( D_{x'} h_1(E_f^u(x')) \neq E_f^s(h_1(x')) \) since we have assumed \( D_{x_1} h_1(E_f^u(x_1)) \neq E_f^s(h_1(x_1)) \). If \( D_{x'} h_1(E_f^u(x')) \neq E_f^s(h_1(x')) \) for some point \( x' \) close to \( x_1 \), we then complete the proof by choosing \( h = h_1 \) and \( x = x' \).

By the transitivity of \( f \), we choose \( x' \) sufficiently close to \( x_1 \) such that

- \( h_1(x') \) lies in a dense backward orbit of \( f \);
- \( (1) \) holds with \( x' \) in place of \( x_1 \);
- \( D_{x'} h_1(E_f^u(x')) = E_f^s(h_1(x')) \).

For any \( N \) we have \( D_{x'} f^{-N} h_1(E_f^u(x')) = E_f^s(f^{-N} h_1(x')) \). Next, by assumption (1) and taking \( N \) sufficiently large we have that \( D_{x'} f^{-N} h_1(E_f^u(x')) \) is as close as we want to \( E_f^s(f^{-N} h_1(x')) \).

By the first bullet point and the assumption (1), we can choose \( N \) such that \( f^{-N} h_1(x') \) and \( x_1 \) are close enough to guarantee that the image under \( D h_1 \) of both \( D_{x'} f^{-N} h_1(E_f^u(x')) \) and \( D_{x'} f^{-N} h_1(E_f^s(x')) \) are transverse to \( E^s(h_1 f^{-N} h_1 x') \) and \( E^u(h_1 f^{-N} h_1 x') \). The proposition is then proved by choosing \( h = h_1 f^{-N} h_1 \) and \( x = x' \).

The argument for the case (b) is similar. We have \( D_{x'} f^N h_1(E_f^s(x')) = E_f^u(f^N h_1(x')) \) for all \( N \), and by (1) in the assumption \( D_{x'} f^N h_1(E_f^s(x')) \) is arbitrarily close to \( E_f^u(f^N h_1(x')) \). Hence we complete the proof by taking \( h = h_2 f^N h_1 \) and apply (2) in the assumption.

The following proposition finishes the proof of Theorem 1.10.

**Proposition 3.2.** Suppose \( f \in \Gamma \) is an Anosov diffeomorphism of \( \mathbb{T}^2 \) and let \( W^s \) and \( W^u \) be the stable and unstable foliations of \( f \). One of the following holds:

1. There exists an index 2-subgroup \( \Gamma_0 \subset \Gamma \) preserving one among \( W^s \) or \( W^u \).
2. There exists \( h \in \Gamma \) and \( x \in \mathbb{T}^2 \) such that each of \( D_x h(E_f^s(x)) \) and \( D_x h(E_f^u(x)) \) intersect transversely both \( E_f^s(h(x)) \) and \( E_f^u(h(x)) \).

**Proof.** If Condition (2) does not hold, then one of the conditions in Proposition 3.1 does not hold, suppose the first condition does not hold, that is, suppose that for every \( h \in \Gamma \) and every \( x \in M \) we have that \( D_x h(E_f^s(x)) \) does not intersect transversely one among \( E_f^s(h(x)) \) and \( E_f^u(h(x)) \), as \( \mathbb{T}^2 \) is 2-dimensional then we must have \( D_x h(E_f^s(x)) \) coincides with \( E_f^s(h(x)) \) or \( E_f^u(h(x)) \) and as \( \mathbb{T}^2 \) is connected, this imply that \( h(W^s) = W^s \) or \( h(W^s) = W^u \), and this implies that the subgroup

\[
\Gamma_0 = \{ h \in \Gamma \text{ such that } h(W^s) = W^s \}
\]

is a subgroup of index 2 in \( \Gamma \) preserving \( W^s \).

□
4. Global rigidity of Anosov ABC actions on $\mathbb{T}^2$: Proof outline

4.1. Outline of the global rigidity. In this section we give an outline of the proof of Theorem 1.1. We will proceed to carry out this outline in the next sections. Recall that for a matrix $B \in \text{SL}_2(\mathbb{Z})$, we let $\Gamma_B := \mathbb{Z} \rtimes_B \mathbb{Z}^2$ be the corresponding Polycyclic Abelian-by-cyclic (ABC) group, which is a semi-direct product where $\mathbb{Z}$ acts in $\mathbb{Z}^2$ by $B$ and for convenience we denote the generator of $\mathbb{Z}$ in $\Gamma_B$ by $B$ as well.

Recall that we are concerned with understanding homomorphisms $\Phi : \Gamma_B \to \text{Diff}(\mathbb{T}^2)$ such that $\Phi(B)$ is an Anosov diffeomorphism. By Franks Theorem [11] there exists a homeomorphism $h \in \text{Homeo}(\mathbb{T}^2)$ such that the conjugate $A := h\Phi(B)h^{-1}$ is an Anosov linear map, we denote by $\lambda_A$ and $\lambda_B$ the larger in modulus eigenvalue of $A$ and $B$ respectively.

The proof of Theorem 1.1 has several steps:

Step 1: By using our ping pong result (Theorem 1.10) and using the fact that $\Gamma_B$ does not contain any free subgroup we can reduce to the case where $\Phi(\Gamma_B)$ preserves the unstable (or stable) foliation of $\Phi(B)$.

Step 2: We must show that for all $v \in \mathbb{Z}^2 \subset \Gamma_B$, $h\Phi(v)h^{-1}$ of $\text{Homeo}(\mathbb{T}^2)$ is a translation. We first use the fact that $\text{PSL}_2(\mathbb{Z})$ is isomorphic to $\mathbb{Z}_2 * \mathbb{Z}_3$ to show that each $\Phi(v)$ is homotopic to identity (Corollary 5.1). Next, we show that for any $v \in \mathbb{Z}^2$, the Misiurewicz-Ziemian rotation set of $\Phi(v)$ consists of a single point (Proposition 5.5). This is carried out as follows. Step 1 allow us to show that for $v \in \mathbb{Z}^2$ the rotation set of a lift of $\Phi(v)$ consists of either a single rotation vector or is a non-trivial segment parallel to the unstable eigenvector of $A$ and so we have to rule out this later possibility. While understanding the rotation set of elements of $\Phi(\mathbb{Z}^2)$, we were lead into understanding the joint rotation set (See Definition 5.9) which is the set of possible average rotation vectors for $\Phi(\mathbb{Z}^2)$-invariant measures, this set has a natural invariance by $A$ and $B$ and by using the hyperbolicity of $A$ and $B$ and basic facts about the joint rotation set, we will conclude that the actual rotation set of $\Phi(v)$ is a singleton for all $v \in \mathbb{Z}^2 \subset \Gamma_B$. The proof that the rotation set is a singleton is much easier in the case where $|\lambda_A| \neq |\lambda_B|$, the case where $|\lambda_A| = |\lambda_B|$ require another argument which is given in Section 7.

After showing that rotation sets consists of singletons, the proof is divided into two cases:

Step 3: The case where $|\lambda_A| \geq |\lambda_B|$: In the case where $|\lambda_A| > |\lambda_B|$, using that rotation set consists of singletons and conjugating by high powers of $B$ one shows that the group $h\Phi(\mathbb{Z}^2)h^{-1}$ is a group of translations by a elementary geometric argument, finishing the proof in this case (Proposition 6.2). The case where $|\lambda_A| = |\lambda_B|$ is more difficult and requires further understanding of the rotation set and a modification of Proposition 6.2.

Step 4: The case where $|\lambda_A| < |\lambda_B|$: This case is much more difficult, we proceed by contradiction and assume such an action exists. By using the classification of the affine actions of $\Gamma_B$ (Section 4.2), we can show that the set of rotation vectors is finite and passing to a finite index subgroup we assume that for all $v \in \mathbb{Z}^2$, $\Phi(v)$ preserves set-wise each of the leaves of the unstable foliation of $\Phi(A)$, obtaining actions of $\mathbb{Z}^2$ in copies of $\mathbb{R}$ (the leaves of the foliation). We then proceed to study such actions and show that away from fixed points, such actions are conjugated to an action by translations in $\mathbb{R}$ that satisfies a specific algebraic (diophantine) condition on the possible translations.

This allows to use one-dimensional Herman-Yoccoz theory to embed the action of $\mathbb{Z}^2$ into a flow $g_t$ that preserves the leaves of the unstable foliation. The vector field $X$ generating such flow must...
satisfy $D(\Phi(B))X = \lambda_B X$ by the group relations of $\Gamma_B$ and then if $\mu$ is the SRB measure preserved by $f$, we must have that $h_{\mu}(f) = \log(|\lambda_B|)$, but as the topological entropy $h_{top}(\Phi(B)) = \log(|\lambda_A|)$ being $\Phi(B)$ conjugate to $A$ we must have that $\log |\lambda_A| \geq \log |\lambda_B|$, contradicting the assumption that $|\lambda_A| < |\lambda_B|$.

We finish this section by discussing the classification of affine actions of our group $\Gamma_B$.

4.2. **Classification of affine actions.** Let us consider affine actions $\Phi : \Gamma_B \to \text{Aff}(\mathbb{T}^2)$ of the form $\Phi(B) = A \in \text{SL}_2(\mathbb{Z})$, and $\Phi(e_i) : x \mapsto x + \rho_i \in \mathbb{R}^2$ where $e_1 = (1,0)$ and $e_2 = (0,1)$, $i = 0,1$. We denote by $\rho$ the matrix $(\rho_1, \rho_2)$ formed by $\rho_1$ and $\rho_2$ as columns. For this affine action, the group relation can be written as (equation (2))

$$A\rho = \rho B \mod \mathbb{Z}^{2 \times 2}.$$ 

The main result of this section is the following.

**Proposition 4.1.** Let $A$ and $B$ be two matrices in $\text{SL}_2(\mathbb{Z})$ satisfying (2). Then

1. If $\text{tr}(A) \neq \text{tr}(B)$, then all solutions $\rho$ of the equation (2) have rational coefficients and therefore the action $\Phi$ is not faithful.
2. If $\text{tr}(A) = \text{tr}(B)$, and $|\text{tr}(A)| > 2$, then
   a. the null space of the map $X \mapsto AX - XB$ is spanned by $u_A \otimes u_B^t$ and $u_{A^{-1}} \otimes u_{B^{-t}}$.

So the general solution to (2) has the form

$$\rho = \rho_C + c_1 u_A \otimes u_B^t + c_2 u_{A^{-1}} \otimes u_{B^{-t}}, \quad c_1, c_2 \in \mathbb{R}$$

where $\rho_C \in \mathbb{Q}^{2 \times 2}$ is a particular solution to $A\rho = \rho B = C, \quad C \in \mathbb{Z}^{2 \times 2}$.

b. Suppose in addition that $A = B$, then the null space of $X \mapsto AX - XA$ is spanned by $\{\text{id}, A\}$. So the general solution to (2) has the form

$$\rho = \rho_C + c_1 \text{id} + c_2 A, \quad c_1, c_2 \in \mathbb{R}.$$ 

We use the following definition of Kronecker product.

**Definition 4.2.** Let $A' = (a_{ij})$ be a $m \times n$ matrix and $B' = (b_{ij})$ be a $p \times q$ matrix. We define the Kronecker product $A' \otimes B'$ as the $mp \times nq$ matrix written in block form whose $(i,j)$-th block is $a_{ij} B'$ where $1 \leq i \leq m, 1 \leq j \leq n$.

The Kronecker product has the following properties:

**Proposition 4.3.**

1. If $A', B'$ are both $n \times n$ with eigenvalues $\{\lambda_i\}$ and $\{\mu_i\}$, respectively, then the eigenvalues of $A' \otimes B'$ are $\{\lambda_i \mu_j\}$, $i, j = 1, \ldots, n$.
2. $(A' \otimes B')(C' \otimes D') = A'C' \otimes B'D'$.
3. $(A' \otimes B')^{-1} = A'^{-1} \otimes B'^{-1}$, if both $A'$ and $B'$ are invertible.
4. ([24], Chapter 4.4) an $n \times n$ matrix $X$ solves $A'X + XB' = C'$ for $A', B', C' n \times n$ matrices, if and only if the $\mathbb{R}^{n^2}$ vector $\tilde{X}$ formed by listing columns of $X$ solves the equation $(\text{id}_n \otimes A' + B' \otimes \text{id}_n)\tilde{X} = \tilde{C}.$

**Proof of Proposition 4.1.** Now for $A, B \in \text{SL}_2(\mathbb{Z})$, there are two cases depending on whether $\text{tr}(A)$ and $\text{tr}(B)$ are equal or not.

Suppose $\text{tr}(A) \neq \text{tr}(B)$. In this case, the set of eigenvalues of $A$ has no intersection with that of $B$, so that $\text{id}_2 \otimes A - B \otimes \text{id}_2$ is nondegenerate. The equation $A\rho = \rho B$ has only trivial solution and
the equation $A\rho = \rho B + C$ where $C \in \mathbb{Z}^{2\times 2}$ has only rational solution. In this case, the affine action generated by (1.1) cannot be faithful.

Suppose $tr(A) = tr(B)$. Nontrivial affine actions exist already in the case of $C = 0$. Indeed, in this case, we have

$$(id_2 \circ A - B \circ id_2) = (id_2 \circ A)(id_4 - B \circ A^{-1}).$$

The matrix $(id_2 \circ A)$ is nondegenerate and the matrix $(id_4 - B \circ A^{-1})$ is degenerate since $B \circ A^{-1}$ has two eigenvalues 1. So the equation $A\rho = \rho B$ has nontrivial solution.

Note that in general we may have $B$ and $A$ with $tr(A) = tr(B)$ without being conjugate in $SL_2(\mathbb{Z})$. For example one can take $B = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$ and $A = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$, but we must point out that for a matrix $A$ in $SL_2(\mathbb{Z})$, the number of conjugacy classes of matrices $B$ with $tr(A) = tr(B)$ is finite and it is related to the class number of the number field defined by the eigenvalues of $A$.

We next prove item (2) in the statement. By assumption, $A$ and $B$ have same eigenvalues $\{\lambda, 1/\lambda\}$ with $\lambda \neq 1/\lambda$. In this case, the null space of $X \mapsto AX - XB$ has dimension 2 by Theorem 4.4.14 of [24]. It is clear that both $u_A \otimes u_B$ and $u_A^{-1} \otimes u_B^{-1}$ are in the null space and are linearly independent, so we obtain item (2.a).

For item (2.b), again the null space has dimension 2 and it is clear that id and $A$ are in the null space and are linearly independent (see also Corollary 4.4.15 of [24]). This gives item (2.b).

To show that $\rho_C$ can be chosen to be rational, we apply Gauss elimination to the equation $(id \circ A - B \circ id)\tilde{\rho} = \tilde{C}$. This completes the proof. \hfill \Box

5. Basic Reductions and Rotation Set Properties

In this section we begin the proof of Theorem 1.1 by first making some reductions and then proceed to study the rotation set of the translation subgroup of $\Gamma_B$. This carries out Step 1 and Step 2 of the outline discussed in Section 4.

5.1. The $\mathbb{Z}^2$ action has trivial homotopy type. To start with the proof of Theorem 1.1, we have from Theorem 1.10, the following corollary:

**Proposition 5.1.** Suppose that $B \in SL_2(\mathbb{Z})$ is Anosov and $\Phi : \Gamma_B \to \text{Diff}(\mathbb{T}^2)$ is such that the diffeomorphism $f := \Phi(B)$ is an Anosov diffeomorphism of $\mathbb{T}^2$. Then there exist a finite index subgroup $\Gamma' \subset \Gamma_B$ isomorphic to $\Gamma_B$ and $g \in \text{Homeo}(\mathbb{T}^2)$ such that the conjugate action $\Phi' : \Gamma' \to \text{Homeo}(\mathbb{T}^2)$ given by $\Phi'(\gamma) := g\Phi(\gamma)g^{-1}$ satisfy:

1. $A = g\Phi(B)g^{-1} \in SL_2(\mathbb{Z})$.
2. Each $\Phi'(\gamma), \gamma \in \Gamma'$, preserves the unstable foliation $\mathcal{F}$ of $A$ (up to replacing $A$ with $A^{-1}$).
3. For each $\gamma \in \mathbb{Z}^2 \cap \Gamma'$, $\Phi'(\gamma)$ is homotopic to the identity.

**Proof.** By Franks Theorem [11], there is a homeomorphism $g \in \mathbb{T}^2$ homotopic to the identity such that the conjugate $g\Phi(B)g^{-1}$ is an Anosov linear map of $\mathbb{T}^2$ homotopic to the action $\Phi(B)_*$ of $\Phi(B)$ on $H_1(\mathbb{T}^2, \mathbb{Z})$. We denote by $A = \Phi(B)_* \in SL_2(\mathbb{Z})$.

By Theorem 1.10, there exists a finite index subgroup of $\Gamma_1 \subset \Gamma_B$ that preserves one of the stable or unstable foliations of $A$ denoted by $\mathcal{F}$, moreover we can assume that such subgroup $\Gamma_1$ contains $B$, because $\Phi(B)$ preserves $\mathcal{F}$.
Now observe that $\Phi$ gives a homomorphism:

$$\Psi : \Gamma_B \to \text{MCG}(\mathbb{T}^2)$$

where $\text{MCG}(\mathbb{T}^2)$ is the mapping class group of $\mathbb{T}^2$. It is well known that $\text{MCG}(\mathbb{T}^2)$ is isomorphic to $\text{SL}_2(\mathbb{Z})$ and $\text{PSL}_2(\mathbb{Z})$ is isomorphic to the free product $\mathbb{Z}_2 \ast \mathbb{Z}_3$ which is virtually free. As $\Gamma_B$ is a solvable group and $\Psi(B)$ has infinite order (because $\Phi(B)$ is Anosov), by using that $B$ normalizes $\mathbb{Z}^2$ and $\text{MCG}(\mathbb{T}^2)$ virtually free, it follows easily that $\Psi(\mathbb{Z}^2)$ must be finite and so we can find a subgroup of the form $\mathbb{Z} \ltimes_B k\mathbb{Z}^2$ for some $k \neq 0$ such that $\Psi(k\mathbb{Z}^2)$ is trivial, we call this group $\Gamma_2$.

The group $\Gamma_3 = \Gamma_1 \cap \Gamma_2$ satisfy both (1) and (2) and contains a subgroup $\Gamma'$ of the form $\mathbb{Z} \ltimes_B k'\mathbb{Z}^2$ for some $k' \neq 0$, which is isomorphic to $\Gamma_B$. $\square$

5.2. Rotation sets. We now discuss rotation sets for toral homeomorphisms and some of its properties.

**Definition 5.2.** For a homeomorphism $f : \mathbb{T}^n \to \mathbb{T}^n$ homotopic to the identity of $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ and a lift $\tilde{f} : \mathbb{R}^n \to \mathbb{R}^n$ of $f$ the Misiurewicz-Ziemian rotation set $\text{rot}(\tilde{f})$ is defined to be the subset of $\mathbb{R}^n$ consisting of limits of sequences of the form

$$\left\{ \frac{\tilde{f}^n(x_i) - x_i}{n_i} \right\}$$

where $x_i \in \mathbb{R}^n$ and $n_i \to \infty$.

The rotation set is an important dynamical invariant and it is known to be compact for arbitrary $n$ and convex in the case $n = 2$, see [34]. Observe also that the rotation set depends on the lift $\tilde{f}$ but it is defined for $f$ up to translation by an element of $\mathbb{Z}^n$. Observe also that if $\text{rot}(\tilde{f})$ consists of only one point, we have that for every $x, y \in \mathbb{R}^n$:

$$\lim_{n \to \infty} \frac{(\tilde{f}^n(x) - x) - (\tilde{f}^n(y) - y)}{n} = 0$$

and moreover such limit converges uniformly with respect to $x$ and $y$ in $\mathbb{R}^n$.

**Definition 5.3.** For an $f$-invariant probability measure $\mu$ on $\mathbb{T}^n$ and a lift $\tilde{f} : \mathbb{R}^n \to \mathbb{R}^n$, one can also define the average rotation vector $\text{rot}_\mu(\tilde{f})$ by

$$\text{rot}_\mu(\tilde{f}) = \int_{\mathbb{T}^n} \tilde{f}(\tilde{x}) - \tilde{x} \, d\mu.$$  

When the measure $\mu$ is ergodic, Birkhoff’s theorem implies that $\text{rot}_\mu(\tilde{f}) \in \text{rot}(\tilde{f})$ and so in the case when $n = 2$, the convexity of $\text{rot}(\tilde{f})$, implies that for any invariant probability measure $\mu$, $\text{rot}_\mu(\tilde{f}) \in \text{rot}(\tilde{f})$. We will use the following fact repeatedly:

**Lemma 5.4.** Let $f : \mathbb{T}^2 \to \mathbb{T}^2$ be a homeomorphism homotopic to the identity and $\tilde{f} : \mathbb{R}^2 \to \mathbb{R}^2$ be a lift of $f$. Let $r$ be an extreme point of the convex set $\text{rot}(\tilde{f})$. Then there exists an invariant ergodic probability measure $\mu$ on $\mathbb{T}^2$ such that $\text{rot}_\mu(\tilde{f}) = r$.

5.3. Trivializing the rotation sets. The main result of this section is the following.

**Proposition 5.5.** Suppose $B \in \text{SL}_2(\mathbb{Z})$ is Anosov and $\Phi : \Gamma_B \to \text{Homeo}(\mathbb{T}^2)$ is such that:
We conclude that \( f \) is linear. Using that \( f \) preserves the unstable foliation \( F \) of \( A \), we have that \( f(s) = s + v \) for some \( v \in \mathbb{R}^2 \) and some leaf \( s \in F \). Therefore, we have \( f(s) = s + v \) for every \( s \in \mathbb{R}^2 \) because \( f \) commutes with \( \mathbb{Z}^2 \) translations. As for the (irrational) foliation \( F \) the leaves \( s + v \) where \( v \in \mathbb{Z}^2 \) are dense in \( \mathbb{R}^2 \).

Therefore for every \( x \in \mathbb{R}^2 \) we have that \( f^n(x) - x \in \mathbb{R}^2 \) must lie in the line \( x' = \{tu | t \in \mathbb{R}\} \). This together with the convexity of rot(\( f \)) imply Proposition 5.6 easily.

Moreover we have the following:

**Proposition 5.7.** The following dichotomy holds:

1. For every \( v \in \mathbb{Z}^2 \setminus \{0\} \), the rotation set of a lift \( f_v \) of \( \Phi(v) \) is a non-trivial segment.
2. There is a finite index subgroup \( L \subset \mathbb{Z}^2 \) such that for every \( v \in \mathbb{Z}^2 \setminus \{0\} \), the rotation set of a lift \( f_v \) of \( \Phi(v) \) is a point.

**Proof.** Assume that for some \( w \in \mathbb{Z}^2 \setminus \{0\} \), the rotation set of a lift \( f_w \) of \( \Phi(w) \) is a point. We will show in that case that there is a finite index subgroup \( L < \mathbb{Z}^2 \) such that for every \( v \in \mathbb{Z}^2 \setminus \{0\} \), the rotation set of a lift \( f_v \) of \( \Phi(v) \) is a point.

Observe that for every \( v \in \mathbb{Z}^2 \) and any integer \( n \), the rotation set of a lift of \( f_v \) of \( \Phi(v) \) consists of a single point if and only if the rotation set of \( f_{nw} \) consists of a single point.

As \( \Phi(B(w)) \) is conjugate to \( \Phi(w) \) by \( A \) using the group relation in \( \Gamma_B \), the rotation set of a lift of \( \Phi(B(w)) \) must also consist of a single point. Observe that as \( B \) is Anosov, the vectors \( w \) and \( B(w) \) generate a group \( L \) of finite index in \( \mathbb{Z}^2 \), and so it is enough to show that the rotation set of \( \Phi(v) \) for any \( v \in \mathbb{Z}^2 \setminus \{0\} \) consists of a single point.

Therefore, let \( v = nw + mB(w) \), then each of the rotation sets of \( f_{nw} \) and \( f_{-mB(w)} \) consists of a point, so does the rotation set of \( f_{nw+mB(w)} \), because if it was a segment, then we obtain a contradiction to the following:

**Proposition 5.8.** Let \( v, w \in \mathbb{Z}^2 \), if the rotation set of a lift \( f_v \) of \( \Phi(v) \) consists of a single point and the rotation set of \( f_w \) is a non-trivial segment, then the rotation set of a lift \( f_{v+w} \) of \( \Phi(v+w) \) is a non-trivial segment.
Proof. Let \( \text{rot}(f_v) = \{ \alpha \} \), and observe that \( f_v f_w \) is a lift of \( \Phi(v + w) \). By the definition of rotation set, if \( \beta_1 \) and \( \beta_2 \) are two different rotation vectors for \( f_w \), we have two sequences \( \{x_i\}, \{y_i\} \) in \( \mathbb{T}^2 \) and two divergent sequences \( \{n_i\} \) and \( \{m_i\} \) of positive integers such that
\[
\lim \frac{f_w^{n_i}(x_i) - x_i}{n_i} = \beta_1 \quad \text{and} \quad \lim \frac{f_w^{m_i}(y_i) - y_i}{m_i} = \beta_2.
\]
Moreover we have that for every \( x \in \mathbb{R}^2 \), \( f_v^n(x) - x = \alpha \) and this convergence must be uniform for all \( x \in \mathbb{R}^2 \). It follows that
\[
\lim \frac{(f_v f_w)'^n(x_i) - x_i}{n_i} = \lim \frac{f_v^n f_w^{n_i}(x_i) - f_w^{n_i}(x_i)}{n_i} + \lim \frac{f_w^{n_i}(x_i) - x_i}{n_i} = \alpha + \beta_1,
\]
where the first limit holds because of the uniform convergence and similarly \( \lim \frac{(f_v f_w)'^m(y_i) - y_i}{m_i} = \alpha + \beta_2 \) and therefore the rotation set of \( f_v f_w \) has two different vectors and by Proposition 5.6 must be a non-trivial segment.

We define the following subset of the set of \( 2 \times 2 \) real matrices \( \mathcal{M}_{2 \times 2}(\mathbb{R}) \) whose study will be crucial to finish the proof of proposition 5.5.

**Definition 5.9.** For a fixed choice of lifts \( f_{e_1} \) and \( f_{e_2} \), the set of joint rotation numbers \( S_{\text{rot}} \subset \mathcal{M}_{2 \times 2}(\mathbb{R}) \) is defined by:
\[
S_{\text{rot}} := \left\{ (\text{rot}_\mu(f_{e_1}), \text{rot}_\mu(f_{e_2})) \in \mathcal{M}_{2 \times 2}(\mathbb{R}) \mid \mu \text{ is } \mathbb{Z}^2\text{-invariant} \right\}.
\]

Observe that \( S_{\text{rot}} \) must be contained in the bounded square \( \text{rot}(f_{e_1}) \times \text{rot}(f_{e_2}) \subset \mathcal{M}_{2 \times 2}(\mathbb{R}) \) since a choice of lifts \( f_{e_1} \) and \( f_{e_2} \) is fixed. Recall that for a subset \( X \) of \( \mathcal{M}_{2 \times 2}(\mathbb{R}) \), \( X - X := \{ x - y \mid x, y \in X \} \).

Proposition 5.5 follows easily from the following key proposition.

**Proposition 5.10.** The set \( S_{\text{rot}} \) satisfies the following properties:

1. \( S_{\text{rot}} \) is a convex and compact set.
2. For every \( n \in \mathbb{Z} \): if \( R_1, R_2 \in S_{\text{rot}} \), then \( A^n(R_1 - R_2) B^{-n} \in S_{\text{rot}} - S_{\text{rot}} \).
3. One of the following holds:
   a. The set \( \text{rot}(f_{e_1}) \times \text{rot}(f_{e_2}) \) consists of a single point.
   b. \( |\lambda_A| = |\lambda_B| \) and \( S_{\text{rot}} \) is one of the diagonals of \( \text{rot}(f_{e_1}) \times \text{rot}(f_{e_2}) \).

We postpone the proof of Proposition 5.10 to the next subsection.

### 5.4. Description of the rotation sets

In this section, we give the proof of Proposition 5.10.

**Lemma 5.11.** The action of \( \mathbb{Z}^2 \) on \( \mathbb{T}^2 \) lifts to an action on the universal covering \( \mathbb{R}^2 \) of \( \mathbb{T}^2 \). In other words, for every \( v \in \mathbb{Z}^2 \), we can choose a lift \( f_v : \mathbb{R}^2 \to \mathbb{R}^2 \) of \( \Phi(v) \), so that we have
\[
f_{v+w} = f_v f_w, \text{ for all } v, w \in \mathbb{Z}^2.
\]
Proof. If $v, w \in \mathbb{Z}^2$ and $f_v, f_w$ are lifts of $\Phi(v), \Phi(w)$ respectively, we have that $f_v f_w f_v^{-1} f_w^{-1}$ is a lift of the identity so it must be a translation by some $k \in \mathbb{Z}^2$, the action of $\mathbb{Z}^2$ lifts if and only if $k = 0$ for all $v, w$.

From the proof of Proposition 5.6, we have $f_v(\ell) = \ell + \alpha_v$ for all $\ell \in \mathcal{F}$ and $\alpha_v \in \mathbb{R}^2$ depending only on $v \in \mathbb{Z}^2$, so we have $f_v^{-1}(\ell) = \ell - \alpha_v$. This implies that $f_v f_w f_v^{-1} f_w^{-1}(\ell) = \ell$ for all $\ell \in \mathcal{F}$. In particular, $f_v f_w f_v^{-1} f_w^{-1}$ preserves the leaf $\ell(0) \in \mathcal{F}$ passing through 0. However, as the foliation is irrational $\ell(0) \cap \mathbb{Z}^2 = \{0\}$, which implies that $k = 0$. \qed

In the following discussion we fix such choice of lifts.

For a $\mathbb{Z}^2$-invariant probability measure $\mu$ on $\mathbb{T}^2$ and each $i = 1, 2$, we consider the vector $\text{rot}_\mu(f_{e_i})$ as a column vector and so we construct a $2 \times 2$ matrix in $M_{2 \times 2}(\mathbb{R})$ given by:

$$(\text{rot}_\mu(f_{e_1}), \text{rot}_\mu(f_{e_2})).$$

Observe also that as $A^{-1} \Phi(e_i) A = \Phi(B^{-1}(e_i))$, then $A^{-1} f_{e_i} A$ is also a lift of $\Phi(B^{-1}(e_i))$ and so there exists $w_i \in \mathbb{Z}^2$ such that

$$f_{e_i} \circ A = T_{w_i} \circ A \circ f_{B^{-1}(e_i)}. \quad (6)$$

One can easily check that for any $C \in \text{SL}_2(\mathbb{Z})$, the linearity equation (5) implies that:

$$(\text{rot}_\mu(f_{C e_1}), \text{rot}_\mu(f_{C e_2})) = (\text{rot}_\mu(f_{e_1}), \text{rot}_\mu(f_{e_2})) C. \quad (7)$$

We have the following:

**Proposition 5.12.** If $\mu$ is a $\mathbb{Z}^2$-invariant probability measure on $\mathbb{T}^2$, then $A_* \mu$ is $\mathbb{Z}^2$-invariant and we have the following:

$$(\text{rot}_{A_* \mu}(f_{e_1}), \text{rot}_{A_* \mu}(f_{e_2})) = A(\text{rot}_\mu(f_{e_1}), \text{rot}_\mu(f_{e_2})) B^{-1} + (w_1, w_2),$$

where $w_1, w_2 \in \mathbb{Z}^2$ are as in equation (6).

**Proof.** We first show that $A_* \mu$ is $\mathbb{Z}^2$-invariant. Since $\mu$ is $\mathbb{Z}^2$ invariant, we have $\Phi(v)_* \mu = \mu$ for all $v \in \mathbb{Z}^2$. So we get

$$\Phi(v)_*(A_* \mu)(E) = A_* \mu(\Phi(v)^{-1}(E)) = \mu(A^{-1} \Phi(v)^{-1}(E)) = \mu(\Phi(B^{-1}v)A^{-1}(E)) = \mu(A^{-1}(E))$$

for all Borel set $E$ and all $v \in \mathbb{Z}^2$. This gives the $\mathbb{Z}^2$-invariance of $A_* \mu$.

We next prove the identity. By definition we have for every $i = 1, 2$:

$$\text{rot}_{A_* \mu}(f_{e_i}) = \int_{\mathbb{T}^2} f_{e_i}(\tilde{A} x) - \tilde{A} x \, d\mu$$

$$= \int_{\mathbb{T}^2} f_{e_i}(A \tilde{x}) - A \tilde{x} \, d\mu$$

$$= \int_{\mathbb{T}^2} T_{w_i} \circ A f_{B^{-1}(e_i)}(\tilde{x}) - A(\tilde{x}) \, d\mu$$

$$= \int_{\mathbb{T}^2} A(f_{B^{-1}(e_i)}(\tilde{x}) - \tilde{x}) \, d\mu + w_i d\mu$$

$$= A \int_{\mathbb{T}^2} (f_{B^{-1}(e_i)}(\tilde{x}) - \tilde{x}) \, d\mu + w_i$$

$$= A(\text{rot}_\mu(f_{B^{-1}(e_i)}) + w_i.$$
And so to finish the proof, we must only show that:

\[
(rot(\mu_f B^{-1} e_1), rot(\mu_f B^{-1} e_2)) = (rot(\mu_f e_1), rot(\mu_f e_2)) B^{-1},
\]

which follows immediately from equation (7).

\[\Box\]

**Proof of Proposition 5.10.** Item (1) follows from the convexity and compactness of \( \mathbb{Z}^2 \) invariant probability measures and the fact that the map \( \text{rot}_\mu(\cdot) \) depends linearly in \( \mu \). Item (2) follows from Proposition 5.12. We remark that the integer vectors \((w_1, w_2)\) in Proposition 5.12 might depend on \( n \) if we iterate \( A^n \) and \( B^n \). However, they disappear after taking difference \( S_{\text{rot}} - S_{\text{rot}} \).

We consider item (3). Suppose that \( r \in \mathbb{R}^2 \) is an extreme point of \( \text{rot}(f_{e_1}) \), we will show that there exist \((r, s) \in S_{\text{rot}} \) for some \( s \in \mathbb{R}^2 \). Let \( \mu_1 \) a probability measure which is \( \Phi(e_1) \)-invariant on \( \mathbb{T}^2 \) such that \( \text{rot}(\mu_f e_1) = r \), if we consider the sequence of measures

\[
\nu_n := \frac{1}{n} \sum_{i=1}^{n} \Phi(e_2)^i(\mu_1)
\]

and \( \nu \) a limit measure of a subsequence of \( \nu_n \), then \( \nu \) is \( \mathbb{Z}^2 \)-invariant and we have \( \text{rot}_\nu (f_{e_1}) = r \). Moreover, if we denote \( s = \text{rot}_\nu (f_{e_2}) \), then we have that \((r, s) \in S_{\text{rot}} \) by definition.

This implies that \( S_{\text{rot}} \) has at least a point in two opposite sides of the square \( \text{rot}(f_{e_1}) \times \text{rot}(f_{e_2}) \), arguing similarly for \( \text{rot}(f_{e_2}) \) one can show there are points in the other pair of opposite sides of the square \( \text{rot}(f_{e_1}) \times \text{rot}(f_{e_2}) \). In conclusion the set \( S_{\text{rot}} \) contains points in each side of the square \( \text{rot}(f_{e_1}) \times \text{rot}(f_{e_2}) \). From the convexity of \( S_{\text{rot}} \) it follows that either \( S_{\text{rot}} \) is a one of the diagonals of \( \text{rot}(f_{e_1}) \times \text{rot}(f_{e_2}) \) or it has non-empty interior and then item (3) follows from the following Proposition 5.13.

\[\Box\]

**Proposition 5.13.** The set \( S_{\text{rot}} \) has empty interior. Moreover if \( |\lambda_A| \neq |\lambda_B| \), the set \( \text{rot}(f_{e_1}) \times \text{rot}(f_{e_2}) \) consists of a single point.

**Proof.** Let’s show that \( S_{\text{rot}} \) has empty interior. Assume by contradiction that \( S_{\text{rot}} \) has non empty interior. As we know that for every \( i = 1, 2 \), points in the set \( \text{rot}(f_{e_i}) \) are of the form \( \alpha e_i + t u \), where \( t \in [t_{e_1}, t_{e_2}] \). If \( S_{\text{rot}} \) has non-empty interior we conclude there exists \( \varepsilon > 0 \) such that for any \( t_1, t_2 \in [-\varepsilon, \varepsilon] \) we have that \((t_1 u, t_2 u) \in S_{\text{rot}} - S_{\text{rot}} \) and so by item (2) in Proposition 5.10 we have \( A^n(t_1 u, t_2 u) B^{-n} \in S_{\text{rot}} - S_{\text{rot}} \). We will show that this cannot be true for \( n \) large enough.

Diagonalizing the matrix \( B \), we have that \( B = Q \text{diag}(\lambda_B, \lambda_B^{-1}) Q^{-1} \) for \( Q = (v_1, v_2) \) where \( v_1 \) and \( v_2 \) are the eigenvectors of \( B \). We then have:

\[
A^n(t_1 u, t_2 u) B^{-n} = (t_1 A^n u, t_2 A^n u) B^{-n} = \lambda_A^n (t_1 u, t_2 u) B^{-n} = \lambda_A^n (t_1 u, t_2 u) Q \text{diag}(\lambda_B^{-n}, \lambda_B^n) Q^{-1} = (t_1 u, t_2 u)(v_1, v_2) \text{diag}((\lambda_A \lambda_B^{-1})^n, (\lambda_A \lambda_B)^n) Q^{-1} = (t_1 u, t_2 u)((\lambda_A \lambda_B^{-1})^n v_1, (\lambda_A \lambda_B)^n v_2) Q^{-1}.
\]
One can show easily that the vector \((t_1 \mathbf{u}, t_2 \mathbf{u})((\lambda A\lambda B)^n \mathbf{v}_2)\) diverges for an appropriate choice of \(t_1, t_2 \in [-\varepsilon, \varepsilon]\) and so \(A^n(t_1 \mathbf{u}, t_2 \mathbf{u})B^{-n}\) cannot lie in the compact set \(S_{\text{rot}} - S_{\text{rot}}\) for every \(n > 0\). Therefore \(S_{\text{rot}}\) must have empty interior.

In the case where \(|\lambda_A| \neq |\lambda_B|\). As we showed that \(S_{\text{rot}}\) must be one of the diagonals of \(\text{rot}(f_{e_1}) \times \text{rot}(f_{e_2})\) we must only show that \(S_{\text{rot}}\) does not consist of a single point. Arguing by contradiction as before if \(S_{\text{rot}}\) was not a single point, then \((t_1 \mathbf{u}, t_2 \mathbf{u}) \in S_{\text{rot}} - S_{\text{rot}}\) for some \(t_1, t_2 \in \mathbb{R}\) where at least one of them is different than zero. From the calculations above we have:

\[
A^n(t_1 \mathbf{u}, t_2 \mathbf{u})B^{-n} = (t_1 \mathbf{u}, t_2 \mathbf{u})((\lambda A\lambda B)^n \mathbf{v}_1, (\lambda_A\lambda B)^n \mathbf{v}_2)Q^{-1},
\]

which implies in particular that \((t_1 \mathbf{u}, t_2 \mathbf{u})((\lambda A\lambda B)^n \mathbf{v}_2)\) is bounded and so by taking \(n\) large we have \((t_1, t_2)\) is orthogonal to \(\mathbf{v}_2\). We also have that \((t_1 \mathbf{u}, t_2 \mathbf{u})((\lambda A\lambda B)^n \mathbf{v}_1)\) is bounded and so taking \(n\) or \(-n\) large depending whether \(|\lambda_1| > |\lambda_2|\) or not, we have that \((t_1, t_2)\) is also orthogonal to \(\mathbf{v}_1\) and so \(t_1 = t_2 = 0\), a contradiction.

\[\square\]

6. Case \(|\lambda_A| > |\lambda_B|\).

In this section, we prove Theorem 1.1 in the case where \(|\lambda_A| > |\lambda_B|\), more concretely, we show the following proposition under the assumption that \(|\lambda_A| > |\lambda_B|\).

**Theorem 6.1.** Suppose that \(B \in \text{SL}_2(\mathbb{Z})\) is Anosov and \(\Phi : \Gamma_B \to \text{Diff}(\mathbb{T}^2)\) is such that the diffeomorphism \(\Phi(B)\) is an Anosov diffeomorphism of \(\mathbb{T}^2\) homotopic to \(A \in \text{SL}_2(\mathbb{Z})\) with \(|\lambda_A| \geq |\lambda_B|\).

Then \(\Phi\) is topologically conjugate to an affine action of \(\Gamma_B\) of the form (1.1) up to finite index. More concretely, there exist a finite index subgroup \(\Gamma' \subset \Gamma_B\) and \(g \in \text{Homeo}(\mathbb{T}^2)\) such that \(g\Phi(\Gamma')g^{-1}\) coincides with an action of the form (1.1).

Theorem 6.1 in the case \(|\lambda_A| > |\lambda_B|\) follows from Proposition 5.5 and the following proposition.

The case in Theorem 6.1 when \(|\lambda_A| = |\lambda_B|\) will be consider in the next section.

**Proposition 6.2.** Suppose \(\Gamma_B = \mathbb{Z} \times \mathbb{Z}^2\) is an ABC group as in Theorem 6.1 and \(\Phi : \Gamma_B \to \text{Homeo}(\mathbb{T}^2)\) is such that:

1. \(A = \Phi(B)\) is an Anosov linear map with \(|\lambda_A| \geq |\lambda_B|\).
2. \(\Phi(e_1)\) and \(\Phi(e_2)\) are homotopic to the identity.
3. For \(i = 1, 2\), if \(f_{e_i} : \mathbb{R}^2 \to \mathbb{R}^2\) is any lift of \(\Phi(e_i)\), the rotation set \(\text{rot}(f_{e_i})\) consists of a single point.

Then \(\Phi\) is an affine action.

**Proof.** It is enough to show that \(f_{e_i}, i = 1, 2,\) is a translation of \(\mathbb{R}^2\). Indeed, for any \(v = (v_1, v_2) \in \mathbb{Z}^2\), we have \(f_v\), the lift of \(\Phi(v)\), equals \(f_{e_1}^{v_1}f_{e_2}^{v_2} + k\), for some \(k \in \mathbb{Z}^2\). Since \(f_{e_i}\) is a translation, we get that \(f_v\) hence \(\Phi(v)\) are also translations.

Suppose that \(x, y \in \mathbb{R}^2\) are such that \(f_{e_1}(x) = x + p\) and \(f_{e_1}(y) = y + q\) for some \(p, q \in \mathbb{R}^2\), we must show that \(p = q\). We have the following:

\[
A^n f_{e_1}(x) = A^n(x) + A^n(p),
\]
\[
A^n f_{e_1}(y) = A^n(y) + A^n(q).
\]
We have that both $A^n f_{e_1} A^{-n}$ and $f_{B^n(e_1)}$ are lifts of $\Phi(B^n(e_1))$, therefore they differ by a translation and so $A^n f_{e_1} = T_w f_{B^n(e_1)} A^n$ for some translation $T_w$ for some $w \in \mathbb{Z}^2$, and so:

$$A^n f_{e_1}(x) = f_{B^n(e_1)} A^n(x) + w,$$

$$A^n f_{e_1}(y) = f_{B^n(e_1)} A^n(y) + w.$$

Therefore from the previous equations:

$$f_{B^n(e_1)} A^n(x) - f_{B^n(e_1)} A^n(y) = A^n(x) + A^n(p) - (A^n(y) + A^n(q))$$

and so we obtain:

$$\frac{(f_{B^n(e_1)} A^n(x) - A^n(x)) - (f_{B^n(e_1)} A^n(y) - A^n(y))}{|B^n(e_1)|} = A^n(p - q). \quad (8)$$

Since we have $f_v(x + k) = f_v(x) + k$, $k \in \mathbb{Z}^2$ by the definition of a lift, in the expression $(f_{B^n(e_1)} A^n(x) - A^n(x))$, we may assume that $A^n: \mathbb{T}^2 \to \mathbb{T}^2$. Suppose $B^n = \begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix}$, then $B^n(e_1) = (a_n, c_n)$ and $f_{B^n(e_1)} = f_{e_1} f_{e_2}^c$. Denoting $z_n = A^n(x)$ or $A^n(y) \in \mathbb{T}^2$, we have

$$f_{B^n(e_1)}(z_n) - z_n = f_{e_1} f_{e_2}^c (z_n) - z_n$$

$$= f_{e_1}^a f_{e_2}^c (z_n) - f_{e_2}^c (z_n) + f_{e_2}^c (z_n) - z_n$$

$$= (f_{e_1}^a \Phi(c_n e_2)(z_n) - \Phi(c_n e_2)(z_n)) + (f_{e_2}^c(z_n) - z_n).$$

The sequences $(a_n), (c_n), (|B^n(e_1)|)$ all go to infinity with a rate proportional to $|\lambda_B|^n$. Using the fact that $f_{e_i}$, $i = 1, 2$, has one point in its rotation set, by equation $(4)$ and the uniformity of the convergence we get

$$\frac{1}{|B^n(e_1)|} [(f_{e_1}^a \Phi(c_n e_2)(x_n) - \Phi(c_n e_2)(x_n)) - (f_{e_1}^a \Phi(c_n e_2)(y_n) - \Phi(c_n e_2)(y_n))] \to 0,$$

$$\frac{1}{|B^n(e_1)|} [(f_{e_2}^c(x_n) - x_n) - (f_{e_2}^c(y_n) - y_n)] \to 0,$$

where $x_n = A^n(x)$, $y_n = A^n(y) \in \mathbb{T}^2$. This shows that the LHS of $(8)$ vanishes. So we have

$$\lim_{n \to \infty} \frac{A^n(p - q)}{\lambda_B^n} = 0.$$

But this can happen only if $p - q$ lies in the one dimensional eigenspace of $\lambda_A^{-1}$ since we have $|\lambda_A| \geq |\lambda_B|$. Applying the same argument with $B^{-1}$ we have that $p - q$ lies in the one dimensional eigenspace of $\lambda_A$, therefore $p - q = 0$ as we wanted. This proves that $f_{e_1}$ is a translation. Similarly, we get that $f_{e_2}$ is a translation.

$\square$

7. Case $|\lambda_A| = |\lambda_B|$

In this section we finish the proof of Theorem 6.1 dealing with the remaining case where $|\lambda_A| = |\lambda_B|$. Recall we assume that $\Phi: \Gamma_B \to \text{Homeo}(\mathbb{T}^2)$ is an action satisfying the following:

1. $A = \Phi(B)$ is an Anosov linear map with $|\lambda_A| = |\lambda_B|$. 

For every $v \in \mathbb{Z}^2$, $\Phi(v)$ is isotopic to the identity and $f_v : \mathbb{R}^2 \to \mathbb{R}^2$ denotes some lift of $\Phi(v)$.

Observe that we cannot use Proposition 6.2 because we have not shown the rotation set of elements of $\mathbb{Z}^2$ consists of a single point in the case $|\lambda_A| = |\lambda_B|$. We will argue by contradiction to show that for all $v \in \mathbb{Z}^2$ the rotation set of $\Phi(v)$ consists of a single point. Therefore, we will suppose that $\text{rot}(f_v)$ is not a single point and in that case, recall we have shown in Proposition 5.6 and 5.7 that $\text{rot}(f_v)$ must be a non-trivial segment for every $v \in \mathbb{Z}^2$. Let $e_1, e_2$ be the standard basis of $\mathbb{Z}^2$ and recall the joint rotation set $S_{\text{rot}}$ defined in Section 5.3. By Proposition 5.10, we can assume that $S_{\text{rot}}$ is one of the diagonals of the square $\text{rot}(f_{e_1}) \times \text{rot}(f_{e_2})$. Moreover, a point in the segment $S_{\text{rot}}$ has the form $\alpha + (au, bu)$, $\alpha \in \mathbb{R}^2 \times \mathbb{R}^2$, $a/b \notin \mathbb{Q}$ arguing as Proposition 5.13. Indeed, if $a/b \in \mathbb{Q}$, then following the proof of Proposition 5.13, we will see that $((au, bu))(\lambda_A \lambda_B)^n v_2$ diverges as $n \to \infty$ since we have $(a, b) \cdot v_2 \neq 0$ due to the irrationality of $v_2$. Then the argument of Proposition 5.13 goes through and we get $\text{rot}(f_{e_1}) \times \text{rot}(f_{e_2})$ is a single point. The irrationality of $a/b$ implies that the segment $S_{\text{rot}}$ cannot form a loop on the torus $\mathbb{T}^2 \times \mathbb{T}^2(> \text{rot}(f_{e_1}) \times \text{rot}(f_{e_2}))$.

We start by showing the existence of a nice $\Gamma_B$-invariant measure $\nu$ where the asymptotic behaviour of the action of $\mathbb{Z}^2$ is given by the rotation number with respect to $\nu$. We will then use this measure together with a modification of Proposition 6.2 to obtain a contradiction.

**Proposition 7.1.** There exists a $\Gamma_B$-invariant probability measure $\nu$ on $\mathbb{T}^2$ such for any $i = 1, 2$, any lift $f_{e_i} : \mathbb{R}^2 \to \mathbb{R}^2$ of $\Phi(e_i)$ and $\nu$ a.e. $x \in \mathbb{T}^2$ we have:

$$\lim_{n \to \infty} \frac{f_{e_i}^n(x) - x}{n} = \text{rot}_{\nu}(f_{e_i}) \mod \mathbb{Z}^2.$$ 

**Proof.** We will prove the proposition for the coordinate vectors $v = e_1, e_2$.

Recall also that if $\mu$ is a $\mathbb{Z}^2$-invariant measure, then $\text{rot}_{\mu}$ is also $\mathbb{Z}^2$-invariant and we have an action of $A$ in $S_{\text{rot}} (\mod \mathbb{Z}^2 \times \mathbb{Z}^2)$ which is described by Proposition 5.12 and states that:

$$(\text{rot}_{A, \mu}(f_{e_1}), \text{rot}_{A, \mu}(f_{e_2})) = A(\text{rot}_{\mu}(f_{e_1}), \text{rot}_{\mu}(f_{e_2})) B^{-1} \mod \mathbb{Z}^2 \times \mathbb{Z}^2.$$ 

Observe that the action in $S_{\text{rot}}$ is linear and therefore is clearly continuous. Therefore the two endpoints of the interval $S_{\text{rot}}$ are either fixed or permuted and by replacing $A$ with $A^2$ (this is the same as replacing $\Gamma_B$ with the finite index subgroup $\Gamma_B^2$) we will assume that the endpoints of $S_{\text{rot}}$ are fixed. We fix an endpoint of $S_{\text{rot}}$ and take a $\mathbb{Z}^2$-ergodic invariant measure $\mu$ such that $(\text{rot}_{\mu}(f_{e_1}), \text{rot}_{\mu}(f_{e_2}))$ is equal to this endpoint, this implies in particular that both $\text{rot}_{\mu}(f_{e_1})$ and $\text{rot}_{\mu}(f_{e_2})$ are extremal points of $\text{rot}(f_{e_1})$ and $\text{rot}(f_{e_2})$ respectively.

Consider the sequence of measures $\mu_n := \frac{1}{n} \sum_{j=1}^{n} A_j^i \mu$. From the equality

$$(\text{rot}_{A, \mu}(f_{e_1}), \text{rot}_{A, \mu}(f_{e_2})) = (\text{rot}_{\mu}(f_{e_1}), \text{rot}_{\mu}(f_{e_2})) \mod \mathbb{Z}^2 \times \mathbb{Z}^2,$$

we get by linearity that $(\text{rot}_{\mu_n}(f_{e_1}), \text{rot}_{\mu_n}(f_{e_2})) = (\text{rot}_{\mu}(f_{e_1}), \text{rot}_{\mu}(f_{e_2}))$. We let $\nu$ be any weak *-limit of $(\mu_n)$. Observe that $\nu$ is $\mathbb{Z}^2$- and $A$-invariant and therefore is $\Gamma_B$-invariant. By continuity we also have

$$(\text{rot}_{\nu}(f_{e_1}), \text{rot}_{\nu}(f_{e_2})) = (\text{rot}_{\mu}(f_{e_1}), \text{rot}_{\mu}(f_{e_2})).$$

It is possible that $\nu$ is not $\Phi(e_1)$-ergodic, but nonetheless for $\nu$-a.e. every $\Phi(e_1)$-ergodic component $\eta$ of $\nu$ we must have $\text{rot}_\eta(f_{e_1}) = \text{rot}_{\nu}(f_{e_1})$ because $\text{rot}_{\nu}(f_{e_1})$ is an extremal point of $\text{rot}(f_{e_1})$. From
Birkhoff ergodic theorem we conclude that:
\[ \lim_{n \to \infty} \frac{f_{e_1}^n(x) - x}{n} = \text{rot}_\nu(f_{e_1}) \mod \mathbb{Z}^2 \]
for \( \eta \) a.e. \( x \in \mathbb{T}^2 \) and because this is true for a.e. every ergodic component \( \eta \), it is also true for \( \nu \) a.e. \( x \in \mathbb{T}^2 \) as we wanted. We can argue similarly for \( \Phi(e_2) \).

\[ \square \]

We will now show that for a measure \( \nu \) as in Proposition 7.1, the action of \( \mathbb{Z}^2 \) restricted to the support of \( \nu \) is by translations, the proof is very similar to the proof of Proposition 6.2. This will lead us later to a contradiction.

**Proposition 7.2.** For every \( i = 1, 2 \) and every \( \delta, \varepsilon > 0 \), there exists a set \( C \) with \( \nu(C) > 1 - \delta \) such that for every \( x, y \in C \) we have:
\[ |(f_{e_i}(x) - x) - (f_{e_i}(y) - y)| < \varepsilon. \]

**Proof.** The proof follows closely the argument in Proposition 6.2. We only prove it for \( \nu = e_1 \).

Suppose that \( p_x, p_y \in \mathbb{R}^2 \) are such that \( f_{e_1}(x) = x + p_x \) and \( f_{e_1}(y) = y + p_y \) for each \( x, y \in \mathbb{R}^2 \). We will find a set \( C \) with \( \mu(C) > 1 - \delta \) such that \( |p_x - p_y| < \varepsilon \) for \( x, y \in C \). Observe that because \( \mathbb{Z}^2 \) preserves the unstable foliation, there exists a fixed \( \alpha \in \mathbb{R}^2 \) and \( t_x, t_y \in \mathbb{R} \) such that \( p_x = \alpha + t_x u \) and \( p_y = \alpha + t_y u \) (see the proof of Proposition 5.6).

We have the following:
\[
A^n f_{e_1}(x) = A^n(x) + A^n(p_x), \\
A^n f_{e_1}(y) = A^n(y) + A^n(p_y).
\]

As in proposition 6.2, we have that for some \( w \in \mathbb{Z}^2 \) we have:
\[
A^n f_{e_1}(x) = f_{B^n(e_1)} A^n(x) + w, \\
A^n f_{e_1}(y) = f_{B^n(e_1)} A^n(y) + w.
\]

Therefore from the previous equations:
\[
f_{B^n(e_1)} A^n(x) - f_{B^n(e_1)} A^n(y) = A^n(x) + A^n(p_x) - (A^n(y) + A^n(p_y))
\]
and if we let \( x_n := A^n(x) \) and \( y_n := A^n(y) \) we obtain:
\[
\left| \frac{(f_{B^n(e_1)}(x_n) - x_n) - (f_{B^n(e_1)}(y_n) - y_n)}{\lambda_B^n} \right| = \left| \frac{A^n(p_x - p_y)}{\lambda_B^n} \right| = \left| \frac{A^n(t_x u - t_y u)}{\lambda_A^n} \right| = |t_x - t_y|. \tag{9}
\]

Suppose \( B^n = \begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix} \), then \( B^n(e_1) = (a_n, c_n) \) and \( f_{B^n(e_1)} = f_{e_1} \circ f_{e_2} \). We can choose a large constant \( M > 0 \) such that the quantities \( \frac{|a_n|}{|\lambda_B|^1}, \frac{|b_n|}{|\lambda_B|^1}, \frac{|c_n|}{|\lambda_B|^1}, \frac{|d_n|}{|\lambda_B|^1} \) are bounded above by \( M \) and below by \( 1/M \).

From Proposition 7.1, we can choose \( N_\varepsilon \) and \( C_\delta \) such that \( \mu(C_\delta) > 1 - \frac{\delta}{3} \), every \( z \in C_\delta \) and every \( |n| > N_\varepsilon \) and \( i = 1, 2 \) we have:
Therefore using (10), the fact that \( |c_n| > N_\varepsilon \) and \( |a_n| > N_\varepsilon \) and define
\[
C := C_\delta \cap A^{-n}(C_\delta) \cap A^{-n}f^{-c_n}(C_\delta).
\]

Observe that \( \mu(C) > 1-\delta \) as we wanted. Let \( x, y \in C \) and define \( x'_n := f^{c_n}_{e_2}(x_n) \) and \( y'_n := f^{c_n}_{e_2}(y_n) \).

We have by definition of \( C \) that the points \( x_n, x'_n, y_n, y'_n \in C \), and easy computation shows that:
\[
f^{B^n(e_1)}(x_n) - x_n = f^{a_n}_{e_1}(x'_n) - x'_n + f^{c_n}_{e_2}(x_n) - x_n
\]
\[
f^{B^n(e_1)}(y_n) - y_n = f^{a_n}_{e_1}(y'_n) - y'_n + f^{c_n}_{e_2}(y_n) - y_n.
\]

Therefore using (10), the fact that \( |a_n|, |c_n| > N_\varepsilon \) and the triangle inequality we obtain:
\[
\left| \frac{f^{B^n(e_1)}(x_n) - x_n - f^{B^n(e_1)}(y_n) - y_n}{|\lambda^n_B|} \right| \leq \frac{M|f^{a_n}_{e_1}(x'_n) - x'_n - f^{a_n}_{e_1}(y'_n) - y'_n|}{|a_n|} + \frac{M|f^{c_n}_{e_2}(x_n) - x_n - f^{c_n}_{e_2}(y_n) - y_n|}{|c_n|} < \varepsilon/2 + \varepsilon/2 = \varepsilon.
\]

Therefore, by equation (9) we obtain \( |(f^{e_1}(x) - x) - (f^{e_1}(y) - y)| = |t_x - t_y| < \varepsilon \) for every \( x, y \in C \) as we wanted. We can argue similarly for \( f^{e_2} \).

\( \square \)

**Proposition 7.3.** Suppose that \( \nu \) is a \( \Gamma_B \)-invariant probability measure as in Proposition 7.1. We have that for \( \nu \) a.e. \( x \in \mathbb{T}^2 \) and every \( v \in \mathbb{Z}^2 \):
\[
f_\nu(x) = x + \text{rot}_\nu(f_\nu).
\]

That is, \( \mathbb{Z}^2 \) acts as a translation for \( \nu \) a.e. \( x \in \mathbb{T}^2 \).

**Proof.** We will only prove it in the case \( v = e_1, e_2 \). It follows from Proposition 7.2 that the function \( f^{e_1}(x) - x \) must be constant for \( \nu \) a.e. \( x \in \mathbb{T}^2 \) and it follows easily by integration that such constant must be equal to \( \text{rot}_\nu(f^{e_1}) \) finishing the proof of Proposition 7.3. \( \square \)

We are now ready to complete the proof of Theorem 6.1.

**Completing the proof of Theorem 6.1.** We will now obtain a contradiction to the non-triviality of the rotation sets and show that \( \mathbb{Z}^2 \) acts by translations on \( \mathbb{T}^2 \). Let \( \nu \) be a measure as in Proposition 7.1. Observe that by continuity, Proposition 7.3 implies the action of \( \mathbb{Z}^2 \) restricted to the support \( \text{supp}(\nu) \) is by translations. Moreover we have that \( \text{supp}(\nu) \) is also invariant by the Anosov linear transformation \( A \) because \( \nu \) is \( \Gamma_B \)-invariant.

If the matrix \( (\text{rot}_\nu(f^{e_1}), \text{rot}_\nu(f^{e_2})) \) contains an irrational entry we have that \( \text{supp}(\nu) \) is invariant by an irrational translation, so \( \text{supp}(\nu) \) is either \( \mathbb{T}^2 \) or contains a circle. In the latter case, the circle intersects the stable manifold of 0 transversally since the eigenvectors of \( A \) have irrational slopes. We apply the \( \lambda \)-lemma to obtain that under iterates the image of the circle will approach the unstable manifold of 0 and get stretched exponentially so will become dense on \( \mathbb{T}^2 \). Using the invariance by
the Anosov $A$ it follows that $\text{supp}(\nu) = \mathbb{T}^2$. Therefore we conclude that $f_{e_1}, f_{e_2}$ must be translations everywhere.

In the case where $(\text{rot}_\nu(f_{e_1}), \text{rot}_\nu(f_{e_2}))$ has all entries rational, recall that in the proof of Proposition 7.1 the probability measure $\nu$ was constructed in such a way that the matrix $(\text{rot}_\nu(f_{e_1}), \text{rot}_\nu(f_{e_2}))$ is one of the endpoints of the interval $S_{\text{rot}}$. We can also construct in the same way another measure $\nu' \in \mathbb{T}^2$ such that $(\text{rot}_{\nu'}(f_{e_1}), \text{rot}_{\nu'}(f_{e_2}))$ is the other endpoint of $S_{\text{rot}}$, but this matrix cannot have all entries being rational because otherwise $S_{\text{rot}}$ is a segment with rational slope and this contradicts the fact that $\text{rot}(f_{e_1}), \text{rot}(f_{e_2})$ are segments with an irrational direction given by the unstable eigenvector $u$ of $A$. Therefore the matrix $(\text{rot}_{\nu'}(f_{e_1}), \text{rot}_{\nu'}(f_{e_2}))$ contains an irrational entry and in that case $\nu'$ is invariant by an irrational translation and we can argue as before. This gives us a contradiction and finishes the proof of Theorem 6.1.

$\square$

8. CASE $|\lambda_A| < |\lambda_B|$.

Proposition 6.1 proves the case of $|\lambda_A| \geq |\lambda_B|$ of Theorem 1.1. In this section, we prove Theorem 1.1 in the case where $|\lambda_B| > |\lambda_A|$, finishing the proof of Theorem 1.1, we will use the same notation as in Section 5. We recall that in this case we want to obtain a contradiction from the fact that the action is faithful, we will indeed show that for all $v \in \mathbb{Z}^2$ we have $\Phi(v) = \text{Id}$. From Proposition 5.5, we can assume that the rotation set for $v \in \mathbb{Z}^2 \subset \Gamma_B$ every diffeomorphisms $f_v$ consists of a single point in $\mathbb{R}^2$ and we will abuse notation and call such vector $\text{rot}(f_v)$. We have from Proposition 5.12 that the following equation holds:

$$(\text{rot}(f_{e_1}), \text{rot}(f_{e_2}))B = A(\text{rot}(f_{e_1}), \text{rot}(f_{e_2})) + (w_1, w_2)$$

for some $w_1, w_2 \in \mathbb{Z}^2$. Therefore from Proposition 4.1, as $\text{tr}(B) \neq \text{tr}(A)$, we have that both $\text{rot}(f_{e_1})$ and $\text{rot}(f_{e_2})$ are rational vectors. Therefore we can pass to a finite cover of $\Gamma_B$ of the form $B \times m\mathbb{Z}^2$ and assume that the rotation set of every $f_v$ is in $\mathbb{Z}^2$ for all $v \in \mathbb{Z}$. Moreover, we will assume that the rotation set of $f_v$ is $\{0\}$ for all $v$, which follows because we can assume that the lifts $f_{e_1}, f_{e_2}$ have rotation set $\{0\}$ and this imply that $f_v$ has rotation set $\{0\}$ for all $v$. We have the following:

**Lemma 8.1.** The action of $\Gamma_B$ on $\mathbb{T}^2$ lifts to an action of $\Gamma_B$ in the universal covering $\mathbb{R}^2$.

**Proof.** As we observed, we can take lifts of the $\mathbb{Z}^2$ action in such a way that every $f_v$ has as rotation set $\{0\}$, then one can check that $Af_vA^{-1}$ has rotation set $\{0\}$ as well and conclude that

$$Af_vA^{-1} = f_{B(v)}$$

which implies the action $\Gamma_B$ lifts.

$\square$

Therefore, we can assume that $\Gamma_B$ acts on $\mathbb{R}^2$. Moreover, every $f_v$ leaves invariant every leaf of the foliation $F$, this follows because all $f_v$ have rotation set $\{0\}$, and each $f_v$ acts in the set of leaves as a translation, this is because the lifts of any leaf of $F$ are dense and as $f_v$ commutes with deck transformations if $f_v(l) = l + \alpha$ for some leaf $l \in F$, $\alpha \in \mathbb{R}$, then $f_v(l) = l + \alpha$ for all $l \in L$, therefore if a leaf of $F$ is not preserved we will have non-zero rotation set. The following two propositions will be proved in the next subsection.
Proposition 8.2. There exists a continuous one parameter subgroup \( g_t : \mathbb{T}^2 \to \mathbb{T}^2 \) such that

1. For each \( t \in \mathbb{R} \), the map \( g_t : \mathbb{T}^2 \to \mathbb{T}^2 \) is measurable;
2. The map \( t \mapsto g_t \) induces a group homomorphism \( \mathbb{R} \to \text{Homeo}(\mathbb{R}) \) which preserves the unstable foliation of \( \Phi(B) \);
3. We have \( g_1 = \Phi(e_1) \) and \( g_c = \Phi(e_2) \) where \((1, c)\) is an unstable eigenvector of \( B \).

Moreover we have

\[
\Phi(B)g_t\Phi(B)^{-1} = g_{\lambda_B t}
\]

where \( \lambda_B \) is the larger in norm eigenvalue of \( B \).

Proposition 8.3. The flow \( \Phi(B) : \mathbb{R} \times \mathbb{T}^2 \to \mathbb{T}^2 \) is differentiable in \( t \in \mathbb{R} \) and measurable in \( p \in \mathbb{T}^2 \).

We should point out that the smoothness of \( g_t(p) \) as function of \( t \) will be crucial for the following discussion. Assuming the previous propositions, we can define the vector field \( X(p) = \frac{dt}{\lambda_B} |_{t=0}(p) \). Observe that \( X(p) \) is a well defined vector field because of 8.3, but it might not vary continuously as a function of \( p \). From the equality \( \Phi(B)g_t\Phi(B)^{-1} = g_{\lambda_B t} \), we have the following equation:

\[
D_p(\Phi(B))X(p) = \lambda_B X(\Phi(B)p), \quad \forall p \in \mathbb{T}^2.
\]

Definition 8.4. (1) Let

\[
C = \{ p \in \mathbb{T}^2 \mid \Phi(v)(p) = p, \forall v \in \mathbb{Z}^2 \}.
\]

Observe that \( C \) is a closed set which is invariant by the action of \( \Gamma_B \).

2. For every real \( M > 0 \), let

\[
X_M = \left\{ p \in \mathbb{T}^2 \mid \frac{1}{M} < \|X(p)\| < M \right\}
\]

and observe that \( \mathbb{T}^2 = C \cup (\bigcup_{M>0} X_M) \).

Let \( \mu \) be the SRB-ergodic measure preserved by \( \Phi(B) \), we have the following:

Proposition 8.5. If \( \mu(X_M) > 0 \) for some \( M > 0 \), then we have \( |\lambda_A| \geq |\lambda_B| \).

Proof. As \( \Phi(B) \) is topologically conjugate to \( A \), the topological entropy of \( \Phi(B) \) is equal to \( \log(|\lambda_A|) \).

Let \( \chi_\mu \) be the positive Lyapunov exponent of \( \Phi(B) \), if \( \mu(X_M) > 0 \), then for a.e. \( p \in X_M \) we have

\[
\chi_\mu = \lim_{n \to \infty} \log \|D_p(\Phi(B)^n(X(p)))\|^{1/n}
\]

but from equation (11) we have that

\[
D_p(\Phi(B)^n)(X(p)) = \lambda_B^n X(\Phi(B)^n p)
\]

and so Poincaré recurrence implies that \( \chi_\mu = \log(|\lambda_B|) \).

From the Ledrappier-Young entropy formula we have

\[
\log(|\lambda_A|) = h_{top}(\Phi(B)) \geq h_\mu(\Phi(B)) = \dim_\mu(W^u) \log(|\lambda_B|) = \log(|\lambda_B|).
\]

The last equality holds because the dimension \( \dim_\mu(W^u) \) of the measure \( \mu \) along the unstable manifolds is equal to one as \( \mu \) is an SRB-measure.

With this proposition, we complete the proof of Theorem 1.1 in the case of \( |\lambda_B| > |\lambda_A| \).
Proof of Theorem 1.1 in the case of $|\lambda_B| > |\lambda_A|$. If $\mu(X_M) > 0$ for some $M > 0$ we are done. If no such $M$ exists, $\mu(\bigcup_{M>0} X_M) = 0$ and so $\mu(C) = 1$, but as the conditional measures of $\mu$ along unstable manifolds are absolutely continuous with respect to Lebesgue by taking a partition subordinate to $\mathcal{F}$ we must have that there exists a closed non-trivial interval $J$ of a leaf of $\mathcal{F}$ such that Lebesgue a.e. point of $J$ is in $C$, but $C$ is closed, and so $J$ must be contained in $C$. Also as $C$ is invariant by $\Phi(B)$, then $\Phi(B^n)(J)$ is contained in $C$, but then as $\bigcup_{n>0} \Phi(B^n)(J)$ is dense we have $C = \mathbb{T}^2$ and we are done. \hfill \Box

8.1. The reduction to circle maps with Diophantine rotation number. The remaining work is to prove Proposition 8.2 and Proposition 8.3. In this section, we first show how the $\mathbb{Z}^2$ action along each leaf of $\mathcal{W}$ is by translations and induces a circle map with Diophantine rotation number (coming from the group structure of $\Gamma_B$). We then apply the Herman-Yoccoz theory for circle maps to show that the flow $g_t$ is smooth in $t \in \mathbb{R}$. We should also point out that Denjoy’s theorem is not enough for concluding the smoothness of $g_t$.

Lemma 8.6. For every $v \in \mathbb{Z}^2$, we have

$$\lim_{n \to \infty} f_{B^{-n}(v)}(p) = p$$

for all $p \in \mathbb{R}^2$ and moreover such limit converges uniformly with respect to $p$.

Proof. Suppose there is $v \in \mathbb{Z}^2$ and a sequence of points $p_k$ in $\mathbb{R}^2$ such that for some constant $K > 0$ and a subsequence $\{n_k\}$ we have $|f_{B^{-n_k}(v)}(p_k) - p_k| \geq K > 0$. This will lead us to a contradiction because then as both $p_k$ and $f_{B^{-n_k}(v)}(p_k)$ are in the same leaf of the foliation $\mathcal{F}$, we have that as $A$ expands the leaves of the foliation, we have that

$$\lim_{n \to \infty} |A^n f_{B^{-n_k}(v)}(p_k) - A^n(p_k)| = \infty$$

But this is the same as $|f_v A^n(p_k) - A^n(p_k)| \to \infty$ and this is not possible because $|f_v(p) - p|$ is bounded independent of $p \in \mathbb{R}^2$ because $f_v$ is a lift of a map in $\mathbb{T}^2$. \hfill \Box

We take a leaf $\ell \in \mathcal{F}$ and consider an open interval $\mathcal{I}_v \subset \ell$ which is invariant under $f_v$, for some $v \in \mathbb{Z}^2 \setminus \{0\}$, and where there is no fixed point for the $f_v$. If this interval does not exist, that means for any open interval $\mathcal{J}_v$ that is invariant under $f_v$, there is a fixed point of $f_v$ in the interval, which breaks $\mathcal{J}_v$ into two subintervals that are invariant under $f_v$ due to the one-dimensionality. Repeating the argument, we get that there is a dense subset of fixed points in $\mathcal{J}_v$. Then the continuity of $f_v$ implies that each point in the entire interval $\mathcal{J}_v$ is fixed by $f_v$. Arguing this for any open interval $\mathcal{J}_v$ and for all $v \in \mathbb{Z}^2$, we conclude that $\mathbb{Z}^2$ fixes $\ell$. As $\ell$ is dense on the torus, we have that the $\mathbb{Z}^2$ action is trivial. So in the following, we consider the above defined interval $\mathcal{I}_v$.

We have the following analogue of Kopell’s lemma.

Proposition 8.7. The group $\mathbb{Z}^2$ acts freely in $\mathcal{I}_v$, more precisely, if $f_w(p) = p$ for some $p \in \mathcal{I}_v$ and some $w \neq v \in \mathbb{Z}^2 \setminus \{0\}$, then $f_w = \text{id}$.

Proof. If $f_w$ has a fixed point $p$ inside $\mathcal{I}_v$, we can assume that $p$ is an endpoint of an interval $J \subset \mathcal{I}_v$ where $f_v$ is conjugate to a translation on $\mathbb{R}$. Since $f_v$ is known to have no fixed point in $\mathcal{I}_v$, we get that $v$ and $w$ are linearly independent. We decompose $mB^{-n}(v) = a_nv + b_nw$ for some $m \in \mathbb{Z}$
depending only on $w, v$ and $a_n, b_n \in \mathbb{Z}$. So we get
\[ f_{mB^{-n}(v)}p = f_{a_n v}f_{b_n w}p = f_{v a_n}^p. \]

For $n$ large, the LHS converges to $p$ by Lemma 8.6, while the RHS moves away from $p$ since $f_v$ is monotone on $I_v$ and each point converges to one of the endpoints under iterations of $f_v$. This contradiction verifies our claim.

\[ \square \]

It follows that the $\mathbb{Z}^2$ action on any interval like $I$ is conjugate to a group of translations (Corollary 4.1.4 and Theorem 4.1.37 of [35]), therefore there is a homeomorphism $h : I \to \mathbb{R}$ such that $T_v := hf_vh^{-1}$ is a translation and suppose that $T_{e_1}(p) = p + 1$ and $T_{e_2}(p) = p + c$ for some $c \in \mathbb{R}$.

**Lemma 8.8.** The vector $(1, c)$ lies in the unstable subspace corresponding to $B$.

**Proof.** Recall that from Proposition 8.6 we have that $T_{B^{-n}(e_1)}(x) \to x$ for $i = 1, 2$.

Therefore if
\[ B^{-n} = \begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix}, \]
we have that $T_{B^{-n}(e_1)}(p) = p + a_n + cc_n$ and $T_{B^{-n}(e_2)}(p) = p + b_n + cd_n$ and so
\[ \begin{pmatrix} a_n & c_n \\ b_n & d_n \end{pmatrix} \begin{pmatrix} 1 \\ c \end{pmatrix} \to \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \]

which implies the fact that $(1, c)$ must be in the unstable subspace corresponding to $B$ (in particular that $c$ is an algebraic irrational number.) \[ \square \]

**Lemma 8.9.** Let $I$ be any open interval which is a connected component of $\ell \setminus C$ for some leaf $\ell$ of the unstable foliation $\mathcal{F}$ of the Anosov diffeomorphism $\Phi(B)$. Then restricted to $I$, we have that $\Phi : \mathbb{Z}^2 \to \text{Diff}(I)$ is conjugate to an action by translations and the conjugacy is smooth.

**Proof.** It remains to show that the conjugacy is smooth. We reduce the action to a circle map and obtain the smoothness of the conjugacy by the Herman-Yoccoz theory for circle maps.

We have that $I$ is a $C^r$ smooth embedded interval and observe that $S^1 \cong I / \Phi(e_1)$ is a smooth circle where $\Phi(e_2)$ acts as a smooth diffeo with rotation number $c$, which is algebraic irrational number and therefore Diophantine (\( \forall \delta, \exists \gamma \text{ such that } |qc - p| > \frac{2}{q^{1 + \delta}}, \forall (p, q) \in \mathbb{Z}, q \neq 0 \), c.f. Chapter VI, Corollary 1E of [36]) and therefore $\Phi(e_2)$ is conjugate to a rotation with angle $c$ (modulo 1) by a conjugacy that is $C^{r-1-2\delta}$ (c.f. [28]).

The action of $\Phi(e_2)$ on $I / \Phi(e_1)$ is defined as follows more explicitly. Pick a fundamental domain $D \subset I$ and $p \in D$ and denote by $\tilde{D}$ the circle obtain from $D$ by identifying two endpoints. Denote by $\tilde{f} : \tilde{D} \to \tilde{D}$ the action. Then
\[ \tilde{f}^n p = \Phi(m(n)e_1)\Phi(ne_2)p, \]

where $m(n)$ is the unique integer such that $\Phi(m(n)e_1)\Phi(ne_2)p \in D$. By Herman-Yoccoz theory, there exists a $C^1 h : \tilde{D} \to S^1$ such that $hf^{-1}p = p + c \mod 1$. That is to say
\[ \Phi(m(n)e_1)\Phi(ne_2)p = h^{-1}(h(p) + nc). \]

Define $D_m := \Phi(me_1)D$ and $\tilde{f}_m : \tilde{D}_m \to \tilde{D}_m$ the action on $\tilde{D}_m$. We get that $h_m\tilde{f}_m h_m^{-1}(p) = p + c$ where $h_m : \tilde{D}_m \to S^1$ is given by $h_m = h\Phi(-me_1)$. 
8.2. Embedding the action into a flow. We next embed the $\mathbb{Z}^2$-action into a flow. For a single diffeomorphism on $(0, 1)$ without fixed points, this is given by Szekeres theorem (Theorem 4.1.11 of [35]). The construction in our setting is simpler.

**Proof of Proposition 8.2.** We introduce the following map $\mathbb{R} \to \Phi(\mathbb{Z}^2)$ via $(1, c) \cdot v \to \Phi(v)$. Using the circle map constructed in the previous proof, we give a more explicit description. Choosing $(-1/2, 1/2)$ as the fundamental domain for $\mathbb{R}/\mathbb{Z}$ and $D$ as a fundamental domain for $I/\Phi(e_1)$. We have a conjugation $h : D \to S^1$ satisfying

$$\Phi(m(n)e_1)\Phi(ne_2)p = h^{-1}(h(p) + nc) = h^{-1}(h(p) + nc + m(n)).$$

Let $v \in \mathbb{Z}^2$ be such that $(1, c) \cdot v \in (-1/2, 1/2)$ then we have $\Phi(v)p = h^{-1}(h(p) + (1, c) \cdot v)$. This defines the flow for $t$ in the dense subset of $\mathbb{R}$. We then extend the domain of definition of the flow to $\mathbb{R}$ by the smoothness of $h$. Explicitly, we obtain a flow $g : \mathbb{R} \to \text{Diff}(I)$ via $g_t(p) = h^{-1}(h(p) + t)$ for all $p \in I$.

This gives us a smooth one parameter flow $g_t : I \to I$ without fixed points such that for every $v \in \mathbb{Z}^2$, $\Phi(v) = g_{tv}$ for $tv = (1, c) \cdot v \in \mathbb{R}$. So we have $g_1 = \Phi(e_1)$ and $g_c = \Phi(e_2)$. We next define a one parameter subgroup of maps $\lambda t : \mathbb{T}^2 \to \mathbb{T}^2$ by defining $\lambda t$ in $F \setminus C$ as above and defined $g_t(p) = p$ for all $p \in C$. By construction $g_t$ is differentiable with respect to $t$, but a priori $g_t(p)$ is not differentiable in $p$.

From the equation $\Phi(B)\Phi(v)\Phi(B^{-1}) = \Phi(B(v))$ it follows that for $tv = (1, c) \cdot v$, we have

$$\Phi(B)g_tv\Phi(B)^{-1} = \Phi(B)\Phi(v)\Phi(B)^{-1} = \Phi(Bv) = g_{(1,c)Bv} = g_{Btv}.$$

Extending the domain of $tv$, we get

$$\Phi(B)g_t\Phi(B)^{-1} = g_{Bt}$$

for every $t \in \mathbb{R}$.

**Proof of Lemma 8.3.** The differentiability in $t$ is seen from the definition of the flow. To see the measurability in $p$, we first notice that $\Phi(v)(p) = g_tv(p)$ where $t_v = (1, c) \cdot v$. This shows that for each $v$, the flow $g_tv(p)$ is differentiable in $p$ for $tv$ fixed. Choosing $v \in \mathbb{Z}^2$, we get that $t_v$ is defined on a dense subset $D$ of $\mathbb{R}$, so this shows that $g_t(p)$ is differentiable in $p$ for $t \in D$. For each $t_v \notin D$, we choose a sequence $t_n \in D$ with $t_n \to t_v$, we get that $g_{tn}(p) \to g_{tv}(p)$ pointwise. This shows that $g_{tv}(p)$ is measurable in $p$.

9. Higher Regularity for Translations Along Leaves

In this section, we prove Theorem 1.6 and 1.7 of higher regularity.

**Proof of Theorem 1.6.** We consider the case $\rho = u_A \otimes u_B^t$, the other case $\rho = u_{A-1} \otimes u_{B-t}$ is similar. It can be verified that (2) is satisfied. Suppose $u_B = (1, c)$, where $c$ is a Diophantine irrational number because $B$ is Anosov. This shows that

$$h\Phi(e_1)h^{-1}(x) = x + u_A, \quad h\Phi(e_2)h^{-1}(x) = x + cu_A.$$
Therefore the maps $h\Phi(e_1)h^{-1}$ and $h\Phi(e_2)h^{-1}$ preserve each leaf of the unstable foliation of $A$ and act as translations on each leaf, and the maps $\Phi(e_1)$ and $\Phi(e_2)$ preserves the unstable foliations of $\Phi(B)$.

In this setting, we recover Proposition 8.2. Therefore we have (11) by differentiation with respect to $t$.

Let $\text{vol}$ be the volume measure which is preserved by $\Phi(B)$, by the argument in the proof of Proposition 8.5 with $\mu$ replaced by $\text{vol}$, the larger Lyapunov exponent of $\Phi(B)$ is $\log \lambda_B$ and we have the same calculation as (12). By assumption we have $\lambda_A = \lambda_B$, therefore $\text{vol}$ is the measure of maximal entropy for $\Phi(B)$. On the other hand $\log \lambda_B = \log \lambda_A$ is the entropy of $A$ with respect to the measure $h_*\text{vol}$. We also know that $\log \lambda_A$ is the topological entropy of $A$, so we conclude that $h_*\text{vol} = \text{vol}$, since $\text{vol}$ is the unique measure of maximal entropy for $A$. Now we apply Theorem 1.3 of [30] to complete the proof (see also Theorem F of [37]).

**Proof of Theorem 1.7.** We first claim:

**Claim:** On the affine action side, along each eigen-direction $u_{A,i}$ of $A$ corresponding to the eigenvalue $\lambda_i$, as in the statement of the theorem, there is a $\mathbb{Z}^2$ action whose generators act by translating $cu_{A,i}$ and $cvu_{A,i}$, where $v$ is a Diophantine number and $c \neq 0$. Therefore, the line along $u_{A,i}$ quotient the $\mathbb{Z}^2$ action gives rise to a circle rotation $x \mapsto x + \nu \mod 1$.

**Proof of the claim.** We cite the following lemma of Katznelson.

**Lemma 9.1** ([27] Lemma 3). Let $A$ be an irreducible $n \times n$ integer matrix and $v$ be an eigenvector of $A$ corresponding to a real eigenvalue of multiplicity 1. Then $v$ is Diophantine with of type $(C, n - 1)$, i.e. there exists $C > 0$ such that for all $p \in \mathbb{Z}^n \setminus \{0\}$ we have $|\langle p, v \rangle| \geq \frac{C}{|p|^{n-1}}$.

Since we have a $\mathbb{Z}^n$ affine action with rotation matrix $\rho = cu_{A,i} \otimes u_{B^i,j}$, each generator is a translation along $u_{A,i}$ with translation length $c$ times each entry of $u_{B^i,j}$ respectively. If $B$ is irreducible, then by the lemma of Katznelson, the eigenvector $u_{B^i,j}$ is Diophantine of type $(C, n - 1)$. Picking any two entries of $u_{B^i,j}$, we get a vector denoted by $c(1, \nu) \in \mathbb{R}^2$ that is Diophantine of type $(C, n - 1)$. The claim is then established in this case. Suppose $B^i$ is reducible. Without loss of generality, we consider the case $B^i = \begin{bmatrix} B_1 & D \\ 0 & B_2 \end{bmatrix} \in \text{SL}_n(\mathbb{Z})$ where $B_i, \ i = 1, 2$ is irreducible, and the cases of more diagonal blocks are similar. Since there is no eigenvalue 1, we have that $B_i \in \mathbb{Z}^{n_i} \times n_i$ satisfies $n_i \geq 2, \ i = 1, 2$. If $u_1$ is an eigenvector of $B_1$, then $(u_1, 0_{1 \times n_2})$ is an eigenvector $B$ with the same eigenvalue. Since $u_1$ is a Diophantine vector of type $(C, n_1 - 1)$ by the above lemma and has at least two entries, we can pick two entries to obtain a Diophantine vector $c(1, \nu)$. If $u_2$ is an eigenvector of $B_2$ with eigenvalue $\lambda$, then we can solve for $v$ from the equation $(B_1 - \lambda I) v = D u_2$ to get that $(v, u_2)$ is an eigenvector of $B$ with eigenvalue $\lambda$ (note that $\lambda$ is not an eigenvalue of $B_1$ by assumption). Then we pick two entries from $u_2$ to get a Diophantine vector. This completes the proof of the claim.

We next work on the nonlinear action $\Phi$. We relabel the eigenvalues of $B$ as follows

$$|\lambda_k^1| < \ldots < |\lambda_1^p| < 1 < |\lambda_1^n| < \ldots < |\lambda_k^n|, \quad k + \ell = n.$$
It is known that for each \( \lambda_{u/s} \) there is a one-dimensional foliation tangent to the invariant distribution corresponding to \( \lambda_{u/s} \) and that the conjugacy \( h \) preserves the weakest stable and unstable leaves (Lemma 6.1-6.3 of [17]). We then apply the argument in the proof of Theorem 1.6 and the above claim to recover Proposition 8.2 and (11), i.e. there is a \( \mathbb{Z}^2 \) subaction of \( \Phi(\Gamma_{B,n}) \) along the one dimensional foliation which can be embedded into a flow. This shows that the Lyapunov exponent of \( \Phi(B) \) with respect to the volume along the weakest stable and unstable foliations agree with the corresponding Lyapunov exponents for the linear Anosov maps. Applying Theorem F of [37], we get that \( h \) and \( h^{-1} \) are \( C^{1+} \) along the weakest stable and weakest unstable foliations. We remark that in the proof of Proposition 8.2, we need to invoke Herman-Yoccoz theory for the regularity of the conjugacy for circle maps induced by the \( \mathbb{Z}^2 \) subaction. Since the Diophantine vector \((1, \nu)\) in the claim is of type \((C, n - 1)\), we get that the conjugacy is \( C^{1+} \) if \( \Gamma_{B,n} \) acts by \( C^r \), \( r > n \) diffeomorphisms.

We next apply Proposition 2.4 of [22] to show that \( h \) preserves the next weak stable and unstable foliations corresponding to Lyapunov exponents \( \lambda_{u} \) and \( \lambda_{s} \) respectively. The same argument as the previous paragraph allows us to show that \( h \) and \( h^{-1} \) are \( C^{1+} \) along these foliations.

Repeating the above arguments, we obtain that \( h \) and \( h^{-1} \) are \( C^{1+} \) along each one-dimensional foliation corresponding to each eigenvalue \( \lambda_{u/s} \). An application of Journé theorem [25] completes the proof.

Note that when we have shown that \( n - 1 \) Lyapunov exponents of \( \Phi(B) \) agrees with the corresponding ones of \( A \), the remaining one agrees automatically. For this reason, in the statement of the theorem, we may allow \( c_i = 0 \) for \( c_i \) corresponding to \( \lambda_{u} \) or \( \lambda_{s} \).

In the three dimensional case, if \( \Phi(B) \) is smooth enough, by the main result of [18], the conjugacy can be as smooth as possible. \( \square \)

10. PROOF OF THEOREM 1.4

In this section, we give the proof of Theorem 1.4.

Proof of Theorem 1.4. Since \( \Phi(B) \) is Anosov and volume-preserving, we see that the Lebesgue measure \( \text{vol} \) is an SRB measure. We next show that \( \text{vol} \) is preserved by \( h_* \) and is the measure of maximal entropy (see Proposition 10.1 below), so we get that the entropy

\[
\log h_{\text{vol}}(\Phi(B)) = \log \lambda_A.
\]

By ergodicity and the Ledrappier-Young entropy formula, the Lyapunov exponent of \( \Phi(B) \) with respect to \( \text{vol} \) is \( \log \lambda_A \). We then complete the proof of Theorem 1.4 by applying Theorem 1.3 of [30] (see also Theorem F of [37]). \( \square \)

Proposition 10.1. Under the assumption of Theorem 1.4, we have that \( h_*\text{vol} = \text{vol} \), and \( \text{vol} \) is the measure of maximal entropy of \( \Phi(B) \).

Proof. Since each \( \Phi(v) \) is volume-preserving by assumption, we get that \( h_*\Phi(v)_*\text{vol} = h_*\text{vol} \). On the other hand, using the conjugacy \( h \), we can express each \( \Phi(v) \), \( v \in \mathbb{Z}^2 \), as

\[
\Phi(v)(x) = h^{-1}(h(x) + pv).
\]

(13)
Therefore
\[ h_\star \Phi(v)_\star \mathrm{vol} = (h_\star \Phi(v)_\star h^{-1}_\star) (h_\star \mathrm{vol}) = (h\Phi(v)h^{-1})_\star (h_\star \mathrm{vol}) = (h_\star \mathrm{vol})(\bullet + \rho v). \]
So each \( h_\star \Phi(v)_\star \mathrm{vol} \) is a translation of \( h_\star \mathrm{vol} \) by \( \rho v \) and we have
\[ h_\star \mathrm{vol} = h_\star \mathrm{vol}(\bullet + \rho v), \quad \forall v \in \mathbb{Z}^2. \]

Proposition 1.2 of [39] shows that the affine action of the form (1.1) is faithful if and only if \( A \) is of infinite order and \( \{ pp \mid p \in \mathbb{Z}^2 \} \) is dense on \( \mathbb{T}^2 \). So the faithfulness assumption in the statement implies that \( h_\star \mathrm{vol} \) is invariant under the group of translations on \( \mathbb{T}^2 \), therefore we get \( h_\star \mathrm{vol} = \mathrm{vol} \). Moreover, since \( \mathrm{vol} \) is the measure of maximal entropy for \( \Phi(B)h^{-1} = A \) and the measure of maximal entropy is preserved under conjugacy, we see that \( \mathrm{vol} \) is the measure of maximal entropy for \( \Phi(B) \).

\[ \square \]

11. PROOF OF THEOREM 1.5

In this section, we prove Theorem 1.5.

**Proof of Theorem 1.5.** Let \( \mu \) be an ergodic SRB measure of \( \Phi(B) \). We consider the following measure
\[ \nu := \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \Phi(B^n(e_1))_\star \mu. \tag{14} \]
We show in the following Proposition 11.1 that \( \nu = h^{-1}_\star \mathrm{vol} \) is well-defined and is both an SRB measure and the measure of maximal entropy of \( \Phi(B) \). By the Ledrappier-Young entropy formula, since \( \nu \) is the measure of maximal entropy, the positive Lyapunov exponent of \( \Phi(B) \) with respect to \( \nu \) is \( \log \lambda_B \). Applying Theorem F of [37] we conclude that the conjugacy \( h \) restricted to the support of \( \nu \) is \( C^{\nu - \varepsilon} \) along each unstable leaf in \( \text{supp} \nu \), and \( h^{-1} \) is \( C^{\nu - \varepsilon} \) along each unstable leaf of \( A \) in \( \text{supp}h_\star \nu = \mathbb{T}^2 \), for all \( \varepsilon \).

Next, by Corollary 11.14 of [5], we know that \( \nu \) as an SRB measure is a Gibbs \( u \)-state and its support contains entire leaves. Since \( \Phi(B) \) is Anosov on \( \mathbb{T}^2 \), an entire leaf of \( \Phi(B) \) is dense on \( \mathbb{T}^2 \), so we conclude that \( \text{supp} \nu \) is the whole \( \mathbb{T}^2 \). This shows that the conjugacy \( h \) and \( h^{-1} \) are \( C^{\nu - \varepsilon} \) along each unstable leaf.

Applying the same argument to \( \Phi(B^{-1}) \), by taking an SRB measure \( \mu' \) of \( \Phi(B^{-1}) \), the averaging procedure \( \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \Phi(B^{-n}e_1)_\star \mu' \) yields the same measure \( \nu = h^{-1}_\star \mathrm{vol} \) (see the proof of Proposition 11.1). Repeating the above argument, we conclude that the conjugacy \( h \) and \( h^{-1} \) are \( C^{\nu - \varepsilon} \) along each stable leaf for all \( \varepsilon \). Then by Journé theorem [25], we get that \( h \) and \( h^{-1} \) are both \( C^{\nu - \varepsilon} \) for all \( \varepsilon \).

\[ \square \]

**Proposition 11.1.** The limit in the definition (14) of \( \nu \) exists, and \( \nu = h^{-1}_\star \mathrm{vol} \) is an SRB measure and the measure of maximal entropy for \( \Phi(B) \).

**Proof.** Consider the following sequence
\[ h_\star \frac{1}{N} \sum_{n=1}^{N} \Phi(B^n(e_1))_\star \mu = \frac{1}{N} \sum_{n=1}^{N} h_\star \Phi(B^n(e_1))_\star \mu. \]
Using (13), the following holds
\[ h_* \Phi(B^n(e_1))_* \mu = (h \Phi(B^n(e_1))h^{-1})_*(h_* \mu) = h_* \mu(\bullet + pB^n e_1). \]

We know from (3) that
\[ \rho = \rho_C + c_1 u_A \otimes u_{B^1} + c_2 u_{A^{-1}} \otimes u_{B^{-1}}, \]
where \( \rho_C \) solves \( A \rho = \rho B + C, \ C \in \mathbb{Z}^{2 \times 2}. \) Then we have \( A^n \rho = \rho B^n + C_n \) with \( C_n \in \mathbb{Z}^{2 \times 2}. \) Note that linear Anosov is ergodic on \( A \).

Then we have
\[ \rho B^n e_1 = A^n \rho e_1 - C_n e_1 = c_1(u_{B^1} \cdot e_1)\lambda_B^n u_A + c_2(u_{B^{-1}} \cdot e_1)\lambda_B^{-n} u_{A^{-1}} \mod \mathbb{Z}^2. \]

Therefore, the map \( \rho e_1 \mapsto \rho B e_1 \) on \( \mathbb{T}^2 \) can be equivalently considered as the map \( \rho_c \mapsto A \rho_c \) with
\[ \rho_c = c_1(u_{B^1} \cdot e_1)u_A + c_2(u_{B^{-1}} \cdot e_1)u_{A^{-1}}. \]

Note that linear Anosov is ergodic on \( \mathbb{T}^2 \) with respect to Lebesgue. We thus obtain the full measure set \( C \) of parameters as the set of \((c_1, c_2)\) such that the corresponding \( \rho_c \) is a generic point for applying Birkhoff ergodic theorem to \( A \).

Therefore by the Ergodic Theorem we get \( \lim \frac{1}{N} \sum_{n=1}^{N} \delta_{x+A^n \rho_c} = \text{vol}, \) where \( \delta_x \) is the Dirac-\( \delta \) supported at \( x \), also the limit is independent of \( x \) by the translation invariance of vol. Therefore we have
\[
\lim \frac{1}{N} \sum_{n=1}^{N} h_* \Phi(B^n(e_1))_* \mu = \lim \frac{1}{N} \sum_{n=1}^{N} h_* \mu(\bullet + A^n \rho_c)
= \int \lim \frac{1}{N} \sum_{n=1}^{N} \delta_{x+A^n \rho_c} d h_* \mu(x)
= \text{vol}.
\]

This shows that \( \nu = h_*^{-1} \text{vol} \) is well-defined and is invariant under \( \Phi(B) \). Since \( \text{vol} \) is the measure of maximal entropy (MME) for \( A \) and the MME is preserved under conjugacy, we conclude that \( \nu \) is the MME for \( \Phi(B) \).

We next show that \( \nu \) is an SRB measure. First, by Corollary 11.14 of [5], we know that the conditional measure of \( \mu \) along each unstable leaf in \( \text{supp} \mu \) is absolutely continuous with respect to Lebesgue with a density \( \rho \) satisfying
\[ \frac{1}{K} \leq \frac{\rho(z_1)}{\rho(z_2)} \leq K \]
for some constant \( K \) uniform for all leaves in \( \text{supp} \mu \) and all \( z_1, z_2 \) in the same leaf with \( d(z_1, z_2) < 10 \) (equation (11.4) of [5]). Next, by the following Lemma 11.2, we know that \( \| D^u_{x} \Phi(B^n(e_1)) \| \) along each unstable leaf is bounded from above and below uniformly with respect to both \( n \in \mathbb{N} \) and \( x \in \mathbb{T}^2 \). This shows that \( \Phi(B^n(e_1))_* \mu \) has conditional measure absolutely continuous with respect to Lebesgue with a density \( \rho_n \) satisfying
\[ \frac{1}{K'} \leq \frac{\rho_n(z_1)}{\rho_n(z_2)} \leq K' \]
where the constant $K'$ is uniform for all leaves in $\text{supp}(\Phi(B_1^n (e_1))) \ast \mu$, all $z_1, z_2$ in the same leaf with $d(z_1, z_2) < 10K'$ and for all $n \in \mathbb{N}$. The averaged measure $\nu$ therefore has the same property hence is SRB. \hfill \Box

The proof is then reduced to establishing a uniform derivative bound of $\Phi(B^n(e_1))$ along the unstable leaves.

**Lemma 11.2.** For each $v \in \mathbb{Z}^2$, there is a constant $K = K_v$ such that

$$1/K \leq \|D^u_x \Phi(B^n(v))\| \leq K, \quad 1/K \leq \|D^s_x \Phi(B^{-n}(v))\| \leq K, \quad \forall \ x \in T^2, \ n \in \mathbb{N}.$$

**Proof.** We prove the statement for $D^s \Phi(B^{-n}(v))$. The other one is analogous.

Note that $h$ is biHölder. Denote by $\alpha$ the Hölder exponent. Fix a small $r$, then we have

$$B_{r\lambda^\alpha}(h(x)) \subset h B_r(x) \subset B_{r\lambda^\alpha}(h(x)).$$

Choose $z \in B_r(x) \cap W^u_x$ such that $d(\Phi(B^{-n}(v))x, \Phi(B^{-n}(v))z) = \delta$ for some $\delta > 0$. Since $h \Phi(B^{-n}(v))h^{-1}$ acts as translation, by the Hölderness of $h$, we get $\delta^{1/\alpha} \leq d(x, z) \leq \delta^\alpha$. Next, by the mean value theorem, we have for some $y$ lying between $x$ and $z$ on the same leaf

$$\delta = d(\Phi(B^{-n}(v))x, \Phi(B^{-n}(v))z) = d(x, z) \cdot \|D^s_y \Phi(B^{-n}(v))\|$$

Next, applying the following Lemma 11.3 we have

$$1/C\delta^{1-\alpha} \leq \|D^s_y \Phi(B^{-n}(v))\| \leq C\delta^{1-1/\alpha}.$$

\hfill \Box

**Lemma 11.3.** For each $v \in \mathbb{Z}^2$, there exists constant $C = C_v > 1$ such that for any two nearby points $x$ and $y$ on the same unstable leaf, we have

$$\left| \log \left| \frac{D^u_x \Phi(B^n(v))}{D^u_y \Phi(B^n(v))} \right| \right| \leq C|x - y|^{\alpha^2}$$

for all $n \in \mathbb{N}$, where $\alpha$ is the Hölder constant of the conjugacy $h$ and $h^{-1}$.

Similarly, we have for all nearby $x, y$ on the same stable leaf

$$\left| \log \left| \frac{D^s_x \Phi(B^{-n}(v))}{D^s_y \Phi(B^{-n}(v))} \right| \right| \leq C|x - y|^{\alpha^2}. $$

for all $n \in \mathbb{N}$.

**Proof.** By the group relation $\Phi(B)^n \Phi(v) \Phi(B)^{-n} = \Phi(B^n(v))$, we get

$$D_{\Phi(v) x_{-n}} \Phi(B)^n D_{x_{-n}} \Phi(v) D_x \Phi(B)^{-n} = D_x \Phi(B^n(v)),$$

where we use the notation $x_i$ to denote the $i$-th orbit point under the iteration $\Phi(B)$. In the following we denote $x'_{ni} := \Phi(v)x_{-n}$ and $y'_{ni} := \Phi(v)y_{-n}$ and by $(x'_{ni});_i$ the orbit point of $x'_{ni}$ under $\Phi(B)$, similarly for $y'_{ni}$.

By assumption, $x$ and $y$ lie on the same unstable leaf of $\Phi(B)$. So we get that $|x_{-i} - y_{-i}| < \lambda^i |x - y|$ for some $0 < \lambda < 1$.
We have the following estimate
\[
\log \frac{D^n\Phi(B^n(v))}{D^n\Phi(B^n(v))}(x) = \log \frac{D^n_x \Phi(v)}{D^n_y \Phi(v)} - \sum_{i=0}^{n-1} \log \frac{D^n_{x_i} \Phi(B^{-1})}{D^n_{y_i} \Phi(B^{-1})} + \sum_{i=0}^{n-1} \log \frac{D^n_{(x_{-i})} \Phi(B)}{D^n_{(y_{-i})} \Phi(B)}
\]
\[
\leq (D^n_{x_{-i}})^2 \Phi(v) \cdot |x_{-i} - y_{-i}|
\]
\[
+ \sum_{i=0}^{n-1} (D^n_{y_i})^2 \Phi(B) \cdot |x_{-i} - y_{-i}|
\]
\[
+ \sum_{i=0}^{n-1} (D^n_{(y_{-i})})^2 \Phi(B) \cdot |(x'_{-i}) - (y'_{-i})|.
\]

Now by the exponential expansion and contraction, the estimate of the second summand on the RHS is bounded by the distance $C \| (D^n_{x_{-i}})^2 \Phi(B) \| |x - y|$. Also, each term of the third summand on the RHS is bounded by $C \| (D^n_{(y_{-i})})^2 \Phi(B) \| |(x'_{-i}) - (y'_{-i})|$.

As the conjugacy $h$, which is known to be bi-Hölder, we have that
\[
\frac{1}{C} |x - y|^{1/\alpha} \leq |h(x) - h(y)| \leq C |x - y|^\alpha.
\]

As
\[
|h((x'_{-i}) - h((y'_{-i})) = |A^{i-n} h(x) - A^{i-n} h(y)|,
\]
one has that
\[
|(x'_{-i}) - (y'_{-i})| \leq C' |x - y|^{\alpha^2} \lambda^{\alpha(n-i)}
\]
and so the third summand in (15) is bounded by $C'' |x - y|^{\alpha^2}$.

\[\square\]

Acknowledgment

We would like to thank Professor A. Wilkinson for getting us interested in the problem and many stimulating discussions. We also would like to thank the referees for pointing out a mistake (corrected in Section 7) in the first version. J. X. is supported by NSFC (Significant project No.11790273) in China and Beijing Natural Science Foundation (Z180003).

References

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

[2] Asaoka, M. (2017). *Rigidity of certain solvable actions on the torus*. In Geometry, Dynamics, and Foliations 2013: In honor of Steven Hurder and Takashi Tsuboi on the occasion of their 60th birthdays (pp. 269-281). Mathematical Society of Japan.

[3] J. Alonso, C. Rivas, and J. Xavier, *Planar Baumslag-Solitar actions*. arXiv:1707.05675.

[4] Alonso, J., Guelman, N., Rivas, C., & Xavier, J. (2017). *On some planar Baumslag-Solitar actions*. arXiv preprint arXiv:1703.09102.

[5] Bonatti, Christian, Lorenzo J. Díaz, and Marcelo Viana. *Dynamics beyond uniform hyperbolicity: A global geometric and probabilistic perspective*. Vol. 102. Springer Science & Business Media, 2006.
[6] B. Farb and J. Franks, Groups of homeomorphisms of one-manifolds, I: actions of nonlinear groups (2011). Arxiv preprint https://arxiv.org/abs/0107085

[7] Bonatti, C., Monteverde, I., Navas, A., & Rivas, C. (2017). Rigidity for $C^1$ actions on the interval arising from hyperbolicity I: solvable groups. Mathematische Zeitschrift, 286(3-4), 919-949.

[8] L. Burslem and A. Wilkinson, Global rigidity of solvable group actions on $S1$, Geometry & Topology 8 (2004), no. 2, 877-924.

[9] Bonatti, Christian, S-H Kim, T. Koberda, M. Triestino. Small $C^1$ actions of semidirect products on compact manifolds. arXiv preprint arXiv:1910.04620 (2019).

[10] Brin, Michael, and Garrett Stuck. Introduction to dynamical systems. Cambridge University Press, 2002.

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

[12] B. Farb and J. Franks, Groups of homeomorphisms of one-manifolds, I: actions of nonlinear groups (2011). Arxiv preprint https://arxiv.org/abs/0107085.

[13] Franks, John, and Michael Handel. Entropy zero area preserving diffeomorphisms of $S^2$. Geometry & Topology 16.4 (2013): 2187-2284.

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

[15] 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.

[16] Fisher, David. Groups acting on manifolds: around the Zimmer program. arXiv preprint arXiv:0809.4849 (2008).

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

[18] Gogolev, Andrey. Bootstrap for local rigidity of Anosov automorphisms on the 3-torus. Communications in Mathematical Physics 352.2 (2017): 439-455.

[19] Guériton, Nancy, and Isabelle Lioussse. $C^1$-actions of Baumslag-Solitar groups on $S^1$. Algebraic & Geometric Topology 11.3 (2011): 1701-1707.

[20] Guériton, Nancy, and Isabelle Lioussse. Actions of Baumslag-Solitar groups on surfaces. arXiv preprint arXiv:1004.2126 (2010).

[21] Guériton, Nancy, and Isabelle Lioussse. Any Baumslag-Solitar action on surfaces with a pseudo-Anosov element has a finite orbit. Ergodic Theory and Dynamical Systems 39.12 (2019): 3353-3364.

[22] Gogolev, Andrey, Boris Kalinin, and Victoria Sadovskaya. Local rigidity for Anosov automorphisms, Communications in Mathematical Physics, June 2017, Volume 352, Issue 2, 439-455

[23] Ghys, Étienne, and Vlad Sergiescu. Sur un groupe remarquable de difféomorphismes du cercle. Commentarii Mathematici Helvetici 62, no. 1 (1987): 185-239.

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

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

[26] Katok, Anatole, and Boris Hasselblatt. Introduction to the modern theory of dynamical systems. Vol. 54. Cambridge University Press, 1997.

[27] Katznelson, Yitzhak. Ergodic automorphisms of $T^n$ are bernoulli shifts, Israel Journal of Mathematics 10, no. 2(1971), 186-195.

[28] 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.

[29] Liu, Qiao. Local rigidity of certain solvable group actions on tori. arXiv preprint arXiv:1910.13541 (2019).

[30] 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.

[31] Ledrappier, François, and L-S. Young. The metric entropy of diffeomorphisms: Part I: Characterization of measures satisfying Pesin’s entropy formula. Annals of Mathematics (1985): 509-539.
[32] Ledrappier, Francois, and L.-S. Young. *The metric entropy of diffeomorphisms: Part II: Relations between entropy, exponents and dimension*. Annals of Mathematics (1985): 540-574.

[33] A. McCarthy, *Rigidity of trivial actions of Abelian-by-cyclic groups*, Proc. Amer. Math. Soc. 138 (2010), no. 4, 1395-1403.

[34] Misiurewicz, Micha, and Krystyna Zieman. *Rotation sets for maps of tori*. Journal of the London Mathematical Society 2.3 (1989): 490-506.

[35] Navas, Andés. *Groups of circle diffeomorphisms*. University of Chicago Press, 2011.

[36] W. Schmidt, *Diophantine Approximation*. Lecture Notes in Math. 785. Springer-Verlag, New York, 1980.

[37] Saghin, Radu, and Jiagang Yang. *Lyapunov exponents and rigidity of Anosov automorphisms and skew products*. arXiv preprint arXiv:1802.08266 (2018).

[38] Wang, Zhenqi Jenny. *New examples of local rigidity of algebraic partially hyperbolic actions*. arXiv preprint arXiv: 1908.09181v3 (2019).

[39] A. Wilkinson, J. Xue, *Rigidity of some abelian-by-cyclic solvable group actions on $\mathbb{T}^N$*, arXiv 1902.01003, to appear in Communications in Mathematical Physics.

[40] Young, Lai-Sang. *What are SRB measures, and which dynamical systems have them?* Journal of Statistical Physics 108.5 (2002): 733-754.

**Department of Mathematics, University of Chicago, Chicago, IL 60637, USA**

*Email address: shurtados@uchicago.edu*

**Yau Mathematical Sciences Center & Department of Mathematics, Tsinghua University, Beijing, China, 100084**

*Email address: jxue@tsinghua.edu.cn*