SMOOTH LOCAL RIGIDITY FOR HYPERBOLIC TORAL AUTOMORPHISMS

BORIS KALININ[1], VICTORIA SADOVSKAYA[2], AND ZHENQI JENNY WANG[3]

Abstract. We study the regularity of a conjugacy H between a hyperbolic toral automorphism A and its smooth perturbation f. We show that if H is weakly differentiable then it is $C^{1+\text{Hölder}}$ and, if A is also weakly irreducible, then H is C^∞. As a part of the proof, we establish results of independent interest on Hölder continuity of a measurable conjugacy between linear cocycles over a hyperbolic system. As a corollary, we improve regularity of the conjugacy to C^∞ in prior local rigidity results.

1. Introduction and local rigidity results

Hyperbolic automorphisms of tori are the prime examples of hyperbolic systems. The action of a matrix $A \in SL(N, \mathbb{Z})$ on $\mathbb{R}^N$ induces an automorphism of the torus $\mathbb{T}^N = \mathbb{R}^N/\mathbb{Z}^N$, which we denote by the same letter. An automorphism A is called *hyperbolic*, or *Anosov*, if the matrix has no eigenvalues on the unit circle. One of the key properties of hyperbolic systems is *structural stability*: any diffeomorphism f of $\mathbb{T}^N$ sufficiently C^1 close to such an A is also hyperbolic and is topologically conjugate to A [A67], that is, there exists a homeomorphism H of $\mathbb{T}^N$, called a *conjugacy*, such that

$$A \circ H = H \circ f. \tag{1.1}$$

Any two conjugacies differ by an affine automorphisms of $\mathbb{T}^N$ commuting with A [Wa70], and hence have the same regularity. Although H is always bi-Hölder continuous, it is usually not even C^1, as there are various obstructions to smoothness. This is in sharp contrast with rigidity for actions of larger groups, where often any perturbation, or even any smooth action, is C^∞ conjugate to the algebraic model.

In the classical case of a single system, the problem of establishing smoothness of the conjugacy H from some weaker assumptions has been extensively studied. It is often described as rigidity, in the sense that weak equivalence implies strong equivalence.

In dimension two, definitive results were obtained in [dlL87, dlLM88, dlL92]. For hyperbolic automorphisms of $\mathbb{T}^2$, and more generally for Anosov diffeomorphisms of $\mathbb{T}^2$, C^∞ smoothness of the conjugacy was obtained from absolute continuity of H and from equality of Lyapunov exponents of A and f at the periodic points.

Key words: Hyperbolic toral automorphism, conjugacy, local rigidity, linear cocycle, iterative method.
[1] Supported in part by Simons Foundation grant 426243.
[2] Supported in part by NSF grant DMS-1823150.
[3] Supported in part by NSF grant DMS-1845416.

The case of higher dimensional systems is much more complicated. In particular, the problem of exact level of regularity of H is subtle: for any $k \in \mathbb{N}$ and any $N \geq 4$ there exists a reducible hyperbolic automorphism A of $\mathbb{T}^N$ and its analytic perturbation f such that the conjugacy H is C^k but is not C^{k+1} [dlL92]. We recall that A is *reducible* if it has a nontrivial rational invariant subspace or, equivalently, if its characteristic polynomial is reducible over $\mathbb{Q}$.

The two-dimensional results were extended in two directions. First, C^∞ conjugacy was obtained for systems that are conformal on full stable and unstable subspaces under various periodic data assumptions which ensured that the perturbed system is also conformal [dlL02, KS03, dlL04, KS09]. Second, for some classes of irreducible A, equality of Lyapunov exponents or similarity of the periodic data were shown to imply $C^{1+\text{Hölder}}$ smoothness of H [GG08, G08, GKS11, SaY, GKS20, dW21]. Irreducibility of A is necessary for these results [dlL92, dlL02, G08]. Low smoothness of H is due to the method of the proof, which establishes regularity of H along natural one or two dimensional f-invariant foliations of $\mathbb{T}^N$, whose leaves are typically only $C^{1+\text{Hölder}}$ smooth. Therefore, different methods are needed to prove higher regularity of H. Outside of the conformal setting, the only result on C^∞ smoothness was obtained for automorphisms of $\mathbb{T}^3$ in [G17].

In this paper we establish general results on bootstrap of regularity of the conjugacy H. We show that for *any* hyperbolic automorphism A, if H is weakly differentiable in a certain sense then it is $C^{1+\text{Hölder}}$ and, if in addition A is *weakly irreducible*, then H is C^∞. As a corollary, we improve the regularity of H from $C^{1+\text{Hölder}}$ to C^∞ in the previous local rigidity results for the irreducible case.

Now we formulate our main results. We denote by $W^{1,q}(\mathbb{T}^N)$ the Sobolev space of L^q functions with L^q weak partial derivatives of first order. The first result holds for an arbitrary hyperbolic automorphism without any irreducibility assumption. We recall that while H satisfying (1.1) is not unique, there is a unique *conjugacy C^0 close to the identity.* This is H in the homotopy class of the identity with $H(p) = 0$, where p is the fixed point of f closest to 0.

Theorem 1.1. *Let A be a hyperbolic automorphism of $\mathbb{T}^N$ and let f be a $C^{1+\text{Hölder}}$ diffeomorphism of $\mathbb{T}^N$ which is C^1 close to A. Suppose that for some conjugacy H between f and A, either H or H^{-1} is in $W^{1,q}(\mathbb{T}^N)$ with $q > N$. Then H is a $C^{1+\text{Hölder}}$ diffeomorphism.*

More precisely, there is a constant $\beta_0 = \beta_0(A)$, $0 < \beta_0 \leq 1$, so that for any $0 < \beta' < \beta_0$ there exist constants $\delta > 0$ and $K > 0$ such that for any $0 < \beta \leq \beta'$ the following holds.

For any $C^{1+\beta}$ diffeomorphism f with $\|f - A\|_{C^1} < \delta$, if some conjugacy between A and f, or its inverse, is in $W^{1,q}(\mathbb{T}^N)$, $q > N$, then any conjugacy is a $C^{1+\beta}$ diffeomorphism. Moreover, for the conjugacy H that is C^0 close to the identity,

$$\|H - I\|_{C^{1+\beta}} \leq K\|f - A\|_{C^{1+\beta}}. \tag{1.2}$$

Remark 1.2. The assumption of being in $W^{1,q}$ with $q > N$ in this and in the next theorem can be replaced with a slightly weaker one that we actually need for the proof: either H^{-1} is in $W^{1,1}$ and its Jacoby matrix of partial derivatives is invertible and gives the differential of H^{-1} for Lebesgue almost every point of $\mathbb{T}^N$, or the same holds for H and f preserves an absolutely continuous probability measure.

In the next theorem we obtain C^∞ smoothness of the conjugacy assuming that A is weakly irreducible, which is defined as follows. Let $\mathbb{R}^N = \oplus_{\rho_i} E^i$ be the splitting where E^i is the sum of generalized eigenspaces of A corresponding to the eigenvalues of modulus ρ_i, and let $\hat{E}^i = \oplus_{\rho_j \neq \rho_i} E^j$. We say that A is *weakly irreducible* if each $\hat{E}^i$ contains no nonzero elements of $\mathbb{Z}^N$. This condition is weaker than irreducibility, see Section 3.3 for details.

Theorem 1.3. *Let A be a weakly irreducible hyperbolic automorphism of $\mathbb{T}^N$. Then there is a constant $\ell = \ell(A) \in \mathbb{N}$ so that for any C^∞ diffeomorphism f which is C^ℓ close to A the following holds. If for some conjugacy H between f and A either H or H^{-1} is in the Sobolev space $W^{1,q}(\mathbb{T}^N)$ with $q > N$, then any conjugacy between f and A is a C^∞ diffeomorphism.*

The constant $\ell = \ell(A)$ is chosen sufficiently large to satisfy the inequalities (8.16).

Applying Theorem 1.3 we improve the regularity of the conjugacy from $C^{1+\text{Hölder}}$ to C^∞ in the strongest local rigidity results for irreducible toral setting [GKS11, GKS20]:

Corollary 1.4. *Let $A : \mathbb{T}^N \to \mathbb{T}^N$ be an irreducible Anosov automorphism such that no three of its eigenvalues have the same modulus. Let f be a C^∞ diffeomorphism which is C^ℓ-close to A such that the derivative $D_p f^n$ is conjugate to A^n whenever $p = f^n(p)$. Then f is C^∞ conjugate to A.*

Corollary 1.5. *Let $A : \mathbb{T}^N \to \mathbb{T}^N$ be an irreducible Anosov automorphism such that no three of its eigenvalues have the same modulus and there are no pairs of eigenvalues of the form $\lambda, -\lambda$ or $i\lambda, -i\lambda$, where $\lambda \in \mathbb{R}$. Let f be a volume-preserving C^∞ diffeomorphism of $\mathbb{T}^N$ sufficiently C^ℓ-close to A. If the Lyapunov exponents of f with respect to the volume are the same as the Lyapunov exponents of A, then f is C^∞ conjugate to A.*

To prove Theorem 1.1, we first establish results of independent interest on Hölder continuity of a measurable conjugacy between linear cocycles over a hyperbolic system. These results are formulated and discussed in Section 2. In the proof of the theorem we apply them to the conjugacy DH between the derivative cocycles Df and A. We note, however, that existence of *some* continuous conjugacy between the derivative cocycles Df and A does not imply in general that H is C^1.

Theorem 1.1 is used as the first step in the proof of Theorem 1.3 to obtain $C^{1+\beta}$ regularity of H. Then to establish its C^∞ smoothness we use an iterative method which is somewhat similar to the traditional KAM scheme. However, KAM is primarily used for elliptic systems and not for hyperbolic ones. Closest to our setting, KAM techniques were used in [DKt10] to prove C^∞ local rigidity for some $\mathbb{Z}^2$ actions by partially hyperbolic toral automorphisms. The main ingredient in the iterative step is obtaining a C^∞ approximate solution for the linearized conjugacy equation. In [DKt10] the higher rank action was used in an essential way to construct such a solution. In our case these methods do not apply, and instead the argument relies on existence of $C^{1+\beta}$ conjugacy with estimate (1.2). This, however, gives us only low regularity data, from which we need to construct a C^∞ approximate solution and obtain suitable estimates. This is done in Section 7 and is one of the key new ingredients, see remarks after Theorem 7.4 for details. Also, in contrast to the higher rank case in [DKt10], the estimate we obtain for the approximate solution is not tame. This creates problems in establishing convergence of the iterative procedure, which we overcome in Section 8.

The paper is structured as follows. In Section 2 we formulate our results on continuity of a measurable conjugacy between linear cocycles over a hyperbolic system, Theorems 2.1 and 2.2. These theorems are proved in Sections 4 and 5, respectively. In Section 3 we summarize basic notations and facts used throughout the paper. In Section 6 we prove Theorem 1.1. In Section 7 we obtain a result on solving a twisted cohomological equation over A, and in Section 8 we complete the proof of Theorem 1.3.

2. Results on continuity of conjugacy between linear cocycles

In this section we consider linear cocycles over a transitive Anosov diffeomorphism f of a compact connected manifold $\mathcal{M}$. We recall that f is *Anosov* if there exist a splitting of the tangent bundle $T\mathcal{M}$ into a direct sum of two Df-invariant continuous subbundles $\tilde{E}^s$ and $\tilde{E}^u$, a Riemannian metric on $\mathcal{M}$, and continuous functions ν and $\hat{\nu}$ such that

$$\|Df_x(\mathbf{v}^s)\| < \nu(x) < 1 < \hat{\nu}(x) < \|Df_x(\mathbf{v}^u)\| \tag{2.1}$$

for any $x \in \mathcal{M}$ and any unit vectors $\mathbf{v}^s \in \tilde{E}^s(x)$ and $\mathbf{v}^u \in \tilde{E}^u(x)$. The diffeomorphism is *transitive* if there is a point $x \in \mathcal{M}$ with dense orbit. All known examples satisfy this property.

Let A be a map from $\mathcal{M}$ to $GL(N,\mathbb{R})$. The $GL(N,\mathbb{R})$-*valued cocycle over f generated by A* is the map $\mathcal{A}: X \times \mathbb{Z} \to GL(N,\mathbb{R})$ defined by $\mathcal{A}(x,0) = \mathrm{Id}$ and for $n \in \mathbb{N}$,

$$\mathcal{A}(x,n) = \mathcal{A}_x^n = A(f^{n-1}x) \circ \cdots \circ A(x) \;\text{ and }\; \mathcal{A}(x,-n) = \mathcal{A}_x^{-n} = (\mathcal{A}_{f^{-n}x}^n)^{-1}.$$

We say that a $GL(d,\mathbb{R})$-valued cocycle $\mathcal{A}$ is β-Hölder continuous if there exists a constant c such that

$$d(\mathcal{A}_x, \mathcal{A}_y) \le c\,\mathrm{dist}(x,y)^\beta \quad \text{for all } x, y \in \mathcal{M},$$

where the metric d on $GL(N,\mathbb{R})$ is given by

$$d(A,B) = \|A - B\| + \|A^{-1} - B^{-1}\|, \quad \text{where } \|\,.\,\| \text{ is the operator norm.}$$

More generally, we consider linear cocycles defined as follows. Let $P: E \to \mathcal{M}$ be a finite dimensional β-Hölder vector bundle over $\mathcal{M}$. A continuous *linear cocycle* over f is a homeomorphism $\mathcal{A}: E \to E$ such that

$$P \circ \mathcal{A} = f \circ P \quad \text{and } \mathcal{A}_x : E_x \to E_{fx} \text{ is a linear isomorphism for each } x \in \mathcal{M}.$$

The linear cocycle $\mathcal{A}$ is called β-Hölder if $\mathcal{A}_x$ depends β-Hölder on x, with proper identification of fibers at nearby points. A detailed description of this setting is given in Section 2.2 of [KS13].

The differential of f and its restrictions to invariant sub-bundles of $T\mathcal{M}$, such as $\tilde{E}^s$ and $\tilde{E}^u$, are prime examples of linear cocycles.

We say that a β-Hölder cocycle $\mathcal{A}$ over an Anosov diffeomorphism f is *fiber bunched* if there exist numbers $\theta < 1$ and c such that for all $x \in \mathcal{M}$ and $n \in \mathbb{N}$,

$$\|\mathcal{A}_x^n\| \cdot \|(\mathcal{A}_x^n)^{-1}\| \cdot (\nu_x^n)^\beta < c\,\theta^n \quad \text{and} \quad \|\mathcal{A}_x^{-n}\| \cdot \|(\mathcal{A}_x^{-n})^{-1}\| \cdot (\hat{\nu}_x^{-n})^\beta < c\,\theta^n, \tag{2.2}$$

where $\nu_x^n = \nu(f^{n-1}x) \cdots \nu(x)$ and $\hat{\nu}_x^{-n} = (\hat{\nu}(f^{-n}x))^{-1} \cdots (\hat{\nu}(f^{-1}x))^{-1}$.

Let μ be an ergodic f-invariant measure on $\mathcal{M}$. We denote by $\lambda_+(\mathcal{A},\mu)$ and $\lambda_-(\mathcal{A},\mu)$ the largest and smallest Lyapunov exponents of $\mathcal{A}$ with respect to μ given by the Oseledets Multiplicative Ergodic Theorem. For μ almost all $x \in \mathcal{M}$, they equal the limits

$$\lambda_+(\mathcal{A},\mu) = \lim_{n\to\infty} n^{-1} \ln \|\mathcal{A}_x^n\| \quad \text{and} \quad \lambda_-(\mathcal{A},\mu) = \lim_{n\to\infty} n^{-1} \ln \|(\mathcal{A}_x^n)^{-1}\|^{-1}. \tag{2.3}$$

We say that a cocycle $\mathcal{A}$ *has one exponent* if for every f-periodic point p the invariant measure μ_p on its orbit satisfies $\lambda_+(\mathcal{A},\mu_p) = \lambda_-(\mathcal{A},\mu_p)$. By Theorem 1.4 in [K11], this condition is equivalent to

$$\lambda_+(\mathcal{A},\mu) = \lambda_-(\mathcal{A},\mu) \quad \text{for every ergodic } f\text{-invariant measure.}$$

We note that if $\mathcal{A}$ has one exponent, then it is fiber bunched [S15, Corollary 4.2].

For $GL(N,\mathbb{R})$ cocycles $\mathcal{A}$ and $\mathcal{B}$ over f, a (measurable or continuous) function $\mathcal{C} : \mathcal{M} \to GL(N,\mathbb{R})$ such that

$$\mathcal{A}_x = \mathcal{C}(fx)\,\mathcal{B}_x\,\mathcal{C}(x)^{-1} \quad \text{for all } x \in \mathcal{M}$$

is called a (measurable or continuous) *conjugacy* or *transfer map* between $\mathcal{A}$ and $\mathcal{B}$. For linear cocycles $\mathcal{A},\mathcal{B} : E \to E$ a conjugacy is defined similarly with $\mathcal{C}(x) \in GL(E_x)$.

The question whether a measurable conjugacy between two cocycles is continuous has been studied in [PaP97, Pa99, Sch99, S13, S15]. An example in [PW01] shows that a measurable conjugacy between two fiber bunched $GL(2,\mathbb{R})$-valued cocycles is not necessarily continuous, moreover, the generators of the cocycles in this example can be chosen arbitrarily close to the identity. Continuity of a measurable conjugacy was proven for cocycles with values in a compact group [PaP97, Pa99] and, somewhat more generally for cocycles with bounded distortion [Sch99], for $GL(2,\mathbb{R})$-valued cocycles with one exponent [S13], and for $GL(N,\mathbb{R})$-valued cocycles such that one is fiber bunched and the other one is uniformly quasiconformal [S15]. The result in [S13] relied on two-dimensionality, and the uniform quasiconformality assumption in [S15] is much stronger than having one exponent. The next theorem establishes continuity of a measurable conjugacy between a fiber bunched cocycle and a cocycle with one exponent.

Theorem 2.1. *Let f be a transitive $C^{1+\text{Hölder}}$ Anosov diffeomorphism of a compact manifold $\mathcal{M}$, and let $\mathcal{A}$ and $\mathcal{B}$ be β-Hölder linear cocycles over f. Suppose that $\mathcal{A}$ has one exponent and $\mathcal{B}$ is fiber bunched.*

Let μ be an ergodic f-invariant measure on $\mathcal{M}$ with full support and local product structure. Then any μ-measurable conjugacy between $\mathcal{A}$ and $\mathcal{B}$ is β-Hölder continuous, i.e., coincides with a β-Hölder continuous conjugacy on a set of full measure.

As we mentioned above, continuity of a measurable conjugacy does not hold in general if $\mathcal{A}$ has more than one exponent, however, we prove it in a special case of a constant $\mathcal{A}$. Moreover, we obtain an estimate of the β-Hölder constant $K_\beta(\mathcal{C})$ of the conjugacy $\mathcal{C}$ in terms of the β-Hölder constant of $\mathcal{B}$.

Theorem 2.2. *Let f and μ be as in Theorem 2.1, and let $\mathcal{A}$ is be a constant $GL(N,\mathbb{R})$-valued cocycle over f. Then for any Hölder continuous $GL(N,\mathbb{R})$-valued cocycle $\mathcal{B}$ sufficiently C^0 close to $\mathcal{A}$, any μ-measurable conjugacy between $\mathcal{A}$ and $\mathcal{B}$ is Hölder continuous.*

More specifically, there exists a constant $\beta_0(A,f)$ so that the following holds. For any $0<\beta'<\beta_0(A,f)$ there is $\delta>0$ and $k>0$ such that for any $0<\beta\le\beta'$ and any β-Hölder $GL(N,\mathbb{R})$-valued cocycle $\mathcal{B}$ over f with $\|\mathcal{B}_x - A\|_{C^0}<\delta$, any μ-measurable conjugacy $\mathcal{C}$ between $\mathcal{A}$ and $\mathcal{B}$ is β-Hölder and its β-Hölder constant satisfies

$$K_\beta(\mathcal{C}) \le k\,\|\mathcal{C}\|_{C^0}\,K_\beta(\mathcal{B}) \qquad \textit{and} \qquad K_\beta(\mathcal{C}^{-1}) \le k\,\|\mathcal{C}^{-1}\|_{C^0}\,K_\beta(\mathcal{B}). \tag{2.4}$$

The constant $\beta_0(A,f)$ is explicitly given by (5.4) in Section 5.

3. Basic notations and facts

3.1. **Norms and Hölder constants.** For $r\in\mathbb{N}\cup\{0\}$ we use $\|\cdot\|_{C^r}$ for the C^r norm of functions with continuous derivatives up to order r on $\mathbb{T}^N$.

For a β-Hölder function g, $0<\beta\le 1$, we denote its β-Hölder constant, or Hölder seminorm, by

$$K_\beta(g)=\|g\|_{C^{0,\beta}} \stackrel{\text{def}}{=} \sup\{\,|g(x)-g(y)|\,d(x,y)^{-\beta}:\ x\ne y\in\mathbb{T}^N\,\}\ <\ \infty.$$

We denote by $C^{1,\beta}$ or $C^{1+\beta}$ the space of functions with β-Hölder first derivative with norm

$$\|f\|_{C^{1+\beta}} \stackrel{\text{def}}{=} \|f\|_{C^1}+K_\beta(Df).$$

3.2. **Invariant subspaces.** For $A\in GL(N,\mathbb{R})$ let $\rho_1<\dots<\rho_L$ be the distinct moduli of its eigenvalues and let

$$\mathbb{R}^N = E^1\oplus\cdots\oplus E^L \tag{3.1}$$

be the corresponding A-invariant splitting, where E^i is the direct sum of generalized eigenspaces corresponding to the eigenvalues with modulus ρ_i. We also denote

$$\hat{E}^i \stackrel{\text{def}}{=} \oplus_{\rho_j\ne\rho_i}E^j, \qquad A_i=A|_{E^i}:E^i\to E^i, \qquad \text{and} \qquad N_i=\dim E^i. \tag{3.2}$$

For the Euclidean norm on $\mathbb{R}^N$ there is a constant K_A such that for each i we have

$$\|A_i^m\|\le K_A\,\rho_i^m\,(|m|+1)^N \quad \text{for all } m\in\mathbb{Z}. \tag{3.3}$$

Also, for any $\epsilon>0$ there is an "adapted" inner product on $\mathbb{R}^N$ such that the direct sum $\oplus E^i$ is orthogonal and for each $1\le i\le L$,

$$(\rho_i-\epsilon)^m\le\|A_i^m u\|\le(\rho_i+\epsilon)^m \ \text{ for any unit vector } u\in E^i \text{ and any } m\in\mathbb{Z}. \tag{3.4}$$

If A is hyperbolic then $\rho_{i_0} < 1 < \rho_{i_0+1}$ for some $1 \le i_0 < L$, and we define the stable and unstable subspaces of A as

$$E^s \stackrel{\text{def}}{=} \oplus_{\rho_i<1} E^i \qquad \text{and} \qquad E^u \stackrel{\text{def}}{=} \oplus_{\rho_i>1} E^i.$$

3.3. **Weak irreducibility.** Recall that $GL(N,\mathbb{Z})$ denotes the integer matrices with determinant ± 1. We say that $A \in GL(N,\mathbb{Z})$ is *weakly irreducible* if each $\hat{E}^i$ contains no nonzero elements of $\mathbb{Z}^N$. Irreducibility over $\mathbb{Q}$ implies weak irreducibility. Indeed, if there is a nonzero integer point $n \in \hat{E}^i$ then $span\{A^m n : m \in \mathbb{Z}\} \subset \hat{E}^i$ is a nontrivial rational invariant subspace. In fact, weak irreducibility is determined by the characteristic polynomial of A as follows.

Lemma 3.1. *A matrix $A \in GL(N,\mathbb{Z})$ is weakly irreducible if and only if there is a set $\Delta \subset \mathbb{R}$ so that for each irreducible over $\mathbb{Q}$ factor of the characteristic polynomial of A the set of moduli of its roots equals Δ.*

Proof. Let $A \in GL(N,\mathbb{Z})$, let p_A be its characteristic polynomial, and let $p_A = \prod_{k=1}^K p_k^{n_k}$ be its prime decomposition over $\mathbb{Q}$. Then we have the corresponding splitting $\mathbb{R}^N = \oplus V_k$ into rational A-invariant subspaces $V_k = \ker p_k^{n_k}(A)$. We also have the (non-rational) A-invariant splitting (3.1), and we set $\Delta = \{\rho_1, \dots, \rho_L\}$. We will show that A is weakly irreducible if and only if Δ is the set of moduli of the roots for each p_k.

If for some $\rho_i \in \Delta$ and $k \in \{1, \dots, K\}$ no root of the irreducible polynomial p_k has modulus ρ_i, then $V_k \subset \hat{E}^i$. Hence A is not weakly irreducible as V_k is a rational subspace and hence it contains nonzero points of $\mathbb{Z}^N$.

Conversely, suppose each p_k has Δ as the set of moduli of its roots. Suppose that for some i there is $0 \ne n \in (\mathbb{Z}^N \cap \hat{E}^i)$. Then for some k its projection n_k to V_k is a nonzero rational vector. We note that $n_k \in \hat{E}^i$ as $\hat{E}^i = \oplus_k(\hat{E}^i \cap V_k)$. Then

$$W = span\{A^m n_k : m \in \mathbb{Z}\}$$

is a rational A-invariant subspace contained in $\hat{E}^i \cap V_k$. Then the characteristic polynomial of $A|_W$ is a power of p_k and hence W contains an eigenvector with eigenvalue of modulus $\rho_i \in \Delta$. Thus $W \cap E^i \ne 0$, contradicting $W \subset \hat{E}^i$. Thus A is weakly irreducible. $\square$

It follows from the lemma that if A is irreducible or weakly irreducible then the following matrices are weakly irreducible

$$\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} A & \mathrm{I} \\ 0 & A \end{pmatrix}.$$

These matrices not irreducible and the latter is not diagonalizable.

4. Proof of Theorem 2.1

Let f be a transitive $C^{1+\text{Hölder}}$ Anosov diffeomorphism of a compact manifold $\mathcal{M}$, let E be a β-Hölder vector bundle over $\mathcal{M}$, and let $\mathcal{F} : E \to E$ be a β-Hölder linear cocycle over f.

In Section 4.1 we recall the definition and properties of holonomies for linear cocycles, in Section 4.2 we prove a preliminary results on twisted cocycles, and in Section 4.3 we give a proof of Theorem 2.1.

4.1. **Holonomies of fiber bunched linear cocycles.** The notion of *holonomies* for linear cocycle was introduced in [BV04, V08] Existence of holonomies was proved in [V08, ASV13] under a stronger "one-step" fiber bunching condition and then extended to bundle setting and weaker fiber bunching (2.2) in [KS13, S15].

Proposition 4.1. *Let $\mathcal{F}$ be a β-Hölder fiber bunched linear cocycle over $(\mathcal{M}, f)$. Then for every $x \in \mathcal{M}$ and $y \in W^s(x)$ the limit*

$$\mathcal{H}^s_{x,y} = \mathcal{H}^{\mathcal{F},s}_{x,y} = \lim_{n\to\infty} (\mathcal{F}^n_y)^{-1} \circ \mathcal{F}^n_x, \tag{4.1}$$

called the stable holonomy, *exists and satisfies*

($\mathcal{H}1$) $\mathcal{H}^s_{x,y}$ *is an invertible linear map from E_x to E_y;*

($\mathcal{H}2$) $\mathcal{H}^s_{x,x} = Id$ *and* $\mathcal{H}^s_{y,z} \circ \mathcal{H}^s_{x,y} = \mathcal{H}^s_{x,z}$, *and hence* $(\mathcal{H}^s_{x,y})^{-1} = \mathcal{H}^s_{y,x}$;

($\mathcal{H}3$) $\mathcal{H}^s_{x,y} = (\mathcal{F}^n_y)^{-1} \circ \mathcal{H}^s_{f^n x,\, f^n y} \circ \mathcal{F}^n_x$ *for all $n \in \mathbb{N}$;*

($\mathcal{H}4$) $\|\mathcal{H}^s_{x,y} - Id\| \le c \cdot d(x,y)^\beta$, *where c is independent of x and $y \in W^s_{loc}(x)$.*

4.2. **Twisted cocycles.** In this section we study the coboundary equation over f twisted by a β-Hölder linear cocycle $\mathcal{F} : E \to E$. We will use its main result, Proposition 4.3, in the inductive process in the proof of Theorem 2.1.

Let $\phi, \eta : \mathcal{M} \to E$ be sections of the bundle E over $\mathcal{M}$. We consider the equation

$$\eta(x) = \phi(x) + (\mathcal{F}_x)^{-1}(\eta(fx)) \quad \text{equivalently} \quad \phi(x) = \eta(x) - (\mathcal{F}_x)^{-1}(\eta(fx)). \tag{4.2}$$

Iterating (4.2) and denoting $\mathcal{F}^n_x = \mathcal{F}_{f^{n-1}x} \circ \cdots \circ \mathcal{F}_{fx} \circ \mathcal{F}_x : E_x \to E_{f^n x}$ we obtain

$$\begin{aligned} \eta(x) &= \phi(x) + (\mathcal{F}_x)^{-1}(\eta(fx)) = \phi(x) + (\mathcal{F}_x)^{-1}[\phi(fx) + \mathcal{F}_{fx}(\eta(f^2x))] = ... \\ &= \phi(x) + (\mathcal{F}_x)^{-1}(\phi(fx)) + \cdots + (\mathcal{F}^{n-1}_x)^{-1}(\phi(f^{n-1}x)) + (\mathcal{F}_{f^{n-1}x})^{-1}(\eta(f^n x)). \end{aligned}$$

Thus

$$\eta(x) = \Phi^n(x) + (\mathcal{F}_{f^{n-1}x})^{-1}(\eta(f^n x)), \quad \text{where} \tag{4.3}$$

$$\Phi^n(x) = \phi(x) + (\mathcal{F}_x)^{-1}(\phi(fx)) + \cdots + (\mathcal{F}^{n-1}_x)^{-1}(\phi(f^{n-1}x)) \in E_x.$$

We say that $\mathcal{F}$ is *uniformly bounded* if there exists K such that $\|\mathcal{F}^n_x\| \le K$ for all $x \in \mathcal{M}$ and $n \in \mathbb{Z}$. A β-Hölder bounded cocycle is fiber-bunched and hence it has stable holonomies $\mathcal{H}^s_{x,y} : E_x \to E_y$ where $y \in W^s(x)$.

Lemma 4.2. *Suppose that ϕ is a β-Hölder section and that $\mathcal{F}$ is a uniformly bounded β-Hölder cocycle. Then for any $x \in \mathcal{M}$ and $y \in W^s(x)$ the following limit exists*

$$\Phi^s_{x,y} = \lim_{n\to\infty} (\Phi^n(x) - \mathcal{H}^s_{y,x}\Phi^n(y)) = \sum_{k=0}^{\infty} \left[\, (\mathcal{F}^k_x)^{-1}(\phi(f^k x)) - \mathcal{H}^s_{y,x}(\mathcal{F}^k_y)^{-1}(\phi(f^k y)) \,\right]$$

and satisfies $\|\Phi^s_{x,y}\| \le K' d(x,y)^\beta$ *with uniform* K' *for all* $x \in \mathcal{M}$ *and* $y \in W^s_{loc}(x)$.

The result holds if instead of being uniformly bounded $\mathcal{F}$ satisfies the following. There exist numbers $\theta < 1$ and L such that for all $x \in \mathcal{M}$ and $n \in \mathbb{N}$,

$$\|(\mathcal{F}^n_x)^{-1}\| \cdot (\nu^n_x)^\beta < L\,\theta^n.$$

Proof. For all $x \in \mathcal{M}$ and $y \in W^s_{loc}(x)$ we have $d(f^k x, f^k y) \le \nu^k_x\, d(x,y)$. As ϕ is β-Hölder we obtain

$$\|\phi(f^k x) - \phi(f^k y)\| \le K_1 (\nu^k_x d(x,y))^\beta,$$

and since $\mathcal{H}^s_{f^k y, f^k x}$ is β-Hölder close to identity by $(\mathcal{H}4)$, we have

$$\|\phi(f^k x) - \mathcal{H}^s_{f^k y, f^k x}\phi(f^k y)\| \le K_2 (\nu^k_x\, d(x,y))^\beta.$$

By uniformly boundedness of $\mathcal{F}$ we have $\|(\mathcal{F}^k_x)^{-1}\| \le K$, and by continuity of ϕ we have $\sup_x \|\phi(x)\| \le K_3$. Therefore,

$$\Phi^n(x) - \mathcal{H}^s_{y,x}\Phi^n(y) = \sum_{k=0}^{n-1} (\mathcal{F}^k_x)^{-1}(\phi(f^k x)) - (\mathcal{H}^s_{y,x} \circ (\mathcal{F}^k_y)^{-1} \circ \mathcal{H}^s_{f^k x, f^k y})(\mathcal{H}^s_{f^k y, f^k x}\phi(f^k y))$$

Since $\mathcal{H}^s_{y,x} \circ (\mathcal{F}^k_y)^{-1} \circ \mathcal{H}^s_{f^k x, f^k y} = (\mathcal{F}^k_x)^{-1}$ by $(\mathcal{H}3)$, the k^{th} term in the sum equals

$$(\mathcal{F}^k_x)^{-1}(\phi(f^k x)) - (\mathcal{F}^k_x)^{-1}(\mathcal{H}^s_{f^k y, f^k x}\phi(f^k y)) = (\mathcal{F}^k_x)^{-1}\,[\phi(f^k x) - \mathcal{H}^s_{f^k y, f^k x}\phi(f^k y)],$$

and we estimate

$$\begin{gathered}\|(\mathcal{F}^k_x)^{-1}\,[\phi(f^k x) - \mathcal{H}^s_{f^k y, f^k x}\phi(f^k y)]\,\| \le \|(\mathcal{F}^k_x)^{-1}\| \cdot \|\phi(f^k x) - \mathcal{H}^s_{f^k y, f^k x}\phi(f^k y)\| \le \\ \|(\mathcal{F}^k_x)^{-1}\| \cdot K_2(\nu^k_x\, d(x,y))^\beta \le K K_2\, \theta^k d(x,y)^\beta \quad \text{for some } \theta < 1.\end{gathered}$$

Hence the series converges and

$$\|\Phi^n(x) - \mathcal{H}^s_{y,x}\Phi^n(y)\| \le \sum_{k=0}^{n-1} K K_2\, \theta^k d(x,y)^\beta \le K' d(x,y)^\beta,$$

so the limit $\Phi^s_{x,y}$ satisfies $\|\Phi^s_{x,y}\| \le K' d(x,y)^\beta$. □

Proposition 4.3. *Let* $\mathcal{F}$ *be a* β*-Hölder uniformly bounded cocycle over an Anosov diffeomorphism* f *(or a hyperbolic system). Let* μ *be an ergodic* f*-invariant measure on* $\mathcal{M}$ *with full support and local product structure.*

Let $\phi : \mathcal{M} \to E$ *be a* β*-Hölder section, and let* $\eta : \mathcal{M} \to E$ *be a* μ*-measurable section satisfying* (4.2). *Then* η *is* β*-Hölder and*

$$\eta(x) = \mathcal{H}^s_{y,x}\,\eta(y) + \Phi^s_{x,y} \quad \textit{for all } x \in X \textit{ and } y \in W^s(x).$$

Proof. Let $x \in \mathcal{M}$ and $y \in W^s(x)$. Using equation (4.3) for $\eta(x)$ and $\eta(y)$ we obtain

$$\eta(x) - \mathcal{H}^s_{y,x}\,\eta(y) = \Phi^n(x) - \mathcal{H}^s_{y,x}\Phi^n(y) + \Delta_n,$$

where

$$\Delta_n = (\mathcal{F}_{f^{n-1}x})^{-1}(\eta(f^n x)) - \mathcal{H}^s_{y,x}(\mathcal{F}_{f^{n-1}y})^{-1}(\eta(f^n y)).$$

By Lemma 4.2, $(\Phi^n(x) - \mathcal{H}^s_{y,x}\Phi^n(y))$ converges to $\Phi^s_{x,y}$.

Now we show that $\|\Delta_n\| \to 0$ along a subsequence for all x, y in a set of full measure. First we note that by property $(\mathcal{H}3)$ we have $\mathcal{H}^s_{y,x}(\mathcal{F}_{f^{n-1}y})^{-1} = (\mathcal{F}_{f^{n-1}x})^{-1} \circ \mathcal{H}^s_{f^ny,f^nx}$. Hence

$$\Delta_n = (\mathcal{F}_{f^{n-1}x})^{-1} \left(\eta(f^nx) - \mathcal{H}^s_{f^ny,f^nx}(\eta(f^ny))\right) = (\mathcal{F}_{f^{n-1}x})^{-1}(\Delta'_n),$$

where $\Delta'_n = \eta(f^nx) - \mathcal{H}^s_{f^ny,f^nx}(\eta(f^ny))$. By uniform boundedness of $\mathcal{F}$ we obtain

$$\|\Delta_n\| \le \|(\mathcal{F}_{f^{n-1}x})^{-1}\| \cdot \|\Delta'_n\| \le K\|\Delta'_n\|.$$

Since the section $\eta : \mathcal{M} \to E$ is μ-measurable, by Lusin's theorem there exists a compact set $S \subset \mathcal{M}$ with $\mu(S) > 1/2$ such that η is uniformly continuous and hence bounded on S. Let Y be the set of points in $\mathcal{M}$ for which the frequency of visiting S equals $\mu(S)$. By Birkhoff Ergodic Theorem, $\mu(Y) = 1$.

If $x, y \in Y$, there exists a subsequence $n_i \to \infty$ such that such that $f^{n_i}x, f^{n_i}y \in S$ for all i. Since $y \in W^s(x)$, $d(f^{n_i}x, f^{n_i}y) \to 0$ and hence $\Delta'_{n_i} \to 0$ by uniform continuity and boundedness of η on S and property $(\mathcal{H}4)$ of $\mathcal{H}^s$. Thus $\Delta_{n_i} \to 0$ and we obtain that

$$\eta(x) = \mathcal{H}^s_{y,x}\, \eta(y) + \Phi^s_{x,y} \quad \text{ for all } x, y \in Y \text{ with } y \in W^s(x).$$

Since $\Phi^s_{x,y}$ is β-Hölder on $W^s_{\text{loc}}(x)$ by Lemma 4.2, we conclude that

$$\|\eta(x) - \mathcal{H}^s_{y,x}\, \eta(y)\| \le K' d(x,y)^\beta \quad \text{ for all } x, y \in Y \text{ with } y \in W^s(x).$$

Since $\mathcal{H}^s_{x,y}$ is β-Hölder by property $(\mathcal{H}4)$, this means that η is essentially β-Hölder along $W^s_{\text{loc}}(x)$.

Similar arguments for $y \in W^u_{loc}(x)$ show that η is also essentially β-Hölder along $W^u_{loc}(x)$. Hence η is β-Hölder by the local product structure of μ and of the stable and unstable manifolds. □

4.3. **Proof of Theorem 2.1.** For convenience, by taking inverse, we will work with a conjugacy $\mathcal{C}$ satisfying

$$\mathcal{B}_x = \mathcal{C}(fx)\, \mathcal{A}_x\, \mathcal{C}(x)^{-1}. \tag{4.4}$$

First we observe that since $\lambda_+(\mathcal{A}, \mu) = \lambda_-(\mathcal{A}, \mu)$ and $\mathcal{B}$ is μ-measurably conjugate to $\mathcal{A}$, the following lemma implies that

$$\lambda_+(\mathcal{B}, \mu) = \lambda_-(\mathcal{B}, \mu).$$

Lemma 4.4. *Let μ be an ergodic f-invariant measure. If $\mathcal{C}$ is a μ-measurable conjugacy between cocycles $\mathcal{A}$ and $\mathcal{B}$, then for μ a.e. x and for each vector $0 \ne u \in E_x$ the forward (resp. backward) Lyapunov exponent of u under $\mathcal{A}$ equals that of $\mathcal{C}_x(u)$ under $\mathcal{B}$.*

Proof. We fix a set of positive measure $Y \subseteq \mathcal{M}$ such that for some K we have $\|\mathcal{C}_x\| \le K$ and $\|(\mathcal{C}_x)^{-1}\| \le K$ for all $x \in Y$. Then we choose an f-invariant set of full measure $X \subseteq \mathcal{M}$ such that for every $x \in X$

(i) the forward and backward Lyapunov exponents under both $\mathcal{A}$ and $\mathcal{B}$ exist for each non-zero vector $v \in E_x$, and

(ii) the frequency of visiting Y under both forward and backward iterates of f equals $\mu(Y) > 0$.

For every $x \in X,\ 0 \neq u \in E_x$, and $n \in \mathbb{Z}$ we have

$$n^{-1} \ln \|\mathcal{B}_x^n(\mathcal{C}_x(u))\| = n^{-1} \ln \|\mathcal{C}_{f^n x}(\mathcal{A}_x^n(u))\|.$$

The limit of the left hand side as $n \to \infty$ (resp. $n \to -\infty$) is the forward (resp. backward) Lyapunov exponent of $\mathcal{C}_x(u)$ under $\mathcal{B}$. On the other hand, by the choice of Y, the limit of the right hand side along a subsequence $n_i \to \infty$ (resp. $n_i \to -\infty$) such that $f^{n_i} x \in Y$ equals the forward (resp. backward) Lyapunov exponent of u under $\mathcal{A}$. □

We use the following results from [KS13]. In the three theorems below, f is a transitive $C^{1+\text{Hölder}}$ Anosov diffeomorphism, $\mathcal{A}, \mathcal{B} : E \to E$ are β-Hölder linear cocycles over f, and μ is an ergodic f-invariant measure with full support and local product structure.

Theorem 4.5. [KS13, Theorem 3.9] *Suppose that for every f-periodic point p the invariant measure μ_p on its orbit satisfies $\lambda_+(\mathcal{A}, \mu_p) = \lambda_-(\mathcal{A}, \mu_p)$. Then there exist a flag of β-Hölder $\mathcal{A}$-invariant sub-bundles*

$$\{0\} = U^0 \subset U^1 \subset ... \subset U^{j-1} \subset U^k = E \tag{4.5}$$

and β-Hölder Riemannian metrics on the quotient bundles U^i/U^{i-1}, $i = 1, ..., k$, so that for some positive β-Hölder function $\phi : \mathcal{M} \to \mathbb{R}$ the quotient-cocycles induced by the cocycle $\phi\mathcal{A}$ on U^i/U^{i-1} are isometries.

Theorem 4.6. [KS13, Theorem 3.1 and Corollary 3.8] *If $\mathcal{B}$ is fiber bunched, then any $\mathcal{B}$-invariant μ-measurable conformal structure on E coincides μ-a.e. with a Hölder continuous conformal structure.*

If a cocycle has more than one Lyapunov exponent, then the corresponding Lyapunov sub-bundles are invariant and measurable, but not continuous in general. For a fiber bunched cocycle with only one Lyapunov exponent measurable invariant sub-bundles are continuous.

Theorem 4.7. [KS13, Theorem 3.3 and Corollary 3.8] *Suppose that $\mathcal{B}$ is fiber bunched and $\lambda_+(\mathcal{B}, \mu) = \lambda_-(\mathcal{B}, \mu)$. Then any μ-measurable $\mathcal{B}$-invariant sub-bundle of $\mathcal{E}$ coincides μ-a.e. with a Hölder continuous one.*

We consider the flag U^i for $\mathcal{A}$ given by Theorem 4.5. Denoting $\mathcal{U}_x^i = \mathcal{C}(x) U_x^i$ we obtain the corresponding flag of measurable $\mathcal{B}$-invariant sub-bundles

$$\{0\} = \mathcal{U}^0 \subset \mathcal{U}^1 \subset \mathcal{U}^2 \subset \cdots \subset \mathcal{U}^k = E.$$

By Theorem 4.7 we may assume that the sub-bundles $\mathcal{U}^i$ are Hölder continuous.

The conformal structure σ_1 on E^1 given by the Riemannian metric in Theorem 4.5 is invariant under $\mathcal{A}$. The push forward of σ_1 by $\mathcal{C}$ gives a measurable $\mathcal{B}$-invariant conformal structure τ_1 on $\mathcal{U}^1$, which is Hölder continuous by Theorem 4.6.

Similarly, we consider Hölder continuous quotient-bundles $\tilde{V}^i = U^i/U^{i-1}$ and $\tilde{\mathcal{V}}^i = \mathcal{U}^i/\mathcal{U}^{i-1}$ over $\mathcal{M}$ with the quotient cocycles $\mathcal{A}^{(i)}$ and $\mathcal{B}^{(i)}$. Since $\mathcal{A}^{(i)}$ preserves a Hölder continuous conformal structure σ_i on $\tilde V^i$, pushing forward by $\mathcal{C}$ we obtain a measurable conformal structure τ_i on $\mathcal{U}^i/\mathcal{U}^{i-1}$ invariant under $\mathcal{B}^{(i)}$, which is Hölder continuous by Theorem 4.6. Thus we obtain a "similar structure" for $\mathcal{B}$.

We fix a β-Hölder Riemannian metric on E. We denote by V^i the orthogonal complement of U^{i-1} in E_i, and we denote by $\mathcal{V}^i$ the orthogonal complement of $\mathcal{U}^{i-1}$ in $\mathcal{U}^i$, $i=1,\dots,k$. Thus $U^i = V^1 \oplus \cdots \oplus V^i$ and $\mathcal{U}^i = \mathcal{V}^1 \oplus \cdots \oplus \mathcal{V}^i$. All these sub-bundles are Hölder continuous but for $i>1$ they are not invariant under $\mathcal{A}$ and $\mathcal{B}$, and $\mathcal{C}$ does not necessarily map V^i to $\mathcal{V}^i$.

We denote by $P^j : E \to V^j$ the projection to the V^j component in the splitting $E = V^1 \oplus \cdots \oplus V^k$ and similarly $\mathcal{P}^j : \mathcal{E} \to \mathcal{V}^j$.

We denote the restriction of $\mathcal{C}$ to V^i by $\mathcal{C}^i$ and we denote by $\mathcal{C}^{j,i}$ its j-component $\mathcal{C}^{j,i} = \mathcal{P}^j \circ \mathcal{C}^i : V^i \to \mathcal{V}^j$. Since $\mathcal{U}^i_x = \mathcal{C}(x) U^i_x$, we have $\mathcal{C}^i : V^i \to \mathcal{U}^i$ and thus $\mathcal{C}^{j,i} = 0$ for $j>i$, that is $\mathcal{C}$ has an upper triangular block structure.

We also define the corresponding blocks $\mathcal{A}^{j,i} : V^i \to V^j$ and $\mathcal{B}^{j,i} : \mathcal{V}^i \to \mathcal{V}^j$ as $\mathcal{A}^{j,i} = P^j \circ \mathcal{A}|_{V^i}$ and similarly for $\mathcal{B}$. The invariance of the flags also yields upper triangular block structures for $\mathcal{A}$ and $\mathcal{B}$: $\mathcal{A}^{j,i} = 0 = \mathcal{B}^{j,i}$ for $j>i$.

We will show inductively that the restriction of $\mathcal{C}$ to U^i is Hölder continuous, $i = 1,\dots,k$. The base case $i=1$ follows from the following result from [S15].

Theorem 4.8. [S15, Theorem 2.7] *Let $\mathcal{A},\mathcal{B} : E \to E$ be β-Hölder linear cocycles over a hyperbolic system. Suppose that $\mathcal{A}$ uniformly quasiconformal and $\mathcal{B}$ is fiber bunched. Let μ be an ergodic invariant measure with full support and local product structure. Then any μ-measurable conjugacy between $\mathcal{A}$ and $\mathcal{B}$ is β-Hölder continuous, i.e. it coincides with a β-Hölder continuous conjugacy on a set of full measure.*

Now we describe the inductive step. Assuming that the restriction of $\mathcal{C}$ to U^{i-1} is β-Hölder continuous we show that so is the restriction to U^i. Since $U^i = V^i \oplus U^{i-1}$, it suffices to show that the restriction $\mathcal{C}^i$ of $\mathcal{C}$ to V^i is also β-Hölder continuous. We will establish this inductively for each of its components $\mathcal{C}^{j,i}$, $j=i,\dots,1$.

First we observe that $\mathcal{C}^{i,i}$ is Hölder continuous for all $i=1,\dots,k$. For this we identify bundles V^i with $\tilde V^i$ and $\mathcal{V}^i$ with $\tilde{\mathcal{V}}^i$ via the projections. Under these identifications the cocycle $\mathcal{A}^{i,i} : V^{i,i} \to V^{i,i}$ corresponds to the quotient cocycle $\mathcal{A}^{(i)}$, the cocycle $\mathcal{B}^{i,i} : \mathcal{V}^{i,i} \to \mathcal{V}^{i,i}$ corresponds to $\mathcal{B}^{(i)}$, and the map $\mathcal{C}^{i,i}$ corresponds to the quotient measurable conjugacy $\mathcal{C}^{(i)}$ between $\mathcal{A}^{(i)}$ and $\mathcal{B}^{(i)}$. Since the quotient cocycles $\mathcal{A}^{(i)}$ and $\mathcal{B}^{(i)}$ are conformal, Theorem 4.8 shows that $\mathcal{C}^{(i)}$ is β-Hölder continuous, and hence so is $\mathcal{C}^{i,i}$.

Now we show that $\mathcal{C}^{i-\ell,i}$ is β-Hölder assuming that $\mathcal{C}^{i-j,i}$ is β-Hölder for $j=0,1,\dots \ell-1$. Using the conjugacy equation

$$\mathcal{B}_x \circ \mathcal{C}_x = \mathcal{C}_{fx} \circ \mathcal{A}_x$$

and equating $(i-\ell,i)$ components we obtain

$$\begin{aligned}&\mathcal{B}_x^{i-\ell,i-\ell} \circ \mathcal{C}_x^{i-\ell,i} + \mathcal{B}_x^{i-\ell,i-\ell+1} \circ \mathcal{C}_x^{i-\ell+1,i} + \cdots + \mathcal{B}_x^{i-\ell,i} \circ \mathcal{C}_x^{i,i}\\ &= \mathcal{C}_{fx}^{i-\ell,i-\ell} \circ \mathcal{A}_x^{i-\ell+1,i} + \mathcal{C}_{fx}^{i-\ell,i-\ell+1} \circ \mathcal{A}_x^{i-\ell+1,i} + \cdots + \mathcal{C}_{fx}^{i-\ell,i} \circ \mathcal{A}_x^{i,i}\end{aligned}$$

and hence

$$\mathcal{C}_x^{i-\ell,i} = (\mathcal{B}_x^{i-\ell,i-\ell})^{-1} \circ \mathcal{C}_{fx}^{i-\ell,i} \circ \mathcal{A}_x^{i,i} + D_x$$

where

$$D_x = (\mathcal{B}_x^{i-\ell,i-\ell})^{-1} \circ (\mathcal{C}_{fx}^{i-\ell,i-\ell} \circ \mathcal{A}_x^{i-\ell+1,i} + \dots + \mathcal{C}_{fx}^{i-\ell,i-1} \circ \mathcal{A}_x^{i-1,i}) - \\ - (\mathcal{B}_x^{i-\ell,i-\ell})^{-1} \circ (\mathcal{B}_x^{i-\ell,i-\ell+1} \circ \mathcal{C}_x^{i-\ell+1,i} + \dots + \mathcal{B}_x^{i-\ell,i} \circ \mathcal{C}_x^{i,i}).$$

We view $\mathcal{C}_x^{i-\ell,i}$ and D_x as sections of the Hölder bundle $L(V^i, \mathcal{V}^{i-\ell})$ whose fiber at x is the space of linear maps $L(V_x^i, \mathcal{V}_x^{i-\ell})$. Thus the last equation is of the form (4.2) with

$$E = L(V^i, \mathcal{V}^{i-\ell}), \quad \phi_x = D_x, \quad \eta_x = \mathcal{C}_x^{i-\ell,i}, \quad \text{and} \quad \mathcal{F}_x(\eta_{fx}) = (\mathcal{B}_x^{i-\ell,i-\ell})^{-1} \circ \eta_{fx} \circ \mathcal{A}_x^{i,i}.$$

We note that D_x is β-Hölder since we inductively know that all its terms are β-Hölder. Also $\mathcal{F}$ is a linear cocycle on the bundle $L(V^i, \mathcal{V}^{i-\ell})$ over f^{-1}, and it is β-Hölder since so are $\mathcal{B}^{i-\ell,i-\ell}$ and $\mathcal{A}^{i,i}$. Moreover, $\mathcal{F}$ is uniformly bounded since cocycles $\mathcal{B}^{i-\ell,i-\ell}$ and $\mathcal{A}^{i,i}$ are conformal and their normalizations are continuously cohomologous. The latter follows since we know that $\mathcal{B}^{i-\ell,i-\ell}$ and $\mathcal{A}^{i-\ell,i-\ell}$ are continuously cohomologous by $\mathcal{C}^{i-\ell,i-\ell}$ and that the normalizations of all $\mathcal{A}^{i,i}$ are given by the same function ϕ^{-1} from Theorem 4.5. Hence we can apply Proposition 4.3 and conclude that $\mathcal{C}^{i-\ell,i}$ is β-Hölder.

The argument above applies $\ell = 1, \dots i-1$ and we conclude that all $\mathcal{C}^{1,i}, \dots, \mathcal{C}^{i,i}$ are Hölder. This proves that the restriction of $\mathcal{C}$ to U^i is Hölder and completes the inductive step. We conclude that $\mathcal{C}$ is Hölder, completing the proof of Theorem 2.1.

5. Proof of Theorem 2.2

In this proof we will also work with a conjugacy $\mathcal{C}$ satisfying (4.4). First, Hölder continuity of $\mathcal{C}$ is deduced from Theorem 2.1 as follows.

Let $A \in GL(N, \mathbb{R})$ be the generator of the constant cocycle $\mathcal{A}$. Let $\rho_1 < \dots < \rho_L$ be the distinct moduli of the eigenvalues of A and let

$$\mathbb{R}^N = E^1 \oplus \dots \oplus E^L \tag{5.1}$$

be the corresponding invariant splitting as in (3.1). In this section we will use the adapted norm on $\mathbb{R}^N$ for which we have estimates (3.4). They imply that for any $\beta > 0$ the cocycle $\mathcal{A}_i$ generated by A_i is fiber bunched if ϵ is sufficiently small.

Let $B(x) = \mathcal{B}_x : \mathcal{M} \to GL(N, \mathbb{R})$ be the generator of the cocycle $\mathcal{B}$. If B is sufficiently C^0 close to A, then $\mathcal{B}$ has Hölder continuous invariant splitting C^0 close to (5.1)

$$\mathbb{R}^N = \mathcal{E}_x^1 \oplus \dots \oplus \mathcal{E}_x^L,$$

so that the restrictions $\mathcal{B}_i = \mathcal{B}|\mathcal{E}^i$ satisfy estimates similar to (3.4)

$$(\rho_i - 2\epsilon)^n \le \|\mathcal{B}_i^n u\| \le (\rho_i + 2\epsilon)^n \quad \text{for any unit vector } u \in \mathcal{E}^i. \tag{5.2}$$

This is well known but also follows from Lemma 5.1, which gives explicit estimates of both Hölder exponent and Hölder constant. We conclude that all restrictions $\mathcal{B}_i$ are β-Hölder and hence are fiber bunched if ϵ is sufficiently small.

Let $\mathcal{C}$ be a measurable conjugacy between $\mathcal{A}$ and $\mathcal{B}$. We claim that $\mathcal{C}$ maps E^i to $\mathcal{E}^i$, that is $\mathcal{C}_x(E^i) = \mathcal{E}^i_x$ for μ a.e. x. Indeed, by Lemma 4.4, for μ a.e. x and for each unit vector $u \in E^i$ the forward and backward Lyapunov exponent of $\mathcal{C}_x(u)$ is $\ln \rho_i$. This yields that $\mathcal{C}_x(u) \in \mathcal{E}^i$, as having a non-zero component in another $\mathcal{E}^j$ would imply having forward or backward Lyapunov exponent under $\mathcal{B}$ different from $\ln \rho_i$ if ϵ is sufficiently small. Then $\mathcal{C}_i = \mathcal{C}|_{E^i}$ is a measurable conjugacy between fiber bunched cocycles $\mathcal{A}_i$ and $\mathcal{B}_i$. By Theorem 2.1 each $\mathcal{C}_i$ is Hölder for all $i = 1, \dots, L$, and hence so is $\mathcal{C}$.

Now we prove the more detailed statement. We denote the Lipschitz constants of f^{-1} and f respectively by

$$\alpha_f = \sup_{x\in\mathcal{M}} \|D_x f^{-1}\| > 1 \quad \text{and} \quad \alpha'_f = \sup_{x\in\mathcal{M}} \|D_x f\| > 1. \tag{5.3}$$

For $1 \le i < L$ we define

$$\beta_i = \frac{\ln(\rho_{i+1}/\rho_i)}{\ln(\alpha_f)} \quad \text{and} \quad \beta'_i = \frac{\ln(\rho_{i+1}/\rho_i)}{\ln(\alpha'_f)},$$

and we choose

$$\beta_0 = \beta_0(A, f) = \min\{1,\ \beta_1, \dots, \beta_{L-1},\ \beta'_1, \dots, \beta'_{L-1}\} > 0. \tag{5.4}$$

Since $\mathcal{B}$ is β-Hölder with $\beta \le \beta' < \beta_0$, Lemma 5.1 shows that the splitting (5.2) is β-Hölder and by Lemma 5.4 so are all restrictions $\mathcal{B}_i$. Then by Theorem 2.1 each restriction $\mathcal{C}_i = \mathcal{C}|_{E^i}$ is β-Hölder and hence so is $\mathcal{C}$. Since $\mathcal{A}_i$ and $\mathcal{B}_i$ are β-fiber bunched for any sufficiently small ϵ, [S15, Proposition 4.5] yields that β-Hölder $\mathcal{C}_i$ intertwines their stable holonomies, that is,

$$\mathcal{H}^{\mathcal{A}_i,s}_{x,y} = \mathcal{C}_i(y) \circ \mathcal{H}^{\mathcal{B}_i,s}_{x,y} \circ \mathcal{C}_i(x)^{-1} \quad \text{for all } x, y \in \mathcal{M} \text{ such that } y \in W^s(x). \tag{5.5}$$

Since for the constant cocycle $\mathcal{A}_i$ the holonomies are all identity, $\mathcal{H}^{\mathcal{A}_i,s}_{x,y} = \mathrm{Id}$, we get

$$\mathcal{C}_i(x) = \mathcal{C}_i(y) \circ \mathcal{H}^{\mathcal{B}_i,s}_{x,y}.$$

Thus using Lemma 5.5 we obtain that for all $y \in W^s(x)$

$$\|\mathcal{C}_i(x) - \mathcal{C}_i(y)\| = \|\mathcal{C}_i(y) \circ (\mathcal{H}^{\mathcal{B}_i,s}_{x,y} - \mathrm{Id})\| \le \|\mathcal{C}_i\|_{C^0} \cdot k_3\, K_\beta(\mathcal{B}) \cdot d_{W^s}(x,y)^\beta.$$

Combining these estimates for all $i = 1, \dots, L$ we conclude that all $y \in W^s(x)$

$$\|\mathcal{C}(x) - \mathcal{C}(y)\| \le \|\mathcal{C}\|_{C^0} \cdot k_4\, K_\beta(\mathcal{B}) \cdot d_{W^s}(x,y)^\beta.$$

Similarly, using the analog of Lemma 5.5 for unstable holonomies, we obtain the same estimate for $y \in W^u(y)$. Then the local product structure of stable and unstable foliations of f implies that the β-Hölder constant of $\mathcal{C}$ can be estimated as

$$K_\beta(\mathcal{C}) \le k\, \|\mathcal{C}\|_{C^0}\, K_\beta(\mathcal{B}).$$

Now, to completes the proof of the second part of the theorem, we state and prove the lemmas used in the above argument.

Lemma 5.1. *For any $0 < \beta' < \beta_0$ there is $\delta > 0$ and $k_1 > 0$ such that for any $0 < \beta \le \beta'$ any β-Hölder $GL(N, \mathbb{R})$ cocycle $\mathcal{B}$ with $\|\mathcal{B}_x - A\|_{C^0} < \delta$ preserves β-Hölder splitting*

$$\mathbb{R}^N = \mathcal{E}_x^1 \oplus \cdots \oplus \mathcal{E}_x^L$$

which is C^0 close to $E^1 \oplus \cdots \oplus E^L$ and for each $1 \le i \le L$ the β-Hölder constant $K_\beta(\mathcal{E}^i)$ of $\mathcal{E}^i$ satisfies

$$K_\beta(\mathcal{E}^i) \le k_1 \, K_\beta(\mathcal{B}) \tag{5.6}$$

Proof. We deduce this lemma from the one below. We fix $1 \le i < L$, and let

$$E' = E^1 \oplus \cdots \oplus E^i \quad \text{and} \quad E = E^{i+1} \oplus \cdots \oplus E^L.$$

Lemma 5.3 below shows that for any $\beta' < \beta_i$ there is $\delta > 0$ and k' such that for any $0 < \beta \le \beta'$ any cocycle $\mathcal{B}$ with $\|\mathcal{B}_x - A\|_{C^0} < \delta$ preserves the bundle $\mathcal{E}$ close to E with the desired estimate for β-Hölder constant. Similarly, for any $\beta' < \beta_i'$ using the inverses of A and f we obtain that $\mathcal{B}$ preserves a bundle $\mathcal{E}'$ close to E' with a similar estimate for its β-Hölder constant. Then for each $1 \le i \le L$ the bundle $\mathcal{E}^i$ is defined as a suitable intersection and hence is also C^0 close to E^i and its β-Hölder constant satisfies (5.6). ☐

Remark 5.2. Lemmas 5.1 and 5.3 do not rely on hyperbolicity of f and use only that it is bi-Lipschitz.

Lemma 5.3. *Let $A \in GL(N, \mathbb{R})$, let $\mathbb{R}^N = E' \oplus E$ be an A-invariant splitting, and let*

$$\xi' = \max\{\, \|Av\| : v \in E',\ \|v\| = 1\,\} = \|A|_{E'}\| \quad \textit{and}$$
$$\xi = \min\{\, \|Av\| : v \in E,\ \|v\| = 1\,\} = \|A^{-1}|_E\|^{-1}.$$

Let $\alpha_f = \sup \|Df^{-1}\| > 1$ be the Lipschitz constant of f^{-1} and let $\beta' > 0$. Suppose that

$$\xi' < \xi \quad \textit{and} \quad \frac{\xi' \alpha_f^{\beta'}}{\xi} < 1, \quad \textit{that is,} \quad \beta' < \frac{\ln(\xi/\xi')}{\ln \alpha_f}.$$

Then there is $\delta > 0$ and k' such that for any $0 < \beta \le \beta'$ any β-Hölder $GL(N, \mathbb{R})$ cocycle $\mathcal{B}$ with $\|\mathcal{B}_x - A\|_{C^0} < \delta$ preserves a β-Hölder sub-bundle $\mathcal{E}$ which is C^0 close to E and its β-Hölder constant $K_\beta(\mathcal{E})$ satisfies

$$K_\beta(\mathcal{E}) \le k' \, K_\beta(\mathcal{B})$$

Proof. The argument is similar to the Hölder version the C^r Section Theorem of M. Hirsch, C. Pugh, and M. Shub (see Theorem 3.8 in [HPS77]), but we give the estimate of the Hölder constant.

We consider the space $\mathcal{L} = \mathcal{L}(E, E')$ of linear operators from E to E' and endow it with the standard operator norm. Since A preserves the splitting $E' \oplus E$ it induces the graph transform action $\hat{A}$ on $\mathcal{L}$ as follows: if $L \in \mathcal{L}$ and $G \subset \mathbb{R}^N$ is its graph then $\hat{A}(L)$ is the operator in $\mathcal{L}$ whose graph is $A(G)$. The map $\hat{A}$ is linear,

$$\hat{A}\,[L] = A|_{E'} \circ L \circ (A|_E)^{-1},$$

so we can estimate its norm as

$$\|\hat{A}\| \le \|A|_{E'}\| \cdot \|(A|_E)^{-1}\| \le \xi'/\xi < 1.$$

Similarly, any linear map $B \in GL(N,\mathbb{R})$ sufficiently close to A induces in the same way the graph transform map $\hat{B}$ on a unit ball $\mathcal{L}_1$ in $\mathcal{L}$. Moreover, $\hat{B}$ is a contraction of $\mathcal{L}_1$ with Lipschitz constant $K(\hat{B})$ close to $K(\hat{A}) = \xi'/\xi < 1$. Indeed, B induces an algebraic map on the Grassmannian of $(\dim E)$-dimensional subspaces which, together with its first derivatives, depends continuously on B. Also, it is easy to see that the map $B \mapsto \hat{B}$ from a small neighborhood of A to $C^0(\mathcal{L}_1, \mathcal{L}_1)$ is Lipschitz with some constant $\hat{L}$.

Now we consider the trivial fiber bundle $\mathcal{V} = \mathcal{M} \times \mathcal{L}_1$. Then any $\mathcal{B}_x$ which is C^0-close to A induces graph transform maps $\hat{\mathcal{B}}_x : \mathcal{V}_x \to \mathcal{V}_{fx}$ and thus the bundle map $\hat{\mathcal{B}} : \mathcal{V} \to \mathcal{V}$ covering f. We consider the space S of continuous sections of $\mathcal{V}$ with the supremum norm, and the induced action $F = F_{\mathcal{B}}$ on S defined for $s \in S$ as $(Fs)(fx) = \mathcal{B}_x(s(x))$. If $K^{\mathcal{B}} := \sup_x K(\hat{\mathcal{B}}_x) < 1$ then F is a contraction on S and hence has a unique fixed point $s_* = Fs_*$. Let $s_0(x) = 0 \in \mathcal{L}$ be the zero section, then we can write $s_* = \lim F^n s_0$ and it follows that s_* is C^0-close to s_0. Denoting the graph of $s(x)$ by $\mathcal{E}_x$ we obtain the unique continuous $\mathcal{B}$-invariant sub-bundle close to E.

Now we will show that s_* is β-Hölder and estimate its β-Hölder constant. For this we will find $M > 0$ such that $K_\beta(s) \le M$ implies $K_\beta(Fs) \le M$. Then $K_\beta(F^n(s_0)) \le M$ for all n and since $s_* = \lim F^n(s_0)$ it will follow that $K_\beta(s_*) \le M$.

Fix points z, z' and let $x = f(z)$, $x' = f(z')$. Then for any β-Hölder $s \in S$ we can estimate, as $\|s(x)\| \le 1$, that

$$\begin{aligned}
&\|Fs(x) - Fs(x')\| = \|\hat{\mathcal{B}}_z s(z) - \hat{\mathcal{B}}_{z'} s(z')\| \\
&\le \|\hat{\mathcal{B}}_z s(z) - \hat{\mathcal{B}}_{z'} s(z)\| + \|\hat{\mathcal{B}}_{z'} s(z) - \hat{\mathcal{B}}_{z'} s(z')\| \\
&\le d_{C^0}(\hat{\mathcal{B}}_z, \hat{\mathcal{B}}_{z'}) + K(\hat{\mathcal{B}}_{z'})\|s(z) - s(z')\| \le \hat{L}\|\mathcal{B}_z - \mathcal{B}_{z'}\| + K^{\mathcal{B}}\|s(z) - s(z')\| \\
&\le \hat{L}\, K_\beta(\mathcal{B})\, d(z,z')^\beta + K^{\mathcal{B}} K_\beta(s)\, d(z,z')^\beta \le [\hat{L}\, K_\beta(\mathcal{B}) + K^{\mathcal{B}} K_\beta(s)]\, (\alpha_f\, d(x,x'))^\beta,
\end{aligned}$$

where α_f is the Lipschitz constant of f^{-1} and $\hat{L}$ is the Lipschitz constant of the map $B \mapsto \hat{B}$ on a neighborhood of A. Hence Fs is also β-Hölder and

$$K_\beta(Fs) \le \hat{L}\, \alpha_f^\beta\, K_\beta(\mathcal{B}) + \alpha_f^\beta\, K^{\mathcal{B}} K_\beta(s).$$

Therefore, $K_\beta(s) \le M$ implies $K_\beta(Fs) \le M$ if we take

$$M = (1 - K^{\mathcal{B}} a_f^\beta)^{-1} \hat{L}\, a_f^\beta\, K_\beta(\mathcal{B}).$$

If $\|\mathcal{B}_x - A\|_{C^0}$ is small then $K^{\mathcal{B}}$ is close to $K(\hat{A}) = \xi'/\xi$. Since $\xi' \alpha^{\beta'}/\xi < 1$ and $\beta \le \beta'$ it follows that $1 - K^{\mathcal{B}} a_f^\beta > 0$ and is separated from 0. Then there is a constant k' which bounds $\hat{L} a_f^\beta\, (1 - K^{\mathcal{B}} a_f^\beta)^{-1}$ for all $0 < \beta \le \beta'$ and all $\mathcal{B}$ with $\|\mathcal{B}_x - A\|_{C^0} < \delta$. Hence,

$$M \le k'\, K_\beta(\mathcal{B}).$$

Finally, since $K_\beta(s_0) = 0$ it follows that $K_\beta(F^n(s_0)) \le M$ for all n and hence for the limit we also have $K_\beta(s_*) \le M \le k' K_\beta(\mathcal{B})$. $\square$

Now we estimate the β-Hölder constants of the restricted cocycles $\mathcal{B}_i = \mathcal{B}|_{\mathcal{E}^i}$.

Lemma 5.4. *For any $0 < \beta' < \beta_0$ there is $\delta > 0$ and $k_2 > 0$ such that for any $0 < \beta \le \beta'$ and any β-Hölder cocycle $\mathcal{B}$ with $\|\mathcal{B}_x - A\|_{C^0} < \delta$ the β-Hölder constant of the cocycle $\mathcal{B}_i$, $i = 1, \dots L$, satisfies*

$$K_\beta(\mathcal{B}_i) \le k_2\, K_\beta(\mathcal{B}).$$

Proof. Denoting $B(x) = \mathcal{B}_x$ and $B_i(x) = \mathcal{B}_x|_{\mathcal{E}^i}$ we need to estimate the distance between $B_i(x)$ and $B_i(y)$. To do this using their difference, we fix β-Hölder identifications $I_{x,y} : \mathcal{E}^i_x \to \mathcal{E}^i_y$, say by translation from x to y in the trivial bundle $\mathcal{M} \times \mathbb{R}^N$ followed by an appropriate rotation. Then for a unit vector $u \in \mathcal{E}^i(x)$ we need to estimate $\|(B_i(x) - B_i(y) \circ I_{x,y})u\|$. We note that

$$\|u - I_{x,y}u\| \le \operatorname{dist}(\mathcal{E}^i_x, \mathcal{E}^i_y) \le K_\beta(\mathcal{E}^i)\, d(x,y)^\beta.$$

Also, since $B(x)$ is β-Hölder have $\|B(x)u - B(y)u\| \le K_\beta(\mathcal{B})\, d(x,y)^\beta$. Hence we obtain that for a unit vector $u \in \mathcal{E}^i(x)$

$$\begin{aligned}\|(B_i(x) - B_i(y) \circ I_{x,y})u\| &\le \|B(x)u - B(y)u\| + \|B(y)\| \cdot \|u - I_{x,y}u\| \\ &\le K_\beta(\mathcal{B})\, d(x,y)^\beta + \|B\|_{C^0}\, K_\beta(\mathcal{E}^i)\, d(x,y)^\beta.\end{aligned}$$

Since $K_\beta(\mathcal{E}^i) \le k_1\, K_\beta(\mathcal{B})$ by (5.6) and $\|B\|_{C^0} \le \|A\| + \|\mathcal{B}_x - A\|_{C^0} \le \|A\| + \delta$ we conclude that

$$\|(B_i(x) - B_i(y) \circ I_{x,y})u\| \le k_2\, K_\beta(\mathcal{B})\, d(x,y)^\beta.$$

Thus $K_\beta(\mathcal{B}_i) \le k_2\, K_\beta(\mathcal{B})$. $\square$

In the next lemma we consider the stable holonomies of cocycles $\mathcal{B}_i = \mathcal{B}|_{\mathcal{E}^i}$, $i = 1, \dots, L$.

Lemma 5.5. *For any $0 < \beta' < \beta_0$ there is $\delta > 0$ and $k_3 > 0$ such that for any $0 < \beta \le \beta'$ and a β-Hölder cocycle $\mathcal{B}$ with $\|\mathcal{B}_x - A\|_{C^0} < \delta$ the holonomies of cocycles $\mathcal{B}_i = \mathcal{B}|_{\mathcal{E}^i}$ satisfy*

$$\|\mathcal{H}^s_{x,y} - Id\| \le k_3\, K_\beta(\mathcal{B})\, d(x,y)^\beta \quad \textit{for any } x \in \mathcal{M} \textit{ and } y \in W^s_{loc}(x).$$

Proof. We fix i and denote $\mathcal{F} = \mathcal{B}_i$. The stable holonomies of $\mathcal{F}$ are given by

$$\mathcal{H}^{\mathcal{F},s}_{x,y} = \lim_{n\to\infty} (\mathcal{F}^n_y)^{-1} \circ \mathcal{F}^n_x. \tag{5.7}$$

The existence is ensured by fiber bunching of $\mathcal{F}$. Indeed, the contraction along W^s is estimated by (2.1) as

$$d(f^n x, f^n y) \le \nu^n d(x,y) \ \text{ for any } x \in \mathcal{M},\ y \in W^s_{\text{loc}}(x),\ n \in \mathbb{N},$$

We also obtain from (5.2) that

$$\|\mathcal{F}^m_x\| \cdot \|(\mathcal{F}^m_y)^{-1}\| \le \prod_{j=0}^{m-1} \|\mathcal{F}_{x_j}\|\, \|(\mathcal{F}_{y_j})^{-1}\| \le \left(\frac{\rho_i + 2\epsilon}{\rho_i - 2\epsilon}\right)^m = \sigma^m \quad \text{for all } x, y \in \mathcal{M}, \tag{5.8}$$

where $\sigma = (\rho_i + 2\epsilon)(\rho_i - 2\epsilon)^{-1}$ is close to 1 when ϵ is small. It follows that

$$\|\mathcal{F}_x^m\| \cdot \|(\mathcal{F}_y^m)^{-1}\| \cdot \nu^{m\beta} \le \sigma^m \cdot \nu^{m\beta} = \theta^m \quad \text{for all } x, y \in \mathcal{M}, \tag{5.9}$$

where $\theta = \sigma\nu^\beta < 1$ if δ and hence ϵ are sufficiently small. In particular, $\mathcal{F}$ is fiber bunched so the limit in (5.7) exits, though this also follows from the proof.

We want to obtain a constant c such that $\|\mathcal{H}_{x,y}^{\mathcal{F},s} - \mathrm{Id}\,\| \le c\, d(x,y)^\beta$ for all $x \in \mathcal{M}$ and $y \in W_{\mathrm{loc}}^s(x)$. Denoting $x_m = f^m(x)$ and $y_m = f^m(y)$, we obtain

$$\begin{aligned}
&(\mathcal{F}_y^n)^{-1} \circ \mathcal{F}_x^n = (\mathcal{F}_y^{n-1})^{-1} \circ \big((\mathcal{F}_{y_{n-1}})^{-1} \circ \mathcal{F}_{x_{n-1}}\big) \circ \mathcal{F}_x^{n-1} \\
&= (\mathcal{F}_y^{n-1})^{-1} \circ (\mathrm{Id} + r_{n-1}) \circ \mathcal{F}_x^{n-1} = (\mathcal{F}_y^{n-1})^{-1} \circ \mathcal{F}_x^{n-1} + (\mathcal{F}_y^{n-1})^{-1} \circ r_{n-1} \circ \mathcal{F}_x^{n-1} \\
&= \cdots = \mathrm{Id} + \sum_{m=0}^{n-1} (\mathcal{F}_y^m)^{-1} \circ r_m \circ \mathcal{F}_x^m, \quad \text{where } r_m = (\mathcal{F}_{y_m})^{-1} \circ \mathcal{F}_{x_m} - \mathrm{Id}.
\end{aligned}$$

Since $\mathcal{F}$ is β-Hölder, denoting $c' = (\rho_i - 2\epsilon)^{-1} K_\beta(\mathcal{F})$, we obtain that for every $m \ge 0$

$$\|r_m\| \le \|(\mathcal{F}_{y_m})^{-1}\| \cdot \|\mathcal{F}_{x_m} - \mathcal{F}_{y_m}\| \le \|\mathcal{F}^{-1}\|_{C^0} K_\beta(\mathcal{F})\, d(x_m, y_m)^\beta \le c'\, d(x,y)^\beta \nu^{m\beta}.$$

Using (5.9) it follows that

$$\|(\mathcal{F}_y^m)^{-1} \circ r_m \circ \mathcal{F}_x^m\| \le \|(\mathcal{F}_y^m)^{-1}\| \cdot \|\mathcal{F}_x^m\| \cdot c'\, d(x,y)^\beta \nu^{m\beta} \le \theta^m\, c'\, d(x,y)^\beta.$$

Therefore, for every $n \in \mathbb{N}$,

$$\|\mathrm{Id} - (\mathcal{F}_y^n)^{-1} \circ \mathcal{F}_x^n\| \le \sum_{i=0}^{n-1} \|(\mathcal{F}_y^i)^{-1} \circ r_i \circ \mathcal{F}_x^i\| \le c'\, d(x,y)^\beta \sum_{i=0}^{n-1} \theta^i \le c\, d(x,y)^\beta,$$

where

$$c = \frac{c'}{1-\theta} \le \frac{(\rho_i - 2\epsilon)^{-1} K_\beta(\mathcal{F})}{1 - \sigma\nu^\beta} = k_3'\, K_\beta(\mathcal{F}) \quad \text{with} \quad k_3' = (\rho_i - 2\epsilon)^{-1}(1 - \sigma\nu^\beta)^{-1}.$$

By (5.7) the sequence $\{(\mathcal{F}_y^n)^{-1} \circ \mathcal{F}_x^n\}$ converges to $\mathcal{H}_{xy}^{\mathcal{F},s}$ (in fact the estimates imply that it is Cauchy) and the limit satisfies

$$\|\mathcal{H}_{x,y}^s - \mathrm{Id}\,\| \le c\, d(x,y)^\beta \;\text{ for any } x \in \mathcal{M} \text{ and } y \in W_{\mathrm{loc}}^s(x).$$

By Lemma 5.4 we have $K_\beta(\mathcal{F}) = K_\beta(\mathcal{B}_i) \le k_2\, K_\beta(\mathcal{B})$ and we conclude that

$$\|\mathcal{H}_{x,y}^s - \mathrm{Id}\,\| \le k_3\, K_\beta(\mathcal{B})\, d(x,y)^\beta \;\text{ for any } x \in \mathcal{M} \text{ and } y \in W_{\mathrm{loc}}^s(x).$$

This completes the proof of Lemma 5.5 $\square$

6. Proof of Theorem 1.1

Any two continuous conjugacies between f and A differ by an element of the centralizer of A. By [Wa70, Corollary 1], any homeomorphism commuting with an ergodic, in particular hyperbolic, automorphism A is an affine automorphism, and hence all conjugacies have the same regularity.

First, using Theorem 2.2 we will show in Section 6.1 that H is a $C^{1+\text{Hölder}}$ diffeomorphism, and moreover the Hölder constant of its derivative satisfies the estimate

$$K_\beta(DH) \le k\,\|DH\|_{C^0}\,\|f-A\|_{C^{1+\beta}}. \tag{6.1}$$

This part does not rely on closeness of H to the identity and the estimate applies to any conjugacy H. Then in Section 6.2 we use (6.1) and an interpolating inequality to obtain the desired estimate (1.2) of $\|H-I\|_{C^{1+\beta}}$ for the conjugacy C^0 close to the identity.

6.1. **Proving that H is a $C^{1+\textbf{Hölder}}$ diffeomorphism.**

First we recall some properties of a map $g \in W^{1,q}(\mathbb{R}^N,\mathbb{R}^N)$, $q>N$, which also extend to the case when $g \in W^{1,q}(\mathbb{T}^N,\mathbb{T}^N)$. It is well known that, as a consequence of Morrey's inequality, for any such g the Jacoby matrix of weak partial derivatives gives the differential D_xg for almost every x with respect to the Lebesgue measure μ. Also, any such g satisfies *Lusin's N-property* [MM73] that $\mu(E)=0$ implies $\mu(g(E))=0$, as well as *Morse-Sard property* [P01] that $\mu(g(\mathcal{C}))=0$ for the set of critical points of g

$$\mathcal{C}_g = \{x \in \mathbb{T}^N :\ D_xg \text{ exists but is not invertible}\},$$

see also [KK18] for sharper results and further references.

Now we assume that $H \in W^{1,q}$ with $q>N$, so that the differential D_xH exists μ-a.e., and for the set

$$G_H = \{x \in \mathbb{T}^N :\ D_xH \text{ exists}\} \ \text{ and its complement } \ E_H = \mathbb{T}^N \setminus G_H$$

we have $\mu(G_H)$=1 and $\mu(E_H)=0$. Further $G_H = \mathcal{C}_H \cup R_H$ is the disjoint union of two measurable sets, the critical set $\mathcal{C}_H$ and the regular set

$$R_H = \{x \in \mathbb{T}^N :\ D_xH \text{ is invertible}\}.$$

Since f and A are diffeomorphisms, it follows from the conjugacy equation $H \circ f = A \circ H$ that the sets G_H, $\mathcal{C}_H$, and R_H are f-invariant. Further, differentiating the equation on the set G_H we obtain

$$D_{fx}H \circ D_xf = A \circ D_xH. \tag{6.2}$$

Denoting $\mathcal{C}(x) = D_xH$ on the set R_H we obtain the conjugacy equation over f

$$A = \mathcal{C}(fx) \circ \mathcal{B}_x \circ \mathcal{C}(x)^{-1} \quad \text{for cocycles } \ \mathcal{B}_x = D_xf \text{ and } \mathcal{A}_x = A. \tag{6.3}$$

Now we show that $\mu(R_H)=1$ and also that f preserves a measure $\tilde\mu$ equivalent to μ. Since $\mu(E_H)=0$, the Lusin's N-property of H yields $\mu(H(E_H))=0$. Also, we have $\mu(H(\mathcal{C}_H))=0$ by the Morse-Sard property. Hence for $R'_H = H(R_H)$ we have $\mu(R'_H)=1$. Now we consider the measure $\tilde\mu = (H^{-1})_*(\mu)$ and note that $\tilde\mu(R_H)=1$ as $\mu(R'_H)=1$. Since H is a topological conjugacy between f and A, $\tilde\mu$ is f-invariant and, in fact, is the Bowen-Margulis measure of maximal entropy for f, since μ is that for A. Indeed, denoting the topological entropy by $\mathbf{h}_{top}$ and metric entropy with respect to $\tilde\mu$ by $\mathbf{h}_{\tilde\mu}$ we get

$$\mathbf{h}_{\tilde\mu}(f) = \mathbf{h}_\mu(A) = \mathbf{h}_{top}(A) = \mathbf{h}_{top}(f).$$

In particular, $\tilde{\mu}$ is ergodic with full support and local product structure. Since $\mathcal{C}$ is a conjugacy between $\mathcal{B}$ and A on R_H with $\tilde{\mu}(R_H) = 1$, by Lemma 4.4 we obtain that the Lyapunov exponents $\lambda_i^{f,\tilde{\mu}}$ of $\tilde{\mu}$ for the cocycle $\mathcal{B} = Df$ are equal to the Lyapunov exponents λ_i^A of A. Hence the sum of positive Lyapunov exponents (counted with multiplicities) for $\tilde{\mu}$ equals its entropy

$$\mathbf{h}_{\tilde{\mu}}(f) = \mathbf{h}_{\mu}(A) = \sum_{\lambda_i^A > 0} \lambda_i^A = \sum_{\lambda_i^{f,\tilde{\mu}} > 0} \lambda_i^{f,\tilde{\mu}}.$$

Thus we have equality in the Pesin-Ruelle formula, which implies that $\tilde{\mu}$ has absolutely continuous conditional measures on the unstable foliation of f [Le84]. Similarly, equality of the negative Lyapunov exponents yields that $\tilde{\mu}$ has absolutely continuous conditional measures on the stable foliation of f. We conclude that $\tilde{\mu}$ itself is absolutely continuous. Moreover, the density $\sigma(x) = \frac{d\tilde{\mu}}{d\mu}$ is smooth and positive as a measurable solution of the coboundary equation $\sigma(fx)\sigma(x)^{-1} = \det Df(x)$. Thus $\tilde{\mu}$ is equivalent to μ, so that $\tilde{\mu}(R_H) = 1$ implies $\mu(R_H) = 1$.

Provided that $\|A - \mathcal{B}_x\|_{C^0} = \|A - D_x f\|_{C^0} \le \|A - f\|_{C^1} < \delta$, where $\delta > 0$ is from Theorem 2.2, we can apply this theorem with f and $\tilde{\mu}$ to obtain that

$$\mathcal{C}(x) = D_x H : \mathbb{T}^N \to GL(N, \mathbb{R})$$

coincides with a Hölder continuous function almost everywhere with respect to $\tilde{\mu}$ and hence μ. Since $H \in W^{1,q}$ we conclude that H is $C^{1+\text{Hölder}}$. Also, since $(D_x H)^{-1} = \mathcal{C}(x)^{-1}$ exists and is also Hölder continuous we see that H is $C^{1+\text{Hölder}}$ diffeomorphism. Further, Theorem 2.2 gives us the estimate (6.1), which we will use to obtain the desired estimate for $\|H - \text{Id}\|_{C^{1+\beta}}$ in Section 6.2. This completes the proof that H is $C^{1+\text{Hölder}}$ diffeomorphism assuming that $H \in W^{1,q}$.

Now we consider the case when $\tilde{H} = H^{-1}$ is in $W^{1,q}$ and hence $D_x\tilde{H}$ exists μ-a.e. We similarly define the sets $G_{\tilde{H}}$, $E_{\tilde{H}}$, $\mathfrak{C}_{\tilde{H}}$, and $R_{\tilde{H}}$, which are measurable and A-invariant. Hence by ergodicity of A the set $R_{\tilde{H}}$ must be null or co-null for μ. If $\mu(R_{\tilde{H}}) = 0$ then $\mu(\tilde{H}(R_{\tilde{H}})) = 0$ by the Lusin's N-property of $\tilde{H}$, but this is impossible since $\mu(\tilde{H}(E_{\tilde{H}})) = 0$ by the Lusin's N-property and $\mu(\tilde{H}(\mathfrak{C}_{\tilde{H}})) = 0$ by the Morse-Sard property. Hence $\mu(R_{\tilde{H}}) = 1$. Then for $R'_{\tilde{H}} = \tilde{H}(R_{\tilde{H}})$ we have $\tilde{\mu}(R'_{\tilde{H}}) = 1$, where as before $\tilde{\mu} = \tilde{H}_*(\mu)$ is the measure of maximal entropy for f. Now the Lusin's N-property of $\tilde{H}$ yields that $\tilde{\mu}$ is absolutely continuous and then equivalent to μ. Hence we also have $\mu(R'_{\tilde{H}}) = 1$. Since $H = \tilde{H}^{-1}$ is a homeomorphism, and $D_x\tilde{H}$ is invertible for $x \in R_{\tilde{H}}$, it follows that $D_y H = (D_x\tilde{H})^{-1}$ is the differential of H for each $y = \tilde{H}(x)$ in $R'_{\tilde{H}}$.

Therefore, we can again differentiate $H \circ f = A \circ H$ to obtain (6.3) and then the conjugacy equation (6.3) with $\mathcal{C}(x) = D_x H$ on the set $R'_{\tilde{H}}$ of full measure for both μ and $\tilde{\mu}$. Then by Theorem 2.2 applied with f and $\tilde{\mu}$ we obtain that $\mathcal{C}(x) = D_x H$ is Hölder on $\mathbb{T}^N$ and hence so is $\mathcal{C}(y)^{-1} = D_x\tilde{H}$. Since $\tilde{H} = H^{-1}$ is in $W^{1,q}$ we conclude that H^{-1} is $C^{1+\text{Hölder}}$ diffeomorphism. In this case we also get (6.1).

6.2. **Estimating $\|H - I\|_{C^{1+\beta}}$.** We showed that any conjugacy H is a $C^{1+\text{Hölder}}$ diffeomorphism satisfying (6.1). Now we prove estimate (1.2) for the conjugacy H that is C^0 close to the identity.

Any two conjugacies in the homotopy class of the identity differ by a composition with an affine automorphism commuting with A, which is translation $T_v(x) = x + v$, where $v \in \mathbb{T}^N$ is a fixed point of A. It is well known that if f is C^1-close to A, then it has a unique fixed point p which is the perturbation of 0. More precisely, there are $0 < \delta(A), r(A) < 1/5$ and $k(A)$ so that for each f satisfying $\|f - A\|_{C^1} < \delta(A)$ there is a unique fixed point $p = f(p)$ with $d(p, 0) < r(A)$ and it satisfies

$$d(p, 0) \le k(A)\|f - A\|_{C^0}.$$

Since H maps fixed points of f to those of A we see that if $\|H - I\|_{C^0} < r(A)$ then it is in the homotopy class of the identity and satisfies $H(p) = 0$.

Replacing f by $\tilde{f} = T_{-p} \circ f \circ T_p$ we can change p to 0. Since for $\tilde{f}(x) = f(x + p) - p$ we have that

$$\|D\tilde{f} - A\|_{C^k} = \|Df - A\|_{C^k} \quad \text{for any } k \ge 0,$$

and so only $\|f - A\|_{C^0}$ is affected by this change. Moreover, if we write $f = A + R$, then

$$\tilde{f}(x) - A(x) = A(x + p) + R(x + p) - p - A(x) = R(x + p) + A(p) - p$$

and hence

$$\|\tilde{f} - A\|_{C^0} \le \|R\|_{C^0} + \|A(p) - p\| = \|f - A\|_{C^0} + \|A(p) - f(p)\| \le 2\|f - A\|_{C^0}.$$

Thus $\|\tilde{f} - A\|_{C^{1+\beta}} \le 2\|f - A\|_{C^{1+\beta}}$. Also, if $\tilde{H}$ is the corresponding conjugacy between $\tilde{f}$ and A then $H(x) = \tilde{H}(x - p)$ and hence

$$\|H - \text{Id}\|_{C^{1+\beta}} \le \|\tilde{H} - \text{Id}\|_{C^{1+\beta}} + d(p, 0) \le \|\tilde{H} - \text{Id}\|_{C^{1+\beta}} + k(A)\|f - A\|_{C^0}$$

Thus the estimate (6.1) for $\tilde{H}$ via $\tilde{f}$ would yield the corresponding estimate for H via f. So without loss of generality we will assume that

$$f(0) = 0 \;\text{ and }\; H(0) = 0.$$

Now we recall how the conjugacy equation $H \circ f = A \circ h$ can be rewritten using lifts. We denote by $\bar{f}$ and $\bar{H}$ the lifts of f and H to $\mathbb{R}^N$ satisfying $\bar{f}(0) = 0$ and $\bar{H}(0) = 0$ so that we have $\bar{H} \circ \bar{f} = A \circ \bar{H}$ where all maps are $\mathbb{R}^N \to \mathbb{R}^N$. Since H is homotopic to the identity and f is homotopic to A we can write

$$\bar{H} = \text{Id} + h \quad \text{and} \quad \bar{f} = A + R,$$

Then the commutation relation on $\mathbb{R}^N$

$$(\text{Id} + h) \circ (A + R) = A \circ (\text{Id} + h) \quad \text{yields} \quad h = A^{-1}(h \circ \bar{f}) + A^{-1}R.$$

Since $h, R : \mathbb{R}^N \to \mathbb{R}^N$ are $\mathbb{Z}^N$-periodic we can view them as

$$h = H - \text{Id} : \; \mathbb{T}^N \to \mathbb{R}^N \quad \text{and} \quad R = f - A \; : \; \mathbb{T}^N \to \mathbb{R}^N$$

and rewrite the conjugacy equation as one for $\mathbb{R}^N$-valued functions on $\mathbb{T}^N$

$$h = A^{-1}(h \circ f) + A^{-1}R. \tag{6.4}$$

Using the A-invariant splitting $\mathbb{R}^N = E^u \oplus E^s$ we define the projections h_* and R_* of h and R to E^*, where $* = s, u$, and obtain

$$h_* = A_*^{-1}(h_* \circ f) + A_*^{-1}R_*, \quad \text{where } A_* = A|_{E^*}. \tag{6.5}$$

Thus h_* is a fixed point of the affine operator

$$T_*(\psi) = A_*^{-1}(\psi \circ f) + A_*^{-1}R_* \tag{6.6}$$

Since $\|A_u^{-1}\| < 1$, the operator T_u is a contraction on the space $C^0(\mathbb{T}^d, E^u)$, and thus h_u is its unique fixed point

$$h_u = \lim_{m\to\infty} T_u^m(0) \;=\sum_{m=0}^{\infty} A_u^{-m}(A_u^{-1}R_u \circ f^m). \tag{6.7}$$

Hence

$$\|h_u\|_{C^0} \;\le\; \sum_{m=0}^{\infty} \|A_u^{-1}\|^{m+1}\|R_u\|_{C^0} \;\le\; k\,\|R_u\|_{C^0} \;\le\; k\,\|A-f\|_{C^0}. \tag{6.8}$$

Similarly, h_s is the unique fixed point of contraction T_s^{-1} and hence satisfies a similar estimate. Combining them we conclude that

$$\|H - \mathrm{Id}\|_{C^0} = \|h\|_{C^0} \le k_0\|R\|_{C^0} = k_0\|A-f\|_{C^0}. \tag{6.9}$$

Now we estimate $\|H - \mathrm{Id}\|_{C^{1+\beta}}$ using (6.9), (6.1), and the following elementary interpolation lemma. We note that $DH = \mathrm{Id} + Dh$, so that $K_\beta(Dh) = K_\beta(DH)$.

Lemma 6.1. *If $h : \mathbb{T}^N \to \mathbb{R}^N$ satisfies $K_\beta(Dh) \le K$ then*

$$\|Dh\|_{C^0} \le 8\,\|h\|_{C^0}^{\beta/(1+\beta)} K^{1/(1+\beta)}.$$

Proof. Denote $b = \|Dh\|_{C^0}$ and choose $x \in \mathbb{T}^N$ such that $\|D_x h\| = b$. Then for some unit vectors $u, v \in \mathbb{R}^N$ we have $(D_x h)u = bv$. For $y \in \mathbb{T}^N$ let $b_y = \langle (D_y h)u, v\rangle$, so $b_x = b$. Then

$$|b - b_y| \le \|(D_x h)u - (D_y h)u\| \le K\, d(x,y)^\beta \le b/2 \quad \text{if} \quad d(x,y) \le (b/2K)^{1/\beta}$$

and hence $b_y \ge b/2$ for such y. Consider $y(t) = x + tu$, with $0 \le t \le t_0 = (b/2K)^{1/\beta}$, and $g(t) = \langle h(y(t)), v\rangle$. Then

$$g'(t) = \langle (D_y h)u, v\rangle = b_{y(t)} \ge b/2,$$

and hence by integrating we get $bt_0/2 \le g(t_0) - g(0)$. Since $|g(t_0) - g(0)| \le 2\|h\|_{C^0}$ we obtain $bt_0 \le 4\|h\|_{C^0}$. Substituting $t_0 = (b/2K)^{1/\beta}$ we obtain

$$b(b/2K)^{1/\beta} \le 4\,\|h\|_{C^0} \;\Rightarrow\; b^{(1+\beta)/\beta} \le 4\,\|h\|_{C^0}(2K)^{1/\beta} \;\Rightarrow\; b \le 8\,\|h\|_{C^0}^{\beta/(1+\beta)} K^{1/(1+\beta)}$$

as $4^{\beta/(1+\beta)}2^{1/(1+\beta)} < 8$. $\square$

We denote $a = \|h\|_{C^0}$, $b = \|Dh\|_{C^0}$, and $d = \|f - A\|_{C^{1+\beta}}$. Then

$$\|DH\|_{C^0} = \|\mathrm{Id} + Dh\|_{C^0} \le 1 + b,$$

and hence (6.1) implies that

$$K = K_\beta(Dh) = K_\beta(DH) \le k(1+b)d. \tag{6.10}$$

Also, by (6.9) we have $a = \|h\|_{C^0} \le k_0 d$. Then Lemma 6.1 gives

$$b \le 8(kd)^{\beta/(1+\beta)}(k(1+b)d)^{1/(1+\beta)} < k_1\, d(1+b)^{1/(1+\beta)}.$$

It follows that b is bounded by some k_2 if $d \le 1$. Then (6.10) implies that

$$K = K_\beta(Dh) \le k_3 d.$$

With this K Lemma 6.1 gives

$$b \le 8(kd)^{\beta/(1+\beta)}(k_3(A)d)^{1/(1+\beta)} < k_4 d.$$

We conclude that

$$b = \|Dh\|_{C^0} < k_4 d, \quad a = \|h\|_{C^0} \le k_0\, d, \;\text{ and }\; K_\beta(Dh) \le k_3 d,$$

so that

$$\|H - \mathrm{Id}\|_{C^{1+\beta}} = \|h\|_{C^{1+\beta}} \le k_5 d = k_5\, \|f - A\|_{C^{1+\beta}}.$$

This completes the proof of Theorem 1.1.

7. Linearized conjugacy equation

In this section we begin the proof of Theorem 1.3, and in the next one we will complete it using an iterative process. In these sections we fix a hyperbolic matrix $A \in SL(N, \mathbb{Z})$. We will use K to denote any constant that depends only on A, and K_x to denote a constant that also depends on a parameter x.

7.1. **Preliminaries.** Set $\tilde{A} = (A^\tau)^{-1}$ where A^τ denotes transpose matrix. We call $\tilde{A}$ the dual map on $\mathbb{Z}^N$. Since A is hyperbolic so is $\tilde{A}$, and we denote its stable and unstable subspaces by $\tilde{E}^s$ and $\tilde{E}^u$. Thus there is $\rho > 1$ ($\rho < \min\{\rho_{i_0+1}, \rho_{i_0}^{-1}\}$) such that

$$\begin{aligned} \|\tilde{A}^k v\| &\ge K\rho^k \|v\|, \quad k \ge 0,\ v \in \tilde{E}^u, \\ \|\tilde{A}^{-k} v\| &\ge K\rho^k \|v\|, \quad k \ge 0,\ v \in \tilde{E}^s. \end{aligned} \tag{7.1}$$

For a subspace V of $\mathbb{R}^N$, we use π_V to denote the (orthogonal) projection to V. For any integer vector $n \in \mathbb{Z}^N$ we write $n_s = \pi_{\tilde{E}^s} n$ and $n_u = \pi_{\tilde{E}^u} n$. Since $\tilde{A} \in SL(N, \mathbb{Z})$ is hyperbolic, for any $0 \ne n \in \mathbb{Z}^N$ both n_s and n_u are nonzero and there is a unique $k_0 = k_0(n) \in \mathbb{Z}$ such that

$$\begin{aligned} \|\tilde{A}^k n_s\| &\ge \|\tilde{A}^k n_u\| \quad \text{for all } k \le k_0 \quad \text{and} \\ \|\tilde{A}^k n_s\| &< \|\tilde{A}^k n_u\| \quad \text{for all } k > k_0. \end{aligned}$$

The corresponding element $\tilde{A}^{k_0(n)}n$ on the orbit of n will be called *minimal* and

$$M = \{\tilde{A}^{k_0(n)}n : \; 0 \neq n \in \mathbb{Z}^N\} \subset \mathbb{Z}^N \backslash 0. \tag{7.2}$$

For any $n \in M$ we have $\|n_s\| \geq \frac{1}{2}\|n\|$ and $\|\tilde{A}n_u\| > \frac{1}{2}\|\tilde{A}n\|$.

For a function $\theta \in L^2(\mathbb{T}^N, \mathbb{C})$ we denote its Fourier coefficients by $\hat{\theta}_n$, $n \in \mathbb{Z}^N$, so that

$$\theta(x) = \sum_{n \in \mathbb{Z}^N} \widehat{\theta}_n e^{2\pi \mathbf{i}\, n \cdot x} \quad \text{in } L^2(\mathbb{T}^N).$$

We say that θ is *excellent* (for A) if $\widehat{\theta}_n = 0$ for all $n \notin M$.

To simplify our estimates, instead of the standard Sobolev spaces we will work the spaces $\mathcal{H}^s(\mathbb{T}^N)$, $s > 0$, defined as follows. A function $\theta \in L^2(\mathbb{T}^N)$ belongs to $\mathcal{H}^s(\mathbb{T}^N)$ if

$$\|\theta\|_s \overset{\text{def}}{=} \sup_n |\widehat{\theta}_n| \|n\|^s + |\widehat{\theta}_0| < \infty.$$

The following relations hold (see, for example, Section 3.1 of [dlL99]). If $\sigma > N + 1$ and $r \in \mathbb{N}$, then for any $\theta \in C^r(\mathbb{T}^N)$ and $\omega \in \mathcal{H}^{r+\sigma}$ we have $\theta \in \mathcal{H}^r$ and $\omega \in C^r(\mathbb{T}^N)$ with estimates

$$\|\theta\|_r \leq K\|\theta\|_{C^r} \quad \text{and} \quad \|\omega\|_{C^r} \leq K\|\omega\|_{r+\sigma}. \tag{7.3}$$

For a vector-valued function $\theta : \mathbb{T}^N \to \mathbb{C}^m$ we denote its coordinate functions by θ_j, $j = 1, \ldots, m$. We say that θ is in $\mathcal{H}^s(\mathbb{T}^N)$ if each θ_j is in $\mathcal{H}^s(\mathbb{T}^N)$ and set

$$\|\theta\|_s \overset{\text{def}}{=} \max_{1 \leq j \leq m} \|\theta_j\|_s, \quad \widehat{\theta}_n \overset{\text{def}}{=} ((\widehat{\theta_1})_n, \ldots, (\widehat{\theta_m})_n) \quad \text{for any } n \in \mathbb{Z}^N$$

We say that θ is excellent if θ_j is excellent for each j.

7.2. **Twisted cohomological equation over A in high regularity.**

A crucial step in the iterative process is solving the twisted cohomological equation

$$A\omega - \omega \circ A = \theta \tag{7.4}$$

over A, which can be viewed as the linearized conjugacy equation. In this section we give preliminary results on solving this equation in high regularity. We start with a scalar cohomological equation over A twisted by $\lambda \in \mathbb{C} \backslash \{0, 1\}$,

$$\lambda\omega - \omega \circ A = \theta. \tag{7.5}$$

The next lemma shows that the obstructions to solving it in C^∞ category are sums of Fourier coefficients of θ along the orbits of $\tilde{A}$. Moreover, for any C^∞ function θ there is a well behaved splitting $\theta = \theta^\iota + \theta^*$, where θ^ι can be view as a projection to the space of twisted coboundaries and θ^* as the error. A similar result was proved for ergodic toral automorphisms in [DKt10] and used for establishing C^∞ local rigidity of some partially hyperbolic $\mathbb{Z}^k$ actions. We prove the result for hyperbolic case to keep our exposition self-contained and get a better constant $\sigma(\lambda)$.

Lemma 7.1. *For a function $\theta : \mathbb{T}^N \to \mathbb{C}$ in $\mathcal{H}^a(\mathbb{T}^N)$ and $\lambda \in \mathbb{C}\backslash\{0,1\}$ we define*

$$D_\theta(n) = \sum_{i=-\infty}^{\infty} \lambda^{-(i+1)} \widehat{\theta}_{\tilde{A}^i n}.$$

Suppose $a \geq \sigma(\lambda) = \frac{|\log|\lambda||}{\log \rho} + 1$, where $\rho > 1$ is the expansion rate of $\tilde{A}$ from (7.1). Then

(i) *The sum $D_\theta(n)$ converges absolutely for any $n \neq 0$; moreover the function*

$$\theta^* \stackrel{def}{=} \sum_{n \in M} D_\theta(n) e^{2\pi \mathbf{i}\, n \cdot x},$$

where M is from (7.2), is in $\mathcal{H}^a(\mathbb{T}^N)$ with the estimate $\|\theta^*\|_a \leq K_{a,\lambda}\|\theta\|_a$.

(ii) *If $D_\theta(n) = 0$ for any $n \neq 0$, then the equation (7.5) has a solution $\omega \in \mathcal{H}^a(\mathbb{T}^N)$ with the estimate*

$$\|\omega\|_a \leq K_{r,\lambda}\|\theta\|_a.$$

(iii) *If the equation (7.5) has a solution $\omega \in \mathcal{H}^{\sigma(\lambda)}(\mathbb{T}^N)$, then $D_\theta(n) = 0$ for any $n \neq 0$.*

(iv) *For $\theta^\iota \stackrel{def}{=} \theta - \theta^*$ the equation:*

$$\lambda\omega - \omega \circ A = \theta^\iota$$

has a solution $\omega \in \mathcal{H}^a(\mathbb{T}^N)$ with the estimate $\|\omega\|_a \leq K_{r,\lambda}\|\theta\|_a$.

Remark 7.2. We emphasize that the existence of θ^* requires a high regularity of θ. In fact, for any $b \leq \sigma(\lambda)$, we have to estimate it as $\|\theta^*\|_b \leq K_\lambda\|\theta\|_{\sigma(\lambda)}$.

Proof. We define

$$D_\theta(n)_+ = \sum_{i \geq 1} \lambda^{-(i+1)} \widehat{\theta}_{\tilde{A}^i n} \quad \text{and} \quad D_\theta(n)_- = -\sum_{i \leq 0} \lambda^{-(i+1)} \widehat{\theta}_{\tilde{A}^i n}.$$

(i). Let $n \in M$. The inequality $\|\pi_{\tilde{E}^s}(n)\| \geq \frac{1}{2}\|n\|$ we obtain

$$|D_\theta(n)_-| \;\leq\; \|\theta\|_a \sum_{i \leq 0} |\lambda|^{-(i+1)} \, \|\tilde{A}^i n\|^{-a} \leq \; \|\theta\|_a \sum_{i \leq 0} |\lambda|^{-(i+1)} \, \|\pi_{\tilde{E}^s}(\tilde{A}^i n)\|^{-a}$$

$$\leq \; \|\theta\|_a C^{-a} \sum_{i \leq 0} |\lambda|^{-(i+1)} \rho^{ia} \, \|\pi_{\tilde{E}^s}(n)\|^{-a} \stackrel{(1)}{\leq} \; K_{a,\lambda}\|\theta\|_a \, \|n\|^{-a}. \tag{7.6}$$

Here in (1) convergence is guaranteed by $a > \frac{|\log|\lambda||}{\log \rho}$. The sum $D_\theta(n)_+$ can be estimated similarly using the inequality $\|\pi_{\tilde{E}^u}(\tilde{A}n)\| \geq \frac{1}{2}\|\tilde{A}n\|$. Hence we get

$$\|\theta^*\|_a \leq K_{a,\lambda}\|\theta\|_a.$$

For any $z \in \mathbb{Z}^N$ and $k \in \mathbb{Z}$, we see that

$$D_\theta(\tilde{A}^k z) = \lambda^k D_\theta(z). \tag{7.7}$$

This shows that $D_\theta(n)$ converges absolutely for any $n \neq 0$.

(ii) In the dual space the equation $\lambda\omega - \omega \circ A = \theta$ has he form

$$\lambda\widehat{\omega}_n - \widehat{\omega}_{\tilde{A}n} = \widehat{\theta}_n, \qquad \forall\, n \in \mathbb{Z}^N.$$

For $n = 0$, we let $\widehat{\omega}_0 = \frac{\widehat{\theta}_0}{\lambda - 1}$. For any $n \neq 0$, let $\widehat{\omega}_n = D_\theta(n)_-$. Then $\omega = \sum_{n\in\mathbb{Z}^N} \widehat{\omega}_n e^{2\pi \mathbf{i}}$ is a formal solution. Next, we obtain its Sobolev estimates. If $\|\pi_{\tilde{E}^s}(n)\| \geq \frac{1}{2}\|n\|$, then from (7.6) we have

$$|\widehat{\omega}_n| \cdot \|n\|^a \leq K_{a,\lambda}. \tag{7.8}$$

If $\|\pi_{\tilde{E}^u}(\tilde{A}n)\| \geq \frac{1}{2}\|\tilde{A}n\|$, then the assumption $D_\theta(n) = 0$ implies that $\widehat{\omega}_n = D_\theta(n)_+$. The arguments in **(i)** show that (7.8) still holds.

(iii) By **(i)** and (7.7) we have: for any $n \neq 0$

$$D_\theta(n) = D_{\lambda\omega - \omega\circ A}(n) = \lambda D_\omega(n) - D_\omega(\tilde{A}n) = \lambda D_\omega(n) - \lambda D_\omega(n) = 0.$$

(iv) It is clear that $D_{\theta^\iota}(n) = D_{\theta - \theta^*}(n) = D_\theta(n) - D_{\theta^*}(n) = 0$ for any $n \neq 0$. Then the result follows from **(ii)**. $\square$

Now we extend Lemma 7.1 to the vector valued case. We consider the equation

$$A_i\omega - \omega \circ A = \theta$$

with the twist given by the restriction $A_i = A|E^i$, where E^i, $i = 1, \dots, L$, is a subspace of the splitting (3.1). We note that any eigenvalue λ of A_i satisfies $|\lambda| = \rho_i$.

Lemma 7.3. *Let $\rho > 1$ be the expansion rate for $\tilde{A}$ from* (7.1) *and let*

$$\sigma = \max_{i=1,\dots,L} \left(\frac{|\log \rho_i|}{\log \rho} + 1 \right) N + N + 2. \tag{7.9}$$

Then for any $i = 1, \dots, L$ and any C^∞ map $\theta : \mathbb{T}^N \to \mathbb{C}^{N_i}$, there is a splitting of θ

$$\theta = \theta^\iota + \theta^*$$

such that the equation:

$$A_i\omega - \omega \circ A = \theta^\iota \tag{7.10}$$

has a C^∞ solution ω with estimates

$$\|\omega\|_{C^r} \leq K_r \|\theta\|_{C^{r+\sigma}}, \qquad \forall\, r \geq 0;$$

and $\theta^* : \mathbb{T}^N \to \mathbb{C}^{N_i}$ is an excellent C^∞ map so that for any $r \geq 0$

$$\|\theta^*\|_{C^r} \leq K_r \|\theta\|_{C^{r+\sigma}} \qquad \textit{and} \qquad \|\theta^*\|_r \leq K_r \|\theta\|_{r+\sigma-2-N}.$$

Proof. If A_i is semisimple, then the conclusion follows directly from Lemma 7.1 as the equation (7.10) splits into finitely many equations of the type

$$\lambda_j\omega_j - \omega_j \circ A = (\theta_j)^\iota$$

where θ_j is a coordinate function of θ and λ_j is the corresponding eigenvalue of A_i.

If A_i is not semisimple, we choose a basis in which A_i is in its Jordan normal form with some nontrivial Jordan blocks. We note that the excellency of maps is preserved under the change of basis. Let $J = (J_{l,j})$ to be an $m \times m$ Jordan block of A_i corresponding to an eigenvalue λ with $|\lambda| = \rho_i$, that is, $J_{l,l} = \lambda$ for all $1 \le l \le m$ and $\lambda_{l,l+1} = 1$ for all $1 \le l \le m-1$. Then equation (7.10) splits into equations of the form

$$J\Omega - \Omega \circ A = \Theta^\iota, \tag{7.11}$$

corresponding to the Jordan blocks J. Each equation (7.11) further splits into the following m equations:

$$\lambda\Omega_j - \Omega_j \circ A + \Omega_{j+1} = (\Theta^\iota)_j, \qquad \text{and}$$
$$\lambda\Omega_m - \Omega_m \circ A = (\Theta^\iota)_m = (\Theta_m)^\iota,$$

$1 \le j \le m-1$. For the m-th equation, Lemma 7.1 gives the splitting

$$\Theta_m = \lambda\Omega_m - \Omega_m \circ A + (\Theta^*)_m$$

where Ω_m, $(\Theta^*)_m = (\Theta_m)^*$, and $(\Theta^\iota)_m = \lambda\Omega_m - \Omega_m \circ A$ are C^∞ functions satisfying the estimates:

$$\max\{\|(\Theta^*)_m\|_r,\ \|\Omega_m\|_r\} \le K_{r,m}\|\Theta\|_{r+\sigma(\rho_i)}, \qquad \forall\, r \ge 0$$

and Θ^*_m is excellent.

Now we proceed by induction. Fix $1 \le k \le m-1$ and assume that for all $k+1 \le j \le m$ we already have the splitting

$$\Theta_j = \lambda\Omega_j - \Omega_j \circ A + \Omega_{j+1} + (\Theta^*)_j$$

where Ω_j, Θ^*_j, and $(\Theta^\iota)_j = \lambda\Omega_j - \Omega_j \circ A + \Omega_{j+1}$ are C^∞ functions satisfying the estimates:

$$\max\{\|\Omega_j\|_r,\ \|(\Theta^*)_j\|_r\} \le K_{r,j}\|\Theta\|_{r+(m-j+1)\sigma(\rho_i)}, \qquad \forall\, r \ge 0 \tag{7.12}$$

and $(\Theta^*)_j$ is excellent. By Lemma 7.1 we obtain the splitting

$$\Theta_k - \Omega_{k+1} = \lambda\Omega_k - \Omega_k \circ A + (\Theta_k - \Omega_{k+1})^*$$

where Ω_k, $(\Theta^*)_k = (\Theta_k - \Omega_{k+1})^*$, and $(\Theta^\iota)_k = \lambda\Omega_k - \Omega_k \circ A + \Omega_{k+1}$ are C^∞ functions satisfying the estimates following from (7.12):

$$\max\{\|\Omega_k\|_r,\ \|(\Theta^*)_k\|_r\} \le K_r\|\Theta_k - \Omega_{k+1}\|_{r+\sigma(\rho_i)} \le K_{r,k}\|\Theta\|_{r+(m-k+1)\sigma(\rho_i)}, \quad \forall\, r \ge 0$$

and $(\Theta^*)_k$ is excellent. Let Ω, Θ^ι and Θ^* be maps with coordinate functions Ω_j, $(\Theta^\iota)_j$ and $(\Theta^*)_j$, $1 \le j \le m$ respectively. Hence we show that there is a splitting of Θ

$$\Theta = \Theta^\iota + \Theta^*$$

such that the equation (7.11) has a C^∞ solution Ω with estimates.

$$\max\{\|\Theta^*\|_r,\ \|\Omega\|_r\} \le K_r\|\Theta\|_{r+m\sigma(\rho_i)}, \quad \forall\, r \ge 0$$

This can be repeated for all corresponding blocks of A. Since the maximal size of a Jordan block is bounded by N, we obtain estimates for the $\|\cdot\|_r$ norms of ω and θ^*. This implies estimates for the $\|\cdot\|_{C^r}$ norms as well by (7.3). $\square$

7.3. **Main result on the linearized equation.** The next theorem is our main result on solving the linearized equation. It plays the crucial role in the inductive step of the iterative process, Proposition 8.3. The goal of the inductive step is, given a C^1 conjugacy H between A and its perturbation f, to construct a smaller perturbation $\tilde{f}$ which is smoothly conjugate to f by $\tilde{H}$. The conjugacy $\tilde{H}$ is constructed in the form $\tilde{H} = I - \omega$, where ω is a C^∞ approximate solution of the linearized equation given by Theorem 7.4. The C^1 conjugacy H is upgraded to C^{1+a} by Theorem 1.1. It yields an *approximate* C^{1+a} solution $\mathfrak{h} = H - I$ of the linearized equation (7.13). This necessitates the introduction of the error term Ψ in the assumption of the theorem.

Theorem 7.4. *Let A be weakly irreducible hyperbolic automorphism of $\mathbb{T}^N$. Suppose that*

$$A\mathfrak{h} - \mathfrak{h} \circ A = \mathcal{R} + \Psi, \tag{7.13}$$

where maps $\mathfrak{h}, \Psi : \mathbb{T}^N \to \mathbb{R}^N$ are C^{1+a} and $\mathcal{R} : \mathbb{T}^N \to \mathbb{R}^N$ is C^∞.

Then there exist C^∞ maps $\omega, \Phi : \mathbb{T}^N \to \mathbb{R}^N$ satisfying the equation

$$\mathcal{R} = A\omega - \omega \circ A + \Phi \tag{7.14}$$

and the estimates

$$\|\omega\|_{C^r} \le K_r \|\mathcal{R}\|_{C^{r+\sigma}}$$

$$\|\Phi\|_{C^0} \le K_{l,a} (\|\Psi\|_{C^{1+a}})^{\frac{l-2-N}{l+N}} (\|\mathcal{R}\|_{C^{l+\sigma}})^{\frac{2N+2}{l+N}}$$

for any $r \ge 0$ and $l > N + 2$, where σ is given by (7.9).

For traditional KAM iteration scheme, the convergence requires the error Φ in solving the twisted coboundary (7.14) to be small compared with $\mathcal{R}$. This is established by showing that Φ is tame with respect to Ψ, which is almost quadratically small with respect to $\mathcal{R}$. Tameness means that the C^r norm of Φ can be bounded by the C^{r+p} norm of Ψ, where r is arbitrarily large while p is a constant.

One difficulty in our setting is that the estimate of Φ depends on Ψ and $\mathcal{R}$ rather than on Ψ only. This results in technical issues in proving convergence of the iterative procedure, and so the traditional KAM scheme fails to work. We resolve this issue by introducing a parameter l when estimating $\|\Phi\|_{C^0}$. If the parameters are well chosen, the constructed approximation behaves as if it were tame.

The main difficulty in estimating Φ in our setting is that low regularity of $\mathfrak{h}$ yields smallness of Ψ only in $C^{1+\text{Hölder}}$ norm, see Lemma 8.4 and equation (8.12). This does not allow us to directly estimate orbit sums of Fourier coefficients and split R into a smooth coboundary $R^\iota = A\omega - \omega \circ A$ and an error term $R^* = \Phi$, see Remark 7.2. To overcome this problem we use the splitting $\mathbb{R}^N = \oplus E^i$ to decompose the equation (7.13) and then differentiate i^{th} component along directions in E^i. This allows us to "balance" the twist (up to a polynomial growth of Jordan blocks) and analyze the differentiated equation using Hölder regularity. This is done in the following Lemma 7.5. After that, we establish Lemma 7.6 to relate Fourier coefficients of a function and its directional derivatives. We then complete the proof of Theorem 7.4 in Section 7.5.

Now we begin the analysis of the differentiated equation (7.13). For any $1 \le i \le L$ and any unit vector $u_0 \in E^i$, we consider unit vectors u_k and scalars a_k, $k \in \mathbb{Z}$, given by

$$u_k = \frac{A_i^k u_0}{\|A_i^k u_0\|} \quad \text{and} \quad a_k = \|A_i u_k\| = \frac{\|A_i^{k+1} u_0\|}{\|A_i^k u_0\|} \quad \text{so that} \quad A_i u_k = a_k u_{k+1}. \tag{7.15}$$

We define a sequence of matrices $P_k \in GL(N_i, \mathbb{R})$ which commute with A_i and satisfy the recursive equation

$$P_{k+1} = a_k A_i^{-1} P_k. \tag{7.16}$$

Specifically, we set

$$P_0 = \mathrm{Id} \quad \text{and} \quad P_k = \begin{cases} a_0 \cdots a_{k-1} A_i^{-k} = \|A_i^k u_0\|\, A_i^{-k}, & k > 0, \\ (a_{-1} \cdots a_{-k})^{-1} A_i^k = \|A_i^{-k} u_0\|\, A_i^k, & k < 0. \end{cases} \tag{7.17}$$

Lemma 7.5. *Let $\varphi_k : \mathbb{T}^N \to \mathbb{R}^{N_i}$ be a sequence of maps in $\mathcal{H}^a(\mathbb{T}^N)$, $a > 0$, satisfying $\|\varphi_k\|_a \le \mathfrak{b}$ for all $k \in \mathbb{Z}$, let $P_k \in GL(N_i, \mathbb{R})$ be as in (7.17), and let*

$$S(n) = \sum_{k \in \mathbb{Z}} P_k\, (\widehat{\varphi_k})_{\tilde{A}^k n}.$$

- **(i)** *For any $n \in M$ the sum $S(n)$ converges absolutely in $\mathbb{C}^{N_i}$ with the estimate $\|S(n)\| \le K_a \mathfrak{b}\, \|n\|^{-a}$.*
- **(ii)** *If $\mathfrak{h}_k : \mathbb{T}^N \to \mathbb{R}^{N_i}$ is another sequence in $\mathcal{H}^a(\mathbb{T}^N)$ so that for all $k \in \mathbb{Z}$ we have $\|\mathfrak{h}_k\|_a \le \mathfrak{c}$ and*

$$A_i \mathfrak{h}_k - a_k \mathfrak{h}_{k+1} \circ A = \varphi_k, \tag{7.18}$$

then $S(n) = 0$ for every $n \in M$.

Proof. **(i)**. Since all eigenvalues of A_i have the same modulus ρ_i, we have (3.3), and so there exists a constant C such that all P_k satisfy the polynomial estimate

$$\|P_k\| \le \|A_i^k\| \cdot \|A_i^{-k}\| \le C(|k| + 1)^{2N} =: p(|k|), \quad \text{for all } k \in \mathbb{Z}. \tag{7.19}$$

Let $n \in M$. We write $\varphi_k = (\varphi_{k,1}, \cdots, \varphi_{k,N_i})$ and set

$$S(n)_+ = \sum_{k \ge 1} P_k\, (\widehat{\varphi_k})_{\tilde{A}^k n} \quad \text{and} \quad S(n)_- = \sum_{k \le 0} P_k\, (\widehat{\varphi_k})_{\tilde{A}^k n}.$$

Using the assumption $\|\varphi_k\|_a \le \mathfrak{b}$, estimates (7.19) and (7.1), and the inequality $\|\pi_{\tilde{E}^s}(n)\| \ge \frac{1}{2}\|n\|$ we obtain

$$\begin{aligned} \|S(n)_-\| \;&\le\; \sum_{k \le 0} \|P_k\| \max_{1 \le j \le m} |(\widehat{\varphi_{k,j}})_{\tilde{A}^k n}| \;\le\; \sum_{k \le 0} \|\varphi_k\|_a\, \|P_k\|\, \|\tilde{A}^k n\|^{-a} \\ &\le\; \mathfrak{b} \sum_{k \le 0} p(|k|)\, \|\pi_{\tilde{E}^s}(\tilde{A}^k n)\|^{-a} \;\le\; \mathfrak{b} C^{-a} \sum_{k \le 0} p(|k|)\, \rho^{ka}\, \|\pi_{\tilde{E}^s}(n)\|^{-a} \\ &\le\; K_a \mathfrak{b}\, \|n\|^{-a}. \end{aligned}$$

The sum $S(n)_+$ can be estimated similarly using the inequality $\|\pi_{\tilde{E}^u}(\tilde{A}n)\| \ge \frac{1}{2}\|\tilde{A}n\|$.

(ii) Let $n \in M$. From the equation (7.18) we obtain that for any $k \in \mathbb{Z}$

$$P_k\, \varphi_k \circ A^k = P_k A_i\, \mathfrak{h}_k \circ A^k - a_k P_k\, \mathfrak{h}_{k+1} \circ A^{k+1}$$

Summing from $-m$ to j and observing that the sum on the right is telescoping as $a_k P_k = A_i P_{k+1} = P_{k+1} A_i$ by the choice of P_k in (7.16), we obtain

$$\sum_{k=-m}^{j} P_k\, \varphi_k \circ A^k = A_i P_{-m} \mathfrak{h}_{-m} \circ A^{-m} - a_j P_j\, \mathfrak{h}_{j+1} \circ A^{j+1}.$$

Taking Fourier coefficients and noting that $(\widehat{\theta \circ A^k})_n = \widehat{\theta}_{\tilde{A}^k n}$ we obtain

$$\sum_{k=-m}^{j} P_k (\widehat{\varphi_k})_{\tilde{A}^k n} = A_i P_{-m} (\widehat{\mathfrak{h}_{-m}})_{\tilde{A}^{-m} n} - a_j P_j (\widehat{\mathfrak{h}_{j+1}})_{\tilde{A}^{j+1} n}.$$

Since the series $\sum_{k\in\mathbb{Z}} P_k (\widehat{\mathfrak{h}_k})_{\tilde{A}^k n}$ converges by part (i), we have $P_k (\widehat{\mathfrak{h}_k})_{\tilde{A}^k n} \to 0$ as $k \to \pm\infty$ and hence, as a_k are bounded,

$$\begin{aligned} &a_j P_j (\widehat{\mathfrak{h}_{j+1}})_{\tilde{A}^{j+1} n} \to 0, \qquad \text{as } j \to \infty; \quad \text{and} \\ &A_i P_m (\widehat{\mathfrak{h}_m})_{\tilde{A}^m n} \to 0, \qquad \text{as } m \to -\infty. \end{aligned}$$

We conclude that $S(n) = 0$. □

7.4. **Directional derivatives.** In this section we establish some estimates for Fourier coefficients of a C^1 function $\theta : \mathbb{T}^N \to \mathbb{R}$ via Fourier coefficients of its directional derivatives along a subspace E^i of the splitting (3.1). This relies on weak irreducibility of A.

For any $v \in \mathbb{R}^N$ with $\|v\| = 1$, we denote the directional derivative of θ along v by θ_v.

Lemma 7.6. *Let A be a weakly irreducible integer matrix and let $v_{i,j}$, $j = 1, \dots, N_i$, be an orthonormal basis of a subspace E^i from (3.1). Then there exists a constant $K = K(A)$ such that for any $i = 1, \dots, L$ and any C^1 function $\theta : \mathbb{T}^N \to \mathbb{R}$,*

$$|\hat{\theta}_n| \,\le\, K \sum_{j=1}^{N_i} |(\widehat{\theta_{v_{i,j}}})_n| \cdot \|n\|^N \quad \textit{for all } n \in \mathbb{Z}^N \backslash 0.$$

Proof. We denote by $\|.\|$ the standard Euclidean norm in $\mathbb{R}^N$. Since θ is C^1, we have

$$2\pi \mathrm{i} (n \cdot v_{i,j}) \hat{\theta}_n = (\widehat{\theta_{v_{i,j}}})_n, \qquad 1 \le j \le N_i.$$

Adding over j we obtain that for any $n \in \mathbb{Z}^N \backslash 0$ we have

$$|\hat{\theta}_n| \,=\, \frac{\sum_{j=1}^{N_i} |(\widehat{\theta_{v_{i,j}}})_n|}{2\pi \sum_{j=1}^{N_i} |n \cdot v_{i,j}|} \,\le\, \frac{\sum_{j=1}^{N_i} |(\widehat{\theta_{v_{i,j}}})_n|}{2\pi \|\pi_{E^i} n\|},$$

since for an orthonormal basis $v_{i,j}$ we have $\sum_{j=1}^{N_i} |n \cdot v_{i,j}| \ge \|\pi_{E^i} n\|$. Since $\|\pi_{E^i} n\| = d(n, (E^i)^\perp)$, to complete the proof it remains to show that $d(n, (E^i)^\perp) \ge K' \|n\|^{-N}$.

Since A is weakly irreducible, so is the transpose A^τ. This follows from Lemma 3.3 which gives an equivalent condition for weak irreducibility in terms of the characteristic polynomial. We denote the splitting (3.1) for A^τ by $\mathbb{R}^N = E^1_\tau \oplus \cdots \oplus E^L_\tau$ and similarly let $\hat{E}^i_\tau = \oplus_{j\neq i} E^i_\tau$. Then we obtain $(E^i)^\perp = \hat{E}^i_\tau$. Indeed, the polynomial

$$p_i(x) = \prod_{|\lambda|=\rho_i} (x-\lambda)^N,$$

where the product is over all eigenvalues of A of modulus ρ_i, is real and

$$(E^i)^\perp = (\ker p_i(A))^\perp = range(p_i(A)^\tau) = range(p_i(A^\tau)) = \hat{E}^i_\tau,$$

since $p_i(A^\tau)$ is invertible on $\hat{E}^i_\tau$. Now the desired inequality

$$d(n,(E^i)^\perp) = d(n, \hat{E}^i_\tau) \geq K' \|n\|^{-N}$$

follows from Katznelson's Lemma below. We apply it to A^τ with the invariant splitting $\mathbb{R}^N = \hat{E}^i_\tau \oplus E^i_\tau$ and note that $\hat{E}^i_\tau \cap \mathbb{Z}^N = \{0\}$ by weak irreducibility of A^τ. □

Lemma 7.7 (Katznelson's Lemma). *Let A be an $N\times N$ integer matrix. Assume that $\mathbb{R}^N$ splits as $\mathbb{R}^N = V_1 \bigoplus V_2$ with V_1 and V_2 invariant under A and such that $A|_{V_1}$ and $A|_{V_2}$ have no common eigenvalues. If $V_1 \cap \mathbb{Z}^N = \{0\}$, then there exists a constant K such that*

$$d(n, V_1) \geq K\|n\|^{-N} \quad \textit{for all } 0 \neq n \in \mathbb{Z}^N,$$

where $\|v\|$ denotes Euclidean norm and d is Euclidean distance.

See e.g. [DKt10, Lemma 4.1] for a proof.

7.5. **Proof of Theorem 7.4.** Using the splitting $\mathbb{R}^N = \oplus E^i$ we decompose (7.13) into equations

$$A_i\mathfrak{h}_i - \mathfrak{h}_i \circ A = \mathcal{R}_i + \Psi_i, \qquad i = 1,\ldots,L \tag{7.20}$$

where $\mathfrak{h}_i$, $\mathcal{R}_i$ and Ψ_i are coordinate maps in the of $\mathfrak{h}$, $\mathcal{R}$ and Ψ respectively.

By Lemma 7.3 there is an excellent C^∞ map $\mathcal{R}^*_i$ with estimates

$$\|\mathcal{R}^*_i\|_{C^r} \leq K_r \|\mathcal{R}_i\|_{C^{r+\sigma}}, \quad \|\mathcal{R}^*_i\|_r \leq K_r \|\mathcal{R}_i\|_{r+\sigma-N-2} \tag{7.21}$$

for any $r \geq 0$, such that the equation:

$$A_i\omega_i - \omega_i \circ A = \mathcal{R}_i + \mathcal{R}^*_i \tag{7.22}$$

has a C^∞ solution ω_i with estimates

$$\|\omega_i\|_{C^r} \leq K_r \|\mathcal{R}_i\|_{C^{r+\sigma}}, \qquad \forall\, r \geq 0.$$

Let ω be the map with coordinate maps ω_i.

We obtain from (7.20) and (7.22) that C^{1+a} maps $\mathfrak{p}_i = \mathfrak{h}_i - \omega_i$ and $\Lambda_i = -\mathcal{R}^*_i + \Psi_i$ satisfy

$$A_i\mathfrak{p}_i - \mathfrak{p}_i \circ A = \Lambda_i.$$

We fix $1 \le i \le L$ and an orthonormal basis $v_{i,j}$ of E^i. We fix $1 \le j \le N_i$ and, as in (7.15), consider unit vectors $u_0 = v_{i,j}$ and $u_k = \frac{A^k u_0}{\|A^k u_0\|}$, and let $a_k = \|Au_k\|$, $k \in \mathbb{Z}$. Taking the derivative of the previous equation in the direction of u_k we obtain equations

$$A_i(\mathfrak{p}_i)_{u_k} - a_k(\mathfrak{p}_i)_{u_{k+1}} \circ A = (\Lambda_i)_{u_k}, \qquad \forall\, k \in \mathbb{Z}.$$

We note that for any $k \in \mathbb{Z}$ the maps $(\mathfrak{p}_i)_{u_k}$ and $(\Lambda_i)_{u_k}$ are in C^a and hence in $\mathcal{H}^a$, as we recall that for any function g by (7.3) we have

$$\|g_{u_k}\|_a \le K\|g_{u_k}\|_{C^a} \le K_1\|g\|_{C^{1+a}}. \tag{7.23}$$

Now we use (ii) of Lemma 7.5 with $\mathfrak{h}_k = (\mathfrak{p}_i)_{u_k}$, $\varphi_k = (\Lambda_i)_{u_k}$, and P_k is as defined in (7.17) to obtain that for any $n \in \mathcal{M}$

$$\sum_{k\in\mathbb{Z}} P_k \widehat{((\Psi_i)_{u_k})}_{\tilde{A}^k n} - \sum_{k\in\mathbb{Z}} P_k \widehat{((\mathcal{R}_i^*)_{u_k})}_{\tilde{A}^k n} = \sum_{k\in\mathbb{Z}} P_k \widehat{((\Lambda_i)_{u_k})}_{\tilde{A}^k n} = 0.$$

Since $(\mathcal{R}_i^*)_{u_k}$ is excellent, for each $k \in \mathbb{Z}$ we have

$$\sum_{k\in\mathbb{Z}} P_k \widehat{((\Psi_i)_{u_k})}_{\tilde{A}^k n} = \sum_{k\in\mathbb{Z}} P_k \widehat{((\mathcal{R}_i^*)_{u_k})}_{\tilde{A}^k n} = \widehat{((\mathcal{R}_i^*)_{u_0})}_n$$

for any $n \in M$, which gives

$$|\widehat{((\mathcal{R}_i^*)_{u_0})}_n| \overset{(1)}{\le} K_a \max_{k\in\mathbb{Z}}\{\|(\Psi_i)_{u_k}\|_a\}\|n\|^{-a} \overset{(2)}{\le} K_{a,1}\|\Psi_i\|_{C^{1+a}}\|n\|^{-a}.$$

Here in (1) we use (i) of Lemma 7.5 and in (2) we use (7.23).

We conclude that for any $v_{i,j}$, $1 \le j \le N_i$, we have

$$|\widehat{((\mathcal{R}_i^*)_{v_{i,j}})}_n| \le K_a\|\Psi_i\|_{C^{1+a}}\|n\|^{-a}, \qquad \forall\, n \in M. \tag{7.24}$$

Finally, using Lemma 7.6 and (7.24), we obtain that for any $n \in M$

$$\begin{aligned} |\widehat{(\mathcal{R}_i^*)}_n| &\le K \sum_{j=1}^{N_i} |\widehat{((\mathcal{R}_i^*)_{v_{i,j}})}_n|\|n\|^N \le K_a\|\Psi_i\|_{C^{1+a}}\|n\|^{N-a} \\ &\le K_a\|\Psi_i\|_{C^{1+a}}\|n\|^N. \end{aligned} \tag{7.25}$$

Now for any $r > N+2$ and any $n \in M$ we can estimate splitting the exponent of the first term as α and $1-\alpha$ in the way to get the total the exponent of $\|n\|$ be zero

$$\begin{aligned} |\widehat{(\mathcal{R}_i^*)}_n|\|n\|^{N+2} &= |\widehat{(\mathcal{R}_i^*)}_n|^{\frac{l-2-N}{l+N}} |\widehat{(\mathcal{R}_i^*)}_n|^{\frac{2N+2}{l+N}} \|n\|^{N+2} \\ &\overset{(1)}{\le} \left(K_a\|\Psi_i\|_{C^{1+a}}\|n\|^N\right)^{\frac{l-2-N}{l+N}} \left(\|n\|^{-l}\|\mathcal{R}_i^*\|_l\right)^{\frac{2N+2}{l+N}} \|n\|^{N+2} \\ &= K_a^{\frac{l-2-N}{l+N}} (\|\Psi_i\|_{C^{1+a}})^{\frac{l-2-N}{l+N}} (\|\mathcal{R}_i^*\|_l)^{\frac{2N+2}{l+N}} \\ &\overset{(2)}{\le} K_{l,a}(\|\Psi_i\|_{C^{1+a}})^{\frac{l-2-N}{l+N}} (\|\mathcal{R}_i^*\|_{C^l})^{\frac{2N+2}{l+N}} \\ &\overset{(3)}{\le} K_{l,a}(\|\Psi_i\|_{C^{1+a}})^{\frac{l-2-N}{l+N}} (\|\mathcal{R}_i\|_{C^{l+\sigma}})^{\frac{2N+2}{l+N}}. \end{aligned}$$

Here in (1) we use that $\mathcal{R}_i^*$ is C^∞ and (7.25); in (2) we use (7.3); in (3) we use (7.21). Then by (7.3) we get

$$\|\mathcal{R}_i^*\|_{C^0} \le C\|\mathcal{R}_i^*\|_{N+2} \le K_{l,a}(\|\Psi_i\|_{C^{1+a}})^{\frac{l-2-N}{l+N}}(\|\mathcal{R}_i\|_{C^{l+\sigma}})^{\frac{2N+2}{l+N}}.$$

Finally, we denote by Φ the map with coordinate maps $\mathcal{R}_i^*$.

8. Proof of Theorem 1.3

In this section we complete the proof of Theorem 1.3 using an iterative process. The main part is the inductive step given by Proposition 8.3. We start with a sufficiently small perturbation f_n of A which is C^1 conjugate to A. We construct a smaller perturbation f_{n+1} which is smoothly conjugate to f_n. The conjugacy $\tilde{H}_{n+1}$ between f_n and f_{n+1} is obtained using Theorem 7.4. Then the iterative process is set up so that f_n converges to A and $\tilde{H}_1 \circ \cdots \circ \tilde{H}_{n+1}$ converge in sufficiently high regularity.

8.1. **Iterative step and error estimate.**

We recall the following results, which will be used the proof of Proposition 8.3.

Lemma 8.1. [dlLO98, Propositions 5.5] *For any $r \ge 1$ there exists a constant M_r such that for any $h, g \in C^r(\mathcal{M})$,*

$$\|h \circ g\|_{C^r} \le M_r\left(1 + \|g\|_{C^1}^{r-1}\right)(\|h\|_{C^1}\|g\|_{C^r} + \|h\|_{C^r}\|g\|_{C^1}) + \|h\|_{C^0}.$$

Lemma 8.2. [La93, Lemma AII.26.] *There is $d > 0$ and such that for any $h \in C^r(\mathcal{M})$, if $\|h - I\|_{C^1} \le d$ then h^{-1} exists with the estimate $\|h^{-1} - I\|_{C^r} \le K_r\|h - I\|_{C^r}$.*

Proposition 8.3. *Let A be a weakly irreducible Anosov automorphism of $\mathbb{T}^N$. Let $\beta = \frac{\beta_0}{2}$, where β_0 is as in Theorem 1.1. There exists $0 < c < \frac{1}{2}$ such that for any C^∞ perturbation f_n of A satisfying*

$$\|f_n - A\|_{C^{\sigma+2}} < c, \text{ where } \sigma \text{ is from Lemma 7.3,}$$

and the conjugacy equation

$$H_n \circ f_n = A \circ H_n \text{ with a function } H_n \in C^1(\mathbb{T}^N) \text{ with } \|H_n - I\|_{C^0} \le c \tag{8.1}$$

the following holds. There exists $\omega_{n+1} \in C^\infty(\mathbb{T}^N)$ so that the functions

$$\tilde{H}_{n+1} = I - \omega_{n+1}, \quad H_{n+1} = H_n \circ \tilde{H}_{n+1}, \quad f_{n+1} = \tilde{H}_{n+1}^{-1} \circ f_n \circ \tilde{H}_{n+1} \tag{8.2}$$

satisfy the new conjugacy equation

$$H_{n+1} \circ f_{n+1} = A \circ H_{n+1},$$

and we have the following estimates.

(i) *For any $r \ge 0$ and any $t > 1$*

$$\|\omega_{n+1}\|_{C^r} \le K_r \min\{t^\sigma\|R_n\|_{C^r}, \|R_n\|_{C^{r+\sigma}}\}. \quad \text{where} \quad R_n = f_n - A.$$

(ii) *For the new error $R_{n+1} = f_{n+1} - A$, we have*

$$\|R_{n+1}\|_{C^0} \le Kt^\sigma \|R_n\|_{C^1} \|R_n\|_{C^0} + K_\ell t^{-\ell} \|R_n\|_{C^\ell}$$
$$+ K_{l,\ell} (t^{-\ell+2} \|R_n\|_{C^\ell} + \|R_n\|_{C^2}^{1+\frac{\beta}{2}})^{\frac{l-2-N}{l+N}} (t^\sigma \|R_n\|_{C^l})^{\frac{2N+2}{l+N}}$$

for any $t > 1$, $\ell \ge 0$ and $l > N+2$; and also for any $r \ge 0$ we have

$$\|R_{n+1}\|_{C^r} \le K_r t^\sigma \|R_n\|_{C^r} + K_r. \tag{8.3}$$

(iii) *For the new conjugacy H_{n+1}, we have*

$$\|H_{n+1} - I\|_{C^0} \le K\|R_n\|_{C^\sigma} + \|H_n - I\|_{C^0} \tag{8.4}$$

Proof. We denote $h_n = H_n - I$ and $R_n = f_n - A$ and, similarly to (6.4), we write the conjugacy equation (8.1) as

$$Ah_n - h_n \circ f_n = R_n$$

We can assume that $c < \delta$, where $\delta = \delta(\beta)$ is from Theorem 1.1, and that $\|H_n - I\|_{C^0} \le c$ yields that H is the conjugacy close to the identity. Then Theorem 1.1 gives the estimate

$$\|h_n\|_{C^{1+\beta}} \le K\|R_n\|_{C^{1+\beta}}. \tag{8.5}$$

We define

$$\Omega_n = Ah_n - h_n \circ A, \quad \text{and} \quad \Theta_n = R_n - \Omega_n = h_n \circ A - h_n \circ f_n. \tag{8.6}$$

Lemma 8.4. $\|\Theta_n\|_{C^{1+\frac{\beta}{2}}} \le K_A \|R_n\|_{C^{1+\beta}}^{1+\frac{\beta}{2}}$.

Proof. We omit index n in the proof of the lemma. We note that

$$\|R\|_{C^{1+\beta}} = \|f - A\|_{C^{1+\beta}} < c < 1.$$

Differentiating at $x \in \mathbb{T}^N$ we get

$$\begin{aligned} D\Theta(x) &\overset{*}{=} Dh(Ax) \circ A - Dh(fx) \circ Df(x) \\ &= Dh(Ax) \circ A - Dh(fx) \circ A + Dh(fx) \circ (A - Df(x)), \end{aligned}$$

and hence

$$\begin{aligned} \|D\Theta\|_{C^0} &\le \|A\| \, \|Dh(Ax) - Dh(fx)\|_{C^0} + \|Dh(fx) \circ DR(x)\|_{C^0} \\ &\le \|A\| \, \|Dh\|_{C^\beta} \|R\|_{C^0}^\beta + \|Dh\|_{C^0} \|DR\|_{C^0} \\ &\le \|A\| \, \|h\|_{C^{1+\beta}} \|R\|_{C^0}^\beta + \|h\|_{C^1} \|R\|_{C^1}. \end{aligned}$$

Since we also have $\|\Theta\|_{C^0} \le \|h\|_{C^1} \|R\|_{C^0}$, we conclude using (8.5) and $\|R\|_{C^{1+\beta}} < 1$ that

$$\|\Theta\|_{C^1} \le \|A\| \, \|h\|_{C^{1+\beta}} \|R\|_{C^0}^\beta + \|h\|_{C^1} \|R\|_{C^1} \le K\|R\|_{C^{1+\beta}}^{1+\beta}. \tag{8.7}$$

To estimate the Hölder norm of $D\Theta$, using equation $*$ for any x, $y \in \mathbb{T}^N$ we have

$$\begin{aligned} &D\Theta(x) - D\Theta(y) \\ &= \big(Dh(Ax) - Dh(Ay)\big) \circ A + Dh(fx) \circ \big(Df(y) - Df(x)\big) \\ &+ \big(Dh(fy) - Dh(fx)\big) \circ Df(y), \end{aligned}$$

and hence

$$\begin{aligned} &\|D\Theta(x) - D\Theta(y)\| \\ &\le \|A\|\,\|Dh(Ax) - Dh(Ay)\| + \|Dh(fx)\|\|Df(y) - Df(x)\| \\ &+ \|Dh(fy) - Dh(fx)\|\|Df(y)\| \\ &\le \|A\|\,\|Dh\|_{C^\beta}\|Ax - Ay\|^\beta + \|h\|_{C^1}\|Df\|_{C^\beta}\|y - x\|^\beta \\ &+ \|f\|_{C^1}\|Dh\|_{C^\beta}\|fx - fy\|^\beta \\ &\le \|A\|^{1+\beta}\,\|h\|_{C^{1+\beta}}\|x - y\|^\beta + \|h\|_{C^1}\|f\|_{C^{1+\beta}}\|y - x\|^\beta \\ &+ \|f\|_{C^1}\|h\|_{C^{1+\beta}}\|f\|^\beta_{C^1}\|x - y\|^\beta. \end{aligned}$$

We conclude using (8.5) and $\|f - A\|_{C^{1+\beta}} < 1$ that

$$\|D\Theta\|_{C^{0,\beta}} \le \|A\|^{1+\beta}\,\|h\|_{C^{1+\beta}} + \|h\|_{C^1}\|f\|_{C^{1+\beta}} + \|h\|_{C^{1+\beta}}\|f\|^{1+\beta}_{C^1} \le K\|R\|_{C^{1+\beta}}. \tag{8.8}$$

Therefore

$$\|\Theta\|_{C^{1+\beta}} \le \|\Theta\|_{C^1} + \|D\Theta\|_{C^{0,\beta}} \le 2K\|R\|_{C^{1+\beta}}. \tag{8.9}$$

Finally, we complete the proof of the lemma using an interpolation inequality

$$\|\Theta\|_{C^{1+\frac{\beta}{2}}} \le K\|\Theta\|^{\frac{1}{2}}_{C^1}\|\Theta\|^{\frac{1}{2}}_{C^{1+\beta}} \le K_A\|R\|^{1+\frac{\beta}{2}}_{C^{1+\beta}}. \tag{8.10}$$

□

We recall that there exists a collection of smoothing operators $\mathfrak{s}_t$, $t > 0$, such that for any $s \ge s_1 \ge 0$ and $s_2 \ge 0$, for any $g \in C^s(\mathbb{T}^N)$ the following holds, see [DKt10] and [Ha82]:

$$\|\mathfrak{s}_t g\|_{C^{s+s_2}} \le K_{s,s_2}\, t^{s_2}\, \|g\|_{C^s}, \quad \text{and} \quad \|(I - \mathfrak{s}_t)g\|_{C^{s-s_1}} \le K_{s,s'}\, t^{-s_1}\, \|g\|_{C^s}. \tag{8.11}$$

We write (8.6) as

$$Ah_n - h_n \circ A = \Omega_n = R_n - \Theta_n = [\mathfrak{s}_t R_n] + [(I - \mathfrak{s}_t)R_n - \Theta_n] =: \mathcal{R} + \Psi \tag{8.12}$$

and apply Theorem 7.4 to get the new splitting and obtain the estimates:

$$\mathfrak{s}_t R_n = A\omega_{n+1} - \omega_{n+1} \circ A + \Phi_n \tag{8.13}$$

where ω_{n+1} and Φ_n are C^∞ maps with the estimates:

$$\|\omega_{n+1}\|_{C^r} \le K_r \|\mathfrak{s}_t(R_n)\|_{C^{r+\sigma}} \overset{(a)}{\le} K_r \min\{t^\sigma \|R_n\|_{C^r}, \|R_n\|_{C^{r+\sigma}}\}, \quad \text{and} \tag{8.14}$$

$$\begin{aligned} \|\Phi_n\|_{C^0} &\le K_l(\|(I-\mathfrak{s}_t)R_n - \Theta_n\|_{C^{1+\frac{\beta}{2}}})^{\frac{l-2-N}{l+N}} (\|\mathfrak{s}_t R_n\|_{C^{l+\sigma}})^{\frac{2N+2}{l+N}} \\ &\overset{(b)}{\le} K_l(\|(I-\mathfrak{s}_t)R_n\|_{C^2} + \|R_n\|_{C^2}^{1+\frac{\beta}{2}})^{\frac{l-2-N}{l+N}} (\|\mathfrak{s}_t R_n\|_{C^{l+\sigma}})^{\frac{2N+2}{l+N}} \\ &\overset{(a)}{\le} K_{l,\ell}(t^{-\ell+2}\|R_n\|_{C^\ell} + \|R_n\|_{C^2}^{1+\frac{\beta}{2}})^{\frac{l-2-N}{l+N}} (t^\sigma \|R_n\|_{C^l})^{\frac{2N+2}{l+N}} \end{aligned} \tag{8.15}$$

for any $r,\, \ell \ge 0$ and any $l > N+2$. Here in (a) we use (8.11) and in (b) we use (8.10).

From equation (8.13) we obtain a C^r estimate for Φ_n with $r \ge 0$

$$\begin{aligned} \|\Phi_n\|_{C^r} &= \|A\omega_{n+1} - \omega_{n+1}\circ A - \mathfrak{s}_t R_n\|_{C^r} \\ &\le K\|\omega_{n+1}\|_{C^r} + \|\mathfrak{s}_t R_n\|_{C^r} \overset{(1)}{\le} K_r t^\sigma \|R_n\|_{C^r}. \end{aligned}$$

Here in (1) we use (8.11) and (8.14).

Let $\tilde{H}_{n+1} = I - \omega_{n+1}$. From (8.14) we can assume that $\|\omega_{n+1}\|_{C^1} < \min\{\frac{1}{2}, d\}$ (see Lemma 8.2) if c is sufficiently small. Hence $\tilde{H}_{n+1}$ is invertible. We estimate the new error

$$R_{n+1} = f_{n+1} - A$$

by using

$$\begin{aligned} f_{n+1} = \tilde{H}_{n+1}^{-1}\circ f_n \circ \tilde{H}_{n+1} &\Rightarrow \tilde{H}_{n+1}\circ f_{n+1} = f_n \circ \tilde{H}_{n+1} \\ &\Rightarrow (I-\omega_{n+1})\circ f_{n+1} = f_n\circ \tilde{H}_{n+1} \\ &\Rightarrow f_{n+1} = \omega_{n+1}\circ f_{n+1} + f_n \circ \tilde{H}_{n+1}. \end{aligned}$$

This gives

$$\begin{aligned} R_{n+1} &= \omega_{n+1}\circ f_{n+1} + f_n\circ \tilde{H}_{n+1} - A \\ &= \omega_{n+1}\circ f_{n+1} + (R_n + A)\circ (I-\omega_{n+1}) - A \\ &= \omega_{n+1}\circ f_{n+1} + R_n \circ (I-\omega_{n+1}) - A\circ \omega_{n+1}. \end{aligned}$$

Hence we see that R_{n+1} has three parts:

$$\begin{aligned} R_{n+1} = &\underbrace{\big(\omega_{n+1}\circ f_{n+1} - \omega_{n+1}\circ A\big)}_{\mathcal{E}_1} + \underbrace{\big(R_n\circ(I-\omega_{n+1}) - R_n\big)}_{\mathcal{E}_2} \\ &+ \underbrace{\big(\omega_{n+1}\circ A - A\circ\omega_{n+1} + R_n\big)}_{\mathcal{E}_3}. \end{aligned}$$

We note that

$$\|\mathcal{E}_1\|_{C^0} \le \|\omega_{n+1}\|_{C^1}\|f_{n+1} - A\|_{C^0} \overset{(0)}{\le} \frac{1}{2}\|R_{n+1}\|_{C^0},$$
$$\|\mathcal{E}_2\|_{C^0} \le K\|R_n\|_{C^1}\|\omega_{n+1}\|_{C^0} \overset{(1)}{\le} Kt^{\sigma}\|R_n\|_{C^1}\|R_n\|_{C^0};$$

and

$$\begin{aligned}\|\mathcal{E}_3\|_{C^0} &= \|\Phi_n + (I - \mathfrak{s}_t)R_n\|_{C^0} \le \|\Phi_n\|_{C^0} + \|(I - \mathfrak{s}_t)R_n\|_{C^0}\\ &\overset{(2)}{\le} \|\Phi_n\|_{C^0} + K_\ell t^{-\ell}\|R_n\|_{C^\ell}\end{aligned}$$

for any $\ell \ge 0$. Here in (0) we recall that $\|\omega_{n+1}\|_{C^1} < \frac{1}{2}$; in (1) we use (8.14); and in (2) we use (8.11).

Hence it follows that

$$\|R_{n+1}\|_{C^0} \le \|\mathcal{E}_1\|_{C^0} + \|\mathcal{E}_2\|_{C^0} + \|\mathcal{E}_3\|_{C^0} \le \frac{1}{2}\|R_{n+1}\|_{C^0} + \|\mathcal{E}_2\|_{C^0} + \|\mathcal{E}_3\|_{C^0},$$

which gives

$$\begin{aligned}\|R_{n+1}\|_{C^0} &\le 2\|\mathcal{E}_2\|_{C^0} + 2\|\mathcal{E}_3\|_{C^0}\\ &\le Kt^{\sigma}\|R_n\|_{C^1}\|R_n\|_{C^0} + K_r t^{-\ell}\|R_n\|_{C^\ell} + \|\Phi_n\|_{C^0}\\ &\overset{(3)}{\le} Kt^{\sigma}\|R_n\|_{C^1}\|R_n\|_{C^0} + K_\ell t^{-\ell}\|R_n\|_{C^\ell}\\ &\quad + K_{l,\ell}(t^{-\ell+2}\|R_n\|_{C^\ell} + \|R_n\|_{C^2}^{1+\frac{\beta}{2}})^{\frac{l-2-N}{l+N}}(t^{\sigma}\|R_n\|_{C^l})^{\frac{2N+2}{l+N}}\end{aligned}$$

for any $l > N + 2$. Here in (3) we use (8.15).

Now we estimate $\|R_{n+1}\|_{C^r}$. We note that

$$R_{n+1} = (I - \omega_{n+1})^{-1} \circ (R_n + A) \circ (I - \omega_{n+1}) - A = (I - \omega_{n+1})^{-1} \circ P - A.$$

By Lemma 8.1 we have

$$\begin{aligned}\|P\|_{C^r} &\le M_r\left(1 + \|I - \omega_{n+1}\|_{C^1}^{r-1}\right)\\ &\quad \cdot (\|R_n + A\|_{C^1}\|I - \omega_{n+1}\|_{C^r} + \|R_n + A\|_{C^r}\|I - \omega_{n+1}\|_{C^1}) + \|R_n + A\|_{C^0}\\ &\overset{(1)}{\le} K_r t^{\sigma}\|R_n\|_{C^r} + K_r, \qquad \text{and} \qquad \|P\|_{C^1} \overset{(1)}{\le} K.\end{aligned}$$

Here in (1) we use the fact that ω_{n+1} satisfies the estimates $\|\omega_{n+1}\|_{C^r} \le K_r t^{\sigma}\|R_n\|_{C^r}$ (see (8.14)) and $\|\omega_{n+1}\|_{C^1} < \frac{1}{2}$. Using Lemma 8.2 this also implies that

$$\|(I - \omega_{n+1})^{-1}\|_{C^r} \le K_r\|\omega_{n+1}\|_{C^r} \le K_{r,1}t^{\sigma}\|R_n\|_{C^r}$$

and $\|(I - \omega_{n+1})^{-1}\|_{C^1} < 2$.

As a direct consequence of Lemma 8.1 and the above discussion we have

$$\begin{aligned}\|R_{n+1}\|_{C^r} &\leq M_r\left(1+\|P\|_{C^1}^{r-1}\right)\left(\|(I-\omega_{n+1})^{-1}\|_{C^1}\|P\|_{C^r}+\|(I-\omega_{n+1})^{-1}\|_{C^r}\|P\|_{C^1}\right)+K\\ &\leq K_r\|P\|_{C^r}+K_r t^\sigma\|R_n\|_{C^r}+K\\ &\leq K_{r,1}t^\sigma\|R_n\|_{C^r}+K_{r,1}.\end{aligned}$$

To get (8.4) we have

$$\begin{aligned}\|H_{n+1}-I\|_{C^0} = \|H_n\circ(I-\omega_{n+1})-I\|_{C^0} &\leq \|H_n\circ(I-\omega_{n+1})-H_n\|_{C^0}+\|H_n-I\|_{C^0}\\ &\leq \|H_n\|_{C^1}\|\omega_{n+1}\|_{C^0}+\|H_n-I\|_{C^0}\\ &\overset{(1)}{\leq} K\|R_n\|_{C^\sigma}+\|H_n-I\|_{C^0}.\end{aligned}$$

Here in (1) we use (8.5) and (8.14). $\square$

8.2. **The iteration scheme.** First we note that by [dlL92, Theorem 6.3] there exists $\sigma_0=\sigma_0(A)\in\mathbb{N}$ such that if H and H^{-1} are C^{σ_0} then H and H^{-1} are C^∞.

To set up the iterative process we take ℓ sufficiently large so that the following holds

$$\begin{aligned}&\ell\geq\max\left\{\frac{3\sigma+10}{1-\frac{\beta}{3}},\ \frac{24\sigma}{\beta},\ 2(5\max\{\sigma_0,\sigma\}+1),\ 2(2\sigma+5)\right\},\\ &\left(1+\frac{\beta}{2}\right)\left(1-\frac{5}{\ell}\right)\left(\frac{\ell-2-N}{\ell+N}\right)-2\frac{2N+2}{\ell+N}\ \geq\ 1+\frac{\beta}{3}.\end{aligned}\tag{8.16}$$

Now we construct R_n, f_n, ω_n and H_n inductively as follows. For $n=0$ we take

$$f_0=f,\quad H_0=H,\quad R_0=f-A,\quad \omega_0=0,\quad \text{and define }\ \epsilon_n=\epsilon^{\gamma^n}$$

where $\gamma=1+\frac{\beta}{4}$ and $\epsilon>0$ is sufficiently small so that the following holds

$$\|R_0\|_{C^0}\leq\epsilon_0=\epsilon,\qquad \|R_0\|_{C^\ell}\leq\epsilon_0^{-1},\qquad \|H_0-I\|_{C^0}<\epsilon_0^{\frac{1}{2}}.$$

We note that $H_0\in C^1(\mathbb{T}^N)$ by Theorem 1.1. Now we assume inductively that $H_n\in C^1(\mathbb{T}^N)$ satisfies the conjugacy equation

$$H_n\circ f_n=A\circ H_n$$

and that H_n and $R_n=f_n-A$ satisfy

$$\|R_n\|_{C^0}\leq\epsilon_n,\qquad \|R_n\|_{C^\ell}\leq\epsilon_n^{-1},\quad \|H_n-I\|_{C^0}<\sum_{i=0}^{n-1}\epsilon_i^{\frac{1}{2}}.$$

By interpolation inequalities we have

$$\|R_n\|_{C^{\sigma+2}}\leq K_\ell\|R_n\|_{C^0}^{\frac{\ell-2-\sigma}{\ell}}\|R_n\|_{C^\ell}^{\frac{2+\sigma}{\ell}}<\epsilon_n^{1-\frac{5+2\sigma}{\ell}}\leq\epsilon_n^{\frac{1}{2}}.\tag{8.17}$$

provided $\ell \geq 2(2\sigma+5)$. Here, and subsequently, we estimate various constants from above by $\epsilon_n^{-\frac{1}{\ell}}$. This can be done since ℓ is fixed, we can take ϵ small enough. We also have

$$\|H_n - I\|_{C^0} < \sum_{i=0}^{n-1} \epsilon_i^{\frac{1}{2}} < \sum_{i=1}^{\infty} (\epsilon^{\frac{1}{4}})^i < 2\epsilon^{\frac{1}{4}}. \tag{8.18}$$

Then (8.17) and (8.18) allow us to use Proposition 8.3 to obtain the new iterates R_{n+1}, f_{n+1}, ω_{n+1} and H_{n+1}. Now we show that these iterates satisfy the inductive assumption and establish appropriate convergence.

8.3. **Inductive estimates and convergence.**

We use Proposition 8.3 with $t_n = \epsilon_n^{-\frac{3}{\ell}}$ and $l = \ell$.

(1) *C^ℓ estimate for R_{n+1}*

$$\begin{aligned}\|R_{n+1}\|_{C^\ell} &\leq K_\ell t_n^\sigma \|R_n\|_{C^\ell} + K_\ell \leq K_\ell \epsilon_n^{-\frac{3\sigma}{\ell}}(\epsilon_n^{-1}+1)\\ &< \epsilon_n^{-1-\frac{\beta}{8}-\frac{3\sigma}{\ell}} \leq \epsilon_n^{-1-\frac{\beta}{4}} = \epsilon_{n+1}^{-1},\end{aligned}$$

provided $\ell \geq \frac{24\sigma}{\beta}$.

(2) *C^0 estimate for R_{n+1}*

$$\begin{aligned}\|R_{n+1}\|_{C^0} &\leq K t_n^\sigma \|R_n\|_{C^2}^2 + K_\ell t_n^{-\ell}\|R_n\|_{C^\ell}\\ &\quad + K_\ell (t_n^{-\ell+2}\|R_n\|_{C^\ell} + \|R_n\|_{C^2}^{1+\frac{\beta}{2}})^{\frac{\ell-2-N}{\ell+N}} (t_n^\sigma \|R_n\|_{C^\ell})^{\frac{2N+2}{\ell+N}}\\ &\overset{(a)}{\leq} K\epsilon_n^{2-\frac{3\sigma+10}{\ell}} + K_\ell \epsilon_n^3 \epsilon_n^{-1}\\ &\quad + K_\ell (\epsilon_n^{\frac{3(\ell-2)}{\ell}}\epsilon_n^{-1} + \epsilon_n^{(1+\frac{\beta}{2})(1-\frac{5}{\ell})})^{\frac{\ell-2-N}{\ell+N}} (\epsilon_n^{-\frac{3\sigma}{\ell}}\epsilon_n^{-1})^{\frac{2N+2}{\ell+N}}\\ &\overset{(b)}{\leq} K\epsilon_n^{2-\frac{3\sigma+10}{\ell}} + K_\ell \epsilon_n^2\\ &\quad + 2K_\ell (\epsilon_n^{(1+\frac{\beta}{2})(1-\frac{5}{\ell})})^{\frac{\ell-2-N}{\ell+N}} (\epsilon_n^{-2})^{\frac{2N+2}{\ell+N}}\\ &\overset{(c)}{<} \epsilon_n^\gamma = \epsilon_{n+1}.\end{aligned}$$

Here in (a) we use interpolation inequalities:

$$\|R_n\|_{C^2} \leq C\|R_n\|_{C^0}^{\frac{\ell-2}{\ell}} \|R_n\|_{C^\ell}^{\frac{2}{\ell}} < \epsilon_n^{1-\frac{5}{\ell}}; \tag{8.19}$$

in (b) we note that

$$(1+\frac{\beta}{2})(1-\frac{5}{\ell}) < 2(1-\frac{5}{\ell}) < 2 - \frac{6}{\ell} \quad \text{and} \quad \frac{3\sigma}{\ell} < 1.$$

Then $\epsilon_n^{-\frac{3\sigma}{\ell}}\epsilon_n^{-1} < \epsilon_n^{-2}$ and

$$\max\{\epsilon_n^{(1+\frac{\beta}{2})(1-\frac{5}{\ell})}, \epsilon_n^{\frac{3(\ell-2)}{\ell}}\epsilon_n^{-1}\} = \epsilon_n^{(1+\frac{\beta}{2})(1-\frac{5}{\ell})};$$

in (c) we use

$$\epsilon_n^{2-\frac{3\sigma+10}{\ell}} < \epsilon_n^{1+\frac{\beta}{3}}, \quad \left(\epsilon_n^{(1+\frac{\beta}{2})(1-\frac{5}{\ell})}\right)^{\frac{\ell-2-N}{\ell+N}} \left(\epsilon_n^{-2}\right)^{\frac{2N+2}{\ell+N}} < \epsilon_n^{1+\frac{\beta}{3}},$$

provided

$$2-\tfrac{3\sigma+10}{\ell} \geq 1+\tfrac{\beta}{3}, \qquad (1+\tfrac{\beta}{2})(1-\tfrac{5}{\ell})(\tfrac{\ell-2-N}{\ell+N}) - 2\tfrac{2N+2}{\ell+N} \geq 1+\tfrac{\beta}{3}.$$

By (8.16) and the assumption all inequalities above are satisfied.

(3) C^{σ_0} *estimate for* ω_{n+1}: By interpolation inequalities we have

$$\|R_n\|_{C^{\sigma_0}} \leq K_\ell \|R_n\|_{C^0}^{\frac{\ell-\sigma_0}{\ell}} \|R_n\|_{C^\ell}^{\frac{\sigma_0}{\ell}} < \epsilon_n^{1-\frac{2\sigma_0+1}{\ell}}.$$

Hence we have

$$\|\omega_{n+1}\|_{C^{\sigma_0}} \leq K t_n^\sigma \|R_n\|_{C^{\sigma_0}} \leq K\epsilon_n^{-\frac{3\sigma}{\ell}} \epsilon_n^{1-\frac{2\sigma_0+1}{\ell}} < \epsilon_n^{\frac{1}{2}}, \tag{8.20}$$

provided

$$-\frac{3\sigma}{\ell} + 1 - \frac{2\sigma_0+1}{\ell} > \frac{1}{2},$$

which is satisfied for $\ell > 2(5\max\{\sigma_0, \sigma\}+1)$.

(4) C^0 *estimate for* H_{n+1}: By (8.17) we have

$$\|H_{n+1} - I\|_{C^0} \leq K\|R_n\|_{C^\sigma} + \|H_n - I\|_{C^0} < K\epsilon_n^{1-\frac{5+2\sigma}{\ell}} + \sum_{i=0}^{n-1} \epsilon_n^{\frac{1}{2}} \leq \epsilon_n^{\frac{1}{2}} + \sum_{i=0}^{n-1} \epsilon_i^{\frac{1}{2}} = \sum_{i=0}^{n} \epsilon_i^{\frac{1}{2}}$$

Consequently, we have

$$\begin{aligned} f_{n+1} &= \tilde{H}_{n+1}^{-1} \circ \tilde{H}_n^{-1} \circ \cdots \circ \tilde{H}_1^{-1} \circ f \circ \tilde{H}_1 \circ \cdots \circ \tilde{H}_{n+1} \\ &= \mathfrak{L}_{n+1}^{-1} \circ f \circ \mathfrak{L}_{n+1} \end{aligned}$$

where $\tilde{H}_i = I - \omega_i$, $1 \leq i \leq n+1$; and $\mathfrak{L}_{n+1} = \tilde{H}_1 \circ \cdots \circ \tilde{H}_{n+1}$.

Finally, (8.20) implies that $\mathfrak{L}_n$ converges in C^{σ_0} topology to a C^{σ_0} diffeomorphism H, which is a conjugacy between f and A. By [dlL92, Theorem 6.3] and the choice of σ_0 we conclude that H is a C^∞ diffeomorphism.

References

[A67] D. Anosov. *Geodesic Flows on Closed Riemannian Manifolds with Negative Curvature.* Proceedings of Steklov Institute of Mathematics, 1967, 90.

[ASV13] A. Avila, J. Santamaria, M. Viana. *Holonomy invariance: rough regularity and applications to Lyapunov exponents.* Asterisque 358 (2013), 13-74.

[BV04] C. Bonatti, M. Viana. *Lyapunov exponents with multiplicity 1 for deterministic products of matrices.* Ergodic Theory Dynam. Systems. 24 (2004) 1295-1330.

[DKt10] D. Damjanović and A. Katok. *Local rigidity of partially hyperbolic actions. I. KAM method and $\mathbb{Z}^k$ actions on the torus.* Annals of Mathematics **172** (2010), 1805–1858.

[dlL87] R. de la Llave. *Invariants for smooth conjugacy of hyperbolic dynamical systems II.* Comm. Math. Phys., 109 (1987), 368-378.

[dlL92] R. de la Llave. *Smooth conjugacy and SRB measures for uniformly and non-uniformly hyperbolic systems.* Comm. Math. Phys., 150 (1992), 289-320.

[dlL99] R. de la Llave. *A tutorial on KAM theory.* In Smooth Ergodic Theory and its Applications (Seattle, WA, 1999), Proc. Sympos. Pure Math. 69, Amer. Math. Soc., Providence, RI, 2001, pp. 175-292.

[dlL02] R. de la Llave. *Rigidity of higher-dimensional conformal Anosov systems.* Ergodic Theory Dynam. Systems, 22 (2002), no. 6, 1845–1870.

[dlL04] R. de la Llave. *Further rigidity properties of conformal Anosov systems.* Ergodic Theory Dynam. Systems, 24 (2004), no. 5, 1425–1441.

[dlLM88] R. de la Llave, R. Moriyón. *Invariants for smooth conjugacy of hyperbolic dynamical systems IV.* Commun. Math. Phys., 116 (1988), 185-192.

[dlLO98] R. de la Llave and R. Obaya. *Regularity of the composition operator in spaces of Hölder functions.* Discrete and Continuous Dynamical Systems. 5 (1999), no. 1, 157-184.

[dW21] Jonathan De Witt *Local Lyapunov spectrum rigidity of nilmanifold automorphisms.* J. Modern Dynamics 17 (2021): 65-109.

[G08] A. Gogolev. *Smooth conjugacy of Anosov diffeomorphisms on higher dimensional tori.* J. Modern Dynamics, 2, no. 4, 645-700 (2008).

[G17] A. Gogolev. *Bootstrap for local rigidity of Anosov automorphisms of the 3-torus.* Comm. Math. Phys., 352, no. 2, 439-455 (2017).

[GG08] A. Gogolev and M. Guysinski. *C^1-differentiable conjugacy of Anosov diffeomorphisms on three dimensional torus.* DCDS-A, 22, no. 1/2, 183-200 (2008).

[GKS11] A. Gogolev, B. Kalinin, V. Sadovskaya. *Local rigidity for Anosov automorphisms.* Math. Research Letters, 18 (2011), no. 05, 843-858.

[GKS20] A. Gogolev, B. Kalinin, V. Sadovskaya. *Local rigidity of Lyapunov spectrum for toral automorphisms.* Israel J. Math., 238 (2020), 389-403.

[Ha82] R. S. Hamilton. *The inverse function theorem of Nash and Moser.* Bull. Amer. Math. Soc. (N.S.), 7 (1982), 65-222.

[HPS77] M. Hirsch, C. Pugh, M. Shub. *Invariant manifolds.* Springer-Verlag, New York, 1977.

[K11] B. Kalinin. *Livšic theorem for matrix cocycles.* Annals of Mathematics, 173 (2011), no. 2, 1025-1042.

[KS03] B. Kalinin, V. Sadovskaya. *On local and global rigidity of quasiconformal Anosov diffeomorphisms.* J. Institute of Mathematics of Jussieu, 2 (2003), no. 4, 567-582.

[KS09] B. Kalinin, V. Sadovskaya. *On Anosov diffeomorphisms with asymptotically conformal periodic data.* Ergodic Theory Dynam. Systems, 29 (2009), 117-136.

[KS13] B. Kalinin, V. Sadovskaya. *Cocycles with one exponent over partially hyperbolic systems.* Geometriae Dedicata, Vol. 167, Issue 1 (2013), 167-188.

[KK18] M. Korobkov and J. Kristensen. *The Trace Theorem, the Luzin N and Morse-Sard Properties for the Sharp Case of Sobolev-Lorentz Mappings.* J Geom Anal (2018) 28:2834-2856.

[La93] V. F. Lazutkin, *KAM Theory and Semiclassical Approximations to Eigenfunctions, Ergeb. Math. Grenzgeb.* 24, Springer-Verlag, New York, 1993.

[Le84] F. Ledrappier. *Propriétés ergodiques des mesures de Sinaï.* Inst. Hautes Etudes Sci. Publ. Math. No. 59 (1984), 163-188.

[Liv72] A. N. Livšic. *Cohomology of dynamical systems.* Math. USSR Izvestija 6, 1278-1301, 1972.

[MM73] M. Marcus and V. J. Mizel. *Transformations by functions in Sobolev spaces and lower semicontinuity for parametric variational problems.* Bull. Amer. Math. Soc. (N.S.) 79 (1973), 790-795.

[P01] L. De Pascale. *The Morse-Sard theorem in Sobolev spaces.* Indiana Univ. Math. J. 50, 1371-1386, 2001.

[PW01] M. Pollicott, C. P. Walkden. *Livšic theorems for connected Lie groups.* Trans. Amer. Math. Soc., 353(7), 2879-2895, 2001.

[Pa99] W. Parry. *The Livšic periodic point theorem for non-Abelian cocycles.* Ergodic Theory Dynam. Systems, 19(3), 687-701, 1999.

[PaP97] W. Parry, M. Pollicott. *The Livšic cocycle equation for compact Lie group extensions of hyperbolic systems.* J. London Math. Soc. (2), 56(2) 405-416, 1997.

[S13] V. Sadovskaya. *Cohomology of $GL(2,\mathbb{R})$-valued cocycles over hyperbolic systems.* Discrete and Continuous Dynamical Systems, vol. 33, no. 5 (2013), 2085-2104.

[S15] V. Sadovskaya. *Cohomology of fiber bunched cocycles over hyperbolic systems.* Ergodic Theory Dynam. Systems, Vol. 35, Issue 8 (2015), 2669-2688.

[SaY] R. Saghin, J. Yang. *Lyapunov exponents and rigidity of Anosov automorphisms and skew products.* Advances in Math., Vol. 355 (2019).

[Sch99] K. Schmidt. *Remarks on Livšic theory for non-Abelian cocycles.* Ergodic Theory Dynam. Systems, 19(3), 703-721, 1999.

[V08] M. Viana. *Almost all cocycles over any hyperbolic system have nonvanishing Lyapunov exponents.* Ann. of Math. (2) 167 (2008), no. 2, 643–680.

[Wa70] P. Walters. *Conjugacy properties of affine transformations of nilmanifolds.* Math. Systems Theory 4 (1970), 327-333.

Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA.

E-mail address: kalinin@psu.edu, sadovskaya@psu.edu

Department of Mathematics, Michigan State University, East Lansing, MI 48824,USA

E-mail address: wangzq@math.msu.edu