COMPUTING THE INVARIANT CIRCLE AND ITS STABLE MANIFOLDS FOR A 2-D MAP BY THE PARAMETERIZATION METHOD:
EFFECTIVE ALGORITHMS AND RIGOROUS PROOFS OF CONVERGENCE

YIAN YAO AND RAFAEL DE LA LLAVE

ABSTRACT. We present and analyze rigorously a quadratically convergent algorithm to compute an invariant circle for 2-dimensional maps along with the corresponding foliation by stable manifolds. The algorithm is based on solving an invariance equation using a quasi-Newton method.

We prove that when the algorithm starts from an initial guess that satisfies the invariance equation very approximately (depending on some condition numbers, evaluated on the approximate solution), then the algorithm converges to a true solution which is close to the initial guess. The convergence is faster than exponential in smooth norms.

We also conclude that (in a smooth norm), the distance from the exact solution and the approximation is bounded by the initial error. This allows validating the numerical approximations (a-posteriori results). It also implies the usual persistence formulations since the exact solutions of the invariance equation for a model are approximate solutions for a similar model.

The algorithm we present works irrespective of whether the dynamics on the invariant circle is a rotation or it is phase-locked. The condition numbers required do not involve any global qualitative properties of the map. They are obtained by evaluating derivatives of the initial guess, derivatives of the map in a neighborhood of the guess, performing algebraic operations and taking suprema.

The proof of the convergence is based on a general Nash-Moser implicit function theorem specially tailored for this problem. The Nash-Moser procedure has unusual properties. As it turns out, the regularity requirements are not very severe (only 2 derivatives suffice). We hope that this implicit function theorem may be of independent interest and have presented it in a self-contained appendix.

The algorithm in this paper is very practical since it converges quadratically, and it requires moderate storage and operation count. Details of the implementation and results of the runs are described in a companion paper [YdlL21].

keywords: invariant circles, isochrons, parameterization method, Nash-Moser implicit function theorem, phase-locked regions [2021] 37M22, 37M21, 37D10, 47J07, 37C86, 46-08

1. INTRODUCTION

In the modern theory of dynamical systems, the study of the invariant manifold and their corresponding stable manifolds plays a key role. The dynamics on these objects organize the dynamics in the whole phase space.

In this paper, we study attractive (or repulsive) invariant circles in 2-dimensional maps as well as the stable (unstable) manifolds of points. The collection of such manifolds forms a foliation in a neighborhood of the torus.

We recall that according to the theory of normally hyperbolic manifolds [Fen72, Pes04], W_x^s, the stable manifolds of a point x in the invariant circle, are the points whose orbits converge *with a fast enough exponential rate* to the orbit of x.

Remark 1. The paper [Win75] defined isochrons as the set of points with the same asymptotic phase on the limit cycle. This is not equivalent to the stable manifolds in the sense of normally hyperbolic theory. In the theory of normally hyperbolic manifolds, the stable leaves are characterized by a fast enough convergence to the limit cycle.

When the dynamics in the invariant circle contains an attractive and a repelling periodic orbit (which are attractive and hyperbolic for the full map), the points in the plane whose orbit is asymptotic to the stable periodic orbit includes an open set. On the other hand, the stable manifolds in the sense of normally hyperbolic theory will be one dimensional manifolds. At the periodic orbit, the stable manifold in the sense of normally hyperbolic theory is the "strong stable manifold" in the theory of invariant manifolds at fixed points.

In this the paper, for the sake of having manageable sentences, we will occasionally use "isochron" to mean *" leaves of the foliation by the stable manifolds in the sense of normally hyperbolic theory"*.

The 2-dimensional maps we consider appear in several applications. For example, as reductions of higher dimensional systems to two-dimensional manifolds after a Neimark-Sacker bifurcation [Sac64, RT71, MM76]. Another case that motivates us is the periodic perturbation of a 2-D ordinary differential equation with a limit cycle. Such examples are very common in practice. For example, when oscillating circuits with a limit cicle are subject to AC forcing [AVK87, Min62] or in Biology when the circadian rhythms are subject to external forcing [Win01]. Also when neurons are subject to the periodic forcing of others [Izh07, ET10]

The interpretation of periodic forcing of limit cycles is useful to keep in mind since the methods we apply are inspired by those in [HdlL13], which considered limit cycles and their manifolds in 2D autonomous ODEs. As in [HdlL13], our goal will be to find a system of coordinates that turns the dynamics in a neighborhood of the limit cycle into a simple one. We will take advantage of several identities to obtain a fast quasi-Newton method.

Remark 2. Passing from 2-D differential equations to 2-D maps (or 3-D differential equations) is non-trivial since new dynamical phenomena appear. The most notorious one is that, for 2-D maps, the dynamics in the invariant circle could be phase-locked. That is, the dynamics restricted to the invariant circle could have an attracting periodic orbit and a repelling one.

Similarly, passing from 2-D maps to 3-D maps involves the new phenomenon of normal resonances, which is briefly discussed in [YdlL21]).

From a more technical point of view in the study of 2-D maps, we do not expect that the invariant circle or the foliation by stable manifolds of points are analytic but only finitely differentiable even if the map is analytic (in this paper, we will consider only analytic mappings) See Section 8.2 of [dlL97] and later in this paper.

On the other hand, each of the stable manifolds of a point will be shown to be analytic. This anisotropic regularity of the parameterizations of the foliation by stable manifolds – one of the unknowns in the invariance equation – has to be taken into account when choosing the spaces for the formulation of the implicit function theorem. It also affects the choices of discretizations in the implementations discussed in the companion paper [YdlL21]. Anisotropic regularity is very typical in the theory of Normally Hyperbolic Invariant Manifolds (NHIM).

It so happens that the NHIM has a regularity limited by ratios of rates of convergence while the stable manifolds of a point have a regularity limited only by the regularity of the map. This anisotropic regularity does not happen in the 2-D ODE case. In [HdlL13] it is shown that for 2D analytic ODE, both the circle and the foliation by stable leaves are analytic. The anisotropic regularity is an important novelty going from 2D ODE to 2D maps.

The goal of this paper is to provide a framework to study these objects (invariant circles and their stable foliations) in 2-D maps in a non-perturbative way which also leads to reliable abd efficient numerical algorithms. The numerical algorithms we present and justify here converge to the true solution faster than exponentially. Hence, mathematical results presented here also allow us to validate the results of the numerical algorithms.

The proof of the convergence of the algorithm is based on an abstract implicit function theorem of Nash-Moser type with some differences from other similar theorems, but which we hope could be useful for several problems in dynamics and related areas. See Section 1.2 for some comparison with other hard implicit function theorems in the literature. The algorithm is based on taking advantage of several cancellations that allow to get better estimates. It is shown in [YdlL21] that the same cancellations that allow to get better estimates, also allow to lower the storage requirements of the algorithm and the operations needed for a step. The storage requirements and the operation count per step are proportional to the number of discretization points.

The numerical algorithms described here have been implemented. Details on the implementation, some numerical results and investigation of phenomena that happen at the boundary of validity of our results are described in a companion paper [YdlL21].

1.1. **Description of the Method.** Following the idea of the parameterization method [CFdlL03a, CFdlL03b, CFdlL05, HCF$^+$16], we formulate an invariance equation (see (2)). This equation has two unknowns:

- a) embedings of the circle and its stable manifolds
- b) the dynamics of the map restricted to the invariant objects (the dynamics on the invariant circle and the dynamics on the leaves of the foliations).

This invariance equation (2) expresses that the circle is invariant, that the stable foliation is invariant (the leaves of the foliation are not invariant but they get sent to another leaf of the foliation by the dynamics).

We prove that, given an approximate solution of (2), we can evaluate some condition numbers on this approximate solution. If the error in (2) is smaller than an explicit function of the condition numbers, then there is a true solution of the invariance equation. Furthermore, the true solution is close to the approximate one. The condition numbers will be obtained by computing several observations of the approximate solution. The condition numbers do not involve any global assumptions on the map beyond some estimates on the derivatives in a neighborhood of the approximate solution. Such results are called *a-posteriori* theorems in the numerical literature [AO00].

A-posteriori results imply the usual persistence results under perturbations of dynamical systems. If one can find a system with these structures (invariant circle and its stable manifolds), then, for a small perturbation of the system, the original invariant objects provide an approximate invariant object for the perturbed system.

The a-posteriori results are also of great use in numerical analysis since they can provide criteria that ensure that the outputs of numerical computations – which are approximate

solutions of the invariance equation – can be trusted if we supplement them with a calculation of the condition numbers. Having very explicit condition numbers and results that allow trusting the calculation is invaluable when studying the phenomena that happen near the breakdown of the invariant objects and elementary tests (reruns, changing discretizations and the like) may get confusing. Furthermore, if the evaluation of the errors and the condition numbers are done taking care of all sources of error (truncation, round off, etc), one obtains a computer-assisted proof. Besides their use in numerical analysis, a-posteriori theorems can be used to validate the results of other non-rigorous techniques such as asymptotic expansions (these sophisticated expansions are useful in the study of degenerate Neimark-Sacker bifurcations [Nei59, Sac09]).

The way that one often proves an a-posteriori theorem is by describing an algorithm that given an approximate solution produces an even more approximate one and then showing that, if one starts from an approximate enough solution, the process converges.

In our case, we will develop a modification of the standard Newton's method to solve the invariance equation both for the parameterization of the invariant circle, the invariant foliation and for their dynamics. We will show that, when started from an approximate enough solution, this quasi-Newton method converges to a true solution.

To obtain the quasi-Newton method, we start with standard Newton method for the functional equation, but take into account that due to the structure of the problems, there are several useful identities. Using these identities coming from the geometry (related to the "group structure" in [Mos66b]) we can obtain an algorithm that is much easier (and much faster and reliable when implemented numerically) than the straightforward Newton method without affecting the essential feature of the Newton method, namely that the error after one step is roughly quadratic with respect to the original error. It is interesting that the same identities that are used to obtain convergence of the rigorous proof lead also to a more efficient and reliable algorithm. We will refer to this iterative method as a *"quasi-Newton"* method.

To prove the convergence of the quasi-Newton method, we rely on a Nash-Moser technique, combining the Newton step with a smoothing step. In the self-contained Appendix A, we present an abstract result, Theorem 13, which we hope could be applicable in similar problems.

As we will see, the equation (2) is underdetermined. This underdetermination is quite useful since it allows to develop more efficient numerical methods. As it is well known, the geometric objects (invariant circle and the stable foliation are locally unique. The under-determinacy, is only about the parameterization. The same geometric object can be given different parameterizations. Some of them will be numerically more efficient.

1.2. **Some Remarks on Comparison of the Nash-Moser Theorem with Other Results.** For the experts in Nash-Moser theory, we point out that Theorem 13 developed in Appendix A has several unusual properties.

Of course, this subsection can be omitted in the first reading, but we hope this could serve as motivation for some of the analysis.

- The linearized equation can be solved without loss of regularity for a range of regularities, but there is no theory of solutions for more regular data.

    This is very different from the Nash-Moser applications in small divisor problems or in PDE, in which one can solve the linearized equation in spaces of functions with any regularity (including analytic) but the solution incurs a loss of regularity.

As a consequence, in our problem, we cannot use usual smoothing techniques of approximating by analytic or C^∞ solutions. The only smoothing technique we can use is approximations by C^r functions (the so-called C^r smoothing).

We found inspiring the abstract implicit function theorems from [Sch60] and [Zeh75].

- We will need to consider spaces of functions with mixed regularity. The functions we will consider are C^r smooth in one of the variables (θ), but analytic in the other variable (s).

  The function spaces we use have two indices, one to measure the number of derivatives in the first variable and another one to measure the size of analyticity domains in the second variable.

  These spaces are indeed forced by the nature of the problem. It is known that the invariant circles could be only finitely differentiable [dlL97] – the degree of differentiability is limited, not just by the regularity of the map, but also by the ratios of the eigenvalues at the periodic orbits. On the other hand, the leaves of the stable foliation are always analytic. It is known that similar anisotropic regularities happen in the theory of normally hyperbolic invariant manifolds. We hope that many of the techniques developed here could have wider applicability.

- The nonlinear operator involved in the functional equation is basically the composition operator – which has very unusual regularity properties in C^r spaces, see [dlLO99].

  This operator maps C^r spaces into themselves. However, it is not differentiable from C^r to C^r but it is differentiable when the domain and the range are given other topologies. See [dlLO99] for an exhaustive study.

  Hence, computing the remainder of the functional after a correction involves losses of derivatives in C^r spaces. On the other hand, when considering Banach spaces of analytic functions, provided that the domains and ranges allow the composition, the composition operator is differentiable (even analytic).

  Since the composition operator appears very commonly in the study of invariance equations in dynamical systems, maybe some of the techniques developed in this paper may have other applications.

- As we will see in the detailed calculations, we will only need to smooth in the finite differentiable variable (θ), but we do not need to smooth in the analytic variable (s).

- The iteration we use takes advantage of some identities obtained by taking derivatives of the invariance equation. (From the practical point of view, the use of these identities is crucial to obtain quadratically convergent algorithms that require small storage and small operation count per step of iteration.)

  This entails that the remainder after applying the Newton method contains a term that involves the derivative of remainder of the starting approximation times the correction. This term is very common in many problems of dynamical systems that are solved taking advantage of *automatic reducibility*. Such terms do not appear in many other abstract Nash-Moser theorems. Some abstract theorems that incorporate the similar terms appear in [Van02, CdlL10, CCdlL13].

- The equation considered is underdetermined so that the linearized equation will have a kernel.

- The loss of regularity incurred in our result: Theorem 13 is much smaller than the loss of regularities in other abstract hard implicit function theorems.

Remark 3. Newton or quasi-Newton methods to compute invariant objects with the parameterization method have been used for a long time in the numerical literature [HdlL06b, CH17a, CH17b, HCF⁺16, Gra17] since they were found empirically to be efficient, and the solutions obtained could be validated using the more conventional methods (either contraction-based methods [BLZ08] or topological methods [CZ15]).

In implementing Newton methods for invariance equation, turns out that out of the box Krylov-Arnoldi, etc. methods doew not to work very well since the spectrum of equations involving invariance problems are invariant under rotations [Mat68, Ado07].

Note also that the Newton or quasi-Newton methods are much more effective than contraction based graph transform methods, especially when the contracting exponents are close to one. This is physically the regime of small friction, which is receiving great attention since in many practical problems, reducing dissipation is a design goal.

In the case that the internal dynamics (denoted by a in later sections) is fixed to be a rotation, one can solve the cohomology equations by Fourier methods so that the computation remains valid even for very weak contraction properties (which would require a large number of iterates by graph transform methods). This case has been studied in the literature several times. [CCdlL13, CH17b]. It is interesting that, in this case, the method requires the use of small divisors. Even if small divisors are not required in the linearized invariance equation, but to keep the internal dynamics being a rotation.

Remark 4. Studying simultaneously the equation for the circle and the foliation is, paradoxically, more efficient than studying first the circle and then the foliation.

The reason for this speedup is that the approximate solutions for the foliation are very powerful preconditioners for the invariance equation for the circle.

Remark 5. Besides using the Nash-Moser method, there are other methods that also lead to an a-posteriori format by using a contraction in C^0 and propagated bounds in higher regularity. [BLZ08]. Such contraction methods may give better regularity results than the Nash-Moser methods presented in this paper,

Remark 6. The models we consider – limit cycles subject to periodic perturbations are known to present regimes of parameters where the phenomena studied here breaks down and some complex behaviors appear: [Lev81, WY02, WY03]. The study of the boundary between the regular behavior presented here and the chaotic behavior is a very interesting mathematical problem [BDV05]. Some elements of the boundary of validity of the results have been explored in [GE88, Ran92a, Ran92b, Ran92c, HdlL06a, HdlL07, BS08, CF12, FH12]. It is clear that there can be several interesting phenomena at play and that a systematic exploration of the boundary will yield a very rich variety of behaviors. Inspired by this paper, the numerical algorithms implemented in [YdlL21] can, in principle, continue the results in the space of parameters to reach arbitrarily close to parameters where the objects described here break down (The precise definition of breakdown is somewhat subtle, please refer to [YdlL21] for more detailed discussions). One can hope that these numerical explorations of the frontier of hyperbolicity – which will require substantial effort – could yield some new ideas. Having mathematical tools that allow being confident of numerical results even if they are unexpected, will be important to discover new phenomena.

Remark 7. In this paper, we will specialize in the case of maps in two dimensions, but many of the techniques that we develop – including the abstract implicit function theorem – applies in any number of dimensions. The adaptation, however, is not completely straightforward

since new phenomena may appear, related to resonances among normal eigenvalues and the dynamics in the stable foliation will have to be more complicated. We hope to come back to this problem, but anticipate that the dynamics in the stable manifold has to involve more parameters.

We also note that in the case of higher dimensional manifolds, there are more complicated foliations defined in a neighborhood. These foliations are, in general not unique, but they have been found useful to describe the behavior in a neighborhood of a normally hyperbolic manifold [BLZ00]. We also call attention to the very interesting numerical paper [CJ15] and its associated numerical package `FOLI8PAK` which deals with similar problems. We hope that the present method can be adapted to the study of these manifolds or even to some non-resonant foliations. The paper [Sza20] points that these invariant objects may be useful in data reduction (see also [dlLK19]). We hope to come back to these problems.

1.3. **Organization of the Paper.** In this paper, we have chosen to present the motivation before the main statement, since the motivation leads to a practical algorithm. Many of the choices in the precise formulation are motivated by the need to give a precise formulation to the calculations. Of course, the readers interested only on the precisely formulated mathematical results can skip to Section 5.

The paper is organized as follows. In Section 2, we formulate an invariance equation by the parameterization method, which is the essential object in this paper.

The algorithm for solving the invariance equation is discussed and motivated in Section 3.

The rigorous result in the convergence of the algorithm (Theorem 4) for the existence of the solution and the convergence for the algorithm is presented in Section 5.

The proof of Theorem 4 is presented in Section 6, where we establish estimates on the ingredients of the algorithm.

The final step of the proof of Theorem 4 is a modified version of a Nash-Moser implicit function theorem (Theorem 13) which we present in the self-contained Appendix A. We hope that Theorem 13 can be of independent interest since it could be applicable to similar problems.

2. SETUP OF THE PROBLEM

In this section, we first briefly introduce the general idea of the parameterization method (Section 2.1). More detailed discussions about this method are in [HCF$^+$16].

Then, in a manner inspired by [HdlL13], we formulate an invariance equation (2) for the invariant circle and stable foliation near it. (Section 2.2)

It is important to notice that the invariance equation (2) is very underdetermined. Taking advantage of this underdetermination, in Section 2.3, we find a version of the invariance equation with extra properties. In Section 2.4, we discuss the stable manifolds.

2.1. **The General Setting for the Parameterization Method for Invariant Objects.** We start by describing the general idea of the parameterization method for finding invariant manifolds. In the later discussion, we will use the generalized version which allows to find also invariant foliations.

In a phase space $\mathscr{A}$, $f : \mathscr{A} \to \mathscr{A}$ is a diffeomorphism that generates a discrete dynamical system, the goal is to find an f-invariant submanifold $\mathscr{K} \subset \mathscr{A}$, i.e. $f(\mathscr{K}) \subseteq \mathscr{K}$. Consider $K : \Theta \to \mathscr{A}$ to be an injective immersion from some model manifold Θ that parameterizes
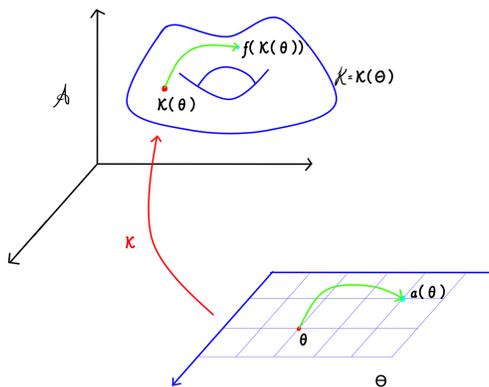
FIGURE 1. Parameterization of an invariant manifold (Figure taken from [HCF$^+$16])

$\mathscr{K}$. We have that $\mathscr{K}$ is f-invariant if and only if the following invariance equation holds:

$$(1) \qquad\qquad\qquad f \circ K(\theta) = K \circ a(\theta),$$

where the diffeomorphism $a : \Theta \to \Theta$ is the internal dynamics on Θ, and θ is the local coordinate in Θ (See Figure 1). The goal now becomes solving Equation (1) with $K(\theta), a(\theta)$ as the unknowns.

There are several methods to solve equation (1) depending on the class of dynamical systems used.

A widely applicable idea (and the one we will be concerned with here) is to apply the Newton (or quasi-Newton) iterative method to find the correction $\Delta_K(\theta)$ and $\Delta_a(\theta)$ that improves approximate $K(\theta)$ and $a(\theta)$. By constructing an adapted frame $P(\theta)$, and representing $\Delta_K(\theta) = P(\theta)\phi(\theta)$, solving the Newton method for the equation (1) amounts solving cohomological equations of the form as in equation (25), which can then be solved under hyperbolicity assumptions. We will present algorithms and establish their convergence.

Remark 8. In the case when the rotation number for the internal dynamics is fixed to be a given Diophantine number ω, $a(\theta) = \theta + \omega$ is no longer an unknown in (1). On the other hand, one has to adjust parameters. See [CH17a, CH17b] for the theory for invariant circles.

An alternative theoretical point of view for the adjustment of parameters is that, if we consider a family with parameters λ, our method will obtain a family of circle mappings a_λ. Adjusting parameters λ as in [Mos66b, Mos66a], we obtain that the map a_λ is smoothly conjugate to a Diophantine rotation.

2.2. **The Invariance Equation of the Invariant Circle and the Stable Foliation.**
Given a smooth diffeomorphism $f : \mathbb{T} \times \mathbb{R} \to \mathbb{T} \times \mathbb{R}$ that generates a discrete dynamical system in $\mathbb{T} \times \mathbb{R}$, we assume that f admits a stable invariant circle. Our goal is to find the invariant circle and the corresponding stable manifolds of points.

More precisely, following a similar approach as in Section 2.1, we are looking for an injective immersion $W : \mathbb{T} \times \mathbb{R} \to \mathbb{T} \times \mathbb{R}$ such that it parameterizes the neighborhood of the invariant circle. Thus, we will consider the following invariance equation:

$$(2) \qquad\qquad\qquad f \circ W(\theta, s) - W(a(\theta), \lambda(\theta, s)) = 0,$$

where $a : \mathbb{T} \to \mathbb{T}$ describes the internal dynamics on the invariant circle, and $\lambda : \mathbb{T} \times \mathbb{R} \to \mathbb{R}$ describes the dynamics on the stable manifolds of points.

In the above equation (2), $W(\theta, s), a(\theta), \lambda(\theta, s)$ are the unknowns, and $f(\theta, s)$ is the only known function.

It is important to emphasize that the unknowns for equation (2) are functions. Dealing with it in this paper will require tools from functional analysis.

Note that when the phase space is $\mathbb{T} \times \mathbb{R}$, there are two topologically different embeddings of circles. One is when the circle is non-contractible in the phase space and the other is when the circle is embedded in a contractible way. This can be seen as boundary conditions on the embedding W. In the non-contractible case, the lift of the embedding satisfies $W(\theta + 1, 0) = W(\theta, 0) + (1, 0)$ and in the contractible case, the lift of the embedding satisfies $W(\theta + 1, 0) = W(\theta, 0)$.

It is reasonable to assume that $W(\theta, 0)$ is the parameterization of the invariant circle, it follows that $\lambda(\theta, 0) = 0$. If one denotes $K(\theta)$ as $W(\theta, 0)$, the invariance equation (2) reduces to equation (1). Moreover, if $\sup_\theta |\partial_s \lambda(\theta, 0)| < 1$, we have the invariant circle is stable. Sharper sufficient conditions for stability will be derived later.

In this paper, we will allow that the internal dynamics $a(\theta)$ is phase-locked (i.e. it has an attractive periodic orbit). In such a case, it can happen (indeed, one expects that this is the most common case in applications) that the invariant circle is only finitely differentiable even if the map f is analytic or even polynomial.

2.3. **Underdetermination of the Invariance Equation.** One nice property of the invariance equation (2) is that it is highly underdeterminate, thus admits many solutions. Hence, depending on the problems, we can impose extra properties that improve the computation. In this section, we will review some of the sources of underdetermination that lead to improvements in the computation.

Clearly, the changes of coordinates in the reference manifold leads the same geometric objects (same circle, same stable leaves) but given different parameterizations. It can be shown that the only lack of local uniqueness of the reference manifold of (2) is these changes of variables in the reference manifold. Any two solutions of (2) close enough are related by a change of variables in the reference manifold and hence describe the same geometric object.

From the numerical point of view, depending on the properties of the system, we can recalibrate the system of coordinates so that the computation is better. Clearly, if our goal is to find a solution, having several solutions available is a very good feature.

In the following, we review the different sources of underdetermination in (2) so that we take advantage of them.

Given $(W(\theta, s), a(\theta), \lambda(\theta, s))$ satisfying (2), we have that

- Conjugacy on θ: For any diffeomorphism $g : \mathbb{T} \to \mathbb{T}$, we have

$$\widetilde{W}(\theta, s) = W(g(\theta), s),$$

$$\widetilde{a}(\theta) = g^{-1} \circ a \circ g(\theta),$$

$$\widetilde{\lambda}(\theta, s) = \lambda(g(\theta), s),$$

is also a solution of (2).

- Conjugacy on s: For any $\widehat{\lambda} : \mathbb{T} \times \mathbb{R} \to \mathbb{R}$, if there exists a differentiable function $h : \mathbb{T} \times \mathbb{R} \to \mathbb{R}$ such that

$$(3) \qquad\qquad h(a(\theta), \widehat{\lambda}(\theta, s)) = \lambda(\theta, h(\theta, s)),$$

we have

$$\widehat{W}(\theta, s) = W(\theta, h(\theta, s)),$$

$$\widehat{a}(\theta) = a(\theta),$$

$$\widehat{\lambda}(\theta, s)$$

is also another solution of (2).

According to Lemma 6 and its remarks, we can see that such $h(\theta, s)$ as in (3) exists in the case that $\widehat{\lambda}(\theta, s)$ equals to the linear term of $\lambda(\theta, s)$ with respect to s, provided that the norm of λ is small enough. We postpone the detailed discussion and the proof to Section 6.1. As remarked there, Lemma 6 is a fibered version of Poincaré-Sternberg theorem on the linearization of contractions.

Benefiting from the second underdetermination property and Lemma 6, instead of considering (2) we can consider

$$(4) \qquad\qquad f \circ W(\theta, s) - W(a(\theta), \lambda(\theta)s) = 0.$$

Our goal now becomes solving for $W(\theta, s), a(\theta)$ and $\lambda(\theta)$ from equation (4).

Note that, if W, a, λ is a solution of (4), clearly W is an invariant foliation with internal dynamics given by a, λ. One could, however, wonder if there are other invariant foliations. The content of Lemma 6 is to show that, if there was an invariant foliation, then, one can obtain a solution of (4) by reparameterizing it. Hence, finding a solution of (4) is not only sufficient for finding invariant foliations but also equivalent.

Remark 9. If $a(\theta)$ conjugates to a Diophantine rotation $\theta + \omega$, one can show that $\lambda(\theta)$ can be reduced to a constant.

In fact, given a tuple $(W(\theta, s), \theta + \omega, \lambda(\theta))$ satisfies Equation 4, we can show that there exists a constant $\bar{\lambda}$ and $h(\theta, s) = r(\theta)s$ such that Identity (3) holds.

To prove this, we start with a function $\widehat{\lambda}(\theta)$, and the goal is to find $r(\theta)$ to reduce such $\widehat{\lambda}(\theta)$ to a constant.

From Identity (3), we have

$$r(\theta + w)\lambda(\theta) = \widehat{\lambda}(\theta)r(\theta),$$

by taking the logrithm, we have

$$(5) \qquad\qquad \log \widehat{\lambda}(\theta) = \log \lambda(\theta) + \log r(\theta) - \log r(\theta + \omega),$$

Standard discussions of cohomological equations in KAM theory [dlL01] show that $\widehat{\lambda}$ can be made into a constant if and only if

$$(6) \qquad\qquad \log \lambda(\theta) + \log r(\theta) - \log r(\theta + \omega) = \int_{\mathbb{T}} \log \lambda(\theta)d\theta$$

holds, in which case $\widehat{\lambda}(\theta) = \exp \int_{\mathbb{T}} \log \lambda(\theta)d\theta \triangleq \bar{\lambda}$.

2.4. **Stable Manifolds of Points (Isochrons).** Notice that the invariance equation (2) contains not only the dynamics of the invariant circle, but also the dynamics in a neighborhood of the invariant circle. In particular, if equation (4) is satisfied, and if $\sup_{\theta \in \mathbb{T}} |\lambda(\theta)| < 1$, we have the set

$$I_\theta = \{W(\theta, s) \mid s \in \mathbb{R}\}.$$

consists of points whose orbits converge exponentially fast (with a high enough rate) to the orbit of $W(\theta, 0)$ since

(7)
$$f^{\circ j}(W(\theta, s)) = W(a^{\circ j}(\theta), \lambda^{[j]}(\theta)s),$$
$$f^{\circ j}(W(\theta, 0)) = W(a^{\circ j}(\theta), 0)$$

where

(8)
$$\lambda^{[j]}(\theta) = \lambda(\theta)\lambda(a(\theta))\lambda(a^{\circ 2}(\theta)) \cdots \lambda(a^{\circ(j-1)}(\theta))$$

and $a^j(\theta)$ denotes $a(\theta)$ composing with itself j times. Note that

(9)
$$\lambda^{[j+k]}(\theta) = \lambda^{[j]}(a^{\circ k}(\theta))\lambda^{[k]}$$

Hence $\sup_\theta |\lambda^{[j+k]}(\theta)| \leq \sup_\theta |\lambda^{[j]}(\theta)| \cdot \sup_\theta |\lambda^{[k]}(\theta)|$.

More specifically, when $\sup_\theta |\lambda(\theta)| < 1$, for all θ, we have $f^k(I_\theta) \to a^k(\theta)$ exponentially fast as $n \to \infty$.

Note that the isochrons are not invariant sets. Nevertheless, they behave well under the map. We have

$$f(I_\theta) \subset I_{a(\theta)}$$

so that the foliation given by all the isochrons is invariant in the sense that if two points are in the same leaf, applying the map to both of them, we obtain another pair of points in the same leaf (different from the original one).

Remark 10. Given $\lambda(\theta)$, we will refer to the quantity

(10)
$$\lambda^* := \lim_{n \to \infty} \left(\|\lambda^{[n]}\|_{C^0} \right)^{\frac{1}{n}}$$

as the *dynamical average*.

Since $\|\lambda^{[n+m]}\|_{C^0} \leq \|\lambda^{[n]}\|_{C^0}\|\lambda^{[n+m]}\|_{C^0}$, the limit in (10) always exists.

The implicit function theorem shows that the set of stable manifolds forms a foliation in a neighborhood of the circle and we can use the equation (2) to show that the set of isochrons is indeed a foliation globally. Note that applying the implicit function theorem requires that the circle is C^1. When the circle is less regular, the implicit function theorem can only conclude that the leaves form a pre-foliation. The conclusion that the isochrons form a foliation is also obtained using more dynamical arguments in [Fen74, Fen77]. It suffices to realize that the relation

$$y \approx \tilde{y} \Leftrightarrow d\left(f^n(y), f^n(\tilde{y})\right) \leq C_{y,\tilde{y}}\lambda^n \quad n > 0$$

is an equivalence relation. In [Fen74, Fen77], it is required that $\lambda < \|Da^{\circ k}\|$ for some $k > 0$ for the persistence of the circle as a C^1 manifold.

The phenomena that happen when this inequality is violated, have been studied in the literature. A discussion can be found in [YdlL21].

3. The Algorithm

In this section, we discuss our algorithm for solving the invariance equation (4). Unfortunately, (4) is hard to solve using the Newton method. Instead of the standard Newton method, we use a modification obtained by omitting terms that are heuristically quadratically small. Omitting these terms makes the equation much easier to solve but, heuristically, does not change the quadratic convergence. These heuristic arguments are rigorously justified later in Section 6.3.

In Section 3.1 and 3.2, we present the details of one step of the quasi-Newton method.

Given an approximate solution, we look for the corrections that so that the new error is quadratic in the original error. The method takes advantage of several identities.

As it turns out, the main ingredient in the method is solving *cohomological equations*. The cohomological equations are solved in Section 3.3. In Section 3.4, we briefly discuss the step-by-step algorithm. In Section 5 we will state a result on the convergence of the algorithm. We will show that the steps can be repeated infinitely often and indeed converges.

The algorithm formulated in this section has been implemented in [YdlL21] and run in examples. We refer to [YdlL21] for details on implementation (how to discretize functions, number of variables used) As often happens, the algorithm is found to work with even in regions beyond the requirement of regularity of the rigorous proof and some new phenomena requiring mathematical explanation have been identified.

3.1. **The quasi-Newton Method.** In this subsection, we formulate one step of the quasi-Newton method to solve equation (4).

Assume that we have an approximate parameterization of the neighborhood of the invariant circle $W(\theta, s)$, an approximate internal dynamics $a(\theta)$ and an approximate dynamics on the isochrons $\lambda(\theta)s$ such that

$$(11) \qquad e(\theta, s) = f \circ W(\theta, s) - W(a(\theta), \lambda(\theta)s),$$

where $e(\theta, s)$ is the error.

The goal of one step of the quasi-Newton method is to compute the corrections $\Delta_W(\theta, s)$, $\Delta_a(\theta)$ and $\Delta_\lambda(\theta)$ such that

$$(12) \qquad f(W + \Delta_W)(\theta, s) - (W + \Delta_W)((a + \Delta_a)(\theta), (\lambda + \Delta_\lambda)(\theta)s) = 0$$

up to an error which is quadratically smaller than the initial error e.

For the moment, we work heuristically and ignore regularities. All these issues will be settled later in Lemma 11.

Using Taylor expansion and omitting higher order terms, Equation (12) becomes

$$\begin{aligned} 0 = & f(W(\theta, s)) + Df(W)(\theta, s)\Delta_W(\theta, s) - W(a(\theta), \lambda(\theta)s) \\ (13) \qquad & - DW(a(\theta), \lambda(\theta)s)\begin{pmatrix} \Delta_a(\theta) \\ \Delta_\lambda(\theta)s \end{pmatrix} - \Delta_W(a(\theta), \lambda(\theta)s) + \text{higher order terms}, \end{aligned}$$

where the term $D[\Delta_W(a(\theta), \lambda(\theta)s)]\begin{pmatrix} \Delta_a(\theta) \\ \Delta_\lambda(\theta)s \end{pmatrix}$ is ignored for now because it is "heuristically" quadratically small. We will make a rigorous argument later in Lemma 11.

Now we have that Equation (13) has become

$$Df(W(\theta,s))\Delta_W(\theta,s) - DW(a(\theta),\lambda(\theta)s)\begin{pmatrix}\Delta_a(\theta)\\ \Delta_\lambda(\theta)s\end{pmatrix}$$

$$(14) \qquad -\Delta_W(a(\theta),\lambda(\theta)s) = -e(\theta,s).$$

Remark 11. Notice that one should treat equation (14) as an equation for $\Delta_W(\theta,s), \Delta_a(\theta)$ and $\Delta_\lambda(\theta)$, with $f(\theta,s)$ given by the problem, and $W(\theta,s), a(\theta)$ and $\lambda(\theta)$ given by the initial approximation as well as the RHS e.

To simplify the above equation (14), we will express $\Delta_W(\theta,s)$ in the frame $DW(\theta,s)$ as follows:

$$(15) \qquad \Delta_W(\theta,s) = DW(\theta,s)\Gamma(\theta,s).$$

Remark 12. Notice that if $DW(\theta,s)$ is invertible, solving for $\Delta_W(\theta,s)$ is equivalent to solving for $\Gamma(\theta,s)$. One can see that if the initial guess of $W(\theta,s)$ is close enough to the true solution and $DW(\theta,s)$ is invertible initially, $DW(\theta,s)$ remains to be invertible for each step of the iteration.

By taking the derivative of equation (11), we have that

$$(16) \qquad De(\theta,s) = Df(W(\theta,s))DW(\theta,s) - DW(a(\theta),\lambda(\theta)s)\begin{pmatrix}Da(\theta) & 0\\ D\lambda(\theta)s & \lambda(\theta)\end{pmatrix}.$$

Then, by substituting (15) and (16) in the quasi-Newton equation (14), we obtain

$$\begin{pmatrix}Da(\theta) & 0\\ D\lambda(\theta)s & \lambda(\theta)\end{pmatrix}\Gamma(\theta,s) - \begin{pmatrix}\Delta_a(\theta)\\ \Delta_\lambda(\theta)s\end{pmatrix} - \Gamma(a(\theta),\lambda(\theta)s)$$

$$= -(DW(a(\theta),\lambda(\theta)s))^{-1}e(\theta,s)$$

$$(17) \qquad \triangleq \widetilde{e}(\theta,s),$$

where the term $De(\theta,s)\Gamma(\theta,s)$ is also omitted for the same reason as in equation (13), and the rigorous justification is again left to Lemma 11.

If we express equation (17) in components, we obtain the following two equations for the unknowns $\Gamma_1(\theta,s)$, $\Gamma_2(\theta,s)$, $\Delta_a(\theta)$ and $\Delta_\lambda(\theta)$.

$$(18) \qquad Da(\theta)\Gamma_1(\theta,s) - \Delta_a(\theta) - \Gamma_1(a(\theta),\lambda(\theta)s) = \widetilde{e}_1(\theta,s),$$

$$(19) \qquad \lambda(\theta)\Gamma_2(\theta,s) - \Delta_\lambda(\theta)s - \Gamma_2(a(\theta),\lambda(\theta)s) = \widetilde{e}_2(\theta,s) - D\lambda(\theta)s\Gamma_1(\theta,s)$$

$$\triangleq M(\theta,s).$$

where $\Gamma_1(\theta,s)$ and $\Gamma_2(\theta,s)$ are the components of $\Gamma(\theta,s)$.

3.2. **Solving $\Gamma_{1,2}, \Delta_\lambda, \Delta_a$ from Equation** (18), (19). In this subsection, we present the details of solving equation (18) and (19). To study those two equations, we will discretize any function from $\mathbb{T}\times\mathbb{R}: g(\theta,s)$ as Taylor series with respect to s:

$$g(\theta,s) = \sum_{j=0}^{\infty} g^{(j)}(\theta)s^j,$$

with the assumption that $g(\theta,s)$ is C^r in θ and real analytic in s, where $g^{(j)}(\theta)\in C^r$ is the coefficient for $s^j, j\geq 0, j\in\mathbb{N}$. In the context of Section 5, $g(\theta,s)\in\mathcal{X}^{r,\delta}$ for some $\delta>0$.

By matching coefficients of s^j on both sides, we can rewrite equation (18) and (19) as a hierarchy of equations provided that $Da(\theta)$ and $\lambda(\theta)$ are not equal to 0 for any $\theta \in \mathbb{T}$.

- For equation (18):
  - For the coefficients of s^0:

$$(20) \qquad Da(\theta)\Gamma_1^{(0)}(\theta) - \Gamma_1^{(0)}(a(\theta)) - \Delta_a(\theta) = \widetilde{e}_1^{(0)}(\theta),$$

  - For the coefficients of s^j, $j \geq 1, j \in \mathbb{N}$:

$$(21) \qquad \Gamma_1^{(j)}(\theta) = \frac{\lambda^j(\theta)}{Da(\theta)}\Gamma_1^{(j)}(a(\theta)) + \frac{\widetilde{e}_1^{(j)}(\theta)}{Da(\theta)}.$$

- For equation (19):
  - For the coefficients of s^0:

$$\lambda(\theta)\Gamma_2^{(0)}(\theta) - \Gamma_2^{(0)}(a(\theta)) = M^{(0)}(\theta),$$

   which, by composing $a^{-1}(\theta)$, can be rewritten as

$$(22) \qquad \Gamma_2^{(0)}(\theta) = \lambda(a^{-1}(\theta))\Gamma_2^{(0)}(a^{-1}(\theta)) - M^{(0)}(a^{-1}(\theta)),$$

  - For the coefficients of s^1:

$$(23) \qquad \lambda(\theta)\Gamma_2^{(1)}(\theta) - \Gamma_2^{(1)}(a(\theta))\lambda(\theta) - \Delta_\lambda(\theta) = M^{(1)}(\theta),$$

  - For the coefficients of s^j, $j \geq 2, j \in \mathbb{N}$:

$$(24) \qquad \Gamma_2^{(j)}(\theta) = \lambda^{j-1}(\theta)\Gamma_2^{(j)}(a(\theta)) + \frac{M^{(j)}(\theta)}{\lambda(\theta)}.$$

The hierarchy of equations above is well known from perturbation expansions. Algorithms for efficient computation of the coefficients can be found in [ZdlL18].

Again, our goal is to solve the above equations for $\Delta_a(\theta), \Delta_\lambda(\theta), \Gamma_1^{(j)}(\theta), \Gamma_2^{(j)}(\theta)$ for $j \geq 0$.

First, notice that Equation (20) and (23) are underdetermined equations, hence the solution is not unique. An interesting question we have not yet pursued is how to choose the solution of (20) and (23) that improves the numerical stability of the algorithm. Intuitively, it seems desirable to design the algorithms so that the a, λ are *"simple"*, but we have not succeeded in making this precise when the inner dynamics is phase-locked. (When $a(\theta)$ is conjugate to a Diophantine rotation, one can use the underdeterminacy to set $a(\theta)$ to be a Diophantine rotation and $\lambda(\theta)$ to be a constant, see Remark 9.)

In this paper, we choose the most obvious solution: For equation (20), we let $\Gamma_1^{(0)}(\theta) = 0$ and thus $\Delta_a(\theta) = -\widetilde{e}_1^{(0)}(\theta)$; for equation (23), we let $\Gamma_2^{(1)}(\theta) = 0$ and thus $\Delta_\lambda(\theta) = -M^{(1)}(\theta)$. This choice of solution guarantees the norm is controled by the error, it is referred as the graph style in [HCF$^+$16].

Notice that Equation (21), (22) and (24) have been reorganized so that are written as cohomological equation of the form:

$$(25) \qquad \phi(\theta) = l(\theta)\phi(a(\theta)) + \eta(\theta).$$

where $\phi(\theta)$ is the unknown and $a(\theta), l(\theta)$ and $\eta(\theta)$ are given.

3.3. **Solving ϕ from the Cohomological Equation** (25)**.** In this subsection, we solve Equation (25) by contraction.

By inductively replacing $\phi(\theta)$ on the right hand side of (25) by the equation itself, we have

$$\phi(\theta) = \eta(\theta) + l(\theta)\eta(a(\theta)) + l(\theta)l(a(\theta))\eta(a^{\circ 2}(\theta))$$
$$+ \ldots + l(\theta)l(a(\theta))l(a^{\circ 2}(\theta))\cdots l(a^{\circ(n-1)}(\theta))\eta(a^{\circ n}(\theta))$$
$$+ l(\theta)l(a(\theta))l(a^{\circ 2}(\theta))\cdots l(a^{\circ n}(\theta))\phi(a^{\circ(n+1)}(\theta))$$

$$(26) \qquad = \sum_{j=0}^{n} l^{[j]}(\theta)\eta(a^{\circ j}(\theta)) + l^{[n+1]}(\theta)\phi(a^{\circ(n+1)}(\theta)),$$

where as in equation (8), $l^{[j]}(\theta) = l(\theta)l(a(\theta))l(a^{\circ 2}(\theta))\cdots l(a^{\circ(j-1)}(\theta))$, and $l^{[0]}(\theta) = 1$.

Note that, if $\|l^{[j]}\|_{C^0} < 1$, and ϕ is bounded, the last term in (26) tends to zero uniformly. Hence, the only possible C^0 solution of (25), is

$$(27) \qquad \phi(\theta) = \sum_{j=0}^{\infty} l^{[j]}(\theta)\eta(a^{\circ j}(\theta)).$$

As proved in Lemma 7 in Section 6.2, given r such that

$$(28) \qquad \|l\|_{C^0}\|Da\|_{C^0}^r < 1,$$

we will show $\|l^{[j]}(\cdot)\eta(a^{\circ j}(\cdot))\|_{C^r} \le C\alpha^j$ for some $C > 0, \alpha < 1$ so that $\sum_{j=0}^{n} l^{[j]}(\theta)\eta(a^{\circ j}(\theta))$ converges absolutely in C^r. Hence, (25) has a C^r solution.

The conditions (28) can be slightly improved to $\|l^{[k]}\|_{C^0}\|D(a^{\circ k})\|_{C^0}^r < 1$ (or even to $\|l^{[k]}D(a^{\circ k})\|_{C^0} < 1$). Nevertheless, there are explicit examples discussed in the remark after Lemma 7, cohomological equation (25) can only be solved for a finite range of r. These examples are rather persistent and they happen in open C^1 neighborhoods of a. So the phenomenon of the quasi-Newton method being defined only on a finite range of regularities has to be considered by the Nash-Moser method we develop in Appendix A.

Remark 13. As we will see in [YdlL21], the right hand side of equation (27) can be implemented very efficiently so that the summation of M terms requires only $\log M$ steps.

3.4. **The Algorithm for One Iteration of the quasi-Newton Method.** By the discussion in Section 3.1, we now summarize the steps for one iteration of the quasi-Newton method derived above. Estimations of the norms will be discussed in Section 6. Given an approximate solution $W(\theta, s)$, $a(\theta)$ and $\lambda(\theta)$, where $W(\theta, s)$ is truncated up to the L-th order in the power series expansion (from the analysts' point of view, $L = \infty$), the correction $\Delta_W(\theta, s)$(with maximal order L), $\Delta_a(\theta)$ and $\Delta_\lambda(\theta)$ are calculated by the algorithm as follows:

Remark 14. Step 6, 10, and 11 are solved based on Equation (27). As stated in Remark 13, in [YdlL21], we present a faster algorithm regarding to this.

In this paper, we will just present the analysis of the algorithm above and show its convergence under the hypothesis that the starting step is close to being a solution.

In [YdlL21] we will discuss the implementation details (discretization, programming considerations, and more importantly, diagnostics of reliability).

We point out that the algorithm is very efficient (it only manipulates functions). At no time one needs to store (much less invert) a matrix with the discretization of the error.

Algorithm 1 One iteration of the algorithm

Input: Initial $W(\theta, s), a(\theta)$ and $\lambda(\theta)$
Output: Solution $W(\theta, s), a(\theta)$ and $\lambda(\theta)$ to the invariance equation (4)

1: $\sum_{j=0}^{L} e^{(j)}(\theta)s^j = e(\theta, s) \leftarrow f \circ W(\theta, s) - W(a(\theta), \lambda(\theta)s)$,
2: Compute $DW(\theta, s)$ and $DW \circ (a(\theta), \lambda(\theta)s)$,
3: $\sum_{j=0}^{L} \widetilde{e}^{(j)}(\theta)s^j = \widetilde{e}(\theta, s) \leftarrow (DW(a(\theta), \lambda(\theta)s))^{-1}e(\theta, s)$,
4: $\Delta_a(\theta) \leftarrow -\widetilde{e}_1^{(0)}(\theta)$,
5: $\Gamma_1^{(0)}(\theta) \leftarrow 0$,
6: Solve $\Gamma_1^{(j)}(\theta)$ from equation (21) for $1 \le j \le L$,
7: $\sum_{j=0}^{L} M^{(j)}(\theta)s^j = M(\theta, s) \leftarrow \widetilde{e}_2(\theta, s) - D\lambda(\theta)s\Gamma_1(\theta, s)$,
8: $\Delta_\lambda(\theta) \leftarrow -M^{(1)}(\theta)$,
9: $\Gamma_2^{(1)}(\theta) \leftarrow 0$,
10: Solve $\Gamma_2^{(0)}(\theta)$ from equation (22),
11: Solve $\Gamma_2^{(j)}(\theta)$ from equation (24) for $2 \le j \le L$,
12: $\sum_{j=0}^{L} \Delta_W^{(j)}(\theta)s^j = \Delta_W(\theta, s) \leftarrow DW(\theta, s)\Gamma(\theta, s)$,
13: $W(\theta, s) \leftarrow W(\theta, s) + \Delta_W(\theta, s)$,
14: $a(\theta) \leftarrow a(\theta) + \Delta_a(\theta)$,
15: $\lambda(\theta) \leftarrow \lambda(\theta) + \Delta_\lambda(\theta)$,
16: Return updated $W(\theta, s), a(\theta)$ and $\lambda(\theta)$.

Hence the storage requirements will be proportional to space taken by the discretization of functions (not the square!) and that the operation count will be roughly proportional to the number of variables used to discretize a function (there may be logarithmic corrections if one uses Fourier methods; see [YdlL21].)

The algorithm is also easy to implement in a preliminary – but workable – form. Note that the algorithm has only 16 steps, each of which can be efficiently implemented in a few lines in a high-level language (or a good scientific library). The most complicated step is solving the cohomology equation, but we have iterative formulas for the solution, along with the quadratic convergence contraction algorithm (more in [YdlL21]).

Of course, developing a high-quality practical algorithm requires developing criteria that ensure correctness and monitor the accuracy. In that respect, having an a-posteriori theorem is an invaluable help.

The proof of the convergence involves alternating the algorithm with smoothing steps. In numerical applications, we have found it convenient to include a low pass filter that smooths the numerical calculations. This seems to provide enough smoothing. See a more detailed discussion in [YdlL21].

4. SCALE OF BANACH SPACES

In this section, we set up the scale of Banach spaces that is needed in Section 5 and Section 6. Since for the problem we are dealing with, functions with domain $\mathbb{T}$ admits only finite regularity (Lemma 7), we first recall the C^r space ($r \in \mathbb{N}^+ + (0, 1)$) [dlLO99] along with some inequalities, and based on that, we proceed to the $\mathcal{X}^{r,\delta}$ space for functions in $\mathbb{T} \times \mathbb{R}$,

and finally the $\mathscr{X}^{r,\delta}$ space that will be used in Section 5. The existence of the smoothing operator in C^r guarantees the existence of such smoothing operator in $\mathcal{X}^{r,\delta}$ and $\mathscr{X}^{r,\delta}$ spaces.

4.1. **Setup of the Scale of Spaces.** In this section, we describe the spaces that we will use. Roughly speaking, the spaces are for functions with domain $(\theta, s) \in \mathbb{T} \times \mathbb{R}$. The functions we are interested in will be finitely differentiable in the θ variable and analytic in the s variable. The spaces will therefore have two indices. One index measuring the – finite order – differentiability in θ and another index measuring the size of the analyticity domain in s.

The most delicate analysis (smoothing, approximation) will happen in the finite differentiable direction. In our case, this will be the circle. The analysis of finite differentiable spaces we present is rather standard. As it is well known in approximation theory, defining a family of regularities indexed by a real parameter becomes subtle for integer values of the parameter. A good reference is [Zeh75, Ste70, dlLO99]. The properties of spaces of functions with mixed regularity used in this paper are built on those.

We first recall the C^r spaces.

4.1.1. *Space for Functions in* $\mathbb{T}$. By standard definitions of the Hölder spaces (as in [dlLO99, dlLW11]), the spaces we will be concerned with for functions defined on $\mathbb{T}$ are:

Definition. Let X be a Banach space.
For $r \in \mathbb{N}$, we define:

$$C^r(\mathbb{T}, X) = \{f : \mathbb{T} \to X, r \text{ times continuously differentiable.}\}$$

We endow C^r with the supremum norm of all the derivatives of order up to r, which makes it into a Banach space.

For $r = n + \alpha \notin \mathbb{N}$ with $n = \lfloor r \rfloor \in \mathbb{N}, \alpha \in (0, 1)$ we define $C^r = C^{n+\alpha}$:

$$C^r(\mathbb{T}, X) = \{f : \mathbb{T} \to X, r \text{ times continuously differentiable}, D^\alpha f \text{ is } \alpha\text{-Hölder.}\}$$

We endow $C^r(\mathbb{T}, X)$ with the norm

$$\|f\|_{C^{n+\alpha}} = \max(\|f\|_{C^n}, H_\alpha(D^n f)),$$

where for a function $\phi : \mathbb{T} \to X$, we set

$$H_\alpha(\phi) = \sup_{x \neq y} \frac{|\phi(x) - \phi(y)|}{d(x, y)^\alpha}.$$

Remark 15. In this paper (excluding Appendix A), we always denote $r \geq 0$ for the regularity, and we always have $n = \lfloor r \rfloor$ and $\alpha = r - n$.

Remark 16. The case $\alpha = 1$ agrees with the Lipschitz constant and is very natural. We have excluded it to avoid complicating the notation since C^{r+1} would be ambiguous when r is an integer.

Remark 17. H_α is a seminorm and $H_\alpha(\phi) = 0$ if and only if ϕ is a constant.

Remark 18. The C^r scale of spaces is very natural and easy to work with since the definitions of the norms are very explicit. As it is well known, the C^r scale of spaces has anomalies when r is an integer (the properties of approximation and smoothing are not as expected, etc). So, it is common in analysis to use the other scales of spaces. (called Λ_r in [Ste70] or $\hat{C}^r$ in [Mos66b, Zeh75]).

In this paper, we will not use the Λ_r spaces (the composition operator plays a role in our study and there does not seem to be in the literature a systematic study of composition in the Λ_α scales) but our results will include some caveats that the spaces in the hypothesis or in the conclusions are not integers. Sometimes, this just amounts to making some inequalities in the range strict.

Remark 19. When $r \in \mathbb{N}$, the C^r spaces can be defined taking values in any manifold Riemannian (or even Finsler) manifold. When $r > 1$ and $r \notin \mathbb{N}$, the definition, in general, is complicated since to define H_α, one needs to compare the values of derivatives at two different points. This requires making explicit some cumbersome choices. In this paper, however, we will only need to deal with $C^r(\mathbb{T}, \mathbb{T})$ or $C^r(\mathbb{T}, \mathbb{R})$. For $\mathbb{T}$, there is a natural identification of all the tangent spaces of different points, so that there is no problem in defining C^r spaces taking values on the torus.

4.1.2. *Space for Functions in $\mathbb{T} \times \mathbb{R}$.* Given $\delta < 1$, we define the space $\mathcal{X}^{r,\delta}$ as follows:

Definition. For a function $u(\theta, s)$ with domain $\mathbb{T} \times [-\delta, \delta]$, we say $u \in \mathcal{X}^{r,\delta}$ if $u(\theta, s) = \sum_{j=0}^{\infty} u^{(j)}(\theta) s^j$ with $u^{(j)}(\theta) \in C^r$ and $\sum_{j=0}^{\infty} \left\| u^{(j)} \right\|_{C^r} \delta^j < \infty$. In other words,

$$\mathcal{X}^{r,\delta} = \left\{ u(\theta, s) = \sum_{j=0}^{\infty} u^{(j)}(\theta) s^j \mid u^{(j)}(\theta) \in C^r, \text{ and } \sum_{j=0}^{\infty} \left\| u^{(j)} \right\|_{C^r} \delta^j < \infty \right\}$$

with norm

$$\|u\|_{\mathcal{X}^{r,\delta}} = \sum_{j=0}^{\infty} \left\| u^{(j)} \right\|_{C^r} \delta^j.$$

Remark 20. It is useful to think of $\mathcal{X}^{r,\delta}$ as a space of C^r functions from the circle to a space of analytic functions on the unit disk.

This corresponds well to the idea of local foliations. We can think of a function that to each of the base points associates a segment of the analytic leaf.

Remark 21. Note that the space $\mathcal{X}^{r,\delta}$ consists of functions that in the variable s have a domain of analyticity which is a disk.

This is, of course, enough when we are considering local foliations, but if we study global foliations, it can well happen that the true domain of analyticity of the leaves is not a disk.

From the numerical point of view, it is natural and efficient to represent functions in a disk using power series and indeed the definition of the norm in $\mathcal{X}^{r,\delta}$ is done to reflect that. On the other hand, one should keep in mind that in the global study of foliations, finding solutions of (4) in $\mathcal{X}^{r,\delta}$ only gives us segments of the leaves. Roughly, we are studying the solution in a circle, which extends to the singularity closest to the origin. If this singularity happens away from the real line, the parameterization may be analytic for real values outside the circle of convergence.

Numerically, this corresponds to the step of *"globalization"*. Once we have obtained a good representation of the function in a neighborhood of the origin using power series, we can use (4) to obtain the parameterization in a larger domain. Some interesting examples of foliation with global computations appear in [BST98].

For notational simplicity, we denote $\mathcal{X}^{r,\delta}$ as $\mathcal{X}^r$ when the δ is understood. We will also not distinguish $\|\cdot\|_{\mathcal{X}^{r,\delta}}$ and $\|\cdot\|_{C^r}$ if the space of the analytic function is understood. Since for $f : \mathbb{T} \to \mathbb{T}$, $f \in C^r$ implies $f \in \mathcal{X}^{r,\delta}$ and we have $\|f\|_{C^r} = \|f\|_{\mathcal{X}^{r,\delta}}$.

4.2. **Basic Properties of C^r and $\mathcal{X}^{r,\delta}$ Spaces.**

4.2.1. *Inequalities for Basic Operations.* In this subsection, we present some basic properties and inequalities in the C^r space.

Lemma 1 (Inequalities in C^r Space). *For $\phi, \psi, a \in C^r$, where $r \geq 1$, and $a : \mathbb{T} \to \mathbb{T}$ is a diffeomorphism, we have the following inequalities [dlLO99]:*

(1) $H_\alpha(\phi \circ a) \leq H_\alpha(\phi)\|Da\|_{C^0}^\alpha$,
(2) $H_\alpha(\phi \cdot \psi) \leq \|\phi\|_{C^0} H_\alpha(\psi) + H_\alpha(\phi)\|\psi\|_{C^0}$,
(3) $\|\phi \cdot \psi\|_{C^r} \leq 2^{2n+1} \|\phi\|_{C^r} \|\psi\|_{C^r}$,
(4) $\|\phi \circ \psi\|_{C^r} \leq M_r \|\phi\|_{C^r} (1 + \|\psi\|_{C^r}^r) \leq 2M_r \|\phi\|_{C^r} \|\psi\|_{C^r}^r$, *where $M_r \geq 1$.*

Remark 22. If $a : \mathbb{T} \to \mathbb{T}$ is only α-Hölder for $\alpha < 1$, the Hölder space is not preserved under composition and the best that we can have is $H_{\alpha\beta}(\phi \circ a) \leq H_\alpha(\phi)H_\beta(a)^\alpha$.

Based on Lemma 1, we can further derive the following inequalities. These inequalities will be used in the estimation in Section 6. We extract them here as an extension to [dlLO99] and we hope they can also be used in other applications.

Lemma 2 (More Inequalities in C^r Space). *For $\phi, \psi, a \in C^r$, where $a : \mathbb{T} \to \mathbb{T}$ is a diffeomorphism. We assume $k, p, q \in \mathbb{N}^+$. The inequalities are as follows:*

(1) $H_\alpha(D^p a \circ a^{\circ k}) \leq H_\alpha(D^p a)\|Da\|_{C^0}^{k\alpha}$,
(2) $H_\alpha(D(a^{\circ k})) \leq k H_\alpha(Da)\|Da\|_{C^0}^{k(\alpha+1)-1}$,
(3) $\|a^{\circ k}\|_{C^r} \leq k^n n! \|Da\|_{C^0}^{r(k-1)} \|a\|_{C^r}^{r+1}$,
(4) $\|\phi(a^{\circ k})\|_{C^r} \leq n! k^{n-1}(n + nk + 1)\|\phi\|_{C^r} \|a\|_{C^r}^{r+1} \|Da\|_{C^0}^{kr}$,
(5) $\|\psi^{[k]}\|_{C^r} \leq k^{n+1}(n+1)!(\|\psi\|_{C^r} + \|a\|_{C^r})^{r+1} \|\psi\|_{C^0}^{\max(0,k-n-1)} \|Da\|_{C^0}^{kr}$, *where $\|\psi\|_{C^0} < 1$, and as in equation 8, $\psi^{[k]} = \psi(a^{\circ(k-1)}) \cdots \psi(a)\psi$,*
(6) *If $\|\psi\|_{C^0} < 1$, $\|\phi(a^{\circ k})\psi^{[k]}\|_{C^r} \leq C_r\|\phi\|_{C^r}(\|\psi\|_{C^r} + \|a\|_{C^r})^{r+1} \|\psi\|_{C^0}^{-n} k^r(\|\psi\|_{C^0}\|Da\|_{C^0}^r)^k$,*
(7) $\|\phi^k\|_{C^r} \leq k^{2(n-1)}\|\phi\|_{C^r}^{\min(k,r)} \|\phi\|_{C^0}^{\max(k-n-1,0)}$.

Proof. By Lemma 1, we have

(1) $H_\alpha(D^p a \circ a^{\circ k}) \leq H_\alpha(D^p a)\|Da^{\circ k}\|_{C^0}^\alpha \leq H_\alpha(D^p a)\|Da\|_{C^0}^{k\alpha}$,
(2) $H_\alpha(Da^{\circ k}) \leq k\|Da\|_{C^0}^{k-1} \max_{0 \leq j \leq k} H_\alpha(Da \circ a^{\circ j}) \leq k H_\alpha(Da)\|Da\|_{C^0}^{k(\alpha+1)-1}$,
(3) Suppose $D^p(a^{\circ k})$ has T_p terms, each term has F_p factors, then by

$$F_{p+1} \leq F_p + k - 1, T_{p+1} \leq T_p F_p \text{ and } F_1 = k, T_1 = 1,$$

we have $F_n \leq nk, T_n \leq k^n(n-1)!$, for the same $n = \lfloor r \rfloor$.

In each term, at most $n(k-1)$ factors are $Da \circ a^{\circ q}$, at most n factors are $D^p(a) \circ a^{\circ q}$, where $0 \leq p \leq n, 0 \leq q \leq k$.

Thus we have

$$\|D^n a^k\|_{C^0} \leq k^n(n-1)!\|a\|_{C^n}^n \|Da\|_{C^0}^{n(k-1)}.$$

We also have

$$H_\alpha(D^n a^{\circ k}) \leq k^n (n-1)! H_\alpha(\text{each term in } D^n a^k)$$

$$\leq k^n (n-1)! \Big(n \|Da\|_{C^0}^{n(k-1)} \|a\|_{C^n}^{n-1} \max_{0 \leq p \leq n, 0 \leq q \leq k} H_\alpha(D^p a \circ a^{\circ q})$$

$$+ n(k-1) \|Da\|_{C^0}^{n(k-1)-1} \max_{0 \leq q \leq k} H_\alpha(Da \circ a^{\circ q}) \|a\|_{C^n}^n \Big)$$

$$\leq k^n (n-1)! \Big(n \|Da\|_{C^0}^{n(k-1)} \|a\|_{C^n}^{n-1} H_\alpha(D^n a) \|Da\|_{C^0}^{k\alpha}$$

$$+ n(k-1) \|Da\|_{C^0}^{n(k-1)-1} H_\alpha(Da) \|Da\|_{C^0}^{k\alpha} \|a\|_{C^n}^n \Big)$$

$$\leq k^{n+1} n! \|Da\|_{C^0}^{kr-n} \|a\|_{C^r}^{r+1}$$

Above all, we have

$$\|a^{\circ k}\|_{C^r} \leq \max \Big(\|a^{\circ k}\|_{C^n}, H_\alpha(D^n a^{\circ k}) \Big) \leq k^n n! \|Da\|_{C^0}^{r(k-1)} \|a\|_{C^r}^{r+1}.$$

(4) By the same notation and same method as (3), we have $F_n \leq (n+1)k, T_n \leq n! k^{n-1}$. In each term, at most n factors of $D^p a \circ a^{\circ q}$, at most nk factors of $Da \circ a^{\circ q}$ and there is a term of $D^p(\phi) \circ a^k$, where $0 \leq p \leq n, 0 \leq q \leq k$.

It follows that

$$\|D^n[\phi(a^{\circ k})]\|_{C^0} \leq n! k^{n-1} \|\phi\|_{C^n} \|a\|_{C^n}^n \|Da\|_{C^0}^{nk},$$

and

$$H_\alpha(D^n[\phi(a^{\circ k})]) \leq n! k^{n-1}(n+nk+1) \|a\|_{C^r}^{r+1} \|Da\|_{C^0}^{kr} \|\phi\|_{C^r}.$$

Thus, we have

$$\|\phi(a^{\circ k})\|_{C^r} \leq n! k^{n-1}(n+nk+1) \|\phi\|_{C^r} \|a\|_{C^r}^{r+1} \|Da\|_{C^0}^{kr}.$$

(5) By running the same analysis on $\psi^{[k]}$, we have $F_n \leq k(n+1), T_n \leq k^n n!$, and for each term, there are at least $\max(k-n, 0)$ factors of ψ, at most n factors of either $D^p a \circ a^{\circ q}$ or $D^p \psi \circ a^{\circ q}$, at most nk factors of $Da \circ a^{\circ q}$, where $0 \leq p \leq n, 0 \leq q \leq k$.

It follows that

$$\|D^n \psi^{[k]}\|_{C^0} \leq k^n n! \|\psi\|_{C^0}^{\max(k-n,0)} (\|\psi\|_{C^n} + \|a\|_{C^n})^n \|Da\|_{C^0}^{nk},$$

and

$$H_\alpha(D^n \psi^{[k]}) \leq k^{n+1}(n+1)! \|\psi\|_{C^0}^{\max(0,k-n-1)} (\|\psi\|_{C^r} + \|a\|_{C^r})^{r+1} \|Da\|_{C^0}^{kr}.$$

Above all,

$$\|\psi^{[k]}\|_{C^r} \leq k^{n+1}(n+1)! (\|\psi\|_{C^r} + \|a\|_{C^r})^{r+1} \|\psi\|_{C^0}^{\max(0,k-n-1)} \|Da\|_{C^0}^{kr}.$$

(6) Since

$$D^n(\phi(a^{\circ k}) \psi^{[k]}) = \sum_{q=0}^{n} \binom{n}{q} D^{n-q} \phi(a^k) D^q \psi^{[k]},$$

and with the previously derived results, we have (with the tedious computation omitted), that

$$\|\phi(a^{\circ k}) \psi^{[k]}\|_{C^r} \leq C_r \|\phi\|_{C^r} (\|\psi\|_{C^r} + \|a\|_{C^r})^{r+1} \|\psi\|_{C^0}^{-n} k^r (\|\psi\|_{C^0} \|Da\|_{C^0}^r)^k,$$

where C_r is formed by only the power series and factorials of r.

(7) As for $\|\phi^k\|_{C^r}$, we know $D^n(\phi^k)$ has k^{n-1} terms, each term has k factors, and each term has at most $\min(k,n)$ factors of $D^p\phi$ with the rest of the terms are ϕ, we have

$$\|\phi^k\|_{C^r} \le k^{2(n-1)}\|\phi\|_{C^r}^{\min(k,r)}\|\phi\|_{C^0}^{\max(k-n-1,0)}.$$

$\square$

Lemma 1 also implies the following inequality in $\mathcal{X}^{r,\delta}$ space.

Lemma 3 (Inequalities in $\mathcal{X}^{r,\delta}$ space). *Given $f, g \in \mathcal{X}^{r,\delta}$ we have*

- $\|f \cdot g\|_{\mathcal{X}^{r,\delta}} \le 2^{2n+1}\|f\|_{\mathcal{X}^{r,\delta}}\|g\|_{\mathcal{X}^{r,\delta}}$,

By [dlLO99], the $C^r(\mathbb{T}, X)$ space, thus the $\mathcal{X}^{r,\delta}$ space we are considering in this paper admits a scale of Banach Spaces with continuous inclusion. In other word, for $0 \le r \le s$, we have $C^s(U, X) \subset C^r(U, X)$ and $\mathcal{X}^{s,\delta} \subset \mathcal{X}^{r,\delta}$.

Remark 23. Generally speaking, the scale of spaces $C^r(U, X)$ does not admits continuous inclusion for general domain U (counterexample can be found in [dlLO99]). The continuous inclusion is guaranteed when U is a compensated open set [dlLO99]. In our case, the domain of the functions is torus, which is a simple compensated open set.

4.2.2. *Smoothing Operators.* To develop the Nash-Moser smoothing technique, for a scale of Banach spaces $\mathcal{X}^{r,\delta}$, we need the existence of a family of smoothing operators defined as follows:

Definition (Smoothing Operator). For a scale of Banach spaces X_r, a family of smoothing operators $\{S_t\}_{t\in\mathbb{R}^+}$ satisfies

$$(29) \qquad \|S_t u\|_\mu \le t^{\mu-\lambda}C_{\lambda,\mu}\|u\|_\lambda \text{ for } u \in X_\lambda$$

and

$$(30) \qquad \|(S_t - I)u\|_\lambda \le t^{-(\mu-\lambda)}C_{\lambda,\mu}\|u\|_\mu \text{ for } u \in X_\mu$$

for $\mu \ge \lambda \ge 0$, where t is the strength of smoothing.

Remark 24. When X is a Banach space, the existence of the C^r-smoothing operator in $C^r(\mathbb{T}, X)$ is studied in [Zeh75].

With such smoothing operator in C^r space, we can define the smoothing operator for a function $u(\theta, s) = \sum_{j=0}^\infty u^{(j)}(\theta)s^j \in \mathcal{X}^{r,\delta}$ by smoothing each of $u^{(j)}(\theta)$ for $j \ge 0$. More precisely, we have

Definition (Smoothing Operator in $\mathcal{X}^{r,\delta}$). For $u(\theta, s) = \sum_{j=0}^\infty u^{(j)}(\theta)s^j \in \mathcal{X}^r$, the smoothing operator S_t is defined as follows:

$$(31) \qquad S_t u(\theta, s) = \sum_{j=0}^\infty \hat{S}_t u^{(j)}(\theta)s^j.$$

where $\hat{S}_t$ is the smoothing operator in C^r space defined in Remark 24.

In our problem, since (4) has unknowns which are triples of functions, (W, a, λ), we will see that the smoothing operators defined so far, lead straightforwardly to smoothing operators in the space of triples. See Section 4.3.

Note that the definition of smoothing in $\mathcal{X}^{r,\delta}$ defined above is the standard C^r smoothing applied spaces of C^r functions taking values in a space of analytic functions as discussed in Remark 20.

It is standard to see that this operator S_t defined in (31) satisfies condition (29) and (30), thus it is indeed a smoothing operator in $\mathcal{X}^{r,\delta}$.

Remark 25. As shown in [Zeh75, dlL01, dlLO99], the existence of the smoothing operators implies the interpolation inequality, which is for any $0 \leq \lambda \leq \mu$, $0 \leq \gamma \leq 1$, and $v = (1 - \gamma)\lambda + \gamma\mu$, we have

$$(32) \qquad\qquad \|u\|_v \leq C_{\gamma,\lambda,\mu} \|u\|_\lambda^{1-\gamma} \|u\|_\mu^\gamma.$$

Obtaining (32) as a corollary of smoothing, leads to the conclusion only in the case that v is not an integer. In [dlLO99], there is a direct proof regarding this in greater generality.

4.3. **The $\mathscr{X}^{r,\delta}$ and $\mathscr{Y}^{r,\delta}$ Space.** Our problem of solving (4) seeks triples of functions (the embedding W, the inner dynamics in the circle a and the dynamics on the stable manifolds λ). We will need spaces of triple of functions. In this section, we specify the topologies we have found useful.

We now can define the scale of spaces $\mathscr{X}^{r,\delta}$ and $\mathscr{Y}^{r,\delta}$ by the product of Banach spaces as follows:

Definition. Define the product space $\mathscr{X}^{r,\delta} = \mathcal{X}^{r,\delta} \times \mathcal{X}^{r,\delta} \times C^r \times C^r$ with norm

$$\|u\|_{\mathscr{X}^{r,\delta}} = \|W_1\|_{\mathcal{X}^{r,\delta}} + \|W_2\|_{\mathcal{X}^{r,\delta}} + \|a\|_{C^r} + \|\lambda\|_{C^r},$$

where $u = (W_1(\theta, s), W_2(\theta, s), a(\theta), \lambda(\theta)) \in \mathscr{X}^{r,\delta}$. Similarly, define space $\mathscr{Y}^{r,\delta} = \mathcal{X}^{r,\delta} \times \mathcal{X}^{r,\delta}$ with norm

$$\|v\|_{\mathscr{Y}^{r,\delta}} = \|W_1\|_{\mathcal{X}^{r,\delta}} + \|W_2\|_{\mathcal{X}^{r,\delta}},$$

where $v = (W_1(\theta, s), W_2(\theta, s)) \in \mathscr{Y}^{r,\delta}$.

Remark 26. $\mathscr{X}^{r,\delta}, \mathscr{Y}^{r,\delta}$ are both scales of Banach spaces with smoothing operators. The smoothing operators comes natually from the smoothing operators in C^r and $\mathcal{X}^r$ spaces.

Remark 27. For the rest of the paper, we will always denote $u(\theta, s) \in \mathscr{X}^{r,\delta}$ to be the triplet $(W(\theta, s), a(\theta), \lambda(\theta))$, and we will not distinguish among $\|\cdot\|_{\mathscr{X}^{r,\delta}}, \|\cdot\|_{\mathscr{Y}^{r,\delta}}, \|\cdot\|_{C^r}$ and $\|\cdot\|_r$ when δ and the dimension of the function are understood.

5. Statement of The Analytical Result

In this section, we present the statement of the main result: Theorem 4.

As anticipated, the proof is obtained through a Nash-Moser method, alternating the quasi-Newton method with some smoothing steps. As discussed in Section 1.2, the problem at hand is somewhat different from other previous applications of Nash-Moser technique. The loss of differentiability in the estimates comes from the operator in the functional. The solutions of the linearized equation do not lose regularity, but they only work for a range of regularities.

Since the Nash-Moser method requires alternating the quasi-Newton method and smoothings, we start formulating the standard setup. This is a scale of Banach spaces. The smoothing operators map the spaces of less regular functions into the spaces of more regular functions and they have quantitative properties.

By the scale of spaces and the smoothing operators in Section 4, we formulate our main result Theorem 4 and proceed to the proof in Section 6.3. Theorem 4 implies rather directly the result for foliations. We just need to verify that the operator entering in equation (4) satisfies the hypotheses of Theorem 4.

As indicated in Section 1.2, the implicit function theorem we use will require some unusual properties in Nash-Moser theory: We need spaces with anisotropic regularity, the linearized equation does not incur any loss of regularity, but can only be applied in a range of regularities. This will require some severe adaptations from the standard expositions and the methods based on analytic or C^∞ smoothing cannot work here.

Recall that our goal is to find $W(\theta, s), a(\theta)$ and $\lambda(\theta)$ satisfying the invariance equation (4). In other words, given $r \geq 0, \delta > 0$, we are looking for the zero of the functional $\mathscr{F} : \mathscr{X}^r \to \mathscr{Y}^r$ where

$$(33) \qquad \mathscr{F}[u] = \mathscr{F}[W, a, \lambda](\theta, s) = f(W(\theta, s)) - W(a(\theta), \lambda(\theta)s),$$

for $u = (W, a, \lambda) \in \mathscr{X}^{r,\delta}$.

Before presenting the main Theorem 4, we first define **Condition-0** as follows:

Definition (Condition-0). For any sufficiently small $\delta, \rho > 0$. Given $m \in \mathbb{R}$, $W : \mathbb{T} \times \mathbb{R} \to \mathbb{T} \times \mathbb{R}$, $a : \mathbb{T} \to \mathbb{T}$ and $\lambda : \mathbb{T} \to \mathbb{R}$, we say that the tuple (m, W, a, λ) satisfies **Condition-0** if the following restrictions hold:

   (1) $\|\lambda\|_{C^0} < 1$,
   (2) $(W, a, \lambda) \triangleq u \in \mathscr{X}^{m+2,\delta}$,
   (3) For $\widetilde{B}_{m+2}(\rho) \subset \mathscr{X}^{m+2}$ is the ball centered at $u = (W, a, \lambda)$ with radius ρ,

$$\min_{u \in \widetilde{B}_{m+2}(\rho)} \min \left(-\frac{\ln \|\lambda\|_{C^0} \|(Da)^{-1}\|_{C^0}}{\ln \|Da\|_{C^0}}, -\frac{\ln \|\lambda\|_{C^0}}{\ln \|D(a^{-1})\|_{C^0}}, -\frac{\ln \|\lambda\|_{C^0}}{\ln \|Da\|_{C^0}} \right) - 2 \geq m \geq 2.$$

Remark 28. Restriction (1) can be generalized to $\lambda^* < 1$, where λ^* is The Dynamical average. If λ^* is used, one also need to adapt condition (3) accordingly (see Remark 37).

Remark 29. Restriction (3) is to guarantee m is bounded above in such a way that the regularity requirement for solving cohomological equations (21), (22) and (24) covers the scale of regularities in Theorem 4. (See Lemma 7 for more details).

Following the scheme derived in Section 3, we present a theorem for the existence of solution for $\mathscr{F}[u] = 0$:

Theorem 4. *For sufficiently small $\delta > 0, \rho > 0$, suppose there exists a tuple (m, W_0, a_0, λ_0) satisfying Condition-0.*

Let $\mathscr{X}^{r,\delta}$ and $\mathscr{Y}^{r,\delta}$ be two scales of Banach spaces for $m \leq r \leq m + 2$.

Consider the functional $\mathscr{F} : \widetilde{B}_r(\rho) \to \mathscr{Y}^r$ defined in (33), where $\widetilde{B}_r(\rho)$ is a ball centered at $u_0 \triangleq (W_0, a_0, \lambda_0) \in \mathscr{X}^{m+2,\delta}$ with radius ρ.

If $\|\mathscr{F}[u_0]\|_{\mathscr{X}^{m-2,\delta}}$ is sufficiently small, then there exists $u^* \in \widetilde{B}_m(\rho)$ such that $\mathscr{F}[u^*] = 0$.

Moreover, such u^* is the limit of the iteration combining with some smoothing operators. The smoothing parameters go to zero, and the specific rates will be given in the proof. Furthermore, the convergence of the iterations to the limit is superexponential.

As a consequence, we have that

$$\|u^* - u_0\|_{\mathscr{X}^{m,\delta}} \leq C \|\mathscr{F}(u_0)\|_{\mathscr{X}^{m-2,\delta}},$$

where C is a finite constant.

Remark 30. More specifically, the restriction for $\|\mathscr{F}[u_0]\|_{m-2}$ to be sufficiently small is:

$$\|\mathscr{F}[u_0]\|_{m-2} < e^{-2\mu\beta},$$

where μ, β are numbers specified in the proof of Appendix A. The converging rate for the iteration scheme is bounded by $\|\mathscr{F}[u_n]\|_{m-2} \leq ve^{-2\mu\beta\kappa^n}$, with the same μ and β, and κ can be picked to be as close to 2 as possible.

Remark 31. It may seem somewhat surprising that the requirement on $\|\mathscr{F}[u_0]\|_{\mathscr{X}^{m-2,\delta}}$ from lower regularity can result in the existence of solution u^* in higher regularity $\mathscr{X}^{m,\delta}$, but it is actually reasonable because of the requirement from even higher regularity that $u_0 \in \mathscr{X}^{m+2,\delta}$.

Remark 32. Since δ prescribes the range of s, picking a larger δ allows us to parameterize a larger neighborhood of the invariant circle provided that the conditions in Theorem 4 are maintained with the increased δ.

Remark 33. One of the consequences of (4) is that given a family of maps f_ε indexed by a parameter ε so that f_0 contains an invariant circle, we can design a continuation method by taking the exact solution for some value of ε as an approximate solution for $\varepsilon + \eta$ for sufficiently small η [YdlL21].

This procedure is guaranteed to continue till some of the non-degeneracy assumptions of Theorem 4 fail. These assumptions are just the regularity of the circle and some version of hyperbolicity. Hence, we know that these numerical methods will continue till the torus becomes irregular, the manifolds have a domain of analyticity smaller than δ or the hyperbolicity is lost. This may entail that the dynamical average gets close to 1 (or undefined) or that the angle between the stable and unstable manifolds becomes zero (the bundle collapse). Of course, several of the possibilities may happen at the same time.

Remark 34. As seen in several examples (e.g. in [dlL97]) one can see that the optimal regularity of the invariant circle may decrease continuously to 0 as the parameters change. For some parameter value, they will stop being C^2, for another parameter they will stop being C^1, and then, they will become Hölder continuity. (The isochrons remain analytic, even if the optimal domain may change).

This indicates that the breakdown of the tori may depend on what regularity one requires to call something a torus. The fact that the destruction of the tori happens in a very gradual way makes the exploration of the boundary be very subtle since the boundary detected depends significantly on the stopping criterion. For example, the destruction of the circles as C^1 manifolds studied in [Mn78] happens at different values of the places where they disappear as C^0 curves or as continua [JK69, CK20]. Detailed discriptions of the breakdown can be found in [YdlL21].

Detailed numerical explorations of the behavior at breakdown of the hyperbolicity [GE88, Ran92a, Ran92b, HdlL06a, HdlL07, CF12, FH12] has uncovered many interesting phenomena (e.g. scaling relations) that deserve detailed mathematical analysis.

Of course, detailed numerical explorations near the boundary are very delicate and it requires having a very good theory (condition numbers and a-posteriori theorems) that ensure that the calculations are correct even when something unexpected is happening.

The proof of Theorem 4 is done by verifying the conditions of Theorem 13. introduced in Appendix A. Details of the proof of Theorem 4 can be found in section 6.3.

5.1. **The Analyticity Radius for** $W(\theta, s)$**: a Digression.** In general, analytic radius for $W(\theta, s)$ is not infinite (for example, systems with more than one limit cycle) since Algorithm 1 computes the invariant circle and the foliations by stable manifold in a small neighborhood of the limit cycle. On the other hand, following [Poi90], if the map f is entire, we have the following result:

Lemma 5. *If the map* $f : \mathbb{T} \times \mathbb{R} \to \mathbb{T} \times \mathbb{R}$ *is an entire function, and given* $(W(\theta, s), a(\theta), \lambda(\theta))$ $\in \mathscr{X}^{r,\delta}$ *satisfies the invariance equation* (4)*, we have that the analytic radis for* $W(\theta, s)$ *in* s *is infinity.*

Proof. It follows from $W(\theta, s) \in \mathscr{X}^{r,\delta}$ that $W(\theta, \cdot)$ is analytic in $B_{\rho(\theta)}$, where $\rho(\theta) \geq \rho_* > 0$. Since f in entire, we also have $f \circ W(\theta, \cdot)$, thus $W(a(\theta), \lambda(\theta) \cdot)$ by Equation (4), is analytic in $B_{\rho(\theta)}$. It follows that $W(a(\theta), \cdot)$ is analytic in $B_{\lambda^{-1}(\theta)\rho(\theta)}$.

By repeating the above process, one can see that

$$W(a^{\circ m}(\theta), \cdot) \text{ is analytic in } B_{\lambda^{[m]}(\theta)^{-1}\rho(\theta)}.$$

It follows that the analyticity radius for $W(\theta, \cdot)$ is infinity.
$\square$

6. Proof of the Analytical Result

This section can be mainly divided into 2 parts. In the first part, we prove two technical results. More specifically, we prove the fibered version of the Poincaré-Sternberg theorem, namely the existence of $h(\theta, s)$ in equation (3) discussed in Section 2.3 (see Section 6.1), and we prove the existence of the solution to the cohomological equation mentioned in (25) (see Section 6.2). In the second half, we present the proof of the Theorem 4. The idea of the proof is presented in Section 5. The proof is achieved by justifying all the non-degeneracy conditions that are listed in a modified version of the Nash-Moser implicit function theorem (Theorem 13), which can be found in Appendix A.

6.1. **The Existence of** $h(\theta, s)$ **in Equation** (3)**.** In this subsection, we prove Lemma 6. As indicated in Section 2.2. Lemma 6 ensures that the study of (4), which clearly is a sufficient condition for the existence of foliation, is also necessary. This result will not be used in subsequent studies of the existence of solutions of (4). Nevertheless, it introduces some techniques that will be used later. It also allows us to make some remarks about the domains of solutions of functional equations.

Our goal is finding $h(\theta, s)$ such that equation (3) holds for given $\lambda(\theta, s)$ and $\widehat{\lambda}(\theta, s)$. In the following Lemma 6, we show the existence of $h(\theta, s)$ when $\widehat{\lambda}(\theta, s)$ equals to the linear term of $\lambda(\theta, s)$ by the contraction mapping theorem.

Lemma 6 (Existence of $h(\theta, s)$). *There exists* $\delta > 0$ *such that for* $\widehat{\lambda} \in \mathscr{X}^{r,\delta}, \widehat{\lambda}(\theta, s) = \lambda(\theta)s + N(\theta, s)$, *where* $N(\theta, s) = \mathcal{O}(s^2)$. *If there exists* $k \in \mathbb{N}^+$ *such that* $\|\lambda^{[k]}\|_{C^0} < 1$ *and* $\|\lambda^{[k]}\|_{C^r} < \gamma_k$ *for some* $k \in \mathbb{N}^+$*, where* γ_k *is specified in the proof, then we have the existence of* $h(\theta, s) \in \mathscr{X}^{r,\delta}$ *such that equation* (3)*:* $h(a(\theta), \lambda(\theta)s) = \widehat{\lambda}(\theta, h(\theta, s))$ *holds.*

Remark 35. The condition $\|\lambda^{[k]}\|_{C^0} < 1$ for some $k \in \mathbb{N}^+$ can be assured when the dynamical average $\lambda^* < 1$, and the condition $\|\lambda^{[k]}\|_{C^r}$ can be maintained with a suitable choice of initial condition u_0 as in Theorem 4.

Remark 36. This Lemma 6 can be viewed as a *"fibered"* version of the Poincaré-Sternberg theorem on linearization of contractions. We can think of s as the dynamic variable but the map sends a fiber indexed by θ into another fiber indexed by $a(\theta)$.

We have prepared a proof following the version of [Ste57] based on formulating as contractions since it leads to concrete estimates. Since the maps are analytic in the dynamical variable, the original proof of Poincaré-Dulac [Poi79, Dul03] based on majorants can also be adapted.

Proof. By substituting the above $\widehat{\lambda}(\theta, s)$ and $\lambda(\theta)$ in equation (3), we have

$$(34) \qquad h(a(\theta), \lambda(\theta)s) = \lambda(\theta)h(\theta, s) + N(\theta, h(\theta, s)).$$

Since we only need the existence of h, we restrict ourselves for finding $h(\theta, s)$ of the following form:

$$(35) \qquad h(\theta, s) = s + \widehat{h}(\theta, s),$$

where $\widehat{h}(\theta, s) = \mathcal{O}(s^2)$. By substituting (35) back into equation (34) and after some simplifications, we have

$$(36) \qquad \widehat{h}(\theta, s) = \lambda^{-1}(\theta)[\widehat{h}(a(\theta), \lambda(\theta)s) - N(\theta, s + \widehat{h}(\theta, s))].$$

Define a Banach space

$$\widetilde{\mathcal{X}}^{r,\delta} = \left\{ u(\theta, s) = \sum_{j=2}^{\infty} u^{(j)}(\theta)s^j \mid u^{(j)}(\theta) \in C^r, \text{ and } \sum_{j=2}^{\infty} \left\| u^{(j)} \right\|_{C^r} \delta^j < \infty \right\}.$$

We know $\widetilde{\mathcal{X}}^{r,\delta}$ is complete as it is a closed subspace of $\mathcal{X}^{r,\delta}$, and $N(\theta, s), \widetilde{h}(\theta, s) \in \widetilde{\mathcal{X}}^{r,\delta}$.

Denote

$$\mathscr{G}[\widehat{h}] = \lambda^{-1}(\theta)[\widehat{h}(a(\theta), \lambda(\theta)s) - N(\theta, s + \widehat{h}(\theta, s))],$$

then $\mathscr{G} : \widetilde{\mathcal{X}}^{r,\delta} \to \widetilde{\mathcal{X}}^{r,\delta}$. The task now is to show the existence of $\widehat{h}$ such that $\mathscr{G}(\widehat{h}) = \widehat{h}$ through a contraction argument.

Instead of showing that $\mathscr{G}$ is a contraction, we show $\mathscr{G}^{\circ k}$ ($\mathscr{G}$ compose with itself k times for some big enough integer k) is a contraction.

By simple calculations, one can see that

$$(37) \qquad \mathscr{G}^{\circ k}[\widehat{h}] = (\lambda^{-1})^{[k]}(\theta)[\widehat{h}(a^{\circ k}(\theta), \lambda^{[k]}(\theta)s) - k\mathcal{O}(s^2)],$$

where the second term $k\mathcal{O}(s^2)$ is formed by the summation of n terms of $N(\cdot, \cdot)$, each is of $\mathcal{O}(s^2)$, which can be controlled to be small by some upper bound δ_o since $|s| < \delta < \delta_0$.

It remains to show that the first term of (37): $(\lambda^{-1})^{[k]}(\theta)\widehat{h}(a(\theta), \lambda^{[k]}(\theta)s) \triangleq \mathscr{L}[\widehat{h}]$ is a contraction. For every $\hat{h}_1, \hat{h}_2 \in \widetilde{\mathcal{X}}^r$, we have $\hat{h}_1(\theta, s) = \sum_{j=2}^{\infty} \hat{h}_1^{(j)}(\theta)s^j$ and $\hat{h}_2(\theta, s) =$

$\sum_{j=2}^{\infty} \hat{h}_2^{(j)}(\theta)s^j$. We have

$$
\begin{aligned}
\left\| \mathscr{L}[\widehat{h}_1] - \mathscr{L}[\widehat{h}_2] \right\|_{\mathcal{X}^{r,\delta}} &= \left\| (\lambda^{-1})^{[k]}(\theta)(\hat{h}_1(a^{\circ k}(\theta), \lambda^{[k]}(\theta)s) - \hat{h}_2(a^{\circ k}(\theta), \lambda^{[k]}(\theta)s))) \right\|_{\mathcal{X}^{r,\delta}} \\
&\leq \left\| \sum_{j=2}^{\infty} (\hat{h}_1^{(j)} - \hat{h}_2^{(j)})(a^{\circ k}(\theta))\lambda^{[k-1](j-1)}(\theta)s^j \right\|_{\mathcal{X}^{r,\delta}} \\
&\leq C_{r,k,\|a\|_{C^r},\|\lambda\|_{C^0}} \left\| \lambda^{[k]} \right\|_{C^r}^r \sum_{j=2}^{\infty} \left\| \hat{h}_1^{(j)} - \hat{h}_2^{(j)} \right\|_{\mathcal{X}^{r,\delta}} \delta^j \\
&\leq \zeta \left\| \hat{h}_1 - \hat{h}_2 \right\|_{\mathcal{X}^{r,\delta}}
\end{aligned}
$$

provided that $\|\lambda^{[k]}\|_{C^r} < (\zeta C_{r,k,\|a\|_{C^r},\|\lambda\|_{C^0}}^{-1})^{\frac{1}{r}} \triangleq \gamma_k$ for any $0 < \zeta < 1$, where the second last inequality is achieved by utilizing Lemma 2 and $C_{r,k,\|a\|_{C^r},\|\lambda\|_{C^0}}^{-1} > 0$ is a constant related to r, k, $\|a\|_{C^r}$ and $\|\lambda\|_{C^0}$ only.

By the above discussion, we have the existence of a unique $\hat{h}^* \in \tilde{X}^r$ such that $\mathscr{G}(\hat{h}^*) = \hat{h}^*$, which finishes the proof.

$\square$

6.2. **Estimates on Solutions of the Cohomological Equation** (25)**.** We use this subsection to take a closer look at the cohomological equation mentioned in (25) with solution (27). The following result in Lemma 7 is used in both Section 3.1 and Section 6.

Lemma 7. *Given $l(\theta), a(\theta)$ and $\eta(\theta) \in C^r$ with $\|l\|_{C^0} < 1$. If $r < -\ln \|l\|_{C^0} / \ln \|Da\|_{C^0}$ (i.e. $\|Da\|_{C^0}^r \|l\|_{C^0} < 1$), then the cohomological equation (25): $\phi(\theta) = l(\theta)\phi(a(\theta)) + \eta(\theta)$ admits a unique C^r solution:*

$$
\phi(\theta) = \sum_{j=0}^{\infty} l^{[j]}(\theta)\eta(a^j(\theta)) \tag{38}
$$

with

$$
\|\phi\|_{C^r} \leq C_{l,a,r} \|\eta\|_{C^r} \leq \infty,
$$

Proof. First, we prove that (38) is a solution to equation (25). Since $\|l\|_{C^0} < 1$ and $\|\eta\|_{C^0}$ is bounded, by noticing that $\sum_{j=0}^{\infty} l^{[j]}(\theta)\eta(a^j(\theta))$ converges uniformly in C^0, one can substitute this infinite sum back in (25) and rearrange terms to show that (38) is indeed a solution.

On top of this, we argue that (38) is the only C^0 solution. More explictly, if there were two solutions, then by the discussion in (26): $\phi(\theta) = \sum_{j=0}^{n} l^{[j]}(\theta)\eta(a^j(\theta)) + l^{[n+1]}(\theta)\phi(a^{n+1}(\theta))$, they would agree on the first n terms, and since the limit of the C^0 norm for the last term goes to 0 as n goes to infinity, the two solutions are the same.

To finish the proof, we now show that $\|\phi\|_{C^r} < \infty$. By Lemma 1 and Lemma 2, we have the following inequalities,

(1) $\|a^{\circ k}\|_{C^r} \leq k^n n! \|Da\|_{C^0}^{r(k-1)} \|a\|_{C^r}^{r+1}$,
(2) $\|\eta(a^{\circ k})\|_{C^r} \leq n! k^{n-1}(n + nk + 1)\|\eta\|_{C^r}\|a\|_{C^r}^{r+1}\|Da\|_{C^0}^{kr}$,
(3) $\|l^{[k]}\|_{C^r} \leq k^{n+1}(n+1)!(\|l\|_{C^r} + \|a\|_{C^r})^{r+1}\|l\|_{C^0}^{max(0,k-n-1)}\|Da\|_{C^0}^{kr}$,
(4) $\|\eta(a^{\circ k})l^{[k]}\|_{C^r} \leq C_r(\|l\|_{C^r} + \|a\|_{C^r})^{r+1}\|l\|_{C^0}^{-n}\left[k^n(\|l\|_{C^0}\|Da\|_{C^0}^k)^k\right]\|\eta\|_{C^r}$

Thus from (27), we have

$$\sum_{j=0}^{\infty} \left\| l^{[j]} \eta(a^{\circ j}) \right\|_{C^r} \leq C_r (\|l\|_{C^r} + \|a\|_{C^r})^{r+1} \|l\|_{C^r}^{-n} (\sum_{j=1}^{\infty} j^n (\|Da\|_{C^0}^r \|l\|_{C^0})^j) \, \|\eta\|_{C^r},$$

thus if $r < -\ln \|l\|_{C^0} / \ln \|Da\|_{C^0}$, we have $\|l\|_{C^0} \|Da\|_{C^0}^r < 1$. It follows that

$$\|\phi\|_{C^r} \leq \sum_{j=0}^{\infty} \|l^{[j]} \eta(a^{\circ j})\|_{C^r} < \infty,$$

which finishes the proof. □

Remark 37. Give $k \in \mathbb{N}^+$, by rewriting equation (25) into the form as in (26), i.e.,

$$\phi(\theta) = l^{[k+1]}(\theta)\phi(a^{\circ(k+1)}(\theta)) + \sum_{j=0}^{k} l^{[j]}(\theta)\eta(a^{\circ j}(\theta)),$$

The requirement for r can be generalized slightly to be $r < -\ln \|l^{[k]}\|_{C^0} / \ln \|D(a^{\circ(k+1)})\|_{C^0}$, we have $\left\| l^{[k]} \right\|_{C^0} \left\| D(a^{\circ(k+1)}) \right\|_{C^0}^r < 1$,

Remark 38. Lemma 7 shows that if

(39) $$\|l\|_{C^0} \|Da\|_{C^0}^r < 1 \ (\text{or} \ \left\| l^{[k]} \right\|_{C^0} \left\| D(a^{\circ(k+1)}) \right\|_{C^0}^r < 1,)$$

then we have that $\phi(\theta) = \sum_{j=0}^{\infty} l^{[j]}(\theta)\eta(a^j(\theta))$ converges absolutely in the C^r sense, thus $\phi \in C^r$.

Note that the condition (39) can only be satisfied for a finite range of regularity r. We now give examples to show that this condition is sharp.

If $a(\theta)$ has an attractive fixed point, which we place at $\theta = 0$. If $a(\theta) = \lambda\theta$ in a neigborhood and, moreover $l(\theta)$ is a constant, we see that (27) becomes

(40) $$\phi(\theta) = \sum_{j=0}^{\infty} l^j \eta(\lambda^j \theta))$$

which is a version of the classical Weierstrass function, which for even polynomial η can be arranged to be finite differentiable, showing that the range claimed in Lemma 7 is optimal in the generality claimed. Indeed the map that in local coordinates has the expression $(x, y) = \lambda x, ly + \eta(x)$ has an invariant circle given by the graph of the function ϕ in (40).

The fact that one can only solve the cohomology equations for a certain range of regularities makes it impossible to use the Nash-Moser methods that are based on approximating solutions of C^∞ or C^ω problems.

6.3. **Proof for Theorem 4.** Following the same notation as in Theorem 4, we now justify the non-degeneracy conditions of the abstract Nash-Moser Theorem 13 one by one.

Lemma 8 (Condition 1). *For $\delta, \widetilde{B}_r(\rho)$ defined in Theorem 4, we have $\mathscr{F}(\widetilde{B}_r(\rho) \cap \mathscr{X}^{r,\delta}) \subset \mathscr{Y}^{r,\delta}$.*

Proof. For every $u(\theta, s) = (W(\theta, s), a(\theta), \lambda(\theta)) \in \widetilde{B}_r(\rho) \cap \mathscr{X}^{r,\delta}$, recall $\widetilde{B}_r(\rho)$ is a ball with radius ρ, we have

$$\mathscr{F}[u](\theta, s) = f \circ W(\theta, s) - W(a(\theta), \lambda(\theta)s).$$

First, we show that $f \circ W(\theta, s) \in \mathscr{Y}^r$. With no loss of generality, we will only consider the first component of $f = (f_1, f_2)$ and show that $\|f_1(W_1, W_2)\|_{\mathscr{X}^{r,\delta}} < \infty$.

Write

$$f_1(\theta, s) = \sum_{j=0}^{\infty} f_1^{(j)}(\theta)s^j, \ W_1(\theta, s) = \sum_{j=0}^{\infty} W_1^{(j)}(\theta)s^j, \ W_2(\theta, s) = \sum_{j=0}^{\infty} W_2^{(j)}(\theta)s^j.$$

Notice that

$$f_1^{(j)}(W_1(\theta, s)) = f_1^{(j)}(W_1^{(0)}(\theta)) + \left(\frac{d}{d\theta}f_1^{(j)}\right)(W_1^{(0)}(\theta))\left(\sum_{j=1}^{\infty} W_1^{(j)}(\theta)s^j\right) + \dots$$

$$+ \frac{1}{k!}\left(\frac{d^k}{d\theta}f_1^{(j)}\right)(W_1^{(0)}(\theta))\left(\sum_{j=1}^{\infty} W_1^{(j)}(\theta)s^j\right)^k + \dots$$

since $f^{(j)}(\theta)$ is analytic, we can treat it as a function in $\mathbb{C}$, and then by Cauchy's estimates for derivatives, we have

$$\left|\frac{d^k}{d\theta}(f_1^{(j)})(W_1^{(0)}(\theta))\right| \le \frac{k!}{R^k}\max_{z\in\gamma_R}\left|f_1^{(j)}(z)\right| = \frac{k!}{R^k}C_R, \forall R > 0.$$

where $\gamma_R = \{z \mid |z - W_1^{(0)}(\theta)| = R\}$. It follows that

$$\left\|f^{(j)}(W_1)\right\|_{\mathscr{X}^{r,\delta}} \le C_R(1 + R^{-1}\left\|W_1\right\|_{\mathscr{X}^{r,\delta}} + R^{-2}\left\|W_1\right\|_{\mathscr{X}^{r,\delta}}^2 + \dots)$$

(41)
$$\le C_R\left(\frac{1}{1 - \frac{\|W_1\|_{\mathscr{X}^{r,\delta}}}{R}}\right) \le C_R\frac{1}{1 - \frac{\rho}{R}}$$

Thus

$$\|f_1(W_1, W_2)\|_{\mathscr{X}^{r,\delta}} = \left\|\sum_{j=0}^{\infty} f_1^{(j)}(W_1)(W_2 s)^j\right\|_{\mathscr{X}^{r,\delta}}$$

$$\le \sum_{j=0}^{\infty} \left\|f^{(j)}(W_1)\right\|_{\mathscr{X}^{r,\delta}}\left\|W_2 s\right\|_{\mathscr{X}^{r,\delta}}^j (2^r)^j$$

$$\le C_R\left(\frac{1}{1 - \frac{\rho}{R}}\right)\sum_{j=0}^{\infty}(2^r\rho\delta)^j$$

$$< \infty,$$

where the third line is because of (41) and

$$\|W_2 s\|_{\mathscr{X}^{r,\delta}} = \left\|\sum_{j=0}^{\infty} W_2^{(j)}s^{j+1}\right\|_{\mathscr{X}^{r,\delta}} = \sum_{j=0}^{\infty}\left\|W_2^{(j)}\right\|_{\mathscr{X}^{r,\delta}}\delta^{(j+1)} = \|W_2\|_{\mathscr{X}^{r,\delta}}\,\delta,$$

and the last line is because of the assumption on ρ in Theorem 4.

It remains to show that $\|W(a, \lambda s)\|_r < \infty$, this is trivial since

$$
\begin{aligned}
\|W(a, \lambda s)\|_{\mathscr{X}^{r,\delta}} &= \left\|\sum_{j=0}^{\infty} W^{(j)}(a) \lambda^j s^j\right\|_{\mathscr{X}^{r,\delta}} \\
&\leq \sum_{j=0}^{\infty} \left\|W^{(j)}(a) \lambda^j\right\|_{C^r} \delta^j \leq \sum_{j=0}^{\infty} (2^r)^j \left\|W^{(j)}(a)\right\|_{C^r} \|\lambda\|_{C^r}^j \delta^j \\
&\leq 2 M_r \|a\|_{C^r}^r 2^{2n+1} \max_{0 \leq k \leq r}(\|\lambda\|_{C^r}^k) \max_{0 \leq j < \infty}(\|\lambda\|_{C^0}^{\max(j-n-1,0)} j^{2(n-1)}) \|W\|_{\mathscr{X}^{r,\delta}} \\
&< \infty.
\end{aligned}
$$

for $(W(\theta, s), a(\theta), \lambda(\theta)) \in \widetilde{B}_r(\rho)$ and $\|\lambda\|_{C^0} < 1$. $\qquad\qquad\square$

Lemma 9 (Condition 2). $\mathscr{F}|_{\tilde{B}_m \cap \mathscr{X}^r} : \widetilde{B}_r(\rho) \cap \mathscr{X}^r \to \mathscr{X}^r$ *has continuous first and second order Fréchet derivatives, and satisfy the following conditions:*

* $\|D\mathscr{F}[u](h)\|_{m-2} \leq C_{r,\widetilde{B}_r(\rho)} \|h\|_{m-2}$ *for* $h \in \mathscr{X}^m$.
* $\|D^2\mathscr{F}[u](h)(k)\|_{m-2} \leq C_{r,\widetilde{B}_r(\rho)} \|h\|_{m-1} \|k\|_{m-1}$ *for* $k, h \in \mathscr{X}^m$.

where $C_{r,\widetilde{B}_r(\rho)}$ *is a constant depends on the regularity and the ball* $\widetilde{B}_r(\rho) \in \mathscr{X}^{r,\delta}$ *only.*

Proof. By some routine calculation, for $h = (h_1, h_2, h_3)$, $k = (k_1, k_2, k_3) \in \mathscr{X}^r$, where $h_1, k_1 \in \mathscr{X}^r \times \mathscr{X}^r$, $h_2, h_3, k_2, k_3 \in C^r$, we can calculate the first and second order Fréchet derivatives as follows:

$$
\begin{aligned}
D\mathscr{F}[u](h) &= Df(W)h_1 - \partial_1 W(a, \lambda)h_2 - \partial_2 W(a, \lambda)h_3 - h_1(a, \lambda), \\
D^2\mathscr{F}[u](k, h) &= D^2 f(W)(k_1, h_1) - \partial_{11} W(a, \lambda)(k_2, h_2) - \partial_{12} W(a, \lambda)(k_3, h_2) \\
&\quad - \partial_1 k_1(a, \lambda)h_2 - \partial_{21} W(a, \lambda)(k_2, h_3) - \partial_{22} W(a, \lambda)(k_3, h_3) \\
&\quad - \partial_2 k_1(a, \lambda)h_3 - \partial_1 h_1(a, \lambda)k_2 - \partial_2 h_1(a, \lambda)k_3
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
\|D\mathscr{F}[u](h)\|_{m-2} &\leq 2^{2n+1}(\|Df(W)\|_{m-2}\|h_1\|_{m-2} + \|\partial_1 W(a, \lambda)\|_{m-2}\|h_2\|_{m-2}) \\
&\quad + \|\partial_2 W(a, \lambda)\|_{m-2}\|h_3\|_{m-2}) + \|h_1(a, \lambda)\|_{m-2} \\
&\leq C_{r,\widetilde{B}_r(\rho)}\|h\|_{m-2}.
\end{aligned}
$$

and

$$
\begin{aligned}
\left\|D^2\mathscr{F}[u](k, h)\right\|_{m-2} &\leq \left\|D^2 f(W)(k_1, h_1)\right\|_{m-2} - \|\partial_{11} W(a, \lambda)(k_2, h_2)\|_{m-2} \\
&\quad - \|\partial_{12} W(a, \lambda)(k_3, h_2)\|_{m-2} - \|\partial_1 k_1(a, \lambda)h_2\|_{m-2} \\
&\quad - \|\partial_{21} W(a, \lambda)(k_2, h_3)\|_{m-2} - \|\partial_{22} W(a, \lambda)(k_3, h_3)\|_{m-2} \\
&\quad - \|\partial_2 k_1(a, \lambda)h_3\|_{m-2} - \|\partial_1 h_1(a, \lambda)k_2\|_{m-2} - \|\partial_2 h_1(a, \lambda)k_3\|_{m-2}. \\
&\leq C_{r,\widetilde{B}_r(\rho)}\big(\|h\|_{m-2}\|k\|_{m-2} + \|h\|_{m-1}\|k\|_{m-2} + \|h\|_{m-2}\|k\|_{m-1} \\
&\quad + \|h\|_{m-1}\|k\|_{m-1}\big) \\
&\leq C_{r,\widetilde{B}_r(\rho)}\|h\|_{m-1}\|k\|_{m-1}
\end{aligned}
$$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Lemma 10 (Condition 3). *For $u \in \widetilde{B}_r(\rho)$ and $r = m - 2,\ m + 2$, we have*

$$\|\eta[u]\mathscr{F}[u]\|_r \leq C_{r,\widetilde{B}_r(\rho)} \|\mathscr{F}[u]\|_r.$$

where $\eta[u]$ serves as the approximate inverse of the derivative of the functional $\mathscr{F}[u]$, which is defined in our algorithm in Section 3.

Proof. Note that we only need to apply $\eta[u]$ on the range of $\mathscr{F}[u]$ and we do not need estimates on the whole space. In contrast with other Nash-Moser implicit function theorems, the operator $\eta[u]$ is bounded from spaces to themselves and does not entail any loss of regularity.

From Lemma 7, by equation (21), (22) and (24), we have

$$\|\Gamma_1\|_r \leq C_{r,\widetilde{B}_r(\rho)} \|\widetilde{e}_1\|_r$$

and

$$\|\Gamma_2\|_r \leq C_{r,\widetilde{B}_r(\rho)}(\|\widetilde{e}_1\|_r + \|\widetilde{e}_2\|_r),$$

from equation (20) and (23), we also have

$$\|\Delta_a\|_r \leq C_{r,\widetilde{B}_r(\rho)} \|\widetilde{e}_1\|_r$$

and

$$\|\Delta_\lambda\|_r \leq C_{r,\widetilde{B}_r(\rho)} \|\widetilde{e}_2\|_r.$$

Together with $\|\widetilde{e}\|_r \leq C_{r,\widetilde{B}_r(\rho)} \|e\|_r$, which can be shown trivialy, we have $\|\Delta_W\|_r \leq C_{r,\widetilde{B}_r(\rho)} \|e\|_r$, $\|\Delta_a\|_r \leq C_{r,\widetilde{B}_r(\rho)} \|e\|_r$, and $\|\Delta_\lambda\|_r \leq C_{r,\widetilde{B}_r(\rho)} \|e\|_r$, which finishes the proof. $\square$

Lemma 11 (Condition 4). *For $u \in \widetilde{B}_m$, we have*

$$\|(D\mathscr{F}[u]\eta[u] - Id)\mathscr{F}[u]\|_{m-2} \leq C_{r,\widetilde{B}_r(\rho)} \|\mathscr{F}[u]\|_m \|\mathscr{F}[u]\|_{m-1}.$$

Proof.

$$
\begin{aligned}
&\mathscr{F}[u] - D\mathscr{F}[u]\eta[u]\mathscr{F}[u] \\
&= \mathscr{F}[u] + D\mathscr{F}[u]\Delta_u \\
&= \mathscr{F}[u + \Delta_u] + \mathscr{O}(\Delta^2) \\
&= f(W + \Delta_W) - (W + \Delta_W)(a + \Delta_a, (\lambda + \Delta_\lambda)s) + \mathscr{O}(\Delta^2) \\
&= f(W) + Df(W)\Delta_W - W(a, \lambda s) - DW(a, \lambda s)\begin{pmatrix} \Delta_a \\ \Delta_\lambda s \end{pmatrix} - \Delta_W(a, \lambda s) \\
&\quad + D\Delta_W(a, \lambda s)\begin{pmatrix} \Delta_a \\ \Delta_\lambda s \end{pmatrix} + \mathscr{O}(\Delta^2) \\
&= -DW(a, \lambda s)\left[\begin{pmatrix} Da & 0 \\ D\lambda s & \lambda \end{pmatrix}\Gamma - \begin{pmatrix} \Delta_a \\ \Delta_\lambda s \end{pmatrix} - \Gamma(a, \lambda s) - \widetilde{e}\right] \\
&\quad - De\Gamma + D\Delta_W(a, \lambda s)\begin{pmatrix} \Delta_a \\ \Delta_\lambda s \end{pmatrix} + \mathscr{O}(\Delta^2) \\
&= -De\Gamma + D\Delta_W(a, \lambda s)\begin{pmatrix} \Delta_a \\ \Delta_\lambda s \end{pmatrix} + \mathscr{O}(\Delta^2).
\end{aligned}
$$

By the proof in Lemma 10, and that $\|De\|_{m-2} \le C_{r,\widetilde{B}_r(\rho)} \|e\|_{m-1}$, $\|D\Delta_W\|_{m-1} \le C_{r,\widetilde{B}_r(\rho)} \|e\|_{m-1}$, we achieve

$$\|(D\mathscr{F}[u]\eta[u] - Id)\mathscr{F}[u]\|_{m-2} \le C_{r,\widetilde{B}_r(\rho)} \|\mathscr{F}(u)\|_{m-1} \|\mathscr{F}(u)\|_m.$$

□

Lemma 12 (Condition 5). *For $u \in \tilde{B}_m \cap \mathscr{X}^{m+2}$, we have $\|\mathscr{F}[u]\|_{m+2} \le C_{r,\widetilde{B}_r(\rho)}(1+\|u\|_{m+2})$.*

Proof. From the proof of Lemma 8 above, we can see that there exists a constant $C > 0$ such that $\|\mathscr{F}[u]\|_r < C_{r,\widetilde{B}_r(\rho)}$ for $u \in \tilde{B}_m \cap \mathscr{X}^{m+2}$, thus the Lemma follows naturally.   □

Since all the constant $C_{r,\widetilde{B}_r(\rho)}$ we get from Lemma 8, 9, 10, 11 and 12 are universal for u in the respective domain $\widetilde{B}_r(\rho)$, we have finished proving all the non-degeneracy conditions required by Theorem 13 in the Appendix A. Thus we have proved Theorem 4.

APPENDIX A. AN ABSTRACT IMPLICIT FUNCTION THEOREM IN SCALES OF BANACH SPACES

In this appendix, we present and prove Theorem 4, which is a modified version of the Nash-Moser implicit function theorem. We have made the assumptions in Theorem 4 to match the inequalities that we can achieve from the algorithm in Section 3. We hope that this theorem can also be applied in some other problems involving invariance equations in the theory of normally hyperbolic systems.

The main idea of the Nash-Moser smoothing technique is to add a smoothing operation inside the Newton steps. That is, even though the Newton (or quasi-Newton) steps lose regularities, the smoothing operator restores them.

As anticipated in Section 1.2, our problem has some unusual properties which make it impossible to use other results. As peculiarities of the analysis our problem we recall:

(1) The functional we are trying to solve is not differentiable from one space to itself (It is basically, the composition operator).

(2) The linearized equation can be solved without loss of regularity, but only for regularities on a range. This range does not change much by smoothing the problem. Hence, the technique of approximating the problem by C^∞ or analytic ones does not produce any results. A result we found inspiring is [Sch60].

(3) The use of identities to simplify the equation leads to an extra term in the error estimates after applying the iterative method. The new error contains a term estimated by a derivative of the original error multiplied by the correction (in appropriate norms). Implicit function theorems with these terms were already considered in [Van02, CdlL10, CCdlL13] but they were treated by analytic or C^∞ smoothing which is not possible for the problem in this paper.

(4) In the problem at hand it is natural to use functions with a mixed regularity: finitely differentiable in one variable and analytic in another.

The statement of the abstract Nash-Moser implicit function theorem we will use is:

Theorem 13. *Let $m > 2$ and $\mathscr{X}^r$, $\mathscr{Y}^r$ for $m \le r \le m+2$ be scales of Banach spaces with smoothing operators. Let B_r be the unit ball in $\mathscr{X}^r$, $\widetilde{B}_r(\rho) = u_0 + \rho B_r$ be the unit ball translated by $u_0 \in \mathscr{X}^r$ with radius scaled by $\rho > 0$, and $B(\mathscr{Y}^r, \mathscr{X}^r)$ is the space of bounded linear operators from $\mathscr{Y}^r$ to $\mathscr{X}^r$. Consider a map*

$$\mathscr{F} : \widetilde{B}_r(\rho) \to \mathscr{Y}^r$$

and

$$\eta : \tilde{B}_r \to B(\mathscr{Y}^r, \mathscr{X}^r)$$

satisfing:

- $\mathscr{F}(\widetilde{B}_r(\rho) \cap \mathscr{X}^r) \subset \mathscr{Y}^r$ *for* $m \le r \le m+2$.
- $\mathscr{F}|_{\tilde{B}_m \cap \mathscr{X}^r} : \tilde{B}_r(\rho) \cap \mathscr{X}^r \to \mathscr{X}^r$ *has two continuous Fréchet derivatives, and satisfy the following bounded conditions:*
    * $\|D\mathscr{F}[u](h)\|_{m-2} \le C \|h\|_{m-2}$ *for* $h \in \mathscr{X}^m$.
    * $\|D^2\mathscr{F}[u](h)(k)\|_{m-2} \le C \|h\|_{m-1} \|k\|_{m-1}$ *for* $k, h \in \mathscr{X}^m$.
- $\|\eta[u]\mathscr{F}[u]\|_r \le C \|\mathscr{F}[u]\|_r$, $u \in \tilde{B}_r(\rho)$ *for* $r = m-2, m+2$.
- $\|(D\mathscr{F}[u]\eta[u] - Id)\mathscr{F}[u]\|_{m-2} \le \|\mathscr{F}[u]\|_m \|\mathscr{F}[u]\|_{m-1}$, $u \in \tilde{B}_m$.
- $\|\mathscr{F}[u]\|_{m+2} \le C(1 + \|u\|_{m+2})$, $u \in \tilde{B}_m \cap \mathscr{X}^{m+2}$.

Then if $\|\mathscr{F}[u_0]\|_{m-2}$ *is sufficiently small, then there exists* $u^* \in \mathscr{X}^m$ *such that* $\mathscr{F}[u^*] = 0$. *Moreover,*

$$\|u_0 - u^*\|_m \le C \|\mathscr{F}[u_0]\|_{m-2}$$

Proof. Let $\kappa > 1, \beta, \mu, \alpha > 0, 0 < v < 1$ be real numbers to be specified later. Consider the sequence u_n such that

(42) $$u_n = u_{n-1} - S_{t_{n-1}}\eta[u_{n-1}]\mathscr{F}[u_{n-1}],$$

where $t_n = e^{\beta\kappa^{n-1}}$. We will prove that this sequence satisfies the following three conditions inductively:

$(P_1 n)$: $u_n \in \tilde{B}_m$,
$(P_2 n)$: $\|\mathscr{F}[u_n]\|_{m-2} \le ve^{-2\mu\beta\kappa^n}$,
$(P_3 n)$: $1 + \|u_n\|_{m+2} \le ve^{2\alpha\beta\kappa^n}$.

First, for $n = 0$, we know $P_1(n = 0)$ is ture automatically. By setting $v = \|\mathscr{F}[u_0]\|_{m-2} e^{2\mu\beta}$ with μ, β be specified later and $\|\mathscr{F}[u_0]\|_{m-2} < e^{-2\mu\beta}$, $P_2(n = 0)$ is true. Given α, we can let β be big enough such that condition $P_3(n = 0)$: $1 + \|u_0\|_{m+2} \le e^{2\alpha\beta}$ holds. Now, suppose P_1, P_2 and P_3 are true for $n - 1$, we will now show that the three conditions are true for n.

By assumption (3) and $P_2(n - 1)$, we have

(43) $$\|\eta[u_n]\mathscr{F}[u_n]\|_{\mathscr{X}^{m-2}} \le C \|\mathscr{F}[u_n]\|_{\mathscr{Y}^{m-2}} \le Cve^{-2\mu\beta\kappa^n},$$

it follows from (42) and (43) that

$$\begin{aligned}
\|u_n - u_{n-1}\|_{\mathscr{X}^m} &= \left\|S_{t_{n-1}}\eta[u_{n-1}]\mathscr{F}[u_{n-1}]\right\|_{\mathscr{X}^m} \\
&\le Ct_{n-1}^2 \|\eta[u_{n-1}]\mathscr{F}[u_{n-1}]\|_{\mathscr{X}^{m-2}} \\
&\le Cve^{2\beta\kappa^{n-1}(1-\mu)}.
\end{aligned}$$

(44)

where the second inequality above comes from (29).

Thus we have

$$\|u_n - u_0\|_{\mathscr{X}^m} \le \sum_{j=1}^{\infty} \|u_j - u_{j-1}\|_{\mathscr{X}^m}$$

$$\le Cv(e^{2\beta(1-\mu)} + e^{2\beta(1-\mu)\kappa} + e^{2\beta(1-\mu)\kappa^2} + e^{2\beta(1-\mu)\kappa^3} + e^{2\beta(1-\mu)\kappa^4} + \sum_{j=6}^{\infty} e^{2\beta(1-\mu)\kappa^j})$$

$$\le Cv(e^{2\beta(1-\mu)} + e^{2\beta(1-\mu)\kappa} + e^{2\beta(1-\mu)\kappa^2} + e^{2\beta(1-\mu)\kappa^3} + e^{2\beta(1-\mu)\kappa^4} + \frac{e^{2\beta(1-\mu)\kappa^6}}{1 - e^{2\beta(1-\mu)\kappa}})$$

$$(45) \qquad\qquad \le \rho.$$

where the third inequality comes from the fact that $\kappa^{j-1} > j\kappa$ for $j \ge 5$ and $\kappa > \sqrt[3]{5}$, and the last inequality can be achieved if $\mu > 1$ and β is large enough. Thus we have proved $P_1(n)$.

In order to prove $P_2(n)$, let us break $\mathscr{F}[u_n]$ as follows:

$$\|\mathscr{F}[u_n]\|_{\mathscr{Y}^{m-2}} \le \left\|\mathscr{F}[u_n] - \mathscr{F}[u_{n-1}] + D\mathscr{F}[u_n]S_{t_{n-1}}\eta[u_{n-1}]\mathscr{F}[u_{n-1}]\right\|_{\mathscr{Y}^{m-2}}$$

$$+ \left\|(Id - D\mathscr{F}[u_{n-1}]\eta[u_{n-1}])\mathscr{F}[u_{n-1}]\right\|_{\mathscr{Y}^{m-2}}$$

$$(46) \qquad\qquad + \left\|D\mathscr{F}[u_{n-1}](Id - S_{t_{n-1}})\eta[u_{n-1}]\mathscr{F}[u_{n-1}]\right\|_{\mathscr{Y}^{m-2}}$$

and estimates the three terms one by one:

For the first line, by Taylor expansion, the induction condition in the second part of assumption (2), (29) and (43), we have

$$l_1 = \left\|\mathscr{F}[u_n] - \mathscr{F}[u_{n-1}] + D\mathscr{F}[u_n]S_{t_{n-1}}\eta[u_{n-1}]\mathscr{F}[u_{n-1}]\right\|_{\mathscr{Y}^{m-2}}$$

$$\le C \left\|D^2\mathscr{F}[u_n](S_{t_{n-1}}\eta[u_{n-1}]\mathscr{F}[u_{n-1}])(S_{t_{n-1}}\eta[u_{n-1}]\mathscr{F}[u_{n-1}])\right\|_{\mathscr{Y}^{m-2}}$$

$$\le C \left\|(S_{t_{n-1}}\eta[u_{n-1}]\mathscr{F}[u_{n-1}])\right\|_{\mathscr{Y}^{m-1}}^2$$

$$\le Ct_{n-1}^2 \left\|\eta[u_{n-1}\mathscr{F}[u_{n-1}]]\right\|_{\mathscr{Y}^{m-2}}^2$$

$$(47) \qquad\qquad \le Cve^{(1-2\mu)2\beta\kappa^{n-1}}$$

For the second line, by assumption(4), (32), $P_2(n-1)$, assumption (5) and $P_3(n-1)$, we have

$$l_2 = \|(Id - D\mathscr{F}[u_{n-1}]\eta[u_{n-1}])\mathscr{F}[u_{n-1}]\|_{\mathscr{Y}^{m-2}}$$

$$\le C \|\mathscr{F}[u_{n-1}]\|_{\mathscr{Y}^{m-1}} \|\mathscr{F}[u_n]\|_{\mathscr{Y}^m}$$

$$\le C \|\mathscr{F}[u_{n-1}]\|_{\mathscr{Y}^{m-2}}^{\frac{5}{4}} \|\mathscr{F}[u_{n-1}]\|_{\mathscr{Y}^{m+2}}^{\frac{3}{4}}$$

$$\le C \|\mathscr{F}[u_{n-1}]\|_{\mathscr{Y}^{m-2}}^{\frac{5}{4}} (1 + \|u_{n-1}\|_{\mathscr{Y}^{m+2}})^{\frac{3}{4}}$$

$$(48) \qquad\qquad \le Cv^{\frac{3}{2}}e^{(-\frac{5}{2}\mu + \frac{3}{2}\alpha)\beta\kappa^{n-1}}$$

For the third line, by (30), (43) and assumption (5), we have

$$l_3 = \left\|D\mathscr{F}[u_{n-1}](Id - S_{t_{n-1}})\eta[u_{n-1}]\mathscr{F}[u_{n-1}]\right\|_{\mathscr{Y}^{m-2}}$$

$$\le Ct_{n-1}^4 \|\eta[u_{n-1}]\mathscr{F}[u_{n-1}]\|_{\mathscr{Y}^{m+2}}$$

$$\le Ct_{n-1}^4(1 + \|u_{n-1}\|_{\mathscr{Y}^{m+2}})$$

$$(49) \qquad\qquad \le Cve^{2\beta\kappa^{n-1}(\alpha-2)}$$

thus, in order to show that (50) is true, we want $l_1 + l_2 + l_3 \leq ve^{-2\mu\beta\kappa^n}$, i.e.

$$(Cve^{(1-2\mu)2\beta\kappa^{n-1}} + Cv^{\frac{3}{2}}e^{(-\frac{5}{2}\mu+\frac{3}{2}\alpha)\beta\kappa^{n-1}} + ve^{2\beta\kappa^{n-1}(\alpha-2)}) < ve^{-2\mu\beta\kappa^n},$$

Thus we need

(50) $$C(e^{(1-2\mu+\mu\kappa)2\beta\kappa^{n-1}} + v^{\frac{1}{2}}e^{(-3\mu+\alpha+\mu\kappa)2\beta\kappa^{n-1}} + ve^{(\alpha-2+\kappa)2\mu\beta\kappa^{n-1}}) < 1,$$

which can be satisfied if μ, κ and α satisfies

(51) $$\begin{cases} 1 - 2\mu + \mu\kappa < 0, \\ -3\mu + \alpha + \mu\kappa < 0, \\ \alpha - 2 + \kappa < 0. \end{cases}$$

and β is picked large enough.

As for $P_3(n)$, by (44), (42), (42), (43), assumption (5) and $P_3(j)$ for $j < n - 1$, we have

$$1 + \|u_n\|_{\mathscr{X}^{m+2}} \leq 1 + \sum_{j=1}^{n} \|u_j - u_{j-1}\|_{m+2}$$

$$\leq 1 + \sum_{j=1}^{n} \left\| S_{t_{j-1}} \eta[u_{j-1}]\mathscr{F}[u_{j-1}] \right\|_{m+2}$$

$$\leq 1 + C\sum_{j=1}^{n} \left\| \eta[u_{j-1}\mathscr{F}[u_{j-1}]] \right\|_{m+2}$$

$$\leq 1 + C\sum_{j=1}^{n} (1 + \|u_{j-1}\|_{\mathscr{X}^{m+2}})$$

(52) $$\leq 1 + C\sum_{j=1}^{n} e^{2\alpha\beta\kappa^{j-1}}$$

we need

$$(1 + \|u_n\|_{m+2})e^{-2\alpha\beta\kappa^n} < 1,$$

that is

(53) $$e^{-2\alpha\beta\kappa^n} + C\sum_{j=1}^{n} e^{(1-\kappa)\alpha\beta\kappa^{j-1}} < v.$$

which is

$$e^{-2\alpha\beta\kappa^n} + C\sum_{j=1}^{n} e^{(1-\kappa)\alpha\beta\kappa^{j-1}} < \|\mathscr{F}[u_0]\|_{m-2}.$$

By the same reason as in (45), we have

$$\sum_{j=1}^{n} e^{(1-\kappa)\alpha\beta\kappa^{j-1}} \leq e^{(1-\kappa)\alpha\beta} + e^{(1-\kappa)\alpha\beta\kappa} + e^{(1-\kappa)\alpha\beta\kappa^2} + e^{(1-\kappa)\alpha\beta\kappa^3} + e^{(1-\kappa)\alpha\beta\kappa^4} + \frac{e^{(1-\kappa)\alpha\beta6\kappa}}{1 - e^{(1-\kappa)\alpha\beta\kappa}}$$

can be achieved if $\kappa > \sqrt[3]{5}$ and β is large enough.

Above all, in oder to make sure that (45), (50) and (53) are true, we nee the following constrictions for κ, α and μ:

$$
(54) \qquad
\begin{cases}
\mu > 1, \\
\kappa > \sqrt[3]{5}, \\
1 - 2\mu + \mu\kappa < 0, \\
-3\mu + \alpha + \mu\kappa < 0, \\
\alpha - 2 + \kappa < 0.
\end{cases}
$$

and β is large enough.

One possible solution for (54) is $\kappa = 1.75, \mu = 5$ and $\alpha = 0.05$ and β is large enough.

Up to this point, we have finished the proof for induction. By letting $n \to \infty$, the second assumption $\|\mathscr{F}[u_n]\|_{\mathscr{X}^{m-2}} \leq v e^{-2\mu\beta\kappa^n}$ leads to a solution $u^* \in \mathscr{X}^{m-2}$ such that $\mathscr{F}[u^*] = 0$, and the convergence is superexponential. Moreover, by the discussion in (45), we have

$$
\|u^* - u_0\|_m \leq Cv = C \|\mathscr{F}[u_0]\|_{m-2},
$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

Remark 39. Although the result $\|u^* - u_0\|_m \leq Cv = C \|\mathscr{F}[u_0]\|_{m-2}$ is a bit surprising in the sense that the higher regularity norm is bounded by the lower one, but this inequality is actually justified by the bounds from even higher regularity required in the assumption.

References

[Ado07]   Raymond A. Adomaitis. The trouble with spurious eigenvalues. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, 17(4):1375–1381, 2007.

[AO00]   Mark Ainsworth and J. Tinsley Oden. *A posteriori error estimation in finite element analysis.* Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2000.

[AVK87]   A. A. Andronov, A. A. Vitt, and S. È. Khaĭkin. *Theory of oscillators.* Dover Publications, Inc., New York, 1987. Translated from the Russian by F. Immirzi, Reprint of the 1966 translation.

[BDV05]   Christian Bonatti, Lorenzo J. Díaz, and Marcelo Viana. *Dynamics beyond uniform hyperbolicity*, volume 102 of *Encyclopaedia of Mathematical Sciences.* Springer-Verlag, Berlin, 2005. A global geometric and probabilistic perspective, Mathematical Physics, III.

[BLZ00]   Peter W. Bates, Kening Lu, and Chongchun Zeng. Invariant foliations near normally hyperbolic invariant manifolds for semiflows. *Trans. Amer. Math. Soc.*, 352(10):4641–4676, 2000.

[BLZ08]   Peter W. Bates, Kening Lu, and Chongchun Zeng. Approximately invariant manifolds and global dynamics of spike states. *Invent. Math.*, 174(2):355–433, 2008.

[BS08]   Kristian Bjerklöv and Maria Saprykina. Universal asymptotics in hyperbolicity breakdown. *Nonlinearity*, 21(3):557–586, 2008.

[BST98]   Henk Broer, Carles Simó, and Joan Carles Tatjer. Towards global models near homoclinic tangencies of dissipative diffeomorphisms. *Nonlinearity*, 11(3):667–770, 1998.

[CCdlL13] Renato C. Calleja, Alessandra Celletti, and Rafael de la Llave. A KAM theory for conformally symplectic systems: efficient algorithms and their validation. *J. Differential Equations*, 255(5):978–1049, 2013.

[CdlL10] Renato Calleja and Rafael de la Llave. A numerically accessible criterion for the breakdown of quasi-periodic solutions and its rigorous justification. *Nonlinearity*, 23(9):2029–2058, 2010.

[CF12]   Renato Calleja and Jordi-Lluís Figueras. Collision of invariant bundles of quasi-periodic attractors in the dissipative standard map. *Chaos*, 22(3):033114, 10, 2012.

[CFdlL03a] Xavier Cabré, Ernest Fontich, and Rafael de la Llave. The parameterization method for invariant manifolds. I. Manifolds associated to non-resonant subspaces. *Indiana Univ. Math. J.*, 52(2):283–328, 2003.

[CFdlL03b] Xavier Cabré, Ernest Fontich, and Rafael de la Llave. The parameterization method for invariant manifolds. II. Regularity with respect to parameters. *Indiana Univ. Math. J.*, 52(2):329–360, 2003.

[CFdlL05] Xavier Cabré, Ernest Fontich, and Rafael de la Llave. The parameterization method for invariant manifolds. III. Overview and applications. *J. Differential Equations*, 218(2):444–515, 2005.

[CH17a] Marta Canadell and Àlex Haro. Computation of quasi-periodic normally hyperbolic invariant tori: algorithms, numerical explorations and mechanisms of breakdown. *J. Nonlinear Sci.*, 27(6):1829–1868, 2017.

[CH17b] Marta Canadell and Àlex Haro. Computation of quasiperiodic normally hyperbolic invariant tori: rigorous results. *J. Nonlinear Sci.*, 27(6):1869–1904, 2017.

[CJ15] Y.-M. Chung and M. S. Jolly. A unified approach to compute foliations, inertial manifolds, and tracking solutions. *Math. Comp.*, 84(294):1729–1751, 2015.

[CK20] Maciej J. Capiński and Hieronim Kubica. Persistence of normally hyperbolic invariant manifolds in the absence of rate conditions. *Nonlinearity*, 33(9):4967–5005, 2020.

[CZ15] Maciej J. Capiński and Piotr Zgliczyński. Geometric proof for normally hyperbolic invariant manifolds. *J. Differential Equations*, 259(11):6215–6286, 2015.

[dlL97] Rafael de la Llave. Invariant manifolds associated to nonresonant spectral subspaces. *J. Statist. Phys.*, 87(1-2):211–249, 1997.

[dlL01] Rafael de la Llave. A tutorial on KAM theory. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 175–292. Amer. Math. Soc., Providence, RI, 2001.

[dlLK19] Rafael de la Llave and Florian Kogelbauer. Global persistence of lyapunov subcenter manifolds as spectral submanifolds under dissipative perturbations. *SIAM J. Appl. Dyn. Syst.*, 18(4):2099–2142, 2019.

[dlLO99] R. de la Llave and R. Obaya. Regularity of the composition operator in spaces of Hölder functions. *Discrete Contin. Dynam. Systems*, 5(1):157–184, 1999.

[dlLW11] Rafael de la Llave and Alistair Windsor. Smooth dependence on parameters of solutions to cohomology equations over Anosov systems with applications to cohomology equations on diffeomorphism groups. *Discrete Contin. Dyn. Syst.*, 29(3):1141–1154, 2011.

[Dul03] Henri Dulac. *Recherches sur les points singuliers des équations differentielles*. Gauthier-Villars, 1903.

[ET10] G. Bard Ermentrout and David H. Terman. *Mathematical foundations of neuroscience*, volume 35 of *Interdisciplinary Applied Mathematics*. Springer, New York, 2010.

[Fen77] Neil Fenichel. Asymptotic stability with rate conditions. II. *Indiana Univ. Math. J.*, 26(1):81–93, 1977.

[Fen72] Neil Fenichel. Persistence and smoothness of invariant manifolds for flows. *Indiana Univ. Math. J.*, 21:193–226, 1971/72.

[Fen74] Neil Fenichel. Asymptotic stability with rate conditions. *Indiana Univ. Math. J.*, 23:1109–1137, 1973/74.

[FH12] Jordi-Lluís Figueras and Àlex Haro. Reliable computation of robust response tori on the verge of breakdown. *SIAM J. Appl. Dyn. Syst.*, 11(2):597–628, 2012.

[GE88] Melinda P. Golden and B. Erik Ydstie. Bifurcation in model reference adaptive control systems. *Systems & Control Letters*, 11(5):413–430, 1988.

[Gra17] Albert Granados. Invariant manifolds and the parameterization method in coupled energy harvesting piezoelectric oscillators. *Phys. D*, 351/352:14–29, 2017.

[HCF+16] Àlex Haro, Marta Canadell, Jordi-Lluís Figueras, Alejandro Luque, and Josep-Maria Mondelo. *The parameterization method for invariant manifolds*, volume 195 of *Applied Mathematical Sciences*. Springer, [Cham], 2016. From rigorous results to effective computations.

[HdlL06a] À. Haro and R. de la Llave. Manifolds on the verge of a hyperbolicity breakdown. *Chaos*, 16(1):013120, 8, 2006.

[HdlL06b] À. Haro and R. de la Llave. A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: numerical algorithms. *Discrete Contin. Dyn. Syst. Ser. B*, 6(6):1261–1300, 2006.

[HdlL07]   A. Haro and R. de la Llave. A parameterization method for the computation of invariant tori
            and their whiskers in quasi-periodic maps: explorations and mechanisms for the breakdown of
            hyperbolicity. *SIAM J. Appl. Dyn. Syst.*, 6(1):142–207, 2007.
[HdlL13]   Gemma Huguet and Rafael de la Llave. Computation of limit cycles and their isochrons: fast
            algorithms and their convergence. *SIAM J. Appl. Dyn. Syst.*, 12(4):1763–1802, 2013.
[Izh07]   Eugene M. Izhikevich. *Dynamical systems in neuroscience: the geometry of excitability and
            bursting.* Computational Neuroscience. MIT Press, Cambridge, MA, 2007.
[JK69]   Jiří Jarník and Jaroslav Kurzweil. On invariant sets and invariant manifolds of differential
            systems. *J. Differential Equations*, 6:247–263, 1969.
[Lev81]   Mark Levi. Qualitative analysis of the periodically forced relaxation oscillations. *Mem. Amer.
            Math. Soc.*, 32(244):vi+147, 1981.
[Mat68]   John N. Mather. Characterization of Anosov diffeomorphisms. *Nederl. Akad. Wetensch. Proc.
            Ser. A 71 = Indag. Math.*, 30:479–483, 1968.
[Min62]   Nicolas Minorsky. *Nonlinear oscillations.* D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-
            London-New York, 1962.
[MM76]   J. E. Marsden and M. McCracken. *The Hopf bifurcation and its applications.* Springer-Verlag,
            New York, 1976. With contributions by P. Chernoff, G. Childs, S. Chow, J. R. Dorroh, J. Guck-
            enheimer, L. Howard, N. Kopell, O. Lanford, J. Mallet-Paret, G. Oster, O. Ruiz, S. Schecter,
            D. Schmidt and S. Smale, Applied Mathematical Sciences, Vol. 19.
[Mn78]   Ricardo Mañé. Persistent manifolds are normally hyperbolic. *Trans. Amer. Math. Soc.*, 246:261–
            283, 1978.
[Mos66a]   Jürgen Moser. A rapidly convergent iteration method and non-linear differential equations. II.
            Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3), 20:499–535, 1966.
[Mos66b]   Jürgen Moser. A rapidly convergent iteration method and non-linear partial differential equa-
            tions. I. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)*, 20:265–315, 1966.
[Nei59]   Ju Neimark. On some cases of periodic motions depending on parameters. *Dokl. Akad. Nauk
            SSSR*, 129(4):736–739, 1959.
[Pes04]   Yakov B. Pesin. *Lectures on partial hyperbolicity and stable ergodicity.* Zurich Lectures in Ad-
            vanced Mathematics. European Mathematical Society (EMS), Zürich, 2004.
[Poi79]   Henri Poincaré. *Sur les propriétés des fonctions définies par les équations aux differences par-
            tielles.* Gauthier Villars, 1879.
[Poi90]   H. Poincaré. Sur une classe nouvelle de transcendentes uniformes. *Jour. de Math.*, 6:313–365,
            1890.
[Ran92a]   D. A. Rand. Existence, nonexistence and universal breakdown of dissipative golden invariant
            tori. I. Golden critical circle maps. *Nonlinearity*, 5(3):639–662, 1992.
[Ran92b]   D. A. Rand. Existence, nonexistence and universal breakdown of dissipative golden invariant
            tori. II. Convergence of renormalization for mappings of the annulus. *Nonlinearity*, 5(3):663–680,
            1992.
[Ran92c]   D. A. Rand. Existence, nonexistence and universal breakdown of dissipative golden invariant
            tori. III. Invariant circles for mappings of the annulus. *Nonlinearity*, 5(3):681–706, 1992.
[RT71]   David Ruelle and Floris Takens. On the nature of turbulence. *Comm. Math. Phys.*, 20:167–192,
            1971.
[Sac64]   Robert John Sacker. *ON INVARIANT SURFACES AND BIFURCATION OF PERIODIC SO-
            LUTIONS OF ORDINARY DIFFERENTIAL EQUATIONS.* ProQuest LLC, Ann Arbor, MI,
            1964. Thesis (Ph.D.)–New York University.
[Sac09]   Robert J. Sacker. On invariant surfaces and bifurcation of periodic solutions of ordinary differen-
            tial equations. Chapter II. Bifurcation-mapping method. *J. Difference Equ. Appl.*, 15(8-9):759–
            774, 2009. Reprinted from New York Univ. Report IMM-NYU 333, October 1964, Courant Inst.,
            New York.
[Sch60]   J. Schwartz. On Nash's implicit functional theorem. *Comm. Pure Appl. Math.*, 13:509–530, 1960.
[Ste57]   Shlomo Sternberg. Local contractions and a theorem of Poincaré. *Amer. J. Math.*, 79:809–824,
            1957.
[Ste70]   Elias M. Stein. *Singular integrals and differentiability properties of functions.* Princeton Math-
            ematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[Sza20]   Robert Szalai. Invariant spectral foliations with applications to model order reduction and synthesis. *Nonlinear Dynamics*, 101(4):2645–2669, 2020.

[Van02]   John Andrew Vano. *A Nash-Moser implicit function theorem with Whitney regularity and applications*. ProQuest LLC, Ann Arbor, MI, 2002. Thesis (Ph.D.)–The University of Texas at Austin.

[Win01]   Arthur T. Winfree. *The geometry of biological time*, volume 12 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, second edition, 2001.

[Win75]   A. T. Winfree. Patterns of phase compromise in biological cycles. *J. Math. Biol.*, 1(1):73–95, 1974/75.

[WY02]   Qiudong Wang and Lai-Sang Young. From invariant curves to strange attractors. *Comm. Math. Phys.*, 225(2):275–304, 2002.

[WY03]   Qiudong Wang and Lai-Sang Young. Strange attractors in periodically-kicked limit cycles and Hopf bifurcations. *Comm. Math. Phys.*, 240(3):509–529, 2003.

[YdlL21]   Yian Yao and Rafael de la Llave. Computing the invariant circle and the foliation by stable manifolds for 2-d maps by the parameterization method: Numerical implementation and results. 2021. Manuscript in progress.

[ZdlL18]   Lei Zhang and Rafael de la Llave. Transition state theory with quasi-periodic forcing. *Commun. Nonlinear Sci. Numer. Simul.*, 62:229–243, 2018.

[Zeh75]   E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. I. *Comm. Pure Appl. Math.*, 28:91–140, 1975.

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, 686 CHERRY ST., ATLANTA GA 30332-160

Email address: `yyao93@gatech.edu`

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, 686 CHERRY ST., ATLANTA GA 30332-160

Email address: `rafael.delallave@gatech.edu`