BOSTON UNIVERSITY
GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

ANALYSIS OF REDUCTION METHODS FOR MULTISCALE PROBLEMS

by

ANTONIOS ZAGARIS
B. Sc., Aristotle University, Thessaloniki, 1998

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
2005

Approved by

First Reader

Tasso J. Kaper, Ph.D.
Professor of Mathematics

Second Reader

Eugene C. Wayne, Ph.D.
Professor of Mathematics

Third Reader

Glenn R. Hall, Ph.D.
Associate Professor of Mathematics

Acknowledgments

Many thanks to Melina, Marina, Rits, Vincent, and Dave for help and support throughout these years. Also, to Stilian Stoev, Oleg Mikitchenko, Harry Olivieris, and Manolis Galenianos for turning this into a proper academic experience; to my advisor, Tasso Kaper, for burdening the task of molding me as a researcher and for rising to the circumstances when needed; to David Rohrlich, Gene Wayne, Dick Hall, and Horacio Rottstein for agreeing to serve as my dissertation defense committee and for all their useful suggestions and stimulating discussions; and to Hans Kaper, for sharing his experience with me in the role of co-advisor. Finally, to Dimitris Hatzis and his Double Book for putting me through this.

ANALYSIS OF REDUCTION METHODS FOR MULTISCALE PROBLEMS

(Order No.)

ANTONIOS ZAGARIS

Boston University Graduate School of Arts and Sciences, 2005

Major Professor: Tasso Kaper, Professor of Mathematics.

ABSTRACT

Many physical problems are modeled by large systems of nonlinear ordinary differential equations (ODE) with multiple time scales. Typically, the fastest modes are exhausted after a brief transient period and become dependent on the slower ones. In terms of the state space dynamics, the fast transient dynamics bring the solutions close to lower-dimensional invariant manifolds where the long-term dynamics play out. Reduction methods aim to identify such manifolds and to determine the reduced system dynamics on them.

This thesis offers analyses for two reduction methods. The first one, the Computational Singular Perturbation (CSP) method, was developed by Lam and Goussis, and it is a method formulated in the tangent bundle of the state space. The second one, the method based on the Zero Derivative Principle (ZDP), was developed by Gear, Kevrekidis, Kaper, and the author, and it is formulated directly in terms of the state space variables. Both methods are analyzed in the context of ODE systems that possess a normally attracting invariant manifold $\mathcal{M}_\varepsilon$ and in which the time scale disparity is measured by an explicitly identified small parameter ε.

Our analysis of CSP shows that successive iterations of the iterative algorithm generate the asymptotic expansion in ε of $\mathcal{M}_\varepsilon$ (through the first step at each iteration) and of the spaces tangent to the fast fibers along which solutions relax to $\mathcal{M}_\varepsilon$ (through the second step at each iteration). Each iteration adds one order of accuracy in ε to these approximations. These results are illustrated in specific examples.

The analysis of ZDP establishes a similar result concerning the approximation to $\mathcal{M}_\varepsilon$. Successive iterations of the iterative algorithm generate its asymptotic expansion, with each iteration adding one order of accuracy in ε. Additionally, we establish the stability properties of this method.

In the final part of this thesis, we develop a unifying framework for reduction methods with geometrical underpinnings. In particular, we explore the duality between dynamics in the state space and in the tangent bundle, and we interpret CSP in terms of the state space variables.

Contents

List of Abbreviations

CSP:	Computational Singular Perturbation
CSPM:	Computational Singular Perturbation Manifold
GSPT:	Geometric Singular Perturbation Theory
ILDM:	Intrinsic Low-Dimensional Manifold
MIM:	Method Of Invariant Manifold
ODE:	Ordinary Differential Equation
PEA:	Partial Equilibrium Approximation
QSSA:	Quasi-Steady State Approximation
ZDP:	Zero-derivative Principle
ZDPM:	Zero-derivative Principle Manifold

Chapter 1

Introduction

1.1 Systems with disparate time scales

Spatially homogeneous physical ensembles may often be modelled by an ordinary differential equation (ODE),

$$x' = g(x), \tag{1.1}$$

where x is the $N-$dimensional vector of variables. Here, $g : \mathbf{R}^N \to \mathbf{R}^N$ is a smooth vector field, and t is time. Very often, these models are of high dimension (large N) — for example, combustion models or models of processes taking place at a microscopic level may involve hundreds or even thousands of variables. This fact alone is enough to make the physics underlying the system impenetrable. It is also often the case that the vector field g be composed of a multitude of terms whose characteristic time scales take on a wide range of values. For example, combustion models involve tens or even hundreds of reactants participating in a multitude of reactions which occur on a spectrum of time scales ranging from microseconds to seconds. This broad disparity of time scales induces stiffness in the (fully nonlinear, in general) ODE system (1.1). This, in combination with the large dimension of the system, makes a numerical investigation of its behavior both arduous and time consuming.

In these and other systems with multiple time scales, it is often the case that the fastest modes are exhausted after a brief transient period and become dependent on the slower ones. In terms of dynamics in the state space, the "fast" transient dynamics bring the orbits close to an invariant, lower-dimensional manifold where the "slow" dynamics take over, and a reduced model can be constructed by constraining the dynamics to this slow manifold. If the dimension of this manifold is much less than that of the state space, then such reduced models are of much smaller dimension and exhibit less disparity in the time scales, so one can achieve significant savings in the study of the long-term behavior of the full system by analyzing the reduced model. In other words, it is up to simplified models that retain the essential characteristics of the full system to facilitate an understanding of its essential physics.

The ideal case where such a reduction can be performed is when the variables fall into two classes, fast and slow, denoted by $x^1 \in \mathbf{R}^{N_s}$ and $x^2 \in \mathbf{R}^{N_f}$, respectively, with $N_s + N_f = N$. In such a case, the Jacobian of the vector field has a spectral gap which, for the sake of the analysis, is usually identified with the inverse of a small parameter $0 < \varepsilon \ll 1$. Then, the characteristic time scales for the slow and fast variables are given by the "slow" time $\tau = \varepsilon t$ and the"fast" time t, respectively. Assuming, without loss of generality, that the entries of x are ordered in such a way that the slow components x^1 come first and the fast components x^2 come second, we conclude that the vector field g has the form

$$g = \begin{pmatrix} \varepsilon g^1 \\ g^2 \end{pmatrix}, \tag{1.2}$$

where $g^1 \in \mathbf{R}^{N_s}$, $g^2 \in \mathbf{R}^{N_f}$. Both g^1 and g^2 may depend on ε, but the entries of these vectors as well as their partial derivatives are all $\mathcal{O}(1)$ as $\varepsilon \downarrow 0$, uniformly in x. The system (1.1) is called a *fast–slow system* of ODEs.

Geometric singular perturbation theory (GSPT) [5, 19] provides a natural framework for the analysis of multiscale problems. If such a system has a compact slow manifold $\mathcal{M}_0$ in the limit as $\varepsilon \downarrow 0$ and this manifold is normally hyperbolic, then GSPT identifies a (usually nonunique) slow manifold $\mathcal{M}_\varepsilon$ for ε sufficiently small. GSPT also gives a complete geometric and analytical description of all solutions near $\mathcal{M}_\varepsilon$, including how trajectories approach $\mathcal{M}_\varepsilon$. A goal of any reduction method is to find $\mathcal{M}_\varepsilon$, if it exists.

1.2 Basis and methodologies of reduction

As mentioned above, reduction methods attempt to exploit the fact that certain modes (*fast modes*) evolve on fast time scales and thus equilibrate after a brief transient period. Once these modes are exhausted, they become enslaved to the slower ones through algebraic relations, and one can determine the state of the system by integrating for the slow modes and obtaining the fast ones from the algebraic relations.

Reduction methods are designed to locate these lower-dimensional manifolds and, hence, to achieve a systematic decrease in the size and complexity of the systems. In the state space, these slow manifolds are parameterized by a subset of all of the variables (commonly known as *reaction progress variables* in the relevant chemical kinetics literature). For a fixed set of values of these variables, the values of the complementary variables are then determined by the corresponding point on the lower-dimensional manifold. As a result, one only needs to track the dynamical evolution of the progress variables on the lower-dimensional manifold; the concentrations of the complementary variables may be found from look-up tables. Moreover, the reaction progress variables may be a predetermined set of state variables or they may be determined in the course of the computation.

Over the years, a large number of reduction methods have been proposed and implemented in computer codes; references can be found in [20, 49, 50]. Generally, they involve the construction of a coordinate system where the slow manifold has a simple functional representation — the ideal system being one where a number of components (the "fast" components) of the state vector are identically zero, as in the Fenichel Normal Form [19]. A partial list of the most commonly used reduction methods is given right below.

In the Quasi-Steady State Approximation (*QSSA*) method [6, 42, 45], chemical intuition is employed to identify reactants whose concentrations equilibrate relatively fast. Then, the right members of the ordinary differential equations for these concentrations are set to zero to yield algebraic equations, and in turn the solution of these equations is an approximation of the slow manifold.

In the Intrinsic Low-Dimensional Manifold (*ILDM*) method [29, 30], the Jacobian of the vector field is partitioned at each point of the phase space into a slow component and its orthogonal complement, and a basis for the slow subspace is generated by means of a Schur decomposition. The ILDM, which approximates the slow manifold, is then defined as the locus of points where the vector field lies entirely in the slow subspace.

The iterative method developed by Fraser [6] and further developed by Roussel and Fraser [40] is derived formally from the invariance equation — a partial differential equation that is satisfied by all invariant manifolds, and thus also by the slow manifold. At each iteration, the terms of the invariance equation involving derivatives are evaluated at the approximation of the slow manifold available from the previous iteration, while all algebraic terms are left evaluated at a general point of the phase space. This reduces the invariance equation to an algebraic equation, which can be solved to yield the new approximation of the slow manifold.

The Method of Invariant Manifold *MIM*) [10, 11, 13, 14, 15] also exploits the invariance equation satisfied by a slow manifold, and it is constructed so that the reduced equations on the approximate slow manifold satisfy a thermodynamic consistency requirement. An initial approximation of the slow manifold (for example, the QSSA or the equilibrium approximation) is used to seed the method. The parametrization of this approximate manifold induces a projection operator onto its tangent space – as well as a complementary projection operator onto the tangent space of the approximate fast dynamics — and the invariance equation is expressed in terms of this complementary projection operator. Then, a Newton iteration is carried out to obtain a more accurate approximation. This cycle is repeated until the desired accuracy is obtained. At the final step, the parametrization of the final approximate slow manifold is chosen so that the reduced equations on it are thermodynamically consistent.

The Computational Singular Perturbation *CSP*) method was first proposed by Lam and Goussis [16, 24, 25, 26, 27] and subsequently developed in [16, 24, 27].

It is widely used, for example, in combustion modeling [17, 28, 31, 32, 46, 48] and atmospheric science [34]. Moreover, it is generally applicable to systems of nonlinear ordinary differential equations with simultaneous fast and slow dynamics where the long-term dynamics evolve on a low-dimensional slow manifold in the phase space. The method is essentially an algorithm to find successive approximations to the slow manifold and match the initial conditions to the dynamics on the slow manifold. In its original formulation, CSP is a two-step iterative technique to generate a coordinate system for the vector field where the system of differential equations reduces to normal form. At the first step of each iteration, one constructs a slow subspace and defines an approximation to the slow manifold as the locus of points where the vector field lies entirely in that slow subspace. Then, at the second step, one constructs a fast subspace which can be used to project initial conditions on the slow manifold.

The Zero-derivative Principle (*ZDP*) is an iterative method based on the bounded-derivative method of Kreiss [23] and is due to Gear and Kevrekidis [7]. Its principle was developed in the context of the equation-free method for multiscale computations [8, 21, 22]. Initially, a subset of the state variables that "contain" the fast directions in the phase space is picked. Then, successive approximations to the slow manifold are obtained by determining the locus of points where time derivatives of successively higher order of these "fast" variables are equal to zero.

1.3 Analysis of reduction methods

Attempts to analyze certain reduction methods for systems with many degrees of freedom in a (more or less) mathematically rigorous setting started only recently. The most successful include those of Fraser and Roussel for the QSSA and the Partial Equilibrium Approximiation (*PEA*) ([40]); of Stiefenhofer ([44]) and of Turanyi, Tomlin, and Pilling for the QSSA ([45]); and the sharp asymptotic estimates on the accuracy of the ILDM method and the iterative method of Fraser and Roussel obtained by Kaper and Kaper ([20]). All of these studies were carried out in the absence of diffusion, *i.e.*, for spatially homogeneous systems. In fact, the modification of existing existing reduction methods so as to include effects of diffusion is an open problem, and, up to this day, attempts to perform such a feat ([17], [31], [32], [43]) have relied on *ad hoc* techniques and considerations.

The purpose of this thesis is threefold: first, to analyze the approximating properties of the CSP method in an asymptotic setting; second, to perform a similar analysis for the ZDP; and third, to place both methods in a broader class of reduction methods with geometric underpinnings and develop tools for the analysis of such methods.

We introduce the CSP method in detail in Chapter 3. We then consider the method in a fast–slow setting (mentioned briefly in Section 1.1 and introduced in detail in Chapter 2) and state, in that same chapter, our two main theorems on the quality of the approximations provided by CSP for the slow manifold and the fast

fibration over it (Theorems 3.3.1 and 3.3.3, respectively). The proof of Theorem 3.3.1 is carried out in Chapter 4 and that of Theorem 3.3.3 in Chapter 5. The validity of these theorems is then exhibited in Chapter 6 by application of the CSP method to the Michaelis–Menten–Henri system of enzyme kinetics.

The ZDP method is introduced in detail in Chapter 7. Theorem 7.2.1 regarding its approximation properties is stated and proved in that same chapter for systems that are not necessarily fast–slow but simply diffeomorphic to some fast–slow system.

Finally we revisit both methods in Chapter 8, and we show that they share many common characteristics. In particular, we show that, although CSP[1] is formulated so as to operate in the tangent bundle, CSP iterations induce an action in the state space. We also show that, conversely, although ZDP is formulated as a method operating in the state space, ZDP iterations induce an action in the tangent bundle. We then proceed to demonstrate that, from a theoretical viewpoint, the two methods differ only in the particulars and can be thought of as sharing the same basis of reduction. In the process of doing so, we develop a streamlining of the CSP method which is of independent interest.

[1]Since ZDP is not designed to address the problem of projecting initial conditions on the slow manifold, our discussion concerns the *one-step CSP* method, a modification of CSP that we introduce in Section 3.2.2 and which only addresses the problem of identifying the slow manifold.

Chapter 2

Geometric Singular Perturbation Theory

2.1 Fast–slow systems

In this section we give an overview of GSPT for fast–slow systems of ODEs. Consider a general system of ordinary differential equations,

$$\frac{dx}{dt} = g(x), \tag{2.1}$$

for a vector-valued function $x \equiv x(t) : \mathbf{R} \to \mathbf{R}^N$ and a smooth vector field g. We refer to $\mathbf{R}^N$ as the state space and think of $x(t)$ as describing the state of a physical system at time t. As mentioned in Chapter 1, for the present analysis we shall assume that N_s components of x evolve on a time scale characterized by the "slow" time $\tau = \varepsilon t$, where $0 < \varepsilon \ll 1$ is a small parameter while the remaining N_f components evolve on a time scale characterized by the "fast" time t, $N_s + N_f = N$. We collect the slow variables in a vector $x^1 \in \mathbf{R}^{N_f}$ and the fast variables in a vector $x^2 \in \mathbf{R}^{N_s}$. Thus, the system (2.1) is equivalent to either the "fast system"

$$\begin{array}{rcll} (x^1)' & = & \varepsilon g^1(x,\varepsilon), & x^1 \in \mathbf{R}^{N_s}, \\ (x^2)' & = & g^2(x,\varepsilon), & x^2 \in \mathbf{R}^{N_s}. \end{array} \tag{2.2}$$

or the "slow system"

$$\begin{array}{rcll} (x^1)' & = & g^1(x,\varepsilon), & x^1 \in \mathbf{R}^{N_s}, \\ \varepsilon(x^2)' & = & g^2(x,\varepsilon), & x^2 \in \mathbf{R}^{N_s}. \end{array} \tag{2.3}$$

(A prime $'$ denotes differentiation with respect to t and a dot $\dot{}$ differentiation with respect to τ.) The fast system is more appropriate for the short-term dynamics and the slow system for the long-term dynamics of the system (2.1).

Remark. This assumption on the divisibility of the chemical species into fast and slow categories, as in Eq. (2.2), is made for convenience. Our analysis of the CSP and the ZDP can also be applied to general chemical systems, where each species may be involved in both fast and slow reactions and for which there is a slow manifold. More details on this issue are given in Chapter 8, where we consider systems that are not necessarily fast–slow but simply diffeomorphic to some fast–slow system.

2.1.1 Slow manifolds

In the limit $\varepsilon \downarrow 0$, the fast system (2.2) reduces formally to a single equation for the fast variable x^2,

$$(x^2)' = g^2(x^1, x^2, 0),$$

where x^1 is a parameter, while the slow system (2.3) reduces to a differential equation for the slow variable x^1,

$$\dot{y} = g^1(x, 0),$$

with the algebraic constraint $g^2(x, 0) = 0$.

We assume that there exists a compact domain $K \subset \mathbf{R}^{N_s}$ and a smooth single-valued function $h_0 : K \to \mathbf{R}^{N_f}$ such that

$$g^2(x^1, h_0(x^1), 0) = 0, \quad x^1 \in K. \tag{2.4}$$

Then the long-time dynamics of the system (2.2) are confined to the *reduced slow manifold* $\mathcal{M}_0$,

$$\mathcal{M}_0 = \left\{ x \in \mathbf{R}^N \,\middle|\, x^2 = h_0(x^1),\ x^1 \in K \right\}. \tag{2.5}$$

We assume, furthermore, that the spectrum of the matrix $(D_{x^2} g^2)(x^1, h_0(x^1), 0)$ lies in the left half-plane, so $\mathcal{M}_0$ is asymptotically stable for the fast system (2.2). Then, the Fenichel theory [5], which applies more generally to normally hyperbolic invariant manifolds, guarantees that $\mathcal{M}_0$ persists as a slow manifold, so for all sufficiently small ε there exists a *slow manifold*, $\mathcal{M}_\varepsilon$, that is invariant under the dynamics of the system (2.2). Moreover, $\mathcal{M}_\varepsilon$ has the same dimension as $\mathcal{M}_0$ and lies near $\mathcal{M}_0$, all nearby solutions relax exponentially fast to $\mathcal{M}_\varepsilon$, and the long-term dynamics of the system (2.2) are governed by an equation on $\mathcal{M}_\varepsilon$.

If ε is positive but arbitrarily small, Fenichel's theory [5, 19] guarantees that there exists a function h_ε whose graph is a locally invariant slow manifold $\mathcal{M}_\varepsilon$. The following theorem is essentially a restatement of [19, Theorem 2].

Theorem 2.1.1 *For all sufficiently small ε, there is a function h_ε such that the graph*

$$\mathcal{M}_\varepsilon = \left\{ x \,\middle|\, z = h_\varepsilon(x^1),\ x^1 \in K \right\}$$

is locally invariant under the dynamics of Eq. (2.2). The function h_ε admits an asymptotic expansion as $\varepsilon \downarrow 0$,

$$h_\varepsilon(\cdot) = h_0(\cdot) + \varepsilon h_1(\cdot) + \varepsilon^2 h_2(\cdot) + \cdots , \tag{2.6}$$

and $h_\varepsilon \in C^r(K)$ for any finite r. The long-term dynamics of the system (2.2) are governed by the equation

$$\dot{x^1} = g^1(x^1, h_\varepsilon(x^2), \varepsilon)$$

on $\mathcal{M}_\varepsilon$, where $\dot{} = d/d\tau$ with $\tau = \varepsilon t$, as long as $x^1 \in K$.

The coefficients $h_1, h_2, \ldots$ appearing in Eq. (2.6) are found from the *invariance equation*,

$$g^2(x^1, h_\varepsilon(x^1), \varepsilon) - \varepsilon (Dh_\varepsilon)(x^1) g^1(x^1, h_\varepsilon(x^1), \varepsilon) = 0, \quad x^1 \in K, \tag{2.7}$$

which we also write as

$$\big(-Dh_\varepsilon(x^1), I_{N_f}\big)\, g(x^1, h_\varepsilon(x^1), \varepsilon) = 0, \tag{2.8}$$

in the following manner. (The invariance equation follows immediately from the chain rule, $(x^2)' = Dh_\varepsilon(x^1)(x^1)'$, and Eq. (2.2).) Each of the functions $g^1(\cdot\,, h_\varepsilon(\cdot), \varepsilon)$ and $g^2(\cdot\,, h_\varepsilon(\cdot), \varepsilon)$ admits a Taylor expansion near $\varepsilon = 0$,

$$g^1(\cdot\,, h_\varepsilon(\cdot), \varepsilon) = \sum_{m=0}^{\infty} \varepsilon^m g^1_m, \quad g^2(\cdot\,, h_\varepsilon(\cdot), \varepsilon) = \sum_{m=0}^{\infty} \varepsilon^m g^2_m, \tag{2.9}$$

with coefficients

$$g^1_m = \sum_{k=0}^{m-1} \sum_{j=1}^{m-k} \frac{1}{k!j!} (D^j_{x^2} D^k_\varepsilon g^1)_0 \sum_{|i|=m-k} (h_{i_1}, ..., h_{i_j}) + \frac{1}{m!} (D^m_\varepsilon g^1)_0, \tag{2.10}$$

$$g^2_m = \sum_{k=0}^{m-1} \sum_{j=1}^{m-k} \frac{1}{k!j!} (D^j_{x^2} D^k_\varepsilon g^2)_0 \sum_{|i|=m-k} (h_{i_1}, ..., h_{i_j}) + \frac{1}{m!} (D^m_\varepsilon g^2)_0, \tag{2.11}$$

see [20]. The notation $(\,\cdot\,)_0$ indicates that the quantity inside the parentheses is evaluated on $\mathcal{M}_0$ — that is, at $(x^1, h_0(x^1), 0)$, for some $x^1 \in K$. Note that $(D^j_{x^2} D^k_\varepsilon g)$ is a multilinear operator, which maps a $j-$form to a vector. The inner sum in Eqs. (2.10)–(2.11) is taken over all multiindices $i = (i_1, \ldots, i_j)$ of j positive integers i_1 through i_j subject to the constraint $|i| = i_1 + \cdots + i_j = m - k$. The expressions (2.10) and (2.11) hold for all m if it is understood that a sum is empty whenever its lower bound exceeds its upper bound. Substituting the expansions (2.10)–(2.11) into the invariance equation, Eq. (2.7), and setting the coefficient of ε^m equal to zero, we obtain an infinite set of equations,

$$g^2_m - \sum_{\ell=0}^{m-1} (Dh_\ell) g^1_{m-1-\ell} = 0, \quad m = 0, 1, \ldots . \tag{2.12}$$

The first few equations are

$$g_0^2 = 0, \tag{2.13}$$
$$(D_{x^2} g^2)_0 h_1 + (D_\varepsilon g^2)_0 - (Dh_0) g_0^1 = 0, \tag{2.14}$$
$$(D_{x^2} g^2)_0 h_2 + \frac{1}{2}(D^2_{x^2} g^2)_0 \,(h_1, h_1) + (D_{x^2} D_\varepsilon g^2)_0 h_1 + \frac{1}{2}(D^2_\varepsilon g^2)_0 - (Dh_1) g_0^1 - (Dh_0)\left((D_{x^2} g^1)_0 h_1 - (D_\varepsilon g^1)_0\right) = 0. \tag{2.15}$$

Equation (2.13) is satisfied identically, Eq. (2.14) yields the coefficient h_1, Eq. (2.15) the coefficient h_2, and so on.

Remark. Typically, the manifold $\mathcal{M}_\varepsilon$ is not unique. Instead, there is a family of slow manifolds, all having the same asymptotic expansion (2.6) to all orders in ε but differing by exponentially small amounts ($\mathcal{O}(e^{-c/\varepsilon})$, $c > 0$).

2.1.2 Fast fibers

Going back to the fast system (2.2) with $\varepsilon = 0$, we see that with each point $p = (x_p^1, h_0(x_p^1)) \in \mathcal{M}_0$ is associated a fast fiber $\mathcal{F}_0^p$,

$$\mathcal{F}_0^p = \left\{ x = (x_p^1, x^2) \in \mathbf{R}^N \;\middle|\; x^2 \in \mathbf{R}^{N_f} \right\}, \quad p \in \mathcal{M}_0.$$

The fiber $\mathcal{F}_0^p$ is invariant under the dynamics associated with the fast system, and the point $(x^1, h_0(x^1)) = \mathcal{M}_0 \cap \mathcal{F}_0$ is a fixed point on it. In fact, this point is asymptotically stable by our assumption on the negativity of the real parts of the eigenvalues of $D_{x^2} g^2(\cdot, h_0(\cdot), 0)$, and thus all solutions on $\mathcal{F}_0^p$ contract exponentially toward p.

If ε is positive but arbitrarily small, Fenichel's theory guarantees that there exists an invariant family $\mathcal{F}_\varepsilon$,

$$\mathcal{F}_\varepsilon = \bigcup_{p \in \mathcal{M}_\varepsilon} \mathcal{F}_\varepsilon^p,$$

of fast stable fibers $\mathcal{F}_\varepsilon^p$ along which solutions relax to $\mathcal{M}_\varepsilon$. The family is invariant in the sense that, if ϕ_t denotes the time-t map associated with Eq. (2.1), then

$$\phi_t(\mathcal{F}_\varepsilon^p) \subset \mathcal{F}_\varepsilon^{\phi_t(p)}, \quad p \in \mathcal{M}_\varepsilon.$$

The collection of fast fibers $\mathcal{F}_\varepsilon^p$ foliates a neighborhood of $\mathcal{M}_\varepsilon$. Hence, the motion of any point on $\mathcal{F}_\varepsilon^p$ decomposes into a fast contracting component along the fiber and a slow component governed by the motion of the base point of the fiber. Also, $\mathcal{F}_\varepsilon^p$ is $\mathcal{O}(\varepsilon)-$close to $\mathcal{F}_0^p$, in the C^0 topology, in any compact neighborhood of $\mathcal{M}_\varepsilon$.

2.2 Coordinates and their dynamics

In this and the next section, we reconsider general systems of nonlinear ODEs, but *without* restricting our attention to fast–slow systems, so that the theory we develop here can be used for more general systems (see Chapters 7–8). To this end, we consider an ODE system

$$z' = q(z), \tag{2.16}$$

which is not, in general, of the fast–slow format (2.2), but is merely *diffeomorphic* to a fast–slow system. That is, we assume that there exists an invertible change of coordinates

$$x = x(z) \tag{2.17}$$

with inverse

$$z = z(x) \tag{2.18}$$

which puts the system (2.16) into the form (2.2). Under the generic assumptions on system (2.2) listed in Section 2.1, the state space $\mathbf{R}^N$ possesses an N_s-dimensional manifold $\mathcal{M}_\varepsilon$ which is slow, invariant, and normally attracting.

2.2.1 Standard coordinates

A useful framework for the study of Eq. (2.1) is the tangent bundle $\mathcal{T}\mathbf{R}^N$ of $\mathbf{R}^N$ — that is, the collection of elements (z, v), where both z and v are arbitrary elements of $\mathbf{R}^N$ [41]. To introduce a frame of reference in $\mathcal{T}\mathbf{R}^N$, we must first select a vector basis in $\mathbf{R}^N$. The standard basis of $\mathbf{R}^N$ consists of the unit vectors $\{e_1, \ldots, e_N\}$, collectively represented by the unit matrix I. In this basis, the vectors z and $q(z)$ are represented by the $N-$tuples of their coordinates,[1]

$$z = \big(z^1, \ldots, z^N\big)^{\mathrm{t}} \in \mathbf{R}^N, \quad q(z) = \big(q^1(z), \ldots, q^N(z)\big)^{\mathrm{t}} \in \mathbf{R}^N.$$

The standard basis of $\mathcal{T}\mathbf{R}^N$ is a combination of the standard basis of $\mathbf{R}^N$ with itself, so the ordered pair $(z, q(z))$ is represented by

$$(z, q(z)) = \big(z^1, \ldots, z^N, q^1(z), \ldots, q^N(z)\big)^{\mathrm{t}} \in \mathcal{T}\mathbf{R}^N.$$

To find the equations governing the evolution of $(z, q(z))$ in $\mathcal{T}\mathbf{R}^N$, we first consider the dynamics of (the coordinates of) z and $q(z)$ separately. The evolution of z in $\mathbf{R}^N$ is governed by Eq. (2.16). Using this equation and the chain rule of differentiation, we obtain the evolution equation for $q(z)$ in $\mathbf{R}^N$,

$$q' = (D_z q)q. \tag{2.19}$$

[1] Our use of z^1 and z^2 (similarly, of q^1 and q^2) to denote the first two components of z (and similarly for q) instead of its first N_s and its last N_f coordinates (as in Eq. (2.2) for x and g) is limited to this and the next section.

Combining these results, we obtain the evolution equation for $(z, q) = (z, q(z))$ in $\mathcal{T}\mathbf{R}^N$,

$$(z, q)' = (q, (D_z q)q). \tag{2.20}$$

This equation tracks both the change of the state variable z with time and the concurrent change of the vector field q along orbits $z(t)$. It provides a unifying framework for the study of reduction methods that have been developed in the context of either the space of vector fields (as is the case for CSP, see Chapter 3) or the space of state variables (as is the case for ZDP, see Chapter 7). Of course, when it comes to the practical implementation, it is sufficient to work in only one context, and one can solve either Eq. (2.16) or Eq. (2.19).

2.2.2 General coordinates

In Eq. (2.20), both the state variable and the vector field are represented in the standard basis of $\mathbf{R}^N$. In this section, we consider a general coordinate system in $\mathbf{R}^N$, which is related to the standard one via a smooth map of $\mathbf{R}^N$ onto itself, and derive the corresponding evolution equations for the state variable and the vector field in the tangent bundle.

Let the state of the system at time t be represented by a new vector, $y(t) = y(z(t))$; inversely, $z(t) = z(y(t))$. The components of y are

$$\left(y^1, \ldots, y^N\right)^{\mathrm{t}} = \left(y^1(z), \ldots, y^N(z)\right)^{\mathrm{t}} \in \mathbf{R}^N. \tag{2.21}$$

If the new coordinates are curvilinear (and they will be in most of the systems one wants to reduce), then they are not associated with any fixed global vector basis that can be used to express $q(z)$ for all z. Rather, at each point $y = y(z)$ of the state space, they are associated with a *local* basis which generally varies from point to point and which consists of the columns of the Jacobian $(D_y z)(y)$. In that basis, we have

$$q = (D_y z)f, \tag{2.22}$$

for some f. Differentiation of both members of $y(z(y)) = y$ with respect to y yields the identity $(D_z y)(D_y z) = I$, which we can use to invert Eq. (2.22),

$$f = (D_z y)q. \tag{2.23}$$

The evolution equation for y follows from Eqs. (2.1), (2.21), and (2.23),

$$y' = f(y). \tag{2.24}$$

This equation is the analog of Eq. (2.1). The analog of Eq. (2.19) is

$$f' = (D_y f)f. \tag{2.25}$$

It will prove useful to express the Jacobian $D_y f$ in terms of the local basis and its dual. Using Eq. (2.23) and the chain rule, we have

$$D_y f = D_z((D_z y)q)(D_y z) = (D_z y)(D_z q)(D_y z) + ((D_z^2 y)q)(D_y z).$$

Differentiation of both members of $(D_z y)(D_y z) = I$ with respect to t along solutions of Eq. (2.1) yields the identity $((D_z^2 y)q)(D_y z) + (D_z y)(D_z(D_y z)q) = 0$, which we use to rewrite the expression for $D_y f$ as

$$D_y f = (D_z y)\left[(D_z q)(D_y z) - (D_z(D_y z))q\right].$$

The expression inside the brackets is the Lie bracket of $D_y z$ (taken column by column) and q. (The Lie bracket o f any two vector fields a and b is $[a, b] = (D_z b)a - (D_z a)b$, see [35].) Therefore, Eq. (2.25) can also be written as

$$f' = \Lambda f, \tag{2.26}$$

where

$$\Lambda = \Lambda(y, f) = D_y f = (D_z y)\ [D_y z, q]. \tag{2.27}$$

Combining Eqs. (2.24) and (2.26), we obtain the evolution equation for $(y, f) = (y, f(y))$ in $\mathcal{T}\mathbf{R}^N$,

$$(y, f)' = (f, \Lambda f).$$

The task now is to construct a suitable basis for $\mathcal{T}\mathbf{R}^N$ where fast and slow variables can be distinguished. We address this issue in the following section.

2.3 Reduction to the slow manifold

An ideal coordinate system is one in which $\mathcal{M}_\varepsilon$ is the locus of points where N_f components of the state vector are identically zero. If the state vector y, defined in Eq. (2.21), corresponds to such an ideal coordinate system, we can partition its components into two groups,

$$y = \begin{pmatrix} y^f \\ y^s \end{pmatrix}, \tag{2.28}$$

with $y^f = (y^1, \ldots, y^{N_f})^{\mathrm{t}}$ and $y^s = (y^{N_f+1}, \ldots, y^N)^{\mathrm{t}}$, and define $\mathcal{M}_\varepsilon$ implicitly by the equation

$$\mathcal{M}_\varepsilon = \left\{ z \in \mathbf{R}^N \;\middle|\; y^f(z) = 0 \right\} \tag{2.29}$$

or in parametric form by the equation

$$\mathcal{M}_\varepsilon = \left\{ z = z(0, y^s) \;\middle|\; y^s \in \mathbf{R}^{N_s} \right\}. \tag{2.30}$$

Equation (2.29), a condition in the state space, can be combined with the invariance of $\mathcal{M}_\varepsilon$ to yield a condition in the tangent bundle. First, we note that the

decomposition (2.28) induces decompositions for the local basis $D_y z$ and the coordinates f,[2]

$$D_y z = (D_f z, D_s z), \quad f = \begin{pmatrix} f^f \\ f^s \end{pmatrix} = \begin{pmatrix} (D_z y^f) q \\ (D_z y^s) q \end{pmatrix}. \tag{2.31}$$

Given these decompositions, we can rewrite Eq. (2.22) as

$$q = (D_f z) f^f + (D_s z) f^s.$$

This equation represents a decomposition of the vector field into a component in the subspace spanned by the columns of $D_f z$ and a component in the subspace spanned by the columns of $D_s z$. Because $\mathcal{M}_\varepsilon$ is invariant, the vector field $q(p)$ at any point $p \in \mathcal{M}$ lies in the tangent space $\mathcal{T}_p\mathcal{M}$. Furthermore, since the variables y^s parameterize $\mathcal{M}_\varepsilon$ (see Eq. (2.30)), $\mathcal{T}_p\mathcal{M} = \text{span}(\text{cols}(D_s z)(p))$. Therefore, $q(p) \in \text{span}(\text{cols}(D_s z)(p))$, and $\mathcal{M}_\varepsilon$ is the locus of points where q has no component along $\text{span}(\text{cols}(D_f z))$,

$$\mathcal{M} = \left\{ z \in \mathbf{R}^N \;\middle|\; f^f(y(z)) = ((D_z y^f) q)(z) = 0 \right\}. \tag{2.32}$$

This equation identifies $\mathcal{M}_\varepsilon$ in terms of the coordinates f of the vector field.

Remark. The identity $(D_z y)(D_y z) = I$ and the decompositions (2.31) imply four identities, which are summarized in the matrix identity

$$\begin{pmatrix} (D_z y^f)(D_f z) & (D_z y^f)(D_s z) \\ (D_z y^s)(D_f z) & (D_z y^s)(D_s z) \end{pmatrix} = \begin{pmatrix} I_{N_f} & 0 \\ 0 & I_{N_s} \end{pmatrix}. \tag{2.33}$$

In an ideal coordinate system, the fast and slow dynamics decouple naturally. Indeed, take any point $p \in \mathcal{M}$ and consider the decomposition

$$\mathcal{T}_p\mathbf{R}^N = \mathcal{T}_p\mathcal{F} \oplus \mathcal{T}_p\mathcal{M} \tag{2.34}$$

of $\mathcal{T}_p\mathbf{R}^N$, the tangent space to $\mathbf{R}^N$ at p, into the direct sum of the tangent space at p to the fast fiber through p, $\mathcal{T}_p\mathcal{F}$, and $\mathcal{T}_p\mathcal{M}$ (which was defined earlier). This decomposition is unique if $\mathcal{M}_\varepsilon$ is normally hyperbolic and compact, because the two linear spaces intersect transversally — that is, they intersect at the single point p and the sum of their dimensions equals the dimension N of $\mathbf{R}^N$. (Note that $\dim\mathcal{T}_p\mathcal{F} = N_f$ and $\dim\mathcal{T}_p\mathcal{M} = N_s$.) The decompositions (2.31) induce a decomposition for the operator Λ,

$$\Lambda = \begin{pmatrix} \Lambda_f^f & \Lambda_s^f \\ \Lambda_f^s & \Lambda_s^s \end{pmatrix}, \tag{2.35}$$

[2] Throughout this section, we use the short-hand notation D_f for D_{y^f} and D_s for D_{y^s}.

where

$$\Lambda_f^f = D_f f^f = (D_z y^f)[D_f z, q], \quad \Lambda_s^f = D_s f^f = (D_z y^f)[D_s z, q],$$
$$\Lambda_f^s = D_f f^s = (D_z y^s)[D_f z, q], \quad \Lambda_s^s = D_s f^s = (D_z y^s)[D_s z, q].$$

For a general coordinate system $y = y(z)$, the off-diagonal blocks Λ_s^f and Λ_f^s of Λ are not zero, so the equations governing the evolution of the coordinates f^1 and f^2 are coupled. Consequently, f^1 and f^2 cannot be identified with the fast and slow coordinates of q globally along trajectories.

The following lemma appeared in [49] and shows that, if

$$\text{span}(\text{cols}(D_f z)(p)) = \mathcal{T}_p\mathcal{F} \quad \text{and} \quad \text{span}(\text{cols}(D_s z)(p)) = \mathcal{T}_p\mathcal{M}, \tag{2.36}$$

then Λ reduces to a block-diagonal form at p,

$$\Lambda(p) \;=\; \begin{pmatrix} \left(D_f f^f\right)(p) & 0 \\ 0 & \left(D_s f^s\right)(p) \end{pmatrix}, \quad p \in \mathcal{M}. \tag{2.37}$$

Thus, a transformation $y = y(z)$ yielding a basis $(D_y z)(y)$ that respects Eq. (2.34) reduces Λ and separates the fast and slow subspaces along trajectories on $\mathcal{M}_\varepsilon$.

Lemma 2.3.1 *([49, Lemma 3.1]) If condition (2.36) is satisfied, then the off-diagonal blocks in the representation (2.35) of Λ are zero at each point $p \in \mathcal{M}_\varepsilon$.*

Proof. We write A_f and A_s for $D_f z$ and $D_s z$, respectively, and also $B^{s\perp}$ and $B^{f\perp}$ for $D_z y^f$ and $D_z y^s$, respectively, to simplify notation. Since $B^{s\perp} A_s = 0$ on $\mathcal{M}_\varepsilon$ by Eq. (2.33), and also since $\mathcal{M}_\varepsilon$ is invariant, we have

$$\frac{d}{dt}\left(B^{s\perp} A_s\right) = D_z(B^{s\perp} A_s)q = (D_z B^{s\perp})(q, A_s) + B^{s\perp}((D_z A_s)q) = 0. \tag{2.38}$$

($D_z B^{s\perp}$ is a symmetric bilinear form; its action on a matrix must be understood as columnwise action.)

Also, $q \in \mathcal{T}\mathcal{M}$, so $B^{s\perp} q = 0$ on $\mathcal{M}_\varepsilon$. Hence, the directional derivative along A_s (taken column by column) at points on $\mathcal{M}_\varepsilon$ also vanishes,

$$D_z(B^{s\perp} q)A_s = (D_z B^{s\perp})(A_s, q) + B^{s\perp}(D_z q)A_s = 0. \tag{2.39}$$

Subtracting Eq. (2.38) from Eq. (2.39), we obtain the identity

$$\Lambda_s^f = B^{s\perp}[A_s, q] = B^{s\perp}\left((D_z q)\, A_s - (D_z A_s)q\right) = 0.$$

The proof for the lower-left block is more involved, since the fast fibers are invariant as a family. Assume that the fiber $\mathcal{F}^p$ at $p \in \mathcal{M}$ is given implicitly by the

equation $F(z;p) = 0$, $z \in \mathcal{F}^p$. Then the rows of $(D_z F)(z;p)$ form a basis for $\mathcal{N}_z\mathcal{F}^p$, so there exists an invertible matrix $C(z)$ such that $B^{f\perp}(z) = (C(z))\,((D_z F)(z;p))$.

Since the rows of $(D_z F)(z;p)$ span $\mathcal{N}_z\mathcal{F}^p$, we have $(D_z F)(z;p)A_f(z) = 0$. This identity holds, in particular, along solutions of Eq. (2.1), so

$$\begin{aligned}\frac{d}{dt}((D_z F)(z;p)A_f(z)) &= \left((D_z^2 F)(z;p)\right)(q(z), A_f(z)) \\ &\quad + ((D_{pz}F)(z;p))\,(q(z), A_f(z)) \\ &\quad + ((D_z F)(z;p))\,(DA_f(z))\,q(z) \\ &= 0. \end{aligned} \tag{2.40}$$

The family of the fast fibers is invariant under the flow associated with Eq. (2.1), so if $F(z;p) = 0$, then also $F(z(t);p(t)) = 0$ and, hence,

$$\frac{dF(z;p)}{dt} = ((D_z F)(z;p))\,q(z) + ((D_p F)(z;p))\,q(p) = 0.$$

Next, we take the directional derivative of both members of this equation along A_f, keeping in mind that $(D_z q)(p)A_f(z) = 0$ because the base point p does not vary along A_f. (Recall that the columns of $A_f(z)$ span $\mathcal{T}_z\mathcal{F}$.) We find that

$$\begin{aligned}&\left((D_z^2 F)(z;p)\right)(A_f(z), q(z)) + ((D_z F)(z;p))\,(D_z q(z))\,A_f(z) \\ &+ ((D_{pz}F)(z;p))\,(A_f(z), q(z)) = 0. \end{aligned} \tag{2.41}$$

But the bilinear forms $D_z^2 F$ and $D_{pz}F$ are symmetric, so subtracting Eq. (2.40) from Eq. (2.41) and letting $z = p$, we obtain the identity

$$[A_f, q](p) = (D_z F)(p;p)\,(((D_z q)(p))\,A_f(p) - ((DA_f)(p))\,q(p))\,(p) = 0.$$

Hence,

$$\Lambda_f^s(p) = \left(B^{f\perp}[A_f, q]\right)(p) = (C(p))\,((D_q F)(p;p))\,([A_f, q](p)) = 0, \tag{2.42}$$

and the proof of the lemma is complete. ∎

In the same coordinate system, the dynamics of the components of y decouple. On $\mathcal{M}_\varepsilon$, the fast variables satisfy the trivial equation, $(y^f)' = f^f = 0$, and the slow variables satisfy the equation

$$(y^s)' = f^s(0, y^s).$$

This is an equation in $\mathbf{R}^{N_s}$. Since $N_s < N$, its dimension is less than the dimension of the original system (2.1), and a reduction has been achieved.

Chapter 3

The CSP Method

3.1 The intuitive idea

The CSP method was initially developed to achieve reduction for ODE systems arising in chemical kinetics. Typically, the vector field g corresponding to such a system is written in a form suggested by the kinetics, namely, as a weighted sum of the stoichiometric vectors with the weights being the associated reaction rates. This representation is in no way unique. In fact,

$$g = \begin{pmatrix} \varepsilon g^1 \\ g^2 \end{pmatrix} = \begin{pmatrix} I_{N_s} & 0 \\ 0 & I_{N_f} \end{pmatrix} \begin{pmatrix} \varepsilon g^1 \\ g^2 \end{pmatrix} \tag{3.1}$$

shows an equivalent representation of g as a weighted sum of the standard basis vectors of $\mathbf{R}^N$, the weights being the coordinates $\varepsilon g^1, \ldots, \varepsilon g^{N_s}, g^{N_s+1}, \ldots, g^N$.[1] The objective of the CSP method is to express g in yet another basis, one that is tuned to the dynamics of the system, where the fast and slow coordinates (*amplitudes*) evolve independently of each other.

The CSP method achieves this objective constructively by successive approximation. Starting with a more or less arbitrary initial basis, one derives the evolution equations for the fast and slow amplitudes and updates the basis iteratively in such a way that the evolution equations for the updated fast and slow amplitudes decouple to increasingly higher order in the small parameter ε. Each iteration consists of two steps. The first step deals with the dependence of the fast amplitudes on the slow amplitudes, the second step with the dependence of the slow amplitudes on the fast amplitudes; details are provided in Section 3.2.

[1] In writing Eq. (3.1), we assumed that the variables were separated into two categories, fast and slow, to allow for an asymptotic analysis and a quantification of the accuracy of the CSP method. We note that the CSP method does not require such an a priori separation of the variables nor that there is an explicitly identified small parameter ε. It only requires that there is a separation of time scales so that some (functions of) species are fast and some are slow.

After each iteration, one identifies the *CSP manifold* (CSPM) as the locus of points where the then-current fast amplitudes vanish and the family of *CSP Fibers* (CSPFs) as the linear space spanned by the fast vectors of the basis generated by CSP to express the vector field. The CSPM is an approximation to the slow manifold $\mathcal{M}_\varepsilon$, while the CSPFs are approximations to the fast fibers $\mathcal{F}_\varepsilon$ along which initial conditions tend toward the slow manifold. The question is: How good are these approximations? In the next three chapters, we analyze the general class of fast-slow systems of ODEs (2.1)–(1.2) and show that the CSP method generates term by term the asymptotic expansion of the slow manifold $\mathcal{M}_\varepsilon$ (Theorem 3.3.1) and of a basis of vectors for the tangent space $\mathcal{TF}_\varepsilon$ to the family of fast fibers $\mathcal{F}_\varepsilon$ (Theorem 3.3.3). After m iterations ($m = 0, 1, 2, \ldots$), the asymptotic expansions of the CSPM and $\mathcal{M}_\varepsilon$ agree up to and including terms of $\mathcal{O}(\varepsilon^m)$; they differ in general at $\mathcal{O}(\varepsilon^{m+1})$. Similarly, the asymptotic expansions of the CSPFs and $\mathcal{TF}_\varepsilon$ agree up to and including terms of $\mathcal{O}(\varepsilon^m)$ and differ in general at $\mathcal{O}(\varepsilon^{m+1})$. The mth application of the CSP algorithm leaves the terms at $\mathcal{O}(1)$ through $\mathcal{O}(\varepsilon^{m-1})$ in these asymptotic expansions invariant. (This observation is important because the lower-order terms have already been determined correctly in the preceding applications.) Exact statements of these theorems are given in Section 3.3 [2].

The CSP method leads not only to an approximation of the slow manifold $\mathcal{M}_\varepsilon$, but also to an approximation of the reduced dynamics on the slow manifold. After m iterations, this approximation is obtained by substituting the fast variables in terms of the slow variables in Eqs. (2.1)–(1.2), using the expression for the CSPM of order m. Thus, one obtains a system of N_s first-order ODEs that depends only on the slow variables. Since N_s is usually much smaller than N_f, the reduced system is much smaller than the full $N-$dimensional system and, hence, computationally much less expensive. Moreover, since the slow manifold $\mathcal{M}_0$ is assumed to be exponentially attracting, it follows from center manifold theory (for example, see [1]) that solutions with initial conditions near the manifold $\mathcal{M}_\varepsilon$ approach solutions of the reduced system exponentially in time. A study of the truncation errors may be carried out based on the results presented here.

Our proof of Theorem 3.3.1 proceeds via an intermediate result for a one-step CSP method. The one-step CSP method is the same as the full two-step CSP method but involves only the first step. Like the full CSP method, it yields a sequence of slow manifolds whose asymptotic behavior as $\varepsilon \downarrow 0$ can be compared with that of the slow manifold $\mathcal{M}_\varepsilon$. We show that, at each application of the one-step CSP algorithm, the dependence of the slow amplitudes on the fast amplitudes is pushed up one order in ε. The result (Theorem 3.3.2) is that m applications of the one-step CSP algorithm yield an approximate slow manifold that agrees asymptotically with $\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^m)$. In other words, the one-step CSP method is as accurate

[2]Similar results have been obtained by Valorani, Goussis, and Najm [47], but only for $m = 1, 2$ and for a model equation due to Davis and Skodje [2].

as the full CSP method; and, to prove the main result for the full CSP method, one needs only to show that the second step does not affect the lower-order terms in the asymptotic expansion of the CSPM.

Summarizing, we conclude that the CSP method provides for the simultaneous approximation of the slow manifold and the tangents to the fast fibers at their base points. If one is interested only in the slow manifold, then it suffices to implement a reduced (one-step) version of the algorithm. On the other hand, if one is interested in both the slow and fast dynamics, then it is necessary to use the full (two-step) CSP algorithm. Moreover, only the full CSP algorithm allows for a linear matching of any initial data with the dynamics on the slow manifold.

It was shown in [20] that the ILDM method yields an approximate slow manifold that is asymptotically accurate up to and including terms of $\mathcal{O}(\varepsilon)$, with an error of $\mathcal{O}(\varepsilon^2)$ proportional to the curvature of $\mathcal{M}_0$. The CSP method, on the other hand, can generate an approximate slow manifold that is asymptotically accurate up to any order. The difference can be traced to two facts, namely, the choice of the fundamental operator governing the dynamics of the system and the retention of the variation of the Jacobian over the manifold $\mathcal{M}_0$. While the ILDM method is designed to transform the Jacobian of the vector field into triangular form (and often also into diagonal form), the CSP method is an algorithm to diagonalize the operator Λ introduced in Sections 2.2 and 2.3 to successively higher orders in ε. The Jacobian is a linear approximation, so the ILDM method never gets beyond a linear approximation. The variation of the Jacobian over $\mathcal{M}_0$ introduces an extra term in Λ. By retaining it, the CSP method preserves the nonlinear character of the operator governing the dynamics of the system. A detailed discussion of the relation between the two methods is given in Section 3.4.

3.2 The iterative algorithm

Writing a single differential equation like (2.1) as a system of equations amounts to choosing a coordinate system in the state space. For example, in Eq. (2.2), the coordinate system is induced by the vector basis consisting of the ordered set of unit vectors in $\mathbf{R}^N$. The coordinates of g relative to this basis are εg^1 and g^2. If we collect the basis vectors in a matrix in the usual way, then we can express the relation between g and its coordinates in the form given in Eq. (3.1). Note that the basis chosen for this representation is the same at every point of the phase space. The CSP method is based on a generalization of this idea, where the basis is allowed to vary from point to point, so it can be tailored to the local dynamics near $\mathcal{M}_\varepsilon$.

Suppose that, instead of a fixed basis for $\mathbf{R}^N$, we choose a (point-dependent) one. Let A be an $N \times N$ matrix, then, whose entries may depend on x and whose columns form a basis for the space $\mathbf{R}^N$ for each x. The relation between the vector field g and

the vector f of its coordinates (*amplitudes*) relative to this basis is

$$g = Af.$$

Conversely,

$$f = Bg, \tag{3.2}$$

where B is the left inverse of A, $BA = I$ on $\mathbf{R}^N$. In the convention of the CSP method, A is a matrix of column vectors (vectors in $\mathbf{R}^N$) and B a matrix of row vectors (functionals on $\mathbf{R}^N$).

The amplitudes can be split into two classes, $f = \begin{pmatrix} f^1 \\ f^2 \end{pmatrix}$, where f^1 is an N_f-vector representing the fast amplitudes and f^2 an N_s-vector representing the slow amplitudes. The splitting suggests that we consider a decomposition of A and a corresponding decomposition of B, namely,

$$A = (A_1, A_2), \quad B = \begin{pmatrix} B^1 \\ B^2 \end{pmatrix}, \tag{3.3}$$

where A_1 is $N \times N_f$, A_2 is $N \times N_s$, B^1 is $N_f \times N$, and B^2 is $N_s \times N$. Thus, $f^1 = B^1 g$ and $f^2 = B^2 g$. Also, the identity $AB = I$ on $\mathbf{R}^N$ implies that $A_1 B^1 + A_2 B^2 = I$, while the identity $BA = I$ on $\mathbf{R}^N$ implies that

$$BA = \begin{pmatrix} B^1 \\ B^2 \end{pmatrix} (A_1, A_2) = \begin{pmatrix} B^1 A_1 & B^1 A_2 \\ B^2 A_1 & B^2 A_2 \end{pmatrix} = \begin{pmatrix} I_{N_f} & 0 \\ 0 & I_{N_s} \end{pmatrix} = I.$$

The fast and slow amplitudes evolve in time. Differentiating Eq. (3.2) along solutions of the system (2.1), we obtain the ordinary differential equation that f satisfies along trajectories of the system (2.1),

$$\frac{df}{dt} = B\frac{dg}{dt} + \frac{dB}{dt}g = B(Dg)g + \frac{dB}{dt}g,$$

where Dg is the Jacobian of g. Hence, f satisfies the nonlinear ODE

$$\frac{df}{dt} = \Lambda f, \tag{3.4}$$

where Λ, the generator of the dynamics for the amplitudes, is given by

$$\Lambda = B(Dg)A + \frac{dB}{dt}A. \tag{3.5}$$

Since $BA = I$ and I is time invariant, A, B, and their time derivatives satisfy the identity

$$(dB/dt)A + B(dA/dt) = 0$$

at all times. Hence, the definition (3.5) is equivalent to

$$\Lambda = B(Dg)A - B\frac{dA}{dt}, \tag{3.6}$$

where $dA/dt = (DA)g$, and thus Λ is the linear operator[3]

$$\Lambda = B(Dg)A + \frac{dB}{dt}A = B(Dg)A - B\frac{dA}{dt} = B[A, g].$$

Here, Dg is the Jacobian of g, $dB/dt = (DB)g$, $dA/dt = (DA)g$, and $[A, g]$ is the Lie bracket of A (considered column by column) and g,

$$[A, g] = ([A_{\cdot,1}, g], \ldots, [A_{\cdot,N}, g]).$$

3.2.1 The full CSP method

It is clear from Eq. (3.4) that the dynamics of f are governed by the operator Λ, and the CSP method focuses on its structure. In general, Λ is not diagonal, and the equations governing the evolution of f^1 and f^2 are coupled. An ideal basis A is one in which Λ is block-diagonalized, so that the ODEs for f^1 and f^2 decouple. The CSP method is a constructive algorithm to approximate such a basis. The algorithm starts from a constant matrix $A^{(0)}$,

$$A^{(0)} = \left(A_1^{(0)},\, A_2^{(0)}\right) = \begin{pmatrix} A_{11}^{(0)} & A_{12}^{(0)} \\ A_{21}^{(0)} & A_{22}^{(0)} \end{pmatrix}.$$

Here, $A_{11}^{(0)}$ is an $m \times n$ matrix, $A_{22}^{(0)}$ an $n \times m$ matrix, and the off-diagonal blocks $A_{12}^{(0)}$ and $A_{21}^{(0)}$ are full-rank square matrices of order m and n, respectively. A common choice is $A_{11}^{(0)} = 0$, so every column vector of $A_1^{(0)}$ lies in the fast subspace. (This choice corresponds to selecting the columns of $A^{(0)}$ from the set of stoichiometric vectors.[4]) We follow this convention and assume, henceforth, that $A_{11}^{(0)} = 0$,

$$A^{(0)} = \left(A_1^{(0)},\, A_2^{(0)}\right) = \begin{pmatrix} 0 & A_{12}^{(0)} \\ A_{21}^{(0)} & A_{22}^{(0)} \end{pmatrix}. \tag{3.7}$$

[3]This is the same operator Λ that we introduced in Sections 2.2 and 2.3, with the difference that the local basis A may be a non-coordinate basis, *i.e.*, it is not necessarily induced by a coordinate system in the state space.

[4]Typically, the number of stoichiometric vectors exceeds N, and they are not linearly independent. If the columns of A are chosen from this set, then one needs to choose a set that forms a basis if possible. In the remaining cases, *i.e.,* when the stoichiometric vectors span only a subspace of $\mathbf{R}^N$, A must be complemented with a basis for the orthogonal complement of this subspace [24].

A more general choice of $A^{(0)}$ is discussed below. The inverse of $A^{(0)}$ is

$$\begin{aligned} B_{(0)} &= (A^{(0)})^{-1} = \begin{pmatrix} B^1_{(0)} \\ B^2_{(0)} \end{pmatrix} = \begin{pmatrix} B^{11}_{(0)} & B^{12}_{(0)} \\ B^{21}_{(0)} & 0 \end{pmatrix} \\ &= \begin{pmatrix} -(A^{(0)}_{21})^{-1} A^{(0)}_{22} (A^{(0)}_{12})^{-1} & (A^{(0)}_{21})^{-1} \\ (A^{(0)}_{12})^{-1} & 0 \end{pmatrix}. \end{aligned} \tag{3.8}$$

The algorithm then proceeds iteratively. For $m = 0, 1, \ldots$, one first defines the operator $\Lambda_{(m)}$ in accordance with Eq. (3.6),

$$\Lambda_{(m)} = B_{(m)}(Dg)A^{(m)} - B_{(m)}\frac{dA^{(m)}}{dt} = \begin{pmatrix} \Lambda^{11}_{(m)} & \Lambda^{12}_{(m)} \\ \Lambda^{21}_{(m)} & \Lambda^{22}_{(m)} \end{pmatrix}, \tag{3.9}$$

and nilpotent matrices $U_{(m)}$ and $L_{(m)}$,

$$U_{(m)} = \begin{pmatrix} 0 & (\Lambda^{11}_{(m)})^{-1}\Lambda^{12}_{(m)} \\ 0 & 0 \end{pmatrix}, \quad L_{(m)} = \begin{pmatrix} 0 & 0 \\ \Lambda^{21}_{(m)}(\Lambda^{11}_{(m)})^{-1} & 0 \end{pmatrix}. \tag{3.10}$$

Then, one updates $A^{(m)}$ and $B_{(m)}$ according to the formulas

$$A^{(m+1)} = A^{(m)}(I - U_{(m)})(I + L_{(m)}), \tag{3.11}$$

$$B_{(m+1)} = (I - L_{(m)})(I + U_{(m)})B_{(m)}, \tag{3.12}$$

and returns to Eq. (3.9) for the next iteration.[5]

At each iteration, one imposes the *CSP condition*,

$$B^1_{(m)}g = 0, \quad m = 0, 1, \ldots, \tag{3.13}$$

to identify those points where the fast reaction rates vanish with respect to the then-current basis. For $m = 0$, $B^1_{(0)}$ is constant and given by Eq. (3.8); for $m = 1, 2, \ldots$, the CSP condition takes the form

$$B^1_{(m)}(x^1, \psi_{(m-1)}(x^1, \varepsilon), \varepsilon)g(x, \varepsilon) = 0, \quad m = 1, 2, \ldots. \tag{3.14}$$

If, for any m, the CSP condition is satisfied by a function $z = \psi_{(m)}(y, \varepsilon)$, then

$$\mathcal{K}^{(m)}_\varepsilon = \left\{ x \,\middle|\, x^2 = \psi_{(m)}(x^1, \varepsilon),\ x^1 \in K \right\}, \quad m = 0, 1, \ldots \tag{3.15}$$

is defined as the *CSP manifold* (CSPM) of order m.

[5]The nonzero entries in $U_{(m)}$ and $L_{(m)}$ have been chosen in such a way that $\Lambda_{(m+1)}$ is closer to block-diagonal form.

Remark. Lam [24] and Lam and Goussis [27] perform the update (3.11)–(3.12) in two steps. The first step corresponds to the postmultiplication of $A^{(m)}$ with $I - U_{(m)}$ and premultiplication of $B_{(m)}$ with $I + U_{(m)}$ and the second step to the subsequent postmultiplication of $A^{(m)}(I - U_{(m)})$ with $I + L_{(m)}$ and premultiplication of $(I + U_{(m)})B_{(m)}$ with $I - L_{(m)}$. The nonzero entries of $U_{(m)}$ and $L_{(m)}$ are chosen so that Λ is block-diagonalized to successively higher orders in ε. In the special case of a linear vector field $g(x)$, the block-diagonalization is commonly referred to as the power method.

Remark. It is useful to see how Λ transforms under a change of basis. If C is an invertible square matrix representing a coordinate transformation in $\mathbf{R}^N$ and $\hat{A} = AC$ and $\hat{B} = C^{-1}B$, then

$$\begin{aligned}
\hat{\Lambda} &= \hat{B}\left((Dg)\hat{A} - D\hat{A}g\right) \\
&= C^{-1}B\left((Dg)AC - DACg\right) \\
&= C^{-1}B\left((Dg)A - (DA)g\right)C - C^{-1}BA(DC)g \\
&= C^{-1}\Lambda C - C^{-1}(DC)g. \qquad (3.16)
\end{aligned}$$

The presence of the term $C^{-1}(DC)g$ in Eq. (3.16) shows that $\hat{\Lambda}$ and Λ are not similar unless C is constant.

3.2.2 The one-step CSP method

The goal of the one-step CSP method is to reduce the matrix Λ to lower block-triangular form — that is, to push the matrix Λ^{12} to increasingly higher order in ε. The method is identical to the full CSP method except for the updating of the matrices A and B. One starts from the same bases, $\tilde{A}^{(0)} = A^{(0)}$ and $\tilde{B}_{(0)} = B_{(0)}$, and, instead of Eqs. (3.11) and (3.12), uses the one-step expressions

$$\tilde{A}^{(m+1)} = \tilde{A}^{(m)}(I - \tilde{U}_{(m)}), \qquad (3.17)$$

$$\tilde{B}_{(m+1)} = (I + \tilde{U}_{(m)})\tilde{B}_{(m)}, \qquad (3.18)$$

where the matrix $\tilde{U}_{(m)}$ is defined as in Eq. (3.10) with Λ replaced by $\tilde{\Lambda}$. (A tilde $\tilde{}$ distinguishes a quantity from its counterpart in the full CSP method.)

The update rule for $\tilde{\Lambda}$ follows immediately from Eq. (3.16),

$$\tilde{\Lambda}_{(m+1)} = (I + \tilde{U}_{(m)})\tilde{\Lambda}_{(m)}(I - \tilde{U}_{(m)}) + (I + \tilde{U}_{(m)})\frac{d\tilde{U}_{(m)}}{dt}. \qquad (3.19)$$

(Note that the identities $\tilde{A}^{(0)} = A^{(0)}$ and $\tilde{B}_{(0)} = B_{(0)}$ imply that $\tilde{\Lambda}_{(0)} = \Lambda_{(0)}$.) The matrix $\tilde{U}_{(m)}$ and its time derivative have the same block structure; only the upper

right block is nonzero, so $\tilde{U}_{(m)}d\tilde{U}_{(m)}/dt = 0$, and Eq. (3.19) reduces to

$$\tilde{\Lambda}_{(m+1)} = (I + \tilde{U}_{(m)})\tilde{\Lambda}_{(m)}(I - \tilde{U}_{(m)}) + \frac{d\tilde{U}_{(m)}}{dt}.$$

In terms of the constituent blocks, we have

$$\tilde{\Lambda}^{11}_{(m+1)} = \tilde{\Lambda}^{11}_{(m)} + \tilde{U}_{(m)}\tilde{\Lambda}^{21}_{(m)}, \tag{3.20}$$

$$\tilde{\Lambda}^{12}_{(m+1)} = \tilde{U}_{(m)}\tilde{\Lambda}^{22}_{(m)} - \tilde{U}_{(m)}\tilde{\Lambda}^{21}_{(m)}\tilde{U}_{(m)} + \frac{d\tilde{U}_{(m)}}{dt}, \tag{3.21}$$

$$\tilde{\Lambda}^{21}_{(m+1)} = \tilde{\Lambda}^{21}_{(m)}, \tag{3.22}$$

$$\tilde{\Lambda}^{22}_{(m+1)} = \tilde{\Lambda}^{22}_{(m)} - \tilde{\Lambda}^{21}_{(m)}\tilde{U}_{(m)}, \tag{3.23}$$

where we have used Eq. (3.10) to simplify Eq. (3.21). Note that we freely use $\tilde{U}_{(m)}$ to denote both the full update matrix and its restriction to the subspace $\mathbf{R}^{N_s}$; the latter is represented by the matrix $(\tilde{\Lambda}^{11}_{(m)})^{-1}\tilde{\Lambda}^{12}_{(m)}$. The appropriate interpretation is clear from the context.

The one-step CSP method generates a sequence of manifolds,

$$\tilde{\mathcal{K}}^{(m)}_{\varepsilon} = \left\{ x \;\middle|\; x^2 = \tilde{\psi}_{(m)}(x^1, \varepsilon),\ x^1 \in K \right\}, \quad m = 0, 1, \ldots \tag{3.24}$$

just like the full CSP method; cf. Eq. (3.15). The functions $\tilde{\psi}_{(m)}$ are defined by the conditions

$$\tilde{B}^1_{(m)} g = 0, \quad m = 0, 1, \ldots, \tag{3.25}$$

where $\tilde{B}^1_{(m)}$ is obtained from Eq. (3.18).

3.3 Approximating properties of the CSP method

This section contains our main results concerning the approximating properties of the CSP method. The proofs of these results are given in the next two chapters.

3.3.1 Approximation of the slow manifold

After m iterations, the CSP condition (3.13),

$$B^1_{(m)} g = 0, \quad m = 0, 1, \ldots,$$

identifies those points where the fast amplitudes vanish with respect to the then current basis. These points define a manifold that is an approximation for the slow manifold $\mathcal{M}_{\varepsilon}$.

For $m = 0$, $B^1_{(0)}$ is constant and given by Eq. (3.8). Hence, the CSP condition (3.13) reduces to the constraint $g^2(x, \varepsilon) = 0$. In general, this constraint is satisfied by a function $x^2 = \psi_{(0)}(x^1, \varepsilon)$. The graph of this function defines $\mathcal{K}^{(0)}_\varepsilon$, the CSP manifold (CSPM) of order zero. Since the constraint reduces at leading order to the equation $g^2(x, 0) = 0$, which is satisfied by the function $x^2 = h_0(x^1)$, $\mathcal{K}^{(0)}_\varepsilon$ may be chosen to coincide with $\mathcal{M}_0$ to leading order; see Eq. (2.5).

For $m = 1, 2, \ldots$, the CSP condition takes the form of Eq. (3.14),

$$B^1_{(m)}(x^1, \psi_{(m-1)}(x^1, \varepsilon), \varepsilon) g(x, \varepsilon) = 0, \quad m = 1, 2, \ldots .$$

The condition is satisfied by a function $x^2 = \psi_{(m)}(x^1, \varepsilon)$, and the manifold

$$\mathcal{K}^{(m)}_\varepsilon = \left\{ x \,\middle|\, x^2 = \psi_{(m)}(x^1, \varepsilon),\ x^1 \in K \right\}, \quad m = 0, 1, \ldots$$

defines the CSPM of order m (recall Eq. (3.15)), which is an approximation of $\mathcal{M}_\varepsilon$. The following theorem regarding the quality of the approximation originally appeared in [49].

Theorem 3.3.1 ([49, Theorem 3.1]) *The asymptotic expansions of the CSPM of order m, $\mathcal{K}^{(m)}_\varepsilon$, and the slow manifold $\mathcal{M}_\varepsilon$ agree up to and including terms of $\mathcal{O}(\varepsilon^m)$,*

$$\psi_{(m)}(\cdot\,, \varepsilon) = \sum_{j=0}^{m} \varepsilon^j h_j(\cdot) + \mathcal{O}(\varepsilon^{m+1}), \quad \varepsilon \downarrow 0, \quad m = 0, 1, \ldots .$$

This theorem shows that each application of the CSP algorithm improves the asymptotic accuracy of the CSPM by one order of ε, so that, after m iterations, $\mathcal{K}^{(m)}_\varepsilon$ is an $\mathcal{O}(\varepsilon^m)-$approximation to the slow manifold $\mathcal{M}_\varepsilon$. Its proof is reproduced in Chapter 4.

An analogous result is proven in [49] for the one-step CSP method.

Theorem 3.3.2 ([49, Theorem 4.1]) *The manifolds $\tilde{\mathcal{K}}^{(m)}_\varepsilon$ and $\mathcal{M}_\varepsilon$ agree asymptotically up to and including terms of $\mathcal{O}(\varepsilon^m)$ for $m = 0, 1, \ldots$,*

$$\tilde{\psi}_{(m)}(\cdot\,, \varepsilon) = \sum_{j=0}^{m} \varepsilon^j h_j(\cdot) + \mathcal{O}(\varepsilon^{m+1}), \quad \varepsilon \downarrow 0, \quad m = 0, 1, \ldots .$$

This theorem shows that the second step at each iteration of the CSP algorithm may be omitted if one is interested only in approximating the slow manifold. Its proof is reproduced in Chapter 4.

Remark. Theorems 3.3.1 and 3.3.2 extend readily to the case where the eigenvectors of the Jacobian Dg are used, instead of the stoichiometric vectors, to form the initial basis $A^{(0)}$. In that case, the slow subspace of the leading-order Jacobian coincides with the tangent space $\mathcal{T}_p\mathcal{M}_0$ at any point $p \in \mathcal{M}_0$, so the columns of $A_2^{(0)}$ are tangent to $\mathcal{M}_0$ to leading order. In turn, this implies that the rows of $B^1_{(0)}(p)$ span the orthogonal complement of the tangent space, also to leading order. As a result, the initial CSPM, the solution of $B^1_{(0)}g = 0$, coincides with $\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon)$, which is one order higher than is the case when $A^{(0)}$ is given by Eq. (3.7). Moreover, for each $m = 1, 2, \ldots$, the proof of Theorem 3.3.2 generalizes directly to this case. The asymptotic expansion of $\tilde{\psi}_{(m)}$ coincides with that of h_ε up to and including terms of $\mathcal{O}(\varepsilon^{m+1})$, which is one order higher than is the case when $A^{(0)}$ is given by Eq. (3.7).

3.3.2 Approximation of the fast fibers

We now turn our attention to the fast fibers. The columns of $A_f(x^1, h_\varepsilon(x^1))$ span the tangent space to the fast fiber with base point $p = (x^1, h_\varepsilon(x^1))$, so we expect that $A_1^{(m)}$ defines an approximation for the same space after m applications of the CSP algorithm. We denote this approximation by $\mathcal{L}_\varepsilon^{(m)}(x^1)$ and refer to it as the CSP fiber (CSPF) of order m at p,

$$\mathcal{L}_\varepsilon^{(m)}(x^1) = \text{span}\left(\text{cols}\left(A_1^{(m)}\left(x^1, \psi_{(m)}(x^1, \varepsilon), \varepsilon\right)\right)\right). \tag{3.26}$$

The following theorem appeared in [50] and establishes the quality of the approximation of the CSPFs to the family of fast fibers.

Theorem 3.3.3 ([50, Theorem 3.2]) *The asymptotic expansions of $\mathcal{L}_\varepsilon^{(m)}(x^1)$ and $\mathcal{T}_p\mathcal{F}_\varepsilon$, where $p = (x^1, h_\varepsilon(x^1)) \in \mathcal{M}_\varepsilon$, agree up to and including terms of $\mathcal{O}(\varepsilon^m)$ for all $x^1 \in K$ and for $m = 0, 1, \ldots$.*

This theorem implies that the family of the CSPFs $\mathcal{L}_\varepsilon^{(m)} \equiv \bigcup_{p\in\mathcal{M}_\varepsilon}(p, \mathcal{L}_\varepsilon^{(m)}(x^1))$ is an $\mathcal{O}(\varepsilon^m)-$approximation to the tangent bundle $\mathcal{TF}_\varepsilon$. Its proof is reproduced in Chapter 5.

Remark. Each application of the CSP algorithm involves two steps. The first step involves U and serves to push the upper-right block of Λ up by one order of ε, and the second step involves L and serves the same purpose for the lower-left block. The two steps are consecutive. At the first step of the mth iteration, one evaluates $B^1_{(m)}$ on $\mathcal{K}_\varepsilon^{(m-1)}$ to find $\mathcal{K}_\varepsilon^{(m)}$ by solving the CSP condition (3.13) for the function $\psi_{(m)}$. One then uses this expression in the second step to update A and B, thus effectively evaluating $A_1^{(m)}$ on $\mathcal{K}_\varepsilon^{(m)}$ rather than on $\mathcal{K}_\varepsilon^{(m-1)}$. If, though, in the second step of the CSP algorithm, $A_1^{(m)}$ were evaluated on $\mathcal{K}_\varepsilon^{(m-1)}$ instead of on $\mathcal{K}_\varepsilon^{(m)}$, the approximation

of $\mathcal{TF}_\varepsilon$ would still be $\mathcal{O}(\varepsilon^m)-$accurate. In section 6.1 we demonstrate this by means of an example; in Appendix B, we give the proof for the general case.

Remark. As was the case with Theorem 3.3.1, the statement of Theorem 3.3.3 as given above is tailored to suit our choice of an initial basis $A^{(0)}$; see Eq. (3.7). Other choices have also been considered for systems in which there is not an explicit small parameter. Hadjinicolaou and Goussis [17] use an arbitrary basis, while Lam and Goussis [26] and Massias and Goussis [32] use the eigenvectors of the Jacobian. These other choices of $A^{(0)}$ introduce modifications only in the degree to which the CSPM and the CSPFs of order m approximate $\mathcal{M}_\varepsilon$ and $\mathcal{TF}_\varepsilon$, respectively, and, for any choice of $A^{(0)}$, statements similar to those of Theorems 3.3.1 and 3.3.3 can be shown to be true.

3.4 Relation between CSPM and ILDM

As noted in Section 3.1, the fundamental difference between the CSP method and the ILDM method can be traced to (1) the choice of the fundamental operator governing the dynamics of the system, and (2) the retention of the variation of the Jacobian over the manifold $\mathcal{M}_0$. The CSP iteration procedure is designed to diagonalize the Lie bracket $[\,\cdot\,, g]$. At each iteration, the then-current basis is updated in such a way that $[\,\cdot\,, g]$ is block-diagonalized to the next-higher order in ε. Thus, each iteration improves the quality of the basis of the orthogonal complement of the tangent space. The CSPM is defined as the locus of points where the vector field is orthogonal to that orthogonal complement.

The ILDM method works, instead, with the Jacobian, Dg, of Eq. (2.3). A Schur decomposition transforms Dg into upper triangular form,

$$Dg = QNQ', \quad N = \begin{pmatrix} N_s & N_{sf} \\ 0 & N_f \end{pmatrix},$$

where $Q = (Q_s, Q_f)$ is unitary. The eigenvalues of Dg appear on the diagonal of N in descending order of their real parts, from least negative in the upper left to most negative in the lower right. The first N_s Schur vectors (the columns of Q_s) form an orthogonal basis of the slow subspace and the remaining N_f Schur vectors (the columns of Q_f) an orthogonal basis of the orthogonal complement of the slow subspace. The vector field g is entirely in the slow subspace if it is orthogonal to this orthogonal complement — that is, if

$$Q_f' g = 0.$$

This equation defines the ILDM; see [20, Section 3].

It was shown in [20] that the ILDM is only a first-order approximation to $\mathcal{M}_\varepsilon$. The error is always $\mathcal{O}(\varepsilon^2)$, unless $\mathcal{M}_0$ is linear. The error can be traced back to the choice of the operator. The tangent space is a left-invariant subspace of the Jacobian only to leading order, so putting Dg in upper triangular form yields the orthogonal complement only to leading order. Since the linearized system is only an approximation of the original ODEs (2.2), this choice does not produce an exact result unless g is linear. The success of the CSP method in approximating the slow manifold is due to the fact that the ODEs for the amplitudes f are equivalent to the ODEs (2.2). That is, the full nonlinearity is retained.

The time-derivative term must be included in the evaluation of Λ, see Eq. (3.6); otherwise, the accuracy of the CSP method is compromised. In fact, such an omission results in implementing the ILDM rather than the CSP method, which may be seen as follows. With our initial choice of a point-independent basis $A^{(0)}$, the matrix $\Lambda_{(0)}$ is similar to Dg; see Eq. (3.9). The omission of the term $(dB_{(m)}/dt)A^{(m)}$ in the calculation of $\Lambda_{(m)}$, for $m = 1, 2, \ldots$, would lead to the formula $\Lambda_{(m)} = (I + \tilde{P}_{(m)})B_{(0)}(Dg)A^{(0)}(I - \tilde{P}_{(m)})$, which would imply that $\Lambda_{(m)}$ is similar to Dg. Therefore, the one-step CSP method would put Dg, rather than Λ, in lower-triangular form, just like the ILDM method. After the second iteration, one would make an error (proportional to the curvature of $\mathcal{M}_0$) at $\mathcal{O}(\varepsilon^2)$, which subsequent iterations would not remove. The MMH example in Section 6.1 illustrates these observations.

Chapter 4

Proof of Theorems 3.3.1 and 3.3.2

Our proof of Theorem 3.3.1 is based on the validity of Theorem 3.3.2; hence, we first prove the latter theorem. In particular, we show that, after m iterations, the one-step CSP method generates a manifold $\tilde{\mathcal{K}}_\varepsilon^{(m)}$, whose asymptotic expansion agrees with that of $\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^m)$ (Section 4.1). In other words, the one-step CSP method is as accurate as the full CSP method is claimed to be in Theorem 3.3.1. To establish Theorem 3.3.1, then, we analyze the modifications introduced by the second step. We show that, at the mth iteration, the second step affects only terms of $\mathcal{O}(\varepsilon^{m+1})$ and higher (Section 4.2). Hence, $\mathcal{K}_\varepsilon^{(m)}$ approximates $\mathcal{M}_\varepsilon$ as accurately as $\tilde{\mathcal{K}}_\varepsilon^{(m)}$, and Theorem 3.3.1 follows.

4.1 Proof of Theorem 3.3.2

The proof of the theorem is by induction. In Section 4.1.1, we formulate the induction hypothesis. In Section 4.1.2, we prove the validity of the theorem for $m = 1$, *i.e.*, after one CSP iteration has been carried out. In Section 4.1.3, we prove the theorem for the case of a general m.

4.1.1 The induction hypothesis

The central idea of the proof of Theorem 3.3.2 is to express the CSP condition (3.25) in a form that resembles that of the invariance equation (2.7) and then to derive the conditions under which the left and right members of the two equations are the same at each order.

We begin by expressing the quantities $\tilde{A}^{(m+1)}$, $\tilde{B}_{(m+1)}$, and $\tilde{\Lambda}_{(m+1)}$ in terms of the original quantities $A^{(0)}$, $B_{(0)}$, and $\Lambda_{(0)}$. Applying the definition (3.17) recursively, we find

$$\tilde{A}^{(m+1)} = A^{(0)} \prod_{j=0}^{m} (I - \tilde{U}_{(j)}).$$

Since each $\tilde{U}_{(j)}$ is nilpotent, it follows that

$$\tilde{A}^{(m+1)} = A^{(0)}(I - \tilde{P}_{(m)}), \tag{4.1}$$

where

$$\tilde{P}_{(q)} = \sum_{j=0}^{q} \tilde{U}_{(j)} = \begin{pmatrix} 0 & \sum_{\ell=0}^{q} (\tilde{\Lambda}_{(\ell)}^{11})^{-1} \tilde{\Lambda}_{(\ell)}^{12} \\ 0 & 0 \end{pmatrix}. \tag{4.2}$$

Similarly,

$$\tilde{B}_{(m+1)} = (I + \tilde{P}_{(m)}) B_{(0)}. \tag{4.3}$$

Substituting Eqs. (4.1) and (4.3) into the transformation formula (3.16), and recalling that $\tilde{\Lambda}_{(0)} = \Lambda_{(0)}$ and $\tilde{P}_{(m)} d\tilde{P}_{(m)}/dt = 0$, we find

$$\tilde{\Lambda}_{(m+1)} = (I + \tilde{P}_{(m)}) \Lambda_{(0)} (I - \tilde{P}_{(m)}) + \frac{d\tilde{P}_{(m)}}{dt}. \tag{4.4}$$

We use these expressions to rewrite Eq. (3.25). Since $B_{(0)}^{22} = 0$, the equation becomes

$$B_{(0)}^{12} g^2 + \varepsilon \left[\tilde{P}_{(m-1)} B_{(0)}^{21} + B_{(0)}^{11} \right] g^1 = 0$$

or, since $B_{(0)}^{12} = (A_{21}^{(0)})^{-1}$,

$$g^2 + \varepsilon A_{21}^{(0)} \left[\tilde{P}_{(m-1)} B_{(0)}^{21} + B_{(0)}^{11} \right] g^1 = 0. \tag{4.5}$$

The last equation has the same form as the invariance equation (2.7). The solution of Eq. (2.7) is $x^2 = h_\varepsilon(x^1)$, which defines $\mathcal{M}_\varepsilon$, while the solution of Eq. (4.5) is $x^2 = \tilde{\psi}_{(m)}(x^1, \varepsilon)$, which defines $\tilde{\mathcal{K}}_\varepsilon^{(m)}$.

We analyze the CSP condition (4.5) order by order, up to and including the terms of $\mathcal{O}(\varepsilon^m)$. We recall that the components of the vector field $g(x, \varepsilon)$ are evaluated at $x^2 = \tilde{\psi}_{(m)}(x^1, \varepsilon)$, the matrix $\tilde{P}_{(m-1)}$ is evaluated at $x^2 = \tilde{\psi}_{(m-1)}(x^1, \varepsilon)$, and the blocks of $A^{(0)}$ and $B_{(0)}$ are constant. Substituting the asymptotic expansion of $\tilde{\psi}_{(m)}$,

$$\tilde{\psi}_{(m)}(x^1, \varepsilon) = \sum_{j=0}^{\infty} \varepsilon^j \tilde{\psi}_{(m,j)}(x^1), \quad \varepsilon \downarrow 0,$$

into Eq. (4.5) and setting the coefficients of $1, \varepsilon, \ldots, \varepsilon^m$ equal to zero, we obtain a set of equations,

$$g_j^2 + A_{21}^{(0)} \left[\tilde{P}_{(m-1,0)} B_{(0)}^{21} + B_{(0)}^{11} \right] g_{j-1}^1 + \sum_{\ell=1}^{j-1} A_{21}^{(0)} \tilde{P}_{(m-1,\ell)} B_{(0)}^{21} g_{j-\ell-1}^1 = 0, \tag{4.6}$$

for $j = 0, 1, \ldots, m$. Here, $\tilde{P}_{(m-1,\ell)}$ is the coefficient of the $\mathcal{O}(\varepsilon^\ell)$ term in the asymptotic expansion of $\tilde{P}_{(m-1)}$.

Equation (4.6) defines $\tilde{\psi}_{(m,j)}$ for $j = 0, 1, \ldots, m$. The leading-order ($j = 0$) equation in the system (4.6) is the same for all m,

$$g^2(\cdot, \tilde{\psi}_{(m,0)}(\cdot), 0) = 0, \quad m = 0, 1, \ldots .$$

This is also the equation defining h_0. Its solution need not be unique, but we can identify each $\tilde{\psi}_{(m,0)}$ with h_0,

$$\tilde{\psi}_{(m,0)}(\cdot) = h_0(\cdot), \quad m = 0, 1, \ldots .$$

Then also $\tilde{\psi}_{(m)}(\cdot\,, 0) = h_0(\cdot)$ for $m = 0, 1, \ldots$, so to leading order each manifold $\tilde{\mathcal{K}}_\varepsilon^{(m)}$ coincides with $\mathcal{M}_0$.

We wish to show that $\tilde{\psi}_{(m,j)} = h_j$ also for $j = 1, 2, \ldots, m$. To this end, we compare Eqs. (2.12) and (4.6). For a fixed j, the two equations match if

$$A_{21}^{(0)} \left[\tilde{P}_{(m-1,0)} B_{(0)}^{21} + B_{(0)}^{11} \right] = -Dh_0, \tag{4.7}$$

$$A_{21}^{(0)} \tilde{P}_{(m-1,\ell)} B_{(0)}^{21} = -Dh_\ell, \quad \ell = 1, \ldots, j-1. \tag{4.8}$$

Conversely, if Eqs. (4.7) and (4.8) hold, then $\tilde{\psi}_{(m,j)} = h_j$. Notice that Eqs. (4.7) and (4.8) are independent of j; hence, they are nested, in the sense that, when j is increased by one, the equations for lower values of j remain the same. Thus, it suffices to prove Eqs. (4.7) and (4.8) for $j = m$. The proof is by induction on m, where the induction hypothesis is

$$\tilde{U}_{(m-1)}(\cdot\,, \tilde{\psi}_{(m-1)}(\cdot, \varepsilon), \varepsilon) = \mathcal{O}(\varepsilon^{m-1}), \tag{4.9}$$

$$A_{21}^{(0)} \left[\tilde{P}_{(m-1)}(\cdot\,, \tilde{\psi}_{(m-1)}(\cdot, \varepsilon), \varepsilon) B_{(0)}^{21} + B_{(0)}^{11} \right] = -\sum_{j=0}^{m-1} \varepsilon^j Dh_j(\cdot) + \mathcal{O}(\varepsilon^m), \tag{4.10}$$

$$\tilde{\psi}_{(m)}(\cdot\,, \varepsilon) = \sum_{j=0}^{m} \varepsilon^j h_j(\cdot) + \mathcal{O}(\varepsilon^{m+1}). \tag{4.11}$$

The validity of these equations for $m = 1$ is shown in Section 4.1.2. The induction step is carried out in Section 4.1.3.

4.1.2 Proof of Theorem 3.3.2 for $m = 1$

We fix $m = 1$ and consider the $\mathcal{O}(\varepsilon)$ terms of Eq. (4.5),

$$(D_{x^2} g^2)_0 \tilde{\psi}_{(1,1)} + (D_\varepsilon g^2)_0 + A_{21}^{(0)} \left[\tilde{P}_{(0,0)} B_{(0)}^{21} + B_{(0)}^{11} \right] g_0^1 = 0. \tag{4.12}$$

The first and second terms in this equation are exactly the same as those in the equation for h_1, see Eq. (2.14). Therefore, we need only to show that the third term equals $-(Dh_0)g_0^1$ in order to prove the theorem for $m = 1$.

According to the definitions (4.2) and (3.10) with $m = 0$, we have

$$\tilde{P}_{(0)} = \tilde{U}_{(0)} = (\tilde{\Lambda}^{11}_{(0)})^{-1}\tilde{\Lambda}^{12}_{(0)} = (\Lambda^{11}_{(0)})^{-1}\Lambda^{12}_{(0)}, \tag{4.13}$$

where $\Lambda_{(0)} = B_{(0)}(Dg)A^{(0)}$, according to the definition in Eq. (3.9). Now, $\Lambda_{(0)}$ admits an asymptotic expansion, $\Lambda_{(0)} = \sum_{j=0}^{\infty}\varepsilon^j\Lambda_{(0,j)}$, and each of the coefficient matrices $\Lambda_{(0,j)}$ consists of four blocks,

$$\begin{aligned}
\Lambda^{11}_{(0,j)} &= \left[B^{12}_{(0)}(D_{x^2}g^2)_j + B^{11}_{(0)}(D_{x^2}g^1)_{j-1}\right]A^{(0)}_{21}, && (4.14)\\
\Lambda^{12}_{(0,j)} &= B^{12}_{(0)}\left[(D_{x^1}g^2)_jA^{(0)}_{12} + (D_{x^2}g^2)_jA^{(0)}_{22}\right] \\
&\quad + B^{11}_{(0)}\left[(D_{x^1}g^1)_{j-1}A^{(0)}_{12} + (D_{x^2}g^1)_{j-1}A^{(0)}_{22}\right], && (4.15)\\
\Lambda^{21}_{(0,j)} &= B^{21}_{(0)}(D_{x^2}g^1)_{j-1}A^{(0)}_{21}, && (4.16)\\
\Lambda^{22}_{(0,j)} &= B^{21}_{(0)}\left[(D_{x^1}g^1)_{j-1}A^{(0)}_{12} + (D_{x^2}g^1)_{j-1}A^{(0)}_{22}\right]. && (4.17)
\end{aligned}$$

The notation $(\,\cdot\,)_j$ indicates the jth term in the asymptotic expansion of the quantity inside the parentheses, and it is understood that such a term is absent if the subscript is negative.

A direct evaluation shows that the blocks $\Lambda^{11}_{(0,0)}$ and $\Lambda^{12}_{(0,0)}$ are nonzero. Therefore, $\Lambda^{11}_{(0)}$ and $\Lambda^{12}_{(0)}$ are both $\mathcal{O}(1)$, and

$$\tilde{P}_{(0,0)} = \tilde{U}_{(0,0)} = (\Lambda^{11}_{(0,0)})^{-1}\Lambda^{12}_{(0,0)} = B^{12}_{(0)}\left[\left(D_{x^2}g^2\right)_0^{-1}\left(D_{x^1}g^2\right)_0A^{(0)}_{12} + A^{(0)}_{22}\right]. \tag{4.18}$$

Here, all the quantities are evaluated on $\mathcal{M}_0$, where the identity

$$\left(D_{x^2}g^2\right)_0^{-1}\left(D_{x^1}g^2\right)_0 = -Dh_0$$

holds. Hence, Eq. (4.18) implies

$$\tilde{P}_{(0,0)} = \tilde{U}_{(0,0)} = B^{12}_{(0)}\left(A^{(0)}_{22} - (Dh_0)A^{(0)}_{12}\right).$$

Finally, substituting this expression for $\tilde{P}_{(0,0)}$ into Eq. (4.12) and using the identity $A^{(0)}_{21}B^{11}_{(0)} = -A^{(0)}_{22}B^{21}_{(0)}$, we obtain

$$(D_{x^2}g^2)_0\tilde{\psi}_{(1,1)} + (D_\varepsilon g^2)_0 - (Dh_0)g_0^1 = 0.$$

This equation for $\tilde{\psi}_{(1,1)}$ is the same as Eq. (2.14) for h_1; hence, $\tilde{\psi}_{(1,1)} = h_1$ and $\tilde{\psi}_{(1)} = h_0 + \varepsilon h_1 + \mathcal{O}(\varepsilon^2)$. This proves the theorem for $m = 1$.

4.1.3 Proof of Theorem 3.3.2 for $m \geq 2$

We assume that Eqs. (4.9)–(4.11) hold for $0, 1, \ldots, m$ and prove that they hold for $m+1$. By our discussion of Eqs. (4.7) and (4.8), Eq. (4.11) follows immediately from Eq. (4.10), so we need only to consider Eqs. (4.9) and (4.10).

Establishing Eq. (4.9.) We first consider Eq. (4.9). The induction hypothesis provides us with the estimate

$$\tilde{U}_{(i)}(\cdot\,, \tilde{\psi}_{(i)}(\cdot, \varepsilon), \varepsilon) = \mathcal{O}(\varepsilon^i), \quad i = 0, 1, \ldots, m-1.$$

Also,

$$\tilde{\psi}_{(m)} = \tilde{\psi}_{(i)} + \mathcal{O}(\varepsilon^{i+1}), \quad i = 0, 1, \ldots, m-1.$$

Hence,

$$\tilde{U}_{(i)}(\cdot\,, \tilde{\psi}_{(m)}(\cdot, \varepsilon), \varepsilon) = \tilde{U}_{(i)}(\cdot\,, \tilde{\psi}_{(i)}(\cdot, \varepsilon), \varepsilon) + \mathcal{O}(\varepsilon^{i+1}), \quad i = 0, 1, \ldots, m-1.$$

from which it follows that

$$\tilde{U}_{(i)}(\cdot\,, \tilde{\psi}_{(m)}(\cdot, \varepsilon), \varepsilon) = \mathcal{O}(\varepsilon^i), \quad i = 0, 1, \ldots, m-1. \tag{4.19}$$

In particular, $\tilde{U}_{(0)}(\cdot\,, \tilde{\psi}_{(m)}(\cdot, \varepsilon), \varepsilon) = \mathcal{O}(1)$, so

$$\tilde{P}_{(m-1)} = \sum_{\ell=0}^{m-1} \tilde{U}_{(\ell)} = \mathcal{O}(1) \text{ on } \tilde{\mathcal{K}}_\varepsilon^{(m)}.$$

This asymptotic estimate can be used to derive asymptotic expansions of the blocks of $\Lambda_{(m)}$. We begin with $\tilde{\Lambda}^{11}_{(m)}$. From Eq. (4.4), we have

$$\tilde{\Lambda}^{11}_{(m)} = \Lambda^{11}_{(0)} + \tilde{P}_{(m-1)}\Lambda^{21}_{(0)}.$$

Since $\Lambda^{21}_{(0)} = \mathcal{O}(\varepsilon)$ by Eq. (4.16), we see immediately that

$$\tilde{\Lambda}^{11}_{(m)} = \Lambda^{11}_{(0,0)} + \mathcal{O}(\varepsilon). \tag{4.20}$$

Next, we examine the block $\tilde{\Lambda}^{12}_{(m)}$. From Eq. (3.21), we have

$$\tilde{\Lambda}^{12}_{(m)} = \tilde{U}_{(m-1)}\tilde{\Lambda}^{22}_{(m-1)} - \tilde{U}_{(m-1)}\tilde{\Lambda}^{21}_{(m-1)}\tilde{U}_{(m-1)} + \frac{d\tilde{U}_{(m-1)}}{dt}.$$

First, $\tilde{U}_{(m-1)}(\cdot\,, \tilde{\psi}_{(m)}(\cdot, \varepsilon), \varepsilon) = \mathcal{O}(\varepsilon^{m-1})$ by Eq. (4.19). Also, $\tilde{\Lambda}^{21}_{(m-1)} = \Lambda^{21}_{(0)} = \mathcal{O}(\varepsilon)$ on $\tilde{\mathcal{K}}_\varepsilon^{(m)}$ by Eqs. (4.4) and (4.16). Moreover, $\tilde{\Lambda}^{22}_{(m-1)} = \Lambda^{22}_{(0)} - \Lambda^{21}_{(0)}\tilde{P}_{(m-2)} = \mathcal{O}(\varepsilon)$ by

Eqs. (4.4) and (4.17). Finally, by applying Lemma 9.2.1 with $V = \tilde{U}_{(m-1)}$, we find that $d\tilde{U}_{(m-1)}/dt$ is $\mathcal{O}(\varepsilon^m)$. Putting these estimates together, we obtain the estimate

$$\tilde{\Lambda}^{12}_{(m)} = \varepsilon^m \tilde{\Lambda}^{12}_{(m,m)} + \mathcal{O}(\varepsilon^{m+1}), \tag{4.21}$$

where we grouped all of the $\mathcal{O}(\varepsilon^m)$ terms into $\varepsilon^m \tilde{\Lambda}^{12}_{(m,m)}$. By combining the definition (3.10) with Eqs. (4.20) and (4.21), we derive the desired estimate, $\tilde{U}_{(m)} = (\tilde{\Lambda}^{11}_{(m)})^{-1}\tilde{\Lambda}^{12}_{(m)} = \mathcal{O}(\varepsilon^m)$.

Remark. While the estimates of $\tilde{\Lambda}^{21}_{(m)}$ and $\tilde{\Lambda}^{22}_{(m)}$ are not needed here, they will be needed in Section 4.2. First, $\tilde{\Lambda}^{21}_{(m)} = \Lambda^{21}_{(0)} = \mathcal{O}(\varepsilon)$ on $\tilde{\mathcal{K}}^{(m)}_{\varepsilon}$, by Eqs. (4.4) and (4.16). Then, $\tilde{\Lambda}^{22}_{(m)} = \Lambda^{22}_{(0)} - \Lambda^{21}_{(0)}\tilde{P}_{(m-1)}$ by Eq. (4.4). Now, $\Lambda^{22}_{(0)} = \mathcal{O}(\varepsilon)$ by Eq. (4.17), and thus the discussion for the size of $\tilde{\Lambda}^{11}_{(m)}$ also yields that $\tilde{\Lambda}^{22}_{(m)} = \mathcal{O}(\varepsilon)$. Putting the estimates of this section together, we obtain

$$\tilde{\Lambda}_{(m)}(\cdot\,, \tilde{\psi}_{(m)}(\cdot, \varepsilon), \varepsilon) = \begin{pmatrix} \Lambda^{11}_{(0,0)} + \mathcal{O}(\varepsilon) & \varepsilon^m \tilde{\Lambda}^{12}_{(m,m)} + \mathcal{O}(\varepsilon^{m+1}) \\ \varepsilon\Lambda^{21}_{(0,1)} + \mathcal{O}(\varepsilon^2) & \varepsilon\tilde{\Lambda}^{22}_{(m,1)} + \mathcal{O}(\varepsilon^2) \end{pmatrix}. \tag{4.22}$$

Establishing Eq. (4.10.) Next, we consider Eq. (4.10). The induction hypothesis gives the estimate

$$A^{(0)}_{21}[\tilde{P}_{(i)}(\cdot\,, \tilde{\psi}_{(i)}(\cdot, \varepsilon), \varepsilon)B^{21}_{(0)} + B^{11}_{(0)}] = -\sum_{j=0}^{i} \varepsilon^j Dh_j(\cdot) + \mathcal{O}(\varepsilon^{i+1}),$$

for $i = 0, 1, \ldots, m-1$. Our goal is to show that this equation also holds for $i = m$. We first show that the terms up to and including $\mathcal{O}(\varepsilon^{m-1})$ in both members of the equation agree for $i = m$. Then we analyze the terms of $\mathcal{O}(\varepsilon^m)$.

By the induction hypothesis, we have the asymptotic expansion

$$A^{(0)}_{21}\left[\tilde{P}_{(m-1)}(\cdot\,, \tilde{\psi}_{(m-1)}(\cdot, \varepsilon), \varepsilon)B^{21}_{(0)} + B^{11}_{(0)}\right] = -\sum_{j=0}^{m-1} \varepsilon^j Dh_j(\cdot) + \mathcal{O}(\varepsilon^m).$$

Also by the induction hypothesis, $\tilde{\psi}_{(m)} = \tilde{\psi}_{(m-1)} + \mathcal{O}(\varepsilon^m)$. Hence,

$$A^{(0)}_{21}\left[\tilde{P}_{(m-1)}(\cdot\,, \tilde{\psi}_{(m)}(\cdot, \varepsilon), \varepsilon)B^{21}_{(0)} + B^{11}_{(0)}\right] = -\sum_{j=0}^{m-1} \varepsilon^j Dh_j(\cdot) + \mathcal{O}(\varepsilon^m).$$

The definition (4.2) of $\tilde{P}_{(m)}$ yields the update formula

$$\tilde{P}_{(m)} = \tilde{P}_{(m-1)} + \tilde{U}_{(m)}. \tag{4.23}$$

We already showed that $\tilde{U}_{(m)}(\cdot\,,\tilde{\psi}_{(m)},\varepsilon) = \mathcal{O}(\varepsilon^m)$, so Eq. (4.23) implies that the asymptotic expansions of $\tilde{P}_{(m)}(\cdot\,,\tilde{\psi}_{(m)}(\cdot,\varepsilon),\varepsilon)$ and $\tilde{P}_{(m-1)}(\cdot\,,\tilde{\psi}_{(m)}(\cdot,\varepsilon),\varepsilon)$ agree up to and including terms of $\mathcal{O}(\varepsilon^{m-1})$. The same, then, holds for the asymptotic expansions of the matrices

$$A_{21}^{(0)}[\tilde{P}_{(m)}(\cdot\,,\tilde{\psi}_{(m)}(\cdot,\varepsilon),\varepsilon)B_{(0)}^{21} + B_{(0)}^{11}] \quad \text{and} \quad A_{21}^{(0)}[\tilde{P}_{(m-1)}(\cdot\,,\tilde{\psi}_{(m)}(\cdot,\varepsilon),\varepsilon)B_{(0)}^{21} + B_{(0)}^{11}].$$

Therefore,

$$A_{21}^{(0)}\left[\tilde{P}_{(m)}(\cdot\,,\tilde{\psi}_{(m)}(\cdot,\varepsilon),\varepsilon)B_{(0)}^{21} + B_{(0)}^{11}\right] = -\sum_{j=0}^{m-1}\varepsilon^j Dh_j(\cdot) + \mathcal{O}(\varepsilon^m).$$

In other words,

$$\begin{aligned} A_{21}^{(0)}\left[\tilde{P}_{(m,0)}B_{(0)}^{21} + B_{(0)}^{11}\right] &= -Dh_0, \\ A_{21}^{(0)}\tilde{P}_{(m,j)}B_{(0)}^{21} &= -Dh_j, \quad \text{for all } \; j = 1,\dots,m-1, \end{aligned}$$

which establishes Eq. (4.10) for all terms up to and including $\mathcal{O}(\varepsilon^{m-1})$.

It remains to show that the terms of $\mathcal{O}(\varepsilon^m)$ in both members of Eq. (4.10) agree, that is,

$$A_{21}^{(0)}\tilde{P}_{(m,m)}B_{(0)}^{21} = -Dh_m.$$

We achieve this by deriving an explicit formula for $A_{21}^{(0)}\tilde{P}_{(m,m)}B_{(0)}^{21}$ and comparing it to that for Dh_m, which is given in the Appendix (Lemma 9.1.1). We proceed in two steps. In step one, we express $A_{21}^{(0)}\tilde{P}_{(m,m)}B_{(0)}^{21}$ in terms of $\tilde{P}_{(m-1,0)},\dots,\tilde{P}_{(m-1,m-1)}$. Then, in step two, we obtain the explicit formula for $A_{21}^{(0)}\tilde{P}_{(m,m)}B_{(0)}^{21}$ in terms of the vector field and of Dh_i, $i = 0,1,\dots,m-1$.

Step 1. Recall the update formula (4.23), $\tilde{P}_{(m)} = \tilde{P}_{(m-1)} + \tilde{U}_{(m)}$. Using the definition (3.10) of $\tilde{U}_{(m)}$ and the explicit formula (4.4) for $\tilde{\Lambda}_{(m)}$, we can express $\tilde{U}_{(m)}$ in terms of $\Lambda_{(0)}$ and $\tilde{P}_{(m-1)}$. In particular, Eq. (4.22) implies that $\tilde{U}_{(m,m)} = (\Lambda_{(0,0)}^{11})^{-1}\tilde{\Lambda}_{(m,m)}^{12}$. Also, Eq. (4.4) gives

$$\tilde{\Lambda}_{(m)}^{12} = \Lambda_{(0)}^{12} - \Lambda_{(0)}^{11}\tilde{P}_{(m-1)} + \tilde{P}_{(m-1)}\Lambda_{(0)}^{22} - \tilde{P}_{(m-1)}\Lambda_{(0)}^{21}\tilde{P}_{(m-1)} + \frac{d\tilde{P}_{(m-1)}}{dt}.$$

It follows that

$$\begin{aligned} \tilde{U}_{(m,m)} &= (\Lambda_{(0,0)}^{11})^{-1}\left[\Lambda_{(0,m)}^{12} - \left(\Lambda_{(0)}^{11}\tilde{P}_{(m-1)}\right)_m + \left(\tilde{P}_{(m-1)}\Lambda_{(0)}^{22}\right)_m \right. \\ &\qquad \left. - \left(\tilde{P}_{(m-1)}\Lambda_{(0)}^{21}\tilde{P}_{(m-1)}\right)_m + \left(\frac{d\tilde{P}_{(m-1)}}{dt}\right)_m\right], \end{aligned} \tag{4.24}$$

where we recall the notational convention that $(\,\cdot\,)_m$ stands for the coefficient of the $\mathcal{O}(\varepsilon^m)$ term in the asymptotic expansion of the quantity in parentheses. Using Lemma 9.2.1 with $V = \tilde{P}_{(m-1)}$ and the fact that $\Lambda^{22}_{(0,0)}$ and $\Lambda^{21}_{(0,0)}$ are both zero, we rewrite Eq. (4.24) as

$$\tilde{U}_{(m,m)} \;=\; (\Lambda^{11}_{(0,0)})^{-1}\left[J_1 + (J_2 - \Lambda^{11}_{(0,0)}\tilde{P}_{(m-1,m)}) + J_3 + J_4 + J_5\right], \qquad (4.25)$$

where

$$J_1 = \Lambda^{12}_{(0,m)}, \qquad J_2 = -\sum_{\ell=0}^{m-1} \Lambda^{11}_{(0,m-\ell)}\tilde{P}_{(m-1,\ell)}, \qquad J_3 = \sum_{\ell=0}^{m-1} \tilde{P}_{(m-1,\ell)}\Lambda^{22}_{(0,m-\ell)},$$
$$J_4 = -\sum_{i=0}^{m-1}\sum_{j=0}^{m-1-i} \tilde{P}_{(m-1,j)}\Lambda^{21}_{(0,m-i-j)}\tilde{P}_{(m-1,i)}, \qquad J_5 = \sum_{\ell=0}^{m-1} \frac{d\tilde{P}_{(m-1,\ell)}}{dy} g^1_{m-1-\ell}. \qquad (4.26)$$

Substituting the expression (4.25) into the update formula (4.23) for $\tilde{P}_{(m)}$, we find

$$A^{(0)}_{21}\tilde{P}_{(m,m)}B^{21}_{(0)} \;=\; A^{(0)}_{21}(\Lambda^{11}_{(0,0)})^{-1}\left[J_1 + J_2 + J_3 + J_4 + J_5\right]B^{21}_{(0)}, \qquad (4.27)$$

Step 2. We rewrite the terms $J_1, \ldots, J_5$ by means of the induction hypothesis and the explicit formulas (4.14)–(4.17) for the blocks of $\Lambda_{(0)}$.

Equation (4.14) and the identity $A^{(0)}_{21}B^{12}_{(0)} = I_{N_f}$ imply that

$$A^{(0)}_{21}\left(\Lambda^{11}_{(0,0)}\right)^{-1} = \left((D_{x^2}g^2)_0\right)^{-1} A^{(0)}_{21}. \qquad (4.28)$$

Here, $(D_{x^2}g^2)_0$ stands for the leading order term in the asymptotic expansion of $(D_{x^2}g^2)(\cdot\,, \tilde{\psi}_{(m)}(\cdot,\varepsilon),\varepsilon)$. Since $\tilde{\psi}_{(m)}$ and h_ε agree up to and including $\mathcal{O}(\varepsilon^m)$ terms by assumption, the asymptotic expansions of the matrices $(D_{x^2}g^2)(\cdot\,, \tilde{\psi}_{(m)}(\cdot,\varepsilon),\varepsilon)$ and $(D_{x^2}g^2)(\cdot\,, h_\varepsilon(\cdot),\varepsilon)$ also agree up to and including $\mathcal{O}(\varepsilon^m)$ terms. For the remainder of this section, it does not matter whether quantities are evaluated on $\tilde{\mathcal{K}}^{(m)}_\varepsilon$ or on $\mathcal{M}_\varepsilon$, since only the coefficients of ε^m or lower appear in our formulas. Accordingly, we make no distinction between the asymptotic expansions of a quantity evaluated on the two manifolds.

Using Eq. (4.15) and the identities $B^{12}_{(0)} = (A^{(0)}_{21})^{-1}$, $B^{21}_{(0)} = (A^{(0)}_{12})^{-1}$, and $B^{11}_{(0)} = -B^{12}_{(0)}A^{(0)}_{22}B^{21}_{(0)}$, we find

$$\begin{aligned} A^{(0)}_{21}J_1B^{21}_{(0)} \;=\;& (D_{x^1}g^2)_m + (D_{x^2}g^2)_m A^{(0)}_{22}B^{21}_{(0)} - A^{(0)}_{22}B^{21}_{(0)}(D_{x^1}g^1)_{m-1} \\ & - A^{(0)}_{22}B^{21}_{(0)}(D_{x^2}g^1)_{m-1}A^{(0)}_{22}B^{21}_{(0)}. \end{aligned} \qquad (4.29)$$

Next, substituting for $A_{21}^{(0)}\tilde{P}_{(m-1,\ell)}B_{(0)}^{21}$ from the induction hypothesis (4.10), we obtain

$$\begin{aligned} A_{21}^{(0)}J_2B_{(0)}^{21} &= \sum_{\ell=0}^{m-1}(D_{x^2}g^2)_{m-\ell}Dh_\ell - (D_{x^2}g^2)_m A_{22}^{(0)}B_{(0)}^{21} \\ &\quad - \sum_{\ell=0}^{m-1} A_{22}^{(0)}B_{(0)}^{21}(D_{x^2}g^1)_{m-1-\ell}Dh_\ell + A_{22}^{(0)}B_{(0)}^{21}(D_{x^2}g^1)_{m-1}A_{22}^{(0)}B_{(0)}^{21}. \quad (4.30) \end{aligned}$$

Then, using Eq. (4.17) and the assumptions of the lemma, we find

$$\begin{aligned} A_{21}^{(0)}J_3B_{(0)}^{21} &= -\sum_{\ell=0}^{m-1} Dh_\ell(D_{x^1}g^1)_{m-1-\ell} - \sum_{\ell=0}^{m-1} Dh_\ell(D_{x^2}g^1)_{m-1-\ell}A_{22}^{(0)}B_{(0)}^{21} \\ &\quad + A_{22}^{(0)}B_{(0)}^{21}(D_{x^1}g^1)_{m-1} + A_{22}^{(0)}B_{(0)}^{21}(D_{x^2}g^1)_{m-1}A_{22}^{(0)}B_{(0)}^{21}. \quad (4.31) \end{aligned}$$

In the same vein, we use the induction hypothesis on J_4,

$$\begin{aligned} A_{21}^{(0)}J_4B_{(0)}^{21} = &-\sum_{i=0}^{m-1}\sum_{j=0}^{m-1-i} Dh_j(D_{x^2}g^1)_{m-1-i-j}Dh_i + \sum_{i=0}^{m-1} A_{22}^{(0)}B_{(0)}^{21}(D_{x^2}g^1)_{m-1-i}Dh_i \\ &+ \sum_{j=0}^{m-1} Dh_j(D_{x^2}g^1)_{m-1-j}A_{22}^{(0)}B_{(0)}^{21} - A_{22}^{(0)}B_{(0)}^{21}(D_{x^2}g^1)_{m-1}A_{22}^{(0)}B_{(0)}^{21}. \quad (4.32) \end{aligned}$$

The terms in Eq. (4.27) containing $A_{22}^{(0)}$ sum to zero, which may be seen as follows. The second and fourth terms in Eq. (4.29) cancel against the second and fourth terms in Eq. (4.30); the third term in Eq. (4.29) cancels against the third term in Eq. (4.31); the third term in Eq. (4.30) cancels against the second term in Eq. (4.32); and the second and fourth terms in Eq. (4.31) cancel against the third and fourth terms in Eq. (4.32). These cancellations were to be expected because the approximation should be independent of the choice of $A^{(0)}$.

Carrying out the same type of calculation as above, we obtain

$$A_{21}^{(0)}J_5B_{(0)}^{21} = -\sum_{\ell=0}^{m-1}(D^2h_\ell)g_{m-1-\ell}^1, \quad (4.33)$$

where we have used the symmetry of the bilinear form D^2h_ℓ.

Equations (4.28)–(4.33), together with the observed cancellations, yield

$$\begin{aligned} &A_{21}^{(0)}\tilde{P}_{(m,m)}B_{(0)}^{21} \\ &= \big((D_{x^2}g^2)_0\big)^{-1}\Bigg[(D_{x^1}g^2)_m + \sum_{\ell=0}^{m-1}(D_{x^2}g^2)_{m-\ell}(Dh_\ell) - \sum_{\ell=0}^{m-1}(D^2h_\ell)g_{m-1-\ell}^1 \\ &\quad - \sum_{\ell=0}^{m-1}(Dh_\ell)(D_{x^1}g^1)_{m-1-\ell} - \sum_{i=0}^{m-1}\sum_{j=0}^{m-1-i}(Dh_j)(D_{x^2}g^1)_{m-1-i-j}Dh_i\Bigg]. \end{aligned}$$

A term-by-term comparison with the expression for $-Dh_m$ given in the Appendix, Eq. (9.1), shows that $A_{21}^{(0)}\tilde{P}_{(m,m)}A_{(0)}^{21} = -Dh_m$. Thus, the proof of Theorem 3.3.2 is complete.

Remark. In general, the error term is nontrivial, as can already be seen at $m = 0$. The equation determining $\tilde{\psi}_{(0,1)}$ is

$$(D_{x^2}g^2)_0\tilde{\psi}_{(0,1)} + (D_\varepsilon g^2)_0 - A_{22}^{(0)}B_{(0)}^{21}g_0^1 = 0. \tag{4.34}$$

This equation is not the same as Eq. (2.14), which determines h_1. Where Eq. (2.14) has the term Dh_0, Eq. (4.34) has the term $A_{22}^{(0)}B_{(0)}^{21}$. When the slow manifold is nonlinear, Dh_0 depends on x^1, whereas $A_{22}^{(0)}B_{(0)}^{21}$ is a constant matrix. Therefore, in general $\tilde{\psi}_{(0,1)} \neq h_1$, and the strongest claim we can make is $\tilde{\psi}_{(0)} = h_0 + \mathcal{O}(\varepsilon)$. A similar argument applies to higher values of m.

4.2 Proof of Theorem 3.3.1

We now turn our attention to the full CSP method and prove Theorem 3.3.1. Since the full CSP method and the one-step CSP method start from the same basis, the conditions (3.13) and (3.25) are the same for $m = 0$,

$$B_{(0)}^1 g = 0.$$

Therefore, we can choose $\psi_{(0)} = \tilde{\psi}_{(0)} = h_0$.

4.2.1 Proof of Theorem 3.3.1 for $m = 1$

In this section, we carry out the first iteration of the full CSP method and determine the resulting approximation $\mathcal{K}_\varepsilon^{(1)}$ of the slow manifold. We then compare $\mathcal{K}_\varepsilon^{(1)}$ and $\tilde{\mathcal{K}}_\varepsilon^{(1)}$.

The update quantities $U_{(0)}$ and $L_{(0)}$ follow from the definition (3.10),

$$U_{(0)} = (\Lambda_{(0)}^{11})^{-1}\Lambda_{(0)}^{12}, \quad L_{(0)} = \Lambda_{(0)}^{21}(\Lambda_{(0)}^{11})^{-1}. \tag{4.35}$$

(We recall that we use the same notation $U_{(0)}$ and $L_{(0)}$ for the full matrix and the nonzero block.) In particular, Eqs. (4.35) and (4.13) imply that $U_{(0)} = \tilde{U}_{(0)}$. Next, we update the matrix $B_{(0)}$. Following Eq. (3.12), we find

$$B_{(1)} = (I - L_{(0)})(I + U_{(0)})B_{(0)}.$$

The upper and lower row blocks of $B_{(1)}$ are

$$B_{(1)}^1 = B_{(0)}^1 + U_{(0)}B_{(0)}^2, \tag{4.36}$$

$$B_{(1)}^2 = (I - L_{(0)}U_{(0)})B_{(0)}^2 - L_{(0)}B_{(0)}^1. \tag{4.37}$$

Since $\tilde{P}_{(0)} = \tilde{U}_{(0)} = U_{(0)}$ and $\psi_{(0)} = \tilde{\psi}_{(0)}$, Eqs. (4.3) and (4.36) imply that

$$B^1_{(1)} = \tilde{B}^1_{(1)},$$

so after the first iteration the CSP condition is the same as for the one-step method. Therefore, $\psi_{(1)} = \tilde{\psi}_{(1)}$ and, by Theorem 3.3.2,

$$\psi_{(1)} = h_0 + \varepsilon h_1 + \mathcal{O}(\varepsilon^2).$$

This proves Theorem 3.3.1 for $m = 1$.

4.2.2 The induction hypothesis

So far, we have established the identities $B^1_{(0)} = \tilde{B}^1_{(0)}$ and $B^1_{(1)} = \tilde{B}^1_{(1)}$, from which we could conclude that $\mathcal{K}^{(0)}_\varepsilon = \tilde{\mathcal{K}}^{(0)}_\varepsilon$ and $\mathcal{K}^{(1)}_\varepsilon = \tilde{\mathcal{K}}^{(1)}_\varepsilon$. In general, though, it is not true that $B^1_{(m)} = \tilde{B}^1_{(m)}$ for higher values of m, as we now demonstrate.

In the one-step CSP method, Eq. (4.3) yields

$$\tilde{B}^1_{(2)} = B^1_{(0)} + (\tilde{U}_{(0)} + \tilde{U}_{(1)})B^2_{(0)}.$$

By contrast, in the full CSP method, we obtain from Eq. (3.12)

$$B^1_{(2)} = \left(I_{N_f} - U_{(1)}L_{(0)}\right) B^1_{(0)} + \left(U_{(0)} + U_{(1)} - U_{(1)}L_{(0)}U_{(0)}\right) B^2_{(0)}. \tag{4.38}$$

The rows of $B^1_{(0)}$ and $B^2_{(0)}$ are linearly independent, as can be seen from Eq. (3.8), so the presence of the premultiplier of $B^1_{(0)}$ in the expression (4.38) implies that $B^1_{(2)} \neq \tilde{B}^1_{(2)}$. A similar argument shows that $B^1_{(m)} \neq \tilde{B}^1_{(m)}$ for $m = 2, 3, \ldots$. Consequently, the proof of Theorem 3.3.1 for $m = 1$ given in Section 4.2.1 does not generalize to higher values of m.

The matrix $\tilde{B}^1_{(m)}$ has an important property. Using Eq. (4.3), we write

$$\tilde{B}^1_{(m)} = \left(\tilde{P}_{(m-1)}B^{21}_{(0)} + B^{11}_{(0)},\, B^{12}_{(0)}\right) = B^{12}_{(0)} \left(A^{(0)}_{21} \left[\tilde{P}_{(m-1)}B^{21}_{(0)} + B^{11}_{(0)}\right],\, I_{N_f}\right).$$

Given the induction hypothesis (4.10), we rewrite this expression once more,

$$\tilde{B}^1_{(m)} = B^{12}_{(0)} \left(- \sum_{j=0}^{m-1} \varepsilon^j D h_j + \mathcal{O}(\varepsilon^m),\, I_{N_f} \right). \tag{4.39}$$

Take any $x^1 \in K$, and let the points $\tilde{Q} \in \tilde{\mathcal{K}}^{(m-1)}_\varepsilon$, $Q \in \mathcal{K}^{(m-1)}_\varepsilon$, and $Q' \in \mathcal{M}_\varepsilon$ be defined by

$$\tilde{Q} = (x^1, \tilde{\psi}_{(m-1)}(x^1, \varepsilon)), \quad Q = (x^1, \psi_{(m-1)}(x^1, \varepsilon)), \quad Q' = (x^1, h_\varepsilon(x^1)).$$

The N_f row vectors of the matrix $(-Dh_\varepsilon(y),\, I_{N_f})$ form an exact basis for $\mathcal{N}_{Q'}\mathcal{M}_\varepsilon$, the space normal to $\mathcal{M}_\varepsilon$ at Q'. Therefore, by Eq. (4.39), $\tilde{B}^1_{(m)}(\tilde{Q})$ is a linear combination of the basis vectors of $\mathcal{N}_{Q'}\mathcal{M}_\varepsilon$, up to and including terms of $\mathcal{O}(\varepsilon^{m-1})$, via the invertible matrix $B^{12}_{(0)}$. Hence, the columns of $\tilde{B}^1_{(m)}(\tilde{Q})$ form a basis for $\mathcal{N}_{Q'}\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^{m-1})$. This property of $\tilde{B}^1_{(m)}(\tilde{Q})$ was central to the proof of Theorem 3.3.2. We seek to prove a similar result for the rows of $B^1_{(m)}(Q)$.

The rows of $B_{(m)}(Q)$ can be written as linear combinations of the rows of $\tilde{B}_{(m)}(\tilde{Q})$,

$$B_{(m)}(Q) = T_{(m)}(x^1, \varepsilon)\tilde{B}_{(m)}(\tilde{Q}), \tag{4.40}$$

because $\tilde{B}_{(m)}(\tilde{Q})$ is invertible (see Eq. (4.3)). In terms of the constituent blocks,

$$\begin{aligned} B^1_{(m)} &= T^{11}_{(m)}\tilde{B}^1_{(m)} + T^{12}_{(m)}\tilde{B}^2_{(m)}, & (4.41)\\ B^2_{(m)} &= T^{21}_{(m)}\tilde{B}^1_{(m)} + T^{22}_{(m)}\tilde{B}^2_{(m)}. & (4.42) \end{aligned}$$

Equation (4.41) shows that the requirement that the rows of $B^1_{(m)}(Q)$ span $\mathcal{N}_{Q'}\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^{m-1})$ is equivalent to the conditions

$$T^{11}_{(m)}(x^1, \varepsilon) = \mathcal{O}(1) \text{ and invertible}, \quad T^{12}_{(m)}(x^1, \varepsilon) = \mathcal{O}(\varepsilon^m). \tag{4.43}$$

Assume for the moment that these conditions are satisfied. Then, the CSP condition (3.13) after the mth iteration can be recast as

$$\Big[T^{11}_{(m)}(y,\varepsilon)\tilde{B}^1_{(m)}(x^1, \psi_{(m-1)}(x^1,\varepsilon),\varepsilon) + T^{12}_{(m)}(x^1,\varepsilon)\tilde{B}^2_{(m)}(x^1, \psi_{(m-1)}(x^1,\varepsilon),\varepsilon)\Big]\, g(x,\varepsilon) = 0,$$

or, since $T^{11}_{(m)}(y,\varepsilon)$ is invertible,

$$\tilde{B}^1_{(m)} g + \left(T^{11}_{(m)}\right)^{-1} T^{12}_{(m)}\tilde{B}^2_{(m)} g = 0. \tag{4.44}$$

The second term is at least of $\mathcal{O}(\varepsilon^m)$, by the second assumption in Eq. (4.43), so the terms of $\mathcal{O}(\varepsilon^j)$ in Eqs. (3.25) and (4.44) are equal for $j = 0, 1, \ldots, m-1$. At $\mathcal{O}(\varepsilon^m)$, the two equations differ by the term $\left(T^{11}_{(m,0)}\right)^{-1} T^{12}_{(m,m)}\tilde{B}^2_{(m,0)} g(x^1, \psi_{(m,0)}(x^1), \varepsilon)$. Since the $\mathcal{O}(1)$ terms of the two equations agree, it follows that $\psi_{(m,0)} = \tilde{\psi}_{(m,0)} = h_0$ and, therefore, $g(x^1, \psi_{(q,0)}(x^1), \varepsilon) = 0$. Hence, Eqs. (4.5) and (4.44) agree up to and including terms of $\mathcal{O}(\varepsilon^m)$, so Eq. (4.44) produces the asymptotic expansion of the slow manifold up to and including terms of $\mathcal{O}(\varepsilon^m)$, by Theorem 3.3.2.

To complete the proof of Theorem 3.3.1, we need to verify the conditions (4.43) for $m = 2, 3, \ldots$, which we do by induction on m. The induction hypothesis is

$$\begin{aligned} T_{(m)}(\cdot\,, \psi_{(m-1)}(\cdot,\varepsilon),\varepsilon) &= \begin{pmatrix} I_{N_f} + \mathcal{O}(\varepsilon^2) & \varepsilon^m T^{12}_{(m,m)} + \mathcal{O}(\varepsilon^{m+1}) \\ \varepsilon T^{21}_{(m,1)} + \mathcal{O}(\varepsilon^2) & I_{N_s} + \mathcal{O}(\varepsilon^2) \end{pmatrix}, & (4.45)\\ \psi_{(m)}(\cdot\,,\varepsilon) &= \sum_{j=0}^{m} \varepsilon^j h_j(\cdot) + \mathcal{O}(\varepsilon^{m+1}). & (4.46) \end{aligned}$$

4.2.3 Proof of Theorem 3.3.1 for $m \geq 2$

In this section, we carry out the induction step of the proof. We assume that Eqs. (4.45) and (4.46) hold for $0, 1, \ldots, m$ and prove that they also hold for $m+1$. It suffices to establish Eq. (4.45); Eq. (4.46) follows immediately from Eq. (4.45) and our discussion of the CSP condition (4.44).

Before carrying out the induction step, we derive an update formula for $T_{(m)}$. Using Eq. (4.40) with m replaced by $m+1$, we obtain

$$T_{(m+1)} = B_{(m+1)}\tilde{A}^{(m+1)}. \tag{4.47}$$

(Here, we used the identity $(\tilde{B}_{(m+1)})^{-1} = \tilde{A}^{(m+1)}$.) Next, we use the update formulas (3.12) and (3.17) for $B_{(m+1)}$ and $\tilde{A}^{(m+1)}$, respectively, to rewrite Eq. (4.47),

$$T_{(m+1)} = \left(I - L_{(m)}\right)\left(I + U_{(m)}\right) T_{(m)}(I - \tilde{U}_{(m)}). \tag{4.48}$$

Equation (4.40) also relates $A^{(m+1)}$ to $\tilde{A}^{(m+1)}$,

$$A^{(m+1)} = \tilde{A}^{(m+1)}\left(T_{(m+1)}\right)^{-1}.$$

Taking $C = \left(T_{(m)}\right)^{-1}$ in Eq. (3.16), we express $\Lambda_{(m)}$ in terms of $\tilde{\Lambda}_{(m)}$,

$$\Lambda_{(m)} = T_{(m)}\tilde{\Lambda}_{(m)}\left(T_{(m)}\right)^{-1} - T_{(m)}\frac{d\left(T_{(m)}\right)^{-1}}{dt}$$

or, equivalently,

$$\Lambda_{(m)} = T_{(m)}\tilde{\Lambda}_{(m)}\left(T_{(m)}\right)^{-1} + \left(T_{(m)}\right)^{-1}\frac{dT_{(m)}}{dt}. \tag{4.49}$$

Next, we estimate the blocks of the matrices in Eq. (4.49). The estimate of $T_{(m)}$ is given in the induction hypothesis (4.45); its inverse satisfies a similar estimate,

$$\left(T_{(m)}\right)^{-1}(\,\cdot\,, \psi_{(m)}(\cdot,\varepsilon),\varepsilon) = \begin{pmatrix} I_{N_f} + \mathcal{O}(\varepsilon^2) & -\varepsilon^m T^{12}_{(m,m)} + \mathcal{O}(\varepsilon^{m+1}) \\ -\varepsilon T^{21}_{(m,1)} + \mathcal{O}(\varepsilon^2) & I_{N_s} + \mathcal{O}(\varepsilon^2) \end{pmatrix}. \tag{4.50}$$

Also, the induction hypothesis (4.46) and Theorem 3.3.2 guarantee that $\psi_{(m)} = \tilde{\psi}_{(m)} + \mathcal{O}(\varepsilon^{m+1})$, so the expansions of $\tilde{\Lambda}_{(m)}(y, \tilde{\psi}_{(m)}, \varepsilon)$ and $\tilde{\Lambda}_{(m)}(y, \psi_{(m)}, \varepsilon)$ are equal up to and including terms of $\mathcal{O}(\varepsilon^m)$. It follows from Eq. (4.22) that

$$\tilde{\Lambda}_{(m)}(\,\cdot\,, \psi_{(m)}(\cdot,\varepsilon),\varepsilon) = \begin{pmatrix} \Lambda^{11}_{(0,0)} + \mathcal{O}(\varepsilon) & \varepsilon^m\tilde{\Lambda}^{12}_{(m,m)} + \mathcal{O}(\varepsilon^{m+1}) \\ \varepsilon\Lambda^{21}_{(0,1)} + \mathcal{O}(\varepsilon^2) & \varepsilon\tilde{\Lambda}^{22}_{(m,1)} + \mathcal{O}(\varepsilon^2) \end{pmatrix}. \tag{4.51}$$

Taking $V = T_{(m)}$ in Lemma 9.2.1, we conclude from Eq. (4.45) that

$$DT_{(m)}g = \begin{pmatrix} \mathcal{O}(\varepsilon^3) & \mathcal{O}(\varepsilon^{m+1}) \\ \mathcal{O}(\varepsilon^2) & \mathcal{O}(\varepsilon^3) \end{pmatrix}. \tag{4.52}$$

The desired estimate of $\Lambda_{(m)}$ now follows immediately from Eqs. (4.45), (4.50), (4.51), and (4.52),

$$\begin{aligned}
\Lambda^{11}_{(m)} &= \Lambda^{11}_{(0,0)} + \mathcal{O}(\varepsilon), && (4.53)\\
\Lambda^{12}_{(m)} &= \varepsilon^m[\tilde{\Lambda}^{12}_{(m,m)} - \Lambda^{11}_{(0,0)}T^{12}_{(m,m)}] + \mathcal{O}(\varepsilon^{m+1}), && (4.54)\\
\Lambda^{21}_{(m)} &= \varepsilon[\Lambda^{21}_{(0,1)} + T^{21}_{(m,1)}\Lambda^{11}_{(0,0)}] + \mathcal{O}(\varepsilon^2), && (4.55)\\
\Lambda^{22}_{(m)} &= \varepsilon\tilde{\Lambda}^{22}_{(m,1)} + \mathcal{O}(\varepsilon^2). && (4.56)
\end{aligned}$$

The definition (3.10) and Eqs. (4.53) and (4.54) imply that $U_{(m)} = \mathcal{O}(\varepsilon^m)$, with the leading-order coefficient given by

$$U_{(m,m)} = \left(\Lambda^{11}_{(m,0)}\right)^{-1}\Lambda^{12}_{(m,m)} = \tilde{U}_{(m,m)} - T^{12}_{(m,m)}. \tag{4.57}$$

Furthermore, the definition (3.10) and Eqs. (4.53) and (4.55) imply that

$$L_{(m)} = \Lambda^{21}_{(m)}\left(\Lambda^{11}_{(m)}\right)^{-1} = \mathcal{O}(\varepsilon). \tag{4.58}$$

Finally, we observe that, to leading order, the blocks of $T_{(m)}(\cdot\,, \psi_{(m)}(\cdot, \varepsilon), \varepsilon)$ are all equal to the corresponding blocks of $T_{(m)}(\cdot\,, \psi_{(m-1)}(\cdot, \varepsilon), \varepsilon)$. The latter are given by the induction hypothesis (4.45).

We are now ready to estimate the size of the blocks of $T_{(m+1)}(\cdot\,, \psi_{(m)}, \varepsilon)$. The update formula (4.48) gives $T^{11}_{(m+1)} = T^{11}_{(m)} + U_{(m)}T^{21}_{(m)}$. According to the induction hypothesis, $T^{11}_{(m)} = I_{N_f} + \mathcal{O}(\varepsilon^2)$ and $T^{21}_{(m)} = \mathcal{O}(\varepsilon)$. Furthermore, $U_{(m)} = \mathcal{O}(\varepsilon^m)$, by Eq. (4.57). Thus, $T^{11}_{(m+1)} = I_{N_f} + \mathcal{O}(\varepsilon^2)$, as desired.

The update formula (4.48) also gives

$$T^{12}_{(m+1)} = T^{12}_{(m)} - T^{11}_{(m)}\tilde{U}_{(m)} + U_{(m)}T^{22}_{(m)} - U_{(m)}T^{21}_{(m)}\tilde{U}_{(m)}.$$

According to the induction hypothesis (4.45), $T^{12}_{(m)} = \mathcal{O}(\varepsilon^m)$, $T^{11}_{(m)} = I_{N_f} + \mathcal{O}(\varepsilon^2)$, $T^{21}_{(m)} = \mathcal{O}(\varepsilon)$, and $T^{22}_{(m)} = I_{N_s} + \mathcal{O}(\varepsilon^2)$. Furthermore, $U_{(m)} = \mathcal{O}(\varepsilon^m)$, by Eq. (4.57), and $\tilde{U}_{(m)} = \mathcal{O}(\varepsilon^m)$, by Eq. (4.9). Thus, the terms in the formula for $T^{12}_{(m+1)}$ are all at least $\mathcal{O}(\varepsilon^m)$. The same is then true for $T^{12}_{(m+1)}$. We will now show that $T^{12}_{(m+1)}$ is, in fact, at least $\mathcal{O}(\varepsilon^{m+1})$ by showing that $T^{12}_{(m+1,m)} = 0$. To leading order, the update formula for $T^{12}_{(m+1)}$ is

$$T^{12}_{(m+1,m)} = T^{12}_{(m,m)} - \tilde{U}_{(m,m)} + U_{(m,m)}. \tag{4.59}$$

Equation (4.57) implies that the right member of (4.59) vanishes. Thus, $T_{(m+1,m)} = 0$, as desired. We emphasize again that the choice of $U_{(m)}$ is central to the working of the CSP method.

Next, the update formula (4.48) gives $T^{21}_{(m+1)} = T^{21}_{(m)} - L_{(m)}U_{(m)}T^{21}_{(m)} - L_{(m)}T^{11}_{(m)}$. According to the induction hypothesis, $T^{21}_{(m)} = \mathcal{O}(\varepsilon)$ and $T^{11}_{(m)} = I_n + \mathcal{O}(\varepsilon^2)$. Furthermore, $U_{(m)} = \mathcal{O}(\varepsilon^m)$ and $L_{(m)} = \mathcal{O}(\varepsilon)$, by Eqs. (4.57) and (4.58). Thus, the terms in the update formula for $T^{21}_{(m+1)}$ are all at least $\mathcal{O}(\varepsilon)$. Hence, $T^{21}_{(m)}$ is also at least $\mathcal{O}(\varepsilon)$, as desired.

Lastly, the update formula (4.48) gives

$$\begin{aligned} T^{22}_{(m+1)} &= T^{22}_{(m)} - L_{(m)}T^{12}_{(m)} - T^{21}_{(m)}\tilde{U}_{(m)} + L_{(m)}T^{11}_{(m)}\tilde{U}_{(m)} \\ &\quad - L_{(m)}U_{(m)}T^{22}_{(m)} + L_{(m)}U_{(m)}T^{21}_{(m)}\tilde{U}_{(m)}. \end{aligned}$$

According to the induction hypothesis, $T^{22}_{(m)} = I_{N_s} + \mathcal{O}(\varepsilon^2)$. The remaining terms have already been shown to be at least $\mathcal{O}(\varepsilon^2)$. Hence, $T^{22}_{(m+1)} = I_{N_s} + \mathcal{O}(\varepsilon^2)$.

The proof of Theorem 3.3.1 is complete. ∎

Chapter 5

Proof Of Theorem 3.3.3

5.1 Proof of Theorem 3.3.3

The proof of Theorem 3.3.3 is by induction on m. Section 5.1.1 contains an auxiliary lemma that shows that each successive application of the CSP algorithm pushes Λ closer to block-diagonal form. The induction hypothesis is formulated in section 5.1.2, the hypothesis is shown to be true for $m = 0$ in section 5.1.3, and the induction step is taken in section 5.1.4. In Section 5.2, we discuss two ways in which tangent information about the fast fibers can be used to project initial conditions on the slow manifold to obtain a reduced model. In Section 5.3, we discuss a variant of the CSP method and show that it achieves the same accuracy as the full CSP method in constructing the tangent spaces to the fast fibers.

5.1.1 Asymptotic estimates of Λ

As stated in Chapter 3, the goal of the CSP method is to reduce Λ to block-diagonal form. This goal is approached by the repeated application of a two-step algorithm. In Chapter 4, we showed that the first step of the algorithm is engineered so that each application increases the asymptotic accuracy of the upper-right block $\Lambda^{12}_{(m)}$ by one order of ε; in particular, $\Lambda^{12}_{(m)} = \mathcal{O}(\varepsilon^m)$ on $\mathcal{K}^{(m)}_\varepsilon$, see Eq. (4.22). We now complete the picture and show that each application of the second step increases the asymptotic accuracy of the lower-left block $\Lambda^{21}_{(m)}$ by one order of ε when the information obtained in the first step of the same iteration is used. In particular, $\Lambda^{21}_{(m)} = \mathcal{O}(\varepsilon^{m+1})$ on $\mathcal{K}^{(m+1)}_\varepsilon$, where $\mathcal{K}^{(m+1)}_\varepsilon$ has been obtained in the first step of the $(m+1)$th refinement.

Lemma 5.1.1 *For $m = 0, 1, \ldots,$*

$$\Lambda_{(m)} = \begin{pmatrix} \Lambda^{11}_{(0,0)} + \mathcal{O}(\varepsilon) & \varepsilon^m \Lambda^{12}_{(m,m)} \\ \varepsilon^{m+1}\Lambda^{21}_{(m,m+1)} & \varepsilon\Lambda^{22}_{(1,1)} + \mathcal{O}(\varepsilon^2) \end{pmatrix}, \tag{5.1}$$

when $\Lambda_{(m)}$ is evaluated on $\mathcal{K}_\varepsilon^{(m+1)}$.

Proof. The proof is by induction. The desired estimates of $\Lambda^{11}_{(m)}$, $\Lambda^{12}_{(m)}$, and $\Lambda^{22}_{(m)}$ on $\mathcal{K}_\varepsilon^{(m)}$ were established in Eqs. (4.53), (4.54), and (4.56). Since the asymptotic expansions of $\mathcal{K}_\varepsilon^{(m+1)}$ and $\mathcal{K}_\varepsilon^{(m)}$ differ only at terms of $\mathcal{O}(\varepsilon^{m+1})$ or higher by Theorem 3.3.1, these estimates of $\Lambda^{11}_{(m)}$, $\Lambda^{12}_{(m)}$, and $\Lambda^{22}_{(m)}$ are true also on $\mathcal{K}_\varepsilon^{(m+1)}$. It remains only to estimate $\Lambda^{21}_{(m)}$.

Consider the case $m = 0$. Let $\Lambda^{21}_{(0,j)}$ be the coefficient of ε^j in the asymptotic expansion of $\Lambda^{21}_{(0)}(\cdot, \psi_{(1)}(\cdot,\varepsilon),\varepsilon)$. The estimate $\Lambda^{21}_{(0)} = \mathcal{O}(\varepsilon)$ on $\mathcal{K}_\varepsilon^{(1)}$ follows if we can show that $\Lambda^{21}_{(0,0)} = 0$. It is already stated in [49, Eq. (4.30)] that $\Lambda^{21}_{(0,0)} = 0$ on $\mathcal{K}_\varepsilon^{(0)}$. Furthermore, Theorem 3.3.1 implies that the asymptotic expansions of $\psi_{(1)}$ and $\psi_{(0)}$ agree to leading order. Thus, the asymptotic expansions of $\Lambda^{21}_{(0)}(\cdot, \psi_{(0)}(\cdot,\varepsilon),\varepsilon)$ and $\Lambda^{21}_{(0)}(\cdot, \psi_{(1)}(\cdot,\varepsilon),\varepsilon)$ also agree to leading order, and the result follows.

Now, assume that the estimate holds for $0, 1, \ldots, m$. From Eq. (3.16) we obtain

$$\begin{aligned}\Lambda^{21}_{(m+1)} &= \Lambda^{21}_{(m)} - L_{(m)}\Lambda^{11}_{(m)} + \Lambda^{22}_{(m)}L_{(m)} - L_{(m)}\Lambda^{12}_{(m)}L_{(m)} - \Lambda^{21}_{(m)}U_{(m)}L_{(m)} \\ &\quad - L_{(m)}U_{(m)}\Lambda^{21}_{(m)} + L_{(m)}\Lambda^{11}_{(m)}U_{(m)}L_{(m)} - L_{(m)}U_{(m)}\Lambda^{22}_{(m)}L_{(m)} \\ &\quad + L_{(m)}U_{(m)}\Lambda^{21}_{(m)}U_{(m)}L_{(m)} + \left(DL_{(m)}\right)g + L_{(m)}\left(\left(DU_{(m)}\right)g\right)L_{(m)}. \quad (5.2)\end{aligned}$$

The first two terms in the right member sum to zero by virtue of the definition (3.10) of $L_{(m)}$. The next seven terms are all $\mathcal{O}(\varepsilon^{m+2})$ or higher by virtue of the induction hypothesis. Finally, the last two terms are also $\mathcal{O}(\varepsilon^{m+2})$ or higher by the induction hypothesis and Lemma 9.2.1. ∎

5.1.2 The induction hypothesis

The CSPF of order m, $\mathcal{L}_\varepsilon^{(m)}(x^1)$, is defined in Eq. (3.26) to be the linear space spanned by the columns of the fast component, $A_1^{(m)}(x^1, \psi_{(m)}(x^1,\varepsilon),\varepsilon)$, of the basis $A^{(m)}$. Thus, to prove Theorem 3.3.3, it suffices to show that the asymptotic expansions of the space spanned by the columns of $(A_1^{(m)}(x^1, \psi_{(m)}(x^1,\varepsilon),\varepsilon))$ and the space tangent to the fast fiber, $\mathcal{T}_p\mathcal{F}_\varepsilon$, agree up to and including terms of $\mathcal{O}(\varepsilon^m)$ for $p = (x^1, h_\varepsilon(x^1))$ and for $m = 0, 1, \ldots$. The central idea of the proof is to show that each successive application of the CSP method pushes the projection of $A_1^{(m)}$ on $\mathcal{TM}_\varepsilon$ along $\mathcal{TF}_\varepsilon$ to one higher order in ε.

We express $A^{(m)}$, generated after m applications of the CSP algorithm, in terms of the basis A,

$$A^{(m)}(x,\varepsilon) = A(x^1, h_\varepsilon(x^1),\varepsilon)Q^{(m)}(x,\varepsilon), \quad m = 0, 1, \ldots . \quad (5.3)$$

Since $B_{(m)}$ and B are the left inverses of $A^{(m)}$ and A, respectively, we also have

$$B_{(m)}(x,\varepsilon) = R_{(m)}(x,\varepsilon)B(x^1, h_\varepsilon(x^1),\varepsilon), \quad m = 0,1,\ldots, \tag{5.4}$$

where $R_{(m)} \equiv (Q^{(m)})^{-1}$. Introducing the block structure of $Q^{(m)}$ and $R_{(m)}$,

$$Q^{(m)} = \begin{pmatrix} Q^{(m)}_{1f} & Q^{(m)}_{2f} \\ Q^{(m)}_{1s} & Q^{(m)}_{2s} \end{pmatrix}, \quad R_{(m)} = \begin{pmatrix} R^{1s\perp}_{(m)} & R^{1f\perp}_{(m)} \\ R^{2s\perp}_{(m)} & R^{2f\perp}_{(m)} \end{pmatrix}, \tag{5.5}$$

we rewrite Eqs. (5.3) and (5.4) as

$$A^{(m)}_1 = A_f Q^{(m)}_{1f} + A_s Q^{(m)}_{1s}, \quad A^{(m)}_2 = A_f Q^{(m)}_{2f} + A_s Q^{(m)}_{2s} \tag{5.6}$$

and

$$B^1_{(m)} = R^{1s\perp}_{(m)} B^{s\perp} + R^{1f\perp}_{(m)} B^{f\perp}, \quad B^2_{(m)} = R^{2s\perp}_{(m)} B^{s\perp} + R^{2f\perp}_{(m)} B^{f\perp} \tag{5.7}$$

for $m = 0, 1, \ldots$.

Equation (5.7) shows that $A_s Q^{(m)}_{1s}$ is the projection of $A^{(m)}_1$ on $\mathcal{TM}_\varepsilon$. Thus, to establish Theorem 3.3.3, we need only to prove the asymptotic estimate $Q^{(m)}_{1s} = \mathcal{O}(\varepsilon^{m+1})$. The proof is by induction on m, where the induction hypothesis is

$$Q^{(m)}(\cdot\,, \psi_{(m)}(\cdot,\varepsilon),\varepsilon) = \begin{pmatrix} \mathcal{O}(1) & \mathcal{O}(\varepsilon^m) \\ \mathcal{O}(\varepsilon^{m+1}) & \mathcal{O}(1) \end{pmatrix}, \tag{5.8}$$

$$R_{(m)}(\cdot\,, \psi_{(m)}(\cdot,\varepsilon),\varepsilon) = \begin{pmatrix} \mathcal{O}(1) & \mathcal{O}(\varepsilon^m) \\ \mathcal{O}(\varepsilon^{m+1}) & \mathcal{O}(1) \end{pmatrix}, \quad m = 0,1,\ldots. \tag{5.9}$$

Remark. Although the estimate of $Q^{(m)}_{1s}$ is sufficient to establish Theorem 3.3.3, we provide the estimates of all the blocks appearing in Eqs. (5.8)–(5.9) because they will be required in the induction step.

5.1.3 Proof of Theorem 3.3.3 for $m = 0$

We fix $m = 0$ and verify the induction hypothesis for $Q^{(0)}$ and $R_{(0)}$. By Eq. (5.3),

$$Q^{(0)} = BA^{(0)}, \tag{5.10}$$

whence

$$Q^{(0)} = \begin{pmatrix} B^{s\perp}A^{(0)}_1 & B^{s\perp}A^{(0)}_2 \\ B^{f\perp}A^{(0)}_1 & B^{f\perp}A^{(0)}_2 \end{pmatrix}. \tag{5.11}$$

It suffices to show that the lower-left block is zero to leading order, since the other blocks are all $\mathcal{O}(1)$. We do this by showing that $Q^{(0,0)}_{1s} = 0$. By Eq. (5.11),

$$Q^{(0,0)}_{1s} = B^{f\perp}_0 A^{(0,0)}_1. \tag{5.12}$$

$B_0^{f\perp}$ spans $\mathcal{N}_p\mathcal{F}_0$ for every $p \in \mathcal{K}_\varepsilon^{(0)}$. Also, z is constant on $\mathcal{N}_p\mathcal{F}_0$, so $B_0^{f\perp} = (B^{1f\perp}, 0)$, where $B^{1f\perp}$ is a full-rank matrix of size m. Last, $A_1^{(0,0)} = A_1^{(0)} = \begin{pmatrix} 0 \\ A_{21}^{(0)} \end{pmatrix}$ by Eq. (3.7). Substituting these expressions for $B_0^{f\perp}$ and $A_1^{(0,0)}$ into Eq. (5.12), we obtain that $Q_{1s}^{(0,0)} = 0$.

The induction hypothesis on $R_{(0)}$ can be verified either by a similar argument or by recalling that $R_{(0)} = (Q^{(0)})^{-1}$, where $Q^{(0)}$ was shown above to be block-triangular to leading order.

5.1.4 Proof of Theorem 3.3.3 for $m \geq 1$

We assume that the induction hypothesis (5.8)–(5.9) holds for $0, 1, \ldots, m$ and show that it holds for $m+1$. The proof proceeds in four steps. In Step 1, we derive explicit expressions for $R_{(m+1)}$ and $Q^{(m+1)}$ in terms of $R_{(m)}$ and $Q^{(m)}$; these expressions also involve $U_{(m)}$ and $L_{(m)}$. In Step 2, we derive the leading-order asymptotics of $U_{(m)}$ and in Step 3 the leading-order asymptotics of $L_{(m)}$. Then, in Step 4, we substitute these results into the expressions derived in Step 1 to complete the induction.

Step 1. We derive the expressions for $Q^{(m+1)}$ and $R_{(m+1)}$. Equations (5.3) and (5.4), together with the update formulas (3.11) for $A^{(m)}$ and (3.12) for $B_{(m)}$, yield

$$\begin{aligned} Q^{(m+1)} &= Q^{(m)}(I - U_{(m)})(I + L_{(m)}), && (5.13) \\ R_{(m+1)} &= (I - L_{(m)})(I + U_{(m)})R_{(m)}. && (5.14) \end{aligned}$$

In terms of the constituent blocks, we have

$$\begin{aligned} Q_{1f}^{(m+1)} &= Q_{1f}^{(m)} + Q_{2f}^{(m)}L_{(m)} - Q_{1f}^{(m)}U_{(m)}L_{(m)}, && (5.15) \\ Q_{2f}^{(m+1)} &= Q_{2f}^{(m)} - Q_{1f}^{(m)}U_{(m)}, && (5.16) \\ Q_{1s}^{(m+1)} &= Q_{1s}^{(m)} + Q_{2s}^{(m)}L_{(m)} - Q_{1s}^{(m)}U_{(m)}L_{(m)}, && (5.17) \\ Q_{2s}^{(m+1)} &= Q_{2s}^{(m)} - Q_{1s}^{(m)}U_{(m)} && (5.18) \end{aligned}$$

and

$$\begin{aligned} R_{(m+1)}^{1s\perp} &= R_{(m)}^{1s\perp} + U_{(m)}R_{(m)}^{2s\perp}, && (5.19) \\ R_{(m+1)}^{1f\perp} &= R_{(m)}^{1f\perp} + U_{(m)}R_{(m)}^{2f\perp}, && (5.20) \\ R_{(m+1)}^{2s\perp} &= R_{(m)}^{2s\perp} - L_{(m)}R_{(m)}^{1s\perp} - L_{(m)}U_{(m)}R_{(m)}^{2s\perp}, && (5.21) \\ R_{(m+1)}^{2f\perp} &= R_{(m)}^{2f\perp} - L_{(m)}R_{(m)}^{1f\perp} - L_{(m)}U_{(m)}R_{(m)}^{2f\perp}. && (5.22) \end{aligned}$$

Step 2. We derive the leading-order asymptotics of the matrix $U_{(m)}$.

Recall that $U_{(m)} = (\Lambda^{11}_{(m)})^{-1}\Lambda^{12}_{(m)}$. Moreover, $\Lambda^{11}_{(m)}$ is strictly $\mathcal{O}(1)$ and $\Lambda^{12}_{(m)}$ is strictly $\mathcal{O}(\varepsilon^m)$ by Lemma 5.1.1. Hence, $U_{(m)} = U_{(m,m)}\varepsilon^m + \mathcal{O}(\varepsilon^{m+1})$, with $U_{(m,m)} = (\Lambda^{11}_{(m,0)})^{-1}\Lambda^{12}_{(m,m)}$. Therefore, it suffices to derive the leading-order asymptotics of these blocks of Λ.

By definition, $\Lambda_{(m)} = B_{(m)}[A^{(m)}, g]$. Therefore,

$$\Lambda_{(m)} = \begin{pmatrix} B^1_{(m)}[A^{(m)}_1, g] & B^1_{(m)}[A^{(m)}_2, g] \\ B^2_{(m)}[A^{(m)}_1, g] & B^2_{(m)}[A^{(m)}_2, g] \end{pmatrix}. \tag{5.23}$$

The individual blocks of $\Lambda_{(m)}$ are obtained by substituting Eqs. (5.6) and (5.7) into (5.23). We observe that one-half of all the terms would vanish, were they to be evaluated on $\mathcal{M}_\varepsilon$, by virtue of Lemma 2.3.1. Since they are evaluated on $\mathcal{K}^{(m+1)}_\varepsilon$, instead, which is $\mathcal{O}(\varepsilon^{m+1})-$close to $\mathcal{M}_\varepsilon$, these terms are $\mathcal{O}(\varepsilon^{m+2})$ and therefore of higher order for each of the blocks; recall Lemma 5.1.1. Thus,

$$\Lambda^{11}_{(m)} = R^{1s\perp}_{(m)}B^{s\perp}[A_f Q^{(m)}_{1f}, g] + R^{1f\perp}_{(m)}B^{f\perp}[A_s Q^{(m)}_{1s}, g], \tag{5.24}$$

$$\Lambda^{12}_{(m)} = R^{1s\perp}_{(m)}B^{s\perp}[A_f Q^{(m)}_{2f}, g] + R^{1f\perp}_{(m)}B^{f\perp}[A_s Q^{(m)}_{2s}, g], \tag{5.25}$$

$$\Lambda^{21}_{(m)} = R^{2s\perp}_{(m)}B^{s\perp}[A_f Q^{(m)}_{1f}, g] + R^{2f\perp}_{(m)}B^{f\perp}[A_s Q^{(m)}_{1s}, g], \tag{5.26}$$

where the remainders of $\mathcal{O}(\varepsilon^{m+2})$ have been omitted for brevity. Recalling the definition of the Lie bracket, we rewrite Eq. (5.24) as

$$\begin{aligned} \Lambda^{11}_{(m)} = {} & R^{1s\perp}_{(m)}B^{s\perp}\left((Dg)A_f Q^{(m)}_{1f} - \frac{d}{dt}\left(A_f Q^{(m)}_{1f}\right)\right) \\ & + R^{1f\perp}_{(m)}B^{f\perp}\left((Dg)A_s Q^{(m)}_{1s} - \frac{d}{dt}\left(A_s Q^{(m)}_{1s}\right)\right), \end{aligned} \tag{5.27}$$

where we recall that all of the quantities are evaluated at $(\cdot, \psi_{(m+1)}(\cdot, \varepsilon), \varepsilon)$. Next, $(Dg)A_s$ and the two time derivatives in Eq. (5.27) are zero to leading order by Lemmas 9.3.1 and 9.2.1, respectively. Therefore, to leading order Eq. (5.27) becomes

$$\Lambda^{11}_{(m,0)} = R^{1s\perp}_{(m,0)}B^{s\perp}_0(Dg)_0 A^0_f Q^{(m,0)}_{1f}. \tag{5.28}$$

Here, $\Lambda^{11}_{(m,0)}$ stands for the leading-order term in the asymptotic expansion of $\Lambda^{11}_{(m)}(\cdot, \psi_{(m+1)}(y)(\cdot, \varepsilon), \varepsilon)$, and the right member is the leading order term in the asymptotic expansion of $(R^{1s\perp}_{(m)}B^{s\perp}(Dg)A_f Q^{(m)}_{1f})(\cdot, h_\varepsilon(\cdot), \varepsilon)$.

We derive a similar formula for $\Lambda^{12}_{(m,m)}$. First, we rewrite Eq. (5.25) as

$$\begin{aligned} \Lambda^{12}_{(m)} = {} & R^{1s\perp}_{(m)}B^{s\perp}\left((Dg)A_f Q^{(m)}_{2f} - \frac{d}{dt}\left(A_f Q^{(m)}_{2f}\right)\right) \\ & + R^{1f\perp}_{(m)}B^{f\perp}\left((Dg)A_s Q^{(m)}_{2s} - \frac{d}{dt}\left(A_s Q^{(m)}_{2s}\right)\right). \end{aligned} \tag{5.29}$$

Next, $Q_{2f}^{(m)} = \mathcal{O}(\varepsilon^m)$, $Q_{2s}^{(m)} = \mathcal{O}(1)$, $R_{(m)}^{1s\perp} = \mathcal{O}(1)$, and $R_{(m)}^{1f\perp} = \mathcal{O}(\varepsilon^m)$ by the induction hypothesis (5.8)–(5.9). Thus, Lemma 9.2.1 implies that the two terms in Eq. (5.29) involving time derivatives are $\mathcal{O}(\varepsilon^{m+1})$ and therefore of higher order. Also, $(Dg)A_s$ is zero to leading order by Lemma 9.3.1, and thus

$$\Lambda_{(m,m)}^{12} = R_{(m,0)}^{1s\perp} B_0^{s\perp} (Dg)_0 A_f^0 Q_{2f}^{(m,m)}. \tag{5.30}$$

We now substitute $\Lambda_{(m,0)}^{11}$ and $\Lambda_{(m,m)}^{12}$ from Eqs. (5.28) and (5.30) in the expression $U_{(m,m)} = (\Lambda_{(m,0)}^{11})^{-1}\Lambda_{(m,m)}^{12}$ to find the desired expression for $U_{(m,m)}$ in terms of $Q^{(m)}$,

$$U_{(m,m)} = \left(Q_{1f}^{(m,0)}\right)^{-1} Q_{2f}^{(m,m)}. \tag{5.31}$$

We also need an expression for $U_{(m,m)}$ in terms of blocks of $R_{(m)}$, which we will use in Eqs. (5.19)–(5.22). Since $R_{(m)}$ has the near block-diagonal structure given by the induction hypothesis (5.8)–(5.9) and $Q^{(m)}$ is its inverse, we find that

$$Q^{(m)} = \begin{pmatrix} (R_{(m,0)}^{1s\perp})^{-1} & -\varepsilon^m (R_{(m,0)}^{1s\perp})^{-1} R_{(m,m)}^{1f\perp} (R_{(m,0)}^{2f\perp})^{-1} \\ -\varepsilon^{m+1} (R_{(m,0)}^{2f\perp})^{-1} R_{(m,m+1)}^{2s\perp} (R_{(m,0)}^{1s\perp})^{-1} & (R_{(m,0)}^{2f\perp})^{-1} \end{pmatrix}$$

to leading order for each of the blocks and for $m = 1, 2, \ldots$. This equation, in conjunction with Eq. (5.31), leads to the desired expression for $U_{(m,m)}$ in terms of $R_{(m)}$,

$$U_{(m,m)} = -R_{(m,m)}^{1f\perp} \left(R_{(m,0)}^{2f\perp}\right)^{-1}. \tag{5.32}$$

Step 3. We derive the leading-order asymptotics of the matrix $L_{(m)}$.

Recall that $L_{(m)} = \Lambda_{(m)}^{21}(\Lambda_{(m)}^{11})^{-1}$. Moreover, by Lemma 5.1.1, $\Lambda_{(m)}^{11}$ is strictly $\mathcal{O}(1)$ and $\Lambda_{(m)}^{21}$ is strictly $\mathcal{O}(\varepsilon^{m+1})$. Hence, $L_{(m)} = L_{(m,m+1)}\varepsilon^{m+1} + \mathcal{O}(\varepsilon^{m+2})$, with $L_{(m,m+1)} = \Lambda_{(m,m+1)}^{21}(\Lambda_{(m,0)}^{11})^{-1}$. An expression for $\Lambda_{(m,0)}^{11}$ was derived in Eq. (5.28), so here we focus on $\Lambda_{(m,m+1)}^{21}$.

Equation (5.26) and the definition of the Lie bracket imply that

$$\begin{aligned} \Lambda_{(m)}^{21} &= R_{(m)}^{2s\perp} B^{s\perp} \left((Dg) A_f Q_{1f}^{(m)} - \frac{d}{dt}\left(A_f Q_{1f}^{(m)}\right)\right) \\ &\quad + R_{(m)}^{2f\perp} B^{f\perp} \left((Dg) A_s Q_{1s}^{(m)} - \frac{d}{dt}\left(A_s Q_{1s}^{(m)}\right)\right). \end{aligned} \tag{5.33}$$

Next, $Q_{1f}^{(m)} = \mathcal{O}(1)$, $Q_{1s}^{(m)} = \mathcal{O}(\varepsilon^{m+1})$, $R_{(m)}^{2s\perp} = \mathcal{O}(\varepsilon^{m+1})$, and $R_{(m)}^{2f\perp} = \mathcal{O}(1)$ by the induction hypothesis. Also, the time derivatives are $\mathcal{O}(\varepsilon)$ by Lemma 9.2.1, and thus the two terms in Eq. (5.33) that involve time derivatives are $\mathcal{O}(\varepsilon^{m+2})$. Last, $(Dg)A_s = \mathcal{O}(\varepsilon)$ by Lemma 9.3.1. Thus, we find that

$$\Lambda_{(m,m+1)}^{21} = R_{(m,m+1)}^{2s\perp} B_0^{s\perp} (Dg)_0 A_f^0 Q_{1f}^{(m,0)}. \tag{5.34}$$

Equations (5.28) and (5.34) yield the desired formula for $L_{(m,m+1)}$ in terms of the blocks of $R_{(m)}$,

$$L_{(m,m+1)} = \Lambda^{21}_{(m,m+1)} \left(\Lambda^{11}_{(m,0)}\right)^{-1} = R^{2s\perp}_{(m,m+1)} \left(R^{1s\perp}_{(m,0)}\right)^{-1}. \tag{5.35}$$

Next, we recast Eq. (5.35) in terms of blocks of $Q^{(m)}$ in order to use it in Eqs. (5.15)–(5.18). The matrix $R_{(m)}$ is the inverse of $Q^{(m)}$ and has the near block-diagonal form given in Eq. (5.9). Thus,

$$R_{(m)} = \begin{pmatrix} (Q^{(m,0)}_{1f})^{-1} & -\varepsilon^m (Q^{(m,0)}_{1f})^{-1} Q^{(m,m)}_{2f} (Q^{(m,0)}_{2s})^{-1} \\ -\varepsilon^{m+1} (Q^{(m,0)}_{2s})^{-1} Q^{(m,m+1)}_{1s} (Q^{(m,0)}_{1f})^{-1} & (Q^{(m,0)}_{2s})^{-1} \end{pmatrix}$$

to leading order for each block and for $m = 1, 2, \ldots$. This equation, in conjunction with Eq. (5.35), leads to the desired expression for $L_{(m,m+1)}$ in terms of the blocks of $Q^{(m)}$,

$$L_{(m,m+1)} = -\left(Q^{(m,0)}_{2s}\right)^{-1} Q^{(m,m+1)}_{1s}. \tag{5.36}$$

Step 4. We substitute the results obtained in Steps 2 and 3 into the formulas (5.15)–(5.22) derived in Step 1.

Equations (5.15) and (5.18), together with the induction hypothesis and the estimates $U_{(m)} = \mathcal{O}(\varepsilon^m)$ and $L_{(m)} = \mathcal{O}(\varepsilon^{m+1})$, imply that $Q^{(m+1)}_{1f}$ and $Q^{(m+1)}_{2s}$ remain $\mathcal{O}(1)$. This concludes the estimation of these blocks.

Next, we show that $Q^{(m+1)}_{2f} = \mathcal{O}(\varepsilon^{m+1})$. First, $Q^{(m+1)}_{2f}$ and $Q^{(m)}_{2f}$ are equal up to and including terms of $\mathcal{O}(\varepsilon^{m-1})$ by Eq. (5.16) and the estimate on $U_{(m)}$. Thus, $Q^{(m+1,i)}_{2f} = 0$ for $i = 0, 1, \ldots, m-1$ by the induction hypothesis on $Q^{(m)}_{2f}$. It suffices to show that $Q^{(m+1,m)}_{2f} = 0$. Equation (5.16) implies that

$$Q^{(m+1,m)}_{2f} = Q^{(m,m)}_{2f} - Q^{(m,0)}_{1f} U_{(m,m)}.$$

The right member of this equation is zero by Eq. (5.31), and the estimation of $Q^{(m+1)}_{2f}$ is complete.

Finally, we show that $Q^{(m+1)}_{1s} = \mathcal{O}(\varepsilon^{m+2})$ to complete the estimates on the blocks of $Q^{(m+1)}$. First, $Q^{(m+1)}_{1s}$ and $Q^{(m)}_{1s}$ are equal up to and including terms of $\mathcal{O}(\varepsilon^m)$ by Eq. (5.17) and the order estimates on $U_{(m)}$ and $L_{(m)}$. Thus, $Q^{(m+1,i)}_{1s} = 0$ for $i = 0, 1, \ldots, m$ by the induction hypothesis on $Q^{(m)}_{1s}$. It suffices to show that $Q^{(m+1,m+1)}_{1s} = 0$. Equation (5.17) implies that

$$Q^{(m+1,m+1)}_{1s} = Q^{(m,m+1)}_{1s} + Q^{(m,0)}_{2s} L_{(m,m+1)},$$

where the right member of this equation is zero by Eq. (5.36). The estimation of $Q^{(m+1)}_{1s}$ is complete.

The blocks of $R_{(m)}$ may be estimated in a similar manner, using Eqs. (5.19)–(5.22) instead of (5.15)–(5.18) and (5.32) and (5.35) instead of (5.31) and (5.36). The proof of Theorem 3.3.3 is complete. ∎

5.2 Linear projection of initial conditions

The main result in this chapter, Theorem 3.3.3, states that after m iterations the CSP method successfully identifies $\mathcal{TF}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^{m+1})$, where this approximation is given explicitly by $A_1^{(m)}$. This information is postprocessed to project the initial conditions on the CSPM of order m. In this section, we discuss the accuracy and limitations of this linear projection.

Geometrically, one knows from Fenichel's theory that any given initial condition x_0 sufficiently close to $\mathcal{M}_\varepsilon$ lies on a (generally nonlinear) fiber $\mathcal{F}_\varepsilon^p$ with base point p on $\mathcal{M}_\varepsilon$. Hence, the ideal projection would be $\pi_F(x_0) = p$ (the subscript F stands for fiber or Fenichel), and this is, in general, a nonlinear projection.

Within the framework of an algorithm that yields only linearized information about the fast fibers, one must ask how best to approximate this ideal. A consistent approach is to identify a point on the slow manifold such that the approximate linearized fiber through it also goes through the given initial condition. This approach was used, for example, by Roberts [39] for systems with asymptotically stable center manifolds, where we note that a different method is first used to approximate the center manifold. Also, this approach is exact in the special case that the perturbed fast fibers are hyperplanes which need not be vertical. In general, if x_0 lies on the linearized fiber $\mathcal{L}_\varepsilon^{p_1}$ and if $\pi_F(x_0) = p_2$, then the error $\|p_1 - p_2\|$ made by projecting linearly is $\mathcal{O}(\varepsilon)$ and proportional to the curvature of the fiber (see also [39]).

For fast-slow systems, there is yet another way to linearly project initial conditions on the slow manifold. One projects along the approximate CSPF to the space $\mathcal{T}_p\mathcal{F}_\varepsilon$, where p is the point on the CSPM that lies on the same $\varepsilon = 0$ fiber as the initial condition. This type of projection is also consistent, in the sense that it yields an exact result for $\varepsilon = 0$ but has an error of $\mathcal{O}(\varepsilon)$ for $\varepsilon > 0$. Moreover, it is algorithmically simpler, since it does not involve a search for the base point of the linearized fiber on which the initial conditions lie. However, it has the disadvantage that the projection is not exact in the special case that the fast fibers are (nonvertical) hyperplanes.

5.3 The CSPFs in a variant of the CSP method

In section 3.3.2, we emphasized that when we construct the CSPFs of order m at the second step of the mth iteration we use the information obtained in the first step of

the same iteration. In particular, we evaluate $A_1^{(m)}$ on $\mathcal{K}_\varepsilon^{(m)}$ and define the CSPF of order m to be its span; see Eq. (3.26).

In this section we examine the asymptotic accuracy of a variant of the CSP method, where we evaluate $A_1^{(m)}$ on $\mathcal{K}_\varepsilon^{(m-1)}$, so that the CSP quantities updated in the mth iteration are all evaluated on the same manifold, namely on the CSPM of order $m-1$. We show that this modification does not reduce the asymptotic accuracy of the CSPFs.

The proof is by induction. We assume that the variant of the CSP method described above yields, at the mth iteration, CSPFs of asymptotic accuracy $\mathcal{O}(\varepsilon^m)$ for $0, 1, \ldots, m$, and we show that the same is true for $m+1$. Recall that the CSPFs of order $m+1$ are constructed at the second step of the $(m+1)$th iteration and that this step is carried out via the update matrix $L_{(m)}$. The idea behind the proof is to show that the modifications introduced in $L_{(m)}$ by replacing $\mathcal{K}_\varepsilon^{(m+1)}$ by $\mathcal{K}_\varepsilon^{(m)}$ are of $\mathcal{O}(\varepsilon^{m+2})$, although $\mathcal{K}_\varepsilon^{(m+1)}$ and $\mathcal{K}_\varepsilon^{(m)}$ differ at terms of $\mathcal{O}(\varepsilon^{m+1})$. (This property can be attributed to the fact that, at each iteration, $L_{(m)}$ is zero to leading order; see Lemma 5.1.1.) Since $L_{(m)}$ is $\mathcal{O}(\varepsilon^{m+1})$ on $\mathcal{K}_\varepsilon^{(m+1)}$ by Lemma 5.1.1, we conclude that $L_{(m)}$ stays unaltered to leading order when evaluated on $\mathcal{K}_\varepsilon^{(m)}$, instead of on $\mathcal{K}_\varepsilon^{(m+1)}$, and thus the CSPFs of order $m+1$, $\mathcal{L}_\varepsilon^{(m+1)}$, retain their asymptotic accuracy of $\mathcal{O}(\varepsilon^{m+1})$.

By definition (3.10), $L_{(m)} \equiv \Lambda_{(m)}^{21}(\Lambda_{(m)}^{11})^{-1}$, and thus Lemma 5.1.1 implies that the leading order coefficient in the asymptotic expansion of $L_{(m)}$ is

$$L_{(m,m+1)} \equiv \Lambda_{(m,m+1)}^{21}(\Lambda_{(m,0)}^{11})^{-1}.$$

Clearly, the modifications introduced in $\Lambda_{(m)}^{11}$ when it is evaluated on $\mathcal{K}_\varepsilon^{(m)}$ instead of on $\mathcal{K}_\varepsilon^{(m+1)}$ are $\mathcal{O}(\varepsilon^{m+1})$, and thus $\Lambda_{(m,0)}^{11}$ remains unaltered. Thus, it remains to determine the effect that the replacement of $\mathcal{K}_\varepsilon^{(m+1)}$ by $\mathcal{K}_\varepsilon^{(m)}$ has on $\Lambda_{(m,m+1)}^{21}$. As stated in the paragraph following Eq. (5.23), the two *homogeneous* terms $R_{(m)}^{2f\perp}B^{f\perp}[A_fQ_{1f}^{(m)}, g]$ and $R_{(m)}^{2s\perp}B^{s\perp}[A_sQ_{1s}^{(m)}, g]$ that appear in the formula for $\Lambda_{21}^{(m)}$ are identically zero, when they are evaluated on $\mathcal{M}_\varepsilon$. In the context of the CSP method, these terms are evaluated on $\mathcal{K}_\varepsilon^{(m+1)}$, which is $\mathcal{O}(\varepsilon^{m+1})-$close to $\mathcal{M}_\varepsilon$, and thus they are at least of $\mathcal{O}(\varepsilon^{m+2})$ and therefore negligible to leading order. In the current context, and since $\mathcal{K}_\varepsilon^{(m)}$ and $\mathcal{M}_\varepsilon$ are only $\mathcal{O}(\varepsilon^m)-$close, it is not immediate that these terms are $\mathcal{O}(\varepsilon^{m+2})$, and they need to be included in Eq. (5.26). Thus, this equation must be replaced by

$$\begin{aligned}\Lambda_{21}^{(m)} &= R_{(m)}^{2s\perp}B^{f\perp}[A_fQ_{1f}^{(m)}, g] + R_{(m)}^{2f\perp}B^{s\perp}[A_sQ_{1s}^{(m)}, g] \\ &\quad + R_{(m)}^{2f\perp}B^{f\perp}[A_fQ_{1f}^{(m)}, g] + R_{(m)}^{2s\perp}B^{s\perp}[A_sQ_{1s}^{(m)}, g], \qquad (5.37)\end{aligned}$$

and each term in the right member of Eq. (5.37) must be estimated to $\mathcal{O}(\varepsilon^{m+1})$. The estimation of the first two terms is identical to the one contained in Section 5.1.4.

the result is given by the right member of Eq. (5.34). Next, we estimate the last two terms and show that they are zero up to and including $\mathcal{O}(\varepsilon^{m+1})$ terms.

We start with $\Lambda^{(m)}_{21,s\perp s} \equiv R^{2s\perp}_{(m)} B^{s\perp}[A_s Q^{(m)}_{1s}, g]$. Let $\Delta x^2 \equiv x^2 - h_\varepsilon(x^1)$. Taylor expanding $\Lambda^{(m)}_{21,f\perp f}(x,\varepsilon)$ around $(x^1, h_\varepsilon(x^1), \varepsilon)$, we obtain

$$\begin{aligned}\Lambda^{(m)}_{21,s\perp s}(x,\varepsilon) &= \Lambda^{(m)}_{21,s\perp s}(x^1, h_\varepsilon(x^1),\varepsilon) \\ &\quad + \left(\left(D_z \Lambda^{(m)}_{21,s\perp s}\right)(x^1, h_\varepsilon(x^1),\varepsilon)\right)(\Delta x^2) + \mathcal{O}\left((\Delta x^2)^2\right). \quad (5.38)\end{aligned}$$

We are interested in the case $x^2 = \psi_{(m)}(x^1,\varepsilon)$. First, the term $\Lambda^{(m)}_{21,s\perp s}(y, h_\varepsilon, \varepsilon)$ is identically zero by Lemma 2.3.1. Then, $\Delta x^2 = \mathcal{O}(\varepsilon^{m+1})$ by Theorem 3.3.1, and thus the $\mathcal{O}\left((\Delta x^2)^2\right)$ terms in the right member of Eq. (5.38) are $\mathcal{O}(\varepsilon^{2m+2})$ and therefore zero to leading order. It only remains to estimate the derivative term to $\mathcal{O}(\varepsilon^{m+1})$. First, recall the definition of $\Lambda^{(m)}_{21,s\perp s}$ and observe that $B^{s\perp}$ and A_s only depend on x^1 (see Eqs. (5.3)–(5.4)). Thus, $D_{x^2} B^{s\perp} = 0$ and $D_{x^2} A_s = 0$, and therefore,

$$\begin{aligned}\left(D_z \Lambda^{(m)}_{21,s\perp s}\right)(\Delta x^2) &= \left(D_z R^{2s\perp}_{(m)}(\Delta x^2)\right) B^{s\perp}\left[A_s Q^{(m)}_{1s}, g\right] \\ &\quad + R^{2s\perp}_{(m)} B^{s\perp}\left[A_s\left(D_z Q^{(m)}_{1s}(\Delta x^2)\right), g\right] \\ &\quad + R^{2s\perp}_{(m)} B^{s\perp}\left[A_s Q^{(m)}_{1s}, (D_{x^2} g)(\Delta x^2)\right], \quad (5.39)\end{aligned}$$

where all the terms are evaluated on $\mathcal{M}_\varepsilon$. The first two terms in the right member of Eq. (5.39) are identically zero on $\mathcal{M}_\varepsilon$ by virtue of Lemma 2.3.1, and hence

$$\begin{aligned}&\left(\left(D_z \Lambda^{(m)}_{21,s\perp s}\right)(x^1, h_\varepsilon(x^1),\varepsilon)\right)(\Delta x^2) = \\ &\quad \left(R^{2s\perp}_{(m)} B^{s\perp}\left[A_s Q^{(m)}_{1s}, (D_{x^2} g)(\Delta x^2)\right]\right)(x^1, h_\varepsilon(x^1),\varepsilon). \quad (5.40)\end{aligned}$$

Next, recall that $R^{2s\perp}_{(m)}$, $Q^{(m)}_{1s}$, and Δx^2 are $\mathcal{O}(\varepsilon^{m+1})$ on $\mathcal{K}^{(m)}_\varepsilon$ and thus also on $\mathcal{M}_\varepsilon$. Thus, the right member of Eq. (5.40) is zero up to and including terms of $\mathcal{O}(\varepsilon^{m+1})$ and $\Lambda^{(m)}_{21,f\perp f}(x^1, \psi_{(m)}(x^1,\varepsilon),\varepsilon)$ contributes nothing to $\Lambda^{(m,m+1)}_{21}$.

Next, we show that $\Lambda^{(m)}_{21,f\perp f} \equiv R^{2f\perp}_{(m)} B^{f\perp}[A_f Q^{(m)}_{1f}, g]$ is also zero up to and including terms of $\mathcal{O}(\varepsilon^{m+1})$ and thus contributes nothing to $\Lambda^{(m,m+1)}_{21}$. Taylor expanding $\Lambda^{(m)}_{21,s\perp s}(y,z,\varepsilon)$ around $(y, h_\varepsilon, \varepsilon)$, we obtain

$$\begin{aligned}\Lambda^{(m)}_{21,f\perp f}(x,\varepsilon) &= \Lambda^{(m)}_{21,f\perp f}(x^1, h_\varepsilon(x^1),\varepsilon) \\ &\quad + \left(\left(D_z \Lambda^{(m)}_{21,f\perp f}\right)(x^1, h_\varepsilon(x^1),\varepsilon)\right)(\Delta x^2) + \mathcal{O}\left((\Delta x^2)^2\right). \quad (5.41)\end{aligned}$$

Let $z = \psi_{(m)}(x^1,\varepsilon)$, as before. The first term in the right member of Eq. (5.41) vanishes by Lemma 2.3.1 and the $\mathcal{O}\left((\Delta x^2)^2\right)$ terms are $\mathcal{O}(\varepsilon^{2m+2})$, as before, so we

focus on the derivative term. Recalling that $D_{x^2}B^{f\perp} = 0$ and $D_{x^2}A_f = 0$, we find

$$\begin{aligned}\left(D_z\Lambda^{(m)}_{21,f\perp f}\right)(\Delta x^2) &= \left(D_z R^{2f\perp}_{(m)}(\Delta x^2)\right) B^{f\perp}\left[A_f Q^{(m)}_{1f}, g\right] \\ &\quad + R^{2f\perp}_{(m)} B^{f\perp}\left[A_f\left(D_z Q^{(m)}_{1f}(\Delta x^2)\right), g\right] \\ &\quad + R^{2f\perp}_{(m)} B^{f\perp}\left[A_f Q^{(m)}_{1f}, (D_{x^2}g)(\Delta x^2)\right], \end{aligned} \tag{5.42}$$

with all the terms evaluated on $\mathcal{M}_\varepsilon$. The first two terms in the right member of Eq. (5.42) vanish on $\mathcal{M}_\varepsilon$, by virtue of Lemma 2.3.1, and hence

$$\begin{aligned}&\left(\left(D_{x^2}\Lambda^{(m)}_{21,f\perp f}\right)(x^1, h_\varepsilon(x^1), \varepsilon)\right)(\Delta x^2) = \\ &\qquad \left(R^{2f\perp}_{(m)} B^{f\perp}\left[A_f Q^{(m)}_{1f}, (D_{x^2}g)(\Delta x^2)\right]\right)(x^1, h_\varepsilon(x^1), \varepsilon). \end{aligned} \tag{5.43}$$

Finally, when $x^2 = \psi_{(m)}(x^1, \varepsilon)$, Δx^2 is $\mathcal{O}(\varepsilon^{m+1})$, $D_{x^2}g$ is $\mathcal{O}(1)$ and fast to leading order, and thus an easy calculation shows that $[A_f Q^{(m)}_{1f}, (D_{x^2}g)(\Delta x^2)]$ is $\mathcal{O}(\varepsilon^{m+1})$ and that the coefficient of ε^{m+1} in the asymptotic expansion of the quantity is fast. Hence, the right member of Eq. (5.43) is zero up to and including terms of $\mathcal{O}(\varepsilon^{m+1})$, and the estimation is complete.

Chapter 6

Application Of The CSP Method To The Michaelis–Menten–Henri Model

6.1 The Michaelis–Menten–Henri reaction

In this chapter, we apply the CSP method to the Michaelis–Menten–Henri (MMH) mechanism of enzyme kinetics [4, 18, 36, 37, 38] to illustrate Theorems 3.3.1, 3.3.2, and 3.3.3. We consider the planar system of ODEs for a slow variable s and a fast variable c,

$$s' = \varepsilon(-s + (s + \kappa - \lambda)c), \tag{6.1}$$
$$c' = s - (s + \kappa)c. \tag{6.2}$$

The parameters satisfy the inequalities $0 < \varepsilon \ll 1$ and $\kappa > \lambda > 0$. Only nonnegative values of s and c are relevant. The system of Eqs. (6.1)–(6.2) is of the form (2.2) with $N_s = N_f = 1$, $x^1 = s$, $x^2 = c$, $g^1 = -s + (s + \kappa - \lambda)c$, and $g^2 = s - (s + \kappa)c$.

6.1.1 Slow manifolds and fast fibers

In the limit as $\varepsilon \downarrow 0$, the dynamics of the MMH equations are confined to the reduced slow manifold,

$$\mathcal{M}_0 = \left\{ (s, c) \,\middle|\, c = \frac{s}{s + \kappa},\ s \geq 0 \right\}.$$

The manifold $\mathcal{M}_0$ is asymptotically stable, so there exists a locally invariant slow manifold $\mathcal{M}_\varepsilon$ for all sufficiently small ε that is $\mathcal{O}(\varepsilon)$ close to $\mathcal{M}_0$ on any compact set. Moreover, $\mathcal{M}_\varepsilon$ is the graph of a function h_ε,

$$\mathcal{M}_\varepsilon = \{(s, c) \mid c = h_\varepsilon(s),\ s \geq 0\},$$

and h_ε admits an asymptotic expansion, $h_\varepsilon = h_0 + \varepsilon h_1 + \varepsilon^2 h_2 + \cdots$. The coefficients are found from the invariance equation,

$$s - (s+\kappa)h_\varepsilon(s) = \varepsilon h_\varepsilon'(s)(-s + (s+\kappa-\lambda)h_\varepsilon(s)). \tag{6.3}$$

The first few coefficients are

$$h_0(s) = \frac{s}{s+\kappa}, \quad h_1(s) = \frac{\kappa\lambda s}{(s+\kappa)^4}, \quad h_2(s) = \frac{\kappa\lambda s(2\kappa\lambda - 3\lambda s - \kappa s - \kappa^2)}{(s+\kappa)^7}. \tag{6.4}$$

In the limit as $\varepsilon \downarrow 0$, each line of constant s is trivially invariant under the dynamics of system (6.1)–(6.2). These are the (one-dimensional) fast fibers $\mathcal{F}_0^p$ with base point $p = (s, h_0(s)) \in \mathcal{M}_0$. All points on $\mathcal{F}_0^p$ contract exponentially fast to p with rate constant $-(s+\kappa)$. The fast fiber $\mathcal{F}_0^p$ perturbs to a curve $\mathcal{F}_\varepsilon^p$ that is $\mathcal{O}(\varepsilon)-$close to $\mathcal{F}_0^p$ in any compact neighborhood of $\mathcal{M}_\varepsilon$. The fast fibers $\mathcal{F}_\varepsilon^p$, $p \in \mathcal{M}_\varepsilon$, form an invariant family.

6.2 Asymptotic expansions of the fast fibers

Asymptotic information about the slow manifold is readily available through the invariance equation (6.3). To derive asymptotic information about the fast fibers, we look for general solutions of system (6.1)–(6.2) that are given by asymptotic expansions,

$$s(t;\varepsilon) = \sum_{i=0} \varepsilon^i s_i(t), \quad c(t;\varepsilon) = \sum_{i=0} \varepsilon^i c_i(t), \tag{6.5}$$

where the coefficients s_i and c_i are determined order by order.

Consider the fast fiber $\mathcal{F}_\varepsilon^p$ with base point $p = (s, h_\varepsilon(s))$, and let (s^A, c^A) and (s^B, c^B) be two points on it; let $\Delta s(t) = s^B(t) - s^A(t)$ and $\Delta c(t) = c^B(t) - c^A(t)$. The distance between any two points on the same fast fiber will contract exponentially fast towards zero at the $\mathcal{O}(1)$ rate, as long as both points are chosen in a neighborhood of $\mathcal{M}_\varepsilon$. We may write

$$\Delta s(t;\varepsilon) = \sum_{i=0} \varepsilon^i \Delta s_i(t), \quad \Delta c(t;\varepsilon) = \sum_{i=0} \varepsilon^i \Delta c_i(t),$$

where $\Delta s_i(t) = s_i^B(t) - s_i^A(t)$ and $\Delta c_i(t) = c_i^B(t) - c_i^A(t)$. The condition on fast exponential decay of $\Delta s(t)$ and $\Delta c(t)$ translates into

$$\Delta s_i(t) = \mathcal{O}(e^{-C_s t}), \quad \Delta c_i(t) = \mathcal{O}(e^{-C_c t}), \quad t \to \infty,$$

for some positive constants C_s and C_c. We let (s^A, c^A) and (s^B, c^B) be infinitesimally close, since we are interested in vectors tangent to the fast fiber.

6.2.1 $\mathcal{O}(1)$ fast fibers

Substituting the expansions (6.5) into (6.1)–(6.2) and equating $\mathcal{O}(1)$ terms, we find

$$\begin{aligned} s_0' &= 0, \\ c_0' &= s_0 - (s_0 + \kappa)c_0. \end{aligned}$$

The equations can be integrated,

$$s_0(t) = s_0(0) = s_0, \tag{6.6}$$

$$c_0(t) = \frac{s_0}{s_0 + \kappa} + \left(c_0(0) - \frac{s_0}{s_0 + \kappa} \right) e^{-(s_0+\kappa)t}. \tag{6.7}$$

Hence,

$$\Delta s_0(t) = \Delta s_0(0), \tag{6.8}$$

$$\Delta c_0(t) = \Delta c_0(0) e^{-(s_0+\kappa)t} + (\partial_{s_0} c_0(t))\, \Delta s_0(0) + \mathcal{O}((\Delta s_0(0))^2). \tag{6.9}$$

The points A and B lie on the same fiber if and only if

$$\Delta s_0(0) = 0. \tag{6.10}$$

Thus, Eq. (6.9) simplifies to

$$\Delta c_0(t) = \Delta c_0(0) e^{-(s_0+\kappa)t},$$

and $\Delta c_0(t)$ decays exponentially towards zero, irrespective of the choice of $\Delta c_0(0)$. Hence, $\Delta c_0(0)$ is a free parameter.

We conclude that, to $\mathcal{O}(1)$, any vector $\begin{pmatrix} 0 \\ \alpha \end{pmatrix}$ with α constant ($\alpha \neq 0$) is tangent to every fast fiber at the base point.

6.2.2 $\mathcal{O}(\varepsilon)$ fast fibers

At $\mathcal{O}(\varepsilon)$, we obtain the equations

$$s_1' = -s_0 + (s_0 + \kappa - \lambda)c_0, \tag{6.11}$$

$$c_1' = s_1 - (s_0 + \kappa)c_1 - s_1 c_0. \tag{6.12}$$

Using Eqs. (6.6) and (6.7), we integrate both members of Eq. (6.11) to obtain

$$s_1(t) = s_1(0) - \frac{\lambda s_0}{s_0 + \kappa} t + \frac{s_0 + \kappa - \lambda}{s_0 + \kappa} \left(c_0(0) - \frac{s_0}{s_0 + \kappa} \right) (1 - e^{-(s_0+\kappa)t}).$$

Therefore, at $\mathcal{O}(\varepsilon)$,

$$\Delta s_1(t) = \Delta s_1(0) + \frac{s_0+\kappa-\lambda}{s_0+\kappa}\Delta c_0(0)(1-e^{-(s_0+\kappa)t}).$$

For the two points to have the same phase asymptotically, it is necessary that $\lim_{t\to\infty}\Delta s_1(t) = 0$. This condition is satisfied if and only if

$$\Delta s_1(0) = -\frac{s_0+\kappa-\lambda}{s_0+\kappa}\Delta c_0(0). \tag{6.13}$$

Next, $c_1(t)$ follows upon integration of both members of Eq. (6.12),

$$\begin{aligned} c_1(t) \;=\; & c_1(0)e^{-(s_0+\kappa)t} \\ & + \frac{\kappa}{(s_0+\kappa)^2}\left(s_1(0) + \frac{s_0+\kappa-\lambda}{s_0+\kappa}\left(c_0(0) - \frac{s_0}{s_0+\kappa}\right)\right)(1-e^{-(s_0+\kappa)t}) \\ & - \left(c_0(0) - \frac{s_0}{s_0+\kappa}\right)\left(s_1(0) + \frac{s_0+\kappa-\lambda}{s_0+\kappa}\left(c_0(0) + \frac{\kappa-s_0}{s_0+\kappa}\right)\right)te^{-(s_0+\kappa)t} \\ & - \frac{s_0+\kappa-\lambda}{(s_0+\kappa)^2}\left(c_0(0) - \frac{s_0}{s_0+\kappa}\right)^2(e^{-2(s_0+\kappa)t} - e^{-(s_0+\kappa)t}) \\ & + \frac{\lambda s_0}{2(s_0+\kappa)}\left(c_0(0) - \frac{s_0}{s_0+\kappa}\right)t^2e^{-(s_0+\kappa)t} \\ & - \frac{\kappa\lambda s_0}{(s_0+\kappa)^4}(e^{-(s_0+\kappa)t} + (s_0+\kappa)t - 1). \end{aligned}$$

We infer from this expression that $\lim_{t\to\infty}\Delta c_1(t) = 0$, as long as Eqs. (6.10) and (6.13) hold. Hence, $\Delta c_1(0)$ is a free parameter, just like $\Delta c_0(0)$, and the only condition that arises at $\mathcal{O}(\varepsilon)$ is (6.13) on $\Delta s_1(0)$.

We conclude that any vector

$$\begin{pmatrix} 0 \\ \alpha \end{pmatrix} + \varepsilon \begin{pmatrix} -\left(1-\frac{\lambda}{s_0+\kappa}\right)\alpha \\ \beta \end{pmatrix}, \tag{6.14}$$

with α and β constant ($\alpha \neq 0$), is tangent to every fast fiber at the base point up to and including terms of $\mathcal{O}(\varepsilon)$. Any such vector may be written as the product of a free parameter and a constant vector (fixed by s_0),

$$(\alpha+\varepsilon\beta)\begin{pmatrix} -\varepsilon\left(1-\frac{\lambda}{s_0+\kappa}\right) \\ 1 \end{pmatrix} + \mathcal{O}(\varepsilon^2).$$

6.2.3 $\mathcal{O}(\varepsilon^2)$ fast fibers

At $\mathcal{O}(\varepsilon^2)$, we obtain the equation

$$s_2' = s_1(c_0 - 1) + (s_0 + \kappa - \lambda)c_1.$$

Direct integration yields

$$\begin{aligned} s_2(t) &= s_2(0) + \left[\frac{\lambda}{(s_0+\kappa)^2}\left(c_0(0) - \frac{s_0}{s_0+\kappa}\right) - \frac{\kappa(s_0+\kappa-\lambda)}{(s_0+\kappa)^3}\right] s_1(0) \\ &\quad - \left[\frac{\kappa(s_0+\kappa-\lambda)(s_0+\kappa-2\lambda)+\lambda^2 s_0}{(s_0+\kappa)^4}\right]\left(c_0(0) - \frac{s_0}{s_0+\kappa}\right) \\ &\quad + \frac{\lambda(s_0+\kappa-\lambda)}{2(s_0+\kappa)^3}\left(c_0(0) - \frac{s_0}{s_0+\kappa}\right)^2 \\ &\quad + \left(1 - \frac{\lambda}{s_0+\kappa}\right)\left(c_1(0) - \frac{\kappa\lambda s_0}{(s_0+\kappa)^4}\right) \\ &\quad - \frac{\kappa\lambda}{(s_0+\kappa)^2}\left[s_1(0) + \frac{s_0+\kappa-\lambda}{s_0+\kappa}\left(c_0(0) - \frac{2s_0}{s_0+\kappa}\right)\right] t \\ &\quad + \frac{\kappa\lambda^2 s_0}{2(s_0+\kappa)^3} t^2 + \mathcal{R}(t), \end{aligned}$$

where the remainder $\mathcal{R}(t)$ involves the functions $e^{-(s_0+\kappa)t}$, $te^{-(s_0+\kappa)t}$, $t^2e^{-(s_0+\kappa)t}$, and $e^{-2(s_0+\kappa)t}$. From this expression we find

$$\begin{aligned} \Delta s_2(t) &= \Delta s_2(0) + (\partial_{s_0} s_2(t))\,\Delta s_0(0) + (\partial_{c_0} s_2(t))\,\Delta c_0(0) \\ &\quad + (\partial_{s_1} s_2(t))\,\Delta s_1(0) + (\partial_{c_1} s_2(t))\,\Delta c_1(0) + \mathcal{O}(2) + \mathcal{O}(e^{-Ct}) \end{aligned} \quad (6.15)$$

for some $C > 0$. Here, ∂_{c_0} is an abbreviation for the partial derivative $\partial_{c_0(0)}$, and so on, and $\mathcal{O}(2)$ denotes quadratic terms in the multivariable Taylor expansion. First, we recall that $\Delta s_0(0) = 0$ by Eq. (6.10). Next, we calculate the partial derivatives in each of the three remaining terms,

$$\begin{aligned} \partial_{c_0} s_2(t) &= \frac{\lambda s_1(0)}{(s_0+\kappa)^2} - \frac{\kappa(s_0+\kappa-\lambda)(s_0+\kappa-2\lambda)+\lambda^2 s_0}{(s_0+\kappa)^4} \\ &\quad + \frac{\lambda(s_0+\kappa-\lambda)}{(s_0+\kappa)^3}\left(c_0(0) - \frac{s_0}{s_0+\kappa}\right) - \frac{\kappa\lambda(s_0+\kappa-\lambda)}{(s_0+\kappa)^3} t, \\ \partial_{s_1} s_2(t) &= \frac{\lambda}{(s_0+\kappa)^2}\left(c_0(0) - \frac{s_0}{s_0+\kappa}\right) - \frac{\kappa(s_0+\kappa-\lambda)}{(s_0+\kappa)^3} \\ &\quad - \frac{\kappa\lambda}{(s_0+\kappa)^2} t, \\ \partial_{c_1} s_2(t) &= 1 - \frac{\lambda}{s_0+\kappa}. \end{aligned}$$

We substitute these expressions into Eq. (6.15), recall Eq. (6.13), and carry out the algebra to obtain

$$\begin{aligned}\Delta s_2(t) &= \Delta s_2(0) + \left(1 - \frac{\lambda}{s_0+\kappa}\right)\Delta c_1(0) \\ &\quad + \frac{\lambda}{(s_0+\kappa)^2}\left(s_1(0) + \frac{\kappa(s_0+\kappa-\lambda)-\lambda s_0}{(s_0+\kappa)^2}\right)\Delta c_0(0) \\ &\quad + \mathcal{O}(2) + \mathcal{O}(e^{-Ct}), \quad C>0. \end{aligned} \tag{6.16}$$

In the limit $t \to \infty$, Eq. (6.16) yields the condition

$$\begin{aligned}\Delta s_2(0) &= -\left(1 - \frac{\lambda}{s_0+\kappa}\right)\Delta c_1(0) \\ &\quad - \frac{\lambda}{(s_0+\kappa)^2}\left(s_1(0) + \frac{\kappa(s_0+\kappa-\lambda)-\lambda s_0}{(s_0+\kappa)^2}\right)\Delta c_0(0). \end{aligned} \tag{6.17}$$

Finally, $\Delta c_2(t)$ vanishes exponentially, as follows directly from the conditions (6.13) and (6.17). Thus, no further conditions besides (6.17) arise at $\mathcal{O}(\varepsilon^2)$.

We conclude that any vector

$$\begin{pmatrix} 0 \\ \alpha \end{pmatrix} + \varepsilon \begin{pmatrix} -\left(1-\frac{\lambda}{s_0+\kappa}\right)\alpha \\ \beta \end{pmatrix} + \varepsilon^2 \begin{pmatrix} -\left(1-\frac{\lambda}{s_0+\kappa}\right)\beta - \frac{\lambda}{(s_0+\kappa)^2}\left(s_1(0) + \frac{\kappa(s_0+\kappa-\lambda)-\lambda s_0}{(s_0+\kappa)^2}\right)\alpha \\ \gamma \end{pmatrix}, \tag{6.18}$$

with α, β, and γ constant ($\alpha \neq 0$), is tangent to every fiber at the base point up to and including terms of $\mathcal{O}(\varepsilon^2)$.

6.3 Application of the one-step CSP method

We initialize both the one-step and the full CSP methods with the same initial basis. We choose the stoichiometric vectors as the basis vectors, so

$$\tilde{A}^{(0)} = A^{(0)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tilde{B}_{(0)} = B_{(0)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{6.19}$$

The CSP condition $B^1_{(0)} g = 0$ is satisfied if $c = h_0(s)$, so the CSP manifolds $\tilde{\mathcal{K}}^{(0)}_\varepsilon$ and $\mathcal{K}^{(0)}_\varepsilon$ coincide with $\mathcal{M}_0$. With this choice of initial basis, we have

$$\tilde{\Lambda}_{(0)} = \Lambda_{(0)} = B_{(0)}(Dg)A^{(0)} = \begin{pmatrix} -(s+\kappa) & -(c-1) \\ \varepsilon(s+\kappa-\lambda) & \varepsilon(c-1) \end{pmatrix}. \tag{6.20}$$

6.3.1 First iteration.

At any point (s, c), we have

$$\tilde{A}^{(1)} = \begin{pmatrix} 0 & 1 \\ 1 & -\frac{c-1}{s+\kappa} \end{pmatrix}, \quad \tilde{B}_{(1)} = \begin{pmatrix} \frac{c-1}{s+\kappa} & 1 \\ 1 & 0 \end{pmatrix}.$$

On $\tilde{\mathcal{K}}_\varepsilon^{(0)}$, these expressions reduce to

$$\tilde{A}^{(1)} = \begin{pmatrix} 0 & 1 \\ 1 & \frac{\kappa}{(s+\kappa)^2} \end{pmatrix}, \quad \tilde{B}_{(1)} = \begin{pmatrix} \frac{-\kappa}{(s+\kappa)^2} & 1 \\ 1 & 0 \end{pmatrix}. \tag{6.21}$$

The CSP condition

$$\tilde{B}^1_{(1)} g = s - (s+\kappa)c - \varepsilon \frac{\kappa(-s + (s+\kappa-\lambda)c)}{(s+\kappa)^2} = 0$$

is satisfied if

$$c = \frac{s}{s+\kappa} + \varepsilon \frac{\kappa\lambda s}{(s+\kappa)^4} - \varepsilon^2 \frac{\kappa^2 \lambda s(s+\kappa-\lambda)}{(s+\kappa)^7} + \mathcal{O}(\varepsilon^3).$$

Comparing this result with Eq. (6.4), we see that the asymptotic expansions of $\tilde{\mathcal{K}}_\varepsilon^{(1)}$ and $\mathcal{M}_\varepsilon$ coincide up to and including $\mathcal{O}(\varepsilon)$ terms, in accordance with Theorem 3.3.2 for $m = 1$; however, the $\mathcal{O}(\varepsilon^2)$ terms differ at this stage.

6.3.2 Second iteration.

The blocks of $\tilde{\Lambda}_{(1)}$ are

$$\begin{aligned}
\tilde{\Lambda}^{11}_{(1)} &= -(s+\kappa) + \varepsilon \frac{(s+\kappa-\lambda)(c-1)}{s+\kappa}, \\
\tilde{\Lambda}^{12}_{(1)} &= \frac{s}{s+\kappa} - c + \varepsilon \frac{(c-1)[\lambda(c-1) - (-s + (s+\kappa-\lambda)c)]}{(s+\kappa)^2}, \\
\tilde{\Lambda}^{21}_{(1)} &= \varepsilon(s+\kappa-\lambda), \quad \tilde{\Lambda}^{22}_{(1)} = \varepsilon \frac{\lambda(c-1)}{s+\kappa}.
\end{aligned}$$

On $\tilde{\mathcal{K}}_\varepsilon^{(1)}$, the blocks reduce to

$$\begin{aligned}
\tilde{\Lambda}^{11}_{(1)} &= -(s+\kappa) - \varepsilon \frac{\kappa(s+\kappa-\lambda)}{(s+\kappa)^2} + \varepsilon^2 \frac{\kappa\lambda s(s+\kappa-\lambda)}{(s+\kappa)^5}, \\
\tilde{\Lambda}^{12}_{(1)} &= \varepsilon \frac{\kappa\lambda(\kappa-2s)}{(s+\kappa)^4} + \varepsilon^2 \frac{\kappa\lambda s(2\kappa(s+\kappa-2\lambda)+\lambda s)}{(s+\kappa)^7}, \\
\tilde{\Lambda}^{21}_{(1)} &= \varepsilon(s+\kappa-\lambda), \quad \tilde{\Lambda}^{22}_{(1)} = -\varepsilon \frac{\kappa\lambda}{(s+\kappa)^2} + \varepsilon^2 \frac{\kappa\lambda^2 s}{(s+\kappa)^5}.
\end{aligned}$$

The second update is

$$\begin{aligned}
\tilde{A}_1^{(2)} &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\
\tilde{A}_2^{(2)} &= \begin{pmatrix} 1 \\ \frac{\kappa}{(s+\kappa)^2} \end{pmatrix} + \varepsilon \begin{pmatrix} 0 \\ \frac{\kappa\lambda(\kappa-3s)}{(s+\kappa)^5} \end{pmatrix} \\
&\quad + \varepsilon^2 \begin{pmatrix} 0 \\ \frac{\kappa\lambda[\kappa(5s-\kappa)(s+\kappa-\lambda)+\lambda s(s-2\kappa)]}{(s+\kappa)^8} \end{pmatrix} + \mathcal{O}(\varepsilon^3), \\
\tilde{B}_{(2)}^1 &= \left(-\frac{\kappa}{(s+\kappa)^2},\, 1\right) + \varepsilon \left(-\frac{\kappa\lambda(\kappa-3s)}{(s+\kappa)^5},\, 0\right) \\
&\quad + \varepsilon^2 \left(-\frac{\kappa\lambda[\kappa(5s-\kappa)(s+\kappa-\lambda)+\lambda s(s-2\kappa)]}{(s+\kappa)^8},\, 0\right) \\
&\quad + \mathcal{O}(\varepsilon^3), \\
\tilde{B}_{(2)}^2 &= (1,\, 0)\,.
\end{aligned}$$

The CSP condition

$$\begin{aligned}
\tilde{B}_{(2)}^1 g &= s - (s+\kappa)c - \varepsilon\frac{\kappa(-s+(s+\kappa-\lambda)c)}{(s+\kappa)^2} \\
&\quad + \varepsilon^2\kappa\lambda\frac{(3s-\kappa)(-s+(s+\kappa-\lambda)c)}{(s+\kappa)^5} + \mathcal{O}(\varepsilon^3) \\
&= 0,
\end{aligned}$$

is satisfied if

$$c = \frac{s}{s+\kappa} + \varepsilon\frac{\kappa\lambda s}{(s+\kappa)^4} + \varepsilon^2\frac{\kappa\lambda s(2\kappa\lambda - 3\lambda s - \kappa s - \kappa^2)}{(s+\kappa)^7} + \mathcal{O}(\varepsilon^3).$$

Comparing this result with Eq. (6.4), we see that the asymptotic expansions of $\tilde{\mathcal{K}}_\varepsilon^{(2)}$ and $\mathcal{M}_\varepsilon$ coincide up to and including $\mathcal{O}(\varepsilon^2)$ terms, in accordance with Theorem 3.3.2 for $m = 2$.

6.4 Application of the full CSP method

6.4.1 First iteration.

We mentioned in the previous section that both the one-step and the full CSP methods are initialized with the same basis given in Eq. (6.19). Thus, $\Lambda_{(0)}$ for the full CSP method is identical to $\Lambda_{(0)}$ for the one-step method, which is given in Eq. (6.20).

First Step. At any point (s, c), we have

$$\begin{aligned} A_1^{(1)} &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \varepsilon \frac{s+\kappa-\lambda}{s+\kappa} \begin{pmatrix} -1 \\ \frac{c-1}{s+\kappa} \end{pmatrix}, \quad A_2^{(1)} = \begin{pmatrix} 1 \\ -\frac{c-1}{s+\kappa} \end{pmatrix}, \\ B_{(1)}^1 &= \left(-A_{22}^{(1)},\, A_{12}^{(1)} \right), \qquad\qquad B_{(1)}^2 = \left(A_{21}^{(1)},\, -A_{11}^{(1)} \right). \end{aligned}$$

In the first step, we evaluate $A_2^{(1)}$ and $B_{(1)}^1$ on $\mathcal{K}_\varepsilon^{(0)}$ to obtain

$$A_2^{(1)} = \begin{pmatrix} 1 \\ \frac{\kappa}{(s+\kappa)^2} \end{pmatrix}, \quad B_{(1)}^1 = \left(-\frac{\kappa}{(s+\kappa)^2},\, 1 \right). \tag{6.22}$$

Equations (6.21) and (6.22) imply that $B_{(1)} = \tilde{B}_{(1)}$, so the CSP condition yields $\psi_{(1)} = \tilde{\psi}_{(1)}$. Thus, after one iteration, the full CSP method also finds the expansion of $\mathcal{M}_\varepsilon$ up to and including $\mathcal{O}(\varepsilon)$ terms.

Second Step. Then, in the second step, the new fast basis vector, $A_1^{(1)}$, and its complement, $B_{(1)}^2$, in the dual basis are evaluated on $\mathcal{K}_\varepsilon^{(1)}$,

$$\begin{aligned} A_1^{(1)} &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \varepsilon \begin{pmatrix} 1 \\ \frac{\kappa(s+\kappa-\lambda)}{(s+\kappa)^3} \end{pmatrix} + \varepsilon^2 \begin{pmatrix} 0 \\ \frac{\kappa\lambda s(s+\kappa-\lambda)}{(s+\kappa)^6} \end{pmatrix} + \mathcal{O}(\varepsilon^3), \\ B_{(1)}^2 &= \left(A_{21}^{(1)},\, -A_{11}^{(1)} \right). \end{aligned}$$

Thus, we see that $A_1^{(1)}$ is tangent to the fast fibers at their base points up to and including terms of $\mathcal{O}(\varepsilon)$, as Eq. (6.14) (with $\alpha = 1$, $\beta = -\frac{\kappa(s+\kappa-\lambda)}{(s+\kappa)^3}$) implies. As a result, $\mathcal{L}_\varepsilon^{(1)}$ approximates $\mathcal{TF}_\varepsilon$ also up to and including terms of $\mathcal{O}(\varepsilon)$.

The matrix relating $B_{(1)}$ to its one-step counterpart $\tilde{B}_{(1)}$ (recall Eq. (4.40)) is

$$T_{(1)} = \begin{pmatrix} 1 & 0 \\ \varepsilon \frac{s+\kappa-\lambda}{s+\kappa} & 1 \end{pmatrix},$$

so $T_{(1)}$ is indeed of the form (4.45) on $\mathcal{K}_\varepsilon^{(0)}$.

Remark. If one evaluates $A_1^{(1)}$ on $\mathcal{K}_\varepsilon^{(0)}$, as opposed to $\mathcal{K}_\varepsilon^{(1)}$ as we did above, then the approximation of $\mathcal{TF}_\varepsilon$ is also accurate up to and including terms of $\mathcal{O}(\varepsilon)$. See also Appendix B.

6.4.2 Second iteration.

The blocks of $\Lambda_{(1)}$ are

$$\begin{aligned}
\Lambda^{11}_{(1)} &= -(s+\kappa)+\varepsilon\frac{(s+\kappa-\lambda)}{s+\kappa}\left[(c-1)+\left(c-\frac{s}{s+\kappa}\right)\right] \\
&\quad +\varepsilon^2\frac{(c-1)(s+\kappa-\lambda)}{(s+\kappa)^3}\Big[-\lambda(c-1)+[(s+\kappa-\lambda)c-s]\Big], \\
\Lambda^{12}_{(1)} &= \frac{s}{s+\kappa}-c+\varepsilon\frac{c-1}{(s+\kappa)^2}\Big[\lambda(c-1)-[(s+\kappa-\lambda)c-s]\Big], \\
\Lambda^{21}_{(1)} &= \frac{\varepsilon^2}{(s+\kappa)^2}\Big[(c-1)(s+\kappa-\lambda)(s+\kappa-2\lambda) \\
&\quad +\lambda[(s+\kappa-\lambda)c-s]+(s+\kappa-\lambda)^2\left(c-\frac{s}{s+\kappa}\right)\Big], \\
\Lambda^{22}_{(1)} &= \frac{\varepsilon}{s+\kappa}\left[\lambda(c-1)+(s+\kappa-\lambda)\left(\frac{s}{s+\kappa}-c\right)\right] \\
&\quad +\varepsilon^2\frac{(c-1)(s+\kappa-\lambda)}{(s+\kappa)^3}\Big[\lambda(c-1)-[(s+\kappa-\lambda)c-s]\Big],
\end{aligned}$$

with remainders of $\mathcal{O}(\varepsilon^3)$. On $\mathcal{K}^{(1)}_\varepsilon$, the blocks reduce to

$$\begin{aligned}
\Lambda^{11}_{(1)} &= -(s+\kappa)-\varepsilon\frac{\kappa(s+\kappa-\lambda)}{(s+\kappa)^2}+\varepsilon^2\frac{\kappa\lambda(s+\kappa-\lambda)(3s-\kappa)}{(s+\kappa)^5}, \\
\Lambda^{12}_{(1)} &= \varepsilon\frac{\kappa\lambda(\kappa-2s)}{(s+\kappa)^4}+\varepsilon^2\frac{\kappa\lambda s(2\kappa(s+\kappa-2\lambda)+\lambda s)}{(s+\kappa)^7}, \\
\Lambda^{21}_{(1)} &= -\varepsilon^2\frac{\kappa(s+\kappa-\lambda)(s+\kappa-2\lambda)+\lambda^2 s}{(s+\kappa)^3}, \\
\Lambda^{22}_{(1)} &= -\varepsilon\frac{\kappa\lambda}{(s+\kappa)^2}-\varepsilon^2\frac{\kappa\lambda((2s-\kappa)(s+\kappa-\lambda)-\lambda s)}{(s+\kappa)^5},
\end{aligned}$$

with errors of $\mathcal{O}(\varepsilon^3)$.

First Step. In the first step, we update $A_2^{(1)}$ and $B_{(1)}^1$ and evaluate the updated quantities on $\mathcal{K}_\varepsilon^{(1)}$ to obtain

$$\begin{aligned}
A_{12}^{(2)} &= 1+\varepsilon^2\frac{\kappa\lambda(2s-\kappa)(s+\kappa-\lambda)}{(s+\kappa)^6},\\
A_{22}^{(2)} &= \frac{\kappa}{(s+\kappa)^2}+\varepsilon\frac{\kappa\lambda(\kappa-3s)}{(s+\kappa)^5}\\
&\quad +\varepsilon^2\frac{\kappa^2\lambda(7s-2\kappa)(s+\kappa-\lambda)+\kappa\lambda^2 s(s-2\kappa)}{(s+\kappa)^8},\\
B_{(2)}^1 &= \left(-A_{22}^{(2)},\, A_{12}^{(2)}\right)
\end{aligned}$$

up to and including terms of $\mathcal{O}(\varepsilon^2)$.

The CSP condition

$$\begin{aligned}
B_{(2)}^1 g &= s-(s+\kappa)c-\varepsilon\frac{\kappa(-s+(s+\kappa-\lambda)c)}{(s+\kappa)^2}\\
&\quad +\varepsilon^2\kappa\lambda\left(\frac{(3s-\kappa)(-s+(s+\kappa-\lambda)c)}{(s+\kappa)^5}\right.\\
&\quad \left.+\frac{(2s-\kappa)(s+\kappa-\lambda)(s-(s+\kappa)c)}{(s+\kappa)^6}\right)+\mathcal{O}(\varepsilon^3)\\
&= 0
\end{aligned}$$

is satisfied if

$$c=\frac{s}{s+\kappa}+\varepsilon\frac{\kappa\lambda s}{(s+\kappa)^4}+\varepsilon^2\frac{\kappa\lambda s(2\kappa\lambda-3\lambda s-\kappa s-\kappa^2)}{(s+\kappa)^7}+\mathcal{O}(\varepsilon^3). \tag{6.23}$$

Equation (6.23) defines $\mathcal{K}_\varepsilon^{(2)}$, the CSPM of order two, which agrees with $\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^2)$; recall Eq. (6.4).

Second Step. Then, in the second step, we update $A_1^{(1)}$ and $B^2_{(1)}$ to obtain

$$\begin{aligned}
A_{11}^{(2)} &= -\varepsilon\frac{s+\kappa-\lambda}{s+\kappa} - \varepsilon^2\frac{1}{(s+\kappa)^3}\Bigg[(s+\kappa-\lambda)(s+\kappa-2\lambda)(c-1) \\
&\quad + (s+\kappa-\lambda)^2\left(c-\frac{s}{s+\kappa}\right) + \lambda[(s+\kappa-\lambda)c-s]\Bigg]. \\
A_{21}^{(2)} &= 1+\varepsilon\frac{(s+\kappa-\lambda)(c-1)}{(s+\kappa)^2} \\
&\quad + \varepsilon^2\frac{1}{(s+\kappa)^4}\Bigg[(s+\kappa-\lambda)\Big[(s+\kappa-2\lambda)(c-1) \\
&\quad + (s+\kappa-\lambda)\left(c-\frac{s}{s+\kappa}\right)+\lambda c\Big] - \lambda s\Bigg]\left(2c-\frac{2s+\kappa}{s+\kappa}\right), \\
B^2_{(2)} &= \left(A_{21}^{(2)},\, -A_{11}^{(2)}\right),
\end{aligned}$$

with remainders of $\mathcal{O}(\varepsilon^3)$. Evaluating these expressions on $\mathcal{K}_\varepsilon^{(2)}$, we obtain

$$\begin{aligned}
A_{11}^{(2)} &= -\varepsilon\frac{s+\kappa-\lambda}{s+\kappa} + \varepsilon^2\frac{\kappa(s+\kappa-2\lambda)(s+\kappa-\lambda)+\lambda^2 s}{(s+\kappa)^4}, \\
A_{21}^{(2)} &= 1-\varepsilon\frac{\kappa(s+\kappa-\lambda)}{(s+\kappa)^3} \\
&\quad + \varepsilon^2\frac{(s+\kappa-\lambda)(\kappa^2(s+\kappa-2\lambda)+\kappa\lambda s)+\kappa\lambda^2 s}{(s+\kappa)^6}, \\
B^2_{(2)} &= \left(A_{21}^{(2)},\, -A_{11}^{(2)}\right),
\end{aligned}$$

with remainders of $\mathcal{O}(\varepsilon^3)$. Therefore, $A_1^{(2)}$ is tangent to the fast fibers at their base points up to and including terms of $\mathcal{O}(\varepsilon^2)$, according to Eq. (6.18) (with $\alpha = 1$, $\beta = -\frac{\kappa(s+\kappa-\lambda)}{(s+\kappa)^3}$, $\gamma = \frac{(s+\kappa-\lambda)(\kappa^2(s+\kappa-2\lambda)+\kappa\lambda s)+\kappa\lambda^2 s}{(s+\kappa)^6}$), and $\mathcal{L}_\varepsilon^{(2)}$ is an $\mathcal{O}(\varepsilon^2)-$accurate approximation to $\mathcal{TF}_\varepsilon$.

Also, on $\mathcal{K}_\varepsilon^{(1)}$,

$$\begin{aligned}
T_{(2)}^{11} &= 1+\varepsilon^2\frac{\kappa\lambda(s+\kappa-\lambda)(2s-\kappa)}{(s+\kappa)^6}, \\
T_{(2)}^{12} &= 0, \\
T_{(2)}^{21} &= \varepsilon\frac{(s+\kappa-\lambda)}{s+\kappa} - \varepsilon^2\frac{\kappa(s+\kappa-\lambda)(s+\kappa-2\lambda)+\lambda^2 s}{(s+\kappa)^4}, \\
T_{(2)}^{22} &= 1-\varepsilon^2\frac{\kappa\lambda(2s-\kappa)(s+\kappa-\lambda)}{(s+\kappa)^8},
\end{aligned}$$

with remainders of $\mathcal{O}(\varepsilon^3)$. Thus, $T_{(2)}$ is indeed of the form (4.45) on $\mathcal{K}_\varepsilon^{(1)}$.

Remark. If one evaluates $A_1^{(2)}$ on $\mathcal{K}_\varepsilon^{(1)}$ instead of on $\mathcal{K}_\varepsilon^{(2)}$, as we did above, then the approximation of $\mathcal{TF}_\varepsilon$ is also accurate up to and including terms of $\mathcal{O}(\varepsilon^2)$, see Section 5.3.

Chapter 7

The ZDP Method

The Zero-derivative Principle (ZDP) focuses on the state variables and is, in essence, an iterative process to generate an ideal coordinate system in the state space (in the sense described in Section 2.3) where the fast and slow dynamics decouple. At each step, one partitions the ZDP coordinates into "fast" and "slow" and constructs a change of coordinates $y = y(x)$ by identifying the new "fast" coordinates with the time derivatives of the old ones. An approximation to the slow manifold is defined as the locus of points in state space where the then current "fast" variables vanish, as in Eq. (2.29). In Section 7.1 we describe ZDP in some detail. Then, in Section 7.2 we state the main theorem, Theorem 7.2.1, concerning the approximating properties of ZDP. The proof of the theorem is given in Section 7.5.

7.1 The ZDP algorithm

The ZDP algorithm is applicable to general ODEs of the form (2.16). We recall here that these systems are not necessarily in fast–slow format, as in Eq. (2.2), but merely *diffeomorphic* to a fast–slow system (see Section 2.2 for details). We also remark that neither the explicit form of the transformation Eqs. (2.17)–(2.18) nor that of Eq. (2.2) need be known for the applicability of the technique.

To apply the ZDP algorithm, one partitions the vector z of state space variables into two parts, $z^1 \in \mathbf{R}^{N_s}$ and $z^2 \in \mathbf{R}^{N_f}$, where the variables z^2 are suspected of "*somehow containing the fast variables*" (this statement will be made precise when we state Theorem 7.2.1). To make this fact explicit, we rewrite the system (2.16) as

$$\begin{aligned} (z^1)' &= q^1(z), \qquad z^1 \in \mathbf{R}^{N_s}, \\ (z^2)' &= q^2(z), \qquad z^2 \in \mathbf{R}^{N_f}. \end{aligned}$$

Next, we initialize the algorithm with the standard coordinate system in $\mathbf{R}^N$,[1]

$$y^1_{(0)}(z) = z^2 \quad \text{and} \quad y^2_{(0)}(z) = z^1,$$

and generate a sequence of state space coordinates,

$$\left\{ y_{(m)}(z) \middle| \ m = 0, 1, \ldots \right\},$$

by means of the update rule

$$y_{(m+1)}(z) = \begin{pmatrix} y^1_{(m+1)}(z) \\ y^2_{(m+1)}(z) \end{pmatrix} = \begin{pmatrix} (y^1_{(m)})'(z) \\ y^2_{(m)}(z) \end{pmatrix}, \quad m = 0, 1, \ldots . \tag{7.1}$$

Time derivatives are taken along solutions of Eq. (2.16), so $d \cdot /dt = (D_z \cdot)q$. An explicit form of the update rule is

$$y_{(m+1)}(z) = \begin{pmatrix} L_{(m)}(z) \\ z^1 \end{pmatrix}, \quad m = 0, 1, \ldots , \tag{7.2}$$

where

$$L_{(m)}(z) \equiv \left(\frac{d^{m+1} z^2}{dt^{m+1}} \right)(z). \tag{7.3}$$

At the $(m+1)$th iteration, an approximate slow manifold is defined by

$$\mathcal{L}^{(m)} = \left\{ z \in \mathbf{R}^N \ \middle| \ y^1_{(m+1)}(z) = L_{(m)}(z) = 0 \right\}, \quad m = 0, 1, \ldots . \tag{7.4}$$

This definition, which is the analog of Eq. (2.29), together with Eq. (7.3) shows why we refer to the present reduction technique as the "Zero-derivative Principle." Also, in analogy with the terminology we introduced fot the CSP method, we call $\mathcal{L}^{(m)}$ the *ZDP manifold* (ZDPM) of order m.

7.2 Approximating properties of the ZDP

We have seen in Section 2.2 that, under standard assumptions on the spectrum of $D_{x^2} g^2$ and on the smoothness of q and of the transformations (2.17) and (2.18), the system (2.2) has a normally attractive manifold

$$\mathcal{M}_\varepsilon = \left\{ x \in \mathbf{R}^N \ \middle| \ x^2 = h_\varepsilon(x^1), \ x^1 \in K \right\}.$$

For the sake of the analysis, we assume that there is a compact set $\tilde{K}$ such that the manifold $\mathcal{M}_\varepsilon$ can be expressed as a graph over $\tilde{K}$ in the original coordinates,

$$\mathcal{M}_\varepsilon = \left\{ z \in \mathbf{R}^N \ \middle| \ z^2 = \phi_\varepsilon(z^1), \ z^1 \in \tilde{K} \right\}, \quad \text{for some} \quad \phi_\varepsilon : \tilde{K} \to \mathbf{R}^{N_f}.$$

[1] The inversion is deliberate; we choose to have the fast variables come first for reasons of comparison with CSP, see Section 8.1.2.

(In Section 9.4 in the Appendix, we prove that this is a natural assumption to make under the hypotheses of our main theorem concerning the ZDP.)

The following theorem contains our main result on the approximation of $\mathcal{M}_\varepsilon$ that the ZDP provides at the $m-$th iteration through the ZDP manifold of order m.

Theorem 7.2.1 *If $D_{x^2}z^2$ and $D_{x^2}q^2$ are non-singular in a neighborhood of $\mathcal{M}_0$, then the $(m+1)$st derivative condition*

$$L_{(m)}(z) = \left(\frac{d^{m+1}z^2}{dt^{m+1}}\right)(z) = 0$$

can be solved for z^2 to yield an N_s-dimensional manifold $\mathcal{L}^{(m)}$ which is the graph of a function $\phi_{(m)} : \tilde{K} \to \mathbf{R}^{N_f}$ over z^1. The functions $\phi_{(m)}$ and ϕ_ε are $\mathcal{O}(\varepsilon^{m+1})-$close,

$$\|\phi_{(m)}(\cdot) - \phi_\varepsilon(\cdot)\| = \mathcal{O}(\varepsilon^{m+1}).$$

Remark. The $(m+1)$st derivative condition has, in general, other solutions as well. The important fact is that, for a given $z_0^1 \in \tilde{K}$, the root $z^2 = \phi_{(m)}(z_0^1)$ of $L_{(m)}(z_0^1, z^2)$ is isolated, and thus the $(m+1)$st derivative condition defines a manifold of the right dimension, namely N_s. Note that, this was automatic for the CSP, but is not in general true for the ZDP if $D_{x^2}z^2$ is singular.

The proof of Theorem 7.2.1 occupies the rest of the chapter. It is based largely on the following analogous result in the $x-$coordinate system.

Proposition 7.2.1 *If $D_{x^2}z^2$ and $D_{x^2}q^2$ are non-singular in a neighborhood of $\mathcal{M}_0$, then the $(m+1)$st derivative condition*

$$L_{(m)}(z(x)) = 0$$

can be solved for x^2 to yield the N_s-dimensional manifold $\mathcal{L}^{(m)}$ as the graph of a function $h_{(m)} : K \to \mathbf{R}^{N_f}$ over x^1. The asymptotic expansions of $h_{(m)}$ and h_ε agree up to and including terms of $\mathcal{O}(\varepsilon^m)$,

$$h_{(m)}(\cdot) = \sum_{i=0}^{\infty} \varepsilon^i h_{(m,i)}(\cdot) = \sum_{i=1}^{m} \varepsilon^i h_i(\cdot) + \mathcal{O}(\varepsilon^{m+1}).$$

Remark. The condition on the non-singularity of $D_{x^2}z^2$ is a mathematically precise restatement of the requirement that the variables z^2 "*somehow contain the fast variables*" found in the beginning of this section.

7.3 Proof of Proposition 7.2.1 for $m = 0, 1$

First, we demonstrate the validity of Proposition 7.2.1 for the cases $m = 0$ and $m = 1$. This will provide us with valuable intuition on proving the proposition for the case of a general m.

7.3.1 The $m = 0$ case

First, we show that, for a given $x^1 \in K$, the function $L_{(0)}(z(x^1, x^2))$ does indeed have a root $x^2 = h_{(0)}(x^1) = h_0(x^1) + \mathcal{O}(\varepsilon)$. For $m = 0$, the definition (7.3), chain rule, and the ODEs (2.2) yield

$$L_{(0)} = \left(\frac{dz^2}{dt}\right) = \left(D_{x^1} z^2\right)(\varepsilon g^1) + \left(D_{x^2} z^2\right) g^2. \tag{7.5}$$

Substituting the asymptotic expansion $x^2 = h_{(0)}(x^1) = \sum_{i=0} \varepsilon^i h_{(0,i)}(x^1)$ into this formula, we obtain to leading order

$$\left(D_{x^2} z^2(x^1, h_{(0,0)}(x^1))\right) g^2(x^1, h_{(0,0)}(x^1), 0) = 0,$$

The invariance equation (2.13), in comparison, yields to leading order

$$g^2\left(x^1, h_0(x^1), 0\right) = 0. \tag{7.6}$$

Thus, $h_{(0,0)}$ can be chosen to be equal to h_0, and $L_{(0)}$ has a root that is $\mathcal{O}(\varepsilon)-$close to the slow manifold.

Now that the proximity of $h_{(0)}$ to h_ε has been established, it remains to show that the condition $L_{(0)} = 0$ can be *actually solved* for x^2 to yield an N_s-dimensional manifold $\mathcal{L}^{(0)}$. Both of these properties follow by the Implicit Function Theorem, and [35, Theorem 1.13], respectively, provided that

$$\det\left(D_{x^2} L_{(0)}\right)(\cdot, h_{(0)}(\cdot)) \neq 0.$$

Using Eq. (7.5), now, we calculate

$$\left(D_{x^2} L_{(0)}\right) = \varepsilon\left[\left(D^2_{x^1 x^2} z^2\right) g^1 + \left(D_{x^1} z^2\right)\left(D_{x^2} g^1\right)\right] + \left(D^2_{x^2} z^2\right) g^2 + \left(D_{x^2} z^2\right)\left(D_{x^2} g^2\right),$$

where all quantities are evaluated at $(\cdot, h_{(0)}(\cdot), \varepsilon)$. The two terms in the square bracket are $\mathcal{O}(\varepsilon)$, as is the third term by virtue of Eq. (7.6). and the equation

$$g^2\left(\cdot, h_{(0)}(\cdot), \varepsilon\right) = g^2\left(\cdot, h_{(0,0)}(\cdot), 0\right) + \mathcal{O}(\varepsilon) = g^2\left(\cdot, h_0(\cdot), 0\right) + \mathcal{O}(\varepsilon)$$

Thus,

$$\left(D_{x^2} L_{(0)}\right)(\cdot, h_{(0)}(\cdot)) = \left[\left(D_{x^2} z^2\right)(\cdot, h_0(\cdot))\right]\left[\left(D_{x^2} g^2\right)(\cdot, h_0(\cdot), 0)\right] + \mathcal{O}(\varepsilon),$$

and so $D_{x^2} L_{(0)}(\cdot, h_{(0)}(\cdot))$ is non-singular for $0 < \varepsilon \ll 1$ by the assumption of Proposition 7.2.1 ($\det(D_{x^2} z^2) \neq 0$) and normal hyperbolicity ($\det(D_{x^2} g^2) \neq 0$).

Remark. Under the assumption that $D_{x^2} z^2$ is invertible in a neighborhood of $\mathcal{M}_0$ (and thus of $\mathcal{L}^{(0)}$), we can rewrite Eq. (7.5) locally as

$$L_{(0)} = \left(D_{x^2} z^2\right) N_{(0)} g, \quad \text{where} \quad N_{(0)} \equiv \left(\left(D_{x^2} z^2\right)^{-1}\left(D_{x^1} z^2\right), I_{N_f}\right). \tag{7.7}$$

This equation is in the form of the invariance equation (2.13). This observation will prove crucial in proving Proposition 7.2.1 for a general value of m.

7.3.2 The $m = 1$ case

For $m = 1$ now, we calculate

$$L_{(1)} = \left(\frac{d^2 z^2}{dt^2}\right) = \frac{d}{dt}\left(\frac{dz^2}{dt}\right) = \frac{d}{dt} L_{(0)}.$$

Using chain rule and recalling the ODEs (2.2) and Eq. (7.5), we rewrite this equation in the form

$$\begin{aligned} L_{(1)} &= \frac{d}{dt}\left[\varepsilon\left(D_{x^1} z^2\right) g^1 + \left(D_{x^2} z^2\right) g^2\right] \\ &= \varepsilon D_{x^1}\left[\varepsilon\left(D_{x^1} z^2\right) g^1 + \left(D_{x^2} z^2\right) g^2\right] g^1 \\ &\quad + D_{x^2}\left[\varepsilon\left(D_{x^1} z^2\right) g^1 + \left(D_{x^2} z^2\right) g^2\right] g^2. \end{aligned} \tag{7.8}$$

To leading order, this equation yields

$$L_{(1)} = \left(D^2_{x^2} z^2\right)(g_0^2, g_0^2) + \left(D_{x^2} g^2\right) g_0^2, \quad \text{where} \quad g_0^2 \equiv g^2(\cdot, 0).$$

and therefore $x^2 = h_0(x^1)$ is a root of $L_{(1)}$ to leading order by Eq. (7.6). Thus, $h_{(1,0)}$ can be selected to be equal to h_0.

Next, we show that $h_{(1,1)} = h_1$. First, we note that the premultiplier of g^2,

$$D_{x^2}\left[\varepsilon\left(D_{x^1} z^2\right) g^1 + \left(D_{x^2} z^2\right) g^2\right],$$

in invertible when evaluated at $(x^1, h_{(1)})$. Indeed, to leading order we calculate

$$D_{x^2}\left[\varepsilon\left(D_{x^1} z^2\right) g^1 + \left(D_{x^2} z^2\right) g^2\right](x^1, h_{(0)}(x^1)) = \left(D^2_{x^2} z^2\right) g^2 + \left(D_{x^2} z^2\right)\left(D_{x^2} g^2\right)$$

with the terms in the right member evaluated at $(x^1, h_{(1,0)}(x^1), 0)$. Recalling that $h_{(1,0)} = h_0$ and Eq. (7.6), then, we obtain

$$\begin{aligned} &D_{x^2}\left[\varepsilon\left(D_{x^1} z^2\right) g^1 + \left(D_{x^2} z^2\right) g^2\right](x^1, h_{(0)}(x^1)) = \\ &\quad \left[\left(D_{x^2} z^2\right)(x^1, h_0(x^1))\right]\left[\left(D_{x^2} g^2\right)(x^1, h_0(x^1), 0)\right] + \mathcal{O}(\varepsilon). \end{aligned} \tag{7.9}$$

Both matrices in the right member of this equation are invertible by virtue of our assumptions, and thus the prefactor is invertible as well for small values of ε. Using this information, we rewrite Eq. (7.8) as

$$L_{(1)} = D_{x^2}\left[\varepsilon(D_{x^1} z^2) g^1 + (D_{x^2} z^2) g^2\right] N_{(1)} g, \tag{7.10}$$

where we have introduced the notation

$$N_{(1)} = \left(\left(D_{x^2}\left[\varepsilon(D_{x^1} z^2) g^1 + (D_{x^2} z^2) g^2\right]\right)^{-1}\left(D_{x^1}\left[\varepsilon(D_{x^1} z^2) g^1 + (D_{x^2} z^2) g^2\right], I_{N_f}\right)\right).$$

Substituting the asymptotic expansion $x^2 = h_{(1)}(x^1)$ into Eq. (7.10), removing the invertible prefactor, and using the above expression for $N_{(1)}$, we obtain

$$\left(N^1_{(1)}, I_{N_f}\right) g = 0, \tag{7.11}$$

where

$$N^1_{(1)} \equiv \left[D_{x^2}\left[\varepsilon(D_{x^1}z^2)g^1 + (D_{x^2}z^2)g^2\right]\right]^{-1}\left[D_{x^1}\left[\varepsilon(D_{x^1}z^2)g^1 + (D_{x^2}z^2)g^2\right]\right]$$

and all quantities are evaluated at $(x^1, h_{(1)}(x^1), \varepsilon)$.

To obtain the $\mathcal{O}(\varepsilon)$ components of Eq. (7.11), we need to expand $N_{(1)}$ and g asymptotically. We start with $N_{(1)}$. Using the asymptotic expansions

$$\begin{aligned}
\left[D_{x^2}\left[\varepsilon(D_{x^1}z^2)g^1 + (D_{x^2}z^2)g^2\right]\right]^{-1}(x^1, h_{(1)}(x^1), \varepsilon) &= \left[D_{x^2}\left[(D_{x^2}z^2)g^2\right]\right]_0^{-1} + \mathcal{O}(\varepsilon),\\
\left[D_{x^1}\left[\varepsilon(D_{x^1}z^2)g^1 + (D_{x^2}z^2)g^2\right]\right](x^1, h_{(1)}(x^1), \varepsilon) &= \left[D_{x^1}\left[(D_{x^2}z^2)g^2\right]\right]_0 + \mathcal{O}(\varepsilon),
\end{aligned}$$

where

$$\begin{aligned}
\left[D_{x^2}\left[(D_{x^2}z^2)g^2\right]\right]_0 &= \left[D_{x^2}\left[(D_{x^2}z^2)g^2\right]\right](x^1, h_{(1,0)}(x^1), 0),\\
\left[D_{x^1}\left[(D_{x^2}z^2)g^2\right]\right]_0 &= \left[D_{x^1}\left[(D_{x^2}z^2)g^2\right]\right](x^1, h_{(1,0)}(x^1), 0),
\end{aligned}$$

we obtain

$$N_{(1)} = \left(\left(D_y\left[(D_y v)g\right]\right)_0^{-1}\left(D_x\left[(D_y v)g\right]\right)_0 + \varepsilon N^{1,1}_{(1)} + \mathcal{O}(\varepsilon^2), I_{N_v}\right). \tag{7.12}$$

(The term $N^{1,1}_{(1)}$ that appears in the expression above is the coefficient at $\mathcal{O}(\varepsilon)$ of the left block of $N_{(1)}$, and we will see later that an explicit formula for it is not needed.) Next, we expand g asymptotically to find

$$g(x^1, h_{(1)}(x^1), \varepsilon) = \begin{pmatrix} 0 \\ g^2_0(x^1) \end{pmatrix} + \varepsilon \begin{pmatrix} g^1_0(x^1) \\ g^2_1(x^1) \end{pmatrix} + \mathcal{O}(\varepsilon^2), \tag{7.13}$$

where

$$\begin{aligned}
g^2_0(x^1) &= g^2(x^1, h_{(1,0)}(x^1), 0),\\
g^1_0(x^1) &= g^1(x^1, h_{(1,0)}(x^1), 0),\\
g^2_1(x^1) &= (D_\varepsilon g^2)(x^1, h_{(1,0)}(x^1), 0) + \left((D_{x^2}g^2)(x^1, h_{(1,0)}(x^1), 0)\right) h_{(1,1)}(x^1).
\end{aligned}$$

We now write out the components of Eq. (7.11) order by order in ε. At $\mathcal{O}(1)$, we obtain

$$\left(\left[D_{x^2}\left[(D_{x^2}z^2)g^2\right]\right]_0^{-1}\left[D_{x^1}\left[(D_{x^2}z^2)g^2\right]\right]_0, I_{N_f}\right)\begin{pmatrix} 0 \\ g_0(x^1) \end{pmatrix} = 0.$$

We have already seen that this equation reduces to

$$g_0(x^1) = g(x^1, h_{(1,0)}(x^1), 0) = 0,$$

thus allowing us to select $h_{(1,0)} = h_0$. This identification allows us to simplify considerably formula (7.12) and therefore also the comparison of the $\mathcal{O}(\varepsilon)$ components of Eqs. (2.8) and (7.11). Indeed, multiplying the identity $g_0 = 0$ through by $(D_{x^2} z^2)_0$, differentiating the resulting equation with respect to x^1, and recalling that the subscript "$_0$" stands for evaluation at $(x^1, h_{(1,0)}(x^1), 0) = (x^1, h_0(x^1), 0)$, we find

$$\left[D_{x^1}\left[(D_{x^2} v) g^2\right]\right]_0 + \left[D_{x^2}\left[(D_{x^2} v) g^2\right]\right]_0 (Dh_0) = 0.$$

Multiplying through by $[D_{x^1}[(D_{x^1} z^2) g^2]]_0^{-1}$ and separating terms, we obtain

$$\left[D_{x^2}\left[(D_{x^2} v) g^2\right]\right]_0^{-1} \left[D_{x^1}\left[(D_{x^2} v) g^2\right]\right]_0 = -Dh_0,$$

and thus Eq. (7.12) becomes

$$N_{(1)} \;=\; \left(-Dh_0 + \varepsilon N^{1,1}_{(1)} + \mathcal{O}(\varepsilon^2), I_{N_f}\right). \tag{7.14}$$

We are now ready to compare the $\mathcal{O}(\varepsilon)$ components of Eqs. (2.8) and (7.11). Combining Eqs. (7.11), (7.13), and (7.14), we obtain for the $\mathcal{O}(\varepsilon)$ component of Eq. (7.11) the expression

$$\left(-Dh_0(x^1), I_{N_f}\right) \begin{pmatrix} g_0^1(x^1) \\ g_1^2(x^1) \end{pmatrix} + \left(N^{1,1}_{(1)}, 0\right) \begin{pmatrix} 0 \\ g_0^2(x^1) \end{pmatrix} = 0.$$

Since $g_0 = 0$, this equation reduces to

$$g_1^2(x^1) - \left((Dh_0)(x^1)\right) g_0^1(x^1) = 0.$$

Similarly, the invariance equation (2.8) yields to $\mathcal{O}(\varepsilon)$

$$\left(-Dh_0(x^1), I_{N_f}\right) \begin{pmatrix} g_0^1(x^1) \\ g_1^2(x^1) \end{pmatrix} + \left(-Dh_1(x^1), 0\right) \begin{pmatrix} 0 \\ g_0^2(x^1) \end{pmatrix} = 0,$$

which also reduces to

$$g_1^2(x^1) - \left((Dh_0)(x^1)\right) g_0^1(x^1) = 0.$$

The two equations agree, and thus $h_{(1,1)} = h_1$. This concludes the proof of the validity of Proposition 7.2.1 for the case $m = 1$.

It only remains to show that the condition $L_{(1)} = 0$ can be solved for x^2 to yield an N_s-dimensional manifold $\mathcal{L}^{(1)}$. Here again, as was the case for $m = 0$, both of

these properties follow by the Implicit Function Theorem, and [35, Theorem 1.13], respectively, provided that

$$\det\left(D_{x^2}L_{(1)}\right)(\cdot, h_{(1)}(\cdot)) \neq 0.$$

Using Eqs. (7.8)–(7.9), as well as the fact that $h_{(1,0)} = h_0$ and Eq. (7.6) to eliminate $\mathcal{O}(\|g^2\|^2)$ terms, we obtain the leading order formula

$$\left(D_{x^2}L_{(1)}\right) = \left(D_{x^2}z^2\right)\left(D_{x^2}g^2\right)^2 + \mathcal{O}(\varepsilon),$$

where all quantities are evaluated at $(\cdot, h_0(\cdot), 0)$. As was the case for $m = 0$, then, $D_{x^2}L_{(1)}(\cdot, h_{(1)}(\cdot))$ is non-singular by the assumption of Proposition 7.2.1 and normal hyperbolicity.

7.4 Proof of Proposition 7.2.1 for $m > 1$

The proof of Proposition 7.2.1 is by induction. The idea behind the proof is to compare the $(m+1)$st derivative condition (7.4) and the invariance equation (2.8) and show that they are the same (possibly modulo an invertible premultiplier) up to and including terms of $\mathcal{O}(\varepsilon^m)$. To enable this comparison, we first recast the former in a form reminiscent of the latter following the pattern that we established for $m = 0, 1$ (see Eqs. (7.7) and (7.10), respectively). Let $m \geq 0$ be arbitrary but fixed; it follows from definition (7.3) that

$$L_{(m)} = \left(D_x L_{(m-1)}\right) g,$$

and thus, the $(m+1)$st derivative condition (7.4) can be rewritten as

$$\left(D_x L_{(m-1)}\right) g = 0. \tag{7.15}$$

The validity of the induction hypothesis for $m = 0$ and $m = 1$ has been shown in Section 7.3. Thus, it remains to perform the induction step. We assume that Proposition 7.2.1 is true for m, and we seek to show that the condition[2]

$$L_{(m+1)}(x) = \left[(D_x L_{(m)})(x)\right] g(x, \varepsilon) = 0 \tag{7.16}$$

can be solved for x^2 to yield $x^2 = h_{(m+1)}(x^1)$, where

$$h_{(m+1)}(\cdot) = \sum_{i=0}^{m+1} \varepsilon^i h_i(\cdot) + \mathcal{O}(\varepsilon^{m+2}).$$

[2]For the remainder of this section, we abuse notation by writing $L_{(m)}(x)$ instead of $L_{(m)}(z(x))$.

The result will be established in two steps. In step 1, we modify Eq. (7.16) so that the gradient matrix $D_x L_{(m)}$ is evaluated on $\mathcal{L}^{(m)}$ (already determined at the mth iteration) instead of on $\mathcal{L}^{(m+1)}$, and we show that the solution to

$$\left[(D_x L_{(m)})(x^1, h_{(m)}(x^1))\right] g(x, \varepsilon) = 0,$$

in which the unknown x^2 only appears in g, approximates h_ε up to and including $\mathcal{O}(\varepsilon^{m+1})$ terms. This motivates the following lemma.

Lemma 7.4.1 *The condition*

$$\left[(D_x L_{(m)})(x^1, h_{(m)}(x^1))\right] g(x, \varepsilon) = 0$$

can be solved for x^2 to yield

$$x^2 = \tilde{h}_{(m+1)}(x^1) = \sum_{i=0}^{m+1} \varepsilon^i h_i(x) + \mathcal{O}(\varepsilon^{m+2}), \quad \textit{for all} \quad x^1 \in K.$$

Then, in step 2, we show that the asymptotic expansions of $h_{(m+1)}$ and $\tilde{h}_{(m+1)}$ agree up to and including terms of $\mathcal{O}(\varepsilon^{m+1})$. In particular, we establish the following lemma.

Lemma 7.4.2 *The condition (7.15),*

$$\left[(D_x L_{(m)})(x, \varepsilon)\right] g(x, \varepsilon) = 0,$$

can be solved for x^2 to yield

$$x^2 = h_{(m+1)}(x^1) = \sum_{i=0}^{m+1} \varepsilon^i \tilde{h}_{(m+1,i)}(x^1) + \mathcal{O}(\varepsilon^{m+2}), \quad \textit{for all} \quad x^1 \in K.$$

Given these lemmata, Proposition 7.2.1 follows as a corollary.

The proofs of Lemmata 7.4.1 and 7.4.2 require the following crucial result that establishes a connection between the order in ε to which a set N of row vectors approximates $\mathcal{NM}_\varepsilon$ and the order to which the solution $\psi(x)$ to the condition $N\,g = 0$ approximates h_ε. In what follows, we write Q for $(x^1, h_{(m)}(x^1))$, $\tilde{Q}$ for $(x^1, \tilde{h}_{(m)}(x^1))$, and Q' for $(x^1, h_\varepsilon(x^1))$.[3]

[3]This proposition is a restatement of the "*one higher order phenomenon*" that we observed in chapters 4 and 5. There, $B^1_{(m)}$ spanned $\mathcal{N}_{Q'}\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^{m-1})$ — see Eqs. (5.7) and (5.9) — while the asymptotic expansion of the CSPM of order m, obtained via the CSP condition $B^1_{(m)} g = 0$, agreed with that of h_ε up to and including terms of $\mathcal{O}(\varepsilon^m)$ — see Theorem 3.3.1.

Proposition 7.4.1 *Let $N(x^1, \varepsilon)$ be an $N_f \times N$ matrix with the property that its rows span $\mathcal{N}_{Q'}\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^m)$, i.e., with the property that there exists a non-singular $N_f \times N_f$ matrix C such that*

$$N(\cdot, \varepsilon) = C \left(-\sum_{i=0}^{m} \varepsilon^i Dh_i(\cdot) + \mathcal{O}(\varepsilon^{m+1}), I_{N_f} \right), \quad \det(C) \neq 0. \tag{7.17}$$

Then, the condition

$$N(x^1, \varepsilon) g(x, \varepsilon) = 0 \tag{7.18}$$

can be solved for x^2 to yield a function $x^2 = \psi(x^1)(x^1)$ the asymptotic expansion of which agrees with that of $h_\varepsilon(x^1)$ up to and including terms of $\mathcal{O}(\varepsilon^{m+1})$,

$$\psi(\cdot) = \sum_{i=0} \varepsilon^i \psi_i(\cdot) = \sum_{i=0}^{m+1} \varepsilon^i h_i(\cdot) + \mathcal{O}(\varepsilon^{m+2}). \tag{7.19}$$

Proof of Proposition 7.4.1. To prove this proposition, we recall that h_i is determined from the $\mathcal{O}(\varepsilon^i)$ terms of the invariance equation and that, similarly, ψ_i is determined from the $\mathcal{O}(\varepsilon^i)$ terms of Eq. (7.18). Thus, to establish Eq. (7.19), it suffices to compare the $\mathcal{O}(1)$ through $\mathcal{O}(\varepsilon^{m+1})$ terms of these two equations and show that they are equal.

At $\mathcal{O}(\varepsilon^i)$, the invariance equation yields

$$\left(-Dh_0, I_{N_f}\right) g_i + \sum_{j=1}^{i} \left(-Dh_j, 0\right) g_{i-j} = 0. \tag{7.20}$$

Now, substituting the hypothesis (7.17) in the condition (7.18) and left-multiplying by C^{-1}, we obtain

$$C^{-1}\, N\, g = \left(-\sum_{i=0}^{m} \varepsilon^i Dh_i + \mathcal{O}(\varepsilon^{m+1}), I_{N_f} \right) g = 0. \tag{7.21}$$

Eq. (7.21) yields

$$\left(-Dh_0, I_{N_f}\right) g_i + \sum_{j=1}^{i} \left(-Dh_j, 0\right) g_{i-j} = 0, \quad \text{for} \quad i = 0, 1, \ldots, m, \tag{7.22}$$

which is identical to Eq. (7.20). Thus, $\psi_i = h_i$, for $i = 0, 1, \ldots, m$.

Last, we look at the $\mathcal{O}(\varepsilon^{m+1})$ terms of the two equations. Eq. (7.20) yields

$$\left(-Dh_0, I_{N_f}\right) g_{m+1} + \sum_{j=1}^{m} \left(-Dh_j, 0\right) g_{m+1-j} + \left(-Dh_{m+1}, 0\right) g_0 = 0. \tag{7.23}$$

Writing $\sum_{i=m+1} \varepsilon^i R_i$ for the remainder of $\mathcal{O}(\varepsilon^{m+1})$ in the expression (7.17), we find that Eq. (7.22) yields, at $\mathcal{O}(\varepsilon^{m+1})$,

$$\left(-Dh_0, I_{N_f}\right) g_{m+1} + \sum_{j=1}^{m} \left(-Dh_j, 0\right) g_{m+1-j} + \left(R_{m+1}, 0\right) g_0 = 0, \tag{7.24}$$

where $R_{m+1} \neq -Dh_{m+1}$, in general. Recalling that the asymptotic terms appearing in Eqs. (7.23) and (7.24) are evaluated at $(x, \psi_0, 0) \equiv (x, h_0, 0)$, though, we conclude that $g_0 = 0$. Thus, Eqs. (7.23) and (7.24) also agree, and therefore $\psi_{m+1} = h_{m+1}$. This completes the proof of the proposition. ∎

Proof of Lemma 7.4.1. To prove the lemma, we will show that the rows of $(D_x L_{(m)})(x^1, h_{(m)}(x^1), \varepsilon)$ span $\mathcal{N}_{Q'}\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^m)$. The result then follows from Proposition 7.4.1.

First, we observe that the manifold $\mathcal{L}^{(m)}$ is the graph of the function $h_{(m)}(\cdot)$. Therefore, it coincides exactly with the zero level set (or at least with a connected component of it) of the function $-h_{(m)}(x^1) + x^2$ and, as a result, the rows of the $N_f \times N$ gradient matrix $(-Dh_{(m)}(x^1), I_{N_f})$ form a basis for $\mathcal{N}_Q\mathcal{L}^{(m)}$. Second, the function $h_{(m)}(\cdot)$ is defined through the equation $L_{(m)}(\cdot, h_{(m)}(\cdot), \varepsilon) = 0$. Therefore, $\mathcal{L}^{(m)}$ is also the zero level set of the function $L_{(m)}(x, \varepsilon)$, and thus the rows of the $N_f \times N$ gradient matrix $(D_x L_{(m)})(x^1, h_{(m)}(x^1), \varepsilon)$ also form a basis for $\mathcal{N}_Q\mathcal{L}^{(m)}$. It follows that there exists a $N_f \times N_f$ non-singular matrix C such that

$$\left(D_x L_{(m)}\right)(\cdot, h_{(m)}(\cdot), \varepsilon) = C\left(-Dh_{(m)}(\cdot), I_{N_f}\right). \tag{7.25}$$

Next, the induction hypothesis implies that the asymptotic expansions of $h_{(m)}$ and h_ε agree up to and including terms of $\mathcal{O}(\varepsilon^m)$,

$$h_{(m)}(\cdot) = \sum_{i=0}^{m} \varepsilon^i h_i(\cdot) + \mathcal{O}(\varepsilon^{m+1}). \tag{7.26}$$

Since the vector field is assumed to be sufficiently smooth, we may differentiate both sides of Eq. (7.26) with respect to x^1 to obtain

$$Dh_{(m)}(\cdot) = \sum_{i=0}^{m} \varepsilon^i Dh_i(\cdot) + \mathcal{O}(\varepsilon^{m+1}). \tag{7.27}$$

Finally, combining Eqs. (7.25) and (7.27) we find

$$\left(D_x L_{(m)}\right)(\cdot, h_{(m)}(\cdot), \varepsilon) = C\left(-\sum_{i=0}^{m} \varepsilon^i Dh_i(\cdot) + \mathcal{O}(\varepsilon^{m+1}), I_{N_f}\right). \tag{7.28}$$

Eq. (7.28) shows that the rows of $(D_x L_{(m)})(x^1, h_{(m)}(x^1), \varepsilon)$ span $\mathcal{N}_{Q'}\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^m)$, we desired. ∎

Proof of Lemma 7.4.2. First, we use [1, Theorem 3] to establish that the condition (7.15) has a solution $x^2 = h_{(m+1)}(x^1)$ with $\tilde{h}_{(m+1)}$ approximating $h_{(m+1)}$ up to and including terms of $\mathcal{O}(\varepsilon^{m+1})$. According to that theorem, it suffices to show that

$$\Big((D_x L_{(m)})(x^1, \tilde{h}_{(m+1)}(x^1), \varepsilon)\Big)\, g(x^1, \tilde{h}_{(m+1)}(x^1), \varepsilon) = \mathcal{O}(\varepsilon^{m+2}).$$

By definition of $\tilde{h}_{(m+1)}$,

$$\big((D_x L_{(m)})(\cdot, h_{(m)}(\cdot), \varepsilon)\big)\, g(\cdot, \tilde{h}_{(m+1)}(\cdot), \varepsilon) = 0.$$

Thus, we may write

$$\begin{aligned}
&\Big((D_x L_{(m)})(\cdot, \tilde{h}_{(m+1)}(\cdot), \varepsilon)\Big)\, g(\cdot, \tilde{h}_{(m+1)}(\cdot), \varepsilon) \\
&\quad = \quad \Big[\Big((D_x L_{(m)})(\cdot, \tilde{h}_{(m+1)}(\cdot), \varepsilon)\Big) - \big((D_x L_{(m)})(\cdot, h_{(m)}(\cdot), \varepsilon)\big)\Big]\, g(\cdot, \tilde{h}_{(m+1)}(\cdot), \varepsilon). \quad (7.29)
\end{aligned}$$

Next, we estimate the asymptotic magnitudes of the two terms in the right member of Eq. (7.29). First,

$$\tilde{h}_{(m+1)} = \sum_{i=0}^{m+1} \varepsilon^i h_i + \mathcal{O}(\varepsilon^{m+2})$$

by Lemma 7.4.1 and also

$$h_{(m)} = \sum_{i=0}^{m} \varepsilon^i h_i + \mathcal{O}(\varepsilon^{m+1})$$

by the induction hypothesis. Thus,

$$\tilde{h}_{(m+1)} - h_{(m)} = \mathcal{O}(\varepsilon^{m+1}),$$

and since $L_{(m)}$ and its derivatives are $\mathcal{O}(1)$, Taylor's Theorem with remainder yields

$$\Big((D_x L_{(m)})(\cdot, \tilde{h}_{(m+1)}(\cdot), \varepsilon)\Big) - \big((D_x L_{(m)})(\cdot, h_{(m)}(\cdot), \varepsilon)\big) = \mathcal{O}(\varepsilon^{m+1}).$$

It remains to estimate the second term, $g(\cdot, \tilde{h}_{(m+1)}(\cdot), \varepsilon)$. Since $\tilde{h}_{(m+1,0)} = h_0$ by Lemma 7.4.1, g is zero to leading order and thus $g(\cdot, \tilde{h}_{(m+1)}(\cdot), \varepsilon) = \mathcal{O}(\varepsilon)$. Combining these estimates, we see that the right member of Eq. (7.29) is $\mathcal{O}(\varepsilon^{m+2})$, which is the desired estimate.

Finally, we use the Implicit Function Theorem and [35, Theorem 1.13] to prove that the condition $L_{(m+1)} = 0$ can be solved for x^2 to yield the N_s-dimensional manifold $\mathcal{L}^{(m+1)}$. It suffices to show that

$$\det\left(D_{x^2} L_{(m+1)}\right)(\cdot, h_{(m+1)}(\cdot)) \neq 0.$$

Using Lemma 9.5.1, and noting that the terms

$$(D^2_{x^2}z^2)\left(\left(D_{x^2}g^2\right)_0^m, g_0^2\right) \quad \text{and} \quad m(D_{x^2}z^2)\left(D^2_{x^2}g^2\right)\left(\left(D_{x^2}g^2\right)_0^{m-1}, g_0^2\right)$$

are $\mathcal{O}(\|g^2\|)$, we calculate

$$\left(D_{x^2}L_{(m)}\right)(x) = \left(\left(D_{x^2}z^2\right)(x)\right)\left(\left(D_{x^2}g^2\right)_0(x)\right)^{m+1} + \mathcal{O}\left(\varepsilon, \|g^2\|\right),$$

at a general point x. Recalling that $h_{(m+1,0)} = h_0$ and Eq. (7.6), then, we obtain, to leading order,

$$\left(D_{x^2}L_{(m)}\right)(\cdot, h_{(m+1)}(\cdot)) = \left(\left(D_{x^2}z^2\right)(\cdot, h_0(\cdot))\right)\left(\left(D_{x^2}g^2\right)(\cdot, h_0(\cdot), 0)\right)^{m+1},$$

with a remainder of $\mathcal{O}(\varepsilon)$. Thus, $(D_{x^2}L_{(m)})(x^1, h_{(m+1)}(x^1))$ is invertible for $0 < \varepsilon \ll 1$ by the induction hypothesis and normal hyperbolicity, and the proof is complete. ∎

7.5 Proof of Theorem 7.2.1

In this section, we prove Theorem 7.2.1. First, in Section 7.5.1 we establish that the solution $z^2 = \phi_{(m)}(z^1)$ to the $(m+1)$st derivative condition is $\mathcal{O}(\varepsilon^{m+1})-$close to $\phi_\varepsilon(z^1)$. Then, in Section 7.5.2, we show that $\mathcal{L}^{(m)}$ has the structure of an N_s-dimensional manifold, and thus the particular solution $z^2 = \phi_{(m)}(z^1)$ to the $(m+1)$st derivative condition is isolated.

7.5.1 Proximity of $\phi_{(m)}$ to ϕ_ε

Let m be a non-negative integer and $z_0^1 \in \tilde{K}$. In this section, we show that the solution $z^2 = \phi_{(m)}(z_0^1)$ to the $(m+1)$st derivative condition (7.4) is $\mathcal{O}(\varepsilon^{m+1})-$close to $\phi_\varepsilon(z_0^1)$.

First, we introduce the notation

$$\begin{array}{rcccl} x_{(m)} & \equiv & x(z_0^1, \phi_{(m)}(z_0^1)) & = & (x^1_{(m)}, h_{(m)}(x^1_{(m)})) \in \mathcal{L}^{(m)} \\ x_* & \equiv & x(z_0^1, \phi_\varepsilon(z_0^1)) & = & (x^1_*, h_\varepsilon(x^1_*)) \in \mathcal{M}_\varepsilon \end{array} \quad , \tag{7.30}$$

and remark that both of these points are on the hypersurface $z^1(x) = z_0^1$. These definitions yield

$$\phi_{(m)}(z_0^1) = z^2(x_{(m)}) \quad \text{and} \quad \phi_\varepsilon(z_0^1) = z^2(x_*),$$

whence

$$\begin{array}{rcl} \|\phi_{(m)}(z_0^1) - \phi_\varepsilon(z_0^1)\| & = & \|z^2(x_{(m)}) - z^2(x_*)\| \\ & \leq & C\left(\|x^1_{(m)} - x^1_*\| + \|x^2_{(m)} - x^2_*\|\right). \end{array} \tag{7.31}$$

Here, C is an $\mathcal{O}(1)$ constant and the inequality holds by virtue of the transformation (2.18) having been assumed to be Lipschitz continuous. Then, using the definitions collected in Eq. (7.30), we rewrite the term in parentheses as

$$\begin{aligned}
&\|x^1_{(m)} - x^1_*\| + \|x^2_{(m)} - x^2_*\| \\
&\quad = \|x^1_{(m)} - x^1_*\| + \|h_{(m)}(x^1_{(m)}) - h_\varepsilon(x^1_*)\| \\
&\quad \leq \|x^1_{(m)} - x^1_*\| + \|h_{(m)}(x^1_{(m)}) - h_\varepsilon(x^1_{(m)})\| + \|h_\varepsilon(x^1_{(m)}) - h_\varepsilon(x^1_*)\|. \qquad (7.32)
\end{aligned}$$

Next, Dh_ε is uniformly bounded by normal hyperbolicity and the compactness of K, and thus it is also Lipschitz continuous. Therefore, there exists an $\mathcal{O}(1)$ constant $\hat{C}$ such that

$$\|h_\varepsilon(x^1_{(m)}) - h_\varepsilon(x^1_*)\| \leq \hat{C}\|x^1_{(m)} - x^1_*\|.$$

Combining this inequality with formula (7.32) and substituting the result in formula (7.31), we obtain

$$\|\phi_{(m)}(z^1_0) - \phi_\varepsilon(z^1_0)\| \leq C\left((1+\hat{C})\|x^1_{(m)} - x^1_*\| + \|h_{(m)}(x^1_{(m)}) - h_\varepsilon(x^1_{(m)})\|\right). \quad (7.33)$$

It now remains to show that both terms in the right member are $\mathcal{O}(\varepsilon^{m+1})$.

The second term is $\mathcal{O}(\varepsilon^{m+1})$ by Proposition 7.2.1, so we focus on the first one. First, we observe that the points x_* and $x_{(m)}$ are the intersections of the hypersurface $z^1(x) = z^1_0$ with $\mathcal{L}^{(m)}$ and $\mathcal{M}_\varepsilon$, respectively (see the definitions collected in Eq. (7.30)). Since $\mathcal{M}_0 = \{z \mid q^2(z) = 0\}$ (see Lemma 9.4.1) and, also, since $D_{z^2}q^2_0$ is non-singular on $\mathcal{M}_0$, the hypersurface $z^1(x) = z^1_0$ intersects transversally with $\mathcal{M}_0$, and thus also with $\mathcal{M}_\varepsilon$, for $0 < \varepsilon \ll 1$. As a result, x_* and $x_{(m)}$ are $\mathcal{O}(\varepsilon^{m+1})-$close on the hypersurface $z^1(x) = z^1_0$, and thus the first term in the right member of inequality (7.33) is also $\mathcal{O}(\varepsilon^{m+1})$, as desired.

7.5.2 Manifold structure of $\mathcal{L}^{(m)}$

In this section, we let again m be a non-negative integer and prove that $\mathcal{L}^{(m)}$ is an N_s-dimensional manifold. As was the case for $m = 0$ and $m = 1$, this result follows from the Implicit Function Theorem and [35, Theorem 1.13] provided that

$$\det\left(\left(D_{z^2}L_{(m)}\right)(z(p_{(m)}))\right) \neq 0, \quad \text{for all} \quad z^1_0 \in \tilde{K}, \qquad (7.34)$$

where $z(p_{(m)}) \equiv (z^1_0, \phi_{(m)}(z^1_0)) \in \mathcal{L}^{(m)}$. Here, we will show that

$$(D_{z^2}L_{(m)})(z^1_0, \phi_{(m)}(z^1_0)) = \left(D_{x^2}z^2\right)_0 \left(D_{x^2}g^2\right)^{m+1}_0 \left(D_{x^2}z^2\right)^{-1}_0 T_0, \qquad (7.35)$$

to leading order in ε, where

$$T_0 = \left[\left(-D\phi_0, I_{N_f}\right)\begin{pmatrix} \left(D_{x^2}z^1\right)_0 \left(D_{x^2}z^2\right)^{-1}_0 \\ I_{N_f} \end{pmatrix}\right]^{-1} \qquad (7.36)$$

and the notation "$(\cdot)_0$" signifies that the parenthesized quantity is evaluated at the point $z(p) \equiv (z_0^1, \phi_0(z_0^1)) \in \mathcal{M}_0$ and for $\varepsilon = 0$. Eq. (7.34) is, then, immediate by virtue of the assumptions of Theorem 7.2.1 and normal hyperbolicity. We establish Eq. (7.35) in four steps.

Remark. The matrix T_0 is well-defined, since

$$\begin{aligned} \operatorname{span}\left(-D\phi_0(z(p)), I_{N_f}\right) &= \mathcal{N}_p\mathcal{M}_0, \\ \operatorname{span}\begin{pmatrix} \left(D_{x^2}z^1\right)_0 \left(D_{x^2}z^2\right)_0^{-1} \\ I_{N_f} \end{pmatrix} &= \operatorname{span}\left(D_{x^2}z\right)_0 = \mathcal{T}_p\mathcal{F}_0, \end{aligned}$$

and thus

$$\det\left[\left(-D\phi_0, I_{N_f}\right)\begin{pmatrix} \left(D_{x^2}z^1\right)_0 \left(D_{x^2}z^2\right)_0^{-1} \\ I_{N_f} \end{pmatrix}\right] \neq 0$$

by transversality of intersection between $\mathcal{T}_p\mathcal{M}_0$ and $\mathcal{T}_f\mathcal{F}_0$.

Step 1. Lemma 9.5.1 provides us with an explicit leading order expression for $L_{(m)}$. Differentiating both members of that expression with respect to z^2 and observing that $D_{z^2}(\|g^2\|^2) = \mathcal{O}(\|g^2\|)$ and also that differentiation of $\mathcal{O}(\varepsilon)$ terms with respect to z^2 does not change their asymptotic magnitude, we obtain

$$\begin{aligned} D_{z^2}L_{(m)} &= \left(D_{x^1}L_{(m)}\right)\left(D_{z^2}x^1\right) + \left(D_{x^2}L_{(m)}\right)\left(D_{z^2}x^2\right) \\ &= \left(D_{x^1}\left[\left(D_{x^2}z^2\right)\left(D_{x^2}g^2\right)^m g^2\right]\right)\left(D_{z^2}x^1\right) \\ &\quad + \left(D_{x^2}\left[\left(D_{x^2}z^2\right)\left(D_{x^2}g^2\right)^m g^2\right]\right)\left(D_{z^2}x^2\right) + \mathcal{O}\left(\varepsilon, \|g^2\|\right). \end{aligned}$$

Next, we expand the derivative terms in the right member and absorb the resulting $\mathcal{O}(\|g\|)$ terms in the $\mathcal{O}(\varepsilon, \|g\|)$ remainder to find

$$\begin{aligned} &D_{z^2}L_{(m)} \\ &= \left(D_{x^2}z^2\right)\left(D_{x^2}g^2\right)^m\left[\left(D_{x^1}g^2\right)\left(D_{z^2}x^1\right) + \left(D_{x^2}g^2\right)\left(D_{z^2}x^2\right)\right] + \mathcal{O}\left(\varepsilon, \|g\|\right). \quad (7.37) \end{aligned}$$

We now evaluate both members of Eq. (7.37) at the point $p_{(m)}$. Combining the equation $z(p_{(m)}) = z(p) + \mathcal{O}(\varepsilon)$ (see Proposition 7.2.1) and Eq. (7.6)), we conclude that $g^2(x(p), \varepsilon) = \mathcal{O}(\varepsilon)$. Thus, Eq. (7.37) yields to leading order

$$\begin{aligned} &\left(D_{z^2}L_{(m)}\right)\left(z(p_{(m)})\right) \\ &= \left(D_{x^2}z^2\right)_0\left(D_{x^2}g^2\right)_0^m\left[\left(D_{x^1}g^2\right)_0\left(D_{z^2}x^1\right)_0 + \left(D_{x^2}g^2\right)_0\left(D_{z^2}x^2\right)_0\right], \quad (7.38) \end{aligned}$$

with a remainder of $\mathcal{O}(\varepsilon)$.

Step 2. Next, we rewrite $D_{z^2}L_{(m)}$ as the restriction on the subspace $x^1(z) = x^1_*$ (that is, on $\mathcal{T}_{z_*}\mathcal{F}$) of a map involving the full Jacobian. In particular, we establish the leading order identity

$$\left(D_{z^2}L_{(m)}\right)\left(z(p_{(m)})\right) = \left(D_x z^2\right)_0 \left(D_x g\right)_0^{m+1} \left(D_{z^2}x\right)_0, \tag{7.39}$$

where we remark that the matrices $D_x z^2$, $D_x g$, and $D_{z^2}x$ are of dimensions $N_f \times N$, $N \times N$, and $N \times N_f$, respectively.

To prove this formula, we start by showing that, to leading order,

$$\left(D_x g\right)^{m+1}\left(x(p)\right) = \begin{pmatrix} 0 & 0 \\ \left(D_{x^2}g^2\right)_0^m \left(D_{x^1}g^2\right)_0 & \left(D_{x^2}g^2\right)_0^{m+1} \end{pmatrix}, \tag{7.40}$$

with an $\mathcal{O}(\varepsilon)$ remainder. This proof is by induction. The result holds for $m = 0$, since Eq. (2.2) yields

$$\left(D_x g\right)\left(x(p)\right) = \begin{pmatrix} \varepsilon\left(D_{x^1}g^1\right)_0 & \varepsilon\left(D_{x^2}g^1\right)_0 \\ \left(D_{x^1}g^2\right)_0 & \left(D_{x^2}g^2\right)_0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \left(D_{x^1}g^2\right)_0 & \left(D_{x^2}g^2\right)_0 \end{pmatrix} + \mathcal{O}(\varepsilon).$$

Assuming, then, that the result holds for m, we calculate

$$\left(D_x g\right)^{m+1}\left(x(p)\right) = \left[\left(D_x g\right)\left(x(p)\right)\right]\left[\left(D_x g\right)^m\left(x(p)\right)\right],$$

and Eq. (7.40) follows simply by carrying out the multiplication.

Then, using Eq. (7.40), we calculate

$$\begin{aligned}
&\left(D_x z^2\right)_0 \left(D_x g\right)_0^{m+1} \left(D_{z^2}x\right)_0 \\
&= \left(D_{x^1}z^2, D_{x^2}z^2\right) \begin{pmatrix} 0 & 0 \\ \left(D_{x^2}g^2\right)_0^m \left(D_{x^1}g^2\right)_0 & \left(D_{x^2}g^2\right)_0^{m+1} \end{pmatrix} \begin{pmatrix} D_{z^2}x^1 \\ D_{x^1}x^2 \end{pmatrix} \\
&= \left(D_{x^2}z^2\right)\left(D_{x^2}g^2\right)_0^m \left(\left(D_{x^1}g^2\right)_0, \left(D_{x^2}g^2\right)_0\right) \begin{pmatrix} D_{z^2}x^1 \\ D_{x^1}x^2 \end{pmatrix} \\
&= \left(D_{x^2}z^2\right)\left(D_{x^2}g^2\right)_0^m \left[\left(D_{x^1}g^2\right)_0 \left(D_{z^2}x^1\right) + \left(D_{x^2}g^2\right)_0 \left(D_{x^1}x^2\right)\right].
\end{aligned}$$

This equation, in conjunction with Eq. (7.38), yields Eq. (7.39) as desired.

Step 3. To obtain Eq. (7.35) from Eq. (7.39), it is necessary to use the fact that $\mathrm{Ker}((D_x g)(x(p)))_0 = \mathcal{T}_p\mathcal{M}_0$ (see Lemma 9.3.1). In order to do that, we must first decompose the columns of the matrix on which $(D_x g)^{m+1}$ acts, namely $(D_{z^2}x)_0$, into their leading order fast and slow components.

Lemma 9.6.1 states that such a leading order fast–slow decomposition is given by

$$\begin{aligned}
\left(D_{z^2}x\right)_0 &= \left(D_z x\right)_0 \begin{pmatrix} 0 \\ I_{N_f} \end{pmatrix} \\
&= \left(D_z x\right)_0 A^0 B_0 \begin{pmatrix} 0 \\ I_{N_f} \end{pmatrix} \\
&= \left(D_z x\right)_0 \left[A_f^0 B_0^{s\perp} \begin{pmatrix} 0 \\ I_{N_f} \end{pmatrix} + A_s^0 B_0^{f\perp} \begin{pmatrix} 0 \\ I_{N_f} \end{pmatrix}\right],
\end{aligned} \tag{7.41}$$

where A^0 and B_0 are defined in that lemma and an explicit formula for A^0 is given in Lemma 9.6.2. Using the formula for B_0 from Lemma 9.6.1, we calculate

$$B_0^{s\perp}\begin{pmatrix} 0 \\ I_{N_f} \end{pmatrix} = I_{N_f} \quad \text{and} \quad B_0^{f\perp}\begin{pmatrix} 0 \\ I_{N_f} \end{pmatrix} = \left(D_{z^2}x^1\right)_0,$$

and thus, Eq. (7.41) becomes

$$(D_{z^2}x)_0 = (D_z x)_0 \left[A_f^0 + A_s^0 \left(D_{z^2}x^1\right)_0\right]. \tag{7.42}$$

Next, we use the formula for A^0 from Lemma 9.6.2 to find

$$\begin{aligned}
(D_z x)_0 A_f^0 &= \begin{pmatrix} (D_z x^1)_0 \\ (D_z x^2)_0 \end{pmatrix} (D_{x^2}z)_0 \left(D_{x^2}z^2\right)_0^{-1} T_0 \\
&= \begin{pmatrix} (D_{x^2}x^1)_0 \\ (D_{x^1}x^1)_0 \end{pmatrix} \left(D_{x^2}z^2\right)_0^{-1} T_0 \\
&= \begin{pmatrix} 0 \\ I_{N_f} \end{pmatrix} \left(D_{x^2}z^2\right)_0^{-1} T_0
\end{aligned} \tag{7.43}$$

and

$$\begin{aligned}
(D_z x)_0 A_s^0 &= \begin{pmatrix} (D_z x^1)_0 \\ (D_z x^2)_0 \end{pmatrix} \left[(D_{x^1}z)_0 + (D_{x^2}z)_0 (Dh_0)\right] \\
&= \begin{pmatrix} (D_{x^1}x^1)_0 \\ (D_{x^1}x^2)_0 \end{pmatrix} + \begin{pmatrix} (D_{x^2}x^1)_0 \\ (D_{x^2}x^2)_0 \end{pmatrix} (Dh_0) \\
&= \begin{pmatrix} I_{N_s} \\ Dh_0 \end{pmatrix}.
\end{aligned} \tag{7.44}$$

Substituting the results of Equations (7.43)–(7.44) into Eq. (7.42), we obtain the desired leading order fast–slow decomposition of $(D_{z^2}x)_0$,

$$(D_{z^2}x)_0 = \begin{pmatrix} 0 \\ I_{N_f} \end{pmatrix} \left(D_{x^2}z^2\right)_0^{-1} T_0 + \begin{pmatrix} I_{N_s} \\ Dh_0 \end{pmatrix}. \tag{7.45}$$

Step 4. In the fourth and final step, we use the decomposition (7.45) derived in Step 3 and Lemma 9.3.1 to rewrite Eq. (7.39).

Substituting for $(D_{z^2}x)_0$ into Eq. (7.39), we obtain

$$\begin{aligned}
&\left(D_{z^2}L_{(m)}\right)\left(z(p_{(m)})\right) \\
&\quad = \left(D_z z^2\right)_0 (D_x g)_0^{m+1} \left[\begin{pmatrix} 0 \\ I_{N_f} \end{pmatrix} \left(D_{x^2}z^2\right)_0^{-1} T_0 + \begin{pmatrix} I_{N_s} \\ Dh_0 \end{pmatrix}\right].
\end{aligned}$$

Since

$$(D_x g)_0^{m+1} \begin{pmatrix} I_{N_s} \\ Dh_0 \end{pmatrix} = 0$$

by Lemma 9.3.1, the equation above simplifies to

$$\left(D_{z^2}L_{(m)}\right)\left(z(p_{(m)})\right) = \left(D_x z^2\right)_0 \left(D_x g\right)_0^{m+1} \begin{pmatrix} 0 \\ I_{N_f} \end{pmatrix} \left(D_{x^2} z^2\right)_0^{-1} T_0. \tag{7.46}$$

Using Eq. (7.40), we calculate

$$\left(D_x g\right)_0^{m+1} \begin{pmatrix} 0 \\ I_{N_f} \end{pmatrix} = \begin{pmatrix} 0 \\ \left(D_{x^2} g^2\right)_0^{m+1} \end{pmatrix},$$

and substituting this result into Eq. (7.46), we obtain

$$\begin{aligned} \left(D_{z^2}L_{(m)}\right)\left(z(p_{(m)})\right) &= \left(\left(D_{x^1} z^2\right)_0, \left(D_{x^2} z^2\right)_0\right) \begin{pmatrix} 0 \\ \left(D_{x^2} g^2\right)_0^{m+1} \end{pmatrix} \left(D_{x^2} z^2\right)_0^{-1} T_0 \\ &= \left(D_{x^2} z^2\right)_0 \left(D_{x^2} g^2\right)_0^{m+1} \left(D_{x^2} z^2\right)_0^{-1} T_0. \end{aligned}$$

This result validates the desired formula (7.35).

Chapter 8

A Unifying Framework For The CSP And The ZDP

In this chapter, we establish two results for CSP (Section 8.1) and one for ZDP (Section 8.2), showing at the same time how the two techniques are related. For CSP, we first show that the update rules, by which the fast and slow components of the vector field become decoupled, can be streamlined. We find that, along with the original vector field and a basis to initiate the iterative process, only the Jacobian of the current vector field is needed to compute the next iterate of the vector field. Second, we show that the update rule for the vector field can induce a sequence of coordinate changes in the state space. Thus, the CSP reduction in the tangent bundle $\mathcal{T}\mathbf{R}^N$ induces a corresponding reduction in the state space $\mathbf{R}^N$ (via a sequence of successively improved approximations to the slow manifold), and this induced reduction is similar to that attained by ZDP. For ZDP, we establish an analogous result, but one that goes in the opposite direction because ZDP works directly in the state space. In particular, we show that the iteration scheme for finding successive approximations of the slow manifold in the state space naturally induces a sequence of coordinate changes in the tangent bundle. This induced sequence, in turn, decouples the dynamics of the fast and slow components of the vector field, in much the same way CSP functions.

Therefore, although CSP and ZDP approach the problem of approximating slow manifolds from opposite directions, they are similar when considered as algorithms to generate coordinate systems in the tangent bundle, an observation that provides a framework to simultaneously track the dynamics in the state space and in the tangent bundle. In fact, from the joint perspective offered by the tangent bundle, CSP and ZDP can also be understood in a more traditional manner as generating, at each iteration, approximate normal forms for the system under examination.

8.1 Computational Singular Perturbation

As we have already mentioned in Chapter 3, the CSP method focuses on the vector field g and is, in essence, an iterative process to reduce the system (2.1) to normal form. At each step of the iterative process, one constructs a transformation $y = y(z)$ and corresponding local bases $D_y z$ and $D_z y$ (cf. Sections 2.2 and 2.3) by requiring that the off-diagonal blocks of the then current approximation of the operator Λ be zero. An approximate slow manifold is then defined via the CSP condition (3.14) as the locus of all points in state space where the vector field lies entirely in the then current approximation of the slow subspace, as in Eq. (2.32).

In the framework of CSP, it is customary to use the notation

$$A = D_y z, \quad B = D_z y$$

and consider A as a matrix of column vectors (vectors in $\mathbf{R}^N$) and B as a matrix of row vectors (functionals on $\mathbf{R}^N$). The dimensions of the fast and slow subspaces are not known a priori, because the fast and slow variables can, in general, not be identified without further knowledge of the system. To emphasize this fact, we will adhere to a practice we effectively introduced in Chapter 2 — we will use the generic sub- and superscripts "1" and "2" as labels for *potentially* fast and slow categories and reserve the subscripts "f" and "s" for ideally fast and slow categories in the sense defined in Section 2.3 (and similarly for the superscripts "$f^\perp$" and "$s^\perp$"). Under this convention, the matrices A and B are partitioned as in Eq. (3.3) (following the partition (2.28) of y), where $A_1 = D_{y^1} z$, $A_2 = D_{y^2} z$, $B^1 = D_z y^1$, and $B^2 = D_z y^2$.

After m iterations, the state of the system is represented by the vector $y_{(m)} = y_{(m)}(z)$, the current local basis $A^{(m)} = D_{y_{(m)}} z$, and its dual $B_{(m)} = D_x y_{(m)}$. The vector field is expressed as $q = A^{(m)} f_{(m)}$, with $f_{(m)} = B_{(m)} q$. The evolution of $y_{(m)}$ is governed by the equation $y'_{(m)} = f_{(m)}$ and that of $f_{(m)}$ by the equation $f'_{(m)} = \Lambda_{(m)} f_{(m)}$, where $\Lambda_{(m)} = D_{y_{(m)}} f_{(m)} = B_{(m)}[A^{(m)}, q]$ see Eqs. (2.24), (2.26) and (2.27). In block form,

$$\Lambda_{(m)} = \begin{pmatrix} \Lambda^{11}_{(m)} & \Lambda^{12}_{(m)} \\ \Lambda^{21}_{(m)} & \Lambda^{22}_{(m)} \end{pmatrix}$$

(cf. Eq. 3.9), with

$$\Lambda^{11}_{(m)} = D_{y^1_{(m)}} f^1_{(m)} = B^1_{(m)}[A^{(m)}_1, q], \quad \Lambda^{12}_{(m)} = D_{y^2_{(m)}} f^1_{(m)} = B^1_{(m)}[A^{(m)}_2, q],$$
$$\Lambda^{21}_{(m)} = D_{y^1_{(m)}} f^2_{(m)} = B^2_{(m)}[A^{(m)}_1, q], \quad \Lambda^{22}_{(m)} = D_{y^2_{(m)}} f^2_{(m)} = B^2_{(m)}[A^{(m)}_2, q].$$

The current approximation to the slow manifold is defined as in Eq. (2.32),

$$\tilde{\mathcal{K}}^{(m)}_\varepsilon = \left\{ z \in \mathbf{R}^N \;\middle|\; f^1_{(m)}(y(z)) = 0 \right\}, \tag{8.1}$$

see also Eq. (3.25).

8.1.1 CSP update rules

In this section, we focus on the vector of "fast" coordinates of q and show that, if one is interested only in the slow manifold, $f^1_{(m+1)}$ may be found recursively in terms of $f^1_{(m)}$ without reference to the basis $A^{(m)}$, its inverse $B^{(m)}$, or the operator $\Lambda_{(m)}$.

In Section 3.2.1, we remarked that the CSP algorithm performs the updating in two steps; the first involves the matrices U, the second the matrices L. Then, in Chapters 3–5, we saw that the two steps play different roles and that, if one is interested solely in approximating the slow manifold, one may as well skip the second step. Since the purpose of this chapter is to compare CSP and ZDP, and the latter focuses entirely on the approximation of the slow manifold, we restrict the following discussion to the one-step CSP (see Section 3.2.2), where the update rule (3.11)–(3.12) is reduced to Eqs. (3.17)–(3.18).

Equations (4.1)–(4.3) yield the explicit rules

$$\tilde{A}^{(m+1)} = \tilde{A}^{(0)} \left(I - \sum_{i=0}^{m} \tilde{U}_{(i)} \right), \quad \tilde{B}_{(m+1)} = \left(I + \sum_{i=0}^{m} \tilde{U}_{(i)} \right) \tilde{B}_{(0)}, \tag{8.2}$$

for $m = 0, 1, \ldots$.[1] The sum $\sum_{i=0}^{m} \tilde{U}_{(i)}$ has only one nonzero block, $\sum_{i=0}^{m} (\tilde{\Lambda}^{11}_{(i)})^{-1} \tilde{\Lambda}^{12}_{(i)}$, in the 12-position. In the remainder of this section, we will use the same notation, $\sum_{i=0}^{m} \tilde{U}_{(i)}$, to denote this nonzero block. Note that the 12-block is rectangular; its size is $N_f \times N_s$.

We partition $\tilde{A}^{(0)}$ as in Eq. (3.3), $\tilde{A}^{(0)} = (\tilde{A}^{(0)}_1, \tilde{A}^{(0)}_2)$, where $\tilde{A}^{(0)}_1$ and $\tilde{A}^{(0)}_2$ are rectangular matrices of size $N \times N_f$ and $N \times N_s$, respectively, and rewrite the expressions (8.2) in the form

$$\tilde{A}^{(m+1)} = \left(\tilde{A}^{(0)}_1, \tilde{A}^{(0)}_2 - \tilde{A}^{(0)}_1 \sum_{i=0}^{m} \tilde{U}_{(i)} \right), \quad m = 0, 1, \ldots, \tag{8.3}$$

$$\tilde{B}_{(m+1)} = \begin{pmatrix} \tilde{B}^1_{(0)} + \left(\sum_{i=0}^{m} \tilde{U}_{(i)} \right) \tilde{B}^2_{(0)} \\ \tilde{B}^2_{(0)} \end{pmatrix}, \quad m = 0, 1, \ldots. \tag{8.4}$$

The sum can be expressed solely in terms of CSP quantities current at the mth iteration. To see this, start from the definition $\tilde{\Lambda}_{(m)} = D_{\tilde{y}_{(m)}} \tilde{f}_{(m)}$, evaluate at $z = z(\tilde{y}_{(m)})$, and use the chain rule to replace the derivative with respect to $\tilde{y}_{(m)}$ by the derivative with respect to z. Since $D_{\tilde{y}_{(m)}} z = A^{(m)}$, we obtain

$$\tilde{\Lambda}_{(m)} = D_{\tilde{y}_{(m)}} \tilde{f}_{(m)} = (D_z \tilde{f}_{(m)})(D_{\tilde{y}_{(m)}} z) = (D_z \tilde{f}_{(m)}) \tilde{A}^{(m)}.$$

[1] For the remainder of our discussion of the CSP method, all one-step CSP quantities will be tilded in accordance with the notation used in Section 3.2.2 and Chapter 5 and also to distinguish them from their ZDP analogs.

If we now substitute $\tilde{A}^{(w)}$ in accordance with Eq. (8.3), we find

$$\begin{aligned}
\tilde{\Lambda}^{11}_{(m)} &= \left(D_z\tilde{f}^1_{(m)}\right)\tilde{A}^{(m)}_1 = \left(D_z\tilde{f}^1_{(m)}\right)\tilde{A}^{(0)}_1, \\
\tilde{\Lambda}^{12}_{(m)} &= \left(D_z\tilde{f}^1_{(m)}\right)\tilde{A}^{(m)}_2 = \left(D_z\tilde{f}^1_{(m)}\right)\tilde{A}^{(0)}_2 - \left(D_z\tilde{f}^1_{(m)}\right)\tilde{A}^{(0)}_1 \sum_{i=0}^{m-1}\tilde{U}_{(i)}.
\end{aligned}$$

Hence,

$$\tilde{U}_{(m)} = \left(\tilde{\Lambda}^{11}_{(m)}\right)^{-1}\tilde{\Lambda}^{12}_{(0)} = \left[\left(D_z\tilde{f}^1_{(m)}\right)\tilde{A}^{(0)}_1\right]^{-1}\left(D_z\tilde{f}^1_{(m)}\right)\tilde{A}^{(0)}_2 - \sum_{i=0}^{m-1}\tilde{U}_{(i)}.$$

It follows that $\sum_{i=0}^{m}\tilde{U}_{(i)}$ is given in terms of $\tilde{f}^1_{(m)}$ by

$$\sum_{i=0}^{m}\tilde{U}_{(i)} = \left[\left(D_z\tilde{f}^1_{(m)}\right)\tilde{A}^{(0)}_1\right]^{-1}\left(D_z\tilde{f}^1_{(m)}\right)\tilde{A}^{(0)}_2.$$

(Note that the matrix inside the square brackets is square ($N_f \times N_f$), but its components are rectangular ($D_z\tilde{f}^1_{(m)}$ is $N_f \times N$, $\tilde{A}^{(0)}_1$ is $N \times N_f$), so the inverse cannot be written as the product of the inverses of the component matrices.) Substituting this result into Eqs. (8.3)–(8.4), we obtain the following expressions for $\tilde{A}^{(m+1)}$ and $\tilde{B}_{(m+1)}$:

$$\tilde{A}^{(m+1)} = \left(\tilde{A}^{(0)}_1\ ,\ \tilde{A}^{(0)}_2 - \tilde{A}^{(0)}_1\left[\left(D_z\tilde{f}^1_{(m)}\right)\tilde{A}^{(0)}_1\right]^{-1}\left(D_z\tilde{f}^1_{(m)}\right)\tilde{A}^{(0)}_2\right), \quad m = 0, 1, \tag{8.5}$$

$$\tilde{B}_{(m+1)} = \begin{pmatrix} \tilde{B}^1_{(0)} + \left[(D_z\tilde{f}^1_{(m)})\tilde{A}^{(0)}_1\right]^{-1}(D_z\tilde{f}^1_{(m)})\tilde{A}^{(0)}_2\tilde{B}^2_{(0)} \\ \tilde{B}^2_{(0)} \end{pmatrix}, \quad m = 0, 1, \ldots . \tag{8.6}$$

If we use the expression (8.6) in the identity $\tilde{f}^1_{(m+1)} = \tilde{B}^1_{(m+1)}q$, we obtain the update rule for the fast coordinates of the vector field,

$$\begin{aligned}
\tilde{f}^1_{(m+1)} &= \left(\tilde{B}^1_{(0)} + \left[\left(D_z\tilde{f}^1_{(m)}\right)\tilde{A}^{(0)}_1\right]^{-1}\left(D_z\tilde{f}^1_{(m)}\right)\tilde{A}^{(0)}_2\tilde{B}^2_{(0)}\right)q \\
&= \left[\left(D_z\tilde{f}^1_{(m)}\right)\tilde{A}^{(0)}_1\right]^{-1}\left(D_z\tilde{f}^1_{(m)}\right)(\tilde{A}^{(0)}_1\tilde{B}^1_{(0)} + \tilde{A}^{(0)}_2\tilde{B}^2_{(0)})q, \quad m = 0, 1, \ldots .
\end{aligned}$$

But $\tilde{A}^{(0)}_1\tilde{B}^1_{(0)} + \tilde{A}^{(0)}_2\tilde{B}^2_{(0)} = \tilde{A}^{(0)}\tilde{B}_{(0)} = I$, so the update rule simplifies to

$$\tilde{f}^1_{(m+1)} = \left[\left(D_z\tilde{f}^1_{(m)}\right)A^{(0)}_1\right]^{-1}\left(D_z\tilde{f}^1_{(m)}\right)q, \quad m = 0, 1, \ldots . \tag{8.7}$$

In the special case where CSP is initialized with the matrices

$$A^{(0)} = \begin{pmatrix} 0 & I_{N_s} \\ I_{N_f} & 0 \end{pmatrix}, \quad B_{(0)} = \begin{pmatrix} 0 & I_{N_f} \\ I_{N_s} & 0 \end{pmatrix}$$

(see the next section for details on this choice), the update rule (8.7) reduces further to

$$\tilde{f}^1_{(m+1)} = (D_{z^2}\tilde{f}^1_{(m)})^{-1}(D_z\tilde{f}^1_{(m)})q, \quad m = 0, 1, \ldots .$$

The rule (8.7) shows that one can obtain $\tilde{f}^1_{(m+1)}$ solely in terms of $\tilde{f}^1_{(m)}$ and q. The same is then true for $\tilde{\mathcal{K}}^{(m+1)}_\varepsilon$, see Eq. (8.1). In other words, one can obtain the sequence $\{\tilde{\mathcal{K}}^{(m)}_\varepsilon | m = 0, 1, \ldots\}$ of approximate manifolds recursively by means of Eqs. (8.1) and (8.7), knowing only $\tilde{A}^{(0)}$ and $\tilde{B}_{(0)}$.

8.1.2 CSP coordinates in state space

We conclude the discussion of CSP by showing that the local basis $\tilde{A}^{(m)}$ can induce a set of coordinates $\tilde{y}_{(m)}$ in the state space. In this way, the CSP reduction of the vector field induces a reduction in the state space.

For simplicity, and to enable a comparison with ZDP, we assume that the CSP method is seeded with the standard coordinate system, $\tilde{y}^1_{(0)}(z) = z^2$ and $\tilde{y}^2_{(0)}(z) = z^1$,[2] and

$$\tilde{A}^{(0)} = \begin{pmatrix} 0 & I_{N_s} \\ I_{N_f} & 0 \end{pmatrix}, \quad \tilde{B}_{(0)} = \begin{pmatrix} 0 & I_{N_f} \\ I_{N_s} & 0 \end{pmatrix}. \tag{8.8}$$

In this case, Eqs. (8.5)–(8.6) reduce to

$$\tilde{A}^{(m+1)} = \begin{pmatrix} 0 & I_{N_s} \\ I_{N_f} & -(D_{z^2}\tilde{f}^1_{(m)})^{-1}(D_{z^1}\tilde{f}^1_{(m)}) \end{pmatrix}, \quad m = 0, 1, \ldots , \tag{8.9}$$

$$\tilde{B}_{(m+1)} = \begin{pmatrix} (D_{z^2}\tilde{f}^1_{(m)})^{-1}(D_{z^1}\tilde{f}^1_{(m)}) & I_{N_f} \\ I_{N_s} & 0 \end{pmatrix}, \quad m = 0, 1, \ldots . \tag{8.10}$$

Furthermore, we assume that $\tilde{\mathcal{K}}^{(m)}_\varepsilon$ is the graph of a function $\tilde{\psi}_{(m)} : K \to \mathbf{R}^{N_f}$ over a compact set $K \subset \mathbf{R}^{N_s}$ (cf. Eq. (3.24)),

$$\tilde{f}^1_{(m)}(y(z^1, \tilde{\psi}_{(m)}(z^1))) = \left(\tilde{B}^1_{(m)}(z^1, \tilde{\psi}_{(m-1)}(z^1))\right) q(z^1, \tilde{\psi}_{(m)}(z^1)) = 0, \quad z^1 \in K. \tag{8.11}$$

[2]The inversion is deliberate; in Chapter 2, we considered fast–slow systems written in such a way that the slow variables come first. The CSP method, on the other hand, puts fast quantities first. Although in this chapter we don't restrict our attention to fast–slow systems, we would like our results to translate directly into something meaningful for the fast–slow case, and thus we treat z^1 as "slow" and z^2 as "fast."

Taking the total derivative of Eq. (8.11) with respect to z^1, we obtain the identity

$$D_{z^1}\tilde{f}^1_{(m)} + (D_{z^2}\tilde{f}^1_{(m)})(D\tilde{\psi}_{(m)}) = 0,$$

where $D\tilde{\psi}_{(m)}$ is the Jacobian of $\tilde{\psi}_{(m)}$ (a rectangular matrix of size $N_f \times N_s$). Hence,

$$(D_{z^2}f^1_{(m)})^{-1}(D_{z^1}f^1_{(m)}) = -D\tilde{\psi}_{(m)} \quad \text{ on } \tilde{\mathcal{K}}^{(m)}_\varepsilon.$$

Consequently, on $\tilde{\mathcal{K}}^{(m)}_\varepsilon$, Eqs. (8.9)–(8.10) reduce to

$$\tilde{A}^{(m+1)} = \begin{pmatrix} 0 & I_{N_s} \\ I_{N_f} & D\tilde{\psi}_{(m)} \end{pmatrix}, \quad \tilde{B}_{(m+1)} = \begin{pmatrix} -D\tilde{\psi}_{(m)} & I_{N_f} \\ I_{N_s} & 0 \end{pmatrix}, \quad m = 0, 1, \ldots . \tag{8.12}$$

We extend these expressions from points on $\tilde{\mathcal{K}}^{(m)}_\varepsilon$ to a neighborhood of $\tilde{\mathcal{K}}^{(m)}_\varepsilon$ by assigning the basis

$$A^{(m+1)}(z) = A^{(m+1)}(z^1, \tilde{\psi}_{(m)}(z^1))$$

to any point $z = (z^1, z^2)$ with $z^1 \in K$ and z^2 close to $\tilde{\psi}_{(m)}(z^1)$.

To see that $\tilde{A}^{(m+1)}$ thus extended induces coordinates in the state space, start from the identity $\tilde{B}_{(m+1)} = D_z y_{(m+1)}$ and substitute $\tilde{B}_{(m+1)}$ from Eq. (8.12),

$$B_{(m+1)} = \begin{pmatrix} D_{z^1}\tilde{y}^1_{(m+1)} & D_{z^2}\tilde{y}^1_{(m+1)} \\ D_{z^1}\tilde{y}^2_{(m+1)} & D_{z^2}\tilde{y}^2_{(m+1)} \end{pmatrix} = \begin{pmatrix} -D\tilde{\psi}_{(m)} & I_{N_f} \\ I_{N_s} & 0 \end{pmatrix}, \quad m = 0, 1, \ldots .$$

These ODEs can be integrated up to additive constants; the latter correspond to a translation of the origin and can be ignored. The result is

$$\begin{pmatrix} \tilde{y}^1_{(m+1)}(z) \\ \tilde{y}^2_{(m+1)}(z) \end{pmatrix} = \begin{pmatrix} z^2 - \tilde{\psi}_{(m)}(z^1) \\ z^1 \end{pmatrix}, \quad m = 0, 1, \ldots .$$

Therefore, each update of the local basis from $\tilde{A}^{(m)}$ to $\tilde{A}^{(m+1)}$ induces a coordinate change in the state space from $y_{(m)}$ to $y_{(m+1)}$.

We emphasize that the above results are particular to the CSP update rules, where the basis $\tilde{A}^{(m)}$ is extended to a neighborhood of $\tilde{\mathcal{K}}^{(m)}_\varepsilon$ as described after Eq. (8.12); otherwise, the basis $\tilde{A}^{(m+1)}$ is not necessarily associated with a coordinate system. (Such a basis is often referred to as a *non-coordinate basis*, see [41]). This can be seen already for $\tilde{A}^{(1)}$. If this basis were associated with some coordinate system (Y^1, Y^2), then Eq. (8.10) would have to be satisfied. By Eq. (8.8), $\tilde{f}^1_{(0)} = q^2$, so Eq. (8.10) yields

$$B_{(1)} = \begin{pmatrix} D_{z^1}Y^1 & D_{z^2}Y^1 \\ D_{z^1}Y^2 & D_{z^2}Y^2 \end{pmatrix} = \begin{pmatrix} (D_{z^2}q^2)^{-1}(D_{z^1}q^2) & I_{N_f} \\ I_{N_s} & 0 \end{pmatrix}.$$

This is a system of $N \times N$ ODEs written in block format. The upper left block yields $Y^1 = z^2 - \omega(z^1)$, for some function ω. Substituting this expression in the upper right

block, we arrive at the equation $D_{z^2}\omega = -(D_{z^2}q^2)^{-1}(D_{z^1}q^2)$. While the left member of this equation for ω depends only on z^1, its right member depends, for a general nonlinear vector field q, on both z^1 and z^2; hence, the equation has, in general, no solution. Therefore, $A^{(1)}$ does not, in general, induce a coordinate system in the state space. A similar argument applies to higher values of m.

8.2 The Zero-Derivative Principle

The zero-derivative principle focuses on the state variables and is, in essence, an iterative process to generate an ideal coordinate system in the state space where the fast and slow dynamics decouple. At each step, one partitions the ZDP coordinates into "fast" and "slow" and constructs a change of coordinates $y = y(z)$ by identifying the new "fast" coordinates with the time derivatives of the old ones. An approximation to the slow manifold is defined as the locus of points in state space where the then-current "fast" variables vanish, as in Eq. (2.29).

As with CSP, the dimensions of the fast and slow subspaces—and therefore also the variables that can be classified as fast and slow—are not known a priori. Therefore, we again use the generic sub- and superscripts "1" and "2", rather than "f" and "s", as labels for potentially fast and slow categories. In the case of a fast–slow system, where the fast and slow variables can be identified a priori and the ratio of the characteristic times of the fast and slow variables is measured by an arbitrarily small parameter ε, ZDP generates the asymptotic expansions of slow manifolds order by order, iterate by iterate [9].

Our main results are contained in Sections 8.2.1 and 8.2.2. In Section 8.2.1 we determine the basis that is induced in the tangent space by the coordinate transformation in the state space. In Section 8.2.2 we derive the update rule for this basis and show that it is similar to the update rule of CSP. Throughout this section, we use the same symbols as in Section 8.1 to denote quantities that have a similar meaning in CSP and ZDP.

8.2.1 Local bases for the ZDP coordinates

In this section, we determine the local basis induced in $\mathcal{T}\mathbf{R}^N$ by the coordinates $y_{(m)}$ and show that, in this basis, $\mathcal{L}^{(m)}$ coincides with the locus of points where the "fast" vector field component current at the mth iteration vanishes, as in Eq. (2.32).

We first consider the case $m = 0$. Here, $y^1_{(0)}(z) = z^2$ and $y^2_{(0)}(z) = z^1$, so

$$D_{y_{(0)}}z = A^{(0)} = \begin{pmatrix} 0 & I_{N_s} \\ I_{N_f} & 0 \end{pmatrix}, \quad D_z y_{(0)} = B^{(0)} = \begin{pmatrix} 0 & I_{N_f} \\ I_{N_s} & 0 \end{pmatrix}.$$

In other words, the initial coordinates $y_{(0)}$ induce the standard basis in the tangent

space. Since $q^2 = (z^2)' = L_{(0)}$, we can write the vector $f_{(0)}$ as

$$f_{(0)} = \begin{pmatrix} L_{(0)} \\ q^1 \end{pmatrix}.$$

For higher values of m, we use the definition (7.2) to find the dual basis,

$$B_{(m)} = \begin{pmatrix} B^1_{(m)} \\ B^2_{(m)} \end{pmatrix} = \begin{pmatrix} D_z y^1_{(m)} \\ D_z y^2_{(m)} \end{pmatrix} = \begin{pmatrix} D_{z^1} L_{(m-1)} & D_{z^2} L_{(m-1)} \\ I_{N_s} & 0 \end{pmatrix}. \tag{8.13}$$

The corresponding basis $A^{(m)}$ is found by inversion,

$$A^{(m)} = \big(B_{(m)}\big)^{-1} = \begin{pmatrix} 0 & I_{N_s} \\ \big(D_{z^2} L_{(m-1)}\big)^{-1} & -\big(D_{z^2} L_{(m-1)}\big)^{-1} \big(D_{z^1} L_{(m-1)}\big) \end{pmatrix}. \tag{8.14}$$

In this basis, the vector field is given by $q = A^{(m)} f_{(m)}$, with

$$f_{(m)} = B_{(m)} q = \begin{pmatrix} (D_z L_{(m-1)}) q \\ q^1 \end{pmatrix}.$$

Since

$$(D_z L_{(m-1)}) q = (L_{(m-1)})' = \left(\frac{d^m z^1}{dt^m}\right)' = \frac{d^{m+1} z^1}{dt^{m+1}} = L_{(m)}, \tag{8.15}$$

it follows that

$$f_{(m)} = \begin{pmatrix} L_{(m)} \\ q^1 \end{pmatrix}, \quad m = 0, 1, \ldots . \tag{8.16}$$

In particular, $f^1_{(m)} = L_{(m)}$, for $m = 0, 1, \ldots$.

Combining this result with the definition (7.4), we obtain yet a third definition of $\mathcal{L}^{(m)}$:

$$\mathcal{L}^{(m)} = \big\{ z \in \mathbf{R}^N \;\big|\; f^1_{(m)}(z) = 0 \big\}, \quad m = 0, 1, \ldots . \tag{8.17}$$

This is the ZDP analog of Eq. (8.1).

We emphasize that the above calculation of $A^{(m)}$ and $B_{(m)}$ and the alternative characterization (8.17) of $\mathcal{L}^{(m)}$ are not part of ZDP; they are introduced here for the sole purpose of comparing ZDP and CSP. In practical implementations of ZDP, the sequence $\{\mathcal{L}^{(m)} | m = 0, 1, \ldots\}$ is obtained by working strictly in the state space, as explained in Section 7.1.

Given the alternative characterization (8.17) of $\mathcal{L}^{(m)}$, one can recursively update f^1 without recourse to computing the bases. Indeed, it follows from Eqs. (8.15) and (8.16) that

$$f^1_{(m+1)} = \big(f^1_{(m)}\big)' = \big(D_z f^1_{(m)}\big)\, q. \tag{8.18}$$

Hence, the "fast" vector field component current at the $(m+1)$th iteration, $f^1_{(m+1)}$, can be obtained in terms of its counterpart at the mth iteration, $f^1_{(m)}$, just as for CSP, see Eq. (8.7). Notice that the update rule (8.18) differs from its CSP analog (8.7) by a scaling factor of Λ^{11}.

Remark. Equations (7.2) and (8.16) show that the update rule (7.1) for the ZDP coordinates can be recast as

$$y_{(m+1)}(z) = \begin{pmatrix} y^1_{(m+1)}(z) \\ y^2_{(m+1)}(z) \end{pmatrix} = \begin{pmatrix} f^1_{(m)}(z) \\ z^1 \end{pmatrix}, \quad m = 0, 1, \ldots .$$

That is, at each iteration, the new "fast" coordinates in the state space are obtained as the old "fast" coordinates of the vector field.

8.2.2 Local basis updates

In this section, we derive the update rules for the ZDP bases $A^{(m)}$ and $B_{(m)}$. They are similar to those of CSP, Eqs. (3.17)–(3.18), and reduce $\Lambda_{(m)}$ to block-diagonal form. The latter is defined as before, $\Lambda_{(m)} = D_{y_{(m)}} f_{(m)} = (D_z f_{(m)}) A^{(m)} = B_{(m)}[A^{(m)}, q]$, and governs the evolution of $f_{(m)}$.

Let $C_{(m)}$ be the invertible matrix updating $A^{(m)}$ to $A^{(m+1)}$,

$$A^{(m+1)} = A^{(m)} C_{(m)}, \quad B^{(m+1)} = (C_{(m)})^{-1} B^{(m)}.$$

Applying $B_{(m)}$ to both sides of the first equation and using the identity $B_{(m)} A^{(m)} = I$, we obtain an expression for the update matrix,

$$C_{(m)} = B_{(m)} A^{(m+1)}.$$

Expressions for $B_{(m)}$ and $A^{(m)}$ are given in Eqs. (8.13) and (8.14), respectively. Carrying out the matrix multiplication, we obtain the following expressions for the blocks of $C_{(m)}$:

$$\begin{aligned}
C^{11}_{(m)} &= \left(D_{z^2} L_{(m-1)}\right) \left(D_{z^2} L_{(m)}\right)^{-1}, \\
C^{12}_{(m)} &= D_{z^1} L_{(m-1)} - \left(D_{z^2} L_{(m-1)}\right) \left(D_{z^2} L_{(m)}\right)^{-1} \left(D_{z^1} L_{(m)}\right), \\
C^{21}_{(m)} &= 0, \\
C^{22}_{(m)} &= I_{N_s}.
\end{aligned}$$

These expressions hold for $m = 0, 1, \ldots$, provided we take

$$L_{(-1)}(z) = \left(\frac{d^0 z^2}{dt^0}\right)(z) = z^2.$$

We now express the blocks of $C_{(m)}$ in terms of the blocks of $\Lambda_{(m)}$. We calculate the latter using the formula $\Lambda_{(m)} = (D_z f_{(m)}) A^{(m)}$, substituting $f_{(m)}$ from Eq. (8.16) and $A^{(m)}$ from Eq. (8.14). Thus, we obtain

$$\begin{aligned}
\Lambda^{11}_{(m)} &= \left(D_{z^2} L_{(m)}\right) \left(D_{z^2} L_{(m-1)}\right)^{-1}, \\
\Lambda^{12}_{(m)} &= D_{z^1} L_{(m)} - \left(D_{z^2} L_{(m)}\right) \left(D_{z^2} L_{(m-1)}\right)^{-1} \left(D_{z^1} L_{(m-1)}\right), \\
\Lambda^{21}_{(m)} &= \left(D_{z^2} q^1\right) \left(D_{z^2} L_{(m-1)}\right)^{-1}, \\
\Lambda^{22}_{(m)} &= D_{z^1} q^1 - \left(D_{z^2} q^1\right) \left(D_{z^2} L_{(m-1)}\right)^{-1} \left(D_{z^1} L_{(m-1)}\right).
\end{aligned}$$

The matrix $C_{(m)}$ can therefore be rewritten as

$$C_{(m)} = \begin{pmatrix} (\Lambda^{11}_{(m)})^{-1} & -(\Lambda^{11}_{(m)})^{-1}\Lambda^{12}_{(m)} \\ 0 & I_{N_s} \end{pmatrix}, \quad m = 0, 1, \ldots, \tag{8.19}$$

and the update rule for the bases becomes

$$A^{(m+1)} = A^{(m)} \begin{pmatrix} (\Lambda^{11}_{(m)})^{-1} & -(\Lambda^{11}_{(m)})^{-1}\Lambda^{12}_{(m)} \\ 0 & I_{N_s} \end{pmatrix}, \quad m = 0, 1, \ldots, \tag{8.20}$$

$$B_{(m+1)} = \begin{pmatrix} \Lambda^{11}_{(m)} & \Lambda^{12}_{(m)} \\ 0 & I_{N_s} \end{pmatrix} B_{(m)}, \quad m = 0, 1, \ldots . \tag{8.21}$$

A comparison of Eqs. (8.20)–(8.21) with their CSP analog, Eqs. (3.17)–(3.18), reveals that they differ only by a scaling factor of Λ^{11}. The same was true for the update rules for $\tilde{f}^1_{(m)}$, Eqs. (8.7) and (8.18).

Chapter 9

Appendix

9.1 The Jacobian of h_m

In this section, we derive a formula for the Jacobian Dh_m in terms of the Jacobians Dh_i and the bilinear forms D^2h_i, $i = 1, \ldots, m-1$, and of quantities pertaining to the vector field g.

Lemma 9.1.1 *The quantity Dh_m is given by the formula*

$$\begin{aligned} Dh_m &= -((D_{x^2}g^2)_0)^{-1}\left[\left(D_{x^1}g^2\right)_m + \sum_{i=0}^{m-1}\left(D_{x^2}g^2\right)_{m-i}(Dh_i) - \sum_{\ell=0}^{m-1}(D^2h_\ell)g^1_{m-\ell-1}\right. \\ &\quad \left. - \sum_{\ell=0}^{m-1}(Dh_\ell)(D_{x^1}g^1)_{m-1-\ell} - \sum_{i=0}^{m-1}\sum_{\ell=0}^{m-1-i}(Dh_\ell)(D_{x^2}g^1)_{m-1-i-\ell}(Dh_i)\right]. \quad (9.1) \end{aligned}$$

Proof. The coefficient h_m is found from the $\mathcal{O}(\varepsilon^m)$ terms in the invariance equation (2.7),

$$g^2_m = \sum_{\ell=0}^{m-1}(Dh_\ell)g^1_{m-\ell-1}. \quad (9.2)$$

Taking the total derivative with respect to x^1 of both sides of Eq. (9.2), we find

$$\frac{d}{dx^1}g^2_m = \sum_{\ell=0}^{m-1}(D^2h_\ell)g^1_{m-\ell-1} + \sum_{\ell=0}^{m-1}(Dh_\ell)\frac{d}{dx^1}g^1_{m-\ell-1}. \quad (9.3)$$

The operations of taking the total derivative with respect to x^1 and expanding with respect to ε commute, because the Fenichel theory guarantees C^r smoothness in ε

and x^1 for each r. Therefore,

$$\frac{d}{dx^1}g^1_m = \left(\frac{dg^2}{dx^1}\right)_m = \left(D_{x^1}g^2\right)_m + \sum_{i=0}^{m}\left(D_{x^2}g^2\right)_{m-i}(Dh_i), \tag{9.4}$$

$$\frac{d}{dx^1}g^1_{m-1-\ell} = \left(\frac{dg^1}{dx^1}\right)_{m-1-\ell} = \left(D_{x^1}g^1\right)_{m-1-\ell} + \sum_{i=0}^{m-1-\ell}\left(D_{x^2}g^1\right)_{m-1-\ell-i}(Dh_i). \tag{9.5}$$

Substituting Eqs. (9.4) and (9.5) into Eq. (9.3), we obtain

$$\begin{aligned}\left(D_{x^1}g^2\right)_m + \sum_{i=0}^{m}\left(D_{x^1}g^2\right)_{m-i}Dh_i &= \sum_{\ell=0}^{m-1}(D^2h_\ell)g^1_{m-\ell-1} + \sum_{\ell=0}^{m-1}(Dh_\ell)(D_{x^1}g^1)_{m-1-\ell} \\ &\quad + \sum_{\ell=0}^{m-1}\sum_{i=0}^{m-1-\ell}(Dh_\ell)(D_{x^2}g^1)_{m-1-\ell-i}(Dh_i). \end{aligned}\tag{9.6}$$

Separating the $i = m$ term in the sum of the left member, changing the order of summation in the last sum of the right member, and solving for Dh_m, we obtain Eq. (9.1). ∎

9.2 Near-invariance of the CSPMs

In this section, we derive an ODE that is satisfied by matrices defined on the CSPM of order m, $\mathcal{K}^{(m)}_\varepsilon$. When the matrix is taken to be the function $\psi_{(m)}$, whose graph is $\mathcal{K}^{(m)}_\varepsilon$, the result is a near-invariance equation for that CSPM.

Lemma 9.2.1 *Let V be a matrix-valued function of x and ε that, together with its first-order derivatives, is smooth and $\mathcal{O}(1)$ as $\varepsilon \downarrow 0$. If $x^2 = \psi_{(m)}(x^1, \varepsilon)$ and*

$$\begin{aligned} V(\cdot\,, \psi_{(m)}(\cdot, \varepsilon), \varepsilon) &= \sum_{\ell=0}^{m}\varepsilon^\ell V_\ell(\cdot) + \mathcal{O}(\varepsilon^{m+1}), \\ g^1(\cdot\,, \psi_{(m)}(\cdot, \varepsilon), \varepsilon) &= \sum_{\ell=0}^{m}\varepsilon^\ell g^1_\ell(\cdot) + \mathcal{O}(\varepsilon^{m+1}), \end{aligned}$$

then,

$$\frac{dV}{dt}(\cdot\,, \psi_{(m)}(\cdot, \varepsilon), \varepsilon) = \sum_{i=0}^{m}\varepsilon^{i+1}\sum_{\ell=0}^{i}\frac{dV_\ell}{dx^1}g^1_{i-\ell} + \mathcal{O}(\varepsilon^{m+1}). \tag{9.7}$$

Proof. A direct computation gives

$$\frac{dV}{dt} = (DV)g = \varepsilon(D_{x^1}V)g^1 + (D_{x^2}V)g^2, \tag{9.8}$$

where all the terms are evaluated at $(\cdot, \psi_{(m)}(\cdot,\varepsilon),\varepsilon)$. Since $\psi_{(m)}$ approximates the slow manifold up to and including $\mathcal{O}(\varepsilon^m)$ terms,

$$g^1(\cdot\,, \psi_{(m)}(\cdot,\varepsilon),\varepsilon) = g^1(\cdot\,, h_\varepsilon(\cdot),\varepsilon) + \mathcal{O}(\varepsilon^{m+1}), \tag{9.9}$$

$$g^2(\cdot\,, \psi_{(m)}(\cdot,\varepsilon),\varepsilon) = g^2(\cdot\,, h_\varepsilon(\cdot),\varepsilon) + \mathcal{O}(\varepsilon^{m+1}), \tag{9.10}$$

and also

$$D\psi_{(m)} = Dh_\varepsilon + \mathcal{O}(\varepsilon^{m+1}). \tag{9.11}$$

Using Eqs. (2.7), (9.9), and (9.11), we rewrite Eq. (9.10) as

$$g^2(\cdot\,, \psi_{(m)}(\cdot,\varepsilon),\varepsilon) = \varepsilon(D\psi_{(m)})g^1(\cdot\,, \psi_{(m)}(\cdot,\varepsilon),\varepsilon) + \mathcal{O}(\varepsilon^{m+1}). \tag{9.12}$$

Equation (9.12) is an equation for $\mathcal{K}_\varepsilon^{(m)}$. We recast it so the right member involves a total derivative with respect to x^1,

$$(DV)g = \varepsilon\left(D_{x^1}V + D_{x^2}VD\psi_{(m)}\right)g^1 + \mathcal{O}(\varepsilon^{m+1}) = \varepsilon\frac{dV}{dx^1}g^1 + \mathcal{O}(\varepsilon^{m+1}) \tag{9.13}$$

or, expanding in powers of ε,

$$(DV)g \;\; = \;\; \sum_{i=0}^{m}\varepsilon^{i+1}\sum_{\ell=0}^{i}\left(\frac{dV}{dx^1}\right)_\ell g^1_{i-\ell} + \mathcal{O}(\varepsilon^{m+1}). \tag{9.14}$$

The operations of taking the total derivative with respect to x^1 and expanding with respect to ε commute, so $(dV/dx^1)_\ell = dV_\ell/dx^1$ and Eq. (9.7) follows. ∎

9.3 Action of the Jacobian on the slow subspace

The spaces $\mathcal{T}_p\mathcal{F}_\varepsilon$ and $\mathcal{T}_p\mathcal{M}_\varepsilon$ depend, in general, on both the point $p \in \mathcal{M}_\varepsilon$ and ε. As a result, the basis A also depends on p and ε, and hence A_f and A_s possess formal asymptotic expansions in terms of ε,

$$A_f = \sum_{i=0}\varepsilon^i A_f^i, \quad A_s = \sum_{i=0}\varepsilon^i A_s^i. \tag{9.15}$$

Next, we compute the action of the Jacobian on A_s to leading order.

Lemma 9.3.1 *For any $p \in \mathcal{M}_0$, $\mathrm{Ker}(Dg(p))_0 = \mathcal{T}_p\mathcal{M}_0$. In particular, $(Dg)_0A_s^0 = 0$.*

Proof. The Jacobian is a linear operator, so it suffices to show that every column vector of a basis for $\mathcal{T}_p\mathcal{M}_0$ vanishes under the left action of the Jacobian. We choose this basis to be the matrix $\begin{pmatrix} I_{N_s} \\ Dh_0 \end{pmatrix}$.

We compute

$$Dg_0 \begin{pmatrix} I_{N_s} \\ Dh_0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ D_{x^1}g^2 & D_{x^2}g^2 \end{pmatrix} \begin{pmatrix} I_{N_s} \\ Dh_0 \end{pmatrix} = \begin{pmatrix} 0 \\ D_{x^1}g^2 + (D_{x^2}g^2)(Dh_0) \end{pmatrix}. \tag{9.16}$$

Differentiating both members of the $\mathcal{O}(1)$ invariance equation $g^2(x^1, h_0(x^1), 0) = 0$ with respect to x^1, we obtain

$$D_{x^1}g^2(x^1, h_0(x^1), 0) + \left(D_{x^2}g^2(x^1, h_0(x^1), 0)\right)\left(Dh_0(x^1)\right) = 0. \tag{9.17}$$

Equations (9.16) and (9.17) yield the desired result,

$$Dg_0 \begin{pmatrix} I_{N_s} \\ Dh_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ on } \mathcal{M}_0. \tag{9.18}$$

Finally, the identity $(Dg)_0 A_s^0 = 0$ follows from the fact that A_s^0 spans $\mathcal{T}_p\mathcal{M}_0$, since $A_s^0 = A_s|_{\varepsilon=0}$ by Eq. (9.15). ∎

9.4 $\mathcal{M}_\varepsilon$ as a graph over the variables z^1

In Section 7.2, we assumed that the manifold $\mathcal{M}_\varepsilon$ is the graph, over some compact set, of some function of z^1. Here, we prove that, under mild assumptions, this is indeed the case.

Lemma 9.4.1 *Assume that*

$$\det\left(D_{x^2}z^2\right) \neq 0 \quad \textit{and} \quad \det\left(D_{z^2}q^2\right) \neq 0$$

everywhere on $\mathcal{M}_0$. Then, $\mathcal{M}_\varepsilon$ can be expressed locally as a graph over the variables z^1,

$$\mathcal{M}_\varepsilon = \left\{ z \in \mathbf{R}^N \;\middle|\; z^2 = \phi_\varepsilon(z^1),\ z^1 \in \tilde{K} \right\}.$$

Proof. It suffices to prove this result to leading order, that is, to prove that $\mathcal{M}_0$ is a graph over z^1,

$$\mathcal{M}_0 = \left\{ z \in \mathbf{R}^N \;\middle|\; z^2 = \phi_0(z^1),\ z^1 \in \tilde{K} \right\}, \tag{9.19}$$

for some ϕ_0 with $D\phi_0$ uniformly bounded over $\tilde{K}$.

First, we show that the zero level sets of both q^2 and g^2 identify $\mathcal{M}_0$, to leading order. We calculate

$$q^2 = (z^2)' = \left(D_{x^1}z^2\right)\left(\varepsilon g^1\right) + \left(D_{x^2}z^2\right)g^2, \tag{9.20}$$

which yields, to leading order,

$$q_0^2 = \left(D_{x^2} z^2\right) g_0^2, \tag{9.21}$$

with a remainder of $\mathcal{O}(\varepsilon)$. Since $D_{x^2} z^2$ is invertible on $\mathcal{M}_0$ by assumption, $q_0^2 = 0$ if and only if $g_0^2 = 0$. Recalling Eqs. (2.4)–(2.5), then, we obtain

$$\mathcal{M}_0 = \left\{ z \in \mathbf{R}^N \;\middle|\; q_0^2(z) = 0 \right\} = \left\{ x \in \mathbf{R}^N \;\middle|\; g_0^2(x) = 0 \right\}. \tag{9.22}$$

By assumption, now, $\det(D_{z^2} q^2) \neq 0$, and thus the Implicit Function Theorem implies that the equation $q_0^2(z) = 0$ can be solved to yield a function $z^2 = \phi_0(z^1)$. The boundedness of $D\phi_0$ follows from the compactness of $\mathcal{M}_0$, and the proof is complete. ∎

Remark. It is interesting to note that the two assumptions of Lemma 9.4.1 are independent of each other, despite the fact that Eqs. (7.35) and (7.36) seem to imply otherwise. Indeed, these two equations yield, for $m = 0$ and since $L_{(0)} = q^2$ (see definition (7.3)),

$$D_{z^2} q^2 = \left(D_{x^2} z^2\right)_0 \left(D_{x^2} g^2\right)_0 \left(D_{x^2} z^2\right)_0^{-1} T_0, \tag{9.23}$$

and thus it appears that the assumption on the invertibility of $D_{x^2} z^2$, together with normal hyperbolicity, is enough to guerantee the invertibility of $D_{z^2} q^2$. Nevertheless, this is not the case; Eq. (9.23) is a special case of Eq. (7.35), and the latter is obtained, in Section 7.5.2, under the explicit assumption that $\mathcal{M}_0$ is the graph of ϕ_0 (see, for example, Eq. (9.28)). This is illustrated by considering the planar system

$$\begin{pmatrix} z^1 \\ z^2 \end{pmatrix}' = \begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} z^1 \\ z^2 \end{pmatrix},$$

which can be put in fast–slow format through the transformation

$$\begin{pmatrix} z^1 \\ z^2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} \tag{9.24}$$

The resulting fast–slow system has the form

$$\begin{pmatrix} x^1 \\ x^2 \end{pmatrix}' = \begin{pmatrix} 0 & 0 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix},$$

and thus, it possesses a line of fixed points which corresponds to the graph of the function $x^2 = -2x^1$. (Under the transformation (9.24), this line maps to the z^2-axis.) Thus, despite the fact that $D_{x^2} z^2 = 1 \neq 0$, we find $q^2 = -z^1$ and thus $D_{z^2} q^2 = 0$ everywhere. Note, also, that the equation $q^2 = 0$ identifies $\mathcal{M}_0$ correctly, but $\mathcal{M}_0$ is not a graph over z^1.

9.5 A leading-order asymptotic formula for $L_{(m)}$

Lemma 9.5.1 *For $m = 0, 1, \ldots$, the function $L_{(m)}$ can be written as*

$$L_{(m)}(z(x)) = \left(\left(D_{x^2} z^2\right)(x)\right)\left(\left(D_{x^2} g^2\right)_0(x)\right)^m g_0^2(x) + \mathcal{O}\left(\varepsilon, \|g^2\|^2\right),$$

where the notation "$(\cdot)_0(x)$" signifies that the parenthesized quantity is evaluated at $(x, 0)$.

Proof. The proof is by induction on m. For $m = 0$, we need to show that

$$L_{(0)} = \left(D_{x^2} z^2\right) g_0^2 + \mathcal{O}(\varepsilon, \|g^2\|^2).$$

We have already shown that

$$L_{(0)} \;=\; \left(D_{x^1} z^2\right)(\varepsilon g^1) + \left(D_{x^2} z^2\right) g^2,$$

see Eq. (7.5). The terms $D_{x^1} v$ and $D_{x^2} z^2$ are independent of ε by Eq. (2.18), while εg^1 is $\mathcal{O}(\varepsilon)$. Expanding g^2 in powers of ε, then, we obtain the desired formula.

We now carry out the inductive step. We assume that

$$L_{(m)} = \left(D_{x^2} z^2\right)\left(D_{x^2} g^2\right)_0^m g_0 + \mathcal{O}\left(\varepsilon, \|g^2\|^2\right), \tag{9.25}$$

and we show that

$$L_{(m+1)} = \left(D_{x^2} z^2\right)\left(D_{x^2} g^2\right)_0^{m+1} g_0 + \mathcal{O}\left(\varepsilon, \|g^2\|^2\right). \tag{9.26}$$

By definition of $L_{(m)}$ (see Eq. (7.3)),

$$L_{(m+1)} = \frac{dL_{(m)}}{dt} = (D_x L_{(m)}) g,$$

and thus

$$L_{(m+1)} = (D_{x^1} L_{(m)})(\varepsilon g^1) + (D_{x^2} L_{(m)}) g^2. \tag{9.27}$$

Next, we substitute Eq. (9.25) into Eq. (9.27). First, we see that an application of the differential operators $(D_{x^1}\cdot)(\varepsilon g^1)$ and $(D_{x^2}\cdot)g^2$ on that same remainder does not alter its asymptotic magnitude. Second, we note that the term $(D_{x^1} L_{(m)})(\varepsilon g^1)$ is $\mathcal{O}(\varepsilon)$ and thus can be absorbed in the $\mathcal{O}(\varepsilon, \|g\|^2)$ remainder. Thus, we find

$$L_{(m+1)} = \left(D_{x^2}\left[\left(D_{x^2} z^2\right)\left(D_{x^2} g^2\right)^m g^2\right]\right) g^2 + \mathcal{O}\left(\varepsilon, \|g\|^2\right).$$

Finally, we use the product rule on the first term in the right member, absorb in the remainder all of the resulting terms that are quadratic in g^2 — namely

$$\left(D_{x^2}^2 z^2\right)\left(\left(D_{x^2} g^2\right)^m g^2, g^2\right) \quad \text{and} \quad m\left(D_{x^2} z^2\right)\left(D_{x^2}^2 g^2\right)\left(\left(D_{x^2} g^2\right)^{m-1} g^2, g^2\right)$$

— and expand all other terms in powers of ε to find the desired formula (9.26). This completes the proof of the lemma. ∎

9.6 Leading order fast–slow decompositions

Lemma 9.6.1 *Let $w(z)$ be the representation in the $z-$coordinate system of a vector field defined on $\mathcal{M}_0$, and define the matrices*

$$B_0 = \begin{pmatrix} B_0^{s\perp} \\ B_0^{f\perp} \end{pmatrix} \equiv \begin{pmatrix} -D\phi_0 & I_{N_f} \\ (D_{z^1}x^1)_0 & (D_{z^2}x^1)_0 \end{pmatrix} \quad \textit{and} \quad A^0 = \left(A_f^0, A_s^0\right) = (B_0)^{-1}, \tag{9.28}$$

where the notation "$(\cdot)_0$" signifies that all parenthesized quantities are evaluated at $z(p)$, with $p \in \mathcal{M}_0$. Then, the equation

$$w = A^0 \begin{pmatrix} w_0^f \\ w_0^s \end{pmatrix}, \quad \textit{where} \quad \begin{pmatrix} w_0^f \\ w_0^s \end{pmatrix} = B_0 w,$$

is a leading order fast–slow decomposition of w, that is, $w_0^f \in \mathcal{TF}_0$ and $w_0^s \in \mathcal{TM}_0$.

Proof. It suffices to show that

$$\text{span}\left(A_f^0(z(p))\right) = \mathcal{T}_p\mathcal{F}_0 \quad \text{and} \quad \text{span}\left(A_s^0(z(p))\right) = \mathcal{T}_p\mathcal{M}_0. \tag{9.29}$$

We begin by observing that

$$\begin{array}{rcccl} B_0^{s\perp} & = & \left(-D\phi_0, I_{N_f}\right) & = & D_z(-\phi_0(z^1) + z^2) \\ B_0^{f\perp} & = & \left((D_{z^1}x^1)_0, (D_{z^2}x^1)_0\right) & = & D_z x^1. \end{array}$$

Since

$$\mathcal{M}_0 = \left\{ z \in \mathbf{R}^N \;\middle|\; \phi_0(z^1) + z^2 = 0 \right\} \quad \text{and} \quad \mathcal{F}_0^p = \left\{ z \in \mathbf{R}^N \;\middle|\; x^1(z) = x^1(p) \right\}, \tag{9.30}$$

and since, also, gradients are normal to level sets, we find that

$$\text{span}\left(B_0^{s\perp}(z(p))\right) = \mathcal{N}_p\mathcal{M}_0 \quad \text{and} \quad \text{span}\left(B_0^{f\perp}(z(p))\right) = \mathcal{N}_p\mathcal{F}_0.$$

Eq. (9.29) now follows. ∎

Lemma 9.6.2 *An explicit expression for the matrix A^0, defined in Lemma 9.6.1, is given by the formula*

$$A^0 = \begin{pmatrix} (D_{x^2}z^1)_0 \left(D_{x^2}z^2\right)_0^{-1} T_0 & (D_{x^1}z^1)_0 + (D_{x^2}z^1)_0 \left(Dh_0\right) \\ T_0 & (D_{x^1}z^2)_0 + (D_{x^2}z^2)_0 \left(Dh_0\right) \end{pmatrix}, \tag{9.31}$$

where

$$T_0 = \left[\left(-D\phi_0, I_{N_f}\right) \begin{pmatrix} (D_{x^2}z^1)_0 \left(D_{x^2}z^2\right)_0^{-1} \\ I_{N_f} \end{pmatrix} \right]^{-1}.$$

All quantities are evaluated at $z(p)$, with $p \in \mathcal{M}_0$.

Proof. To calculate $A^0 \equiv (B_0)^{-1}$, we work as follows. First, we define the matrix

$$\hat{A}^0 = \left(\hat{A}^0_f, \hat{A}^0_s\right) \equiv \begin{pmatrix} 0 & I_{N_s} \\ I_{N_f} & Dh_0 \end{pmatrix} \tag{9.32}$$

with inverse

$$\hat{B}_0 = \begin{pmatrix} \hat{B}^{s\perp}_0 \\ \hat{B}^{f\perp}_0 \end{pmatrix} \equiv \begin{pmatrix} -Dh_0 & I_{N_f} \\ I_{N_s} & 0 \end{pmatrix}.$$

For each $p \in \mathcal{M}_0$, we consider the columns of the matrix $A^0(x(p))$ as the representation in the $x-$coordinate system of a vector basis for $\mathcal{T}_p\mathbf{R}^N$.[1] Under this convention, and observing that

$$\begin{array}{rcccl} \hat{B}^{s\perp}_0 & = & \left(-Dh_0, I_{N_f}\right) & = & D_x(-h_0(x^1) + x^2), \\ \hat{B}^{f\perp}_0 & = & (I_{N_s}, 0) & = & D_x x^1 \end{array}$$

and also that

$$\mathcal{M}_0 = \left\{ x \in \mathbf{R}^N \;\middle|\; -h_0(x^1) + x^2 = 0 \right\} \quad \text{and} \quad \mathcal{F}^p_0 = \left\{ x \in \mathbf{R}^N \;\middle|\; x^1 = x^1(p) \right\}, \tag{9.33}$$

we conclude that

$$\text{span}\left(\hat{B}^{s\perp}_0(x(p))\right) = \mathcal{N}_p\mathcal{M}_0 \quad \text{and} \quad \text{span}\left(\hat{B}^{f\perp}_0(x(p))\right) = \mathcal{N}_p\mathcal{F}_0.$$

Therefore, it is also the case that

$$\text{span}\left(\hat{A}^0_f(x(p))\right) = \mathcal{T}_p\mathcal{F}_0 \quad \text{and} \quad \text{span}\left(\hat{A}^0_s(x(p))\right) = \mathcal{T}_p\mathcal{M}_0. \tag{9.34}$$

Next, we observe that, by virtue of see Eq. (2.22), the representation in the $z-$coordinate system of the vector basis $\hat{A}^0$ is given by

$$\tilde{A}^0 = \left(\tilde{A}^0_f, \tilde{A}^0_s\right) \equiv (D_x z)_0 \, \hat{A}^0$$

with

$$\begin{array}{rcl} \text{span}\left(\tilde{A}^0_f(x(p))\right) & = & \text{span}\left(\hat{A}^0_f(z(p))\right) \\ \text{span}\left(\tilde{A}^0_s(x(p))\right) & = & \text{span}\left(\hat{A}^0_s(z(p))\right). \end{array} \tag{9.35}$$

Combining Eqs. (9.29, (9.34), and (9.35), we conclude that

$$\text{span}\left(\tilde{A}^0_f(x(p))\right) = \text{span}\left(A^0_f(x(p))\right) \quad \text{and} \quad \text{span}\left(\tilde{A}^0_f(x(p))\right) = \text{span}\left(A^0_f(x(p))\right),$$

[1] Readers that are not interested in the geometry of the problem can skip the rest of this paragraph all the way to Eq. (9.36) and regard the latter merely as the definition of the matrix C.

for all $p \in \mathcal{M}_0$. As a result, there is an $N \times N$, block-diagonal, non-singular matrix C such that

$$A^0 = \tilde{A}^0 C = (D_x z)_0 \, \hat{A}^0 C. \tag{9.36}$$

If, then, we can find a formula for C, we can find a formula for A^0 through Eq. (9.36).

To obtain the desired formula for C, we invert both members of Eq. (9.36) to obtain

$$B_0 = C^{-1} \hat{B}_0 \, (D_x z)_0^{-1} \, ,$$

whence

$$C^{-1} = B_0 \, (D_x z)_0 \, \hat{A}^0. \tag{9.37}$$

Substituting for B_0 and $\hat{A}^0$ from Eqs. (9.28) and (9.32), respectively, we calculate

$$C^{-1} = \begin{pmatrix} -D\phi_0 & I_{N_f} \\ (D_{z^1} x^1)_0 & (D_{z^2} x^1)_0 \end{pmatrix} \begin{pmatrix} (D_{x^1} z^1)_0 & (D_{x^2} z^1)_0 \\ (D_{x^1} z^2)_0 & (D_{x^2} z^2)_0 \end{pmatrix} \begin{pmatrix} 0 & I_{N_s} \\ I_{N_f} & Dh_0 \end{pmatrix}. \tag{9.38}$$

Now,

$$\begin{aligned} &\begin{pmatrix} -D\phi_0 & I_{N_f} \\ (D_{z^1} x^1)_0 & (D_{z^2} x^1)_0 \end{pmatrix} \begin{pmatrix} (D_{x^1} z^1)_0 & (D_{x^2} z^1)_0 \\ (D_{x^1} z^2)_0 & (D_{x^2} z^2)_0 \end{pmatrix} \\ &= \begin{pmatrix} -(D\phi_0)(D_{x^1} z^1)_0 + (D_{x^1} z^2)_0 & -(D\phi_0)(D_{x^2} z^1)_0 + (D_{x^2} z^2)_0 \\ I_{N_s} & 0 \end{pmatrix}. \end{aligned}$$

Substituting this result in Eq. (9.38), we find

$$\begin{aligned} \left(C^{-1}\right)_1^1 &= -(D\phi_0)\left(D_{x^2} z^1\right)_0 + \left(D_{x^2} z^2\right)_0 , \\ \left(C^{-1}\right)_2^1 &= -(D\phi_0)\left(D_{x^1} z^1\right)_0 + \left(D_{x^1} z^2\right)_0 + \left[-(D\phi_0)\left(D_{x^2} z^1\right)_0 + \left(D_{x^2} z^2\right)_0\right](Dh_0), \\ \left(C^{-1}\right)_1^2 &= 0, \\ \left(C^{-1}\right)_2^2 &= I_{N_s}. \end{aligned}$$

To complete the derivation of the formula for C^{-1}, we show that $(C^{-1})_2^1 = 0$. Combining Eqs. (9.30) and (9.33) for $\mathcal{M}_0$, we obtain

$$-\phi_0(z^1(x^1, h_0(x^1))) + z^2(x^1, h_0(x^1))) = 0.$$

Taking the total derivative with respect to x^1 of both members, we obtain

$$-(D\phi_0)\left[\left(D_{x^1} z^1\right)_0 + \left(D_{x^2} z^1\right)_0 (Dh_0)\right] + \left(D_{x^1} z^2\right)_0 + \left(D_{x^2} z^2\right)_0 (Dh_0) = 0,$$

and therefore $(C^{-1})_2^1 = 0$ as desired. Thus,

$$C^{-1} = \begin{pmatrix} -(D\phi_0)(D_{x^2} z^1)_0 + (D_{x^2} z^2)_0 & 0 \\ 0 & I_{N_s} \end{pmatrix} = \begin{pmatrix} (T_0)^{-1} (D_{x^2} z^2)_0 & 0 \\ 0 & I_{N_s} \end{pmatrix},$$

and so

$$C = \begin{pmatrix} (D_{x^2} z^2)_0^{-1} T_0 & 0 \\ 0 & I_{N_s} \end{pmatrix}. \tag{9.39}$$

Next, we return to Eq. (9.36) and substitute for $\hat{A}_0$ and C from Eqs. (9.32) and (9.39), respectively, to find

$$\begin{aligned} A^0 &= \begin{pmatrix} (D_{x^1} z^1)_0 & (D_{x^2} z^1)_0 \\ (D_{x^1} z^2)_0 & (D_{x^2} z^2)_0 \end{pmatrix} \begin{pmatrix} 0 & I_{N_s} \\ I_{N_f} & Dh_0 \end{pmatrix} \begin{pmatrix} (D_{x^2} z^2)_0^{-1} T_0 & 0 \\ 0 & I_{N_s} \end{pmatrix} \\ &= \begin{pmatrix} (D_{x^1} z^1)_0 & (D_{x^2} z^1)_0 \\ (D_{x^1} z^2)_0 & (D_{x^2} z^2)_0 \end{pmatrix} \begin{pmatrix} 0 & I_{N_s} \\ (D_{x^2} z^2)_0^{-1} T_0 & Dh_0 \end{pmatrix} \\ &= \begin{pmatrix} (D_{x^2} z^1)_0 (D_{x^2} z^2)_0^{-1} T_0 & (D_{x^1} z^1)_0 + (D_{x^2} z^1)_0 (Dh_0) \\ T_0 & (D_{x^1} z^2)_0 + (D_{x^2} z^2)_0 (Dh_0) \end{pmatrix} \end{aligned}$$

This completes the proof. ∎

Bibliography

[1] J. Carr, *Applications of Centre Manifold Theory*, Applied Mathematical Sciences, **35**, Springer-Verlag, New York, 1981

[2] M. J. Davis and R. T. Skodje, Geometric investigation of low-dimensional manifolds in systems approaching equilibrium, *Journal of Chemical Physics*, **111** (1999) 859–874

[3] B. A. Dubrovin, A. T. Fomenko, and S. P. Novikov, *Modern Geometry – Methods and Applications, Vol. 2*, Graduate Texts in Mathematics, **104**, Springer-Verlag, New York, 1985

[4] A. Fehrst, *Enzyme Structure and Mechanisms*, 2nd ed., W. F. Freeman, New York, 1975

[5] N. Fenichel, Geometric singular perturbation theory for ordinary differential equations, *Journal of Differential Equations*, **31** (1979) 53–98

[6] S. J. Fraser, The steady state and equilibrium approximations: A geometrical picture, *Journal of Chemical Physics*, **88** (1988) 4732–4738

[7] C. W. Gear and I. G. Kevrekidis, Constraint-defined Manifolds: A Legacy-code Approach to Low-dimensional Computation, *SIAM Journal on Scientific Computation*, to appear

[8] C. W. Gear, T. J. Kaper, I. G. Kevrekidis, A. Zagaris, Projecting to a Slow Manifold: Singularly Perturbed Systems and Legacy Codes, *SIAM Journal on Scientific Computation*, to appear

[9] C. W. Gear, T. J. Kaper, I. G. Kevrekidis, A. Zagaris, Projecting to a Slow Manifold: Singularly Perturbed Systems and Legacy Codes, Part 2 (working title), in preparation

[10] A. N. Gorban and I. V. Karlin, Thermodynamic Parameterization, *Physica A*, **190** (1992) 393–404

[11] A. N. Gorban and I. V. Karlin, Method of Invariant Manifolds and regularization of acoustic spectra, *Transport Theory and Statistical Physics*, **23** (1994) 559–632

[12] A. N. Gorban and I. V. Karlin, Method of invariant manifolds for chemical kinetics, *arXiv:cond-mat/0207231* (9 Jul 2002)

[13] A. N. Gorban and I. V. Karlin, Method of Invariant Manifolds for chemical kinetics, *Chemical Engineering Science*, **58** (2003) 4751–4768

[14] A. N. Gorban, I. V. Karlin, and A. Yu. Zinovyev, Constructive methods for invariant manifolds for kinetic problems, *IHES preprint*, July 2003

[15] A. N. Gorban, I. V. Karlin, and A. Yu. Zinovyev, Invariant grids for reaction kinetics, *Physica A*, **333** (2004) 106–154

[16] D. A. Goussis and S. H. Lam, A study of homogeneous methanol oxidation kinetics using CSP, in: *Twenty-Fourth Symposium (International) on Combustion, The University of Sydney, Sydney, Australia, July 5–10, 1992*, The Combustion Institute, Pittsburgh, 1992, pp. 113–120

[17] M. Hadjinicolaou and D. A. Goussis, Asymptotic solutions of stiff PDEs with the CSP method: The reaction diffusion equation, *SIAM Journal on Scientific Computation*, **20** (1999) 781–810

[18] F. Heineken, H. Tsuchiya, and R. Aris, On the mathematical status of the pseudo-steady-state hypothesis of biochemical kinetics, *Mathematical Biosciences*, **1** (1967) 95–113

[19] C. K. R. T. Jones, Geometric singular perturbation theory, in: *Dynamical Systems, Montecatini Terme,* L. Arnold, Lecture Notes in Mathematics, **1609**, Springer-Verlag, Berlin, 1994, pp. 44–118

[20] H. G. Kaper and T. J. Kaper, Asymptotic analysis of two reduction methods for systems of chemical kinetics, *Physica D*, **165** (2002) 66–93

[21] I. G. Kevrekidis, C. W. Gear, and G. Hummer, Equation-free: The Computer-aided Analysis of Complex Multiscale Systems, *AIChE Journal*, **50** (2004) 1438–1452

[22] I. G. Kevrekidis, C. W. Gear, J. M. Hyman, P. G. Kevrekidis, O. Runborg, and K. Theodoropoulos, Equation-free Coarse-grained Multiscale Computation: Enabling Microscopic Simulators to Perform System-level Tasks, *Communications in Mathematical Sciences*, **1** (2003) 715–762

[23] H-O. Kreiss, Problems with Different Time Scales, in *Multiple Time Scales*, J. H. Brackbill and B. I. Cohen, eds., Academic Press, 1985, pp. 29-57

[24] S. H. Lam, Using CSP to understand complex chemical kinetics, *Combustion Science and Technology*, **89** (1993) 375–404

[25] S. H. Lam and D. A. Goussis, Understanding complex chemical kinetics with computational singular perturbation, in *Twenty-Second Symposium (International) on Combustion, The University of Washington, Seattle, Washington, August 14–19, 1988,* The Combustion Institute, Pittsburgh, 1988, pp. 931–941

[26] S. H. Lam and D. A. Goussis, Conventional asymptotics and computational singular perturbation theory for simplified kinetics modeling, in *Reduced Kinetic Mechanisms and Asymptotic Approximations for Methane-Air Flames*, M. Smooke, editor, Lecture Notes in Physics **384**, Springer-Verlag, New York, 1991, Chapter 10

[27] S. H. Lam and D. A. Goussis, The CSP method for simplifying kinetics, *International Journal of Chemical Kinetics*, **26** (1994) 461–486

[28] T. F. Lu, Y. G. Ju, and C. K. Law, Complex CSP for chemistry reduction and analysis, *Combustion and Flame*, **126** (2001) 1445–1455

[29] U. Maas and S. B. Pope, Implementation of simplified chemical kinetics based on intrinsic low-dimensional manifolds, in: *Twenty-Fourth Symposium (International) on Combustion, The University of Sydney, Sydney, Australia, July 5–10, 1992*, The Combustion Institute, Pittsburgh, 1992, pp. 103–112

[30] U. Maas and S. B. Pope, Simplifying chemical kinetics: Intrinsic low-dimensional manifolds in composition space, *Combustion and Flame*, **88** (1992) 239–264

[31] A. Massias, D. Diamantis, E. Mastorakos, and D. Goussis, Global reduced mechanisms for methane and hydrogen combustion with nitric oxide formation constructed with CSP data, *Combustion Theory and Modelling*, **3** (1999) 233–257

[32] A. Massias and D. A. Goussis, On the manifold of stiff reaction-diffusion PDE's: The effects of diffusion, preprint (2001)

[33] K. D. Mease, Geometry of computational singular perturbations, in *Nonlinear Control System Design,* vol. 2, A. J. Kerner and D. Q. M. Mayne, editors, Pergamon Press, Oxford, U.K., 1996, pp. 855–861

[34] M. K. Neophytou, D. A. Goussis, M. van Loon, and E. Mastorakos, Reduced Chemical Mechanisms for Atmospheric Pollution Using Computational Singular Perturbation Analysis, *Atmospheric Environment* **38**, 3661–3673 (2004)

[35] P. J. Olver, *Applications of Lie Groups to Differential Equations*, Graduate Texts in Mathematics, **107**, Springer-Verlag, New York, 1986

[36] R. E. O'Malley, Jr., *Singular Perturbation Methods for Ordinary Differential Equations*, Springer-Verlag, New York, 1991

[37] B. O. Palsson, On the dynamics of the irreversible Michaelis–Menten reaction mechanism, *Chemical Engineering Science*, **42** (1987) 447–458

[38] B. O. Palsson and E. N. Lightfoot, Mathematical modelling of dynamics and control in metabolic networks. I. On Michaelis–Menten kinetics, *Journal of Theoretical Biology*, **111** (1984) 273–302

[39] A. J. Roberts, Computer algebra derives correct initial conditions for low-dimensional dynamical systems, arXiv: chao-dyn/9901010

[40] M. R. Roussel and S. J. Fraser, Geometry of the steady-state approximation: Perturbation and accelerated convergence methods, *Journal of Chemical Physics*, **93** (1990) 1072–1081

[41] B. Schutz, *Geometrical Methods of Mathematical Physics* (Cambridge University Press, 1980)

[42] L. A. Segel and M. Slemrod, The quasi-steady-state assumption: a case study in perturbation, *SIAM Review*, **31** (1989) 446–477

[43] S. Singh, J. M. Powers, and S. Paolucci, On slow manifolds of chemically reactive systems, *Journal of Chemical Physics*, **117** (2002) 1482–1496

[44] M. Stiefenhofer, Quasi-steady-state approximation for chemical reaction networks, *Journal of Mathematical Biology*, **36** (1998) 593–609

[45] T. Turányi, A. S. Tomlin, and M. J. Pilling, On the error of the quasi-steady-state approximation, *Journal of Physical Chemistry*, **97** (1993) 163–172

[46] M. Valorani and D. A. Goussis, Explicit time-scale splitting algorithm for stiff problems: auto-ignition of gaseous mixtures behind a steady shock, *Journal of Computational Physics*, **169** (2001) 44–79

[47] M. Valorani, D. A. Goussis, and H. Najm, personal communication (2002)

[48] M. Valorani, H. M. Najm, and D. A. Goussis, CSP analysis of a transient flame-vortex interaction: time scales and manifolds, *Combustion and Flame*, **134** (2003) 35–53

[49] A. Zagaris, H. G. Kaper, and T. J. Kaper, Analysis of the Computational Singular Perturbation reduction method for chemical kinetics, *Journal of Nonlinear Science*, **14** (2004) 59–91, also available at arXiv: math.DS/0305355

[50] A. Zagaris, H. G. Kaper, and T. J. Kaper, Fast and Slow Dynamics for the Computational Singular Perturbation Method, *Multiscale Modeling and Simulation*, **2** (2004) 613–638

Curriculum Vitae

Antonios Zagaris

Boston University	Office: MCS 162
Department of Mathematics	Phone: (+1)(617) 353 3923
111 Cummington st.	Fax: (+1)(617) 353 8100
Boston, MA 02215	Email: azagaris@math.bu.edu
USA	http://math.bu.edu/people/azagaris

PERSONAL DATA

Date of birth: February 8 1976
Place of birth: Veria, Greece
Citizenship: Greek

RESEARCH INTERESTS

Singularly Perturbed Systems, Reaction–Diffusion Equations, Numerical Algorithms for ODE and PDE, Reduction Methods, Statistical Physics

EDUCATION

09/2001 – 01/2005: *Ph.D. (Mathematics)*, Boston University, MA, USA (expected)

09/1999 – 05/2001: *M.Sc. (Mathematics)*, Boston University, MA, USA (GPA: 4.00/4.00)

09/1993 - 07/1998: *B.Sc. (Physics)*, Aristotelian University of Thessaloniki, Greece (GPA: 8.88/10.00)

TEACHING EXPERIENCE

01/2002 – 05/2002: Teaching Fellow, Multivariate Calculus, Boston University, Boston, MA

09/2001 – 12/2001: Teaching Fellow, Multivariate Calculus, Boston University, Boston, MA

01/2001 – 05/2001: Teaching Fellow, Differential Equations, Boston University, Boston, MA

09/2000 – 12/2000: Teaching Fellow, Calculus I, Boston University, Boston, MA

05/2000 – 06/2000: Instructor, Discrete Mathematics, Boston University, Boston, MA

PUBLICATIONS

- A. Zagaris, H. G. Kaper, and T. J. Kaper, *Analysis of the Computational Singular Perturbation reduction method for chemical kinetics*, J. Nonlin. Sci. **14** (2004) 59–91, also available at arXiv: math.DS/0305355
- A. Zagaris, H. G. Kaper, and T. J. Kaper, *Fast and slow dynamics for the Computational Singular Perturbation method*, Multiscale Model. Sim. **4(2)** (2004) 613–638
- A. Zagaris, H. G. Kaper, and T. J. Kaper, *Two perspectives on reduction of ordinary differential equations*, Mathematische Nachrichten, submitted
- C. W. Gear, T. J. Kaper, I. G. Kevrekidis, and A. Zagaris, *Projecting to a slow manifold: Singularly perturbed systems and legacy codes*, SIADS, submitted
- C. W. Gear, T. J. Kaper, I. G. Kevrekidis, and A. Zagaris, *Projecting to a slow manifold: Singularly perturbed systems and legacy codes part II*, in preparation

PRESENTATIONS

- *Asymptotic analysis of the Computational Singular Perturbation reduction method for chemical kinetics*, Poster, Thirtieth International Conference on Combustion, Chicago, IL, USA, July 25–30 2004
- *Analysis of reduction methods for large systems of ODEs*, Invited Talk, Cluster of Modelling, Analysis and Simulation, CWI, Amsterdam, The Netherlands, June 30 2004

- *Analysis of the Computational Singular Perturbation (CSP) reduction method for chemical kinetics*, Poster, Tenth International Conference on Numerical Combustion, Sedona, AZ, USA, May 9–12 2004
- *Analysis of the CSP reduction method of Lam & Goussis for chemical kinetics*, Invited Talk, SIAM Conference on Applications of Dynamical Systems, Snowbird, UT, USA, May 27-31 2003
- *Analysis of the ILDM and CSP reduction methods for chemical kinetics*, Invited Talk, Department of Mathematics and Statistics, Boston University, Boston, MA, May 5 2003

AWARDS

06/2002–present: Research Fellowship, Boston University, Boston, MA

06/2001 – 08/2001: Research Fellowship, Boston University, Boston, MA

09/2000 – 05/2002: Teaching Fellowship, Boston University, Boston, MA

09/1999 – 05/2000: Presidential University Graduate Fellowship, Boston University, Boston, MA

09/1993 – 07/1997: Greek National Fellowships Foundation (I.K.Y.) Award for academic excellence, Thessaloniki, Greece

07/1993: Greek Physical Society Award for excellence in the national Physics contest for high school students

GRADUATE COURSEWORK

Ordinary Differential Equations, Partial Differential Equations, Applied Nonlinear Dynamics, Discrete Dynamics, Partial Differential Equations Seminar, Seminar on Geometric Singular Perturbation Theory, Soliton Theory I & II (directed study), Symplectic Integrators (directed study), Celestial Mechanics, Real Analysis, Complex Analysis, Functional Analysis, Scientific Computing, Differential Topology, Probability

LANGUAGES

English: Spoken/Written fluently
French: Spoken/Written quite well
Greek: Native

References

Prof. Tasso Kaper
Boston University
Dept. of Mathematics
and Statistics
Boston, MA, USA
(+1)(617) 353 9552
tasso@math.bu.edu

Prof. Gene Wayne
Boston University
Dept. of Mathematics
and Statistics
Boston, MA, USA
(+1)(617) 353 1495
cew@math.bu.edu

Prof. Ioannis Kevrekidis
Princeton University
Dept. of Chemical
Engineering
Princeton, NJ, USA
(+1)(609) 258 2818
yannis@princeton.edu