PERSISTENCE OF INVARIANT OBJECTS UNDER DELAY PERTURBATIONS

A Dissertation
Presented to
The Academic Faculty

By

Jiaqi Yang

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

Georgia Institute of Technology

August 2021

PERSISTENCE OF INVARIANT OBJECTS UNDER DELAY PERTURBATIONS

Thesis committee:

Dr. Luca Dieci
School of Mathematics
Georgia Institute of Technology

Dr. Yao Yao
School of Mathematics
Georgia Institute of Technology

Dr. Rafael de la Llave
School of Mathematics
Georgia Institute of Technology

Dr. Chongchun Zeng
School of Mathematics
Georgia Institute of Technology

Dr. Erik Verriest
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date approved: May 6, 2021

To my loved ones.

ACKNOWLEDGMENTS

First, I would like to express my sincere appreciation to my thesis advisor, Prof. Rafael de la Llave, for his constant guidance, support, and patience. I cherish all the discussions and the happy times we had. His enthusiasm, wisdom, and kindness will continuously inspire me in my life and career. I am more than lucky to have him as my advisor.

I would like to acknowledge my thesis committee members, professors Luca Dieci, Erik Verriest, Yao Yao, Chongchun Zeng, for their insightful questions and suggestions for my thesis, along with their help and support throughout my graduate study.

Special thanks to my dear friend Joan Gimeno, for years of collaboration and company. To professors Àlex Haro, Jean-Philippe Lessard, and Jason Mireles-James, I am grateful for the wonderful time working with them.

I would like to express my gratitude to Prof. Yingfei Yi for professional advice and help. I also appreciate all the help of Prof. Enid Steinbart offered for my teaching and career. I am grateful for the valuable suggestions on my research from Prof. Carmen Chicone. Thank Ms. Klara Grodzinsky for her help and advice on my teaching. I am thankful to Prof. Mohammad Ghomi, Prof. Rachel Kuske, and Prof. Xingxing Yu for their support.

To other members in our group: Adrián Pérez Bustamante, Hongyu Cheng, Christopher Dupre, Jorge L. Gonzalez, Xiaolong He, Bhanu Kumar, Jieun Seong, Fenfen Wang, Pingyuan Wei, Xiaodan Xu, Yian Yao, and Lei Zhang. To my friends: Jiangning Chen, Renyi Chen, Sally Collins, Zhibo Dai, Jaewoo Jung, Qianli Hu, Kisun Lee, Changong Li, Hangfan Li, Ruilin Li, Shasha Liao, Xiao Liu, Shu Liu, Hyunki Min, Jaemin Park, Thomas Rodewald, Longmei Shu, Haodong Sun, Xiaowen Sun, Xin Wang, Yuqing Wang, Xin Xing, Xiaofan Yuan, Haoyan Zhai, Weiwei Zhang. Thank you for the friendship and all the fun we had.

Last but not the least, I would like to thank my parents and my husband for their unconditional love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

There has been much effort on understanding the behavior of functional differential equations, e.g, equations with delays, in particular when the delays have complicated formats. Developing the general theory for equations with involved delays is a challenging task. This work is dedicated to studying the invariant objects of equations with delay perturbations, for example, the delay-related parts are led by small parameters, or the delays are small.

CHAPTER 1

INTRODUCTION AND BACKGROUND

1.1 Functional Differential Equations

Many causes in the natural sciences take some time to generate effects. If one incorporates these delays in the models, one is led to descriptions of systems in which the derivatives of states are functions of the states at previous times. These are commonly called delay differential equations (DDEs). DDEs arise naturally in models from electrodynamics, control theory, biology, neuroscience, and economics, see [1, 2, 3, 4, 5, 6, 7, 8, 9, 10] and references therein.

When the delay is a positive constant h, one can start with a function defined on $[-h, 0]$ as the initial condition and integrate forward. Note that the natural phase space is the space of continuous functions defined on $[-h, 0]$, which is infinite-dimensional. There is a rather satisfactory theory of existence and uniqueness and even a qualitative theory for this case, see [3, 11, 12, 13].

However, equations from many real-life applications require non-constant delays, e.g., the delay may depend on time, the state at current time, a segment on the trajectory (state-dependent delay). In some cases, the delay involve integration (distributed delay), or it is defined implicitly. Moreover, the delays do not have to be literally delays, they can be negative, the so-called advances, where the derivatives of states depend on the states at some future times. Sometimes several delays are involved in one equation, with different forms. The mathematical theory for these kinds of functional differential equations (FDEs) in general has complications, and there is still a lot to explore.

1.2 Invariant Objects

Invariant objects are important landmarks in dynamical systems which govern the behavior of the solutions. Solutions start on the invariant objects will stay there for all time. Examples of invariant objects include periodic orbits, invariant tori, stable manifolds, unstable manifolds, center manifolds, etc.

There has been many interests in studying the invariant objects both in theory and practice. The parameterization method, developed in [14, 15, 16, 17] attracts attention thanks to its advantages proven in theoretical as well as numerical study.

1.3 Goal and Methodology

The goal of this work is to study the effect of delay-related singular perturbations. These singular perturbations will produce a lot of new solutions which are not present prior to perturbations. At the same time, some solutions in the unperturbed equations will persist, and indeed, will guide our understanding of the perturbed systems. More precisely, we start with ordinary differential equations (ODEs) or partial differential equations (PDEs) which have some invariant objects, and show that when the delay-related perturbations satisfy some conditions, the invariant objects will persist. We allow the delays to admit various complicated forms, or to be advances, as long as the conditions are satisfied. In particular, we analyze equations with implicitly defined small delays appearing in electrodynamics.

Our approach relies on the parameterization method and functional analysis. We deal with functional equations coming from the parameterization method and construct their solutions using tools from functional analysis. We provide results in "a posteriori" format: given an approximate solution of the functional equation, which has some good condition numbers, we prove that there is a true solution close to the approximate one. Thus, our result can be used to validate approximate solutions produced even by non-rigorous methods, e.g. formal power expansions in the delay, or the results of numerical computations. The proofs are constructive and lead to practical algorithms.

A philosophy similar to that of this thesis has also been used in other papers. [18, 19, 20] develop functional equations for quasi-periodic solutions in several contexts and study them using KAM theory.

This thesis is based on a series of joint works with Dr. Joan Gimeno and Dr. Rafael de la Llave.

1.4 Prospectives

The results and proofs here are suitable for computer-assisted proofs. One can find with confidence the size of the perturbative parameter which ensures that the invariant objects are preserved.

The methods introduced here constitute a powerful toolkit that we hope can produce results in other problems. For example, we expect to get persistence and higher regularity of the center manifolds for state-dependent delay equations (SDDEs), which is essential for applications of the center manifold reduction to bifurcation theory [5]. We can also study more dynamical objects, like hyperbolic sets, under delay perturbations. Of course, removing the perturbative setting remains a long term goal, but this seems to pass through refining the theory of existence and regularity of [21]. Similar ideas can be used to investigate effects of localized perturbations of dynamical systems on their invariant objects (center manifold, normally hyperbolic invariant manifold).

1.5 Organization

In Chapter 2, we study an SDDE resulting from adding a state-dependent delay perturbation to a planar ODE, which is based on [22]. There we study stable periodic orbits and their stable foliations. Chapter 3, based on [23], investigates FDEs with some delays or advances which are close to ODEs or evolutionary PDEs with periodic orbits. Numerical implementation of results in Chapter 2 is explained in Chapter 4 [24].

CHAPTER 2

LIMIT CYCLES AND ISOCHRONS FOR FDES NEAR A PLANAR ODE

In the case where the delay in an equation is not a constant and depends on the state, one needs to consider state-dependent delay equations (SDDEs). In contrast with the constant delay case, more challenges present when developing the mathematical theory of SDDEs. The paper [21] made important progress for the appropriate phase space for SDDEs. We refer to [4] for a very comprehensive survey of the mathematical theory and the applications.

In this chapter, we consider a simple model (two-dimensional ODE with a limit cycle) and show that all solutions close to the limit cycle present in this model persist (in some appropriate sense) when we add a state-dependent delay perturbation. Models of the form considered here (see equation (2.4)) appear in several concrete problems in the natural sciences (circuits, neuroscience, and population dynamics), where small delay effects are added, see [4].

The result is subtle to formulate since the perturbation of adding a state-dependent delay is very singular, it changes the nature of the equation: the unperturbed case is finite-dimensional while the perturbed case is infinite-dimensional. The basic idea is that we establish the existence of some finite-dimensional families of solutions (in the phase space of the SDDE), which resemble (in an appropriate sense) the solutions of the original ODE. This allows to establish many other properties (e.g. dependence on parameters) which may be false for solutions of SDDEs in general.

The method of proof bypasses the theory of SDDEs based on the evolution operator. We consider the class of functions of time that have a well defined behavior (e.g. periodic, or asymptotic to periodic) and derive functional equations which impose that they are solutions of the SDDE. These functional equations are studied using functional analysis methods. The method of proof also leads to algorithms which have been implemented, see Chapter 4.

One advantage of the method presented is that it allows to obtain smooth dependence on parameters for the periodic solutions and their slow stable manifolds without studying the smoothness of the flow (which seems to be problematic for SDDEs, for now the optimal result on smoothness of the flow is C^1). Therefore, one can obtain higher than C^1 dependence on parameters.

We hope that the method can be extended in several directions. For example, we hope to produce higher dimensional families, families with other behaviors, and to consider more complicated models. The conjectural picture is that in SDDEs, even if the dynamics in a full Banach space of solutions is problematic, one can find a very rich set of solutions organized in families. The families may not fit together well and may leave gaps, so that a general theory may have problems [25].

2.1 Overview of the Method

It is known that in a neighborhood of a limit cycle of a 2-dimensional ODE, we can find $K : \mathbb{T} \times [-1, 1] \to \mathbb{R}^2$, and ω_0 and λ_0 in such a way that for any θ, s, the function given by

$$x(t) = K(\theta + \omega_0 t, se^{\lambda_0 t}) \tag{2.1}$$

solves the ODE, see [26]. The fact that all the functions of the form equation (2.1) are solutions of the original ODE is equivalent to a functional equation for K, ω_0 and λ_0, which we call "invariance equation" (2.6). Efficient methods to study the resulting functional equation were presented in [26]. We will, henceforth, assume that K, ω_0, λ_0 are known.

Similarly, for the perturbed case, when we impose that for fixed θ, s the function of the form

$$x(t) = K \circ W(\theta + \omega t, se^{\lambda t}) \tag{2.2}$$

is a solution of our delay differential equation, we obtain a functional equation for W, ω, λ (see equation (2.8)). Note that the unknowns in equation (2.8) are the embedding W and the numbers ω, λ.

Our goal will be to solve equation (2.8) using techniques of functional analysis. The equation is rather degenerate and our treatment has several steps. We first find some asymptotic expansions in powers of s to a finite order, and then, we formulate a fixed point problem for the remainder. Due to the delay, information at previous times is needed. We anticipate a technical problem is that the domain of definition of the unknown have to depend on the details of the unknown. Similar problems appear in the theory of center manifolds [27]. Here we have to resort to cut-offs and extensions. After this process, we get a prepared equation, equation (2.9), which has the same format as equation (2.8), and agrees with equation (2.8) in a neighborhood. Solutions of the prepared equation which stay in the neighborhood will be solutions of the original problem.

The main result of this chapter is Theorem 7, which establishes that with respect to some condition numbers of the problem, verified for small enough ε, given an approximate solution of the extended invariance equation equation (2.9), one obtain a true solution nearby (*"a posteriori"* format). Then as a corollary, Theorem 10 answers the question of smooth dependence on parameters.

As in the case of center manifolds, the family of solutions found to the original problem may depend on the extension considered.

2.2 Formulation of the Problem

We consider an ordinary differential equation in the plane

$$\dot{x}(t) = X_0(x(t)), \tag{2.3}$$

where $x(t) \in \mathbb{R}^2$, $X_0 : \mathbb{R}^2 \to \mathbb{R}^2$ is analytic. We assume above equation (2.3) admits a limit cycle. Clearly, there is a two dimensional family of solutions to this ODE. This family can be parameterized e.g. by the initial conditions, but as we will see, there are more efficient parameterizations near the limit cycle.

We study an SDDE that is a "small" modification of equation (2.3) in which we add some small term for the derivative that depends on some previous time. Adding some dependence on the solution at previous times, arises naturally in many problems. Limit cycles appear in feedback loops and if the feedback loops have a delayed effect, which depends on the present state, to incorporate them in the model, we are led to:

$$\dot{x}(t) = X(x(t), \varepsilon x(t - r(x(t)))), \qquad 0 \leqslant \varepsilon \ll 1. \tag{2.4}$$

Where $x(t) \in \mathbb{R}^2$, $X : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ is analytic, the state-dependent delay function $r : \mathbb{R}^2 \to [0, h]$ is as smooth as we need. The equation (2.4) is formally a perturbation of equation (2.3) with $X(x, 0) = X_0(x)$.

We can rewrite equation (2.4) as

$$\dot{x}(t) = X(x(t), 0) + \varepsilon P(x(t), x(t - r(x(t))), \varepsilon), \tag{2.5}$$

where we define

$$\varepsilon P(x(t), x(t - r(x(t))), \varepsilon) = X(x(t), \varepsilon x(t - r(x(t)))) - X(x(t), 0).$$

We will find a two dimensional family of solutions of equation (2.4), which resembles the two-dimensional family of solutions of equation (2.3). This is a much simpler problem than developing a general theory of existence of solutions for an SDDE, which is a rather difficult. Nevertheless, persistence of some family of solutions is of physical interest.

Since the perturbed problem is infinite-dimensional, the precise meaning of the continuation of the unperturbed solutions into solutions of the perturbed problem is somewhat subtle.

2.2.1 Limit cycles and isochrons for ODEs

Under our assumption, there exists a limit cycle (stable periodic orbit) in the unperturbed equation (2.3). In a neighborhood of the limit cycle, points have asymptotic phases (see [28, 29]). The points sharing the same asymptotic phase as point p on the limit cycle is the stable manifold for point p. The stable manifold of the limit cycle is foliated by the stable manifolds for points on the limit cycle (sometimes referred as stable foliations). The stable manifolds for points on the limit cycle are also called isochrons in the biology literature, see [28, 29].

According to [26], we can find a parameterization of the limit cycle and the isochrons in a neighborhood of the limit cycle. More precisely, there exists real numbers $\omega_0 > 0$, $\lambda_0 < 0$, and an analytic local diffeomorphism $K : \mathbb{T} \times [-1, 1] \to \mathbb{R}^2$, such that

$$X_0(K(\theta, s)) = DK(\theta, s) \begin{pmatrix} \omega_0 \\ \lambda_0 s \end{pmatrix}, \tag{2.6}$$

where K is periodic in θ, i.e. $K(\theta+1, s) = K(\theta, s)$. Saying that K solves equation (2.6) is equivalent to saying that for fixed parameters θ and s, the function $x(t) = K(\theta+\omega_0 t, se^{\lambda_0 t})$ solves equation (2.3) for all t such that $|se^{\lambda_0 t}| < 1$. Notice that when $s = 0$, $K(\theta, 0)$ parameterizes the limit cycle, and for a fixed θ with varying s, we get the local stable manifold of the point $K(\theta, 0)$.

Note that geometrically, K can be viewed as a change of coordinates, under which the original vector field is equivalent to the vector field $X_0'(\theta, s) = (\omega_0, \lambda_0 s)$ in the space $\mathbb{T} \times [-1, 1]$. We could have started with this vector field X_0' and then added some perturbation to it. However, to keep contact with applications, we decided not to do this.

Remark 1. *As pointed out in [26], the K solving equation* (2.6) *can never be unique. If*

$K(\theta, s)$ is a solution of equation (2.6)*, then for any θ_0, $b \neq 0$, $K(\theta + \theta_0, bs)$ will also be a solution of equation* (2.6)*. [26] also shows that this is the only source of non-uniqueness. We will call such b scaling factor, and such θ_0 phase shift. Note that by using a different b, we can change the domain of K. However, no matter how the domain changes, s has to lie in a finite interval.*

In this chapter, for the equation after perturbation (2.4), we will show that if ε is small enough, the limit cycle and its isochrons for the unperturbed equation persist as limit cycle and its slow stable manifolds of the delayed model. We will use the name isochrons to denote the slow stable manifolds and distinguish them from the infinite dimensional stable manifolds similar to the one established in [12]. Meanwhile, we will find a parameterization of them. More precisely, we will find $\omega > 0$, $\lambda < 0$, and W which maps a subset of $\mathbb{T} \times \mathbb{R}$ to a subset of $\mathbb{T} \times \mathbb{R}$, such that for small s, $K \circ W(\theta, s)$ gives us a parameterization of the limit cycle as well as its isochrons in a neighborhood. We assume that W can be lifted to a function from $\mathbb{R}^2$ to $\mathbb{R}^2$ (we will use the same letter to denote the function before and after the lift) which satisfies the periodicity condition:

$$W(\theta + 1, s) = W(\theta, s) + \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right). \tag{2.7}$$

We remark that $K \circ W$ being a parameterization of the limit cycle and its isochrons is the same as for given θ, and s in the domain of W, $x(t) = K \circ W(\theta + \omega t, se^{\lambda t})$ solving equation (2.4) for $t \geqslant 0$.

2.2.2 The invariance equation and the prepared invariance equation

Substitute $x(t) = K \circ W(\theta + \omega t, se^{\lambda t})$ into equation (2.5), let $t = 0$, use the fact that DK is invertible, we get that $x(t) = K \circ W(\theta + \omega t, se^{\lambda t})$ solves equation (2.4) if and only if W

satisfies

$$DW(\theta,s)\begin{pmatrix}\omega\\ \lambda s\end{pmatrix}=\begin{pmatrix}\omega_0\\ \lambda_0 W_2(\theta,s)\end{pmatrix}+\varepsilon Y(W(\theta,s),\widetilde{W}(\theta,s),\varepsilon), \tag{2.8}$$

where $W_2(\theta,s)$ is the second component of $W(\theta,s)$, $\widetilde{W}$ is the term caused by the delay:

$$\widetilde{W}(\theta,s)=W(\theta-\omega r\circ K(W(\theta,s)),se^{-\lambda r\circ K(W(\theta,s))}),$$

and

$$Y(W(\theta,s),\widetilde{W}(\theta,s),\varepsilon)=(DK(W(\theta,s)))^{-1}P(K(W(\theta,s)),K(\widetilde{W}(\theta,s)),\varepsilon).$$

Note that even if $\widetilde{W}$ is typographically convenient, $\widetilde{W}$ is a very complicated function of W, it involves several compositions.

Now we need to look at equation (2.8) more closely and specify the domain and range of W. One cannot define W on $\mathbb{T}\times[-b,b]$, where $b>0$ is a constant. Indeed, observing the second component in expression of $\widetilde{W}$, $se^{-\lambda r\circ K(W(\theta,s))}$, one will note that $|se^{-\lambda r\circ K(W(\theta,s))}|>|s|$. This will drive us out of the domain of W since the second component of W lies in a finite interval. Therefore, W has to be defined for all s on the real line. So we let $W:\mathbb{T}\times\mathbb{R}\to\mathbb{T}\times\mathbb{R}$. There is another technical issue as pointed out in the following Remark 2.

Remark 2. *When ε is small, we expect W to be close to the identity map. Then for s far from 0, $W(\theta,s)$ does not lie in the domain of K, thus the invariance equation is not well defined.*

Similar to the study of center manifolds. We will use cut-off functions to resolve the above issues.

We will transform our original equation (2.8) into a well-defined equation of the same format:

$$DW(\theta,s)\begin{pmatrix}\omega\\ \lambda s\end{pmatrix}=\begin{pmatrix}\omega_0\\ \lambda_0 W_2(\theta,s)\end{pmatrix}+\varepsilon\overline{Y}(W(\theta,s),\widetilde{W}(\theta,s),\varepsilon), \tag{2.9}$$

where $\overline{Y}$ is defined on $(\mathbb{T}\times\mathbb{R})^2\times\mathbb{R}_+$, and $\overline{r\circ K}$ is defined on $\mathbb{T}\times\mathbb{R}$, with slight abuse of notation, we still denote the term caused by the delay as $\widetilde{W}$:

$$\widetilde{W}(\theta,s)=W(\theta-\omega\overline{r\circ K}(W(\theta,s)),se^{-\lambda\overline{r\circ K}(W(\theta,s))}).$$

We follow standard practice in the theory of center manifolds of differential equations, see [27], and introduce the extensions:

- For $r\circ K$ which is defined only on $\mathbb{T}\times[-1,1]$, we define a function $\overline{r\circ K}$ on $\mathbb{T}\times\mathbb{R}$, which agrees with $r\circ K$ on $\mathbb{T}\times[-\frac{1}{2},\frac{1}{2}]$, and is zero outside of $\mathbb{T}\times[-1,1]$.
- For $Y:(\mathbb{T}\times[-1,1])^2\times\mathbb{R}_+\to\mathbb{R}^2$, we define $\overline{Y}:(\mathbb{T}\times\mathbb{R})^2\times\mathbb{R}_+\to\mathbb{R}^2$, which agrees with Y on the set $(\mathbb{T}\times[-\frac{1}{2},\frac{1}{2}])^2\times\mathbb{R}_+$, and is zero outside $(\mathbb{T}\times[-1,1])^2\times\mathbb{R}_+$.

To achieve above extensions, let $\phi:\mathbb{R}\to[0,1]$ be a C^∞ cut-off function:

$$\phi(x)=\begin{cases}1 & \text{if} \quad |x|\leqslant\frac{1}{2},\\ 0 & \text{if} \quad |x|>1.\end{cases} \tag{2.10}$$

We define

$$\overline{r\circ K}(\theta,s)=r\circ K(\theta,s)\phi(s),$$

and,

$$\overline{Y}(W(\theta,s),\widetilde{W}(\theta,s),\varepsilon)=Y(W(\theta,s),\widetilde{W}(\theta,s),\varepsilon)\phi(W_2(\theta,s))\phi(\widetilde{W}_2(\theta,s)).$$

After these extensions, equation (2.8) is turned into the well-defined equation (2.9). Note that, $\overline{Y}$, $\overline{r\circ K}$ defined above have bounded derivatives in their domains up to any order.

Remark 3. *In the definition of cut-off function, one can let ϕ to vanish for $|x| > c_1$ where the constant $c_1 < 1$, and let $\phi = 1$ for $|x| \leqslant c_2$ where the constant $c_2 < c_1$.*

Remark 4. *The use of the cut-off function here is very similar to the use of cut-offs in the study of center manifolds in the literature, if we choose a different cut-off function ϕ, we will possibly end up with a different W, which solves equation* (2.9) *with the new cut-off function ϕ.*

Remark 5. *If instead of having a stable periodic orbit, the unperturbed ODE has an unstable periodic orbit, then λ_0 in equation* (2.6) *is positive. Analogous results to Theorems 6 and 7 will give us the parameterization of the periodic orbit and the unstable manifold for small ε. The same proof, only with minor modifications, will work. At the same time, the invariance equation* (2.8) *will be well-defined for a suitably chosen domain for W, we do not need to do extensions. Similarly, the idea here will also work for advanced equations, which have the same format as equation* (2.4)*, with $r : \mathbb{R}^2 \to [-h, 0]$. We omit the details for these cases.*

2.2.3 Representation of the unknown function

In order to study the functional equation equation (2.9), we consider W of the form

$$W(\theta, s) = \sum_{j=0}^{N-1} W^j(\theta)s^j + W^>(\theta, s) \tag{2.11}$$

solving equation (2.9). Where $W^0(\theta)$ is the zeroth order term in s, $W^j(\theta)s^j$ is the j-th order term in s, $W^>(\theta, s)$ is of order at least N in s. $W^j : \mathbb{T} \to \mathbb{T} \times \mathbb{R}$, and $W^> : \mathbb{T} \times \mathbb{R} \to \mathbb{T} \times \mathbb{R}$. As we will see, the truncation number N could be chosen as any integer larger than 1 to obtain the main result of this chapter. From now on, we will use superscripts to denote corresponding orders, and subscripts, as we did before, to denote corresponding components.

We consider lifts of $W^0(\theta)$, $W^j(\theta)$, and $W^>(\theta, s)$, which will be functions from $\mathbb{R}$ or $\mathbb{R}^2$ to $\mathbb{R}^2$. We will not distinguish notations for the functions before or after lifts. According to the periodicity condition for W in equation (2.7), the lifted functions satisfy the following periodicity conditions:

$$W^0(\theta+1) = W^0(\theta) + \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), \tag{2.12}$$

$$W^j(\theta+1) = W^j(\theta), \tag{2.13}$$

$$W^>(\theta+1, s) = W^>(\theta, s). \tag{2.14}$$

Based on the form of W in equation (2.11), we can match coefficients of different powers of s in the invariance equation (2.9). Thus, the invariance equation (2.9) is equivalent to a sequence of equations. As we will see, the equations for W^0 and W^1 are special. The equation for W^0 is very nonlinear, the equation for W^1 is a relative eigenvector equation. The equations for W^j's are all similar. The equation for $W^>$ is hard to study, it has 2 variables. As we will see later, for small enough ε, $W^>$ is the only case where we need the cut-off.

2.2.3.1 Invariance equation for zero order term

Matching zero order terms of s in equation (2.9), we obtain the equation for the unknowns ω and W^0:

$$\omega \frac{d}{d\theta} W^0(\theta) - \begin{pmatrix} \omega_0 \\ \lambda_0 W_2^0(\theta) \end{pmatrix} = \varepsilon \overline{Y}(W^0(\theta), \widetilde{W}^0(\theta;\omega), \varepsilon), \tag{2.15}$$

where

$$\widetilde{W}^0(\theta;\omega) = W^0\left(\theta - \omega \overline{r \circ K}(W^0(\theta))\right)$$

is the function caused by delay.

2.2.3.2 *Invariance equation for first order term*

Equating the coefficients of s^1 in equation (2.9), we obtain the equation for the unknowns λ and W^1:

$$\omega\frac{d}{d\theta}W^1(\theta) + \lambda W^1(\theta) - \begin{pmatrix} 0 \\ \lambda_0 W_2^1(\theta) \end{pmatrix} = \varepsilon\overline{Y}^1(\theta, \lambda, W^0, W^1, \varepsilon), \tag{2.16}$$

where $\overline{Y}^1(\theta, \lambda, W^0, W^1, \varepsilon)$ is the coefficient of s in $\overline{Y}$. $\overline{Y}^1(\theta, \lambda, W^0, W^1, \varepsilon)$ is linear in W^1. We will specify it later in equation (2.44).

2.2.3.3 *Invariance equation for the j-th order term*

For $2 \leqslant j \leqslant N-1$, matching the coefficients of s^j, the equation for the unknown W^j is:

$$\omega\frac{d}{d\theta}W^j(\theta) + \lambda j W^j(\theta) - \begin{pmatrix} 0 \\ \lambda_0 W_2^j(\theta) \end{pmatrix} = \varepsilon\overline{Y}^j(\theta, \lambda, W^0, W^j, \varepsilon) + R^j(\theta), \tag{2.17}$$

where $\overline{Y}^j(\theta, \lambda, W^0, W^j, \varepsilon)$ is the coefficient of s^j in $\overline{Y}$. $\overline{Y}^j(\theta, \lambda, W^0, W^j, \varepsilon)$ is linear in W^j, which will be specified in equation (2.56), and R^j is a function of θ which depends only on W^0, W^1,..., W^{j-1}.

Having $W^0, \ldots, W^{N-1}$, we are ready to consider $W^>$.

2.2.3.4 *Invariance equation for higher order term*

Note that $W^>(\theta, s)$ solves the equation:

$$(\omega\partial_\theta + s\lambda\partial_s)W^>(\theta, s) - \begin{pmatrix} 0 \\ \lambda_0 W_2^>(\theta, s) \end{pmatrix} = \varepsilon Y^>(W^>, \theta, s, \varepsilon) \tag{2.18}$$

where $Y^{>}(W^{>}, \theta, s, \varepsilon)$ is the term of order at least N in s of $\overline{Y}$, which will be specified later in equation (2.64).

2.3 Main Results

2.3.1 Results for prepared equations

Under the assumption that the map $\overline{Y} : (\mathbb{T} \times \mathbb{R})^2 \times \mathbb{R}_+ \to \mathbb{R}^2$ has bounded derivatives up to any order, $\overline{r \circ K} : \mathbb{T} \times \mathbb{R} \to [0, h]$ has bounded derivatives up to any order, we have:

Theorem 6 (Zero Order)**.** *For any given integer $L > 0$, there is $\varepsilon_0 > 0$ such that when $0 \leqslant \varepsilon < \varepsilon_0$, there exist an $\omega > 0$ and an L times differentiable map $W^0 : \mathbb{T} \to \mathbb{T} \times \mathbb{R}$, with L-th derivative Lipschitz, which solve equation* (2.15)*.*

Moreover, for initial guess ω^0, and $W^{0,0}(\theta)$ satisfying the periodicity condition equation (2.12)*. If they satisfy the invariance equation* (2.15) *with error $E^0(\theta)$, then there exist unique ω, $W^0(\theta)$ (satisfying the periodic condition equation* (2.12)*) closed by solving the same equation exactly, with*

$$\|W^{0,0} - W^0\|_{C^l} \leqslant C\|E^0\|_{C^0}^{1-\frac{l}{L}}, \quad 0 \leqslant l < L \tag{2.19}$$

$$|\omega^0 - \omega| \leqslant C\|E^0\|_{C^0}, \tag{2.20}$$

for some constant C, where C may depend on ε, ω_0, λ_0, l, L, and prior bound for $\|W^{0,0}\|_{L+\mathrm{Lip}}$. In fact, W^0 has derivatives up to any order.

Moreover,

Theorem 7 (All Orders)**.** *For any given integers $N \geqslant 2$, and $L \geqslant 2 + N$, there is $\varepsilon_0 > 0$ such that when $0 \leqslant \varepsilon < \varepsilon_0$, there exist $\omega > 0$, $\lambda < 0$, and $W : \mathbb{T} \times \mathbb{R} \to \mathbb{T} \times \mathbb{R}$ of the form*

$$W(\theta, s) = \sum_{j=0}^{N-1} W^j(\theta)s^j + W^{>}(\theta, s) \tag{2.21}$$

which solve the equation (2.9) *in a neighborhood of* $s = 0$.

Where $W^0 : \mathbb{T} \to \mathbb{T} \times \mathbb{R}$ *is* L *times differentiable with Lipschitz* L*-th derivative. For* $1 \leqslant j \leqslant N-1$, $W^j : \mathbb{T} \to \mathbb{T} \times \mathbb{R}$ *is* $(L-1)$ *times differentiable with Lipschitz* $(L-1)$*-th derivative, and* $W^>$ *is of order at least N in* s *and is jointly* $(L-2-N)$ *times differentiable in* θ *and* s*, with* $(L-2-N)$*-th derivative Lipschitz.*

Moreover, if ω^0, $W^{0,0}(\theta)$, λ^0, $W^{1,0}(\theta)$, $W^{j,0}(\theta)$, *and* $W^{>,0}(\theta, s)$ *satisfy the invariance equations* (2.15)*,* (2.16)*,* (2.17)*, and* (2.18)*, with errors* $E^0(\theta)$, $E^1(\theta)$, $E^j(\theta)$, *and* $E^>(\theta, s)$*, respectively, then there are* ω, $W^0(\theta)$, λ, $W^1(\theta)$, $W^1(\theta)$, *and* $W^>(\theta, s)$ *which solve equations* (2.15)*,* (2.16)*,* (2.17)*, and* (2.18)*. Therefore, equation* (2.9) *is solved by* ω, λ, *and* $W(\theta, s)$ *of above form* (2.21)*. For* $0 \leqslant l \leqslant L-2-N$*, we have*

$$\begin{aligned} &\|W(\theta, s) - \sum_{j=0}^{N-1} W^{j,0}(\theta)s^j - W^{>,0}(\theta, s)\|_{C^l} \\ &\qquad \leqslant C\Big(\sum_{j=0}^{N-1} \|E^j\|_{C^0}|s|^j + \|E^>\|_{0,N}|s|^N\Big)^{1-\frac{l}{(L-2-N)}}, \end{aligned} \tag{2.22}$$

$$\begin{aligned} |\omega - \omega^0| &\leqslant C(\|E^0\|_{C^0}), \\ |\lambda - \lambda^0| &\leqslant C(\|E^1\|_{C^0}), \end{aligned} \tag{2.23}$$

for some constant C *depending on* ε, ω_0, λ_0, N, l, L, *prior bounds for* $\|W^{0,0}\|_{L+\mathrm{Lip}}$, $\|W^{j,0}\|_{L-1+\mathrm{Lip}}$, $j = 1, \ldots, N-1$*, and derivatives of* $W^{>,0}$.

Remark 8. *In Theorem 6,* $W^0(\theta)$ *is unique up to a phase shift.*

Remark 9. *The above theorems are in a posteriori format. The main input needed are some functions that satisfy the invariance equations approximately. This can be numerical computations (that indeed produce good approximate solutions) or Lindstedt series, see for example [30].*

Notice that with these Theorems, we do not need to analyze the procedure used to *produce the approximate solutions. The only thing that we need to establish is that the solutions produced satisfy the invariance equations up to small errors.*

The a posteriori format of the theorem leads to automatic Hölder dependence of the solution W^0 on ε and Y.

It suffices to observe that if we consider W^0 solving the invariance equation for some ε_1, Y_1, it will solve the invariance equation for ε_2, Y_2 up to an error which is bounded in the C^l norm by $C\left(|\varepsilon_1 - \varepsilon_2| + \|Y_1 - Y_2\|_{C^0}\right)^{1-\frac{l}{L}}$

Our approach leads very easily to smooth dependence on parameters.

Theorem 10. *Consider a family of functions Y_η, r_η as above, where η lies in an open interval $I \subset \mathbb{R}$. Assume that Y_η and r_η are smooth in their inputs as well as in η, with bounded derivatives.*

Then for any positive integer L, there is a small enough positive ε_0 such that when $\varepsilon < \varepsilon_0$, for each $\eta \in I$ we can find ω_η, W_η^0 solving equation (2.15).

Furthermore, the $W_\eta^0(\theta)$ is jointly $C^{L+\mathrm{Lip}}$ in η, θ.

Theorem 11. *Under the same assumption as in Theorem 10, for any given integers $N \geqslant 2$, and $L \geqslant 2 + N$, there is a small enough positive ε_0 such that when $\varepsilon < \varepsilon_0$, for each $\eta \in I$, we can find ω_η, W_η^0, λ_η, W_η^j, $j = 1, \ldots, N-1$, and $W_\eta^>(\theta, s)$, which solve the invariance equations* (2.15), (2.16), (2.17), *and* (2.18).

Furthermore, $W_\eta^0(\theta)$ is jointly $C^{L+\mathrm{Lip}}$ in η, θ; $W_\eta^j(\theta)$, $j = 1, \ldots, N-1$, are jointly $C^{L-1+\mathrm{Lip}}$ in η, θ; $W_\eta^>(\theta, s)$ is jointly $C^{L-2-N+\mathrm{Lip}}$ in η, θ, and s.

Note that the regularity conclusions of Theorem 10 can be interpreted in a more functional form as the mapping that to η associates W_η^0 is $C^{\ell+\mathrm{Lip}}$ when the space of embedding W is given the $C^{L-\ell}$ topology. Similar interpretation can be made for Theorem 11. This functional point of view is consistent with the point of view of RFDE where the phase space is infinite dimensional.

2.3.2 Results for original problem in a neighborhood of the limit cycle

Note that to find the low order terms, W^j ($j = 1, \dots, N-1$), for small ε, the extensions are not needed. Heuristically, the low order terms are *infinitesimals*. Hence, to compute them, it suffices to know the expansion of the vector field.

More precisely, if we take the initial guess for zero order term as $W^{0,0}(\theta) = \left(\begin{smallmatrix}\theta \\ 0\end{smallmatrix}\right)$, the error is of order ε. Then by Theorem 6, the true solution W^0 is within a distance of order ε from $W^{0,0}(\theta)$. Therefore, with ε being small enough, we have $\sup_{\theta\in\mathbb{T}} |W_2^0(\theta)| < \frac{1}{2}$, we are reduced to the case without extension:

$$\begin{aligned}
&\overline{r \circ K}(W^0(\theta)) = r \circ K(W^0(\theta)),\\
&\overline{Y}(W^0(\theta), \widetilde{W}^0(\theta;\omega), \varepsilon) = Y(W^0(\theta), \widetilde{W}^0(\theta;\omega), \varepsilon),
\end{aligned}$$

where,

$$\widetilde{W}^0(\theta;\omega) = W^0(\theta - \omega r \circ K(W^0(\theta))).$$

Then we can rewrite the invariance equation for W^0, equation (2.15), as:

$$\omega \frac{d}{d\theta} W^0(\theta) - \begin{pmatrix} \omega_0 \\ \lambda_0 W_2^0(\theta) \end{pmatrix} = \varepsilon Y(W^0(\theta), \widetilde{W}^0(\theta;\omega), \varepsilon). \tag{2.24}$$

Similar arguments apply for the equations for W^1 and W^j's ($2 \leqslant j \leqslant N-1$) if we look at expressions of $\overline{Y}^1$ in equation (2.44), $\overline{Y}^j$ in equation (2.56), and form of R^j.

We can find $0 < s_0 < \frac{1}{2}$, such that $W(\mathbb{T} \times [-s_0, s_0]) \subset \mathbb{T} \times [-\frac{1}{2}, \frac{1}{2}]$, and $\widetilde{W}(\mathbb{T} \times [-s_0, s_0]) \subset \mathbb{T} \times [-\frac{1}{2}, \frac{1}{2}]$. Therefore, the original problem is solved in a neighborhood of the limit cycle by applying the results in section 2.3.1.

For the original problem in section 2.2, we have

Corollary 12 (Limit Cycle)**.** *When $\varepsilon < \varepsilon_0$ in Theorem 6 is so small that $\sup_{\theta\in\mathbb{T}} |W_2^0(\theta)| < \frac{1}{2}$, equation* (2.4) *admits a limit cycle close to the limit cycle of the unperturbed equation.*

If ω, W^0 solve the invariance equation (2.24)*, then $K \circ W^0(\theta)$ gives a parameterization of the limit cycle of equation* (2.4)*, i.e. for any θ, $K \circ W^0(\theta + \omega t)$ solves equation* (2.4) *for all t.*

We can also find a 2-parameter family of solutions close to the limit cycle:

Corollary 13 (Isochrons)**.** *For small ε as in Corollary 12, there are isochrons for the limit cycle of equation* (2.4)*. If ω, λ, and $W : \mathbb{T} \times \mathbb{R} \to \mathbb{T} \times \mathbb{R}$ solve the extended invariance equation* (2.9)*, then there exists $0 < s_0 < \frac{1}{2}$, such that $K \circ W(\theta, s)$, $|s| \leqslant s_0$ gives a parameterization of the limit cycle with its isochrons in a neighborhood, i.e. for any θ, and s, with $|s| \leqslant s_0$, $K \circ W(\theta + \omega t, se^{\lambda t})$ solves equation* (2.4) *for all $t \geqslant 0$.*

One can formulate dependence on parameters results using Theorems 10 and 11. The cut-offs and extensions should be carried out in a way that preserves the smoothness with respect to parameters, which can be done by applying the bump functions in the same way for all the elements in the family. Note that only the higher order term $W^>$ requires extension. We omit the precise formulations here.

2.3.3 Comparison with results on RFDEs based on time evolution

The persistence of a periodic solution under perturbation for retarded functional differential equation (RFDE) is presented in Chapter 10 of [12], notably Theorem 4.1. In this section, we present some remarks that can help the specialists to compare our results with those obtainable considering the time evolution of RFDEs.

The set up presented there does not seem to apply to our case since the phase space considered in [12] is the space of continuous functions on an interval, namely, $C^0[-h, 0]$, and they require differentiability properties of the equation which are not satisfied in our case. Note also that we can obtain smooth dependence on parameters (see Theorem 10). Obtaining such smooth dependence using the methods based on the evolutionary approach would require obtaining regularity of the evolution operator, which does not seem to be available.

More precisely, if we employ the notation x_t as a function defined on $[-h, 0]$, with

$$x_t(s) = x(t+s)$$

for $s \in [-h, 0]$, we can write our SDDE equation (2.4) as

$$\dot{x}(t) = F(x_t, \varepsilon),$$

where we define $F(\phi, \varepsilon) := X(\phi(0), \varepsilon\phi(-r(\phi(0))))$. For $\varepsilon = 0$, we have an ODE, which can be viewed as a delay equation, with a non-degenerate periodic orbit (see [12]). However, above F cannot be continuously differentiable in ϕ if ϕ is only continuous. This obstructs application of Theorem 4.1 for RFDE in [12].

It is very interesting to study whether a similar method to the one in [12] can be extended to our case with some variations of the phase space (solution manifold, see [21]). However, since only C^1 regularity of the evolution has been proved ([21]), (higher regularity of the evolution in SDDE seems problematic), one cannot hope to obtain more than C^1 dependence on parameters. On the other hand, the method in this chapter allows to get rather straightforwardly higher smoothness with respect to parameters. See Theorem 10. We mention that some progress in continuation of periodic orbits is in [31, 32].

Considering RFDEs as evolutions in infinite dimensional phase spaces, [12] establishes the existence of infinite-dimensional strong stable manifolds for periodic orbits corresponding to the Floquet multipliers smaller than a number.

Again, we remark that there are some technical issues of regularity of evolutions in the phase space of SDDE to define stable manifolds and even stability. We hope that these regularity issues of the evolution can be made precise (using techniques as in [21, 33, 34]).

Nevertheless, there is a very fundamental difference between the manifolds we consider and those in [12].

If we consider the unperturbed ODE as an RFDE in an infinite-dimensional phase space, the Floquet multipliers are 1 with multiplicity 1, $\exp\left(\frac{\lambda_0}{\omega_0}\right)$ with multiplicity 1, and 0 (with infinite multiplicity). With C^1-smoothness of the evolution as in [21], under small perturbation, the new Floquet multipliers are closed by (one exactly 1, one close to $\exp(\frac{\lambda_0}{\omega_0})$ and infinitely many near 0).

The theory developed in [12] attaches an infinite-dimensional manifold to the most stable part of the spectrum in the case of RFDEs. That is the strong stable manifold.

Although the stability for all the solutions in a neighborhood of the limit cycle is out of the scope of the present paper, heuristically, the manifold that we consider here, in the infinite-dimensional phase space, is attached to the least stable Floquet multiplier, hence it is a slow stable manifold from the infinite-dimensional point of view.

We think that the finite-dimensional manifold we obtain is more practically relevant than the strong stable manifold. We expect that infinitely many modes will die out very fast and, therefore, be hard to observe. All the solutions of the full problem will be asymptotically similar to the solutions we consider. In summary, solutions close to the limit cycle will converge to the limit cycle tangent to the slow stable manifolds described here. One problem to make all this precise is that the evolution is only known to be C^1.

Our motivation is to obtain solutions which resemble solutions of the ODE, in accordance with the physical intuition that the solutions in the perturbed problem – in spite of the singular nature of the perturbation – look similar to those of the unperturbed problem (this is the reason why relativity and its delays were hard to discover).

One of the features of the formalism in this chapter is that it allows to describe in a unified way the solutions of the SDDE in an infinite-dimensional space and the solutions of the unperturbed finite-dimensional ODE.

Of course in this chapter, we only deal with models of a very special kind, (we indeed have the hope that the range of applicability of the method can be expanded; the models considered here are a proof of concept) but we obtain rather smooth invariant manifolds and smooth dependence on parameters with high degree of differentiability. Furthermore, the proof presented here leads to algorithms to compute the limit cycles and their manifolds. These algorithms are practical and have been implemented, see Chapter 4.

It is also interesting to investigate whether evolution based methods lead to computational algorithms [35] and compare them with the algorithms based on functional equations.

2.4 Overview of the Proof

In equation (2.15), ω and W^0 are the unknowns. Under a choice of the phase, we define an operator such that its fixed point solves equation (2.15). We will show that when ε is small enough, the operator is a "C^0" contraction and maps a $C^{L+\mathrm{Lip}}$ ball to itself. Then one obtains the existence of the fixed point (ω, W^0), and that W^0 in the fixed point has some regularity. Therefore, equation (2.15) is solved.

In equation (2.16), λ and W^1 are the unknowns. We will impose an appropriate normalization when defining the operator to make sure the solution is uniquely found, and that W is close to the identity map with appropriate scaling factor. Then similar to above case, for small enough ε, this operator has a fixed point (λ, W^1) in which W^1 has some regularity.

In equation (2.17), W^j is the only unknown. We define an operator which is a contraction for small enough ε. The operator has a fixed point with certain regularity solving the equation.

In equation (2.18), $W^>$ is an unknown function of 2 variables. We will define an operator on a function space with a weighted norm, then prove that for small ε, this operator has a fixed point in this function space, which solves the equation (2.18).

We emphasize again that for small enough ε, the equation for $W^>$ is the only place where extension is needed. (Recall section 2.3.2)

There are finitely many smallness conditions for ε, so there are ε's which satisfy all the smallness conditions.

Same idea will be used for proving the smooth dependence on parameters.

2.5 Proof of the Main Results

2.5.1 Zero order solution

In this section, we prove our first result, Theorem 6.

Recall equation (2.15), invariance equation for ω and W^0, which is obtained by setting $s = 0$ in equation (2.9).

Componentwise, $W^0 = (W_1^0, W_2^0)$, and $\overline{Y} = (\overline{Y}_1, \overline{Y}_2)$, we have the equations as:

$$\omega \frac{d}{d\theta} W_1^0(\theta) - \omega_0 = \varepsilon \overline{Y}_1(W^0(\theta), \widetilde{W}^0(\theta;\omega), \varepsilon), \tag{2.25}$$

and

$$\omega \frac{d}{d\theta} W_2^0(\theta) - \lambda_0 W_2^0(\theta) = \varepsilon \overline{Y}_2(W^0(\theta), \widetilde{W}^0(\theta;\omega), \varepsilon). \tag{2.26}$$

Taking periodicity condition equation (2.12) into account, we define an operator Γ^0 as follows:

$$\begin{aligned} \Gamma^0 \begin{pmatrix} a \\ Z_1 \\ Z_2 \end{pmatrix}(\theta) &= \begin{pmatrix} \Gamma_1^0(a, Z) \\ \Gamma_2^0(a, Z)(\theta) \\ \Gamma_3^0(a, Z)(\theta) \end{pmatrix} \\ &= \begin{pmatrix} \omega_0 + \varepsilon \int_0^1 \overline{Y}_1(Z(\theta), \widetilde{Z}(\theta; a), \varepsilon) d\theta \\ \frac{1}{\Gamma_1^0(a,Z)} \left(\omega_0 \theta + \varepsilon \int_0^\theta \overline{Y}_1(Z(\sigma), \widetilde{Z}(\sigma; a), \varepsilon) d\sigma\right) \\ \varepsilon \int_0^\infty e^{\lambda_0 t} \overline{Y}_2(Z(\theta - at), \widetilde{Z}(\theta - at; a), \varepsilon) dt \end{pmatrix}, \end{aligned} \tag{2.27}$$

Notice that if Γ^0 has a fixed point (a^*, Z^*), then equation (2.15) are solved by $\omega = a^*$ and $W^0 = Z^*$, at the same time, periodic condition equation (2.12) is satisfied.

Remark 14. *As we can see, the operator Γ^0 depends on ε, however, to simplify the expression, we do not include ε in the notation of the operator Γ^0.*

Remark 15. *Similar to Remark 1, we will not have uniqueness of the solution to invariance equation* (2.15)*. Once we have a solution $W^0(\theta)$ to the equation, for any $\theta_0 \neq 0$, $W^0(\theta+\theta_0)$ will also solve the equation, which is called phase shift. This is indeed the only source of non-uniqueness.*

By considering the operator equation (2.27)*, we fix a phase by $\Gamma_2^0(a, Z)(0) = 0$.*

For the domain of Γ^0, we consider the closed interval $I^0 = \{a : |a - \omega_0| \leqslant \frac{\omega_0}{2}\}$. For fixed positive integer L and positive constant B^0, define a subset of the space of functions which are L times differentiable, with Lipschitz L-th derivative as follows (see more details about regularity properties in Section A):

$$\begin{aligned}\mathscr{C}_0^{L+\mathrm{Lip}} = \{f \mid f\colon \mathbb{T} \to \mathbb{T} \times \mathbb{R}, f &\text{ can be lifted to a function from } \mathbb{R} \text{ to } \mathbb{R}^2,\\ &\text{still denoted as } f, \text{which satisfies } f(\theta+1) = f(\theta) + \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right),\\ &f_1(0) = 0, \|f\|_{L+\mathrm{Lip}} \leqslant B^0\},\end{aligned} \tag{2.28}$$

where

$$\|f\|_{L+\mathrm{Lip}} = \max_{i=1,2,k=0,\dots,L} \left\{ \sup_{\theta\in[0,1]} \|f_i^{(k)}(\theta)\|,\ \mathrm{Lip}(f_i^{(L)}) \right\}.$$

Define $D^0 = I^0 \times \mathscr{C}_0^{L+\mathrm{Lip}}$, then Γ^0 is defined on D^0. We have the following:

Lemma 16. *There exists $\varepsilon^0 > 0$, such that when $\varepsilon < \varepsilon^0$, $\Gamma^0(D^0) \subset D^0$.*

Proof. For $(a, Z) \in D^0$, by assumption, we have that $\overline{Y}_1(Z(\theta), \widetilde{Z}(\theta; a), \varepsilon)$ is bounded by a constant which is independent of choice of (a, Z) in D^0. Then, one can choose ε small enough such that $\Gamma_1^0(a, Z) = \omega_0 + \varepsilon \int_0^1 \overline{Y}_1(Z(\theta), \widetilde{Z}(\theta; a), \varepsilon)d\theta$ is in I^0.

Now consider $\Gamma_2^0(a, Z)(\theta) = \frac{1}{\Gamma_1^0(a,Z)}\big(\omega_0\theta + \varepsilon \int_0^\theta \overline{Y}_1(Z(\sigma), \widetilde{Z}(\sigma; a), \varepsilon)d\sigma\big)$. First we observe that

$$\Gamma_2^0(a, Z)(\theta + 1) = \Gamma_2^0(a, Z)(\theta) + 1.$$

Then we need to check bounds for the derivatives

$$\frac{d}{d\theta}\Gamma_2^0(a, Z)(\theta) = \frac{1}{\Gamma_1^0(a, Z)}\big(\omega_0 + \varepsilon\overline{Y}_1(Z(\theta), \widetilde{Z}(\theta; a), \varepsilon)\big).$$

By Faá di Bruno's formula in Lemma 86, for $2 \leqslant n \leqslant L$, $\frac{d^n}{d\theta^n}\Gamma_2^0(a, Z)(\theta)$ will be a polynomial of a common factor $\frac{\varepsilon}{\Gamma_1^0(a,Z)}$, each term will contain products of derivatives of $\overline{Y}_1$, Z, and $\overline{r \circ K}$ up to order $(n-1)$. By assumption on $\overline{Y}_1$ and $\overline{r \circ K}$, for $(a, Z) \in D^0$, if we choose B^0 to be larger than 2, then for small enough ε, $\Gamma_2^0(a, Z)(\theta)$ on $[0, 1]$ has derivatives up to order L bounded by B^0 and $L-th$ derivative Lipschitz with Lipschitz constant less than B^0.

For $\Gamma_3^0(a, Z)(\theta) = \varepsilon \int_0^\infty e^{\lambda_0 t}\overline{Y}_2(Z(\theta - at), \widetilde{Z}(\theta - at; a), \varepsilon)dt$. It satisfies

$$\Gamma_3^0(a, Z)(\theta + 1) = \Gamma_3^0(a, Z)(\theta).$$

To establish bounds for the derivatives of $\Gamma_3^0(a, Z)(\theta)$, we apply a similar argument as above. Notice that for $n \leqslant L$, $\frac{\partial^n}{\partial\theta^n}\overline{Y}_2(Z(\theta - at), \widetilde{Z}(\theta - at; a), \varepsilon)$ will be a polynomial with each term a product of derivatives of $\overline{Y}_2$, Z, and $\overline{r \circ K}$ up to order n. With regularity of $\overline{Y}_2$, and $\overline{r \circ K}$, for $(a, Z) \in D^0$, $|\frac{\partial^n}{\partial\theta^n}\overline{Y}_2(Z(\theta - at), \widetilde{Z}(\theta - at), \varepsilon)|$ will be bounded. Therefore, for small enough ε, $\Gamma_3^0(a, Z)$ has derivatives up to order L bounded by B^0 and its $L - th$ derivative is Lipschitz with Lipschitz constant less than B^0.

If we take ε^0 such that above conditions are satisfied at the same time, then for $\varepsilon < \varepsilon^0$, we have $\Gamma^0(D^0) \subset D^0$. □

We now define a distance on D^0, which is essentially C^0 distance. Under this distance, the space D^0 is complete. For (a, Z) and (a', Z') in D^0,

$$d((a, Z), (a', Z')) \coloneqq |a - a'| + \|Z - Z'\|, \tag{2.29}$$

where

$$\|Z - Z'\| = \max\left\{\sup_{\theta}|Z_1(\theta) - Z_1'(\theta)|, \sup_{\theta}|Z_2(\theta) - Z_2'(\theta)|\right\}. \tag{2.30}$$

Lemma 17. *There exists $\varepsilon^0 > 0$, such that when $\varepsilon < \varepsilon^0$, under above choice of distance equation* (3.22) *on D^0, the operator Γ^0 is a contraction.*

Proof. We will show that for ε small enough, (the explicit form of smallness conditions will become clear along the proof), we can find a constant $\mu_0 < 1$ such that for distance defined in equation (3.22)

$$d(\Gamma^0(a, Z), \Gamma^0(a', Z')) < \mu_0 \cdot d((a, Z), (a', Z')). \tag{2.31}$$

Note that

$$\begin{aligned} d(\Gamma^0(a, Z), \Gamma^0(a', Z')) = &\left|\Gamma_1^0(a, Z) - \Gamma_1^0(a', Z')\right| \\ &+ \|(\Gamma_2^0(a, Z), \Gamma_3^0(a, Z)) - (\Gamma_2^0(a', Z'), \Gamma_3^0(a', Z'))\| \end{aligned} \tag{2.32}$$

More explicitly, above distance is

$$\begin{aligned} &\varepsilon\left|\int_0^1 \overline{Y}_1(Z(\theta), \widetilde{Z}(\theta; a), \varepsilon)d\theta - \int_0^1 \overline{Y}_1(Z'(\theta), \widetilde{Z}'(\theta; a'), \varepsilon)d\theta\right| \\ &\quad + \max\left\{\sup_{\theta}\left|\frac{1}{\Gamma_1^0(a, Z)}\big(\omega_0\theta + \varepsilon\int_0^\theta \overline{Y}_1(Z(\sigma), \widetilde{Z}(\sigma; a), \varepsilon)d\sigma\big)\right.\right. \\ &\qquad\qquad \left. - \frac{1}{\Gamma_1^0(a', Z')}\big(\omega_0\theta + \varepsilon\int_0^\theta \overline{Y}_1(Z'(\sigma), \widetilde{Z}'(\sigma; a'), \varepsilon)d\sigma\big)\right|, \\ &\qquad\quad \varepsilon\sup_{\theta}\left|\int_0^\infty e^{\lambda_0 t}\overline{Y}_2(Z(\theta - at), \widetilde{Z}(\theta - at; a), \varepsilon)dt\right. \\ &\qquad\qquad\qquad \left.\left. - \int_0^\infty e^{\lambda_0 t}\overline{Y}_2(Z'(\theta - a't), \widetilde{Z}'(\theta - a't; a'), \varepsilon)dt\right|\right\} \end{aligned} \tag{2.33}$$

Now we consider each term of above expression equation (2.33). Note that in the above expression, it suffices to take the supremums for $\theta \in [0, 1]$, which follows from periodicity condition equation (2.12). By adding and subtracting terms, we have

$$
\begin{aligned}
&\left|\overline{Y}_1(Z(\theta),\widetilde{Z}(\theta;a),\varepsilon)-\overline{Y}_1(Z'(\theta),\widetilde{Z}'(\theta;a'),\varepsilon)\right|\\
&=\left|\overline{Y}_1(Z(\theta),Z(\theta-a\overline{r\circ K}(Z(\theta))),\varepsilon)-\overline{Y}_1(Z'(\theta),Z'(\theta-a'\overline{r\circ K}(Z'(\theta))),\varepsilon)\right|\\
&\leqslant\left|\overline{Y}_1(Z(\theta),Z(\theta-a\overline{r\circ K}(Z(\theta))),\varepsilon)-\overline{Y}_1(Z'(\theta),Z(\theta-a\overline{r\circ K}(Z(\theta))),\varepsilon)\right|\\
&\quad+\left|\overline{Y}_1(Z'(\theta),Z(\theta-a\overline{r\circ K}(Z(\theta))),\varepsilon)-\overline{Y}_1(Z'(\theta),Z'(\theta-a\overline{r\circ K}(Z(\theta))),\varepsilon)\right|\\
&\quad+\left|\overline{Y}_1(Z'(\theta),Z'(\theta-a\overline{r\circ K}(Z(\theta))),\varepsilon)-\overline{Y}_1(Z'(\theta),Z'(\theta-a'\overline{r\circ K}(Z(\theta))),\varepsilon)\right|\\
&\quad+\left|\overline{Y}_1(Z'(\theta),Z'(\theta-a'\overline{r\circ K}(Z(\theta))),\varepsilon)-\overline{Y}_1(Z'(\theta),Z'(\theta-a'\overline{r\circ K}(Z'(\theta))),\varepsilon)\right|.
\end{aligned}
$$

By the mean value theorem, and the fact that (a,Z) and (a',Z') are in D^0, we have

$$
\begin{aligned}
\big|\overline{Y}_1(Z(\theta),&\widetilde{Z}(\theta;a),\varepsilon)-\overline{Y}_1(Z'(\theta),\widetilde{Z}'(\theta;a'),\varepsilon)\big|\\
&\leqslant 2\|D\overline{Y}_1\|\|Z-Z'\|+\|D\overline{Y}_1\|\|DZ'\|\|\overline{r\circ K}\||a-a'|\\
&\quad+|D\overline{Y}_1\|\|DZ'\||a'|\|D(\overline{r\circ K})\|\|Z-Z'\|\\
&\leqslant\|D\overline{Y}_1\|\left(2+B^0|a'|\|D(\overline{r\circ K})\|\right)\|Z-Z'\|\\
&\quad+\|D\overline{Y}_1\|B^0\|\overline{r\circ K}\||a-a'|.
\end{aligned}
\tag{2.34}
$$

Where the norms are supremum norms on $\mathbb{R}$ or $\mathbb{R}^2$ (defined as above in equation (2.30)), and

$$
\|D\overline{Y}_1\|=\max\{\|D_1\overline{Y}_1\|,\|D_2\overline{Y}_1\|\},\tag{2.35}
$$

where $\|D_i\overline{Y}_1\|$, $i=1,2$, is the supremum of the operator norm corresponding to the infinity norm defined on $\mathbb{R}^2$.

By equation (2.34),

$$
\begin{aligned}
\left|\Gamma_1^0(a,Z)-\Gamma_1^0(a',Z')\right|&\leqslant\varepsilon\|D\overline{Y}_1\|\left(2+B^0|a'|\|D(\overline{r\circ K})\|\right)\|Z-Z'\|\\
&\quad+\varepsilon B^0\|D\overline{Y}_1\|\|\overline{r\circ K}\||a-a'|
\end{aligned}
\tag{2.36}
$$

Now consider the first component of the maximum, for $\theta\in[0,1]$ in equation (2.33), by adding and subtracting terms, we have:

$$
\begin{aligned}
&\left|\Gamma_2^0(a,Z)-\Gamma_2^0(a',Z')\right| \\
&\quad\leqslant \frac{\varepsilon}{|\Gamma_1^0(a,Z)|}\int_0^1\left|\overline{Y}_1(Z(\theta),\widetilde{Z}(\theta),\varepsilon)d\theta-\overline{Y}_1(Z'(\theta),\widetilde{Z}'(\theta),\varepsilon)\right|d\theta \\
&\qquad+\frac{\varepsilon\int_0^1\left|\overline{Y}_1(Z'(\theta),\widetilde{Z}'(\theta;a'),\varepsilon)\right|d\theta}{|\Gamma_1^0(a,Z)\Gamma_1^0(a',Z')|}\left|\Gamma_1^0(a,Z)-\Gamma_1^0(a',Z')\right| \\
&\qquad+\frac{|\omega_0|}{|\Gamma_1^0(a,Z)\Gamma_1^0(a',Z')|}\left|\Gamma_1^0(a,Z)-\Gamma_1^0(a',Z')\right| \\
&\quad\leqslant \frac{\varepsilon}{|\Gamma_1^0(a,Z)|}\int_0^1\left|\overline{Y}_1(Z(\theta),\widetilde{Z}(\theta),\varepsilon)d\theta-\overline{Y}_1(Z'(\theta),\widetilde{Z}'(\theta),\varepsilon)\right|d\theta \\
&\qquad+\frac{|\omega_0|+\varepsilon\|\overline{Y}_1\|}{|\Gamma_1^0(a,Z)\Gamma_1^0(a',Z')|}\left|\Gamma_1^0(a,Z)-\Gamma_1^0(a',Z')\right|.
\end{aligned}
\tag{2.37}
$$

By equation (2.34) and equation (2.36), with $\Gamma_1^0(a,Z),\Gamma_1^0(a',Z')\in I^0$, we have,

$$
\begin{aligned}
&\left|\Gamma_2^0(a,Z)-\Gamma_2^0(a',Z')\right| \\
&\quad\leqslant \frac{\varepsilon|\omega_0|+\varepsilon^2\|\overline{Y}_1\|+\varepsilon|\Gamma_1^0(a',Z')|}{|\Gamma_1^0(a,Z)\Gamma_1^0(a',Z')|}\Bigg(\|D\overline{Y}_1\|B^0\|\overline{r\circ K}\||a-a'| \\
&\qquad\qquad+\|D\overline{Y}_1\|\left(2+B^0|a'|\|D(\overline{r\circ K})\|\right)\|Z-Z'\|\Bigg)
\end{aligned}
\tag{2.38}
$$

For the third term, similar to before, we add and subtract terms, then use the mean value theorem to get the estimate

$$
\begin{aligned}
&\left|\overline{Y}_2(Z(\theta-at),\widetilde{Z}(\theta-at;a),\varepsilon)-\overline{Y}_2(Z'(\theta-a't),\widetilde{Z}'(\theta-a't;a'),\varepsilon)\right| \\
&\quad\leqslant 2\|D\overline{Y}_2\|\|Z-Z'\|+2t\|D\overline{Y}_2\|\|DZ'\||a-a'|+\|D\overline{Y}_2\|\|DZ'\|\|r\circ K\||a-a'| \\
&\qquad+\|D\overline{Y}_2\|\|DZ'\||a'|\|D(\overline{r\circ K})\|\|Z-Z'\| \\
&\qquad+t\|D\overline{Y}_2\|\|DZ'\|^2|a'|\|D(\overline{r\circ K})\||a-a'| \\
&\quad\leqslant \|D\overline{Y}_2\|\big(2+B^0|a'|\|D(\overline{r\circ K})\|\big)\|Z-Z'\| \\
&\qquad+B^0\|D\overline{Y}_2\|\|\overline{r\circ K}\||a-a'|+tB^0\|D\overline{Y}_2\|\big(2+B^0|a'|\|D(\overline{r\circ K})\|\big)|a-a'|.
\end{aligned}
\tag{2.39}
$$

Where $\|D\overline{Y}_2\|$ is defined similarly to equation (2.35). Then,

$$
\begin{aligned}
&\big|\Gamma_3^0(a,Z),-\Gamma_3^0(a',Z')\big| \\
&\qquad \leqslant \varepsilon\|D\overline{Y}_2\|B^0\big(\frac{1}{\lambda_0^2}(2+B^0|a'|\|D(\overline{r\circ K})\|)-\frac{\|\overline{r\circ K}\|}{\lambda_0}\big)|a-a'| \\
&\qquad\quad -\frac{\varepsilon}{\lambda_0}\|D\overline{Y}_2\|\big(2+B^0|a'|\|D(\overline{r\circ K})\|\big)\|Z-Z'\|.
\end{aligned}
\tag{2.40}
$$

With above estimates for each terms equation (2.36), equation (2.38), and equation (2.40), we have that for the distance defined in equation (3.22), $d\big(\Gamma^0(a,Z),\Gamma^0(a',Z')\big)$ is smaller than the sums of the right hand sides of equation (2.36), equation (2.38), and equation (2.40). More precisely,

$$
d\big(\Gamma^0(\omega,Z),\Gamma^0(\omega_2,Z')\big)\leqslant c_1|a-a'|+c_2\|Z-Z'\|
$$

Where

$$
\begin{aligned}
c_1 = \varepsilon B^0\|\overline{r\circ K}\| &\left(\|D\overline{Y}_1\|\big(1+\frac{|\omega_0|+\varepsilon\|\overline{Y}_1\|+|\Gamma_1^0(a',Z')|}{|\Gamma_1^0(a,Z)\Gamma_1^0(a',Z')|}\big)-\frac{\|D\overline{Y}_2\|}{\lambda_0}\right) \\
&+\varepsilon\frac{B^0}{\lambda_0^2}\|D\overline{Y}_2\|\big(2+B^0|a'|\|D(\overline{r\circ K})\|\big)
\end{aligned}
$$

and

$$
c_2=\varepsilon\big(2+B^0|a'|\|D(\overline{r\circ K})\|\big)\left(\|D\overline{Y}_1\|\big(1+\frac{|\omega_0|+\varepsilon\|\overline{Y}_1\|+|\Gamma_1^0(a',Z')|}{|\Gamma_1^0(a,Z)\Gamma_1^0(a',Z')|}\big)-\frac{\|D\overline{Y}_2\|}{\lambda_0}\right).
$$

Since a, a', $\Gamma_1^0(a,Z)$, and $\Gamma_1^0(a',Z')$ are all in I^0, we have

$$
\begin{aligned}
c_1 \leqslant \varepsilon B^0\|\overline{r\circ K}\| &\left(\|D\overline{Y}_1\|\big(1+\frac{4|\omega_0|+4\varepsilon\|\overline{Y}_1\|+6|\omega_0|}{|\omega_0|^2}\big)-\frac{\|D\overline{Y}_2\|}{\lambda_0}\right) \\
&+\varepsilon\frac{B^0}{\lambda_0^2}\|D\overline{Y}_2\|\big(2+B^0|a'|\|D(\overline{r\circ K})\|\big),
\end{aligned}
$$

and

$$c_2 \leqslant \varepsilon\big(2 + B^0|a'|\|D(\overline{r \circ K})\|\big)\left(\|D\overline{Y}_1\|\big(1 + \frac{4|\omega_0| + 4\varepsilon\|\overline{Y}_1\| + 6|\omega_0|}{|\omega_0|^2}\big) - \frac{\|D\overline{Y}_2\|}{\lambda_0}\right)$$

Because c_1 and c_2 are bounded by ε multiplied by some constants, they can be as small as we want when ε is small. Therefore, for sufficiently small ε, we can find a constant $\mu_0 < 1$, such that equation (2.31) is true, Γ^0 is a contraction. $\square$

Taking any initial guess $(\omega^0, W^{0,0}(\theta)) \in D^0$. For example, one can take $\omega = \omega_0$, $W^{0,0}(\theta) = \binom{\theta}{0}$. Iterations of this initial guess under Γ^0 will have a limit by Lemma 17. Then Lemma 16 and Lemma 92 ensure that the limit is in D^0. Therefore, we have a fixed point of Γ^0 in D^0, that is, there exist $\omega > 0$ and W^0 in $\mathscr{C}_0^{L+\mathrm{Lip}}$ such that equation (2.15) is solved. Moreover, by the contraction argument, we know that the solution is unique. Therefore, ω is unique, W^0 is unique in the $\mathscr{C}_0^{L+\mathrm{Lip}}$ space for the fixed phase $W_1^0(0) = 0$.

Now we prove the a posteriori estimation part of Theorem 6. Since Γ^0 is a contraction on D^0, we know that

$$\begin{aligned} d\big((\omega^0, W^{0,0}), (\omega, W^0)\big) &= \lim_{k\to\infty} d\big((\omega^0, W^{0,0}), (\Gamma^0)^k(\omega^0, W^{0,0})\big) \\ &\leqslant \sum_{k=0}^{\infty} (\mu_0)^k d\big((\omega^0, W^{0,0}), \Gamma^0(\omega^0, W^{0,0})\big) \\ &\leqslant \frac{1}{1-\mu_0} d\big((\omega^0, W^{0,0}), \Gamma^0(\omega^0, W^{0,0})\big). \end{aligned} \tag{2.41}$$

It remains to estimate $d\big((\omega^0, W^{0,0}), \Gamma^0(\omega^0, W^{0,0})\big)$ by $\|E^0\|$, where the norm is the maximum norm defined in equation (2.30). We have

$$E^0(\theta) = \omega^0 \frac{d}{d\theta} W^{0,0}(\theta) - \begin{pmatrix} \omega_0 \\ \lambda_0 W_2^{0,0}(\theta) \end{pmatrix} - \varepsilon Y(W^{0,0}(\theta), \widetilde{W}^{0,0}(\theta; \omega^0), \varepsilon),$$

that is,

$$\begin{pmatrix} E_1^0(\theta) \\ E_2^0(\theta) \end{pmatrix} = \begin{pmatrix} \omega^0 \frac{d}{d\theta} W_1^{0,0}(\theta) - \omega_0 - \varepsilon \overline{Y}_1(W^{0,0}(\theta), \widetilde{W}^{0,0}(\theta;\omega^0), \varepsilon) \\ \omega^0 \frac{d}{d\theta} W_2^{0,0}(\theta) - \lambda_0 W_2^{0,0}(\theta) - \varepsilon \overline{Y}_2(W^{0,0}(\theta), \widetilde{W}^{0,0}(\theta;\omega^0), \varepsilon) \end{pmatrix},$$

and,

$$\begin{aligned}
&d\big((\omega^0, W^{0,0}), \Gamma^0(\omega^0, W^{0,0})\big) \\
&\quad \leqslant \left|\omega_0 + \varepsilon \int_0^1 \overline{Y}_1(W^{0,0}(\theta), \widetilde{W}^{0,0}(\theta;\omega^0), \varepsilon) d\theta - \omega^0\right| \\
&\qquad + \sup_\theta \left| \frac{1}{\Gamma_1^0(\omega^0, W^0)} \big(\omega_0 \theta + \varepsilon \int_0^\theta \overline{Y}_1(W^{0,0}(\sigma), \widetilde{W}^{0,0}(\sigma;\omega^0), \varepsilon) d\sigma\big) - W_1^{0,0}(\theta)\right| \\
&\qquad + \sup_\theta \left| \varepsilon \int_0^\infty e^{\lambda_0 t} \overline{Y}_2(W^{0,0}(\theta - \omega^0 t), \widetilde{W}^{0,0}(\theta - \omega^0 t;\omega^0), \varepsilon) dt - W_2^{0,0}(\theta)\right| \\
&\quad \leqslant \left|\int_0^1 E_1^0(\theta) d\theta\right| + \left|\int_0^\infty e^{\lambda_0 t} E_2^0(\theta - \omega^0 t) dt\right| \\
&\qquad + \frac{1}{|\Gamma_1^0(\omega^0, W^0)|} \left(\left|\int_0^\theta E_1^0(\sigma) d\sigma\right| + \|W_1^{0,0}\| \left|\int_0^1 E_1^0(\theta) d\theta\right| \right) \\
&\quad \leqslant (1 + \frac{2B^0}{|\omega_0|}) \left|\int_0^1 E_1^0(\theta) d\theta\right| + \frac{2}{|\omega_0|} \left|\int_0^\theta E_1^0(\sigma) d\sigma\right| + \left|\int_0^\infty e^{\lambda_0 t} E_2^0(\theta - \omega^0 t) dt\right|
\end{aligned}$$

For $\theta \in [0,1]$, we have

$$d\big((\omega^0, W^{0,0}), \Gamma^0(\omega^0, W^{0,0})\big) \leqslant \left(1 + \frac{2 + 2B^0}{|\omega_0|}\right) \|E_1^0\| - \frac{1}{\lambda_0} \|E_2^0\|.$$

Combine this with the inequality equation (2.41), we have

$$d\big((\omega^0, W^{0,0}), (\omega, W^0)\big) \leqslant \frac{1}{1 - \mu_0} \left[\left(1 + \frac{2 + 2B^0}{|\omega_0|}\right) \|E_1^0\|_{C_0} - \frac{1}{\lambda_0} \|E_2^0\|_{C_0} \right]. \tag{2.42}$$

By definition of the norm, equation (2.20) and $l = 0$ case of equation (2.19) are true for a constant C, which depends on ε, B^0, ω_0, λ_0.

For other values of l, one can use interpolation inequality in Lemma 91, to get

$$\begin{aligned} \|W_1^{0,0} - W_1^0\|_{C^l} &\leqslant c(l,L)\|W_1^{0,0} - W_1^0\|_{C^0}^{1-\frac{l}{L}} \|W_1^{0,0} - W_1^0\|_{C^L}^{\frac{l}{L}} \\ &\leqslant c(l,L)\|W_1^{0,0} - W_1^0\|_{C^0}^{1-\frac{l}{L}} (2B^0)^{\frac{l}{L}}. \end{aligned} \tag{2.43}$$

Similar estimates can be done for the second component, this finishes the proof of the estimations in theorem 6.

For solution of the equation (2.15), note that $K \circ W^0(\theta + \omega t)$ solves the equation (2.4):

$$\frac{d}{dt} K \circ W^0(\theta + \omega t) = X(K \circ W^0(\theta + \omega t), K \circ W^0(\theta + \omega(t - r(K \circ W^0(\theta + \omega t))))).$$

If W^0 is L times differentiable, then right hand side of above equation is L times differentiable, so is the left hand side. Using the fact that K is an analytic local diffeomorphism, one can conclude that W^0 is (L+1) times differentiable. A bootstrap argument can be used to see W^0 is differentiable up to any order.

2.5.2 Proof of Theorem 7

With Theorem 6, ω and W^0 are known to us. To prove Theorem 7, we have to consider the equations for the first order term, j-th order term, and then higher order term in s. We will obtain λ, W^1 solving the first order equation (2.16), W^j solving equation (2.17), and then $W^>$ which solves equation (2.18).

2.5.2.1 First-order Equation

Recall that for the first order term, we got an invariance equation (2.16), see also below:

$$\omega \frac{d}{d\theta} W^1(\theta) + \lambda W^1(\theta) - \begin{pmatrix} 0 \\ \lambda_0 W_2^1(\theta) \end{pmatrix} = \varepsilon \overline{Y}^1(\theta, \lambda, W^0, W^1, \varepsilon),$$

where

$$\overline{Y}^1(\theta, \lambda, W^0, W^1, \varepsilon) = A(\theta)W^1(\theta) + B(\theta; \lambda)W^1(\theta - \omega\overline{r \circ K}(W^0(\theta))), \tag{2.44}$$

$$\begin{aligned} A(\theta) = &-\omega D_2\overline{Y}(W^0(\theta), \widetilde{W}^0(\theta), \varepsilon)DW^0(\theta - \omega\overline{r \circ K}(W^0(\theta)))D(\overline{r \circ K})(W^0(\theta)) \\ &+ D_1\overline{Y}(W^0(\theta), \widetilde{W}^0(\theta), \varepsilon), \end{aligned} \tag{2.45}$$

and

$$B(\theta; \lambda) = e^{-\lambda\overline{r \circ K}(W^0(\theta))}D_2\overline{Y}(W^0(\theta), \widetilde{W}^0(\theta), \varepsilon).$$

Note that in the expression of A and B above, we suppressed the ω in the expression of $\widetilde{W}^0$. We do this to simplify the notation, since ω is already known from Theorem 6.

Remark 18. *Since $\overline{Y}^1(\theta, \lambda, W^0, W^1, \varepsilon)$ in equation* (2.44) *is linear in W^1, equation* (2.16) *for W^1, is linear and homogenous in W^1. Hence if $W^1(\theta)$ solves equation* (2.16)*, so does any scalar multiple of $W^1(\theta)$.*

Componentwise, we have the following two equations:

$$\omega\frac{d}{d\theta}W_1^1(\theta) + \lambda W_1^1(\theta) = \varepsilon\overline{Y}_1^1(\theta, \lambda, W^0, W^1, \varepsilon), \tag{2.46}$$

$$\omega\frac{d}{d\theta}W_2^1(\theta) + (\lambda - \lambda_0)W_2^1(\theta) = \varepsilon\overline{Y}_2^1(\theta, \lambda, W^0, W^1, \varepsilon). \tag{2.47}$$

As already pointed out, for the unperturbed case, W could be chosen as the identity map. So after adding a small perturbation, $W^1(\theta) \approx \binom{0}{1}$. We will be able to find a unique W^1 close to $\binom{0}{1}$ solving above equation (2.16), by considering the following normalization:

$$\int_0^1 W_2^1(\theta)d\theta = 1. \tag{2.48}$$

Remark 19. *It is natural to choose above normalization equation* (2.48)*, since under small perturbation, we have $W^1(\theta) \approx \binom{0}{1}$. Meanwhile, one can show that λ does not depend on* *the choice of normalization as long as* $\int_0^1 W_2^1(\theta)d\theta \neq 0$.

From now on, since W^0 is already known to us, we will omit W^0 from $\overline{Y}^1(\theta,\lambda,W^0,W^1,\varepsilon)$, and denote it as $\overline{Y}^1(\theta,\lambda,W^1,\varepsilon)$. We define an operator Γ^1 as follows:

$$\begin{aligned}
\Gamma^1\begin{pmatrix} b \\ F_1 \\ F_2 \end{pmatrix}(\theta) &= \begin{pmatrix} \Gamma_1^1(b,F) \\ \Gamma_2^1(b,F)(\theta) \\ \Gamma_3^1(b,F)(\theta) \end{pmatrix} \\
&= \begin{pmatrix} \lambda_0 + \varepsilon\int_0^1 \overline{Y}_2^1(\theta,b,F,\varepsilon)d\theta \\ -\varepsilon\int_0^\infty e^{bt}\overline{Y}_1^1(\theta+\omega t,b,F,\varepsilon)dt \\ C(b,F) + \frac{\varepsilon}{\omega}\int_0^\theta \overline{Y}_2^1(\sigma,b,F,\varepsilon) - (\int_0^1 \overline{Y}_2^1(\theta,b,F,\varepsilon)d\theta)F_2(\sigma)d\sigma \end{pmatrix},
\end{aligned} \tag{2.49}$$

where

$$\begin{aligned}
C(b,F) = 1 &- \frac{\varepsilon}{\omega}\int_0^1\int_0^\theta \overline{Y}_2^1(\sigma,b,F,\varepsilon)d\sigma d\theta \\
&+ \frac{\varepsilon}{\omega}(\int_0^1 \overline{Y}_2^1(\theta,b,F,\varepsilon)d\theta)\int_0^1\int_0^\theta F_2(\sigma)d\sigma d\theta
\end{aligned} \tag{2.50}$$

is a constant chosen to ensure that $\Gamma_3^1(b,F)$ also satisfies the normalization condition equation (2.48), i.e. $\int_0^1 \Gamma_3^1(b,F)(\theta)d\theta = 1$.

Similar to previous section, section 2.5.1, for the domain of Γ^1, we consider the closed interval $I^1 = \{b : |b-\lambda_0| \leqslant \frac{|\lambda_0|}{3}\}$, as well as the function space

$$\begin{aligned}
\mathscr{C}_1^{L-1+\mathrm{Lip}} = \{f \mid f\colon \mathbb{T}\to\mathbb{T}\times\mathbb{R},\ &f \text{ can be lifted to a function from } \mathbb{R} \text{ to } \mathbb{R}^2, \\
&\text{still denoted as } f, \text{which satisfies } f(\theta+1)=f(\theta), \\
&\|f\|_{L-1+\mathrm{Lip}} \leqslant B^1, \text{and } \int_0^1 f_2(\theta)d\theta = 1\},
\end{aligned}$$

where

$$\|f\|_{L-1+\mathrm{Lip}} = \max_{i=1,2,k=0,\ldots,L-1}\left\{\sup_{\theta\in[0,1]}\|f_i^{(k)}(\theta)\|, \mathrm{Lip}(f_i^{(L-1)})\right\},$$

L is the same as in Theorem 6, and B^1 is a positive constant.

Let $D^1 \coloneqq I^1 \times \mathscr{C}_1^{L-1+\mathrm{Lip}}$ be the domain of Γ^1. We have the following:

Lemma 20. *If ε is small enough, $\Gamma^1(D^1) \subset D^1$.*

Proof. Since $\overline{Y}_2^1(\theta, b, F, \varepsilon)$ is bounded, for small ε, we have $\Gamma_1^1(b, F) \in I^1$.

Now consider $\Gamma_2^1(b, F)(\theta)$, we first have to show that

$$\Gamma_2^1(b, F)(\theta + 1) = \Gamma_2^1(b, F)(\theta).$$

This follows from the fact that $\overline{Y}_1^1(\theta + 1, b, F, \varepsilon) = \overline{Y}_1^1(\theta, b, F, \varepsilon)$, which is true by periodicity of W^0 as in equation (2.12), of F, and of $\overline{r \circ K}$ with respect to its first component.

Now we check $\frac{d^n}{d\theta^n}\Gamma_2^1(b, F)(\theta)$, $0 \leqslant n \leqslant L-1$, is bounded. Notice that

$$\frac{d^n}{d\theta^n}\Gamma_2^1(b, F)(\theta) = -\varepsilon \int_0^\infty e^{bt} \frac{\partial^n}{\partial\theta^n}\overline{Y}_1^1(\theta + \omega t, b, F, \varepsilon)dt.$$

By dominated convergence theorem, it suffices to check that $\frac{\partial^n}{\partial\theta^n}\overline{Y}_1^1(\theta + \omega t, b, F, \varepsilon)$ is bounded. Using Faà di Bruno's formula in Lemma 86, boundedness of $\frac{\partial^n}{\partial\theta^n}\overline{Y}_1^1(\theta+\omega t, b, F, \varepsilon)$ is ensured by assumptions on $\overline{Y}$, $\overline{r \circ K}$, and W^0, as well as $F \in \mathscr{C}_1^{L-1+\mathrm{Lip}}$. Then for ε small enough, the derivatives can be bounded by B^1. Bound for Lipschitz constant of $\frac{d^{L-1}}{d\theta^{L-1}}\Gamma_2^1(b, F)(\theta)$ also follows.

For $\Gamma_3^1(b, F)(\theta)$, we first show that it is periodic. Notice that

$$\frac{d}{d\theta}\Gamma_3^1(b, F)(\theta) = \frac{\varepsilon}{\omega}\overline{Y}_2^1(\theta, b, F, \varepsilon) - \frac{\varepsilon}{\omega}\left(\int_0^1 \overline{Y}_2^1(\theta, b, F, \varepsilon)d\theta\right) F_2(\theta) \tag{2.51}$$

is periodic. Hence, to show periodicity of $\Gamma_3^1(b, F)(\theta)$, it suffices to see that $\Gamma_3^1(b, F)(0) = \Gamma_3^1(b, F)(1)$, which is true because $\int_0^1 F_2(\theta)d\theta = 1$. The choice of the constant $C(b, F)$ ensures that the normalization condition $\int_0^1 \Gamma_3^1(b, F)(\theta)d\theta = 1$ is also verified.

Taking derivatives of equation (2.51), we have for $2 \leqslant n \leqslant L-1$

$$\frac{d^n}{d\theta^n}\Gamma_3^1(b, F)(\theta) = \frac{\varepsilon}{\omega}\left(\frac{d^{(n-1)}}{d\theta^{(n-1)}}\overline{Y}_2^1(\theta, b, F, \varepsilon) - \left(\int_0^1 \overline{Y}_2^1(\theta, b, F, \varepsilon)d\theta\right)\frac{d^{(n-1)}}{d\theta^{(n-1)}}F_2(\theta)\right),$$

which will be $\frac{\varepsilon}{\omega}$ multiplied by bounded functions due to the assumptions on $\overline{Y}$, $\overline{r \circ K}$, and W^0, as well as $F \in \mathscr{C}_1^{L-1+\mathrm{Lip}}$. When ε is small, they will all be bounded by B^1. Similar for Lipschitz constant of $\frac{d^{L-1}}{d\theta^{L-1}}\Gamma_3^1(b,F)(\theta)$.

Hence for ε small enough, where the smallness condition depends on bounds of the derivatives of $\overline{Y}$, $\overline{r \circ K}$, B^0, and B^1, but not on the specific choice of $(b,F) \in D^1$, we have that $(\Gamma_2^1(b,F), \Gamma_3^1(b,F)) \in \mathscr{C}_1^{L-1+\mathrm{Lip}}$. This finishes the proof. □

Recall the distance introduced in equation (3.22):

$$d((a,Z),(a',Z')) = |a-a'| + \|Z-Z'\|,$$

where

$$\|Z-Z'\| = \max\left\{\sup_\theta |Z_1(\theta) - Z_1'(\theta)|, \sup_\theta |Z_2(\theta) - Z_2'(\theta)|\right\}.$$

Lemma 21. *Under above definition of distance on D^1, for small enough ε, Γ^1 is a contraction.*

Proof. We will show that for ε small enough, we can find a constant $0 < \mu_1 < 1$ such that

$$d(\Gamma^1(b,F), \Gamma^1(b',F')) < \mu_1 \cdot d((b,F),(b',F')). \tag{2.52}$$

Note that

$$
\begin{aligned}
&d(\Gamma^1(b,F),\Gamma^1(b',F'))\\
&\quad\leqslant \varepsilon\left|\int_0^1 \overline{Y}_2^1(\theta,b,F,\varepsilon)-\overline{Y}_2^1(\theta,b',F',\varepsilon)d\theta\right|\\
&\qquad+\varepsilon\sup_\theta\left|\int_0^\infty e^{bt}\overline{Y}_1^1(\theta+\omega t,b,F,\varepsilon)-e^{b't}\overline{Y}_1^1(\theta+\omega t,b',F',\varepsilon)dt\right|\\
&\qquad+\frac{\varepsilon}{|\omega|}\sup_\theta\left|\int_0^\theta \overline{Y}_2^1(\sigma,b,F,\varepsilon)-\big(\int_0^1\overline{Y}_2^1(\theta,b,F,\varepsilon)d\theta\big)F_2(\sigma)d\sigma\right.\\
&\qquad\qquad\left.-\int_0^\theta \overline{Y}_2^1(\sigma,b',F',\varepsilon)+\big(\int_0^1\overline{Y}_2^1(\theta,b',F',\varepsilon)d\theta\big)F_2'(\sigma)d\sigma\right|\\
&\qquad+|C(F,b)-C(F',b')|
\end{aligned}
\tag{2.53}
$$

As before, we will consider each term of the right hand side of the above inequality equation (2.53).

Recall that $\overline{Y}^1$ has the form equation (2.44)

$$\overline{Y}^1(\theta,\lambda,W^1,\varepsilon)=A(\theta)W^1(\theta)+B(\theta;\lambda)W^1(\theta-\omega\overline{r\circ K}(W^0(\theta))).$$

If we use notation:

$$A(\theta)=\begin{pmatrix}A_{11}(\theta) & A_{12}(\theta)\\ A_{21}(\theta) & A_{22}(\theta)\end{pmatrix},\qquad B(\theta;\lambda)=\begin{pmatrix}B_{11}(\theta;\lambda) & B_{12}(\theta;\lambda)\\ B_{21}(\theta;\lambda) & B_{22}(\theta;\lambda)\end{pmatrix},$$

then

$$
\begin{aligned}
\overline{Y}_1^1(\theta,\lambda,W^1,\varepsilon)=&A_{11}(\theta)W_1^1(\theta)+A_{12}(\theta)W_2^1(\theta)\\
&+B_{11}(\theta;\lambda)W_1^1(\theta-\omega\overline{r\circ K}(W^0(\theta)))\\
&+B_{12}(\theta;\lambda)W_2^1(\theta-\omega\overline{r\circ K}(W^0(\theta))),
\end{aligned}
$$

and

$$\begin{aligned}\overline{Y}_2^1(\theta, \lambda, W^1, \varepsilon) =& A_{21}(\theta)W_1^1(\theta) + A_{22}(\theta)W_2^1(\theta) \\ &+ B_{21}(\theta;\lambda)W_1^1(\theta - \omega\overline{r \circ K}(W^0(\theta))) \\ &+ B_{22}(\theta;\lambda)W_2^1(\theta - \omega\overline{r \circ K}(W^0(\theta))).\end{aligned}$$

We estimate

$$|B(\theta;b)| \leqslant e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|D_2\overline{Y}\|,$$

and

$$|B(\theta;b) - B(\theta;b')| \leqslant \|D_2\overline{Y}\|e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|\overline{r \circ K}\||b - b'|.$$

Also, if we define $\|A\| = \max_\theta \|A(\theta)\|$, where $\|A(\theta)\|$ is the operator norm corresponding to the maximum norm $\|\cdot\|$ defined in equation (2.30). Then,

$$\begin{aligned}&|\overline{Y}_1^1(\theta, b, F, \varepsilon) - \overline{Y}_1^1(\theta, b', F', \varepsilon)| \\ &\quad\leqslant \|A\|\|F - F'\| + \|B(\theta;b)\|\|F - F'\| + \|B(\theta;b) - B(\theta;b')\|\|F'\| \\ &\quad\leqslant (\|A\| + e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|D_2\overline{Y}\|)\|F - F'\| + B^1\|D_2\overline{Y}\|e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|\overline{r \circ K}\||b - b'|,\end{aligned}$$

and similarly,

$$\begin{aligned}&|\overline{Y}_2^1(\theta, b, F, \varepsilon) - \overline{Y}_2^1(\theta, b', F', \varepsilon)| \\ &\quad\leqslant (\|A\| + e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|D_2\overline{Y}\|)\|F - F'\| + B^1\|D_2\overline{Y}\|e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|\overline{r \circ K}\||b - b'|.\end{aligned}$$

Note also that

$$|\overline{Y}_1^1(\theta, b, F, \varepsilon)| \leqslant B^1(\|A\| + e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|D_2\overline{Y}\|),$$

similarly,

$$|\overline{Y}_2^1(\theta, b, F, \varepsilon)| \leqslant B^1(\|A\| + e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|D_2\overline{Y}\|).$$

Now for the first term in equation (2.53), we have

$$\begin{aligned}\left|\Gamma_1^1(b,F) - \Gamma_1^1(b',F')\right| \leqslant \varepsilon(\|A\| + e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|D_2\overline{Y}\|)\|F - F'\| \\ + \varepsilon B^1\|D_2\overline{Y}\|e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|\overline{r\circ K}\||b-b'|.\end{aligned}$$

For the second term in equation (2.53), we have for all θ,

$$\begin{aligned}&\left|\Gamma_2^1(b,F) - \Gamma_2^1(b',F')\right| \leqslant \\ &\quad -\frac{3\varepsilon}{2\lambda_0}(\|A\| + e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|D_2\overline{Y}\|)\|F-F'\| \\ &\quad -\frac{3B^1\varepsilon}{2\lambda_0}\left(e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|D_2\overline{Y}\|\big(\|\overline{r\circ K}\| - \frac{3}{2\lambda_0}\big) - \frac{3}{2\lambda_0}\|A\|\right)|b-b'|\end{aligned}$$

For the third term in equation (2.53), we have

$$\begin{aligned}\left|\Gamma_3^1(b,F) - \Gamma_3^1(b',F')\right| \leqslant &\frac{\varepsilon}{|\omega|}(1+2B^1)(\|A\| + e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|D_2\overline{Y}\|)\|F-F'\| \\ &+ \frac{B^1\varepsilon}{|\omega|}(1+B^1)\|D_2\overline{Y}\|e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|\overline{r\circ K}\||b-b'|\end{aligned}$$

Similar holds for the last part in equation (2.53),

$$\begin{aligned}|C(F,b) - C(F',b')| \leqslant &\frac{\varepsilon}{|\omega|}(1+2B^1)(\|A\| + e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|D_2\overline{Y}\|)\|F-F'\| \\ &+ \frac{B^1\varepsilon}{|\omega|}(1+B^1)\|D_2\overline{Y}\|e^{-\frac{4}{3}\lambda_0\|\overline{r\circ K}\|}\|\overline{r\circ K}\||b-b'|\end{aligned}$$

Combine all the estimations above, we can find constants c_1, c_2 such that,

$$d(\Gamma^1(b,F), \Gamma^1(b',F')) \leqslant \varepsilon(c_1|b-b'| + c_2\|F-F'\|).$$

Therefore, for small enough ε, we have that Γ^1 is a contraction, i.e., we can find a constant μ_1 such that equation (2.52) is true. $\square$

Taking any initial guess $(\lambda^0, W^{1,0}) \in D^1$, we could take $\lambda^0 = \lambda_0$ and $W^{1,0}(\theta) = \binom{0}{1}$, the sequence $(\Gamma^1)^n(\lambda^0, W^{1,0})$ has a limit in D^1, which we denote by (λ, W^1). (λ, W^1) is the fixed point of operator Γ^1, hence it solves equation (2.16). Since the operator is a contraction, λ is unique, W^1 is unique in C^0 sense under the normalization condition equation (2.48).

Similar to what we have done in estimation equation (2.41) in section 2.5.1, notice that

$$d\big((\lambda^0, W^{1,0}), (\lambda, W^1)\big) \leqslant \frac{1}{1-\mu_1} d\big((\lambda^0, W^{1,0}), \Gamma^1(\lambda^0, W^{1,0})\big). \tag{2.54}$$

We will estimate $d\big((\lambda^0, W^{1,0}), \Gamma^1(\lambda^0, W^{1,0})\big)$ by $\|E^1\|$. If we write $E^1(\theta)$ in matrix form, we have

$$\begin{pmatrix} E_1^1(\theta) \\ E_2^1(\theta) \end{pmatrix} = \begin{pmatrix} \omega \frac{d}{d\theta} W_1^{1,0}(\theta) + \lambda^0 W_1^{1,0}(\theta) - \varepsilon \overline{Y}_1^1(\theta, \lambda^0, W^{1,0}, \varepsilon) \\ \omega \frac{d}{d\theta} W_2^{1,0}(\theta) + (\lambda^0 - \lambda_0) W_2^{1,0}(\theta) - \varepsilon \overline{Y}_2^1(\theta, \lambda^0, W^{1,0}, \varepsilon) \end{pmatrix}.$$

Therefore,

$$
\begin{aligned}
d\big((\lambda^0, W^{1,0}), &\Gamma^1(\lambda^0, W^{1,0})\big) \\
\leqslant & |\lambda_0 + \varepsilon \int_0^1 \overline{Y}_2^1(\theta, \lambda^0, W^{1,0}, \varepsilon) d\theta - \lambda^0| \\
& + \sup_\theta \left| W_1^{1,0}(\theta) + \varepsilon \int_0^\infty e^{\lambda^0 t} \overline{Y}_1^1(\theta + \omega t, \lambda^0, W^{1,0}, \varepsilon) dt \right| \\
& + \sup_\theta \Bigg| C(\lambda^0, W^{1,0}) + \frac{\varepsilon}{\omega} \int_0^\theta \overline{Y}_2^1(\sigma, \lambda^0, W^{1,0}, \varepsilon) \\
& \qquad - \left(\int_0^1 \overline{Y}_2^1(\theta, \lambda^0, W^{1,0}, \varepsilon) d\theta \right) W_2^{1,0}(\sigma) d\sigma - W_2^{1,0}(\theta) \Bigg| \\
\leqslant & \left| \int_0^1 E_2^1(\theta) d\theta \right| + \left| \int_0^\infty e^{\lambda^0 t} E_1^1(\theta + \omega t) dt \right| + \frac{2 + 2B^1}{|\omega|} \|E_2^1\| \\
\leqslant & \frac{1}{|\lambda^0|} \|E_1^1\| + \left(1 + \frac{2 + 2B^1}{|\omega|} \right) \|E_2^1\| \\
\leqslant & \frac{3}{2|\lambda_0|} \|E_1^1\| + \left(1 + \frac{4 + 4B^1}{\omega_0} \right) \|E_2^1\|.
\end{aligned}
$$

Then

$$
d\big((\lambda^0, W^{1,0}), (\lambda, W^1)\big) \leqslant \frac{1}{1 - \mu_1} \left[\frac{3}{2|\lambda_0|} \|E_1^1\| + \left(1 + \frac{4 + 4B^1}{\omega_0} \right) \|E_2^1\| \right]. \tag{2.55}
$$

Therefore, we can find a constant C, depending on ε, B^1, ω_0 and λ_0 such that $|\lambda - \lambda^0| \leqslant C\|E^1\|$. This proves equation (2.23).

2.5.2.2 Equation for jth order terms

For each $j \geqslant 2$, we can proceed in a similar manner to find W^j. With ω, λ, W^0, and W^1 known, equations for W^j's are easier to analyze.

Remark 22. *As we will see, for theoretical result, we can stop at order 1 and start to deal with the higher order term. We include here the discussion for W^j's for numerical interests.*

Assume now that we have already obtained $W^0, \dots, W^{j-1}$, and ω, λ, we are going to find $W^j(\theta)$. To obtain the invariance equation satisfied by W^j, mentioned in equation

(2.17). We consider the j-th order term in equation (2.9). Note that there are only two terms in the coefficient of s^j in $\widetilde{W}(\theta, s)$ which contain W^j:

$$-\omega DW^0(\theta - \omega\overline{r \circ K}(W^0(\theta)))D(\overline{r \circ K})(W^0(\theta))W^j(\theta),$$

and

$$e^{-\lambda j \overline{r\circ K}(W^0(\theta))}W^j(\theta - \omega\overline{r \circ K}(W^0(\theta))).$$

Therefore, $\overline{Y}^j$ is of the form:

$$\overline{Y}^j(\theta, \lambda, W^0, W^j, \varepsilon) = A(\theta)W^j(\theta) + B_j(\theta)W^j(\theta - \omega\overline{r \circ K}(W^0(\theta))), \tag{2.56}$$

where $A(\theta)$ is the same as in equation (2.45),

$$\begin{aligned} A(\theta) = &-\omega D_2\overline{Y}(W^0(\theta), \widetilde{W}(\theta), \varepsilon)DW^0(\theta - \omega\overline{r \circ K}(W^0(\theta))D(\overline{r \circ K})(W^0(\theta)) \\ &+ D_1\overline{Y}(W^0(\theta), \widetilde{W}(\theta), \varepsilon), \end{aligned}$$

and

$$B_j(\theta) := e^{-\lambda j \overline{r\circ K}(W^0(\theta))}D_2\overline{Y}(W^0(\theta), \widetilde{W}^0(\theta), \varepsilon).$$

We also note that $R^j(\theta)$ will be some expression in the derivatives of $\overline{Y}$ evaluated at $(W^0(\theta), \widetilde{W}(\theta), \varepsilon)$, multiplied with $W^0, \dots, W^{j-1}$. Therefore, $R^j(\theta)$ will have the same regularity as W^{j-1}. We will show inductively by the following argument that W^j is $(L-1)$ times differentiable with $(L-1)$-th derivative Lipschitz.

From now on, we will write $\overline{Y}^j$ as $\overline{Y}^j(\theta, W^j, \varepsilon)$, for that λ and W^0 are known to us. Componentwisely, W^j should satisfy

$$\omega\frac{d}{d\theta}W_1^j(\theta) + \lambda j W_1^j(\theta) = \varepsilon\overline{Y}_1^j(\theta, W^j, \varepsilon) + R_1^j(\theta), \tag{2.57}$$

$$\omega\frac{d}{d\theta}W_2^j(\theta) + (\lambda j - \lambda_0)W_2^j(\theta) = \varepsilon\overline{Y}_2^j(\theta, W^j, \varepsilon) + R_2^j(\theta). \tag{2.58}$$

Consider functions in the space

$$
\begin{aligned}
\mathscr{C}_j^{L-1+\mathrm{Lip}} = \{f \mid f\colon \mathbb{T} \to \mathbb{T}\times\mathbb{R},\ & f \text{ can be lifted to a function from } \mathbb{R} \text{ to } \mathbb{R}^2,\\
& \text{still denoted as } f, \text{which satisfies } f(\theta+1) = f(\theta),\\
& \|f\|_{L-1+\mathrm{Lip}} \leqslant B^j\},
\end{aligned}
$$

where B^j's are positive constants, and

$$
\|f\|_{L-1+\mathrm{Lip}} = \max_{i=1,2,k=0,\ldots,L-1}\{\sup_{\theta\in[0,1]}\|f_i^{(k)}(\theta)\|, Lip(f_i^{(L-1)})\}.
$$

Similar to what we have done above, define an operator on space $\mathscr{C}_j^{L-1+\mathrm{Lip}}$:

$$
\Gamma^j(G)(\theta) = \begin{pmatrix} -\varepsilon\int_0^\infty e^{\lambda jt}\left(\overline{Y}_1^j(\theta+\omega t, G, \varepsilon) + R_1^j(\theta+\omega t)\right)dt \\ -\varepsilon\int_0^\infty e^{(\lambda j-\lambda_0)t}\left(\overline{Y}_2^j(\theta+\omega t, G, \varepsilon) + R_2^j(\theta+\omega t)\right)dt \end{pmatrix} \tag{2.59}
$$

Assume that we have already obtained W^k in $\mathscr{C}_k^{L-1+\mathrm{Lip}}$ for $k = 0,\ldots,j-1$, we have the following:

Lemma 23. *For small enough ε, we have $\Gamma^j(\mathscr{C}_j^{L-1+\mathrm{Lip}}) \subset \mathscr{C}_j^{L-1+\mathrm{Lip}}$.*

This follows from $\lambda < 0$ and $(\lambda j - \lambda_0) < 0$ for $j \geqslant 2$ and the regularity of $W^0,\ldots,W^j$, $\overline{Y}^j$, and R^j. Moreover, we have ε in front of the expression. Since this is very similar to the analysis of W^0 and W^1, we will omit the detailed proof here.

We also know that Γ^j is a C^0 contraction for small ε.

Lemma 24. *For small enough ε, Γ^j is a contraction in C^0 distance.*

This follows easily from that $\lambda < 0$ and $(\lambda j - \lambda_0) < 0$ for $j \geqslant 2$, and $\overline{Y}^j$ is linear in W^j.

If we define norm as before

$$
\|G\| = \max\{\sup_\theta |G_1(\theta)|, \sup_\theta |G_2(\theta)|\},
$$

above lemma tells us that, if ε is small enough, then one can find $0 < \mu_j < 1$, such that

$$\|\Gamma(G) - \Gamma(G')\| \leqslant \mu_j \|G - G'\|.$$

Taking any initial guess $W^{j,0} \in \mathscr{C}_j^{L-1+\mathrm{Lip}}$, we would take $W^{j,0}(\theta) = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$, the sequence $(\Gamma^j)^n(W^{j,0})$ has a limit in $\mathscr{C}_j^{L-1+\mathrm{Lip}}$, we denote it by W^j. W^j is the fixed point of operator Γ^j, so it solves equation (2.17). W^j is close to the initial guess, and is unique in the sense of C^0 by the contraction argument. We will see quantitative estimates below.

We know that

$$\|W^j - W^{j,0}\| \leqslant \frac{1}{1-\mu_j} \|W^{j,0} - \Gamma^j(W^{j,0})\|. \tag{2.60}$$

With similar argument as in the error estimation of W^0 and W^1, we have

$$\begin{aligned} |W_1^{j,0}(\theta) - \Gamma_1^j(W^{j,0})(\theta)| &\leqslant -\frac{1}{j\lambda}\|E_1^j\|, \\ |W_2^{j,0}(\theta) - \Gamma_2^j(W^{j,0})(\theta)| &\leqslant -\frac{1}{j\lambda - \lambda_0}\|E_2^j\|. \end{aligned}$$

Therefore, we have

$$\|W^j - W^{j,0}\| \leqslant \frac{1}{1-\mu_j}\left(-\frac{1}{j\lambda}\|E_1^j\| - \frac{1}{j\lambda - \lambda_0}\|E_2^j\|\right) \leqslant C\|E^j\|. \tag{2.61}$$

We stress that above C depends on j, ε, λ, B^j, and the SDDE, however, it does not depend on the choice of $W^{j,0}$ in space $\mathscr{C}_j^{L-1+\mathrm{Lip}}$.

2.5.2.3 *Equation of Higher Order Term*

Now we have already found ω, λ, $W^0, \ldots, W^{N-1}$. It remains to consider the higher order term. We will solve equation (2.18) locally in this section, which will establish the existence in Theorem 7. From now on, we will write:

$$W(\theta, s) = W^{\leqslant}(\theta, s) + W^{>}(\theta, s), \tag{2.62}$$

where $W^{\leqslant}(\theta, s) = \sum_{j=0}^{N-1} W^j(\theta)s^j$. To make the analysis feasible, we do a cut-off to the equation satisfied by $W^>$ in equation (2.18):

$$(\omega\partial_\theta + s\lambda\partial_s)W^>(\theta, s) = \begin{pmatrix} 0 \\ \lambda_0 W_2^>(\theta, s) \end{pmatrix} + \varepsilon Y^>(W^>, \theta, s, \varepsilon)\phi(s), \tag{2.63}$$

where

$$Y^>(W^>, \theta, s, \varepsilon) = \overline{Y}(W(\theta, s), \widetilde{W}(\theta, s), \varepsilon) - \sum_{i=0}^{N-1} \overline{Y}^i(\theta)s^i, \tag{2.64}$$

$$\overline{Y}^i(\theta) = \frac{1}{i!}\frac{\partial^i}{\partial s^i}(\overline{Y}(W(\theta, s), \widetilde{W}(\theta, s), \varepsilon))\Big|_{s=0},$$

and recall the C^∞ cut-off function $\phi : \mathbb{R} \to [0, 1]$ introduced in equation (2.10):

$$\phi(x) = \begin{cases} 1 & \text{if} \quad |x| \leqslant \frac{1}{2}, \\ 0 & \text{if} \quad |x| > 1. \end{cases}$$

Remark 25. *A cut-off is needed in our method. We note that similar to before, the boundaries for the cut-off function above (*$\frac{1}{2}$ *and* 1*) could be changed to any positive numbers* $a_1 < a_2$.

Adding a cut-off is not too restrictive. Indeed, we only get local results for the original problem near the limit cycle. Since we have used extensions to get the prepared equation (2.9)*, what happens for* s *with large absolute value will not matter.*

Now let $c(t) = (\theta + \omega t, se^{\lambda t})$ be the characteristics, we define an operator:

$$\Gamma^>(H)(\theta, s) = -\varepsilon \int_0^\infty \begin{pmatrix} 1 & 0 \\ 0 & e^{-\lambda_0 t} \end{pmatrix} Y^>(H, c(t), \varepsilon)\phi(se^{\lambda t})dt. \tag{2.65}$$

If there is a fixed point of $\Gamma^>$ which has some regularity, it will solve the modified invariance equation (2.63). For the domain of $\Gamma^>$, assuming that $L^>$ is a positive integer, we consider $D^>$, the space of functions $H : \mathbb{T}\times\mathbb{R} \to \mathbb{T}\times\mathbb{R}$, where $\partial_\theta^l \partial_s^m H_i(\theta,s)$, $i = 1,2$, exists if $l + m \leqslant L^>$, with $\|\cdot\|_{L^>,N}$ norm bounded by a constant B:

$$\|H\|_{L^>,N} := \max_{l+m\leqslant L^>, i=1,2} \begin{cases} \sup_{(\theta,s)\in\mathbb{T}\times\mathbb{R}} |\partial_\theta^l \partial_s^m H_i(\theta,s)||s|^{-(N-m)} & \text{if} \quad m \leqslant N, \\ \sup_{(\theta,s)\in\mathbb{T}\times\mathbb{R}} |\partial_\theta^l \partial_s^m H_i(\theta,s)| & \text{if} \quad m > N. \end{cases} \tag{2.66}$$

Using the notation introduced in (6.38), we have

$$\begin{aligned} \widetilde{W}(\theta,s) &= W(\theta - \omega\overline{r\circ K}(W(\theta,s)), se^{-\lambda\overline{r\circ K}(W(\theta,s))}) \\ &= W^{\leqslant}(\theta - \omega\overline{r\circ K}(W(\theta,s)), se^{-\lambda\overline{r\circ K}(W(\theta,s))}) \\ &\quad + W^{>}(\theta - \omega\overline{r\circ K}(W(\theta,s)), se^{-\lambda\overline{r\circ K}(W(\theta,s))}). \end{aligned}$$

We define

$$\widetilde{W}^{>}(\theta,s) = W^{>}(\theta - \omega\overline{r\circ K}((W^{\leqslant}+W^{>})(\theta,s)), se^{-\lambda\overline{r\circ K}((W^{\leqslant}+W^{>})(\theta,s))}). \tag{2.67}$$

Lemma 26. *If ε is small enough, $\Gamma^>(D^>) \subset D^>$.*

Proof. For $H \in D^>$, we need to prove that for $i = 1,2$, and $l+m \leqslant L^>$, $\partial_\theta^l \partial_s^m \Gamma_i^>(H)(\theta,s)$ exists, and that $\|\Gamma^>(H)\|_{L^>,N}$ is bounded by B. Using definition in equation (2.67)

$$\widetilde{H}(\theta,s) = H(\theta - \omega\overline{r\circ K}((W^{\leqslant}+H)(\theta,s)), se^{-\lambda\overline{r\circ K}((W^{\leqslant}+H)(\theta,s))})$$

We first claim that for $\|H\|_{L^>,N} \leqslant B$, we can find constant C, which does not depend on the choice of H, such that for $l+m \leqslant L^>$, $i = 1,2$, $(\theta,s) \in \widetilde{\mathbb{T}} \times [-1,1]$:

$$\begin{cases} |\partial_\theta^l \partial_s^m \widetilde{H}_i(\theta,s)| \leqslant C|s|^{(N-m)} & \text{if} \quad m \leqslant N, \\ |\partial_\theta^l \partial_s^m \widetilde{H}_i(\theta,s)| \leqslant C & \text{if} \quad m > N. \end{cases} \tag{2.68}$$

Note that within the proof of this lemma, C may vary from line to line. Finally, we will take C to be the maximum of all C's appeared in this proof.

To prove above claim, notice that $\|H\|_{L^>,N} \leqslant B$ implies that

$$\begin{cases} |\partial_\theta^l \partial_s^m H_i(\theta, s)| \leqslant B|s|^{(N-m)} & \text{if} \quad m \leqslant N, \\ |\partial_\theta^l \partial_s^m H_i(\theta, s)| \leqslant B & \text{if} \quad m > N. \end{cases}$$

for $l + m \leqslant L^>$, $i = 1, 2$, and $(\theta, s) \in \mathbb{T} \times \mathbb{R}$. Then

$$|\widetilde{H}_i(\theta, s)| \leqslant B|s|^N e^{-\lambda N \overline{r \circ K}((W^{\leqslant} + H)(\theta, s))}.$$

By boundedness of $\overline{r \circ K}$, we have that $|\widetilde{H}_i(\theta, s)| \leqslant C|s|^N$. Note that

$$\begin{aligned} \frac{\partial}{\partial \theta} \widetilde{H}_i(\theta, s) = \partial_\theta H_i &\left(\theta - \omega \overline{r \circ K}((W^{\leqslant} + H)(\theta, s)), s e^{-\lambda \overline{r \circ K}((W^{\leqslant} + H)(\theta, s))}\right) \cdot \\ &\cdot \left(1 - \omega D(\overline{r \circ K})((W^{\leqslant} + H)(\theta, s)) \partial_\theta (W^{\leqslant} + H)(\theta, s)\right) \\ &+ \partial_s H_i \left(\theta - \omega \overline{r \circ K}((W^{\leqslant} + H)(\theta, s)), s e^{-\lambda \overline{r \circ K}((W^{\leqslant} + H)(\theta, s))}\right) \cdot \\ &\cdot s(-\lambda) D(\overline{r \circ K})((W^{\leqslant} + H)(\theta, s)) \partial_\theta (W^{\leqslant} + H)(\theta, s) e^{-\lambda \overline{r \circ K}((W^{\leqslant} + H)(\theta, s))} \end{aligned}$$

Then, we have

$$\begin{aligned} \left| \frac{\partial}{\partial \theta} \widetilde{H}_i(\theta, s) \right| \leqslant & B|s|^N e^{-\lambda N \|\overline{r \circ K}\|} (1 + |\omega| \|D(\overline{r \circ K})\| \|\partial_\theta (W^{\leqslant} + H)\| \\ & + B|s|^{N-1} e^{-\lambda (N-1) \|\overline{r \circ K}\|} |s| |\lambda| \|D(\overline{r \circ K})\| e^{-\lambda \|r \circ K\|} \|\partial_\theta (W^{\leqslant} + H)\|. \end{aligned}$$

By boundedness of $W^{\leqslant}$, H, $\overline{r \circ K}$, and their derivatives, we have

$$\left| \frac{\partial}{\partial \theta} \widetilde{H}_i(\theta, s) \right| \leqslant C|s|^N.$$

Above C depends on B, but it will not depend on the choice of $H \in D^>$.

Similarly,

$$\begin{aligned}\frac{\partial}{\partial s}\widetilde{H}_i(\theta,s) = \partial_\theta H_i(\theta - \omega\overline{r\circ K}((W^{\leqslant}+H)(\theta,s)), se^{-\lambda\overline{r\circ K}((W^{\leqslant}+H)(\theta,s))})\cdot \\ \cdot(-\omega)D(\overline{r\circ K})((W^{\leqslant}+H)(\theta,s))\partial_s(W^{\leqslant}+H)(\theta,s) \\ +\partial_s H_i(\theta - \omega\overline{r\circ K}((W^{\leqslant}+H)(\theta,s)), se^{-\lambda\overline{r\circ K}((W^{\leqslant}+H)(\theta,s))})\cdot \\ \cdot\left(1+s(-\lambda)D(\overline{r\circ K})((W^{\leqslant}+H)(\theta,s))\partial_s(W^{\leqslant}+H)(\theta,s)\right)e^{-\lambda\overline{r\circ K}((W^{\leqslant}+H)(\theta,s))}.\end{aligned}$$

Then,

$$\begin{aligned}\left|\frac{\partial}{\partial s}\widetilde{H}_i(\theta,s)\right| \leqslant & B|s|^{N-1}e^{-\lambda(N-1)\|\overline{r\circ K}\|}\left(1+|s||\lambda|\|D(\overline{r\circ K})\|e^{-\lambda\|\overline{r\circ K}\|}\|\partial_s(W^{\leqslant}+H)\|\right) \\ & + B|s|^N e^{-\lambda N\|\overline{r\circ K}\|}|\omega|\|D(\overline{r\circ K})\|\|\partial_s(W^{\leqslant}+H)\|.\end{aligned}$$

Since we have $|s|\leqslant 1$, regularity of $W^{\leqslant}$ and H,

$$\left|\frac{\partial}{\partial s}\widetilde{H}_i(\theta,s)\right| \leqslant C|s|^{N-1}.$$

The C will not depend on the choice of H as long as $\|H\|_{L^>,N}\leqslant B$. The proof of the claim is then finished by induction.

Now we observe that we can bound the integrand in the operator $\Gamma^>$.

Claim: There exists a constant C, such that $\|Y(H,\theta,s,\varepsilon)\phi(s)\|_{L^>,N}\leqslant C$ when $\|H\|_{L^>,N}\leqslant B$.

Note that by definition of the cut-off function ϕ, it suffices to consider $s\in[-1,1]$.

$$Y^>(H,\theta,s,\varepsilon) = \overline{Y}((W^{\leqslant}+H)(\theta,s),(\widetilde{W^{\leqslant}+H})(\theta,s),\varepsilon) - \sum_{i=0}^{N-1}\overline{Y}^i(\theta)s^i,$$

where

$$\overline{Y}^i(\theta) = \frac{1}{i!}\frac{\partial^i}{\partial s^i}(\overline{Y}((W^{\leqslant}+H)(\theta,s),(\widetilde{W^{\leqslant}+H})(\theta,s),\varepsilon))\Big|_{s=0}.$$

One can add and subtract terms in above expression,

$$
\begin{aligned}
Y^{>}(H,\theta,s,\varepsilon) =& \overline{Y}((W^{\leqslant}+H)(\theta,s),\widetilde{(W^{\leqslant}+H)}(\theta,s),\varepsilon) \\
& -\overline{Y}(W^{\leqslant}(\theta,s),\widetilde{W^{\leqslant}}(\theta,s,H),\varepsilon) \\
& +\overline{Y}(W^{\leqslant}(\theta,s),\widetilde{W^{\leqslant}}(\theta,s,H),\varepsilon) \\
& -\overline{Y}(W^{\leqslant}(\theta,s),W^{\leqslant}(\theta-\omega\overline{r\circ K}(W^{\leqslant}(\theta,s)),se^{-\lambda\overline{r\circ K}(W^{\leqslant}(\theta,s))}),\varepsilon) \\
& +\overline{Y}(W^{\leqslant}(\theta,s),W^{\leqslant}(\theta-\omega\overline{r\circ K}(W^{\leqslant}(\theta,s)),se^{-\lambda\overline{r\circ K}(W^{\leqslant}(\theta,s))}),\varepsilon) \\
& -\sum_{i=0}^{N-1}\overline{Y}^{i}(\theta)s^{i},
\end{aligned}
\tag{2.69}
$$

where we used the notation

$$
\widetilde{W^{\leqslant}}(\theta,s;H)=W^{\leqslant}(\theta-\omega\overline{r\circ K}((W^{\leqslant}+H)(\theta,s)),se^{-\lambda\overline{r\circ K}((W^{\leqslant}+H)(\theta,s))}).
$$

We group the first two lines, the two lines in the middle, and the last two lines in equation (2.69), and denote them as ℓ_1, ℓ_2, and ℓ_3, respectively. Then for ℓ_1:

$$
\begin{aligned}
\ell_1 = & \int_0^1 D_1\overline{Y}((1-t)W^{\leqslant}(\theta,s)+t(W^{\leqslant}+H)(\theta,s),\widetilde{(W^{\leqslant}+H)}(\theta,s),\varepsilon)H(\theta,s)dt \\
& +\int_0^1 D_2\overline{Y}(W^{\leqslant}(\theta,s),(1-t)\widetilde{W^{\leqslant}}(\theta,s;H)+t\widetilde{(W^{\leqslant}+H)}(\theta,s),\varepsilon)\widetilde{H}(\theta,s)dt
\end{aligned}
$$

By the regularity of Y and $W^{\leqslant}$, $\|H\|_{L^{>},N}\leqslant B$, and that $\widetilde{H}$ satisfies equation (2.68), we know that $\|\ell_1\phi(s)\|_{L^{>},N}\leqslant C$.

Similarly, ℓ_2 is

$$
\begin{aligned}
& \int_0^1 D_2\overline{Y}(W^{\leqslant}(\theta,s),W^{\leqslant}(\theta-\omega\overline{r\circ K}((W^{\leqslant}+tH)(\theta,s)),se^{-\lambda\overline{r\circ K}((W^{\leqslant}+tH)(\theta,s))}),\varepsilon)\cdot \\
& \quad [\partial_\theta W^{\leqslant}(\cdot)(-\omega)D(\overline{r\circ K})(\cdot)+\partial_s W^{\leqslant}(\cdot)se^{-\lambda\overline{r\circ K}(\cdot)}D(\overline{r\circ K})(\cdot)(-\lambda)]H(\theta,s)dt,
\end{aligned}
$$

we have that $\|\ell_2\phi(s)\|_{L^>,N} \leqslant C$.

For ℓ_3, notice that $\sum_{i=0}^{N-1}\overline{Y}^i(\theta)s^i$ is the Taylor expansion at $s=0$ for

$$\overline{Y}(W^{\leqslant}(\theta,s), W^{\leqslant}(\theta-\omega\overline{r\circ K}(W^{\leqslant}(\theta,s)), se^{-\lambda\overline{r\circ K}(W^{\leqslant}(\theta,s))}),\varepsilon), \tag{2.70}$$

According to Taylor's Formula with remainder, see [36], we just need to show that for $m \leqslant N$

$$\frac{\partial^{N-m}}{\partial s^{N-m}}\frac{\partial^l}{\partial\theta^l}\frac{\partial^m}{\partial s^m}(2.70),$$

and for $m > N$,

$$\frac{\partial^m}{\partial s^m}\frac{\partial^l}{\partial\theta^l}(\ell_3),$$

are bounded for all θ, $|s| \leqslant 1$, and $l+m \leqslant L^>$. This is true if we assume that the lower order term has more regularity, more precisely, $L-1 \geqslant L^> + N$. We will take $L^> = L-1-N$ to optimize regularity. Therefore, we have $\|\ell_3\phi(s)\|_{L^>,N} \leqslant C$, and the claim is proved.

Hence, according to equation (2.65), if $m \leqslant N$, for small ε, we have that

$$|\partial_\theta^l\partial_s^m\Gamma_i^>(H)(\theta,s)| \leqslant \varepsilon\left|\int_0^\infty e^{-\lambda_0 t}C|s|^{N-m}e^{\lambda(N-m)t}e^{\lambda mt}dt\right| \leqslant B|s|^{N-m}, \tag{2.71}$$

if $m > N$, for small ε, we have that

$$|\partial_\theta^l\partial_s^m\Gamma_i^>(H)(\theta,s)| \leqslant \varepsilon\left|\int_0^\infty e^{-\lambda_0 t}Ce^{\lambda mt}dt\right| \leqslant B, \tag{2.72}$$

Therefore, for small ε, $\|\Gamma_i^>(H)\|_{L^>,N} \leqslant B$ when $\|H\|_{L^>,N} \leqslant B$. $\square$

Lemma 27. *If ε is small enough, $\Gamma^>$ is a contraction in $\|\cdot\|_{0,N}$.*

Proof. Recall that $\|H\|_{0,N} = \sup_{(\theta,s)\in\mathbb{T}\times\mathbb{R}} |H(\theta,s)||s|^{-N}$. We consider

$$\begin{aligned}\Gamma^{>}(H)(\theta,s) &- \Gamma^{>}(H')(\theta,s)\\ &= -\varepsilon\int_0^\infty \begin{pmatrix}1 & 0\\ 0 & e^{-\lambda_0 t}\end{pmatrix}\left(Y^{>}(H,c(t),\varepsilon) - Y^{>}(H',c(t),\varepsilon)\right)\phi(se^{\lambda t})dt \quad (2.73)\end{aligned}$$

Given the low order terms, denote $W = W^{\leqslant} + H$ and $W' = W^{\leqslant} + H'$, we have

$$\begin{aligned}Y^{>}(H,c(t),\varepsilon) &- Y^{>}(H',c(t),\varepsilon)\\ &= \overline{Y}(W(c(t)),\widetilde{W}(c(t)),\varepsilon) - \overline{Y}(W'(c(t)),\widetilde{W}'(c(t)),\varepsilon). \quad (2.74)\end{aligned}$$

Note that for all θ and s,

$$|W(\theta,s) - W'(\theta,s)| = |H(\theta,s) - H'(\theta,s)| \leqslant \|H - H'\|_{0,N}|s|^N. \quad (2.75)$$

Then for $\widetilde{W}(\theta, s) - \widetilde{W'}(\theta, s)$, by adding and subtracting terms, we have for all θ and s,

$$\begin{aligned}
|\widetilde{W}(\theta, s) - \widetilde{W'}(\theta, s)| =& \Big| W(\theta - \omega\overline{r \circ K}(W(\theta, s)), se^{-\lambda\overline{r\circ K}(W(\theta,s))}) \\
& - W'(\theta - \omega\overline{r \circ K}(W'(\theta, s)), se^{-\lambda\overline{r\circ K}(W'(\theta,s))}) \Big| \\
\leqslant& \Big| W(\theta - \omega\overline{r \circ K}(W(\theta, s)), se^{-\lambda\overline{r\circ K}(W(\theta,s))}) \\
& - W'(\theta - \omega\overline{r \circ K}(W(\theta, s)), se^{-\lambda\overline{r\circ K}(W(\theta,s))}) \Big| \\
& + \Big| W'(\theta - \omega\overline{r \circ K}(W(\theta, s)), se^{-\lambda\overline{r\circ K}(W(\theta,s))}) \\
& - W'(\theta - \omega\overline{r \circ K}(W'(\theta, s)), se^{-\lambda\overline{r\circ K}(W(\theta,s))}) \Big| \\
& + \Big| W'(\theta - \omega\overline{r \circ K}(W'(\theta, s)), se^{-\lambda\overline{r\circ K}(W(\theta,s))}) \\
& - W'(\theta - \omega\overline{r \circ K}(W'(\theta, s)), se^{-\lambda\overline{r\circ K}(W'(\theta,s))}) \Big| \\
\leqslant& M_1 \|H - H'\|_{0,N} |s|^N,
\end{aligned}$$

where

$$M_1 = e^{-\lambda N \|\overline{r\circ K}\|} + (\|DW^{\leqslant}\| + B)\|D(\overline{r \circ K})\|(|\omega| + |\lambda||s|e^{-\lambda\|\overline{r\circ K}\|}).$$

Then,

$$|\Gamma^{>}(H)(\theta, s) - \Gamma^{>}(H')(\theta, s)| \leqslant \varepsilon \|H - H'\|_{0,N} |s|^N \int_0^\infty e^{(\lambda N - \lambda_0)t} M\phi(se^{\lambda t}) dt,$$

where

$$M = \|D_1\overline{Y}\| + \|D_2\overline{Y}\| M_1.$$

Now, notice that by definition of D^1, we have that $\lambda \in [\frac{4\lambda_0}{3}, \frac{2\lambda_0}{3}]$, then $\lambda N - \lambda_0 < 0$ if

$N \geqslant 2$. Under this assumption, we have for all θ, s,

$$|\Gamma^{>}(H)(\theta, s) - \Gamma^{>}(H')(\theta, s)| \leqslant -\frac{\varepsilon M}{\lambda N - \lambda_0}\|H - H'\|_{0,N}|s|^N.$$

If ε is small enough, we have for all θ, s,

$$|\Gamma^{>}(H)(\theta, s) - \Gamma^{>}(H')(\theta, s)| \leqslant \mu\|H - H'\|_{0,N}|s|^N.$$

Hence for small enough ε,

$$\|\Gamma^{>}(H) - \Gamma^{>}(H')\|_{0,N} \leqslant \mu\|H - H'\|_{0,N},$$

$\Gamma^{>}$ is a contraction. Note that the smallness condition for ε depends on N, B^j, $j = 0, \dots, N-1$, B, ω_0, λ_0, $\overline{Y}$, and $\overline{r \circ K}$. $\square$

Now for any initial guess $W^{<,0}$, the sequence $(\Gamma^{>})^n(W^{>,0})$, in the function space $D^{>}$, will converge pointwise to a function $W^{>}$, which is the fixed point of $\Gamma^{>}$. By Lemma 92, we know that $W^{>}$ is $(L^{>} - 1)$ times differentiable, with $(L^{>} - 1)$-th derivative Lipschitz.

It remains to do the error analysis in this case. Notice that

$$E^{>}(\theta, s) = (\omega\partial_\theta + s\lambda\partial_s)W^{>,0}(\theta, s) - \begin{pmatrix} 0 \\ \lambda_0 W_2^{>,0}(\theta, s) \end{pmatrix} - \varepsilon Y^{>}(W^{>,0}, \theta, s, \varepsilon)\phi(s),$$

along the characteristics, we have

$$\begin{aligned} E^{>}(c(t)) = (\omega\partial_\theta + se^{\lambda t}\lambda\partial_s)W^{>,0}(c(t)) - \begin{pmatrix} 0 \\ \lambda_0 W_2^{>,0}(c(t)) \end{pmatrix} \\ - \varepsilon Y^{>}(W^{>,0}, c(t), \varepsilon)\phi(se^{\lambda t}). \end{aligned}$$

Hence,

$$\Gamma^{>}(W^{>,0})(\theta,s) - W^{>,0}(\theta,s) = \int_0^{\infty} \begin{pmatrix} 1 & 0 \\ 0 & e^{-\lambda_0 t} \end{pmatrix} E^{>}(c(t))dt.$$

The proof of Lemma 26 implies that $\|E^{>}\|_{0,N}$ is bounded, therefore, for the maximum norm,

$$\left\|\Gamma^{>}(W^{>,0}) - W^{>,0}\right\| \leqslant \frac{1}{\lambda_0 - \lambda N}\|E^{>}\|_{0,N}|s|^N,$$

and then

$$\|W^{>} - W^{>,0}\| \leqslant \frac{1}{1-\mu}\left\|\Gamma^{>}(W^{>,0}) - W^{>,0}\right\| \leqslant \frac{1}{(1-\mu)(\lambda_0 - \lambda N)}\|E^{>}\|_{0,N}|s|^N. \quad (2.76)$$

Combining error estimations in equation (2.42),equation (2.55), equation (2.61), and equation (2.76), we see that the $l = 0$ case of equation (2.22) is proved. Inequalities in equation (2.22) for $l \neq 0$ is obtained using interpolation inequalities.

2.5.3 Proof of Theorem 10 and Theorem 11

The proofs of Theorem 10 and Theorem 11 are obtained by considering the functions $W_\eta^j(\theta)$ as functions of two variables η and θ, denoted as $\tilde{W}^j(\eta,\theta)$. We can straightforwardly lift the operators Γ^0, Γ^1, and Γ^j defined in equation (2.27), equation (3.15), and equation (2.59) to operators acting on functions of two variables. We denote these operators acting on two-variable functions by $\tilde{\Gamma}^0$, $\tilde{\Gamma}^1$, and $\tilde{\Gamma}^j$, respectively. At the same time, we lift the operator $\Gamma^{>}$ to an operator acting on functions of three variables, denoted as $\tilde{\Gamma}^{>}$.

To prove Theorem 10, given a function $\tilde{W}^0(\eta,\theta)$ of the variables η,θ, we treat η as a parameter and take into account that now, Y and r depend also on η, in a smooth way.

We use the same strategy as in the proof of Theorem 6. We first show the propagated bounds property, similar to Lemma 16, and then, show that the operator is a contraction under a C^0-type distance, similar to Lemma 17. The distance here is quite analogue to the distance defined in equation (3.22). It is given by the sum of the C^0 distance of the two-variable functions and the difference between the frequencies. Then, the desired result, Theorem 10 follows by an application of Lemma 92.

In order to get the propagated bounds property, the key is to show that if $\|\tilde{W}\|_{L+\text{Lip}} \leqslant \tilde{B}^0$, for $\varepsilon < \varepsilon_0$, we have that the $C^{L+\text{Lip}}$ norms of the function components of $\tilde{\Gamma}^0(\tilde{W})$ are also bounded by $\tilde{B}^0$. This proof is rather straightforward and identical to the proof as before. More precisely, we apply Faá di Bruno formula in Lemma 86, and observe that the derivatives of order up to L of the function components of $\tilde{\Gamma}(\tilde{W}^0)$, are polynomials in the derivatives of $\tilde{W}^0$ of order up to L whose coefficients are derivatives of Y, r and combinatorial constants. Similarly, we can estimate the Lipschitz constants because upper bounds for the Lipschitz constants satisfy an analogue of Faá di Bruno formula.

To obtain the proof of the contraction, we just need to observe that the proof of the contraction in Lemma 17 only uses very few properties of Y and r. The properties hold uniformly for all η. Hence, one can obtain the contraction in the uniform norm on both variables.

Analogous arguments as above for the operators $\tilde{\Gamma}^j$ and $\tilde{\Gamma}^>$, using similar methods as in Sections 2.5.2.1, 2.5.2.2, 2.5.2.3, complete the proof for Theorem 11.

CHAPTER 3

PERIODIC ORBITS IN FDES CLOSE TO AN ODE OR AN EVOLUTIONARY PDE

Periodic orbits are important landmarks in dynamical systems. There has been interest in studying periodic orbits in DDEs, see [37, 38, 31, 39]. Some studies in the setting of SDDEs are in [40, 32, 34, 41]. Some numerical works are in [42, 43, 44].

In this Chapter, we first present a systematic approach to the study of periodic orbits of FDEs which are singular perturbations of smooth ODEs in $\mathbb{R}^n$.

We formulate functional equations satisfied by parameterizations of the periodic orbits and their frequencies in appropriate spaces of smooth functions. We solve the functional equations using a fixed point approach, and obtain existence of smooth solutions and dependence on parameters with high regularity.

Then, using a similar but more elaborate proof, we get results on periodic orbits for equations with small delays, which have applications in electrodynamics.

Finally, we extend the results to perturbations of PDEs. We can consider PDEs which have good forward (but not backward) evolutions such as parabolic equations as well as some ill-posed equations (e.g. Boussinesq equation in water waves, which even if ill posed, admits many physically interesting solutions).

We note that results on persistence of non-degenerate periodic orbits and dependence on parameters for FDEs with constant delays was proven by studying the evolution operator, see [12, 45], and [46]. This method is difficult to apply to SDDEs for example, for regularities higher than C^1, since one would need to extend the regularity theory of the evolution [21] to higher regularities. The paper [47] also studies functional equations satisfied by periodic orbits, but treats them using topological methods, which do not allow to study regularity. See also the excellent surveys [48, 49].

3.1 Formulation of the Problem

Consider an n-dimensional ODE

$$\dot{x}(t) = f(x(t)), \tag{3.1}$$

where, for the moment, $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is a C^∞ vector field (later we will assume less regularity).

We assume that equation (3.1) has a periodic orbit with frequency $\omega_0 \neq 0$. The existence of periodic solutions for ODEs will not be discussed here. (We note however that the same methods discussed here can be used to produce periodic solutions of the ODEs perturbatively.)

We consider singular perturbation of equation (3.1) to FDEs with parameter γ :

$$\dot{x}(t) = f(x(t)) + \varepsilon P(x_t, \gamma), \tag{3.2}$$

where $P\colon \mathscr{R}[-h,h] \times O \to \mathbb{R}^n$, h is a positive constant. $\mathscr{R}[-h,h]$ is a space of regular functions from $[-h,h]$ to $\mathbb{R}^n$. The precise regularity of the functions in $\mathscr{R}[-h,h]$ will be specified later. The *"history segment"* $x_t \in \mathscr{R}[-h,h]$ is defined as $x_t(s) = x(t+s)$ for $s \in [-h,h]$. And $\gamma \in O$ is a parameter, where O is a bounded open set in $\mathbb{R}^m$. Note that we allow our history segments to involve also the future, so that the theory we will develop applies not just to delay equations but to equations that involve the future.

In many treatments of delay equations it is customary to think of $\mathscr{R}[-h,h]$ as the phase space in which one sets initial conditions and defines an evolution. For example, in the case of constant delay equations, it is customary to impose initial conditions in $C^0[-h,0]$, with constant h being the delay. Nevertheless, in the case of SDDEs, this space includes many functions which cannot satisfy the equations and, therefore, have no physical meaning. As it will be clear later, our treatment bypasses the consideration of the evolution defined by the FDE, so that we will not think of $\mathscr{R}[-h, h]$ as the phase space of the evolution.

Under nondegeneracy condition on the periodic orbit of equation (3.1) and some mild assumptions on P, see more details in the definition of $\mathscr{P}$ in (3.4) and assumptions (H2.1), (H3.1), (H2.2), and (H3.2), we show that for small enough ε, there exists periodic orbit for FDE (3.2). We also show that the periodic orbits for equation (3.2) depend on γ smoothly.

From now on, we will identify the periodic orbit for FDE (3.2) in a function space with a periodic function having values in $\mathbb{R}^n$. Under this identification, we will see that the periodic orbit for FDE (3.2) is close to the periodic orbit for equation (3.1) for small ε.

3.2 Parameterization of Periodic Orbits

Let $K_0 \colon \mathbb{T} \to \mathbb{R}^n$ be a parameterization of the periodic orbit of equation (3.1), where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. This means that for any fixed θ, $x(t) = K_0(\theta + \omega_0 t)$ solves equation (3.1). Equivalently, K_0 satisfies the functional equation (invariance equation):

$$\omega_0 DK_0(\theta) = f(K_0(\theta)). \tag{3.3}$$

Note that such K_0 is unique up to a phase shift. In this case, K_0 is C^∞ since f is C^∞.

We aim to find $K \colon \mathbb{T} \to \mathbb{R}^n$ and $\omega > 0$, such that for any θ, $x(t) = K(\theta + \omega t)$ solves equation (3.2). And we say such K parameterizes the periodic orbit of FDE (3.2).

The expression $x(t) = K(\theta + \omega t)$ solving equation (3.2) is equivalent to K satisfying the functional equation:

$$\omega DK(\theta) = f(K(\theta)) + \varepsilon \mathscr{P}(K, \omega, \gamma, \theta), \tag{3.4}$$

where $\mathscr{P}(K, \omega, \gamma, \theta)$ results from substituting $x(t) = K(\theta + \omega t)$ into $P(x_t, \gamma)$ in equation (3.2) and letting $t = 0$. See Sections 3.5, 3.6, and 3.7 for explicit formulations of $\mathscr{P}$ in some specific examples.

The equation (3.4) will be the centerpiece of our treatment. We will see that, using different methods of analysis, we can give results on existence of solutions of (3.4). Note that this analysis produces periodic solutions of (3.2) without discussing a general theory of existence and dependence on parameters of the solutions for FDEs.

3.3 Main Results

3.3.1 Assumptions

For a given $\theta_0 \in \mathbb{T}$, let $\Phi(\theta; \theta_0)$ be the fundamental solution of the variational equation of the ODE (3.1), i.e.,

$$\omega_0 \frac{d}{d\theta}\Phi(\theta; \theta_0) = Df(K_0(\theta))\Phi(\theta; \theta_0), \quad \Phi(\theta_0; \theta_0) = Id. \tag{3.5}$$

We need to assume that the periodic orbit of (3.1) is nondegenerate, that is, we impose the following assumption on $\Phi(\theta_0 + 1; \theta_0)$:

(H1) $\Phi(\theta_0 + 1; \theta_0)$ has a simple eigenvalue 1 whose eigenspace is generated by $DK_0(\theta_0)$.

Note that, because of the existence and uniqueness of the solutions of (3.5), and the periodicity of K_0, we have that

$$\begin{aligned} &\Phi(\theta_2; \theta_0) = \Phi(\theta_2; \theta_1)\Phi(\theta_1; \theta_0); \\ &\Phi(\theta_1 + 1; \theta_0 + 1) = \Phi(\theta_1; \theta_0). \end{aligned}$$

As a consequence,

$$\Phi(\theta_0 + 1; \theta_0) = \Phi(\theta_0 + 1; 1)\Phi(1; 0)\Phi(0; \theta_0) = \Phi(0; \theta_0)^{-1}\Phi(1; 0)\Phi(0; \theta_0).$$

So that the spectrum of $\Phi(\theta_0 + 1; \theta_0)$, commonly called the Floquet multipliers, is independent of the starting point θ_0.

Under assumption (H1), there exists an $(n-1)$-dimensional linear space E_{θ_0} at $K_0(\theta_0)$, $\big($the spectral complement of Span$\{DK_0(\theta_0)\}$, corresponding to the eigenvalues of $\Phi(\theta_0 +$

$1;\theta_0)$ other than 1, $\mathbb{R}^n = E_{\theta_0} \oplus \text{Span}\{DK_0(\theta_0)\}\big)$, on which the matrix $[Id - \Phi(\theta_0 + 1; \theta_0)]$ is invertible. We denote the projections onto Span$\{DK_0(\theta_0)\}$ and E_{θ_0} as $\Pi^{\top}_{\theta_0}$ and $\Pi^{\perp}_{\theta_0}$, respectively.

Remark 28. *An equivalent formulation of* (H1) *in terms of functional analysis is* (H1'). *Define the operator* $\mathscr{L} \colon C^1(\mathbb{T}, \mathbb{R}^n) \to C^0(\mathbb{T}, \mathbb{R}^n)$*:*

$$\mathscr{L}(v)(\theta) = \omega_0 Dv(\theta) - Df(K_0(\theta))v(\theta).$$

(H1') $Range(\mathscr{L})$ *is of co-dimension 1,* $Range(\mathscr{L}) \oplus$ *Span*$\{DK_0\} = C^0(\mathbb{T}, \mathbb{R}^n)$.

The proofs of the Theorems in the next section imply the equivalence of (H1) *and* (H1').

To show the persistence of periodic orbit for a fixed $\gamma \in O$, the following assumptions on $\mathscr{P}$ are crucial. The assumption (H2.1) is about smoothness of $\mathscr{P}$ and expresses that $\mathscr{P}$ maps $C^{\ell+\text{Lip}}$ balls around zero into $C^{\ell-1+\text{Lip}}$ balls around zero (see Definition 84 for $C^{\ell+\text{Lip}}$ spaces). (H3.1) is about Lipschitz property of $\mathscr{P}$ in C^0 for smooth K's. These properties are verified in the examples we study in Sections 3.5 and 3.6. For example, when the functional P is evaluation on $x(t - r(x(t)))$, the regularity is a consequence of the fact that we can control the C^ℓ norm of $f \circ g$ by the C^ℓ norm of f, g. (We can even loose a derivative). The C^0 Lipschitz property results from the mean value theorem ($\|f \circ g_1 - f \circ g_2\|_{C^0} \leqslant \|f\|_{C^1} \|g_1 - g_2\|_{C^0}$).

In the following, ℓ is an arbitrarily fixed positive integer.

Let U_ρ be the ball of radius ρ in the space $C^{\ell+\text{Lip}}(\mathbb{T}, \mathbb{R}^n)$ centered at K_0, and let B_δ be the interval in $\mathbb{R}$ with radius δ centered at ω_0.

(H2.1) If $K \in U_\rho$ and $\omega \in B_\delta$, then $\mathscr{P}(K, \omega, \gamma, \cdot) \colon \mathbb{T} \to \mathbb{R}^n$ is $C^{\ell-1+\text{Lip}}$, with

$$\|\mathscr{P}(K, \omega, \gamma, \cdot)\|_{C^{\ell-1+\text{Lip}}} \leqslant \phi_{\rho,\delta},$$

where $\phi_{\rho,\delta}$ is a positive constant that may depend on ρ and δ. See (A.1) for definition of $C^{\ell+\mathrm{Lip}}$ norm.

(H3.1) For $K,\ K' \in U_\rho$, and $\omega,\ \omega' \in B_\delta$, there exists constant $\alpha_{\rho,\delta} > 0$, such that for all $\theta \in \mathbb{T}$,

$$|\mathscr{P}(K,\omega,\gamma,\theta) - \mathscr{P}(K',\omega',\gamma,\theta)| \leqslant \alpha_{\rho,\delta} \max\left\{|\omega-\omega'|, \|K-K'\|\right\}, \quad (3.6)$$

where $\|K-K'\|$ is the C^0-norm of $K-K'$ under the Euclidean distance on $\mathbb{R}^n$.

To show that the periodic orbits of the FDE (3.2) depend on the parameter γ smoothly, one needs to consider K, as a function of θ and γ, and ω as a function of γ. (H2.2) and (H3.2) are similar to (H2.1) and (H3.1), respectively. Note that we have slightly abused the notations K and ω.

We let $\mathcal{U}_\rho$ be the ball of radius ρ in the space $C^{\ell+\mathrm{Lip}}(\mathbb{T}\times O, \mathbb{R}^n)$ centered at K_0, and let $\mathcal{B}_\delta$ be the ball in $C^{\ell+\mathrm{Lip}}(O,\mathbb{R})$ with radius δ centered at constant function ω_0.

(H2.2) If $K \in \mathcal{U}_\rho$ and $\omega \in \mathcal{B}_\delta$, then $\mathscr{P}(K,\omega,\cdot,\cdot)\colon \mathbb{T}\times O \to \mathbb{R}^n$ is $C^{\ell+\mathrm{Lip}}$ in γ, and $C^{\ell-1+\mathrm{Lip}}$ in θ, with

$$\|\mathscr{P}(K,\omega,\cdot,\theta)\|_{C^{\ell+\mathrm{Lip}}} \leqslant \phi_{\rho,\delta},$$
$$\|\mathscr{P}(K,\omega,\gamma,\cdot)\|_{C^{\ell-1+\mathrm{Lip}}} \leqslant \phi_{\rho,\delta},$$

where $\phi_{\rho,\delta}$ is a positive constant.

(H3.2) For $K,\ K' \in \mathcal{U}_\rho$ and $\omega,\ \omega' \in \mathcal{B}_\delta$, there exists constant $\alpha_{\rho,\delta} > 0$, such that for all $\theta \in \mathbb{T}$ and $\gamma \in O$,

$$|\mathscr{P}(K,\omega,\gamma,\theta) - \mathscr{P}(K',\omega',\gamma,\theta)| \leqslant \alpha_{\rho,\delta} \max\left\{\|\omega-\omega'\|, \|K-K'\|\right\},$$

where $\|\omega-\omega'\|$ is the C^0-norm of $\omega-\omega'$.

Remark 29. *Note that our results work exactly the same if the perturbation depends on* ε*, i.e. we have* $P(x_t, \gamma, \varepsilon)$ *instead of* $P(x_t, \gamma)$ *in* (3.2)*. We can get* $\mathscr{P}(K, \omega, \gamma, \varepsilon, \theta)$ *in this case. We need assumptions on* $\mathscr{P}$ *to hold uniformly in* ε *for all small* ε*.*

More specifically, (H2.1)*,* (H3.1) *can be reformulated as:*

(H2.1') *If* $K \in U_\rho$ *and* $\omega \in B_\delta$*, then* $\mathscr{P}(K, \omega, \gamma, \varepsilon, \cdot) \colon \mathbb{T} \to \mathbb{R}^n$ *is* $C^{\ell-1+\mathrm{Lip}}$*, with*

$$\|\mathscr{P}(K, \omega, \gamma, \varepsilon, \cdot)\|_{C^{\ell-1+\mathrm{Lip}}} \leqslant \phi_{\rho,\delta}(\varepsilon).$$

Function $\phi_{\rho,\delta}$ *satisfies that* $\varepsilon\phi_{\rho,\delta}(\varepsilon)$ *converges to zero as* $\varepsilon \to 0$*.*

(H3.1') *For* $K,\ K' \in U_\rho$*, and* $\omega,\ \omega' \in B_\delta$*, there exists positive function* $\alpha_{\rho,\delta}$*, such that for all* $\theta \in \mathbb{T}$*,*

$$|\mathscr{P}(K, \omega, \gamma, \varepsilon, \theta) - \mathscr{P}(K', \omega', \gamma, \varepsilon, \theta)| \leqslant \alpha_{\rho,\delta}(\varepsilon) \max\{|\omega - \omega'|, \|K - K'\|\},$$

function $\alpha_{\rho,\delta}$ *satisfies that* $\varepsilon\alpha_{\rho,\delta}(\varepsilon)$ *converges to zero as* $\varepsilon \to 0$*.*

The assumptions similar to (H2.2)*,* (H3.2) *can be formulated similarly.*

Remark 30. *The assumptions we use are similar to assumptions in invariant manifold theory. For example in [50], the* (H2.1) *is called* propagated bounds.

Remark 31. *We call attention to the fact that in Section 3.7 we will weaken substantially the assumption* (H3.1) *to be able to deal with equations with small delays.*

3.3.2 Main theorems

Let $\mathbb{N}$ denote the set of positive numbers.

Theorem 32 (Persistence)**.** *For a given* $\ell \in \mathbb{N}$*, assume that* f *in* (3.2) *is* $C^{\ell+\mathrm{Lip}}$*, and that* (H1), (H2.1)*, and* (H3.1) *are satisfied for a given* $\gamma \in O$*. Then, there exists* $\varepsilon_0 > 0$*, such* *that when $\varepsilon < \varepsilon_0$, the FDE* (3.2) *has a periodic orbit, which is parameterized by a $C^{\ell+\mathrm{Lip}}$ map $K\colon \mathbb{T} \to \mathbb{R}^n$. The smallness condition of ε_0 depends on ℓ, f, and P.*

The frequency ω for the periodic orbit of equation (3.2) *is close to ω_0, the frequency of the periodic orbit of equation* (3.1)*. $\|K - K_0\|_{C^\ell}$ is small under a suitable choice of the phases.*

Theorem 33 (Smooth Dependence on Parameter)**.** *For a given $\ell \in \mathbb{N}$, assume that f in* (3.2) *is $C^{\ell+\mathrm{Lip}}$, and that* (H1), (H2.2), *and* (H3.2) *are satisfied. Then, there is $\varepsilon_0 > 0$, such that if $\varepsilon < \varepsilon_0$, one can find $K_\gamma(\theta)$ which parameterizes the periodic orbit of FDE* (3.2) *persisted from the periodic orbit of* (3.1)*. The smallness condition of ε_0 depends on ℓ, f, and P.*

K_γ has frequency ω_γ. $K_\gamma(\theta)$ is jointly $C^{\ell+\mathrm{Lip}}$ in θ and γ, ω_γ is $C^{\ell+\mathrm{Lip}}$ in γ.

3.3.3 Some comments on the Theorems 32 and 33

Remark 34. *One physically important case where assumption* (H1) *fails is when there is a conserved quantity (for example, the energy in mechanical systems). We are not able to deal with this case by the method of this Chapter, but we hope to come back to this problem.*

Remark 35. *Note that K will not be unique. If $K(\theta)$ parameterizes the periodic orbit, then for any given θ_1, $K(\theta+\theta_1)$ also parameterizes the periodic orbit, with a shifted phase. Hence, in Theorem 32, the smallness of $\|K-K_0\|_{C^{\ell-1}}$ is interpreted under a suitable choice of the phases.*

This is the only source of non-uniqueness since the proofs of Theorems 32 and 33 are based on contraction mapping argument, the parameterizations we found are locally unique up to phase shifts.

Remark 36. *The smallness of ε depends on ℓ, hence, the method cannot get a C^∞ result directly. Note, however, that in some cases, e.g. state-dependent delay perturbations in equation* (3.28)*, one can bootstrap the regularity from C^1 to C^∞.*

Remark 37. *Our results apply to several types of FDEs, especially to many DDEs, see Sections 3.5 and 3.6. We only need that* (H1), (H2.1), (H2.2), (H3.1), *and* (H3.2) *are satisfied. Indeed, we allow several terms in the equation which may involve forward and backward delays.*

Remark 38. *Our method allows to bypass the propagation of discontinuity in DDEs. Moreover, it has no restriction on the relation between the period of the periodic orbits and the size of the delay.*

Remark 39. *The proofs we present are constructive, hence they can be implemented numerically. Indeed, we formulate the problem as a fixed point of a contractive operator, which concatenates several elementary operators. Implementations of these elementary operators for a 2D model are addressed in a numerical toolkit developed in [24].*

The proofs, based on fixed point approach, also lead to results in an a posteriori format, which state that if there is an approximate solution (satisfying some mild assumptions), then there is a true solution which is close to the approximate one. See more details in Section 3.4.5.

Remark 40. *A posteriori theorems justify asymptotic expansions where solutions are written as formal expansions in terms of the small parameters, see [51, 30]. Truncations of the formal power series provide approximate solutions. The a posteriori theorem shows that there is one true solution close by.*

A posteriori theorems are also the base of computer-assisted proofs. Numerical methods produce approximate solutions. If one can estimate rigorously the error and the non-degeneracy conditions, then one has established the existence of the solution. The verification of the error in the approximation is a finite (but long) calculation which can be done using computers taking care of round-off and truncation. Some cases where computer-assisted proofs have been used in constant delay equations for periodic orbits and unstable manifolds are [52, 53].

3.4 Proofs

The proofs of Theorems 32 and 33 are based on fixed point approach. We will provide the detailed proof of Theorem 32. The proof of Theorem 33 follows in the same manner by adding the parameters in the unknowns, see Section 3.4.6.

The proof consist of several steps. First, we define an operator in an appropriate space of smooth functions. Then, we show that (i) the operator maps a ball in this space into itself (Section 3.4.3); (ii) the operator is a contraction in a C^0 type of distance (Section 3.4.4). The existence of fixed point in desired space is hence ensured using a generalization of contraction mapping [50].

3.4.1 Invariance equations

In this section, we reformulate the invariance equation (3.4). Since we expect that the solutions K, ω will be small perturbations of the unperturbed ones, it is natural to reformulate (3.4) as an equation for the corrections from the unperturbed ones. In Section 3.4.2, we will manipulate the equation for the corrections into a fixed point problem.

Let

$$\begin{aligned} K(\theta) &\coloneqq K_0(\theta) + \widehat{K}(\theta), \\ \omega &\coloneqq \omega_0 + \widehat{\omega}, \end{aligned} \tag{3.7}$$

where $\widehat{K}\colon \mathbb{T} \to \mathbb{R}^n$ and $\widehat{\omega} \in \mathbb{R}$ are corrections to the parameterization and frequency of the periodic orbit of the unperturbed equation. Our goal is to find $\widehat{K}$ and $\widehat{\omega}$ so that K and ω satisfy the functional equation (3.4).

Using the notation in (3.7) and the invariance equation (3.3) for K_0 and ω_0, we are led to the following functional equation for $\widehat{K}$ and $\widehat{\omega}$,

$$\omega_0 D\widehat{K}(\theta) - Df(K_0(\theta))\widehat{K}(\theta) = B^\varepsilon(\widehat{K}, \widehat{\omega}, \gamma, \theta) - \widehat{\omega} DK_0(\theta), \tag{3.8}$$

where

$$B^{\varepsilon}(\widehat{K},\widehat{\omega},\gamma,\theta) \coloneqq N(\theta,\widehat{K}) + \varepsilon\mathscr{P}(K,\omega,\gamma,\theta) - \widehat{\omega}D\widehat{K}(\theta), \tag{3.9}$$
$$N(\theta,\widehat{K}) \coloneqq f(K_0(\theta)+\widehat{K}(\theta)) - f(K_0(\theta)) - Df(K_0(\theta))\widehat{K}(\theta).$$

The basic idea for this regrouping is that since K and ω are expected to be close to K_0 and ω_0 respectively, we only need to find the corrections.

3.4.2 The operator

Recall $\Phi(\theta;\theta_0)$ introduced in (3.5) as the flow of the variational equations. Using the variation of parameters formula, equation (3.8) for $\widehat{K}$ and $\widehat{\omega}$ is equivalent to:

$$\widehat{K}(\theta) = \Phi(\theta;\theta_0)\left\{u_0 + \frac{1}{\omega_0}\int_{\theta_0}^{\theta}\Phi(s;\theta_0)^{-1}\big(B^{\varepsilon}(\widehat{K},\widehat{\omega},\gamma,s) - \widehat{\omega}DK_0(s)\big)ds\right\}, \tag{3.10}$$

where the initial condition $\widehat{K}(\theta_0) = u_0$ is to be found imposing that $\widehat{K}$ is periodic. This will be discussed in the Section 3.4.2.1.

We can think of (3.10) as a fixed point equation. The right hand side is an operator in $\widehat{K}$, see Section 3.4.2.3. We start with a given $\widehat{K}$, choose $\widehat{\omega}$ following Section 3.4.2.1 and we substitute them in right hand side of (3.10).

3.4.2.1 Periodicity Condition

Since the right hand side of equation (3.8) is periodic, $\widehat{K}$ is periodic if and only if $\widehat{K}(\theta_0) = \widehat{K}(\theta_0+1)$, i.e.,

$$\begin{aligned}[Id-\Phi(\theta_0+1;\theta_0)]u_0 = {} & \frac{1}{\omega_0}\Phi(\theta_0+1;\theta_0)\int_{\theta_0}^{\theta_0+1}\Phi(s;\theta_0)^{-1}B^{\varepsilon}(\widehat{K},\widehat{\omega},\gamma,s)ds \\ & -\frac{\widehat{\omega}}{\omega_0}\Phi(\theta_0+1;\theta_0)\int_{\theta_0}^{\theta_0+1}\Phi(s;\theta_0)^{-1}DK_0(s)ds.\end{aligned} \tag{3.11}$$

Since K_0 solves (3.3), and Φ satisfies (3.5), we have

$$\Phi(s;\theta_0)DK_0(\theta_0) = DK_0(s).$$

Then, the periodicity condition (3.11) becomes

$$\begin{aligned}[Id - \Phi(\theta_0+1;\theta_0)]u_0 = \frac{1}{\omega_0}\int_{\theta_0}^{\theta_0+1} \Phi(\theta_0+1;s)B^\varepsilon(\widehat{K},\widehat{\omega},\gamma,s)ds \\ -\frac{\widehat{\omega}}{\omega_0}DK_0(\theta_0). \end{aligned} \tag{3.12}$$

One is able to solve for u_0 if the right hand side of equation (3.12) is in the range of $Id-\Phi(\theta_0+1;\theta_0)$. Thanks to assumption (H1), this can be achieved by choosing the correct $\widehat{\omega}$. The choice of $\widehat{\omega}$ is unique.

3.4.2.2 Spaces

Let $a > 0$ and define interval $I_a = [-a, a]$, let

$$\begin{aligned}\mathscr{B}_\beta = \Big\{ g\colon \mathbb{T} \to \mathbb{R}^n \Big|\ g \text{ is } C^{\ell+\mathrm{Lip}},\ \Big\| \frac{d^i}{d\theta^i} g(\theta)\Big\| \leqslant \beta_i,\ i = 0,1,\ldots,\ell, \\ \mathrm{Lip}\left(\frac{d^\ell}{d\theta^\ell}g(\theta)\right) \leqslant \beta_\ell^{\mathrm{Lip}} \Big\},\end{aligned} \tag{3.13}$$

where $\beta = (\beta_0, \beta_1, \ldots, \beta_\ell, \beta_\ell^{\mathrm{Lip}})$. The constants a, β_i, $i = 0,1,\ldots,\ell$, and $\beta_\ell^{\mathrm{Lip}}$ will be chosen in the proof.

3.4.2.3 Definition of the Operator

Define the operator Γ^ε on $I_a \times \mathscr{B}_\beta$,

$$\Gamma^\varepsilon(\widehat{\omega},\widehat{K}) = \begin{pmatrix} \Gamma_1^\varepsilon(\widehat{\omega},\widehat{K}) \\ \Gamma_2^\varepsilon(\widehat{\omega},\widehat{K}) \end{pmatrix}. \tag{3.14}$$

Componentwise,

$$\Gamma_1^\varepsilon(\widehat{\omega}, \widehat{K}) = \frac{\left\langle \int_{\theta_0}^{\theta_0+1} \Pi_{\theta_0}^\top \Phi(\theta_0+1; s) B^\varepsilon(\widehat{K}, \widehat{\omega}, \gamma, s) ds, DK_0(\theta_0) \right\rangle}{|DK_0(\theta_0)|^2}, \tag{3.15}$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^n$.

$$\begin{aligned}\Gamma_2^\varepsilon(\widehat{\omega}, \widehat{K})(\theta) = {} & \Phi(\theta; \theta_0) u_0 \\ & + \frac{1}{\omega_0} \int_{\theta_0}^{\theta} \Phi(\theta; s) \big(B^\varepsilon(\widehat{K}, \widehat{\omega}, \gamma, s) - \Gamma_1^\varepsilon(\widehat{\omega}, \widehat{K}) DK_0(s) \big) ds,\end{aligned} \tag{3.16}$$

where $u_0 \in E$ satisfies

$$\begin{aligned}[Id - \Phi(\theta_0+1; \theta_0)] u_0 = {} & \frac{1}{\omega_0} \int_{\theta_0}^{\theta_0+1} \Phi(\theta_0+1; s) B^\varepsilon(\widehat{K}, \widehat{\omega}, \gamma, s) ds \\ & - \frac{\Gamma_1^\varepsilon(\widehat{\omega}, \widehat{K})}{\omega_0} DK_0(\theta_0) \\ = {} & \frac{1}{\omega_0} \int_{\theta_0}^{\theta_0+1} \Pi_{\theta_0}^\perp \Phi(\theta_0+1; s) B^\varepsilon(\widehat{K}, \widehat{\omega}, \gamma, s) ds.\end{aligned} \tag{3.17}$$

Remark 41. *The definition of* $\Gamma_1^\varepsilon(\widehat{\omega}, \widehat{K})$ *ensures the right hand side of* (3.17) *to be in the range of* $Id - \Phi(\theta_0+1; \theta_0)$*. Since the kernel of* $Id - \Phi(\theta_0+1; \theta_0)$ *is Span*$\{DK_0(\theta_0)\}$*, equation* (3.17) *has infinitely many solutions, all of them are the same up to constant multiples of* $DK_0(\theta_0)$*. In the definition of the operator* Γ^ε*, we have chosen the solution for equation* (3.17) *which lies in the space* E*. If we choose a different* u_0 *solving* (3.17)*, we will get another parameterization of the periodic orbit corresponding to a different phase, see Remark 35.*

Our goal is to find the fixed point $(\widehat{\omega}^*, \widehat{K}^*)$ of the operator Γ^ε in a ball $I_a \times \mathscr{B}_\beta$, which will solve the equation (3.8). Hence $\omega = \omega_0 + \widehat{\omega}^*$ and $K = K_0 + \widehat{K}^*$ satisfy (3.4), K parameterizes the periodic orbit of (3.2) with frequency ω.

To this end, under the assumptions (H1), (H2.1) and (H2.2), we show in Section 3.4.3

that for small ε, we can choose a and β so that Γ^ε maps $I_a \times \mathscr{B}_\beta$ back into itself.

In Section 3.4.4 we show that Γ^ε is a contraction in a C^0-like distance. The desired result of existence of a locally unique fixed point follows from a fixed point result in the literature that we have collected as Lemma 92.

3.4.3 Propagated bounds for Γ^ε

In this section, we will prove the following Lemma.

Lemma 42. *Assume ε is small enough, then a and β can be chosen such that $\Gamma^\varepsilon\colon I_a \times \mathscr{B}_\beta \to I_a \times \mathscr{B}_\beta$.*

Proof. Note that

$$\|N(\theta, \widehat{K})\| \leqslant \frac{1}{2}\operatorname{Lip}(Df)\|\widehat{K}\|^2,$$

where $\|\cdot\|$ means C^0-norm. Indeed, here and later in this proof we only need the Lipschitz constant of $Df(x)$ in a neighborhood of the periodic orbit of the unperturbed ODE, i.e. $K_0(\mathbb{T})$.

Using the integration by parts formula, for $\theta \in [\theta_0, \theta_0 + 1]$, we have

$$\left|\int_{\theta_0}^{\theta} \Phi(\theta_0 + 1; s)\widehat{\omega} D\widehat{K}(s)ds\right| \leqslant \left(2\|\Phi(\theta_0 + 1; \theta)\| + \left\|\frac{d}{d\theta}\Phi(\theta_0 + 1; \theta)\right\|\right)|\widehat{\omega}|\|\widehat{K}\|.$$

where

$$\|\Phi(\theta_0 + 1; \theta)\| \coloneqq \max_{\theta\in[\theta_0,\theta_0+1]} |\Phi(\theta_0 + 1; \theta)|,$$

and

$$\left\|\frac{d}{d\theta}\Phi(\theta_0 + 1; \theta)\right\| \coloneqq \max_{\theta\in[\theta_0,\theta_0+1]} \left|\frac{d}{d\theta}\Phi(\theta_0 + 1; \theta)\right|,$$

and $|\cdot|$ denotes the operator norm of the matrix. We will use similar conventions for norms from now on.

Since $(\widehat{\omega}, \widehat{K}) \in I_a \times \mathscr{B}_\beta$, we have

$$|\Gamma_1^\varepsilon(\widehat{\omega}, \widehat{K})| \leqslant \frac{\|\Pi_{\theta_0}^\top\|}{\big|DK_0(\theta_0)\big|}\Big[\|\Phi(\theta_0+1;s)\|\Big(\frac{1}{2}\operatorname{Lip}(Df)\beta_0^2 + \varepsilon\|\mathscr{P}(K,\omega,\gamma,\theta)\|\Big) \\ + \Big(2\|\Phi(\theta_0+1;\theta)\| + \Big\|\frac{d}{d\theta}\Phi(\theta_0+1;\theta)\Big\|\Big)a\beta_0\Big], \tag{3.18}$$

and,

$$\begin{aligned}\|\Gamma_2^\varepsilon(\widehat{\omega}, \widehat{K})\| \leqslant \|\Phi(\theta;\theta_0)\| M\Big[&\|\Phi(\theta_0+1;s)\|(\frac{1}{2}\operatorname{Lip}(Df)\beta_0^2 + \varepsilon\|\mathscr{P}(K,\omega,\gamma,\theta)\|) \\ &+ \Big(2\|\Phi(\theta_0+1;\theta)\| + \Big\|\frac{d}{d\theta}\Phi(\theta_0+1;\theta)\Big\|\Big)a\beta_0\Big] \\ + \frac{1}{\omega_0}\Big[&\|\Phi(\theta;s)\|(\frac{1}{2}\operatorname{Lip}(Df)\beta_0^2 + \varepsilon\|\mathscr{P}(K,\omega,\gamma,\theta)\|) \\ &+ \Big(2\|\Phi(\theta;\theta_0)\| + \Big\|\frac{d}{ds}\Phi(\theta;s)\Big\|\Big)a\beta_0\Big] \\ + \frac{\|\Phi(\theta;s)\|}{\omega_0}&\Big\|DK_0(s)\Big\||\Gamma_1^\varepsilon(\widehat{\omega}, \widehat{K})|,\end{aligned} \tag{3.19}$$

where

$$\begin{aligned}\|\Phi(\theta;s)\| &:= \max_{\theta\in[\theta_0,\theta_0+1]}\max_{s\in[\theta_0,\theta]}|\Phi(\theta;s)|, \\ \Big\|\frac{d}{ds}\Phi(\theta;s)\Big\| &:= \max_{\theta\in[\theta_0,\theta_0+1]}\max_{s\in[\theta_0,\theta]}\Big|\frac{d}{ds}\Phi(\theta;s)\Big|,\end{aligned}$$

and

$$M := \frac{\|[Id - \Phi(\theta_0+1;\theta_0)]^{-1}\|\|\Pi_{\theta_0}^\perp\|}{\omega_0}. \tag{3.20}$$

We have used $[Id - \Phi(\theta_0+1;\theta_0)]^{-1}$ to denote the inverse of $[Id - \Phi(\theta_0+1;\theta_0)]$ in the $(n-1)$-dimensional space E_{θ_0} introduced in Section 3.3.1.

Note that for the right hand sides of the inequalities (3.18) and (3.19) above, each term is either quadratic in a, β_0 or has a factor ε. Under smallness assumptions of a, β_0, and ε, we will have $|\Gamma_1^\varepsilon(\widehat{\omega}, \widehat{K})| \leqslant a$ and $\|\Gamma_2^\varepsilon(\widehat{\omega}, \widehat{K})\| \leqslant \beta_0$.

Now we consider the derivatives of $\Gamma_2^\varepsilon(\widehat{\omega}, \widehat{K})$.

The first derivative $\frac{d}{d\theta}\Gamma_2^\varepsilon(\widehat{\omega},\widehat{K})(\theta)$ has the expression:

$$\Big(\frac{d}{d\theta}\Phi(\theta;\theta_0)\Big)\Big\{u_0+\frac{1}{\omega_0}\int_{\theta_0}^{\theta}\Phi(s;\theta_0)^{-1}\big(B^\varepsilon(\widehat{K},\widehat{\omega},\gamma,s)-\Gamma_1^\varepsilon(\widehat{\omega},\widehat{K})DK_0(s)\big)ds\Big\}$$
$$+\frac{1}{\omega_0}\Big\{B^\varepsilon(\widehat{K},\widehat{\omega},\gamma,\theta)-\Gamma_1^\varepsilon(\widehat{\omega},\widehat{K})DK_0(\theta)\Big\}.$$

Recall that $\Phi(\theta;\theta_0)$ solves equation (3.5). Therefore,

$$\begin{aligned}\Big\|\frac{d}{d\theta}\Gamma_2^\varepsilon(\widehat{\omega},\widehat{K})\Big\| \leqslant&\frac{1}{\omega_0}\|Df(K_0(\theta))\|\|\Gamma_2^\varepsilon(\widehat{\omega},\widehat{K})\|+\frac{|\Gamma_1^\varepsilon(\widehat{\omega},\widehat{K})|}{\mathrm{Lip}}(Df)\omega_0\Big\|DK_0(\theta)\Big\|\\
&+\frac{1}{\omega_0}\Big(\frac{1}{2}\,\mathrm{Lip}(Df)\|\widehat{K}\|^2+\varepsilon\|\mathscr{P}(K,\omega,\gamma,\theta)\|+|\widehat{\omega}|\Big\|D\widehat{K}(\theta)\Big\|\Big)\\
\leqslant&\frac{1}{\omega_0}\|Df(K_0(\theta))\|\beta_0+\frac{a}{\omega_0}\Big\|DK_0(\theta)\Big\|\\
&+\frac{1}{\omega_0}\Big(\frac{1}{2}\,\mathrm{Lip}(Df)\beta_0^2+\varepsilon\|\mathscr{P}(K,\omega,\gamma,\theta)\|+a\beta_1\Big).\end{aligned}$$

If ε, a, and β_0 are small enough, we can choose β_1 to ensure that $\Big\|\frac{d}{d\theta}\Gamma_2^\varepsilon(\widehat{\omega},\widehat{K})\Big\|\leqslant\beta_1$.

Now we proceed inductively, for $n\geqslant 2$, $\frac{d^n}{d\theta^n}\Gamma_2^\varepsilon(\widehat{\omega},\widehat{K})$ is an expression involving Φ, K_0, and their derivatives up to order n, as well as B^ε and its derivatives up to order $n-1$. Within this expression, K_0 and its derivatives are always multiplied by the small factor $\Gamma_1^\varepsilon(\widehat{\omega},\widehat{K})$, which has absolute value bounded by constant a. It remains to consider B^ε and its derivatives.

Recall the definition of B^ε in (3.9), we now consider the three terms in B^ε separately:

- For derivatives of $N(\theta,\widehat{K})$, we use the Faa di Bruno formula. The j-th derivative of N is an expression which contains derivatives of f up to order $j+1$, derivatives of $\widehat{K}$ up to order j. All the terms in derivatives of N can be controlled taking advantage of the fact that N is of order at least 2 in $\widehat{K}$.

- Derivatives of $\mathscr{P}$ are bounded thanks to the assumption (H2.1). Moreover, note that in B^ε, $\mathscr{P}$ has the perturbation parameter ε as its coefficient. Hence, this term is less crucial.

- For the last term, $\widehat{\omega} D\widehat{K}(\theta)$, its j-th derivative is $\widehat{\omega} D^{j+1}\widehat{K}(\theta)$. All are under control since $|\widehat{\omega}| < a$. Notice that the $(n-1)$-th derivative of this term is $\widehat{\omega} D^n\widehat{K}(\theta)$, which is the only place that $D^n\widehat{K}(\theta)$ appears.

Taking all the terms above into consideration and using the triangle inequality, we obtain bounds

$$\left\| \frac{d^n}{d\theta^n}\Gamma_2^\varepsilon(\widehat{\omega}, \widehat{K}) \right\| \leqslant P_n(a, \beta_0, \dots, \beta_{n-1}) + \alpha\beta_n, \tag{3.21}$$

where for each n, P_n is a polynomial expression with positive coefficients, and $\alpha < 1$. The coefficients of P_n are combinatorial numbers multiplied by derivatives of K_0, $\mathscr{P}$, f, and $\Phi(\theta;\theta_0)$. Therefore, we can choose recursively the β_i's such that right hand side of inequality (3.21) is bounded by β_n.

Similar estimation can be obtained for the Lipschitz constant of $\frac{d^\ell}{d\theta^\ell}\Gamma_2^\varepsilon(\widehat{\omega}, \widehat{K})$. Hence, we can choose a, β such that $\Gamma^\varepsilon\colon I_a \times \mathscr{B}_\beta \to I_a \times \mathscr{B}_\beta$. $\square$

Remark 43. *Note that for ε sufficiently small, we can choose constant a and each component of β to be as small as we want.*

Remark 44. *Note that $I_a \times \mathscr{B}_\beta \subset \mathbb{R} \times C(\mathbb{T}, \mathbb{R}^n)$ is compact and convex, and it is obvious that $\Gamma^\varepsilon\colon I_a \times \mathscr{B}_\beta \to I_a \times \mathscr{B}_\beta$ is continuous, so one could apply Schauder's fixed point Theorem to obtain existence of the fixed point. Indeed, weaker assumptions than assumption* (H3.1) *on $\mathscr{P}$ could also suffice to ensure continuity of Γ^ε.*

We will later prove that Γ^ε is a contraction in C^0 topology, which will give local uniqueness of the fixed point and a posteriori estimates on the difference between an initial guess and the fixed point.

In principle, the Banach contraction theorem provides estimates of the difference in C^0 norm, but, taking into account the propagated bounds, we can use interpolation inequalities (Lemma 91) to obtain estimates in norms with higher regularity. See Section 3.4.5.

3.4.4 Contraction properties of Γ^ε

Define C^0-type distance on $I_a \times \mathscr{B}_\beta$:

$$d\left((\widehat{\omega}, \widehat{K}), (\widehat{\omega}', \widehat{K}')\right) \coloneqq \max\left\{|\widehat{\omega} - \widehat{\omega}'|, \|\widehat{K} - \widehat{K}'\|\right\}. \tag{3.22}$$

Lemma 45. *For small enough ε, a, and β_0 (as in β), the operator in* (3.14) *is a contraction on $I_a \times \mathscr{B}_\beta$ with distance* (3.22)*, i.e., there exists $0 < \mu < 1$, such that*

$$d\left(\Gamma^\varepsilon(\widehat{\omega}, \widehat{K}), \Gamma^\varepsilon(\widehat{\omega}', \widehat{K}')\right) < \mu \cdot d\left((\widehat{\omega}, \widehat{K}), (\widehat{\omega}', \widehat{K}')\right). \tag{3.23}$$

Proof. The proof of this lemma consists basically in adding and subtracting and estimating by the mean value theorem.

We first list some useful inequalities for proving this lemma:

$$\|N(\theta, \widehat{K}) - N(\theta, \widehat{K}')\| \leqslant \frac{1}{2} \operatorname{Lip}(Df)\big(\|\widehat{K}\| + \|\widehat{K}'\|\big)\|\widehat{K} - \widehat{K}'\|,$$

where $\operatorname{Lip}(Df)$ is still interpreted as the Lipschitz constant of $Df(x)$ in a neighborhood of the periodic orbit of the unperturbed ODE, as in the proof of Lemma 42.

For $\theta \in [\theta_0, \theta_0 + 1]$,

$$\begin{aligned}
&\left|\int_{\theta_0}^{\theta} \Phi(\theta_0 + 1; s)\widehat{\omega} D\widehat{K}(s)ds - \int_{\theta_0}^{\theta} \Phi(\theta_0 + 1; s)\widehat{\omega}' D\widehat{K}'(s)ds\right| \\
&\qquad \leqslant \left(2\|\Phi(\theta_0 + 1; \theta)\| + \left\|\frac{d}{d\theta}\Phi(\theta_0 + 1; \theta)\right\|\right)\left(\|\widehat{K}\||\widehat{\omega} - \widehat{\omega}'| + |\widehat{\omega}'|\|\widehat{K} - \widehat{K}'\|\right).
\end{aligned}$$

Define

$$\omega' = \omega_0 + \widehat{\omega}', \quad K' = K_0 + \widehat{K}',$$

similar to (3.7).

By assumption (H3.1),

$$|\mathscr{P}(K,\omega,\gamma,\theta)-\mathscr{P}(K',\omega',\gamma,\theta)|\leqslant \alpha_{\rho,\delta}\max\{|\omega-\omega'|,\|K-K'\|\}.$$

Then,

$$\begin{aligned}
&|\Gamma_1^\varepsilon(\widehat{\omega},\widehat{K})-\Gamma_1^\varepsilon(\widehat{\omega}',\widehat{K}')| \\
&\quad\leqslant\frac{\|\Pi_{\theta_0}^\top\|\Big(2\|\Phi(\theta_0+1;\theta)\|+\big\|\frac{d}{d\theta}\Phi(\theta_0+1;\theta)\big\|\Big)}{\big|DK_0(\theta_0)\big|}\big(\beta_0|\widehat{\omega}-\widehat{\omega}'|+a\|\widehat{K}-\widehat{K}'\|\big) \\
&\qquad+\frac{\|\Pi_{\theta_0}^\top\|\|\Phi(\theta_0+1;\theta)\|}{\big|DK_0(\theta_0)\big|}\Big[\beta_0\operatorname{Lip}(Df)\|\widehat{K}-\widehat{K}'\|+\varepsilon\alpha_{\rho,\delta}d\big((\widehat{\omega},\widehat{K}),(\widehat{\omega}',\widehat{K}')\big)\Big].
\end{aligned}\tag{3.24}$$

The initial conditions in both cases are:

$$\begin{aligned}
u_0&=\frac{1}{\omega_0}[Id-\Phi(\theta_0+1;\theta_0)]^{-1}\int_{\theta_0}^{\theta_0+1}\Pi_{\theta_0}^\perp\Phi(\theta_0+1;s)B^\varepsilon(\widehat{K},\widehat{\omega},\gamma,s)ds;\\
u_0'&=\frac{1}{\omega_0}[Id-\Phi(\theta_0+1;\theta_0)]^{-1}\int_{\theta_0}^{\theta_0+1}\Pi_{\theta_0}^\perp\Phi(\theta_0+1;s)B^\varepsilon(\widehat{K}',\widehat{\omega}',\gamma,s)ds.
\end{aligned}$$

As before, $[Id-\Phi(\theta_0+1;\theta_0)]^{-1}$ denotes the inverse of $[Id-\Phi(\theta_0+1;\theta_0)]$ in the $(n-1)$-dimensional space E_{θ_0} introduced in Section 3.3.1.

Therefore,

$$\begin{aligned}
&|u_0-u_0'| \\
&\quad\leqslant M\Big(2\|\Phi(\theta_0+1;\theta)\|+\Big\|\frac{d}{d\theta}\Phi(\theta_0+1;\theta)\Big\|\Big)\big(\beta_0|\widehat{\omega}-\widehat{\omega}'|+a\|\widehat{K}-\widehat{K}'\|\big) \\
&\qquad+M\|\Phi(\theta_0+1;\theta)\|\Big[\beta_0\operatorname{Lip}(Df)\|\widehat{K}-\widehat{K}'\|+\varepsilon\alpha_{\rho,\delta}d\big((\widehat{\omega},\widehat{K}),(\widehat{\omega}',\widehat{K}')\big)\Big],
\end{aligned}\tag{3.25}$$

where M is defined as in (3.20). Therefore,

$$\begin{aligned}
&\|\Gamma_2^\varepsilon(\widehat{\omega}, \widehat{K}) - \Gamma_2^\varepsilon(\widehat{\omega}', \widehat{K}')\| \\
&\quad \leqslant \|\Phi(\theta;\theta_0)\| |u_0 - u_0'| + \frac{\|\Phi(\theta;s)\|}{\omega_0} \|DK_0(\theta)\| |\Gamma_1^\varepsilon(\widehat{\omega}, \widehat{K}) - \Gamma_1^\varepsilon(\widehat{\omega}', \widehat{K}')| \\
&\quad + \frac{\|\Phi(\theta;\theta_0)\| + \left\|\frac{d}{ds}\Phi(\theta;s)\right\| + 1}{\omega_0} \left(\beta_0 |\widehat{\omega} - \widehat{\omega}'| + a\|\widehat{K} - \widehat{K}'\|\right) \\
&\quad + \frac{\|\Phi(\theta;s)\|}{\omega_0} \Big[\beta_0 \operatorname{Lip}(Df)\|\widehat{K} - \widehat{K}'\| + \varepsilon \alpha_{\rho,\delta} d\big((\widehat{\omega}, \widehat{K}), (\widehat{\omega}', \widehat{K}')\big)\Big].
\end{aligned} \tag{3.26}$$

Combining (3.24), (3.25), and (3.24), if ε is sufficiently small, a and β_0 are chosen to be sufficiently small, we can find μ such that (3.23) is true, Γ^ε is a contraction. □

3.4.5 Conclusion of the proof of Theorem 32

There exists a fixed point $(\widehat{\omega}^*, \widehat{K}^*)$ of contraction Γ^ε. According to Arzela-Ascoli Theorem (see Lemma 92 in Appendix), $(\widehat{\omega}^*, \widehat{K}^*) \in I_a \times \mathscr{B}_\beta$, hence is a solution of the functional equation (3.8) with desired regularity. Then, $K = K_0 + \widehat{K}^*$ gives a parameterization of the periodic orbit of (3.28).

The proof based on fixed point approach leads to a posteriori type of results. Suppose we start with initial guess $(\widehat{\omega}^0, \widehat{K}^0)$ for $(\widehat{\omega}, \widehat{K})$, since Γ^ε is contractive, see equation (3.23), we have

$$d\left((\widehat{\omega}^0, \widehat{K}^0), (\widehat{\omega}^*, \widehat{K}^*)\right) < \frac{1}{1-\mu} d\left((\widehat{\omega}^0, \widehat{K}^0), \Gamma^\varepsilon(\widehat{\omega}^0, \widehat{K}^0)\right). \tag{3.27}$$

Therefore, if we have a good choice of initial guess such that the error in the fixed point equation, $d\big((\widehat{\omega}^0, \widehat{K}^0), \Gamma^\varepsilon(\widehat{\omega}^0, \widehat{K}^0)\big)$, is small, then we know the fixed point is close to the initial guess.

Using interpolation inequalities in Lemma 91, we also have

$$\|\widehat{K}^0 - \widehat{K}^*\|_{C^m} \leqslant C\|\widehat{K}^0 - \widehat{K}^*\|_{C^0}^{1-\frac{m}{\ell}}$$

for $0 \leqslant m \leqslant \ell$, where the constant C depends on $m,\ \ell$, and β. In particular, the distance between the initial guess $(\widehat{\omega}^0, \widehat{K}^0) = (0,0)$ and the fixed point $(\widehat{\omega}^*, \widehat{K}^*)$ is of order ε, therefore, $\|\widehat{K}^*\|_{C^m}$ is small for $0 \leqslant m \leqslant \ell$. This finishes the proof of Theorem 32.

3.4.6 Comments on proof of Theorem 33

A very similar method proves Theorem 33. Now we view $\widehat{\omega}$ as a function of γ, and $\widehat{K}$ as a function of θ and γ. Define operator $\widetilde{\Gamma}^\varepsilon$ of the same format as in (3.15) and (3.16) on the space $\mathcal{I} \times \mathcal{F}$, where $\mathcal{I}$ contains $C^{\ell+\mathrm{Lip}}$ functions from set O to $\mathbb{R}$ and $\mathcal{F}$ contains $C^{\ell+\mathrm{Lip}}$ functions from $\mathbb{T} \times O$ to $\mathbb{R}^n$, with bounded derivatives similar to (3.13). We can then prove that for small enough ε, and suitable choices for bounds of derivatives, $\widetilde{\Gamma}^\varepsilon$ maps $\mathcal{I} \times \mathcal{F}$ to itself using assumption (H2.2), and is a contraction in C^0 norm, taking advantage of assumption (H3.2). Therefore, there exists a fixed point for $\widetilde{\Gamma}^\varepsilon$ in the space $\mathcal{I} \times \mathcal{F}$ solving equation (3.8). Same as above, Theorem 33 is proved.

3.5 Delay Perturbation to Autonomous ODE

In this section we show how several concrete examples fit into our general result, Theorem 32 and Theorem 33. In all the cases, we will show how to construct the operators $\mathscr{P}$ and to verify the properties in assumptions (H2) and (H3) (see 3.3.1).

3.5.1 State-dependent delay perturbation

An important class of equations that one can consider is DDEs with state-dependent delays (backward or forward or mixed):

$$\dot{x}(t) = f(x(t)) + \varepsilon P\Big(x(t), x\big(t - r(x(t))\big), \gamma\Big), \tag{3.28}$$

where $P\colon \mathbb{R}^n \times \mathbb{R}^n \times O \to \mathbb{R}^n$ is a C^∞ map, $r\colon \mathbb{R}^n \to [-h, h]$ is C^∞, h is a positive constant.

Note that in this case, the operator $\mathscr{P}$ is,

$$\mathscr{P}(K, \omega, \gamma, \theta) = P(K(\theta), \widetilde{K}(\theta), \gamma) \tag{3.29}$$

where $\widetilde{K}(\theta) \coloneqq K(\theta - \omega r(K(\theta)))$ is caused by the delay.

Remark 46. *Note that the operator $\mathscr{P}$ involves the composition operator, whose differentiability properties are very complicated (See [54] for a systematic study). Hence, using the standard strategy of studying variational equations etc. to study regularity of the evolution will be rather complicated. Indeed, it will be hard to go beyond the first derivative.*

On the other hand, the present strategy, only requires much simpler results. We only need to get bounds on the derivatives of $\mathscr{P}$ assuming bounds on the derivatives of K.

Applying the composition Lemma 87 repeatedly, we know that $\mathscr{P}$ above satisfies (H2.1). With the standard adding and subtracting terms method, one gets that $\mathscr{P}$ satisfies (H3.1). Similarly, $\mathscr{P}$ satisfies (H2.2) and (H3.2). Thus, Theorems 32 and 33 can be applied.

Note also that for the above equation (3.28), we are able to prove that the operator Γ^{ε} is a contraction under $C^{\ell-1+\mathrm{Lip}}$ norm in the second component, by using Lemma 90.

We may improve the regularity conclusion of Theorem 32 for this case. Indeed, once we have that the parameterization K of the periodic orbit is C^1 in θ, we can use the standard bootstrapping argument to conclude higher regularity of K based on the smoothness of the equation, see Remark 50.

We can also consider more general state-dependent delays:

$$\dot{x}(t) = f(x(t)) + \varepsilon P\Big(x(t), x\big(t - r\left(x_t, \gamma\right)\big), \gamma\Big), \tag{3.30}$$

where $r \colon \mathscr{R}[-h, h] \times O \to \mathbb{R}$, positive constant h is an upper bound for $|r|$.

In this case,

$$\mathscr{P}(K,\omega,\gamma,\theta) = P\left(K(\theta), K\big(\theta - \omega r(K_{\theta,\omega},\gamma)\big), \gamma\right), \tag{3.31}$$

where $K_{\theta,\omega} \colon [-h, h] \to \mathbb{R}^n$ is defined by

$$K_{\theta,\omega}(s) \coloneqq K(\theta + \omega s). \tag{3.32}$$

If r is chosen such that (H2) and (H3) are verified, Theorems 32 and 33 can be applied.

3.5.2 Distributed delay perturbation

Our results apply to models with distributed delays as well

$$\dot{x}(t) = f(x(t)) + \varepsilon P\Big(x(t), \int_{-r}^{0} x_t(s)d\mu(s), \gamma\Big), \tag{3.33}$$

where $P \colon \mathbb{R}^n \times \mathbb{R}^n \times O \to \mathbb{R}^n$ is a C^∞ map, r is a constant, and μ is a signed Borel measure. In this case,

$$\mathscr{P}(K,\omega,\gamma,\theta) = P\left(K(\theta), \int_{-r}^{0} K_{\theta,\omega}(s)d\mu(s), \gamma\right), \tag{3.34}$$

where $K_{\theta,\omega}$ is defined in (3.32).

Above $\mathscr{P}$ verifies (H2), since we only care about derivatives with respect to θ. It is not hard to see that (H3) is also satisfied in this case. Therefore, Theorems 32 and 33 apply.

3.5.3 Remarks on further applicability of Theorem 32

Remark 47. *It is straightforward to see that our results could be applied to systems similar to above systems with multiple forward or backward delays.*

Remark 48. *In some applications, the delays are defined by some implicit relations from the full trajectory.*

Theorems 32 and 33 can be applied if we can justify (H2.1)*,* (H3.1)*,* (H2.2)*, and* (H3.2)*. Notice that we only need to justify these hypothesis when $\widehat{K}$ lies in ball in a space of differentiable functions. In such a case, we can often use the implicit function theorem.*

Remark 49. *The results so far do not include the models in which the perturbation is just adding a small delay. This small delay perturbation is more singular and seems to require extra assumptions and slightly different proofs. The extension of the results to the small delay case is done in Section 3.7.*

Remark 50. *In the case of state dependent delay or distributed delay with a smooth f and r, it is automatic to show that if K is C^ℓ, the right hand side of* (3.4) *is C^ℓ, hence, looking at the left hand side of* (3.4)*, K is $C^{\ell+1}$. The bootstrap stops only when we do not have any more regularity of f, P, or r.*

So, in case that f, P, and r are C^∞, we obtain that the K is C^∞.

One natural question that deserves more study is whether in the case that f, P, and r are analytic, the K is analytic. The remarkable paper [40] contains obstructions that equations with time dependent delays – heuristically better behaved than the ones considered here, may fail to have analytic solutions. In view of these results, it is natural to conjecture that the periodic solutions produced here, could fail to be analytic even if f, P, and r are analytic.

3.6 Delay Perturbation to Non-Autonomous Periodic ODE

Time periodic systems appear in many problems in physics, for example, see Section 3.8. And when there are conserved quantities in the ODE systems, periodic orbits cannot satisfy the assumption (H1). These are the motivations to consider a non-autonomous ODE:

$$\dot{x}(t) = f(x(t), t), \tag{3.35}$$

where $f\colon \mathbb{R}^n \times \frac{1}{\omega_0}\mathbb{T} \to \mathbb{R}^n$ (f is periodic in t with period $\frac{1}{\omega_0}$). Add the perturbation:

$$\dot{x}(t) = f(x(t), t) + \varepsilon P(x_t, \gamma), \tag{3.36}$$

Using the standard method of adding an extra variable to equation (3.35) to make it autonomous, we see that we can reduce the problem to the previous case. The autonomous equation corresponding to (3.35) is

$$\begin{pmatrix} \dot{x}(t) \\ \dot{\tau}(t) \end{pmatrix} = g(x, \tau) \coloneqq \begin{pmatrix} f(x, \tau) \\ 1 \end{pmatrix}. \tag{3.37}$$

Denote Ψ as the solution of the variational equation for the periodic orbit of (3.37):

$$\omega_0 \frac{d}{d\theta}\Psi(\theta; \theta_0) = Dg\left(K_0(\theta), \frac{\theta}{\omega_0}\right)\Psi(\theta; \theta_0), \quad \Psi(\theta_0; \theta_0) = Id. \tag{3.38}$$

Since

$$Dg\left(K_0(\theta), \frac{\theta}{\omega_0}\right) = \begin{pmatrix} D_1 f\left(K_0(\theta), \frac{\theta}{\omega_0}\right) & D_2 f\left(K_0(\theta), \frac{\theta}{\omega_0}\right) \\ 0 & 0 \end{pmatrix},$$

we have

$$\Psi(\theta; \theta_0) = \begin{pmatrix} \Phi(\theta; \theta_0) & * \\ 0 & 1 \end{pmatrix}, \tag{3.39}$$

where

$$\omega_0 \frac{d}{d\theta}\Phi(\theta; \theta_0) = D_1 f\left(K_0(\theta), \frac{\theta}{\omega_0}\right)\Phi(\theta; \theta_0). \tag{3.40}$$

If Ψ satisfies assumption (H1), then 1 is not an eigenvalue of $\Phi(\theta_0 + 1; \theta_0)$.

Equivalently, we could start the discussion in this section directly with the following assumption on Φ defined in (3.40):

(H1”) 1 is not an eigenvalue of $\Phi(\theta_0 + 1; \theta_0)$.

Under either assumption (H1) on Ψ or assumption (H1”) on Φ, we are able to solve the invariance equation (3.4) without adjusting the frequency. More precisely, (3.4) becomes:

$$\omega_0 DK(\theta) = f\left(K(\theta), \frac{\theta}{\omega_0}\right) + \varepsilon \mathscr{P}(K, \omega_0, \gamma, \theta), \tag{3.41}$$

Let $K = K_0 + \widehat{K}$ as in (3.7), we are led to

$$\omega_0 D\widehat{K}(\theta) - D_1 f\left(K_0(\theta), \frac{\theta}{\omega_0}\right)\widehat{K}(\theta) = B^\varepsilon(\theta, \omega_0, \widehat{K}, \gamma), \tag{3.42}$$

where

$$\begin{aligned}
&B^\varepsilon(\theta, \omega_0, \widehat{K}, \gamma) \coloneqq N(\theta, \widehat{K}) + \varepsilon \mathscr{P}(K, \omega_0, \gamma, \theta),\\
&N(\theta, \widehat{K}) \coloneqq f\left(K(\theta), \frac{\theta}{\omega_0}\right) - f\left(K_0(\theta), \frac{\theta}{\omega_0}\right) - D_1 f\left(K_0(\theta), \frac{\theta}{\omega_0}\right)\widehat{K}(\theta).
\end{aligned}$$

Now we can define an operator Υ^ε on the space $\mathscr{B}_\beta$ (see (3.13)) very similar to the second component of Γ^ε introduced in section 3.4.2.

$$\Upsilon^\varepsilon(\widehat{K})(\theta) \coloneqq \Phi(\theta; \theta_0)u_0 + \frac{1}{\omega_0}\int_{\theta_0}^{\theta} \Phi(\theta; s) B^\varepsilon(s, \omega_0, \widehat{K}, \gamma)\, ds, \tag{3.43}$$

where

$$u_0 = \frac{1}{\omega_0}[Id - \Phi(\theta_0 + 1; \theta_0)]^{-1} \int_{\theta_0}^{\theta_0+1} \Phi(\theta_0 + 1; s) B^\varepsilon(s, \omega_0, \widehat{K}, \gamma)\, ds. \tag{3.44}$$

We have employed that, in the periodic case, the matrix $[Id - \Phi(\theta_0 + 1; \theta_0)]$ is invertible.

Under the assumption that $\mathscr{P}$ satisfies (H2.1), (H3.1), (H2.2) and (H3.2), we can prove that Υ^ε has a fixed point $\widehat{K}^* \in \mathscr{B}_\beta$ by proving $\Upsilon^\varepsilon \colon \mathscr{B}_\beta \to \mathscr{B}_\beta$ (similar to Lemma 42) and Υ^ε is a contraction (similar to Lemma 45). The periodic orbit of (3.36) is parameterized by $K = K_0 + \widehat{K}^*$. The analysis of the operator Υ^ε in (3.43) is actually simpler that the analysis presented for the operator Γ^ε in (3.14) because we do not need to adjust the frequency.

Remark 51. *Similarly, we can also consider a non-autonomous perturbation $P(t, x_t, \gamma)$, we need that P to be periodic in t with the same period $\frac{1}{\omega_0}$.*

3.7 The Case of Small Delays

Many problems in the literature lead to equations of the form:

$$\begin{aligned} y'(t) &= g(y(t - \varepsilon r)) \\ y'(t) &= f(y(t - \varepsilon r), t) \end{aligned} \tag{3.45}$$

where r could be either a constant, an explicit function of t, a function of $y(t)$, or y_t, or an implicit function, and may depend on ε; and f is periodic of period 1 in t. Indeed, our results apply also to variants of (3.45) with perturbations involving several forward or backward delays.

In problems which present feedback loops, the feedback takes some time to start acting. The problems (3.45) correspond to the feedback taking a short time to start acting.

Equations of the form (3.45) play an important role in electrodynamics, where the small parameter $\varepsilon = \frac{1}{c}$ is the inverse of the speed of light and the delay r is a functional that depends on the trajectory. Given the physical importance of electrodynamics, we devote Section 3.8 to give more details and to show that it can be studied applying the main result of this section, Theorem 53.

Introducing a small delay to the ODE is a very singular perturbation, since the phase space becomes infinite dimensional. The limit is mathematically harder because the effect of a small delay is similar to adding an extra term containing the derivative $y(t - \varepsilon r) \approx y(t) + \varepsilon y'(t) r$. This shows that, heuristically, the perturbation is of the same order as the equation.

Remark 52. *In the physical literature, one can find the use of higher order expansions to obtain heuristically even higher order equations, see [55]. As a general theory for all the* *solutions of the equations, these theories have severe paradoxes (e.g.* preacceleration*). The results of this section show, however that the non-degenerate periodic solutions produced in many of these expansions, since they are very approximate solutions of the invariance equation, approximate true periodic solutions of the full system.*

As a reflection of the extra difficulty of the small delay problem compared with the previous ones, the main result of this section, Theorem 53, requires a more delicate proof than Theorem 32 and we need stronger regularity to obtain the C^0 contraction.

An important mathematical paper on the singular problem of small delay is [51]. We also point out that, there is a considerable literature in the formal study of $\frac{1}{c}$ limit in electrodynamics and in gravity [56, 57, 58, 59]. Many famous consequences of relativity theory (e.g. the precession of the perihelion of Mercury) are only studied by formal perturbations.

Formal expansions of periodic and quasiperiodic solutions for small delays were considered in [30]. The results of this section establish that the formal expansions of periodic orbits obtained in [30] correspond to true periodic orbits and are asymptotic to the true periodic solutions in a very strong sense.

In this section, we establish results on persistence of periodic orbits for the models in (3.45), see Theorem 53. As we will see, when we perform the detailed discussion, we will not be able to reduce Theorem 53 to be a particular case of Theorem 32. The proof of Theorem 53 will be very similar to that of Theorem 32 and which is based on the study of operator Γ^ε very similar to those in (3.14). Nevertheless, the analysis of the operator Γ^ε in the current case will require to take advantage of an extra cancellation.

3.7.1 Formulation of the results

Our main result for the small delay problem (3.45) is as follows. Without specifying the delay functional r, we will use $r(\omega, K, \varepsilon)$ to denote the expression after substituting $K(\theta + \omega t)$ into r and letting $t = 0$.

Theorem 53. *For integer $\ell \geqslant 3$, assume that the function g (resp. f) in* (3.45) *is $C^{\ell+\mathrm{Lip}}$.*

Assume that for $\varepsilon = 0$, the ordinary differential equation $y' = g(y)$ has a periodic orbit satisfying (H1).*(resp. $y' = f(y,t)$ has a periodic orbit satisfying* (H1”) *). We denote by K_0 the parameterization of this periodic orbit with frequency ω_0.*

Recall that U_ρ is the ball of radius ρ in $C^{\ell+\mathrm{Lip}}(\mathbb{T},\mathbb{R}^n)$ centered at K_0, and B_δ is the interval with radius δ centered at ω_0.

Recall distance d defined in (3.22). *Assume that the delay functional r satisfies*

$$\|r(\omega,K,\varepsilon)\|_{C^{\ell-1+\mathrm{Lip}}(\mathbb{T},\mathbb{R}^n)} \leqslant \phi_{\rho,\delta}(\varepsilon) \quad \forall K \in U_\rho,\ \omega \in B_\delta \tag{3.46}$$

for some $\phi_{\rho,\delta}(\varepsilon) > 0$, with $\varepsilon\phi_{\rho,\delta}(\varepsilon) \to 0$ as $\varepsilon \to 0$. And for any $K_1, K_2 \in U_\rho$, $\omega_1, \omega_2 \in B_\delta$, there is $\alpha_{\rho,\delta}(\varepsilon) > 0$, with $\varepsilon\alpha_{\rho,\delta}(\varepsilon) \to 0$ as $\varepsilon \to 0$, such that

$$\|r(\omega_1,K_1,\varepsilon) - r(\omega_2,K_2,\varepsilon)\|_{C^0} \leqslant \alpha_{\rho,\delta}(\varepsilon) d\big((\omega_1,K_1),(\omega_2,K_2)\big). \tag{3.47}$$

Then, there exists $\varepsilon_0 > 0$ such that for $\varepsilon < \varepsilon_0$, the problem (3.45) *admits a periodic solution. There is a $C^{\ell+\mathrm{Lip}}$ parameterization K of the periodic orbit which is close to K_0 in the sense of C^ℓ.*

Remark 54. *As before, the requirements of smallness in ε for Theorem 53 depend on the regularity considered.*

In many applied situations, the g and f considered are C^∞ or even analytic. (for example in the electrodynamics applications considered in Section 3.8). In such a case, we can consider any ℓ by assuming ε is small enough.

This allows us to obtain the a posteriori estimates in more regular spaces as ε goes to zero.

Hence, the formal power series in [30] are asymptotic in the strong sense that the error in the truncation is bounded by a power of ε, where a stronger norm can be used for smaller ε.

We leave for the reader the formulation of a corresponding result for the smooth dependence on parameters similar to Theorem 33. The proof requires only small modifications from discussion in Section 3.7.2, see comments in Section 3.4.6.

The proof of Theorem 53 will be given in Section 3.7.2. We first find the operator in this case. Then for the operator, we prove Lemma 42 in Section 3.7.2.1, and prove Lemma 45 in Section 3.7.2.2. The existence of fixed point of the operator is thus established. As it turns out, the analysis of the operator requires more care than in the case of Theorem 32.

3.7.2 Existence of fixed point

The equations (3.45) can be rearranged as

$$\begin{aligned} y'(t) &= g(y(t)) + [g(y(t-\varepsilon r)) - g(y(t))] \\ &= g(y(t)) - \varepsilon \int_0^1 \big[Dg\big(y(t-s\varepsilon r)\big) Dy(t-s\varepsilon r) r\big]\, ds \\ y'(t) &= f(t, y(t)) + [f(t, y(t-\varepsilon r)) - f(t, y(t))] \\ &= f(t, y(t)) - \varepsilon \int_0^1 \big[D_1 f\big(y(t-s\varepsilon r), t\big) Dy(t-s\varepsilon r) r\big]\, ds \end{aligned} \tag{3.48}$$

For typographical convenience, we will discuss only the autonomous case, which is the most complicated. We refer the reader to Section 3.6 to see how the discussion simplifies in the periodic case (the most relevant case for applications to electrodynamics).

Note that (3.48) is in the form of (3.2), with the operator P defined as

$$P(y_t) \coloneqq - \int_0^1 \big[Dg\big(y(t-s\varepsilon r)\big) Dy(t-s\varepsilon r) r\big]\ ds. \tag{3.49}$$

Then,

$$\mathscr{P}(K, \omega, \gamma, \theta) \coloneqq - \int_0^1 \big[Dg\big(K(\theta - \varepsilon s\omega r)\big) DK(\theta - \varepsilon s\omega r)\omega r\big]\, ds, \tag{3.50}$$

where the r's are $r(\omega, K, \varepsilon)$, the delay functional evaluated on the periodic orbit.

We define operator Γ^ε in the same way as in Section 3.4, substituting the expression of $\mathscr{P}$ in (3.50) into the general formula in (3.14).

In this section, we will proceed as before and show Lemmas 42 and 45 are true for the resulting operator Γ^ε with $\mathscr{P}$ defined in (3.50).

Lemma 42 is proven in this case, same as above, by noticing $\mathscr{P}$ satisfies assumption (H2.1). The proof for Lemma 45 is slightly different from before. In Section 3.4, we only needed to take advantage of the Lipschitz property of the operator $\mathscr{P}$ (assumption (H3.1)). In the present case, we will have to take into account that the operator Γ^ε involves not only $\mathscr{P}$, but also an integral, which has nice properties that compensate the bad properties of $\mathscr{P}$.

3.7.2.1 Propagated bounds

We observe that if $K \in U_\rho$ and $\omega \in B_\delta$, by the assumption (3.46), $r(\omega, K, \varepsilon)$ is in a $C^{\ell-1+\text{Lip}}$ ball of size $\phi_{\rho,\delta}(\varepsilon)$ and, using the estimates on composition, Lemma 87, so is $K(t-\varepsilon s\omega r)$. If $g \in C^{\ell+\text{Lip}}$, then $Dg \in C^{\ell-1+\text{Lip}}$ and we conclude that $Dg \circ K(t-\varepsilon s\omega r)$ is contained in a $C^{\ell-1+\text{Lip}}$ ball.

We also have that if K is in a $C^{\ell+\text{Lip}}$ ball, $DK(t-\varepsilon s\omega r)\omega r$ is in a $C^{\ell-1+\text{Lip}}$ ball whose size is a function of ρ and $\phi_{\rho,\delta}(\varepsilon)$.

Putting it all together we obtain that (H2.1) is true for $\mathscr{P}$ defined in (3.50). Therefore, Lemma 42 is proven in this case.

3.7.2.2 Contraction in C^0

Before estimating $\Gamma^\varepsilon(\widehat{\omega}, \widehat{K}) - \Gamma^\varepsilon(\widehat{\omega}', \widehat{K}')$, we estimate $\mathscr{P}(K, \omega, \gamma, \theta) - \mathscr{P}(K', \omega', \gamma, \theta)$ (we denote by r, r' the two delay terms corresponding to ω, K, and ω', K' respectively).

As usual, adding and subtracting, we obtain that the difference in the integrands in $\mathscr{P}$,

$$Dg\big(K(\theta-\varepsilon s\omega r)\big)DK(\theta-\varepsilon s\omega r)\omega r - Dg\big(K'(\theta-\varepsilon s\omega' r')\big)DK'(\theta-\varepsilon s\omega' r')\omega' r'$$

can be written as a sum of 8 differences in which only one of the objects changes, see (3.51) below. As it turns out, 7 of them will be straightforward to estimate and only one of them will require some effort. We give the details.

$$
\begin{aligned}
&[Dg\big(K(\theta - \varepsilon s \omega r)\big) - Dg\big(K'(\theta - \varepsilon s \omega r)\big)]DK(\theta - \varepsilon s \omega r)\omega r \\
&+ [Dg\big(K'(\theta - \varepsilon s \omega r)\big) - Dg\big(K'(\theta - \varepsilon s \omega' r)\big)]DK(\theta - \varepsilon s \omega r)\omega r \\
&+ [Dg\big(K'(\theta - \varepsilon s \omega' r)\big) - Dg\big(K'(\theta - \varepsilon s \omega' r')\big)]DK(\theta - \varepsilon s \omega r)\omega r \\
&+ Dg\big(K'(\theta - \varepsilon s \omega' r')\big)[DK - DK'](\theta - \varepsilon s \omega r)\omega r \\
&+ Dg\big(K'(\theta - \varepsilon s \omega' r')\big)[DK'(\theta - \varepsilon s \omega r) - DK'(\theta - \varepsilon s \omega' r)]\omega r \\
&+ Dg\big(K'(\theta - \varepsilon s \omega' r')\big)[DK'(\theta - \varepsilon s \omega' r) - DK'(\theta - \varepsilon s \omega' r')]\omega r \\
&+ Dg\big(K'(\theta - \varepsilon s \omega' r')\big)DK'(\theta - \varepsilon s \omega' r')(\omega - \omega')r \\
&+ Dg\big(K'(\theta - \varepsilon s \omega' r')\big)DK'(\theta - \varepsilon s \omega' r')\omega'(r - r')
\end{aligned}
\tag{3.51}
$$

All the terms except for the 4th term are straightforward to estimate in C^0 by some constant multiple of $d\big((\widehat{\omega}, \widehat{K}), (\widehat{\omega}', \widehat{K}')\big)$, keeping in mind bounds on the $C^{\ell+\mathrm{Lip}}$ norms of g, K, K', and r $\big($see assumption (3.46)$\big)$, and the assumption (3.47). We consider the first term for an example, the rest is similar.

$$
\begin{aligned}
&\|[Dg\big(K(\theta - \varepsilon s \omega r)\big) - Dg\big(K'(\theta - \varepsilon s \omega r)\big)]DK(\theta - \varepsilon s \omega r)\omega r\|_{C^0} \\
&\qquad\qquad\qquad\qquad \leqslant \omega \|D^2 g\| \|DK\| \|r\|_{C^0} \|\widehat{K} - \widehat{K}'\|_{C^0}
\end{aligned}
\tag{3.52}
$$

Observe the form of the operator Γ^ε in (3.14). Note that if we have a bound of

$$
\int_{\theta_0}^{\theta} \Phi(\theta; s)\big(\mathscr{P}(K, \omega, \gamma, s) - \mathscr{P}(K', \omega', \gamma, s)\big) ds.
$$

by a multiple of $d\big((\widehat{\omega}, \widehat{K}), (\widehat{\omega}', \widehat{K}')\big)$, we prove Lemma 45.

All terms except the 4th one in (3.51) are controlled using estimates similar to (3.52). Hence, to complete the proof, we just need to estimate the part coming from the 4th term in (3.51). We will take advantage of the integral which is an operator that improves the bounds.

We use integration by parts to get:

$$\begin{aligned}
&\int_{\theta_0}^{\theta} \Phi(\theta; s) \int_0^1 Dg\big(K'(s-\varepsilon\tau\omega' r')\big)[DK - DK'](s-\varepsilon\tau\omega r)\omega r d\tau ds \\
&= \int_0^1 \int_{\theta_0}^{\theta} \Phi(\theta; s) Dg(\cdot) \frac{\omega r}{1-\varepsilon\tau\omega\frac{dr}{ds}}(1-\varepsilon\tau\omega\frac{dr}{ds})[DK - DK'](s-\varepsilon\tau\omega r) ds d\tau \\
&= \int_0^1 \Big[\Phi(\theta; s) Dg(\cdot) \frac{\omega r}{1-\varepsilon\tau\omega\frac{dr}{ds}}[K-K'](s-\varepsilon\tau\omega r)\Big|_{s=\theta_0}^{s=\theta} \\
&\qquad - \int_{\theta_0}^{\theta} \frac{d}{ds}\Big(\Phi(\theta; s) Dg(\cdot)\frac{\omega r}{1-\varepsilon\tau\omega\frac{dr}{ds}}\Big)[K-K'](s-\varepsilon\tau\omega r) ds\Big] d\tau
\end{aligned}$$

The C^0 norm of the above expression is bounded by a multiple of $\|K - K'\|_{C^0} \leqslant d\big((\widehat{\omega}, \widehat{K}), (\widehat{\omega}', \widehat{K}')\big)$, we have proved Lemma 45 in this case. The proof of Theorem 53 is finished.

Remark 55. *Note that we need to differentiate r along the periodic orbit twice in the above expression, that is why we required $\ell \geqslant 3$ in Theorem 53, so that $r(\omega, K, \varepsilon)$ is more than C^2.*

3.8 Delays Implicitly Defined by the Solution. Applications to Electrodynamics

In this section, we show how to deal with delays that depend implicitly on the solution. The main motivation is electrodynamics, so we deal with this case in detail, but we formulate a more general mathematical result in Section 3.8.3.

We point out that implicitly defined delays appear naturally in other problems in which the delay of the effect is related to the state of the system. As we indicate later, the explicit state dependent delays appeared in (3.28) are often approximations of implicitly defined delays. One corollary of our treatment is a justification of the fact that the periodic solutions of this approximation are an approximation to the true periodic solutions.

3.8.1 Motivation from electrodynamics

One of the original motivations for the whole field of delay equations was the study of forces in electrodynamics. The forces among charged particles, depend on the positions of the particles. Since the signals from a particle take time to reach another particle, this leads to a delay equation. Notice that the delay depends on the position (at a previous time) so that the delay is obtained by an implicit equation on the trajectory. This formulation was proposed very explicitly in [1], which we will follow.

Remark 56. *An alternative description of electrodynamics uses the concept of fields. One problem of the concept of fields is to explain why particles do not interact with their own fields. We refer to [60] for a very lucid physical discussion of the paradoxes faced by a coherent formulation of classical electrodynamics.*

Remark 57. *Many Physicists object to [1] that it does not make clear what is the phase space and what are the initial conditions.*

In this Chaper, we show that one does not need to answer these question to construct a theory of periodic solutions. We hope that similar results hold for other types of solutions. So that one can have a systematic theory of many solutions that resemble the classical ones.

Of course, it should also be possible to construct other solutions that are completely different from those of the systems without delays.

Remark 58. *Even if one can have a rich theory of perturbative solutions, It is not clear that these solutions fit together in a smooth manifold. The paper [30] develops asymptotic expansions, which suggests that the resulting solutions may be difficult to fit together in a manifold.*

We speculate that this may give a way to reconcile the successes of predictive mechanics *[58] with the no-interaction theorems [25]. It could well happen that the results* *of predictive mechanics apply to the abundant solutions we construct, but, according to the no-interaction theorem, this set cannot be all the initial conditions. Of course, these speculations are far from being theorems.*

3.8.2 Mathematical formulation

If we consider (time-dependent) external and magnetic fields as prescribed, the equations of a system of N particles in $\mathbb{R}^3$ are, denoting by $q_i(t)$ the position of the i-th particle.

$$q_i''(t) = A_{\text{ext}}(t, q_i(t), q_i'(t)) + \sum_{j \neq i} A_{i,j}\left(q_i(t), q_i'(t), q_j(t - \tau_{ij}), q_j'(t - \tau_{ij})\right) \tag{3.53}$$

where the time delay is defined implicitly by (c is the speed of light)

$$\tau_{ij}(t) = \frac{1}{c}\left|q_i(t) - q_j(t - \tau_{ij}(t))\right|. \tag{3.54}$$

For more explicit expressions, we refer to [1, 61, 3]. We just remark that (3.53) is the usual equation of acceleration equals force divided by the mass. The relativistic mass has some complicated expression depending on the velocity.

The term A_{ext} denotes the external force. The terms A_{ij} correspond to the Coulomb and Lorenz forces of the fields obtained from Liénard–Wiechert potentials. This is a standard calculation which is classical in electrodynamics, see [56, 62, 63]. Roughly, they are the Coulomb and Ampere (electric and magnetic) forces at previous times but some derivative terms appear.

We observe that (3.53) is in the form imposed by the principle of relativity, and that any force which is relativistically invariant should have the form (3.53) with, of course, different expressions for the terms A_{ij}. Hence, the treatment discussed here should apply not only to electrodynamics but also to any forces subject to the rules of special relativity.

The exact form of the equations does not play an important role here. We point out some properties that play a role:

1. The expressions defining the forces are algebraic expressions. They have singularities when there are collisions ($q_i(t) = q_j(t)$ for some $i \neq j$) or when some particle reach the speed of light ($|q_i'(t)| = c$ for some i).

2. The delays τ_{ij} as in (3.54) are subtle. The expression of τ_{ij} involve a small parameter $\varepsilon \coloneqq 1/c$, and the delays can be approximated in first order as:

$$\tau_{ij}(t) = \varepsilon\big|q_i(t) - q_j(t)\big| + O(\varepsilon^2). \tag{3.55}$$

   Keeping only the first order approximation in (3.55) makes (3.53) an SDDE, but with (3.54), the delay depends implicitly on the trajectory.

   Note that it is not true that $\tau_{ij} = \tau_{ji}$ even if this symmetry is true in the first order approximation (3.55).

3. In the case that $\tau_{ij} = 0$ and that the external forces are autonomous, the energy is conserved. This has two consequences:

   - In the autonomous case, the periodic orbits do not satisfy the hypothesis (H1). Hence, we will only make precise statements in the case of time periodic external fields. In this case (very well studied in accelerator physics, plasma, etc.), there are many examples of periodic orbits satisfying assumption (H1"), so that the results presented here are not vacuous.
   - If the external potential and external magnetic fields are bounded, the periodic orbits of finite energy and away from collisions satisfy $|q_i'(t)| \leqslant \xi_1 c$ $\big($where $\xi_1 \in (0,1)\big)$and $|q_i(t) - q_j(t)| \geqslant \xi_2 > 0$, $i \neq j$. We will assume these two properties.

Denoting $y(t) \coloneqq (q_1(t), \dots, q_N(t), q_1'(t), \dots q_N'(t))$, we can write the equation (3.53) in the form of (3.45) with the delays being implicitly defined. Note that there are $N(N-1)$ delays in total.

Remark 59. *Even if we formulate the result for the retarded potentials, we point out that the mathematical treatment of Maxwell equations admits also advanced potentials.*

It is customary to take only the retarded potentials because of "physical reasons" which are relegated to footnotes in most classical electrodynamics books. More detailed discussions appear in [61, 60]. Note, that selecting only retarded potentials breaks, even at the classical level, the time-reversibility present in Maxwell's and Newton's equations. Mathematically any combination of advanced and retarded potentials would make sense from Maxwell equations. Indeed, [64] proposes a theory with half advanced and half retarded potentials.

We do not want to enter now into the physical arguments, which should be decided by experiment (we are not aware of explicit experimentation of these points). We just point out that the mathematical theory here and the asymptotic expansions [30] applies to retarded, advanced, or combination of advanced and retarded potentials.

3.8.3 Mathematical results for electrodynamics

In this section, we will collect the ideas we have been establishing and formulate our main result for the model (3.53). Note that we formulate the result only for periodic external fields, since when the external fields are time-independent, energy is conserved which prevents periodic orbits from satisfying assumption (H1).

We will assume that there exists $0 < \xi_1 < 1$, and $\xi_2 > 0$, such that for all t:

$$
\begin{aligned}
&|q_j'(t)| \leqslant \xi_1 c \\
&|q_i(t) - q_j(t)| \geqslant \xi_2
\end{aligned}
\tag{3.56}
$$

Note that (3.56) implies that the internal forces and the masses are analytic around the trajectory. Therefore, the regularity assumptions for the equation concern only the external fields.

Theorem 60. *Denote $\varepsilon = 1/c$. Consider the model* (3.53) *with the delays defined in* (3.54).

Assume that for $\varepsilon = 0$, the resulting time periodic ODE has periodic solution satisfying hypothesis (H1") *as well as* (3.56). *Assume that the external fields A_{ext} are $C^{\ell+\mathrm{Lip}}$.*

Then, for small enough ε, we can find a $C^{\ell+\mathrm{Lip}}$ periodic solution of (3.53).

Similarly, in the case that the external fields are jointly $C^{\ell+\mathrm{Lip}}$ in time, position, velocity, and in a parameter γ, the periodic solutions are jointly $C^{\ell+\mathrm{Lip}}$ as functions of the variable of the parameterization and the parameter γ.

The proof follows the steps of Theorem 53 once we have the estimates on delays (3.46) and (3.47), which will be discussed in the next section.

Remark 61. *Since the fully relativistic equations are cumbersome to handle, there are many approximations in the literature. [65, 66, 67, 68, 69] approximate the relativistic equations up to $O(c^{-m})$. Theorem 60 ensures that the non-degenerate periodic orbits of the case when $c = \infty$ persist in these models. Furthermore, due to the a posteriori format of Theorem 60, we obtain that these periodic orbits are $O(c^{-m})$ close to periodic orbits of the relativistic model.*

3.8.4 Some preliminary results on the regularity of the delay

In this section, we study (3.54) as an equation for $\tau_{ij}(t)$ when we prescribe the trajectories q_i and q_j. This makes precise the notion that the delay is a functional of the whole trajectory.

In the following proposition, we collect the proofs of estimates that establish (3.46) and (3.47). Both follow rather straightforwardly from considering (3.54) as a contraction mapping.

Proposition 62. *Let q_i and q_j be continuously differentiable trajectories that satisfy* (3.56).

Then, for each $t \in \mathbb{R}$, we can find a unique $\tau_{ij}(t) > 0$ solving (3.54).

Moreover:

If the trajectories q_i and q_j are $C^{\ell+\mathrm{Lip}}$, then the τ_{ij} is $C^{\ell+\mathrm{Lip}}$. There is an explicit expression

$$\|\tau_{ij}\|_{C^{\ell+\mathrm{Lip}}} \leqslant \phi(\|q_i\|_{C^{\ell+\mathrm{Lip}}}, \|q_j\|_{C^{\ell+\mathrm{Lip}}}, \xi_1, \xi_2). \tag{3.57}$$

Let $q_i, q_j, \bar{q}_i$, and $\bar{q}_j$ be trajectories satisfying (3.56). *Denote by τ_{ij} and $\bar{\tau}_{ij}$ the solutions of* (3.54) *corresponding to q_i, q_j and to $\bar{q}_i, \bar{q}_j$, respectively. Then we have:*

$$\|\tau_{ij} - \bar{\tau}_{ij}\|_{C^0} \leqslant C(\xi_1, \xi_2) \left(\|q_i - \bar{q}_i\|_{C^0} + \|q_j - \bar{q}_j\|_{C^0}\right). \tag{3.58}$$

Proof. Fix t and, hence, $q_i(t)$.

We treat (3.54) as a fixed point problem for the – long named – unknown $\tau_{ij}(t)$ with the functions q_i, q_j as well as the number t fixed.

The first part of the asumption (3.56) implies that the RHS of (3.54), as a function of $\tau_{ij}(t)$ has derivative with modulus bounded by $\xi_1 < 1$. Hence, we can apply the contraction mapping principle. This establishes existence and uniqueness.

Moreover, we can apply the implicit function theorem and obtain that $\tau_{ij}(t)$ is as differentiable on t as the RHS of (3.54). Furthermore, we can get expressions for $\frac{d^k}{dt^k}\tau_{ij}(t)$ which are algebraic expressions involving derivatives with respect to t of $q_i(t),\ q_j(t)$ up to order k, and derivatives of $\tau_{ij}(t)$ up to order $k-1$. The exact combinatorial formulas are very well known. Using recurrence in the order of derivatives, we obtain (3.57).

To prove (3.58), we observe that since the contraction we used before is uniform in t, we can consider the RHS of (3.54) as a contraction in C^0.

We evaluate the RHS of (3.54) corresponding to $\bar{q}_i$ and $\bar{q}_j$ on τ_{ij}, note that

$$\begin{aligned}\bar{q}_i(\cdot) - \bar{q}_j(\cdot - \tau_{ij}(\cdot)) = {} & (\bar{q}_i(\cdot) - q_i(\cdot)) + (q_j(\cdot - \tau_{ij}(\cdot)) - \bar{q}_j(\cdot - \tau_{ij}(\cdot)) \\ & + (q_i(\cdot) - q_j(\cdot - \tau_{ij}(\cdot)))\,.\end{aligned}$$

Hence,

$$\left\| \frac{1}{c}\left|\bar{q}_i(\cdot) - \bar{q}_j(\cdot - \tau_{ij}(\cdot))\right| - \tau_{ij}(\cdot)\right\|_{C^0} \leqslant \frac{1}{c}\|q_i - \bar{q}_i\|_{C^0} + \frac{1}{c}\|q_j - \bar{q}_j\|_{C^0}$$

From this, (3.58) follows from the Banach contraction mapping. □

Remark 63. *Notice that the delays τ_{ij}'s contain small factor $\frac{1}{c}$, so are the right hands of the inequalities* (3.57) *and* (3.58)*, as we can see in the proof above. We can view $\tau_{ij} := \frac{1}{c} r_{ij}$ to fit in the case of small delays.*

3.9 The Case of Hyperbolic Periodic Orbits

Our main result Theorems 32 and 33 are based on the assumption (H1), which is automatically satisfied when the periodic orbit of the unperturbed equation is hyperbolic. Hence, the main results of this section can be viewed as corollaries of Theorems 32 and 33. In fact, we need slightly stronger assumptions in the regularity in this section.

In this section, we will introduce an operator, see (3.63), which is slightly different from the one introduced in Section 3.4.2.

Even if the operator considered in this section requires more regularity in the finite dimensional case, it generalizes our results to perturbations of PDEs, see Section 3.10, to perturbations of Delay Differential Equations, and to other solutions that we will not discuss here (quasi-periodic, normally hyperbolic manifolds). We also note that the corrections needed in this section can be independent of the period. This makes it possible to develop a theory of aperiodic hyperbolic sets. We hope to come back to this problem.

3.9.1 Dynamical definition of hyperbolic periodic orbits

It is a standard notion that a periodic orbit of the ODE $\dot{x} = f(x)$ is hyperbolic when the following strengthening of (H1) holds.

With the same notation as in Section 3.3.1, we say that a periodic orbit is hyperbolic if:

(H1.1) $\Phi(\theta_0 + 1; \theta_0)$ has a simple eigenvalue 1 whose eigenspace is generated by $DK_0(\theta_0)$. Moreover, all the other eigenvalues of $\Phi(\theta_0 + 1; \theta_0)$ have modulus different from 1.

The assumption (H1.1) is equivalent to the following evolutionary formulation (H1.1') in terms of invariant decompositions. In the finite dimensional case, this formulation is easily obtained by taking the stable and unstable spaces of the monodromy matrix and propagating them by the variational equations. In the infinite dimensional cases, similar formulations are obtained using semi-group theory under appropriate spectral assumptions.

(H1.1') For every $\theta \in \mathbb{T}$ there is a decomposition

$$\mathbb{R}^n = E^s_\theta \oplus E^u_\theta \oplus E^c_\theta, \quad E^c_\theta = \text{Span}\{DK_0(\theta)\}, \tag{3.59}$$

depending continuously on θ such that E^s_θ is forward invariant, E^u_θ is backward invariant under the variational equation. Moreover, the forward semiflow (resp. backward semiflow) of the variational equation is contractive on E^s_θ (resp. E^u_θ).

More explicitly, we can find families of linear operators

$$\begin{aligned} &\{U^s_\theta(t)\}_{\theta\in\mathbb{T},t\in\mathbb{R}_+}, \quad U^s_\theta(t)\colon E^s_\theta \to E^s_{\theta+\omega_0 t} \quad t \in \mathbb{R}_+, \\ &\{U^u_\theta(t)\}_{\theta\in\mathbb{T},t\in\mathbb{R}_-}, \quad U^u_\theta(t)\colon E^u_\theta \to E^u_{\theta+\omega_0 t} \quad t \in \mathbb{R}_-, \end{aligned}$$

satisfying for all $\theta \in \mathbb{T}$

$$\begin{aligned} &\partial_t U^\sigma_\theta(t) = Df(K_0(\omega_0 t + \theta))U^\sigma_\theta(t) \quad \sigma \in \{s, u\} \\ &U^\sigma_\theta(0) = \text{Id}\big|_{E^\sigma_\theta}, \end{aligned} \tag{3.60}$$

and

$$U^\sigma_\theta(t + \tau) = U^\sigma_{\omega_0 t+\theta}(\tau). \tag{3.61}$$

Moreover, there exist $C > 0$, $\mu_s > 0$, $\mu_u > 0$ such that

$$\begin{aligned} \|U^s_\theta(t)\| &\leqslant Ce^{-\mu_s t} \qquad t \geqslant 0, \\ \|U^u_\theta(t)\| &\leqslant Ce^{-\mu_u |t|} \quad\; t \leqslant 0. \end{aligned} \tag{3.62}$$

We can also define an evolution operator $U^c_\theta(t)$ in the E^c direction. Note that $U^c_\theta(1) =$

$\mathrm{Id}|_{E^c_\theta}$.

3.9.2 Main result in hyperbolic case

The first result in this case is that Theorem 32 is true if assumption (H1) is changed to assumption (H1.1) or (H1.1'), and assumption (H2.1) is strengthened to (H2.1.1) as follows:

(H2.1.1) If $K \in U_\rho$ and $\omega \in B_\delta$, then $\mathscr{P}(K, \omega, \gamma, \cdot) \colon \mathbb{T} \to \mathbb{R}^n$ is $C^{\ell+\mathrm{Lip}}$, with

$$\|\mathscr{P}(K, \omega, \gamma, \cdot)\|_{C^{\ell+\mathrm{Lip}}} \leqslant \phi_{\rho,\delta},$$

where $\phi_{\rho,\delta}$ is a positive constant that may depend on ρ and δ.

Recall that U_ρ is the ball of radius ρ in the space $C^{\ell+\mathrm{Lip}}(\mathbb{T}, \mathbb{R}^n)$ centered at K_0, and B_δ is the interval in $\mathbb{R}$ with radius δ centered at ω_0.

The second result is that the results in Theorem 33 is true if assumption (H1) is substituted by assumption (H1.1) or (H1.1'), and assumption (H2.2) is strengthened to (H2.2.1) as follows:

(H2.2.1) If $K \in \mathcal{U}_\rho$ and $\omega \in \mathcal{B}_\delta$, then $\mathscr{P}(K, \omega, \cdot, \cdot) \colon \mathbb{T} \times O \to \mathbb{R}^n$ is $C^{\ell+\mathrm{Lip}}$, with

$$\|\mathscr{P}(K, \omega, \cdot, \cdot)\|_{C^{\ell+\mathrm{Lip}}} \leqslant \phi_{\rho,\delta},$$

where $\phi_{\rho,\delta}$ is a positive constant that may depend on ρ and δ.

Recall that $\mathcal{U}_\rho$ is the ball of radius ρ in the space $C^{\ell+\mathrm{Lip}}(\mathbb{T} \times O, \mathbb{R}^n)$ centered at K_0, and $\mathcal{B}_\delta$ is the ball in $C^{\ell+\mathrm{Lip}}(O, \mathbb{R})$ with radius δ centered at constant function ω_0.

Remark 64. *We emphasize that the results in this section are weaker than Theorem 32 and Theorem 33, however, we want to introduce a different operator in the proof which has applications in ill-posed PDEs, see Section 3.10. Modification of the operator will be useful in the study of other dynamical objects.*

3.9.3 Proof

We proceed as in Section 3.4.2 and manipulate (3.8) as a fixed point problem taking advantage of the geometric structures assumed in (H1.1').

Given the decomposition as (3.59),we define projections $\Pi^s_\theta, \Pi^u_\theta, \Pi^c_\theta$ over the spaces $E^s_\theta, E^u_\theta, E^c_\theta$. We also use the notation

$$\widehat{K}^\sigma(\theta) \coloneqq \Pi^\sigma_\theta \widehat{K}(\theta), \qquad \sigma \in \{s, u\}.$$

Taking projections along the spaces of the decomposition, using the variation of parameters formula, and taking the initial conditions to infinity (this procedure is standard since [70]), we see that (3.8) implies

$$\begin{aligned}
\widehat{\omega} &= \omega_0 \frac{\langle \Pi^c_{\theta_0} \int_0^{\frac{1}{\omega_0}} U^c_{\theta_0+\omega_0 t}(\frac{1}{\omega_0} - t) B^\varepsilon(\widehat{K}, \widehat{\omega}, \gamma, \theta_0 + \omega_0 t) dt, DK_0(\theta_0)\rangle}{|DK_0(\theta_0)|^2}, \\
\widehat{K}^s(\theta) &= \int_{-\infty}^0 U^s_{\theta+\omega_0 t}(-t) \Pi^s_{\theta+\omega_0 t} B^\varepsilon(\widehat{K}, \widehat{\omega}, \gamma, \theta_0 + \omega_0 t)\, dt, \\
\widehat{K}^u(\theta) &= -\int_0^\infty U^u_{\theta+\omega_0 t}(-t) \Pi^u_{\theta+\omega_0 t} B^\varepsilon(\widehat{K}, \widehat{\omega}, \gamma, \theta_0 + \omega_0 t)\, dt.
\end{aligned} \tag{3.63}$$

Define the right hand side of (3.63) as an operator of $(\widehat{\omega}, \widehat{K}^s, \widehat{K}^u)$, one can get lemmas which are similar to Lemmas 42 and 45. Hence we can get a fixed point of the operator in this case.

When the solutions of (3.63) are smooth enough and decay fast enough that we can take derivatives inside of the integral sign (which will be the case of the fixed points that we produce), it is possible to show, taking derivatives of both sides of (3.63) and reversing the algebra that the well behaved fixed points of (3.63) indeed are solutions of (3.8).

The remarkable aspect of (3.63) is that we only need U^s_θ for positive times, and U^u_θ for negative times. Hence, the assumed bounds (3.62) imply that the indefinite integrals in (3.63) converge uniformly in the $C^{\ell+\mathrm{Lip}}$ sense. At the same time, we pay the price of requiring one more derivative of $\mathscr{P}$ while using this operator.

Another important feature of the operator (3.63) is that it does not require many assumptions on the long term evolution of the solutions (in Section 3.4.2 we use heavily that the solutions we seek are periodic). This makes it possible to use analogues of (3.63) in several other problems. We hope to come back to these questions in the near future.

3.10 Evolutionary Equations with Delays

In this section we extend the results on ODEs in the previous sections to PDEs and other evolutionary equations (e.g. equations involving fractional operators or integral operators).

The key observation is that, the previous treatments of periodic solutions do not use much that the functions we are seeking take values in a finite dimensional space. For example, the Lemma 92 is valid for functions taking values in Banach spaces. Hence, we will show that the methods developed in the previous sections can be applied without much change to a wide class of PDEs.

Indeed, since one of the points of the previous theory was to avoid the discussion of the evolutions, the theory applies easily to PDEs using only very simple results on the evolution of the PDE.

Remark 65. *In this section, we will not discuss the existence of periodic solutions of evolutionary equations before adding the delays. There is already a large literature in this area.*

We point, however that in studying the periodic solutions of a PDE (which lie in an infinite dimensional space), it is natural to consider the periodic solutions of a finite dimensional truncation (e.g. a Galerkin approximation). The problem of going from the periodic solutions of a finite dimensional problem to the periodic solutions in an infinite dimensional space, has some similarity with the problems dealt with in the first parts of this Chaper.

A framework that systematizes the passing from periodic solutions of the Galerkin approximations to periodic solutions of the PDEs is in [71]. The methods of [71] have some points in common with the methods used in this Chapter. It bypasses the study of evolutionary equations and just studies the functional equations satisfied by a parametrization of a periodic orbit. The methods in [71] lead to computer-assisted proofs that have been implemented in [72, 73]. Since the methods of [71] and this Chapter have points in common, one can hope to combine them and go from a periodic solution of Galerkin truncation of the PDE to a periodic solution of the delay perturbation of the PDE.

3.10.1 Formulation of the problem and preliminary results

We use the standard set up of evolutionary equations (see [74, 75]).

Consider problem of the form

$$\partial_t u(t) = \mathscr{F}(u(t)) + \varepsilon P(u(t), u_t; \gamma), \tag{3.64}$$

where $u(t)$, is the unknown and lies in a space X consisting of functions on a domain Ω. The points in Ω will be given the coordinate x, so that we can also consider $u(t,x)$ as a function on $\mathbb{R} \times \Omega$.

The function space X encodes regularity properties of the functions as well as boundary conditions. In particular, changing the boundary conditions, changes the space X and therefore, the functional analysis properties (e.g. spectra) of the operators acting on it.

The operator $\mathscr{F}$ is a (possibly nonlinear) differential (or fractional differential etc.) operator.

As before (and contrary to the standard use in PDEs where u_t denotes partial derivative), we use u_t to denote a segment of the solution, which can be related with history or future. For $s \in [-h, h]$, $u_t(s) = u(t+s)$, so that $u_t \in \mathscr{R}([-h, h], X)$, a space of regular functions on $[-h, h]$ with values in X. To denote derivatives with respect to time we will always use

$\partial_t u$.

We consider $P\colon X \times \mathscr{R}([-h,h],X) \times \mathbb{R}^m \to X$.

It is useful to think heuristically of

$$\partial_t u(t) = \mathscr{F}(u(t)) \tag{3.65}$$

as a differential equation in X and indeed, our results will be based on this heuristic principle. To make sense of this heuristic principle we have to overcome the problem that in the interesting applications (see e.g. Section 3.10.3), $\mathscr{F}$ is highly discontinuous (involving derivatives) and not defined everywhere so that the standard tools for smooth ODEs do not apply, but this is a well studied problem.

A research program which became specially prominent in the 60's shows that one can recover many of the results (existence, dependence on initial conditions, etc.) for the equation (3.65) by assuming functional analysis properties of the operator $\mathscr{F}$, see [74, 76, 77, 75, 78, 79, 80]. Of course, the verification of the functional analysis assumptions in concrete examples, requires some hard analysis. One of the subtle points of this program is that the notion of solutions may be redefined to be weak or mild solutions.

Even if we will use the language and some material from the above program, we will take a different point of view.

- We will not be interested in the theory of existence and well-posedness for **ALL** the possible initial conditions.

- Indeed, because we are not going to discuss the initial value problems, we can consider situations where the set of initial conditions for the delay problems are not clear. Nevertheless, we can get existence of smooth solutions.

- Since we are only aiming to produce some particular solutions, one gets stronger results by taking more reduced spaces so that the solutions are more regular and can be understood in the classical sense. In particular, in all the cases we will consider, the functions and their derivatives will be bounded. (This happens, e.g. if X is a Sobolev space of high enough order.)

This is in contrast with the general theory of existence and uniqueness, where the figure of merit is considering a more general space of initial conditions.

- A more elaborate set-up for existence of evolutions that includes also FDEs is in [81]. In this Chapter, however, we will avoid discussing the evolution of the FDEs and need only some results on the evolution of the PDE.

3.10.2 Overview of the method

Roughly, we will formulate analogues of the operator Γ^ε in (3.15) and (3.16) as well as the operator in (3.63) and verify that similar contraction argument can be carried out.

The requirements of the above program on the theory of existence are very mild. The operator Γ^ε only requires the existence of solutions of the variational equation for finite time. The operators formulated in (3.63) only require the existence of partial evolutions (forward and backward evolutions in complementary spaces), which allows to consider ill-posed equations, see Section 3.10.5. Moreover, the smoothness requirements on the delay terms are very mild.

3.10.3 Examples

In this section, we will present some examples which are representative of the results we establish and which have appeared in applications.

Even if we hope that this section can serve as motivation, from the purely logical point of view, it can be skipped. Of course, our results apply to many more models and this section is not meant to be an exhaustive list but to provide some intuition.

3.10.3.1 *Delay Perturbations*

One example of delay perturbation which considers long range interaction is

$$P(u(t), u_t; \gamma) = \int_{\mathbb{R}^d} K(x, y) \cdot u(t, x) \cdot u(t - \frac{1}{c}|x - y|, y)dy. \tag{3.66}$$

This models a situation in which the position x interacts with position y with a strength $K(x, y)$, with the interaction taking some time (proportional to the distance) to propagate. In (3.66) we have denoted by c the speed of propagation of the signal, which is assumed to be constant.

Note that the interaction term could be more general than quadratic, and may involve higher spacial derivatives thanks to the smoothing property of solutions. Meanwhile, the speed of propagation of the signal may not be constant (the propagation of signals may depend on their strength).

Another example

$$P(u(t), u_t; \gamma) = \int_0^\infty G(s, u(t - s, x))\, ds, \tag{3.67}$$

treating non-local interaction, is very typical in the modeling of materials with memory effects (for example *thixotropic* materials) where the properties of the materials depend on the history. The effect of the previous state at present time often decrease when the time delay grows. This is reflected on the function $G(s, u)$ decreasing when s (the delay in the effect) increases.

Of course, the mathematical theory that will be developed accommodates more complicated effects such as G depending on spatial derivatives of u.

There are many other $P(u(t), u_t; \gamma)$ that we can consider. We only need P to satisfy some assumptions on regularity and Lipschitz property, see (H2.1*), (H3.1*), and (H2.1.1*), where we actually allow loss of regularity in the space variable.

In the coming sections, we see examples of unperturbed equations (3.65).

3.10.3.2 Parabolic equations

Consider the equation for $u\colon \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$:

$$\begin{aligned} &\partial_t u = \Delta u + N(x, u, \nabla u) \\ &u(t, x) = u(t, x + e) \quad \forall e \in \mathbb{Z}^d \end{aligned} \tag{3.68}$$

with N vanishing to quadratic order. For simplicity, we have imposed periodic boundary conditions in space.

Notice that we have not imposed initial conditions at $t = 0$ in example (3.68). Indeed, the initial conditions needed require some thought.

As we will see, our treatment overcomes other possible complications not mentioned explicitly so far. We mention them because they are natural in modeling and eliminating them from the literature may be motivated by the need to have a more mathematically treatable problem.

Let us just mention briefly some small modifications.

- The unknown u could take values in $\mathbb{R}^d$. Note that considering systems rather than scalar equations makes a big difference in some PDE treatments (based on maximum principle), but it is not an issue in our case.

- The papers [82, 83] consider damped wave equations with a delay. From the functional analysis point of view, the damped wave equations are similar to (3.68).

3.10.3.3 Kuramoto-Sivashinsky equations

The model below is called the Kuramoto-Sivashinsky equation.

$$\begin{aligned} &\partial_t u = \Delta u + \Delta^2 u + \mu \partial_x (u^2) \\ &u(t,x) = u(t,x+e) \quad \forall e \in \mathbb{Z}^d \end{aligned} \tag{3.69}$$

The Kuramoto-Sivashinsky equations appear as *amplitude equations* for many problems arising in a variety of applications (water waves, chemical reactions, interactive populations, etc.).

From the mathematical point of view, when $d = 1$ (reduction of models with more variables), the equation is known to have an inertial manifold (all the solutions converge to a finite dimensional manifold), which can be analyzed by finite dimensional methods. The equation (3.69) is known to have many periodic solutions. A very large number was identified by non-rigorous, but reliable methods in [84]. Rigorous periodic solutions have been established in many papers, including bifurcations in [85, 86]. From the point of view of this Chapter, it is interesting to note that [73, 72] use computer assisted proofs to establish the existence of periodic orbits.

The equations discussed in the previous two sections are parabolic PDEs so that indeed, the evolution is well defined and the solutions gain smoothness. The linearized operator Φ that enters in (3.15) and (3.16) is also smoothing. Of course, for large solutions, there could be finite time blow ups, but we are in the regime of periodic solutions, which are well behaved.

3.10.3.4 The Boussinesq equations in long wave approximation for water waves

In this section we present some physical equations that are ill-posed in the sense that it is impossible to define an evolution for every initial condition. On the other hand, these equations may possess many interesting and physically relevant solutions.

Since one of the main ideas of our treatment of FDEs is to bypass the evolution, we obtain results on delay perturbations of ill-posed equations. This indeed highlights the difference of the present method with the methods in evolution equations.

The material of this section is somewhat more sophisticated than the rest of the Chapter and does not affect any of the other results.

Consider the equation for $u\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}$, derived in [87], as a long wave approximation for water waves.

$$\partial_t^2 u = \mu \partial_x^4 u + \partial_x^2 u + (u^2)_x \quad u(t, x+1) = u(t, x) \tag{3.70}$$

This equation (3.70) can be written as an evolution equation of the form (3.65) as follows:

$$\begin{aligned} &\partial_t u = v \\ &\partial_t v = \mu \partial_x^4 u + \partial_x^2 u + (u^2)_x \\ &u(t, x+1) = u(t, x); \quad v(t, x+1) = v(t, x) \end{aligned} \tag{3.71}$$

The linear part of the evolution is

$$\begin{aligned} &\partial_t u = v \\ &\partial_t v = \mu \partial_x^4 u + \partial_x^2 u \end{aligned} \tag{3.72}$$

Equations similar to (3.70) have also appeared in other contexts. In water wave theory, $\mu > 0$, which leads to (3.70) being ill-posed. Indeed, consider the linear part of the equation, the coefficient of the k-th Fourier mode $\hat{u}_k$ satisfies $\frac{d}{dt^2}\hat{u}_k = (\mu k^4 - k^2)\hat{u}_k$, which leads to exponentially growing solutions either in the future or in the past.

Nevertheless, it is well known that the Boussinesq equation contains many physically interesting solutions, including traveling waves and other periodic and quasi-periodic solutions that are not traveling waves. Notably, it contains a finite dimensional manifold (local center manifold) which is locally invariant and on which solutions can be defined till they leave the local center manifold [88, 89, 90]. In particular, the periodic and quasi-periodic solutions in the local center manifold are defined for all times.

For our purposes, the Boussinesq equation (3.70) is Hamiltonian, so that all the periodic solutions have monodromy with eigenvalues 1 – corresponding to the conservation of the energy – which make them unsuitable for the present version of our theory. Hence, we will consider, for $u\colon \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, mainly time periodic perturbations of higher dimensional version of (3.70), which following the notation in [90], we write as:

$$\begin{cases} \partial_t \theta = \omega \\ \partial_t^2 u = \mu \Delta^2 u + \Delta u + N_1(\theta, x) + N_2(\theta, x) u + N_3(\theta, x, u, \nabla u, \Delta u), \\ t \in \mathbb{R}, \;\; \theta \in \mathbb{T}, \quad x \in \mathbb{T}^d, \end{cases} \tag{3.73}$$

The model (3.73) can be a long wave approximation of a water wave model perturbed periodically. These are physically sensible long wave approximations of a water wave subject to periodic forcing (e.g. waves in the ocean subject to tides or water waves in a vibrating table – Faraday experiment).

The result of [90] implies, under very mild regularity assumptions on N_1, N_2, N_3, that there is a finite dimensional local center manifold of (3.73) which is locally invariant.

This local center manifold is modeled on $\mathbb{T} \times \mathbb{R}^n$. The periodic solutions in the manifold are defined for all time. For specific forms of N, it is possible to prove the existence of periodic orbits of (3.73), which are non-degenerate in the center manifold.

A natural space to consider (3.71) is $(u, v) \in X \coloneqq H^r \times H^{r-1}$ for sufficiently large r. Even if it is impossible to define an evolution of the linear part (3.72) in the full space X, it is easy to show using Fourier analysis that there are two complementary spaces in which one can define the evolution forwards and backwards. A remarkable result in [89, 90] is that this splitting with partial evolution operators persists in the linearization near periodic orbits, provided that they stay close to the origin.

3.10.4 Result for well-posed PDE

The Theorem 66 will be our main result for well-posed PDEs. Essentially, the assumptions of the theorem are that we can formulate the functional equation in (3.15) and (3.16) and that the delay term prossesses enough regularity so that the argument we used to prove Theorem 32 goes through unchanged.

Therefore, the proof of Theorem 66 is a trivial walk-through. On the other hand, the fact that the assumptions are satisfied in the cases (3.68), (3.69) for some choices of spaces X is not trivial and will be discussed in Section 3.10.4.5. Of course, similar verifications can be done in other models.

The only subtlety is that we will use the *two spaces approach* of [78]. (See also [91, 92] for a more streamlined and refined version.) This allows to consider perturbations which are unbounded but of lower order than the evolution operator. For example in (3.68), the nonlinearity involves the first derivatives taking advantage of the fact that the main evolution operator is of second order. In the case of (3.69), since the linear term is a fourth order elliptic operator, the nonlinearity could involve terms of order up to three. As we will see, the two space approach also allows to lower the regularity requirements of the delay term. (See hypotheses in Theorem 66.)

3.10.4.1 The two spaces approach

The basic idea of the two spaces approach is that we study the evolution equation using two spaces X, Y consisting of functions with different regularity. In applications to PDEs, often $X = H^{r+k}, Y = H^r$ with H^r the standard Sobolev spaces or the product of these spaces. In our case, we will take r large enough so that the solutions are classical, and the space H^r enjoys properties that it is a Banach algebra and the composition operator is smooth.

Differential operators, which are unbounded from a space to itself become bounded from X to Y. Then, the main evolution operator, smooths things out, such that it maps Y

to X in a bounded way. Of course, the bound of the evolution as an operator from the rough space Y to the smooth space X depends on the time that the evolution has been acting and becomes singular as the time goes to zero, but we assume that there are bounds for the negative powers, which ensures integrability.

3.10.4.2 Setup of the result

Consider the evolutionary PDE (3.65). Let X, Y be Banach spaces consisting of smooth enough functions satisfying the boundary conditions imposed on (3.65). We will assume that Y consists of less smooth functions, such that $\mathscr{F}$ is a differentiable map from space X to space Y. One consequence is that X has a compact embedding into Y.

Let $K_0\colon \mathbb{T} \to X$ be a parameterization of the periodic orbit of (3.65). As in Section 3.4.1, we use the notation $K(\theta) = K_0(\theta) + \widehat{K}(\theta)$ with $\widehat{K}\colon \mathbb{T} \to X$, and we derive formally the equation (3.74).

$$\omega_0 D\widehat{K}(\theta) - D\mathscr{F}(K_0(\theta))\widehat{K}(\theta) = B^\varepsilon(K, \omega, \gamma, \theta) - \widehat{\omega} DK_0(\theta), \tag{3.74}$$

where

$$\begin{aligned} B^\varepsilon(K, \omega, \gamma, \theta) &\coloneqq N(\theta, \widehat{K}) + \varepsilon \mathscr{P}(K, \omega, \gamma, \theta) - \widehat{\omega} D\widehat{K}(\theta), \\ N(\theta, \widehat{K}) &\coloneqq \mathscr{F}(K_0(\theta) + \widehat{K}(\theta)) - \mathscr{F}(K_0(\theta)) - D\mathscr{F}(K_0(\theta))\widehat{K}(\theta). \end{aligned} \tag{3.75}$$

3.10.4.3 Statement of the result

We first formulate an abstract result, Theorem 66, whose proof is almost identical to the proof of Theorem 32. The deep result is to verify that the hypotheses of Theorem 66 hold in examples of interest. In Section 3.10.4.5, we show that the examples in Section 3.10.3 verify the hypotheses. We leave the verification in other models of interest to the readers.

Theorem 66. *Assume that when $\varepsilon = 0$, the equation* (3.64) *has a periodic orbit which* *satisfies:*

- *The linearized equation around the periodic orbit admits a solution. That is, for any $\theta_0 \in \mathbb{T}$ and $\theta_0 < \theta \in \mathbb{T}$, there is an operator $\Phi(\theta;\theta_0)$ mapping from Y to X solving*

$$\omega_0 \frac{d}{d\theta}\Phi(\theta;\theta_0) = D\mathscr{F}(K_0(\theta))\Phi(\theta;\theta_0); \tag{3.76}$$

-
  - $1 \in Spec(\Phi(1;0), X)$ *is a simple eigenvalue.*
  - *The spectral projection on $Spec(\Phi(1;0), X)\backslash\{1\}$ in X is bounded.*

- *The family of operators Φ is smoothing in the sense that it satisfies*

$$\|\Phi(t;\theta_0)\|_{Y,X} \leqslant C(t-\theta_0)^{-\alpha} \quad 0 < \alpha < 1, \tag{3.77}$$

  where $\|\cdot\|_{Y,X}$ is the norm of an operator mapping from Y to X, C is a constant.

We also need the following two assumptions on the delay perturbation. Let $\ell > 0$ be an integer. Denote the ball of radius ρ in the space $C^{\ell+\mathrm{Lip}}(\mathbb{T}, X)$ centered at K_0 as $\mathcal{U}_\rho$, and the interval in $\mathbb{R}$ centered at ω_0 with radius δ as B_δ.

(H2.1)* *If $K \in \mathcal{U}_\rho$ and $\omega \in B_\delta$, then $\mathscr{P}(K,\omega,\gamma,\cdot)\colon \mathbb{T} \to Y$ is $C^{\ell-1+\mathrm{Lip}}$, with*

$$\|\mathscr{P}(K,\omega,\gamma,\cdot)\|_{C^{\ell-1+\mathrm{Lip}}(\mathbb{T},Y)} \leqslant \phi_{\rho,\delta},$$

where $\phi_{\rho,\delta}$ is a positive constant that may depend on ρ and δ.

(H3.1)* *For $K,\ K' \in \mathcal{U}_\rho$, and $\omega,\ \omega' \in B_\delta$, there exists constant $\alpha_{\rho,\delta} > 0$, such that for all $\theta \in \mathbb{T}$,*

$$\|\mathscr{P}(K,\omega,\gamma,\theta) - \mathscr{P}(K',\omega',\gamma,\theta)\|_Y \leqslant \alpha_{\rho,\delta} \max\left\{|\omega-\omega'|, \|K-K'\|_{C^0(\mathbb{T},X)}\right\}.$$

Then, for small enough ε, the equation (3.64) *has a periodic orbit, which is parameterized by a $C^{\ell+\mathrm{Lip}}$ map $K\colon \mathbb{T} \to X$. Moreover, K is close to K_0 in the sense of $C^\ell(\mathbb{T}, X)$.*

The proof of Theorem 66 is very easy. It suffices to observe that, thanks to the hypotheses of the theorem, the operator Γ^ε, defined in the same way as before, sends a ball in the space $\mathbb{R} \times C^{\ell+\mathrm{Lip}}(\mathbb{T}, X)$ to itself and that in this ball, Γ^ε is a contraction under the norm of $\mathbb{R} \times C^0(\mathbb{T}, X)$. Then, we apply Lemma 92.

Similar to before, one can get smooth dependence on parameters result.

3.10.4.4 Some remarks

Remark 67. *The assumption that equation* (3.76) *admits solutions with the bounds in* (3.77) *is rather nontrivial and its verification in concrete examples requires PDE techniques.*

Remark 68. *Thanks to* (3.77)*, $\Phi(1; 0)$ is bounded from Y to X and, hence compact from Y to Y. Therefore, the spectrum away from zero is characterized by the existence of finite dimensional eigenspaces.*

However, for an operator A acting on two spaces $X \subset Y$, there is no relation of $Spec(A, X)$ and $Spec(A, Y)$ in general.

Remark 69. *In our case, for the operator $\Phi(1; 0)$, its point spectrum in space X agrees with its point spectrum in space Y. This is not hard to see from the eigenvector equation and the smoothing effect of the operator $\Phi(1; 0)$.*

3.10.4.5 Verification of the assumptions of Theorem 66 in some examples

For the parabolic equations (3.68) and (3.69), a very elegant formalism is developed in [78]. The case (3.69) will be simpler than (3.68) since the linearized operator being higher order leads to stronger smoothing properties of the evolution.

The space Y will be H^r, a Sobolev space of high enough order. We emphasize once again that for our purposes, the results are stronger if the space is more restrictive.

The semigroup theory tells us that we can solve the equation (3.76) and that the solution is smoothing in the sense that

$$\|\Phi(\theta;\theta_0)\|_{H^r,\ H^{r+a}} \leqslant C(a)|\theta-\theta_0|^{-a}. \tag{3.78}$$

3.10.5 Result for ill-posed PDE

In this section, we show how one can get existence of periodic solutions for delay perturbations of ill-posed PDEs.

We just need to assume that the linearized equation admits partial evolutions (one evolution forward in time and another one backward in time) defined in complementary spaces. If these evolutions are smoothing, the methods of Section 3.9 apply without change.

Again the deeper part is to show that the concrete examples satisfy the assumptions. In the case of the periodically forced Boussinesq equation (3.73) with a periodic solution which is hyperbolic, we will show that the periodic solution persists under delay perturbation. The assumption that (3.73) has a hyperbolic periodic orbit is a non-trivial – but easily verifiable in concrete models – assumption. We note that the time independent Boussinesq equation (3.70) does not have hyperbolic periodic orbits due to energy conservation. Our results require delicate regularity properties of the periodic orbits, which are verified for all the bounded small solutions in [90].

Since the partial evolutions involve smoothing properties, we still use the two spaces approach summarized in Section 3.10.4.1. We have used the same set up as [90] to help the reader check for the applications.

Remark 70. *When the non-linear terms N in* (3.73) *are analytic, the periodic orbits are analytic. As mentioned in Remark 50, we do not expect that the periodic orbits of the perturbed equations are analytic. So, we follow [90] and deduce the regularity of the periodic orbits from the C^r regularity of the center manifold.*

3.10.5.1 Abstract setup for the study of ill-posed equations

We will assume that there is a periodic solution of the evolution equation (3.73), which satisfies the following Definition 71. Definition 71 can be verified for the linear part of (3.73), and is shown to be stable under perturbations (which can be unbounded) in [89, 90]. (Related notions of splittings and their stability using a different functional analysis set up appear also in [93, 94]. We have found that the two spaces approach is more concrete and easier to adapt to the delay case.)

Definition 71 is motivated by an analogue of hyperbolicity for ill-posed equations. We do not assume that the linearized equations define an evolution such as Φ, but we assume that there are two evolutions (one in the future and one in the past) defined in complementary spaces. This is enough to follow the set up introduced in Section 3.9 and formulate a fixed point equation for the periodic orbit of the perturbed equation.

Let us make some remarks about some subtle technical points.

• We assume that when these evolutions are defined, they are *smoothing*. That is, they take functions of a certain degree of differentiability (in x) and map them into functions with more derivatives. As shown in [89, 90], this allows to show that these structures are stable under perturbations, which can be unbounded but are of lower order. This generality is important in the treatment of examples such as (3.70) since it allows to show that the periodic solutions constructed in the above papers satisfy Definition 71.

• It is important to note that Definition 71 only needs to be applied to the periodic orbits of the problem without the delay. In this section the unperturbed problem will be a PDE, which is exactly the case discussed in [89, 90]. As in Section 3.9, the invariant splitting will be used to set up a functional equation and it will remain fixed, so that once we verify the existence in the unperturbed case, it does not get updated.

• Both [89, 90] consider situations more general than periodic orbits. The paper [89] considers quasi-periodic orbits and [90] considers bounded orbits. In the case of quasi-periodic (in particular periodic) orbits, it is natural in the examples considered to assume that the bundles are analytic. For orbits with a time-dependence more complicated than periodic, it is natural to assume only finite regularity. In this Section we have adopted the definition in [89], which includes analyticity, since it applies to the examples we have in mind. Notice, however that the solutions of the delay equation will only be shown to be finitely differentiable and depend regularly on parameters in finite differentiable topologies. Indeed, we do not expect that the solutions of the delay problem will be analytic. See Remark 50.

Definition 71. *Let $X \subset Y$ be two Banach spaces. We say that an embedding $K_0 \colon \mathbb{T}_\rho \to X$ is spectrally nondegenerate if for every θ in $\mathbb{T}$, we can find splittings:*

$$\begin{aligned} X &= X_\theta^s \oplus X_\theta^c \oplus X_\theta^u \\ Y &= Y_\theta^s \oplus Y_\theta^c \oplus Y_\theta^u \end{aligned} \tag{3.79}$$

with associated bounded projections on X and Y. (We will abuse the notation and use $\Pi_\theta^{s,c,u}$ to denote the projections as maps in $L(X,X)$ or in $L(Y,Y)$.) The projections depend analytically on $\theta \in \mathbb{T}_\rho \coloneqq \{z \in \mathbb{C}/\mathbb{Z} : |Im\ z| < \rho\}$, and have continuous extensions to the closure of $\mathbb{T}_\rho$. Spaces $X_\theta^{s,c,u}$ and $Y_\theta^{s,c,u}$ have the following properties.

- *We can find families of operators*

$$\begin{aligned} &U_\theta^s(t) \colon Y_\theta^s \to X_{\theta+\omega_0 t}^s, \quad t > 0, \\ &U_\theta^u(t) \colon Y_\theta^u \to X_{\theta+\omega_0 t}^u, \quad t < 0, \\ &U_\theta^c(t) \colon Y_\theta^c \to X_{\theta+\omega_0 t}^c, \quad t \in \mathbb{R}. \end{aligned}$$

- *The operators $U_\theta^{s,c,u}(t)$ are cocycles over the rotation satisfying*

$$U_{\theta+\omega_0 t}^{s,c,u}(\tau) U_\theta^{s,c,u}(t) = U_\theta^{s,c,u}(\tau + t). \tag{3.80}$$

- *The operators $U_\theta^{s,c,u}(t)$ are smoothing in the time direction where they can be defined* *and they satisfy assumptions in the quantitative rates. There exist constants* $\alpha_1, \alpha_2 \in [0,1)$, $\beta_1, \beta_2, \beta_3^+, \beta_3^- > 0$ *with* $\beta_1 > \beta_3^-$, *and* $\beta_2 > \beta_3^+$, *and* $C > 1$, *independent of* θ, *such that the evolution operators satisfy the following rate conditions:*

$$\|U_\theta^s(t)\|_{\rho,Y,X} \leqslant Ce^{-\beta_1 t} t^{-\alpha_1}, \quad t > 0, \tag{3.81}$$

$$\|U_\theta^u(t)\|_{\rho,Y,X} \leqslant Ce^{-\beta_2 |t|} |t|^{-\alpha_2}, \quad t < 0, \tag{3.82}$$

and

$$\begin{aligned} \|U_\theta^c(t)\|_{\rho,Y,X} &\leqslant Ce^{\beta_3^+ t}, \quad t \geqslant 0, \\ \|U_\theta^c(t)\|_{\rho,Y,X} &\leqslant Ce^{\beta_3^- |t|}, \quad t \leqslant 0. \end{aligned} \tag{3.83}$$

- *The operators* $U_\theta^{s,c,u}(t)$ *are solutions of the variational equations in the sense that*

$$\begin{aligned} U_\theta^s(t) &= Id + \int_0^t D\mathscr{F}^s(K_0(\theta + \omega_0 \tau)) U_\theta^s(\tau) d\tau, \quad t > 0, \\ U_\theta^u(t) &= Id + \int_0^t D\mathscr{F}^u(K_0(\theta + \omega_0 \tau)) U_\theta^u(\tau) d\tau, \quad t < 0, \\ U_\theta^c(t) &= Id + \int_0^t D\mathscr{F}^c(K_0(\theta + \omega_0 \tau)) U_\theta^c(\tau) d\tau, \quad t \in \mathbb{R}. \end{aligned} \tag{3.84}$$

In this Section, we will also need:

- *The space* X^c *is unidimensional and it is spanned by the direction of the evolution along the periodic orbit.*

Recall that $\mathcal{U}_\rho \subset C^{\ell+\mathrm{Lip}}(\mathbb{T}, X)$ is the ball of radius ρ centered at K_0, and $B_\delta \subset \mathbb{R}$ is the interval centered at ω_0 with radius δ. Compared with the hypothesis for well-posed equations in (H2.1*), we make similar but slightly stronger assumption on the delay term:

(H2.1.1*) If $K \in \mathcal{U}_\rho$ and $\omega \in B_\delta$, then $\mathscr{P}(K, \omega, \gamma, \cdot)\colon \mathbb{T} \to Y$ is $C^{\ell+\mathrm{Lip}}$, with

$$\|\mathscr{P}(K, \omega, \gamma, \cdot)\|_{C^{\ell+\mathrm{Lip}}(\mathbb{T},Y)} \leqslant \phi_{\rho,\delta}$$

where $\phi_{\rho,\delta}$ is a positive constant that may depend on ρ and δ.

3.10.5.2 Statement of the result

Theorem 72. *Assume that we have an evolution equation* (3.65) *that admits a periodic solution satisfying Definition 71, and that we perturb by delay terms satisfying assumptions (H2.1.1*) and (H3.1*).*

Then, for sufficiently small ε, the equation (3.64) *has a periodic solution of frequency ω, which is parameterized by a $C^{\ell+\mathrm{Lip}}$ map $K\colon \mathbb{T} \to X$. Moreover, K is close to K_0 in the sense of $C^{\ell}(\mathbb{T}, X)$.*

The proof of Theorem 72 follows the same line as in Section 3.9.3. We work with the fixed point equation (3.63). Using that we have evolution U_θ^c and partial evolutions U_θ^s and U_θ^u for the linearized equations satisfying Definition 71, we can find solution to equation (3.63), with $\widehat{\omega} \in \mathbb{R}$ and $\widehat{K}^s$, $\widehat{K}^u \in C^{\ell+\mathrm{Lip}}(\mathbb{T}, X)$.

As we have discussed, the regularity properties are verified for concrete examples in [90] for the time-perturbed Boussinesq equation (3.73).

CHAPTER 4

NUMERICAL COMPUTATION OF LIMIT CYCLES AND ISOCHRONS FOR FDES NEAR A PLANAR ODE

We present algorithms and their implementation for the results discussed in Chapter 2. Namely, we compute limit cycles and their isochrons (slow stable manifolds) for SDDEs perturbed from a planar ODE.

The numerical methods developed here, produce an approximate solution and provide estimates of the condition numbers. Therefore, thanks to the *a posteriori* results in Chapter 2, we are confident that the solutions produced by our numerical methods correspond to true solutions.

The algorithms consist in specifying discretizations for all the functional analysis steps in Chapter 2. We choose a systematic way to approximate functions by a finite set of numbers (Taylor-Fourier series) and develop a toolkit of algorithms that implement the operators – notably composition – that enter into the theory.

We do not present rigorous estimates on the effects of discretizations (they are in principle applications of standard estimates), but we present analysis of running times. We implement the algorithms above and report the results of running them in some representative examples. In our examples, one can indeed obtain very accurate solutions in a few minutes in a standard today's laptop.

The results in this chapter are presented in an increasing level of details, from the general steps of the algorithms to more specialized and the hardest steps of them. The algorithms that allow to solve the invariance equation (2.8) are fully detailed in section 4.2. Section 4.3 explains the numerical composition of periodic mappings as well as its computational complexity in Fourier representation. In section 4.4 we report the results in some examples.

Our results take the parameterization of the unperturbed limit cycle and its isochrons as input. They can be obtained from standard ODE techniques. For completeness, we summarize in Appendix B the steps and add practical comments of the parameterization method described in [26].

Our numerical representation for periodic orbit is going to be one-dimensional Fourier expansion. See Appendix C for a summary of possibly well-known results of Fourier representation and how they are managed and packed from a programming point of view.

4.1 Non-Uniqueness of the Solution

Recall that our goal is to find constants ω, λ, and map W of the form (2.11), which solve equation (2.8). We will only compute W up to a finite order, and will not consider $W^{>}(\theta, s)$. That is, we only compute the limit cycle and a finite Taylor expansion of the isochrons. The error of the reminder of the Taylor expansion is indeed very small (much smaller than other sources of numerical error, which are already small).

The equation (2.8) is underdetermined, i.e., if W, ω, and λ solve equation (2.8), then $W_{\sigma,\eta}, \omega$, and λ also solve the same equation with

$$W_{\sigma,\eta}(\theta, s) = W(\theta + \sigma, \eta s). \tag{4.1}$$

The parameters σ and η correspond to choosing a different origin in the angle coordinate θ and a different scale of the parameter s, respectively. All these solutions in (4.1) are mathematically equivalent, we introduce two normalizations to fix one solution in this family.

A convenient way to fix the origin of θ is to require

$$\int_0^1 \left[\partial_\theta W_1^0(\theta, 0)\, W_1(\theta, 0)\right] d\theta = a \tag{4.2}$$

where W^0 is an initial approximation and a is a real number, typically it is close to 1. This normalization is easy to compute and is rather sensitive since, when we move in the family

(4.1), the derivative with respect to the shift is a positive number.

The normalization of the origin of coordinates, has no numerical consequences except for the possibility of comparing the solutions in different runs. The solutions corresponding to different normalizations have very similar properties. The numerical algorithm 74 in its step *5* has a small drift in the normalization in each iteration, but it is guaranteed to converge to one of the solutions in (4.1).

The second normalization is just a choice of the eigenvector of an operator. We find it convenient to take

$$\int_0^1 \partial_s W_2(\theta, 0)\, d\theta = \rho \tag{4.3}$$

with a real $\rho \neq 0$.

We anticipate that changing the value of ρ is equivalent to changing s into bs where b is commonly named *scaling factor*.

The choice of this normalization affects the numerical accuracy dramatically. Notice that if we change s into bs, the coefficients $W^j(\theta)$ in (2.11) change into $W^j(\theta)b^j$. So, different choices of b may cause the Taylor coefficients to be very large or very small, which makes the computations with them very susceptible to roundoff error. It is numerically advantageous to choose the scale in such a way that the Taylor coefficients have a comparable size. In our problem, we are also going to use the scaling to ensure that the second component of W lie in the domain of K so that $K \circ W$ is well-defined.

In practice, we run the calculations twice. First we do a preliminary calculation whose only purpose is to compute an approximation of the scale that makes the coefficients remain more or less of the same size. Then, a more definitive calculation can be run. The latter running is more numerically reliable.

Remark 73. *In standard implementation of the Newton method for the fixed point of a functional, say Ψ, the fact that the space of solutions is two dimensional results in $D\Psi - Id$ having a two-dimensional kernel, and not be invertible.*

In our case, we will develop a very explicit and fast algorithm that produces an approximate linear right inverse. This linear right inverse leads to convergence to an element of the family (4.1).

4.2 Computation of (W, ω, λ) – Perturbed Case

4.2.1 Fixed point approach

We compute all the coefficients $W^j(\theta)$ of the truncated expression $W(\theta, s)$ in (2.11) order by order. The zero and first orders require a special attention due to the fact that the values ω and λ are obtained in the equation (2.8) matching coefficients of s^0 and s^1 respectively. The condition that allows to obtain ω comes from the periodicity condition (2.12). The mapping W^0 is not a periodic function. But we can use it to get a periodic one defined by $\hat{W}^0(\theta) \coloneqq W^0(\theta) - \left(\begin{smallmatrix}\theta\\0\end{smallmatrix}\right)$. The condition for λ is given by the normalization condition (4.3). We can use a scaling factor, which allows to set the value of ρ in (4.3) to 1.

Algorithm 74 sketches the fixed-point procedure to get ω and W^0 whose periodicity condition is ensured in step (*5*). In this case the initial condition will be ω_0 (the value for $\varepsilon = 0$) for ω and $\left(\begin{smallmatrix}\theta\\0\end{smallmatrix}\right)$ for $W^0(\theta)$ since $W(\theta, s)$ is close to the identity.

Algorithm 74 (s^0 case)**.** *Let* $\widetilde{W^0}(\theta) \coloneqq W^0\big(\theta - \omega r \circ K(W^0(\theta))\big)$.

- ⋆ `Input`***:*** $\dot{x} = X(x) + \varepsilon P(x, \tilde{x}, \varepsilon)$*,* $0 < \varepsilon \ll 1$*,* $K(\theta, s) = \sum_{j=0}^{m-1} K^j(\theta)(b_0 s)^j$*,* $b_0 > 0$*,* $\omega_0 > 0$*,* $\lambda_0 < 0$ *and a tolerance tol.*

- ⋆ `Output`***:*** $\hat{W}^0 \colon \mathbb{T} \to \mathbb{R}^2$ *and* $\omega > 0$*.*

1. $\hat{W}^0(\theta) \leftarrow 0$ *and* $\omega \leftarrow \omega_0$*.*

2. $W^0(\theta) \leftarrow \begin{pmatrix}\theta\\0\end{pmatrix} + \hat{W}^0(\theta)$*.*

3. *Solve* $DK \circ W^0(\theta)\eta(\theta) = \varepsilon P(K \circ W^0(\theta), K \circ \widetilde{W^0}(\theta), \varepsilon)$*. Let* $\eta \equiv (\eta_1, \eta_2)$*.*

4. $\alpha \leftarrow \int_0^1 \eta_1(\theta)\, d\theta$ *and* $\omega \leftarrow \omega_0 + \alpha$.

5. *Solve* $\omega\partial_\theta \hat{W}_1^0(\theta) = \eta_1(\theta) - \alpha$ *imposing* $\int_0^1 \hat{W}_1^0(\theta)\, d\theta = 0$.

6. *Solve* $(\omega\partial_\theta - \lambda_0)\hat{W}_2^0(\theta) = \eta_2(\theta)$.

7. *Iterate from* (2) *to* (6) *until convergence in* W^0 *and* ω *with tolerance* tol.

Algorithm 75 sketches the steps to compute (W^1, λ) and W^n for $n \geqslant 2$. The initial guesses are λ_0 for λ, $\binom{0}{1}$ for W^1 and $\binom{0}{0}$ for W^n.

Algorithm 75 (s^1 case and s^n case with $n \geqslant 2$)**.**
Let $\widetilde{W}(\theta, s) \coloneqq W\big(\theta - \omega r \circ K(W(\theta, s)), s e^{-\lambda r \circ K(W(\theta,s))}\big)$.

- $\star$ `Input`*:* $\dot{x} = X(x) + \varepsilon P(x, \widetilde{x}, \varepsilon)$, $0 < \varepsilon \ll 1$, $K(\theta, s) = \sum_{j=0}^{m-1} K^j(\theta)(b_0 s)^j$, $b_0 > 0$, $\omega_0 > 0$, $\lambda_0 < 0$, $\hat{W}^0(\theta)$, $W^j(\theta)$ *for* $0 < j < n$, $b > 0$, $\omega > 0$ *and a tolerance* tol.

- $\star$ `Output`*: either* $W^1 \colon \mathbb{T} \to \mathbb{T} \times \mathbb{R}$ *and* $\lambda < 0$ *or* $W^n \colon \mathbb{T} \to \mathbb{T} \times \mathbb{R}$.

1. $W^n(\theta) \leftarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

s^1 *If* $n = 1$, $W^1(\theta) \leftarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ *and* $\lambda \leftarrow \lambda_0$.

2. $W(\theta, s) \leftarrow \begin{pmatrix} \theta \\ 0 \end{pmatrix} + \hat{W}^0(\theta) + \sum\limits_{j=1}^{n} W^j(\theta)(bs)^j$.

3. $Y(W(\theta, s)) \leftarrow DK \circ W(\theta, s)^{-1} P(K \circ W(\theta, s), K \circ \widetilde{W}(\theta, s), \varepsilon)$.

4. $\eta(\theta) \leftarrow \varepsilon \frac{\partial^n Y}{\partial s^n}(W(\theta, s))_{|s=0}$. *Let* $\eta \equiv (\eta_1, \eta_2)$.

s^1 *If* $n = 1$, *then* $\lambda \leftarrow \lambda_0 + \int_0^1 \eta_2(\theta)\, d\theta$.

5. *Solve* $(\omega\partial_\theta + n\lambda)W_1^n(\theta) = \eta_1(\theta)$.

6. *Solve* $(\omega\partial_\theta + n\lambda - \lambda_0)W_2^n(\theta) = \eta_2(\theta)$.

7. *Iterate from* (2) *to* (6) *until convergence with tolerance* tol. *Then undo the scaling* b.

Both algorithms 74 and 75 have non-trivial parts, such as, the effective computation of $\widetilde{W}$, the numerical composition of K with W and also with $\widetilde{W}$ (see 4.3), the effective computation of the step *4* in Algorithm 75 and the choice of the scaling factor (see 4.2.3). On the other hand, there are steps that we can use the same methods in the unperturbed case (see B), such as, the solution of linear systems like step *3* in Algorithm 75 via Lemma 77 or the solutions of the cohomological equations by Proposition 76.

In the next sections we address each of these parts.

4.2.2 Stopping criterion

Algorithms 74 and 75 require to respectively stop when the prescribed tolerance has been reached.

Alternatively, one can stop when the invariance equation is satisfied up to the given tolerance.

4.2.3 Scaling factor for orders $n \geqslant 1$

As we discussed, if $W(\theta, s)$ is a solution, then $W(\theta + \theta_0, bs)$ will also be a solution for any θ_0 and b. A difference with the unperturbed case is that now $K \circ W$ and $K \circ \widetilde{W}$ are required to be well-defined. That is, the second components of W and $\widetilde{W}$ must lie in $[-1, 1]$. Stronger conditions are

$$p(s) \coloneqq \sum_{j\geqslant 0} \|W_2^j(\theta)\| |s|^j \leqslant 1 \qquad \text{and} \qquad \widetilde{p}(s) \coloneqq \sum_{j\geqslant 0} \|\widetilde{W_2^j}(\theta)\| |s|^j \leqslant 1.$$

In the iterative scheme of Algorithm 75, these series become finite sums and an upper bound for the value $b > 0$ is $\min\{s^*, \widetilde{s}^*\}$, where $s^* > 0$ is the value so that $p(s^*) = 1$ and, similarly, $\widetilde{s}^* > 0$ is the value verifying $\widetilde{p}(\widetilde{s}^*) = 1$. Notice that, the solutions s^* and $\widetilde{s}^*$ exist because $\|W_2^0(\theta)\| < 1$ and $\|\widetilde{W_2^0}(\theta)\| < 1$.

4.2.4 Solutions of the cohomology equations in Fourier representation

Under the Fourier representation (see Section C) we can solve the cohomological equations in the steps *5* and *6* in Algorithm 74 as well as in steps *5* and *6* in Algorithm 75.

Proposition 76 (Fourier version, [26]). *Let* $E(\theta, s) = \sum_{j,k} E_{jk} e^{2\pi \boldsymbol{i} k\theta} s^j$.

- *If* $E_{00} = 0$, *then* $(\omega\partial_\theta + \lambda s \partial_s) u(\theta, s) = E(\theta, s)$ *has solution* $u(\theta, s) = \sum_{j,k} u_{jk} e^{2\pi \boldsymbol{i} k\theta} s^j$ *and*
$$u_{jk} = \begin{cases} \frac{E_{jk}}{\lambda j + 2\pi \boldsymbol{i} \omega k} & \text{if } (j,k) \neq (0,0) \\ \alpha & \text{otherwise.} \end{cases}$$
for all real α. *Imposing* $\int_0^1 u(\theta, 0)\, d\theta = 0$, *then* $\alpha = 0$.

- *If* $E_{10} = 0$, *then* $(\omega\partial_\theta + \lambda s \partial_s - \lambda) u(\theta, s) = E(\theta, s)$ *has solution* $u(\theta, s) = \sum_{j,k} u_{jk} e^{2\pi \boldsymbol{i} k\theta} s^j$ *and*
$$u_{jk} = \begin{cases} \frac{E_{jk}}{\lambda (j-1) + 2\pi \boldsymbol{i} \omega k} & \text{if } (j,k) \neq (1,0) \\ \alpha & \text{otherwise.} \end{cases}$$
for all real α. *Imposing* $\int_0^1 \partial_s u(\theta, 0)\, d\theta = 0$, *then* $\alpha = 0$.

The paper [26] also presents a solution in terms of integrals. Those integral formulas for the solution are independent of the discretization and work for discretizations such as Fourier series, splines and collocations methods. Indeed, the integral formulas are very efficient for discretizations in splines or in collocation methods which could be preferable in some regimes where the limit cycles are bursting. In this chapter, we will not use the integral formulas, since we will discretize functions in Fourier series, and for this discretization, the methods described in Proposition 76 are more efficient.

4.2.5 Treatment of the step *3* in Algorithm 75

To solve the linear system in the step *3* of Algorithm 75, we can use Lemma 77, whose proof is a direct power matching.

Lemma 77. *Let $A(\theta,s)x(\theta,s) = b(\theta,s)$ be a linear system of equations for each given (θ,s). Explicitly:*

$$\left(\sum_{k\geqslant 0} A_k(\theta)s^k\right)\sum_{k\geqslant 0}\boldsymbol{x}_k(\theta)s^k = \sum_{k\geqslant 0}\boldsymbol{b}_k(\theta)s^k.$$

Then, the coefficients $\boldsymbol{x}_k(\theta)$ are obtained recursively by solving

$$A_0(\theta)\boldsymbol{x}_k(\theta) = \boldsymbol{b}_k(\theta) - \sum_{j=1}^{k} A_j(\theta)\boldsymbol{x}_{k-j}(\theta).$$

which can be done provided that $A_0(\theta)$ is invertible and that one knows how to multiply and add periodic functions of θ.

4.2.6 Use of polynomials for elementary operations

We also recall that composition of a polynomial in the left with exponential, trigonometric functions, powers, logarithms (or any function that satisfies an easy differential equation) can be done very efficiently using algorithms that are reviewed in [17] which goes back to [95].

We present here the case of the exponential which can be used in Algorithm 75 for the computation of $\widetilde{W}$.

If P is a given polynomial – or a power series – with coefficients P_j, we see that $E(s) = \exp P(s)$ satisfies

$$\frac{d}{ds}E(s) = E(s)\frac{d}{ds}P(s),$$

with Taylor coefficients E_j at $s = 0$. Equating like powers on both sides, it leads to

$E_0 = \exp P(0)$, and the recursion:

$$E_j = \frac{1}{j} \sum_{k=0}^{j-1} (j-k) P_{j-k} E_k, \qquad j \geqslant 1,$$

Note that this can also be done if the coefficients of P are periodic functions of θ (or polynomials in other variables). In modern languages supporting overloading or arithmetic functions, all this can be done in an automatic manner.

Note that if the polynomial has degree n_s, the computation up to degree n_s takes $\Theta(n_s^2)$ operations of multiplications of the coefficients.

4.3 Numerical Composition of Periodic Maps

The goal of this section is to deeply discuss how we can numerically compute $\widetilde{W}$ and the compositions of K with $W(\theta, s)$ and $\widetilde{W}(\theta, s)$ only having a numerical representation (or approximation) of K and W in the algorithms 74 and 75.

There are a variety of methods that can be employed to numerically get the composition of a periodic mapping with another (or the same) mapping. Some of these methods depend strongly on the representation of the periodic mapping and others only depend on specific parts of the algorithm.

We start the discussion from the general methods to those that strongly depend on the numerical representation. One expects that the general ones will have a bigger numerical complexity or they will be less accurate.

Before discussing the algorithms, it is important to stress again that for functions of two variables $(\theta, s) \in \mathbb{T} \times [-1, 1]$, there are two complementary ways of looking at them. We can think of them as functions that given θ produce a polynomial in s – this polynomial valued function will be periodic in θ – or we can think of them as polynomials in s taking values in spaces of periodic functions (of the variable θ). Of course, the periodic functions that appear in our interpretation can be discretized either by the values in a grid of points or by the Fourier transform.

Each of these equivalent interpretations will be useful in some algorithms. In the second interpretation, we can "overload" algorithms for standard polynomials to work with polynomials whose coefficients are periodic functions (in particular Horner schemes). In the first interpretation, we can easily parallelize algorithms for polynomials for each of the values of θ using the grid discretization of periodic functions.

Possibly the hardest part of algorithms 74 and 75 is the compositions between K with W and with $\widetilde{W}$. Due to the step *4* of Algorithm 75 the composition should be done so that the output is still a polynomial in s with coefficients that are periodic functions of θ. In our implementation, we use the Automatic Differentiation (AD) approach [17, 96].

If $W(\theta, s) = (W_1(\theta, s), W_2(\theta, s))$ is a function of two variables taking values in $\mathbb{R}^2$, then

$$K \circ W(\theta, s) = \sum_{j=0}^{m-1} K^j(W_1(\theta, s)) \left(b_0 W_2(\theta, s)\right)^j , \tag{4.4}$$

which can be evaluated with $m-1$ polynomial products and $m-1$ polynomial sums using Horner scheme, once we have computed $K^j \circ W_1(\theta, s)$.

The problem of composing a periodic function with a periodic polynomial in s – to produce a polynomial in s taking values in the space of periodic functions – is what we consider now. In particular, we are going to discuss three different approaches and their computational complexities.

The first one is the most general one and it is based on a dynamic programming technique. It assumes some given information to build a table from where the composition can be extracted. In this case the numerical representation is in the part that is assumed to be given.

The second one exploits the Fourier representation in the inputs of the dynamic programming to provide the final full complexity of the composition.

Finally, the third approach also uses the Fourier representation but rather than using the dynamic programming technique, it uses the recurrences in Automatic Differentiation for the sine and cosine functions.

4.3.1 Composition via dynamic programming

The most general method considers S a periodic function, the K^j in (4.4), and $q(s) = \sum_{j=0}^{k} q_j s^j$ a polynomial of a fixed order $k \geqslant 0$ where the q_j are periodic functions of θ that we consider discretized by their values in a grid (the W_1 in (4.4)).

We want to compute the polynomial $p \coloneqq S \circ q$ up to order k. Assume that $\frac{d^j}{d\theta^j} S(q_0)$ for $0 \leqslant j \leqslant k$ are given as input and that they have been previously computed in a bounded computational cost. These inputs in a computer strongly depend on the numerical representation of the periodic function S. In further sections we will consider the Fourier series as a representation which will lead to two different algorithms.

The chain rule gives us a procedure to compute the coefficients of $p(s) = \sum_{j=0}^{k} p_j s^j$. Indeed, one can build a table, whose entries are polynomials in s, like Table 4.1 and which follows the generation rule in Figure 4.1.

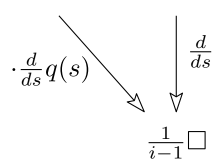

Figure 4.1. Generation rule for $i = 2, \ldots, k+1$ Table 4.1 entries.

The inputs of Table 4.1 are $a_{i,1} = 0$ for $i \neq 1$ and $a_{2,2} = \frac{d}{ds} q(s)$. Then the entries a_{ij} with $2 \leqslant j \leqslant i \leqslant k+1$ are given by

$$a_{ij}(s) = \frac{1}{i-1} \left(\frac{d}{ds} a_{i-1,j}(s) + a_{i-1,j-1}(s) \frac{d}{ds} q(s) \right). \tag{4.5}$$

Thus, the coefficients of $p(s)$ are $p_j = \sum_{l=0}^{k} a_{jl}(0) \frac{d^l}{d\theta^l} S(q_0)$ for $0 \leqslant j \leqslant k$.

Note that it is enough to store in memory k entries of the Table 4.1 to compute all the coefficients p_j.

Moreover, for each entry in the ith row with $i = 2, \ldots, k+1$, one only needs to consider

Table 4.1. Composition of a function with a polynomial.

	$S(q_0)$	$\frac{d}{d\theta}S(q_0)$	$\frac{d^2}{d\theta^2}S(q_0)$	$\cdots$	$\frac{d^{k-1}}{d\theta^{k-1}}S(q_0)$	$\frac{d^k}{d\theta^k}S(q_0)$
p_0	1	0				
p_1	0	$\frac{d}{ds}q(s)$	0			
p_2	0	$\frac{1}{2}\square$	$\frac{1}{2}\square$			
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	0	
p_{k-1}	0	$\frac{1}{k-1}\square$	$\frac{1}{k-1}\square$	$\cdots$	$\frac{1}{k-1}\square$	0
p_k	0	$\frac{1}{k}\square$	$\frac{1}{k}\square$	$\cdots$	$\frac{1}{k}\square$	$\frac{1}{k}\square$

polynomials of degree $k+1-i$. Overall the memory required is at most $\frac{1}{2}k(k+1)$. The number of arithmetic operations following the rule (4.5) are given by the Proposition 78.

Proposition 78. *Let S be a real-periodic function and let $q(s)$ be a real polynomial of degree k. Given $\frac{d^j}{d\theta^j}S(q(0))$ for $j = 0, \ldots, k$. The polynomial $S \circ q$ can be performed using Table 4.1 with $\frac{1}{2}k(k+1)$ units of memory and $\Theta(k^4)$ multiplications and additions.*

Proof. Note that $k(k+1)$ multiplications and $(k+1)^2$ additions are needed to perform the product of two polynomials of degree k. Also k multiplications are needed to perform the derivative of a polynomial of degree k multiplied by a scalar. To bound the number of operations we must distinct three different situations of the Table 4.1.

1. The column $a_{3..k,2}$. $\sum_{i=1}^{k-2}(k-i+1) = \frac{1}{2}(k^2+k-6)$ multiplications.
2. The diagonal $a_{3..k,3..k}$.
   - $\sum_{j=1}^{k-2}(k-j-1)(k-j+1)+1 = \frac{1}{6}(2k^3-3k^2+k-6)$ multiplications.
   - $\sum_{j=1}^{k-2}(k-j-1)^2+1 = \frac{1}{6}(2k^3-9k^2+19k-18)$ additions.
3. The rest.

- $\sum_{j=1}^{k-2}\sum_{i=j+1}^{k-2}(k-i-1)(k-i+1)+(k-i-2)+1 = \frac{1}{12}(7k^4-56k^3+71k^2+38k-24)$ multiplications.
- $\sum_{j=1}^{k-2}\sum_{i=j+1}^{k-2}(k-i-1)^2+(k-i)+1 = \frac{1}{12}(5k^4-36k^3+85k^2-102k+72)$ additions.

Overall $\frac{7}{12}k^4+\Theta(k^3)$ multiplications and $\frac{5}{12}k^4+\Theta(k^3)$ additions. $\square$

The next Theorem 79 summarizes the previous explanations and it provides the complexities to numerically compute $K \circ W$ in (4.4). It assumes that $\frac{d^i}{d\theta^i}S(q_0)$ of Table 4.1, which are the $\frac{d^i}{d\theta^i}K^j(W_1(\theta,0))$ in $K \circ W$, are given as input. These inputs are the only elements in Table 4.1 that depend on the numerical representation of the periodic functions (i.e. the K^j in $K \circ W$) and makes the result in Theorem 79 independent of how periodic functions are represented.

Theorem 79. *For a fixed θ, the computational complexity to compute the compositions of $K(\theta,s) = \sum_{j=0}^{m-1}K^j(\theta)(b_0s)^j$ with $W(\theta,s) = \sum_{j=0}^{k-1}W^j(\theta)(bs)^j$ and $\widetilde{W}(\theta,s)$ using Table 4.1 is $\Theta(mk^4)$ and space $\Omega(k^2)$ assuming $\frac{d^i}{d\theta^i}K^j(W_1^0(\theta))$ as input for $i = 0,\dots,k-1$.*

Remark 80. *In general, if n_θ denotes the mesh size of the variable θ, we will have $k \leqslant m \ll n_\theta$. That is, the mesh size will be much larger than the degree (in s) of $K(\theta,s)$. That means that the parallelization in n_θ will be more advantageous.*

Theorem 79 has an important assumption involving $\frac{d^i}{d\theta^i}K^j(W_1^0(\theta))$ which can have a big impact in the complexity of $K \circ W(\theta,s)$. However, such an impact strongly depends on the numerical representation of K^j and it will be discussed in the Fourier representation case.

4.3.2 Composition in Fourier

Theorem 79 reduces the problem of computing $K \circ W(\theta,s)$ in (4.4) to the problem of computing composition of a periodic function with another one.

In the case of a Fourier representation (see C) of an arbitrary mapping $S\colon \mathbb{T} \to \mathbb{R}$ (the K^j's in (4.4)) , such a composition between Fourier truncated series may require to know the values not in the standard equispaced mesh $\{k/n_\theta\}_{k=0}^{n_\theta-1}$ of θ which hampers the use of the FFT. Indeed, the FFT which states a fast way to biject $\{S(k/n_\theta)\}_{k=0}^{n_\theta-1} \subset \mathbb{R}$ to $\{\widehat{S}^j\}_{j=0}^{n_\theta-1} \subset \mathbb{C}$ such that,

$$S(k/n_\theta) = \sum_{j=0}^{n_\theta-1} \widehat{S}^j e^{2\pi \boldsymbol{i} jk/n_\theta} \quad \text{and} \quad \widehat{S}^j = \frac{1}{n_\theta} \sum_{k=0}^{n_\theta-1} S(k/n_\theta) e^{-2\pi \boldsymbol{i} jk/n_\theta}. \tag{4.6}$$

assumes the mesh of θ to be equispaced. However the $S(q_0)$ may request to evaluate S outside the mesh.

A direct composition of real Fourier series requires a computational complexity $\Theta(n_\theta^2)$. However, nowadays recent algorithms with a $\Theta(n_\theta \log n_\theta)$ complexity efficiently solve this possible bottleneck in the performance of our algorithms. See, for instance, the NFFT3 in [97] or FINUFFT in [98, 99]. The package NFFT3 allows to express $S\colon \mathbb{T} \to \mathbb{R}$ with the same coefficients in (4.6) and perform its evaluation in an even number of non-equispaced nodes $(\theta_k)_{k=0}^{n_\theta-1} \subset \mathbb{T}$ by

$$S(\theta_k) = \sum_{j=0}^{n_\theta-1} \widehat{S}^j e^{-2\pi \boldsymbol{i}(j-n_\theta/2)(\theta_k-1/2)}. \tag{4.7}$$

The corrections of θ_k in (4.7) is because NFFT3 considers $\mathbb{T} \simeq [-1/2, 1/2)$ rather than the other standard equispaced discretization in $[0, 1)$. NFFT3 uses some window functions for a first approximation as a cut-off in the frequency domain and also for a second approximation as a cut-off in time domain. It takes under control (by bounds) these approximations to ensure the solution is a good approximation. Joining these results with Proposition 78 and Theorem 79, we have

Theorem 81. *The computational complexity to compute in Algorithm 75 the compositions of* $K(\theta, s) = \sum_{j=0}^{m-1} K^j(\theta)(b_0 s)^j$ *with the maps* $W(\theta, s) = \sum_{j=0}^{k-1} W^j(\theta)(bs)^j$ *and*

$\widetilde{W}(\theta, s) = \sum_{j=0}^{k-1} \widetilde{W}^j(\theta)(bs)^j$ using Table 4.1 and NFFT3*, and assuming that K^j, W^j and $\widetilde{W}^j$ are expressed with n_θ Fourier coefficients is $\Theta(mk^4 n_\theta + mkn_\theta \log n_\theta)$. The space complexity is $\Omega(kn_\theta + k^2)$.*

Remark 82. *The remark 80 also applies to Theorem 81 in terms of the parallelization of n_θ due to the fact that in general $k \leqslant m \ll n_\theta$. However, in the parallelism case, the space complexity increase to $\Omega(kn_\theta + k^2 n_p)$ with n_p the number of processes although the part corresponding to kn_θ can be shared memory.*

In particular, the NFFT3 *can also be used for the zero order W^0 of Algorithm 74 giving in that case the same complexity as Theorem 81 but with $k = 1$.*

4.3.3 Automatic Differentiation in Fourier

Theorem 79 tells us that the composition $K \circ W(\theta, s)$ can numerically be done independently of the periodic mapping representation. Nevertheless, differentiation is a notoriously ill-posed problem due to the lack of information in the discretized problem. Thus, Table 4.1 is a good option when no advantage of the computer periodic representation exists or $k \ll m$.

Using the representation (C.3), we can use the Taylor expansion of the sine and cosine by recurrence [95, 17]. That is, if $q(s)$ is a polynomial, then $\sin q(s)$ and $\cos q(s)$ are given by $s_0 = \sin q_0$, $c_0 = \cos q_0$ and for $j \geqslant 1$,

$$s_j = \frac{1}{j} \sum_{k=0}^{j-1} (j-k) q_{j-k} c_k, \qquad c_j = -\frac{1}{j} \sum_{k=0}^{j-1} (j-k) q_{j-k} s_k. \tag{4.8}$$

Therefore the computational cost to obtain the sine and cosine of a polynomial is linear with respect to its degree.

Theorem 83 says that the composition of K with W or $\widetilde{W}$ are rather than $\Theta(mk^4 n_\theta + mkn_\theta \log n_\theta)$ like in Theorem 81 just $\Theta(mkn_\theta^2)$. Therefore if $k \ll m$ and n_θ is large, the approach given by Theorem 81 has a better complexity although Theorem 83 will be more stable for larger k.

Theorem 83. *The computational complexity to compute in Algorithm 75 the compositions of* $K(\theta, s) = \sum_{j=0}^{m-1} K^j(\theta)(b_0 s)^j$ *with the maps* $W(\theta, s) = \sum_{j=0}^{k-1} W^j(\theta)(bs)^j$ *and* $\widetilde{W}(\theta, s) = \sum_{j=0}^{k-1} \widetilde{W}^j(\theta)(bs)^j$ *using Automatic Differentiation and assuming that* K^j, W^j *and* $\widetilde{W}^j$ *are expressed with* n_θ *Fourier coefficients is* $\Theta(mkn_\theta^2)$.

4.4 Numerical Results

The van der Pol oscillator [100] is an oscillator with a non-linear damping governed by a second-order differential equation.

As an example, we consider the state-dependent perturbation of the van der Pol oscillator like in [101], which has the form

$$\begin{aligned} \dot{x}(t) &= y(t), \\ \dot{y}(t) &= \mu(1 - x(t)^2)y(t) - x(t) + \varepsilon x(t - r(x(t))), \end{aligned} \tag{4.9}$$

with $\mu > 0$ and $0 < \varepsilon \ll 1$. For the delay function $r(x(t))$ we are going to consider two cases. A pure state-dependent delay case $r(x(t)) = 0.006e^{x(t)}$ or just a constant delay case $r(x(t)) = 0.006$.

The first step consists in computing the change of coordinate K, the frequency ω_0 of the limit cycle and its stability value $\lambda_0 < 0$ for $\varepsilon = 0$. By standard methods of computing periodic orbits and their first variational equations, we compute the limit cycles close to $(x, y) = (2, 0)$ for different values of μ. Table 4.2 shows the values of ω_0 and λ_0 for each of those values of the parameter μ.

The computation of $K(\theta, s)$, following Algorithm 96, up to order 16 in s and with a Fourier mesh size of 1024 allows to plot the isochrons in Figure 4.2.

In the case of ODEs, the isochrons computed by evaluating the expansion can be globalized by integration of the ODE (4.9) forward and backward in time, see [26]. In the case of

Table 4.2. Values of ω_0 and λ_0 for different values of the parameter μ in eq. (4.9) with $\varepsilon = 0$.

μ	ω_0	λ_0
0.25	0.1585366857025485	-0.2509741760777654
0.5	0.1567232109993800	-0.5077310891698608
1	0.1500760842377394	-1.0593769948418550
1.5	0.1409170454968141	-1.6837946490433340

the SDDE, $\varepsilon \neq 0$, propagating backwards is not possible. We hope that this limitation can be overcome, but this will require some new rigorous developments and more algorithms. We think that this is a very interesting problem.

A relevant indicator for engineers is the power spectrum, i.e. the square of the modulus of the complex Fourier coefficients. In Figure 4.3 we illustrate the power spectrum for K^0, since K^0 is the one that is commonly observed in a circuit system.

Due to the quadratic convergence of the Algorithm 96, see [26], the computation of Table 4.2 and Figure 4.2 are performed in less than one minute in a today standard laptop. However, we notice that for values of $\mu > 1.5$ the method may not converge for the unperturbed case, the scaling factor and the Fourier mesh size need to be smaller due to spikes, especially for the high orders in s, i.e. $K^j(\theta)$ for large j. This is an inherent drawback of the numerical representation of periodic functions that can be emphasized with the model involved.

4.4.1 Perturbed case

Let us analyze the case of $\mu = 1.5$ for two different types of delay functions; a constant one $r(x(t)) = 0.006$ and a state-dependent one $r(x(t)) = 0.006e^{x(t)}$.

The two cases have some advantages to be exploited. For instance, in the constant case $\widetilde{W}(\theta, s) = W(\theta - \omega\beta, se^{-\lambda\beta})$ is easier to compute than in the state-dependent case. Since in both cases W and $\widetilde{W}$ must be composed by K, the use of automatic differentiation for the step *4* in Algorithm 75 is still needed. In particular, for the Algorithm 74 and the composition via Theorem 81, the NFFT3 can be used to perform the numerical composition

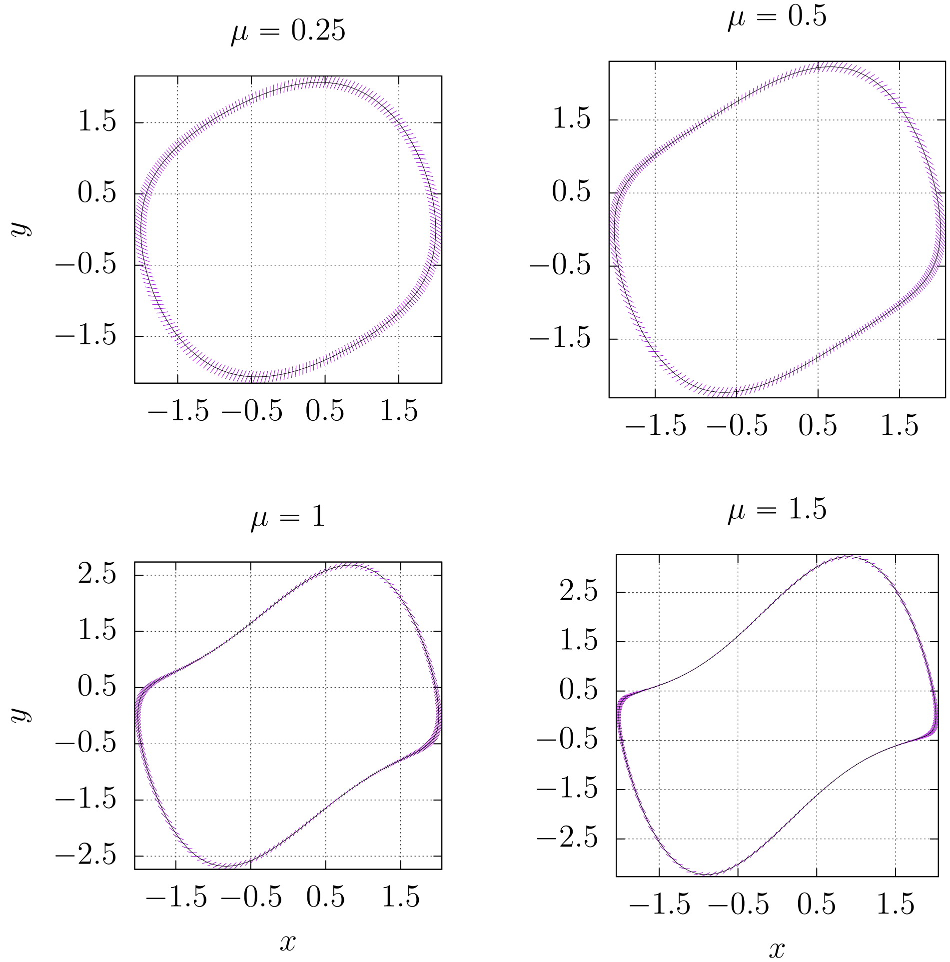

Figure 4.2. Limit cycles and their isochrons for different values of the parameter μ in the unperturbed eq. (4.9).

of K with W and $\widetilde{W}$.

The first steps of our method get ω and λ which we distinguish their values depending on the delay function and the parameter ε. Again here we are assuming $\mu = 1.5$. These values are summarized respectively in Tables 4.3 and 4.4. They were computed fixing a tolerance for the stopping criterion of 10^{-10} in double-precision. Because the result is perturbative, these values are close to those in Table 4.2 and are further as ε increase. Moreover we report a speed factor around 2.25 using the NFFT3 with respect to a direct implementation of the Fourier composition.

Figure 4.4 shows, for different values of ε in eq. (4.9), the logarithmic error of invari-

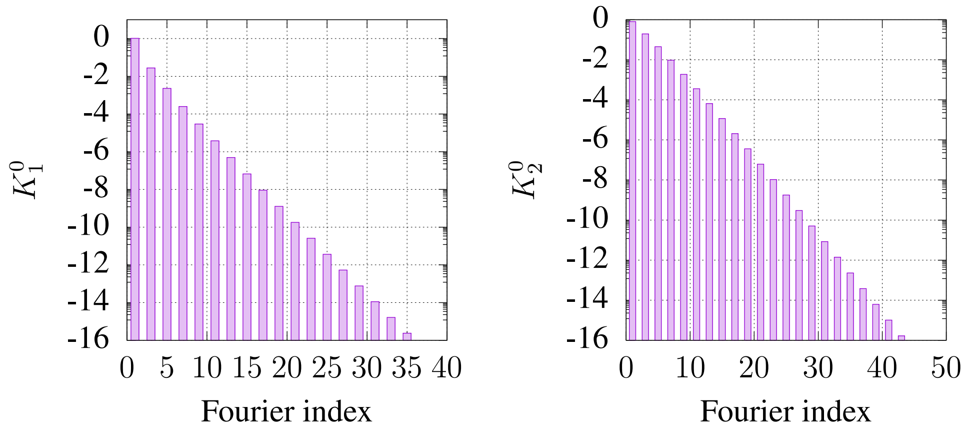

Figure 4.3. Logscale of the power spectrum of $K^0 \equiv (K_1^0, K_2^0)$ for $\mu = 1.5$ and $\varepsilon = 0$ in eq. (4.9).

Table 4.3. Values of ω for different values of ε in eq. (4.9) with $\mu = 1.5$ obtained by Algorithm 74. ω_s corresponds to the state-dependent delay and ω_c the constant delay.

ε	ω_s	ω_c
10^{-4}	0.140908673246532	0.140908547470887
10^{-3}	0.140833302396846	0.140832045466042
10^{-2}	0.140077545298062	0.140065058638519

ance equation for each of the different orders $j \geqslant 0$. That is, the finite system of invariance equations obtained after plugging $W(\theta, s) = \sum W^j(\theta)s^j$ into eq. (2.8) and matching terms of the same order. The state-dependent case requires ε to be smaller than the constant delay case, the reason can be seen from our proofs in Chapter 2.

Figures 4.5 shows the difference between the isochrons for the perturbed and unperturbed case. As one expects from the theorems in Chapter 2, the error is smaller as the perturbation parameter value ε becomes smaller.

An important point in Algorithm 75 is the well-definedness of the composition of K with W and $\widetilde{W}$. Because the state-dependent delays consider much more situations than just the constant delay, one expects that potentially smaller scaling factor compared to the

Table 4.4. Values of λ for different values of ε in eq. (4.9) with $\mu = 1.5$ obtained by Algorithm 75. λ_s corresponds to the state-dependent delay and λ_c the constant delay.

ε	λ_s	λ_c
10^{-4}	-1.6838123845562083	-1.6838091880373793
10^{-3}	-1.6839721186835845	-1.6839401491442914
10^{-2}	-1.6855808865357260	-1.6852607528946115

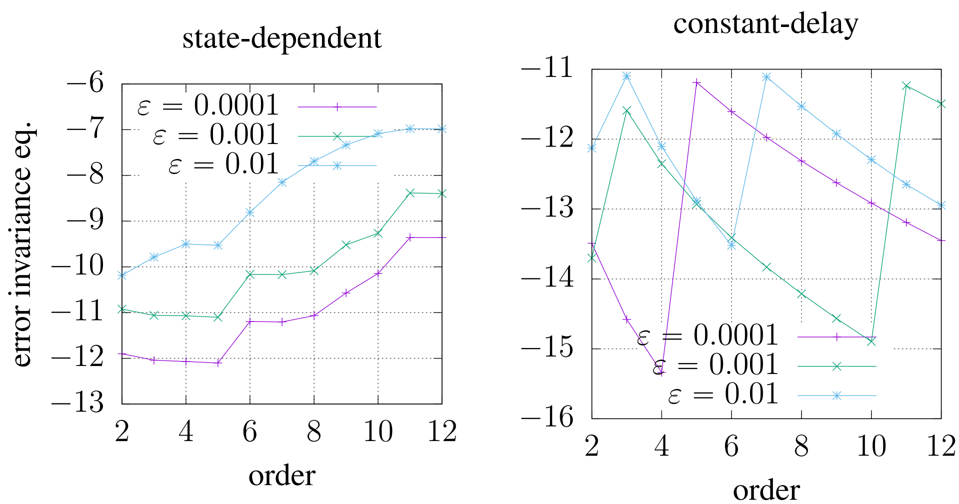

Figure 4.4. Log10 scale of the 2-norm of the error in the invariance equation.

constant delay will be needed as large order is computed. Figure 4.6 shows if ε is large, the scaling factor will need to be small. We also see that for the constant case, it is enough to use a constant scaling factor, and for the state-dependent case, the scaling factor decreases drastically in the first orders.

To illustrate the physical observation, the Figures 4.7 and 4.8 shows the power spectra of the limit cycles after the perturbations. More concretely, Figure 4.7 displays the power spectrum of $(K \circ W)^0$ for the pure state-dependent delay case and $\varepsilon = 0.01$. In contrast with Figure 4.3, we observe that for the even indexes they have non-zero values in the double-precision arithmetic sense. On the other hand, Figure 4.8 shows that these non-zero values in the even indexes are not present in the constant delay case and the power spectrum for the case $\varepsilon > 0$ is away from that when $\varepsilon = 0$ as ε increase.

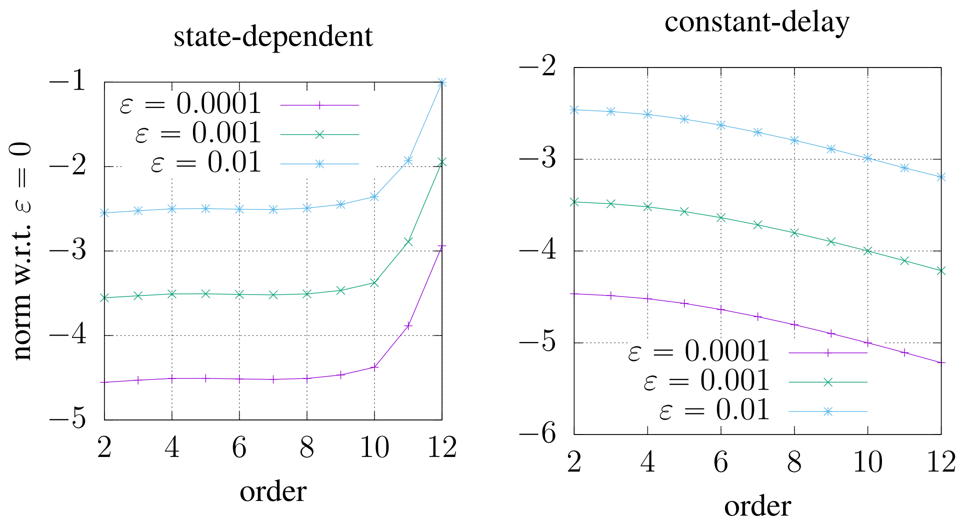

Figure 4.5. Log10 scale of the 2-norm of the difference between the perturbed and unperturbed cases. That is, $\|K^j - (K \circ W)^j\|$.

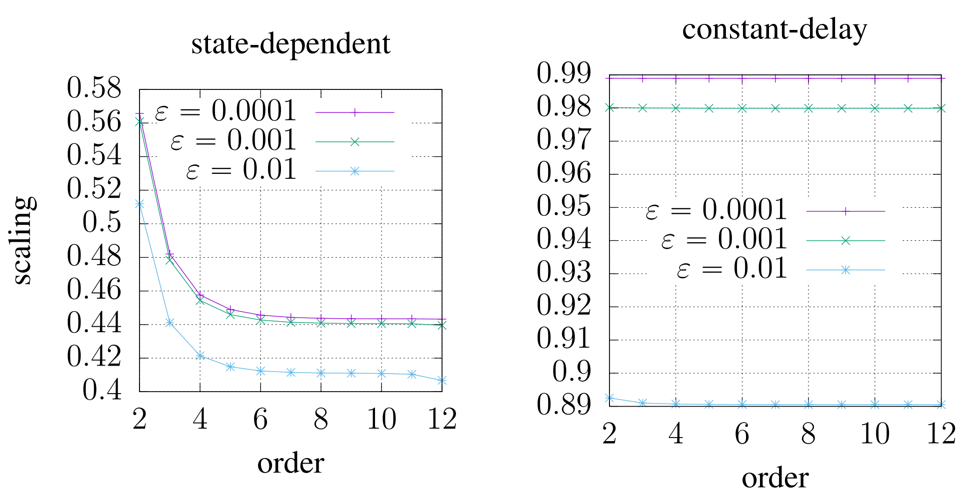

Figure 4.6. Scaling factor to ensure that the composition of K with W and with $\widetilde{W}$ in Algorithm 75 are well-defined.

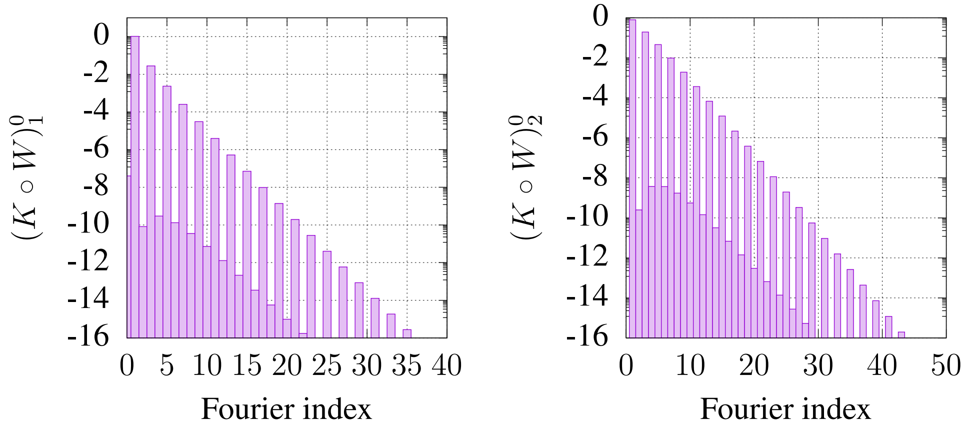

Figure 4.7. Log10 scale of the power spectrum of $(K \circ W)^0$ for $\mu = 1.5$, $\varepsilon = 0.01$ and the state-dependent delay $r(x(t)) = 0.006e^{x(t)}$ in eq. (4.9).

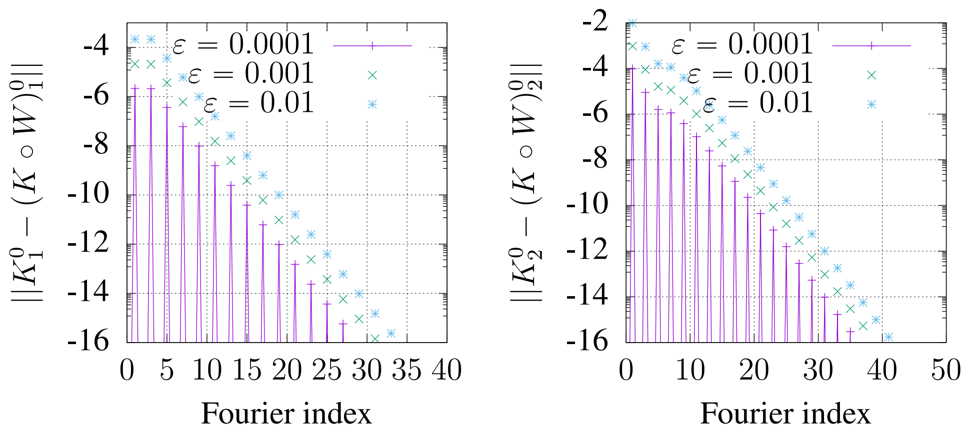

Figure 4.8. Log10 scale of the difference between the power spectrum of K^0 and the power spectrum of $(K \circ W)^0$ for $\mu = 1.5$, different values of ε and constant delay $r = 0.006$ in eq. (4.9).

Appendices

APPENDIX A
REGULARITY PROPERTIES

One of the sources of complication in the study of delay equations – especially state dependent delay equations – is that the equations involve compositions, which have many surprising properties. In this appendix we collect a few of them. A systematic study of the composition operator in C^r spaces which are the most natural for our problem is in [54].

A.1 Function Spaces

Let ℓ be a positive integer, let X be a Banach space and $U \subset X$ be a an open set. For functions on U taking values in another Banach space Y, we can define derivatives [102, 103], and Lipschitz and Hölder regularity of the derivatives.

We recall that the j derivative is a j-multilinear function from $X^{\otimes j}$ to Y and that there is a natural norm for multilinear functions (supremum of the norm of the values when the arguments have norm 1).

We denote by $C^\ell(U, Y)$ the space of all functions with uniformly bounded continuous derivatives up to order ℓ. We endow $C^\ell(U, Y)$ with the norm

$$\|f\|_{C^\ell} = \max_{0 \leqslant j \leqslant \ell} \sup_{\xi \in U} \|D^j f(\xi)\|_{X^{\otimes j} \to Y},$$

so that $C^\ell(U, Y)$ is a Banach space. We denote by

$$\mathrm{Lip}\,(F) = \sup_{x,y \in U, x \neq y} \|F(x) - F(y)\|_Y / \|x - y\|_X$$

Definition 84. *We say that $K\colon U \to Y$ is in $C^{\ell+\mathrm{Lip}}(U, Y)$ when K has ℓ derivatives and the ℓ derivative is Lipschitz.*

We endow $C^{\ell+\mathrm{Lip}}(U,Y)$ with the norm:

$$\|K\|_{C^{\ell+\mathrm{Lip}}} = \max\left\{\|K\|_{C^\ell}, \mathrm{Lip}(D^\ell K)\right\} \tag{A.1}$$

which makes $C^{\ell+\mathrm{Lip}}$ into a Banach space.

A similar definition can be written when U is a Riemannian manifold. In this thesis we will use the case that $U = \mathbb{T}$ or $U = \mathbb{T} \times \mathbb{R}$.

Remark 85. *We note that Definition 84 assumes uniform bounds of the derivatives in the whole domain. There are other very standard definitions of differentiable sets that only assume continuity and bounds in compact subsets of U. Even when $U = \mathbb{R}^n$ these definitions (e.g. Whitney topology, very natural in differential geometry) do not lead to $C^{\ell+\mathrm{Lip}}$ being a Banach space and we will not use them.*

A.2 Faà di Bruno formula

We quote Faà di Bruno formula, which deals with the derivatives of composition of two functions.

Lemma 86. *Let $g(x)$ be defined on a neighborhood of x^0 in a Banach space E, and have derivatives up to order n at x^0. Let $f(y)$ be defined on a neighborhood of $y^0 = g(x^0)$ in a Banach space F, and have derivatives up to order n at y^0. Then, the nth derivative of the composition $h(x) = f[g(x)]$ at x^0 is given by the formula*

$$h_n = \sum_{k=1}^{n} f_k \sum_{p(n,k)} n! \prod_{i=1}^{n} \frac{g_i^{\lambda_i}}{(\lambda_i!)(i!)^{\lambda_i}}. \tag{A.2}$$

In the above expression, we set

$$h_n = \frac{d^n}{dx^n} h(x^0), \quad f_k = \frac{d^k}{dy^k} f(y^0), \quad g_i = \frac{d^i}{dx^i} g(x^0),$$

and

$$p(n,k) = \left\{ (\lambda_1, \ldots, \lambda_n) : \lambda_i \in \mathbb{N}, \sum_{i=1}^{n} \lambda_i = k, \sum_{i=1}^{n} i\lambda_i = n \right\}.$$

This can be proved by the Chain Rule and induction. See [104] for a proof.

A.3 Simple estimates on Composition

We will need the following property of the composition operator, one can refer to [54] for more details.

Lemma 87. *Let X, Y, Z be Banach spaces. Let $E \subset X$, $F \subset Y$ be open subsets.*

Assume that: $g \in C^{\ell+\mathrm{Lip}}(E, Y)$, $f \in C^{\ell+\mathrm{Lip}}(F, Z)$ and that $g(E) \subset F$ so that $f \circ g$ can be defined. Then, $f \circ g \in C^{\ell+\mathrm{Lip}}(E, Z)$, and

$$\|f \circ g\|_{C^{\ell+\mathrm{Lip}}(E,Z)} \leqslant M_\ell \|f\|_{C^{\ell+\mathrm{Lip}}(F,Z)} \big(1 + \|g\|^{\ell+1}_{C^{\ell+\mathrm{Lip}}(E,Y)}\big) \tag{A.3}$$

The proof of Lemma 87 just uses the Faà di Bruno formula in Lemma 86 for the derivatives of the composition. To control the Lipschitz constant of the ℓ derivative, we use that the Lipschitz constant of product and composition satisfy the same formulas as those of the derivative with an inequality in place of equality.

In (A.3) we can take any set F that contains $g(E)$. The results are sharper when we take F as small as possible.

A.4 The mean value theorem

Definition 88. *We say that an open set $U \subset X$ is a compensated domain when it is connected, and there is $C > 0$ such that for any $x, y \in U$, there is a C^1 path $\gamma \subset U$ such that*

$$\textit{length}(\gamma) \leqslant C\|x - y\|$$

In particular, a convex domain is compensated with $C = 1$.

We also recall the fundamental theorem of calculus.

Theorem 89. *Assume that $U \subset X$ is open connected, $F\colon U \to X$ is a C^1 function, $x, y \in U$ and that γ is a C^1 path joining x, y. Then*

$$F(x) - F(y) = \int_0^1 DF(\gamma(t))\, D\gamma(t)\, dt$$

As a corollary of Theorem 89 we have that

$$\|F(x) - F(y)\| \leqslant \|DF\|_{C^0} \cdot \text{length}(\gamma) \leqslant \|F\|_{C^1} \cdot \text{length}(\gamma)$$

If the domain U is compensated, we obtain that

$$\|F(x) - F(y)\| \leqslant C\|F\|_{C^1}\|x - y\|$$

In particular, C^1 functions on compensated domains are Lipschitz.

The conclusion that C^1 implies Lipschitz, is not true if the domain is not compensated. It is not difficult to obtain examples of domains where C^1 functions are not continuous even when $X = \mathbb{R}^2$.

Lemma 90. *Assume that for some $\ell \geqslant 1$, $\|f\|_{C^{\ell+\mathrm{Lip}}} \leqslant A$, $\|g_1\|_{C^{\ell-1+\mathrm{Lip}}}$, $\|g_2\|_{C^{\ell-1+\mathrm{Lip}}} \leqslant B$. Then:*

$$\|f \circ g_1 - f \circ g_2\|_{C^{\ell-1+\mathrm{Lip}}} \leqslant C(A, B)\|g_1 - g_2\|_{C^{\ell-1+\mathrm{Lip}}} \tag{A.4}$$

Proof. By the fundamental theorem of calculus we have pointwise

$$f \circ g_1 - f \circ g_2 = \int_0^1 Df(g_2 + t(g_1 - g_2))(g_1 - g_2)\, dt$$

If we interpret the above as identity among functions we have

$$\|f \circ g_1 - f \circ g_2\|_{C^{\ell-1+\mathrm{Lip}}} \leqslant \int_0^1 C\|Df(g_2 + t(g_1 - g_2))\|_{C^{\ell-1+\mathrm{Lip}}} \cdot \|(g_1 - g_2)\|_{C^{\ell-1+\mathrm{Lip}}}\, dt$$

Using Lemma 87, $\|Df(g_2 + t(g_1 - g_2))\|_{C^{\ell-1+\mathrm{Lip}}}$ is bounded by a function of A and B, we are done. $\square$

A.5 Interpolation

We quote the following result from [105, 106]. See [54] for a modern, very simple proof valid for functions on compensated domains in Banach spaces.

Lemma 91. *Let U be a convex and bounded open subset of a Banach space E, F be a Banach space. Let r, s, t be positive numbers, $0 \leqslant r < s < t$, and $\mu = \frac{t-s}{t-r}$. There is a constant $M_{r,t}$, such that if $f \in C^t(U, F)$, then*

$$\|f\|_{C^s} \leqslant M_{r,t} \|f\|_{C^r}^{\mu} \|f\|_{C^t}^{1-\mu}.$$

A.6 Closure Properties of $C^{\ell+\mathrm{Lip}}$ ball

We quote a very practical result which appears as Lemma 2.4 in [50]. (This paper is largely reproduced as a chapter in [107]. See Lemma (2.5) on p. 39.) A related notion, Quasi-Banach space, was used in [108].

Lemma 92. *Let $U \subset X$ be a compensated domain.*

Denote by $\mathbf{B}$ a closed ball in $C^{\ell+\mathrm{Lip}}(U, Y)$. Let $\{u_n\}_{n\in\mathbb{N}} \subset \mathbf{B}$ be such that u_n converges pointwise weakly to u. Then, $u \in \mathbf{B}$.

Furthermore, the derivatives of u_n of order up to ℓ converge weakly to the derivatives of u.

We note that the hypothesis of Lemma 92 are easy to verify in operators that involve composition. The propagated bounds just amount to proving that the size of derivatives of composition of two functions can be estimated by the sizes of the derivatives of the original functions. The contraction properties are done under the assumption that the functions are smooth so that one can use the mean value theorem.

A similar result to Lemma 92 is the following, which appears as Lemma 6.1.6 in [78, p. 151].

Lemma 93. *Let $U \subset X$ be an open set. Denote by $\mathbf{B}$ a closed ball in $C^{\ell+\mathrm{Lip}}(U, Y)$. Let $\{u_n\}_{n\in\mathbb{N}} \subset \mathbf{B}$ be such that u_n converges uniformly to u. Then, $u \in \mathbf{B}$.*

Furthermore the derivatives of u_n of order up to ℓ converge uniformly to the derivatives of u away from the boundary of U.

Both Lemma 92 and Lemma 93 remain true when we replace the spaces of $C^{\ell+\mathrm{Lip}}$ functions by Hölder spaces.

Remark 94. *It is instructive to compare the proofs of Lemma 92 and Lemma 93 in their original references.*

The proof of [50] is based on considering restrictions to lines. Then, one can apply Arzela-Ascoli theorem and extract converging subsequences. The assumption of a weak pointwise limit ensures that the limit is unique. The uniformity of the $C^{\ell+\mathrm{Lip}}$ norms of the functions ensures the existence of derivatives and the convergence.

The proof of [78] goes along different lines. It shows that there are bounds on the derivatives by the C^0 norms and the size of the ball. An alternative argument is to use interpolation inequalities in Lemma 91, which provides uniform convergence of the derivatives on U (also near the boundary).

As a consequence of Lemma 92, we have the following version of the contraction mapping.

Lemma 95. *With the same notation of Lemma 92.*

Assume $\mathscr{T} : \mathbf{B} \to \mathbf{B}$ satisfies that there exists $\kappa < 1$ such that

$$\|\mathscr{T}(u) - \mathscr{T}(v)\|_{C^0} \leqslant \kappa \|u - v\|_{C^0} \quad \forall u, v \in \mathbf{B}$$

Then, $\mathscr{T}$ has a unique fixed point u^* in $\mathbf{B}$.

For any $u \in \mathbf{B}$, and $0 \leqslant j \leqslant \ell$

$$\|\mathscr{T}^n(u) - u^*\|_{C^{j+\mathrm{Lip}}} \leqslant C\kappa^{n\frac{\ell-j}{\ell+1}} \|\mathscr{T}(u) - u\|_{C^0}^{\frac{\ell-j}{\ell+1}}$$

where C is a constant that depends on the radius of the ball $\mathbf{B}$ and j.

Furthermore,

$$\|u - u^*\|_{C^{j+\mathrm{Lip}}} \leqslant C(1-\kappa)^{-\frac{\ell-j}{\ell+1}} \|\mathscr{T}(u) - u\|_{C^0}^{\frac{\ell-j}{\ell+1}}$$

Proof. When X is finite dimensional (or just separable), Lemma 92 is a corollary of Ascoli-Arzela theorem. For any subsequence of u_n we can extract a sub-subsequence that converges in C^ℓ sense. The limit of this sub-subsequence has to be u. It follows that the u_n converges to u in C^ℓ sense. It then follows that the ℓ-derivative is Lipschitz.

If X is infinite dimensional, one can repeat the above argument restricting to lines. The uniform regularity assumed on u_n translates to uniform regularity of u restricted to lines.

We refer to [50] for more details. Indeed, [50] only needs to assume that the sequence converges weakly pointwise. The convergence properties are only used to guarantee the uniqueness of the limit obtained through compactness (The paper [50] is written when the domain U is the whole space, but this is not used).

Once we have the closure property, the existence of the unique fixed point is as in Banach contraction. We observe that for any $u \in \mathbf{B}$,

$$\|\mathscr{T}^{n+1}(u) - \mathscr{T}^n(u)\|_{C^0} \leqslant \kappa^n \|\mathscr{T}(u) - u\|_{C^0}$$

Using the interpolation inequalities Lemma 91 and that the $C^{\ell+\mathrm{Lip}}$ norms of the iterates are bounded, we obtain

$$\|\mathscr{T}^{n+1}(u) - \mathscr{T}^n(u)\|_{C^{j+\mathrm{Lip}}} \leqslant C\kappa^{n\frac{\ell-j}{\ell+1}} \|\mathscr{T}(u) - u\|_{C^0}^{\frac{\ell-j}{\ell+1}} \tag{A.5}$$

From this one obtains that $\mathscr{T}^n(u) - u = \sum_{k=1}^{n} (\mathscr{T}^k(u) - \mathscr{T}^{k-1}(u))$ is an absolutely convergent series in the $C^{j+\mathrm{Lip}}$ sense. Let u^* be the fixed point. Using (A.5) to estimate the series, we obtain:

$$\|u - u^*\|_{C^{j+\mathrm{Lip}}} \leqslant C(1 - \kappa^{\frac{\ell-j}{\ell+1}})^{-1} \|\mathscr{T}(u) - u\|_{C^0}^{\frac{\ell-j}{\ell+1}}.$$

On the other hand, from the standard Banach fixed point theory, we obtain that $\|u - u^*\|_{C^0} \leqslant (1-\kappa)^{-1} \|\mathscr{T}(u) - u\|_{C^0}$. By Lemma 91 we obtain

$$\|u - u^*\|_{C^{j+\mathrm{Lip}}} \leqslant C(1 - \kappa)^{-\frac{\ell-j}{\ell+1}} \|\mathscr{T}(u) - u\|_{C^0}^{\frac{\ell-j}{\ell+1}}.$$

It is easy to see that this bound is better than the previously obtained one summing the series. $\square$

APPENDIX B

COMPUTATION OF (K, ω_0, λ_0) – UNPERTURBED CASE

For completeness, we quote the Algorithm 4.4 in [26] adding some practical comments. That algorithm allows us to numerically compute ω_0, λ_0 and $K\colon \mathbb{T} \times [-1, 1] \to \mathbb{R}^2$ in (2.6). We note that the algorithm has quadratic convergence as it was proved in [26].

Algorithm 96. *Quasi-Newton method*

- ⋆ `Input`*:* $\dot{x} = X(x)$ *in* $\mathbb{R}^2$, $K(\theta, s) = K^0(\theta) + K^1(\theta) b_0 s$, $\omega_0 > 0$, $\lambda_0 \in \mathbb{R}$, *scaling factor* $b_0 > 0$ *and a tolerance* tol.

- ⋆ `Output`*:* $K(\theta, s) = \sum_{j=0}^{m-1} K^j(\theta)(b_0 s)^j$, ω_0 *and* λ_0 *such that* $\|E\| \ll 1$.

1. $E \leftarrow X \circ K - (\omega_0 \partial_\theta + \lambda_0 s \partial_s) K$.

2. *Solve* $DK\tilde{E} = E$ *and denote* $\tilde{E} \equiv (\tilde{E}_1, \tilde{E}_2)$.

3. $\sigma \leftarrow \int_0^1 \tilde{E}_1(\theta, 0)\, d\theta$ *and* $\eta \leftarrow \int_0^1 \partial_s \tilde{E}_2(\theta, 0)\, d\theta$.

4. $E_1 \leftarrow \tilde{E}_1 - \sigma$ *and* $E_2 \leftarrow \tilde{E}_2 - \eta s$.

5. *Solve* $(\omega_0 \partial_\theta + \lambda_0 s \partial_s) S_1 = E_1$ *imposing*

$$\int_0^1 S_1(\theta, 0)\, d\theta = 0. \tag{B.1}$$

6. *Solve* $(\omega_0 \partial_\theta + \lambda_0 s \partial_s) S_2 - \lambda_0 S_2 = E_2$ *imposing*

$$\int_0^1 \partial_s S_2(\theta, 0)\, d\theta = 0. \tag{B.2}$$

7. $S \equiv (S_1, S_2)$.

8. *Update:* $K \leftarrow K + DKS$, $\omega_0 \leftarrow \omega_0 + \sigma$ *and* $\lambda_0 \leftarrow \lambda_0 + \eta$.

9. *Iterate from* (*1*) *to* (*8*) *until convergence with tolerance* tol *in* K, ω *and* λ. *Then undo the scaling* b_0.

Algorithm 96 requires some practical considerations:

i. *Initial guess.* $K^0 \colon \mathbb{T} \to \mathbb{R}^2$ will be a parameterization of the periodic orbit of the ODE with frequency ω_0. It can be obtained, for instance, by a Poincaré section method, continuation of integrable systems or Lindstedt series. An approximation for $K^1 \colon \mathbb{T} \to \mathbb{R}^2$ and λ can be obtained by solving the variational equation

$$DX \circ K^0(\theta)U(\theta) = \omega_0 \frac{d}{d\theta} U(\theta),$$
$$U(0) = Id_2.$$

Hence if $(e^{\lambda_0/\omega_0}, K^1(0))$ is the eigenpair of $U(1)$ such that $\lambda_0 < 0$, then $K^1(\theta) = U(\theta)K^1(0)e^{-\lambda_0\theta/\omega_0}$.

ii. *Stopping criteria.* As any Newton method, a possible condition to stop the iteration can be when either $\|E\|$ or $\max\{\|DKS\|, |\sigma|, |\eta|\}$ is smaller than a given tolerance.

Note that the *a posteriori* theorems in [26] give a criterion of smallness on the error depending on properties of the function K. If these criteria are satisfied, one can ensure that there is a true solution close to the numerical one.

iii. *Uniqueness.* Note that in the steps *5* and *6*, which involve solving the cohomology equations, the solutions are determined only up to adding constants in the zeroth or first order terms. We have adopted the conventions (B.1), (B.2). These conventions make the solution operator linear (which matches well the standard theory of Nash-Moser methods since it is easy to estimate the norm of the solutions).

As it is shown in [26], the algorithm converges quadratically fast to a solution, but since the problem is underdetermined, we have to be careful when comparing solutions of different discretization. In [26] there is discussion of the uniqueness, but for our purposes in Chapter 4, any of the solutions will work. The uniqueness of the solutions considered in Chapter 4 is discussed in section 4.1.

iv. *Convergence.* It has been proved in [26] that the quasi-Newton method still has quadratic convergence.

Note that it is remarkable that we can implement a Newton like method without having to store – much less invert – any large matrix. Note also that we can get a Newton method even if the derivative of the operator in the fixed point equation has eigenvalues 1. See Remark 73.

v. *Cohomological equations.* The most delicate steps of above algorithm are *5* and *6*, which are often called *cohomology equations*. These steps involve solving PDEs whereas the other steps are much simpler. In case of a Fourier representation (see C), they can be addressed by using Proposition 76.

vi. *Linear system.* Step *2* can be addressed by Lemma 77.

APPENDIX C

FOURIER DISCRETIZATION OF PERIODIC FUNCTIONS

As mentioned before, the key step of Algorithm 96 is to solve the equations in steps *5* and *6*. Their numerical resolution will be particularly efficient when the functions are discretized in Fourier-Taylor series. This is the only discretization we consider in Chapter 4 providing a deep discussion.

Recall that a function $S\colon \mathbb{R} \to \mathbb{R}$ is called periodic when $S(\theta+1) = S(\theta)$ for all θ.

To get a computer representation of a periodic function, we can either take a mesh in θ, i.e. $(\theta_k)_{k=0}^{n_\theta-1}$ and store the values of S at these points: $\breve{S} = (\breve{S}_k)_{k=0}^{n_\theta-1} \in \mathbb{R}^{n_\theta}$ with $\breve{S}_k = S(\theta_k)$ or we can take advantage of the periodicity and represent it in a trigonometric basis.

The Discrete Fourier Transform (DFT), and also its inverse, allows to switch between the two representations above. If we fix a mesh of points of size n_θ uniformly distributed in $[0,1)$, i.e. $\theta_k = k/n_\theta$, the DFT is:

$$\widehat{S} = (\widehat{S}_k)_{k=0}^{n_\theta-1} \in \mathbb{C}^{n_\theta}$$

so that

$$\breve{S}_k = \sum_{j=0}^{n_\theta-1} \widehat{S}_j e^{2\pi \boldsymbol{i} jk/n_\theta} \tag{C.1}$$

or equivalently

$$\widehat{S}_k = \frac{1}{n_\theta} \sum_{j=0}^{n_\theta-1} \breve{S}_j e^{-2\pi \boldsymbol{i} jk/n_\theta}. \tag{C.2}$$

In the case of a real valued function, $\widehat{S}_0$ is real and the complex numbers $\widehat{S}$ satisfy Hermitian symmetry, i.e. $\widehat{S}_k = \widehat{S}^*_{n_\theta-k}$ (denoting by * the complex conjugate), which implies $\widehat{S}_{n_\theta/2}$ real when n_θ is even. Then, we define real numbers $(a_0; a_k, b_k)_{k=1}^{\lceil n_\theta/2 \rceil -1}$ if n_θ is odd, here $\lceil\cdot\rceil$ denotes the ceiling function, otherwise $(a_0, a_{n_\theta/2}; a_k, b_k)_{k=1}^{n_\theta/2-1}$ defined by

$$a_0 = 2\widehat{S}_0,\, a_{n_\theta/2} = 2\widehat{S}_{n_\theta/2},\, a_k = 2\,\mathrm{Re}\,\widehat{S}_k \text{ and } b_k = -2\,\mathrm{Im}\,\widehat{S}_k$$

with $1 \leqslant k < \lceil n_\theta/2 \rceil$.

Thus, S can be approximated by

$$S(\theta) = \frac{a_0}{2} + \frac{a_{n_\theta/2}}{2}\cos(\pi n_\theta \theta) + \sum_{k=1}^{\lceil n_\theta/2\rceil-1} a_k \cos(2\pi k\theta) + b_k \sin(2\pi k\theta) \tag{C.3}$$

where the coefficient $a_{n_\theta/2}$ only appears when n_θ is even and it refers to the aliasing notion in signal theory.

Therefore (C.3) is equivalent to (C.1) but rather than $2n_\theta$ real numbers, only half of them are needed.

Henceforth, all real periodic functions S can be represented in a computer by an array of length n_θ whose values are either the values of S on a grid or the Fourier coefficients. These two representations are, for all practical purposes equivalent since there is a well known algorithm, Fast Fourier Transform (FFT), which allows to go from one to the other in $\Theta(n_\theta \log n_\theta)$ operations. The FFT has very efficient implementations so that the theoretical estimates on time are realistic (we can use FFTW3 [109], which optimizes the use of the hardware).

We can also think of functions of two variables $W(\theta, s)$ where one variable θ is periodic and the other variable s is a real variable. In the numerical implementations, the variable s will be discretized as a polynomial. Thus $W(\theta, s)$ can be thought as a function of θ taking values in polynomials of length n_s. Hence, a function of two variables with periodicity as above will be discretized by an array $n_\theta \times n_s$. The meaning could be that it is a polynomial for each value of θ in a mesh or that it a polynomial of whose coefficients are Fourier coefficients. Alternatively, we could think of $W(\theta, s)$ as a polynomial in s taking values in a space of periodic functions.

This mixed representation of Fourier series in one variable and power series in another variable, is often called Fourier-Taylor series and has been used in celestial mechanics for a long time, dates back to [110] or earlier. We note that, modern computer languages allow to overload the arithmetic operations among different types in a simple way.

It is important to note that all the operations in Algorithm 96 are fast either on the Fourier representation or in the values of a mesh representation. For example, the product of two functions or the composition on the left with a known function are fast in the representation by values in a mesh. More importantly for us, as we will see, the solution of cohomology equations is fast in the Fourier representation. On the other hand, there are other steps of Algorithm 96, such as adding, are fast in both representations.

Similar consideration of the efficiency of the steps will apply to the algorithms needed to solve our problem. The main novelty of the algorithms in Chapter 4 compared with those of [26] is that we will need to compose some of the unknown functions (in [26] the unknowns are only composed on the left with a known function). The algorithms we use to deal with composition is presented in Section 4.3. The composition operator will be the most delicate numerical aspect, which was to be expected, since it was also the most delicate step in the analysis in Chapter 2. The composition operator is analytically subtle [54, 111].

Remark 97. *Fourier series are extremely efficient for smooth functions which do not have very pronounced spikes. For rather smooth functions – a situation that appears often in practice – it seems that Fourier Taylor series is better than other methods.*

It should be noted, however that in several models of interest in electronics and neuroscience, the solutions move slowly during a large fraction of the period, but there is a fast movement for a short time (bursting). In these situations, the Fourier scheme has the disadvantage that the coefficients decrease slowly and that the discretization method does not allow to put more effort in describing the solutions during the times that they are indeed changing fast. Hence, the Fourier methods become unpractical when the limit cycles are *bursting. In such cases, one can use other methods of discretization. In this thesis, we will not discuss alternative numerical methods, but note that the theoretical estimates in Chapter 2 are independent of the discretization. We hope to come back to implementing the results here in other discretizations.*

Remark 98. *One of the confusing practical aspects of the actual implementation is that the coefficients of the Fourier arrays are often stored in a complicated order to optimize the operations and the access during the FFT.*

For example, the coefficients a_k's and b_k's in (C.3)*, in* FFTW3*, the* `fftw_plan_r2r_1d` *uses the following order of the Fourier coefficients in a real array $(v_0, \dots, v_{n_\theta-1})$.*

$$\begin{aligned} v_0 &= a_0, \\ v_k &= 2a_k \text{ and } v_{n_\theta-k} = -2b_k \quad \text{for } 1 \leqslant k < \lceil n_\theta/2 \rceil, \\ v_{n_\theta/2} &= a_{n_\theta/2} \end{aligned}$$

where the index $n_\theta/2$ is taken into consideration if and only if n_θ is even. Another standard order in other packages is just $(a_0, a_{n_\theta/2}; a_k, b_k)$ in sequential order or $(a_0; a_k, b_k)$ if n_θ is odd.

To measure errors and size of functions represented by Fourier series, we have found useful to deal with weighted norms involving the Fourier coefficients.

$$\begin{aligned} \|S\|_{w\ell^1,n} &= 2(n_\theta/2)^n |\widehat{S}_{n_\theta/2}| + \sum_{k=1}^{\lceil n_\theta/2 \rceil - 1} ((n_\theta - k)^n + k^n) |\widehat{S}_k| \\ &= (n_\theta/2)^n |a_{n_\theta/2}| + \frac{1}{2} \sum_{k=1}^{\lceil n_\theta/2 \rceil - 1} ((n_\theta - k)^n + k^n)(a_k^2 + b_k^2)^{1/2}. \end{aligned}$$

where, again, the term for $n_\theta/2$ only appears if n_θ is even.

The smoothness of S can be measured by the speed of decay of the Fourier coefficients and indeed, the above norms give useful regularity classes that have been studied by harmonic analysts.

Remark 99. *The relation of the above regularity classes with the most common C^m is not straightforward, as it is well known by Harmonic analysts [112].*

Riemann-Lebesgue's Lemma tells us that if S is continuous and periodic, $\widehat{S}_k \to 0$ as $k \to \infty$ and in general if S is m times differentiable, then $|\widehat{S}_k||k|^m$ tends to zero. In particular, $|\widehat{S}_k| \leqslant C/|k|^m$ for some constant $C > 0$.

In the other direction, from $|\widehat{S}_k| \leqslant C/|k|^m$ we cannot deduce that $S \in C^m$.

One has to use more complicated methods. In [113] it was found that one could find a practical method based on Littlewood-Paley theorem (see [112]) which states that the function S is in α-Hölder space with $\alpha \in \mathbb{R}_+$ if and only if, for each $\eta \geqslant 0$ there is constant $C > 0$ such that for all $t > 0$.

$$\left\| \left(\frac{\partial}{\partial t} \right)^{\eta} e^{-t\sqrt{-\Delta}\theta} \right\|_{L^\infty(\mathbb{T})} \leqslant Ct^{\alpha-\eta}.$$

The above formula is easy to implement if one has the Fourier coefficients, as it is the case in our algorithms.

REFERENCES

[1] J. A. Wheeler and R. P. Feynmann, “Classical electrodynamics in terms of direct interparticle action,” *Rev. Mod. Phys.*, vol. 21, no. 3, pp. 425–433, 1949.

[2] R. D. Driver, “A two-body problem of classical electrodynamics: The one-dimensional case,” *Annals of Physics*, vol. 21, no. 1, pp. 122 –142, 1963.

[3] R. D. Driver, “A neutral system with state-dependent delay,” *J. Differential Equations*, vol. 54, no. 1, pp. 73–86, 1984.

[4] F. Hartung, T. Krisztin, H.-O. Walther, and J. Wu, “Functional differential equations with state-dependent delays: Theory and applications,” in *Handbook of differential equations: ordinary differential equations. Vol. III*, ser. Handb. Differ. Equ. Elsevier/North-Holland, Amsterdam, 2006, pp. 435–545.

[5] R. C. Calleja, A. R. Humphries, and B. Krauskopf, “Resonance phenomena in a scalar delay differential equation with two state-dependent delays,” *SIAM J. Appl. Dyn. Syst.*, vol. 16, no. 3, pp. 1474–1513, 2017.

[6] M. Kloosterman, S. A. Campbell, and F. J. Poulin, “A closed NPZ model with delayed nutrient recycling,” *J. Math. Biol.*, vol. 68, no. 4, pp. 815–850, 2014.

[7] G. Fan, S. A. Campbell, G. S. K. Wolkowicz, and H. Zhu, “The bifurcation study of $1:2$ resonance in a delayed system of two coupled neurons,” *J. Dynam. Differential Equations*, vol. 25, no. 1, pp. 193–216, 2013.

[8] W. G. Aiello, H. I. Freedman, and J. Wu, “Analysis of a model representing stage-structured population growth with state-dependent time delay,” *SIAM J. Appl. Math.*, vol. 52, no. 3, pp. 855–869, 1992.

[9] H.-O. Walther, “On a model for soft landing with state-dependent delay,” *J. Dynam. Differential Equations*, vol. 19, no. 3, pp. 593–622, 2007.

[10] M. C. Mackey, “Commodity price fluctuations: Price dependent delays and nonlinearities as explanatory factors,” *Journal of Economic theory*, vol. 48, no. 2, pp. 497–509, 1989.

[11] J. Hale, *Theory of functional differential equations*, Second. Springer-Verlag, New York-Heidelberg, 1977, pp. x+365, Applied Mathematical Sciences, Vol. 3.

[12] J. K. Hale and S. M. Verduyn Lunel, *Introduction to functional-differential equations*, ser. Applied Mathematical Sciences. Springer-Verlag, New York, 1993, vol. 99, pp. x+447, ISBN: 0-387-94076-6.

[13] O. Diekmann, S. A. van Gils, S. M. Verduyn Lunel, and H.-O. Walther, *Delay equations*, ser. Applied Mathematical Sciences. Springer-Verlag, New York, 1995, vol. 110, pp. xii+534, Functional, complex, and nonlinear analysis, ISBN: 0-387-94416-8.

[14] X. Cabré, E. Fontich, and R. de la Llave, "The parameterization method for invariant manifolds. I. Manifolds associated to non-resonant subspaces," *Indiana Univ. Math. J.*, vol. 52, no. 2, pp. 283–328, 2003.

[15] ——, "The parameterization method for invariant manifolds. II. Regularity with respect to parameters," *Indiana Univ. Math. J.*, vol. 52, no. 2, pp. 329–360, 2003.

[16] ——, "The parameterization method for invariant manifolds. III. Overview and applications," *J. Differential Equations*, vol. 218, no. 2, pp. 444–515, 2005.

[17] Haro, À., Canadell, M., Figueras, J.-L., Luque, A., and Mondelo, J.-M., *The parameterization method for invariant manifolds*, ser. Applied Mathematical Sciences. Springer, [Cham], 2016, vol. 195, pp. xvi+267, From rigorous results to effective computations, ISBN: 978-3-319-29660-9; 978-3-319-29662-3.

[18] X. Li and R. de la Llave, "Construction of quasi-periodic solutions of delay differential equations via KAM techniques," *J. Differential Equations*, vol. 247, no. 3, pp. 822–865, 2009.

[19] X. He and R. de la Llave, "Construction of quasi-periodic solutions of state-dependent delay differential equations by the parameterization method I: Finitely differentiable, hyperbolic case," *J. Dynam. Differential Equations*, vol. 29, no. 4, pp. 1503–1517, 2017.

[20] ——, "Construction of quasi-periodic solutions of state-dependent delay differential equations by the parameterization method II: Analytic case," *J. Differential Equations*, vol. 261, no. 3, pp. 2068–2108, 2016.

[21] H.-O. Walther, "The solution manifold and c^1-smoothness for differential equations with state-dependent delay," *Journal of Differential Equations*, vol. 195, no. 1, pp. 46 –65, 2003.

[22] J. Yang, J. Gimeno, and R. D. la Llave, "Parameterization method for state-dependent delay perturbation of an ordinary differential equation," *SIAM Journal on Mathematical Analysis*, vol. 53, no. 4, pp. 4031–4067, 2021. eprint: https://doi.org/10.1137/20M1311430.

[23] J. Yang, J. Gimeno, and R. de la Llave, "Persistence and smooth dependence on parameters of periodic orbits in functional differential equations close to an ode or an evolutionary pde," –, 2021, https://arxiv.org/abs/2103.05203.

[24] 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," *SIAM J. Appl. Dyn. Syst.*, 2021, to appear, https://arxiv.org/abs/2005.06086.

[25] D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, "Relativistic invariance and Hamiltonian theories of interacting particles," *Rev. Modern Phys.*, vol. 35, pp. 350–375, 1963.

[26] G. Huguet and R. de la Llave, "Computation of limit cycles and their isochrons: Fast algorithms and their convergence," *SIAM J. Appl. Dyn. Syst.*, vol. 12, no. 4, pp. 1763–1802, 2013.

[27] J. Carr, *Applications of centre manifold theory*, ser. Applied Mathematical Sciences. Springer-Verlag, New York-Berlin, 1981, vol. 35, pp. vi+142, ISBN: 0-387-90577-4.

[28] A. T. Winfree, "Patterns of phase compromise in biological cycles," *J. Math. Biol.*, vol. 1, no. 1, pp. 73–95, 1974/75.

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

[30] A. Casal, L. Corsi, and R. de la Llave, "Expansions in the delay of quasi-periodic solutions for state dependent delay equations," *Journal of Physics A: Mathematical and Theoretical*, vol. 53, no. 23, p. 235 202, 2020.

[31] J. Mallet-Paret and R. D. Nussbaum, "Global continuation and asymptotic behaviour for periodic solutions of a differential-delay equation," *Annali di Matematica Pura ed Applicata*, vol. 145, no. 1, pp. 33–128, 1986.

[32] J. Mallet-Paret, R. D. Nussbaum, and P. Paraskevopoulos, "Periodic solutions for functional-differential equations with multiple state-dependent time lags," *Topol. Methods Nonlinear Anal.*, vol. 3, no. 1, pp. 101–162, 1994.

[33] I. Maroto, C. Núñez, and R. Obaya, "Exponential stability for nonautonomous functional differential equations with state-dependent delay," *Discrete Contin. Dyn. Syst. Ser. B*, vol. 22, no. 8, pp. 3167–3197, 2017.

[34] J. Mallet-Paret and R. D. Nussbaum, "Stability of periodic solutions of state-dependent delay-differential equations," *J. Differential Equations*, vol. 250, no. 11, pp. 4085–4103, 2011.

[35] J. Gimeno, "Effective methods for recurrence solutions in delay differential equations," PhD thesis, Universitat de Barcelona, 2019.

[36] X. Li and R. de la Llave, "Convergence of differentiable functions on closed sets and remarks on the proofs of the "converse approximation lemmas"," *Discrete Contin. Dyn. Syst. Ser. S*, vol. 3, no. 4, pp. 623–641, 2010.

[37] R. D. Nussbaum, "Periodic solutions of analytic functional differential equations are analytic," *Michigan Math. J.*, vol. 20, pp. 249–255, 1973.

[38] J. L. Kaplan and J. A. Yorke, "On the stability of a periodic solution of a differential delay equation," *SIAM J. Math. Anal.*, vol. 6, pp. 268–282, 1975.

[39] J. Jaquette, J.-P. Lessard, and K. Mischaikow, "Stability and uniqueness of slowly oscillating periodic solutions to Wright's equation," *J. Differential Equations*, vol. 263, no. 11, pp. 7263–7286, 2017.

[40] J. Mallet-Paret and R. D. Nussbaum, "Analyticity and nonanalyticity of solutions of delay-differential equations," *SIAM J. Math. Anal.*, vol. 46, no. 4, pp. 2468–2500, 2014.

[41] J. Sieber, "Finding periodic orbits in state-dependent delay differential equations as roots of algebraic equations," *Discrete Contin. Dyn. Syst.*, vol. 32, no. 8, pp. 2607–2651, 2012.

[42] J. B. van den Berg, C. Groothedde, and J.-P. Lessard, "A general method for computer-assisted proofs of periodic solutions in delay differential problems," *Journal of Dynamics and Differential Equations*, 2020.

[43] R. Szczelina, "A computer assisted proof of multiple periodic orbits in some first order non-linear delay differential equation," *Electron. J. Qual. Theory Differ. Equ.*, Paper No. 83, 19, 2016.

[44] R. Szczelina and P. Zgliczyński, "Algorithm for rigorous integration of delay differential equations and the computer-assisted proof of periodic orbits in the Mackey-Glass equation," *Found. Comput. Math.*, vol. 18, no. 6, pp. 1299–1332, 2018.

[45] J. K. Hale, "Solutions near simple periodic orbits of functional differential equations," *J. Differential Equations*, vol. 7, pp. 126–138, 1970.

[46] J. K. Hale and M. Weedermann, "On perturbations of delay-differential equations with periodic orbits," *J. Differential Equations*, vol. 197, no. 2, pp. 219–246, 2004.

[47] J. Mawhin, "Équations intégrales et solutions périodiques des systèmes différentiels non linéaires," *Acad. Roy. Belg. Bull. Cl. Sci. (5)*, vol. 55, pp. 934–947, 1969.

[48] ——, "Periodic solutions of nonlinear functional differential equations," *J. Differential Equations*, vol. 10, pp. 240–261, 1971.

[49] R. D. Nussbaum, "Periodic solutions of nonlinear autonomous functional differential equations," in *Functional differential equations and approximation of fixed points (Proc. Summer School and Conf., Univ. Bonn, Bonn, 1978)*, ser. Lecture Notes in Math. Vol. 730, Springer, Berlin, 1979, pp. 283–325.

[50] O. E. Lanford, "Bifurcation of periodic solutions into invariant tori: The work of ruelle and takens," in *Nonlinear problems in the physical sciences and biology*, I. Stakgold, D. D. Joseph, and D. H. Sattinger, Eds., Lecture Notes in Mathematics, Vol. 322, Berlin: Springer-Verlag, 1973, pp. 159–192.

[51] C. Chicone, "Inertial and slow manifolds for delay equations with small delays," *J. Differential Equations*, vol. 190, no. 2, pp. 364–406, 2003.

[52] G. Kiss and J.-P. Lessard, "Computational fixed-point theory for differential delay equations with multiple time lags," *J. Differential Equations*, vol. 252, no. 4, pp. 3093–3115, 2012.

[53] C. M. Groothedde and J. D. Mireles James, "Parameterization method for unstable manifolds of delay differential equations," *J. Comput. Dyn.*, vol. 4, no. 1-2, pp. 21–70, 2017.

[54] 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.

[55] P. A. M. Dirac, "Classical theory of radiating electrons.," *Proc. R. Soc. Lond., Ser. A*, vol. 167, pp. 148–169, 1938.

[56] L. D. Landau and E. M. Lifshitz, *The classical theory of fields*, ser. Revised second edition. Course of Theoretical Physics, Vol. 2. Translated from the Russian by Morton Hamermesh. Pergamon Press, Oxford-London-Paris-Frankfurt; Addison-Wesley Publishing Co. Inc. Reading, Mass., 1962, pp. ix+404.

[57] G. N. Plass, "Classical electrodynamic equations of motion with radiative reaction," *Rev. Mod. Phys.*, vol. 33, pp. 37–62, 1961.

[58] L. Bel, "Predictive relativistic mechanics," *Ann. Inst. H. Poincaré Sect. A (N.S.)*, vol. 14, pp. 189–203, 1971.

[59] F. Rohrlich, "The equations of motion of classical charges," *Ann. Physics*, vol. 13, pp. 93–109, 1961.

[60] H. Spohn, *Dynamics of charged particles and their radiation field*. Cambridge University Press, Cambridge, 2004, pp. xvi+360, ISBN: 0-521-83697-2.

[61] F. Rohrlich, *Classical charged particles*, Third. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2007, pp. xvi+305, ISBN: 978-981-270-004-9; 981-270-004-8.

[62] J. D. Jackson, *Classical electrodynamics*. John Wiley & Sons, 2007.

[63] A. Zangwill, *Modern electrodynamics*. Cambridge University Press, 2013.

[64] J. A. Wheeler and R. P. Feynmann, "Interaction with the absorber as the mechanism of radiation," *Rev. Mod. Phys.*, vol. 17, no. 2/3, pp. 157–181, 1945.

[65] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*. W. H. Freeman and Co., San Francisco, Calif., 1973, ii+xxvi+1279+iipp.

[66] E. I. Verriest, "The principle of "borrowed feedback" and application to control and observation for systems with implicit state dependent delay," in *Time delay systems*, ser. Adv. Delays Dyn. Vol. 7, Springer, Cham, 2017, pp. 47–61.

[67] ——, "Post-Newtonian gravitation," in *Mathematical and computational approaches in advancing modern science and engineering*, Springer, [Cham], 2016, pp. 153–164.

[68] J. Martin and J. L. Sanz, "Slow motion approximation in predictive relativistic mechanics. I. Approximated dynamics up to order c^{-4}," *J. Math. Phys.*, vol. 19, no. 9, pp. 1887–1891, 1978.

[69] C. Chicone and B. Mashhoon, "Nonlocal Lagrangians for accelerated systems," *Ann. Phys.*, vol. 16, no. 12, pp. 811–823, 2007.

[70] O. Perron, "Über Stabilität und asymptotisches Verhalten der Integrale von Differentialgleichungssystemen," *Math. Z.*, vol. 29, no. 1, pp. 129–160, 1929.

[71] J.-L. Figueras, M. Gameiro, J.-P. Lessard, and R. de la Llave, "A framework for the numerical computation and a posteriori verification of invariant objects of evolution equations," *SIAM J. Appl. Dyn. Syst.*, vol. 16, no. 2, pp. 1070–1088, 2017.

[72] M. Gameiro and J.-P. Lessard, "A posteriori verification of invariant objects of evolution equations: Periodic orbits in the Kuramoto-Sivashinsky PDE," *SIAM J. Appl. Dyn. Syst.*, vol. 16, no. 1, pp. 687–728, 2017.

[73] J.-L. Figueras and R. de la Llave, "Numerical computations and computer assisted proofs of periodic orbits of the Kuramoto-Sivashinsky equation," *SIAM J. Appl. Dyn. Syst.*, vol. 16, no. 2, pp. 834–852, 2017.

[74] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod; Gauthier-Villars, Paris, 1969, pp. xx+554.

[75] R. E. Showalter, *Monotone operators in Banach space and nonlinear partial differential equations*, ser. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 1997, vol. 49, pp. xiv+278, ISBN: 0-8218-0500-2.

[76] L. Bers, F. John, and M. Schechter, *Partial differential equations*, ser. Lectures in Applied Mathematics, Vol. III. Interscience Publishers John Wiley & Sons, Inc. New York-London-Sydney, 1964, pp. xiii+343.

[77] R. E. Showalter, *Hilbert space methods for partial differential equations*. Electronic Monographs in Differential Equations, San Marcos, TX, 1994, iii+242 pp.; available at http://ejde.math.swt.edu/mtoc.html, Electronic reprint of the 1977 original.

[78] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, ser. Lecture Notes in Math. Berlin: Springer-Verlag, 1981, vol. 840, pp. iv+348, ISBN: 3-540-10557-3.

[79] G. R. Sell and Y. You, *Dynamics of evolutionary equations*, ser. Applied Mathematical Sciences. Springer-Verlag, New York, 2002, vol. 143, pp. xiv+670, ISBN: 0-387-98347-3.

[80] I. D. Chueshov, *Introduction to the Theory of Infinite-Dimensional Dissipative Systems*. AKTA Scientific Publishing House, Kharkiv, 2002.

[81] J. Wu, *Theory and applications of partial functional-differential equations*, ser. Applied Mathematical Sciences. Springer-Verlag, New York, 1996, vol. 119, pp. x+429, ISBN: 0-387-94771-X.

[82] N. Kosovalić and B. Pigott, "Self-excited vibrations for damped and delayed 1-dimensional wave equations," *J. Dynam. Differential Equations*, vol. 31, no. 1, pp. 129–152, 2019.

[83] ——, "Self-excited vibrations for damped and delayed higher dimensional wave equations," *Discrete Contin. Dyn. Syst.*, vol. 39, no. 5, pp. 2413–2435, 2019.

[84] Y. Lan and P. Cvitanović, "Variational method for finding periodic orbits in a general flow," *Phys. Rev. E (3)*, vol. 69, no. 1, pp. 016217, 10, 2004.

[85] G. Arioli and H. Koch, "Families of periodic solutions for some Hamiltonian PDEs," *SIAM J. Appl. Dyn. Syst.*, vol. 16, no. 1, pp. 1–15, 2017.

[86] P. Zgliczyński, "Rigorous numerics for dissipative partial differential equations. II. Periodic orbit for the Kuramoto-Sivashinsky PDE—a computer-assisted proof," *Found. Comput. Math.*, vol. 4, no. 2, pp. 157–185, 2004.

[87] J. Boussinesq, "Théorie des ondes et des remous qui se propagent le long d'un canal rectangulaire horizontal, en communiquant au liquide contenu dans ce canal des vitesses sensiblement pareilles de la surface au fond," *J. Math. Pures Appl. (2)*, vol. 17, pp. 55–108, 1872.

[88] R. de la Llave, "A smooth center manifold theorem which applies to some ill-posed partial differential equations with unbounded nonlinearities," *J. Dynam. Differential Equations*, vol. 21, no. 3, pp. 371–415, 2009.

[89] R. de la Llave and Y. Sire, "An a posteriori KAM theorem for whiskered tori in Hamiltonian partial differential equations with applications to some ill-posed equations," *Arch. Ration. Mech. Anal.*, vol. 231, no. 2, pp. 971–1044, 2019.

[90] H. Cheng and R. de la Llave, "Stable manifolds to bounded solutions in possibly ill-posed PDEs," *J. Differential Equations*, vol. 268, no. 8, pp. 4830–4899, 2020.

[91] M. E. Taylor, *Partial differential equations III. Nonlinear equations*, Second, ser. Applied Mathematical Sciences. Springer, New York, 2011, vol. 117, pp. xxii+715, ISBN: 978-1-4419-7048-0.

[92] I. Chueshov, *Dynamics of quasi-stable dissipative systems*, ser. Universitext. Springer, 2015, pp. xvii+390, ISBN: 978-3-310-22902-7; 978-3-319-22903-4.

[93] S.-N. Chow and H. Leiva, "Unbounded perturbation of the exponential dichotomy for evolution equations," *J. Differential Equations*, vol. 129, no. 2, pp. 509–531, 1996.

[94] Y. Latushkin, S. Montgomery-Smith, and T. Randolph, "Evolutionary semigroups and dichotomy of linear skew-product flows on locally compact spaces with Banach fibers," *J. Differential Equations*, vol. 125, no. 1, pp. 73–116, 1996.

[95] Knuth, D. E., *The art of computer programming. Vol. 2*, Second. Addison-Wesley Publishing Co., Reading, Mass., 1981, pp. xiii+688, Seminumerical algorithms, Addison-Wesley Series in Computer Science and Information Processing, ISBN: 0-201-03822-6.

[96] Griewank, A. and Walther, A., *Evaluating derivatives*, Second. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008, pp. xxii+438, Principles and techniques of algorithmic differentiation, ISBN: 978-0-898716-59-7.

[97] Keiner, J., Kunis, S., and Potts, D., "Using NFFT 3—a software library for various nonequispaced fast Fourier transforms," *ACM Trans. Math. Software*, vol. 36, no. 4, Art. 19, 30, 2009.

[98] A. H. Barnett, J. Magland, and L. af Klinteberg, "A parallel nonuniform fast Fourier transform library based on an "exponential of semicircle" kernel," *SIAM J. Sci. Comput.*, vol. 41, no. 5, pp. C479–C504, 2019.

[99] A. H. Barnett, "Aliasing error of the $\exp(\beta\sqrt{1-z^2})$ kernel in the nonuniform fast Fourier transform," *Appl. Comput. Harmon. Anal.*, vol. 51, pp. 1–16, 2021.

[100] van der Pol, B., "A theory of the amplitude of free and forced triode vibrations," *Radio Review.*, vol. 1, no. 701-710, pp. 754–762, 1920.

[101] Hou, A. and Guo, S., "Stability and Hopf bifurcation in van der Pol oscillators with state-dependent delayed feedback," *Nonlinear Dynam.*, vol. 79, no. 4, pp. 2407–2419, 2015.

[102] J. Dieudonné, *Foundations of modern analysis*. Academic Press, New York-London, 1969, pp. xviii+387, Enlarged and corrected printing, Pure and Applied Mathematics, Vol. 10-I.

[103] L. H. Loomis and S. Sternberg, *Advanced calculus*. Jones and Bartlett Publishers, Boston, MA, 1990, pp. xii+580, ISBN: 0-86720-122-3.

[104] R. Abraham and J. Robbin, *Transversal mappings and flows*, ser. An appendix by Al Kelley. W. A. Benjamin, Inc., New York-Amsterdam, 1967, pp. x+161.

[105] J. Hadamard, "Sur le module maximum d'une fonction et de ses derives," *Bull. Soc. Math. France*, vol. 42, pp. 68–72, 1898.

[106] A. Kolmogoroff, "On inequalities between the upper bounds of the successive derivatives of an arbitrary function on an infinite interval," *Amer. Math. Soc. Translation*, vol. 1949, no. 4, p. 19, 1949.

[107] 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.

[108] F. Hartung and J. Turi, "On differentiability of solutions with respect to parameters in state-dependent delay equations," *J. Differential Equations*, vol. 135, no. 2, pp. 192–237, 1997.

[109] Frigo, M. and Johnson, S. G., "The design and implementation of FFTW3," *Proceedings of the IEEE*, vol. 93, no. 2, pp. 216–231, 2005, Special issue on "Program Generation, Optimization, and Platform Adaptation".

[110] Broucke, R. and Garthwhite, K., "A programming system for analytical series expansions on a computer," *Celestial mechanics*, vol. 1, no. 2, pp. 271–284, 1969.

[111] Appell, J. and Zabrejko, P. P., *Nonlinear superposition operators*, ser. Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1990, vol. 95, pp. viii+311, ISBN: 0-521-36102-8.

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

[113] de la Llave, R. and Petrov, N. P., "Regularity of conjugacies between critical circle maps: An experimental study," *Experiment. Math.*, vol. 11, no. 2, pp. 219–241, 2002.