JOURNAL OF MODERN DYNAMICS
VOLUME 11, 2017, 341–368
doi: 10.3934/jmd.2017014

NORMAL FORMS FOR NON-UNIFORM CONTRACTIONS

BORIS KALININ AND VICTORIA SADOVSKAYA
(Communicated by Giovanni Forni)

ABSTRACT. Let f be a measure-preserving transformation of a Lebesgue space (X,μ) and let $\mathscr{F}$ be its extension to a bundle $\mathscr{E} = X\times\mathbb{R}^m$ by smooth fiber maps $\mathscr{F}_x : \mathscr{E}_x \to \mathscr{E}_{fx}$ so that the derivative of $\mathscr{F}$ at the zero section has negative Lyapunov exponents. We construct a measurable system of smooth coordinate changes $\mathscr{H}_x$ on $\mathscr{E}_x$ for μ-a.e. x so that the maps $\mathscr{P}_x = \mathscr{H}_{fx}\circ\mathscr{F}_x\circ\mathscr{H}_x^{-1}$ are sub-resonance polynomials in a finite dimensional Lie group. Our construction shows that such $\mathscr{H}_x$ and $\mathscr{P}_x$ are unique up to a sub-resonance polynomial. As a consequence, we obtain the centralizer theorem that the coordinate change $\mathscr{H}$ also conjugates any commuting extension to a polynomial extension of the same type. We apply our results to a measure-preserving diffeomorphism f with a non-uniformly contracting invariant foliation W. We construct a measurable system of smooth coordinate changes $\mathscr{H}_x : W_x \to T_xW$ such that the maps $\mathscr{H}_{fx}\circ f\circ\mathscr{H}_x^{-1}$ are polynomials of sub-resonance type. Moreover, we show that for almost every leaf the coordinate changes exist at each point on the leaf and give a coherent atlas with transition maps in a finite dimensional Lie group.

1. INTRODUCTION

The theory of normal forms for smooth maps originated in the works of Poincaré and Sternberg [26], and normal forms at fixed points and invariant manifolds have been extensively studied [4]. More recently, non-stationary normal form theory was developed in the context of a diffeomorphism f contracting a foliation W. The goal is to obtain a family of diffeomorphisms $\mathscr{H}_x : W_x \to T_xW$ such that the maps

$$\tilde{f}_x = \mathscr{H}_{fx}\circ f\circ\mathscr{H}_x^{-1} : T_xW \to T_{fx}W \tag{1.1}$$

are as simple as possible, for example, linear maps or polynomial maps in a finite dimensional Lie group. Such a map $\tilde{f}_x$ is called a normal form of f on W_x.

The non-stationary normal form theory started with the linearization along one-dimensional foliations obtained by Katok and Lewis [15]. In a more general

Received May 12, 2016; revised March 3, 2017.
2010 *Mathematics Subject Classification:* Primary: 37D10, 37D25; Secondary: 37D30, 34C20.
Key words and phrases: Normal form, contracting foliation, non-uniform hyperbolicity, Lyapunov exponents, polynomial map, homogeneous structure.
BK: Supported in part by Simons Foundation grant 426243.
VS: Supported in part by NSF grant DMS-1301693.

setting of contractions with narrow band spectrum, it was developed by Guysinsky and Katok [11, 10], and a differential geometric point of view was presented by Feres [7]. For the linearization, further results were obtained by the second author in [25], and it was shown in [19] that the coordinates $\mathscr{H}_x$ give a consistent affine atlas on each leaf of W. In [20] we extended these results to the general narrow band case. More precisely, we gave a construction of $\mathscr{H}_x$ that depend smoothly on x along the leaves and proved that they define an atlas with transition maps in a finite dimensional Lie group. Non-stationary normal forms were used extensively in the study of rigidity of uniformly hyperbolic dynamical systems and group actions, see, for example, [17, 18, 19, 5, 6, 9, 8].

To obtain applications for non-uniformly hyperbolic systems and actions, one needs a similar theory of non-stationary normal forms for non-uniform contractions. The existence and centralizer theorems were stated without proof in [12] along with a program of potential applications. The theory, however, was not developed for quite a while. The linearization of a $C^{1+\alpha}$ diffeomorphism along a one-dimensional non-uniformly contracting foliation was constructed in [13] and used in the study of measure rigidity in [13, 14]. Similar results for higher dimensional foliations with pinched exponents were obtained by Katok and Rodriguez Hertz in [16]. The existence of $\mathscr{H}_x$ for a general contracting C^∞ extension was proved by Li and Lu [22] in the setting of random dynamical systems. Some results, such as existence of Taylor polynomial or formal series for $\mathscr{H}_x$, can be obtained for extensions more general than contractions, see [2, 1, 22].

In this paper we develop the theory of non-stationary polynomial normal forms for smooth extensions of measure preserving transformations by non-uniform contractions, described in the beginning of Section 2. This is a convenient general setting for the construction. The foliation setting reduces to it by locally identifying the leaf W_x with its tangent space $\mathscr{E}_x = T_x W$ and viewing $\mathscr{F}_x = f|_{W_x} : \mathscr{E}_x \to \mathscr{E}_{fx}$ as an extension of the base system $f : \mathcal{M} \to \mathcal{M}$ by smooth maps on the bundle $\mathscr{E} = TW$. The base system can then be viewed as just a measure preserving one. In the extension setting, the map $\mathscr{H}_x$ is a coordinate change on $\mathscr{E}_x$, and we denote

$$\mathscr{P}_x = \mathscr{H}_{fx} \circ \mathscr{F}_x \circ \mathscr{H}_x^{-1} : \mathscr{E}_x \to \mathscr{E}_{fx}.$$

In Theorem 2.3 we construct coordinate changes $\mathscr{H}_x$ for μ almost every x so that $\mathscr{P}_x$ is a sub-resonance polynomial. For any regularity of $\mathscr{F}$ above the critical level, we obtain $\mathscr{H}$ in the same regularity class.

Our construction allows us to describe the exact extent of non-uniqueness in $\mathscr{H}_x$ and $\mathscr{P}_x$. Essentially, they are defined up to a sub-resonance polynomial. As a consequence of this, we obtain the centralizer theorem that the coordinate change $\mathscr{H}$ also conjugates any commuting extension to a normal form of the same type. We just learned of similar results in differential geometric formulations by Melnick [23]. The approach in [23] is different from ours and it relies on ergodic theorems for higher jets of $\mathscr{F}_x$. Our results assume only temperedness of the higher derivatives of $\mathscr{F}_x$ rather than certain integrability

required in [23]. This allows us to obtain applications to the foliation setting without any assumptions on transverse regularity of the foliation. In particular, we consider a diffeomorphism f which preserves an ergodic measure with some negative Lyapunov exponents and take W to be any strong part of the stable foliation. In this setting Theorem 2.5 gives sub-resonance normal forms for f along the leaves of W. Moreover, we show that for almost every leaf the normal form coordinates $\mathscr{H}_x$ exist at each point on the leaf and give a coherent atlas with transition maps in a finite dimensional Lie group G determined by sub-resonance polynomials. This yields an invariant structure of a G homogeneous space on almost every leaf.

We expect these results to be useful in the study of non-uniformly hyperbolic systems and group actions.

2. STATEMENTS OF RESULTS

ASSUMPTIONS 2.1. In this paper,

- (X,μ) is a Lebesgue probability space,
- $f : X \to X$ is an invertible ergodic measure-preserving transformation of (X,μ),
- $\mathscr{E} = X \times \mathbb{R}^m$ is a finite dimensional vector bundle over X,
- $\mathscr{V}$ is a neighborhood of the zero section in $\mathscr{E}$,
- $\mathscr{F} : \mathscr{V} \to \mathscr{E}$ is a measurable extension of f that preserves the zero section,
- $F : \mathscr{E} \to \mathscr{E}$ is the derivative of $\mathscr{F}$ at zero section, $F_x = D_0\mathscr{F}_x : \mathscr{E}_x \to \mathscr{E}_{fx}$,
- F and F^{-1} exist and satisfy $\log\|F_x\| \in L^1(X,\mu)$ and $\log\|F_x^{-1}\| \in L^1(X,\mu)$,
- the Lyapunov exponents of F are negative: $\chi_1 < \cdots < \chi_\ell < 0$.

Sub-resonance polynomials. Let $\chi_1 < \cdots < \chi_\ell < 0$ be the distinct Lyapunov exponents of F and let $\mathscr{E}_x = \mathscr{E}_x^1 \oplus \cdots \oplus \mathscr{E}_x^\ell$ be the splitting of $\mathscr{E}_x$ for $x \in \Lambda$ into the Lyapunov subspaces given by the Multiplicative Ergodic Theorem 3.1.

We say that a map between vector spaces is *polynomial* if each component is given by a polynomial in some, and hence every, bases. We consider a polynomial map $P : \mathscr{E}_x \to \mathscr{E}_y$ with $P(0_x) = 0_y$ and split it into components $(P_1(t), \ldots, P_\ell(t))$, where $P_i : \mathscr{E}_x \to \mathscr{E}_y^i$. Each P_i can be written uniquely as a linear combination of polynomials of specific homogeneous types: we say that $Q : \mathscr{E}_x \to \mathscr{E}_y^i$ has homogeneous type $s = (s_1, \ldots, s_\ell)$ if for any real numbers $a_1, \ldots, a_\ell$ and vectors $t_j \in \mathscr{E}_x^j$, $j = 1, \ldots, \ell$, we have

$$Q(a_1 t_1 + \cdots + a_\ell t_\ell) = a_1^{s_1} \cdots a_\ell^{s_\ell} \cdot Q(t_1 + \cdots + t_\ell). \tag{2.1}$$

DEFINITION 2.2. We say that a polynomial map $P : \mathscr{E}_x \to \mathscr{E}_y$ is *sub-resonance* if each component P_i has only terms of homogeneous types $s = (s_1, \ldots, s_\ell)$ satisfying *sub-resonance relations*

$$\chi_i \le \sum s_j \chi_j, \quad \text{where } s_1, \ldots, s_\ell \text{ are non-negative integers.} \tag{2.2}$$

We denote by $\mathscr{S}_{x,y}$ the space of all sub-resonance polynomial maps from $\mathscr{E}_x$ to $\mathscr{E}_y$.

Clearly, for any sub-resonance relation we have that $s_j = 0$ for $j < i$ and that $\sum s_j \le \chi_1/\chi_\ell$. It follows that sub-resonance polynomial maps have degree at most

$$d = d(\chi) = \lfloor \chi_1/\chi_\ell \rfloor. \tag{2.3}$$

Sub-resonance polynomial maps $P : \mathscr{E}_x \to \mathscr{E}_x$ with $P(0) = 0$ with invertible derivative at the origin form a group with respect to composition [11]. We will denote this finite-dimensional Lie group by G_x^χ. All groups G_x^χ are isomorphic, moreover, any map $P \in \mathscr{S}_{x,y}$ with $P(0_x) = 0_y$ and invertible derivative at 0_x induces an isomorphism between G_x^χ and G_y^χ by conjugation.

We denote by $B_{x,\sigma(x)}$ the closed ball of radius $\sigma(x)$ centered at $0 \in \mathscr{E}_x$. For $N \ge 1$ and $0 < \alpha \le 1$ we denote by $C^{N,\alpha}(B_{x,\sigma(x)}) = C^{N,\alpha}(B_{x,\sigma(x)}, \mathscr{E}_x)$ the space of functions from $B_{x,\sigma(x)}$ to $\mathscr{E}_x$ with continuous derivatives up to order $N \ge 1$ on $B_{x,\sigma(x)}$ and with N^{th} derivative satisfying α-Hölder condition at 0:

$$\|D^{(N)}R\|_\alpha = \sup\{\|D_t^{(N)}R - D_0^{(N)}R\| \cdot \|t\|^{-\alpha} : 0 \ne t \in B_{x,\sigma(x)}\} < \infty. \tag{2.4}$$

We call $\|D^{(N)}R\|_\alpha$ the α-Hölder constant of $D^{(N)}R$ at 0. We equip the space $C^{N,\alpha}(B_{x,\sigma(x)})$ with the norm

$$\|R\|_{C^{N,\alpha}(B_{x,\sigma(x)})} = \max\{\|R\|_0, \|D^{(1)}R\|_0, \dots, \|D^{(N)}R\|_0, \|D^{(N)}R\|_\alpha\}, \tag{2.5}$$

where $\|D^{(k)}R\|_0 = \sup\{\|D_t^{(k)}R\| : t \in B_{x,\sigma(x)}\}$.

We say that a non-negative real-valued function K on X is *ε-tempered at x* if

$$\sup\{K(f^n x)\, e^{-\varepsilon n} : n \in \mathbb{N}\} < \infty, \tag{2.6}$$

and that K is *ε-tempered on a set* if it is ε-tempered at each of its points.

We consider an extension $\mathscr{F}$ satisfying Assumptions 2.1 and denote by Λ the set of regular points and by $\chi_1 < \dots < \chi_\ell < 0$ the Lyapunov exponents of F given by the Multiplicative Ergodic Theorem 3.1. For N and α as above we define

$$\kappa = 1 + 3/\alpha \text{ if } N = 1 \text{ and } \kappa = 4 \text{ if } N \ge 2. \tag{2.7}$$

If $N \ge 2$ we allow $\alpha = 0$, in which case we understand $C^{N,\alpha}$ as C^N.

THEOREM 2.3 (Normal forms for non-uniformly contracting extensions)**.** *Let $\mathscr{F}$ be an extension of f satisfying Assumptions 2.1. Suppose that*

$$N \ge 1, \quad 0 \le \alpha \le 1 \quad \textit{and} \quad N + \alpha > \chi_1/\chi_\ell. \tag{2.8}$$

Then there exist positive constants $L = L(N,\alpha)$ and $\varepsilon* = \varepsilon_*(N,\alpha,\chi_1,\dots,\chi_\ell)$ so that for any $0 < \varepsilon \le \varepsilon_*$ the following holds. If there exists a positive measurable function $\sigma : \Lambda \to \mathbb{R}$ so that $1/\sigma$ is ε-tempered on Λ and $\mathscr{F}_x$ is $C^{N,\alpha}(B_{x,\sigma(x)})$ for all $x \in \Lambda$ with the derivatives measurable in x and with $\|\mathscr{F}_x\|_{C^{N,\alpha}}$ ε-tempered on Λ then

(1) There exists a positive measurable function $\rho : \Lambda \to \mathbb{R}$ so that $1/\rho$ is $\kappa\varepsilon$-tempered on Λ and a measurable family $\{\mathscr{H}_x\}_{x\in\Lambda}$ of $C^{N,\alpha}$ diffeomorphisms $\mathscr{H}_x : B_{x,\rho(x)} \to \mathscr{E}_x$ satisfying $\mathscr{H}_x(0) = 0$ and $D_0\mathscr{H}_x = Id$ which conjugate $\mathscr{F}$ to a sub-resonance polynomial extension $\mathscr{P}$:

$$\mathscr{H}_{fx} \circ \mathscr{F}_x = \mathscr{P}_x \circ \mathscr{H}_x, \text{ where } \mathscr{P}_x \in \mathscr{S}_{x,fx} \text{ for all } x \in \Lambda.$$

Moreover, $\|\mathscr{H}_x\|_{C^{N,\alpha}(B_{x,\rho(x)})}$ *is* $L\varepsilon$*-tempered on* Λ *and* $\|D_0^{(n)}\mathscr{H}_x\|$ *is* $n^2\varepsilon$*-tempered on* Λ *for* $n=1,\dots,N$*, with respect to the* ε*-Lyapunov metric* (3.2).

(2) *Suppose* $\tilde{\mathscr{H}}=\{\tilde{\mathscr{H}}_x\}_{x\in\Lambda}$ *is another measurable family of diffeomorphisms as in (1) conjugating* $\mathscr{F}$ *to a sub-resonance polynomial extension* $\tilde{\mathscr{P}}$*. Then for all* $x\in\Lambda$ *there exists* $G_x\in G_x^\chi$ *which is measurable and tempered in* x *such that* $\mathscr{H}_x=G_x\circ\tilde{\mathscr{H}}_x$*. Moreover, if* $D_0^{(n)}\tilde{\mathscr{H}}_x=D_0^{(n)}\mathscr{H}_x$ *for all* $n=2,\dots,d$ *with* $d=\lfloor\chi_1/\chi_\ell\rfloor$*, then* $\mathscr{H}_x=\tilde{\mathscr{H}}_x$ *for all* $x\in\Lambda$*. In particular,* $\{\mathscr{H}_x\}_{x\in\Lambda}$ *is unique if* $d=1$*.*

(3) *Let* $g:X\to X$ *be an invertible map commuting with* f *and let* Λ' *be a subset of* Λ *which is both* f *and* g *invariant. Let* $\mathscr{G}(x,t)=(g(x),\mathscr{G}_x(t))$ *be an extension of* g *to* $\mathscr{E}$ *which preserves the zero section and commutes with* $\mathscr{F}$*. Suppose that* $\mathscr{G}_x$ *is* $C^{N,\alpha}(B_{x,\sigma(x)})$ *for all* $x\in\Lambda'$ *with the derivatives measurable in* x*, and that* $\|\mathscr{G}_x\|_{C^{N,\alpha}}$ *and* $\|(D_0\mathscr{G}_x)^{-1}\|$ *are* ε*-tempered on* Λ'*. Then* $\mathscr{H}_{gx}\circ\mathscr{G}_x\circ\mathscr{H}_x^{-1}\in S_{x,fx}$ *for all* $x\in\Lambda'$*.*

COROLLARY 2.4. *Suppose that* $\mathscr{F}_x$ *is* $C^\infty(B_{x,\sigma(x)})$ *and that* $1/\sigma$ *and* $\|\mathscr{F}_x\|_{C^N}$ *for each* $N\in\mathbb{N}$ *are* ε*-tempered on* Λ *for each* $\varepsilon>0$*. Then* $\mathscr{H}_x$ *in part (1) of Theorem 2.3 is* $C^\infty(B_{x,\rho(x)})$*.*

Normal forms on stable manifolds. Let $\mathcal{M}$ be a compact smooth manifold and let f be a diffeomorphism of $\mathcal{M}$ preserving an ergodic Borel probability measure μ. We assume that f is $C^{N,\alpha}$, that is C^N with N^{th} derivative α-Hölder on $\mathcal{M}$. We denote by Λ the full measure set of Lyapunov regular points for (Df,μ). Let $\chi_1<\cdots<\chi_{\ell'}$ be the Lyapunov exponents of (Df,μ) and suppose ℓ is such that $\chi_\ell<0$. Then for each $x\in\Lambda$ there exists the (strong) stable manifold W_x tangent to $\mathcal{E}_x=\mathcal{E}_x^1\oplus\cdots\oplus\mathcal{E}_x^\ell$ [24, Theorem 6.1].

THEOREM 2.5 (Normal forms on stable manifolds)**.** *Let* $\mathcal{M}$ *be a compact smooth manifold and let* f *be a* $C^{N,\alpha}$ *diffeomorphism of* $\mathcal{M}$ *preserving an ergodic Borel probability measure* μ*. Suppose that* $N\geq 1$*,* $0\leq\alpha\leq 1$ *and* $N+\alpha>\chi_1/\chi_\ell$*. Then there exist a full measure set* X *which consists of full stable manifolds* W_x *and a measurable family* $\{\mathscr{H}_x\}_{x\in X}$ *of* $C^{N,\alpha}$ *diffeomorphisms*

$$\mathscr{H}_x:W_x\to\mathcal{E}_x=T_xW_x$$

such that

(i) $\mathscr{P}_x=\mathscr{H}_{fx}\circ f\circ\mathscr{H}_x^{-1}:\mathcal{E}_x\to\mathcal{E}_{fx}$ *is a sub-resonance polynomial map for each* $x\in X$*,*

(ii) $\mathscr{H}_x(x)=0$ *and* $D_x\mathscr{H}_x$ *is the identity map for each* $x\in X$*,*

(iii) $\|\mathscr{H}_x\|_{C^{N,\alpha}}$ *is tempered on* X*,*

(iv) $\mathscr{H}_y\circ\mathscr{H}_x^{-1}:\mathcal{E}_x\to\mathcal{E}_y$ *is a sub-resonance polynomial map for all* $x\in X$ *and* $y\in W_x$*,*

(v) *If* $g:\mathcal{M}\to\mathcal{M}$ *is a* $C^{N,\alpha}$ *diffeomorphism commuting with* f *which preserves the measure class of* μ *then* $\mathscr{H}_{gx}\circ g\circ\mathscr{H}_x^{-1}:\mathcal{E}_x\to\mathcal{E}_{gx}$ *is a sub-resonance polynomial map for all* x *in a full measure set* X' *which consists of full stable manifolds.*

Another way to interpret (iv) is to view $\mathscr{H}_x$ as a coordinate chart on W_x identifying it with $\mathscr{E}_x$. In this coordinate chart, (iv) yields that all transition maps $\mathscr{H}_y \circ \mathscr{H}_x^{-1}$ for $y \in W_x$ are in the group $\bar{G}_x^\chi$ generated by G_x^χ and the translations of $\mathscr{E}_x$. Thus $\mathscr{H}_x$ gives the leaf a structure of homogeneous space $W_x \sim \bar{G}_x^\chi / G_x^\chi$, which is consistent with other coordinate charts $\mathscr{H}_y$ for $y \in W_x$ and is preserved by the normal form $\mathscr{P}_x$ by (i).

COROLLARY 2.6. *Under the assumptions of the Theorem 2.5, if $d = \lfloor \chi_1/\chi_\ell \rfloor = 1$, i.e., $2\chi_\ell < \chi_1$, then $\mathscr{P}_x$ is the linear map $Df|_{\mathscr{E}_x}$, the family $\{\mathscr{H}_x\}_{x\in X}$ satisfying (ii) and (iii) is unique, the maps $\mathscr{H}_y \circ \mathscr{H}_x^{-1} : \mathscr{E}_x \to \mathscr{E}_y$ are affine for all $x \in X$ and $y \in W_x$, and $\mathscr{H}_y$ depends C^N-smoothly on y along the stable manifolds.*

3. LYAPUNOV EXPONENTS AND LYAPUNOV NORM

In this section we review some basic definitions and facts of the Oseledets theory of linear extensions. We use [3] as a general reference. For a linear extension F of a map f we will use the notation

$$F_x^n = F_{f^{n-1}x} \circ \cdots \circ F_{fx} \circ F_x. \tag{3.1}$$

THEOREM 3.1 (Oseledets Multiplicative Ergodic Theorem, see [3] Theorem 3.4.3)**.** *Let f be an invertible ergodic measure-preserving transformation of a Lebesgue probability space (X,μ), and let F be a measurable linear extension satisfying $\log\|F_x\|, \log\|F_x^{-1}\| \in L^1(X,\mu)$. Then there exist numbers $\chi_1 < \cdots < \chi_\ell$, an f-invariant set Λ with $\mu(\Lambda) = 1$, and an F-invariant Lyapunov decomposition*

$$\mathscr{E}_x = \mathscr{E}_x^1 \oplus \cdots \oplus \mathscr{E}_x^\ell \ \text{ for } x \in \Lambda$$

such that

(i) $\lim\limits_{n\to\pm\infty} n^{-1}\log\|F_x^n v\| = \chi_i$ *for any $i = 1,\dots,\ell$ and any $0 \neq v \in \mathscr{E}_x^i$, and*

(ii) $\lim\limits_{n\to\pm\infty} n^{-1}\log|\det F_x^n| = \sum_{i=1}^\ell m_i\chi_i$, *where $m_i = \dim \mathscr{E}_x^i$.*

The numbers $\chi_1,\dots,\chi_\ell$ are called the *Lyapunov exponents* of F and the points of the set Λ are called *regular*.

We denote the standard scalar product in $\mathbb{R}^m$ by $\langle\cdot,\cdot\rangle$. For a fixed $\varepsilon > 0$ and a regular point x, the *ε-Lyapunov scalar product (or metric)* $\langle\cdot,\cdot\rangle_{x,\varepsilon}$ in $\mathscr{E}_x = \mathbb{R}^m$ is defined as follows. For $u \in \mathscr{E}_x^i$ and $v \in \mathscr{E}_x^j$ with $i \neq j$, $\langle u, v\rangle_{x,\varepsilon} := 0$, and for $i = 1,\dots,\ell$ and $u, v \in \mathscr{E}_x^i$,

$$\langle u, v\rangle_{x,\varepsilon} = m \sum_{n\in\mathbb{Z}} \langle F_x^n(u), F_x^n(v)\rangle \exp(-2\chi_i n - \varepsilon|n|). \tag{3.2}$$

Note that the series converges exponentially for any regular x. The constant m in front of the conventional formula is introduced for more convenient comparison with the standard scalar product. Usually, ε will be fixed and we will denote $\langle\cdot,\cdot\rangle_{x,\varepsilon}$ simply by $\langle\cdot,\cdot\rangle_x$ and call it the *Lyapunov scalar product*. The norm generated by this scalar product is called the *Lyapunov norm* and is denoted by $\|\cdot\|_{x,\varepsilon}$ or $\|\cdot\|_x$.

Below we summarize the basic properties of the Lyapunov scalar product and norm, for more details see [3, Sections 3.5.1-3.5.3]. A direct calculation shows [3, Theorem 3.5.5] that for any regular x and any $u \in \mathscr{E}^i_x$

$$\exp(n\chi_i - \varepsilon|n|)\,\|u\|_{x,\varepsilon} \le \|F^n_x(u)\|_{f^n x,\varepsilon} \le \exp(n\chi_i + \varepsilon|n|)\,\|u\|_{x,\varepsilon} \text{ for all } n \in \mathbb{Z} \tag{3.3}$$

and

$$\exp(n\chi_\ell - \varepsilon n) \le \|F^n_x\|_{f^n x \leftarrow x} \le \exp(n\chi_\ell + \varepsilon n) \text{ for all } n \in \mathbb{N}, \tag{3.4}$$

where $\|\cdot\|_{f^n x \leftarrow x}$ is the operator norm with respect to the Lyapunov norms. It is defined for any points $x, y \in \Lambda$ and any linear map $F : \mathscr{E}_x \to \mathscr{E}_y$ as follows:

$$\|F\|_{y \leftarrow x} = \sup\{\|Fu\|_{y,\varepsilon} :\ u \in \mathscr{E}_x,\ \|u\|_{x,\varepsilon} = 1\}.$$

We emphasize that Lyapunov scalar product and norm are defined only for regular points and depend measurably on the point. Thus, a comparison with the standard norm is important. The uniform lower bound follows easily from the definition: $\|u\|_{x,\varepsilon} \ge \|u\|$. The upper bound is not uniform, but it changes slowly along the regular orbits [3, Proposition 3.5.8]: there exists a measurable function $K_\varepsilon(x)$ defined on Λ such that

$$\|u\| \le \|u\|_{x,\varepsilon} \le K_\varepsilon(x) \cdot \|u\| \text{ for all } x \in \Lambda \text{ and } u \in \mathscr{E}_x, \quad \text{where } K_\varepsilon(x) \ge 1, \tag{3.5}$$

and

$$K_\varepsilon(x) e^{-\varepsilon|n|} \le K_\varepsilon(f^n x) \le K_\varepsilon(x) e^{\varepsilon|n|} \quad \text{for all } x \in \Lambda \text{ and } n \in \mathbb{Z}. \tag{3.6}$$

Using (3.5) we obtain that for any point $x, y \in \Lambda$ and any linear map $F : \mathscr{E}_x \to \mathscr{E}_y$

$$K_\varepsilon(x)^{-1} \cdot \|F\| \le \|F\|_{y \leftarrow x} \le K_\varepsilon(y) \cdot \|F\|. \tag{3.7}$$

When ε is fixed we will usually omit it and write $K(x) = K_\varepsilon(x)$ and $\|u\|_x = \|u\|_{x,\varepsilon}$.

Similarly, we will consider the Lyapunov norm of a homogeneous polynomial map $R : \mathscr{E}_x \to \mathscr{E}_y$ of degree n defined as

$$\|R\|_{y \leftarrow x} = \sup\{\|R(u)\|_{y,\varepsilon} :\ u \in \mathscr{E}_x,\ \|u\|_{x,\varepsilon} = 1\}. \tag{3.8}$$

It follows that

$$\|R \circ P\| \le \|R\| \cdot \|P\|^n. \tag{3.9}$$

For a homogeneous polynomial map $R : \mathscr{E}_x \to \mathscr{E}_y$ of degree n we have

$$K_\varepsilon(x)^n \cdot \|R\| \le \|R\|_{y \leftarrow x} \le K_\varepsilon(y) \cdot \|R\|. \tag{3.10}$$

This formula allows us to switch between the standard and Lyapunov norms in spaces of polynomials and smooth functions.

4. PROOF OF THEOREM 2.3

We note that (2.8) implies $N \geq d$. We give the proof for the case $\alpha > 0$. The proof for $\alpha = 0$, and hence $N \geq 2$, is similar but avoids difficulties of estimating the Hölder constant at 0. We will use the notation $\mathscr{F}_x^n = \mathscr{F}_{f^{n-1}x} \circ \ldots \circ \mathscr{F}_{fx} \circ \mathscr{F}_x$. Now we define constants that will be used throughout the proof. We set

$$L \geq L(N,\alpha) = \max\{\kappa, N^3 + 3N^2 + 1\}, \quad \text{where } \kappa \text{ is given by (2.7).} \tag{4.1}$$

We define $\lambda < 0$ as the largest value of $-\chi_i + \sum_{j=1}^{\ell} s_j \chi_j$ over all $i \in \{1,\ldots,\ell\}$ and non-negative integers $s_1,\ldots,s_\ell$ such that sub-resonance condition (2.2) is *not* satisfied:

$$\lambda = \max\Big\{-\chi_i + \sum s_j \chi_j < 0\Big\}. \tag{4.2}$$

The maximum exists since there are only finitely many values of $-\chi_i + \sum s_j \chi_j$ greater than any given number. Next we recall that $N + \alpha > \chi_1/\chi_\ell$ and set

$$\nu = \chi_1 - (N+\alpha)\chi_\ell > 0. \tag{4.3}$$

The proof of part (1) of Theorem 2.3 works for any $L \geq L(N,\alpha)$ and any $\varepsilon < \varepsilon_0$, where

$$\varepsilon_0 = \min\{\nu/(2L + 4(N+1+\alpha)),\ -\chi_\ell/(2NL+3),\ -\lambda/(N^2+N+1)\} > 0. \tag{4.4}$$

For parts (2) and (3) of Theorem 2.3 we will use smaller bounds on ε, namely, $\varepsilon_1 = \varepsilon_0/(N+1)$ and $\varepsilon_* = \varepsilon_0/3(N+1)$, respectively.

We fix $L \geq L(N,\alpha)$ and $0 < \varepsilon < \varepsilon_0$ and let $K = K_\varepsilon$ be as in (3.5). Since $\|\mathscr{F}_x\|_{C^{N,\alpha}}$ is ε-tempered, there is a function $C : \Lambda \to [1,\infty)$ such that for all $x \in \Lambda$ and $n \in \mathbb{N}$

$$\|\mathscr{F}_x\|_{C^{N,\alpha}} \leq C(x) \quad \text{and} \quad C(f^n x) \leq e^{n\varepsilon} C(x). \tag{4.5}$$

Similarly, replacing σ by a smaller function if necessary, we can assume that it satisfies

$$\sigma : \Lambda \to (0,1] \quad \text{and} \quad \sigma(f^n x) \geq e^{-n\varepsilon}\sigma(x). \tag{4.6}$$

LEMMA 4.1. *Under the assumptions of Theorem 2.3, there exists a function $\rho : \Lambda \to (0,1]$ so that for all $x \in \Lambda$, $n \in \mathbb{N}$, and $t \in B_{x,\rho(x)} \subset \mathscr{E}_x$, we have $\rho(x) < \sigma(x) \leq 1$ and*

(1) $\rho(f^n x) \geq e^{-\kappa\varepsilon n}\rho(x)$, where κ is given by (2.7),
(2) $\|D_t \mathscr{F}_x^n\|_{f^n x \leftarrow x} \leq e^{(\chi_\ell + 2\varepsilon)n}$,
(3) $\|D_t \mathscr{F}_x^n\| \leq K(x)\, e^{(\chi_\ell+2\varepsilon)n}$,
(4) $\|\mathscr{F}_x^n(t)\| \leq K(x)\, e^{(\chi_\ell+2\varepsilon)n}\|t\|$,
(5) $\|\mathscr{F}_x^n(t)\|_{f^n x} \leq e^{(\chi_\ell+2\varepsilon)n}\|t\|_x$.

Proof. We take $\beta = 1$ if $N \geq 2$ and $\beta = \alpha > 0$ if $N = 1$. For each $x \in \Lambda$ we define

$$\rho(x) = \sigma(x)[\varepsilon\, e^{\chi_\ell}(C(x)K(x)^2)^{-1}]^{1/\beta}. \tag{4.7}$$

Then (1) follows from (4.5), (4.6), and (3.6); (5) follows from (2) by the mean value theorem since $\mathscr{F}_x^n(0) = 0$. We prove (2), (3), and (4) by induction. The

statements are clear for $n=0$, suppose they hold for n. Note that (2) implies (3) by (3.7), and (3) implies (4). We observe that

$$\|D_t\mathscr{F}_x^{n+1}\|_{f^{n+1}x\leftarrow x} \le \|D_{t'}\mathscr{F}_{f^nx}\|_{f^{n+1}x\leftarrow f^nx}\cdot\|D_t\mathscr{F}_{f^nx}\|_{f^nx\leftarrow x}, \text{ where } t'=\mathscr{F}_x^n(t).$$

Then (2) follows from the inductive assumption and

$$\|D_t\mathscr{F}_{f^nx}\|_{f^{n+1}x\leftarrow f^nx} \le e^{\chi_\ell+2\varepsilon}. \tag{4.8}$$

To prove (4.8) we denote $\Delta = D_{t'}\mathscr{F}_{f^nx} - D_0\mathscr{F}_{f^nx}$. By the choice of β and (4.5), the β-Hölder constant of $D_s\mathscr{F}_{f^nx}$ at 0 is at most $C(f^nx)$, so using (3.7) we obtain

$$\|\Delta\|_{f^{n+1}x\leftarrow f^nx} \le K(f^{n+1}x)\|\Delta\| \le K(f^{n+1}x)C(f^nx)\|t'\|^\beta,$$

and using (4.5), (3.6) and the inductive assumption (4) we get that this is at most

$$K(x)\,C(x)\,e^{(2n+1)\varepsilon}K(x)^\beta\,e^{\beta(\chi_\ell+2\varepsilon)n}\,\|t\|^\beta \le e^\varepsilon\,C(x)\,K(x)^2\,e^{[2\varepsilon+\beta(\chi_\ell+2\varepsilon)]n}\,\|t\|^\beta.$$

Since $\|t\|\le\rho(x)$ and $\beta\chi_\ell+2(1+\beta)\varepsilon\le 0$ we obtain

$$\|\Delta\|_{f^{n+1}x\leftarrow f^nx} \le e^\varepsilon\,C(x)\,K(x)^2\,\rho(x)^\beta \le \varepsilon\,e^{\chi_\ell+\varepsilon}\sigma(x)^\beta \le \varepsilon\,e^{\chi_\ell+\varepsilon}.$$

Since

$$D_0\mathscr{F}_{f^nx} = F_{f^nx} \quad\text{and}\quad \|F_{f^nx}\|_{f^{n+1}x\leftarrow f^nx} \le e^{\chi_\ell+\varepsilon}$$

by (3.4), we conclude that

$$\begin{aligned}\|D_{t'}\mathscr{F}_{f^nx}\|_{f^{n+1}x\leftarrow f^nx} &\le \|\Delta\|_{f^{n+1}x\leftarrow f^nx} + \|F_{f^nx}\|_{f^{n+1}x\leftarrow f^nx}\\ &\le \varepsilon\,e^{\chi_\ell+\varepsilon} + e^{\chi_\ell+\varepsilon} \le e^{\chi_\ell+2\varepsilon}.\end{aligned}$$

□

4.1. **Construction of $\mathscr{P}$ and of the Taylor polynomial for $\mathscr{H}$.** For each $x\in\Lambda$ and map $\mathscr{F}_x:\mathscr{E}_x\to\mathscr{E}_{fx}$ we consider the Taylor polynomial at $t=0$:

$$\mathscr{F}_x(t) \sim \sum_{n=1}^{N} F_x^{(n)}(t). \tag{4.9}$$

As a function of t, $F_x^{(n)}(t):\mathscr{E}_x\to\mathscr{E}_{fx}$ is a homogeneous polynomial map of degree n. First we construct the Taylor polynomials at $t=0$ for the desired coordinate change $\mathscr{H}_x(t)$ and the polynomial extension $\mathscr{P}_x(t)$. We use similar notations for these Taylor polynomials:

$$\mathscr{H}_x(t) \sim \sum_{n=1}^{N} H_x^{(n)}(t) \quad\text{and}\quad \mathscr{P}_x(t) = \sum_{n=1}^{d} P_x^{(n)}(t).$$

For the first derivative we choose

$$H_x^{(1)} = \mathrm{Id}:\mathscr{E}_x\to\mathscr{E}_x \quad\text{and}\quad P_x^{(1)} = F_x \quad\text{for all } x\in\Lambda.$$

We will inductively construct the terms $H_x^{(n)}$ and $P_x^{(n)}$ for all in $x\in\Lambda$ so that $P_x^{(n)}$ is of sub-resonance type and they are measurable in x and $n^2\varepsilon$-tempered, i.e.,

$$\sup_{k\in\mathbb{N}} e^{-n^2\varepsilon k}\|H_{f^kx}^{(n)}\|_{f^kx\leftarrow f^kx} < \infty \quad\text{and}\quad \sup_{k\in\mathbb{N}} e^{-n^2\varepsilon k}\|P_{f^kx}^{(n)}\|_{f^{k+1}x\leftarrow f^kx} < \infty. \tag{4.10}$$

The base of the induction is the linear terms chosen above. Now we assume that the terms of order less than n are constructed. Using these notations in the conjugacy equation $\mathscr{H}_{fx} \circ \mathscr{F}_x = \mathscr{P}_x \circ \mathscr{H}_x$ we write

$$\left(\mathrm{Id} + \sum_{i=2}^{N} H_{fx}^{(i)}\right) \circ \left(F_x + \sum_{i=2}^{N} F_x^{(i)}\right) \sim \left(F_x + \sum_{i=2}^{d} P_x^{(i)}\right) \circ \left(\mathrm{Id} + \sum_{i=2}^{N} H_x^{(i)}\right).$$

Considering the terms of degree $N \geq n \geq 2$, we obtain

$$F_x^{(n)} + H_{fx}^{(n)} \circ F(x) + \sum H_{fx}^{(i)} \circ F_x^{(j)} = F_x \circ H_x^{(n)} + P_x^{(n)} + \sum P_x^{(j)} \circ H_x^{(i)},$$

where the summations are over all i and j such that $ij = n$ and $1 < i, j < n$. We rewrite the equation as

$$F_x^{-1} \circ P_x^{(n)} = -H_x^{(n)} + F_x^{-1} \circ H_{fx}^{(n)} \circ F_x + Q_x, \tag{4.11}$$

where

$$Q_x = F_x^{-1}\left(F_x^{(n)} + \sum_{ij=n,\ 1<i,j<n} H_{fx}^{(i)} \circ F_x^{(j)} - P_x^{(j)} \circ H_x^{(i)}\right). \tag{4.12}$$

We note that Q_x is composed only of terms $H^{(i)}$ and $P^{(i)}$ with $1 < i < n$, which are already constructed, and terms $F^{(i)}$ with $1 < i \leq n$, which are given. Thus by the inductive assumption Q_x is defined for all $x \in \Lambda$ and measurable. We will show later that they are also suitably tempered in x.

Let $\mathscr{R}_x^{(n)}$ be the space of all homogeneous polynomial maps on $\mathscr{E}_x$ of degree n, and let $\mathscr{S}_x^{(n)}$ and $\mathscr{N}_x^{(n)}$ be the subspaces of sub-resonance and non sub-resonance polynomials respectively. We seek $H_x^{(n)}$ so that the right side of (4.11) is in $\mathscr{S}_x^{(n)}$, and hence so is $P_x^{(n)}$ when defined by this equation.

Projecting (4.11) to the factor bundle $\mathscr{R}^{(n)}/\mathscr{S}^{(n)}$, our goal is to solve the equation

$$0 = -\bar{H}_x^{(n)} + F_x^{-1} \circ \bar{H}_{fx}^{(n)} \circ F_x + \bar{Q}_x, \tag{4.13}$$

where $\bar{H}^{(n)}$ and $\bar{Q}$ are the projections of $H^{(n)}$ and Q respectively.

We consider the bundle automorphism $\Phi : \mathscr{R}^{(n)} \to \mathscr{R}^{(n)}$ covering $f^{-1} : \mathcal{M} \to \mathcal{M}$ given by the maps $\Phi_x : \mathscr{R}_{fx}^{(n)} \to \mathscr{R}_x^{(n)}$

$$\Phi_x(R) = F_x^{-1} \circ R \circ F_x. \tag{4.14}$$

Since F preserves the splitting $\mathscr{E} = \mathscr{E}^1 \oplus \cdots \oplus \mathscr{E}^\ell$, it follows from the definition that the sub-bundles $\mathscr{S}^{(n)}$ and $\mathscr{N}^{(n)}$ are Φ-invariant. We denote by $\bar{\Phi}$ the induced automorphism of $\mathscr{R}^{(n)}/\mathscr{S}^{(n)}$ and conclude that (4.13) is equivalent to

$$\bar{H}_x^{(n)} = \tilde{\Phi}_x(\bar{H}_{fx}^{(n)}), \quad \text{where } \tilde{\Phi}_x(R) = \bar{\Phi}_x(R) + \bar{Q}_x. \tag{4.15}$$

Thus a solution of (4.13) is a $\tilde{\Phi}$-invariant section of $\mathscr{R}^{(n)}/\mathscr{S}^{(n)}$. We will show that $\tilde{\Phi}$ is a nonuniform contraction and that it has a unique measurable tempered invariant section. First, for polynomials of specific homogeneous type the exponent of Φ is determined by the exponents of F as follows.

LEMMA 4.2. *For a polynomial $R : \mathscr{E}_{fx} \to \mathscr{E}^i_{fx}$ of homogeneous type $s = (s_1, \dots, s_\ell)$ with $s_1 + \cdots + s_\ell = n$,*

$$\|\Phi_x(R)\|_{x\leftarrow x} \le e^{-\chi_i + \sum s_j\chi_j + (n+1)\varepsilon} \cdot \|R\|_{fx\leftarrow fx}. \tag{4.16}$$

Proof. Suppose that $v = v_1 + \cdots + v_\ell$, where $v_j \in \mathscr{E}^j_x$, and $\|v\|_x = 1$. We denote $a_j = \|F|_{\mathscr{E}^j_x}\|_{fx\leftarrow x}$ and observe that $F_x(v_j) = a_j v'_j \in \mathscr{E}^j_{fx}$ with $\|v'_j\|_{fx} \le \|v_j\|_x$. Since R has homogeneous type $s = (s_1, \dots, s_\ell)$ we obtain by (2.1) that

$$(R \circ F_x)(v) = R(a_1 v'_1 + \cdots + a_\ell v'_\ell) = a_1^{s_1} \cdots a_\ell^{s_\ell} \cdot R(v'_1 + \cdots + v'_\ell). \tag{4.17}$$

where $v' = v'_1 + \cdots + v'_\ell$ has $\|v'\|_{fx} \le \|v\|_x = 1$ by orthogonality of the splitting in the Lyapunov metric. Thus

$$\|(R \circ F_x)(v)\|_{fx} = a_1^{s_1} \cdots a_\ell^{s_\ell} \cdot \|R(v')\|_{fx} \le a_1^{s_1} \cdots a_\ell^{s_\ell} \cdot \|R\|_{fx\leftarrow fx}$$

for any $v \in \mathscr{E}_x$ with $\|v\|_x = 1$, so we obtain $\|R \circ F_x\|_{x\leftarrow fx} \le a_1^{s_1} \cdots a_\ell^{s_\ell} \cdot \|R\|_{fx\leftarrow fx}$ by definition (3.8). Now (3.9) yields

$$\begin{aligned}
\|\Phi_x(R)\|_{x\leftarrow x} &= \|F|^{-1}_{\mathscr{E}^i_x} \circ R \circ F_x\|_{x\leftarrow x} \\
&\le \|F|^{-1}_{\mathscr{E}^i_x}\|_{x\leftarrow fx} \cdot \|R \circ F_x\|_{x\leftarrow fx} \\
&\le \|F|^{-1}_{\mathscr{E}^i_x}\|_{x\leftarrow fx} \cdot a_1^{s_1} \cdots a_\ell^{s_\ell} \cdot \|R\|_{fx\leftarrow fx} \\
&\le e^{-\chi_i+\varepsilon} \cdot \prod_j (e^{\chi_j+\varepsilon})^{s_j} \cdot \|R\|_{fx\leftarrow fx}.
\end{aligned}$$

Since $a_j = \|F|_{\mathscr{E}^j_x}\|_{fx\leftarrow x} \le e^{\chi_j+\varepsilon}$ and $\|F|^{-1}_{\mathscr{E}^i_x}\|_{x\leftarrow fx} \le e^{-\chi_i+\varepsilon}$ by (3.3). □

REMARK 4.3. Similarly, one can show that $\|\Phi_x^{-1}(R)\|_{fx\leftarrow fx} \le e^{\chi_i - \sum s_j\chi_j + (n+1)\varepsilon} \cdot \|R\|_{x\leftarrow x}$. Since this holds for any $\varepsilon > 0$, using (3.10) to compare the Lyapunov and standard norms, one can conclude that the Lyapunov exponent of Φ on R is

$$\lim_{k\to\pm\infty} k^{-1} \log \|\Phi^k(R)\| = -\chi_i + \sum s_j\chi_j.$$

For all non sub-resonance homogeneous types we have $-\chi_i + \sum s_j\chi_j \le \lambda$ by the definition (4.2) of λ. Thus Lemma 4.2 yields the following lemma.

LEMMA 4.4. *The map $\Phi : \mathscr{N}^{(n)} \to \mathscr{N}^{(n)}$ given by (4.14) is a nonuniform contraction over f^{-1}, and hence so is $\tilde{\Phi} : \mathscr{R}^{(n)}/\mathscr{S}^{(n)} \to \mathscr{R}^{(n)}/\mathscr{S}^{(n)}$ given by (4.15). More precisely, $\|\Phi_x(R)\|_{x\leftarrow x} \le e^{\lambda+(n+1)\varepsilon} \cdot \|R\|_{fx\leftarrow fx}$.*

Proof. The statement about $\tilde{\Phi}$ follows since the linear part $\bar{\Phi}$ of $\tilde{\Phi}$ is given by Φ when $\mathscr{R}^{(n)}/\mathscr{S}^{(n)}$ is naturally identified with $\mathscr{N}^{(n)}$. By the choice of ε, we have $\lambda + (n+1)\varepsilon < 0$. □

It follows from the previous remark that λ is the maximal Lyapunov exponent of Φ over f^{-1} on the space of non sub-resonant polynomials and that all Lyapunov exponents of $\Phi|_{\mathscr{S}^{(n)}}$ are non-negative.

Now we construct a $\tilde{\Phi}$-invariant measurable section of $\mathscr{B} = \mathscr{R}^{(n)}/\mathscr{S}^{(n)}$ and study its properties. The construction is orbit-wise. We fix a point $x \in \Lambda$, consider its positive orbit $\{x_k = f^k x : k \geq 0\}$, and define the Banach space

$$\mathscr{B}^x = \{R = (R_k)_{k=0}^{\infty} : R_k \in \mathscr{B}_{x_k}, \|R\| < \infty\}, \text{ where } \|R\| = \sup_{k\geq 0} e^{-\varepsilon n^2 k} \|R_k\|_{x_k \leftarrow x_k}$$

and $\|R_k\|_{x_k \leftarrow x_k}$ is the norm induced on $\mathscr{B}_{x_k}$ by the Lyapunov norm $\|\cdot\|_{x_k}$ on $\mathscr{E}_x$. We denote $\tilde{Q} = (\bar{Q}_{x_k})_{k=0}^{\infty}$ and claim that it is in $\mathscr{B}^x$. For this we need to estimate the growth of the Lyapunov norm of (4.12) along the trajectory:

$$\begin{aligned} \|Q_{x_k}\|_{x_k \leftarrow x_k} \leq \|F_{x_k}^{-1}\|_{x_k \leftarrow x_{k+1}} \cdot (&\|F_{x_k}^{(n)}\|_{x_{k+1} \leftarrow x_k} \\ &+ \sum_{ij=n,\ 2\leq i,j\leq n/2} \|H_{x_{k+1}}^{(i)}\|_{x_{k+1} \leftarrow x_{k+1}} \|F_{x_k}^{(j)}\|_{x_{k+1} \leftarrow x_k}^{i} \\ &+ \|P_{x_k}^{(j)}\|_{x_{k+1} \leftarrow x_k} \|H_{x_k}^{(i)}\|_{x_k \leftarrow x_k}^{j}). \end{aligned} \tag{4.18}$$

First, $\|F_{x_k}^{-1}\|_{x_k \leftarrow x_{k+1}} \leq e^{-\chi_\ell + \varepsilon}$ for all x and k by (3.4). The exponential growth rate in k of $\|F_{x_k}^{(n)}\|_{x_{k+1} \leftarrow x_k}$ is at most 2ε. Indeed, using (3.10) and (3.6) we can obtain from (4.5) the corresponding estimate for $C^{N,\alpha}$ norm with respect to the Lyapunov metric on $\mathscr{E}_{x_k}$:

$$\|\mathscr{F}_{x_k}\|_{C^{N,\alpha},x_k} \leq K(fx_k)\|\mathscr{F}_{x_k}\|_{C^{N,\alpha}} \leq K(x_{k+1})C(x_k) \leq e^{(2k+1)\varepsilon} K(x)C(x). \tag{4.19}$$

Then using the inductive assumption (4.10) for the terms of order $i, j < n$, we can estimate the exponential growth rate of the two terms in the sum respectively as $(i^2+2i)\varepsilon$ and $(j^2+i^2 j)\varepsilon$, which are at most $((n/2)^2 + in)\varepsilon < n^2\varepsilon$. So the exponential growth rate of $\|Q_{x_k}\|_{x_k \leftarrow x_k}$ can be estimated by $n^2\varepsilon$ and thus $\|\tilde{Q}\| < \infty$.

Then $\tilde{\Phi}^x$ induces an operator on $\mathscr{B}^x$ by $(\tilde{\Phi}^x(R))_k = \bar{\Phi}_{x_k}(R_{k+1}) + \tilde{Q}_k$ and we have

$$\begin{aligned} \|\tilde{\Phi}^x(R) - \tilde{\Phi}^x(R')\| &= \sup_{k\geq 0} e^{-\varepsilon n^2 k} \|\bar{\Phi}_{x_k}(R_{k+1} - R'_{k+1})\|_{x_k \leftarrow x_k} \\ &\leq \sup_{k\geq 0} e^{-\varepsilon n^2 k} e^{\lambda + (n+1)\varepsilon} \|R_{k+1} - R'_{k+1}\|_{x_{k+1} \leftarrow x_{k+1}} \\ &\leq e^{\lambda + (n^2+n+1)\varepsilon} \cdot \left(\sup_{k\geq 0} e^{-\varepsilon n^2 (k+1)} \|(R_{k+1} - R'_{k+1})\|_{x_{k+1} \leftarrow x_{k+1}} \right) \\ &\leq e^{\lambda + (n^2+n+1)\varepsilon} \|R - R'\|. \end{aligned}$$

Since $\lambda + (n^2+n+1)\varepsilon < 0$ by the choice of ε (4.4), $\tilde{\Phi}^x$ is a contraction and thus has a unique fixed point $R^x \in \mathscr{B}^x$. We claim that $\bar{H}_x^{(n)} = R_0^x$ is a measurable function which is a unique solution of (4.15) or equivalently (4.13). Measurability follows from the fact that the fixed point can be explicitly written as a series

$$\bar{H}_x^{(n)} = \sum_{k=0}^{\infty} \left(F_x^k\right)^{-1} \circ \bar{Q}_{x_k} \circ F_x^k. \tag{4.20}$$

Invariance is clear since $\left(R_{k+1}^x\right)_{k=0}^{\infty}$ is a fixed point of $\tilde{\Phi}^{fx}$ which coincides with $\left(R_k^{fx}\right)_{k=0}^{\infty}$ by uniqueness and thus $R_1^x = R_0^{fx}$. More generally, $\bar{H}_{x_k}^{(n)} = R_0^{x_k} = R_k^x$,

and since $R^x \in \mathscr{B}^x$, the exponential growth rate of $\|\bar{H}^{(n)}_{x_k}\|_{x_k \leftarrow x_k}$ is at most $n^2\varepsilon$. Now we can choose $H^{(n)}_x$ as a lift of $\bar{H}^{(n)}_x$ to $\mathscr{R}^{(n)}_x$ which is measurable in x and satisfies (4.10). Then we define $P^{(n)}_x$ by equation (4.11). It also satisfies (4.10) as so do H and Q and as $\|F_x\|_{x \leftarrow fx}$ and $\|F_x^{-1}\|_{fx \leftarrow x}$ are uniformly bounded. This completes the inductive step and the construction of $H^{(n)}$ and $P^{(n)}$, $n = 1, \dots, N$, satisfying (4.10).

Thus we have constructed the N^{th} Taylor polynomial for the coordinate change

$$\mathscr{H}^N_x(t) = \sum_{n=1}^{N} H^{(n)}_x(t) \quad \text{of degree } \ N \geq d = \lfloor \chi_1/\chi_\ell \rfloor \tag{4.21}$$

and the polynomial map $\mathscr{P}_x(t) = \sum_{n=1}^{d} P^{(n)}_x(t)$.

4.2. **Construction of the coordinate change $\mathscr{H}$.** We rewrite the conjugacy equation $\mathscr{H}_{fx} \circ \mathscr{F}_x = \mathscr{P}_x \circ \mathscr{H}_x$ in the form

$$\mathscr{H}_x = \mathscr{P}_x^{-1} \circ \mathscr{H}_{fx} \circ \mathscr{F}_x. \tag{4.22}$$

A solution $\mathscr{H} = \{\mathscr{H}_x\}$ of this equation is a fixed point of the operator T given by

$$T(\mathscr{H})_x = \mathscr{P}_x^{-1} \circ \mathscr{H}_{fx} \circ \mathscr{F}_x. \tag{4.23}$$

We will find $\mathscr{H}$ in the form $\mathscr{H} = \mathscr{H}^N + R$, where $\mathscr{H}^N$ is given by (4.21). We denote

$$R = \mathscr{H} - \mathscr{H}^N \quad \text{and} \quad \tilde{T}(R) = T(\mathscr{H}^N + R) - \mathscr{H}^N \tag{4.24}$$

and observe that $T(\mathscr{H}) = \mathscr{H}$ if and only if $\tilde{T}(R) = R$. We will find R_x using the fixed point of a contraction $\tilde{T}^x$ induced by $\tilde{T}$ on a certain space $\mathscr{C}^x$ of sequences of functions along the orbit of x. Now we define the space $\mathscr{C}^x$.

By the construction of $\mathscr{H}^N$ and $\mathscr{P}$, $\mathscr{H}^N$ and $T(\mathscr{H}^N)$ have the same derivatives at the zero section up to order N, so we consider functions with vanishing derivatives at the zero section up to order N. First we describe the space of functions at each regular point x. For any $x \in \Lambda$ we denote by $B_{x,r}$ the ball centered at 0 in $\mathscr{E}_x$ of radius $r < \rho(x) < 1$ in the Lyapunov norm $\|\cdot\|_x$. We define

$$\mathscr{C}_{x,r} = \{R \in C^{N,\alpha}(B_{x,r}, \mathscr{E}_x) : \ D_0^{(k)} R = 0, \ k = 0, \dots, N\}.$$

Throughout this section we use the $C^{N,\alpha}$ norms *with respect to the Lyapunov metric* on $\mathscr{E}_x$. They are estimated through the norms for the standard metric (2.5) in (4.19). In particular, we use the α-Hölder constant (2.4) of $D^{(N)}R$ at 0 with respect to the Lyapunov metric, which for any $R \in \mathscr{C}_{x,r}$ is given by

$$\|D^{(N)}R\|_{x,\alpha} = \sup\{\|D^{(N)}_t R\|_{x \leftarrow x} \cdot \|t\|_x^{-\alpha} : \ 0 \neq t \in B_{x,r}\}. \tag{4.25}$$

For any $R \in \mathscr{C}_{x,r}$ lower derivatives can be estimated by the mean value theorem as

$$\|D^{(n)}_t R\|_{x \leftarrow x} \leq \|t\|_x^{N-n} \cdot \sup\{\|D^{(N)}_s R\|_{x \leftarrow x} : \ \|s\|_x \leq \|t\|_x\}, \tag{4.26}$$

so using the above Hölder constant we obtain that for any $0 \leq n < N$ and $t \in B_{x,r}$,

$$\|D^{(n)}_t R\|_{x \leftarrow x} \leq \|t\|_x^{1+\alpha} \cdot \|D^{(N)}R\|_{x,\alpha}. \tag{4.27}$$

Thus the norms of all derivatives are dominated by the Hölder constant and hence

$$\|R\|_{C^{N,\alpha}(B_{x,r})} = \|D^{(N)}R\|_{x,\alpha}. \tag{4.28}$$

It follows that $\mathscr{C}_{x,r}$ equipped with the norm $\|D^{(N)}R\|_{x,\alpha}$ is a Banach space.

We will choose a small $r = r(x) < \rho(x)$ satisfying (4.40) and set

$$r_k = re^{-2NLk\varepsilon}, \tag{4.29}$$

where L is given by (4.1). We define $\mathscr{C}^x$ as the following Banach space of sequences of functions along the orbit $x_k = f^k x$.

$$\mathscr{C}^x = \{\bar{R} = (R_k)_{k=0}^{\infty} : R_k \in \mathscr{C}_{x_k,r_k},\ \|\bar{R}\|_{\mathscr{C}^x} < \infty\}, \tag{4.30}$$

where

$$\|\bar{R}\|_{\mathscr{C}^x} = \sup\{e^{-Lk\varepsilon}\|D^{(N)}R_k\|_{x_k,\alpha} :\ k \geq 0\} \tag{4.31}$$

with the norm $\|\cdot\|_{x_k,\alpha}$ defined as in (4.25) and satisfying (4.28). We consider the operator $\tilde{T}^x$ induced by $\tilde{T}$ on $\mathscr{C}^x$:

$$(\tilde{T}^x(\bar{R}))_k = (\mathscr{P}_{x_k})^{-1} \circ (\mathscr{H}^N_{x_{k+1}} + R_{k+1}) \circ \mathscr{F}_{x_k} - \mathscr{H}^N_{x_k}. \tag{4.32}$$

Now we estimate the growth of $C^{N,\alpha}$ norms of $\mathscr{H}^N_x$ and $T(\mathscr{H}^N)_x = \mathscr{P}_x^{-1} \circ \mathscr{H}^N_{fx} \circ \mathscr{F}_x$ along the orbit to verify that $\tilde{T}^x(\bar{0})$ is in $\mathscr{C}^x$. We recall that

$$D_0^{(1)}(\mathscr{H}_{x_k}) = \mathrm{Id} \quad \text{and} \quad D_0^{(1)}(\mathscr{P}_{x_k}) = P^{(1)}_{x_k} = F_{x_k}$$

by the construction, and the latter satisfies

$$\|F_{x_k}\|_{x_{k+1} \leftarrow x_k} \leq e^{\chi_\ell + \varepsilon} \quad \text{and} \quad \|F^{-1}_{x_k}\|_{x_k \leftarrow x_{k+1}} \leq e^{-\chi_1 + \varepsilon}.$$

Also, for $2 \leq n \leq d$, the Lyapunov norms of $D_0^{(n)}(\mathscr{P}_{x_k}) = P^{(n)}_{x_k}$ and $D_0^{(n)}(\mathscr{H}_{x_k}) = H^{(n)}_{x_k}$ grow at most at the exponential growth rate $n^2\varepsilon$ in k by (4.10).

Recall that the inverse of $\mathscr{P}_{x_k}$ is also a sub-resonance polynomial $\mathscr{P}^{-1}_{x_k} = \sum_{n=1}^{d} (\mathscr{P}^{-1}_{x_k})^{(n)}$. We now show that the Lyapunov norms of $(\mathscr{P}^{-1}_{x_k})^{(n)}$ also grows at most at the exponential rate $n^2\varepsilon$ in k. First, its linear term $(\mathscr{P}^{-1}_{x_k})^{(1)} = (P^{(1)}_{x_k})^{-1} = F^{-1}_{x_k}$ has bounded Lyapunov norm. Inductively, we consider terms of order $n > 1$ in the equation $\mathscr{P} \circ \mathscr{P}^{-1} = \mathrm{Id}$ and obtain

$$\mathscr{P}^{(1)}_{x_k} \circ (\mathscr{P}^{-1}_{x_k})^{(n)} + \mathscr{P}^{(n)}_{x_k} \circ (\mathscr{P}^{-1}_{x_k})^{(1)} + \sum_{i,j<n,\, ij=n} \mathscr{P}^{(i)}_{x_k} \circ (\mathscr{P}^{-1}_{x_k})^{(j)} = 0.$$

The terms in the sum can be estimated as $\|\mathscr{P}^{(i)}_{x_k} \circ (\mathscr{P}^{-1}_{x_k})^{(j)}\| \leq \|\mathscr{P}^{(i)}_{x_k}\| \cdot \|(\mathscr{P}^{-1}_{x_k})^{(j)}\|^i$ and hence are $(i^2 + j^2 i)\varepsilon$-tempered by the inductive assumption. Since $i, j \leq \frac{n}{2}$ we obtain $i^2 + j^2 i = i^2 + ni \leq \frac{n^2}{4} + \frac{n^2}{2} < n$. Multiplying the equation by bounded $\left(P^{(1)}_{x_k}\right)^{-1}$ we conclude that $\left(\mathscr{P}^{-1}_{x_k}\right)^{(n)}$ is $n^2\varepsilon$-tempered, completing the induction.

Now we estimate $C^{N,\alpha}$ norms of polynomials $\mathscr{H}^N$ and $\mathscr{P}^{-1}$. For $\mathscr{H}^N$, the derivative of order N is constant $H^{(N)}_x$ on $\mathscr{E}_x$, and the lower derivatives on $B_{x,\rho(x)}$

can be inductively estimated similarly to (4.26),

$$\begin{aligned}\|D_t^{(N-1)}\mathscr{H}_x\|_{x\leftarrow x} &\le \|D_0^{(N-1)}\mathscr{H}_x\|_{x\leftarrow x} + \|t\|_x \|H_x^{(N)}\|_{x\leftarrow x}\\ &\le \|H_x^{(N-1)}\|_{x\leftarrow x} + \|H_x^{(N)}\|_{x\leftarrow x}\end{aligned}$$

yielding the same estimate of the exponential rate as for $H_x^{(N)}$,

$$\|\mathscr{H}_{x_k}\|_{C^{N,\alpha}(B\,x_k,\rho(x_k))} \le c_1(x)e^{N^2k\varepsilon} \quad \text{for all } k\ge 0. \tag{4.33}$$

Similarly for $\mathscr{P}_{x_k}^{-1}$ the derivative of order $d \le N$ is constant on $\mathscr{E}_{x_k}$, higher derivatives are zero, and the lower derivatives can be estimated as for $\mathscr{H}$, so we obtain

$$\|(\mathscr{P}_{x_k})^{-1}\|_{C^{N,\alpha}(B\,x_k,\rho(x_k))} \le c_2(x)e^{N^2k\varepsilon} \quad \text{for all } k\ge 0. \tag{4.34}$$

To obtain estimates for $(T(\mathscr{H}^N))_x = \mathscr{P}_x^{-1}\circ\mathscr{H}^N_{fx}\circ\mathscr{F}_x$ we use the following lemma.

LEMMA 4.5. *If Q is a polynomial of degree at most N and $\mathscr{F}$ is $C^{N,\alpha}$ then $Q\circ\mathscr{F}$ is $C^{N,\alpha}$ and $\|Q\circ\mathscr{F}\|_{C^{N,\alpha}} \le c_N\|Q\|_{C^N}\|\mathscr{F}\|^N_{C^{N,\alpha}} + \|Q\|_{C^0}$, where c_N depends on N only.*

Proof. Since Q is C^∞ it is clear that $Q\circ\mathscr{F}$ is C^N. For the N^{th} derivative we have

$$D_t^{(N)}(Q\circ\mathscr{F}) = D_{\mathscr{F}(t)}Q\circ D_t^{(N)}\mathscr{F} + \sum_{kj=N,\, j<N} D^{(k)}_{\mathscr{F}(t)}Q\circ D_t^{(j)}\mathscr{F}.$$

First we estimate α-Hölder constant at 0 of the first term. As DQ is linear, we get

$$\begin{aligned}&D_{\mathscr{F}(t)}Q\circ D_t^{(N)}\mathscr{F} - D_0Q\circ D_0^{(N)}\mathscr{F}\\ &\qquad = (D_{\mathscr{F}(t)}Q - D_0Q)\circ D_t^{(N)}\mathscr{F} + D_0Q\circ(D_t^{(N)}\mathscr{F} - D_0^{(N)}\mathscr{F})\end{aligned}$$

whose norm can be estimated by

$$\begin{aligned}&\|D_{\mathscr{F}(t)}Q - D_0Q\|\cdot\|D_t^{(N)}\mathscr{F}\| + \|D_0Q\|\cdot\|D_t^{(N)}\mathscr{F} - D_0^{(N)}\mathscr{F}\|\\ &\qquad \le \|Q\|_{C^2}\cdot\|\mathscr{F}(t)\|\cdot\|\mathscr{F}\|_{C^{N,\alpha}}\|Q\|_{C^1}\cdot\|\mathscr{F}\|_{C^{N,\alpha}}\cdot\|t\|^\alpha\\ &\qquad \le \|Q\|_{C^2}\cdot\|\mathscr{F}\|_{C^{N,\alpha}}\cdot\|\mathscr{F}\|_{C^1}\cdot\|t\|\,\|Q\|_{C^1}\cdot\|\mathscr{F}\|_{C^{N,\alpha}}\cdot\|t\|^\alpha.\end{aligned}$$

So the α-Hölder constant at 0 of $D_{\mathscr{F}(t)}Q\circ D_t^{(N)}\mathscr{F}$ is estimated by $2\|Q\|_{C^N}\|\mathscr{F}\|^2_{C^{N,\alpha}}$. The other terms in the sum are C^1 and hence are Lipschitz with constant bounded by supremum norms of their derivatives. These norms, along with the norms of lower derivatives of $Q\circ\mathscr{F}$ can be estimated as a sum of termss of the type

$$\|D^{(k)}_{\mathscr{F}(t)}Q\circ D_t^{(j)}\mathscr{F})\| \le \|D^{(k)}_{\mathscr{F}(t)}Q\|\cdot\|D_t^{(j)}\mathscr{F}\|^k \le \|Q\|_{C^N}\|\mathscr{F}\|^N_{C^{N,\alpha}}. \tag{4.35}$$

We conclude that $\|Q\circ\mathscr{F}\|_{C^{N,\alpha}} \le c_N\|Q\|_{C^N}\|\mathscr{F}\|^N_{C^{N,\alpha}} + \|Q\|_{C^0}$. □

Later we will also need a similar result for the case when Q is not a polynomial.

LEMMA 4.6. *If Q and $\mathscr{F}$ are $C^{N,\alpha}$, then $Q\circ\mathscr{F}$ is $C^{N,\alpha}$ and*

$$\|Q\circ\mathscr{F}\|_{C^{N,\alpha}} \le c''_N\|Q\|_{C^{N,\alpha}}\|\mathscr{F}\|^{N+\alpha}_{C^{N,\alpha}} + \|Q\|_{C^0},$$

where c''_N depends on N only.

Proof. The proof is the same as in Lemma 4.5 except that, since $D^{(N)}Q$ is only Hölder, we also need to estimate the α-Hölder constant at 0 of the term $D^{(N)}_{\mathscr{F}(t)}Q\circ D_t\mathscr{F}$ in

$$D^{(N)}_t(Q\circ\mathscr{F}) = D^{(N)}_{\mathscr{F}(t)}Q\circ D_t\mathscr{F} + D_{\mathscr{F}(t)}Q\circ D^{(N)}_t\mathscr{F} + \sum_{kj=N,\, j,k<N} D^{(k)}_{\mathscr{F}(t)}Q\circ D^{(j)}_t\mathscr{F}.$$

We consider

$$\begin{aligned} &D^{(N)}_{\mathscr{F}(t)}Q\circ D_t\mathscr{F} - D^{(N)}_0Q\circ D_0\mathscr{F}\\ &\quad = (D^{(N)}_{\mathscr{F}(t)}Q - D^{(N)}_0Q)\circ D_t\mathscr{F} + D^{(N)}_0Q\circ D^{(N)}_t\mathscr{F} - D^{(N)}_0Q\circ D^{(N)}_0\mathscr{F} \end{aligned}$$

and estimate its norm as

$$\begin{aligned} &\|Q\|_{C^{N,\alpha}}\|\mathscr{F}(t)\|^\alpha\cdot\|D_t\mathscr{F}\|^N + Lip\big(D^{(N)}_0Q\big)\cdot\|D_t\mathscr{F} - D_0\mathscr{F}\|\\ &\quad\le \|Q\|_{C^{N,\alpha}}\cdot(\|\mathscr{F}\|_{C^1}\|t\|)^\alpha\cdot\|\mathscr{F}\|^N_{C^1} + c'_N\|D^{(N)}_0Q\|\,\|\mathscr{F}\|^{N-1}_{C^1}\cdot\|\mathscr{F}\|_{C^{1,\alpha}}\cdot\|t\|^\alpha\\ &\quad\le \|t\|^\alpha\big(\|Q\|_{C^{N,\alpha}}\cdot\|\mathscr{F}\|^{N+\alpha}_{C^1} + c'_N\|Q\|_{C^N}\cdot\|\mathscr{F}\|^{N-1}_{C^1}\cdot\|\mathscr{F}\|_{C^{1,\alpha}}\big). \end{aligned}$$

Here we estimated the Lipschitz constant $Lip\big(D^{(N)}_0Q\big)$ of the homogeneous polynomial N-form $D^{(N)}_0Q$ on a ball of radius $R = \|\mathscr{F}\|_{C^1}$ by the supremum of its derivative on that ball, which is a homogeneous polynomial $(N-1)$-form whose norm can be estimated by $\|D^{(N)}_0Q\|$ with some constant c'_N depending on N only.

So the α-Hölder constant at 0 of $D^{(N)}_{\mathscr{F}(t)}Q\circ D_t\mathscr{F}$ is estimated by

$$\|Q\|_{C^{N,\alpha}}(\|\mathscr{F}\|^{N+\alpha}_{C^1} + c'_N\|\mathscr{F}\|^N_{C^{1,\alpha}}) \le (c'_N+1)\|Q\|_{C^{N,\alpha}}\|\mathscr{F}\|^{N+\alpha}_{C^{N,\alpha}}.$$

We conclude as in Lemma 4.5 that $\|Q\circ\mathscr{F}\|_{C^{N,\alpha}} \le c''_N\|Q\|_{C^{N,\alpha}}\|\mathscr{F}\|^{N+\alpha}_{C^{N,\alpha}} + \|Q\|_{C^0}$. □

We apply Lemma 4.5 with $Q = \mathscr{H}^N$ and then with $Q = \mathscr{P}^{-1}$. We conclude that $T(\mathscr{H}^N)$ is $C^{N,\alpha}$. Moreover, since $\|\mathscr{F}\|^N_{C^{N,\alpha}}$ is 2ε-tempered by (4.19), using (4.34) and (4.33) we can estimate the growth rate for $T(\mathscr{H}^N)$ by $N^2\varepsilon + N(N^2\varepsilon + N2\varepsilon)$ and obtain

$$\|T(\mathscr{H}^N)_k\|_{C^{N,\alpha}(Bx_k,\rho(x_k))} \le c_3(x)e^{(N^3+3N^2)k\varepsilon}. \tag{4.36}$$

Recall that $L \ge \max\{\kappa, N^3+3N^2+1\}$ by (4.1). Using $r < \rho(x)$, Lemma 4.1(1), and (4.29) we obtain that that for all $k \ge 0$

$$r_k = re^{-2NL\varepsilon k} < re^{-L\varepsilon k} < \rho(x)e^{-\kappa\varepsilon k} \le \rho(x_k). \tag{4.37}$$

Finally, $L > N^3+3N^2$ ensures that $\tilde{T}(\bar{0}) \in \mathscr{C}^x$ as $\|\tilde{T}(\bar{0})\|_{\mathscr{C}^x}$ is at most

$$\gamma' = \sup_{k\ge0} e^{-Lk\varepsilon}\big(\|T(\mathscr{H}^N)_k\|_{C^{N,\alpha}(Bx_k,\rho(x_k))} + \|\mathscr{H}^N\|_{C^{N,\alpha}(Bx_k,\rho(x_k))}\big) < \infty. \tag{4.38}$$

We recall that $\nu > 0$ is given by (4.3) and define

$$\theta = (1-e^{-\nu/2})/2, \quad 0<\theta<1, \quad \text{and} \quad \gamma = \gamma'/\theta. \tag{4.39}$$

We choose $r = r(x) < \rho(x) < 1$ satisfying

$$r \le \varepsilon/(2c_2(x)\gamma), \quad r < \rho(x)/(2\gamma), \quad r \le \theta/(c_5(x)\gamma^N e^{(M+L)\varepsilon}), \tag{4.40}$$

where θ, γ, $c_2(x)$, $c_5(x)$, M are given by (4.39), (4.34), (4.46).

We denote by $B^x(\gamma)$ the closed ball in $\mathscr{C}^x$ of radius γ. Our goal is to show that $\tilde{T}^x$ is a $(1-\theta)$-contraction and that $\tilde{T}^x(B^x(\gamma)) \subset B^x(\gamma)$. Since $\|\tilde{T}(\bar{0})\|_{\mathscr{C}^x} \le \gamma'$ it suffices to prove that the differential of $\tilde{T}^x$ at $\bar{R}$ is a $(1-\theta)$-contraction for each $\bar{R} \in B^x(\gamma)$.

First we check that the compositions in $\big(\tilde{T}(\bar{R})\big)_k$ are well-defined. We take $t \in B_{x_k,r_k}$ and show that $t' = \mathscr{F}_{x_k}(t)$ is in $B_{x_{k+1},r_{k+1}}$. Since by (4.37) t is in the ball $B_{x_k,\rho(x_k)}$ in standard metric, the estimates in Lemma 4.1 hold for any k. In particular, by (2),(5)

$$\|D_t^{(1)}\mathscr{F}_{x_k}\|_{x_{k+1}\leftarrow x_k} \le e^{\chi_\ell+2\varepsilon} \quad \text{and} \quad \|t'\|_{x_{k+1}} = \|\mathscr{F}_{x_k}(t)\|_{x_{k+1}} \le e^{\chi_\ell+2\varepsilon}\|t\|_{x_k}. \tag{4.41}$$

Since $\|t\|_{x_k} \le r_k = re^{-2NLk\varepsilon}$, this yields

$$\|t'\|_{x_{k+1}} \le e^{\chi_\ell+2\varepsilon} re^{-2NLk\varepsilon} < re^{-2NL(k+1)\varepsilon} = r_{k+1}, \tag{4.42}$$

since by the choice of ε we have $\chi_\ell + 2\varepsilon + 2NL\varepsilon < 0$.

Now we estimate $t'' = (\mathscr{H}^N_{x_{k+1}} + R_{k+1})(t')$ to show that it is in $B_{x_{k+1},\rho(x_{k+1})}$ where we have estimates for $(\mathscr{P}_{x_k})^{-1}$. Using the mean value theorem, (4.38), and the inequality $\gamma' < \gamma$ we obtain

$$\begin{aligned}\|\mathscr{H}^N_{x_{k+1}}(t')\|_{x_{k+1}} &\le \|t'\|_{x_{k+1}}\|\mathscr{H}^N_{x_{k+1}}\|_{C^1} \le \|t'\|_{x_{k+1}}\|\mathscr{H}^N_{x_{k+1}}\|_{C^{N,\alpha}(B\,x_{k+1},\rho(x_{k+1}))}\\ &\le r_{k+1}e^{L(k+1)\varepsilon}\gamma' \le re^{(1-2N)L(k+1)\varepsilon}\gamma \le r\gamma\, e^{-L(k+1)\varepsilon}\end{aligned}$$

as $2N-1 \ge 1$. Using (4.28) we obtain similarly that for any $\bar{R} \in B^x(\gamma)$,

$$\|R_{k+1}(t')\|_{x_{k+1}} \le \|t'\|_{x_{k+1}} \cdot \|D^{(N)}R_{k+1}\|_{x_{k+1},\alpha} \le r_{k+1}e^{L(k+1)\varepsilon} \cdot \|\bar{R}\|_{\mathscr{C}^x} \le r\gamma\, e^{-L(k+1)\varepsilon}.$$

Since $\rho(x_k) \ge \rho(x)e^{-\kappa\varepsilon k} \ge \rho(x)e^{-L\varepsilon k}$ and $2\gamma r < \rho(x)$ by (4.40), we obtain

$$\begin{aligned}\|t''\|_{x_{k+1}} = \|\big(\mathscr{H}^N_{x_{k+1}} + R_{k+1}\big)(t')\|_{x_{k+1}} &\le 2\gamma r e^{-L(k+1)\varepsilon}\\ &< \rho(x)e^{-L(k+1)\varepsilon} \le \rho(x_{k+1}).\end{aligned} \tag{4.43}$$

Now we show that $\tilde{T}^x$ is a contraction on $B^x(\gamma)$ by estimating its differential. For any $\bar{R}, \bar{S} \in B^x(\gamma)$ we can write

$$\begin{aligned}&\big(\tilde{T}^x(\bar{R}+\bar{S}) - \tilde{T}^x(\bar{R})\big)_k\\ &\quad = \big(\mathscr{P}_{x_k}\big)^{-1} \circ \big(\mathscr{H}^N_{x_{k+1}} + R_{k+1} + S_{k+1}\big) \circ \mathscr{F}_{x_k} - \big(\mathscr{P}_{x_k}\big)^{-1} \circ \big(\mathscr{H}^N_{x_{k+1}} + R_{k+1}\big) \circ \mathscr{F}_{x_k}.\end{aligned}$$

Differentiating $(\mathscr{P}_{x_k})^{-1}$ and denoting

$$y(t) = \big(\mathscr{H}^N_{x_{k+1}} + R_{k+1}\big)(\mathscr{F}_{x_k}(t)) \quad \text{and} \quad z(t) = S_{k+1}(\mathscr{F}_{x_k}(t))$$

we obtain

$$\big(\tilde{T}^x(\bar{R}+\bar{S}) - \tilde{T}^x(\bar{R})\big)_k(t) = D_{y(t)}(\mathscr{P}_{x_k})^{-1}\, z(t) + E(z(t)),$$

where E is a polynomial with terms of degree at least two. Thus $\|E(z(t))\|_{\mathscr{C}^x} = O(\|\bar{S}\|^2_{\mathscr{C}^x})$ and so the differential of $\tilde{T}^x$ is given by

$$\big([D_{\bar{R}}\tilde{T}^x]\bar{S}\big)_k(t) = D_{y(t)}(\mathscr{P}_{x_k})^{-1}\, S_{k+1}(\mathscr{F}_{x_k}(t)) = A_k(y(t))z(t),$$

where $A_k(s) = D_s(\mathscr{P}_{x_k})^{-1}$. To estimate the norm we consider the derivative of order N. Since $A_k(y(t))$ is a linear operator on z, the product rule yields

$$D^{(N)}[A_k(y(t))z(t)] = A_k(y(t))D^{(N)}z(t) + \sum c_{m,l}\, D^{(m)}A_k(y(t))D^{(l)}z(t), \tag{4.44}$$

where $m+l=N$ and $l<N$ for all terms in the sum. Differentiating $z(t)$ we get

$$D^{(l)}z(t) = D^{(l)}S_{k+1}(\mathscr{F}_{x_k}(t)) = \sum D^{(i)}_{t'}S_{k+1} \circ D^{(j)}_t \mathscr{F}_{x_k},$$

where $ij=l$ and $t' = \mathscr{F}_{x_k}(t)$. Only the first term in (4.44) contains $D^{(N)}S_{k+1}$ so

$$D^{(N)}_t\big([D_{\bar{R}}\tilde{T}^x]\bar{S}\big)_k = D^{(1)}_{y(t)}(\mathscr{P}_{x_k})^{-1} \circ D^{(N)}_{t'}S_{k+1} \circ D^{(1)}_t \mathscr{F}_{x_k} + J_k, \tag{4.45}$$

where J_k consists of a fixed number of terms of the type

$$D^{(m)}_t A_k(y(t))\big(D^{(i)}_{t'}S_{k+1} \circ D^{(j)}_t \mathscr{F}_{x_k}\big), \quad i<N,\ m+ij=N,$$

whose Lyapunov norms can be estimated by

$$\|A_k(y(t))\|_{C^N,x_k} \cdot \|D^{(i)}_{t'}S_{k+1}\|_{x_{k+1}} \cdot \|\mathscr{F}_{x_k}\|^{N-1}_{C^N,x_k}.$$

We use (4.19) to estimate the last term: $\|\mathscr{F}_{x_k}\|_{C^{N,\alpha},x_k} \le K(x)C(x)e^{(2k+1)\varepsilon}$. For the middle term by (4.27) we have as $i<N$

$$\|D^{(i)}_{t'}S_{k+1}\|_{x_{k+1}} \le \|t'\|^{1+\alpha}_{x_{k+1}} \cdot \|D^{(N)}S_{k+1}\|_{x_{k+1},\alpha} < \|t\|^{1+\alpha}_{x_k} \cdot \|\bar{S}\|_{\mathscr{C}^x} e^{L(k+1)\varepsilon}$$

since $\|t'\|_{x_{k+1}} \le e^{\chi_\ell+2\varepsilon}\|t\|_{x_k} < \|t\|_{x_k}$ by (4.41).

Since $(\mathscr{P}_{x_k})^{-1}$ is polynomial of degree at most N, using (4.34) we obtain

$$\begin{aligned}\|A_k\|_{C^N(B\,x_k,\rho(x_k))} &= \|D(\mathscr{P}_{x_k})^{-1}\|_{C^N(B\,x_k,\rho(x_k))}\\ &\le \|(\mathscr{P}_{x_k})^{-1}\|_{C^N(B\,x_k,\rho(x_k))}\\ &\le c_2(x)e^{N^2k\varepsilon}.\end{aligned}$$

Finally, since $y(t) = (\mathscr{H}^N_{x_{k+1}} + R_{k+1})(\mathscr{F}_{x_k}(t)) = t'' \in B_{x_{k+1},\rho(x_{k+1})}$ by (4.43), the first term in (4.44) can be estimated using Lemma 4.5 and equation (4.35)

$$\begin{aligned}\|A_k(y(t))\|_{C^N,x_k} &\le \|(\mathscr{P}_{x_k})^{-1}\|_{C^N(B\,x_k,\rho(x_k))} \cdot \|y(t)\|^N_{C^N,x_k}\\ &\le c_2(x)\, e^{N^2k\varepsilon} \cdot \Big(\|\mathscr{H}^N_{x_{k+1}} + R_{k+1}\|_{C^N,x_{k+1}} \cdot \|\mathscr{F}_{x_k}(t)\|^N_{C^N,x_k}\Big)^N\\ &\le c_2(x)\, e^{N^2k\varepsilon} \cdot \big(2\gamma e^{L(k+1)\varepsilon}\big)^N \cdot \big(K(x)C(x)e^{(2k+1)\varepsilon}\big)^{N^2}\\ &\le c_4(x)\,\gamma^N \cdot e^{(NL+3N^2)(k+1)\varepsilon},\end{aligned}$$

since we have $\|\bar{R}\|_{\mathscr{C}^x} \le \gamma$ and $\mathscr{H}^N_{x_{k+1}}$ term is estimated similarly from (4.38). Thus we obtain the following for the norm of J_k in (4.45):

$$\begin{aligned}\|J_k\| &< c_5(x)\,\gamma^N \cdot e^{(NL+3N^2)(k+1)\varepsilon} \cdot \|t\|^{1+\alpha}_{x_k} \cdot \|\bar{S}\|_{\mathscr{C}^x} \cdot e^{L(k+1)\varepsilon} \cdot e^{N(2k+1)\varepsilon}\\ &< c_5(x)\,\gamma^N r_k\, e^{(M+L)(k+1)\varepsilon} \cdot \|t\|^{\alpha}_{x_k} \cdot \|\bar{S}\|_{\mathscr{C}^x},\end{aligned} \tag{4.46}$$

where $M = NL+3N^2+2N$. For the first term in (4.45) we claim that for $t'' = y(t)$

$$\|D^{(1)}_{t''}(\mathscr{P}_{x_k})^{-1}\|_{x_k \leftarrow x_{k+1}} \le e^{-\chi_1+2\varepsilon}. \tag{4.47}$$

Indeed, we recall that

$$\|D_0^{(1)}(\mathscr{P}_{x_k})^{-1}\|_{x_k\leftarrow x_{k+1}} = \|F_{x_k}^{-1}\|_{x_k\leftarrow x_{k+1}} \le e^{-\chi_1+\varepsilon}.$$

If $d=1$, then $D^{(1)}(\mathscr{P}_{x_k})$ is constant and (4.47) follows. If $d\ge 2$ then $N\ge 2$ and the Lipschitz constant of $D^{(1)}(\mathscr{P}_{x_k})^{-1}$ is at most $c_2(x)e^{N^2k\varepsilon}$ by (4.34). So using (4.43) we obtain

$$\begin{aligned}\|D_{t''}^{(1)}(\mathscr{P}_{x_k})^{-1} - D_0^{(1)}(\mathscr{P}_{x_k})^{-1}\|_{x_k\leftarrow x_{k+1}} &\le c_2(x)e^{N^2k\varepsilon}\cdot\|t''\|_{x_{k+1}}\\ &\le c_2(x)e^{N^2k\varepsilon}\cdot 2\gamma r\, e^{-L(k+1)\varepsilon} \le c_2(x)2\gamma r\, e^{(N^2-L)(k+1)\varepsilon} < c_2(x)2\gamma r \le \varepsilon,\end{aligned}$$

where the last two inequalities hold since $N^2<L$ and $r\le \varepsilon/(2c_2(x)\gamma)$ by (4.40). Then (4.47) follows from $e^{-\chi_1+\varepsilon}+\varepsilon < e^{-\chi_1+\varepsilon}(1+\varepsilon) < e^{-\chi_1+2\varepsilon}$.

Now we estimate the main term in (4.45) using (4.47) and (4.41):

$$\begin{aligned}&\|D_{t''}^{(1)}(\mathscr{P}_{x_k})^{-1}\circ D_{t'}^{(N)}S_{k+1}\circ D_t^{(1)}\mathscr{F}_{x_k}\|_{x_k\leftarrow x_k}\\ &\quad\le \|D_{t''}^{(1)}(\mathscr{P}_{x_k})^{-1}\|_{x_k\leftarrow x_{k+1}}\cdot\|D^{(N)}S_{k+1}\|_{x_{k+1},\alpha}\,\|t'\|_{x_{k+1}}^{\alpha}\cdot\|D_t^{(1)}\mathscr{F}_{x_k}\|_{x_{k+1}\leftarrow x_k}^{N}\\ &\quad\le e^{-\chi_1+2\varepsilon}\cdot\|\bar S\|_{\mathscr{C}^x}e^{L(k+1)\varepsilon}\cdot e^{\alpha(\chi_\ell+2\varepsilon)}\|t\|_{x_k}^{\alpha}\cdot e^{N(\chi_\ell+2\varepsilon)}\\ &\quad= e^{-\nu+L'\varepsilon}\|t\|_{x_k}^{\alpha}\|\bar S\|_{\mathscr{C}^x}e^{Lk\varepsilon},\end{aligned}\tag{4.48}$$

where $\nu=-(N+\alpha)\chi_\ell+\chi_1>0$ and $L'=2+L+2(N+\alpha)$. Since $\varepsilon\le\varepsilon_0\le\nu/(2L')$ by the choice of ε_0, we obtain that $e^{-\nu+L'\varepsilon}\le e^{-\nu/2}=1-2\theta$ by (4.39).

Finally we estimate (4.45) combining (4.46) and (4.48). For any $\bar R\in B^x(\gamma)$ we have

$$\|D_t^{(N)}\big([D_{\bar R}\tilde T^x]\bar S\big)_k\|_{x_k\leftarrow x_k} \le \|t\|_{x_k}^{\alpha}\cdot\|\bar S\|_{\mathscr{C}^x}\cdot e^{Lk\varepsilon}\big(1-2\theta+c_5(x)\gamma^N r_k e^{((M+L)(k+1)-Lk)\varepsilon}\big).$$

Since $r_k = re^{-2NLk\varepsilon}$ and $2NL\ge M$, as $L\ge N^3+3N^2+1$, we see that for all $k\ge 0$

$$\begin{aligned}c_5(x)\gamma^N r_k\, e^{((M+L)(k+1)-Lk)\varepsilon} &\le c_5(x)\gamma^N\, re^{((M-2NL)k+M+L)\varepsilon}\\ &\le c_5(x)\gamma^N re^{(M+L)\varepsilon}\le\theta\end{aligned}$$

as $r\le\theta/(c_5(x)\gamma^N e^{(M+L)\varepsilon})$ by (4.40). Then for all $\bar R\in B^x(\gamma)$ we obtain

$$\|D_t^{(N)}\big([D_{\bar R}\tilde T^x]\bar S\big)_k\|_{x_k\leftarrow x_k} \le \|t\|_{x_k}^{\alpha}\cdot\|\bar S\|_{\mathscr{C}^x}\cdot e^{Lk\varepsilon}(1-\theta),$$

hence

$$\|D^{(N)}\big([D_{\bar R}\tilde T^x]\bar S\big)_k\|_{x_k,\alpha} \le (1-\theta)\cdot\|\bar S\|_{\mathscr{C}^x}\cdot e^{Lk\varepsilon},$$

and so

$$\|[D_{\bar R}\tilde T^x]\bar S\|_{\mathscr{C}^x} = \sup_k e^{-Lk\varepsilon}\|D^{(N)}(\tilde T^x(\bar S))_k\|_{x_k,\alpha} \le (1-\theta)\cdot\|\bar S\|_{\mathscr{C}^x}.$$

Thus $\|D_{\bar R}\tilde T^x\|\le 1-\theta$ for all $\bar R\in B^x(\gamma)$. Since $\|\tilde T^x(0)\|_{\mathscr{C}^x}\le\gamma'=\theta\gamma$, the operator $\tilde T^x$ is a contraction from $B^x(\gamma)$ to $B^x(\gamma)$. Thus $\tilde T^x$ has a unique fixed point $\bar R^x\in B^x(\gamma)$ which depends measurably on x. As in the construction of Taylor coefficients, the uniqueness implies that $R_x := (\bar R^x)_0$ is $L\varepsilon$-tempered and solves the equation $\tilde T(R)=R$ where $\tilde T$ is given by (4.24). We conclude that the measurable family of $C^{N,\alpha}$ maps $\mathscr{H}_x=\mathscr{H}_x^N+R_x$, is $L\varepsilon$-tempered and satisfies (4.23), i.e.,

conjugates $\mathscr{P}_x$ and $\mathscr{F}_x$. Then the maps $\mathscr{H}_x$ defined on $B_{x,r(x)}$ can be uniquely extended to $C^{N,\alpha}$ diffeomorphisms on $B_{x,\rho(x)}$ by the invariance

$$\mathscr{H}_x(t) = \left(\mathscr{P}_x^k\right)^{-1} \circ \mathscr{H}_{f^k x} \circ \mathscr{F}_x^k(t)$$

since for each $t \in B_{x,\rho(x)}$ we have $\mathscr{F}_x^k(t) \in B_{x_k,r_k}$ for some k. Indeed, $\mathscr{F}_x^k(t)$ is contracted by Lemma 4.1(5) at a faster rate than r_k by the choice (4.4) of ε_0: $\chi_\ell + 2\varepsilon < -2NL\varepsilon$.

This completes the proof of the first part of the theorem.

4.3. **Prove of part (2): "uniqueness" of $\mathscr{H}$.** This essentially follows from the "uniqueness" of the construction. Starting with $\mathscr{H}_1 = \tilde{\mathscr{H}}$ we inductively construct coordinate changes $\mathscr{H}_k = \{\mathscr{H}_{k,x}\}$ for $k = 1, \dots, N$ and show that they satisfy the same temperedness condition as $\mathscr{H}$. We denote their Taylor series by

$$\mathscr{H}_{k,x}(t) = \sum_{n=1}^{\infty} H_{k,x}^{(n)}(t).$$

The base of the induction is $\mathscr{H}_1 = \tilde{\mathscr{H}}$, which is tempered by the assumption and whose linear term satisfies $H_{1,x}^{(1)} = H_x^{(1)} = \mathrm{Id}$. Suppose $\mathscr{H}_{k-1}$, $k \geq 2$, is constructed so that

$$H_{k-1,x}^{(n)} \text{ are } n^2\varepsilon\text{-tempered for } n = 1, \dots, N, \quad H_x^{(n)} = H_{k-1,x}^{(n)} \text{ for } n = 1, \dots, k-1,$$

and the corresponding normal form $\mathscr{P}_{k-1,x}$ is of sub-resonance type. It follows that $\mathscr{P}$ and $\mathscr{P}_{k-1}$ have the same terms up to order $k-1$. Hence $H_{k-1,x}^{(k)}$ and $H_x^{(k)}$ satisfy the same equation (4.13) when projected to the factor-bundle $\mathscr{R}^{(k)}/\mathscr{S}^{(k)}$. Indeed, the Q term defined by (4.12) is composed only of $F^{(i)}$ and terms $H^{(i)}$ and $P^{(i)}$ with $1 < i \leq k-1$, which are the same for $\mathscr{H}_{k-1}$ and $\mathscr{H}$. By uniqueness we obtain that

$$H_x^{(k)} = H_{k-1,x}^{(k)} + \Delta_x^{(k)}, \text{ where } \Delta_x^{(k)} \in \mathscr{S}_x^{(k)}.$$

Then the coordinate change $\mathscr{H}_{k,x} = \left(\mathrm{Id} + \Delta_x^{(k)}\right) \circ \mathscr{H}_{k-1,x}$ has the same Taylor terms as $\mathscr{H}$ up to order k, and, since the polynomial $\mathrm{Id} + \Delta_x^{(k)}$ is in G_χ, $\mathscr{H}_k$ conjugates $\mathscr{F}$ to a sub-resonance normal form $\mathscr{P}_{k,x} = \left(\mathrm{Id} + \Delta_{fx}^{(k)}\right) \circ \mathscr{P}_{k-1,x} \circ \left(\mathrm{Id} + \Delta_x^{(k)}\right)^{-1}$. To complete the inductive step we need to show that $\|H_{k,x}^{(n)}\|$ is $n^2\varepsilon$-tempered. It suffices to show this for $\|R^{(n)}\|$ where $R = \mathscr{H}_{k,x} - \mathscr{H}_{k-1,x} = \Delta_x^{(k)} \circ \mathscr{H}_{k-1,x}$. Since $\Delta_x^{(k)}$ is homogeneous of degree k, R has only homogeneous terms of degrees $n = jk$. We estimate them as

$$\|R^{(n)}\| = \|\Delta_x^{(k)} \circ \mathscr{H}_{k-1,x}^{(j)}\| \leq \|\Delta_x^{(k)}\| \cdot \|\mathscr{H}_{k-1,x}^{(j)}\|^k,$$

which is $(k^2 + j^2 k)\varepsilon$-tempered by the inductive assumption and the definition of $\Delta_x^{(k)}$. Since $j \leq n/2$ as $k \geq 2$ we get $j^2 k = jn \leq n^2/2$. Also, if $j \geq 2$ then $k^2 \leq n^2/4$, and we obtain $n^2\varepsilon$-temperedness. If $j = 1$, then $n = k$ and $R^{(k)} = \Delta^{(k)}$ is also $k^2\varepsilon$-tempered.

Thus in N steps we obtain the coordinate change

$$\mathscr{H}_{N,x} = G_x \circ \tilde{\mathscr{H}}_x, \text{ where } G_x = \left(\mathrm{Id} + \Delta_x^{(N)}\right) \circ \ldots \circ \left(\mathrm{Id} + \Delta_x^{(2)}\right) \in G_x^\chi,$$

which has the same Taylor terms at 0 as $\mathscr{H}$ up to order N. In fact, for $n > d$ we have $\mathscr{S}^{(n)} = 0$ and hence $\Delta^{(n)} = 0$, so that $\mathscr{H}_N = \mathscr{H}_d$. Now we show that $\mathscr{H} = \mathscr{H}_N$, which also proves the last statement in part (2) of the theorem. The equality follows from the uniqueness in the final step of the construction. Indeed, for $\mathscr{H}^N$ given by (4.21), both differences $R = \mathscr{H} - \mathscr{H}^N$ and $R' = \mathscr{H}_N - \mathscr{H}^N$ are fixed points of operator $\tilde{T}$ given by (4.24). Hence $R = R'$ by uniqueness of the fixed point in the appropriate space $\mathscr{C}_{r,x}$ on which $\tilde{T}$ induces a contraction. To ensure that the sequence (R'_{x_k}) is in $\mathscr{C}_{r,x}$ we need to estimate temperedness of α-Hölder constant at 0 for $\mathscr{H}_N^{(N)}$. As above one can see that all terms in the polynomial G_x are $N^2\varepsilon$-tempered. Then using Lemma 4.5 and the assumption on $\tilde{\mathscr{H}}$ we obtain that $\|\mathscr{H}_{N,x}\|_{C^{N,\alpha}}$ is $\tilde{L}\varepsilon$-tempered for $\tilde{L} = (N^2 + NL) < (N+1)L$ and hence (R'_{x_k}) is in $\mathscr{C}_{r,x}$ with $\tilde{L}$ in place of L. Since the proof of part (1) is for any $L \geq L(N,\alpha)$, we conclude that $\tilde{T}$ induces a contraction in such $\mathscr{C}_{r,x}$ provided that $\varepsilon < \varepsilon_1 = \varepsilon_0/(N+1)$, which is less than ε_0 with $\tilde{L}$ in pace of L in (4.4). Thus $R = R'$ and hence $\mathscr{H} = \mathscr{H}_N$.

4.4. **Proof of Corollary 2.4.** By part (2) of Theorem 2.3, if we fix a choice of Taylor polynomials of degree d for $\mathscr{H}_x$, then the family $\mathscr{H}_x$ is unique. Then for each $N > d$ we can do the construction in part (1) with this fixed choice of Taylor polynomials and obtain the family of C^N diffeomorphisms $\mathscr{H}_x$. By uniqueness, all these families coincide and hence $\mathscr{H}_x$ are C^∞ diffeomorphisms.

4.5. **Proof of part (3): Centralizer of $\mathscr{H}$.** First we prove that the derivative of $\mathscr{G}$ at zero section, $\Gamma_x = D_0\mathscr{G}_x$, is sub-resonance. Since Γ_x is linear, this is equivalent to the fact that Γ_x preserves the fast flag associated with the Lyapunov splitting

$$\mathscr{E}^1_x = \mathscr{V}^1_x \subset \mathscr{V}^2_x \subset \cdots \subset \mathscr{V}^l_x = \mathscr{E}_x, \quad \text{where } \mathscr{V}^i_x = \mathscr{E}^1_x \oplus \cdots \oplus \mathscr{E}^i_x. \tag{4.49}$$

Suppose to the contrary that for some $x \in \Lambda$ and some $i < j$ we have a vector t in $\mathscr{E}^i_x$ such that $t' = \Gamma_x(t)$ has nonzero component t'_j in $\mathscr{E}^j_{gx}$. Then

$$\|(F^n_{gx} \circ \Gamma_x)(t)\|_{f^n gx} \geq \|F^n_{gx}(t'_j)\|_{f^n gx} \geq e^{(\chi_j - \varepsilon)n}\, \|t'_j\|_{gx}$$

and on the other hand

$$\begin{aligned} \|(F^n_{gx} \circ \Gamma_x)(t)\|_{f^n gx} &= \|\Gamma_{f^n x}(F^n_x\, t)\|_{gf^n x} \\ &\leq \|\Gamma_{f^n x}\|_{gf^n x \leftarrow f^n x} \cdot e^{(\chi_i + \varepsilon)n}\, \|t\|_x \leq C e^{(\chi_i + 3\varepsilon)n}, \end{aligned}$$

which is impossible for large n since ε is small. Here we used the fact that the $C^{N,\alpha}$ norm $\|\mathscr{G}_x\|_{C^{N,\alpha},x}$ on $\mathscr{E}_x$ is 2ε-tempered with respect to the Lyapunov metric (3.2) for F. This follows as in (4.19) since $\|\mathscr{G}_x\|_{C^{N,\alpha}}$ in standard norm is ε-tempered by assumption.

We conclude that Γ_x is sub-resonance for each $x \in \Lambda$. Now we consider a new family of coordinate changes

$$\tilde{\mathscr{H}}_x = \Gamma_x^{-1} \circ \mathscr{H}_{gx} \circ \mathscr{G}_x$$

which also satisfies $\tilde{\mathcal{H}}_x(0) = 0$ and $D_0\tilde{\mathcal{H}}_x = \mathrm{Id}$. A direct calculation shows that

$$\begin{aligned}\tilde{\mathcal{H}}_{fx} \circ \mathcal{F}_x \circ \tilde{\mathcal{H}}_x^{-1} &= \Gamma_{fx}^{-1} \circ \mathcal{H}_{fgx} \circ \mathcal{G}_{fx} \circ \mathcal{F}_x \circ \mathcal{G}_x^{-1} \circ \mathcal{H}_{gx}^{-1} \circ \Gamma_x \\ &= \Gamma_{fx}^{-1} \circ \mathcal{H}_{fgx} \circ \mathcal{F}_{gx} \circ \mathcal{H}_{gx}^{-1} \circ \Gamma_x = \Gamma_{fx}^{-1} \circ \mathcal{P}_{gx} \circ \Gamma_x = \tilde{\mathcal{P}}_x,\end{aligned}$$

where $\tilde{\mathcal{P}}_x$ is a sub-resonance polynomial as a composition of sub-resonance polynomials. Now we would like to to apply the uniqueness part of the theorem, which would give $\tilde{\mathcal{H}}_x = G_x\mathcal{H}_x$ for some tempered function $G_x \in G_\chi$. Then it follows from the definition of $\tilde{\mathcal{H}}_x$ that

$$\mathcal{H}_{gx} \circ \mathcal{G}_x = \Gamma_x \circ \tilde{\mathcal{H}}_x = (\Gamma_x G_x) \circ \mathcal{H}_x$$

so that $\mathcal{H}_{gx} \circ \mathcal{G}_x \circ \mathcal{H}_x^{-1} = \Gamma_x G_x$, which is again a sub-resonance polynomial, as claimed.

To complete the proof it remains to show that $\tilde{\mathcal{H}}_x$ is suitably tempered to obtain uniqueness. The n^{th} Taylor term at 0, $\tilde{\mathcal{H}}_x^{(n)}$, is the sum of the terms of the form $\Gamma_x^{-1} \circ \mathcal{H}_{gx}^{(k)} \circ \mathcal{G}_x^{(j)}$ with $n = kj$, whose Lyapunov norms as we can estimate as before

$$\|\Gamma_x^{-1} \circ \mathcal{H}_{gx}^{(k)} \circ \mathcal{G}_x^{(j)}\|_{x \leftarrow x} \le \|\Gamma_x^{-1}\|_{x \leftarrow gx} \cdot \|\mathcal{H}_{gx}^{(k)}\|_{gx \leftarrow gx} \cdot \|\mathcal{G}_x^{(j)}\|_{gx \leftarrow x}^k.$$

Thus we obtain that $\tilde{\mathcal{H}}_x^{(n)}$ is $m\varepsilon$-tempered with $m \le 2 + k^2 + 2k < 3n^2$ for $n \ge 2$. Since $\|\mathcal{H}\|_{C^{N,\alpha}}$ is $L\varepsilon$-tempered, using Lemma 4.6 with $Q = \mathcal{H}$ and $\mathcal{F} = \mathcal{G}$ we obtain that $\|\mathcal{H} \circ \mathcal{G}\|_{C^{N,\alpha}}$ is $(L + 2(N+\alpha))\varepsilon$-tempered. Then Lemma 4.5 implies that $\|\tilde{\mathcal{H}}\|_{C^{N,\alpha}}$ is $(2 + L + 2(N+\alpha))\varepsilon$-tempered and hence $3L\varepsilon$-tempered since $L \ge N + 2$. So the uniqueness result in part (2) of the theorem applies for $\varepsilon < \varepsilon_* = \varepsilon_1/3 = \varepsilon_0/3(N+1)$.

This completes the proof of Theorem 2.3. □

5. PROOF OF THEOREM 2.5

5.1. **Proof of (i), (ii), (iii), (v).** We will apply Theorem 2.3. First we note that the integrability condition for the derivative in Theorem 2.3 was used in the proof only to obtain the Lyapunov splitting and the Lyapunov metric. So while the restriction $Df|_{\mathcal{E}}$ may not satisfy this integrability condition, the Lyapunov splitting and the Lyapunov metric are obtained in this case from the results for the full differential Df.

The centralizer part (v) will follow directly from part (3) of Theorem 2.3 since $X' = \bigcap_{n\in\mathbb{Z}} g^n(X)$ is the desired invariant set of full measure as g preserves the measure class of μ. Moreover, $g(W_x) = W_{gx}$ since g is a diffeomorphism commuting with f, so that X' is also saturated by the stable manifolds.

Parts (i), (ii), (iii) essentially follow from Theorem 2.3, which is formulated so as to apply to this setting. First we consider the regular set Λ for (Df, μ). We fix a family of local (strong) stable manifolds $W_{x,r(x)}$ for $x \in \Lambda$ of sufficiently small size $r(x)$. Identifying $W_{x,r(x)}$ by an exponential map with a neighborhood of 0 in $\mathcal{E}_x$ we obtain the extension $\mathcal{F} = \{\mathcal{F}_x\}$ of f. Then the properties of local stable manifolds ensure that $\mathcal{F}$ satisfies the assumptions of Theorem 2.3. Indeed, they are given by $C^{N,\alpha}$ embeddings so that the $C^{N,\alpha}$ norm and $1/r(x)$ are ε-tempered

for any $\varepsilon > 0$ (see [3] for a general reference and [16, Theorem 5] for a convenient statement of the stable manifold theorem). Hence Theorem 2.3 yields existence of the desired family of local diffeomorphisms $\mathscr{H}_x$, $x \in \Lambda$, which can be uniquely extended by invariance

$$\mathscr{H}_x(t) = \left(\mathscr{P}_x^k\right)^{-1} \circ \mathscr{H}_{f^k x} \circ f^k(t)$$

to the global stable leaf W_x, which consists of those $t \in \mathcal{M}$ for which $f^k t$ is in the local stable leaf of $f^k x$ for some k. Now we define $X = \bigcup_{x\in\Lambda} W_x$ and explain the construction of $\mathscr{H}_y$ for any $y \in X$. By iterating it forward we may assume that $y \in W_{x,r(x)}$. While the individual Lyapunov spaces $\mathscr{E}^i$ may not be defined for all points $y \in W_{x,r(x)}$, the flag $\mathcal{V}$ of fast subspaces (4.49) is defined for each $\mathscr{E}_y = T_y W_{x,r(x)}$; moreover, the subspaces $\mathcal{V}_y^i$ depend Hölder continuously, and in fact $C^{N-1,\alpha}$, on y along W_x [24, Theorem 6.3].

The key observation is that the notion of sub-resonance polynomial depends only on the fast flag $\mathcal{V}$ [20, Proposition 3.2], not on the individual Lyapunov spaces $\mathscr{E}^i$, and thus is well-defined for $\mathscr{E}_y$. Then the sub-bundle $\mathscr{S}^{(n)}$ of sub-resonance polynomials of degree n is well-defined, invariant under Df, and Hölder continuous in y along W, and hence so is the factor bundle $\mathscr{R}^{(n)}/\mathscr{S}^{(n)}$. Then for each $y \in W_{x,r(x)}$ we can define $\mathscr{H}_y$ using the construction in Theorem 2.3. Indeed, first we constructed the Taylor term of degree n using the contraction $\tilde{\Phi}$ on the bundle $\mathscr{R}^{(n)}/\mathscr{S}^{(n)}$ from Lemma 4.4 with linear part estimated as $\|\Phi_x(R)\|_{\varepsilon,x} \le e^{\lambda+(n+1)\varepsilon} \cdot \|R\|_{\varepsilon,fx}$. Then Φ_y, the corresponding map at y, is Hölder close to Φ_x. We note that since $W_{x,r(x)}$ are $C^{N,\alpha}$ embedded, the derivatives $F_y^{(n)} = D_0^{(n)}\mathscr{F}_y$ of all orders $n \le N$ depend α-Hölder continuously on y in $W_{x,r(x)}$. In fact, the linear operator Φ_y depends only on the first derivative. Using the Lyapunov norm at x as the reference norm, we obtain that Φ_y is also a contraction with similar estimate for all $y \in W_{x,r(x)}$ provided that $r(x)$ is sufficiently small. Since $f^k y \in W_{f^k x, r(f^k x)}$ by the contraction property of $W_{x,r(x)}$, the closeness persists along the forward trajectory. This argument is similar to the proof of Lemma 4.1. Then we obtain that the operator $\tilde{\Phi}_y$ on the sequence space is also a contraction. Thus we can define $\tilde{\mathscr{H}}_y^{(n)}$ as before using the unique fixed point in the space of sequences. The last step of the construction can be carried out similarly as it involves only the estimates of the derivatives on the full space $\mathscr{E}$ and does not depend on the splitting. This completes the proof of (i), (ii), (iii).

REMARK 5.1. Any measurable choice of transversals $\tilde{\mathscr{E}}^i$ to $\mathcal{V}^{i-1}$ inside $\mathcal{V}^i$, with $i = 2,\dots,\ell$, yields a transversal $\tilde{\mathscr{N}}^{(n)}$ to $\mathscr{S}^{(n)}$ inside $\mathscr{R}^{(n)}$. The latter gives a preferred choice of the lift. The fixed point of the contraction $\tilde{\mathscr{H}}_y^{(n)}$ depends Hölder continuously, and even smoothly by appropriate C^r section theorem as in [20], on y along $W_{x,r(x)}$ if the same holds for the data $\tilde{Q}$ obtained in the previous step of the construction. To complete the inductive step we need a Hölder lift $\mathscr{H}_y^{(n)}$ to $\mathscr{R}^{(n)}$. If there is a consistent choice which is Hölder on the full leaves of W, then we can obtain a family $\{\mathscr{H}_x\}$ which is Hölder along the

leaves of W. In contrast to the uniform setting of [20], it is not clear that such a choice exists. However, this can be done locally on $W_{x,r(x)}$. Therefore, one can fix a Ledrappier-Young partition subordinate to the leaves of W [21, Def. 1.4.1 and Lemma 3.1.1] and obtain Hölder continuity of $\mathscr{H}_x$ on each element.

5.2. **Consistency of the fast foliations.** The leaf W_x is subfoliated by unique foliations U^k tangent to $\mathcal{V}^k_y$. We denote by $\bar{W}^k$ the corresponding foliations of $\mathscr{E}_x$ obtained by the identification $\mathscr{H}_x : W_x \to \mathscr{E}_x$. Thus we obtain the foliations $\bar{W}^k$ of $\mathscr{E}$ which are invariant under the polynomial extension $\mathscr{P}$. Since the maps $\mathscr{H}_x$ are diffeomorphisms, $\bar{W}^k$ are also the unique fast foliations with the same contraction rates. They are characterized by

$$\text{for } y,z\in\mathscr{E}_x,\quad z\in\bar{W}^k(y) \text{ if and only if } \lim_{n\to\infty}\frac{1}{n}\log\operatorname{dist}(\mathscr{P}^n_x(y),\mathscr{P}^n_x(z))<\chi_{k+1}.$$

It follows from Definition 2.2 that sub-resonance polynomials $R\in\mathscr{S}_{x,y}$ are *block triangular* in the sense that $\mathscr{E}^i$ component does not depend on $\mathscr{E}^j$ components for $j<i$ or, equivalently, it maps the subspaces $\mathcal{V}^i_x$ of the fast flag in $\mathscr{E}_x$ to those in $\mathscr{E}_y$.

It is easy to see that all derivatives of a sub-resonance polynomial are sub-resonance polynomials. In particular, the derivative $D_y\mathscr{P}_x$ at *any* point $y\in\mathscr{E}_x$ is sub-resonance and hence is block triangular. Thus it maps subspaces parallel to $\mathcal{V}^k_x$ to subspaces parallel to $\mathcal{V}^k_{fx}$. Hence the foliation of $\mathscr{E}$ by subspaces parallel to $\mathcal{V}^k_x$ in $\mathscr{E}_x$ is invariant under the extension $\mathscr{P}$ and hence coincides with $\bar{W}^k$ by uniqueness of the fast foliation.

REMARK 5.2. This implies that the fast subfoliations U^k are as smooth along the leaf W_x as the diffeomorphism $\mathscr{H}_x$ which maps them to linear subfoliations of $\mathscr{E}_x$.

It follows that for any $x\in X$ and any $y\in W_x$ the diffeomorphism

$$\mathscr{G}_{x,y}=\mathscr{H}_y\circ\mathscr{H}_x^{-1}:\mathscr{E}_x\to\mathscr{E}_y \tag{5.1}$$

maps the fast flag of linear foliations of $\mathscr{E}_x$ to that of $\mathscr{E}_y$. In particular, the same holds for its derivative $D_0\mathscr{G}_{x,y}=D_x\mathscr{H}_y:\mathscr{E}_x\to\mathscr{E}_y$ and we conclude that $D_0\mathscr{G}_{x,y}$ is block triangular and thus is a sub-resonance linear map.

5.3. **Proof of (iv): Consistency of normal form coordinates.** We need to show that the map $\mathscr{G}_{x,y}$ in (5.1) is a sub-resonance polynomial map for all $x\in X$ and $y\in W_x$. It suffices to consider $x\in\Lambda$ and, using invariance, we may assume that $y\in W_x$ is sufficiently close to x. First we note that

$$\mathscr{G}_{x,y}(0)=\mathscr{H}_y(x)=:\bar{x}\in\mathscr{E}_y\quad\text{and}\quad D_0\mathscr{G}_{x,y}=D_x\mathscr{H}_y.$$

Since $\mathscr{H}^{-1}_{f^nx}\circ\mathscr{P}^n_x\circ\mathscr{H}_x=f^n=\mathscr{H}^{-1}_{f^ny}\circ\mathscr{P}^n_y\circ\mathscr{H}_y$ we obtain that

$$\mathscr{H}_{f^ny}\circ\mathscr{H}^{-1}_{f^nx}\circ\mathscr{P}^n_x=\mathscr{H}_{f^ny}\circ f^n\circ\mathscr{H}_x^{-1}=\mathscr{P}^n_y\circ\mathscr{H}_y\circ\mathscr{H}_x^{-1}$$

and hence

$$\mathscr{G}_{f^nx,f^ny}\circ\mathscr{P}^n_x=\mathscr{P}^n_y\circ\mathscr{G}_{x,y}. \tag{5.2}$$

Now we consider the Taylor polynomial for $\mathscr{G}_{x,y}:\mathscr{E}_x\to\mathscr{E}_y$ at $t=0\in\mathscr{E}_x$:

$$\mathscr{G}_{x,y}(t)\sim G_{x,y}(t)=\bar{x}+\sum_{m=1}^{N}G_{x,y}^{(m)}(t).$$

Our first goal is to show that all its terms are sub-resonance polynomials. We proved in Section 5.2 that the first derivative $G_{x,y}^{(1)}=D_x\mathscr{H}_y$ is a sub-resonance linear map.

Inductively, we assume that $G_{x,y}^{(m)}$ has only sub-resonance terms for $m=1,\dots,k-1$ and show that the same holds for $G_{x,y}^{(k)}$. Suppose for the contrary that $G_{x,y}^{(k)}$ is not a sub-resonance polynomial and consider order k terms in the Taylor polynomial at $0\in\mathscr{E}_x$ for (5.2). The Taylor polynomial for $\mathscr{P}_x^n$ at 0 coincides with itself, $\mathscr{P}_x^n(t)=\sum_{m=1}^{d}P_x^{(m)}(t)$. We also consider the Taylor polynomial for $\mathscr{P}_y^n$ at $\mathscr{G}_{x,y}(0)=\bar{x}\in\mathscr{E}_y$:

$$\mathscr{P}_y^n(z)=\bar{x}_n+\sum_{m=1}^{d}Q_y^{(m)}(z-\bar{x}),\quad\text{where }\bar{x}_n=\mathscr{P}_y^n(\bar{x}).$$

All terms $Q^{(m)}$ are sub-resonance as the derivatives of a sub-resonance polynomial. Consider the Taylor polynomial for

$$\mathscr{G}_{f^nx,f^ny}(t)\sim G_{f^nx,f^ny}(t)=\bar{x}_n+\sum_{m=1}^{N}G_{f^nx,f^ny}^{(m)}(t).$$

Now we obtain from (5.2) the coincidence of the terms up to degree N in

$$\bar{x}_n+\sum_{j=1}^{N}G_{f^nx,f^ny}^{(j)}\left(\sum_{m=1}^{d}P_x^{(m)}(t)\right)\sim\bar{x}_n+\sum_{m=1}^{d}Q_y^{(m)}\left(\sum_{j=1}^{N}G_{x,y}^{(j)}(t)\right).$$

Since any composition of sub-resonance polynomials is again sub-resonance, the inductive assumption gives that all terms of order k in the above equation must be sub-resonance polynomials except for

$$G_{f^nx,f^ny}^{(k)}\left(P_x^{(1)}(t)\right)\quad\text{and}\quad Q_y^{(1)}\left(G_{x,y}^{(k)}(t)\right).$$

Multiplying these terms on the left by sub-resonance linear map $\left(D_0\mathscr{G}_{f^nx,f^ny}\right)^{-1}=\left(D_{f^nx}\mathscr{H}_{f^ny}\right)^{-1}$ and using the fact that $P_x^{(1)}=F_x^n=Df^n|_{\mathscr{E}_x}$ and

$$Q_y^{(1)}=D_{\bar{x}}\mathscr{P}_y^n=D_{f^nx}\mathscr{H}_{f^ny}\circ F_x^n\circ(D_x\mathscr{H}_y)^{-1}$$

we obtain that the following maps from $\mathscr{E}_x$ to $\mathscr{E}_{f^nx}$ agree modulo sub-resonance terms:

$$\left(\left(D_{f^nx}\mathscr{H}_{f^ny}\right)^{-1}\circ G_{f^nx,f^ny}^{(k)}\right)\circ F_x^n\cong F_x^n\circ\left((D_x\mathscr{H}_y)^{-1}\circ G_{x,y}^{(k)}\right)\mod\mathscr{S}_{x,f^nx}.$$

Since $x,f^nx\in\Lambda$ and thus the spaces $\mathscr{E}_x$ and $\mathscr{E}_{f^nx}$ have Lyapunov splittings, we can decompose these polynomial maps into sub-resonance and non sub-resonance terms. Taking non sub-resonance terms on both sides we obtain the equality

$$N_{f^nx}\circ F_x^n=F_x^n\circ N_x\tag{5.3}$$

where $N_{f^n x}$ and N_x denote the *non sub-resonance terms* in $\left(D_{f^n x}\mathscr{H}_{f^n y}\right)^{-1} \circ G^{(k)}_{f^n x, f^n y}$ and $(D_x\mathscr{H}_y)^{-1} \circ G^{(k)}_{x,y}$ respectively. If the latter had only sub-resonance terms then so would $G^{(k)}_{x,y}$, contradicting the assumption. Hence $N_x \neq 0$. We decompose N_x into components $N_x = (N_x^1, \dots, N_x^\ell)$ and let i be the largest index so that $N_x^i \neq 0$, i.e., there exists $t' \in \mathscr{E}_x$ so that $z' = N(t')$ has non-zero component in $\mathscr{E}^i_y$, which we denote by z'_i. Then by (3.3) we obtain

$$\|F_x^n \circ N_x(t')\|_{f^n x} = \|F_x^n(z')\|_{f^n x} \geq e^{n(\chi_i - \varepsilon)}\|z'_i\|_x. \tag{5.4}$$

Now we estimate the norm of the i component of the left-hand side of (5.3) at t'. For each componet t'_j of t' we have $\|F_x^n(t'_j)\|_{f^n x} \leq e^{n(\chi_j+\varepsilon)}\|t'_j\|_x$ by (3.3). Let $N^s_{f^n x}$ be a term of homogeneity type $s = (s_1, \dots, s_\ell)$ in the component $N^i_{f^n x}$. Then we obtain as in Lemma 4.2 that

$$\|N^s_{f^n x}\left(F_x^n(t')\right)\|_{f^n x} \leq \|N_{f^n x}\|_{f^n x} \cdot \|t'\|_x^k \cdot e^{n\sum s_j(\chi_j+\varepsilon)}. \tag{5.5}$$

For non sub-resonance N^s we have $\chi_i > \sum s_j\chi_j$ and hence (5.5) decays faster than (5.4). Since there are no sub-resonance terms in $N^i_{f^n x}$, this contradicts (5.3) for large n if ε is sufficiently small since $\|N_{f^n x}\|_{f^n x}$ is tempered. The latter follows from temperedness of $G^{(k)}_{f^n x, f^n y}$ and the fact that $D_{f^n x}\mathscr{H}_{f^n y}$ is Hölder close to the identity and so the norm of its inverse is bounded in Lyapunov metric.

We conclude that for all $x \in X$ and $y \in W_x$ the Taylor polynomial $G_{x,y}$ of $\mathscr{G}_{x,y}$ contains only sub-resonance terms. Now we will show that $\mathscr{G}_{y,x}$ coincides with its Taylor polynomial. Again it suffices to consider $x \in \Lambda$ and $y \in W_x$ which is sufficiently close to x. In addition to (5.2) we have the same relation for their Taylor polynomials

$$G_{f^n y, f^n x} \circ \mathscr{P}_y^n = \mathscr{P}_x^n \circ G_{y,x}. \tag{5.6}$$

Indeed, the two sides must have the same terms up to order N, but these are sub-resonance polynomials and thus have no terms of degree higher than $d \leq N$.

Denoting $\Delta_n = \mathscr{G}_{f^n y, f^n x} - G_{f^n y, f^n x}$ we obtain from (5.2) and (5.6) that

$$\Delta_n \circ \mathscr{P}_y^n = \mathscr{P}_x^n \circ \mathscr{G}_{y,x} - \mathscr{P}_x^n \circ G_{y,x}. \tag{5.7}$$

We denote $\Delta = \mathscr{G}_{y,x} - G_{y,x} : \mathscr{E}_y \to \mathscr{E}_x$ and suppose that $\Delta \neq 0$. Let i be the largest index for which the i component of Δ is nonzero. Then there exist arbitrarily small $t' \in \mathscr{E}_y$ such that the i component z'_i of $z' = \Delta(t')$ is nonzero. Since $\mathscr{P}_x^n$ is a sub-resonance polynomial, the nonlinear terms in its i component can depend only on j components of the input with $j > i$, which are the same for $\mathscr{G}_{y,x}$ and $G_{y,x}$ by the choice of i. Thus the i component of the right side of (5.7) is $F_x^n(z'_i)$ since the linear part of $\mathscr{P}_x^n$ is F_x^n and it preserves the Lyapunov splitting. So by (3.3) we can estimate the right side of (5.7)

$$\begin{aligned}\|\left(\mathscr{P}_x^n \circ \mathscr{G}_{y,x} - \mathscr{P}_x^n \circ G_{y,x}\right)(t')\|_{f^n x} &\geq \|F_x^n(z'_i)\|_{f^n x} \\ &\geq e^{n(\chi_i-\varepsilon)}\|z'_i\|_x \geq e^{n(\chi_1-\varepsilon)}\|z'_i\|_x.\end{aligned} \tag{5.8}$$

Now we estimate the left side of (5.7). Since $\mathscr{G}_{f^n y, f^n x}$ is $C^{N,\alpha}$ there exists $C_n(x)$ determined by $\|\mathscr{G}_{f^n x, f^n y}\|_{C^{N,\alpha}}$ such that

$$\|\Delta_n(t)\| \le C_n(x) \cdot \|t\|^{N+\alpha} \tag{5.9}$$

for all sufficiently small $t \in \mathscr{E}_{f^n x}$. To estimate $\mathscr{P}_y^n$ we note that $D_0\mathscr{P}_y^n = F_y^n = Df^n|_{\mathscr{E}_y}$ and its norm for y close to x can be estimated using Lemma 4.1(3). Then $\mathscr{P}_y^n$ itself can be estimated as in that lemma:

$$\|\mathscr{P}_y^n(t)\| \le Ke^{n(\chi_\ell+3\varepsilon)}\|t\|$$

for all sufficiently small $t \in \mathscr{E}_y$. Combining this with (5.9) we obtain

$$\|\big(\Delta_n \circ \mathscr{P}_y^n\big)(t')\| \le C_n(x) \cdot \|\mathscr{P}_y^n(t')\|^{N+\alpha} \le C_n(x) \cdot (K\|t'\|)^{N+\alpha} e^{n(N+\alpha)(\chi_\ell+3\varepsilon)}.$$

This contradicts (5.7) and (5.8) for large n if ε is sufficiently small. Indeed $(N+\alpha)\chi_\ell < \chi_1$ while $C_n(x)$ is tempered and the Lyapunov norm satisfies $\|u\| \ge K(x)e^{-n\varepsilon}\|u\|_{f^n x}$. Thus, $\Delta = 0$, i.e., the map $\mathscr{G}_{y,x}$ coincides with its Taylor polynomial.

This completes the proof of Theorem 2.5. □

5.4. **Proof of Corollary 2.6.** If $d = 1$ then all sub-resonance polynomials are linear, the maps $\mathscr{H}_y \circ \mathscr{H}_x^{-1} : \mathscr{E}_x \to \mathscr{E}_y$ are affine, and the family $\{\mathscr{H}_x\}_{x\in X}$ is unique by part (2) of Theorem 2.3. If we identify W_x with $\mathscr{E}_x$ by $\mathscr{H}_x$, then $\mathscr{H}_y$ for $y \in W_x$ becomes an affine map $\mathscr{E}_x \to T_y\mathscr{E}_x$ with identity differential and $\mathscr{H}_y(y) = 0$. Thus it depends C^N on y as the coordinate system $\mathscr{H}_x$ is C^N. □

Acknowledgments. We thank the referees for careful reading of the paper and useful suggestions.

REFERENCES

[1] L. Arnold, *Random Dynamical Systems*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 1998.

[2] L. Arnold and X. Kedai, Normal forms for random diffeomorphisms, *J. of Dynamics and Differential Equations,* **4** (1992), 445–483.

[3] L. Barreira and Ya. Pesin, *Nonuniformly Hyperbolicity: Dynamics of Systems with Nonzero Lyapunov Exponents*, Encyclopedia of Mathematics and Its Applications, 115, Cambridge Univ. Press, Cambridge, 2007.

[4] I. U. Bronstein and A. Ya. Kopanskii, *Smooth Invariant Manifolds and Normal Forms*, World Scientific, 1994.

[5] Y. Fang, On the rigidity of quasiconformal Anosov flows, *Ergodic Theory Dynam. Systems,* **27** (2007), 1773–1802.

[6] Y. Fang, P. Foulon and B. Hasselblatt, Zygmund strong foliations in higher dimension, *J. of Modern Dynamics,* **4** (2010), 549–569.

[7] R. Feres, A differential-geometric view of normal forms of contractions, in *Modern Dynamical Systems and Applications*, Cambridge University Press, 2004, 103–121.

[8] D. Fisher, B. Kalinin and R. Spatzier, Totally nonsymplectic Anosov actions on tori and nilmanifolds, *Geometry and Topology,* **15** (2011), 191–216.

[9] A. Gogolev, B. Kalinin and V. Sadovskaya, Local rigidity for Anosov automorphisms, (With appendix by R. de la Llave) *Math. Research Letters,* **18** (2011), 843–858.

[10] M. Guysinsky, The theory of non-stationary normal forms, *Ergodic Theory Dynam. Systems*, **22** (2002), 845–862.

[11] M. Guysinsky and A. Katok, Normal forms and invariant geometric structures for dynamical systems with invariant contracting foliations, *Math. Research Letters,* **5** (1998), 149–163.
[12] B. Kalinin and A. Katok, Invariant measures for actions of higher rank abelian groups, *Proceedings of Symposia in Pure Mathematics,* **69** (2001), 593–637.
[13] B. Kalinin and A. Katok, Measure rigidity beyond uniform hyperbolicity: Invariant measures for Cartan actions on tori, *J. of Modern Dynamics,* **1** (2007), 123–146.
[14] B. Kalinin, A. Katok and F. Rodriguez-Hertz, Nonuniform measure rigidity, *Annals of Math.,* **174** (2011), 361–400.
[15] A. Katok and J. Lewis, Local rigidity for certain groups of toral automorphisms, *Israel J. Math.,* **75** (1991), 203–241.
[16] A. Katok and F. Rodriguez-Hertz, Arithmeticity and topology of higher rank actions of Abelian groups, *J. of Modern Dynamics,* **10** (2016), 135–172.
[17] A. Katok and R. Spatzier, Differential rigidity of Anosov actions of higher rank abelian groups and algebraic lattice actions, *Tr. Mat. Inst. Steklova,* **216** (1997), Din. Sist. i Smezhnye Vopr., 292–319; translation in *Proc. Steklov Inst. Math.,* **216** (1997), 287–314.
[18] B. Kalinin and V. Sadovskaya, On local and global rigidity of quasiconformal Anosov diffeomorphisms, *J. Inst. Math. Jussieu,* **2** (2003), 567–582.
[19] B. Kalinin and V. Sadovskaya, Global rigidity for totally nonsymplectic Anosov $\mathbb{Z}^k$ actions, *Geometry and Topology,* **10** (2006), 929–954.
[20] B. Kalinin and V. Sadovskaya, Normal forms on contracting foliations: Smoothness and homogeneous structure, *Geometriae Dedicata,* **183** (2016), 181–194.
[21] F. Ledrappier and L.-S. Young, The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula, *Annals of Math.,* **122** (1985), 509–539.
[22] W. Li and K. Lu, Sternberg theorems for random dynamical systems, *Communications on Pure and Applied Mathematics,* **58** (2005), 941–988.
[23] K. Melnick, Nonstationary smooth geometric structures for contracting measurable cocycles, `http://www.math.umd.edu/~kmelnick/`.
[24] D. Ruelle, Ergodic theory of differentiable dynamical systems, *Publications Mathématiques de l'I.H.É.S.,* **50** (1979), 27–58.
[25] V. Sadovskaya, On uniformly quasiconformal Anosov systems, *Math. Research Letters,* **12** (2005), 425–441.
[26] S. Sternberg, Local contractions and a theorem of Poincaré, *Amer. J. of Math.,* **79** (1957), 809–824.

BORIS KALININ <`kalinin@psu.edu`>: Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA

VICTORIA SADOVSKAYA <`sadovskaya@psu.edu`>: Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA