SIAM J. MATH. ANAL.
Vol. 53, No. 4, pp. 4031–4067




# PARAMETERIZATION METHOD FOR STATE-DEPENDENT DELAY PERTURBATION OF AN ORDINARY DIFFERENTIAL EQUATION*

JIAQI YANG†, JOAN GIMENO‡, AND RAFAEL DE LA LLAVE†

**Abstract.** We consider state-dependent delay equations (SDDEs) obtained by adding delays to a planar ODE with a limit cycle. Situations of this type appear in models of several physical processes, where small delay effects are added. Even if the delays are small, they are very singular perturbations since the natural phase space of an SDDE is an infinite-dimensional space. We show that for the SDDE, there are initial values which lead to solutions similar to those of the ODE. That is, there exist a periodic solution and a two parameter family of solutions whose evolution converges to the periodic solution (in the ODE case, these are called the isochrons). The method of proof bypasses the theory of existence, uniqueness, dependence on parameters of SDDE. 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. We provide a result 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 the results of numerical computations or formal expansions. The method of proof also leads to practical algorithms. In a companion paper, we present the implementation details and representative results. One feature of the method presented here is that it allows us 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 $C^1$).

**Key words.** state-dependent delay equations, limit cycle, slow stable manifolds, perturbation

**AMS subject classifications.** 34K19, 39A23, 39A60, 39A30, 37G15

**DOI.** 10.1137/20M1311430

## 1. Introduction.

Many causes in the natural sciences take some time to generate effects. If one incorporates this delay into 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.

In the case that the delay is constant (say 1), one can prescribe the data in an interval $[-1, 0]$ and then propagate the differential equation. This leads to a rather satisfactory theory of existence and uniqueness and even a qualitative theory [Dri84, Hal77, HVL93, DvGVLW95]. Note that the natural phase space is a space of functions on $[-1, 0]$. This is an infinite-dimensional space.

*Received by the editors January 8, 2020; accepted for publication (in revised form) March 15, 2021; published electronically July 19, 2021.
https://doi.org/10.1137/20M1311430
**Funding:** The work of the first and third authors was partially supported by the NSF grant DMS-1800241. The work of the second author was supported by the Spanish Ministry of Economy and Competitiveness, through the María de Maeztu Programme for Units of Excellence in R&D (MDM-2014-0445), the Spanish grant PGC2018-100699-B-I00 (MCIU/AEI/FEDER, UE), and the Catalan grant 2017 SGR 1374. The work of the first and second authors was supported by the H2020-MCA-RISE grant 734577, which supported visits of the first author to the University of Barcelona and the second author to Georgia Institute of Technology.
†School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160 USA (yangjiaqi@zoho.com, rafael.delallave@math.gatech.edu).
‡Department of Mathematics, University of Rome Tor Vergata, Via della Ricerca Scientifica 1, 00133 Rome Italy (joan@maia.ub.es).



© 2021 Jiaqi Yang, Joan Gimeno, and Rafael De La Llave

When the delay is not a constant and depends on the state, one needs to consider state-dependent delay equations (SDDEs for short). In contrast with the constant delay case, the mathematical theory of SDDEs has complications. The papers [Wal03] made important progress for the appropriate phase space for SDDE. We refer to [HKWW06] for a very comprehensive survey of the mathematical theory and the (rather numerous) applications.

In this paper, we consider a simple model (two-dimensional ordinary differential equation 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 in this paper (see (2.2)) appear in several concrete problems in the natural sciences (circuits, neuroscience, and population dynamics); see [HKWW06].

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 an ODE and the perturbed case is an infinite-dimensional problem. 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 us to establish many other properties (e.g., dependence on parameters) which may be false for general solutions of SDDE. 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 treat 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 leave gaps, so that a general theory giving a description for all the solutions may have problems [CJS63].

**1.1. Overview.** Let us start by an informal overview of the method. It is known that in a neighborhood of a limit cycle of a two-dimensional ODE, we can find $K : \mathbb{T} \times [-1,1] \to \mathbb{R}^2$, and $\omega_0$ and $\lambda_0$ in such a way that for any $\theta, s$, the function given by

$$x(t) = K(\theta + \omega_0 t, s e^{\lambda_0 t}) \tag{1.1}$$

solves the ODE; see [HdlL13]. The fact that all the functions of the form (1.1) are solutions of the original ODE is equivalent to a functional equation for $K$, $\omega_0$, and $\lambda_0$, which we call the "invariance equation." Efficient methods to study the resulting functional equation were presented in [HdlL13]. We will, henceforth, assume that $K$, $\omega_0$, $\lambda_0$ are known.

Similarly, for the perturbed case, when we impose that for fixed $\theta$, $s$ the function of the form

$$x(t) = K \circ W(\theta + \omega t, s e^{\lambda t}) \tag{1.2}$$

is a solution of our delay differential equation, we obtain a functional equation for $W$, $\omega$, $\lambda$ (see (2.6)). Note that the unknowns in (2.6) are the embedding $W$ and the numbers $\omega, \lambda$.

Our goal will be to solve (2.6) 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 in that the domain of definition of the unknown has to depend on the details of the unknown. Similar problems appear in the theory of center manifolds [Car81]. Here we have to resort to cut-offs and extensions. After this process, we get a prepared equation, (2.7), which has the same format as (2.6), and agrees with (2.6) 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 paper is Theorem 10, which establishes that with respect to some condition numbers of the problem, verified for small enough $\varepsilon$, given an approximate solution of the extended invariance equation (2.7), one obtain a true solution nearby. (This is sometimes referred as "a posteriori" format.)

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

**1.2. Some comments on the results.** In a geometric language, we can describe our procedure as that we are finding an embedding of the phase space of the ODE into the phase space of the SDDE in such a way that the range of the embedding is foliated by solutions of the SDDE and that the flow in this manifold is similar to the flow of the ODE. Note that this bypasses the need for developing a general theory of solutions of the SDDE. We only construct a two-dimensional manifold of solutions of the SDDE. For these solutions, it is possible to comfortably discuss many desirable properties such as smooth dependence on the model, etc.

Philosophies similar to that of this paper (finding solutions of functional equations that imply the existence of solutions of special kinds) have already been used in [HdlL17, HDlL16, CCdlL20] to study quasi-periodic solutions of SDDEs. For constant delay equations, we can find [KY74, Les10, KL12, vdBGL20] for the study of periodic solutions. The paper [KL17] studies specific models similar to ours for constant delay perturbations. The paper [LdlL09] studies quasi-periodic solutions analytically; [GMJ17] studies numerically unstable manifolds near fixed points. The papers [Sie17, CHK17, HBC+16, MKW14] study normal forms and numerical computations of periodic and quasi-periodic solutions of SDDEs and obtain bifurcations and numerical solutions. Even if the evolutions of the SDDEs considered above are difficult to define as smooth evolutions, we believe that the results above can be understood as suggesting the existence of a subsystem of the evolution which indeed experiences bifurcations. The careful numerical solutions of [CHK17] can presumably be validated.

By solving the invariance equation, (2.7), one actually obtains a parameterization of the limit cycle and its isochrons (two-dimensional slow stable manifolds of the limit cycle). In other words, $K \circ W(\theta, 0)$ parameterizes the limit cycle, and for fixed $\theta$, we have $K \circ W(\theta, s)$ parameterizes the local slow stable manifold of the point $K \circ W(\theta, 0)$ on the limit cycle. We remark that in some previous work, e.g., Chapter 10 of [HVL93], persistence of limit cycles was studied with a different method in the setting of retarded functional differential equations (RFDEs). They have also studied infinite-dimensional stable manifolds of periodic orbits of RFDE. In this paper, we study SDDE, and get a parameterization of a submanifold of the infinite-dimensional stable manifold, which corresponds to the eigenvalue of the time-T map with largest modulus (dominating the evolution). In this sense, we think that the manifold in this paper is practically more relevant than the infinite-dimensional manifolds. For a more detailed comparison of the results and approach of this paper with the study of SDDEs as evolutionary equations, see section 4.3.

Of course, the notions of approximate solutions and that of solutions close to the approximate ones, require us to specify a norm in the space of functions. In [HdlL13], it was natural to specify analytic norms. In this paper, however, we use spaces of finitely differentiable functions. Indeed, we conjecture that the solutions we produce are not more than finitely differentiable.

The a posteriori format of Theorem 10 allows us to validate approximate solutions produced even by nonrigorous methods. In that respect, we note that the related paper [GYdlL19] develops numerical methods that produce approximate solutions. Some papers that study formal expansions in the delay are [CF80] for periodic solutions and bifurcations, mostly with constant delay, and [CCdlL20] which studies periodic and quasi-periodic solutions for SDDEs (and even more general models such as those appearing in electrodynamics).

Using Theorem 10, we obtain that the numerical solutions produced in [GYdlL19], have true solutions nearby and that the formal expansions produced in [CCdlL20] are not just formal expansions but are asymptotic to true solutions. For an earlier example of related philosophies, we mention that asymptotic expansions for equations with small constant delay was produced and validated in the paper [Chi03].

A rather subtle point is that we do not obtain uniqueness of the solution. The reason is that the nature of the problem involves cutting off the perturbation and the solution produced may depend on the cut-off function used. Both the finite regularity and the lack of uniqueness due to the introduction of a cut-off are reminiscent to effects found in the study of center manifolds [Car81, Lan73]. Of course, since one of the goals of the paper is to remedy the paucity of solutions of SDDEs, having many solutions is a feature not a bug. The dependence of the solutions on the cut-offs has to be small as the delay tends to zero (note that the asymptotic expansions in [CCdlL20] do not depend on the cut-off), but we expect that they are small in other senses similar to the situation in center manifolds [Sij85]. We will not formulate here results making precise this intuition.

We hope that the methods of this paper can be extended to prove the existence of other finite-dimensional families of solutions that are not close to families of solutions of the unperturbed ODE.

### 1.3. Organization of the paper.

We introduce the problem and formulate the equations to be solved in section 2. In section 3 we present some notations and some classical results in functional analysis which will be used in the proof. We state our main results in section 4. We give an overview of the proof in section 5. Detailed proofs of the results are given in section 6.

## 2. Formulation of the problem.

We consider an ODE in the plane

$$\dot{x}(t) = X_0(x(t)), \tag{2.1}$$

where $x(t) \in \mathbb{R}^2$, $X_0 : \mathbb{R}^2 \to \mathbb{R}^2$ is analytic. We assume the above equation (2.1) 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.

The goal of this paper is to study an SDDE that is a “small” modification of (2.1) 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 into the model, we are led to

$$\dot{x}(t) = X(x(t), \varepsilon x(t - r(x(t)))), \qquad 0 \le \varepsilon \ll 1. \tag{2.2}$$

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. Equation (2.2) is formally a perturbation of (2.1) with $X(x, 0) = X_0(x)$.

We can rewrite (2.2) as

$$\dot{x}(t) = X(x(t), 0) + \varepsilon P(x(t), x(t - r(x(t))), \varepsilon), \tag{2.3}$$

where we define

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

The question we want to address in this paper is to find a two-dimensional family of solutions of (2.2) which resembles the two-dimensional family of solutions of (2.1). This is a much simpler problem than developing a general theory of the existence of solutions to an SDDE, which is a rather difficult problem. Nevertheless, persistence of some family of solutions is of physical interest.

Note that, when $\varepsilon > 0$, (2.3) is an SDDE, which is an equation of a very different nature than the equation when $\varepsilon = 0$, which is an ODE. Hence, we are facing a very singular perturbation in which the nature of the problem changes drastically from an ODE—whose phase space is $\mathbb{R}^2$—to an SDDE—whose natural phase space is a space of functions. The precise meaning of the continuation of the unperturbed solutions into solutions of the perturbed problem is somewhat subtle.

**2.1. Limit cycles and isochrons for ODEs.** Under our assumption, there exists a limit cycle (stable periodic orbit) in the unperturbed equation (2.1). In a neighborhood of the limit cycle, points have asymptotic phases (see [Win75, Guc75]). 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 [Win75, Guc75].

According to [HdlL13], we can find a parameterization of the limit cycle and the isochrons in a neighborhood of the limit cycle. More precisely, there exist 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

$$X_0(K(\theta, s)) = DK(\theta, s) \begin{pmatrix} \omega_0 \\ \lambda_0 s \end{pmatrix}, \tag{2.4}$$

where $K$ is periodic in $\theta$, i.e., $K(\theta + 1, s) = K(\theta, s)$. Saying that $K$ solves (2.4) is equivalent to saying that for fixed parameters $\theta$ and $s$, the function $x(t) = K(\theta + \omega_0 t, se^{\lambda_0 t})$ solves (2.1) 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 $\theta$ 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 [HdlL13], the $K$ solving (2.4) can never be unique. If $K(\theta, s)$ is a solution of (2.4), then for any $\theta_0$, $b \neq 0$, $K(\theta + \theta_0, bs)$ will also be a solution of (2.4). [HdlL13] also shows that this is the only source of nonuniqueness.

We will call such a $b$ the scaling factor, and such a $\theta_0$ the 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 paper, for the equation after perturbation (2.2), we will show that if $\varepsilon$ is small enough, the limit cycle and its isochrons for the unperturbed equation persist as the 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 [HVL93]. 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:

$$W(\theta + 1, s) = W(\theta, s) + \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right). \tag{2.5}$$

We remark that $K\circ W$ being a parameterization of the limit cycle and its isochrons is the same as for the given $\theta$, and $s$ in the domain of $W$, $x(t) = K \circ W(\theta + \omega t, se^{\lambda t})$, solving (2.2) for $t \geq 0$.

**2.2. The invariance equation and the prepared invariance equation.** Substitute $x(t) = K \circ W(\theta + \omega t, se^{\lambda t})$ into (2.3), 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 (2.2) if and only if $W$ satisfies

$$DW(\theta, s)\begin{pmatrix}\omega\\ \lambda s\end{pmatrix} = \begin{pmatrix}\omega_0\\ \lambda_0 W_2(\theta, s)\end{pmatrix} + \varepsilon Y(W(\theta, s), \widetilde{W}(\theta, s), \varepsilon), \tag{2.6}$$

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 $W$, as it involves several compositions.

Now we need to look at (2.6) 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 the 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 Remark 2.

*Remark* 2. When $\varepsilon$ 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.

Similarly to the study of center manifolds, we will use cut-off functions to resolve the above issues.

We will transform our original equation (2.6) into a well-defined equation of the same format:

$$DW(\theta,s)\begin{pmatrix}\omega\\ \lambda s\end{pmatrix}=\begin{pmatrix}\omega_0\\ \lambda_0 W_2(\theta,s)\end{pmatrix}+\varepsilon\overline{Y}(W(\theta,s),\widetilde{W}(\theta,s),\varepsilon), \tag{2.7}$$

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 [Car81]), and introduce the following 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 the above extensions, let $\phi:\mathbb{R}\to[0,1]$ be a $C^\infty$ cut-off function:

$$\phi(x)=\begin{cases}1 & \text{if} \quad |x|\le\frac{1}{2},\\ 0 & \text{if} \quad |x|>1.\end{cases} \tag{2.8}$$

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, the main equation (2.6) is turned into the well-defined equation (2.7). 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 $\phi$ vanish for $|x|>c_1$, where the constant $c_1<1$, and let $\phi=1$ for $|x|\le c_2$, 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 $\phi$, we will possibly end up with a different $W$, which solves (2.7) with the new cut-off function $\phi$.

*Remark* 5. If instead of having a stable periodic orbit, the unperturbed ODE has an unstable periodic orbit, then $\lambda_0$ in (2.4) is positive. Analogous results to Theorems 9 and 10 will give us the parameterization of the periodic orbit and the unstable manifold for small $\varepsilon$. The same proof, only with minor modifications, will work. At the same time, since the invariance equation (2.6) 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 (2.2), with $r:\mathbb{R}^2\to[-h,0]$. We omit the details for these cases.

**2.3. Representation of the unknown function.** In order to study the functional equation (2.7), we consider $W$ of the form

$$W(\theta, s) = \sum_{j=0}^{N-1} W^j(\theta) s^j + W^>(\theta, s) \tag{2.9}$$

solving (2.7). 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 paper. 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 (2.5), the lifted functions satisfy the following periodicity conditions:

$$W^0(\theta + 1) = W^0(\theta) + \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), \tag{2.10}$$

$$W^j(\theta + 1) = W^j(\theta), \tag{2.11}$$

$$W^>(\theta + 1, s) = W^>(\theta, s). \tag{2.12}$$

Based on the form of $W$ in (2.9), we can match coefficients of different powers of $s$ in the invariance equation (2.7). Thus, the invariance equation (2.7) 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 and the equation for $W^1$ is a relative eigenvector equation. The equations for the $W^j$'s are all similar. The equation for $W^>$ is hard to study, it has 2 variables. As we will see later, for small enough $\varepsilon$, $W^>$ is the only case where we need the cut-off.

**2.3.1. Invariance equation for the zeroth order term.** Matching zero order terms of $s$ in (2.7), we obtain the equation for the unknowns $\omega$ and $W^0$:

$$\omega \frac{d}{d\theta} W^0(\theta) - \begin{pmatrix} \omega_0 \\ \lambda_0 W_2^0(\theta) \end{pmatrix} = \varepsilon \overline{Y}(W^0(\theta), \widetilde{W}^0(\theta; \omega), \varepsilon), \tag{2.13}$$

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.

**2.3.2. Invariance equation for the first order term.** Equating the coefficients of $s^1$ in (2.7), we obtain the equation for the unknowns $\lambda$ and $W^1$:

$$\omega \frac{d}{d\theta} W^1(\theta) + \lambda W^1(\theta) - \begin{pmatrix} 0 \\ \lambda_0 W_2^1(\theta) \end{pmatrix} = \varepsilon \overline{Y}^1(\theta, \lambda, W^0, W^1, \varepsilon), \tag{2.14}$$

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 (6.20).

**2.3.3. Invariance equation for the $j$th order term.** For $2 \le j \le N - 1$, matching the coefficients of $s^j$, the equation for the unknown $W^j$ is

$$\omega \frac{d}{d\theta} W^j(\theta) + \lambda j W^j(\theta) - \begin{pmatrix} 0 \\ \lambda_0 W_2^j(\theta) \end{pmatrix} = \varepsilon \overline{Y}^j(\theta, \lambda, W^0, W^j, \varepsilon) + R^j(\theta), \tag{2.15}$$

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 (6.32), and $R^j$ is a function of $\theta$ which depends only on $W^0$, $W^1,\dots,W^{j-1}$.

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

**2.3.4. Invariance equation for the higher order term.** We note that $W^>(\theta,s)$ solves the equation:

$$(\omega\partial_\theta + s\lambda\partial_s)W^>(\theta,s) - \begin{pmatrix} 0 \\ \lambda_0 W_2^>(\theta,s)\end{pmatrix} = \varepsilon Y^>(W^>,\theta,s,\varepsilon), \tag{2.16}$$

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 (6.40).

**3. Some basic definitions and basic results on function spaces.** In this section, we collect some standard results on the spaces of continuously differentiable functions that we will use.

For a given positive integer $L$, we will denote by $C^L(Y,X)$ the space of all functions from (an open subset of) a Banach space $Y$ to a Banach space $X$, with uniformly bounded continuous derivatives up to order $L$. We endow $C^L(Y,X)$ with the norm

$$\|f\|_{C^L} = \max_{0\le j\le L}\sup_{\xi\in Y}\|D^j f(\xi)\|_{Y^{\otimes j}\to X},$$

so that $C^L(Y,X)$ is a Banach space.

Note that we include in the definition that the derivatives are uniformly bounded. This is not the same as the Whitney topology on spaces of $L$ times differentiable functions in a $\sigma$-compact manifold [GG73, p. 40], which is a Fréchet topology. Even more general definitions appear in [KM97].

We use $C^L_B(Y,X)$ to denote the closed subset of $C^L(Y,X)$ which consists of functions with $\|\cdot\|_{C^L}$ norm bounded by the constant $B$.

We will also denote $C^{L+Lip}(Y,X)$ as the space of $C^L$ functions with $L$th derivative Lipschitz. We define

$$\mathrm{Lip}(D^L f) = \sup_{\xi_1\ne\xi_2}\frac{\|D^L f(\xi_1) - D^L f(\xi_2)\|_{Y^{\otimes L}\to X}}{\|\xi_1-\xi_2\|_Y},$$

and the norm $\|\cdot\|_{C^{L+Lip}(Y,X)}$ as the maximum of the $\|\cdot\|_{C^L}$ norm and $\mathrm{Lip}(D^L f)$.

Define $C_B^{L+Lip}(Y,X)$ to be the closed subset of the space $C^{L+Lip}(Y,X)$ consisting of all functions with norm $\|\cdot\|_{C^{L+Lip}(Y,X)}$ bounded by the constant $B$.

**3.1. Closure of $C^r$ balls in weak topology.** We quote proposition A2 in [Lan73], as it will be used several times throughout this paper. It can be interpreted as $C_1^{L+Lip}(Y,X)$ is closed under pointwise weak topology on $X$. A related notion, quasi-Banach space, was used in [HT97].

LEMMA 6 (Lanford). *Let $(u_n)_{n\in\mathbb{N}}$ be a sequence of functions on a Banach space $Y$ with values on a Banach space $X$. Assume that for all $n$, $y$,*

$$\|D^j u_n(y)\| \le 1 \quad j = 0,1,2,\dots,k,$$

*and that each $D^k u_n$ is Lipschitz with Lipschitz constant 1. Assume also that for each $y$, the sequence $(u_n(y))$ converges weakly (i.e., in the weak topology of $X$) to $u(y)$. Then,*

(a) *$u$ has a Lipschitz $k$th derivative with Lipschitz constant 1;*

(b) *$D^j u_n(y)$ converges weakly to $D^j u(y)$ for all $y$ and $j = 1,2,\dots,k$.*

Note that the assumption of weak convergence of $(u_n(y))$, and part (b) in the conclusion imply that $\|D^j u(y)\| \leq 1$ for all $y$ and $j = 0, 1, 2, \ldots, k$.

As mentioned in [Lan73], if $X$ and $Y$ are finite dimensional, the above lemma is just an application of the Arzela–Ascoli theorem. This is actually the only case we need. For the proof of the above lemma in the general case, we refer to [Lan73].

**3.2. Faà di Bruno formula.** We also quote the Faà di Bruno formula, which deals with the derivatives of the composition of two functions.

LEMMA 7. *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 $n$th derivative of the composition $h(x) = f[g(x)]$ at $x^0$ is given by the formula*

$$h_n = \sum_{k=1}^{n} f_k \sum_{p(n,k)} n! \prod_{i=1}^{n} \frac{g_i^{\lambda_i}}{(\lambda_i!)(i!)^{\lambda_i}}. \tag{3.1}$$

*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 [AR67] for a proof.

**3.3. Interpolation.** The interpolation inequalities will also be used many times. One can refer to [Had98, Kol49, dlLO99] for some related work. We quote the following result from [dlLO99].

LEMMA 8. *Let $U$ be a convex and bounded open subset of a Banach space $E$ and $F$ be a Banach space. Let $r$, $s$, $t$ be positive numbers, $0 \leq r < s < t$, and $\mu = \frac{t-s}{t-r}$. There is a constant $M_{r,t}$, such that if $f \in C^t(U, F)$, then*

$$\|f\|_{C^s} \leq M_{r,t} \|f\|_{C^r}^{\mu} \|f\|_{C^t}^{1-\mu}.$$

## 4. Main results.

**4.1. 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 the following.

THEOREM 9 (zero order). *For any given integer $L > 0$, there is $\varepsilon_0 > 0$ such that when $0 \leq \varepsilon < \varepsilon_0$, 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* (2.13).

*Moreover, for initial guess $\omega^0$, and $W^{0,0}(\theta)$ satisfying the periodicity condition* (2.10), *if they satisfy the invariance equation* (2.13) *with error $E^0(\theta)$, then there exist unique $\omega$, $W^0(\theta)$(satisfying the periodic condition* (2.10)*) closed by solving the same equation exactly with*

$$\|W^{0,0} - W^0\|_{C^l} \leq C \|E^0\|_{C^0}^{1-\frac{l}{L}}, \quad 0 \leq l < L, \tag{4.1}$$

$$|\omega^0 - \omega| \leq C \|E^0\|_{C^0} \tag{4.2}$$

*for some constant $C$, where $C$ may depend on $\varepsilon$, $\omega_0$, $\lambda_0$, $l$, $L$, and a prior bound for $\|W^{0,0}\|_{L+Lip}$. In fact, $W^0$ has derivatives up to any order.*

Moreover, we have the following.

THEOREM 10 (all orders). *For any given integers $N \geq 2$ and $L \geq 2+N$, there is $\varepsilon_0 > 0$ such that when $0 \leq \varepsilon < \varepsilon_0$, there exist $\omega > 0$, $\lambda < 0$, and $W : \mathbb{T}\times\mathbb{R} \to \mathbb{T}\times\mathbb{R}$ of the form*

$$W(\theta, s) = \sum_{j=0}^{N-1} W^j(\theta)s^j + W^>(\theta, s) \tag{4.3}$$

*which solve (2.7) 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 $1 \leq j \leq N-1$, $W^j : \mathbb{T} \to \mathbb{T}\times\mathbb{R}$ 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 $\theta$ and $s$, with $(L-2-N)$th derivative Lipschitz.*

*Moreover, if $\omega^0$, $W^{0,0}(\theta)$, $\lambda^0$, $W^{1,0}(\theta)$, $W^{j,0}(\theta)$, and $W^{>,0}(\theta, s)$ satisfy the invariance equations (2.13), (2.14), (2.15), and (2.16), with errors $E^0(\theta)$, $E^1(\theta)$, $E^j(\theta)$, and $E^>(\theta, s)$, respectively, then there are $\omega$, $W^0(\theta)$, $\lambda$, $W^1(\theta)$, $W^1(\theta)$, and $W^>(\theta, s)$ which solve (2.13), (2.14), (2.15), and (2.16). Therefore, (2.7) is solved by $\omega$, $\lambda$, and $W(\theta, s)$ of above form (4.3). For $0 \leq l \leq L-2-N$, we have*

$$\begin{aligned} &\|W(\theta, s) - \sum_{j=0}^{N-1} W^{j,0}(\theta)s^j - W^{>,0}(\theta, s)\|_{C^l} \\ &\qquad \leq C\Big(\sum_{j=0}^{N-1} \|E^j\|_{C^0}|s|^j + \|E^>\|_{0,N}|s|^N\Big)^{1-\frac{l}{(L-2-N)}}, \end{aligned} \tag{4.4}$$

$$|\omega - \omega^0| \leq C(\|E^0\|_{C^0}),$$

$$|\lambda - \lambda^0| \leq C(\|E^1\|_{C^0}) \tag{4.5}$$

*for some constant $C$ depending on $\varepsilon$, $\omega_0$, $\lambda_0$, $N$, $l$, $L$, prior bounds for $\|W^{0,0}\|_{L+Lip}$, $\|W^{j,0}\|_{L-1+Lip}$, $j = 1, \ldots, N-1$, and derivatives of $W^{>,0}$.*

*Remark* 11. In Theorem 9, $W^0(\theta)$ is unique up to a phase shift.

*Remark* 12. The above theorems are in a posteriori format. The main input needed are some functions that satisfy the invariance equations approximately. These can be numerical computations (that indeed produce good approximate solutions) or Lindstedt series; see, for example, [CCdlL20].

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 $\varepsilon$ and $Y$.

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

As a matter of fact, one of the advantages of our approach is that it leads very easily to smooth dependence on parameters.

THEOREM 13. *Consider a family of functions $Y_\eta, r_\eta$ as above, where $\eta$ lies in an open interval $I \subset \mathbb{R}$. Assume that $Y_\eta$ and $r_\eta$ are smooth in their inputs as well as in $\eta$, with bounded derivatives.*

*Then for any positive integer $L$, there is a small enough positive $\varepsilon_0$ such that when $\varepsilon < \varepsilon_0$, for each $\eta \in I$, we can find $\omega_\eta$, $W^0_\eta$ solving* (2.13).

*Furthermore, $W^0_\eta(\theta)$ is jointly $C^{L+Lip}$ in $\eta$, $\theta$.*

THEOREM 14. *Under the same assumption as in Theorem* 13, *for any given integers $N \geq 2$, and $L \geq 2 + N$, there is a small enough positive $\varepsilon_0$ such that when $\varepsilon < \varepsilon_0$, for each $\eta \in I$, we can find $\omega_\eta$, $W^0_\eta$, $\lambda_\eta$, $W^j_\eta$, $j = 1, \ldots, N-1$, and $W^>_\eta(\theta, s)$, which solve the invariance equations* (2.13), (2.14), (2.15), *and* (2.16).

*Furthermore, $W^0_\eta(\theta)$ is jointly $C^{L+Lip}$ in $\eta$, $\theta$; $W^j_\eta(\theta)$, $j = 1, \ldots, N-1$, is jointly $C^{L-1+Lip}$ in $\eta$, $\theta$; $W^>_\eta(\theta, s)$ is jointly $C^{L-2-N+Lip}$ in $\eta$, $\theta$, and $s$.*

Note that the regularity conclusions of Theorem 13 can be interpreted in a more functional form as the mapping that to $\eta$ associates $W^0_\eta$ 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 14. This functional point of view is consistent with the point of view of RFDE where the phase space is infinite dimensional.

**4.2. Results for original problem in a neighborhood of the limit cycle.** Note that to find the low order terms, $W^j$ ($j = 1, \ldots, N-1$), for small $\varepsilon$, 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 the zero order term as $W^{0,0}(\theta) = \binom{\theta}{0}$, the error for this initial guess is of order $\varepsilon$. Then by Theorem 9, the true solution $W^0$ is within a distance of order $\varepsilon$ from $W^{0,0}(\theta)$. Therefore, with $\varepsilon$ being small enough, we have $\sup_{\theta\in\mathbb{T}} |W^0_2(\theta)| < \frac{1}{2}$, 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

$$ga) = W^0(\theta - \omega r \circ K(W^0(\theta))).$$

Then we can rewrite the invariance equation for $W^0$, (2.13), as

$$\omega \frac{d}{d\theta} W^0(\theta) - \begin{pmatrix} \omega_0 \\ \lambda_0 W^0_2(\theta) \end{pmatrix} = \varepsilon Y(W^0(\theta), \widetilde{W}^0(\theta;\omega), \varepsilon). \tag{4.6}$$

Similar arguments apply for the equations for $W^1$ and $W^j$'s ($2 \leq j \leq N-1$) if we look at expressions of $\overline{Y}^1$ in (6.20), $\overline{Y}^j$ in (6.32), and the form of $R^j$.

We can find $0 < s_1 < \frac{1}{2}$ such that $W(\mathbb{T} \times [-s_1, s_1) \subset \mathbb{T} \times [-\frac{1}{2}, \frac{1}{2}]$, and $\widetilde{W}(\mathbb{T} \times [-s_1, s_1]) \subset \mathbb{T} \times [-\frac{1}{2}, \frac{1}{2}]$. Therefore, the original problem is solved in a neighborhood of the limit cycle by applying the results in section 4.1.

For the original problem in section 2, we have the following.

COROLLARY 15 (limit cycle). *When $\varepsilon < \varepsilon_0$ in Theorem* 9 *is so small that* $\sup_{\theta\in\mathbb{T}} |W^0_2(\theta)| < \frac{1}{2}$, (2.2) *admits a limit cycle close to the limit cycle of the unperturbed equation. If $\omega$, $W^0$ solve the invariance equation* (4.6), *then $K \circ W^0(\theta)$ gives* *a parameterization of the limit cycle of* (2.2)*, i.e., for any* $\theta$*,* $K \circ W^0(\theta + \omega t)$ *solves* (2.2) *for all* $t$*.*

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

COROLLARY 16 (isochrons). *For small* $\varepsilon$ *as in Corollary* 15*, there are isochrons for the limit cycle of* (2.2)*. If* $\omega$*,* $\lambda$*, and* $W : \mathbb{T} \times \mathbb{R} \to \mathbb{T} \times \mathbb{R}$ *solve the extended invariance equation* (2.7)*, then there exists* $0 < s_0 < \frac{1}{2}$ *such that* $K \circ W(\theta, s)$*,* $|s| \leq s_0$*, gives a parameterization of the limit cycle with its isochrons in a neighborhood, i.e., for any* $\theta$*, and* $s$*, with* $|s| \leq s_0$*,* $K \circ W(\theta + \omega t, se^{\lambda t})$ *solves* (2.2) *for all* $t \geq 0$*.*

One can formulate dependence on parameter results using Theorems 13 and 14. 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.

**4.3. Comparison with results on RFDE based on time evolution.** The persistence of a periodic solution under perturbation for RFDE is presented in Chapter 10 of [HVL93], 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 setup presented there does not seem to apply to our case since the phase space considered in [HVL93] 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 13). 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 (2.2) 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 nondegenerate periodic orbit (see [HVL93]). However, the above $F$ cannot be continuously differentiable in $\phi$ if $\phi$ is only continuous. This obstructs application of Theorem 4.1 for RFDE in [HVL93].

It is very interesting to study whether a similar method to the one in [HVL93] can be extended to our case with some variations of the phase space (solution manifold; see [Wal03]). However, since only $C^1$ regularity of the evolution has been proved [Wal03] (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 paper allows one to get rather straightforwardly higher smoothness with respect to parameters for the special solutions considered here. See Theorem 13. We mention that some progress in continuation of periodic orbits is in [MPN86, MPNP94].

Considering RFDEs as evolutions in infinite-dimensional phase spaces, [HVL93] 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 [Wal03, MNnO17, MPN11]).

Nevertheless, there is a very fundamental difference between the manifolds we consider and those in [HVL93].

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 [Wal03], 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 [HVL93] 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 paper is that it allows one 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 paper, 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 in this paper 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 [GYdlL19].

It is also interesting to investigate whether evolution based methods lead to computational algorithms [Gim19] and compare them with the algorithms based on functional equations as in [GYdlL19].

## 5. Overview of the proof.

In (2.13), $\omega$ and $W^0$ are the unknowns. Under a choice of the phase, we define an operator such that its fixed point solves (2.13). We will show that when $\varepsilon$ is small enough, the operator is a "$C^0$" contraction and maps a $C^{L+Lip}$ ball to itself. Then one obtains the existence of the fixed point $(\omega, W^0)$, and that $W^0$ in the fixed point has some regularity. Therefore, (2.13) is solved.

In (2.14), $\lambda$ 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 similarly to the above case, for small enough $\varepsilon$, this operator has a fixed point $(\lambda, W^1)$ in which $W^1$ has some regularity.

In (2.15), $W^j$ is the only unknown. We define an operator which is a contraction for small enough $\varepsilon$. The operator has a fixed point with certain regularity solving the equation.

In (2.16), $W^>$ 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 $\varepsilon$, this operator has a fixed point in this function space, which solves (2.16).

We emphasize again that for small enough $\varepsilon$, the equation for $W^>$ is the only place where extension is needed (recall section 4.2).

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

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

## 6. Proof of the main results.

**6.1. Zero order solution.** In this section, we prove our first result, Theorem 9.

Recall (2.13), the invariance equation for $\omega$ and $W^0$, which is obtained by setting $s = 0$ in (2.7).

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

$$\omega \frac{d}{d\theta} W_1^0(\theta) - \omega_0 = \varepsilon \overline{Y}_1(W^0(\theta), \widetilde{W}^0(\theta;\omega)\varepsilon) \tag{6.1}$$

and

$$\omega \frac{d}{d\theta} W_2^0(\theta) - \lambda_0 W_2^0(\theta) = \varepsilon \overline{Y}_2(W^0(\theta), \widetilde{W}^0(\theta;\omega), \varepsilon). \tag{6.2}$$

Taking periodicity condition (2.10) into account, we define an operator $\Gamma^0$ as follows:

$$\begin{aligned} \Gamma^0 \begin{pmatrix} a \\ Z_1 \\ Z_2 \end{pmatrix} (\theta) &= \begin{pmatrix} \Gamma_1^0(a, Z) \\ \Gamma_2^0(a, Z)(\theta) \\ \Gamma_3^0(a, Z)(\theta) \end{pmatrix} \\ &= \begin{pmatrix} \omega_0 + \varepsilon \int_0^1 \overline{Y}_1(Z(\theta), \widetilde{Z}(\theta; a), \varepsilon) d\theta \\ \frac{1}{\Gamma_1^0(a,Z)} \left(\omega_0 \theta + \varepsilon \int_0^\theta \overline{Y}_1(Z(\sigma), \widetilde{Z}(\sigma; a), \varepsilon) d\sigma\right) \\ \varepsilon \int_0^\infty e^{\lambda_0 t} \overline{Y}_2(Z(\theta - at), \widetilde{Z}(\theta - at; a), \varepsilon) dt \end{pmatrix}. \end{aligned} \tag{6.3}$$

Notice that if $\Gamma^0$ has a fixed point $(a^*, Z^*)$, then (2.13) is solved by $\omega = a^*$ and $W^0 = Z^*$; at the same time, periodic condition (2.10) is satisfied.

*Remark* 17. As we can see, the operator $\Gamma^0$ depends on $\varepsilon$; however, to simplify the expression, we do not include $\varepsilon$ in the notation of the operator $\Gamma^0$.

*Remark* 18. Similarly to Remark 1, we will not have uniqueness of the solution to invariance equation (2.13). 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 nonuniqueness.

By considering the operator (6.3), we fix a phase by $\Gamma_2^0(a, Z)(0) = 0$.

For the domain of $\Gamma^0$, we consider the closed interval $I^0 = \{a : |a - \omega_0| \le \frac{\omega_0}{2}\}$. 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:

$$\begin{aligned} \mathcal{C}_0^{L+Lip} = \{f \mid f\colon \mathbb{T} \to \mathbb{T}\times\mathbb{R}, & f \text{ can be lifted to a function from } \mathbb{R} \text{ to } \mathbb{R}^2, \\ & \text{still denoted as } f, \text{which satisfies } f(\theta+1) = f(\theta) + \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), \\ & f_1(0) = 0, \|f\|_{L+Lip} \le B^0\}, \end{aligned} \tag{6.4}$$

where

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

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

LEMMA 19. *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 $\varepsilon$ 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 Faà di Bruno's formula in Lemma 7, for $2 \le n \le L$, $\frac{d^n}{d\theta^n}\Gamma_2^0(a, Z)(\theta)$ 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 $\varepsilon$, $\Gamma_2^0(a, Z)(\theta)$ 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 $n \le L$, $\frac{\partial^n}{\partial\theta^n}\overline{Y}_2(Z(\theta - at), \widetilde{Z}(\theta - at; a), \varepsilon)$ 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 $\varepsilon$, $\Gamma_3^0(a, Z)$ 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 $\varepsilon^0$ such that the 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 $D^0$,

$$d((a, Z), (a', Z')) := |a - a'| + \|Z - Z'\|, \tag{6.5}$$

where

$$\|Z - Z'\| = \max\left\{\sup_{\theta} |Z_1(\theta) - Z_1'(\theta)|, \sup_{\theta} |Z_2(\theta) - Z_2'(\theta)|\right\}. \tag{6.6}$$

LEMMA 20. *There exists $\varepsilon^0 > 0$ such that when $\varepsilon < \varepsilon^0$, under the above choice of distance (6.5) on $D^0$, the operator $\Gamma^0$ is a contraction.*

*Proof.* We will show that for $\varepsilon$ small enough (the explicit form of smallness conditions will become clear during the proof), we can find a constant $\mu_0 < 1$ such that for distance defined in (6.5)

$$d(\Gamma^0(a, Z), \Gamma^0(a', Z')) < \mu_0 \cdot d((a, Z), (a', Z')). \tag{6.7}$$

Note that

$$\begin{aligned} d(\Gamma^0(a, Z), \Gamma^0(a', Z')) = &\left|\Gamma_1^0(a, Z) - \Gamma_1^0(a', Z')\right| \\ &+ \|(\Gamma_2^0(a, Z), \Gamma_3^0(a, Z)) - (\Gamma_2^0(a', Z'), \Gamma_3^0(a', Z'))\|. \end{aligned} \tag{6.8}$$

More explicitly, the above distance is

$$\begin{aligned} &\varepsilon\left|\int_0^1 \overline{Y}_1(Z(\theta), \widetilde{Z}(\theta; a), \varepsilon)d\theta - \int_0^1 \overline{Y}_1(Z'(\theta), \widetilde{Z'}(\theta; a'), \varepsilon)d\theta\right| \\ &+ \max\Bigg\{\sup_{\theta}\Bigg|\frac{1}{\Gamma_1^0(a, Z)}\big(\omega_0\theta + \varepsilon\int_0^\theta \overline{Y}_1(Z(\sigma), \widetilde{Z}(\sigma; a), \varepsilon)d\sigma\big) \\ &\qquad - \frac{1}{\Gamma_1^0(a', Z')}\big(\omega_0\theta + \varepsilon\int_0^\theta \overline{Y}_1(Z'(\sigma), \widetilde{Z'}(\sigma; a'), \varepsilon)d\sigma\big)\Bigg|, \\ &\qquad \varepsilon\sup_{\theta}\Bigg|\int_0^\infty e^{\lambda_0 t}\overline{Y}_2(Z(\theta - at), \widetilde{Z}(\theta - at; a), \varepsilon)dt \\ &\qquad\qquad - \int_0^\infty e^{\lambda_0 t}\overline{Y}_2(Z'(\theta - a't), \widetilde{Z'}(\theta - a't; a'), \varepsilon)dt\Bigg|\Bigg\}. \end{aligned} \tag{6.9}$$

Now we consider each term of the above expression (6.9). Note that in the above expression, it suffices to take the supremums for $\theta \in [0, 1]$, which follows from periodicity condition (2.10). By adding and subtracting terms, we have

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

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

$$\begin{aligned} &\left|\overline{Y}_1(Z(\theta), \widetilde{Z}(\theta; a), \varepsilon) - \overline{Y}_1(Z'(\theta), \widetilde{Z'}(\theta; a'), \varepsilon)\right| \\ &\qquad \le 2\|D\overline{Y}_1\|\|Z - Z'\| + \|D\overline{Y}_1\|\|DZ'\|\|\overline{r \circ K}\||a - a'| \\ &\qquad\quad + |D\overline{Y}_1\|\|DZ'\||a'|\|D(\overline{r \circ K})\|\|Z - Z'\| \\ &\qquad \le \|D\overline{Y}_1\|\left(2 + B^0|a'|\|D(\overline{r \circ K})\|\right)\|Z - Z'\| \\ &\qquad\quad + \|D\overline{Y}_1\|B^0\|\overline{r \circ K}\||a - a'|, \end{aligned} \tag{6.10}$$

where the norms are supremum norms on $\mathbb{R}$ or $\mathbb{R}^2$ (defined as above in (6.6)), and

$$\|D\overline{Y}_1\| = \max\{\|D_1\overline{Y}_1\|, \|D_2\overline{Y}_1\|\}, \tag{6.11}$$

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 (6.10),

$$\begin{aligned}\left|\Gamma_1^0(a, Z) - \Gamma_1^0(a', Z')\right| &\le \varepsilon\|D\overline{Y}_1\| \left(2 + B^0|a'|\|D(\overline{r\circ K})\|\right) \|Z - Z'\| \\ &\quad + \varepsilon B^0\|D\overline{Y}_1\|\|\overline{r\circ K}\||a - a'|.\end{aligned} \tag{6.12}$$

Now consider the first component of the maximum for $\theta \in [0, 1]$ in (6.9); by adding and subtracting terms, we have

$$\begin{aligned}&\left|\Gamma_2^0(a, Z) - \Gamma_2^0(a', Z')\right| \\ &\quad\le \frac{\varepsilon}{|\Gamma_1^0(a, Z)|}\int_0^1 \left|\overline{Y}_1(Z(\theta), \widetilde{Z}(\theta), \varepsilon)d\theta - \overline{Y}_1(Z'(\theta), \widetilde{Z'}(\theta), \varepsilon)\right| d\theta \\ &\qquad + \frac{\varepsilon\int_0^1 \left|\overline{Y}_1(Z'(\theta), \widetilde{Z'}(\theta; a'), \varepsilon)\right| d\theta}{|\Gamma_1^0(a, Z)\Gamma_1^0(a', Z')|}\left|\Gamma_1^0(a, Z) - \Gamma_1^0(a', Z')\right| \\ &\qquad + \frac{|\omega_0|}{|\Gamma_1^0(a, Z)\Gamma_1^0(a', Z')|}\left|\Gamma_1^0(a, Z) - \Gamma_1^0(a', Z')\right| \\ &\quad\le \frac{\varepsilon}{|\Gamma_1^0(a, Z)|}\int_0^1 \left|\overline{Y}_1(Z(\theta), \widetilde{Z}(\theta), \varepsilon)d\theta - \overline{Y}_1(Z'(\theta), \widetilde{Z'}(\theta), \varepsilon)\right| d\theta \\ &\qquad + \frac{|\omega_0| + \varepsilon\|\overline{Y}_1\|}{|\Gamma_1^0(a, Z)\Gamma_1^0(a', Z')|}\left|\Gamma_1^0(a, Z) - \Gamma_1^0(a', Z')\right|.\end{aligned} \tag{6.13}$$

By (6.10) and (6.12), with $\Gamma_1^0(a, Z)$, $\Gamma_1^0(a', Z') \in I^0$, we have

$$\begin{aligned}&\left|\Gamma_2^0(a, Z) - \Gamma_2^0(a', Z')\right| \\ &\quad\le \frac{\varepsilon|\omega_0| + \varepsilon^2\|\overline{Y}_1\| + \varepsilon|\Gamma_1^0(a', Z')|}{|\Gamma_1^0(a, Z)\Gamma_1^0(a', Z')|}\Big(\|D\overline{Y}_1\|B^0\|\overline{r\circ K}\||a - a'| \\ &\qquad + \|D\overline{Y}_1\| \left(2 + B^0|a'|\|D(\overline{r\circ K})\|\right) \|Z - Z'\|\Big).\end{aligned} \tag{6.14}$$

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

$$\begin{aligned}&\left|\overline{Y}_2(Z(\theta - at), \widetilde{Z}(\theta - at; a), \varepsilon) - \overline{Y}_2(Z'(\theta - a't), \widetilde{Z'}(\theta - a't; a'), \varepsilon)\right| \\ &\quad\le 2\|D\overline{Y}_2\|\|Z - Z'\| + 2t\|D\overline{Y}_2\|\|DZ'\||a - a'| + \|D\overline{Y}_2\|\|DZ'\|\|r\circ K\||a - a'| \\ &\qquad + \|D\overline{Y}_2\|\|DZ'\||a'|\|D(\overline{r\circ K})\|\|Z - Z'\| \\ &\qquad + t\|D\overline{Y}_2\|\|DZ'\|^2|a'|\|D(\overline{r\circ K})\||a - a'| \\ &\quad\le \|D\overline{Y}_2\|\big(2 + B^0|a'|\|D(\overline{r\circ K})\|\big)\|Z - Z'\| \\ &\qquad + B^0\|D\overline{Y}_2\|\|\overline{r\circ K}\||a - a'| + tB^0\|D\overline{Y}_2\|\big(2 + B^0|a'|\|D(\overline{r\circ K})\|\big)|a - a'|,\end{aligned} \tag{6.15}$$

where $\|D\overline{Y}_2\|$ is defined similarly to (6.11). Then,

$$\begin{aligned}&\left|\Gamma_3^0(a, Z), -\Gamma_3^0(a', Z')\right| \\ &\quad \leq \varepsilon \|D\overline{Y}_2\| B^0 \left( \frac{1}{\lambda_0^2}(2 + B^0|a'|\|D(\overline{r \circ K})\|) - \frac{\|\overline{r \circ K}\|}{\lambda_0} \right) |a - a'| \\ &\qquad - \frac{\varepsilon}{\lambda_0} \|D\overline{Y}_2\| \big(2 + B^0|a'|\|D(\overline{r \circ K})\|\big) \|Z - Z'\|.\end{aligned} \tag{6.16}$$

With the above estimates for each term (6.12), (6.14), and (6.16), we have that for the distance defined in (6.5), $d\big(\Gamma^0(a, Z), \Gamma^0(a', Z')\big)$ is smaller than the sums of the right-hand sides of (6.12), (6.14), and (6.16). More precisely,

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

where

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

and

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

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

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

and

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

Because $c_1$ and $c_2$ are bounded by $\varepsilon$ multiplied by some constants, they can be as small as we want when $\varepsilon$ is small. Therefore, for sufficiently small $\varepsilon$, we can find a constant $\mu_0 < 1$ such that (6.7) is true and $\Gamma^0$ is a contraction. ∎

Take 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 $\Gamma^0$ will have a limit by Lemma 20. Then Lemmas 19 and 6 ensure that the limit is in $D^0$. Therefore, we have a fixed point of $\Gamma^0$ in $D^0$, that is, there exist $\omega > 0$ and $W^0$ in $\mathcal{C}_0^{L+Lip}$ such that (2.13) is solved. Moreover, by the contraction argument, we know that the solution is unique. Therefore, $\omega$ is unique, $W^0$ is unique in the $\mathcal{C}_0^{L+Lip}$ space for the fixed phase $W_1^0(0) = 0$.

Now we prove the a posteriori estimation part of Theorem 9. Since $\Gamma^0$ is a contraction on $D^0$, we know that

$$\begin{aligned} d\big((\omega^0, W^{0,0}), (\omega, W^0)\big) &= \lim_{k\to\infty} d\big((\omega^0, W^{0,0}), (\Gamma^0)^k(\omega^0, W^{0,0})\big) \\ &\le \sum_{k=0}^{\infty} (\mu_0)^k d\big((\omega^0, W^{0,0}), \Gamma^0(\omega^0, W^{0,0})\big) \\ &\le \frac{1}{1-\mu_0} d\big((\omega^0, W^{0,0}), \Gamma^0(\omega^0, W^{0,0})\big). \end{aligned} \tag{6.17}$$

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 (6.6). 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

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

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

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

Combining this with inequality (6.17), we have

$$d\big((\omega^0, W^{0,0}), (\omega, W^0)\big) \le \frac{1}{1-\mu_0} \left[ \left(1 + \frac{2 + 2B^0}{|\omega_0|}\right) \|E_1^0\|_{C_0} - \frac{1}{\lambda_0} \|E_2^0\|_{C_0} \right]. \tag{6.18}$$

By the definition of the norm, (4.2) and the $l = 0$ case of (4.1) are true for a constant $C$, which depends on $\varepsilon$, $B^0$, $\omega_0$, $\lambda_0$.

For other values of $l$, one can use the interpolation inequality in Lemma 8, to get

$$\begin{aligned}\|W_1^{0,0} - W_1^0\|_{C^l} &\le c(l,L)\|W_1^{0,0} - W_1^0\|_{C^0}^{1-\frac{l}{L}} \|W_1^{0,0} - W_1^0\|_{C^L}^{\frac{l}{L}} \\ &\le c(l,L)\|W_1^{0,0} - W_1^0\|_{C^0}^{1-\frac{l}{L}} (2B^0)^{\frac{l}{L}}.\end{aligned} \tag{6.19}$$

Similar estimates can be done for the second component; this finishes the proof of the estimations in Theorem 9.

For the solution of (2.13), note that $K \circ W^0(\theta + \omega t)$ solves the equation (2.2):

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

If $W^0$ is L times differentiable, then the right-hand side of the above equation is L times differentiable, as 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.

**6.2. Proof of Theorem 10.** With Theorem 9, $\omega$ and $W^0$ are known to us. To prove Theorem 10, we have to consider the equations for the first order term, the $j$th order term, and then higher order term in $s$. We will obtain $\lambda$, $W^1$ solving the first order equation (2.14), $W^j$ solving (2.15), and then $W^>$ which solves (2.16).

**6.2.1. First order equation.** Recall that for the first order term, we have an invariance equation (2.14); 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

$$\overline{Y}^1(\theta, \lambda, W^0, W^1, \varepsilon) = A(\theta) W^1(\theta) + B(\theta; \lambda) W^1(\theta - \omega \overline{r \circ K}(W^0(\theta))), \tag{6.20}$$

$$\begin{aligned} A(\theta) = &-\omega D_2 \overline{Y}(W^0(\theta), \widetilde{W}^0(\theta), \varepsilon) DW^0(\theta - \omega \overline{r \circ K}(W^0(\theta))) D(\overline{r \circ K})(W^0(\theta)) \\ &+ D_1 \overline{Y}(W^0(\theta), \widetilde{W}^0(\theta), \varepsilon), \end{aligned} \tag{6.21}$$

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 expressions of $A$ and $B$ above, we suppressed $\omega$ in the expression of $\widetilde{W}^0$. We do this to simplify the notation, since $\omega$ is already known from Theorem 9.

*Remark* 21. Since $\overline{Y}^1(\theta, \lambda, W^0, W^1, \varepsilon)$ in (6.20) is linear in $W^1$, (2.14) for $W^1$ is linear and homogenous in $W^1$. Hence if $W^1(\theta)$ solves (2.14), so does any scalar multiple of $W^1(\theta)$.

Componentwise, we have the following two equations:

$$\omega \frac{d}{d\theta} W_1^1(\theta) + \lambda W_1^1(\theta) = \varepsilon \overline{Y}_1^1(\theta, \lambda, W^0, W^1, \varepsilon), \tag{6.22}$$

$$\omega \frac{d}{d\theta} W_2^1(\theta) + (\lambda - \lambda_0) W_2^1(\theta) = \varepsilon \overline{Y}_2^1(\theta, \lambda, W^0, W^1, \varepsilon). \tag{6.23}$$

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 (2.14) by considering the following normalization:

$$\int_0^1 W_2^1(\theta)d\theta = 1. \tag{6.24}$$

*Remark* 22. It is natural to choose normalization (6.24), since, under small perturbations, we have $W^1(\theta) \approx \binom{0}{1}$. Meanwhile, one can show that $\lambda$ 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 $\Gamma^1$ as follows:

(6.25)
$$\Gamma^1 \begin{pmatrix} b \\ F_1 \\ F_2 \end{pmatrix}(\theta) = \begin{pmatrix} \Gamma_1^1(b,F) \\ \Gamma_2^1(b,F)(\theta) \\ \Gamma_3^1(b,F)(\theta) \end{pmatrix}$$
$$= \begin{pmatrix} \lambda_0 + \varepsilon \int_0^1 \overline{Y}_2^1(\theta, b, F, \varepsilon)d\theta \\ -\varepsilon \int_0^\infty e^{bt}\overline{Y}_1^1(\theta + \omega t, b, F, \varepsilon)dt \\ C(b,F) + \frac{\varepsilon}{\omega}\int_0^\theta \overline{Y}_2^1(\sigma, b, F, \varepsilon) - (\int_0^1 \overline{Y}_2^1(\theta, b, F, \varepsilon)d\theta)F_2(\sigma)d\sigma \end{pmatrix},$$

where

$$\begin{aligned} C(b,F) = 1 &- \frac{\varepsilon}{\omega}\int_0^1\int_0^\theta \overline{Y}_2^1(\sigma, b, F, \varepsilon)d\sigma d\theta \\ &+ \frac{\varepsilon}{\omega}\left(\int_0^1 \overline{Y}_2^1(\theta, b, F, \varepsilon)d\theta\right)\int_0^1\int_0^\theta F_2(\sigma)d\sigma d\theta \end{aligned} \tag{6.26}$$

is a constant chosen to ensure that $\Gamma_3^1(b,F)$ also satisfies the normalization condition (6.24), i.e., $\int_0^1 \Gamma_3^1(b,F)(\theta)d\theta = 1$.

Similarly to the previous section, section 6.1, for the domain of $\Gamma^1$, we consider the closed interval $I^1 = \{b : |b - \lambda_0| \le \frac{|\lambda_0|}{3}\}$, as well as the function space

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

where

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

$L$ is the same as in Theorem 9, and $B^1$ is a positive constant.

Let $D^1 := I^1 \times \mathcal{C}_1^{L-1+Lip}$ be the domain of $\Gamma^1$. We have the following.

LEMMA 23. *If $\varepsilon$ is small enough, $\Gamma^1(D^1) \subset D^1$.*

*Proof.* Since $\overline{Y}_2^1(\theta, b, F, \varepsilon)$ is bounded, for small $\varepsilon$, 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 (2.10), of $F$, and of $\overline{r \circ K}$ with respect to its first component.

Now we check if $\frac{d^n}{d\theta^n}\Gamma_2^1(b, F)(\theta)$, $0 \le n \le L - 1$, 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 the 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 7, 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 $F \in \mathcal{C}_1^{L-1+Lip}$. Then for $\varepsilon$ small enough, the derivatives can be bounded by $B^1$. The 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

$$\frac{d}{d\theta}\Gamma_3^1(b, F)(\theta) = \frac{\varepsilon}{\omega}\overline{Y}_2^1(\theta, b, F, \varepsilon) - \frac{\varepsilon}{\omega}\left(\int_0^1 \overline{Y}_2^1(\theta, b, F, \varepsilon)d\theta\right) F_2(\theta) \tag{6.27}$$

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 (6.27), we have for $2 \le n \le L - 1$

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

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 $F \in \mathcal{C}_1^{L-1+Lip}$. When $\varepsilon$ is small, they will all be bounded by $B^1$; similarly for the Lipschitz constant of $\frac{d^{L-1}}{d\theta^{L-1}}\Gamma_3^1(b, F)(\theta)$.

Hence for $\varepsilon$ small enough, where the smallness condition depends on the 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 $(\Gamma_2^1(b, F), \Gamma_3^1(b, F)) \in \mathcal{C}_1^{L-1+Lip}$. This finishes the proof. ∎

Recall the distance introduced in (6.5):

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

where

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

LEMMA 24. *Under the above definition of distance on $D^1$, for small enough $\varepsilon$, $\Gamma^1$ is a contraction.*

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

$$d(\Gamma^1(b, F), \Gamma^1(b', F')) < \mu_1 \cdot d((b, F), (b', F')). \tag{6.28}$$

Note that

$$
\begin{aligned}
& d(\Gamma^1(b,F),\Gamma^1(b',F')) \\
& \quad \le \varepsilon \left| \int_0^1 \overline{Y}_2^1(\theta,b,F,\varepsilon) - \overline{Y}_2^1(\theta,b',F',\varepsilon) d\theta \right| \\
& \quad + \varepsilon \sup_\theta \left| \int_0^\infty e^{bt}\overline{Y}_1^1(\theta+\omega t,b,F,\varepsilon) - e^{b't}\overline{Y}_1^1(\theta+\omega t,b',F',\varepsilon) dt \right| \\
& \quad + \frac{\varepsilon}{|\omega|} \sup_\theta \left| \int_0^\theta \overline{Y}_2^1(\sigma,b,F,\varepsilon) - \left( \int_0^1 \overline{Y}_2^1(\theta,b,F,\varepsilon) d\theta \right) F_2(\sigma) d\sigma \right. \\
& \qquad\qquad \left. - \int_0^\theta \overline{Y}_2^1(\sigma,b',F',\varepsilon) + \left( \int_0^1 \overline{Y}_2^1(\theta,b',F',\varepsilon) d\theta \right) F_2'(\sigma) d\sigma \right| \\
& \quad + |C(F,b) - C(F',b')|.
\end{aligned} \tag{6.29}
$$

As before, we will consider each term of the right-hand side of the above inequality (6.29).

Recall that $\overline{Y}^1$ has the form (6.20)

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

If we use the 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

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

and

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

Also, if we define $\|A\| = \max_\theta \|A(\theta)\|$, where $\|A(\theta)\|$ is the operator norm corresponding to the maximum norm $\|\cdot\|$ defined in (6.6), then,

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

and, similarly,

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

Note also that

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

similarly,

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

Now for the first term in (6.29), we have

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

For the second term in (6.29), we have for all $\theta$,

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

For the third term in (6.29), we have

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

Similarly holds for the last part in (6.29),

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

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

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

Therefore, for small enough $\varepsilon$, we have that $\Gamma^1$ is a contraction, i.e., we can find a constant $\mu_1$ such that (6.28) is true. □

Taking any initial guess $(\lambda^0,W^{1,0})\in D^1$, we could take $\lambda^0=\lambda_0$ and $W^{1,0}(\theta)=\binom{0}{1}$, then sequence $(\Gamma^1)^n(\lambda^0,W^{1,0})$ has a limit in $D^1$, which we denote by $(\lambda,W^1)$. $(\lambda,W^1)$ is the fixed point of operator $\Gamma^1$, hence it solves (2.14). Since the operator is a contraction, $\lambda$ is unique and $W^1$ is unique in the $C^0$ sense under the normalization condition (6.24).

Similarly to what we have done in estimation (6.17) in section 6.1, notice that

$$d\big((\lambda^0, W^{1,0}), (\lambda, W^1)\big) \le \frac{1}{1-\mu_1} d\big((\lambda^0, W^{1,0}), \Gamma^1(\lambda^0, W^{1,0})\big). \tag{6.30}$$

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,

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

Then

$$d\big((\lambda^0, W^{1,0}), (\lambda, W^1)\big) \le \frac{1}{1-\mu_1} \left[ \frac{3}{2|\lambda_0|} \|E_1^1\| + \left( 1 + \frac{4 + 4B^1}{\omega_0} \right) \|E_2^1\| \right]. \tag{6.31}$$

Therefore, we can find a constant $C$, depending on $\varepsilon$, $B^1$, $\omega_0$, and $\lambda_0$ such that $|\lambda - \lambda^0| \le C\|E^1\|$. This proves (4.5).

**6.2.2. Equation for $j$th order terms.** For each $j \ge 2$, we can proceed in a similar manner to find $W^j$. With $\omega$, $\lambda$, $W^0$, and $W^1$ known, equations for the $W^j$'s are easier to analyze.

*Remark* 25. As we will see, for a 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 interest.

Assume now that we have already obtained $W^0, \dots, W^{j-1}$, and $\omega$, $\lambda$, we are going to find $W^j(\theta)$. To obtain the invariance equation satisfied by $W^j$, mentioned in (2.15), we consider the $j$th order term in (2.7). 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

$$\overline{Y}^j(\theta, \lambda, W^0, W^j, \varepsilon) = A(\theta)W^j(\theta) + B_j(\theta)W^j(\theta - \omega\overline{r \circ K}(W^0(\theta))), \tag{6.32}$$

where $A(\theta)$ is the same as in (6.21):

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

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 by $W^0, \dots, W^{j-1}$. 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 $\lambda$ and $W^0$ are known to us. Componentwisely, $W^j$ should satisfy

$$\omega\frac{d}{d\theta}W_1^j(\theta) + \lambda j W_1^j(\theta) = \varepsilon\overline{Y}_1^j(\theta, W^j, \varepsilon) + R_1^j(\theta), \tag{6.33}$$

$$\omega\frac{d}{d\theta}W_2^j(\theta) + (\lambda j - \lambda_0) W_2^j(\theta) = \varepsilon\overline{Y}_2^j(\theta, W^j, \varepsilon) + R_2^j(\theta). \tag{6.34}$$

Consider functions in the space

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

where $B^j$'s are positive constants, and

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

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

$$\Gamma^j(G)(\theta) = \begin{pmatrix} -\varepsilon\int_0^\infty e^{\lambda j t}\left(\overline{Y}_1^j(\theta+\omega t, G, \varepsilon) + R_1^j(\theta+\omega t)\right)dt \\ -\varepsilon\int_0^\infty e^{(\lambda j - \lambda_0)t}\left(\overline{Y}_2^j(\theta+\omega t, G, \varepsilon) + R_2^j(\theta+\omega t)\right)dt \end{pmatrix}. \tag{6.35}$$

Assuming that we have already obtained $W^k$ in $\mathcal{C}_k^{L-1+Lip}$ for $k = 0, \dots, j-1$, we have the following.

LEMMA 26. *For small enough $\varepsilon$, we have $\Gamma^j(\mathcal{C}_j^{L-1+Lip}) \subset \mathcal{C}_j^{L-1+Lip}$.*

This follows from $\lambda < 0$ and $(\lambda j - \lambda_0) < 0$ for $j \geq 2$ and the regularity of $W^0, \dots, W^j$, $\overline{Y}^j$, and $R^j$. Moreover, we have $\varepsilon$ 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 $\Gamma^j$ is a $C^0$ contraction for small $\varepsilon$.

LEMMA 27. *For small enough $\varepsilon$, $\Gamma^j$ is a contraction in the $C^0$ distance.*

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

If we define the norm as before,

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

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

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

Taking any initial guess $W^{j,0} \in \mathcal{C}_j^{L-1+Lip}$, we would take $W^{j,0}(\theta) = \binom{0}{0}$. The sequence $(\Gamma^j)^n(W^{j,0})$ has a limit in $\mathcal{C}_j^{L-1+Lip}$; we denote it by $W^j$. $W^j$ is the fixed point of operator $\Gamma^j$, so it solves (2.15). $W^j$ 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

$$\|W^j - W^{j,0}\| \leq \frac{1}{1-\mu_j}\|W^{j,0} - \Gamma^j(W^{j,0})\|. \tag{6.36}$$

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

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

Therefore, we have

$$\|W^j - W^{j,0}\| \leq \frac{1}{1-\mu_j}\left(-\frac{1}{j\lambda}\|E_1^j\| - \frac{1}{j\lambda - \lambda_0}\|E_2^j\|\right) \leq C\|E^j\|. \tag{6.37}$$

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

**6.2.3. Equation of the higher order term.** Now we have already found $\omega$, $\lambda$, $W^0, \dots, W^{N-1}$. It remains to consider the higher order term. We will solve equation (2.16) locally in this section, which will establish the existence in Theorem 10. From now on, we will write

$$W(\theta, s) = W^{\leq}(\theta, s) + W^{>}(\theta, s), \tag{6.38}$$

where $W^{\leq}(\theta, s) = \sum_{j=0}^{N-1} W^j(\theta)s^j$. To make the analysis feasible, we do a cut-off to the equation satisfied by $W^>$ in (2.16):

$$(\omega\partial_\theta + s\lambda\partial_s)W^>(\theta, s) = \begin{pmatrix} 0 \\ \lambda_0 W_2^>(\theta, s)\end{pmatrix} + \varepsilon Y^>(W^>, \theta, s, \varepsilon)\phi(s), \tag{6.39}$$

where

$$Y^{>}(W^{>}, \theta, s, \varepsilon) = \overline{Y}(W(\theta, s), \widetilde{W}(\theta, s), \varepsilon) - \sum_{i=0}^{N-1} \overline{Y}^i(\theta) s^i, \tag{6.40}$$

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

and recall the $C^\infty$ cut-off function $\phi : \mathbb{R} \to [0, 1]$ introduced in (2.8):

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

*Remark* 28. A cut-off is needed in our method. We note that similarly 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 (2.7), what happens for $s$ with large absolute value will not matter.

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

$$\Gamma^{>}(H)(\theta, s) = -\varepsilon \int_0^\infty \begin{pmatrix} 1 & 0 \\ 0 & e^{-\lambda_0 t} \end{pmatrix} Y^{>}(H, c(t), \varepsilon) \phi(se^{\lambda t}) dt. \tag{6.41}$$

If there is a fixed point of $\Gamma^{>}$ which has some regularity, it will solve the modified invariance equation (6.39). 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 $l + m \le L^{>}$ with $\|\cdot\|_{L^{>},N}$ norm bounded by a constant $B$:

(6.42)
$$\|H\|_{L^{>},N} := \max_{l+m\le L^{>}, i=1,2} \begin{cases} \sup_{(\theta,s)\in\mathbb{T}\times\mathbb{R}} |\partial_\theta^l \partial_s^m H_i(\theta, s)| |s|^{-(N-m)} & \text{if} \quad m \le N, \\ \sup_{(\theta,s)\in\mathbb{T}\times\mathbb{R}} |\partial_\theta^l \partial_s^m H_i(\theta, s)| & \text{if} \quad m > N. \end{cases}$$

Using the notation introduced in (6.38), we have

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

We define

$$\widetilde{W}^{>}(\theta, s) = W^{>}(\theta - \omega\overline{r \circ K}((W^{\le} + W^{>})(\theta, s)), se^{-\lambda\overline{r\circ K}((W^{\le}+W^{>})(\theta,s))}). \tag{6.43}$$

LEMMA 29. *If $\varepsilon$ is small enough, $\Gamma^{>}(D^{>}) \subset D^{>}$.*

*Proof.* For $H \in D^{>}$, we need to prove that for $i = 1, 2$ and $l + m \le L^{>}$, $\partial_\theta^l \partial_s^m \Gamma_i^{>}(H)(\theta, s)$ exists, and that $\|\Gamma^{>}(H)\|_{L^{>},N}$ is bounded by $B$. Using the definition in (6.43)

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

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

$$\begin{cases} |\partial_\theta^l \partial_s^m \widetilde{H}_i(\theta,s)| \leq C|s|^{(N-m)} & \text{if} \quad m \leq N, \\ |\partial_\theta^l \partial_s^m \widetilde{H}_i(\theta,s)| \leq C & \text{if} \quad m > N. \end{cases} \tag{6.44}$$

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 which appeared in this proof.

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

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

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

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

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

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

Then, we have

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

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

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

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

Similarly,

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

Then,

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

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

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

The $C$ will not depend on the choice of $H$ as long as $\|H\|_{L^>,N} \le B$. 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 $\|Y(H,\theta,s,\varepsilon)\phi(s)\|_{L^>,N} \le C$ when $\|H\|_{L^>,N} \le B$.

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

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

where

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

One can add and subtract terms in the above expression:

(6.45)

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

where we used the notation

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

We group the first two lines, the two lines in the middle, and the last two lines in (6.45), and denote them as $\ell_1$, $\ell_2$, and $\ell_3$, respectively. Then for $\ell_1$:

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

By the regularity of $Y$ and $W^{\le}$, $\|H\|_{L^>,N} \le B$, and that $\widetilde{H}$ satisfies (6.44), we know that $\|\ell_1\phi(s)\|_{L^>,N} \le C$.

Similarly, $\ell_2$ is

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

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

For $\ell_3$, notice that $\sum_{i=0}^{N-1} \overline{Y}^i(\theta)s^i$ is the Taylor expansion at $s = 0$ for

$$\overline{Y}(W^{\leq}(\theta, s), W^{\leq}(\theta - \omega\overline{r \circ K}(W^{\leq}(\theta, s)), se^{-\lambda\overline{r\circ K}(W^{\leq}(\theta,s))}), \varepsilon). \tag{6.46}$$

According to Taylor's formula with remainder (see [LdlL10]), we just need to show that for $m \leq N$

$$\frac{\partial^{N-m}}{\partial s^{N-m}} \frac{\partial^l}{\partial \theta^l} \frac{\partial^m}{\partial s^m}(6.46),$$

and for $m > N$,

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

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

Hence, according to (6.41), if $m \leq N$, for small $\varepsilon$, we have that

$$|\partial_\theta^l \partial_s^m \Gamma_i^>(H)(\theta, s)| \leq \varepsilon \left| \int_0^\infty e^{-\lambda_0 t} C|s|^{N-m} e^{\lambda(N-m)t} e^{\lambda m t} dt \right| \leq B|s|^{N-m}; \tag{6.47}$$

if $m > N$, for small $\varepsilon$, we have that

$$|\partial_\theta^l \partial_s^m \Gamma_i^>(H)(\theta, s)| \leq \varepsilon \left| \int_0^\infty e^{-\lambda_0 t} C e^{\lambda m t} dt \right| \leq B. \tag{6.48}$$

Therefore, for small $\varepsilon$, $\|\Gamma_i^>(H)\|_{L^>,N} \leq B$ when $\|H\|_{L^>,N} \leq B$. ∎

LEMMA 30. *If $\varepsilon$ 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

$$\begin{aligned} &\Gamma^>(H)(\theta, s) - \Gamma^>(H')(\theta, s) \\ &\quad = -\varepsilon \int_0^\infty \begin{pmatrix} 1 & 0 \\ 0 & e^{-\lambda_0 t} \end{pmatrix} \left( Y^>(H, c(t), \varepsilon) - Y^>(H', c(t), \varepsilon) \right) \phi(se^{\lambda t}) dt. \end{aligned} \tag{6.49}$$

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

$$\begin{aligned} &Y^>(H, c(t), \varepsilon) - Y^>(H', c(t), \varepsilon) \\ &\quad = \overline{Y}(W(c(t)), \widetilde{W}(c(t)), \varepsilon) - \overline{Y}(W'(c(t)), \widetilde{W'}(c(t)), \varepsilon). \end{aligned} \tag{6.50}$$

Note that for all $\theta$ and $s$,

$$|W(\theta, s) - W'(\theta, s)| = |H(\theta, s) - H'(\theta, s)| \leq \|H - H'\|_{0,N}|s|^N. \tag{6.51}$$

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

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

where

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

Then,

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

where

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

Now, notice that by the 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 $N \ge 2$. Under this assumption, we have for all $\theta$, $s$,

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

If $\varepsilon$ is small enough, we have for all $\theta$, $s$,

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

Hence for small enough $\varepsilon$,

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

$\Gamma^{>}$ is a contraction. Note that the smallness condition for $\varepsilon$ depends on $N$, $B^j$, $j = 0, \dots, N-1$, $B$, $\omega_0$, $\lambda_0$, $\overline{Y}$, and $\overline{r\circ K}$. ∎

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 6, we know that $W^>$ is $(L^> - 1)$ times differentiable, with the $(L^> - 1)$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 29 implies that $\|E^>\|_{0,N}$ is bounded; therefore, for the maximum norm,

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

and then

$$\|W^> - W^{>,0}\| \leq \frac{1}{1-\mu}\left\|\Gamma^>(W^{>,0}) - W^{>,0}\right\| \leq \frac{1}{(1-\mu)(\lambda_0 - \lambda N)}\|E^>\|_{0,N}|s|^N. \tag{6.52}$$

Combining error estimations in (6.18), (6.31), (6.37), and (6.52), we see that the $l = 0$ case of (4.4) is proved. Inequalities in (4.4) for $l \neq 0$ are obtained using interpolation inequalities.

**6.3. Proof of Theorems 13 and 14.** The proofs of Theorems 13 and 14 are obtained by considering the functions $W_\eta^j(\theta)$ as functions of two variables $\eta$ and $\theta$, denoted as $\tilde{W}^j(\eta, \theta)$. We can straightforwardly lift the operators $\Gamma^0$, $\Gamma^1$, and $\Gamma^j$ defined in (6.3), (6.25), and (6.35) 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 13, given a function $\tilde{W}^0(\eta, \theta)$ of the variables $\eta, \theta$, we treat $\eta$ as a parameter and take into account that now $Y$ and $r$ depend also on $\eta$ in a smooth way.

We use the same strategy as in the proof of Theorem 9. We first show the propagated bounds property, similarly to Lemma 19, and then, show that the operator is a contraction under a $C^0$-type distance, similarly to Lemma 20. The distance here is quite an analogue to the distance defined in (6.5). 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 13 follows by an application of Lemma 6.

In order to get the propagated bounds property, the key is to show that if $\|\tilde{W}\|_{L+\mathrm{Lip}} \leq \tilde{B}^0$ for $\varepsilon < \varepsilon_0$, we have that the $C^{L+Lip}$ 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 Faà di Bruno formula in Lemma 7, 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 the Faà di Bruno formula.

To obtain the proof of the contraction, we just need to observe that the proof of the contraction in Lemma 20 only uses very few properties of $Y$ and $r$. The properties hold uniformly for all $\eta$. 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 6.2.1, 6.2.2, 6.2.3, complete the proof for Theorem 14.

**Acknowledgments.** The authors thank professors A. Humphries and R. Calleja for several suggestions, and professors A. Haro, T. M-Seara, E. I. Verriest, and C. Zeng for discussions. We also thank the anonymous referee for very careful reading and detailed suggestions.

## REFERENCES

[AR67] R. Abraham and J. Robbin, *Transversal Mappings and Flows*, W. A. Benjamin, New York, 1967.

[Car81] J. Carr, *Applications of Centre Manifold Theory*, Appl. Math. Sci. 35, Springer, New York, 1981.

[CCdlL20] A. Casal, L. Corsi, and R. de la Llave, *Expansions in the delay of quasi-periodic solutions for state dependent delay equations*, J. Phys. A, 53 (2020), 235202.

[CF80] A. Casal and M. Freedman, *A Poincaré-Lindstedt approach to bifurcation problems for differential-delay equations*, IEEE Trans. Automat. Control, 25 (1980), pp. 967–973.

[Chi03] C. Chicone, *Inertial and slow manifolds for delay equations with small delays*, J. Differential Equations, 190 (2003), pp. 364–406.

[CHK17] 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., 16 (2017), pp. 1474–1513.

[CJS63] D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, *Relativistic invariance and Hamiltonian theories of interacting particles*, Rev. Modern Phys., 35 (1963), pp. 350–375.

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

[Dri84] R. D. Driver, *A neutral system with state-dependent delay*, J. Differential Equations, 54 (1984), pp. 73–86.

[DvGVLW95] O. Diekmann, S. A. van Gils, S. M. Verduyn Lunel, and H.-O. Walther, *Delay Equations*, Functional, Complex, and Nonlinear Analysis, Appl. Math. Sci. 110, Springer, New York, 1995.

[GG73] M. Golubitsky and V. Guillemin, *Stable Mappings and Their Singularities*, Grad. Texts in Math. 14, Springer, New York, 1973.

[Gim19] J. Gimeno, *Effective Methods for Recurrence Solutions in Delay Differential Equations*, Ph.D. thesis, Universitat de Barcelona, Barcelona, 2019.

[GMJ17] C. M. Groothedde and J. D. Mireles James, *Parameterization method for unstable manifolds of delay differential equations*, J. Comput. Dyn., 4 (2017), pp. 21–70.

[Guc75] J. Guckenheimer, *Isochrons and phaseless sets*, J. Math. Biol., 1 (1975), pp. 259–273.

[GYdlL19] 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. Systems, (2020), https://doi.org/10.1007/s10884-020-09908-6.

[Had98] J. Hadamard, *Sur le module maximum d'une fonction et de ses derives*, Bull. Soc. Math. France, 42 (1898), pp. 68–72.

[Hal77] J. Hale, *Theory of Functional Differential Equations*, 2nd ed., Appl. Math. Sci. 3, Springer, New York, 1977.

[HBC+16] A. R. Humphries, D. A. Bernucci, R. C. Calleja, N. Homayounfar, and M. Snarski, *Periodic solutions of a singularly perturbed delay differential equation with two state-dependent delays*, J. Dynam. Differential Equations, 28 (2016), pp. 1215–1263.

[HdlL13] G. Huguet and R. de la Llave, *Computation of limit cycles and their isochrons: Fast algorithms and their convergence*, SIAM J. Appl. Dyn. Syst., 12 (2013), pp. 1763–1802.

[HDlL16] X. He and R. De la Llave, *Construction of quasi-periodic solutions of state-dependent delay differential equations by the parameterization method* II. *Analytic case*, J. Differential Equations, 261 (2016), pp. 2068–2108.

[HdlL17] 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, 29 (2017), pp. 1503–1517.

[HKWW06] 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, Handb. Differential Equations, Elsevier/North-Holland, Amsterdam, 2006, pp. 435–545.

[HT97] F. Hartung and J. Turi, *On differentiability of solutions with respect to parameters in state-dependent delay equations*, J. Differential Equations, 135 (1997), pp. 192–237.

[HVL93] J. K. Hale and S. M. Verduyn Lunel, *Introduction to Functional-Differential Equations*, Appl. Math. Sci. 99, Springer, New York, 1993.

[KL12] G. Kiss and J.-P. Lessard, *Computational fixed-point theory for differential delay equations with multiple time lags*, J. Differential Equations, 252 (2012), pp. 3093–3115.

[KL17] G. Kiss and J.-P. Lessard, *Rapidly and slowly oscillating periodic solutions of a delayed van der Pol oscillator*, J. Dynam. Differential Equations, 29 (2017), pp. 1233–1257.

[KM97] A. Kriegl and P. W. Michor, *The Convenient Setting of Global Analysis*, Math. Surveys Monogr. 53, American Mathematical Society, Providence, RI, 1997.

[Kol49] A. Kolmogoroff, *On inequalities between the upper bounds of the successive derivatives of an arbitrary function on an infinite interval*, Amer. Math. Soc. Transl., 1949 (1949), 19.

[KY74] J. L. Kaplan and J. A. Yorke, *Ordinary differential equations which yield periodic solutions of differential delay equations*, J. Math. Anal. Appl., 48 (1974), pp. 317–324.

[Lan73] 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, Springer, 1973, pp. 159–192.

[LdlL09] X. Li and R. de la Llave, *Construction of quasi-periodic solutions of delay differential equations via KAM techniques*, J. Differential Equations, 247 (2009), pp. 822–865.

[LdlL10] 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, 3 (2010), pp. 623–641.

[Les10] J.-P. Lessard, *Recent advances about the uniqueness of the slowly oscillating periodic solutions of Wright's equation*, J. Differential Equations, 248 (2010), pp. 992–1016.

[MKW14] F. M. G. Magpantay, N. Kosovalić, and J. Wu, *An age-structured population model with state-dependent delay: Derivation and numerical integration*, SIAM J. Numer. Anal., 52 (2014), pp. 735–756.

[MNnO17] I. Maroto, C. Núñez, and R. Obaya, *Exponential stability for nonautonomous functional differential equations with state-dependent delay*, Discrete Contin. Dyn. Syst. Ser. B, 22 (2017), pp. 3167–3197.

[MPN86] 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 (4), 145 (1986), pp. 33–128.

[MPN11] J. Mallet-Paret and R. D. Nussbaum, *Stability of periodic solutions of state-dependent delay-differential equations*, J. Differential Equations, 250 (2011), pp. 4085–4103.

[MPNP94] 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., 3 (1994), pp. 101–162.

[Sie17] J. Sieber, *Local bifurcations in differential equations with state-dependent delay*, Chaos, 27 (2017), 114326.

[Sij85] J. Sijbrand, *Properties of center manifolds*, Trans. Amer. Math. Soc., 289 (1985), pp. 431–469.

[vdBGL20] J. Bouwe van den Berg, C. Groothedde, and J.-P. Lessard, *A general method for computer-assisted proofs of periodic solutions in delay differential problems*, J. Dynam. Differential Equations, (2020), https://doi.org/10.1007/s10884-020-09908-6.

[Wal03] H.-O. Walther, *The solution manifold and $c^1$-smoothness for differential equations with state-dependent delay*, J. Differential Equations, 195 (2003), pp. 46–65.

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