FAST ALGORITHM FOR INVARIANT CIRCLE AND THEIR STABLE MANIFOLDS: RIGOROUS RESULTS AND EFFICIENT IMPLEMENTATIONS

A Dissertation
Presented to
The Academic Faculty

By

Yian Yao

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mathematics

Georgia Institute of Technology

August 2021

FAST ALGORITHM FOR INVARIANT CIRCLE AND THEIR STABLE MANIFOLDS: RIGOROUS RESULTS AND EFFICIENT IMPLEMENTATIONS

Thesis committee:

Prof. Rafael de la Llave, Advisor
School of Mathematics
Georgia Institute of Technology

Prof. Yingjie Liu
School of Mathematics
Georgia Institute of Technology

Prof. Molei Tao
School of Mathematics
Georgia Institute of Technology

Prof. Àlex Haro Provinciale
Departament de Matemàtiques i Informàtica
Universitat de Barcelona

Prof. Sung Ha Kang
School of Mathematics
Georgia Institute of Technology

Date approved: July 09, 2021

For my family and friends

ACKNOWLEDGMENTS

I would like to take this opportunity to express my sincere appreciation to all the people I came across during my graduate study.

First and foremost, I would like to express my profound gratitude to my Ph.D. advisor Prof. Rafael de la Llave. Nothing would be possible without his warm support and wholehearted help. In the past five years, he led me into this beautiful world of dynamical systems, and encouraged me to make explorations of my own choice. His professionalism, wisdom, encouragement, and patience have guided me through this long and occasionally touch Ph.D. journey, and will continue inspiring me for my future life and career. I will always cherish the lessons I learned from him as well as all the happy moments. Prof. Rafael de la Llave is both a fantastic advisor and a very dear friend. Meeting him is one of the best things that happened in my life. I feel truly blessed.

I would like to express special thanks to my committee members, Prof. Àlex Haro, Prof. Sung Ha Kang, Prof. Yingjie Liu, and Prof. Molei Tao. Many thanks to Prof. Àlex Haro for his great suggestions on my research, and for all his encouragement and support; I appreciate Prof. Sung Ha Kang for her enthusiasm, and for her help with my CSE master application; I thank Prof. Yingjie Liu for the rewarding discussions and I am grateful for Prof. Molei Tao for being one of the role models I always look up to.

I am also very grateful to Prof. Srinivas Aluru, Prof. Federico Bonetto, Prof. John Etnyre, Prof. Christopher Heil, Prof. Sung Ha Kang, Prof. Wenjing Liao, Prof. Yingjie Liu, Prof. Heinrich Matzinger, Prof. Le Song, Prof. Molei Tao, Prof. Howard Weiss, Prof. Yao Yao, Prof. Josephine Yu, Prof. Chongchun Zeng, Prof. Xiuwei Zhang and Prof. Haomin Zhou for the amazing courses and seminars. I thank Prof. Luca Dieci for his help and suggestions during my service as a board member of the SIAM student chapter. I thank Dr. Morag Burke, Prof. Xuyan Chen, Ms. Klara Grodzinsky, and Prof Gregory Mayer for all the helpful teaching advice. I thank Ms. Stephanie Cook, Mr. Mitchell Everett, Ms. Iris

Hamilton, Ms. Sharon McDowell, and the IT group for their constant support. Finally, I would like to express my genuine thanks to Prof. Mohammad Ghomi and Prof. Xingxing Yu for all the warm support and sincere help.

I would like to thank all my friends for all these good days we spent together at Georgia Tech. It is truly my great pleasure and honor to meet the amazing individuals: Adrián, Fan, Hongyu, Jiaqi, Joan, Lei, Qianli, Renyi, Ruilin, Shu, Tongzhou, Weiwei, Xianda, Xiao, Xiaofan, Xiaonan, Xin, and Zili. They gave me the strength and courage to keep moving forward, keep surpassing myself, and they have shaped me tremendously into a better version of myself. I will always cherish all the fruitful discussions, all the exciting adventures, and all the happy moments. Furthermore, I would like to thank Bhanu, Christopher, Dantong, Fenfen, Guangyu, Haiyu, Hao, Haodong, Hassan, Jaewoo, Jieun, Jorge, Kewen, Longmei, Mark, Mengyi, Pingyuan, Qinbo, Shasha, Shijie, Victor, Xia, Xiaodan, Xiaolong, Xin, Xingyu, Yifeng, Yuqing, Zhibo and many others for making my life much more colorful and enjoyable. I will miss you all.

Last but not least, I would like to express my heartfelt appreciation to my family, especially to my parents, for their unconditional love and support. In a small town northeast of China, there is a big family full of love and joy. Every time I think of them, my heart smiles.

How I wish the time could stop at this moment, at which we are passionate and energetic, we share courage and ambition, and we have an exciting journey ahead of us. Oh well, I just hope everyone stays happy and healthy, and everyone is getting what his/her heart desires. From the bottom of my heart, I wish you all the very best.

It has been such a fantastic adventure!

TABLE OF CONTENTS

LIST OF TABLES

SUMMARY

In this thesis, we present, analyze, and implement a quadratically convergent algorithm to compute the invariant circle and the foliation by stable manifolds for 2-dimensional maps. The 2-dimensional maps we are considering are motivated by oscillators subject to periodic perturbation.

The algorithm is based on solving an invariance equation using a quasi-Newton method, and the algorithm works irrespective of whether the dynamics on the invariant circle conjugates to a rotation or is phase-locked, and thus we expect only finite regularity on the invariant circle but analytic on the stable manifolds.

The thesis is divided into the following two parts:

In Chapter 2, we derive our quasi-Newton algorithm and prove that starting from an initial guess that satisfies the invariance equation very approximately, the algorithm converges quadratically to a true solution which is close to the initial guess. The proof of the convergence is based on an abstract Nash-Moser Implicit Function Theorem specially tailored for this problem.

In Chapter 3, we discuss some implementation details regarding our algorithm and implemented it on the dissipative standard map. We follow different continuation paths along the perturbation and drift parameter and explore the "bundle merging" scenario when the hyperbolicity of the map losses due to the increase of the perturbation. For non-resonant eigenvalues, we also generalize the algorithm to 3-dimension and implemented it on the 3-D Fattened Arnold Family.

CHAPTER 1

INTRODUCTIONS

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

In this thesis, we study attractive (or repulsive) invariant circles in 2-dimensional maps as well as the stable (unstable) manifolds of points.

We recall, that according to the theory of normally hyperbolic manifolds [1, 2] 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. *In the case of 2-D flows, the quantitative rate is superflous and convergence to an asymptotic phase implies fast convergence. So it is common to call these stable manifolds* isochrones. *[3, 4].*

In two-dimensional flows, the topological notion of convergence gives a characterization of the stable manifolds. This is not true in the case we are considering in this paper and the notion of stable manifolds involves quantitative estimates on the speed of convergence.

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 [5, 6, 7]. Another case that motivated us is the periodic perturbation of a 2-D ode with a limit cycle. Such examples are very common in practice. For example, when oscillating circuits are subject to AC forcing [8, 9] or in Biology when the circadian rhythms are subject to external forcing [10]. Also when neurons are subject to the periodic forcing of others [11, 12].

The interpretation of periodic forcing of limit cycles is helpful, since the methods we apply are inspired by those in [13]. As in [13], our goal will be to find a system of coor-

dinates 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, including that the dynamics in the invariant circle is phase-locked. (Similarly, getting to 3-D maps involves a new phenomenon of normal resonances, which is briefly discussed in [14]).*

From a more technical point of view, we do not expect that the invariant circle or the foliation by stable manifolds of points are analytic as in [13] but only finitely differentiable even if the map considered is analytic (in this paper, we will consider only analytic mappings). Even arbitrarily small perturbations break down the analytic regularity of the invariant circles and one can have open sets of analytic perturbations in which the circles are only finitely differentiable (Section 8.2 of [15]). On the other hand, each of the stable manifolds of a point will be analytic. This anisotropic regularity of the parameterizations of the foliation by stable manifolds of points – our unknown – has to be taken into account when formulating the spaces in the implicit function theorem and in the choice of discretizations [14].

1.1 The Goal of the Thesis

The goal of this thesis is to provide a framework to study these objects (invariant circles and their stable foliations) in a non-perturbative way which also leads to reliable numerical algorithms, which converge to the true solution faster than exponentially. The 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 1.3 for some comparison with other hard implicit function theorems in the literature.

The implementation of the algorithm on some concrete examples is also included in this

thesis, of which we present the implementation details along with some numerical results and investigation of phenomena that happen at the boundary of the validity of our result.

1.2 Description of the Method

Following the idea of the parameterization method [16], [17], [18], [19], we formulate an invariance equation (see (2.2)). This equation has two unknowns:

- a) embeddings of the torus and its stable manifolds,

- b) the dynamics of the map restricted to the invariant objects.

This invariance equation (2.2) 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 by the dynamics). As we will see, the equation (2.2) is undetermined. This undetermination is quite useful since it allows to develop more efficient numerical methods.

We prove that, given an approximate solution of (2.2), we can evaluate some condition numbers on this approximate solution. If the error in (2.2) 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 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 [20].

A-posteriori results imply the usual persistence results under perturbations of dynamical systems. If one can find indeed a system with these structures (invariant circle and its stable manifolds), then, given 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 approxi-

mate 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).

The way that one often proves an a-posteriori theorem is 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 [21]) 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 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 Section 2.6, we present an abstract result: Theorem 12, which we hope could be applicable in similar problems.

The "a-posteriori" result also implies the persistence of the invariant objects under perturbation, as well as the smooth dependency of the solution on parameters. In fact, if an additional small perturbation is applied to the system, the non-degeneracy conditions will remain valid (unless when the parameter is close to the breakdown), and the original invariant object can be viewed as an initial approximation for the algorithm to start with. This provides a theoretical backup for the continuation method.

The low time and space requirement of our quasi-Newton Algorithm makes it suitable for the computation of the continuation of parameters even with desktop computers. Based on [22, 23], we propose a continuation scheme: Algorithm 3. Taking the dissipative standard map as a toy example, we follow different paths in the parameter space, we explored the "bundle merging" scenario that causes the loss of the hyperbolicity, and the "Devil's staircase" phenomenon for the rotation number of the internal dynamics as the drift parameter varies.

The existence domain in the parameter space and the estimation of the breakdown is a fascinating topic, of which many efforts have been made [22, 23]. The breakdown of the invariant circles has many possible definitions. The most common one is the C^1 breakdown, which, according to [24], is caused by the loss of hyperbolicity. In the case when the rotation number of the internal dynamics is rational but the internal dynamics is not a rotation, phase-locked phenomenon appears, and the regularity of the invariant circle can drop continuously, and every choice of $\alpha < 1$ admits a value for the C^α breakdown. In this thesis, we choose an arbitrary continuation path, where the rotation number changes from rational to irrational, and both the scenarios appears alternatively. In this case, We find empirically that the angle between the invariant circe and the isochrons drops to 0 following the same power-law with universal exponent as in [23] when it is close to the

breakdown.

Although the map we consider in this paper is in 2-dimensions, we remark that most of the statements and methods remain valid when the dimensionality of the isochrons increases. In fact, the only scenario that causes problems is when the eigen-directions of the isochrons resonate, in which case our version of the Sternberg's Lemma (see [25]) fails, and the invariance equation is no longer as simple as Equation (2.4).

1.3 Some Remarks on Comparison with Other Theoretical Results

For the experts in Nash-Moser theory, we point out that Theorem 12 developed in Appendix 2.6 has several unusual properties. (Of course, this can be omitted in a first reading, but could serve as motivation for some of the analysis). The guiding principle is that, motivated by the numerics in the literature, we want to justify Newton methods for the problem but they happen to have some unusual properties. We recall that we will be dealing only with analytic mappings.

- 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 situation 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 techniques we can use are approximations by C^r functions (the so-called C^r smoothing).

  We also cannot use the *"double smoothing"* technique of approximating the problem by smoother ones and obtain smooth solutions. A hard implicit function theorem we found inspiring is [26] and [27].

- A technical difference with the standard Nash-Moser theorem is that we will consider

spaces with mixed regularity. We will consider functions of two variables, which 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 [15] – the degree of differentiability is limited, not just by the regularity of the map. On the other hand, the leaves of the stable foliation are always analytic.

Such mixed regularities are very common in the theory of Normally Hyperbolic Invariant Manifolds(NHIM). 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 [28].

  This operator maps C^r spaces into themselves, but it is not differentiable from C^r to C^r but it is differentiable in other topologies [28]. Hence, computing the remainder of the functional after a correction, involve losses of derivatives. 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). This is enough for us since we are only studying analytic dynamical systems in the plane. Considering finitely differentiable maps of the plane would require some extra considerations.

  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).

Since the cohomology equations that drive the analysis of the Newton method, have only solutions for a finite range of regularities, we cannot use C^∞ or C^ω smoothing and we can only use the finite regularity smoothing [27].

- 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 contains a term that involves the remainder of the solution. This term is very common in many problems of dynamical systems that are solved taking advantage of *automatic reducibility*. Abstract implicit function theorems taking into account this extra term – but not the problem of the range of regularities in the solution operator – has been developed in [29, 30].

- As indicated above, the equation considered is undetermined so that the linearized equation will have a kernel.

- The loss of regularity incurred in our result Theorem 12 is much smaller than on 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 [31, 32, 33, 19, 34] since they were found empirically to be efficient, and the solutions obtained could be validated using the more conventional methods (either contraction based methods [35] or topological methods [36]).*

Note also that the Newton or quasi-Newton methods are much more effective than contraction based graph transform methods when the contracting exponents are close to one. In such a case, the contraction properties of graph transform methods are weak and one needs to iterate many times than the very nonlinear graph transform.

When hyperbolicity is weak, the iterations used to solve the cohomology equations in the quasi-Newton method also become slow, but much less than the (very non-linear) graph transform method.

In the case that the internal dynamics a is fixed to be a rotation, one can solve the cohomology equations by Fourier methods so that there is no slow-down even for very weak contraction properties. This case has been studied in the literature several times. [30, 33].

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, much 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 and the identities that allow a speedup is true only when we involve the foliation.

Remark 5. *Besides using the Nash-Moser method, there are other methods that also lead to an a-posterior format by using a contraction in C^0 and propagated bounds in higher regularity. [35].*

Such contraction methods give better regularity results than the Nash-Moser methods presented in this paper. On the other hand, since the Newton methods have been successful in applications (they are known to be fast and reliable), it seems desirable to have a mathematical justification, which we present here. The justification developed here takes advantage of identities that lead to even faster numerical algorithms. Identifying appropriate condition numbers allows one to get confident that the numerical solutions are not spurious even when computing close to the breakdown.

In [16], it is also remarked that one could prove the stable foliations theorems by reducing them to the stable manifold theorem to an infinite dimensional lift of the problem (this

argument only provides a prefoliation, but one can get a foliation observing that $y \approx \tilde{y}$ when $d(f^n(y), f^n(\tilde{y})) \leq C_{y,\tilde{y}} \lambda^n$ is an equivalence relation).

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 a rather complex behaviors appear in [37, 38, 39]. The study of the boundary between the regular behavior presented here and the chaotic behavior is a very interesting mathematical problem [40]. Some elements of the boundary of validity of the results have been explored in [41, 42, 43, 44, 45, 46, 47, 48, 49]. 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 the theoretical result, the numerical algorithms can, in principle, continue the results in the space of parameters to reach arbitrarily close to parameters where the objects described here break down. One can hope that these numerical explorations of the border – 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 apply 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 different 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 [50]. We also call attention to the very interesting numerical paper [51] 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 [52] points that these invariant objects may be useful in data reduction (see also [53]). We hope to come back to these problems.

1.4 Some Remarks on the Numerical Implementations

One advantage of the parameterization method is that inverting large matrix is not required in the algorithm. In fact, for each step of the iteration, the algorithm only requires $\mathcal{O}(N \times L)$ storage and $\mathcal{O}(N \times L)$ operations, where N is the size of the grid of choice, and L is the maximum order of truncation in the parameterization of the isochrons. With the superlinear convergence rate, the computation runs in seconds in standard laptops.

Recently, there have been many efforts towards the computation for the invariant circle and the corresponding stable (unstable) manifolds. [13] studies planar vector fields when the dynamics on the limit cycle conjugates to a rotation; [54, 31, 22] developed and implemented algorithms for the computation of invariant tori and their whiskers in quasi-periodic maps; [51] developed algorithms for computing the leaves in the foliation of an ODE near hyperbolic fixed point using contraction methods.

1.5 Organization of the Thesis

In Chapter 2, we derive our quasi-Newton algorithm and prove that quadractic convergence of the algorithm provided that there exists a approximate solution. The proof is based on a modified Nash-Moser Theorem. This chapter is based on [25]. In Chapter 3, we implement the algorithm on the dissipative standard map. We following different continuation paths and explore the "bundle merging" scenario. For non-resonant eigenvalues, we present a 3-dimensional version of the algorithm and implemented on the 3-D Fattened Arnold Family. This chapter is based on [14].

CHAPTER 2

EFFECTIVE ALGORITHM AND RIGOROUS PROOF OF CONVERGENCE

We present and analyze rigorously a quadratically convergent algorithm to compute an invariant circle and the foliation by stable manifolds for 2-dimensional maps. 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 set it in a self-contained appendix.

The algorithm in this chapter are very practical since it converge quadratically, require

moderate storage and operation count. Details of the implementation and results of the runs are described in Chapter 3.

The chapter is organized as follows. In Section 2.1, 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 1. 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 2.4. The proof of Theorem 4 is presented in Section 2.5, 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 12) which we present in Section 2.6. We hope that Theorem 12 can be of independent interest since it could be applicable to similar problems.

2.1   Setup of the Problem

In this section, we first briefly introduce the general idea of the parameterization method (Section 2.1.1). More detailed discussions about this method are in [19].

Then, in a manner inspired by [13], we formulate an invariance equation (2.2) for the invariant circle and stable foliation near it. (Section 2.1.2)

It is important to notice that the invariance equation (2.2) is very underdetermined. Taking advantage of this underdetermination, in Section 2.1.3, we find a version of the invariance equation with extra properties. In Section 2.1.4, we discuss the stable manifolds.

2.1.1   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 $\mathscr{A}$, given $f : \mathscr{A} \to \mathscr{A}$ 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 $\mathscr{K}$, we have that $\mathscr{K}$ is f-invariant if and only if the following invariance equation holds:

$$f \circ K(\theta) = K \circ a(\theta), \tag{2.1}$$

where the diffeomorphism $a : \Theta \to \Theta$ is the internal dynamics on Θ, and θ is the local coordinate in Θ (See Figure 2.1). The goal now becomes solving Equation (2.1) with $K(\theta), a(\theta)$ as the unknowns.

There are several methods to solve equation (2.1) 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 (2.1) amounts solving cohomological equations of the form as in equation (2.23), 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 (2.1). On the other hand, one has to adjust parameters. See [32, 33] 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 [21, 55], we obtain that the map a_λ is smoothly conjugate to a rotation.
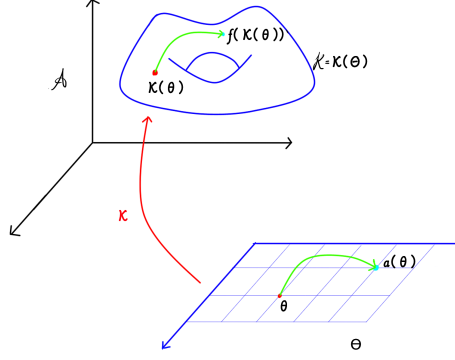
Figure 2.1: Parameterization of an invariant manifold (Figure taken from [19])

2.1.2   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 2.1.1, 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:

$$f \circ W(\theta, s) - W(a(\theta), \lambda(\theta, s)) = 0, \tag{2.2}$$

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

It is important to emphasize that the unknowns for equation (2.2) 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

15

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 (2.2) reduces to equation (2.1). 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.

2.1.3 Underdetermination of the Invariance Equation

One nice property of the invariance equation (2.2) 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 (2.2) is these changes of variables. Any two solutions of (2.2) close enough are related by a change of variables 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 (2.2) so that

16

we take advantage of them in numerical implementations.

Given $(W(\theta, s), a(\theta), \lambda(\theta, s))$ satisfying (2.2), 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 (2.2).

- 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

$$h(a(\theta), \widehat{\lambda}(\theta, s)) = \lambda(\theta, h(\theta, s)), \tag{2.3}$$

  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 (2.2).

According to Lemma 5 and its remarks, we can see that such $h(\theta, s)$ as in (2.3) 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 2.5.1. As remarked there, the existence of h satisfying (2.3) is a fibered version of Poincaré-Sternberg theorem on the linearization of contractions.

Benefiting from the second underdetermination property and Lemma 5, instead of considering (2.2) we can consider

$$f \circ W(\theta, s) - W(a(\theta), \lambda(\theta)s) = 0. \tag{2.4}$$

Our goal now becomes solving for $W(\theta, s), a(\theta)$ and $\lambda(\theta)$ from equation (2.4).

Note that, if W, a, λ is a solution of (2.4), 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 5 is to show that, if there was an invariant foliation, then, one can obtain a solution of (2.4) by reparameterizing it. Hence, finding a solution of (2.4) is not only sufficient for finding invariant foliations but also equivalent.

Remark 9. *If $a(\theta)$ conjugates to an irrational rotation $\theta + \omega$, $\lambda(\theta)$ is reduced to a constant. In fact, with $\lambda(\theta, s) = \lambda(\theta)s, h(\theta, s) = r(\theta)s$ equation (2.3) can be reduced to*

$$r(\theta + \omega)\lambda(\theta) = \lambda(\theta)r(\theta),$$

it follows that

$$\log(\lambda)(\theta) = \log(r)(\theta) - \log(r)(\theta + \omega)$$

forms a cohomological equation. To assure the existence of this equation, we can set $\log(\lambda)(\theta)$ to the average $\int_{\theta \in \mathbb{T}} \log(r)(\theta) - \log(r)(\theta + \omega)d\theta$, which is a constant.

2.1.4 Stable Manifolds of Points (Isochrons)

Notice that the invariance equation (2.2) contains not only the dynamics of the invariant circle, but also the dynamics in a neighborhood of the invariant circle. In particular, if equation (2.4) 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

$$
\begin{aligned}
f^{\circ j}(W(\theta, s)) &= W(a^{\circ j}(\theta), \lambda^{[j]}(\theta)s), \\
f^{\circ j}(W(\theta, 0)) &= W(a^{\circ j}(\theta), 0)
\end{aligned}
\tag{2.5}
$$

where

$$
\lambda^{[j]}(\theta) = \lambda(\theta)\lambda(a(\theta))\lambda(a^{\circ 2}(\theta)) \cdots \lambda(a^{\circ(j-1)}(\theta))
\tag{2.6}
$$

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

$$
\lambda^{[j+k]}(\theta) = \lambda^{[j]}(a^{\circ k}(\theta))\lambda^{[k]}
\tag{2.7}
$$

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*

$$
\lambda^* := \lim_{n \to \infty} \left(\|\lambda^{[n]}\|_{C^0} \right)^{\frac{1}{n}}
\tag{2.8}
$$

as the dynamical average.

Since $\|\lambda^{[n+m]}\|_{C^0} \leq \|\lambda^{[n]}\|_{C^0}\|\lambda^{[n+m]}\|_{C^0}$, the limit in (2.8) 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 (2.2) 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 [56, 57]. It suffices to realize that the relation

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

is an equivalence relation.

Remark 11. *Note that all points in the isochron have the same asymptotic behavior and that the convergence to this asymptotic behavior is reached exponentially with a rate λ^*. In the theory of normal hyperbolicity, it is standard that the stable manifold of a point x is the set of points whose asymptotic behavior gets closer to that of x faster than a certain exponential rate. It is not enough to require that they just get close. For example, in our case, if the dynamics in the invariant circle has an attractive periodic point, then, the points whose orbits are asymptotic to the periodic orbit is an open set. On the other hand, those that converge fast enough is the strong stable manifold of the periodic orbit. One very interesting regime is when the two eigenvalues of a periodic point are the same. In such a case, one expects that the circle stops existing as a C^1 manifold [41, 58], but it may persist as a topological circle or as a continuum [59, 60]. The phenomena that happen near the breakdown of normal hyperbolicity is still a challenging area.*

2.2 The Algorithm

In this section, we discuss our algorithm for solving the invariance equation (2.4). Unfortunately, (2.4) 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 2.5.3.

In Section 2.2.1, we present the details of one step of the quasi-Newton approach. Given an approximate solution, we look for the corrections such that it reduces the starting error superlinearly. The main task is solving some cohomological equations. The cohomological equations are solved in Section 2.2.2. we postpone the discussion of the step-by-step algorithm to Chapter 3. Later, we will show that the steps can be repeated infinitely often and indeed converges.

The algorithm formulated in this section has been implemented in Chapter 3 [14]. In Section 2.4 we will state a result on the convergence of the algorithm. As often happens, the algorithm is found to work with even in regions beyond the requirement of the rigorous proof.

2.2.1 The quasi-Newton Method

In this subsection, we perform one step of the quasi-Newton method to solve equation (2.4).

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

$$e(\theta, s) = f \circ W(\theta, s) - W(a(\theta), \lambda(\theta)s). \tag{2.9}$$

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

$$f(W + \Delta_W)(\theta, s) - (W + \Delta_W)((a + \Delta_a)(\theta), (\lambda + \Delta_\lambda)(\theta)s) = 0 \tag{2.10}$$

21

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 10.

Using Taylor expansion and omitting higher order terms, Equation (2.10) becomes

$$
\begin{aligned}
0 = & f(W(\theta, s)) + Df(W)(\theta, s)\Delta_W(\theta, s) - W(a(\theta), \lambda(\theta)s) \\
& - DW(a(\theta), \lambda(\theta)s) \begin{pmatrix} \Delta_a(\theta) \\ \Delta_\lambda(\theta)s \end{pmatrix} - \Delta_W(a(\theta), \lambda(\theta)s) + \text{higher order terms}, \quad (2.11)
\end{aligned}
$$

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 10.

Now we have that Equation (2.11) has become

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

$$
-\Delta_W(a(\theta), \lambda(\theta)s) = -e(\theta, s). \quad (2.12)
$$

Remark 12. *Notice that one should treat equation* (2.12) *as an equation of* $\Delta_W(\theta, s), \Delta_a(\theta)$ *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 (2.12), we will express $\Delta_W(\theta, s)$ in the frame $DW(\theta, s)$ as follows:

$$
\Delta_W(\theta, s) = DW(\theta, s)\Gamma(\theta, s). \quad (2.13)
$$

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 by Theorem 12 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 (2.9), we have that

$$De(\theta, s) = Df(W(\theta, s))DW(\theta, s) - DW(a(\theta), \lambda(\theta)s) \begin{pmatrix} Da(\theta) & 0 \\ D\lambda(\theta)s & \lambda(\theta) \end{pmatrix}. \quad (2.14)$$

Then, by substituting (2.13) and (2.14) in the quasi-Newton equation (2.12), we obtain

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

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

$$\triangleq \widetilde{e}(\theta, s), \quad (2.15)$$

where the term $De(\theta, s)\Gamma(\theta, s)$ is also omitted for the same reason as in equation (2.11), and the rigorous justification is again left to Lemma 10.

If we express equation (2.15) in components, we obtain the following two equations for the unknowns $\Gamma_1(\theta, s)$, $\Gamma_2(\theta, s)$, $\Delta_a(\theta)$ and $\Delta_\lambda(\theta)$.

$$Da(\theta)\Gamma_1(\theta, s) - \Delta_a(\theta) - \Gamma_1(a(\theta), \lambda(\theta)s) = \widetilde{e}_1(\theta, s), \quad (2.16)$$

$$\lambda(\theta)\Gamma_2(\theta, s) - \Delta_\lambda(\theta)s - \Gamma_2(a(\theta), \lambda(\theta)s) = \widetilde{e}_2(\theta, s) - D\lambda(\theta)s\Gamma_1(\theta, s) \quad (2.17)$$

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

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

2.2.2 Solving $\Gamma_{1,2}, \Delta_\lambda, \Delta_a$ from Equation (2.16), (2.17)

In this subsection, we present the details of solving equation (2.16) and (2.17). 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 2.4, $g(\theta, s) \in \mathcal{X}^{r,\delta}$ for some $\delta > 0$.

By matching coefficients of s^j on both sides, we can rewrite equation (2.16) and (2.17) 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 (2.16):

  ○ For the coefficients of s^0:

$$Da(\theta)\Gamma_1^{(0)}(\theta) - \Gamma_1^{(0)}(a(\theta)) - \Delta_a(\theta) = \widehat{e}_1^{(0)}(\theta), \qquad (2.18)$$

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

$$\Gamma_1^{(j)}(\theta) = \frac{\lambda^j(\theta)}{Da(\theta)}\Gamma_1^{(j)}(a(\theta)) + \frac{\widehat{e}_1^{(j)}(\theta)}{Da(\theta)}. \qquad (2.19)$$

- For equation (2.17):

  ○ 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

$$\Gamma_2^{(0)}(\theta) = \lambda(a^{-1}(\theta))\Gamma_2^{(0)}(a^{-1}(\theta)) - M^{(0)}(a^{-1}(\theta)), \qquad (2.20)$$

○ For the coefficients of s^1:

$$\lambda(\theta)\Gamma_2^{(1)}(\theta) - \Gamma_2^{(1)}(a(\theta))\lambda(\theta) - \Delta_\lambda(\theta) = M^{(1)}(\theta), \qquad (2.21)$$

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

$$\Gamma_2^{(j)}(\theta) = \lambda^{j-1}(\theta)\Gamma_2^{(j)}(a(\theta)) + \frac{M^{(j)}(\theta)}{\lambda(\theta)}. \qquad (2.22)$$

The hierarchy of equations above is well known from perturbation expansions. Algorithms for efficient computation of the coefficients can be found in [61].

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 (2.18) and (2.21) are underdetermined equations, hence the solution is not unique. An interesting question we have not yet pursued is how to choose the solution of (2.18) and (2.21) 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.

In this chapter, we choose the most obvious solution: For equation (2.18), we let $\Gamma_1^{(0)}(\theta) = 0$ and thus $\Delta_a(\theta) = -\widetilde{e}_1^{(0)}(\theta)$; for equation (2.21), 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 [19].

Notice that Equation (2.19), (2.20) and (2.22) have been reorganized so that are written as cohomological equation of the form:

$$\phi(\theta) = l(\theta)\phi(a(\theta)) + \eta(\theta). \qquad (2.23)$$

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

2.2.3 Solving ϕ from the Cohomological Equation (2.23)

In this subsection, we solve Equation (2.23) by contraction.

By inductively replacing $\phi(\theta)$ on the right hand side of (2.23) by the equation itself, we have

$$
\begin{aligned}
\phi(\theta) &= \eta(\theta) + l(\theta)\eta(a(\theta)) + l(\theta)l(a(\theta))\eta(a^{\circ 2}(\theta)) \\
&\quad + \ldots + l(\theta)l(a(\theta))l(a^{\circ 2}(\theta))\cdots l(a^{\circ(n-1)}(\theta))\eta(a^{\circ n}(\theta)) \\
&\quad + l(\theta)l(a(\theta))l(a^{\circ 2}(\theta))\cdots l(a^{\circ n}(\theta))\phi(a^{\circ(n+1)}(\theta)) \\
&= \sum_{j=0}^{n} l^{[j]}(\theta)\eta(a^{\circ j}(\theta)) + l^{[n+1]}(\theta)\phi(a^{\circ(n+1)}(\theta)),
\end{aligned}
\tag{2.24}
$$

where as in equation (2.6), $l^{[j]}(\theta) = l(\theta)l(a(\theta))l(a^{\circ 2}(\theta))\cdots l(a^{\circ(j-1)}(\theta))$, and $l^{[0]}(\theta) = 1$.

Note that, if $\|l^{[j]}\|_{C^0} < 1$, and ϕ is bounded, the last term in (2.24) tends to zero uniformly. Hence, the only possible C^0 solution of (2.23), is

$$
\phi(\theta) = \sum_{j=0}^{\infty} l^{[j]}(\theta)\eta(a^{\circ j}(\theta)).
\tag{2.25}
$$

As proved in Lemma 6 in Section 2.5.2, given r such that

$$
\|l\|_{C^0}\|Da\|_{C^0}^r < 1,
\tag{2.26}
$$

we will show $\|l^{[j]}(\cdot)\eta(a^{\circ j}(\cdot))\|_{C^r} \leq C\alpha^j$ 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, (2.23) has a C^r solution.

The conditions (2.26) 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 6, cohomological equation (2.23) 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 Section 2.6.

In this chapter, 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 [14] (Chapter 3) we will discuss implementation delays (discretization, programing considerations, and, more importantly diagnostics of reliability.

2.3 Scale of Banach Spaces

In this section, we set up the scale of Banach spaces that is needed in Section 2.4 and Section 2.5. Since for the problem we are dealing with, functions with domain $\mathbb{T}$ admits only finite regularity (Lemma 6), we first recall the C^r space ($r \in \mathbb{R}^+$) [28] 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 2.4. 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.

2.3.1 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 [27, 62, 28]. The properties of spaces of functions with mixed regularity used in this paper are built on those.

27

Before build up the $\mathcal{X}^r$ space, we first recall the C^r space.

Space for Functions in $\mathbb{T}$

Following [28, 63], the Hölder 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 13. *In this Chapter (excluding Section 2.6), we always denote $r \geq 0$ for the regularity, and we always have $n = \lfloor r \rfloor$ and $\alpha = r - n$.*

Remark 14. *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 15. *H_α is a seminorm and $H_\alpha(\phi) = 0$ if and only if ϕ is a constant.*

Remark 16. *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 [62] or $\hat{C}^r$ in [21, 27]).*

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 17. *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.*

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,

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

29

with norm

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

Remark 18. *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 19. *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 (2.4) 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 (2.4) to obtain the parameterization in a larger domain.

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 $\|f\|_{C^r} = \|f\|_{\mathcal{X}^{r,\delta}}$.

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

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 [28]:

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 20. *If $a : \mathbb{T} \to \mathbb{T}$ is only of α-Hölder continuity for $\alpha < 1$, the Hölder space is not preserved 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 2.5. We extract them here as an extension to [28] 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. $H_\alpha(D(a^{\circ k})) \leq kH_\alpha(Da)\|Da\|_{C^0}^{k(\alpha+1)-1},$

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 2.6, $\psi^{[k]} = \psi(a^{\circ(k-1)}) \cdots \psi(a)\psi,$

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

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

Proof. By Lemma 1, we have

1. $H_\alpha(D^p a \circ a^{\circ k}) \le H_\alpha(D^p a)\|Da^{\circ k}\|_{C^0}^\alpha \le H_\alpha(D^p a)\|Da\|_{C^0}^{k\alpha}$,

2. $H_\alpha(Da^{\circ k}) \le k\|Da\|_{C^0}^{k-1}\max_{0\le j\le k}H_\alpha(Da\circ a^{\circ j}) \le kH_\alpha(Da)\|Da\|_{C^0}^{k(\alpha+1)-1}$,

3. Suppose $D^p(a^{\circ k})$ has T_p terms, each term has F_p factors, then by

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

we have $F_n \le nk, T_n \le k^n(n-1)!$, 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 $0 \le p \le n, 0 \le q \le k$.

Thus we have

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

We also have

$$\begin{aligned}
H_\alpha(D^n a^{\circ k}) &\le k^n(n-1)!H_\alpha(\text{each term in } D^n a^k) \\
&\le k^n(n-1)!\Big(n\|Da\|_{C^0}^{n(k-1)}\|a\|_{C^n}^{n-1}\max_{0\le p\le n, 0\le q\le k}H_\alpha(D^p a\circ a^{\circ q}) \\
&\quad + n(k-1)\|Da\|_{C^0}^{n(k-1)-1}\max_{0\le q\le k}H_\alpha(Da\circ a^{\circ q})\|a\|_{C^n}^n\Big) \\
&\le k^n(n-1)!\Big(n\|Da\|_{C^0}^{n(k-1)}\|a\|_{C^n}^{n-1}H_\alpha(D^n a)\|Da\|_{C^0}^{k\alpha} \\
&\quad + n(k-1)\|Da\|_{C^0}^{n(k-1)-1}H_\alpha(Da)\|Da\|_{C^0}^{k\alpha}\|a\|_{C^n}^n\Big) \\
&\le k^{n+1}n!\|Da\|_{C^0}^{kr-n}\|a\|_{C^r}^{r+1}
\end{aligned}$$

Above all, we have

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

4. By the same notation and same method as (3), we have $F_n \le (n+1)k, T_n \le n! k^{n-1}$. 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 $0 \le p \le n, 0 \le q \le k$.

It follows that

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

and

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

Thus, we have

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

5. By running the same analysis on $\psi^{[k]}$, we have $F_n \le k(n+1), T_n \le k^n n!$, 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 $0 \le p \le n, 0 \le q \le k$.

It follows that

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

and

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

33

Above all,

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

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

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

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

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

□

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 [28], 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 21. *Generally speaking, the scale of spaces $C^r(U, X)$ does not admits continuous inclusion for general domain U (counterexample can be found in [28]). The continuous inclusion is guaranteed when U is a compensated open set [28].*

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

$$\|S_t u\|_\mu \leq t^{\mu-\lambda} C_{\lambda,\mu} \|u\|_\lambda \text{ for } u \in X_\lambda \tag{2.27}$$

and

$$\|(S_t - I)u\|_\lambda \leq t^{-(\mu-\lambda)} C_{\lambda,\mu} \|u\|_\mu \text{ for } u \in X_\mu \tag{2.28}$$

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

Remark 22. *The existence of the C^r-smoothing operator in $C^r(\mathbb{T}, X)$ is guaranteed by [28, 27].*

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 $j \geq 0$. 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:

$$S_t u(\theta, s) = \sum_{j=0}^\infty \hat{S}_t u^{(j)}(\theta)s^j. \tag{2.29}$$

where $\hat{S}_t$ is the smoothing operator in C^r space defined in Remark 22.

In our problem, since (2.4) 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 2.3.3.

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 18.

It is standard to see that this operator S_t defined in (2.29) satisfies condition (2.27) and (2.28), thus it is indeed a smoothing operator in $\mathcal{X}^{r,\delta}$.

Remark 23. *As shown in [27, 64, 28], 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*

$$\|u\|_v \leq C_{\gamma,\lambda,\mu} \|u\|_\lambda^{1-\gamma} \|u\|_\mu^\gamma. \tag{2.30}$$

Obtaining (2.30) *as a corollary of smoothing, leads to the conclusion only in the case that v is not an integer. In [28], there is a direct proof (even in greater generality).*

2.3.3 The $\mathscr{X}^{r,\delta}$ and $\mathscr{Y}^{r,\delta}$ Space

Our problem of solving (2.4) 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 $\mathscr{Y}^{r,\delta} = \mathcal{X}^{r,\delta} \times$

$\mathscr{X}^{r,\delta}$ with norm

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

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

Remark 24. *$\mathscr{X}^{r,\delta}, \mathscr{Y}^{r,\delta}$ are both scales of Banach spaces with smoothing operators. The smoothing operators comes natually from the smoothing operators in C^r and $\mathscr{X}^r$ spaces.*

Remark 25. *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.*

2.4 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 1.3, 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 2.3, we formulate our main result Theorem 4 and proceed to the proof in Section 2.5.3. Theorem 4 implies rather directly the result for foliations. We just need to verify that the operator entering in equation (2.4) satisfies the hypotheses of Theorem 4.

As indicated in Section 1.3, 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 (2.4). In other words, given $r \geq 0, \delta > 0$, we are looking for the zero of the funcional $\mathscr{F} : \mathscr{X}^r \to \mathscr{Y}^r$ where

$$\mathscr{F}[u] = \mathscr{F}[W, a, \lambda](\theta, s) = f(W(\theta, s)) - W(a(\theta), \lambda(\theta)s), \qquad (2.31)$$

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 26. *Restriction (1) can be generalized to be $\lambda^* < 1$, where λ^* is the dynamical average. If λ^* is used, one also need to adapt condition (3) accordingly (see Remark 35).*

Remark 27. *Restriction (3) is to guarantee m is bounded above in such a way that the regularity requirement for solving cohomological equations* (2.19)*,* (2.20) *and* (2.22) *covers the scale of regularities in Theorem 4. (See Lemma 6 for more details).*

Following the scheme derived in Section 1, we present theorem for the existence of solution for $\mathscr{F}[u] = 0$ as follows:

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 (2.31), 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 28. *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 Section 2.6. 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 29. *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 30. *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 31. *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 [14].*

This procedure is guaranteed to continue till some of the non-degeneracy assumptions of Theorem 4 fail. These assumptions are just 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 32. *As seen in several examples (e.g. in [15]) one can see that the optimal regularity of the invariant circle may decrease continuously to 0 as the parameters changes. For some parameter value, they will stop being C^2, for another parameter they will stop being C^1, etc. (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 [58] happens at different values of the places where they disappear as C^0 curves or as continua [59, 60].

Detailed numerical explorations of the behavior at breakdown of the hyperbolicity [41,

42, 43, 45, 46, 48, 49] has uncovered many interesting phenomena (e.g. scaling relations) that deserve detailed mathematical analysis.

Of course, detailed numerical explorations near the boundary is 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 12. introduced in Appendix 2.6. Details of the proof of Theorem 4 can be found in section 2.5.3.

2.5 Proof for the Analytical Result

This section can be mainly divided into 2 parts. In the first half, we include the proof for the existence of $h(\theta, s)$ in equation (2.3) discussed in Section 2.1.3 (see Section 2.5.1), and the discussion of the cohomological equation mentioned in (2.23) (see Section 2.5.2). In the second half, we present the proof of the Theorem 4. The idea of the proof is presented in Section 2.4. 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 12), which can be found in Appendix 2.6.

2.5.1 The Existence of $h(\theta, s)$ in Equation (2.3)

In this subsection, we prove Lemma 5. As indicated in Section 2.1.2. Lemma 5 ensures that the study of (2.4), 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 (2.4). 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 (2.3) holds for given $\lambda(\theta, s)$ and $\widehat{\lambda}(\theta, s)$. In the following Lemma 5, 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 5 (Existence of $h(\theta, s)$). There exists $\delta > 0$ such that for $\widehat{\lambda} \in \mathcal{X}^{r,\delta}, \widehat{\lambda}(\theta, s) = \lambda(\theta)s + N(\theta, s)$, 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 $h(\theta, s) \in \mathcal{X}^{r,\delta}$ such that equation (2.3): $h(a(\theta), \lambda(\theta)s) = \widehat{\lambda}(\theta, h(\theta, s))$ holds.

Remark 33. *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 34. *This Lemma 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 [65] 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 [66, 67] based on majorants can also be adapted.

Proof. By substituting the above $\widehat{\lambda}(\theta, s)$ and $\lambda(\theta)$ in equation (2.3), we have

$$h(a(\theta), \lambda(\theta)s) = \lambda(\theta)h(\theta, s) + N(\theta, h(\theta, s)). \tag{2.32}$$

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

$$h(\theta, s) = s + \widehat{h}(\theta, s), \tag{2.33}$$

where $\widehat{h}(\theta, s) = \mathcal{O}(s^2)$. By substituting (2.33) back into equation (2.32) and after some simplifications, we have

$$\widehat{h}(\theta, s) = \lambda^{-1}(\theta)[\widehat{h}(a(\theta), \lambda(\theta)s) - N(\theta, s + \widehat{h}(\theta, s))]. \tag{2.34}$$

Define 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 contraction.

Instead of directly 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

$$\mathscr{G}^{\circ k}[\widehat{h}] = (\lambda^{-1})^{[k]}(\theta)[\widehat{h}(a^{\circ k}(\theta), \lambda^{[k]}(\theta)s) - k\mathcal{O}(s^2)], \tag{2.35}$$

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 (2.35): $(\lambda^{-1})^{[k]}(\theta)\widehat{h}(a(\theta), \lambda^{[k]}(\theta)s) \triangleq \mathscr{L}[\widehat{h}]$ 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

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

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

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

$\square$

2.5.2 Estimates on Solutions of the Cohomological Equation (2.23)

We use this subsection to take a closer look at the cohomological equation mentioned in (2.23) with solution (2.25). The following result in Lemma 6 is used in both Section 2.2.1 and Section 2.5.

Lemma 6. 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. $\|Da\|^r_{C^0} \|l\|_{C^0} < 1$), then the cohomological equation (2.23): $\phi(\theta) = l(\theta)\phi(a(\theta)) + \eta(\theta)$ admits a unique C^r solution:

$$\phi(\theta) = \sum_{j=0}^{\infty} l^{[j]}(\theta)\eta(a^j(\theta)) \qquad (2.36)$$

with

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

Proof. First, we prove that (2.36) is a solution to equation (2.23). 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 plug this infinite sum back in (2.23) and rearrange terms to show that (2.36) is indeed a solution.

On top of this, we argue that (2.36) is the only C^0 solution. More explictly, if there were two solutions, then by the discussion in (2.24): $\phi(\theta) = \sum_{j=0}^{n} l^{[j]}(\theta)\eta(a^j(\theta)) + l^{[n+1]}(\theta)\phi(a^{n+1}(\theta))$, 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.

44

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

1. $\|a^{\circ k}\|_{C^r} \le k^n n! \|Da\|_{C^0}^{r(k-1)} \|a\|_{C^r}^{r+1}$,

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

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

4. $\|\eta(a^{\circ k})l^{[k]}\|_{C^r} \le C_r(\|l\|_{C^r}+\|a\|_{C^r})^{r+1}\|l\|_{C^0}^{-n}\left[k^n(\|l\|_{C^0}\|Da\|_{C^0}^r)^k\right]\|\eta\|_{C^r}$

Thus from (2.25), we have

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

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

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

which finishes the proof. $\qquad\square$

Remark 35. *Give $k \in \mathbb{N}^+$, by rewriting equation (2.23) into the form as in (2.24), 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 36. *Lemma 6 shows that if*

$$\|l\|_{C^0}\|Da\|_{C^0}^r < 1 \text{ (or } \left\|l^{[k]}\right\|_{C^0}\left\|D(a^{\circ(k+1)})\right\|_{C^0}^r < 1,) \qquad (2.37)$$

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 (2.37) can only be satisfied for a finite range of regularity r. It is not difficult to 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 (2.25) becomes

$$\phi(\theta) = \sum_{j=0}^{\infty} l^j \eta(\lambda^j\theta)) \tag{2.38}$$

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 6 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 (2.38).

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 equations by solutions of C^{∞} or C^{ω} problems.

2.5.3 Proof for Theorem 4

Following the same notation as in Theorem 4, we now justify the non-degeneracy conditions of the absolute Nash-Moser Theorem 12 one by one.

Lemma 7 (Condition 1). For δ, $\widetilde{B}_r(\rho)$ 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

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

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

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

$$\leq C_R \left(\frac{1}{1 - \frac{\| W_1 \|_{\mathscr{X}^{r,\delta}}}{R}} \right) \leq C_R \frac{1}{1 - \frac{\rho}{R}} \tag{2.39}$$

47

Thus

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

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

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

$$< \infty,$$

where the third line is because of (2.39) 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

$$\|W(a, \lambda s)\|_{\mathscr{X}^{r,\delta}} = \left\|\sum_{j=0}^{\infty} W^{(j)}(a)\lambda^j s^j\right\|_{\mathscr{X}^{r,\delta}}$$

$$\leq \sum_{j=0}^{\infty} \left\|W^{(j)}(a)\lambda^j\right\|_{C^r} \delta^j \leq \sum_{j=0}^{\infty} (2^r)^j \left\|W^{(j)}(a)\right\|_{C^r} \|\lambda\|_{C^r}^j \delta^j$$

$$\leq 2M_r \|a\|_{C^r}^r 2^{2n+1} \max_{0 \leq k \leq r}(\|\lambda\|_{C^r}^k) \max_{0 \leq j < \infty}(\|\lambda\|_{C^0}^{\max(j-n-1,0)} j^{2(n-1)}) \|W\|_{\mathscr{X}^{r,\delta}}$$

$$< \infty.$$

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

Lemma 8 (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 $h_1, k_1 \in \mathcal{X}^r \times \mathcal{X}^r$, $h_2, h_3, k_2, k_3 \in C^r$, we can calculate the first and second order Fréchet derivatives as follows:

$$D\mathscr{F}[u](h) = Df(W)h_1 - \partial_1 W(a, \lambda)h_2 - \partial_2 W(a, \lambda)h_3 - h_1(a, \lambda),$$

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

Thus we have

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

and

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

$\square$

Lemma 9 (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 1.

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 6, by equation (2.19), (2.20) and (2.22), 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 (2.18) and (2.21), 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 $\|\Delta_W\|_r \leq C_{r,\widetilde{B}_r(\rho)} \|e(\theta,s)\|_r$, $\|\Delta_a\|_r \leq C_{r,\widetilde{B}_r(\rho)} \|e\|_r$, and $\|\Delta_\lambda\|_r \leq C_{r,\widetilde{B}_r(\rho)} \|e\|_r$, which finishes the proof. $\qquad\square$

Lemma 10 (Condition 4). For $u \in \tilde{B}_m$, we have

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

Proof.

$$\mathscr{F}[u] - D\mathscr{F}[u]\eta[u]\mathscr{F}[u]$$

$$= \mathscr{F}[u] + D\mathscr{F}[u]\Delta_u$$

$$= \mathscr{F}[u + \Delta_u] + \mathscr{O}(\Delta^2)$$

$$= f(W + \Delta_W) - (W + \Delta_W)(a + \Delta_a, (\lambda + \Delta_\lambda)s) + \mathscr{O}(\Delta^2)$$

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

$$\quad + D\Delta_W(a, \lambda s)\begin{pmatrix} \Delta_a \\ \Delta_\lambda s \end{pmatrix} + \mathscr{O}(\Delta^2)$$

$$= -DW(a, \lambda s)\left[\begin{pmatrix} Da & 0 \\ D\lambda s & \lambda \end{pmatrix}\Gamma - \begin{pmatrix} \Delta_a \\ \Delta_\lambda s \end{pmatrix} - \Gamma(a, \lambda s) - \tilde{e}\right]$$

$$\quad - De\Gamma + D\Delta_W(a, \lambda s)\begin{pmatrix} \Delta_a \\ \Delta_\lambda s \end{pmatrix} + \mathscr{O}(\Delta^2)$$

$$= -De\Gamma + D\Delta_W(a, \lambda s)\begin{pmatrix} \Delta_a \\ \Delta_\lambda s \end{pmatrix} + \mathscr{O}(\Delta^2).$$

By the proof in Lemma 9, and that $\|De\|_{m-2} \leq C_{r,\tilde{B}_r(\rho)} \|e\|_{m-1}$, $\|D\Delta_W\|_{m-1} \leq C_{r,\tilde{B}_r(\rho)} \|e\|_{m-1}$, we achieve

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

$\square$

Lemma 11 (Condition 5). For $u \in \tilde{B}_m \cap \mathscr{X}^{m+2}$, we have $\|\mathscr{F}[u]\|_{m+2} \leq C_{r,\widetilde{B}_r(\rho)}(1 + \|u\|_{m+2})$.

Proof. From the proof of Lemma 7 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. $\square$

Since all the constant $C_{r,\widetilde{B}_r(\rho)}$ we get from Lemma 7, 8, 9, 10 and 11 are universal for u in the respective domain $\widetilde{B}_r(\rho)$, we have finished proving all the non-degeneracy conditions required by Theorem 12 in the Appendix 2.6. Thus we have proved Theorem 4.

2.6   An Abstract Implicit Function Theorem in Scales of Banach Spaces

In this section, 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 1. We hope that this theorem can also be applied in some other problems.

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 1.3, 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 [26].

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 [68, 29, 30] 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 12. Let $m > 2$ and $\mathscr{X}^r$, $\mathscr{Y}^r$ for $m \leq r \leq m+2$ 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:

- $\mathscr{F}\big(\widetilde{B}_r(\rho) \cap \mathscr{X}^r\big) \subset \mathscr{Y}^r$ for $m \leq r \leq m+2$.

- $\mathscr{F}|_{\tilde{B}_m \cap \mathscr{X}^r} : \widetilde{B}_r(\rho) \cap \mathscr{X}^r \to \mathscr{X}^r$ has two continuous Fréchet derivatives, and satisfy the following bounded conditions:

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

- $\|\eta[u]\mathscr{F}[u]\|_r \leq C \|\mathscr{F}[u]\|_r$, $u \in \widetilde{B}_r(\rho)$ for $r = m-2, m+2$.

- $\|(D\mathscr{F}[u]\eta[u] - Id)\mathscr{F}[u]\|_{m-2} \leq \|\mathscr{F}[u]\|_m \|\mathscr{F}[u]\|_{m-1}$, $u \in \tilde{B}_m$.

53

- $\|\mathscr{F}[u]\|_{m+2} \leq C(1 + \|u\|_{m+2}), u \in \tilde{B}_m \cap \mathscr{X}^{m+2}$.

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,

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

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

$$u_n = u_{n-1} - S_{t_{n-1}} \eta[u_{n-1}] \mathscr{F}[u_{n-1}], \tag{2.40}$$

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

$(P_1 n)$: $u_n \in \tilde{B}_m$,

$(P_2 n)$: $\|\mathscr{F}[u_n]\|_{m-2} \leq v e^{-2\mu\beta\kappa^n}$,

$(P_3 n)$: $1 + \|u_n\|_{m+2} \leq v e^{2\alpha\beta\kappa^n}$.

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 $P_3(n = 0)$: $1 + \|u_0\|_{m+2} \leq e^{2\alpha\beta}$ 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

$$\|\eta[u_n]\mathscr{F}[u_n]\|_{\mathscr{X}^{m-2}} \leq C \|\mathscr{F}[u_n]\|_{\mathscr{Y}^{m-2}} \leq C v e^{-2\mu\beta\kappa^n}, \tag{2.41}$$

it follows from (2.40) and (2.41) that

$$
\begin{aligned}
\|u_n - u_{n-1}\|_{\mathscr{X}^m} &= \left\|S_{t_{n-1}}\eta[u_{n-1}]\mathscr{F}[u_{n-1}]\right\|_{\mathscr{X}^m} \\
&\leq Ct_{n-1}^2 \left\|\eta[u_{n-1}]\mathscr{F}[u_{n-1}]\right\|_{\mathscr{X}^{m-2}} \\
&\leq Cve^{2\beta\kappa^{n-1}(1-\mu)}.
\end{aligned}
\tag{2.42}
$$

where the second inequality above comes from (2.27).

Thus we have

$$
\begin{aligned}
\|u_n - u_0\|_{\mathscr{X}^m} &\leq \sum_{j=1}^{\infty} \|u_j - u_{j-1}\|_{\mathscr{X}^m} \\
&\leq Cv(e^{2\beta(1-\mu)} + e^{2\beta(1-\mu)\kappa} + e^{2\beta(1-\mu)\kappa^2} + e^{2\beta(1-\mu)\kappa^3} + e^{2\beta(1-\mu)\kappa^4} + \sum_{j=6}^{\infty} e^{2\beta(1-\mu)\kappa^j}) \\
&\leq Cv(e^{2\beta(1-\mu)} + e^{2\beta(1-\mu)\kappa} + e^{2\beta(1-\mu)\kappa^2} + e^{2\beta(1-\mu)\kappa^3} + e^{2\beta(1-\mu)\kappa^4} + \frac{e^{2\beta(1-\mu)\kappa^6}}{1 - e^{2\beta(1-\mu)\kappa}}) \\
&\leq \rho.
\end{aligned}
\tag{2.43}
$$

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

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

$$
\begin{aligned}
\|\mathscr{F}[u_n]\|_{\mathscr{Y}^{m-2}} &\leq \left\|\mathscr{F}[u_n] - \mathscr{F}[u_{n-1}] + D\mathscr{F}[u_n]S_{t_{n-1}}\eta[u_{n-1}]\mathscr{F}[u_{n-1}]\right\|_{\mathscr{Y}^{m-2}} \\
&+ \left\|(Id - D\mathscr{F}[u_{n-1}]\eta[u_{n-1}])\mathscr{F}[u_{n-1}]\right\|_{\mathscr{Y}^{m-2}} \\
&+ \left\|D\mathscr{F}[u_{n-1}](Id - S_{t_{n-1}})\eta[u_{n-1}]\mathscr{F}[u_{n-1}]\right\|_{\mathscr{Y}^{m-2}}
\end{aligned}
\tag{2.44}
$$

and estimates the three terms one by one:

For the first line, by taylor expansion, the boundary condition in the second part of

55

assumption(2), (2.27) and (2.41), we have

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

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

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

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

$$\leq Cve^{(1-2\mu)2\beta\kappa^{n-1}} \tag{2.45}$$

For the second line, by assumption(4), (2.30), $P_2(n-1)$, assumption (5) and $P_3(n-1)$, we have

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

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

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

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

$$\leq Cv^{\frac{3}{2}}e^{(-\frac{5}{2}\mu+\frac{3}{2}\alpha)\beta\kappa^{n-1}} \tag{2.46}$$

For the third line, by (2.28), (2.41) and assumption (5), we have

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

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

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

$$\leq Cve^{2\beta\kappa^{n-1}(\alpha-2)} \tag{2.47}$$

thus, in order to show that (2.48) 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

$$C(e^{(1-2\mu+\mu\kappa)2\beta\kappa^{n-1}} + v^{\frac{1}{2}}e^{(-3\mu+\alpha+\mu\kappa)2\beta\kappa^{n-1}} + ve^{(\alpha-2+\kappa)2\mu\beta\kappa^{n-1}}) < 1, \qquad (2.48)$$

which can be satisfied if μ, κ and α satisfies

$$\begin{cases} 1 - 2\mu + \mu\kappa < 0, \\ -3\mu + \alpha + \mu\kappa < 0, \\ \alpha - 2 + \kappa < 0. \end{cases} \qquad (2.49)$$

and β is picked big enough.

As for $P_3(n)$, by (2.42), (2.40), (2.40), (2.41), assumption (5) and $P_3(j)$ for $j < n - 1$, we have

$$\begin{aligned} 1 + \|u_n\|_{\mathscr{X}^{m+2}} &\leq 1 + \sum_{j=1}^{n} \|u_j - u_{j-1}\|_{m+2} \\ &\leq 1 + \sum_{j=1}^{n} \|S_{t_{j-1}}\eta[u_{j-1}]\mathscr{F}[u_{j-1}]\|_{m+2} \\ &\leq 1 + C\sum_{j=1}^{n} \|\eta[u_{j-1}\mathscr{F}[u_{j-1}]]\|_{m+2} \\ &\leq 1 + C\sum_{j=1}^{n} (1 + \|u_{j-1}\|_{\mathscr{X}^{m+2}}) \\ &\leq 1 + C\sum_{j=1}^{n} e^{2\alpha\beta\kappa^{j-1}} \qquad (2.50) \end{aligned}$$

we need

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

that is

$$e^{-2\alpha\beta\kappa^n} + C\sum_{j=1}^{n} e^{(1-\kappa)\alpha\beta\kappa^{j-1}} < v. \qquad (2.51)$$

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 (2.43), we have

$$\sum_{j=1}^{n} e^{(1-\kappa)\alpha\beta\kappa^{j-1}} \le e^{(1-\kappa)\alpha\beta} + e^{(1-\kappa)\alpha\beta\kappa} + e^{(1-\kappa)\alpha\beta\kappa^2} + e^{(1-\kappa)\alpha\beta\kappa^3} + e^{(1-\kappa)\alpha\beta\kappa^4} + \frac{e^{(1-\kappa)\alpha\beta 6\kappa}}{1 - e^{(1-\kappa)\alpha\beta\kappa}}$$

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

Above all, in oder to make sure that (2.43), (2.48) and (2.51) are true, we nee the following constrictions for κ, α and μ:

$$
\begin{cases}
\mu > 1, \\
\kappa > \sqrt[3]{5}, \\
1 - 2\mu + \mu\kappa < 0, \\
-3\mu + \alpha + \mu\kappa < 0, \\
\alpha - 2 + \kappa < 0.
\end{cases}
\tag{2.52}
$$

and β is big enough.

One possible solution for (2.52) is $\kappa = 1.75, \mu = 5$ and $\alpha = 0.05$ and β is big enough.

Up to this point, we have finished the proof for induction. By letting $n \to \infty$, the second assumption $\|\mathscr{F}[u_n]\|_{\mathscr{X}^{m-2}} \le v e^{-2\mu\beta\kappa^n}$ 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 (2.43), we have

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

which completes the proof. $\qquad\square$

Remark 37. *Although the result $\|u^* - u_0\|_m \le Cv = C\|\mathscr{F}[u_0]\|_{m-2}$ 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.

CHAPTER 3

IMPLEMENTATION DETAILS AND NUMERICAL EXPLORATIONS

In this Chapter, we present and implement the algorithm that is developed in [25] (As in Chapter 2). Namely, we compute the invariant circle and the corresponding stable manifolds for 2-dimensional maps. The algorithm is based on the parameterization method, and it is backed up by an a-posterior theorem established in [25](Chapter 2).

The algorithm we present works irrespective of whether the internal dynamics in the invariant circle is a rotation or it is phase-locked. The algorithm converges quadratically with linear operations and linear memory requirement.

We implement the algorithm on the dissipative standard map, follow different continuation paths along the perturbation and the drift parameter. We explore a breakdown phenomenon when the hyperbolicity of the map losses due to the merging of the bundles when the rotation number of the internal dynamics is not prescribed. We find empirically that as the perturbation increases, the minimum angle between the invariant circle and the corresponding stable manifolds decreases following a power law with universal exponent.

We also discussed the generalization of the algorithm to 3 dimensions, and implement it on the 3-dimensional Fattened Arnold Family (3D-FAF) map with non-resonant eigenvalues.

This chapter is organized as follows: In Section 3.1, we summarize the main algorithm developed in [25], and present the iterative steps in an algorithmic form (Algorithm 1). In Section 3.2, we discuss some implementation details regarding the algorithm. Section 3.3 is about applying the algorithm on parameter-dependent problems, where we present Algorithm 3 to identify the boundary (breakdown value) of the existing domain for parameters based on the continuation method. We use Section 3.4 to discuss numerical results regarding Algorithm 1 and Algorithm 3 for the dissipative standard map, along with some further

discussions regarding the corresponding properties. In Section 3.5, we briefly discuss the breakdown of the invariant circle when the perturbation is large. In Section 3.6, we generalize Algorithm 1 to 3-dimensional maps and present some numerical results regarding the 3D-FAF map for non-resonant eigenvalues.

3.1 The Numerical Algorithm

Given the approximate $W(\theta, s), a(\theta)$ and $\lambda(\theta)$, we now summarize the pseudo-code of the algorithm in Chapter 2 to compute $\Delta_W(\theta, s), \Delta_a(\theta)$ and $\Delta_\lambda(\theta)$, see Algorithm 1. We truncate functions in $\mathcal{X}^{r,\delta}$ up to the L-th order in the power series expansion w.r.t. s (from the analysists's point of view, $L = \infty$).

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 (2.4)

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: $\sum_{j=0}^{L} \widetilde{e}^{(j)}(\theta)s^j = \widetilde{e}(\theta, s) \leftarrow (DW(a(\theta), \lambda(\theta)s))^{-1}e(\theta, s)$ (Section 3.2.2),
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 (2.19) for $1 \leq j \leq L$ using Algorithm 2,
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 (2.20) using Algorithm 2,
11: Solve $\Gamma_2^{(j)}(\theta)$ from equation (2.22) for $2 \leq j \leq L$ using Algorithm 2,
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 $\lambda(\theta)$.

To solve the invariance equation (2.4), one simply iterates the steps in Algorithm 1 until either the error $\|e\|$ is small enough (the algorithm converges) or $\max\{\|W\|, \|a\|, \|\lambda\|\}$ exceed some certain value (the algorithm fails to converge).

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. 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.)

In fact, both the time and space requirement of Algorithm 1 are $\mathcal{O}(N \times L)$ for a single step of the iteration, where, again, N is the size of the grid for $\mathbb{T}^1$, and L is the order of truncation of $W(\theta, s)$. This is because all the steps in the algorithm are about summation, multiplication, division, and spline interpolation.

Remark 38. *In Algorithm 1, cohomological equations are solved using Algorithm 2. We remark that the while loop in Algorithm 2 will only repeat finite times (bounded above). In our implementation (Section 3.4.1), such while loop only repeat at most 10 times, which is equivalent to applying the contraction on the cohomological equation 2^{10} times, which is sufficient in most of the cases even when the tolerance is close to round-off error or when the contraction is slow.*

Remark 39. *In the case when $a(\theta)$ admits diophantine rotation number, Fourier Transform is commonly used [13, 61, 69, 19]. In these scenarios, despite the operations required becomes $\mathcal{O}(N \log N)$ for each iteration, the constant is smaller and the implementation is indeed faster than the spline interpolation.*

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 quadratically convergence contraction algorithm (more in [14]).

Of course, developing a high-quality practical algorithm requires developing criteria

that ensure correctness and monitor the accuracy. In that respect, having an a-posterior theorem is an invaluable help.

The proof of the convergence in Chapter 2 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.

3.2 Some Implementation Details

The implementation of Algorithm 1 requires some practical considerations in terms of function representation and functional operations. In this section, we provide some implementation details for the algorithm. We start with the representation of the functions, followed by the discussion of some basic functional operations: composition, inverse, etc. We then discuss the algorithm for solving the cohomological equation. At the end of this section, We also propose a draft method for parallel implementation.

3.2.1 Function Representation

The first thing we shall do is to determine a way of discretizing the functions. In order to perform the algorithm, there are two types of functions we need to deal with: Type-1: $f : \mathbb{T} \to \mathbb{R} \in C^r$; Type-2: $g : \mathbb{T} \times [-\delta, \delta] \to \mathbb{R} \in \mathcal{X}^{r,\delta}$.

Type-1 Functions

There are two major methods for discretizing functions of Type-1:

- Method 1: Discretizing $\mathbb{T}$ to a grid of points and storing $f(\theta)$ by the values on the grid. In this case, one can obtain function evaluations and derivatives through interpolation techniques (more specifically, periodic splines);

- Method 2: Representing the function under an orthonormal basis and storing the co-

efficients (for example, Fourier coefficients). To store functions under the spectral representation, one only needs to truncate the series to a suitable order. The evaluation and the derivatives of functions can then be computed accordingly.

In this paper, we will use method 1 to store $f(\theta)$ for the following reasons:

- Since the functions we are dealing with may have spikes or may lose regularities in the neighborhood of certain points when the parameters of the map come close to the breakdown value, using splines can allow us to partition $\mathbb{T}$ in a non-even manner to cope with these kinds of situations, which is not easily achieved by Fourier Transform.

- Solving (2.16) and (2.17) requires computing the composition of two functions, say $f \circ a$, where the internal dynamics, $a(\theta)$, in general does not conjugate to a rotation, thus it is complicated to use Fourier Transform (it is still doable, one can refer to [69] for further discussions). On the other hand, storing f and a via grid points produces a simpler and more reliable way to perform the composition $f \circ a$.

Remark 40. *One delicate point of using grid points representation for functions is when the regularity of the function drops below the order of splines used in interpolation. This happens when the perturbation is close to the breakdown.*

Remark 41. *Classic results ([70]) for the error of cubic spline approximation shows that*

$$\left\| (g - \widehat{g})^{(r)} \right\|_{C^\infty} \leq C_r \left\| g^{(4)} \right\|_{C^\infty} \left(\frac{1}{N} \right)^{4-r},$$

where $g \in C^4$, $\widehat{g}$ is the cubic spline approximation, $C_0 = \frac{5}{384}$, $C_1 = \frac{1}{24}$ and $C_2 = \frac{3}{8}$. Thus the accuracy drops as the regularity increases. The norm in Definition 2.3.1 is hence affected by this round-off error (as indicated in Subsection 3.4.1).

Type-2 Functions

Following Definition 2.3.1, function $g \in \mathcal{X}^{r,\delta}$ of Type-2, $g(\theta, s)$ can be written as the Taylor's series w.r.t. s, i.e. $g(\theta, s) = \sum_{j=0}^{\infty} g^{(j)}(\theta) s^j$, we can truncate $g(\theta, s)$ up to the L-th order in order to store it, provided that L is big enough so that $\sum_{j=0}^{L} g^{(j)}(\theta) s^j$ is a good approximation of $g(\theta, s)$. In this case, storing $g(\theta, s)$ is equivalent as storing $L+1$ functions of Type-1 as $g^{(j)}(\theta) \in C^r(\mathbb{T}^1, \mathbb{R}^1)$ for $j = 0, 1, \ldots, L$, which is essentially storing a 2-d array with size $N \times L$.

3.2.2   Composition Between Functions

As indicated in the algorithm, we need to cope with the composition between functions both in the C^r space and the $X^{r,\delta}$ space .

Coping with Function Composition in C^r

In Algorithm 1, the composition between two C^r functions is required when computing the error of the invariance equation and when deriving and solving the cohomological equations. In these scenarios, such operation can be abstracted as computing $f \circ g$, where $f : \mathbb{T}^1 \to \mathbb{R}$ can be functions either of index 0 or 1, and $g : \mathbb{T} \to \mathbb{T}$ is always of index 1. Recall index 0 function f satisfies $f(\theta + 1) = f(\theta)$, i.e. which is equivalent as periodic function, and index 1 function f satisfies $f(\theta + 1) = f(\theta) + 1$.

If $f(\theta)$ has index 0, $f \circ g$ can be calculated by splines with periodic boundary conditions. If $f(\theta)$ has index 1, $f = id + \hat{f}$, where $\hat{f}(\theta)$ is periodic, and thus $f \circ g = g + \hat{f} \circ g$.

Remark 42. *In the unknowns of equation (2.4), only $W_1^{(0)}(\theta)$ and $a(\theta)$ has index 1, while the other functions: $W_1^{(j)}(\theta)$ for $j = 1, \ldots L$, $W_1^{(j)}(\theta)$ for $j = 0, \ldots, L$ and $\lambda(\theta)$ have index 0.*

Remark 43. *One problem of using splines for the composition is the choice of order for the spline. More specifically, as indicated in Section 3.4.2, the invariant circle losses regularity*

when the parameter value is close to the breakdown, in which case we have to switch from cubic splines to Akima, quadratic or linear splines.

Remark 44. *During the composition between functions, for example, $W^{(0)} \circ a(\theta)$, the monotonicity of $a(\theta)$ is required. We remark that although the cubic spline does not guarantee such monotonicity, the results in [25] indicate that the $a(\theta)$ we are using in the implementation is accurate enough to assure such monotonicity. Nonetheless, although not used in this paper, one may use the Steffen's interpolation method [71] if necessary.*

Coping with Function Composition in $\mathcal{X}^{r,\delta}$

Let the same $f(\theta, s), W(\theta, s)$ as in (2.4) be our functions of our consideration, and furthermore, assume that $f(\theta, s)$ is only formed by the basic operations $(+, -, *, /)$ and compositions by exponential, logarithmic and trigonometric functions. The manipulation of the Taylor series of $f \circ W(\theta, s)$ is achieved by an online algorithm with the aid of radial derivative, also known as automatic differentiation.

We now take the example of computing $\sin(W_1(\theta, s))$ where $W_1(\theta, s) = \sum_{j=0}^{L} W_1^{(j)}(\theta)s^j$ as we will need this result in the example of Section 3.4. One can refer to section 2.3 in [19] and [69] for more detailed discussions in more general cases.

Let L be the maximum order of s we want to calculate, i.e., we start with $W_1(\theta, s) = \sum_{j=0}^{L} W_i^{(j)}(\theta)s^j$, and the goal is to find $\sin(W_1(\theta, s))$ up to the L-th order. Denote $S(\theta, s) \triangleq \sin(w_1(\theta, s)) = \sum_{j=0}^{L} S^{(j)}(\theta)s^j$ and $C(\theta, s) \triangleq \cos(w_1(\theta, s)) = \sum_{j=0}^{L} C^{(j)}(\theta)s^j$. By differentiating $S(\theta, s)$ and $C(\theta, s)$ with respect to s, and by noticing that

$$\partial_s S(\theta, s) = C(\theta, s)\partial_s W_1(\theta, s),$$

$$\partial_s C(\theta, s) = -S(\theta, s)\partial_s W_1(\theta, s),$$

we can have the following iterating formulae for $S^{(j)}(\theta), C^{(j)}(\theta)$:

$$S^{(j)}(\theta) = \frac{1}{j} \sum_{k=0}^{j-1} (j-k) W_1^{(j-k)}(\theta) C^{(k)}(\theta),$$

$$C^{(j)}(\theta) = -\frac{1}{j} \sum_{k=0}^{j-1} (j-k) W_1^{(j-k)}(\theta) S^{(k)}(\theta).$$

Starting with

$$S^{(0)}(\theta) = \sin(W_1(\theta, 0)) = \sin(W_1^{(0)}(\theta)),$$

$$C^{(0)}(\theta) = \cos(W_1(\theta, 0)) = \cos(W_1^{(0)}(\theta)).$$

we are now able to calculate the coefficient of $\sin(W_1)(\theta, s)$ up to the L-th order.

The Computation of $\widetilde{e}(\theta, s)$

For step (4) in Algorithm 1, instead of calculating the inverse of

$$DW \circ (a(\theta), \lambda(\theta)s) \triangleq \begin{pmatrix} \beta_{11}^{(0)}(\theta) & \beta_{12}^0(\theta) \\ \beta_{21}^0(\theta) & \beta_{22}^0(\theta) \end{pmatrix},$$

we will solve the following linear system:

$$(DW(a(\theta), \lambda(\theta)s))\widetilde{e}(\theta, s) = e(\theta, s),$$

which, after some routine calculation, is just to solve the following linear system inductively from $k = 0$ to L to get $\widetilde{e}^{(k)}(\theta)$.

$$\begin{pmatrix} \beta_{11}^{(0)}(\theta) & \beta_{12}^0(\theta) \\ \beta_{21}^0(\theta) & \beta_{22}^0(\theta) \end{pmatrix} \begin{pmatrix} \widetilde{e}_1^{(k)}(\theta) \\ \widetilde{e}_2^{(k)}(\theta) \end{pmatrix} = \begin{pmatrix} \hat{e}_1^{(k)}(\theta) \\ \hat{e}_2^{(k)}(\theta) \end{pmatrix}.$$

67

where

$$\hat{e}_1^{(k)}(\theta) = e_1^{(k)}(\theta) - \sum_{j=0}^{k}(\beta_{11}^{(k-j)}(\theta)\widetilde{e}_1^{(j)}(\theta) + \beta_{12}^{(k-j)}(\theta)\widetilde{e}_2^{(j)}(\theta)),$$

$$\hat{e}_2^{(k)}(\theta) = e_2^{(k)}(\theta) - \sum_{j=0}^{k}(\beta_{21}^{(k-j)}(\theta)\widetilde{e}_1^{(j)}(\theta) + \beta_{22}^{(k-j)}(\theta)\widetilde{e}_2^{(j)}(\theta));$$

Thus,

$$\begin{pmatrix}\widetilde{e}_1^{(k)}(\theta) \\ \widetilde{e}_2^{(k)}(\theta)\end{pmatrix} = \frac{1}{\delta(\theta)}\begin{pmatrix}\beta_{22}^{(0)}(\theta) & -\beta_{12}^{(0)}(\theta) \\ -\beta_{21}^{(0)}(\theta) & \beta_{11}^{(0)}(\theta)\end{pmatrix}\begin{pmatrix}\hat{e}_1^{(k)}(\theta) \\ \hat{e}_2^{(k)}(\theta)\end{pmatrix},$$

where $\delta(\theta) = \beta_{11}^{(0)}(\theta)\beta_{22}^{(0)}(\theta) - \beta_{21}^{(0)}(\theta)\beta_{12}^{(0)}(\theta)$.

3.2.3   Computation of the Approximate Inverse of the Internal Dynamics

Deriving equation (2.20) involves the computation of the inverse of the internal dynamics $a(\theta)$. In other words, given a, Δ_a, a^{-1}, the goal is to find the inverse of $a^+ \equiv a + \Delta_a$, of which we denote as a^-. We remark that computing a^- is a well-posed problem because of the standard implicit function theorem.

In this subsection, we present four methods for computing a^- as follows. Method 1 compute a^- directly via flipping the graph and interpolation. Method 2, 3, and 4 focus on computing the correction of a^{-1}, denoted by $\Delta_{a^{-1}}$ by solving some objective functions up to quadratic error. In method 2, 3, and 4, $a^- = a^{-1} + \Delta_{a^{-1}}$.

- Method 1: Reflecting the graph.

  This method takes advantage of the fact that the homeomorphism $a : \mathbb{T}^1 \to \mathbb{T}^1$ is strictly increasing and of index 1, and thus $a(\theta) = T_{a(0)} \circ \hat{a}(\theta)$, where $T_\alpha(x) = x + \alpha$ and $\hat{a}(\theta)$ is a strictly increasing function with $\hat{a}(0) = 0$ and $\hat{a}(1) = 1$, thus $\hat{a}$ is invertible. It follows that $a^{-1}(\theta) = \hat{a}^{-1} \circ T_{a(0)}^{-1}(\theta)$, where $T_\alpha(x) = x - \alpha$. To compute $\hat{a}^{-1}(\theta)$, where θ is a grid of points in $\mathbb{T}$, we can reflect $\hat{a}(\theta)$ over the line $\hat{a}(\theta) = \theta$ by treating $\hat{a}(\theta)$ as the new grid of $\mathbb{T}$, treating the original grid of θ as the corresponding values for $\hat{a}^{-1}(\theta)$, and then using splines to evaluate $\hat{a}^{-1}(\theta)$ at the original grid of θ.

This method has great performance provided the original grid is sufficiently large such that the values for $\hat{a}^{-1}(\theta)$ are spread over $\mathbb{T}^1$. If $a(\theta)$ has large slope in certain parts of $\mathbb{T}$, one can use adaptive grid and put more points on those parts to ensure the performance.

- Method 2: Compute "Left" inverse of a.

  This method, as discussed in [19], is trying to minimize the error when a^- is applied on the left of a^+. From

  $$\left(a^{-1} + \Delta_{a^{-1}}\right) \circ a^+(\theta) - \theta = 0,$$

  and by omitting the quadratically small terms, we have $\Delta_{a^{-1}}(\theta) = -e \circ a^{-1}(\theta)$, where $e(\theta) = a^{-1}(\theta) \circ a^+(\theta) - \theta$. It follows that the updated inverse of $a(\theta)$ is $a^{-1}(\theta) + \Delta_{a^{-1}}(\theta)$.

- Method 3: Compute "Right" inverse of a.

  Similar to Method 2, we now optimize the right side inverse by solving the following objective function:

  $$a^+ \circ \left(a^{-1} + \Delta_{a^{-1}}\right)(\theta) - \theta = 0,$$

  which, after omitting quadratically small term, gives $\Delta_{a^{-1}} = \frac{\theta - a^+ \circ (a^{-1})}{Da^+ \circ (a^{-1})}$.

- Method 4:

  By combining method 2 and 3, one may naturally aim to optimize

  $$\left(a^{-1} + \Delta_{a^{-1}}\right) \circ a^+(\theta) - a^+ \circ \left(a^{-1} + \Delta_{a^{-1}}\right)(\theta) = 0. \tag{3.1}$$

  However, this leads to a cohomological equation

  $$\Delta_{a^{-1}} \circ a^+(\theta) - \left(D(a^+) \circ a^{-1}\right)\Delta_{a^{-1}}(\theta) = a^+ \circ a^{-1}(\theta) - a^{-1} \circ a^+(\theta),$$

that is hard to solve through contraction.

To resolve this, we modify the objective function (3.1) as

$$(a^{-1} + \Delta_{a^{-1}}) \circ a^+(\theta) - \beta a^+ \circ (a^{-1} + \Delta_{a^{-1}})(\theta) + (\beta - 1)\theta = 0, \qquad (3.2)$$

where $\beta = \text{ceil}(\left\| \frac{1}{Da^+(a^{-1})} \right\|_{C^0})$. Equation (3.2) can then be rewritten as a cohomological equation

$$\Delta_{a^{-1}}(\theta) = \frac{1}{\beta Da^+(a^{-1})}(\theta)\Delta_{a^{-1}} \circ a(\theta) + \frac{e}{\beta Da^+(a^{-1})}(\theta),$$

where

$$e(\theta) = a^{-1} \circ a^+(\theta) - \beta a^+(a^{-1})(\theta) + (\beta - 1)\theta,$$

from which $\Delta_{a^{-1}}(\theta)$ can then be solved via Algorithm 2.

Generally speaking, method 1 tends to have a better performance in the early iterations, while method 2, 3, and 4 work better in the later iterations, as these methods requires Δ_a to be small, where Δ_a is of the same order as the error for the cohomological equation. In practice, we try all the methods and pick the one with the minimal error. One can refer to Section 3.4.1 for some numerical examples for the performance.

3.2.4   The Algorithm for Solving Cohomological Equations

Calculating $\Gamma_{1,2}^{(j)}(\theta)$ from step (7) and (10) requires solving the cohomological equation of the form in equation (2.23):

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

Remark 45. *The cohomological equation* (2.23) *occurs very often in KAM theory, usually in the case when $a(\theta)$ conjugates to a Diophantine rotation, in which case a formal*

70

solution can be achieved through Fourier series. In our case, however, $a(\theta)$ may have non-diophantine rotation number, or even may not conjugate to a rotation, thus we use grid interpolation to avoid small divisor issues. Moreover, when $a(\theta)$ does not conjugate to a rotation, equation (2.23) is hard to tackle using Fourier Transform.

Recall the discussion in (2.24), $\sum_{j=0}^{M} l^{[j]}(\theta)\eta(a^{\circ j}(\theta))$ is a good approximation of the unknown $\phi(\theta)$ provided the series converges (i.e., when the dynamical average $\lambda^* < 1$) and M is big enough. We can use the following algorithm to get to $\sum_{j=0}^{M} l^{[j]}(\theta)\eta(a^{\circ j}(\theta))$ with $\log M$ iterations:

Algorithm 2 Solving the cohomological equation (2.23)

Input: $l(\theta)$, $a(\theta)$, $\eta(\theta)$ and $tolerance$
Output: The solution of equation (2.23): $\phi(\theta)$

1:   $\phi(\theta) \leftarrow \eta(\theta)$,
2:   $L(\theta) \leftarrow l(\theta)$,
3:   $A(\theta) \leftarrow a(\theta)$,
4:   **while** $\|\phi(\theta) - l(\theta)\phi(a(\theta)) - \eta(\theta)\| > tolerance$ **do**
5:       $\phi(\theta) \leftarrow \phi(\theta) + L(\theta)\phi \circ A(\theta)$
6:       $L(\theta) \leftarrow L(\theta)L \circ A(\theta)$
7:       $A(\theta) \leftarrow A \circ A(\theta)$
8:   **end while**
9:   Return $\phi(\theta)$

Remark 46. *We emphasis that Algorithm 2 allows us to make the summation of M terms in $\log M$ steps. We hope such idea can be applied in more general applications.*

Remark 47. *We remark that the main idea of Algorithm 2 is similar to the binary expansion. More spercifically, in the begining of the $(k+1)$-th iteration, we have*

$$\phi(\theta) = \sum_{j=0}^{2^k} l^{[j]}(\theta)\eta(a^{\circ j}), L(\theta) = l^{[2^k]}(\theta), \text{ and } A(\theta) = a^{\circ 2^k}(\theta),$$

then inside the while loop, $\phi(\theta)$, $L(\theta)$ and $A(\theta)$ got updated to

$$\phi(\theta) = \sum_{j=0}^{2^{k+1}} l^{[j]}(\theta)\eta(a^{\circ j}), L(\theta) = l^{[2^{k+1}]}(\theta), \text{ and } A(\theta) = a^{\circ 2^{k+1}}(\theta).$$

Remark 48. *Because of the uniqueness of equation (2.23), solving equation (2.23) is the same as solving*

$$\phi(\theta) = \sum_{j=0}^{n} l^{[j]}(\theta)\eta(a^{\circ j}(\theta)) + l^{[n+1]}(\theta)\phi(a^{\circ(n+1)}(\theta)), \tag{3.3}$$

for any $n > 0$. Solving equation (3.3) for a suitable n gives better contraction and can make Algorithm 2 converges faster. However, one have to be careful with the trade-off between the speed up from the improved contraction and the extra compositions and summations required.

Remark 49. *Due to the accumulation of the truncation error and round-off errors, Algorithm 2 may not be able to reach the required accuracy. To resolve this, one can repeatly apply this algorithm for*

$$\Delta_\phi(\theta) = l(\theta)\Delta_\phi(a(\theta)) + \widehat{e}(\theta),$$

where $\widehat{e}(\theta) = \phi(\theta) - l(\theta)\phi(a(\theta)) - \eta(\theta)$ is the error for the cohomological equation, and the solution for such cohomological equation now becomes $\phi(\theta) + \Delta_\phi(\theta)$.

Remark 50. *From the discussion in [25], we have*

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

where the regularity r is bounded above:

$$r < -\frac{\ln \|l\|_{C^0}}{\ln \|Da\|_{C^0}},$$

indicating that such cohomological equation can only solved for a finite range of regularities.

3.2.5   Truncation Versus Smoothing Operator

In the proof of the hard implicit function theorem in [25], the corrections from the algorithm of $W(\theta, s), a(\theta)$ and $\lambda(\theta)$ are followed by a smoothing operator.

In our practical implementation, we did not use any smoothing operator. Instead, we argue that when the solution is approximate enough, the smoothing required by our Inverse Function Theorem is smaller than the numerical smoothing, which is caused both by the truncation of the function in $\mathbb{T} \times \mathbb{R}$ up to some finite order L, and by the choice of the representation of functions: grid of points with splines. One caveat is that we have to check that the spline have enough grid points so that the discretization is not an issue.

3.2.6   Validation of the Correctness of the Solution

The most naive way of validating the correctness of the solution is to check the norm of the error from the invariance equation and see if it is of the same order as the round-off error. However, this is not a thorough approach.

As mentioned in Section 3.2.1, one of the drawbacks of using splines is that it cannot capture all the information of the function, but only for a selected grid of points. It is possible, especially when the perturbation is close to the breakdown, that localized singularity appears in the solution with less regularity and with a higher frequency of oscillation. Depending on the choice of the grid, the splines may be smooth (when the localized singularity is in the gap of the grid) or not (when the singularity is close to one of the grid point), and the latter scenario leads to a disaster, especially when the regularity drops below the order of choice of our spline.

More specifically, in the case when the rotation number is rational, the localized singularity happens at two hyperbolic points when their corresponding stable and unstable

manifolds meet. This admits a C^α manifold, where $0 < \alpha < 1$ is regulated by the eigen-values at the attractive periodic orbit.

By the above discussions, even when the solution we computed induces small error, it is still not guaranteed that the solution is the true solution. We considered several methods of validating the correctness of the solution as follows:

- Method 1: Grid shifting: To guarantee that the grid of choice captures the information of the invariant circle and isochrons, i.e., no singularity points in the gap of the grid, we can either shift the grid, or increase/double the size of the grid (and maybe repeatedly doing this) and check if the absolute error for the invariance equation remains to be small. In our actural implementation, we choose the grid adaptively and use the grid size doubling method.

- Method 2: Monitor the norm:

  According to the discussion regarding the existence of the solution in [25], we know that the solution we get through iterations lies inside a neighborhood of the initial condition. Thus our theoretical result guarantees that the norm of $W(\theta, s), a(\theta)$ and $\lambda(\theta)$ can not explode.

- Method 3: Monitor the error of the invariance equation:

  Another aspect comes from the proof of our theoritical result in [25] is that the error for the invariance equation: $e(\theta, s)$ has to satisfy $\|e(\theta, s)\|_{m-2} \leq v e^{-2\mu\beta\kappa^n}$ for some prescribed positive $v, \mu, \beta > 0$, and $\kappa > 1$. Thus, if the solution is valid, the convergence rate of our iteration process has to be superlinear.

We have included a brief example regarding the above discussion in Section 3.2.6

3.2.7   A Proposal to a Parallel Implementation

Although the parallelization is not implemented in this paper, we remark that Algorithm 1 can be parallelized using a 2-dimensional mesh as the communication network. In this

subsection, we propose a draft framework when multiple processors are available.

Two-dimensional communication network (with not wraparound link) is an embedding in the way that the processors are arranged to a matrix, where each processor can communicate with their nearest neighbors in the same row or column. This type of embedding is commonly used in matrix manipulation.

Function Representation in 2-dim Network

In our algorithm, the functions we are dealing with are either of Type-1 (see Section 3.2.1): i.e. $a(\theta), \lambda(\theta)$, which are stored as a 1-D array, or of Type-2: i.e. $W(\theta, s)$, which is stored as a matrix of size $N \times L$. We can block distribute Type-1 functions onto the first row of the grid, and block distribute Type-2 functions on the 2-D grid.

More precisely, suppose the 2-D network has size $P_N \times P_L$, and N, L are integer multiples of P_N, P_L, respectively, then each processor on the first row stores $\frac{N}{P_N}$ points for functions of Type-1, and each processor stores $\frac{N}{P_N} \times \frac{L}{P_L}$ points for functions of Type-2.

Storing functions blockwise in this way allows the basic functional operations (for example, summation, subtractions, Type-1 function multiplication) to be processed within each processor in a pointwise manner without communication. The delicate part is the operations when communications between processors are needed. In our algorithm, such operations include function evaluation (for composition) and Taylor polynomial multiplication.

Cubic Spline Interpolation in Parallel

Interpolations are required in our algorithm whenever composition between functions or derivative of functions are needed. Because of the way functions are stored, cubic spline interpolations need to be performed on each row of the communication network.

The computation for the coefficients of cubic splines is essentially achieved by solving a linear system, where the matrix involved is only a tridiagonal matrix. Thus the goal is to

solve this tridiagonal system on a 1-D network.

Following [72], such system can be reduced to a single variable equation recursively in $\log_3 N$ steps, which, when implemented parallelly, has time complexity $\mathcal{O}((\tau + \mu) \log P_N)$ (for communication) and $\mathcal{O}(\log \frac{N}{P_N})$ (for computation), where τ is the latency time, μ is the inverse of the bandwidth and m is the size of the message communicating between processors (in this case is just a constant). Pipelining techniques [73] can be used speedup this computation.

To perform composition, we need to first use the hypercubic AllGather technique to have all the processors in the same row gather all the pieces of the spline (with communication time complexity $\mathcal{O}(\tau \log P_N + \mu N)$). For example, when computing $W(a(\theta), \lambda(\theta)s)$, one need to first BroadCast every piece of $a(\theta)$ from the first row to the rest rows along the corresponding columns (with communication time complexity $\mathcal{O}((\tau + \mu) \log P_N)$), compute $\lambda^i(\theta)$ for row i using parallel prefix sum technique (with communication time $\mathcal{O}(\tau \log P_L + \mu \frac{N}{P_N} \log P_L)$ and computation time $\mathcal{O}(\frac{L \times N}{P_L \times P_N} + \frac{N}{P_N} \log P_L)$), perform the interpolation for each row for the corresponding $W^{(i)}(\theta)$, $i = 0, \ldots, L$, AllGather the spline on each processor, and then evaluate on $a(\theta)$, and then multiply it by the corresponding $\lambda^i(\theta)$ pointwisely.

Type-2 Function Multiplication in Parallel

In our algorithm, products between Type-2 functions are also required. Given $g_1(\theta, s) = \sum_{i=0}^{L} g_1^{(i)}(\theta) s^i$, $g_2(\theta, s) = \sum_{i=0}^{L} g_2^{(i)}(\theta) s^i$, we need to compute $g_1 \times g_2$ up to the L-th order. To achieve this, one can AllGather g_2 for each column, compute the corresponding coefficients, and then AllReduce (with time complexity $\mathcal{O}(\tau \log P_L) + \mu \frac{N}{P_N} \log P_L$) along each column to update the coefficients for the product.

3.3 Continuation Method

By the previous discussions, for a given map $f : \mathbb{T} \times \mathbb{R} \to \mathbb{T} \times \mathbb{R}$, starting with some initial approximation, we are now ready to apply the iterative Algorithm 1 several times until the error is close to the round off error, with the solution for the invariance equation (2.4): $W(\theta, s)$, $a(\theta)$ and $\lambda(\theta)$.

In this section, we do some further discussions of applying our algorithm to the parameter-dependent problems. The idea is based on the standard continuation method.

3.3.1 The Continuation Method

In the parameter-dependent problems, it is a common approach to start with a simple (unperturbed) scenario and perform the continuation method to moving the parameters gradually towards the desired value.

Basic Idea

To have a nice initial approximation of the solution to apply our algorithm, it is common to start from the integrable system with no perturbation, in which case the invariant circle is easier to compute. In each step of the continuation, one can use the result from the previous step to calculate the starting approximation.

Let f_ϵ be a parameter family of diffeomorphim, the goal is to find W_ϵ, a_ϵ and λ_ϵ such that the invariance equation:

$$f_\epsilon \circ W_\epsilon(\theta, s) - W_\epsilon(a_\epsilon(\theta), \lambda_\epsilon(\theta)s) = 0 \tag{3.4}$$

holds, starting from $\epsilon = 0$.

Given the solution for equation (3.4) for some ϵ: $W_\epsilon(\theta, s)$, $a_\epsilon(\theta)$ and $\lambda_\epsilon(\theta)$, the goal is

to find the starting approximation: $W_{\epsilon+h}(\theta, s)$, $a_{\epsilon+h}(\theta)$ and $\lambda_{\epsilon+h}(\theta)$ for

$$f_{\epsilon+h} \circ W_{\epsilon+h}(\theta, s) - W_{\epsilon+h}(a_{\epsilon+h}(\theta), \lambda_{\epsilon+h}(\theta)s) = 0. \tag{3.5}$$

The naive choice would be the 0-th order approximation, i.e. $W_{\epsilon+h}(\theta, s) = W_{\epsilon}(\theta, s)$, $a_{\epsilon+h}(\theta) = a_{\epsilon}(\theta)$ and $\lambda_{\epsilon+h}(\theta) = \lambda_{\epsilon}(\theta)$. We remark that this choice is already good enough if the increment of the perturbation h is small.

One can also try to look for 1-st order approximations, which is looking for $\frac{\partial W_{\epsilon}}{\partial \epsilon}(\theta, s)$, $\frac{\partial a_{\epsilon}}{\partial \epsilon}(\theta)$ and $\frac{\partial \lambda_{\epsilon}}{\partial \epsilon}(\theta)$ such that

$$W_{\epsilon+h}(\theta, s) = W_{\epsilon}(\theta, s) + \frac{\partial W_{\epsilon}}{\partial \epsilon}(\theta, s)h,$$

$$a_{\epsilon+h}(\theta) = a_{\epsilon}(\theta) + \frac{\partial a_{\epsilon}}{\partial \epsilon}(\theta)h,$$

$$\lambda_{\epsilon+h}(\theta) = \lambda_{\epsilon}(\theta) + \frac{\partial \lambda_{\epsilon}}{\partial \epsilon}(\theta)h \tag{3.6}$$

satisfies equation (3.5) up to quadratic error $\mathcal{O}(h^2)$.

The procedure of this computation is similar to the derivation in Section 1: by (3.4) and (3.5), and omitting quadratically small terms, we have

$$Df_{\epsilon} \circ W_{\epsilon}(\theta, s)\frac{\partial W_{\epsilon}}{\partial \epsilon}(\theta, s) - DW_{\epsilon}(a_{\epsilon}(\theta), \lambda_{\epsilon}(\theta)s)\begin{pmatrix} \frac{\partial a_{\epsilon}}{\partial \epsilon}(\theta) \\ \frac{\partial \lambda}{\partial \epsilon}(\theta)s \end{pmatrix} - \frac{\partial W_{\epsilon}}{\partial \epsilon}(a_{\epsilon}(\theta), \lambda_{\epsilon}(\theta)s)$$

$$= -\frac{\partial f_{\epsilon}}{\partial \epsilon}(W_{\epsilon})(\theta, s) \triangleq E_{\epsilon}(\theta, s) \tag{3.7}$$

By $\frac{\partial W_{\epsilon}}{\partial \epsilon}(\theta, s) = DW_{\epsilon}(\theta, s)\eta_{\epsilon}(\theta, s)$ and by differentiating (3.4), we end up with

$$\begin{pmatrix} Da_{\epsilon}(\theta) & 0 \\ D\lambda_{\epsilon}(\theta)s & \lambda_{\epsilon}(\theta) \end{pmatrix}\eta_{\epsilon}(\theta, s) - \begin{pmatrix} \frac{\partial a_{\epsilon}}{\partial \epsilon}(\theta) \\ \frac{\partial \lambda_{\epsilon}}{\partial \epsilon}(\theta)s \end{pmatrix} - \eta_{\epsilon}(a(\theta), \lambda(\theta)s) = E_{\epsilon}(\theta, s), \tag{3.8}$$

By solving equation (3.8) with the same technique used in Section 3.2.4, we can have valid

$\frac{\partial W_\epsilon}{\partial \epsilon}(\theta, s)$, $\frac{\partial a_\epsilon}{\partial \epsilon}(\theta)$ such that (3.6) admits the first-order starting approximation for equation (3.5).

Remark 51. *Similar to what happened in [19], the first-order continuation does not produce a significant improvement on the initial approximation than the 0-th order continuation, both when the perturbation is small, and when the perturbation is close to the breakdown. In fact, when the perturbation is small, the gain from the first-order continuation is not as much as the gain from one step of our quadratic convergence algorithm; when the perturbation is close to the breakdown, the error from the first-order approximation is actually very large because of the blow-up of the norm for $W(\theta, s)$. For these reasons, we mainly used the 0-th order continuation in our implementation.*

The Continuation Algorithm

In terms of implementation, the increment step size h in Section 3.3.1 is chosen dynamically from coarse to fine. Inspired by [23], Algorithm 3 is the continuation algorithm we used for the family of maps f_ϵ with increasing parameters.:

Algorithm 3 Continuation algorithm before the breakdown

Input: $W_{\epsilon_0}(\theta, s), a_{\epsilon_0}(\theta), \lambda_{\epsilon_0}(\theta)$ for the integrable case
Input: $\Delta \epsilon$: The initial increment of the parameter ϵ
 1: **while** both $\Delta \epsilon$ and $\|(W, a, \lambda)\|$ are acceptable **do**
 2:     **if** the Algorithm 1 does not converge **then**
 3:        Move back to the solution (W, a, λ) before the increment of the parameters,
 4:        Decrease the increment: $\Delta \epsilon$,
 5:     **else**
 6:        Update the newly computed (W, a, λ),
 7:        **if** $\|f_\epsilon \circ W_\epsilon - W_\epsilon(a_\epsilon, \lambda_\epsilon s)\| > tolerance$ **then**
 8:           Move back to the solution (W, a, λ) before the increment of the parameters,
 9:           Decrease the increment: $\Delta \epsilon$,
10:        **end if**
11:        **if** $\|(W, a, \lambda)\|$ exceeds a certain value **then**
12:           Double the size of the grid points,
13:        **end if**
14:     **end if**
15: **end while**

Adaptative Grid for Continuation

It is essential that one check for the correctness of the solution before doing continuations. As discussed in Section 3.2.6, we need to find the appropriate grid under which the function has a better representation. This can be done by performing the line search for different grid sizes and find the one with the minimum error while applying our quasi-Newton algorithm.

Remark 52. *An aspect that requires extra caution is that the existence theorem in [25] does not guarantee the local uniqueness of the solution. Indeed, the solution is only unique under conjugacy (Remark 2.1.3). This results the drift of the solution as the parameter changes.*

3.4 Numerical Explorations

In this section, we take the dissipative standard map as an example to run the algorithm and explore some of the properties.

The dissipative standard map is a family of maps $f_{\eta,\gamma,k} : \mathbb{T} \times \mathbb{R} \to \mathbb{T} \times \mathbb{R}$ such that

$$\begin{pmatrix} \theta_{n+1} \\ p_{n+1} \end{pmatrix} \triangleq f_{\eta,\gamma,k}(\theta_n, p_n) = \begin{pmatrix} \theta_n + p_{n+1} + \eta, \\ \gamma p_n + \gamma k V'(\theta_n) \end{pmatrix}, \tag{3.9}$$

with $(\theta_n, s_n) \in \mathbb{T} \times \mathbb{R}$, $\gamma \in (0, 1)$ is the dissipative parameter, $k > 0$ is the perturbation parameter, η is the drift parameter, and $V(\theta)$ is an analytic, periodic function representing the kick from the kicked rotator. In this example, we shall consider the case when $V(\theta) = -\frac{1}{(2\pi)^2} \cos(2\pi\theta)$, then $V'(\theta) = \frac{1}{2\pi} \sin(2\pi\theta)$.

Remark 53. *If $\gamma = 1$, it can be shown that $\eta = 0$, and the (3.9) is reduced to be the Chirikov standard map.*

Remark 54. *One can easily verify that when $k = 0$, the solution to the invariance equation*

(2.4) *for $f_{\eta,\gamma,k}$ is*

$$W_{\eta,\gamma,k}(\theta,s) = \begin{pmatrix} \theta \\ \frac{\gamma}{\gamma-1}s \end{pmatrix}, \ a_{\eta,\gamma,k}(\theta) = \theta + \eta, \ \lambda_{\eta,\gamma,k}(\theta) = \gamma.$$

For a given choice of the parameters, we follow our quasi-Newton Algorithm 1 to solve the invariance equation (2.4) for $W_{\eta,\gamma,k}$, $a_{\eta,\gamma,k}$ and $\lambda_{\eta,\gamma,k}$. For the change of parameters, we can start with the unperturbed case ($k = 0$) and follow the continuation algorithm Algorithm 3.

In this example, we first discuss the behavior of the algorithm regarding the aspects pointed out in Section 3.2 through an instance of choice of parameters, and then we follow the following three continuation paths for further explorations:

1. Continuation with respect to the perturbation parameter k, with fixed γ and η;

2. Continuation with respect to the drift parameter η, with fixed k and γ;

3. Continuation with fixed rotation number for the internal dynamics $a_{\eta,\gamma,k}$, with fixed k, and γ, with η is tuned to ensure the preservance of the rotation.

Remark 55. *In practice, we choose the step size and grid size for the continuation of the parameters dynamically according to Algorithm 3.*

3.4.1 Example Solution

In this subsection, we will show the numerical performance of Algorithm 1 regarding the aspects discussed in Section 3.2. In the following examples, we set $\gamma = 0.5$, $\eta = 0.3$.

Convergence of the quasi-Newton Iteration

To demonstrate the convergence rate, we set $k = 0.3$ and use the solution when $k = 0$ (Remark 54) as the initial approximation, and iterate the algorithm several times until the

error $e(\theta, s)$ for the invariance equation meets the tolerance (see Table 3.1). In this example, $N = 1024$, $L = 10$ and $\delta = 0.001$.

Table 3.1: Convergence of the quasi-Newton Iteration

Number of Iteration	$\|e\|_{\mathcal{X}^{0,\delta}}$	$\|e\|_{\mathcal{X}^{1,\delta}}$	$\|e\|_{\mathcal{X}^{2,\delta}}$
1	9.710402e-03	6.101232e-02	3.833525e-01
2	2.860761e-04	3.274913e-03	5.806850e-02
3	5.587798e-06	9.021029e-05	2.031524e-03
4	4.152389e-10	7.825378e-09	2.376669e-07
5	2.645506e-14	9.540554e-13	2.810590e-09
6	3.889196e-16	3.030427e-13	2.737512e-09

Remark 56. *We remark that the choice of $k = 0.3$ and the initial approximation $(k = 0)$ are only used to demonstrate the convergence of the algorithm. In practice, we usually have a much smaller step size (less than 10^{-3}) and the initial error usually is of order 10^{-6}. The step size gets even much smaller when the perturbation is big and when k is close to the breakdown.*

Remark 57. *As discussed in Remark 43, we mainly used cubic splines in our implementation. Remark 41 illustrates the reason for the increase of the round-off error for $\|\cdot\|_1$ and $\|\cdot\|_2$ in Table 3.1.*

Validation of Correctness of the Solution

In this part, we continue with the solution achieved in Section 3.4.1 (when $k = 0.3$). Following the discussions in Section 3.2.6, we validate the solution through monitoring both the error when the grid size is doubled, and the norm of the solution. For the grid size equals to 1024, the corresponding data are recorded in Table 3.2. Since the error remains to be relatively small, we are more confident to say that the solution we achieved is indeed the true solution.

Remark 58. *As indicated in Table 3.2, the norm for the solution is relatively small. This*

Table 3.2: The performance when the grid size doubled.

	$\|\cdot\|_{\mathcal{X}^{0,\delta}}$	$\|\cdot\|_{\mathcal{X}^{1,\delta}}$	$\|\cdot\|_{\mathcal{X}^{2,\delta}}$
Error for the invariance equation	1.622146e-13	1.494237e-10	3.059292e-06
Norm for the solution	1.07970	1.225545	1.768379

happens because the choice of our perturbation is small. In fact, as discussed in Section 3.4.2, such norm will blow up when k is close to the breakdown value.

Methods for Computing the Inverse of $a(\theta)$

In this paragraph, we present some numerical results regarding the methods proposed in inverting the internal dynamics $a(\theta)$ (see Section 3.2.3). Again, the data we achieved is from the same experiment as the previous sections, with exactly the same convergence as Table 3.1. The detailed data is presented in Figure 3.1.
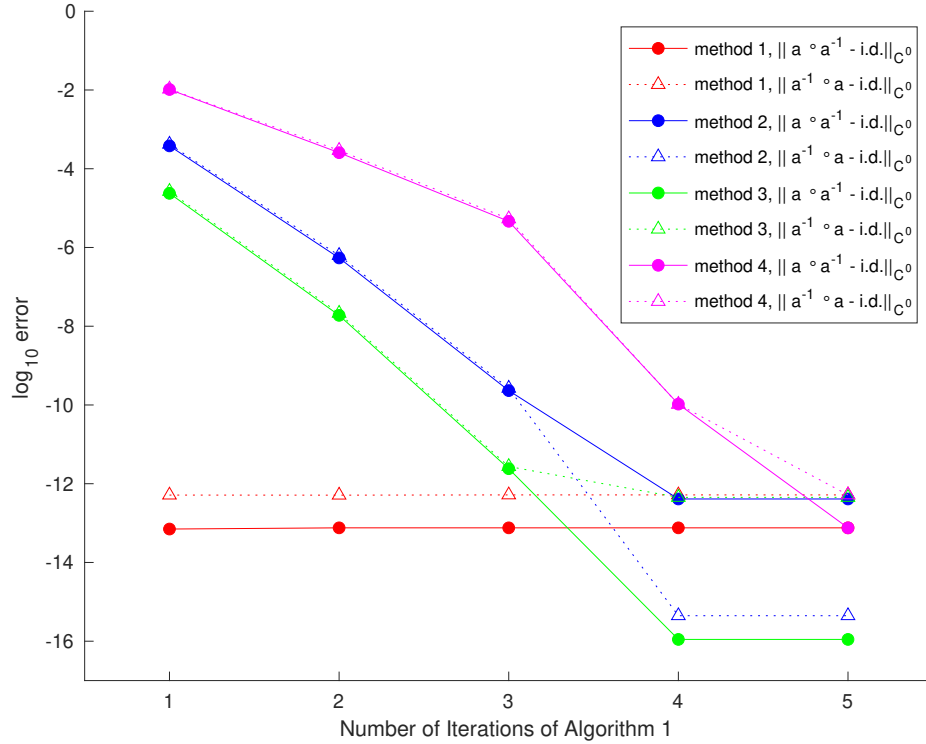
Figure 3.1: Comparison of the methods in Section 3.2.3
Method 1: Graph Reflection; Method 2: "Left" Inverse; Method 3: "Right" Inverse; Method 4: Inverse through Solving Cohomological Equation

Remark 59. *As indicated in Figure 3.1, Method 1 always has a great performance as it is independent to the error of the cohomological equation, while Method 2, 3, 4 slowly get better as the convergence of the algorithm. Despite that Method 4 is designed for optimizing both $a \circ a^{-1}$ and $a^{-1} \circ a$, it turns out this method, in general, does not outperform the other methods. This is also true when k is bigger or even near the breakdown.*

As stated in Section 3.2.3, in practice, we try all of the methods and use the one with the best performance. More specifically, through out the iterations of the algorithm, we always start by method 1, and then replace it with method 2 or 3.

Run Time Analysis

As discussed in Section 3.1, both the time and space complexity for Algorithm 1 are $\mathcal{O}(N \times L)$.

For the same choice of parameters as in Section 3.4.1, the average time for running one iteration of Algorithm 1 can be found in Table 3.3. The code is written in C using the GNU Scientific Library(GSL), and this set of data in Table 3.3 is generated by a Mid 2014 13-inch Macbook Pro with 2.6 GHz Dual-Core Intel Core i5 Processor and 8 GB 1600 MHz DDR3 Memory.

Table 3.3: Average run time (in seconds) for one iteration of Algorithm 1

L	N	Avg Time	L	N	Avg Time
2	1024	0.062405	5	1024	0.132773
	4096	0.284577		4096	0.641300
	16384	1.682731		16384	3.674957
	65536	9.136261		65536	18.175884
	262144	55.717758		262144	122.279080
10	1024	0.392576	20	1024	0.779805
	4096	1.862491		4096	4.708662
	16384	9.303865		16384	21.924758
	65536	44.038226		65536	99.157648
	262144	226.787753		262144	635.228799

Plot of the Invariant Circle and Isochrons

From Remark 54, the isochrons are approximately linear in the neighborhood of the invariant circle when the perturbation is small. To achieve a relatively nontrivial solution, we consider $k = 1.1037$ and $\eta = 0.30532$ where η is tuned manually in order to guarantee a rational rotation number for $a(\theta)$ using the brent algorithm (see Section 3.4.4). As one will see in Section 3.4.1, such cases requires more consideration in terms of globalizing the isochrons. The solution here is computed through continuation method with step size 10^{-3} and error tolerance 10^{-14} in C^0.

(a) $\theta_0 = 0.797$ (b) $\theta_0 = 0.78$

where I_{θ_0} corresponds to the blue isochron.

(c) λ and λ^* (d) rational rotation
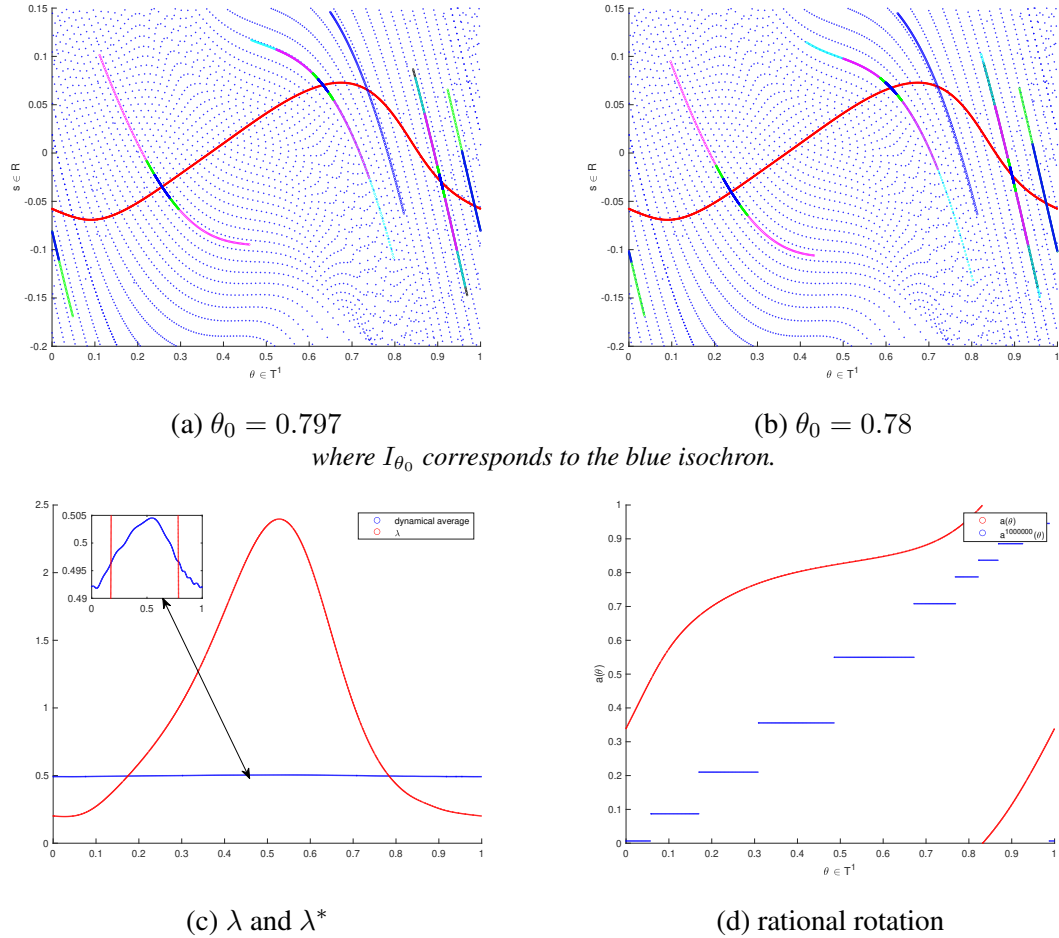
Figure 3.2: Invariant Manifold and Isochrons for the dissipative standard map (3.9)
The last figure is only here to indicate that the rotation number for $a(\theta)$ is indeed rational. Please refer to Section 3.4.3 for further reasons.

In Figure 3.2a, we present the invariant circle (in red) and some of the isochrons (in

the order: blue, green, magenta, cyan, black along the internal dynamics). Notice how the isochrons contracts when the dissipative standard map applies. More specifically, the blue isochron ($I_{\theta=0.797}$) got mapped to the interior of the green isochron, and then the green-blue isochron got mapped to the interior of the magenta one, and then the cyan isochron followed by the black one.

With a different starting point on the grid, Figure 3.2b describes a case when the isochrons does not contract for every single step (indeed, one may find that the cyan isochron is actually expanded rather than contracted). This is perfectly normal: As discussed in Lemma 6, our algorithm allows $\|\lambda\|_0$ to be bigger than one, as long as the dynamical average (discussed in Remark 10) $\lambda^* < 1$ (as shown Figure 3.2c).

Remark 60. *In the case when $\eta = 0.3$, while the other parameters remains unchanged, one can observe that the dynamical average is actually not a constant function, on the contrary to the irrational rotation case (for the same reason as in Remark 9).*

Globalization
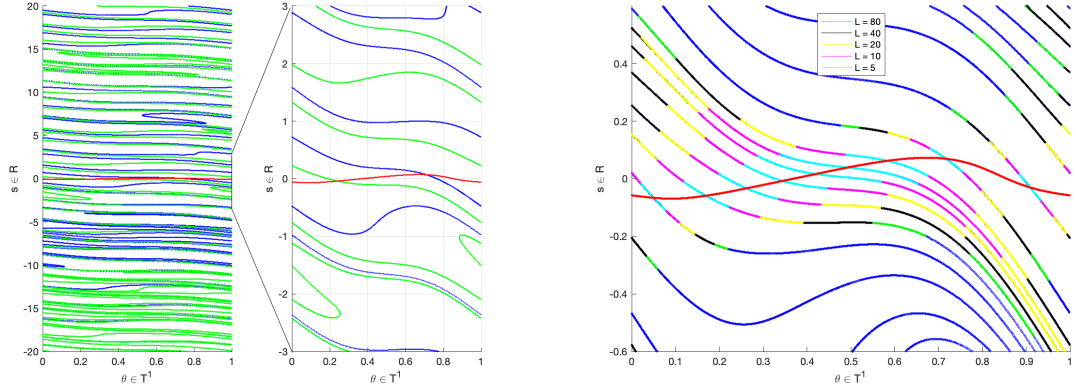
(a) Globalization through Backward Propagation (b) Globalization through Increasing Approximation Order

Figure 3.3: Globalization of the Isochrons

Due to the numerical truncation we made when storing and processing the parameterization $W(\theta, s) = \sum_{i=0}^{L} W^{(i)}(\theta)s^i$, we can only approximate the isochron for a small range
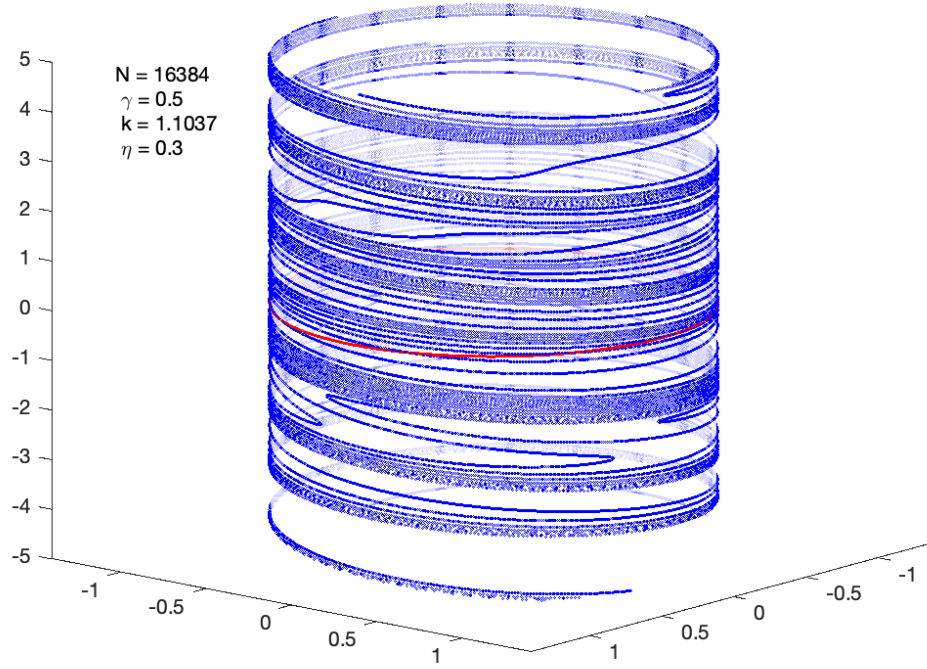
Figure 3.4: Invariant Circles and the Corresponding Isochrons in 3D

of s. In order to increase the region regarding s, we need to perform some globalization techniques.

Since we are dealing with dissipative maps (or any maps with eigenvalues away from 1), the most straightforward way for globalization is to propagate the map backward (or forward).

Remark 61. *When the internal dynamics $a(\theta)$ has rational rotation number (thus so does $a^{-1}(\theta)$), the phase-locking phenomenon occurs. In this case, the globalized isochrons will accumulate near the periodic orbit instead of distributed in the whole $\mathbb{T}^1$.*

To resolve this, one can start an even grid on $\mathbb{T}^1$, first apply $a(\theta)$ n times, compute the stable manifolds, then do the backward propagation of the map n times on the computed stable manifolds to finish the globalization.

Figure 3.3a indicates the globalization of two isochrons (in blue and green) using the above method.

Another way to enlarge the validating region of s is via increasing the truncation order L. To do so, we use the same idea as in the continuation method, while instead of changing a certain parameter, we gradually increase L. More specifically, in order to calculate $W(\theta, s)$ up to the $(l+1)$-th order, we use the previously computed $W(\theta, s)$ with order l as the initial approximation, and perform the same quasi-Newton steps as in Algorithm 1, we also need to extend the range of s accordingly while maintaining the error (see Algorithm 4. Figure 3.3b presents the isochrons up to order 80.

Algorithm 4 Increasing the Order of Parameterization

Input: Solution for Equation (2.4) $W(\theta, s), a(\theta)$ and $\lambda(\theta)$ with truncation order L_{start}
Input: Tolerance for the error of Equation (2.4) in $\mathcal{X}^{\delta_{start}, r}$ norm, with some precribed δ_{start}
Output: Solution $W(\theta, s), a(\theta)$ and $\lambda(\theta)$ to the invariance equation (2.4) with truncation order L_{end}
1: $\lambda(\theta) \leftarrow \lambda(\theta) + \Delta_\lambda(\theta)$,
2: **for** $l \leftarrow L_{start}, L_{end}$ **do**
3:     Find the biggest δ (using any root finding method: Brent, Bisect, etc.) that essures the $\mathcal{X}^{\delta, r}$ norm of the error for Equation (2.4) is within the tolerance.
4: **end for**
5: Return updated $W(\theta, s), a(\theta)$ and $\lambda(\theta)$.

Above all, we present Figure 3.3b with the invariant circle (in red) and 10 isochrons (in blue), and as indicated in Figure 3.2d, the internal dynamics have rational rotation number.

3.4.2 Continuation w.r.t. k

In this subsection, we aim to perform the continuation scheme as in Algorithm 3 for different k, with fixed γ and η. We start at $k = 0$, with Remark 54 as the initial approximation, and keep computing until the quasi-Newton Algorithm 1 stops. The invariant circle and the corresponding stable manifolds are demonstrated in Figure 3.5.
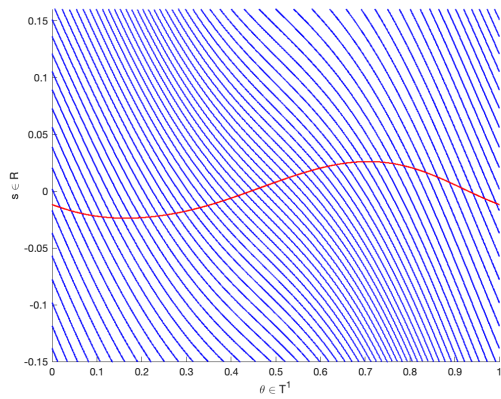
As the perturbation strength k gets larger, the invariant objects become more irregular, thus smaller continuation steps and larger grid size is chosen adaptively according to
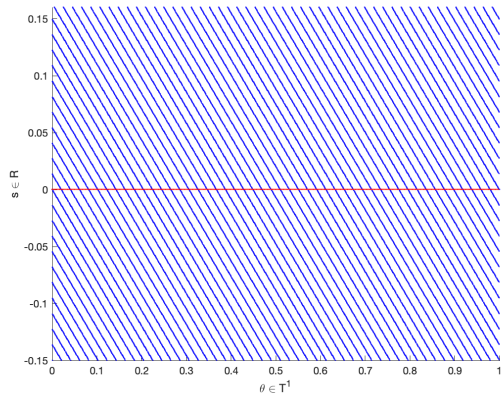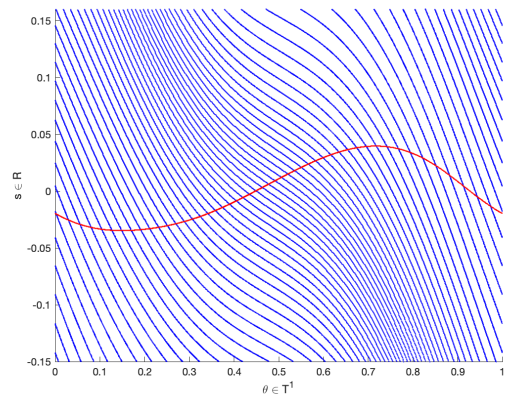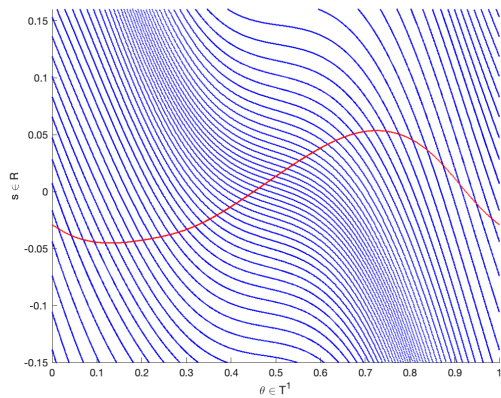
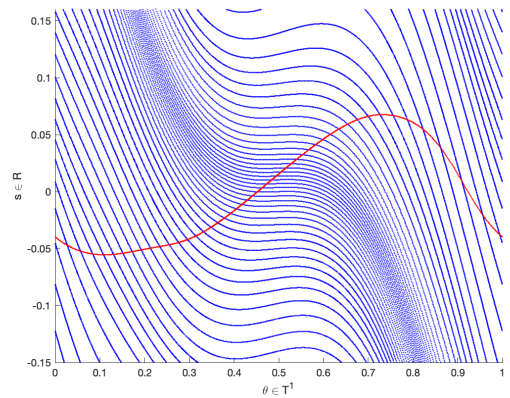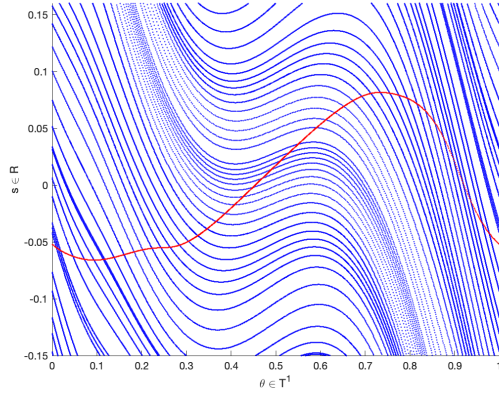(a) $k = 0.0000$
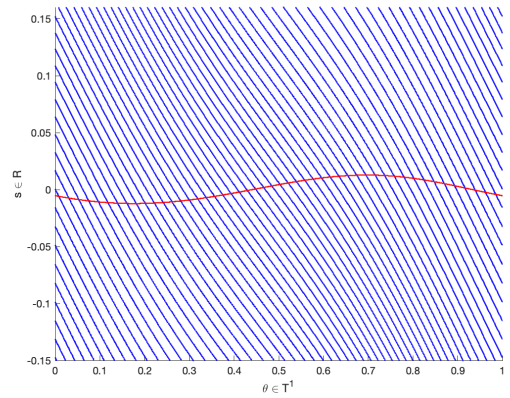
(b) $k = 0.2000$
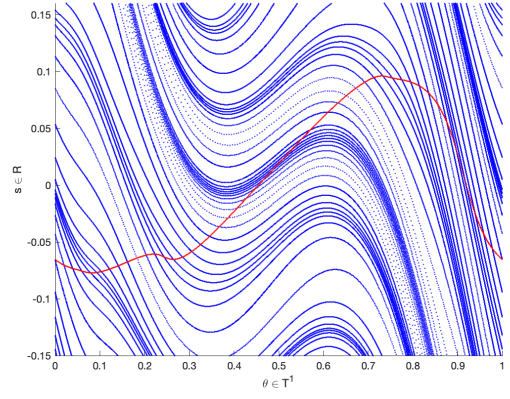
(c) $k = 0.4000$

(d) $k = 0.6000$
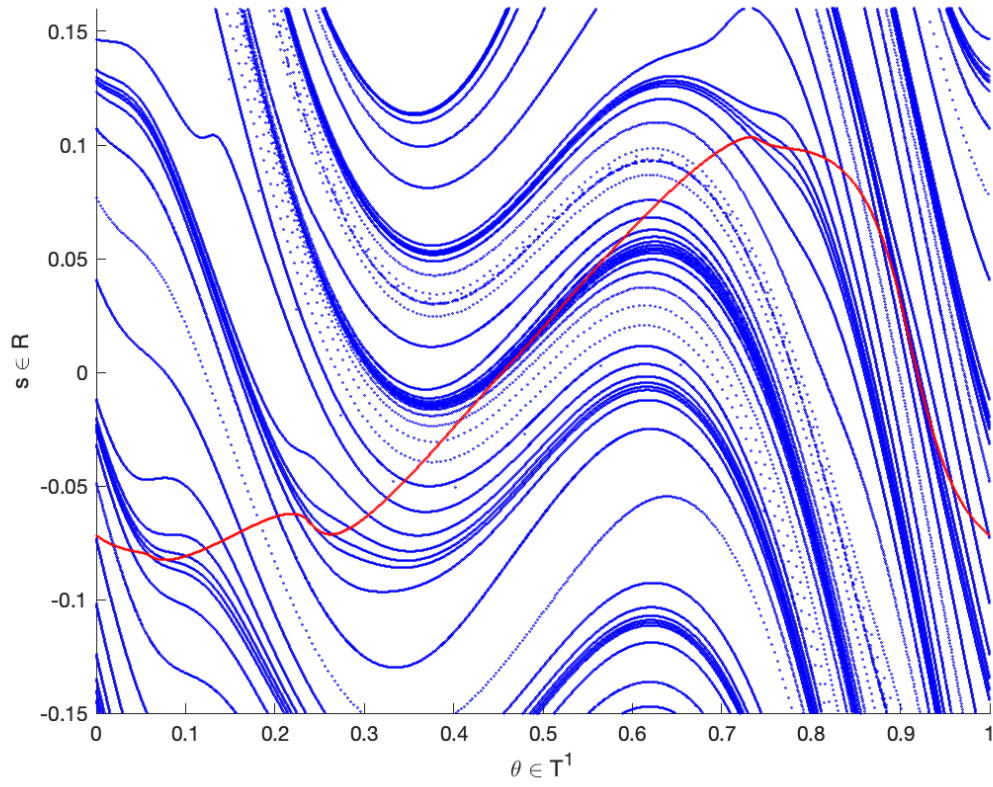
(e) $k = 0.8000$

(f) $k = 1.0000$

(g) $k = 1.2000$

(h) $k = 1.4000$

(i) $k = 1.4927$

Figure 3.5: Invariant Manifold and Isochrons for the dissipative standard map (3.9) where $\gamma = 0.6$, $\eta = 0.4$ and k varies from 0 to 1.4927

Algorithm 3.

As one can observe, while the invariant circle is getting fluctuated as the perturbation increases, the isochrons becomes more irregular, and the minimum angle between the invariant circle and the isochrons is getting closer to 1. We postpone the detailed discussions regarding this "bundle collapsing" scenario to Section 3.5.

3.4.3 Continuation w.r.t. η

An interesting aspect one can observe from fixing k and γ while changing η is to explore how the change of such drift parameter affects the rotation number of the internal dynamics $a_{\eta,\gamma,k}(\theta)$. Figure 3.6 presents such τ_a, the rotation number of $a(\theta)$ as a function of η.
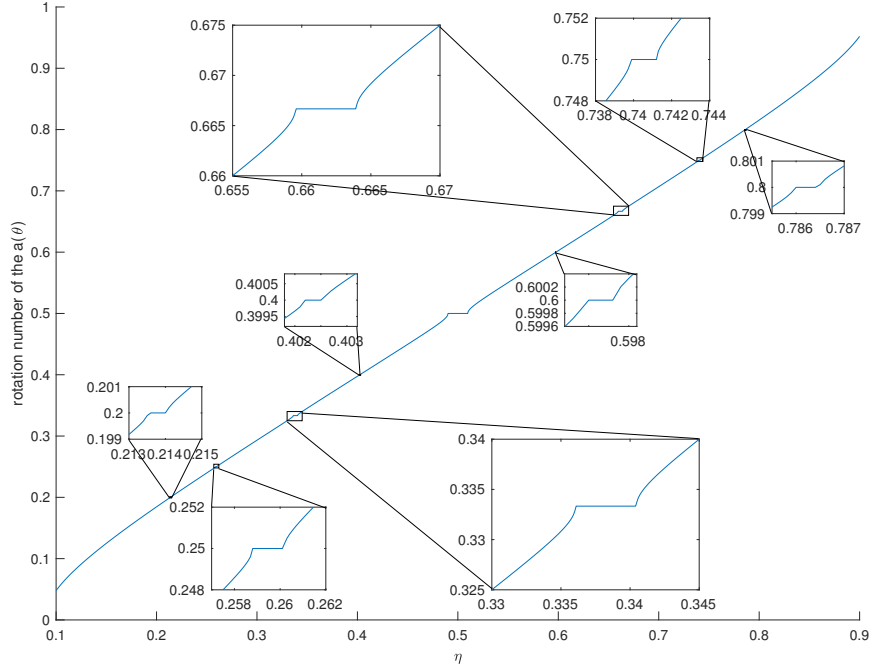
Figure 3.6: Rotation number of $a(\theta)$ w.r.t. η
In this example, $k = 5$ and $\gamma = 0.1$.

The familar Devil's staircase in Figure 3.6 is from the classic results in rotation numbers [74], where $\tau_a(\eta)$ is an monotone function (since our map $f_{\eta,\gamma,k}$ forms a family of orientation-preserving homeomorphisms), and the stairs (where $\tau_a(\eta)$ have constant val-

91

ues) are corresponding to the rational rotation numbers, in which case $a(\theta)$ have periodic orbits.

Computation of the Rotation Number

The most naive way of computing the rotation number is by the definition:

$$\tau_a = \lim_{M \to \infty} \frac{a^{\circ M}(x) - x}{M} \mod 1 \tag{3.10}$$

where $x \in \mathbb{T}$. By computing such rotation number with the help of the Birkhoff average, one can show that such sequence admits $\mathcal{O}(\frac{1}{M})$ convergence rate [75].

In this paper, we follow the approach in [76] and [77]. The main idea is to use the weighted Birkhoff average, which gives the super-polynomial convergence when the rotation number is Diophantine. For every positive integer m, there exists $C_m > 0$ such that

$$\left| \frac{1}{A_M} \sum_{n=0}^{M-1} w(\frac{n}{M})(a^{\circ(n+1)} - a^{\circ n}) \mod 1 - \tau_a \right| \leq C_m M^{-m},$$

where $w(t)$ is the exponential weighting function

$$w(t) = \begin{cases} \exp(\frac{1}{(t(t-1))^p}) & t \in (0, 1), \\ 0 & \text{elsewhere.} \end{cases},$$

$A_N = \sum_{n=0}^{M-1} w(\frac{n}{M})$ and our choice of p is 2.

*Distinguishing Irrational Rotations from Rational Rotations*

By classical results [74], there are three types of orbits for the internal dynamics with rational rotation number τ_a:

1. Periodic orbits with the same period and same order as the rotation $\theta + \tau_a$;

2. Homoclinic orbit;

92

3. Heteroclinic orbit.

Thus by applying the internal dynamics $a(\theta)$ M times on $\mathbb{T}$ (where M is picked to be a large enough number in case the period has large denominator), $a^{\circ M}(\theta), \theta \in \mathbb{T}$ is either a piecewise constant function (as shown in Figure 3.7) or conjugate to a rational rotation (where the latter case rarely happens).

On the other hand, if τ_a is irrational, by the Poincaré Classification theorem, $a^{\circ M}(\theta)$ remains to be a homeomorphisms (as shown in Figure 3.8).
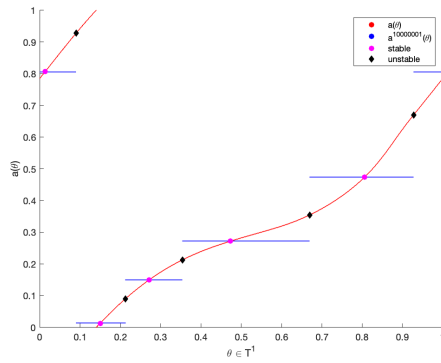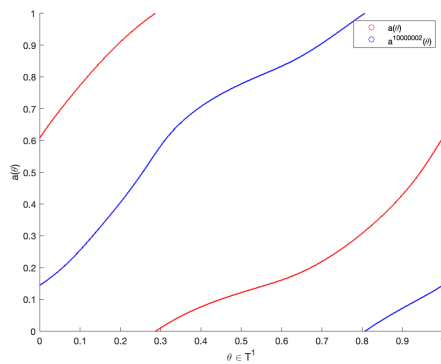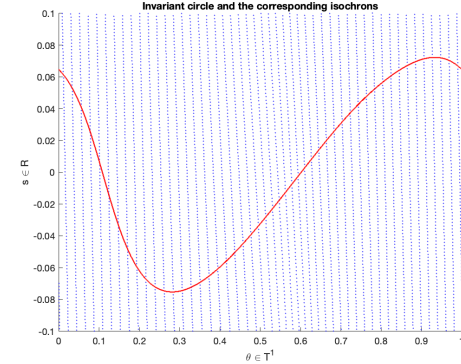

(a) Internal Dynamics
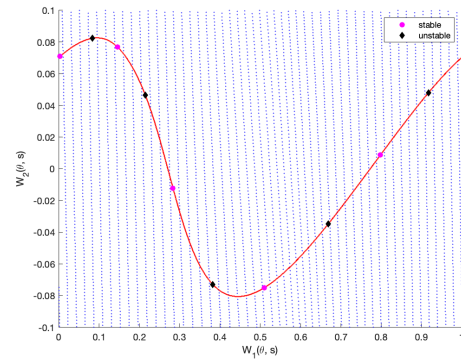
(b) Invariant Circle and Isochrons

Figure 3.7: $a^{\circ 1e+7}$ when a has rational rotation number but doesn't conjugate to a rotation

(a) Internal Dynamics

(b) Invariant Circle and Isochrons

Figure 3.8: $a^{\circ 1e+7}$ when a conjugates an irrational rotation

93

3.4.4 Continuation with Prescribed Rotation Number

Many applications (quasiperiodic attractors, oscillators [13, 61]), require the rotation frequency to be a fixed diophantine number, for example, the golden mean. In such cases, as stated in Remark 9, $a(\theta) = \theta + \omega$, $\lambda(\theta) = \lambda$, and the phace locked phenomenon does not appear. Extensive studies have been made regarding this in [13], [61].

In order to cope with the fixed rotation number, we can vary η using any root finding algorithms (for example, the Brent Method). The plot for η as k varies is in Figure 3.9.
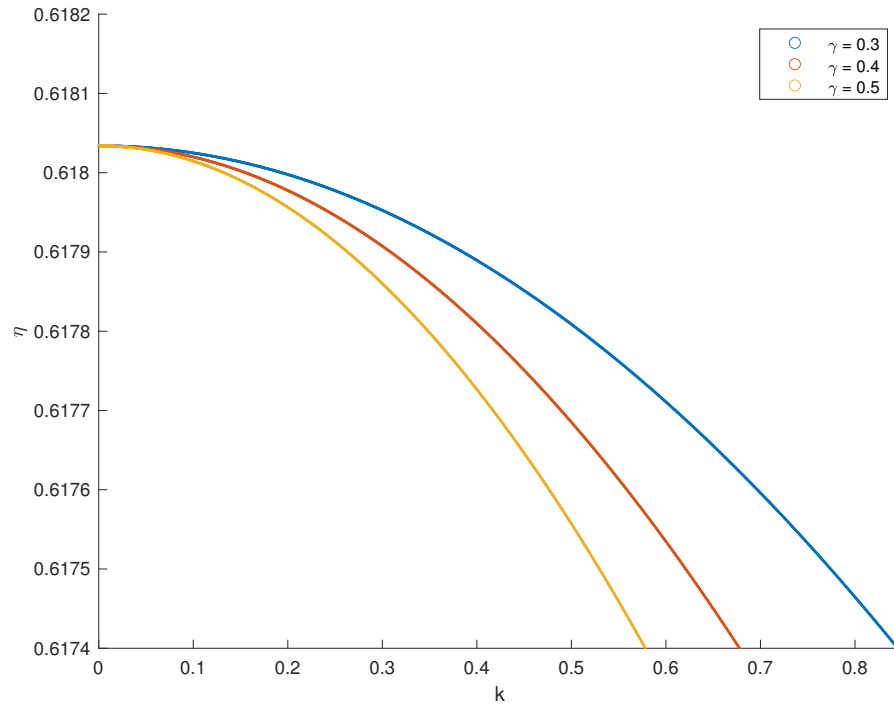

Figure 3.9: η vs k when the rotation number is prescribed by the golden mean, with fixed γ

In principle, with prescribed rotation number, one can follow Algorithm 3, increase the range for k in Figure 3.9 and make explorations when k is close to the breakdown.

3.5 Some Explorations on the Breakdown

Given an algorithm it is quite natural to study its limits of applicability. Of course, the mathematical phenomenon of interest may itself breakdown. The two breakdowns can, of course, happen at different places.

In this section, we will present some remarks about the mathematical formulation of breakdown of invariant circles. We will also present some numerical results exploring the phenomena. We will present two mathematical scenarios of breakdown which certainly happen in the two parameter family (3.9) (Think of the two parameter family as a 1-parameter family of 1-parameter families). We will also present numerical results for a family inside (3.9) which does not satisfy either of the mathematical scenarios presented.

We have tried to separate the results that are rigorously established from the ones that are results of numerical experiments or theoretical studies (renormalization group predictions). Nevertheless, in this area, the boundary between rigorous and numerical is moving.

3.5.1   Some Remarks on the Mathematical Definition of Breakdown

The breakdown of invariant circles is mathematically a very subtle phenomenon. Indeed, there are several possible definitions.

The most widespread theory of persistence of invariant circles is the theory of Normally Hyperbolic Manifolds [1]. This theory ensures that, a C^1 invariant manifold satisfying the assumptions of normal hyperbolicity, persists a C^1 manifold. Conversely, [24] showed that if a map has invariant C^1 manifold that does not satisfy the normally hyperbolic assumptions, one can make C^1 arbitrarily small perturbations that do not possess a C^1 invariant manifold near the original one.

Note that the beautiful result of [24] leaves open several questions that are worth addressing. Note that some of them are related.

Persistence in a Different Sense than C^1 Manifolds, Persistence as Lower Regularity Sets.

As we will see, planar invariant circles that fail to satisfy the NHIM conditions may persist as $C^{0.99999}$ circles. Indeed, $C^{0.999999}$ circles are analytic except at a finite number of points.

There are scenarios, discussed in detail in Section 3.5.1, in which a family of maps (chosen in an open sets of families) has the following properties: f_λ possess invariant circles of regularity $C^\alpha(\lambda)$. where α is a continuous decreasing function $\alpha(\lambda^*) = 0$. Indeed, for $\lambda \in (\lambda^*, \lambda^* + \epsilon)$, there is an invariant set that, even if not the graph of a continuous function is an indecomposable continuum. There are even more general invariant sets.

Therefore, the regularity of the invariant set decreases in a regular way and indeed, the set persists naturally from a curve to a continuum. Each choice of a regularity leads to a breakdown value for this regularity. Mathematically, C^1 is a natural breakdown for differentiable manifold theory (natural regularity for tangent spaces, and differentiable, C^0 is a natural breakdown for topological methods (natural regularity to define rotation number). From the numerical point of view C^k are natural for the use of splines of order k.

Weaker Notions of Normal Hyperbolicity. Some Recent Mathematical Results

The standard notion of normal hyperbolicity includes several ingredients. See [1, 56, 57, 78, 2]

1. Existence of bundles spanning the normal bundle in which there is contraction/expansion

2. The exponential rates of contraction expansion in the bundles dominate the rates of contraction and expansion in the tangent space.

3. There are constants affecting the exponential rates.

4. The contractive and expansive directions are at a fixed angle.

In many theoretical treatments, items 3) 4) above are ignored since the "adapted norm" is used. In this adapted norm. the bundles are orthogonal and the constants are 1. As

96

we will see, there are cases when the breakdown happens when 3) or 4) fail. Note that when 3), 4) fail, the adapted norm becomes singular with respect to the usual norm, so that the stability domains established using the adapted norm become very small in the standard metric. Eventually, the stability results may stop applying for finite values of the parameters. In Section 3.5.1, we will present results in a scenario of collapse of tori in which hyperbolicity is lost, even if the rates remain separated, but the angle between the tangent and the stable bundles goes to zero. See [48] for examples in the family of maps (3.9).

The loss of hyperbolicity because the rates of contraction/expansion in the normal directions get close to 1 but the bundles remain well separated, has been studied in [79, 80, 81].

The remarkable recent paper [82] uses topological methods to show persistence of sets using topological methods. This notion does not require rates conditions. On the other hand, the objects that persist could be much more complicated than a continuous image of a manifold. They are just ensured to be continua. The papers [83, 84] establish lower bounds on the Czeck cohomology of these sets.

The papers [85, 86] establish a statistical theory of persistence of attractors. Even if the attractors no not need to persist under small perturbations, any smooth family of maps containing them has to admit attractors for a large set of parameter values. This mathematical theory has been written in a very constructive form

The earlier paper [87] contains a more topological version of the persistecne of complicated behaviour.

Breakdown of Phase-locked invariant Circles

We consider a family f_λ of analytic maps of the plane or of $\mathbb{T} \times \mathbb{R}$.

We assume that for all the values of the parameter, the family has an invariant circle which contains two periodic orbits. One orbit A will be hyperbolic, the other orbit B will

start being attractive but, at some value λ_* acquires an eigenvalue 1, We will assume that for all parameter values, the unstable manifold of the hyperbolic fixed point A intersects transversely the strong stable manifold of the point B.

Note that the above situation is stable under small C^1 perturbations.

One concrete example of this situation is:

$$f_\lambda(\theta, r) = \begin{pmatrix} \theta + 0.01\sin(2\pi\theta)(\bmod 1) \\ \frac{1}{2}r + \exp(-100\sin(\pi\theta - \frac{\pi}{2})^2)(\lambda r) \end{pmatrix},$$

Note that away from $\theta = \frac{1}{2}$, the exponential factor is very small.

In this simple example, the point $A = (0, 0)$ is hyperbolic with eigenvalues $\approx (1 + 0.02 \cdot \pi, \frac{1}{2})$ and the point $B = (\frac{1}{2}, 0)$ is attractive with eigenvalues $(1 - 0.02 \cdot \pi, \frac{1}{2} + \lambda)$.

We make the following observations:

Any invariant continuous curve that goes through the hyperbolic fixed point has to be either the stable or the unstable manifold and, therefore be analytic.

If we iterate a segment of the unstable manifold of A we will eventually obtain a segment in the small neighborhood of B, The further iterations of the segment will converge towards the stable point.

Therefore, the unstable manifold of the orbit A, together with with the orbit B give an invariant circle.

This invariant circle will be analytic in any segment that does not contain a point in the orbit B. If the eigenvalues of the derivative at the orbit B are $0 < \alpha < \beta$, the invariant circle will be C^r for any $r < \log(\alpha)/log(\beta) \equiv r_*$.

It is also shown in [15] that the strong stable manifold of the point B is analytic and that any C^r curve with $r > r_*)$ has to be the strong stable manifold.

It then follows that, if the unstable manifold of A, does not agree with the strong stable manifold of B, then the invariant circle will be $C^{r_*-\epsilon}$ but not $C^{r_*+\epsilon}$.

Remark 62. *A motivating model to keep in mind is the case when B is a fixed point and the map is linear in a neighborhood of B.*

In such a case, the strong stable manifold is x axis. If we have that W_A^u is given in a segment $(\beta x_0, x_0)$ as the graph of a function ϕ, we obtain that the unstable manifold is given in $\beta^{n+1} x_0, \beta^n x_0$ as the graph of $\phi^n \equiv \alpha^n \phi \circ \beta^{-n}$.

We see that if ϕ is not identically zero,

$$D^j \phi^n \approx \alpha^n \beta^{-jn} \tag{3.11}$$

and similar assymptotics happen for the Hölder seminorms in place of derivatives.

The assumption that the map is linear, can be arranged by the Sternberg linearization theorem whenever $\alpha^n \neq \beta$. We also point out that the conclusions (3.11) *happen also for nonlinear mappings.*

Note that the disagreement of W_A^u, W_B^{ss} is an open phenomenon in a smooth topology in the space of mappings[1]. Therefore, the regularity property of the invariant circles indicated above holds in open sets of maps.

Furthermore, we argue that the disagreement indeed is the *generic* property. (We leave the exact formulation of this property to the reader interested in generic properties). We just note that local segments of the unstable manifold of A and the strong stable manifold of B can be obtained by analyzing the map in very small neighborhoods of A, B. Then, we consider N large enough so that $f^N(W_{A,\text{loc}}^u)$ lies in the neighborhood V of B where W_B^{ss} is defined.

We assume without loss of generality that $f^{N-1}(U) \cap B = \emptyset$. We note that if we do a perturbation of f supported in $f^{N-1}(U)$, it affects $W_A^u|_V$, but not $W^s s_B$. Hence, if for some map the two invariant manifolds agree, one can break the agreement by an arbitrarily

---

[1]Consider the mapping that to the mapping f associates the unstable manifold of A and the strong stable manifold of B. This mapping is continuous – or even differentiable – if we give the spaces of manifolds some C^ℓ topologies and the space of mappings a C^s topology. See [17].

small perturbation.

Lemma 13. For $\lambda, \mu < 1$. Let $f(x, y) = (\lambda x, \mu y)$ in a neighborhood of $(0, 0)$. Consider γ a curve $(p, f(p))$. Assume that $\gamma \cup f(\gamma)$ is C^∞, consider $\gamma^* = \cup_{n \geq 0} f^n \gamma$, then $\gamma^* \subset C^{\ln \lambda / \ln \mu}$. Furthermore, consider the set of C^∞ curves, satisfying, there is a residual set S such that if $\gamma \in S$, γ^* is not $C^{\ln \lambda / \ln \mu + \delta}$ at the origin.

The bundle merging scenario

We consider families of mappings for which the rotation number on the invariant circle is constant and a Diophantine number (as shown in [30]. this can be arranged by adjusting the drift parameter η in (3.9).

The papers [54, 31] discovered that the breakdown of circles can happen when the rates remain safely away but the angles become arbitrarily small. This scenario has been studied numerically in the dissipative standard map in [48].

In [54] there is one specific example where this is shown rigorously to happen. (One uses Herman's subharminicity trick to bound the Lyapunov exponents in the whole family and shows that they are safely in the hyperbolicity region. On the other hand, by studying the maps in different ends of the family, we find that the stable and unstable manifolds change topology. Therefore, hyperbolicity has to be lost, but it cannot be because of degeneration of the rates.

Similar phenomena, have been discovered in other mappings [88, 89]. Some rigorous results establishing the scenario in simpler mappings are in [47, 90].

In the dissipative standard map, since the determinant is constant γ, if we have an invariant circle, C^1 conjugated to a rotation, the stable exponent has to be γ. Using Fenichel theory, the circle has to be very smooth and by a bootstrap argument in [30], it has to be analytic.

The only way that the torus can loose normal hyperbolicity is if the stable bundle becomes tangent to the tangent mapping. The paper [48] presents an empirical study for

the dissipative standard map and discovers numerically several intriguing scaling relations, which require more mathematical study.

3.5.2 Some Numerical Explorations

In the previous sections, we have presented some different scenarios leading to breakdown. Note that both of them assume that the rotation number of the invariant circle is maintained constant till the breakdown (for the phase locked maps, this can happen on open sets in families, but for Diophantine irrational rotation numbers, it is a codimension 1 phenomenon.

In this section, we study numerically the breakdown of a family inside (3.9), chosen arbitrarily. In this family, the rotation number changes and goes from rational to irrational. So, the scenarios Sections 3.5.1 and 3.5.1 will alternate and the phenomena observed will be a combination of the two scenarios.

This section can be considered as testing ground for the numerical algorithms and to raise the need of a more detailed mathematical theory.

For invariant attractors, the Lyapunov multiplier is always smaller than 1, thus in the dissipative standard map (3.9), we are expecting either the "bundle merging" scenario or the "rate meeting" scenario (when $\lambda^s = \mu^s$ as in Section 3.5.1).

Following a random continuation path as computed in Section 3.4.2, where the rotation number is not prescribed, we have Figure 3.10.

To quantify the separation between bundles, we measure the minimum angle between the invariant circle and the corresponding stable manifolds. Figure 3.10a indicates the angles for the k values as in Figure 3.5. Figure 3.10b plots the minimum angle with respect to k, and Figure 3.10c plots the corresponding $\theta \in \mathbb{T}$ such that the minimum angle is reached.

As indicated in Figure 3.10b, the slope of the blue curve (minimum angle) is approaching negative infinity (it is not obvious in the plot due to the different scales in the horizontal and vertical axis). As a matter of fact, the later portion of the minimum angle curve can be
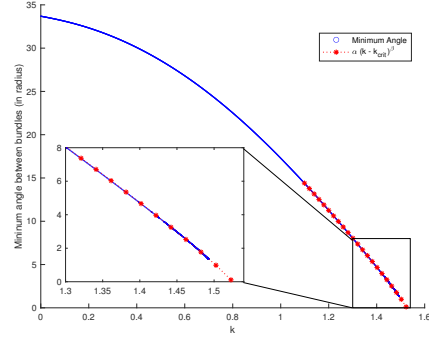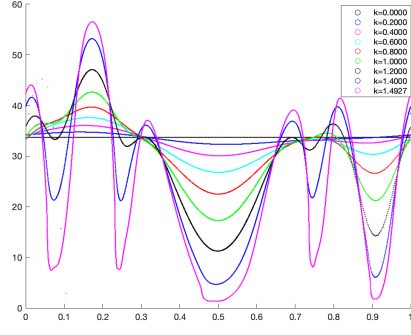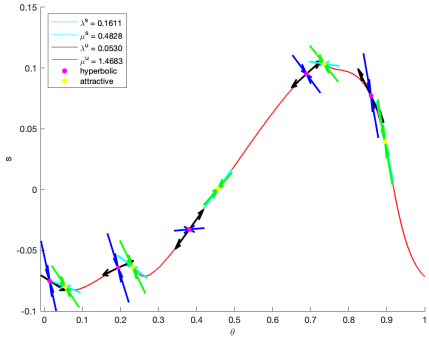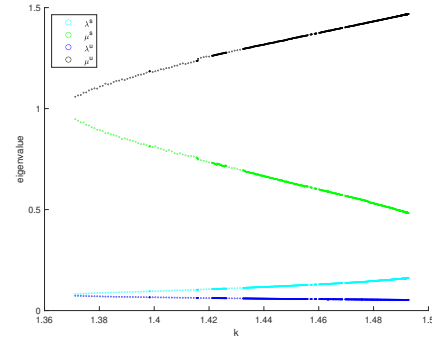
(a) Angles between the tangent and normal bundles for Figure 3.5

(b) The minimum angle between the tangent and normal bundles
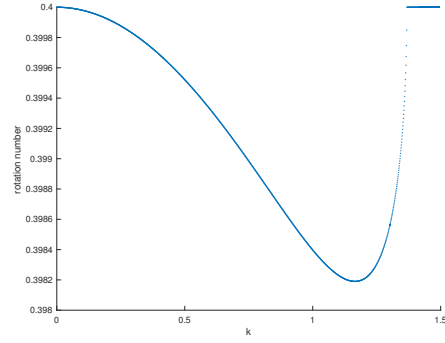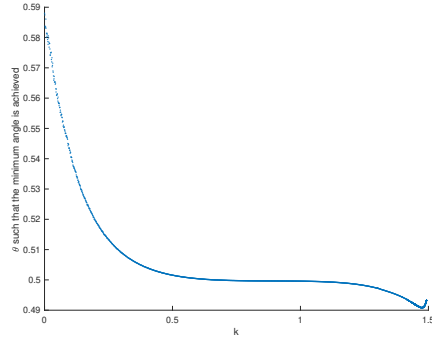
(c) The position such that the minimum angle in Figure 3.10b is achieved

(d) Rotation number w.r.t. the perturbation parameter k

(e) Eigenvalues for $f^{\circ 5}$ of the attractive and hyperbolic periodic orbit when $k > 1.3712$

(f) Stability of the attractive and hyperbolic periodic orbit when $k = 1.49271$

Figure 3.10: "Bundle merging" Explorations

fitted by the following asymptotic expression using nonlinear regression:

$$\alpha(k - k_{crit})^\beta, \qquad (3.12)$$

where α is the scaling constant, β is the scaling exponent, and k_{crit} is the breakdown threshold. In fact, $\alpha = 45.8879$, $\beta = 0.9085$ and $k_{crit} = 1.5247$ gives us an estimation for the critical value such that the tangent and normal bundle collapse.

Remark 63. *The asymptotic expression* (3.12) *appears in [23] when it comes to estimating the blow-up of the Sobolev norm of the parameterization of the invariant circle. Empirically, Figure 3.10b shows that the minimum angle also approaches 0 following such power laws with universal exponents.*

Since the rotation number in this example is not prescribed, it may change along the continuation. In our example, the change of the rotation number as k increases is presented in Figure 3.10d.

Remark 64. *Note that the continuous family of orientation-preserving circle homeomorphism $f_{\eta,\gamma,k}$ is not monotone, in the sense that the lift of $f_{\eta,\gamma,k}$ can not be compared universally on $\mathbb{T}$ [74], thus the rotation number in Figure 3.10d is not expected to be monotone.*

Remark 65. *We are still expecting "stairs" (intervals for k in which case the rotation number is rational). In Figure 3.10d, the only obvious one is when the rotation number is $\frac{2}{5}$, while the rest of the "stairs" (for example, corresponding to rotation number 0.3999) is too small to be plotted.*

As indicated in Figure 3.10d, the rotation number is 0.4 when $k > 1.3712$. In this region, as discussed in Section 3.5.1, there are two periodic orbit, the attractive one with eigenvalues λ^s, μ^s, and the hyperbolic one with eigenvalues λ^u, μ^u. In Figure 3.10e, one can see how the eigenvalue changes as k increases. The seperation between λ^s and μ^s indicates that the contraction rate in the tangential and normal direction will never meet in this example. Figure 3.10f present the eigenvectors on each periodic point.

3.6 NUMEIRICAL ALGORITHM IN 3-D CASES

In this section, we further generalize our algorithm to three dimensional maps which possesses one-dimensional invariant circles.

Following the same idea introduced previously, we first briefly discuss the derivation of the algorithm, and then implement such algorithm to the 3-dimensional Fattened Arnold Family (3D-FAF) maps.

3.6.1   Basic Derivation of the Algorithm

Given a map $f : \mathbb{T}^1 \times \mathbb{R}^2 \to \mathbb{T}^1 \times \mathbb{R}^2$ that induces an one-dimensional invariant circle, the goal in this subsection is to derive an algorithm computing both the invariant circle and the corresponding isochrons.

Because of the increase of dimension on the isochrons, the invariance equation (2.4) now becomes

$$f \circ W(\theta, s_1, s_2) = W(a(\theta), \lambda_1(\theta)s, \lambda_2(\theta)s), \tag{3.13}$$

where $W : \mathbb{T}^1 \times \mathbb{R}^2 \to \mathbb{T}^1 \times \mathbb{R}^2$, $a : \mathbb{T}^1 \to \mathbb{T}^1$, $\lambda_1 : \mathbb{T}^1 \to \mathbb{T}^1$ and $\lambda_2 : \mathbb{T}^1 \to \mathbb{T}^1$ are the unknowns. Again, a is the internal dynamics on the invariant circle, and for any given $\theta_0 \in \mathbb{T}$, $W(\theta_0, s_1, s_2)$ parameterizes the isochron, $\lambda_1(\theta_0)s_1, \lambda_2(\theta_0)s_2$ are the dynamics on the isochron along the eigen-directions.

Following the same procedure as in Section 3.1, we are looking for $\Delta_a(\theta)$, $\Delta_{\lambda_1}(\theta)$, $\Delta_{\lambda_2}(\theta)$ and $\Delta_W(\theta, s_1, s_2) = DW(\theta, s_1, s_2)\Gamma(\theta, s_1, s_2)$ satisfing the following three equations:

$$Da(\theta)\Gamma_1(\theta, s_1, s_2) - \Delta_a(\theta) - \Gamma_1(a(\theta), \lambda_1(\theta)s_1, \lambda_2(\theta)s_2) = M_1(\theta, s_1, s_2), \tag{3.14}$$

$$\lambda_1(\theta)\Gamma_2(\theta, s_1, s_2) - \Delta_{\lambda_1}(\theta)s_1 - \Gamma_2(a(\theta), \lambda_1(\theta)s_1, \lambda_2(\theta)s_2) = M_2(\theta, s_1, s_2), \quad (3.15)$$

$$\lambda_2(\theta)\Gamma_3(\theta, s_1, s_2) - \Delta_{\lambda_2}(\theta)s_2 - \Gamma_3(a(\theta), \lambda_1(\theta)s_1, \lambda_2(\theta)s_2) = M_3(\theta, s_1, s_2), \quad (3.16)$$

where

$$M(\theta, s_1, s_2) = \widetilde{e}(\theta, s_1, s_2) - \begin{pmatrix} 0 \\ D\lambda_1(\theta)s_1\Gamma_1(\theta, s_1, s_2) \\ D\lambda_2(\theta)s_1\Gamma_1(\theta, s_1, s_2) \end{pmatrix},$$

and

$$\widetilde{e}(\theta, s_1, s_2) = -[DW(a(\theta), \lambda_1(\theta)s_1, \lambda_2(\theta)s_2)]^{-1}e(\theta, s_1, s_2).$$

By discretizing function g with three variables (θ, s_1, s_2) as

$$g(\theta, s_1, s_2) = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} g^{(x,y)}(\theta)s_1^x s_2^y, \quad (3.17)$$

we again can further discritize Equation (3.14), (3.15) and (3.16) order by order as follows:

- Equation (3.14):

  - Order $(0,0)$:

  $$Da(\theta)\Gamma_1^{(0,0)}(\theta) - \Delta a(\theta) - \Gamma_1^{(0,0)}(a(\theta)) = M_1^{(0,0)}(\theta), \quad (3.18)$$

  - Order (x, y):

  $$\Gamma_1^{(x,y)}(\theta) = \frac{\lambda_1^x(\theta)\lambda_2^y(\theta)}{Da(\theta)}\Gamma_1^{(x,y)}(a(\theta)) + \frac{1}{Da(\theta)}M_1^{(x,y)}(\theta) \quad (3.19)$$

105

- Equation (3.15):

  - Order $(0,0)$:
  $$\lambda_1(\theta)\Gamma_2^{(0,0)}(\theta) - \Gamma_2^{(0,0)}(a(\theta)) = M_2^{(0,0)}(\theta), \tag{3.20}$$

  - Order $(1,0)$:
  $$\lambda_1(\theta)\Gamma_2^{(1,0)}(\theta) - \Delta_{\lambda_1}(\theta) - \Gamma_2^{(1,0)}(a(\theta))\lambda_1(\theta) = M_2^{(1,0)}(\theta), \tag{3.21}$$

  - Order (x,y):
  $$\Gamma_2^{(x,y)}(\theta) = \lambda_1^{x-1}(\theta)\lambda_2^y(\theta)\Gamma_2^{(x,y)}(a(\theta)) + \frac{M_2^{(x,y)}(\theta)}{\lambda_1(\theta)}, \tag{3.22}$$

- Equation (3.16):

  - Order $(0,0)$:
  $$\lambda_2(\theta)\Gamma_3^{(0,0)}(\theta) - \Gamma_3^{(0,0)}(a(\theta)) = M_3^{(0,0)}(\theta), \tag{3.23}$$

  - Order $(0,1)$:
  $$\lambda_2(\theta)\Gamma_3^{(1,0)}(\theta) - \Delta_{\lambda_2}(\theta) - \Gamma_3^{(1,0)}(a(\theta))\lambda_2(\theta) = M_3^{(1,0)}(\theta), \tag{3.24}$$

  - Order (x,y):
  $$\Gamma_3^{(x,y)}(\theta) = \lambda_1^x(\theta)\lambda_2^{y-1}(\theta)\Gamma_3^{(x,y)}(a(\theta)) + \frac{M_3^{(x,y)}(\theta)}{\lambda_2(\theta)}. \tag{3.25}$$

Equations (3.18), (3.21) and (3.24) are underdetermined equations that can be solved by letting

$$\Gamma_1^{(0,0)}(\theta) = \Gamma_2^{(1,0)}(\theta) = \Gamma_3^{(0,1)}(\theta) = 0,$$

and thus $\Delta_a(\theta) = -M^{(0,0)}(\theta)$, $\Delta_{\lambda_1}(\theta) = -M^{(1,0)}(\theta)$ and $\Delta_{\lambda_2}(\theta) = -M^{(0,1)}(\theta)$.

Equations (3.19), (3.20), (3.22), (3.23), (3.25) can be written in the format

$$\phi(\theta) = l(\theta)\phi(a(\theta)) + \eta(\theta),$$

if the dynamical average $l^* < 1$ (Remark 10), or be rewritten in the format

$$\phi(\theta) = \frac{1}{l(a^{-1})(\theta)}\phi(a^{-1}(\theta)) - \frac{\eta(a^{-1}(\theta))}{l(a^{-1}(\theta))},$$

if $\left(\frac{1}{l(a^{-1})}\right)^* < 1$, and such equations can then be solved through contraction as in (2.24).

Remark 66. *In the 3-dim case, our quasi-Newton method does not work if $\lambda_1(\theta)$ and $\lambda_2(\theta)$ are resonant, in the sense $\frac{\ln(\lambda_1(\theta))}{\ln(\lambda_2(\theta))} \in \mathbb{N}$. More spercifically, in this case, the fibered version of the Sternberg theorem fails. According to Chen's Theorem, for $|\lambda_2|^q < |\lambda_1|$, we are expecting a polynomial $p(\theta, s)$ with degree $\leq q$ such that the invariance equation becomes*

$$f(W(\theta, s)) - W(a(\theta), p(\theta, s)) = 0.$$

The same discussions remain valid for higher dimensional cases. We hope to come back to this problem.

3.6.2   Numerical Exploration: 3D-Fattened Arnold Family

Inspired by [19], in this subsection, we implement the algorithm discussed in Section 3.6.1 on a 3-dimensional Fattened Arnold Family (3D-FAF) $f_{\alpha,\epsilon} : \mathbb{T}^1 \times \mathbb{R}^2 \to \mathbb{T}^1 \times \mathbb{R}^2$:

$$f_{\alpha,\epsilon}(x, y, z) = \begin{pmatrix} x + \alpha + \frac{\epsilon}{2\pi}(\sin(2\pi x) + y + \frac{z}{2}) \\ \beta(\sin(2\pi x) + y) \\ \gamma(\sin(2\pi x) + y + z) \end{pmatrix}, \tag{3.26}$$

where ϵ is the perturbation parameter, α is the drift parameter and β, γ are the eigenvalues. We have implemented the algorithm for the stable (i.e. $\beta < 1$, $\gamma < 1$) case, unstable

$(\beta > 1, \gamma > 1)$ case and the saddle (i.e. $\beta < 1 < \gamma$) case.

The Unperturbed Case

In order to apply the continuation method, we start with the unperturbed case where $\epsilon = 0$. As discussed in [19], in such case, the solution to the invariance equation (3.13) is:

$$W_1(\theta, s_1, s_2) = \theta,$$

$$W_2(\theta, s_1, s_2) = -S(\alpha, \beta)\cos(2\pi\theta) + (C(\alpha, \beta) - 1)\sin(2\pi\theta) + s_1,$$

$$W_3(\theta, s_1, s_2) = \frac{\gamma}{\beta}\left(S(\alpha, \beta)(C(-\alpha, \gamma^{-1}) - 1) + (C(\alpha, \beta) - 1)S(-\alpha, \gamma^{-1})\right)\cos(2\pi\theta)$$

$$\frac{\gamma}{\beta}\left((C(\alpha, \beta) - 1)(C(-\alpha, \gamma^{-1}) - 1) - S(\alpha, \beta)S(-\alpha, \gamma^{-1})\right)\sin(2\pi\theta)$$

$$+ \frac{\gamma}{\beta - \gamma}s_1 + s_2,$$

$$a(\theta) = \theta + \alpha, \ \lambda_1(\theta) = \beta, \ \lambda_2(\theta) = \gamma,$$

where

$$S(x, y) = \frac{y\sin(2\pi x)}{1 - 2y\cos(2\pi x) + y^2}, \quad C(x, y) = \frac{1 - y\cos(2\pi x)}{1 - 2y\cos(2\pi x) + y^2}.$$

The Perturbed Case

Following the same continuation method as in Section 3.3, the invariant circle and the corresponding isochrons are computed for all the stable, unstable and saddle choice of the parameters. For the same color choice as in Figure 3.2a, we still start with the blue isochron, which is mapped to the green isochron, followed by magenta, cyan and black, correspondingly (see Figure 3.11, 3.12 and 3.13).
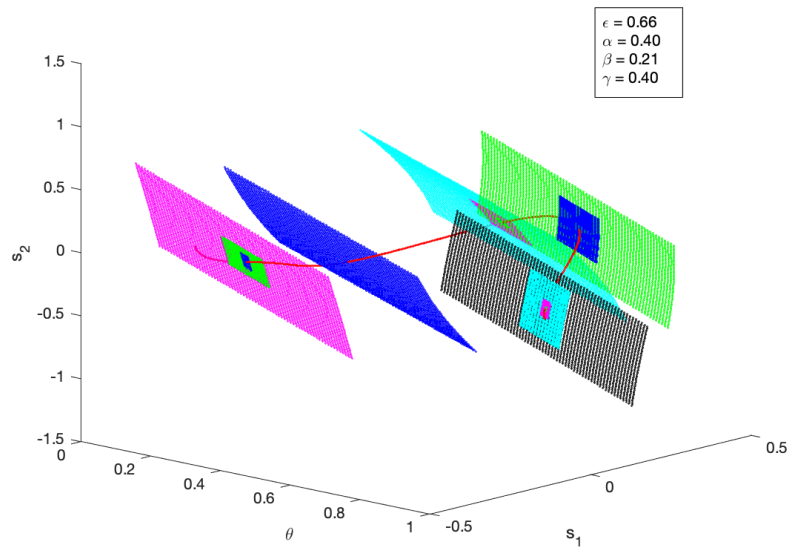
Figure 3.11: 3D-FAF map: Invariant Circle and Stable Manifolds
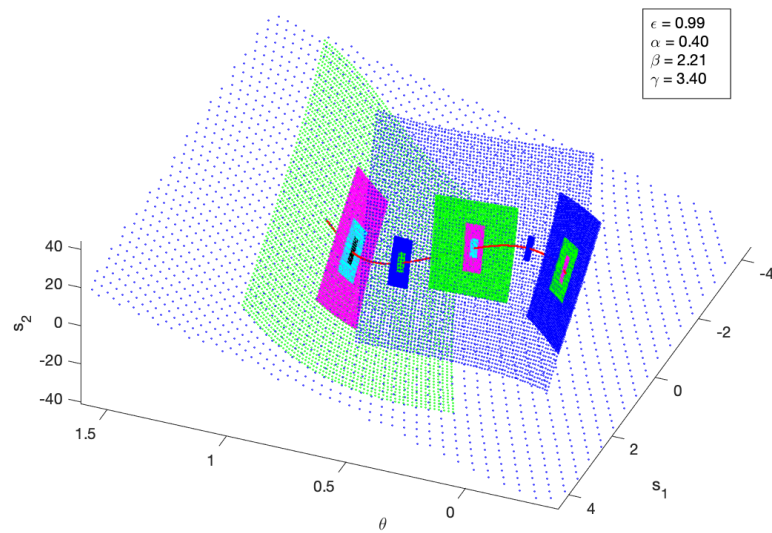*The isochrons contract along both the eigen-directions.*

Figure 3.12: 3D-FAF map: Invariant Circle and Unstable Manifolds
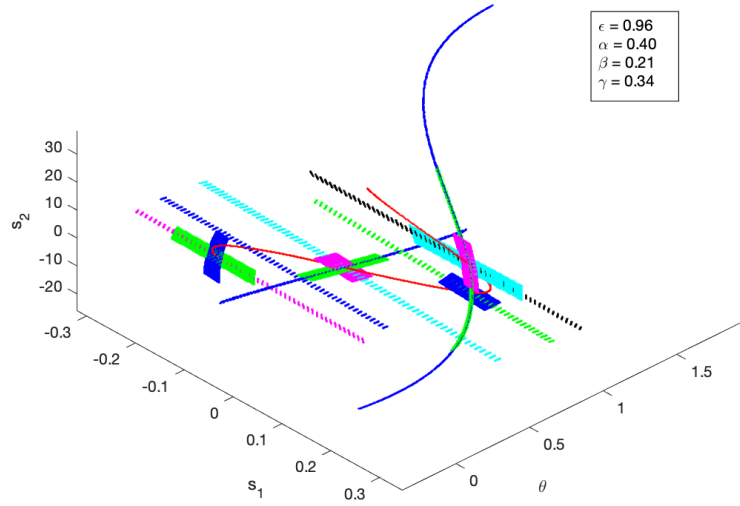*The isochrons expand along both the eigen-directions.*

109

Figure 3.13: 3D-FAF map: Invariant Circle and Stable/Unstable Manifolds for Saddles
The isochrons contract along the stable direction and expand along the unstable direction.

REFERENCES

[1] N. Fenichel, "Persistence and smoothness of invariant manifolds for flows," *Indiana Univ. Math. J.*, vol. 21, pp. 193–226, 1971/72.

[2] Y. B. Pesin, *Lectures on partial hyperbolicity and stable ergodicity*, ser. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2004, pp. vi+122, ISBN: 3-03719-003-5.

[3] A. Winfree, "Patterns of phase compromise in biological cycles," *Journal of Mathematical Biology*, vol. 1, pp. 73–93, May 1974.

[4] J. Guckenheimer, "Isochrons and phaseless sets," *J. Math. Biol.*, vol. 1, no. 3, pp. 259–273, 1974/75.

[5] R. J. Sacker, *ON INVARIANT SURFACES AND BIFURCATION OF PERIODIC SOLUTIONS OF ORDINARY DIFFERENTIAL EQUATIONS*. ProQuest LLC, Ann Arbor, MI, 1964, p. 98, Thesis (Ph.D.)–New York University.

[6] D. Ruelle and F. Takens, "On the nature of turbulence," *Comm. Math. Phys.*, vol. 20, pp. 167–192, 1971.

[7] J. E. Marsden and M. McCracken, *The Hopf bifurcation and its applications*. Springer-Verlag, New York, 1976, pp. xiii+408, With contributions by P. Chernoff, G. Childs, S. Chow, J. R. Dorroh, J. Guckenheimer, 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.

[8] A. A. Andronov, A. A. Vitt, and S. È. Khaɪkin, *Theory of oscillators*. Dover Publications, Inc., New York, 1987, pp. xxxiv+815, Translated from the Russian by F. Immirzi, Reprint of the 1966 translation, ISBN: 0-486-65508-3.

[9] N. Minorsky, *Nonlinear oscillations*. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London-New York, 1962, pp. xviii+714.

[10] A. T. Winfree, *The geometry of biological time*, Second, ser. Interdisciplinary Applied Mathematics. Springer-Verlag, New York, 2001, vol. 12, pp. xxvi+777, ISBN: 0-387-98992-7.

[11] E. M. Izhikevich, *Dynamical systems in neuroscience: the geometry of excitability and bursting*, ser. Computational Neuroscience. MIT Press, Cambridge, MA, 2007, pp. xvi+441, ISBN: 978-0-262-09043-8; 0-262-09043-0.

[12]   G. B. Ermentrout and D. H. Terman, *Mathematical foundations of neuroscience*, ser. Interdisciplinary Applied Mathematics. Springer, New York, 2010, vol. 35, pp. xvi+422, ISBN: 978-0-387-87707-5.

[13]   G. Huguet and R. de la Llave, "Computation of limit cycles and their isochrons: Fast algorithms and their convergence," *SIAM journal on applied dynamical systems.*, vol. 12, pp. 1763–1802, 2013.

[14]   Y. Yao and R. de la Llave, "Algorithms for computation of attractors and isochrones for 2-d maps: Implementation and numerical results," 2021, Manuscript in progress.

[15]   R. de la Llave, "Invariant manifolds associated to nonresonant spectral subspaces," *J. Statist. Phys.*, vol. 87, no. 1-2, pp. 211–249, 1997.

[16]   X. Cabre, E. Fontich, and R. de la Llave, "The parameterization method for invariant manifolds i: Manifolds associated to non-resonant subspaces," *Indiana University Mathematics Journal*, vol. 52, pp. 283–328, 2003.

[17]   ——, "The parameterization method for invariant manifolds ii: Regularity with respect to parameters," *Indiana University Mathematics Journal*, vol. 52, pp. 329–360, 2003.

[18]   X. Cabré, E. Fontich, and R. de la Llave, "The parameterization method for invariant manifolds iii: Overview and applications," *Journal of Differential Equations*, vol. 218, no. 2, pp. 444–515, 2005.

[19]   À. Haro, M. Canadell, J. Figueras, A. Luque, and J. Mondelo, *The Parameterization Method for Invariant Manifolds: From Rigorous Results to Effective Computations*, ser. Applied Mathematical Sciences. Springer International Publishing, 2016, ISBN: 9783319296623.

[20]   M. Ainsworth and J. T. Oden, *A posteriori error estimation in finite element analysis*, ser. Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2000, pp. xx+240, ISBN: 0-471-29411-X.

[21]   J. Moser, "A rapidly convergent iteration method and non-linear partial differential equations. I," *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)*, vol. 20, pp. 265–315, 1966.

[22]   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 Journal on Applied Dynamical Systems*, vol. 6, no. 1, 2007.

[23]   R. Calleja and A. Celletti, "Breakdown of invariant attractors for the dissipative standard map," *Chaos (Woodbury, N.Y.)*, vol. 20, p. 013 121, Mar. 2010.

[24] R. Mañé, "Persistent manifolds are normally hyperbolic," *Transactions of the American Mathematical Society*, vol. 246, pp. 261–283, 1978.

[25] Y. Yao and R. de la Llave, "Algorithms for computation of attractors and isochrones for 2-d maps: Effective algorithms and rigorous proof of convergence," 2021, Manuscript in progress.

[26] J. Schwartz, "On Nash's implicit functional theorem," *Comm. Pure Appl. Math.*, vol. 13, pp. 509–530, 1960.

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

[28] R. de la Llave and R. Obaya, "Regularity of the composition operator in spaces of Hölder functions," *Discrete Contin. Dynam. Systems*, vol. 5, no. 1, pp. 157–184, 1999.

[29] R. Calleja and R. de la Llave, "A numerically accessible criterion for the breakdown of quasi-periodic solutions and its rigorous justification," *Nonlinearity*, vol. 23, no. 9, pp. 2029–2058, 2010.

[30] R. C. Calleja, A. Celletti, and R. de la Llave, "A KAM theory for conformally symplectic systems: Efficient algorithms and their validation," *J. Differential Equations*, vol. 255, no. 5, pp. 978–1049, 2013.

[31] A. 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 and Continuous Dynamical Systems - B*, vol. 6, p. 1261, 2006.

[32] M. Canadell and À. Haro, "Computation of quasi-periodic normally hyperbolic invariant tori: Algorithms, numerical explorations and mechanisms of breakdown," *J. Nonlinear Sci.*, vol. 27, no. 6, pp. 1829–1868, 2017.

[33] ——, "Computation of quasiperiodic normally hyperbolic invariant tori: Rigorous results," *J. Nonlinear Sci.*, vol. 27, no. 6, pp. 1869–1904, 2017.

[34] A. Granados, "Invariant manifolds and the parameterization method in coupled energy harvesting piezoelectric oscillators," *Phys. D*, vol. 351/352, pp. 14–29, 2017.

[35] P. W. Bates, K. Lu, and C. Zeng, "Approximately invariant manifolds and global dynamics of spike states," *Invent. Math.*, vol. 174, no. 2, pp. 355–433, 2008.

[36] M. J. Capiński and P. Zgliczyński, "Geometric proof for normally hyperbolic invariant manifolds," *J. Differential Equations*, vol. 259, no. 11, pp. 6215–6286, 2015.

[37] M. Levi, "Qualitative analysis of the periodically forced relaxation oscillations," *Mem. Amer. Math. Soc.*, vol. 32, no. 244, pp. vi+147, 1981.

[38] Q. Wang and L.-S. Young, "From invariant curves to strange attractors," *Comm. Math. Phys.*, vol. 225, no. 2, pp. 275–304, 2002.

[39] ——, "Strange attractors in periodically-kicked limit cycles and Hopf bifurcations," *Comm. Math. Phys.*, vol. 240, no. 3, pp. 509–529, 2003.

[40] C. Bonatti, L. J. Dı'az, and M. Viana, *Dynamics beyond uniform hyperbolicity*, ser. Encyclopaedia of Mathematical Sciences. Springer-Verlag, Berlin, 2005, vol. 102, pp. xviii+384, A global geometric and probabilistic perspective, Mathematical Physics, III, ISBN: 3-540-22066-6.

[41] M. P. Golden and B. E. Ydstie, "Bifurcationin a model-reference adaptive control system," *Syst. Contr. Lett.*, vol. 11, pp. 413–430, 1988.

[42] D. A. Rand, "Existence, nonexistence and universal breakdown of dissipative golden invariant tori. I. Golden critical circle maps," *Nonlinearity*, vol. 5, no. 3, pp. 639–662, 1992.

[43] ——, "Existence, nonexistence and universal breakdown of dissipative golden invariant tori. II. Convergence of renormalization for mappings of the annulus," *Nonlinearity*, vol. 5, no. 3, pp. 663–680, 1992.

[44] ——, "Existence, nonexistence and universal breakdown of dissipative golden invariant tori. III. Invariant circles for mappings of the annulus," *Nonlinearity*, vol. 5, no. 3, pp. 681–706, 1992.

[45] À. Haro and R. de la Llave, "Manifolds on the verge of a hyperbolicity breakdown," *Chaos*, vol. 16, no. 1, pp. 013120, 8, 2006.

[46] 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.*, vol. 6, no. 1, pp. 142–207, 2007.

[47] K. Bjerklöv and M. Saprykina, "Universal asymptotics in hyperbolicity breakdown," *Nonlinearity*, vol. 21, no. 3, pp. 557–586, 2008.

[48] R. Calleja and J.-L. Figueras, "Collision of invariant bundles of quasi-periodic attractors in the dissipative standard map," *Chaos*, vol. 22, no. 3, pp. 033114, 10, 2012.

[49] J.-L. Figueras and À. Haro, "Reliable computation of robust response tori on the verge of breakdown," *SIAM J. Appl. Dyn. Syst.*, vol. 11, no. 2, pp. 597–628, 2012.

[50] P. W. Bates, K. Lu, and C. Zeng, "Invariant foliations near normally hyperbolic invariant manifolds for semiflows," *Trans. Amer. Math. Soc.*, vol. 352, no. 10, pp. 4641–4676, 2000.

[51] Y.-M. Chung and M. S. Jolly, "A unified approach to compute foliations, inertial manifolds, and tracking solutions," *Math. Comp.*, vol. 84, no. 294, pp. 1729–1751, 2015.

[52] R. Szalai, "Invariant spectral foliations with applications to model order reduction and synthesis," *ArXiv*, no. arXiv:1912.03655v2, 2019.

[53] R. de la Llave and F. Kogelbauer, "Global persistence of lyapunov subcenter manifolds as spectral submanifolds under dissipative perturbations," *SIAM J. Appl. Dyn. Syst.*, vol. 18, no. 4, pp. 2099–2142, 2019.

[54] A. Haro and R. de la Llave, "A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: Rigorous results," *Journal of Differential Equations*, vol. 228, no. 2, pp. 530–579, 2006.

[55] J. Moser, "A rapidly convergent iteration method and non-linear differential equations. II," *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)*, vol. 20, pp. 499–535, 1966.

[56] N. Fenichel, "Asymptotic stability with rate conditions," *Indiana Univ. Math. J.*, vol. 23, pp. 1109–1137, 1973/74.

[57] ——, "Asymptotic stability with rate conditions. II," *Indiana Univ. Math. J.*, vol. 26, no. 1, pp. 81–93, 1977.

[58] R. Mañé, "Persistent manifolds are normally hyperbolic," *Trans. Amer. Math. Soc.*, vol. 246, pp. 261–283, 1978.

[59] J. Jarník and J. Kurzweil, "On invariant sets and invariant manifolds of differential systems," *J. Differential Equations*, vol. 6, pp. 247–263, 1969.

[60] M. J. Capiński and H. Kubica, "Persistence of normally hyperbolic invariant manifolds in the absence of rate conditions," *Nonlinearity*, vol. 33, no. 9, pp. 4967–5005, 2020.

[61] L. Zhang and R. de la Llave, "Transition state theory with quasi-periodic forcing," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 62, pp. 229–243, 2018.

[62] E. M. Stein, *Singular integrals and differentiability properties of functions*, ser. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970, pp. xiv+290.

[63] R. de la Llave and A. Windsor, "Smooth dependence on parameters of solutions to cohomology equations over anosov systems with applications to cohomology equations on diffeomorphism groups," *Discrete and Continuous Dynamical Systems - A*, vol. 29, p. 1141, 2011.

[64] R. de la Llave, *A tutorial on KAM theory*. Providence, R.I. : American Mathematical Society, 2001.

[65] S. Sternberg, "Local contractions and a theorem of Poincaré," *Amer. J. Math.*, vol. 79, pp. 809–824, 1957.

[66] H. Poincaré, "Sur les propriétés des fonctions définies par les équations aux differences partielles," 1879.

[67] H. Dulac, "Recherches sur les points singuliers des équations differentielles," 1903.

[68] J. A. Vano, *A Nash-Moser implicit function theorem with Whitney regularity and applications*. ProQuest LLC, Ann Arbor, MI, 2002, p. 180, Thesis (Ph.D.)–The University of Texas at Austin, ISBN: 978-0496-56073-8.

[69] J. Gimeno, J. Yang, and R. de la Llave, "Numerical computation of periodic orbits and isochrons for state-dependent delay perturbation of an ODE in the plane," *CoRR*, vol. abs/2005.06086, 2020. arXiv: 2005.06086.

[70] C. A. Hall and W. Meyer, "Optimal error bounds for cubic spline interpolation," *Journal of Approximation Theory*, vol. 16, no. 2, pp. 105–122, 1976.

[71] M. Steffen, "A simple method for monotonic interpolation in one dimension.," *Astronomy and Astrophysics*, vol. 239, pp. 443–450, 1990.

[72] P. K. Mohanty, M. Reza, P. Kumar, and P. Kumar, "Implementation of cubic spline interpolation on parallel skeleton using pipeline model on cpu-gpu cluster," in *2016 IEEE 6th International Conference on Advanced Computing (IACC)*, IEEE, 2016, pp. 747–751, ISBN: 9781467382861.

[73] P. D. Michailidis and K. G. Margaritis, "Parallel direct methods for solving the system of linear equations with pipelining on a multicore using openmp," *Journal of Computational and Applied Mathematics*, vol. 236, no. 3, pp. 326–341, 2011, Aspects of Numerical Algorithms, Parallelization and Applications.

[74] A. B. Katok, *Introduction to the modern theory of dynamical systems*. Cambridge : Cambridge University Press, 1995.

[75] U. Krengel, "On the speed of convergence in the ergodic theorem," *Monatshefte für Mathematik*, vol. 86, no. 1, pp. 3–6, 1978.

[76] S. Das, Y. Saiki, E. Sander, and J. Yorke, *Quasiperiodicity: Rotation Numbers, Proceeding of The Foundations of Chaos Revisited: From Poincare to Recent Advancements, Chapter 7*. Switzerland: Springer Complexity, 2016.

[77] ——, "Solving the babylonian problem of quasiperiodic rotation rates," *Discrete and Continuous Dynamical Systems - S*, vol. 12, pp. 2279–2305, Jan. 2018.

[78] M. W. Hirsch, C. C. Pugh, and M. Shub, *Invariant manifolds*, ser. Lecture Notes in Mathematics, Vol. 583. Springer-Verlag, Berlin-New York, 1977, pp. ii+149.

[79] A. Chenciner and G. Iooss, "Bifurcations de tores invariants," *Arch. Rational Mech. Anal.*, vol. 69, no. 2, pp. 109–198, 1979.

[80] ——, "Persistance et bifurcation de tores invariants," *Arch. Rational Mech. Anal.*, vol. 71, no. 4, pp. 301–306, 1979.

[81] G. R. Sell, "Bifurcation of higher-dimensional tori," *Arch. Rational Mech. Anal.*, vol. 69, no. 3, pp. 199–230, 1979.

[82] M. J. Capiński and H. Kubica, "Persistence of normally hyperbolic invariant manifolds in the absence of rate conditions," *Nonlinearity*, vol. 33, no. 9, pp. 4967–5005, 2020.

[83] A. Floer, "A refinement of the Conley index and an application to the stability of hyperbolic invariant sets," *Ergodic Theory Dynam. Systems*, vol. 7, no. 1, pp. 93–103, 1987.

[84] ——, "A topological persistence theorem for normally hyperbolic manifolds via the Conley index," *Trans. Amer. Math. Soc.*, vol. 321, no. 2, pp. 647–657, 1990.

[85] Q. Wang and L.-S. Young, "Strange attractors in periodically-kicked limit cycles and Hopf bifurcations," *Comm. Math. Phys.*, vol. 240, no. 3, pp. 509–529, 2003.

[86] ——, "Strange attractors with one direction of instability," *Comm. Math. Phys.*, vol. 218, no. 1, pp. 1–97, 2001.

[87] M. Levi, "Qualitative analysis of the periodically forced relaxation oscillations," *Mem. Amer. Math. Soc.*, vol. 32, no. 244, pp. vi+147, 1981.

[88] J.-L. Figueras and À. Haro, "Triple collisions of invariant bundles," *Discrete Contin. Dyn. Syst. Ser. B*, vol. 18, no. 8, pp. 2069–2082, 2013.

[89] ——, "Different scenarios for hyperbolicity breakdown in quasiperiodic area preserving twist maps," *Chaos*, vol. 25, no. 12, pp. 123119, 16, 2015.

[90] T. Ohlson Timoudas, "Power law asymptotics in the creation of strange attractors in the quasi-periodically forced quadratic family," *Nonlinearity*, vol. 30, no. 12, pp. 4483–4522, 2017.