SYMMETRY BREAKING, AVERAGING AND ELASTIC RODS WITH HIGH INTRINSIC TWIST

THÈSE N° 2646 (2002)

PRÉSENTÉE A LA FACULTÉ SCIENCES DE BASE

SECTION DE MATHÉMATIQUES

ÉCOLE POLYTECHNIQUE FÉDERALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Stéphane Rey

ingénieur physicien diplômé EPF
de nationalité suisse et originaire de Lens (VS)

acceptée sur proposition du jury:

Prof. Anthony C. Davison, président du jury
Prof. John H. Maddocks, directeur de thèse
Prof. Marcus Grote, rapporteur
Prof. Alexander Mielke, rapporteur
Prof. Charles A. Stuart, rapporteur

Lausanne, EPFL
2002

Abstract

This dissertation develops a theory describing the equilibrium configurations of elastic rods with an anisotropic bending law and a high, but finite, intrinsic twist, under various imposed two-point boundary conditions. The model adopted allows the rod to have an intrinsic curvature, and to undergo large deformations in space by suffering bending and twist. The governing equations can be formulated as a canonical Hamiltonian two-point boundary value problem depending on two different scales, with a periodic dependence in the short scale corresponding to high intrinsic twist. An averaging technique involving a near-identity canonical transformation is developed to treat the problem, leading to the introduction of an effective rod. The equilibria of the effective rod, which are generally easier to compute, are candidate leading order approximations to the equilibria of the rod with high intrinsic twist. In general, the original problem has no continuous symmetry, but after averaging we distinguish two cases:

(a) The effective problem has no continuous symmetry and admits isolated solutions. This case typically occurs when the rod is intrinsically curved. The averaging technique together with the implicit function theorem show that, in general, all equilibria of the effective rod are then approximations to equilibria of the rod with high intrinsic twist.

(b) The effective problem has a continuous symmetry associated with isotropy and admits one-parameter families of solutions. This case typically occurs when the rod is intrinsically straight. In general, only some of the equilibria of the effective rod are approximations to equilibria of the rod with high intrinsic twist. To identify which equilibria are pertinent, we use the averaging technique together with a symmetry breaking argument based upon a reduction of the problem to a situation where the Crandall-Rabinowitz theorem on bifurcation from a simple eigenvalue applies. This leads to the definition of a Melnikov function indicating which of the solutions to the effective problem approximate solutions to the original problem. In general, due to the topology of the families of equilibria of the effective rod, at least two equilibria in each family approximate equilibria of the rod with high intrinsic twist.

In both cases (a) and (b), we present and justify a perturbation method to compute approximations (of any order in the small parameter corresponding to the inverse of the value of the intrinsic twist) to an equilibrium of the rod with high intrinsic twist. The symmetry breaking occuring in case (b) is illustrated by the explicit computation of

the Melnikov function in examples involving a rod with a standard diagonal quadratic strain energy, and subject to various boundary conditions.

We also examine the problem of a rod with an anisotropic bending law, a small intrinsic curvature and a high intrinsic twist. The corresponding effective rod is isotropic, and a symmetry breaking as in case (b) above typically occurs. Here the Melnikov function is the sum of two terms, one associated to intrinsic curvature and the other to intrinsic twist, which gives rise to a competition between the different effects in the determination of symmetry breaking. This competition is illustrated by the explicit computation of the Melnikov function in an example involving a rod with a standard diagonal quadratic strain energy.

For an intrinsically twisted strut with an anisotropic bending law, we obtain an expression determining the buckling loads, for all values of the intrinsic twist. Then, for large values of the intrinsic twist, we estimate the error associated with the approximation of the first of these buckling loads by the first buckling load of the corresponding effective strut. This error is proportional both to the inverse of the intrinsic twist and to a quantity measuring the anisotropy of the bending law.

Version abrégée

Dans ce travail, on développe une théorie permettant de décrire les configurations d'équilibre d'une tige élastique ayant une loi de flexion anisotrope et une grande torsion (géométrique) naturelle, sous l'influence de diverses conditions au bord. Le modèle utilisé permet à la tige d'être courbée dans sa configuration naturelle, et de se déformer dans l'espace en subissant de la flexion et de la torsion. Les équations d'équilibre de la tige sont écrites sous la forme d'un problème Hamiltonien canonique à deux échelles, avec une dépendence périodique par rapport à la petite échelle liée à la grande torsion naturelle. Une méthode de moyennisation basée sur une transformation canonique proche de l'identité est mise au point afin de traiter le problème; elle donne lieu à l'introduction d'une tige effective. Les équilibres de la tige effective, plus aisés à calculer, sont des approximations potentielles pour les équilibres de la tige avec grande torsion naturelle. En général, le problème original n'a pas de symétrie continue mais, après moyennisation, on distingue deux cas:

(a) *Le problème effectif n'a pas de symétrie continue et possède des solutions isolées.* Cette situation est typique d'une tige naturellement courbée. La méthode de moyennisation couplée avec le théorème de la fonction implicite montre que, en général, tous les équilibres de la tige effective sont alors des approximations d'équilibres de la tige avec grande torsion naturelle.

(b) *Le problème effectif a une symétrie continue associée à la propriété d'isotropie et possède des familles de solutions à un paramètre.* Cette situation est typique d'une tige naturellement droite. En général, seuls quelques équilibres de la tige effective sont des approximations d'équilibres de la tige avec grande torsion naturelle. Afin d'identifier les équilibres en question, on couple la méthode de moyennisation avec une technique d'analyse de brisure de symétrie. Cette technique consiste à réduire le problème sous une forme telle que l'on puisse utiliser le théorème de Crandall-Rabinowitz de bifurcation à partir d'une valeur propre simple. Cela conduit à la définition d'une fonction de Melnikov indiquant lesquelles parmi les solutions du problème effectif sont des approximations de solutions du problème original. En général, en raison de la topologie des familles d'équilibres de la tige effective, au moins deux équilibres par famille sont des approximations d'équilibres de la tige avec grande torsion naturelle.

Dans les cas (a) et (b), on présente et justifie une méthode de perturbation permettant de calculer des approximations (à n'importe quel ordre en le petit paramètre correspondant à l'inverse de la valeur de la torsion naturelle) d'un équilibre de la tige avec grande torsion

naturelle. La brisure de symétrie ayant lieu dans le cas (b) est illustrée par le calcul explicite de la fonction de Melnikov dans des exemples faisant intervenir une tige avec une énergie élastique quadratique en les déformations, et sous l'influence de conditions au bord diverses.

On examine également le problème d'une tige avec une loi de flexion anisotrope, une petite courbure naturelle et une grande torsion naturelle. La tige effective correspondante est isotrope, et une brisure de symétrie intervient comme dans le cas (b). Dans ce cas, la fonction de Melnikov est la somme de deux termes, un associé à la courbure naturelle, et l'autre à la torsion naturelle. Cela donne lieu à une concurrence entre ces différents effets pour la détermination de la brisure de symétrie. Cette concurrence est illustrée par le calcul explicite de la fonction de Melnikov dans un exemple de tige avec une énergie élastique quadratique en les déformations.

Pour une tige avec torsion naturelle et loi de flexion anisotrope, on obtient une expression déterminant les charges critiques de flambage, et cela pour n'importe quelle valeur de la torsion naturelle. Pour de larges valeurs de cette torsion naturelle, on estime l'erreur associée à l'approximation de la première charge critique par la première charge critique de la tige effective correspondante. Cette erreur est à la fois inversément proprtionnelle à la torsion naturelle, et proportionnelle à une quantité mesurant l'anisotropie de la loi de flexion.

Acknowledgments

I am indebted to my advisor, Prof. John H. Maddocks, who proposed the fascinating subject of this dissertation. His numerous questions and ideas constituted a precious guide for my research. I am grateful to Profs. Oscar Gonzalez and Boris Buffoni for their interest in this work. They provided insightful remarks and guidance which were invaluable in the advancement of my research. I wish to acknowledge Profs. Antonio Ambrosetti, Joseph B. Keller, Robert S. Manning, Alexander Mielke, Jacobo Pejsachowicz and Charles A. Stuart for fruitful discussions.

It is a pleasure to thank all former and present members of the *Chaire d'Analyse Appliquée* for their help and support. I wish to thank my parents and parents-in-law for their encouragement, and my wife, Cathy, for her interest and patience.

This research was supported by the Swiss National Fund for Scientific Research and the EPFL.

Contents

Chapter 1

Introduction

The basic question that we address is to understand how an elastic rod with an anisotropic bending law and a high, but finite, intrinsic twist can be approximated by an effective rod with an isotropic bending law. In Kehrbaum [27] and Kehrbaum and Maddocks [29], the effective properties of a rod with high intrinsic twist are determined and, in the case of a pure initial value problem, it is shown that the (unique) equilibrium of a rod with high intrinsic twist can be approximated by the (unique) equilibrium of the effective rod. In this dissertation, we generalize these results to the more natural problems for rods, namely two-point bounday value problems. In this case, the set of equilibria of the rod with high intrinsic twist cannot always be approximated by the set of equilibria of the effective rod, as is illustrated by the following example (cf. Rey and Maddocks [45]).

Consider a rod whose unstressed shape is a uniformly twisted prism: the centerline is straight and the evolution of the asymmetric cross-section with respect to arclength consists of a rotation about the centerline with a constant spin rate (cf. figure 1.1). Such a rod is said to be intrinsically straight and to have a uniform intrinsic twist. We measure the intrinsic twist by means of a parameter τ equal to 2π times the total number of rotations of the cross-section in the rod. At each point along the rod, the

$\tau = 0$ $\qquad\qquad$ $\tau = 4\pi$

Figure 1.1: *Unstressed shapes of a prismatic rod* (left) *and of a rod with an intrinsic twist* (right). *The tube represents the centerline and the ribbon indicates the principal axis of the cross-section lying in the stiff plane for bending.*

tangent to the centerline together with one of the principal axes of the cross-section span the *soft plane* for bending, while the tangent to the centerline together with the other principal axis span the *stiff plane* for bending. Assume the rod has a standard diagonal quadratic strain energy characterized by a twist stiffness K_3 and two bending stiffnesses K_1 and K_2, with $K_1 < K_2$, associated with bending in the soft plane and in the stiff plane respectively. Now impose *clamped-free strut* boundary conditions on the rod: the lower end is vertically clamped, while the upper end is loaded by a vertical force of magnitude λ. Figure 1.2 shows a deformed configuration of the rod under such end conditions for different values of the intrinsic twist τ. When $\tau = 0$, the rod buckles in the soft plane at the clamped end. When $\tau \neq 0$, the buckled configuration is no longer planar. However, as τ increases, the equilibrium approaches a vertical plane making a $45°$ angle with the soft plane at the clamped end.

In this example, the effective rod is identical to the rod under consideration, except that it has two equal bending stiffnesses, with common value equal to the harmonic average of K_1 and K_2. Due to the symmetry of isotropy, the effective rod can buckle in any vertical plane. When applied to this example, the general results of this dissertation show that, for large, but finite, values of the intrinsic twist, the buckled equilibria of the rod can be approximated by those buckled equilibria of the effective rod lying in the two vertical planes making a $45°$ angle with the soft plane at the clamped end (cf. section 6.6). Remarkably enough, this result is independent of the values of the stiffnesses K_1, K_2, K_3 and of the load λ.

In the above example, the effective problem possesses a continuous symmetry associated with isotropy, and therefore has more solutions than the original problem. It follows that only some of the equilibria of the effective rod constitute approximations to equilibria of the rod with high intrinsic twist. This phenomenon is not exceptional; it is actually typical when the rod is intrinsically straight. In this dissertation, we develop a general averaging method involving a Melnikov function to resolve this symmetry breaking, i.e. to predict which of the equilibria of the effective rod approximate equilibria of the rod with high intrinsic twist. This method is illustrated by the explicit computation of the Melnikov function in a few particular examples similar to the one presented above, with various boundary conditions.

The situation where the effective problem possesses no continuous symmetry is also analyzed. It is shown that, in general, all equilibria of the effective rod constitute approximations to equilibria of the rod with high intrinsic twist. This case usually occurs when the rod is intrinsically curved.

In his study of a rod with high intrinsic twist, Kehrbaum [27] reformulates the equilibrium equations as a canonical Hamiltonian problem depending on two scales, with a periodic dependence in the short scale, and uses a standard averaging result for initial value problems. We adopt essentially the same two-scale Hamiltonian formulation of the equilibrium equations, and devote a substantial part of the dissertation to generalizing this standard averaging result to two-point boundary value problems, both for the regular case and for the symmetry breaking case. We present and justify a perturbation method to compute approximations (of any order in the small variable corresponding to the inverse of the value of the intrinsic twist) to a solution to the two-scale problem.

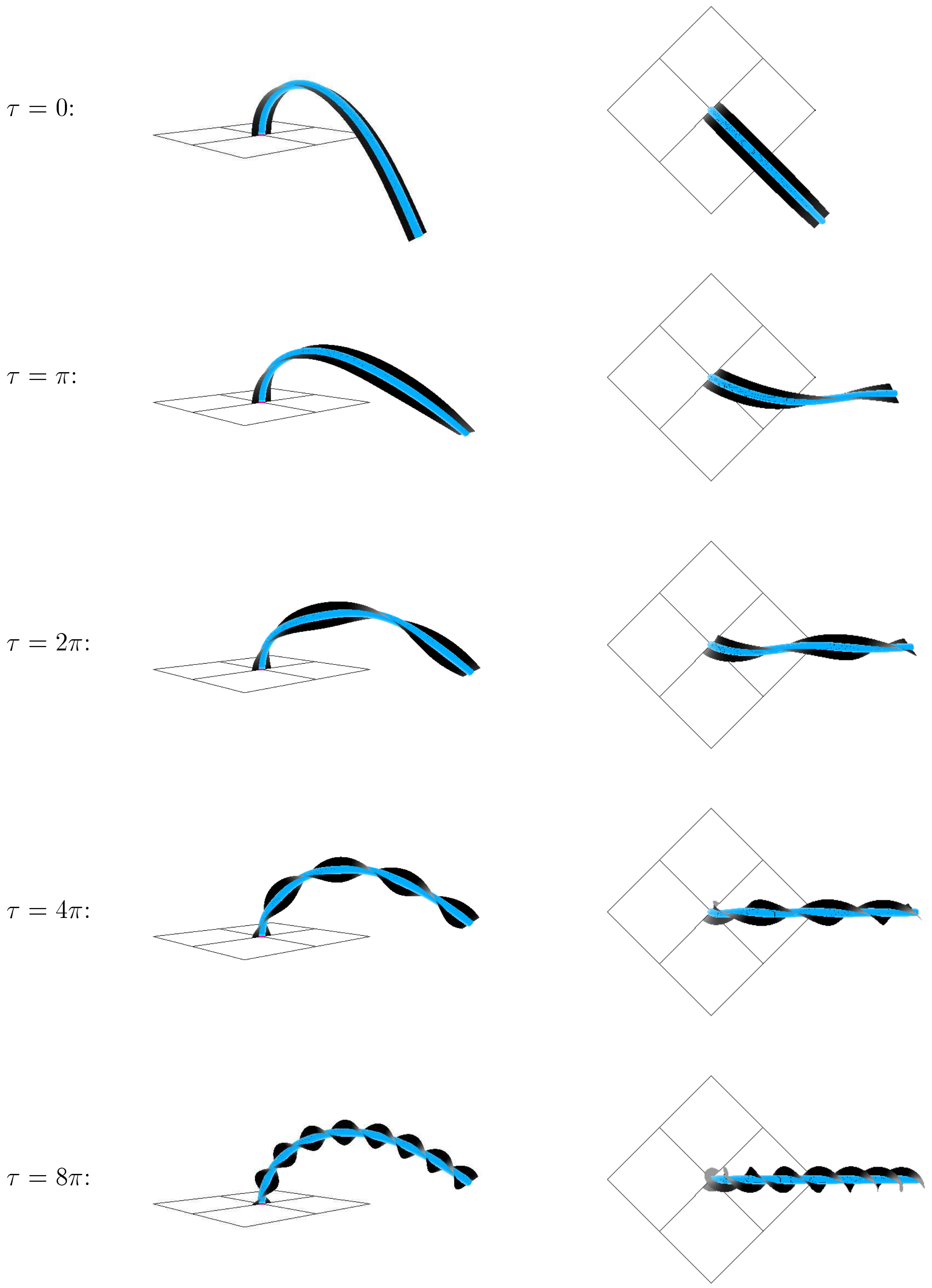


Figure 1.2: *For different values of the intrinsic twist τ, side view (left) and top view (right) of a buckled configuration of a clamped-free strut with stiffness constants $K_1 = 0.1$, $K_2 = 0.4$, $K_3 = 0.4$, and load $\lambda = 0.75$. These configurations are solutions to the governing Hamiltonian boundary value problem presented in chapter 5. They were computed by performing a parameter continuation with the software AUTO (cf. Doedel et al. [17]), starting from unbuckled solutions.*

The problem of a rod with an anisotropic bending law, a small intrinsic curvature and a high intrinsic twist is also treated, under the assumption that the intrinsic curvature and the inverse of the value of the intrinsic twist are of the same (small) order of magnitude. The associated effective rod is isotropic, and the problem usually gives rise to a symmetry breaking. The method developed for the determination of the symmetry breaking in the case of an intrinsically straight rod with high intrinsic twist also applies here. The corresponding Melnikov function is the sum of two terms associated to intrinsic curvature and intrinsic twist respectively. We illustrate this situation by the explicit computation of the Melnikov function in an example involving a rod with a standard diagonal quadratic strain energy.

One motivation for the study of elastic rods with high intrinsic twist is the use of rod models in the computation of deformations of DNA molecules containing N base pairs, where N is typically larger than 100 (cf. the review articles by Schlick [48] and Olson [44], for example). At the base pair level, the elastic properties of DNA are anisotropic: the molecule is easier to bend in the direction of the grooves than in the direction of the backbones. DNA also possesses an intrinsic curvature related to the particular sequence of base pairs, and an intrinsic twist associated to the winding of the two backbones about each other in a double helix. This intrinsic twist, corresponding to approximatively one complete turn every 10.5 base pairs, takes the value $\tau \simeq \frac{2\pi N}{10.5}$, and is typically larger than 60 for the molecules under consideration. One of the questions in DNA modeling is to determine the conditions under which the rod model with anisotropic bending law and high intrinsic twist can be replaced by an effective model obtained through averaging. The advantage of the effective model is that it obeys an isotropic bending law, and its equilibria are easier to compute. The *ring* boundary value problem (cf. section 4.3) is important for the modeling of DNA minicircles (cf. Manning *et al.* [38], for example).

This dissertation is structured as follows. In chapter 2, a general Hamiltonian two-point boundary value problem depending on a small parameter is examined. It is shown that an isolated solution to the unperturbed problem usually persists in the perturbed problem, and that, among a continuous family of solutions to the unperturbed problem, those members persisting in the perturbed problem can generally be identified by means of a Melnikov function. Perturbation expansion techniques for the computation of approximations to solutions to the perturbed problem are presented and justified. In chapter 3, an averaging method for a general Hamiltonian two-point boundary value problem depending on two different scales is developed. The associated averaged problem is introduced. By combining this averaging technique with the perturbation method of chapter 2, it is established that an isolated solution to the averaged problem usually gives rise to a solution to the two-scale problem, and that, among a continuous family of solutions to the averaged problem, those members giving rise to solutions to the two-scale problem can generally be identified by means of a suitable Melnikov function. Perturbation expansion techniques for the computation of approximations to solutions to the two-scale problem are presented and justified. Chapter 4 introduces the elastic rod model adopted in this dissertation, namely the unshearable, inextensible Cosserat model, and reviews some of its properties. Chapter 5 presents a canonical Hamiltonian formulation of the rod equilibrium equations. In chapter 6, elastic rods with high intrinsic twist subject to various boundary conditions are analyzed. The corresponding problem is written in a two-scale form. The effective rod is defined by the averaged problem, and the main results follow from the study of the symmetries of the effective problem together with the application of the averaging method of chapter 3. A summary of the results of the dissertation is presented in chapter 7, along with a discussion of alternative approaches and possible extensions. Appendix A contains a collection of standard theorems used in the thesis. Appendix B presents a variational formulation of Hamiltonian two-point boundary value problems, and in particular of the rod boundary value problems considered in the dissertation. Finally, appendix C discusses the planar equilibria of an isotropic, uniform rod under two different strut boundary conditions. The rod under consideration arises as an effective rod in our study of rods with high intrinsic twist.

Chapter 2

Perturbations in Hamiltonian boundary value problems

This chapter deals with a general canonical Hamiltonian ordinary differential equation subject to separated two-point boundary conditions. The Hamiltonian function is assumed to depend on a perturbation parameter ϵ. The perturbed problem (with ϵ small but nonzero) differs only slightly from the unperturbed problem (with $\epsilon = 0$), but is generally more complicated to solve. We establish criteria for the existence (or absence) of a solution to the perturbed problem in the vicinity of a given solution to the unperturbed problem. We treat the case where the given solution to the unperturbed problem is isolated, as well as the case where it is part of a curve of solutions. The content of the conclusions presented here is certainly not original; it can easily be obtained formally by standard asymptotic expansion techniques. The purpose of this chapter is to expose both precise statements and rigorous proofs of these conclusions. In the case where the unperturbed problem possesses a curve of solutions, we introduce a Melnikov function, which greatly simplifies the statement of the results. As a benefit of our treatment, we obtain a rigorous justification of the use of asymptotic expansion techniques for the computation of approximations to solutions to the perturbed problem.

Section 2.1 introduces the parameter-dependent Hamiltonian boundary value problem to be studied in the rest of the chapter. In section 2.2 we show that, under some nondegeneracy condition, an *isolated* solution $\boldsymbol{x}_0$ to the unperturbed problem *persists* in the perturbed problem. By that we mean that, for ϵ sufficiently small, the perturbed problem admits a (unique) solution $\boldsymbol{\xi}(\epsilon)$ in some neighborhood of $\boldsymbol{x}_0$, where $\boldsymbol{\xi}(\epsilon)$ is a smooth function of ϵ and $\boldsymbol{\xi}(0) = \boldsymbol{x}_0$. Moreover we show that approximations to $\boldsymbol{\xi}(\epsilon)$ can be computed via a perturbation method.

In section 2.3 we assume that the boundary conditions at each end-point define a Lagrangian manifold (with some nondegeneracy condition). We suppose that the unperturbed problem admits a one-parameter family Γ of solutions which is a smooth curve. In the general case, only some of the solutions of Γ persist in the perturbed problem. We give the name of *symmetry breaking* to this phenomenon since the presence of a family Γ of solutions is usually due to a continuous symmetry of the unperturbed problem, and it is because the symmetry is broken by the perturbation that only some of the solutions of Γ persist in the perturbed problem. To study this symmetry breaking we introduce a Melnikov function which associates a real number to each solution in Γ. Under some nondegeneracy conditions, this function indicates which of the solutions in Γ persist in the perturbed problem: if $\boldsymbol{x}^* \in \Gamma$ corresponds to a nonzero value of the Melnikov function, then, for ϵ sufficiently small, the perturbed problem does not admit any solution in some neighborhood of $\boldsymbol{x}^*$; contrariwise, if $\boldsymbol{x}^* \in \Gamma$ corresponds to a simple zero of the Melnikov function, then, for ϵ sufficiently small, the perturbed problem admits a (unique) solution $\boldsymbol{\xi}(\epsilon)$ in some neighborhood of $\boldsymbol{x}^*$, where $\boldsymbol{\xi}(\epsilon)$ is a smooth function of ϵ and $\boldsymbol{\xi}(0) = \boldsymbol{x}^*$. In the latter case we show that approximations to $\boldsymbol{\xi}(\epsilon)$ can be computed via a perturbation method.

2.1 Hamiltonian boundary value problem

In this chapter we study the canonical Hamiltonian boundary value problem (BVP)

$$\begin{cases} \boldsymbol{q}'(s) = H_{\boldsymbol{p}}(\boldsymbol{q}(s), \boldsymbol{p}(s), s, \epsilon), \quad \boldsymbol{p}'(s) = -H_{\boldsymbol{q}}(\boldsymbol{q}(s), \boldsymbol{p}(s), s, \epsilon), \quad \forall\, s \in (0,1), \\ \boldsymbol{b}(\boldsymbol{q}(0), \boldsymbol{p}(0)) = \boldsymbol{0}, \quad \boldsymbol{c}(\boldsymbol{q}(1), \boldsymbol{p}(1)) = \boldsymbol{0}, \end{cases} \tag{2.1}$$

where $\boldsymbol{q}, \boldsymbol{p} : [0,1] \to \mathbb{R}^n$ are the two conjugate variables, ϵ denotes a small parameter, and $H : \mathbb{R}^{2n} \times [0,1] \times \mathbb{R} \to \mathbb{R}$ is the Hamiltonian. The (possibly nonlinear) boundary conditions comprise n scalar conditions imposed at $s = 0$ via the mapping $\boldsymbol{b} : \mathbb{R}^{2n} \to \mathbb{R}^n$ and n scalar conditions imposed at $s = 1$ via the mapping $\boldsymbol{c} : \mathbb{R}^{2n} \to \mathbb{R}^n$.

Defining the phase variable $\boldsymbol{x} = (\boldsymbol{q}, \boldsymbol{p})$ and the canonical symplectic matrix

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I} \\ -\boldsymbol{I} & \boldsymbol{0} \end{pmatrix}, \tag{2.2}$$

where $\boldsymbol{0}$ and $\boldsymbol{I}$ are the $n \times n$ zero and identity matrices, BVP (2.1) can be rewritten as

$$\begin{cases} \boldsymbol{x}'(s) = \boldsymbol{J} H_{\boldsymbol{x}}(\boldsymbol{x}(s), s, \epsilon), \quad \forall\, s \in (0,1), \\ \boldsymbol{b}(\boldsymbol{x}(0)) = \boldsymbol{0}, \quad \boldsymbol{c}(\boldsymbol{x}(1)) = \boldsymbol{0}. \end{cases} \tag{2.3}$$

Let the norm of $\boldsymbol{z} = (z_1, \ldots, z_{2n}) \in \mathbb{R}^{2n}$ be taken as

$$\|\boldsymbol{z}\|_{\mathbb{R}^{2n}} = \max_{1 \leqslant i \leqslant 2n} |z_i|, \tag{2.4}$$

and let $\mathcal{C}^0$ and $\mathcal{C}^1$ denote the Banach spaces $C^0([0,1], \mathbb{R}^{2n})$ and $C^1([0,1], \mathbb{R}^{2n})$ with norms defined by

$$\|\boldsymbol{x}\|_{\mathcal{C}^0} = \max_{0 \leqslant s \leqslant 1} \|\boldsymbol{x}(s)\|_{\mathbb{R}^{2n}} \quad \text{and} \quad \|\boldsymbol{x}\|_{\mathcal{C}^1} = \max(\|\boldsymbol{x}\|_{\mathcal{C}^0}, \|\boldsymbol{x}'\|_{\mathcal{C}^0}). \tag{2.5}$$

We call $\boldsymbol{x}$ a *solution* to BVP (2.3) if $\boldsymbol{x}$ is in $\mathcal{C}^1$ and satisfies (2.3).

2.2 Regular perturbation

In this section we assume $\boldsymbol{x}_0$ is an isolated solution to BVP (2.3) when $\epsilon = 0$. We study the existence of a solution close to $\boldsymbol{x}_0$ when ϵ is small. Our treatment consists essentially in a careful application of the implicit function theorem.

Assume the functions H, $\boldsymbol{b}$ and $\boldsymbol{c}$ satisfy the regularity conditions

$$H \in C^{k+1}(\mathbb{R}^{2n} \times [0,1] \times \mathbb{R}, \mathbb{R}), \quad \boldsymbol{b}, \boldsymbol{c} \in C^k(\mathbb{R}^{2n}, \mathbb{R}^n), \quad \text{with } k \geqslant 1. \tag{\textbf{2.A1}}$$

Suppose

$$\boldsymbol{x}_0 \text{ is a solution to (2.3) when } \epsilon = 0 \tag{\textbf{2.A2}}$$

and consider the corresponding homogeneous linearized BVP

$$\begin{cases} \boldsymbol{v}'(s) - \boldsymbol{J} H_{\boldsymbol{xx}}(\boldsymbol{x}_0(s), s, 0)\boldsymbol{v}(s) = \boldsymbol{0}, \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}_0(0))\boldsymbol{v}(0) = \boldsymbol{0}, \quad \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}_0(1))\boldsymbol{v}(1) = \boldsymbol{0}. \end{cases} \tag{2.6}$$

Finally assume

$$\text{BVP (2.6) admits only the trivial solution } \boldsymbol{v} = \boldsymbol{0}, \tag{\textbf{2.A3}}$$

which is a nondegeneracy condition. As will be shown by theorem 2.2, conditions (**2.A1**)-(**2.A3**) imply that $\boldsymbol{x}_0$ is an *isolated* solution to the unperturbed problem (2.3) with $\epsilon = 0$, in the sense that there exists a $\mathcal{C}^1$-neighborhood of $\boldsymbol{x}_0$ containing no other solution.

Lemma 2.1 *Let conditions (**2.A1**)-(**2.A3**) be satisfied. Then, for any $\boldsymbol{\eta} \in \mathcal{C}^0$ and $\boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{R}^n$, the inhomogeneous linear BVP*

$$\begin{cases} \boldsymbol{v}'(s) - \boldsymbol{J} H_{\boldsymbol{xx}}(\boldsymbol{x}_0(s), s, 0)\boldsymbol{v}(s) = \boldsymbol{\eta}(s), \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}_0(0))\boldsymbol{v}(0) = \boldsymbol{\beta}, \quad \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}_0(1))\boldsymbol{v}(1) = \boldsymbol{\gamma}, \end{cases} \tag{2.7}$$

has a unique solution.

Proof By the variation of constants formula (cf. Hartman [24, sect. IV 2]), the general solution to the inhomogeneous linear differential equation in (2.7) can be expressed as

$$\boldsymbol{v}(s) = \boldsymbol{Z}(s)\boldsymbol{v}(0) + \boldsymbol{Z}(s) \int_0^s \boldsymbol{Z}^{-1}(\tau)\boldsymbol{\eta}(\tau)\, d\tau, \tag{2.8}$$

where $\boldsymbol{Z}(s)$ is the $2n \times 2n$ fundamental matrix satisfying

$$\boldsymbol{Z}'(s) - \boldsymbol{J} H_{\boldsymbol{xx}}(\boldsymbol{x}_0(s), s, 0)\boldsymbol{Z}(s) = \boldsymbol{0}, \quad \boldsymbol{Z}(0) = \boldsymbol{I}. \tag{2.9}$$

In particular, we have

$$\boldsymbol{v}(1) = \boldsymbol{Z}(1)\boldsymbol{v}(0) + \boldsymbol{Z}(1) \int_0^1 \boldsymbol{Z}^{-1}(\tau)\boldsymbol{\eta}(\tau)\, d\tau, \tag{2.10}$$

and BVP (2.7) reduces to the matrix equation

$$\boldsymbol{M}\boldsymbol{v}(0) = \boldsymbol{a}, \tag{2.11}$$

where

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}_0(0)) \\ \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}_0(1))\boldsymbol{Z}(1) \end{pmatrix}, \quad \boldsymbol{a} = \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} - \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}_0(1))\boldsymbol{Z}(1) \int_0^1 \boldsymbol{Z}^{-1}(\tau)\boldsymbol{\eta}(\tau)\, d\tau \end{pmatrix}. \tag{2.12}$$

Setting $(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = \boldsymbol{0}$ and using hypothesis (**2.A3**) shows that in the case $\boldsymbol{a} = \boldsymbol{0}$, equation (2.11) admits only the trivial solution $\boldsymbol{v}(0) = \boldsymbol{0}$. Thus the $2n \times 2n$ matrix $\boldsymbol{M}$ has full rank and is invertible. Therefore equation (2.11) admits a unique solution for any $(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma})$, and so does BVP (2.7). ■

Theorem 2.2 *Let conditions (**2.A1**)-(**2.A3**) be satisfied. Then there exist a positive number ϵ_0, a $\mathcal{C}^1$-neighborhood $\mathcal{X}$ of $\boldsymbol{x}_0$, and a function $\boldsymbol{\xi} \in C^k((-\epsilon_0, \epsilon_0), \mathcal{X})$ such that $\boldsymbol{\xi}(0) = \boldsymbol{x}_0$ and, for any ϵ satisfying $|\epsilon| < \epsilon_0$, $\boldsymbol{x} = \boldsymbol{\xi}(\epsilon)$ is the unique solution to BVP (2.3) in $\mathcal{X}$.*

Proof Define the operator version $h : \mathcal{C}^1 \times \mathbb{R} \to \mathcal{C}^0$ of the Hamiltonian by

$$h(\boldsymbol{x}, \epsilon)(s) = H(\boldsymbol{x}(s), s, \epsilon), \quad \forall\, s \in [0, 1], \tag{2.13}$$

and the operator $\boldsymbol{\Phi} \in C^k(\mathcal{C}^1 \times \mathbb{R}, \mathcal{C}^0 \times \mathbb{R}^{2n})$ by

$$\boldsymbol{\Phi}(\boldsymbol{x}, \epsilon) = \begin{pmatrix} \boldsymbol{x}' - \boldsymbol{J} h_{\boldsymbol{x}}(\boldsymbol{x}, \epsilon) \\ \boldsymbol{b}(\boldsymbol{x}(0)) \\ \boldsymbol{c}(\boldsymbol{x}(1)) \end{pmatrix}, \tag{2.14}$$

so that $\boldsymbol{x} \in \mathcal{C}^1$ is a solution to BVP (2.3) if and only if $\boldsymbol{\Phi}(\boldsymbol{x}, \epsilon) = \boldsymbol{0}$. By assumption (**2.A2**), $\boldsymbol{\Phi}(\boldsymbol{x}_0, 0) = \boldsymbol{0}$. The Fréchet partial derivative $\boldsymbol{\Phi}_{\boldsymbol{x}}(\boldsymbol{x}_0, 0)$ is defined by

$$\boldsymbol{\Phi}_{\boldsymbol{x}}(\boldsymbol{x}_0, 0)\boldsymbol{v} = \begin{pmatrix} \boldsymbol{v}' - \boldsymbol{J} h_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}_0, 0)\boldsymbol{v} \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}_0(0))\boldsymbol{v}(0) \\ \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}_0(1))\boldsymbol{v}(1) \end{pmatrix}. \tag{2.15}$$

By condition (**2.A3**), $\boldsymbol{\Phi}_{\boldsymbol{x}}(\boldsymbol{x}_0, 0)$ is injective, and, by lemma (2.1), it is surjective. Thus the implicit function theorem (cf. appendix A) applies, and the conclusion immediately follows. ■

By Taylor's theorem (cf. appendix A), under the assumptions of theorem 2.2,

$$\boldsymbol{\xi}(\epsilon) = \boldsymbol{x}_0 + O(\epsilon), \quad \text{as } \epsilon \to 0, \tag{2.16}$$

and we call $\boldsymbol{x}_0$ the *leading order approximation* to the solution $\boldsymbol{\xi}(\epsilon)$. Note that, by definition of the $\mathcal{C}^1$-norm, the above equation means that

$$\begin{aligned} &\boldsymbol{\xi}(\epsilon)(s) = \boldsymbol{x}_0(s) + O(\epsilon) \quad \text{and} \quad \boldsymbol{\xi}'(\epsilon)(s) = \boldsymbol{x}_0'(s) + O(\epsilon) \\ &\text{uniformly in } s \in [0, 1], \text{ as } \epsilon \to 0. \end{aligned} \tag{2.17}$$

A standard method to compute a better approximation to a solution to BVP (2.3) is a perturbation expansion of the solution in powers of ϵ. The rest of this section shows that, under the assumptions of theorem 2.2 and for $k \geqslant 2$, this formal method can always be applied and is rigorously justified.

By Taylor's theorem (cf. appendix A), the solution $\boldsymbol{\xi}(\epsilon)$ can be expanded about $\boldsymbol{x}_0$ as

$$\boldsymbol{\xi}(\epsilon) = \boldsymbol{x}_0 + \epsilon \boldsymbol{x}_1 + \ldots + \epsilon^{k-1} \boldsymbol{x}_{k-1} + O(\epsilon^k), \quad \text{as } \epsilon \to 0. \tag{2.18}$$

We now show that the functions $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k-1}$ can be uniquely determined. Substituting Taylor series (2.18) into BVP (2.3), expanding in powers of ϵ, and equating coefficients of like powers of ϵ, we find, for example, that $\boldsymbol{x}_1$ has to satisfy

$$\begin{cases} \boldsymbol{x}_1'(s) - \boldsymbol{J} H_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}_0(s), s, 0)\boldsymbol{x}_1(s) = \boldsymbol{J} H_{\boldsymbol{x}\epsilon}(\boldsymbol{x}_0(s), s, 0), \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}_0(0))\boldsymbol{x}_1(0) = \boldsymbol{0}, \quad \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}_0(1))\boldsymbol{x}_1(1) = \boldsymbol{0}, \end{cases} \tag{2.19}$$

and $\boldsymbol{x}_2$ has to satisfy

$$\begin{cases} \boldsymbol{x}_2'(s) - \boldsymbol{J}H_{\boldsymbol{xx}}(\boldsymbol{x}_0(s),s,0)\boldsymbol{x}_2(s) = \frac{1}{2}\boldsymbol{J}H_{\boldsymbol{xxx}}(\boldsymbol{x}_0(s),s,0)[\boldsymbol{x}_1(s),\boldsymbol{x}_1(s)] \\ \qquad\qquad + \boldsymbol{J}H_{\boldsymbol{xx}\epsilon}(\boldsymbol{x}_0(s),s,0)\boldsymbol{x}_1(s) + \frac{1}{2}\boldsymbol{J}H_{\boldsymbol{x}\epsilon\epsilon}(\boldsymbol{x}_0(s),s,0), \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}_0(0))\boldsymbol{x}_2(0) = -\frac{1}{2}\boldsymbol{b}_{\boldsymbol{xx}}(\boldsymbol{x}_0(0))[\boldsymbol{x}_1(0),\boldsymbol{x}_1(0)], \\ \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}_0(1))\boldsymbol{x}_2(1) = -\frac{1}{2}\boldsymbol{c}_{\boldsymbol{xx}}(\boldsymbol{x}_0(1))[\boldsymbol{x}_1(1),\boldsymbol{x}_1(1)]. \end{cases} \tag{2.20}$$

In fact it is easy to show by induction that, for $i \in \{1,\ldots,k-1\}$, $\boldsymbol{x}_i$ has to satisfy an inhomogeneous linear BVP of the form

$$\begin{cases} \boldsymbol{x}_i'(s) - \boldsymbol{J}H_{\boldsymbol{xx}}(\boldsymbol{x}_0(s),s,0)\boldsymbol{x}_i(s) = \boldsymbol{J}\boldsymbol{\eta}_i(\boldsymbol{x}_0(s),\ldots,\boldsymbol{x}_{i-1}(s),s), \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}_0(0))\boldsymbol{x}_i(0) = \boldsymbol{\beta}_i(\boldsymbol{x}_0(0),\ldots,\boldsymbol{x}_{i-1}(0)), \\ \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}_0(1))\boldsymbol{x}_i(1) = \boldsymbol{\gamma}_i(\boldsymbol{x}_0(1),\ldots,\boldsymbol{x}_{i-1}(1)), \end{cases} \tag{2.21}$$

where $\boldsymbol{\eta}_i$, $\boldsymbol{\beta}_i$ and $\boldsymbol{\gamma}_i$ are known functions. This set of BVPs can be solved successively, starting with $i=1$. Lemma 2.1 ensures the existence and uniqueness of the solutions $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{k-1}$. In view of expansion (2.18), we call $\boldsymbol{x}_1$ the *first correction* to the solution $\boldsymbol{\xi}(\epsilon)$, $\boldsymbol{x}_2$ the *second correction*, and so on.

2.3 Symmetry breaking

In this section we suppose that the boundary conditions of problem (2.3) satisfy a certain property. This property, which is relatively easy to verify in applications, is equivalent to the fact that the boundary conditions define Lagrangian manifolds at $s=0$ and at $s=1$ together with some nondegeneracy conditions. We assume the unperturbed problem (2.3) with $\epsilon=0$ admits a smooth curve Γ of solutions, and consider a particular solution $\boldsymbol{x}^* \in \Gamma$. Under some nondegeneracy conditions which force Γ to be an isolated curve of solutions to the unperturbed problem (at least in some neighborhood of $\boldsymbol{x}^*$), we present criteria both for the persistence and for the nonpersistence of $\boldsymbol{x}^*$ in the perturbed problem (2.3).

The main ingredient in this study is a method used by Stuart [51] for the treatment of a problem of bifurcation of homoclinic orbits. In that context, the method yields criteria (expressed in terms of a Melnikov function) for the persistence of a nontrivial homoclinic orbit under nonperiodic perturbations. This method consists in performing a suitable projection of the dependent variable to reduce the problem to a situation where the persistence of a solution can be directly established by the Crandall-Rabinowitz theorem (cf. appendix A) on bifurcation from a simple eigenvalue. As discussed in chapter 7, the results presented in this section are also related to the variational results of Ambrosetti and Badiale [3, 4].

Here we require the functions H, $\boldsymbol{b}$ and $\boldsymbol{c}$ to be slightly more regular than in the previous section by assuming condition (**2.A1**) with $k \geqslant 2$, namely

$$H \in C^{k+1}(\mathbb{R}^{2n}\times[0,1]\times\mathbb{R},\mathbb{R}), \quad \boldsymbol{b},\boldsymbol{c} \in C^k(\mathbb{R}^{2n},\mathbb{R}^n), \quad \text{with } k \geqslant 2. \tag{\textbf{2.A4}}$$

Consider the sets

$$L_0 = \left\{\boldsymbol{\mu} \in \mathbb{R}^{2n} : \boldsymbol{b}(\boldsymbol{\mu}) = \boldsymbol{0}\right\}, \qquad L_1 = \left\{\boldsymbol{\sigma} \in \mathbb{R}^{2n} : \boldsymbol{c}(\boldsymbol{\sigma}) = \boldsymbol{0}\right\}, \tag{2.22}$$

and assume the functions $\boldsymbol{b}$ and $\boldsymbol{c}$ defining the boundary conditions are such that

$$\begin{aligned} &\forall \boldsymbol{\mu} \in L_0, \quad \boldsymbol{\nu}_1, \boldsymbol{\nu}_2 \in \text{Ker } \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu}) \Longrightarrow \boldsymbol{\nu}_1^T \boldsymbol{J} \boldsymbol{\nu}_2 = 0, \\ &\forall \boldsymbol{\sigma} \in L_1, \quad \boldsymbol{\nu}_1, \boldsymbol{\nu}_2 \in \text{Ker } \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{\sigma}) \Longrightarrow \boldsymbol{\nu}_1^T \boldsymbol{J} \boldsymbol{\nu}_2 = 0. \end{aligned} \tag{\textbf{2.A5}}$$

In appendix B we present a variational problem associated to a large class of Hamiltonian BVPs of the form (2.3), and show that the BVPs contained in that class automatically inherit property (**2.A5**). In particular the rod BVPs considered in this dissertation are all in that class.

Lemma 2.3 *Suppose conditions (**2.A4**) and (**2.A5**) are satisfied. Then* $\mathbf{0} \in \mathbb{R}^n$ *is a* regular value *of* $\boldsymbol{b}$ *in the sense that, for any* $\boldsymbol{\mu} \in L_0$, $\boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu}) : \mathbb{R}^{2n} \to \mathbb{R}^n$ *is surjective. Similarly,* $\mathbf{0} \in \mathbb{R}^n$ *is a* regular value *of* $\boldsymbol{c}$ *in the sense that, for any* $\boldsymbol{\sigma} \in L_1$, $\boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{\sigma}) : \mathbb{R}^{2n} \to \mathbb{R}^n$ *is surjective. It follows that, for any* $\boldsymbol{\mu} \in L_0$ *and* $\boldsymbol{\sigma} \in L_1$*, the restrictions*

$$\boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu}) \big|_{(Ker\, \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu}))^\perp} : \big(Ker\, \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu})\big)^\perp \to \mathbb{R}^n \tag{2.23}$$

and

$$\boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{\sigma}) \big|_{(Ker\, \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{\sigma}))^\perp} : \big(Ker\, \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{\sigma})\big)^\perp \to \mathbb{R}^n \tag{2.24}$$

are bijective.

Proof Let $l = \dim \text{Ker } \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu})$ and $\{\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_l\}$ be an orthogonal basis for Ker $\boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu})$. Then, for $i \neq j$,

$$(\boldsymbol{J}\boldsymbol{\nu}_i)^T(\boldsymbol{J}\boldsymbol{\nu}_j) = \boldsymbol{\nu}_i^T(\boldsymbol{J}^T\boldsymbol{J})\boldsymbol{\nu}_j = \boldsymbol{\nu}_i^T\boldsymbol{\nu}_j = 0, \tag{2.25}$$

so that $\{\boldsymbol{J}\boldsymbol{\nu}_1, \ldots, \boldsymbol{J}\boldsymbol{\nu}_l\}$ is an orthogonal family of vectors. Therefore, by assumption (**2.A5**), $\{\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_l, \boldsymbol{J}\boldsymbol{\nu}_1, \ldots, \boldsymbol{J}\boldsymbol{\nu}_l\}$ constitutes an orthogonal family of $2l$ vectors in $\mathbb{R}^{2n}$. It follows that $l \leqslant n$ and, by the rank theorem,

$$\text{rank } \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu}) = 2n - l \geqslant n. \tag{2.26}$$

On the other hand, Range $\boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu}) \subset \mathbb{R}^n$, so that rank $\boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu}) \leqslant n$. We conclude that rank $\boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu}) = n$, and $\boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu})$ is surjective. It follows that linear mapping (2.23) is also surjective. Moreover the problem

$$\boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu}) \big|_{(\text{Ker } \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu}))^\perp} \boldsymbol{\nu} = \mathbf{0} \tag{2.27}$$

implies

$$\boldsymbol{\nu} \in \text{Ker } \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu}) \cap \big(\text{Ker } \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu})\big)^\perp = \{\mathbf{0}\}, \tag{2.28}$$

so that linear mapping (2.23) is injective as well, and therefore bijective. The surjectivity of $\boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{\sigma})$ and the bijectivity of (2.24) are verified in the same way. ■

Definition 2.4 *A submanifold* L *of* $\mathbb{R}^{2n}$ *is called* Lagrangian *if it has dimension* n *and satisfies the property*

$$\forall \boldsymbol{\mu} \in L, \quad \boldsymbol{\nu}_1, \boldsymbol{\nu}_2 \in T_{\boldsymbol{\mu}}L \Longrightarrow \boldsymbol{\nu}_1^T \boldsymbol{J} \boldsymbol{\nu}_2 = 0, \tag{2.29}$$

where $T_{\boldsymbol{\mu}}L$ *denotes the tangent space to* L *at* $\boldsymbol{\mu}$.

Lemma 2.5 *Let assumption (**2.A4**) be satisfied. Then property (**2.A5**) is equivalent to the fact that L_0 and L_1 are Lagrangian submanifolds of $\mathbb{R}^{2n}$ with tangent spaces given by*

$$T_{\boldsymbol{\mu}} L_0 = Ker\, \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu}) \quad and \quad T_{\boldsymbol{\sigma}} L_1 = Ker\, \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{\sigma}), \tag{2.30}$$

for all $\boldsymbol{\mu} \in L_0$ and $\boldsymbol{\sigma} \in L_1$.

Equalities (2.30) are a form of nondegeneracy conditions. Note that if L_0 and L_1 defined by (2.22) are submanifolds, then the inclusions

$$T_{\boldsymbol{\mu}} L_0 \subset \mathrm{Ker}\, \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu}) \quad \text{and} \quad T_{\boldsymbol{\sigma}} L_1 \subset \mathrm{Ker}\, \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{\sigma}), \tag{2.31}$$

are automatically satisfied, but the equalities (2.30) are not.

Proof Suppose property (**2.A5**) is satisfied. Then, by lemma 2.3 and the submersion theorem (cf. appendix A), L_0 and L_1 are C^k-submanifolds of dimension n of $\mathbb{R}^{2n}$, with tangent spaces given by (2.30). Hence, by property (**2.A5**), L_0 and L_1 are Lagrangian. The converse is immediate. ∎

Let I be an open interval in $\mathbb{R}$ and $\theta^* \in I$. Suppose

$$\begin{aligned} &\text{BVP (2.3) has a family of solutions } \Gamma = \boldsymbol{\zeta}(I) \text{ when } \epsilon = 0, \\ &\text{with } \boldsymbol{\zeta} \in C^k(I, \mathcal{C}^1), \text{ and } \boldsymbol{\zeta}_\theta(\theta^*) \neq \mathbf{0}. \end{aligned} \tag{\textbf{2.A6}}$$

We will use the notation

$$\boldsymbol{x}^* = \boldsymbol{\zeta}(\theta^*), \quad \boldsymbol{x}^*_\theta = \boldsymbol{\zeta}_\theta(\theta^*), \quad \text{and} \quad \boldsymbol{x}^*_{\theta\theta} = \boldsymbol{\zeta}_{\theta\theta}(\theta^*). \tag{2.32}$$

By assumption (**2.A6**), for any $\theta \in I$, we have

$$\begin{cases} \boldsymbol{\zeta}(\theta)'(s) - \boldsymbol{J} H_{\boldsymbol{x}}(\boldsymbol{\zeta}(\theta)(s), s, 0) = \mathbf{0}, \\ \boldsymbol{b}(\boldsymbol{\zeta}(\theta)(0)) = \mathbf{0}, \quad \boldsymbol{c}(\boldsymbol{\zeta}(\theta)(1)) = \mathbf{0}, \end{cases} \tag{2.33}$$

where the prime denotes differentiation with respect to s. Differentiating this BVP with respect to θ and setting $\theta = \theta^*$ shows that $\boldsymbol{x}^*_\theta$ is a nontrivial solution to the homogeneous linear BVP

$$\begin{cases} \boldsymbol{v}'(s) - \boldsymbol{J} H_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{v}(s) = \mathbf{0}, \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))\boldsymbol{v}(0) = \mathbf{0}, \quad \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\boldsymbol{v}(1) = \mathbf{0}. \end{cases} \tag{2.34}$$

In fact we assume $\boldsymbol{x}^*_\theta$ is the *only* nontrivial solution to this problem (up to a multiplicative constant), i.e.

$$\text{the set of solutions to BVP (2.34) is } \mathrm{Span}\{\boldsymbol{x}^*_\theta\}. \tag{\textbf{2.A7}}$$

The following projection lemma (cf. Stuart [51, remark 2, page 1155]) is essential in our analysis. It shows that any $\boldsymbol{x}$ in a $\mathcal{C}^1$-neighborhood of $\boldsymbol{x}^*$ can be uniquely decomposed as

$$\boldsymbol{x} = \boldsymbol{\zeta}(\theta) + \boldsymbol{w}, \tag{2.35}$$

with θ near θ^* and $\boldsymbol{w} \in W$, where

$$W = \left\{ \boldsymbol{v} \in \mathcal{C}^1 : \int_0^1 \boldsymbol{v}(s)^T \boldsymbol{x}^*_\theta(s)\, ds = 0 \right\}. \tag{2.36}$$

Lemma 2.6 *Let condition (**2.A6**) be satisfied. Define the mapping* $\boldsymbol{\psi} : I \times W \to \mathcal{C}^1$ *by*

$$\boldsymbol{\psi}(\theta, \boldsymbol{w}) = \boldsymbol{\zeta}(\theta) + \boldsymbol{w}. \tag{2.37}$$

Then there exist an open neighborhood $I^* \times W^*$ *of* $(\theta^*, \mathbf{0})$ *in* $I \times W$ *and a neighborhood* X^* *of* $\boldsymbol{x}^*$ *in* $\mathcal{C}^1$ *such that* $\boldsymbol{\psi}$ *defines a* C^k*-diffeomorphism between* $I^* \times W^*$ *and* X^*.

Proof The mapping $\boldsymbol{\psi} \in C^k(I \times W, \mathcal{C}^1)$ is such that $\boldsymbol{\psi}(\theta^*, \mathbf{0}) = \boldsymbol{x}^*$. Its derivative $\boldsymbol{\psi}_{(\theta,\boldsymbol{w})}(\theta^*, \mathbf{0}) : \mathbb{R} \times W \to \mathcal{C}^1$ is defined by

$$\boldsymbol{\psi}_{(\theta,\boldsymbol{w})}(\theta^*, \mathbf{0})(\nu, \boldsymbol{v}) = \nu \boldsymbol{x}^*_\theta + \boldsymbol{v}. \tag{2.38}$$

Using the direct sum decomposition $\mathcal{C}^1 = \mathrm{Span}\{\boldsymbol{x}^*_\theta\} \oplus W$, we conclude that this derivative is bijective. Then, by the inverse mapping theorem (cf. appendix A), there exist a neighborhood U_0 of $(\theta^*, \mathbf{0})$ in $I \times W$ and a neighborhood V_0 of $\boldsymbol{x}^*$ in $\mathcal{C}^1$ such that $\boldsymbol{\psi}$ defines a C^k-diffeomorphism between U_0 and V_0.

Now choose I^* and W^* such that $I^* \times W^*$ is an open neighborhood of $(\theta^*, \mathbf{0})$ in U_0 and let $X^* = \boldsymbol{\psi}(I^* \times W^*)$. Then $\boldsymbol{\psi}$ defines a C^k-diffeomorphism between $I^* \times W^*$ and X^*. Moreover X^* is a neighborhood of $\boldsymbol{x}^*$. Indeed, X^* contains $\boldsymbol{x}^*$ and is an open set since it corresponds to the inverse image of the open set $I^* \times W^*$ by the continuous mapping $\boldsymbol{\psi}^{-1}$. ∎

For $(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathcal{C}^0 \times \mathbb{R}^n \times \mathbb{R}^n$, the inhomogeneous linear BVP

$$\begin{cases} \boldsymbol{v}'(s) - \boldsymbol{J} H_{\boldsymbol{xx}}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{v}(s) = \boldsymbol{\eta}(s), \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))\boldsymbol{v}(0) = \boldsymbol{\beta}, \quad \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\boldsymbol{v}(1) = \boldsymbol{\gamma}, \end{cases} \tag{2.39}$$

obeys the following version of the Fredholm alternative.

Lemma 2.7 *Let conditions (**2.A4**), (**2.A6**) and (**2.A7**) be satisfied. Denote by* $\mathcal{R} \subset \mathcal{C}^0 \times \mathbb{R}^{2n}$ *the set of* $(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ *for which BVP (2.39) is solvable. Then* $\mathcal{R}$ *is a closed subspace of* $\mathcal{C}^0 \times \mathbb{R}^{2n}$ *and there exists* $\boldsymbol{\rho} \in \mathcal{C}^0 \times \mathbb{R}^{2n}$, $\boldsymbol{\rho} \neq \mathbf{0}$, *such that*

$$\mathcal{C}^0 \times \mathbb{R}^{2n} = \mathcal{R} \oplus \mathit{Span}\{\boldsymbol{\rho}\}. \tag{2.40}$$

Moreover, in the case where BVP (2.39) is solvable, there is a unique solution $\hat{\boldsymbol{v}} \in W$, *and the set of solutions in* $\mathcal{C}^1$ *is* $\{\hat{\boldsymbol{v}} + \tau \boldsymbol{x}^*_\theta : \tau \in \mathbb{R}\}$.

Proof We start by showing that there exists $(\boldsymbol{\mu}, \boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \in \mathcal{C}^0 \times \mathbb{R}^n \times \mathbb{R}^n$, $(\boldsymbol{\mu}, \boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \neq \mathbf{0}$, such that

$$\mathcal{R} = \left\{ (\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathcal{C}^0 \times \mathbb{R}^n \times \mathbb{R}^n : \int_0^1 \boldsymbol{\mu}(s)^T \boldsymbol{\eta}(s)\, ds + \boldsymbol{\nu}_1^T \boldsymbol{\beta} + \boldsymbol{\nu}_2^T \boldsymbol{\gamma} = 0 \right\}. \tag{2.41}$$

According to the proof of lemma 2.1, BVP (2.39) is solvable if and only if the matrix equation

$$\boldsymbol{M}\boldsymbol{v}(0) = \boldsymbol{a} \tag{2.42}$$

is solvable, where

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0)) \\ \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\boldsymbol{Z}(1) \end{pmatrix}, \quad \boldsymbol{a} = \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} - \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\boldsymbol{Z}(1) \int_0^1 \boldsymbol{Z}^{-1}(s)\boldsymbol{\eta}(s)\, ds \end{pmatrix}, \tag{2.43}$$

and $\boldsymbol{Z}$ is some fundamental matrix. Setting $(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = \mathbf{0}$ and using hypothesis (**2.A7**) shows that, in the case $\boldsymbol{a} = \mathbf{0}$, the solution set to matrix equation (2.42) is spanned by $\boldsymbol{x}^*_\theta(0)$. Therefore $\dim \operatorname{Ker} \boldsymbol{M} = 1$ and, by the rank theorem, $\operatorname{codim} \operatorname{Range} \boldsymbol{M} = 1$. Thus there exists $\boldsymbol{\nu} = (\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \in \mathbb{R}^{2n}$, $\boldsymbol{\nu} \neq \mathbf{0}$, such that (2.42) is solvable if and only if $\boldsymbol{\nu}^T \boldsymbol{a} = 0$. At this point, $\boldsymbol{a}$ can be replaced by its expression given in (2.43) to conclude that BVP (2.39) is solvable if and only if

$$\int_0^1 \boldsymbol{\mu}(s)^T \boldsymbol{\eta}(s)\, ds + \boldsymbol{\nu}_1^T \boldsymbol{\beta} + \boldsymbol{\nu}_2^T \boldsymbol{\gamma} = 0, \tag{2.44}$$

where $\boldsymbol{\mu}(s) = -\left(\boldsymbol{\nu}_2^T \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\boldsymbol{Z}(1)\boldsymbol{Z}^{-1}(s)\right)^T$. This proves equality (2.41).

We now show that $\mathcal{R}$ is closed. To do so, suppose $(\bar{\boldsymbol{\eta}}, \bar{\boldsymbol{\beta}}, \bar{\boldsymbol{\gamma}})$ is the limit of a sequence in $\mathcal{R}$. In other words, suppose that, for any $\delta > 0$, there exists $(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathcal{R}$ such that $\|(\bar{\boldsymbol{\eta}}, \bar{\boldsymbol{\beta}}, \bar{\boldsymbol{\gamma}}) - (\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma})\| < \delta$. Then

$$\begin{aligned} &\left| \int_0^1 \boldsymbol{\mu}(s)^T \bar{\boldsymbol{\eta}}(s)\, ds + \boldsymbol{\nu}_1^T \bar{\boldsymbol{\beta}} + \boldsymbol{\nu}_2^T \bar{\boldsymbol{\gamma}} \right| \\ &\quad = \left| \int_0^1 \boldsymbol{\mu}(s)^T \bar{\boldsymbol{\eta}}(s)\, ds + \boldsymbol{\nu}_1^T \bar{\boldsymbol{\beta}} + \boldsymbol{\nu}_2^T \bar{\boldsymbol{\gamma}} - \int_0^1 \boldsymbol{\mu}(s)^T \boldsymbol{\eta}(s)\, ds - \boldsymbol{\nu}_1^T \boldsymbol{\beta} - \boldsymbol{\nu}_2^T \boldsymbol{\gamma} \right| \\ &\quad \leqslant 2n\big(\|\boldsymbol{\mu}\|\|\bar{\boldsymbol{\eta}} - \boldsymbol{\eta}\| + \|\boldsymbol{\nu}_1\|\|\bar{\boldsymbol{\beta}} - \boldsymbol{\beta}\| + \|\boldsymbol{\nu}_2\|\|\bar{\boldsymbol{\gamma}} - \boldsymbol{\gamma}\|\big) \\ &\quad \leqslant 2n\big(\|\boldsymbol{\mu}\| + \|\boldsymbol{\nu}_1\| + \|\boldsymbol{\nu}_2\|\big) \big\|(\bar{\boldsymbol{\eta}}, \bar{\boldsymbol{\beta}}, \bar{\boldsymbol{\gamma}}) - (\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma})\big\| \\ &\quad < 2n\big(\|\boldsymbol{\mu}\| + \|\boldsymbol{\nu}_1\| + \|\boldsymbol{\nu}_2\|\big)\delta, \end{aligned} \tag{2.45}$$

for any $\delta > 0$. Therefore

$$\int_0^1 \boldsymbol{\mu}(s)^T \bar{\boldsymbol{\eta}}(s)\, ds + \boldsymbol{\nu}_1^T \bar{\boldsymbol{\beta}} + \boldsymbol{\nu}_2^T \bar{\boldsymbol{\gamma}} = 0, \tag{2.46}$$

so that the limit $(\bar{\boldsymbol{\eta}}, \bar{\boldsymbol{\beta}}, \bar{\boldsymbol{\gamma}})$ of the sequence is in $\mathcal{R}$. Thus $\mathcal{R}$ is closed.

We now establish equality (2.40), with $\boldsymbol{\rho} = (\mathbf{0}, \boldsymbol{\nu}_1, \boldsymbol{\nu}_2)$. To do so we have to verify that any $(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathcal{C}^0 \times \mathbb{R}^{2n}$ can be uniquely decomposed as

$$(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = (\boldsymbol{\eta}_{\mathcal{R}}, \boldsymbol{\beta}_{\mathcal{R}}, \boldsymbol{\gamma}_{\mathcal{R}}) + \tau(\mathbf{0}, \boldsymbol{\nu}_1, \boldsymbol{\nu}_2), \tag{2.47}$$

where $(\boldsymbol{\eta}_{\mathcal{R}}, \boldsymbol{\beta}_{\mathcal{R}}, \boldsymbol{\gamma}_{\mathcal{R}}) \in \mathcal{R}$ and $\tau \in \mathbb{R}$. In fact

$$(\boldsymbol{\eta}_{\mathcal{R}}, \boldsymbol{\beta}_{\mathcal{R}}, \boldsymbol{\gamma}_{\mathcal{R}}) = (\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) - \tau(\mathbf{0}, \boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \tag{2.48}$$

is in $\mathcal{R}$ if and only if

$$-\left(\boldsymbol{\nu}_1^T \boldsymbol{\nu}_1 + \boldsymbol{\nu}_2^T \boldsymbol{\nu}_2\right)\tau + \int_0^1 \boldsymbol{\mu}(s)^T \boldsymbol{\eta}(s)\, ds + \boldsymbol{\nu}_1^T \boldsymbol{\beta} + \boldsymbol{\nu}_2^T \boldsymbol{\gamma} = 0. \tag{2.49}$$

Since $(\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \neq \mathbf{0}$, the coefficient of τ in this equation is nonzero, and the equation admits a unique solution $\tau = \hat{\tau} \in \mathbb{R}$. Hence $(\boldsymbol{\eta}_{\mathcal{R}}, \boldsymbol{\beta}_{\mathcal{R}}, \boldsymbol{\gamma}_{\mathcal{R}})$ is given by (2.48) with $\tau = \hat{\tau}$, and (2.40) is proven.

Finally we study the solution set to BVP (2.39) in the case where it is solvable. Suppose $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are two solutions. Then $\boldsymbol{v}_1 - \boldsymbol{v}_2$ is solution to homogeneous BVP (2.34) and, by assumption (**2.A7**), $\boldsymbol{v}_1 - \boldsymbol{v}_2$ is a scalar multiple of $\boldsymbol{x}_\theta^*$. Thus the solution set to BVP (2.39) in $\mathcal{C}^1$ can be written as $\{\boldsymbol{v}_1 + t\boldsymbol{x}_\theta^* : t \in \mathbb{R}\}$. Exactly one of these solutions is in W. Indeed, the solution $\hat{\boldsymbol{v}} = \boldsymbol{v}_1 + t\boldsymbol{x}_\theta^*$ is in W if and only if

$$t \int_0^1 \boldsymbol{x}_\theta^*(s)^T \boldsymbol{x}_\theta^*(s)\, ds + \int_0^1 \boldsymbol{v}_1(s)^T \boldsymbol{x}_\theta^*(s)\, ds = 0, \tag{2.50}$$

and this equation has a unique solution $t = \hat{t} \in \mathbb{R}$ since the integral multiplying t is nonzero. The solution set to BVP (2.34) in $\mathcal{C}^1$ can then be rewritten as $\{\hat{\boldsymbol{v}} + \tau\boldsymbol{x}_\theta^* : \tau \in \mathbb{R}\}$. ■

The following proposition shows that, in a neighborhood of $\boldsymbol{x}^*$, Γ constitutes an *isolated* curve of solutions to the unperturbed problem.

Proposition 2.8 *Let conditions (**2.A4**), (**2.A6**) and (**2.A7**) be satisfied. Then there exists a neighborhood $\mathcal{X}$ of $\boldsymbol{x}^*$ in $\mathcal{C}^1$ such that, for $\epsilon = 0$, the set of solutions to BVP (2.3) in $\mathcal{X}$ is $\Gamma \cap \mathcal{X}$.*

Proof Define the operator version $h : \mathcal{C}^1 \times \mathbb{R} \to \mathcal{C}^0$ of the Hamiltonian by

$$h(\boldsymbol{x}, \epsilon)(s) = H(\boldsymbol{x}(s), s, \epsilon), \quad \forall\, s \in [0,1], \tag{2.51}$$

and the operator $\Xi \in C^k(I^* \times W^*, \mathcal{C}^0 \times \mathbb{R}^{2n})$ by

$$\Xi(\theta, \boldsymbol{w}) = \begin{pmatrix} \boldsymbol{\zeta}(\theta)' + \boldsymbol{w}' - \boldsymbol{J} h_{\boldsymbol{x}}(\boldsymbol{\zeta}(\theta) + \boldsymbol{w}, 0) \\ \boldsymbol{b}(\boldsymbol{\zeta}(\theta)(0) + \boldsymbol{w}(0)) \\ \boldsymbol{c}(\boldsymbol{\zeta}(\theta)(1) + \boldsymbol{w}(1)) \end{pmatrix}. \tag{2.52}$$

Then, according to lemma 2.6, solving BVP (2.3) with $\epsilon = 0$ in the neighborhood X^* of $\boldsymbol{x}^*$ is equivalent to solving

$$\Xi(\theta, \boldsymbol{w}) = \mathbf{0} \tag{2.53}$$

in the neighborhood $I^* \times W^*$ of $(\theta^*, \mathbf{0})$. By hypothesis (**2.A6**),

$$\Xi(\theta, \mathbf{0}) = \mathbf{0}, \quad \forall\, \theta \in I^*. \tag{2.54}$$

The key of this proof is to apply a sort of Lyapunov-Schmidt reduction (cf. Ambrosetti and Prodi [5]) to show that, in some neighborhood of $(\theta^*, \mathbf{0})$, all solutions to (2.53) are of the form $(\theta, \mathbf{0})$, with $\theta \in I^*$.

The derivative $\Xi_{\boldsymbol{w}}(\theta^*, \mathbf{0}) : W \to \mathcal{C}^0 \times \mathbb{R}^{2n}$ is expressed by

$$\Xi_{\boldsymbol{w}}(\theta^*, \mathbf{0})\boldsymbol{v} = \begin{pmatrix} \boldsymbol{v}' - \boldsymbol{J} h_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*, 0)\boldsymbol{v} \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))\boldsymbol{v}(0) \\ \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\boldsymbol{v}(1) \end{pmatrix}. \tag{2.55}$$

According to lemma 2.7,

$$\operatorname{Range} \Xi_{\boldsymbol{w}}(\theta^*, \mathbf{0}) = \mathcal{R}. \tag{2.56}$$

Let $\boldsymbol{\Pi} : \mathcal{C}^0 \times \mathbb{R}^{2n} \to \mathcal{R}$ be the projection operator associated to the direct sum (2.40). By the proof of lemma 2.7 (cf. equations (2.48) and (2.49)), this projection operator is expressed by

$$\boldsymbol{\Pi}(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = (\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) - \hat{\tau}(\mathbf{0}, \boldsymbol{\nu}_1, \boldsymbol{\nu}_2), \tag{2.57}$$

where

$$\hat{\tau} = \frac{\int_0^1 \boldsymbol{\mu}(s)^T \boldsymbol{\eta}(s)\, ds + \boldsymbol{\nu}_1^T \boldsymbol{\beta} + \boldsymbol{\nu}_2^T \boldsymbol{\gamma}}{\boldsymbol{\nu}_1^T \boldsymbol{\nu}_1 + \boldsymbol{\nu}_2^T \boldsymbol{\nu}_2}, \tag{2.58}$$

and $(\boldsymbol{\mu}, \boldsymbol{\nu}_1, \boldsymbol{\nu}_2)$ is some fixed element of $\mathcal{C}^0 \times \mathbb{R}^{2n}$. It follows that $\boldsymbol{\Pi} \in C^1(\mathcal{C}^0 \times \mathbb{R}^{2n})$.

Define the mapping $\boldsymbol{\Lambda} \in C^1(I^* \times W^*, \mathcal{R})$ by

$$\boldsymbol{\Lambda}(\theta, \boldsymbol{w}) = \boldsymbol{\Pi}\, \Xi(\theta, \boldsymbol{w}). \tag{2.59}$$

By (2.54), we have

$$\boldsymbol{\Lambda}(\theta, \mathbf{0}) = \mathbf{0}, \quad \forall\, \theta \in I^*. \tag{2.60}$$

Define the set

$$\mathcal{S} = \{\boldsymbol{v} \in W : \|\boldsymbol{v}\| = 1\}. \tag{2.61}$$

By definition of the Fréchet derivative, for any $\delta > 0$ and $\boldsymbol{v} \in \mathcal{S}$,

$$\Xi(\theta^*, \delta \boldsymbol{v}) = \Xi(\theta^*, \mathbf{0}) + \delta\, \Xi_{\boldsymbol{w}}(\theta^*, \mathbf{0})\boldsymbol{v} + \delta \boldsymbol{\rho}(\delta, \boldsymbol{v}), \tag{2.62}$$

where the quantity $\boldsymbol{\rho}(\delta, \boldsymbol{v})$ is such that

$$\lim_{\delta \to 0} \|\boldsymbol{\rho}(\delta, \boldsymbol{v})\| = 0 \quad \text{uniformly for any } \boldsymbol{v} \in \mathcal{S}. \tag{2.63}$$

Then, for any $\delta > 0$ and $\boldsymbol{v} \in \mathcal{S}$,

$$\begin{aligned} \boldsymbol{\Lambda}(\theta^*, \delta \boldsymbol{v}) &= \boldsymbol{\Pi}\, \Xi(\theta^*, \mathbf{0}) + \delta\, \boldsymbol{\Pi}\, \Xi_{\boldsymbol{w}}(\theta^*, \mathbf{0})\boldsymbol{v} + \delta\, \boldsymbol{\Pi}\boldsymbol{\rho}(\delta, \boldsymbol{v}) \\ &= \boldsymbol{\Lambda}(\theta^*, \mathbf{0}) + \delta\, \Xi_{\boldsymbol{w}}(\theta^*, \mathbf{0})\boldsymbol{v} + \delta\, \boldsymbol{\Pi}\boldsymbol{\rho}(\delta, \boldsymbol{v}). \end{aligned} \tag{2.64}$$

We used the linearity of $\boldsymbol{\Pi}$ to obtain the first equality and the definition of $\boldsymbol{\Pi}$ together with equation (2.56) to derive the second. Moreover

$$\|\boldsymbol{\Pi}\boldsymbol{\rho}(\delta, \boldsymbol{v})\| \leqslant \|\boldsymbol{\rho}(\delta, \boldsymbol{v})\|, \quad \forall \delta > 0,\ \boldsymbol{v} \in \mathcal{S}, \tag{2.65}$$

so that

$$\lim_{\delta \to 0} \|\boldsymbol{\Pi}\boldsymbol{\rho}(\delta, \boldsymbol{v})\| = 0 \quad \text{uniformly for any } \boldsymbol{v} \in \mathcal{S}. \tag{2.66}$$

We conclude from equations (2.64) and (2.66) that the derivative $\boldsymbol{\Lambda}_{\boldsymbol{w}}(\theta^*, \mathbf{0}) : W \to \mathcal{R}$ is expressed by

$$\boldsymbol{\Lambda}_{\boldsymbol{w}}(\theta^*, \mathbf{0})\boldsymbol{v} = \Xi_{\boldsymbol{w}}(\theta^*, \mathbf{0})\boldsymbol{v}. \tag{2.67}$$

Comparing this equation with equality (2.56) shows that the derivative $\boldsymbol{\Lambda}_{\boldsymbol{w}}(\theta^*, \mathbf{0})$ is surjective. By assumption (**2.A7**), the equation $\Xi_{\boldsymbol{w}}(\theta^*, \mathbf{0})\boldsymbol{v} = \mathbf{0}$ admits only the trivial solution in W, so that the derivative $\boldsymbol{\Lambda}_{\boldsymbol{w}}(\theta^*, \mathbf{0})$ is also injective.

Having verified its hypotheses, we can apply the implicit function theorem (cf. appendix A) to conclude that there exists a neighborhood $I_0 \times W_0$ of $(\theta^*, \mathbf{0})$ in $I^* \times W^*$ such that, for any $\theta \in I_0$, equation

$$\boldsymbol{\Lambda}(\theta, \boldsymbol{w}) = \mathbf{0} \tag{2.68}$$

admits a unique solution in W_0. In fact, by (2.60), the set of solutions to equation (2.68) in $I_0 \times W_0$ is $\{(\theta, \mathbf{0}) : \theta \in I_0\}$. Now, by definition of $\boldsymbol{\Lambda}$, the set of solutions to equation (2.53) is included in the set of solutions to equation (2.68). We can conclude, with the help of (2.54), that the set of solutions to equation (2.53) in $I_0 \times W_0$ is $\{(\theta, \mathbf{0}) : \theta \in I_0\}$. Finally, after defining the set $\mathcal{X} = \boldsymbol{\psi}(I_0 \times W_0)$, the conclusion of the proposition follows from lemma 2.6. ■

Assumption (**2.A5**), which was involved neither in lemmas 2.6, 2.7, nor in proposition 2.8, comes back into play now. According to lemma 2.3, it is possible to define the linear mappings

$$\boldsymbol{B} = \left(\boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))\,\big|_{(\mathrm{Ker}\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0)))^\perp}\right)^{-1} : \mathbb{R}^n \to \left(\mathrm{Ker}\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))\right)^\perp \tag{2.69}$$

and

$$\boldsymbol{C} = \left(\boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\,\big|_{(\mathrm{Ker}\ \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1)))^\perp}\right)^{-1} : \mathbb{R}^n \to \left(\mathrm{Ker}\ \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\right)^\perp. \tag{2.70}$$

With this we are in a position to give an explicit expression for the solvability condition for BVP (2.39).

Lemma 2.9 *Let conditions (**2.A4**)-(**2.A7**) be satisfied. Then BVP (2.39) is solvable if and only if*

$$\int_0^1 \boldsymbol{x}_\theta^*(s)^T \boldsymbol{J}\boldsymbol{\eta}(s)\ ds + \boldsymbol{x}_\theta^*(0)^T \boldsymbol{J}\boldsymbol{B}\boldsymbol{\beta} - \boldsymbol{x}_\theta^*(1)^T \boldsymbol{J}\boldsymbol{C}\boldsymbol{\gamma} = 0. \tag{2.71}$$

Thus the set $\mathcal{R}$ defined in lemma 2.7 can be expressed as

$$\mathcal{R} = \left\{(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathcal{C}^0 \times \mathbb{R}^n \times \mathbb{R}^n : (\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \text{ satisfies (2.71)}\right\}. \tag{2.72}$$

Proof We start by showing that (2.71) is a necessary condition for BVP (2.39) to be solvable. Suppose BVP (2.39) is solvable and $\boldsymbol{v}$ is a solution. Left-multiplying the differential equation in (2.39) by $\boldsymbol{x}_\theta^*(s)^T \boldsymbol{J}$ and integrating, we obtain

$$\begin{aligned}
\int_0^1 \boldsymbol{x}_\theta^*(s)^T \boldsymbol{J}\boldsymbol{\eta}(s)\, ds &= \int_0^1 \left(\boldsymbol{x}_\theta^*(s)^T \boldsymbol{J}\boldsymbol{v}'(s) + \boldsymbol{x}_\theta^*(s)^T H_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{v}(s)\right) ds \\
&= \int_0^1 \left(\boldsymbol{v}(s)^T \boldsymbol{J}\left(\boldsymbol{x}_\theta^{*\prime}(s) - \boldsymbol{J} H_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{x}_\theta^*(s)\right) ds \right.\\
&\qquad + \boldsymbol{x}_\theta^*(1)^T \boldsymbol{J}\boldsymbol{v}(1) - \boldsymbol{x}_\theta^*(0)^T \boldsymbol{J}\boldsymbol{v}(0) \\
&= \boldsymbol{x}_\theta^*(1)^T \boldsymbol{J}\boldsymbol{v}(1) - \boldsymbol{x}_\theta^*(0)^T \boldsymbol{J}\boldsymbol{v}(0),
\end{aligned} \tag{2.73}$$

where the second equality is the consequence of an integration by parts and the last equality follows from the fact that $\boldsymbol{x}_\theta^*$ is a solution to BVP (2.34). To compute the boundary term $\boldsymbol{x}_\theta^*(0)^T\boldsymbol{J}\boldsymbol{v}(0)$, introduce the decomposition $\boldsymbol{v}(0) = \boldsymbol{\nu} + \boldsymbol{\kappa}$, with $\boldsymbol{\nu} \in \operatorname{Ker} \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))$ and $\boldsymbol{\kappa} \in (\operatorname{Ker} \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0)))^\perp$. By the supposition that $\boldsymbol{v}$ is a solution to BVP (2.39),

$$\boldsymbol{\beta} = \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))\boldsymbol{v}(0) = \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))\boldsymbol{\kappa}. \tag{2.74}$$

Then $\boldsymbol{\kappa} = \boldsymbol{B}\boldsymbol{\beta}$ by definition of $\boldsymbol{B}$. Thus

$$\boldsymbol{x}_\theta^*(0)^T\boldsymbol{J}\boldsymbol{v}(0) = \boldsymbol{x}_\theta^*(0)^T\boldsymbol{J}\boldsymbol{\nu} + \boldsymbol{x}_\theta^*(0)^T\boldsymbol{J}\boldsymbol{B}\boldsymbol{\beta} = \boldsymbol{x}_\theta^*(0)^T\boldsymbol{J}\boldsymbol{B}\boldsymbol{\beta}, \tag{2.75}$$

where the second equality follows from the fact that $\boldsymbol{x}_\theta^*(0)$, $\boldsymbol{\nu} \in \operatorname{Ker} \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))$ together with assumption (**2.A5**). In the same way, we find

$$\boldsymbol{x}_\theta^*(1)^T\boldsymbol{J}\boldsymbol{v}(1) = \boldsymbol{x}_\theta^*(1)^T\boldsymbol{J}\boldsymbol{C}\boldsymbol{\gamma}. \tag{2.76}$$

Collecting results (2.73), (2.75) and (2.76) shows that (2.71) is a necessary condition for BVP (2.39) to be solvable. Moreover, defining the set

$$\widetilde{\mathcal{R}} = \left\{(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathcal{C}^0 \times \mathbb{R}^n \times \mathbb{R}^n : (\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \text{ satisfies (2.71)}\right\}, \tag{2.77}$$

we conclude that

$$\mathcal{R} \subset \widetilde{\mathcal{R}}, \tag{2.78}$$

where $\mathcal{R}$ is defined in lemma 2.7.

The rest of this proof is devoted to establishing the equality $\mathcal{R} = \widetilde{\mathcal{R}}$, from which the conclusion of the lemma immediately follows. We start by showing that

$$\mathcal{C}^0 \times \mathbb{R}^{2n} = \widetilde{\mathcal{R}} \oplus \operatorname{Span}\{\tilde{\boldsymbol{\rho}}\}, \tag{2.79}$$

with $\tilde{\boldsymbol{\rho}} = (\boldsymbol{J}\boldsymbol{x}_\theta^*, \boldsymbol{0}, \boldsymbol{0}) \neq \boldsymbol{0}$. For that we have to verify that any $(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathcal{C}^0 \times \mathbb{R}^{2n}$ can be uniquely decomposed as

$$(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = (\boldsymbol{\eta}_{\widetilde{\mathcal{R}}}, \boldsymbol{\beta}_{\widetilde{\mathcal{R}}}, \boldsymbol{\gamma}_{\widetilde{\mathcal{R}}}) + \tau(\boldsymbol{J}\boldsymbol{x}_\theta^*, \boldsymbol{0}, \boldsymbol{0}), \tag{2.80}$$

where $(\boldsymbol{\eta}_{\widetilde{\mathcal{R}}}, \boldsymbol{\beta}_{\widetilde{\mathcal{R}}}, \boldsymbol{\gamma}_{\widetilde{\mathcal{R}}}) \in \widetilde{\mathcal{R}}$ and $\tau \in \mathbb{R}$. In fact

$$(\boldsymbol{\eta}_{\widetilde{\mathcal{R}}}, \boldsymbol{\beta}_{\widetilde{\mathcal{R}}}, \boldsymbol{\gamma}_{\widetilde{\mathcal{R}}}) = (\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) - \tau(\boldsymbol{J}\boldsymbol{x}_\theta^*, \boldsymbol{0}, \boldsymbol{0}) \tag{2.81}$$

is in $\widetilde{\mathcal{R}}$ if and only if

$$\tau \int_0^1 \boldsymbol{x}_\theta^*(s)^T \boldsymbol{x}_\theta^*(s)\, ds + \int_0^1 \boldsymbol{x}_\theta^*(s)^T \boldsymbol{J}\boldsymbol{\eta}(s)\, ds + \boldsymbol{x}_\theta^*(0)^T\boldsymbol{J}\boldsymbol{B}\boldsymbol{\beta} - \boldsymbol{x}_\theta^*(1)^T\boldsymbol{J}\boldsymbol{C}\boldsymbol{\gamma} = 0. \tag{2.82}$$

Since $\boldsymbol{x}_\theta^* \neq \boldsymbol{0}$, the integral multiplying τ is nonzero so that this equation admits a unique solution $\tau = \hat{\tau} \in \mathbb{R}$. Hence $(\boldsymbol{\eta}_{\widetilde{\mathcal{R}}}, \boldsymbol{\beta}_{\widetilde{\mathcal{R}}}, \boldsymbol{\gamma}_{\widetilde{\mathcal{R}}})$ is given by (2.81), with $\tau = \hat{\tau}$. This establishes (2.79).

We now suppose $\mathcal{R} \neq \widetilde{\mathcal{R}}$. Then, by (2.78), there exists $\boldsymbol{\kappa} \in \widetilde{\mathcal{R}} \backslash \mathcal{R}$. Hence, by (2.40), $\boldsymbol{\kappa} = \boldsymbol{\kappa}_{\mathcal{R}} + \tau\boldsymbol{\rho}$, with $\boldsymbol{\kappa}_{\mathcal{R}} \in \mathcal{R}$ and $\tau \neq 0$. Then, by (2.78), $\boldsymbol{\kappa}_{\mathcal{R}} \in \widetilde{\mathcal{R}}$, so that

$$\boldsymbol{\rho} = \frac{1}{\tau}(\boldsymbol{\kappa} - \boldsymbol{\kappa}_{\mathcal{R}}) \in \widetilde{\mathcal{R}}. \tag{2.83}$$

Finally, by (2.40), (2.78) and (2.83),

$$\mathcal{C}^0 \times \mathbb{R}^{2n} = \mathcal{R} \oplus \mathrm{Span}\,\{\boldsymbol{\rho}\} \subset \widetilde{\mathcal{R}}, \tag{2.84}$$

which contradicts (2.79). We conclude $\mathcal{R} = \widetilde{\mathcal{R}}$. ■

The following theorem shows that the solution $\boldsymbol{x}^*$ persists in the perturbed problem if θ^* corresponds to a simple zero of the function $\mathcal{M} \in C^1(I, \mathbb{R})$ defined by

$$\begin{aligned} \mathcal{M}(\theta) &= \int_0^1 \boldsymbol{\zeta}_\theta(\theta)(s)^T H_{\boldsymbol{x}\epsilon}(\boldsymbol{\zeta}(\theta)(s), s, 0)\, ds \\ &= \frac{d}{d\theta} \int_0^1 H_\epsilon(\boldsymbol{\zeta}(\theta)(s), s, 0)\, ds. \end{aligned} \tag{2.85}$$

Note that the simple zeroes of $\mathcal{M}(\theta)$ coincide with the nondegenerate critical points of the function

$$\mathcal{P}(\theta) = \int_0^1 H_\epsilon(\boldsymbol{\zeta}(\theta)(s), s, 0)\, ds. \tag{2.86}$$

$\mathcal{M}$ is called a *Melnikov function*, and $\mathcal{P}$ a *Poincaré function* (cf. Stuart [51] and Ambrosetti and Badiale [3]).

Theorem 2.10 *Let conditions (**2.A4**)-(**2.A7**) be satisfied, and consider the Melnikov function $\mathcal{M}(\theta)$ defined by (2.85).*

- *(a) If $\mathcal{M}(\theta^*) \neq 0$, then there exist a positive number ϵ_0 and a $\mathcal{C}^1$-neighborhood $\mathcal{X}$ of $\boldsymbol{x}^*$ such that, for ϵ satisfying $0 \neq |\epsilon| < \epsilon_0$, BVP (2.3) does not admit any solution in $\mathcal{X}$.*

- *(b) If $\mathcal{M}(\theta^*) = 0$ and $\mathcal{M}_\theta(\theta^*) \neq 0$, then there exist a positive number ϵ_0, a $\mathcal{C}^1$-neighborhood $\mathcal{X}$ of $\boldsymbol{x}^*$ and a function $\boldsymbol{\xi} \in C^{k-1}((-\epsilon_0, \epsilon_0), \mathcal{X})$, such that $\boldsymbol{\xi}(0) = \boldsymbol{x}^*$ and, for ϵ satisfying $0 \neq |\epsilon| < \epsilon_0$, $\boldsymbol{x} = \boldsymbol{\xi}(\epsilon)$ is the unique solution to BVP (2.3) in $\mathcal{X}$.*

Proof Define the operator $\boldsymbol{\Phi} \in C^k(I^* \times \mathbb{R} \times W^*, \mathcal{C}^0 \times \mathbb{R}^{2n})$ by

$$\boldsymbol{\Phi}(\theta, \epsilon, \boldsymbol{w}) = \begin{pmatrix} \boldsymbol{\zeta}(\theta)' + \boldsymbol{w}' - \boldsymbol{J} h_{\boldsymbol{x}}(\boldsymbol{\zeta}(\theta) + \boldsymbol{w}, \epsilon) \\ \boldsymbol{b}(\boldsymbol{\zeta}(\theta)(0) + \boldsymbol{w}(0)) \\ \boldsymbol{c}(\boldsymbol{\zeta}(\theta)(1) + \boldsymbol{w}(1)) \end{pmatrix}, \tag{2.87}$$

where I^* and W^* are introduced in lemma 2.6, and the operator version of the Hamiltonian h is defined in (2.51). According to lemma 2.6, solving BVP (2.3) in the neighborhood X^* of $\boldsymbol{x}^*$ is equivalent to solving

$$\boldsymbol{\Phi}(\theta, \epsilon, \boldsymbol{w}) = \boldsymbol{0}. \tag{2.88}$$

By hypothesis (**2.A6**),

$$\boldsymbol{\Phi}(\theta, 0, \boldsymbol{0}) = \boldsymbol{0}, \quad \forall\, \theta \in I^*, \tag{2.89}$$

i.e. $I^* \times \{0\} \times \{\mathbf{0}\}$ is a trivial branch of solutions. We will show that $(\theta^*, 0, \mathbf{0})$ does not constitute a bifurcation point in case (a) of the theorem, but it does in case (b).

For $(e, \boldsymbol{v}) \in \mathbb{R} \times W$, we compute

$$\boldsymbol{\Phi}_{(\epsilon,\boldsymbol{w})}(\theta, \epsilon, \boldsymbol{w})(e, \boldsymbol{v}) = \begin{pmatrix} \boldsymbol{v}' - \boldsymbol{J} h_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{\zeta}(\theta) + \boldsymbol{w}, \epsilon)\boldsymbol{v} - e\boldsymbol{J} h_{\boldsymbol{x}\epsilon}(\boldsymbol{\zeta}(\theta) + \boldsymbol{w}, \epsilon) \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\zeta}(\theta)(0) + \boldsymbol{w}(0))\boldsymbol{v}(0) \\ \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{\zeta}(\theta)(1) + \boldsymbol{w}(1))\boldsymbol{v}(1) \end{pmatrix}. \tag{2.90}$$

The operator $\boldsymbol{A} : \mathbb{R} \times W \to \mathcal{C}^0 \times \mathbb{R}^{2n}$ defined by

$$\boldsymbol{A} = \boldsymbol{\Phi}_{(\epsilon,\boldsymbol{w})}(\theta^*, 0, \mathbf{0}) \tag{2.91}$$

is expressed by

$$\boldsymbol{A}(e, \boldsymbol{v}) = \begin{pmatrix} \boldsymbol{v}' - \boldsymbol{J} h_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*, 0)\boldsymbol{v} - e\boldsymbol{J} h_{\boldsymbol{x}\epsilon}(\boldsymbol{x}^*, 0) \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))\boldsymbol{v}(0) \\ \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\boldsymbol{v}(1) \end{pmatrix}. \tag{2.92}$$

By assumption (**2.A7**), and since $\boldsymbol{x}^*_\theta \notin W$, a nonzero nullvector of $\boldsymbol{A}$ cannot be of the form $(0, \boldsymbol{v})$. Thus, by linearity of $\boldsymbol{A}$,

$$\text{Ker } \boldsymbol{A} = \text{Span}\big\{(1, \boldsymbol{v}) \in \text{Ker } \boldsymbol{A}\big\}. \tag{2.93}$$

The key to the proof of part (a) of the theorem is an application of the implicit function theorem. Under the hypotheses of part (a), and according to lemma 2.9, the equation $\boldsymbol{A}(1, \boldsymbol{v}) = \mathbf{0}$ does not admit any solution since the corresponding solvability condition $\mathcal{M}(\theta^*) = 0$ is not satisfied. Thus Ker $\boldsymbol{A} = \{\mathbf{0}\}$, and $\boldsymbol{A}$ is injective. We show now that $\boldsymbol{A}$ is also surjective, i.e. for any $(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathcal{C}^0 \times \mathbb{R}^{2n}$, the problem

$$\begin{cases} \boldsymbol{v}'(s) - \boldsymbol{J} H_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{v}(s) - e\boldsymbol{J} H_{\boldsymbol{x}\epsilon}(\boldsymbol{x}^*(s), s, 0) = \boldsymbol{\eta}(s), \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))\boldsymbol{v}(0) = \boldsymbol{\beta}, \quad \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\boldsymbol{v}(1) = \boldsymbol{\gamma}, \end{cases} \tag{2.94}$$

admits a solution in $\mathbb{R} \times W$. The associated solvability condition is

$$-e\,\mathcal{M}(\theta^*) + \int_0^1 \boldsymbol{x}^*_\theta(s)^T \boldsymbol{J}\boldsymbol{\eta}(s)\, ds + \boldsymbol{x}^*_\theta(0)^T \boldsymbol{J}\boldsymbol{B}\boldsymbol{\beta} - \boldsymbol{x}^*_\theta(1)^T \boldsymbol{J}\boldsymbol{C}\boldsymbol{\gamma} = 0. \tag{2.95}$$

Since $\mathcal{M}(\theta^*) \neq 0$, this equation admits a solution $e = \hat{e} \in \mathbb{R}$. Then, by lemma 2.7, BVP (2.94) with $e = \hat{e}$ is solvable in W, and surjectivity is proven.

We can now apply the implicit function theorem (cf. appendix A) to conclude that there exists a positive number ϵ_0, a neighborhood W_0 of $\mathbf{0}$ in W^* and a neighborhood I_0 of θ^* in I^* such that, for any $\theta \in I_0$, equation (2.88) admits a unique solution in $(-\epsilon_0, \epsilon_0) \times W_0$. In fact this solution is given by equation (2.89). Setting $\mathcal{X} = \boldsymbol{\psi}(I_0 \times W_0)$, where $\boldsymbol{\psi}$ is defined in lemma 2.6, the conclusion of part (a) of the theorem directly follows.

The proof of part (b) of the theorem is similar to the proof of theorem 3.2 in Stuart [51]; it is an application of the Crandall-Rabinowitz theorem (cf. appendix A). Before applying that theorem, we must check the following hypotheses:

(i) Ker $\boldsymbol{A} = \text{Span}\,\{(1, \boldsymbol{v}_1^*)\}$, with $\boldsymbol{v}_1^* \in W$,

(ii) Range $\boldsymbol{A}$ is closed and codim Range $\boldsymbol{A} = 1$,

(iii) $\boldsymbol{\Phi}_{(\epsilon,\boldsymbol{w})\theta}(\theta^*, 0, \boldsymbol{0})(1, \boldsymbol{v}_1^*) \notin \text{Range}\,\boldsymbol{A}$.

For (i) we note that, according to lemma 2.7, the equation $\boldsymbol{A}(1, \boldsymbol{v}) = \boldsymbol{0}$ admits a unique solution $\boldsymbol{v}_1^* \in W$ since the corresponding solvability condition $\mathcal{M}(\theta^*) = 0$ is satisfied. Thus, by (2.93), property (i) is satisfied.

For (ii) we note that $(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathcal{C}^0 \times \mathbb{R}^{2n}$ is in Range $\boldsymbol{A}$ if and only if there exists $(e, \boldsymbol{v}) \in \mathbb{R} \times W$ satisfying BVP (2.94), i.e. if and only if there exists $e \in \mathbb{R}$ such that solvability condition (2.95) is satisfied. Since $\mathcal{M}(\theta^*) = 0$, it follows from equation (2.72) that

$$\text{Range}\,\boldsymbol{A} = \mathcal{R}. \tag{2.96}$$

Then, by lemma 2.7, property (ii) is satisfied.

The proof of (iii) requires more effort. In particular, it involves the use of properties of Lagrangian submanifolds. We start by computing

$$\boldsymbol{\Phi}_{(\epsilon,\boldsymbol{w})\theta}(\theta^*, 0, \boldsymbol{0})(1, \boldsymbol{v}_1^*) = \begin{pmatrix} -\boldsymbol{J} h_{\boldsymbol{xxx}}(\boldsymbol{x}^*, 0)[\boldsymbol{v}_1^*, \boldsymbol{x}_\theta^*] - \boldsymbol{J} h_{\boldsymbol{xx}\epsilon}(\boldsymbol{x}^*, 0)\boldsymbol{x}_\theta^* \\ \boldsymbol{b}_{\boldsymbol{xx}}(\boldsymbol{x}^*(0))[\boldsymbol{v}_1^*(0), \boldsymbol{x}_\theta^*(0)] \\ \boldsymbol{c}_{\boldsymbol{xx}}(\boldsymbol{x}^*(1))[\boldsymbol{v}_1^*(1), \boldsymbol{x}_\theta^*(1)] \end{pmatrix}. \tag{2.97}$$

By (2.96), property (iii) is equivalent to $F_1 \neq 0$, where the quantity F_1 is defined as

$$\begin{aligned} F_1 = &\int_0^1 \boldsymbol{x}_\theta^*(s)^T \Big(H_{\boldsymbol{xxx}}(\boldsymbol{x}^*(s), s, 0)[\boldsymbol{v}_1^*(s), \boldsymbol{x}_\theta^*(s)] + H_{\boldsymbol{xx}\epsilon}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{x}_\theta^*(s) \Big)\, ds \\ &+ \boldsymbol{x}_\theta^*(0)^T \boldsymbol{J}\boldsymbol{B}\boldsymbol{b}_{\boldsymbol{xx}}(\boldsymbol{x}^*(0))[\boldsymbol{v}_1^*(0), \boldsymbol{x}_\theta^*(0)] \\ &- \boldsymbol{x}_\theta^*(1)^T \boldsymbol{J}\boldsymbol{C}\boldsymbol{c}_{\boldsymbol{xx}}(\boldsymbol{x}^*(1))[\boldsymbol{v}_1^*(1), \boldsymbol{x}_\theta^*(1)], \end{aligned} \tag{2.98}$$

We show next that

$$F_1 = \mathcal{M}_\theta(\theta^*), \tag{2.99}$$

so that property (iii) immediately follows from hypothesis $\mathcal{M}_\theta(\theta^*) \neq 0$.

The quantity $\mathcal{M}_\theta(\theta^*)$ is expressed as

$$\mathcal{M}_\theta(\theta^*) = \int_0^1 \boldsymbol{x}_{\theta\theta}^*(s)^T H_{\boldsymbol{x}\epsilon}(\boldsymbol{x}^*(s), s, 0)\, ds + \int_0^1 \boldsymbol{x}_\theta^*(s)^T H_{\boldsymbol{xx}\epsilon}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{x}_\theta^*(s)\, ds. \tag{2.100}$$

We divide the proof of (2.99) into two steps. We start the first step by differentiating BVP (2.33) with respect to θ twice and setting $\theta = \theta^*$. We find

$$\begin{cases} \boldsymbol{x}_{\theta\theta}^{*\prime}(s) - \boldsymbol{J} H_{\boldsymbol{xx}}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{x}_{\theta\theta}^*(s) = \boldsymbol{J} H_{\boldsymbol{xxx}}(\boldsymbol{x}^*(s), s, 0)[\boldsymbol{x}_\theta^*(s), \boldsymbol{x}_\theta^*(s)], \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))\boldsymbol{x}_{\theta\theta}^*(0) = -\boldsymbol{b}_{\boldsymbol{xx}}(\boldsymbol{x}^*(0))[\boldsymbol{x}_\theta^*(0), \boldsymbol{x}_\theta^*(0)], \\ \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\boldsymbol{x}_{\theta\theta}^*(1) = -\boldsymbol{c}_{\boldsymbol{xx}}(\boldsymbol{x}^*(1))[\boldsymbol{x}_\theta^*(1), \boldsymbol{x}_\theta^*(1)]. \end{cases} \tag{2.101}$$

It follows from the fact that $\boldsymbol{v}_1^*(0) \in \text{Ker}\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))$, assumption (**2.A5**), the definition of $\boldsymbol{B}$ and the second line in (2.101) that

$$\boldsymbol{v}_1^*(0)^T \boldsymbol{J}\boldsymbol{x}_{\theta\theta}^*(0) = -\boldsymbol{v}_1^*(0)^T \boldsymbol{J}\boldsymbol{B}\boldsymbol{b}_{\boldsymbol{xx}}(\boldsymbol{x}^*(0))[\boldsymbol{x}_\theta^*(0), \boldsymbol{x}_\theta^*(0)]. \tag{2.102}$$

In the same way,

$$\boldsymbol{v}_1^*(1)^T \boldsymbol{J} \boldsymbol{x}_{\theta\theta}^*(1) = -\boldsymbol{v}_1^*(1)^T \boldsymbol{J} \boldsymbol{C} \boldsymbol{c}_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(1))[\boldsymbol{x}_\theta^*(1), \boldsymbol{x}_\theta^*(1)]. \tag{2.103}$$

We now compute

$$\begin{aligned}
&\int_0^1 \boldsymbol{x}_\theta^*(s)^T H_{\boldsymbol{x}\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(s), s, 0)[\boldsymbol{v}_1^*(s), \boldsymbol{x}_\theta^*(s)]\, ds \\
&\quad = \int_0^1 \boldsymbol{v}_1^*(s)^T H_{\boldsymbol{x}\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(s), s, 0)[\boldsymbol{x}_\theta^*(s), \boldsymbol{x}_\theta^*(s)]\, ds \\
&\quad = -\int_0^1 \boldsymbol{v}_1^*(s)^T \boldsymbol{J} \boldsymbol{x}_{\theta\theta}^{*\prime}(s)\, ds - \int_0^1 \boldsymbol{v}_1^*(s)^T H_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(s), s, 0) \boldsymbol{x}_{\theta\theta}^*(s)\, ds \\
&\quad = \boldsymbol{v}_1^*(0)^T \boldsymbol{J} \boldsymbol{x}_{\theta\theta}^*(0) - \boldsymbol{v}_1^*(1)^T \boldsymbol{J} \boldsymbol{x}_{\theta\theta}^*(1) \\
&\qquad - \int_0^1 \boldsymbol{x}_{\theta\theta}^*(s)^T \boldsymbol{J} \Big(\boldsymbol{v}_1^{*\prime}(s) - \boldsymbol{J} H_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(s), s, 0) \boldsymbol{v}_1^*(s) \Big)\, ds \\
&\quad = -\boldsymbol{v}_1^*(0)^T \boldsymbol{J} \boldsymbol{B} \boldsymbol{b}_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(0))[\boldsymbol{x}_\theta^*(0), \boldsymbol{x}_\theta^*(0)] + \boldsymbol{v}_1^*(1)^T \boldsymbol{J} \boldsymbol{C} \boldsymbol{c}_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(1))[\boldsymbol{x}_\theta^*(1), \boldsymbol{x}_\theta^*(1)] \\
&\qquad + \int_0^1 \boldsymbol{x}_{\theta\theta}^*(s)^T H_{\boldsymbol{x}\epsilon}(\boldsymbol{x}^*(s), s, 0)\, ds,
\end{aligned} \tag{2.104}$$

where the second equality comes from (2.101), the third equality is a consequence of an integration by parts, and the last equality follows from (2.102), (2.103) and the definition of $\boldsymbol{v}_1^*$.

We continue with the second step of the proof of equality (2.99). By lemma 2.5, L_0 is a submanifold of $\mathbb{R}^{2n}$ with dimension n. Therefore, by definition A.7, there exist an open neighborhood U of $\boldsymbol{x}^*(0)$ in $\mathbb{R}^{2n}$, an open set $V \times Z$ in $\mathbb{R}^n \times \mathbb{R}^n$ and a C^k-diffeomorphism $\boldsymbol{\chi} : U \to V \times Z$ such that $\boldsymbol{\chi}(L_0 \cap U) = V \times \{\boldsymbol{0}\}$. Let $(\boldsymbol{y}^*, \boldsymbol{0}) = \boldsymbol{\chi}(\boldsymbol{x}^*(0))$, and define $\boldsymbol{\varphi} : V \to L_0 \cap U$ by $\boldsymbol{\varphi}(\boldsymbol{y}) = \boldsymbol{\chi}^{-1}(\boldsymbol{y}, \boldsymbol{0})$. Then $\boldsymbol{\varphi} \in C^k(V)$ and $\boldsymbol{\varphi}(\boldsymbol{y}^*) = \boldsymbol{x}^*(0)$. By definition of the tangent space,

$$T_{\boldsymbol{x}^*(0)} L_0 = \operatorname{Range} \boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*), \tag{2.105}$$

so that, by equation (2.30),

$$\operatorname{Ker} \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0)) = \operatorname{Range} \boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*). \tag{2.106}$$

By definition of L_0, $\boldsymbol{b}(\boldsymbol{\varphi}(\boldsymbol{y})) = \boldsymbol{0}$. Differentiating this equation twice and evaluating at $\boldsymbol{y} = \boldsymbol{y}^*$, we obtain

$$\boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0)) \boldsymbol{\varphi}_{\boldsymbol{y}\boldsymbol{y}}(\boldsymbol{y}^*)[\boldsymbol{y}, \boldsymbol{z}] = -\boldsymbol{b}_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(0))[\boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{y}, \boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{z}], \tag{2.107}$$

for any $\boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$. On the other end, from equality (2.106) and assumption (**2.A5**),

$$\big(\boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{y}_1\big)^T \boldsymbol{J} \boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{y}_2 = 0, \tag{2.108}$$

for any $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^n$. Differentiating this equation, we find

$$\big(\boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{y}_2\big)^T \boldsymbol{J} \boldsymbol{\varphi}_{\boldsymbol{y}\boldsymbol{y}}(\boldsymbol{y}^*)[\boldsymbol{y}_1, \boldsymbol{z}] = \big(\boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{y}_1\big)^T \boldsymbol{J} \boldsymbol{\varphi}_{\boldsymbol{y}\boldsymbol{y}}(\boldsymbol{y}^*)[\boldsymbol{y}_2, \boldsymbol{z}]. \tag{2.109}$$

By assumption (**2.A5**), equation (2.107), and the definition of $\boldsymbol{B}$, this equation becomes

$$\begin{aligned}\big(\boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{y}_2\big)^T \boldsymbol{J}\boldsymbol{B}\boldsymbol{b}_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(0))[\boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{y}_1, \boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{z}]& \\ = \big(\boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{y}_1\big)^T \boldsymbol{J}\boldsymbol{B}\boldsymbol{b}_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(0))[\boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{y}_2, \boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{z}].&\end{aligned} \tag{2.110}$$

By (2.106), it is possible to choose $\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{z} \in \mathbb{R}^n$ such that

$$\boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{y}_1 = \boldsymbol{v}_1^*(0), \quad \boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{y}_2 = \boldsymbol{\varphi}_{\boldsymbol{y}}(\boldsymbol{y}^*)\boldsymbol{z} = \boldsymbol{x}_\theta^*(0). \tag{2.111}$$

We conclude that

$$\boldsymbol{x}_\theta^*(0)^T \boldsymbol{J}\boldsymbol{B}\boldsymbol{b}_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(0))[\boldsymbol{v}_1^*(0), \boldsymbol{x}_\theta^*(0)] = \boldsymbol{v}_1^*(0)^T \boldsymbol{J}\boldsymbol{B}\boldsymbol{b}_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(0))[\boldsymbol{x}_\theta^*(0), \boldsymbol{x}_\theta^*(0)]. \tag{2.112}$$

Similarly,

$$\boldsymbol{x}_\theta^*(1)^T \boldsymbol{J}\boldsymbol{C}\boldsymbol{c}_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(1))[\boldsymbol{v}_1^*(1), \boldsymbol{x}_\theta^*(1)] = \boldsymbol{v}_1^*(1)^T \boldsymbol{J}\boldsymbol{C}\boldsymbol{c}_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}^*(1))[\boldsymbol{x}_\theta^*(1), \boldsymbol{x}_\theta^*(1)]. \tag{2.113}$$

Combining equations (2.98), (2.100), (2.104), (2.112) and (2.113) establishes equality (2.99). This achieves the verification of the hypotheses (i), (ii) and (iii) of the Crandall-Rabinowitz theorem.

Note that

$$\mathbb{R} \times W = \operatorname{Span}\big\{(1, \boldsymbol{v}_1^*)\big\} \oplus \big(\{0\} \times W\big), \tag{2.114}$$

i.e. $\{0\} \times W$ is a complement of Ker $\boldsymbol{A}$ in $\mathbb{R} \times W$. Therefore, by the Crandall-Rabinowitz theorem (cf. appendix A), there exist a positive number ϵ_0, a neighborhood $I_0 \times W_0$ of $(\theta^*, \mathbf{0})$ in $I^* \times W^*$ and functions

$$\phi \in C^{k-1}((-\epsilon_0, \epsilon_0), I_0), \quad \boldsymbol{\eta} \in C^{k-1}((-\epsilon_0, \epsilon_0), W) \tag{2.115}$$

such that

$$\phi(0) = \theta^*, \quad \boldsymbol{\eta}(0) = \mathbf{0}, \quad \boldsymbol{\eta}_\epsilon(0) = \mathbf{0}, \tag{2.116}$$

and the set of solutions to $\boldsymbol{\Phi}(\theta, \epsilon, \boldsymbol{w}) = \mathbf{0}$ in $I_0 \times (-\epsilon_0, \epsilon_0) \times W_0$ is

$$\big\{(\theta, 0, \mathbf{0}) : \theta \in I_0\big\} \cup \big\{\big(\phi(\epsilon), \epsilon, \epsilon \boldsymbol{v}_1^* + \boldsymbol{\eta}(\epsilon)\big) : |\epsilon| < \epsilon_0\big\}. \tag{2.117}$$

Choosing $\mathcal{X} = \boldsymbol{\psi}(I_0 \times W_0)$ as neighborhood of $\boldsymbol{x}^*$ in $\mathcal{C}^1$, we conclude by lemma 2.6 that, for any ϵ satisfying $0 \neq |\epsilon| < \epsilon_0$, the unique solution to BVP (2.3) in $\mathcal{X}$ is given by $\boldsymbol{x} = \boldsymbol{\xi}(\epsilon)$, where $\boldsymbol{\xi} \in C^{k-1}((-\epsilon_0, \epsilon_0), \mathcal{X})$ is defined by

$$\boldsymbol{\xi}(\epsilon) = \boldsymbol{\psi}\big(\phi(\epsilon), \epsilon \boldsymbol{v}_1^* + \boldsymbol{\eta}(\epsilon)\big) = \boldsymbol{\zeta}(\phi(\epsilon)) + \epsilon \boldsymbol{v}_1^* + \boldsymbol{\eta}(\epsilon). \tag{2.118}$$

■

It follows from Taylor's theorem (cf. appendix A) that, under the assumptions of part (b) of theorem 2.10,

$$\boldsymbol{\xi}(\epsilon) = \boldsymbol{x}^* + O(\epsilon), \quad \text{as } \epsilon \to 0, \tag{2.119}$$

and we call $\boldsymbol{x}^*$ the leading order approximation to the solution $\boldsymbol{\xi}(\epsilon)$. A standard method to compute a better approximation to a solution to BVP (2.3) is a perturbation expansion of the solution in powers of ϵ. The remainder of this section shows that, under the assumptions of part (b) of theorem 2.10 and for $k \geqslant 4$, this formal method can always be applied and is rigorously justified.

By Taylor's theorem (cf. appendix A), the solution $\boldsymbol{\xi}(\epsilon)$ can be expanded about $\boldsymbol{x}^*$ as

$$\boldsymbol{\xi}(\epsilon) = \boldsymbol{x}^* + \epsilon \boldsymbol{x}_1 + \ldots + \epsilon^{k-2}\boldsymbol{x}_{k-2} + O(\epsilon^{k-1}), \quad \text{as } \epsilon \to 0. \tag{2.120}$$

We now show that the functions $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k-3}$ can be uniquely determined. Substituting Taylor series (2.120) into BVP (2.3), expanding in powers of ϵ, and equating coefficients of like powers of ϵ, we find that $\boldsymbol{x}_1$ has to satisfy

$$\begin{cases} \boldsymbol{x}_1'(s) - \boldsymbol{J}H_{\boldsymbol{xx}}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{x}_1(s) = \boldsymbol{J}H_{\boldsymbol{x}\epsilon}(\boldsymbol{x}^*(s), s, 0), \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))\boldsymbol{x}_1(0) = \boldsymbol{0}, \quad \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\boldsymbol{x}_1(1) = \boldsymbol{0}, \end{cases} \tag{2.121}$$

$\boldsymbol{x}_2$ has to satisfy

$$\begin{cases} \boldsymbol{x}_2'(s) - \boldsymbol{J}H_{\boldsymbol{xx}}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{x}_2(s) = \tfrac{1}{2}\boldsymbol{J}H_{\boldsymbol{xxx}}(\boldsymbol{x}^*(s), s, 0)[\boldsymbol{x}_1(s), \boldsymbol{x}_1(s)] \\ \qquad\qquad + \boldsymbol{J}H_{\boldsymbol{xx}\epsilon}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{x}_1(s) + \tfrac{1}{2}\boldsymbol{J}H_{\boldsymbol{x}\epsilon\epsilon}(\boldsymbol{x}^*(s), s, 0), \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))\boldsymbol{x}_2(0) = -\tfrac{1}{2}\boldsymbol{b}_{\boldsymbol{xx}}(\boldsymbol{x}^*(0))[\boldsymbol{x}_1(0), \boldsymbol{x}_1(0)], \\ \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\boldsymbol{x}_2(1) = -\tfrac{1}{2}\boldsymbol{c}_{\boldsymbol{xx}}(\boldsymbol{x}^*(1))[\boldsymbol{x}_1(1), \boldsymbol{x}_1(1)], \end{cases} \tag{2.122}$$

and, for $i \in \{3, \ldots, k-2\}$, $\boldsymbol{x}_i$ has to satisfy an inhomogeneous linear BVP of the form

$$\begin{cases} \boldsymbol{x}_i'(s) - \boldsymbol{J}H_{\boldsymbol{xx}}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{x}_i(s) = \boldsymbol{J}H_{\boldsymbol{xxx}}(\boldsymbol{x}^*(s), s, 0)[\boldsymbol{x}_1(s), \boldsymbol{x}_{i-1}(s)] \\ \qquad\qquad + \boldsymbol{J}H_{\boldsymbol{xx}\epsilon}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{x}_{i-1}(s) + \boldsymbol{J}\boldsymbol{\eta}_i(\boldsymbol{x}^*(s), \ldots, \boldsymbol{x}_{i-2}(s), s), \\ \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))\boldsymbol{x}_i(0) = -\tfrac{1}{2}\boldsymbol{b}_{\boldsymbol{xx}}(\boldsymbol{x}^*(0))[\boldsymbol{x}_1(0), \boldsymbol{x}_{i-1}(0)] + \boldsymbol{\beta}_i(\boldsymbol{x}^*(0), \ldots, \boldsymbol{x}_{i-2}(0)), \\ \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\boldsymbol{x}_i(1) = -\tfrac{1}{2}\boldsymbol{c}_{\boldsymbol{xx}}(\boldsymbol{x}^*(1))[\boldsymbol{x}_1(1), \boldsymbol{x}_{i-1}(1)] + \boldsymbol{\gamma}_i(\boldsymbol{x}^*(1), \ldots, \boldsymbol{x}_{i-2}(1)), \end{cases} \tag{2.123}$$

where $\boldsymbol{\eta}_i$, $\boldsymbol{\beta}_i$ and $\boldsymbol{\gamma}_i$ are known functions. Equation (2.123) is easily proven by induction. We now show that the set of BVPs (2.121)-(2.123) can be solved successively to obtain $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k-3}$.

The solvability condition for BVP (2.121), $\mathcal{M}(\theta^*) = 0$, is satisfied so that, by lemma 2.7, the solution to BVP (2.121) is of the form

$$\boldsymbol{x}_1 = \boldsymbol{v}_1^* + \sigma_1 \boldsymbol{x}_\theta^*, \tag{2.124}$$

where $\boldsymbol{v}_1^* \in W$ is uniquely defined and $\sigma_1 \in \mathbb{R}$ is a free parameter. However $\boldsymbol{x}_2$ is a solution to BVP (2.122), so that the associated solvability condition must be satisfied, which fixes the value of σ_1, as we show now. Substituting (2.124) into the solvability condition for BVP (2.122) and grouping the terms of like powers of σ_1, we obtain

$$\tfrac{1}{2}F_2\sigma_1^2 + F_1\sigma_1 + F_0 = 0, \tag{2.125}$$

where

$$
\begin{aligned}
F_2 = \int_0^1 & \boldsymbol{x}_\theta^*(s)^T H_{\boldsymbol{xxx}}(\boldsymbol{x}^*(s), s, 0)[\boldsymbol{x}_\theta^*(s), \boldsymbol{x}_\theta^*(s)]\, ds \\
& + \boldsymbol{x}_\theta^*(0)^T \boldsymbol{JB} \boldsymbol{b}_{\boldsymbol{xx}}(\boldsymbol{x}^*(0))[\boldsymbol{x}_\theta^*(0), \boldsymbol{x}_\theta^*(0)] \\
& - \boldsymbol{x}_\theta^*(1)^T \boldsymbol{JC} \boldsymbol{c}_{\boldsymbol{xx}}(\boldsymbol{x}^*(1))[\boldsymbol{x}_\theta^*(1), \boldsymbol{x}_\theta^*(1)],
\end{aligned} \tag{2.126}
$$

F_1 is defined in (2.98), and

$$
\begin{aligned}
F_0 = \int_0^1 & \boldsymbol{x}_\theta^*(s)^T \Big(\tfrac{1}{2} H_{\boldsymbol{xxx}}(\boldsymbol{x}^*(s), s, 0)[\boldsymbol{v}_1^*(s), \boldsymbol{v}_1^*(s)] \\
& \qquad + H_{\boldsymbol{xx}\epsilon}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{v}_1^*(s) + \tfrac{1}{2} H_{\boldsymbol{x}\epsilon\epsilon}(\boldsymbol{x}^*(s), s, 0) \Big)\, ds \\
& + \tfrac{1}{2} \boldsymbol{x}_\theta^*(0)^T \boldsymbol{JB} \boldsymbol{b}_{\boldsymbol{xx}}(\boldsymbol{x}^*(0))[\boldsymbol{v}_1^*(0), \boldsymbol{v}_1^*(0)] \\
& - \tfrac{1}{2} \boldsymbol{x}_\theta^*(1)^T \boldsymbol{JC} \boldsymbol{c}_{\boldsymbol{xx}}(\boldsymbol{x}^*(1))[\boldsymbol{v}_1^*(1), \boldsymbol{v}_1^*(1)].
\end{aligned} \tag{2.127}
$$

Differentiating BVP (2.33) with respect to θ twice and setting $\theta = \theta^*$ shows that $\boldsymbol{x}_{\theta\theta}^*$ is a solution to the BVP

$$
\begin{cases}
\boldsymbol{v}'(s) - \boldsymbol{J} H_{\boldsymbol{xx}}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{v}(s) = \boldsymbol{J} H_{\boldsymbol{xxx}}(\boldsymbol{x}^*(s), s, 0)[\boldsymbol{x}_\theta^*(s), \boldsymbol{x}_\theta^*(s)], \\
\boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{x}^*(0))\boldsymbol{v}(0) = -\boldsymbol{b}_{\boldsymbol{xx}}(\boldsymbol{x}^*(0))[\boldsymbol{x}_\theta^*(0), \boldsymbol{x}_\theta^*(0)], \\
\boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{x}^*(1))\boldsymbol{v}(1) = -\boldsymbol{c}_{\boldsymbol{xx}}(\boldsymbol{x}^*(1))[\boldsymbol{x}_\theta^*(1), \boldsymbol{x}_\theta^*(1)].
\end{cases} \tag{2.128}
$$

Thus, by lemma 2.9, the associated solvability condition is satisfied, i.e.

$$
F_2 = 0. \tag{2.129}
$$

Therefore, by equality (2.99), equation (2.125) simplifies to

$$
\mathcal{M}_\theta(\theta^*)\sigma_1 + F_0 = 0. \tag{2.130}
$$

By transversality condition $\mathcal{M}_\theta(\theta^*) \neq 0$, this equation has a unique solution σ_1^*, and the first correction

$$
\boldsymbol{x}_1 = \boldsymbol{v}_1^* + \sigma_1^* \boldsymbol{x}_\theta^*, \tag{2.131}
$$

is completely determined. Now substitute (2.131) into BVP (2.122). By lemma 2.7 its solution is of the form

$$
\boldsymbol{x}_2 = \boldsymbol{v}_2^* + \sigma_2 \boldsymbol{x}_\theta^*, \tag{2.132}
$$

where $\boldsymbol{v}_2^* \in W$ is uniquely defined and $\sigma_2 \in \mathbb{R}$ still has to be determined.

We claim that for any fixed $i \in \{3, \ldots, k-2\}$, if

$$
\boldsymbol{x}_j = \boldsymbol{v}_j^* + \sigma_j^* \boldsymbol{x}_\theta^*, \quad j = 1, \ldots, i-2, \tag{2.133}
$$

are known and

$$
\boldsymbol{x}_{i-1} = \boldsymbol{v}_{i-1}^* + \sigma_{i-1} \boldsymbol{x}_\theta^* \tag{2.134}
$$

is known up to a constant $\sigma_{i-1} \in \mathbb{R}$, then the solvability condition for BVP (2.123) uniquely defines the value σ_{i-1}^* of the constant, and $\boldsymbol{x}_i$ is of the form

$$\boldsymbol{x}_i = \boldsymbol{v}_i^* + \sigma_i \boldsymbol{x}_\theta^*, \tag{2.135}$$

where $\boldsymbol{v}_i^* \in W$ is uniquely defined and $\sigma_i \in \mathbb{R}$. To prove the claim we start by substituting (2.133) and (2.134) into the solvability condition for BVP (2.123) and group the terms of like powers of σ_{i-1}. We obtain

$$(F_2\sigma_1^* + F_1)\sigma_{i-1} + G_i = 0, \tag{2.136}$$

where F_2 and F_1 are defined in (2.126) and (2.98), and

$$\begin{aligned} G_i = \; & \int_0^1 \boldsymbol{x}_\theta^*(s)^T \Big(H_{\boldsymbol{xxx}}(\boldsymbol{x}^*(s), s, 0)[\boldsymbol{v}_1^*(s) + \sigma_1^* \boldsymbol{x}_\theta^*(s), \boldsymbol{v}_{i-1}^*(s)] \\ & \qquad + H_{\boldsymbol{xx}\epsilon}(\boldsymbol{x}^*(s), s, 0)\boldsymbol{v}_{i-1}^*(s) + \boldsymbol{\eta}_i(\boldsymbol{x}^*(s), \dots, \boldsymbol{v}_{i-2}^*(s) + \sigma_{i-2}^*\boldsymbol{x}_\theta^*(s), s) \Big)\, ds \\ & + \boldsymbol{x}_\theta^*(0)^T \boldsymbol{JB} \Big(\boldsymbol{b}_{\boldsymbol{xx}}(\boldsymbol{x}^*(0))[\boldsymbol{v}_1^*(0) + \sigma_1^* \boldsymbol{x}_\theta^*(0), \boldsymbol{v}_{i-1}^*(0)] \\ & \qquad + \boldsymbol{\beta}_i(\boldsymbol{x}^*(0), \dots, \boldsymbol{v}_{i-2}^*(0) + \sigma_{i-2}^*\boldsymbol{x}_\theta^*(0)) \Big) \\ & + \boldsymbol{x}_\theta^*(1)^T \boldsymbol{JC} \Big(\boldsymbol{c}_{\boldsymbol{xx}}(\boldsymbol{x}^*(1))[\boldsymbol{v}_1^*(1) + \sigma_1^* \boldsymbol{x}_\theta^*(1), \boldsymbol{v}_{i-1}^*(1)] \\ & \qquad + \boldsymbol{\gamma}_i(\boldsymbol{x}^*(1), \dots, \boldsymbol{v}_{i-2}^*(1) + \sigma_{i-2}^*\boldsymbol{x}_\theta^*(1)) \Big). \end{aligned} \tag{2.137}$$

By (2.129) and (2.99), equation (2.136) simplifies to

$$\mathcal{M}_\theta(\theta^*)\sigma_{i-1} + G_i = 0. \tag{2.138}$$

By transversality condition $\mathcal{M}_\theta(\theta^*) \neq 0$, this equation has a unique solution σ_{i-1}^*, and

$$\boldsymbol{x}_{i-1} = \boldsymbol{v}_{i-1}^* + \sigma_{i-1}^* \boldsymbol{x}_\theta^*. \tag{2.139}$$

Substituting (2.133) and (2.139) into BVP (2.123), and using lemma 2.7 shows that $\boldsymbol{x}_i$ is of the form

$$\boldsymbol{x}_i = \boldsymbol{v}_i^* + \sigma_i \boldsymbol{x}_\theta^*, \tag{2.140}$$

where $\boldsymbol{v}_i^* \in W$ is uniquely defined and $\sigma_i \in \mathbb{R}$. This concludes the proof of the claim.

From the previous argument by induction we conclude that the BVPs (2.121)-(2.123) uniquely define the corrections $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{k-3}$ in the Taylor expansion (2.120). Note that $\boldsymbol{x}_{k-2}$ can only be determined up to a scalar multiple of $\boldsymbol{x}_\theta^*$ by BVPs (2.121)-(2.123).

Chapter 3

Averaging in Hamiltonian boundary value problems

Consider the canonical Hamiltonian boundary value problem (BVP)

$$\begin{cases} \boldsymbol{x}'(s) = \boldsymbol{J} H_{\boldsymbol{x}}\big(\boldsymbol{x}(s), s, \tfrac{s}{\epsilon}, \epsilon\big), \quad \forall\, s \in (0,1), \\ \boldsymbol{b}(\boldsymbol{x}(0)) = \boldsymbol{0}, \quad \boldsymbol{c}(\boldsymbol{x}(1)) = \boldsymbol{0}, \end{cases} \tag{3.1}$$

where $\boldsymbol{x}(s) \in \mathbb{R}^{2n}$, $\boldsymbol{J}$ is the canonical symplectic matrix (2.2), ϵ is a small (nonzero) parameter, and

$$\begin{aligned} &H : \mathbb{R}^{2n} \times [0,1] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad \boldsymbol{b}, \boldsymbol{c} : \mathbb{R}^{2n} \to \mathbb{R}^n \quad \text{are smooth}, \\ &\exists\, T > 0 \ \text{ such that } \ H(\,\cdot\,, \cdot\,, \sigma + T, \cdot\,) = H(\,\cdot\,, \cdot\,, \sigma, \cdot\,), \quad \forall\, \sigma \in \mathbb{R}. \end{aligned} \tag{\textbf{3.A1}}$$

According to these hypotheses, the Hamiltonian in (3.1) has a T-periodic dependence in $\frac{s}{\epsilon}$. This dependence adds a rapid oscillation (of period ϵT) to the variation of the Hamiltonian in the independent variable s. Problem (3.1) is then characterized by two widely different scales defined by the *long scale* variable s and the *short scale* variable $\frac{s}{\epsilon}$. Consider the associated *averaged BVP*

$$\begin{cases} \boldsymbol{y}'(s) = \boldsymbol{J} \langle H \rangle_{\boldsymbol{y}}(\boldsymbol{y}(s), s), \\ \boldsymbol{b}(\boldsymbol{y}(0)) = \boldsymbol{0}, \quad \boldsymbol{c}(\boldsymbol{y}(1)) = \boldsymbol{0}, \end{cases} \tag{3.2}$$

where the averaged Hamiltonian $\langle H \rangle$ is defined by

$$\langle H \rangle(\boldsymbol{y}, s) = \frac{1}{T} \int_0^T H(\boldsymbol{y}, s, t, 0)\, dt. \tag{3.3}$$

The purpose of this chapter is to establish rigorous criteria for the existence (or absence) of a solution to the two-scale problem (3.1) in the vicinity of a given solution to the averaged problem (3.2). We treat the case where the given solution to the averaged problem is isolated, as well as the case where it is part of a curve of solutions. As a benefit of our study, we obtain rigorously justified techniques to compute approximations (of any given order in ϵ) to solutions to the two-scale problem.

One of the standard problems treated in *averaging theory* (cf. Murdock [41, chap. 6], Guckenheimer and Holmes [21, chap. 4], Sanders and Verhulst [47, chap. 3]) is the initial value problem

$$\boldsymbol{z}'(\sigma) = \epsilon \boldsymbol{f}(\boldsymbol{z}(\sigma), \sigma, \epsilon), \quad \boldsymbol{z}(0) = \boldsymbol{z}_0, \tag{3.4}$$

where $\sigma \in \mathbb{R}$, $\boldsymbol{z}(\sigma) \in \mathbb{R}^m$, the function $\boldsymbol{f}(\boldsymbol{\mu}, \sigma, \epsilon)$ is smooth and T-periodic in σ, ϵ is a small parameter, and $\boldsymbol{z}_0 \in \mathbb{R}^m$ is a given initial condition. The associated averaged problem is

$$\boldsymbol{z}'(\sigma) = \epsilon \frac{1}{T} \int_0^T \boldsymbol{f}(\boldsymbol{z}(\sigma), t, 0)\, dt, \quad \boldsymbol{z}(0) = \boldsymbol{z}_0. \tag{3.5}$$

If the solution $\bar{\boldsymbol{z}}(\sigma)$ to the averaged problem exists for σ in an expanding interval $[0, \Sigma/\epsilon]$, with $0 < \Sigma < \infty$, then, for ϵ sufficiently small, the solution $\boldsymbol{z}(\sigma)$ to the original problem (3.4) also exists for $\sigma \in [0, \Sigma/\epsilon]$. Moreover, $\|\boldsymbol{z}(\sigma) - \bar{\boldsymbol{z}}(\sigma)\| = O(\epsilon)$ uniformly for $\sigma \in [0, \Sigma/\epsilon]$, and $\bar{\boldsymbol{z}}(\sigma)$ constitutes a leading order approximation to $\boldsymbol{z}(\sigma)$. In fact, the averaging theory proposes a technique to construct approximations to $\boldsymbol{z}(\sigma)$ of any order in ϵ. A careful discussion of these standard results can be found in Murdock [41, thms. 6.2.2 and 6.5.2]. The key to the proof of these results is a *near-identity transformation*, i.e. a change of variable which reduces to the identity in the limit $\epsilon \to 0$. The purpose of this transformation is to remove the explicit σ-dependence in the problem, i.e. to make the problem autonomous (up to some order in ϵ). The averaged problem is obtained by keeping only the $O(1)$ and $O(\epsilon)$ terms from the transformed problem. In the case where problem (3.4) is Hamiltonian, the near-identity transformation can be chosen to be canonical (cf. Kevorkian and Cole [30, sect. 5.2]).

Under the change of independent variable $s \to \sigma = \frac{s}{\epsilon}$ and the change of dependent variable $\boldsymbol{x} \to \boldsymbol{z} = (\boldsymbol{x}, s)$, the two-scale differential equation in (3.1) takes the form of the differential equation in (3.4). Thus, in the case of a pure initial value problem, the standard averaging results for problem (3.4) immediately yields averaging results for the two-scale differential equation in (3.1). This fact was exploited by Kehrbaum [27, sect. 8.3] in the context of rods with high intrinsic twist.

The change of variable $s \to \sigma$ is not as helpful when applied to the full two-point BVP (3.1), since it transforms the boundary conditions at $s = 1$ into boundary conditions at $\sigma = \frac{1}{\epsilon}$, which is sent to infinity in the limit $\epsilon \to 0$. To avoid this difficulty, we work with the long scale variable s as independent variable to establish averaging results for BVP (3.1). Inspired by the standard theory of averaging, we perform a near-identity canonical change of variable to remove the $\frac{s}{\epsilon}$-dependence in the Hamiltonian (up to some order in ϵ). Unlike the original Hamiltonian, the transformed one is a smooth function of ϵ at $\epsilon = 0$, so that the perturbation theorems of chapter 2 apply on the transformed BVP, and yield the desired results. In some sense, the averaged BVP (3.2) plays the role of the unperturbed problem in chapter 2.

Two important features distinguish our treatment of BVP (3.1) from the standard averaging theory for initial value problems. The first difference is that solutions to the averaged BVP (3.2) may be nonisolated. We treat the case of an isolated solution to the averaged problem, as well as the symmetry breaking occurring in a situation where the averaged problem has a curve of solutions. The second difference is that, due to the two-point boundary conditions, our near-identity transformation can be used only for values of the small parameter ϵ belonging to the set

$$\mathbb{E}_T = \left\{ \frac{1}{\nu T} : \nu \in \mathbb{Z}\backslash\{0\} \right\}. \tag{3.6}$$

However, this is not an essential restriction since, as is shown at the end of the chapter, a problem of the form (3.1) with $\epsilon \notin \mathbb{E}_T$ can easily be turned into a problem of the same form with a small parameter belonging to $\mathbb{E}_T$. Note that a similar argument has been exploited in Mielke *et al.* [40] for the study of modulated systems with periodic boundary conditions.

This chapter is organized as follows. In section 3.1, we construct and analyze a general transformation depending on s, ϵ and a parameter $\sigma \in \mathbb{R}$. In section 3.2, we introduce the near-identity canonical change of variable for BVP (3.1), and verify that the transformed problem is sufficiently smooth for the theorems of chapter 2 to apply. In doing this, we exploit the fact that the near-identity canonical change of variable can be obtained by setting $\sigma = \frac{s}{\epsilon}$ in the general transformation of section 3.1. In section 3.3 it is shown that, under some nondegeneracy condition, there exists a solution to the two-scale BVP (3.1) in some neighborhood of a given isolated solution to the averaged BVP (3.2) when $\epsilon \in \mathbb{E}_T$ is sufficiently small. A general method to compute approximations (of any given order in ϵ) to the solution to BVP (3.1) is presented. In section 3.4 we assume the boundary conditions of BVP (3.1) define a Lagrangian manifold at each end-point, and the averaged BVP (3.2) admits a smooth curve Γ of solutions. We exhibit a Melnikov function which, under some nondegeneracy conditions, indicates which of the solutions in Γ give rise to solutions to the two-scale BVP (3.1): if $\boldsymbol{y}^* \in \Gamma$ is associated to a nonzero value of the Melnikov function then, for $\epsilon \in \mathbb{E}_T$ sufficiently small, BVP (3.1) does not admit any solution in some neighborhood of $\boldsymbol{y}^*$; contrariwise, if $\boldsymbol{y}^* \in \Gamma$ is associated to a simple zero of the Melnikov function then, for $\epsilon \in \mathbb{E}_T$ sufficiently small, there exists a solution to BVP (3.1) in some neighborhood of $\boldsymbol{y}^*$. In the latter case we present a general method to compute approximations (of any given order in ϵ) to the solution to BVP (3.1). Finally, section 3.5 shows that a problem of the form (3.1) with $\epsilon \notin \mathbb{E}_T$ can easily be turned into a problem of the same form with a small parameter belonging to $\mathbb{E}_T$. As a consequence, a problem of the form (3.1) with $\epsilon \notin \mathbb{E}_T$ can also be treated by the results of sections 3.3 and 3.4. However we do not insist on this case since we will study only BVPs of the form (3.1) with $\epsilon \in \mathbb{E}_T$ in the rest of this dissertation (cf. chapter 6).

3.1 Preliminary results

In this section, we construct and analyze a general transformation depending on s, ϵ and a parameter $\sigma \in \mathbb{R}$. The properties of this transformation are used in section 3.2 to derive properties of the near-identity canonical change of variable for problem (3.1).

Consider the transformation $\boldsymbol{x} = (\boldsymbol{q}, \boldsymbol{p}) \to \boldsymbol{y} = (\boldsymbol{Q}, \boldsymbol{P})$ defined implicitly by

$$\boldsymbol{q} = \boldsymbol{Q} + \sum_{i=1}^{2k+3} \epsilon^i \, W_{\boldsymbol{p}}^{(i)}(\boldsymbol{Q}, \boldsymbol{p}, s, \sigma), \quad \boldsymbol{p} = \boldsymbol{P} - \sum_{i=1}^{2k+3} \epsilon^i \, W_{\boldsymbol{Q}}^{(i)}(\boldsymbol{Q}, \boldsymbol{p}, s, \sigma), \tag{3.7}$$

where $k \geqslant 1$ and, for $i = 1, \ldots, 2k+3$,

$$W^{(i)} : \mathbb{R}^n \times \mathbb{R}^n \times [0,1] \times \mathbb{R} \to \mathbb{R} \text{ is smooth,} \quad W^{(i)}(\,\cdot\,,\,\cdot\,,\sigma+T) = W^{(i)}(\,\cdot\,,\,\cdot\,,\sigma). \tag{3.8}$$

Note that we apply the adjective *smooth* to any function that is continuously differentiable and that has as many derivatives as are needed to make the associated mathematical processes valid.

Lemma 3.1 *Let $R > 0$ and define $\mathcal{B}_R = \{\boldsymbol{y} \in \mathbb{R}^{2n} : \|\boldsymbol{y}\| \leqslant R\}$. Then there exists a positive number ϵ_R and a mapping $\boldsymbol{\psi}$ satisfying*

$$\begin{aligned} &\boldsymbol{\psi} : \mathcal{B}_R \times [0,1] \times \mathbb{R} \times (-\epsilon_R, \epsilon_R) \to \mathcal{B}_{2R} \textit{ is smooth,} \\ &\boldsymbol{\psi}(\,\cdot\,,\,\cdot\,,\sigma+T,\,\cdot\,) = \boldsymbol{\psi}(\,\cdot\,,\,\cdot\,,\sigma,\,\cdot\,), \end{aligned} \tag{3.9}$$

such that, under the constraints $\boldsymbol{x} \in \mathcal{B}_{2R}$, $\boldsymbol{y} \in \mathcal{B}_R$, $\epsilon \in (-\epsilon_R, \epsilon_R)$, transformation (3.7) is equivalent to the explicit transformation

$$\boldsymbol{x} = \boldsymbol{\psi}(\boldsymbol{y}, s, \sigma, \epsilon), \tag{3.10}$$

which defines an isomorphism between $\mathcal{B}_R$ in $\boldsymbol{y}$-space and $\boldsymbol{\psi}(\mathcal{B}_R, s, \sigma, \epsilon)$ in $\boldsymbol{x}$-space, with $\mathcal{B}_{R/2} \subset \boldsymbol{\psi}(\mathcal{B}_R, s, \sigma, \epsilon) \subset \mathcal{B}_{2R}$.

Proof The key to this proof is the Banach fixed-point theorem. Define the functions

$$\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}, s, \sigma, \epsilon) = \sum_{i=1}^{2k+3} \epsilon^{i-1} \begin{pmatrix} W^{(i)}_{\boldsymbol{p}}(\boldsymbol{Q}, \boldsymbol{p}, s, \sigma) \\ -W^{(i)}_{\boldsymbol{Q}}(\boldsymbol{Q}, \boldsymbol{p}, s, \sigma) \end{pmatrix}, \quad \boldsymbol{\phi}(\boldsymbol{x}, \boldsymbol{y}, s, \sigma, \epsilon) = \boldsymbol{y} + \epsilon \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}, s, \sigma, \epsilon), \tag{3.11}$$

and the numbers

$$M = \max_{\substack{\boldsymbol{x} \in \mathcal{B}_{2R},\ \boldsymbol{y} \in \mathcal{B}_R, \\ s \in [0,1],\ \sigma \in [0,T], \\ \epsilon \in [-1,1]}} \|\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}, s, \sigma, \epsilon)\|, \qquad L = \max_{\substack{\boldsymbol{x} \in \mathcal{B}_{2R},\ \boldsymbol{y} \in \mathcal{B}_R, \\ s \in [0,1],\ \sigma \in [0,T], \\ \epsilon \in [-1,1]}} \|\boldsymbol{g}_{\boldsymbol{x}}(\boldsymbol{x}, \boldsymbol{y}, s, \sigma, \epsilon)\|, \tag{3.12}$$

$\epsilon_1 = \min\left(1, \frac{R}{M}, \frac{1}{2L}\right) > 0$. Transformation (3.7) can be rewritten as

$$\boldsymbol{x} = \boldsymbol{\phi}(\boldsymbol{x}, \boldsymbol{y}, s, \sigma, \epsilon). \tag{3.13}$$

For $\boldsymbol{x}, \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{B}_{2R}$, $\boldsymbol{y} \in \mathcal{B}_R$, $s \in [0,1]$, $\sigma \in \mathbb{R}$, $\epsilon \in (-\epsilon_1, \epsilon_1)$,

$$\|\boldsymbol{\phi}(\boldsymbol{x}, \boldsymbol{y}, s, \sigma, \epsilon)\| \leqslant \|\boldsymbol{y}\| + |\epsilon|\, \|\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}, s, \sigma, \epsilon)\| \leqslant R + |\epsilon|\, M \leqslant 2R, \tag{3.14}$$

so that the function $\boldsymbol{x} \mapsto \boldsymbol{\phi}(\boldsymbol{x}, \boldsymbol{y}, s, \sigma, \epsilon)$ maps $\mathcal{B}_{2R}$ into itself, and

$$\begin{aligned} \|\boldsymbol{\phi}(\boldsymbol{x}_1, \boldsymbol{y}, s, \sigma, \epsilon) - \boldsymbol{\phi}(\boldsymbol{x}_2, \boldsymbol{y}, s, \sigma, \epsilon)\| &\leqslant |\epsilon|\, \|\boldsymbol{g}(\boldsymbol{x}_1, \boldsymbol{y}, s, \sigma, \epsilon) - \boldsymbol{g}(\boldsymbol{x}_2, \boldsymbol{y}, s, \sigma, \epsilon)\| \\ &\leqslant |\epsilon|\, L\, \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \\ &\leqslant \tfrac{1}{2}\, \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|, \end{aligned} \tag{3.15}$$

so that the function $\boldsymbol{x} \mapsto \boldsymbol{\phi}(\boldsymbol{x}, \boldsymbol{y}, s, \sigma, \epsilon)$ is a contraction. It follows from the Banach fixed-point theorem (cf. appendix A) that, for any $\boldsymbol{y} \in \mathcal{B}_R$, $s \in [0,1]$, $\sigma \in \mathbb{R}$, $\epsilon \in$

$(-\epsilon_1, \epsilon_1)$, equation (3.13) admits a unique solution for $\boldsymbol{x}$ in $\mathcal{B}_{2R}$. Denote this solution by $\boldsymbol{\psi}(\boldsymbol{y}, s, \sigma, \epsilon)$.

We claim the existence of ϵ_2 satisfying $0 < \epsilon_2 \leqslant \epsilon_1$ and such that, for any fixed $s \in [0,1]$, $\sigma \in \mathbb{R}$, $\epsilon \in (-\epsilon_2, \epsilon_2)$, the transformation $\boldsymbol{y} \mapsto \boldsymbol{\psi}(\boldsymbol{y}, s, \sigma, \epsilon)$ defines a bijection from $\mathcal{B}_R$ onto its image $\boldsymbol{\psi}(\mathcal{B}_R, s, \sigma, \epsilon) \subset \mathcal{B}_{2R}$. The surjectivity is immediate. An application of the Banach fixed-point theorem similar to the above one shows that there exists ϵ_2 satisfying $0 < \epsilon_2 \leqslant \epsilon_1$ and such that, for any $\boldsymbol{x} \in \mathcal{B}_{2R}$, $s \in [0,1]$, $\sigma \in \mathbb{R}$, $\epsilon \in (-\epsilon_2, \epsilon_2)$, equation (3.7) admits a unique solution for $\boldsymbol{y}$ in $\mathcal{B}_{3R}$. This proves the injectivity, and the claim is established.

Applying the Banach fixed-point theorem once again in a similar manner shows that there exists ϵ_R satisfying $0 < \epsilon_R \leqslant \epsilon_2$ and such that, for any $\boldsymbol{x} \in \mathcal{B}_{R/2}$, $s \in [0,1]$, $\sigma \in \mathbb{R}$, $\epsilon \in (-\epsilon_R, \epsilon_R)$, equation (3.7) admits a solution for $\boldsymbol{y}$ in $\mathcal{B}_R$. It follows that $\mathcal{B}_{R/2} \subset \boldsymbol{\psi}(\mathcal{B}_R, s, \sigma, \epsilon)$.

We still have to establish properties (3.9). The periodicity of the mapping $\boldsymbol{\psi}$ with respect to the variable σ immediately follows from the periodicity of the functions $W^{(i)}$. The rest of the proof is devoted to establishing the smoothness of $\boldsymbol{\psi}$.

Adopt the notation $\boldsymbol{\gamma} = (\boldsymbol{y}, s, \sigma, \epsilon)$, and use the equality $\boldsymbol{\psi}(\boldsymbol{\gamma}) = \boldsymbol{\phi}(\boldsymbol{\psi}(\boldsymbol{\gamma}), \boldsymbol{\gamma})$ to write

$$\begin{aligned}
\|\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h}) - \boldsymbol{\psi}(\boldsymbol{\gamma})\| &= \|\boldsymbol{\phi}(\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h}), \boldsymbol{\gamma}+\boldsymbol{h}) - \boldsymbol{\phi}(\boldsymbol{\psi}(\boldsymbol{\gamma}), \boldsymbol{\gamma})\| \\
&\leqslant \|\boldsymbol{\phi}(\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h}), \boldsymbol{\gamma}+\boldsymbol{h}) - \boldsymbol{\phi}(\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h}), \boldsymbol{\gamma})\| \\
&\quad + \|\boldsymbol{\phi}(\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h}), \boldsymbol{\gamma}) - \boldsymbol{\phi}(\boldsymbol{\psi}(\boldsymbol{\gamma}), \boldsymbol{\gamma})\| \\
&\leqslant \|\boldsymbol{\phi}(\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h}), \boldsymbol{\gamma}+\boldsymbol{h}) - \boldsymbol{\phi}(\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h}), \boldsymbol{\gamma})\| \\
&\quad + \tfrac{1}{2} \|\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h}) - \boldsymbol{\psi}(\boldsymbol{\gamma})\|,
\end{aligned} \tag{3.16}$$

where the last inequality follows from (3.15). Hence

$$\|\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h}) - \boldsymbol{\psi}(\boldsymbol{\gamma})\| \leqslant 2\, \|\boldsymbol{\phi}(\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h}), \boldsymbol{\gamma}+\boldsymbol{h}) - \boldsymbol{\phi}(\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h}), \boldsymbol{\gamma})\|. \tag{3.17}$$

By continuity of $\boldsymbol{\phi}$, the right hand side goes to zero as $\|\boldsymbol{h}\| \to 0$, so that $\boldsymbol{\psi}$ is continuous.

For $\boldsymbol{\gamma} = (\boldsymbol{y}, s, \sigma, \epsilon) \in \mathcal{B}_R \times [0,1] \times \mathbb{R} \times (-\epsilon_R, \epsilon_R)$,

$$\|\boldsymbol{\phi}_{\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}), \boldsymbol{\gamma})\| = |\epsilon|\, \|\boldsymbol{g}_{\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}), \boldsymbol{\gamma})\| \leqslant |\epsilon|\, L < \tfrac{1}{2}, \tag{3.18}$$

where L is defined in (3.12) and the last inequality follows from the fact that $|\epsilon| < |\epsilon_R| \leqslant \frac{1}{2L}$. We conclude that the continuous linear operator

$$\boldsymbol{I} - \boldsymbol{\phi}_{\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}), \boldsymbol{\gamma}) : \mathbb{R}^{2n} \to \mathbb{R}^{2n}, \tag{3.19}$$

where $\boldsymbol{I}$ denotes the identity operator, is invertible. The inverse is given by the absolutely convergent Neumann series

$$\Big(\boldsymbol{I} - \boldsymbol{\phi}_{\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}), \boldsymbol{\gamma})\Big)^{-1} = \sum_{n=0}^{\infty} \Big(\boldsymbol{\phi}_{\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}), \boldsymbol{\gamma})\Big)^{n}, \tag{3.20}$$

and is a continuous function of $\boldsymbol{\gamma}$. With this result, we are ready to establish the differentiability of $\boldsymbol{\psi}$.

By differentiability of $\boldsymbol{\phi}$ and continuity of $\boldsymbol{\psi}$,

$$\begin{aligned}\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h})-\boldsymbol{\psi}(\boldsymbol{\gamma}) &= \boldsymbol{\phi}(\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h}),\boldsymbol{\gamma}+\boldsymbol{h})-\boldsymbol{\phi}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})\\ &= \boldsymbol{\phi}_{\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})\big(\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h})-\boldsymbol{\psi}(\boldsymbol{\gamma})\big)+\boldsymbol{\phi}_{\boldsymbol{\gamma}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})\boldsymbol{h}+o(\|\boldsymbol{h}\|),\end{aligned} \tag{3.21}$$

as $\|\boldsymbol{h}\|\to 0$. Therefore

$$\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{h})-\boldsymbol{\psi}(\boldsymbol{\gamma}) = \big(\boldsymbol{I}-\boldsymbol{\phi}_{\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})\big)^{-1}\boldsymbol{\phi}_{\boldsymbol{\gamma}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})\boldsymbol{h}+o(\|\boldsymbol{h}\|), \tag{3.22}$$

so that the derivative $\boldsymbol{\psi}_{\boldsymbol{\gamma}}$ exists and is defined by

$$\boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\gamma})\boldsymbol{h} = \big(\boldsymbol{I}-\boldsymbol{\phi}_{\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})\big)^{-1}\boldsymbol{\phi}_{\boldsymbol{\gamma}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})\boldsymbol{h}. \tag{3.23}$$

Moreover, $\boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\gamma})$ is a continuous function of $\boldsymbol{\gamma}$, so that $\boldsymbol{\psi}\in C^1$.

From (3.23),

$$\begin{aligned}&\big(\boldsymbol{I}-\boldsymbol{\phi}_{\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{k}),\boldsymbol{\gamma}+\boldsymbol{k})\big)\boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\gamma}+\boldsymbol{k})\boldsymbol{h}-\big(\boldsymbol{I}-\boldsymbol{\phi}_{\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})\big)\boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\gamma})\boldsymbol{h}\\ &\quad= \boldsymbol{\phi}_{\boldsymbol{\gamma}}(\boldsymbol{\psi}(\boldsymbol{\gamma}+\boldsymbol{k}),\boldsymbol{\gamma}+\boldsymbol{k})\boldsymbol{h}-\boldsymbol{\phi}_{\boldsymbol{\gamma}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})\boldsymbol{h},\end{aligned} \tag{3.24}$$

and, since $\boldsymbol{\phi}\in C^2$ and $\boldsymbol{\psi}\in C^1$,

$$\begin{aligned}&\big(\boldsymbol{I}-\boldsymbol{\phi}_{\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})\big)\big(\boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\gamma}+\boldsymbol{k})\boldsymbol{h}-\boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\gamma})\boldsymbol{h}\big)\\ &\quad= \boldsymbol{\phi}_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})[\boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\gamma})\boldsymbol{k},\boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\gamma})\boldsymbol{h}]+\boldsymbol{\phi}_{\boldsymbol{x}\boldsymbol{\gamma}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})[\boldsymbol{k},\boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\gamma})\boldsymbol{h}]\\ &\qquad+\boldsymbol{\phi}_{\boldsymbol{\gamma}\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})[\boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\gamma})\boldsymbol{k},\boldsymbol{h}]+\boldsymbol{\phi}_{\boldsymbol{\gamma}\boldsymbol{\gamma}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})[\boldsymbol{k},\boldsymbol{h}]+o(\|\boldsymbol{k}\|),\end{aligned} \tag{3.25}$$

as $\|\boldsymbol{k}\|\to 0$. It follows that the second derivative $\boldsymbol{\psi}_{\boldsymbol{\gamma}\boldsymbol{\gamma}}$ exists and is defined by

$$\begin{aligned}\boldsymbol{\psi}_{\boldsymbol{\gamma}\boldsymbol{\gamma}}(\boldsymbol{\gamma})[\boldsymbol{k},\boldsymbol{h}] &= \big(\boldsymbol{I}-\boldsymbol{\phi}_{\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})\big)^{-1}\Big(\boldsymbol{\phi}_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})[\boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\gamma})\boldsymbol{k},\boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\gamma})\boldsymbol{h}]\\ &\quad+\boldsymbol{\phi}_{\boldsymbol{x}\boldsymbol{\gamma}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})[\boldsymbol{k},\boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\gamma})\boldsymbol{h}]+\boldsymbol{\phi}_{\boldsymbol{\gamma}\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})[\boldsymbol{\psi}_{\boldsymbol{\gamma}}(\boldsymbol{\gamma})\boldsymbol{k},\boldsymbol{h}]\\ &\quad+\boldsymbol{\phi}_{\boldsymbol{\gamma}\boldsymbol{\gamma}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})[\boldsymbol{k},\boldsymbol{h}]\Big).\end{aligned} \tag{3.26}$$

Moreover, $\boldsymbol{\psi}_{\boldsymbol{\gamma}\boldsymbol{\gamma}}(\boldsymbol{\gamma})$ is a continuous function of $\boldsymbol{\gamma}$, so that $\boldsymbol{\psi}\in C^2$.

A succession of similar arguments shows that $\boldsymbol{\psi}\in C^3$, $\boldsymbol{\psi}\in C^4$, and so on. The key is that, at each step, the invertible operator $\boldsymbol{I}-\boldsymbol{\phi}_{\boldsymbol{x}}(\boldsymbol{\psi}(\boldsymbol{\gamma}),\boldsymbol{\gamma})$ appears in the same manner in the expression containing $\boldsymbol{\psi}_{\boldsymbol{\gamma}\cdots\boldsymbol{\gamma}}(\boldsymbol{\gamma})$. ■

In the following, we assume $\boldsymbol{x}\in\mathcal{B}_{2R}$, $\boldsymbol{y}\in\mathcal{B}_R$ and $\epsilon\in(-\epsilon_R,\epsilon_R)$. Since $\boldsymbol{x}=(\boldsymbol{q},\boldsymbol{p})$, setting $\boldsymbol{\psi}=(\boldsymbol{\psi}_1,\boldsymbol{\psi}_2)$, we can rewrite the explicit transformation (3.10) as

$$\boldsymbol{q}=\boldsymbol{\psi}_1(\boldsymbol{y},s,\sigma,\epsilon),\quad \boldsymbol{p}=\boldsymbol{\psi}_2(\boldsymbol{y},s,\sigma,\epsilon). \tag{3.27}$$

The function defined by

$$\begin{aligned}\mathcal{H}(\boldsymbol{y},s,\sigma,\epsilon) &= H\big(\boldsymbol{\psi}_1(\boldsymbol{y},s,\sigma,\epsilon),\boldsymbol{\psi}_2(\boldsymbol{y},s,\sigma,\epsilon),s,\sigma,\epsilon\big)\\ &\quad-\sum_{i=1}^{2k+3}\epsilon^i\,W_s^{(i)}\big(\boldsymbol{Q},\boldsymbol{\psi}_2(\boldsymbol{y},s,\sigma,\epsilon),s,\sigma\big)\\ &\quad-\sum_{i=1}^{2k+3}\epsilon^{i-1}\,W_\sigma^{(i)}\big(\boldsymbol{Q},\boldsymbol{\psi}_2(\boldsymbol{y},s,\sigma,\epsilon),s,\sigma\big),\end{aligned} \tag{3.28}$$

where $\boldsymbol{y} = (\boldsymbol{Q}, \boldsymbol{P})$, will play a key role in section 3.2. This function is smooth and T-periodic in σ. We proceed by showing that it is possible to remove the σ-dependence in $\mathcal{H}(\boldsymbol{y}, s, \sigma, \epsilon)$ up to some order in ϵ by a suitable choice of the functions $W^{(i)}$ generating transformation (3.7). By Taylor's theorem (cf. appendix A), $\boldsymbol{\psi}_1$ and $\boldsymbol{\psi}_2$ can be expanded in powers of ϵ as

$$\begin{aligned} \boldsymbol{\psi}_1(\boldsymbol{y}, s, \sigma, \epsilon) &= \sum_{i=0}^{2k+2} \epsilon^i \, \boldsymbol{u}_i(\boldsymbol{y}, s, \sigma) + O(\epsilon^{2k+3}), \\ \boldsymbol{\psi}_2(\boldsymbol{y}, s, \sigma, \epsilon) &= \sum_{i=0}^{2k+2} \epsilon^i \, \boldsymbol{v}_i(\boldsymbol{y}, s, \sigma) + O(\epsilon^{2k+3}). \end{aligned} \tag{3.29}$$

To find $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$, substitute the expansions given by (3.27) and (3.29) for $\boldsymbol{q}$ and $\boldsymbol{p}$ into equation (3.7), and expand in powers of ϵ. We obtain, for example,

$$\begin{aligned} &\boldsymbol{u}_0 = \boldsymbol{Q}, \quad \boldsymbol{u}_1 = W^{(1)}_{\boldsymbol{p}}, \quad \boldsymbol{u}_2 = -W^{(1)}_{\boldsymbol{pp}} W^{(1)}_{\boldsymbol{Q}} + W^{(2)}_{\boldsymbol{p}}, \\ &\boldsymbol{v}_0 = \boldsymbol{P}, \quad \boldsymbol{v}_1 = -W^{(1)}_{\boldsymbol{Q}}, \quad \boldsymbol{v}_2 = W^{(1)}_{\boldsymbol{Qp}} W^{(1)}_{\boldsymbol{Q}} - W^{(2)}_{\boldsymbol{Q}}. \end{aligned} \tag{3.30}$$

Similarly, by Taylor's theorem (cf. appendix A), $\mathcal{H}$ can be expanded as

$$\mathcal{H}(\boldsymbol{y}, s, \sigma, \epsilon) = \sum_{i=0}^{2k+2} \epsilon^i \mathcal{F}^{(i)}(\boldsymbol{y}, s, \sigma) + \mathcal{R}(\boldsymbol{y}, s, \sigma, \epsilon), \tag{3.31}$$

where

$$\mathcal{R}(\boldsymbol{y}, s, \sigma, \epsilon) = \frac{\epsilon^{2k+3}}{(2k+2)!} \int_0^1 (1-\tau)^{2k+2} \mathcal{H}_{\epsilon^{2k+3}}(\boldsymbol{y}, s, \sigma, \tau\epsilon) \, d\tau, \tag{3.32}$$

and $\mathcal{H}_{\epsilon^{2k+3}}$ denotes the $(2k+3)^{\text{rd}}$ partial derivative of $\mathcal{H}$ with respect to ϵ. The functions $\mathcal{F}^{(i)}$ are obtained by inserting (3.29) into (3.28), expanding in powers of ϵ, and comparing with (3.31). It is convenient to introduce the functions $\mathcal{H}^{(i)}$, $i = 0, \ldots, 2k+2$, defined by the decompositions

$$\mathcal{F}^{(i)}(\boldsymbol{y}, s, \sigma) = \mathcal{H}^{(i)}(\boldsymbol{y}, s, \sigma) - W^{(i+1)}_{\sigma}(\boldsymbol{y}, s, \sigma). \tag{3.33}$$

Computing $\mathcal{F}^{(0)}$ and $\mathcal{F}^{(1)}$, for example, we deduce the equalities

$$\mathcal{H}^{(0)}(\boldsymbol{y}, s, \sigma) = H(\boldsymbol{y}, s, \sigma, 0), \tag{3.34}$$

$$\begin{aligned} \mathcal{H}^{(1)}(\boldsymbol{y}, s, \sigma) = {} & H_{\boldsymbol{q}}(\boldsymbol{y}, s, \sigma, 0)^T W^{(1)}_{\boldsymbol{p}}(\boldsymbol{y}, s, \sigma) - H_{\boldsymbol{p}}(\boldsymbol{y}, s, \sigma, 0)^T W^{(1)}_{\boldsymbol{Q}}(\boldsymbol{y}, s, \sigma) \\ & + W^{(1)}_{\sigma\boldsymbol{p}}(\boldsymbol{y}, s, \sigma)^T W^{(1)}_{\boldsymbol{Q}}(\boldsymbol{y}, s, \sigma) + H_{\epsilon}(\boldsymbol{y}, s, \sigma, 0) - W^{(1)}_{s}(\boldsymbol{y}, s, \sigma). \end{aligned} \tag{3.35}$$

In general, $\mathcal{H}^{(i)}$ depends on the functions $H, W^{(1)}, W^{(2)}, \ldots, W^{(i)}$.

Given a function of $\boldsymbol{y}$, s and σ, with a T-periodic dependence in σ, denote its mean over σ by $\langle \cdot \rangle$. For example,

$$\langle \mathcal{H}^{(i)} \rangle(\boldsymbol{y}, s) = \frac{1}{T} \int_0^T \mathcal{H}^{(i)}(\boldsymbol{y}, s, \sigma) \, d\sigma. \tag{3.36}$$

The σ-dependence in $\mathcal{H}$ can be removed up to some order in ϵ by imposing the relation

$$W^{(i+1)}(\boldsymbol{y}, s, \sigma) = \int_0^\sigma \Big(\mathcal{H}^{(i)}(\boldsymbol{y}, s, t) - \langle \mathcal{H}^{(i)} \rangle (\boldsymbol{y}, s) \Big)\, dt, \tag{3.37}$$

for $i = 0, \ldots, 2k+2$. Indeed, combining this equation with (3.33) and (3.31) yields

$$\mathcal{H}(\boldsymbol{y}, s, \sigma, \epsilon) = \sum_{i=0}^{2k+2} \epsilon^i \langle \mathcal{H}^{(i)} \rangle (\boldsymbol{y}, s) + \mathcal{R}(\boldsymbol{y}, s, \sigma, \epsilon), \tag{3.38}$$

and $\mathcal{H}$ is free of σ-dependence up to $O(\epsilon^{2k+2})$. Relation (3.37) together with the above construction of the quantities $\mathcal{H}^{(i)}$ constitute a definition for the functions $W^{(i)}$ and $\mathcal{H}^{(i)}$. Indeed, $\mathcal{H}^{(0)}, W^{(1)}, \mathcal{H}^{(1)}, W^{(2)}, \ldots, \mathcal{H}^{(2k+2)}, W^{(2k+3)}$ can successively be computed explicitly in the following manner. $\mathcal{H}^{(0)}$ is given by (3.34). $W^{(1)}$ is then obtained from (3.37):

$$W^{(1)}(\boldsymbol{y}, s, \sigma) = \int_0^\sigma \Big(H(\boldsymbol{y}, s, t, 0) - \langle H \rangle (\boldsymbol{y}, s) \Big)\, dt. \tag{3.39}$$

$\mathcal{H}^{(1)}$ is given by (3.35), which simplifies with the help of (3.39) to

$$\begin{aligned} \mathcal{H}^{(1)}(\boldsymbol{y}, s, \sigma) = {} & H_{\boldsymbol{q}}(\boldsymbol{y}, s, \sigma, 0)^T W^{(1)}_{\boldsymbol{p}}(\boldsymbol{y}, s, \sigma) - \langle H \rangle_{\boldsymbol{p}}(\boldsymbol{y}, s)^T W^{(1)}_{\boldsymbol{Q}}(\boldsymbol{y}, s, \sigma) \\ & + H_\epsilon(\boldsymbol{y}, s, \sigma, 0) - W^{(1)}_s(\boldsymbol{y}, s, \sigma). \end{aligned} \tag{3.40}$$

$W^{(2)}$ is then obtained using (3.37), and so on. Note that the resulting functions $W^{(i)}$ satisfy properties (3.8).

For further reference, we record that, according to (3.34) and (3.40),

$$\langle \mathcal{H}^{(0)} \rangle (\boldsymbol{y}, s) = \langle H \rangle (\boldsymbol{y}, s) \tag{3.41}$$

and

$$\begin{aligned} \langle \mathcal{H}^{(1)} \rangle (\boldsymbol{y}, s) = {} & \langle H_{\boldsymbol{q}}^T W^{(1)}_{\boldsymbol{p}} \rangle (\boldsymbol{y}, s) - \langle H \rangle_{\boldsymbol{p}}^T \langle W^{(1)} \rangle_{\boldsymbol{Q}}(\boldsymbol{y}, s) \\ & + \langle H \rangle_\epsilon (\boldsymbol{y}, s) - \langle W^{(1)} \rangle_s (\boldsymbol{y}, s). \end{aligned} \tag{3.42}$$

Extending the above computations to the next order in ϵ shows that

$$\begin{aligned} \langle \mathcal{H}^{(2)} \rangle = {} & \langle \mathcal{H}^{(1)T}_{\boldsymbol{p}} W^{(1)}_{\boldsymbol{Q}} \rangle + \tfrac{1}{2} \langle W^{(1)T}_{\boldsymbol{p}} H_{\boldsymbol{qq}} W^{(1)}_{\boldsymbol{p}} \rangle - \tfrac{1}{2} \langle W^{(1)T}_{\boldsymbol{Q}} \langle H \rangle_{\boldsymbol{pp}} W^{(1)}_{\boldsymbol{Q}} \rangle \\ & + \langle H^T_{\boldsymbol{q}\epsilon} W^{(1)}_{\boldsymbol{p}} \rangle + \langle H^T_{\boldsymbol{q}} W^{(2)}_{\boldsymbol{p}} \rangle - \langle H \rangle^T_{\boldsymbol{p}} \langle W^{(2)} \rangle_{\boldsymbol{Q}} + \tfrac{1}{2} \langle H \rangle_{\epsilon\epsilon} - \langle W^{(2)} \rangle_s. \end{aligned} \tag{3.43}$$

3.2 Near-identity canonical transformation

In this section, we exploit the results of section 3.1 to exhibit a near-identity canonical change of variable for problem (3.1), with the property that the transformed problem is smooth enough for the theorems of chapter 2 to apply.

Consider the transformation $\boldsymbol{x} = (\boldsymbol{q}, \boldsymbol{p}) \rightarrow \boldsymbol{y} = (\boldsymbol{Q}, \boldsymbol{P})$ defined implicitly by

$$\boldsymbol{q} = \boldsymbol{Q} + \sum_{i=1}^{2k+3} \epsilon^i\, W^{(i)}_{\boldsymbol{p}}(\boldsymbol{Q}, \boldsymbol{p}, s, \tfrac{s}{\epsilon}), \quad \boldsymbol{p} = \boldsymbol{P} - \sum_{i=1}^{2k+3} \epsilon^i\, W^{(i)}_{\boldsymbol{Q}}(\boldsymbol{Q}, \boldsymbol{p}, s, \tfrac{s}{\epsilon}). \tag{3.44}$$

This change of variable depends on s and ϵ, and can be obtained by setting $\sigma = \frac{s}{\epsilon}$ in transformation (3.7). By lemma 3.1, for $\epsilon \in (\epsilon_R, \epsilon_R)\backslash\{0\}$, this transformation defines an isomorphism between $\mathcal{B}_R$ in $\boldsymbol{y}$-space and $\boldsymbol{\psi}(\mathcal{B}_R, s, \frac{s}{\epsilon}, \epsilon)$ in $\boldsymbol{x}$-space, and therefore constitutes a valid change of variable if $\boldsymbol{y}$ and $\boldsymbol{x}$ are restricted to these sets.

By periodicity with respect to the last variable, the functions $W^{(i)}_{\boldsymbol{p}}(\boldsymbol{Q}, \boldsymbol{p}, s, \frac{s}{\epsilon})$ and $W^{(i)}_{\boldsymbol{Q}}(\boldsymbol{Q}, \boldsymbol{p}, s, \frac{s}{\epsilon})$ stay bounded as $\epsilon \to 0$. Thus transformation (3.44) is close to the identity when ϵ is small, and is called a *near-identity transformation.*

Change of variable (3.44) is also a *canonical transformation* (cf. Goldstein [22, sect. 9-1]) since it can be written in the form

$$\boldsymbol{q} = W_{\boldsymbol{p}}(\boldsymbol{Q}, \boldsymbol{p}, s, \epsilon), \quad \boldsymbol{P} = W_{\boldsymbol{Q}}(\boldsymbol{Q}, \boldsymbol{p}, s, \epsilon), \tag{3.45}$$

where the generating function is

$$W\big(\boldsymbol{Q}, \boldsymbol{p}, s, \epsilon\big) = \boldsymbol{Q}\boldsymbol{p} + \sum_{i=1}^{2k+3} \epsilon^i \, W^{(i)}\big(\boldsymbol{Q}, \boldsymbol{p}, s, \tfrac{s}{\epsilon}\big). \tag{3.46}$$

Then, according to the standard theory of canonical transformations (cf. Goldstein [22, sect. 9-1]), the Hamiltonian structure of the differential equation in (3.1) is conserved under change of variable (3.44), and the new Hamiltonian is given by $H - W_s$. The definition (3.28) of the function $\mathcal{H}$ has been chosen such that

$$\mathcal{H}\big(\boldsymbol{Q}, \boldsymbol{P}, s, \tfrac{s}{\epsilon}, \epsilon\big) = H\big(\boldsymbol{q}, \boldsymbol{p}, s, \tfrac{s}{\epsilon}, \epsilon\big) - W_s(\boldsymbol{Q}, \boldsymbol{p}, s, \epsilon), \tag{3.47}$$

so that $\mathcal{H}(\boldsymbol{y}, s, \frac{s}{\epsilon}, \epsilon)$ corresponds to the new Hamiltonian. We conclude that the differential equation in (3.1) becomes

$$\boldsymbol{y}'(s) = \boldsymbol{J}\mathcal{H}_{\boldsymbol{y}}\big(\boldsymbol{y}(s), s, \tfrac{s}{\epsilon}, \epsilon\big). \tag{3.48}$$

in the new variables.

By (3.37), change of variable (3.44) reduces to the identity when $s = 0$, i.e. $\boldsymbol{x}(0) = \boldsymbol{y}(0)$. If $\epsilon \in \mathbb{E}_T \cap (-\epsilon_R, \epsilon_R)$, where $\mathbb{E}_T$ is defined by (3.6), then, by T-periodicity of W with respect to the last variable and definition (3.37),

$$W^{(i)}\big(\boldsymbol{y}(1), 1, \tfrac{1}{\epsilon}\big) = W^{(i)}\big(\boldsymbol{y}(1), 1, 0\big) = 0, \tag{3.49}$$

so that change of variable (3.44) also reduces to the identity when $s = 1$, i.e. $\boldsymbol{x}(1) = \boldsymbol{y}(1)$. The next lemma summarizes the results we have obtained.

Lemma 3.2 *Under the above hypotheses, and for $\epsilon \in \mathbb{E}_T \cap (-\epsilon_R, \epsilon_R)$, BVP (3.1) becomes*

$$\begin{cases} \boldsymbol{y}'(s) = \boldsymbol{J}\mathcal{H}_{\boldsymbol{y}}\big(\boldsymbol{y}(s), s, \tfrac{s}{\epsilon}, \epsilon\big), \\ \boldsymbol{b}(\boldsymbol{y}(0)) = \boldsymbol{0}, \quad \boldsymbol{c}(\boldsymbol{y}(1)) = \boldsymbol{0}. \end{cases} \tag{3.50}$$

in the new variables defined by transformation (3.44).

We now extend problem (3.50) to all $\epsilon \in (-\epsilon_R, \epsilon_R)$ by considering the BVP

$$\begin{cases} \boldsymbol{y}'(s) = \boldsymbol{J}\bar{\mathcal{H}}_{\boldsymbol{y}}(\boldsymbol{y}(s), s, \epsilon), \\ \boldsymbol{b}(\boldsymbol{y}(0)) = \boldsymbol{0}, \quad \boldsymbol{c}(\boldsymbol{y}(1)) = \boldsymbol{0}, \end{cases} \tag{3.51}$$

where $\bar{\mathcal{H}}$ is defined by

$$\bar{\mathcal{H}}(\boldsymbol{y}, s, \epsilon) = \begin{cases} \mathcal{H}\big(\boldsymbol{y}, s, \frac{s}{\epsilon}, \epsilon\big) & \text{if } \epsilon \neq 0, \\ \langle H\rangle(\boldsymbol{y}, s) & \text{if } \epsilon = 0. \end{cases} \tag{3.52}$$

Lemma 3.3 *Under the above assumptions,*

$$\bar{\mathcal{H}} \in C^{k+1}\big(\mathcal{B}_R \times [0,1] \times (-\epsilon_R, \epsilon_R), \mathbb{R}\big). \tag{3.53}$$

Proof We start by proving that $\bar{\mathcal{H}}$ is C^1 on $\mathcal{B}_R \times [0,1] \times (-\epsilon_R, \epsilon_R)$. From the smoothness of $\mathcal{H}$ and H, $\bar{\mathcal{H}}$ is smooth both on $\mathcal{B}_R \times [0,1] \times (-\epsilon_R, \epsilon_R)\backslash\{0\}$ and on $\mathcal{B}_R \times [0,1] \times \{0\}$, so that it suffices to show that, for any $\boldsymbol{y} \in \mathcal{B}_R$, $s \in [0,1]$,

(i) the derivative $\bar{\mathcal{H}}_{(\boldsymbol{y},s,\epsilon)}(\boldsymbol{y}, s, 0)$ exists,

(ii) $\lim_{\epsilon\to 0} \bar{\mathcal{H}}_{(\boldsymbol{y},s,\epsilon)}(\boldsymbol{y}, s, \epsilon) = \bar{\mathcal{H}}_{(\boldsymbol{y},s,\epsilon)}(\boldsymbol{y}, s, 0)$.

Let us establish (i). For $\tilde{\epsilon} \neq 0$, we compute

$$\begin{aligned} &\bar{\mathcal{H}}(\boldsymbol{y}+\tilde{\boldsymbol{y}}, s+\tilde{s}, \tilde{\epsilon}) - \bar{\mathcal{H}}(\boldsymbol{y}, s, 0) \\ &\quad = \langle H\rangle(\boldsymbol{y}+\tilde{\boldsymbol{y}}, s+\tilde{s}) - \langle H\rangle(\boldsymbol{y}, s) + \sum_{i=1}^{2k+2} \tilde{\epsilon}^i \langle \mathcal{H}^{(i)}\rangle(\boldsymbol{y}+\tilde{\boldsymbol{y}}, s+\tilde{s}) \\ &\quad\quad + \frac{\tilde{\epsilon}^{2k+3}}{(2k+2)!}\int_0^1 (1-\tau)^{2k+2}\mathcal{H}_{\epsilon^{2k+3}}\big(\boldsymbol{y}+\tilde{\boldsymbol{y}}, s+\tilde{s}, \tfrac{s+\tilde{s}}{\tilde{\epsilon}}, \tau\tilde{\epsilon}\big)\, d\tau \\ &\quad = \langle H\rangle_{\boldsymbol{y}}(\boldsymbol{y}, s)\tilde{\boldsymbol{y}} + \langle H\rangle_s(\boldsymbol{y}, s)\tilde{s} + \langle \mathcal{H}^{(1)}\rangle(\boldsymbol{y}, s)\tilde{\epsilon} + o(\|(\tilde{\boldsymbol{y}}, \tilde{s}, \tilde{\epsilon})\|), \end{aligned} \tag{3.54}$$

where the first equality is a consequence of (3.52), (3.38), (3.32), (3.41), and the second equality follows from the smoothness of H, $\mathcal{H}$, $\mathcal{H}^{(i)}$ and from the periodicity of $\mathcal{H}$ with respect to σ. Similarly, in the case $\tilde{\epsilon} = 0$, we compute

$$\begin{aligned} \bar{\mathcal{H}}(\boldsymbol{y}+\tilde{\boldsymbol{y}}, s+\tilde{s}, 0) - \bar{\mathcal{H}}(\boldsymbol{y}, s, 0) &= \langle H\rangle(\boldsymbol{y}+\tilde{\boldsymbol{y}}, s+\tilde{s}) - \langle H\rangle(\boldsymbol{y}, s) \\ &= \langle H\rangle_{\boldsymbol{y}}(\boldsymbol{y}, s)\tilde{\boldsymbol{y}} + \langle H\rangle_s(\boldsymbol{y}, s)\tilde{s} + o(\|(\tilde{\boldsymbol{y}}, \tilde{s})\|). \end{aligned} \tag{3.55}$$

It follows that $\bar{\mathcal{H}}$ is differentiable at $(\boldsymbol{y}, s, 0)$, and the derivative is expressed by

$$\bar{\mathcal{H}}_{(\boldsymbol{y},s,\epsilon)}(\boldsymbol{y}, s, 0)(\tilde{\boldsymbol{y}}, \tilde{s}, \tilde{\epsilon}) = \langle H\rangle_{\boldsymbol{y}}(\boldsymbol{y}, s)\tilde{\boldsymbol{y}} + \langle H\rangle_s(\boldsymbol{y}, s)\tilde{s} + \langle \mathcal{H}^{(1)}\rangle(\boldsymbol{y}, s)\tilde{\epsilon}. \tag{3.56}$$

Let us prove (ii) now. By (3.52), (3.38) and (3.32), we obtain, for $\epsilon \neq 0$,

$$
\begin{aligned}
&\bar{\mathcal{H}}_{(\boldsymbol{y},s,\epsilon)}(\boldsymbol{y},s,\epsilon)(\tilde{\boldsymbol{y}},\tilde{s},\tilde{\epsilon}) \\
&\quad = \mathcal{H}_{\boldsymbol{y}}\big(\boldsymbol{y},s,\tfrac{s}{\epsilon},\epsilon\big)\tilde{\boldsymbol{y}} + \mathcal{H}_{s}\big(\boldsymbol{y},s,\tfrac{s}{\epsilon},\epsilon\big)\tilde{s} + \mathcal{H}_{\epsilon}\big(\boldsymbol{y},s,\tfrac{s}{\epsilon},\epsilon\big)\tilde{\epsilon} + \mathcal{H}_{\sigma}\big(\boldsymbol{y},s,\tfrac{s}{\epsilon},\epsilon\big)\big(\tfrac{\tilde{s}}{\epsilon} - \tfrac{s\tilde{\epsilon}}{\epsilon^2}\big) \\
&\quad = \sum_{i=0}^{2k+2} \epsilon^i \langle \mathcal{H}^{(i)}\rangle_{\boldsymbol{y}}(\boldsymbol{y},s)\tilde{\boldsymbol{y}} + \frac{\epsilon^{2k+3}}{(2k+2)!}\int_0^1 (1-\tau)^{2k+2}\mathcal{H}_{\epsilon^{2k+3}\boldsymbol{y}}\big(\boldsymbol{y},s,\tfrac{s}{\epsilon},\tau\epsilon\big)\,d\tau\,\tilde{\boldsymbol{y}} \\
&\qquad + \sum_{i=0}^{2k+2} \epsilon^i \langle \mathcal{H}^{(i)}\rangle_{s}(\boldsymbol{y},s)\tilde{s} + \frac{\epsilon^{2k+3}}{(2k+2)!}\int_0^1 (1-\tau)^{2k+2}\mathcal{H}_{\epsilon^{2k+3}s}\big(\boldsymbol{y},s,\tfrac{s}{\epsilon},\tau\epsilon\big)\,d\tau\,\tilde{s} \\
&\qquad + \sum_{i=1}^{2k+2} i\,\epsilon^{i-1} \langle \mathcal{H}^{(i)}\rangle(\boldsymbol{y},s)\tilde{\epsilon} + \frac{\epsilon^{2k+3}}{(2k+2)!}\int_0^1 \tau(1-\tau)^{2k+2}\mathcal{H}_{\epsilon^{2k+4}}\big(\boldsymbol{y},s,\tfrac{s}{\epsilon},\tau\epsilon\big)\,d\tau\,\tilde{\epsilon} \\
&\qquad + \frac{2k+3}{(2k+2)!}\,\epsilon^{2k+2}\int_0^1 (1-\tau)^{2k+2}\mathcal{H}_{\epsilon^{2k+3}}\big(\boldsymbol{y},s,\tfrac{s}{\epsilon},\tau\epsilon\big)\,d\tau\,\tilde{\epsilon} \\
&\qquad + \frac{1}{(2k+2)!}\int_0^1 (1-\tau)^{2k+2}\mathcal{H}_{\epsilon^{2k+3}\sigma}\big(\boldsymbol{y},s,\tfrac{s}{\epsilon},\tau\epsilon\big)\,d\tau\big(\epsilon^{2k+2}\tilde{s} - s\epsilon^{2k+1}\tilde{\epsilon}\big).
\end{aligned}
\tag{3.57}
$$

By smoothness of $\mathcal{H}$ and $\mathcal{H}^{(i)}$, periodicity of $\mathcal{H}$, and equality (3.41), we find

$$
\lim_{\epsilon\to 0}\bar{\mathcal{H}}_{(\boldsymbol{y},s,\epsilon)}(\boldsymbol{y},s,\epsilon)(\tilde{\boldsymbol{y}},\tilde{s},\tilde{\epsilon}) = \langle H\rangle_{\boldsymbol{y}}(\boldsymbol{y},s)\tilde{\boldsymbol{y}} + \langle H\rangle_{s}(\boldsymbol{y},s)\tilde{s} + \langle \mathcal{H}^{(1)}\rangle(\boldsymbol{y},s)\tilde{\epsilon}. \tag{3.58}
$$

A comparison with (3.56) yields equality (ii). This concludes the proof of the fact that $\bar{\mathcal{H}}$ is C^1 on $\mathcal{B}_R\times[0,1]\times(-\epsilon_R,\epsilon_R)$.

A similar argument using the expression of the first derivative of $\bar{\mathcal{H}}$ given by (3.56) and (3.57) yields an expression for the second derivative and shows that $\bar{\mathcal{H}}$ is C^2 on $\mathcal{B}_R\times[0,1]\times(-\epsilon_R,\epsilon_R)$. Repeating this process proves successively that $\bar{\mathcal{H}}$ is $C^3,\dots,C^{k+1}$ on $\mathcal{B}_R\times[0,1]\times(-\epsilon_R,\epsilon_R)$. The key is that, for $l=2,\dots,k+1$, among the terms appearing in the expression of $\bar{\mathcal{H}}_{(\boldsymbol{y},s,\epsilon)^l}(\boldsymbol{y},s,\epsilon)$, with $\epsilon\neq 0$, the ones depending on $\frac{s}{\epsilon}$ are all of the form $\epsilon^j\boldsymbol{r}(\boldsymbol{y},s,\frac{s}{\epsilon},\epsilon)$, where $j\geqslant 1$ and $\boldsymbol{r}(\boldsymbol{y},s,\sigma,\epsilon)$ is a smooth function, T-periodic in σ. These terms therefore converge to zero as $\epsilon\to 0$, and the l^{th} derivative of $\bar{\mathcal{H}}$ is continuous at $(\boldsymbol{y},s,0)$. Note that the exponent j reaches the limit value 1 in the following term that appears in the expression of $\bar{\mathcal{H}}_{(\boldsymbol{y},s,\epsilon)^{k+1}}(\boldsymbol{y},s,\epsilon)$:

$$
\begin{aligned}
&\frac{(-s)^{k+1}}{\epsilon^{2(k+1)}}\mathcal{H}_{\sigma^{k+1}}\big(\boldsymbol{y},s,\tfrac{s}{\epsilon},\epsilon\big)[\tilde{\epsilon}_1,\dots,\tilde{\epsilon}_{k+1}] \\
&\qquad = \epsilon\,\frac{(-s)^{k+1}}{(2k+2)!}\int_0^1 (1-\tau)^{2k+2}\mathcal{H}_{\epsilon^{2k+3}\sigma^{k+1}}\big(\boldsymbol{y},s,\tfrac{s}{\epsilon},\tau\epsilon\big)\,d\tau\,[\tilde{\epsilon}_1,\dots,\tilde{\epsilon}_{k+1}].
\end{aligned}
\tag{3.59}
$$

■

In the next two sections we establish averaging results for BVP (3.1) by applying the results of chapter 2 to BVP (3.51). Note that the $\epsilon = 0$ case of BVP (3.51) corresponds to the averaged BVP (3.2).

3.3 Averaging in the regular case

In this section we assume $\boldsymbol{y}_0$ is an isolated solution to the averaged BVP (3.2), and study the existence of a solution to the two-scale BVP (3.1) in the vicinity of $\boldsymbol{y}_0$ when ϵ belongs to $\mathbb{E}_T$ and is small. We establish the existence of a solution to BVP (3.51) by the regular perturbation theorem 2.2, and then deduce the existence of a solution to BVP (3.1) by the results of section 3.2.

Here we make the same assumptions on the averaged problem (3.2) as are made on the unperturbed problem in section 2.2. Suppose

$$\boldsymbol{y}_0 \text{ is a solution to the averaged BVP (3.2),} \tag{\textbf{3.A2}}$$

and consider the corresponding homogeneous linearized BVP

$$\begin{cases} \boldsymbol{v}'(s) - \boldsymbol{J}\langle H\rangle_{\boldsymbol{yy}}(\boldsymbol{y}_0(s),s)\boldsymbol{v}(s) = \boldsymbol{0}, \\ \boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}_0(0))\boldsymbol{v}(0) = \boldsymbol{0}, \quad \boldsymbol{c}_{\boldsymbol{y}}(\boldsymbol{y}_0(1))\boldsymbol{v}(1) = \boldsymbol{0}. \end{cases} \tag{3.60}$$

Assume moreover

$$\text{BVP (3.60) admits only the trivial solution } \boldsymbol{v} = \boldsymbol{0}. \tag{\textbf{3.A3}}$$

Under this assumption, $\boldsymbol{y}_0$ is an isolated solution to the averaged BVP (3.2).

Theorem 3.4 *Suppose conditions (**3.A1**)-(**3.A3**) hold, and fix $k \geqslant 1$, where k controls the number of terms involved in the near-identity canonical transformation (3.44). Then there exists a positive number ϵ_0 such that, for $\epsilon \in \mathbb{E}_T \cap (-\epsilon_0, \epsilon_0)$, BVP (3.1) admits a solution $\boldsymbol{\xi}(\epsilon)$ expressed by*

$$\boldsymbol{\xi}(\epsilon)(s) = \boldsymbol{\psi}\big(\boldsymbol{\eta}(\epsilon)(s), s, \tfrac{s}{\epsilon}, \epsilon\big), \tag{3.61}$$

where $\boldsymbol{\psi}$ is defined in lemma (3.1), $\boldsymbol{\eta} \in C^k\big((-\epsilon_0,\epsilon_0), \mathcal{C}^1\big)$ is such that $\boldsymbol{\eta}(0) = \boldsymbol{y}_0$ and $\boldsymbol{\eta}(\epsilon)$ is a solution to BVP (3.51).

Proof Theorem 2.2 is a local result and, in hypothesis (**2.A1**), it is in fact sufficient to suppose that the Hamiltonian is defined on $\hat{\mathcal{B}} \times [0,1] \times (-\hat{\epsilon}, \hat{\epsilon})$, where $\hat{\epsilon} > 0$ and $\hat{\mathcal{B}} \subset \mathbb{R}^{2n}$ is a closed ball containing $\boldsymbol{x}_0(s)$, $\forall s \in [0,1]$, in its interior. If we choose the value $2\|\boldsymbol{y}_0\|_{\mathcal{C}^0}$ for the number R introduced in lemma 3.1 and set $\hat{\mathcal{B}} = \mathcal{B}_R$ and $\hat{\epsilon} = \epsilon_R$, then, by lemma 3.3 and assumptions (**3.A1**)-(**3.A3**), theorem 2.2 applies on BVP (3.51). Hence there exist ϵ_0 satisfying $0 < \epsilon_0 \leqslant \epsilon_R$ and $\boldsymbol{\eta} \in C^k\big((-\epsilon_0,\epsilon_0), \mathcal{C}^1\big)$ such that $\boldsymbol{\eta}(0) = \boldsymbol{y}_0$ and, for $\epsilon \in (-\epsilon_0, \epsilon_0)$, $\boldsymbol{\eta}(\epsilon)$ is a solution to BVP (3.51) and $\boldsymbol{\eta}(\epsilon)(s) \in \mathcal{B}_R$ for all $s \in [0,1]$. The conclusion of the theorem then follows from lemmas 3.2 and 3.1. ■

By properties of $\boldsymbol{\psi}$ and theorem 3.4, we conclude that

$$\boldsymbol{\xi}(\epsilon)(s) = \boldsymbol{y}_0(s) + O(\epsilon) \;\text{ uniformly in } s \in [0,1], \;\text{ as } \epsilon \to 0, \tag{3.62}$$

and $\boldsymbol{y}_0$ is the leading order approximation to the solution $\boldsymbol{\xi}(\epsilon)$. Under the hypotheses of theorem 3.4, it is possible to compute an expansion of the solution $\boldsymbol{\xi}(\epsilon)$ in powers of ϵ. For that we choose $k \geqslant 2$. The solution $\boldsymbol{\eta}(\epsilon)$ to BVP (3.51) can be expanded as

$$\boldsymbol{\eta}(\epsilon)(s) = \boldsymbol{y}_0(s) + \epsilon\boldsymbol{y}_1(s) + \ldots + \epsilon^{k-1}\boldsymbol{y}_{k-1}(s) + O(\epsilon^k), \tag{3.63}$$

where, for $i = 1, \ldots, k-1$, $\boldsymbol{y}_i$ is the unique solution to the linear BVP obtained by substituting (3.63) into BVP (3.51), expanding in powers of ϵ, and equating coefficients of ϵ^i. For example, $\boldsymbol{y}_1$ is solution to

$$\begin{cases} \boldsymbol{y}_1'(s) - \boldsymbol{J}\langle H\rangle_{\boldsymbol{yy}}(\boldsymbol{y}_0(s), s)\boldsymbol{y}_1(s) = \boldsymbol{J}\langle \mathcal{H}^{(1)}\rangle_{\boldsymbol{y}}(\boldsymbol{y}_0(s), s), \\ \boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}_0(0))\boldsymbol{y}_1(0) = \boldsymbol{0}, \quad \boldsymbol{c}_{\boldsymbol{y}}(\boldsymbol{y}_0(1))\boldsymbol{y}_1(1) = \boldsymbol{0}, \end{cases} \tag{3.64}$$

where $\langle \mathcal{H}^{(1)}\rangle$ is given in (3.42). Then, for $\epsilon \in \mathbb{E}_T \cap (-\epsilon_0, \epsilon_0)$, the solution $\boldsymbol{\xi}(\epsilon)$ to BVP (3.1) can be expanded as

$$\boldsymbol{\xi}(\epsilon)(s) = \boldsymbol{y}_0(s) + \epsilon \boldsymbol{x}_1\big(s, \tfrac{s}{\epsilon}\big) + \ldots + \epsilon^{k-1}\boldsymbol{x}_{k-1}\big(s, \tfrac{s}{\epsilon}\big) + O(\epsilon^k) \tag{3.65}$$

by substituting (3.63), (3.27) and (3.29) into (3.61). For example, the first correction is

$$\boldsymbol{x}_1\big(s, \tfrac{s}{\epsilon}\big) = \boldsymbol{y}_1(s) + \boldsymbol{J}W_{\boldsymbol{y}}^{(1)}\big(\boldsymbol{y}_0(s), s, \tfrac{s}{\epsilon}\big), \tag{3.66}$$

where $W^{(1)}$ is explicitly given by (3.39). For $i = 1, \ldots, k-1$, the function $\boldsymbol{x}_i\big(s, \sigma\big)$ has a T-periodic dependence in σ, and $\boldsymbol{x}_i\big(s, \tfrac{s}{\epsilon}\big) = O(1)$ as $\epsilon \to 0$. Note that, in general, the s-derivative of the solution $\boldsymbol{\xi}(\epsilon)(s)$ does not have a limit as $\epsilon \to 0$. Therefore, $\boldsymbol{\xi}(\epsilon) \to \boldsymbol{y}_0$ as $\epsilon \to 0$ in the $\mathcal{C}^0$-norm, but not in the $\mathcal{C}^1$-norm.

3.4 Averaging and symmetry breaking

In this section we assume the boundary conditions of BVP (3.1) define a Lagrangian manifold at each end-point, and the averaged BVP (3.2) admits a smooth curve Γ of solutions. We show that the elements of Γ which give rise to a solution to the two-scale BVP (3.1) when ϵ belongs to $\mathbb{E}_T$ and is small can be identified by means of a Melnikov function. These results are obtained by applying the symmetry breaking theorem 2.10 on BVP (3.51), and then using the results of section 3.2.

Here we make the same assumptions on the averaged problem (3.2) as are made on the unperturbed problem in section 2.3. Consider the sets

$$L_0 = \left\{\boldsymbol{\mu} \in \mathbb{R}^{2n} : \boldsymbol{b}(\boldsymbol{\mu}) = \boldsymbol{0}\right\}, \qquad L_1 = \left\{\boldsymbol{\rho} \in \mathbb{R}^{2n} : \boldsymbol{c}(\boldsymbol{\rho}) = \boldsymbol{0}\right\}, \tag{3.67}$$

and assume the functions $\boldsymbol{b}$ and $\boldsymbol{c}$ are such that

$$\begin{aligned} &\forall \boldsymbol{\mu} \in L_0, \quad \boldsymbol{\nu}_1, \boldsymbol{\nu}_2 \in \operatorname{Ker} \boldsymbol{b}_{\boldsymbol{x}}(\boldsymbol{\mu}) \Longrightarrow \boldsymbol{\nu}_1^T \boldsymbol{J}\boldsymbol{\nu}_2 = 0, \\ &\forall \boldsymbol{\rho} \in L_1, \quad \boldsymbol{\nu}_1, \boldsymbol{\nu}_2 \in \operatorname{Ker} \boldsymbol{c}_{\boldsymbol{x}}(\boldsymbol{\rho}) \Longrightarrow \boldsymbol{\nu}_1^T \boldsymbol{J}\boldsymbol{\nu}_2 = 0. \end{aligned} \tag{\textbf{3.A4}}$$

By lemma 2.5, this condition implies that L_0 and L_1 are Lagrangian submanifolds of $\mathbb{R}^{2n}$. Let I be an open interval in $\mathbb{R}$ and $\theta^* \in I$. Suppose

$$\begin{aligned} &\text{the averaged BVP (3.2) has a family of solutions } \Gamma = \boldsymbol{\zeta}(I), \\ &\text{where } \boldsymbol{\zeta} : I \to \mathcal{C}^1 \text{ is smooth and } \boldsymbol{\zeta}_\theta(\theta^*) \neq \boldsymbol{0}. \end{aligned} \tag{\textbf{3.A5}}$$

We will use the notation

$$\boldsymbol{y}^* = \boldsymbol{\zeta}(\theta^*) \quad \text{and} \quad \boldsymbol{y}_\theta^* = \boldsymbol{\zeta}_\theta(\theta^*). \tag{3.68}$$

It follows from assumption (**3.A5**) that $\boldsymbol{y}_\theta^*$ is a solution to the homogeneous linear BVP

$$\begin{cases} \boldsymbol{v}'(s) - \boldsymbol{J}\langle H\rangle_{\boldsymbol{yy}}(\boldsymbol{y}^*(s),s)\boldsymbol{v}(s) = \boldsymbol{0}, \\ \boldsymbol{b_y}(\boldsymbol{y}^*(0))\boldsymbol{v}(0) = \boldsymbol{0}, \quad \boldsymbol{c_y}(\boldsymbol{y}^*(1))\boldsymbol{v}(1) = \boldsymbol{0}. \end{cases} \tag{3.69}$$

In fact we suppose

$$\text{the set of solutions to BVP (3.69) is } \mathrm{Span}\{\boldsymbol{y}_\theta^*\}. \tag{\textbf{3.A6}}$$

Under this assumption, and in a neighborhood of $\boldsymbol{y}^*$, Γ constitutes an isolated curve of solutions to the averaged problem (3.2).

From (3.56), $\bar{\mathcal{H}}_\epsilon(\boldsymbol{y},s,0) = \langle\mathcal{H}^{(1)}\rangle(\boldsymbol{y},s)$, so that, according to equation (2.85), the Melnikov function associated with BVP (3.51) is

$$\mathcal{M}(\theta) = \frac{d}{d\theta}\int_0^1 \langle\mathcal{H}^{(1)}\rangle(\boldsymbol{\zeta}(\theta)(s),s)\,ds, \tag{3.70}$$

where $\langle\mathcal{H}^{(1)}\rangle$ is given in (3.42).

Theorem 3.5 *Suppose conditions (**3.A1**), (**3.A4**)-(**3.A6**) hold, fix $k \geqslant 2$, where k controls the number of terms involved in the near-identity canonical transformation (3.44), and consider the Melnikov function $\mathcal{M}(\theta)$ defined by (3.70).*

(a) If $\mathcal{M}(\theta^*) \neq 0$, then there exist a positive number ϵ_0 and a $\mathcal{C}^0$-neighborhood $\mathcal{X}$ of $\boldsymbol{y}^*$ such that, for $\epsilon \in \mathbb{E}_T \cap (-\epsilon_0,\epsilon_0)$, BVP (3.1) does not admit any solution in $\mathcal{X}$.

(b) If $\mathcal{M}(\theta^*) = 0$ and $\mathcal{M}_\theta(\theta^*) \neq 0$, then there exists a positive number ϵ_0 such that, for $\epsilon \in \mathbb{E}_T \cap (-\epsilon_0,\epsilon_0)$, BVP (3.1) admits a solution $\boldsymbol{\xi}(\epsilon)$ expressed by

$$\boldsymbol{\xi}(\epsilon)(s) = \boldsymbol{\psi}\big(\boldsymbol{\eta}(\epsilon)(s), s, \tfrac{s}{\epsilon}, \epsilon\big), \tag{3.71}$$

where $\boldsymbol{\psi}$ is defined in lemma (3.1), $\boldsymbol{\eta} \in C^{k-1}\big((-\epsilon_0,\epsilon_0),\mathcal{C}^1\big)$ is such that $\boldsymbol{\eta}(0) = \boldsymbol{y}^*$ and $\boldsymbol{\eta}(\epsilon)$ is a solution to BVP (3.51).

Proof Theorem 2.10 is a local result and, in hypothesis (**2.A4**), it is in fact sufficient to suppose that the Hamiltonian is defined on $\hat{\mathcal{B}} \times [0,1] \times (-\hat{\epsilon},\hat{\epsilon})$, where $\hat{\epsilon} > 0$ and $\hat{\mathcal{B}} \subset \mathbb{R}^{2n}$ is a closed ball containing $\boldsymbol{x}^*(s)$, $\forall s \in [0,1]$, in its interior. If we choose the value $2\|\boldsymbol{y}_0\|_{\mathcal{C}^0}$ for the number R introduced in lemma 3.1 and set $\hat{\mathcal{B}} = \mathcal{B}_R$ and $\hat{\epsilon} = \epsilon_R$, then, by lemma 3.3 and assumptions (**3.A1**), (**3.A4**)-(**3.A6**), theorem 2.10 applies on BVP (3.51).

In case (a) of the theorem, it follows from theorem 2.10 that there exist ϵ_2 and δ_2 satisfying $0 < \epsilon_2 \leqslant \epsilon_R$ and $0 < \delta_2 \leqslant \frac{R}{2}$ such that

$$\begin{aligned} &\text{for } \epsilon \in (-\epsilon_2,\epsilon_2)\backslash\{0\}, \text{ BVP (3.51) does not admit} \\ &\text{any solution } \boldsymbol{y} \text{ satisfying } \|\boldsymbol{y}-\boldsymbol{y}^*\|_{\mathcal{C}^1} \leqslant \delta_2. \end{aligned} \tag{3.72}$$

Define the numbers

$$L = \max_{\substack{\boldsymbol{z}\in\mathbb{R}^{2n}\,:\,\|\boldsymbol{z}\|\leqslant\delta_2, \\ s\in[0,1],\ \epsilon\in[-\epsilon_2,\epsilon_2]}} \big\|\bar{\mathcal{H}}_{\boldsymbol{y}(\boldsymbol{y},\epsilon)}(\boldsymbol{y}^*(s)+\boldsymbol{z},s,\epsilon)\big\| < \infty, \tag{3.73}$$

$$\epsilon_1 = \min\left(\epsilon_2, \frac{\delta_2}{2L}\right), \quad \delta_1 = \min\left(\delta_2, \frac{\delta_2}{2L}\right). \tag{3.74}$$

We claim that

$$\begin{aligned}&\text{for } \epsilon \in (-\epsilon_1, \epsilon_1)\backslash\{0\}, \text{ BVP (3.51) does not admit}\\ &\text{any solution } \boldsymbol{y} \text{ satisfying } \|\boldsymbol{y} - \boldsymbol{y}^*\|_{\mathcal{C}^0} \leqslant \delta_1.\end{aligned} \tag{3.75}$$

Suppose this is false, i.e. there exists a solution $\boldsymbol{y}$ to BVP (3.51) with $\epsilon \in (-\epsilon_1, \epsilon_1)\backslash\{0\}$, such that $\|\boldsymbol{y} - \boldsymbol{y}^*\|_{\mathcal{C}^0} \leqslant \delta_1$. Then, for any $s \in [0,1]$,

$$\begin{aligned}\|\boldsymbol{y}'(s) - \boldsymbol{y}^{*\prime}(s)\| &\leqslant \big\|\boldsymbol{J}\bar{\mathcal{H}}_{\boldsymbol{y}}(\boldsymbol{y}(s), s, \epsilon) - \boldsymbol{J}\bar{\mathcal{H}}_{\boldsymbol{y}}(\boldsymbol{y}^*(s), s, 0)\big\|\\ &\leqslant L\big(\|\boldsymbol{y}(s) - \boldsymbol{y}^*(s)\| + |\epsilon|\big)\\ &\leqslant L(\delta_1 + \epsilon_1),\end{aligned} \tag{3.76}$$

so that

$$\begin{aligned}\|\boldsymbol{y} - \boldsymbol{y}^*\|_{\mathcal{C}^1} &= \max\big(\|\boldsymbol{y} - \boldsymbol{y}^*\|_{\mathcal{C}^0}, \|\boldsymbol{y}' - \boldsymbol{y}^{*\prime}\|_{\mathcal{C}^0}\big)\\ &\leqslant \max\big(\delta_1, L(\delta_1 + \epsilon_1)\big)\\ &\leqslant \delta_2.\end{aligned} \tag{3.77}$$

This contradicts (3.72), and (3.75) is established.

By lemma 3.1, for any $s \in [0,1]$ and $\epsilon \in (-\epsilon_1, \epsilon_1)\backslash\{0\}$, change of variable (3.44) transforms the ball $B_{\delta_1}(s) = \big\{\boldsymbol{z} \in \mathbb{R}^{2n} : \|\boldsymbol{z} - \boldsymbol{y}^*(s)\| \leqslant \delta_1\big\}$ in $\boldsymbol{y}$-space into $\boldsymbol{\psi}\big(B_{\delta_1}(s), s, \frac{s}{\epsilon}, \epsilon\big)$ in $\boldsymbol{x}$-space. Moreover, applying the Banach fixed-point theorem as in the proof of lemma 3.1 shows the existence of ϵ_0 satisfying $0 < \epsilon_0 \leqslant \epsilon_1$ and such that

$$B_{\delta_1/2}(s) \subset \boldsymbol{\psi}\big(B_{\delta_1}(s), s, \tfrac{s}{\epsilon}, \epsilon\big), \quad \forall\, s \in [0,1],\ \epsilon \in (-\epsilon_0, \epsilon_0)\backslash\{0\}. \tag{3.78}$$

For $\epsilon \in \mathbb{E}_T \cap (-\epsilon_0, \epsilon_0)$, we have the following. By (3.75), BVP (3.51) does not admit any solution $\boldsymbol{y}$ satisfying $\boldsymbol{y}(s) \in B_{\delta_1}(s)$ for all $s \in [0,1]$ so that, by (3.78) and lemma 3.2, BVP (3.1) does not admit any solution $\boldsymbol{x}$ satisfying $\boldsymbol{x}(s) \in B_{\delta_1/2}(s)$ for all $s \in [0,1]$. Finally, setting $\mathcal{X} = \big\{\boldsymbol{x} \in \mathcal{C}^1 : \|\boldsymbol{x} - \boldsymbol{y}^*\|_{\mathcal{C}^0} \leqslant \frac{\delta_1}{2}\big\}$ establishes the conclusion of part (a) of the theorem.

In case (b) of the theorem, it follows from theorem 2.10 that there exist ϵ_0 satisfying $0 < \epsilon_0 \leqslant \epsilon_R$ and $\boldsymbol{\eta} \in C^{k-1}\big((-\epsilon_0, \epsilon_0), \mathcal{C}^1\big)$ such that $\boldsymbol{\eta}(0) = \boldsymbol{y}^*$ and, for $\epsilon \in (-\epsilon_0, \epsilon_0)$, $\boldsymbol{\eta}(\epsilon)$ is a solution to BVP (3.51) and $\boldsymbol{\eta}(\epsilon)(s) \in \mathcal{B}_R$ for all $s \in [0,1]$. The conclusion of part (b) of the theorem then follows from lemmas 3.2 and 3.1. ∎

Suppose the assumptions of part (b) of theorem 3.5 are satisfied. Then, by properties of $\boldsymbol{\psi}$,

$$\boldsymbol{\xi}(\epsilon)(s) = \boldsymbol{y}^*(s) + O(\epsilon) \ \text{ uniformly in } s \in [0,1], \ \text{ as } \epsilon \to 0, \tag{3.79}$$

and $\boldsymbol{y}^*$ is the leading order approximation to the solution $\boldsymbol{\xi}(\epsilon)$. Under the assumptions of part (b) of theorem 3.5, it is possible to compute an expansion of the solution $\boldsymbol{\xi}(\epsilon)$ in powers of ϵ. For that we choose $k \geqslant 4$. The solution $\boldsymbol{\eta}(\epsilon)$ to BVP (3.51) admits the expansion

$$\boldsymbol{\eta}(\epsilon)(s) = \boldsymbol{y}^*(s) + \epsilon\boldsymbol{y}_1(s) + \ldots + \epsilon^{k-3}\boldsymbol{y}_{k-3}(s) + O(\epsilon^{k-2}), \tag{3.80}$$

where the functions $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{k-3}$ can be uniquely determined by the perturbation method described in section 2.3. For example, $\boldsymbol{y}_1$ is the unique solution to problem

$$\begin{cases} \boldsymbol{y}_1'(s) - \boldsymbol{J}\langle H\rangle_{\boldsymbol{yy}}(\boldsymbol{y}^*(s), s)\boldsymbol{y}_1(s) = \boldsymbol{J}\langle \mathcal{H}^{(1)}\rangle_{\boldsymbol{y}}(\boldsymbol{y}^*(s), s), \\ \boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}^*(0))\boldsymbol{y}_1(0) = \boldsymbol{0}, \quad \boldsymbol{c}_{\boldsymbol{y}}(\boldsymbol{y}^*(1))\boldsymbol{y}_1(1) = \boldsymbol{0} \end{cases} \tag{3.81}$$

making problem

$$\begin{cases} \boldsymbol{y}_2'(s) - \boldsymbol{J}\langle H\rangle_{\boldsymbol{yy}}(\boldsymbol{y}^*(s), s)\boldsymbol{y}_2(s) = \frac{1}{2}\boldsymbol{J}\langle H\rangle_{\boldsymbol{yyy}}(\boldsymbol{y}^*(s), s)[\boldsymbol{y}_1(s), \boldsymbol{y}_1(s)] \\ \qquad\qquad + \boldsymbol{J}\langle \mathcal{H}^{(1)}\rangle_{\boldsymbol{yy}}(\boldsymbol{y}^*(s), s)\boldsymbol{y}_1(s) + \boldsymbol{J}\langle \mathcal{H}^{(2)}\rangle_{\boldsymbol{y}}(\boldsymbol{y}^*(s), s), \\ \boldsymbol{b}_{\boldsymbol{y}}(\boldsymbol{y}^*(0))\boldsymbol{y}_2(0) = -\frac{1}{2}\boldsymbol{b}_{\boldsymbol{yy}}(\boldsymbol{y}^*(0))[\boldsymbol{y}_1(0), \boldsymbol{y}_1(0)], \\ \boldsymbol{c}_{\boldsymbol{y}}(\boldsymbol{y}^*(1))\boldsymbol{y}_2(1) = -\frac{1}{2}\boldsymbol{c}_{\boldsymbol{yy}}(\boldsymbol{y}^*(1))[\boldsymbol{y}_1(1), \boldsymbol{y}_1(1)] \end{cases} \tag{3.82}$$

solvable. The functions $\langle \mathcal{H}^{(1)}\rangle$ and $\langle \mathcal{H}^{(2)}\rangle$ appearing in these problems are expressed in equations (3.42) and (3.43). It follows from (2.130) and (2.131) that $\boldsymbol{y}_1$ can be written as

$$\boldsymbol{y}_1 = \boldsymbol{v}_1^* - \frac{F_0}{\mathcal{M}_\theta(\theta^*)}\boldsymbol{y}_\theta^*, \tag{3.83}$$

where $\boldsymbol{y}_\theta^*$ is defined in (3.68), $\boldsymbol{v}_1^*$ is the unique solution to (3.81) satisfying

$$\int_0^1 \boldsymbol{v}_1^*(s)^T \boldsymbol{y}_\theta^*(s) ds = 0, \tag{3.84}$$

and

$$\begin{aligned} F_0 = \int_0^1 & \boldsymbol{y}_\theta^*(s)^T \Big(\tfrac{1}{2}\langle H\rangle_{\boldsymbol{yyy}}(\boldsymbol{y}^*(s), s)[\boldsymbol{v}_1^*(s), \boldsymbol{v}_1^*(s)] \\ & + \langle \mathcal{H}^{(1)}\rangle_{\boldsymbol{yy}}(\boldsymbol{y}^*(s), s)\boldsymbol{v}_1^*(s) + \langle \mathcal{H}^{(2)}\rangle_{\boldsymbol{y}}(\boldsymbol{y}^*(s), s)\Big)\, ds \\ & + \tfrac{1}{2}\boldsymbol{y}_\theta^*(0)^T \boldsymbol{J}\boldsymbol{B}\boldsymbol{b}_{\boldsymbol{yy}}(\boldsymbol{y}^*(0))[\boldsymbol{v}_1^*(0), \boldsymbol{v}_1^*(0)] \\ & - \tfrac{1}{2}\boldsymbol{y}_\theta^*(1)^T \boldsymbol{J}\boldsymbol{C}\boldsymbol{c}_{\boldsymbol{yy}}(\boldsymbol{y}^*(1))[\boldsymbol{v}_1^*(1), \boldsymbol{v}_1^*(1)], \end{aligned} \tag{3.85}$$

with $\boldsymbol{B}$ and $\boldsymbol{C}$ defined in (2.69) and (2.70).

Then, for $\epsilon \in \mathbb{E}_T \cap (-\epsilon_0, \epsilon_0)$, the solution $\boldsymbol{\xi}(\epsilon)$ to BVP (3.1) can be expanded as

$$\boldsymbol{\xi}(\epsilon)(s) = \boldsymbol{y}^*(s) + \epsilon \boldsymbol{x}_1\big(s, \tfrac{s}{\epsilon}\big) + \ldots + \epsilon^{k-3}\boldsymbol{x}_{k-3}\big(s, \tfrac{s}{\epsilon}\big) + O(\epsilon^{k-2}) \tag{3.86}$$

by substituting (3.80), (3.27) and (3.29) into (3.71). For example, the first correction is

$$\boldsymbol{x}_1\big(s, \tfrac{s}{\epsilon}\big) = \boldsymbol{y}_1(s) + \boldsymbol{J}W_{\boldsymbol{y}}^{(1)}\big(\boldsymbol{y}^*(s), s, \tfrac{s}{\epsilon}\big), \tag{3.87}$$

where $W^{(1)}$ is explicitly given by (3.39). For $i = 1, \ldots, k-3$, the function $\boldsymbol{x}_i\big(s, \sigma\big)$ has a T-periodic dependence in σ, and $\boldsymbol{x}_i\big(s, \frac{s}{\epsilon}\big) = O(1)$ as $\epsilon \to 0$. Note that, in general, the s-derivative of the solution $\boldsymbol{\xi}(\epsilon)(s)$ does not have a limit as $\epsilon \to 0$. Therefore, $\boldsymbol{\xi}(\epsilon) \to \boldsymbol{y}^*$ as $\epsilon \to 0$ in the $\mathcal{C}^0$-norm, but not in the $\mathcal{C}^1$-norm.

3.5 Case of an unconstrained small parameter

The averaging results presented in the previous sections are valid only in the case the small parameter ϵ appearing in the two-scale problem belongs to the set $\mathbb{E}_T$. In this section it is shown that if ϵ is not in $\mathbb{E}_T$, then the two-scale problem can easily be turned into a problem of the same form with a small parameter belonging to $\mathbb{E}_T$, so that the above averaging results can still be used.

Consider the two-scale BVP (3.1) with ϵ small, but not restricted to the set $\mathbb{E}_T$. Defining the number

$$\omega = \frac{1}{\epsilon} \pmod{T} \in [0, T) \tag{3.88}$$

and the small parameter

$$\delta = \frac{\epsilon}{1 - \omega\epsilon} \in \mathbb{E}_T, \tag{3.89}$$

BVP (3.1) can be rewritten as

$$\begin{cases} \boldsymbol{x}'(s) = \boldsymbol{J}\check{H}_{\boldsymbol{x}}\big(\boldsymbol{x}(s), s, \frac{s}{\delta}, \delta, \omega\big), \\ \boldsymbol{b}(\boldsymbol{x}(0)) = \mathbf{0}, \quad \boldsymbol{c}(\boldsymbol{x}(1)) = \mathbf{0}, \end{cases} \tag{3.90}$$

where $\check{H}(\boldsymbol{x}, s, \sigma, \delta, \omega) = H\big(\boldsymbol{x}, s, \sigma + \omega s, \frac{\delta}{1+\delta\omega}\big)$. If we regard ω as a fixed parameter, then, under assumption (**3.A1**), this BVP has the same structure and the same properties as BVP (3.1). In particular, $\check{H}(\boldsymbol{x}, s, \sigma, \delta, \omega)$ is a T-periodic function of σ. Moreover, the small parameter δ belongs to $\mathbb{E}_T$. The averaging results of the previous sections therefore apply to this problem. The corresponding averaged BVP is exactly (3.2); it does not depend on the parameter ω.

Suppose $\boldsymbol{y}_0$ is an isolated solution to the averaged BVP and the nondegeneracy condition (**3.A3**) holds. Then, from the results of section 3.3, if δ is small enough, BVP (3.90) admits a solution $\boldsymbol{\xi}(\delta, \omega)$ which can be expanded as

$$\boldsymbol{\xi}(\delta, \omega)(s) = \boldsymbol{y}_0(s) + \delta \boldsymbol{x}_1\big(s, \tfrac{s}{\delta}, \omega\big) + \ldots + \delta^{k-1}\boldsymbol{x}_{k-1}\big(s, \tfrac{s}{\delta}, \omega\big) + O(\delta^k), \tag{3.91}$$

where, for $i = 1, \ldots, k-1$, the function $\boldsymbol{x}_i\big(s, \sigma, \omega\big)$ is smooth, and T-periodic in σ. It follows that, for ϵ small enough, $\boldsymbol{\xi}\big(\frac{\epsilon}{1-\omega\epsilon}, \frac{1}{\epsilon} \pmod{T}\big)$ is a solution to BVP (3.1). Note that this solution converges to $\boldsymbol{y}_0$ as $\epsilon \to 0$.

Suppose the boundary conditions satisfy property (**3.A4**), and the averaged BVP admits a one-parameter family of solutions $\Gamma = \boldsymbol{\zeta}(I)$, where I is an open interval in $\mathbb{R}$, $\boldsymbol{\zeta} : \mathbb{R} \to \mathcal{C}^1$ is smooth and, for any $\theta \in I$, $\boldsymbol{\zeta}_\theta(\theta) \neq \mathbf{0}$. In this case the corresponding Melnikov function is a smooth function of θ and ω. Let $\theta^*(\omega)$ be a simple zero of the Melnikov function, adopt the notation $\boldsymbol{y}^*(\omega) = \boldsymbol{\zeta}(\theta^*(\omega))$, and assume the nondegeneracy condition (**3.A6**) is satisfied. Then, from the results of section 3.4, if δ is small enough, BVP (3.90) admits a solution $\boldsymbol{\xi}(\delta, \omega)$ which can be expanded as

$$\boldsymbol{\xi}(\delta, \omega)(s) = \boldsymbol{y}^*(\omega)(s) + \delta \boldsymbol{x}_1\big(s, \tfrac{s}{\delta}, \omega\big) + \ldots + \delta^{k-3}\boldsymbol{x}_{k-3}\big(s, \tfrac{s}{\delta}, \omega\big) + O(\delta^{k-2}), \tag{3.92}$$

where, for $i = 1, \ldots, k-3$, the function $\boldsymbol{x}_i\big(s, \sigma, \omega\big)$ is smooth, and T-periodic in σ. It follows that, for ϵ small enough, $\boldsymbol{\xi}\big(\frac{\epsilon}{1-\omega\epsilon}, \frac{1}{\epsilon} \pmod{T}\big)$ is a solution to BVP (3.1). In the special case where the simple zero $\theta^*(\omega)$ of the Melnikov function is independent of ω, i.e. $\theta^*(\omega) = \hat{\theta}$, $\forall\, \omega \in [0, T)$, this solution to (3.1) converges to $\hat{\boldsymbol{y}} = \boldsymbol{\zeta}(\hat{\theta})$ as $\epsilon \to 0$. However, in general, the simple zeroes of the Melnikov function vary with ω, i.e. $\theta^*(\omega)$ cannot be chosen constant, and this solution to (3.1) does not have a limit as $\epsilon \to 0$. An example of this general situation is given by the clamped-sleeved strut with high intrinsic twist in section 6.6.

Chapter 4

Elastic rods

We adopt the special Cosserat model of a thin, inextensible, unshearable, elastic rod. This model allows the rod to undergo large deformations in space by suffering bending and twist. The present chapter introduces the model and discusses some of the properties of the rod that will be used in subsequent chapters. This material is fairly standard.

The equilibrium equations for the rod are presented in section 4.1. We examine the effects of a general, arclength-dependent change of framing on these equations in section 4.2. The boundary conditions corresponding to a clamped-free strut, a clamped-sleeved strut and a ring are introduced in section 4.3. The properties of isotropy and uniformity for rods are discussed in section 4.4 and 4.5 respectively. It is seen that, in the problem of a rod subject to one of the sets of boundary conditions mentioned above, isotropy automatically gives rise to a continuous symmetry, while uniformity only gives rise to a continuous symmetry in special cases. Finally, we consider an isotropic, uniform rod in section 4.6. We give expressions for the planar equilibria of the rod under the clamped-free and the clamped-sleeved strut boundary conditions, as well as for the circular equilibria of the rod under the ring boundary conditions.

4.1 Equilibrium equations

This section introduces the special Cosserat model of a thin, inextensible, unshearable rod with a quadratic energy, and presents the corresponding equilibrium equations. A comprehensive discussion of the special Cosserat theory can be found in Antman [6, chap. VIII], and the exact model and notation adopted here is more fully described in Li and Maddocks [33].

Let $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ be a fixed, right-handed, orthonormal frame in $\mathbb{R}^3$. We will sometimes identify vectors of $\mathbb{R}^3$ with their triples of coordinates relative to the basis $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$.

The rod is parametrized by arclength s. For convenience, we choose a length-scale such that $s \in [0, 1]$. A configuration of the rod is specified by the vector functions $\boldsymbol{r} \in C^2([0,1], \mathbb{R}^3)$ and $\boldsymbol{d}_1 \in C^1([0,1], \mathbb{R}^3)$, where the curve $\boldsymbol{r}(s)$ represents the centerline or axis of the rod, and $\boldsymbol{d}_1(s)$ is a unit vector normal to the centerline that determines the orientation of the cross-section at arclength s (see figure 4.1). The vector $\boldsymbol{d}_1(s)$ can be regarded as indicating the direction of a material line of the cross-section.

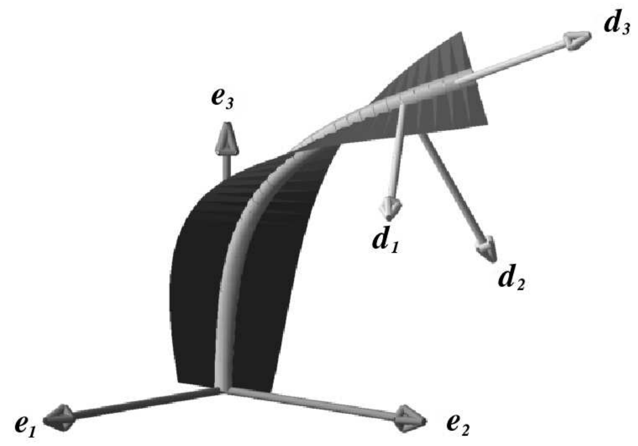


Figure 4.1: *An elastic rod. The centerline $\boldsymbol{r}(s)$ is drawn as a tube, and the normal vector $\boldsymbol{d}_1(s)$ indicating the orientation of the cross-section is represented by a ribbon.*

For any configuration of the rod, the vector defined as

$$\boldsymbol{d}_3(s) = \boldsymbol{r}'(s) \tag{4.1}$$

has unit norm by the condition of inextensibility. Moreover $\boldsymbol{d}_1(s)$ is perpendicular to $\boldsymbol{d}_3(s)$ by the condition of unshearability. Then, defining

$$\boldsymbol{d}_2(s) = \boldsymbol{d}_3(s) \times \boldsymbol{d}_1(s), \tag{4.2}$$

we obtain a right-handed, orthonormal frame $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ characterized by relation (4.2) and

$$\boldsymbol{d}_i(s)^T \boldsymbol{d}_j(s) = \delta_{ij}, \quad i, j = 1, 2, 3. \tag{4.3}$$

The frame $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ is called the *framing* of the rod, and the vectors $\boldsymbol{d}_i(s)$ are called *directors*. Orthonormality of the directors imply the existence of a *strain vector* $\boldsymbol{u}(s)$ satisfying

$$\boldsymbol{d}_i'(s) = \boldsymbol{u}(s) \times \boldsymbol{d}_i(s), \quad i = 1, 2, 3. \tag{4.4}$$

The components $u_i(s) = \boldsymbol{u}(s)^T \boldsymbol{d}_i(s)$ of the vector $\boldsymbol{u}(s)$ with respect to the director frame are the *bending* $(i = 1, 2)$ and *twist* $(i = 3)$ *strains*. Equation (4.4) leads to

$$u_1(s) = -\boldsymbol{d}_3'(s)^T \boldsymbol{d}_2(s), \quad u_2(s) = \boldsymbol{d}_3'(s)^T \boldsymbol{d}_1(s), \quad u_3(s) = \boldsymbol{d}_1'(s)^T \boldsymbol{d}_2(s). \tag{4.5}$$

We define the triple

$$\mathbf{u}(s) = \big(u_1(s), u_2(s), u_3(s)\big). \tag{4.6}$$

Note that the (scalar) curvature of the centerline is given by $\kappa(s) = \sqrt{u_1^2(s) + u_2^2(s)}$.

The stresses acting across the material cross-section at s are equivalent to a resultant force $\boldsymbol{n}(s)$ and a resultant moment $\boldsymbol{m}(s)$. Assuming the external stresses are all applied at the ends of the rod, balance of forces and moments implies

$$\boldsymbol{n}'(s) = \mathbf{0}, \quad \boldsymbol{m}'(s) + \boldsymbol{r}'(s) \times \boldsymbol{n}(s) = \mathbf{0}. \tag{4.7}$$

Assume the rod is hyperelastic with a stored energy density function of the quadratic form

$$W(\mathbf{u}(s), s) = \frac{1}{2}\big(\mathbf{u}(s) - \hat{\mathbf{u}}(s)\big)^T \boldsymbol{K}(s)\big(\mathbf{u}(s) - \hat{\mathbf{u}}(s)\big), \tag{4.8}$$

where $\boldsymbol{K}(s)$ is a 3×3, positive definite, symmetric matrix and $\hat{\mathbf{u}}(s) \in \mathbb{R}^3$. We suppose $\boldsymbol{K}(s)$ and $\hat{\mathbf{u}}(s)$ are continuous functions of s. We call $\boldsymbol{K}(s)$ the *stiffness matrix* and $\boldsymbol{K}^{-1}(s)$ the *flexibility matrix*. The triple $\hat{\mathbf{u}}(s) = (\hat{u}_1(s), \hat{u}_2(s), \hat{u}_3(s))$ corresponds to the strains of the rod in the *intrinsic* configuration, i.e. the configuration taken by the rod in absence of any applied load. The quantities $\hat{u}_i(s)$ are called *intrinsic curvatures* $(i = 1, 2)$ and *intrinsic twist* $(i = 3)$.

The components $m_i(s) = \boldsymbol{m}(s)^T \boldsymbol{d}_i(s)$ of the moment with respect to the director frame are the *bending* $(i = 1, 2)$ and *twist* $(i = 3)$ *moments*, and we define the triple

$$\mathbf{m}(s) = \big(m_1(s), m_2(s), m_3(s)\big). \tag{4.9}$$

The constitutive relation linking the moment to the strain takes the form

$$\mathbf{m}(s) = W_{\mathbf{u}}(\mathbf{u}(s), s) = \boldsymbol{K}(s)\big(\mathbf{u}(s) - \hat{\mathbf{u}}(s)\big). \tag{4.10}$$

In the particular case where the stiffness matrix is diagonal, i.e.

$$\boldsymbol{K}(s) = \operatorname{diag}\big(K_1(s), K_2(s), K_3(s)\big), \tag{4.11}$$

the constitutive relation simplifies to

$$m_i(s) = K_i(s)\big(u_i(s) - \hat{u}_i(s)\big), \quad i = 1, 2, 3. \tag{4.12}$$

Here $K_i(s)$ are called the *bending* $(i = 1, 2)$ and *twist* $(i = 3)$ *stiffnesses*.

To summarize, the equilibrium equations for the rod consist of the constraints (4.1)-(4.3), the kinematic equations (4.4) or (4.5), the balance laws (4.7), and the constitutive relation (4.10) to be satisfied for $s \in (0, 1)$. These equations are invariant under rigid translations and rotations of the rod in space.

We call *intrinsic properties* of a rod its stiffnesses and its intrinsic shape. The stiffnesses are encoded in a stiffness matrix $\boldsymbol{K}(s)$, while the intrinsic shape is described either by intrinsic strains $\hat{\mathbf{u}}(s) = (\hat{u}_1(s), \hat{u}_2(s), \hat{u}_3(s))$ or, equivalently, by an intrinsic framing $(\hat{\boldsymbol{d}}_1(s), \hat{\boldsymbol{d}}_2(s), \hat{\boldsymbol{d}}_3(s))$. Note that, for a given rod, these quantities are not unique; they depend on the choice of the framing, i.e. on the choice of the material lines attached to $\boldsymbol{d}_1(s)$ in the cross-sections (cf. section 4.2). The intrinsic framing associated with given intrinsic strains is defined by the differential equation

$$\hat{\boldsymbol{d}}_i{}'(s) = \hat{\boldsymbol{u}}(s) \times \hat{\boldsymbol{d}}_i(s) = \sum_{j=1}^{3} \hat{u}_j(s)\, \hat{\boldsymbol{d}}_j(s) \times \hat{\boldsymbol{d}}_i(s), \quad i = 1, 2, 3, \tag{4.13}$$

up to a global rotation (corresponding to the arbitrary choice of the initial conditions $(\hat{\boldsymbol{d}}_1(0), \hat{\boldsymbol{d}}_2(0), \hat{\boldsymbol{d}}_3(0))$, for example), while the intrinsic strains associated with a given intrinsic framing are expressed by

$$\hat{u}_1(s) = -\hat{\boldsymbol{d}}_3{}'(s)^T \hat{\boldsymbol{d}}_2(s), \quad \hat{u}_2(s) = \hat{\boldsymbol{d}}_3{}'(s)^T \hat{\boldsymbol{d}}_1(s), \quad \hat{u}_3(s) = \hat{\boldsymbol{d}}_1{}'(s)^T \hat{\boldsymbol{d}}_2(s). \tag{4.14}$$

4.2 Change of framing

For a given rod, different framings are associated with different choices of the material lines carrying $\boldsymbol{d}_1(s)$ in the cross-sections. The transformation of variables generated by a change of framing is examined in this section. The form of the equilibrium equations is shown to be preserved under such a transformation of variables.

Pick a function $\chi \in C^1([0,1],\mathbb{R})$, and rotate the original framing $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ about $\boldsymbol{d}_3(s)$ through an angle $\chi(s)$ to obtain a new set of directors

$$\big(\boldsymbol{D}_1(s), \boldsymbol{D}_2(s), \boldsymbol{D}_3(s)\big) = \big(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s)\big)\, \boldsymbol{\Omega}^T(\chi(s)), \tag{4.15}$$

where $\boldsymbol{\Omega}$ is the rotation matrix defined by

$$\boldsymbol{\Omega}(t) = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.16}$$

The constraints (4.1)-(4.3) on the original framing persist on the new framing, i.e.

$$\begin{gathered} \boldsymbol{D}_3(s) = \boldsymbol{r}'(s), \quad \boldsymbol{D}_2(s) = \boldsymbol{D}_3(s) \times \boldsymbol{D}_1(s), \\ \boldsymbol{D}_i(s)^T \boldsymbol{D}_j(s) = \delta_{ij}, \quad i,j = 1,2,3. \end{gathered} \tag{4.17}$$

The new frame has an associated strain vector $\boldsymbol{U}(s)$ satisfying

$$\boldsymbol{D}_i{}'(s) = \boldsymbol{U}(s) \times \boldsymbol{D}_i(s), \quad i = 1,2,3. \tag{4.18}$$

From (4.15) and (4.4),

$$\begin{aligned} \boldsymbol{D}_1{}'(s) = {} & \cos\chi(s)\, \boldsymbol{d}_1{}'(s) - \sin\chi(s)\, \boldsymbol{d}_2{}'(s) \\ & - \chi'(s)\big(\sin\chi(s)\, \boldsymbol{d}_1(s) + \cos\chi(s)\, \boldsymbol{d}_2(s)\big) \\ = {} & \boldsymbol{u}(s) \times \big(\cos\chi(s)\, \boldsymbol{d}_1(s) - \sin\chi(s)\, \boldsymbol{d}_2(s)\big) - \chi'(s)\, \boldsymbol{D}_2(s) \\ = {} & \boldsymbol{u}(s) \times \boldsymbol{D}_1(s) - \chi'(s)\, \boldsymbol{D}_3(s) \times \boldsymbol{D}_1(s) \\ = {} & \big(\boldsymbol{u}(s) - \chi'(s)\, \boldsymbol{d}_3(s)\big) \times \boldsymbol{D}_1(s) \end{aligned} \tag{4.19}$$

and, similarly,

$$\boldsymbol{D}_i{}'(s) = \big(\boldsymbol{u}(s) - \chi'(s)\, \boldsymbol{d}_3(s)\big) \times \boldsymbol{D}_i(s), \quad i = 2,3, \tag{4.20}$$

so that the new stain vector is expressed by

$$\boldsymbol{U}(s) = \boldsymbol{u}(s) - \chi'(s)\, \boldsymbol{d}_3(s). \tag{4.21}$$

Denote by $\mathbf{U}(s)$ the triple of components $U_i(s) = \boldsymbol{U}(s)^T \boldsymbol{D}_i(s)$, $i = 1,2,3$. Then

$$\begin{aligned} U_1(s) &= \big(\boldsymbol{u}(s) - \chi'(s)\, \boldsymbol{d}_3(s)\big)^T \big(\cos\chi(s)\, \boldsymbol{d}_1(s) - \sin\chi(s)\, \boldsymbol{d}_2(s)\big) \\ &= \cos\chi(s)\, u_1(s) - \sin\chi(s)\, u_2(s) \end{aligned} \tag{4.22}$$

and, similarly,

$$U_2(s) = \sin\chi(s)\, u_1(s) + \cos\chi(s)\, u_2(s), \quad U_3(s) = u_3(s) - \chi'(s). \tag{4.23}$$

Combining these equalities shows that the new strains are related to the old ones through

$$\mathbf{U}(s) = \mathbf{\Omega}(\chi(s))\big(\mathbf{u}(s) - \chi'(s)\,\mathbf{e}_3\big), \tag{4.24}$$

where $\mathbf{e}_3 = (0, 0, 1)$. The same relation holds in the unstressed configuration, i.e.

$$\hat{\mathbf{U}}(s) = \mathbf{\Omega}(\chi(s))\big(\hat{\mathbf{u}}(s) - \chi'(s)\,\mathbf{e}_3\big). \tag{4.25}$$

The force and moment vectors $\boldsymbol{n}(s)$ and $\boldsymbol{m}(s)$ are unperturbed by the change of framing; they still satisfy the balance laws

$$\boldsymbol{n}'(s) = \mathbf{0}, \quad \boldsymbol{m}'(s) + \boldsymbol{r}'(s) \times \boldsymbol{n}(s) = \mathbf{0}. \tag{4.26}$$

Denote by $\mathbf{M}(s)$ the triple of components $M_i(s) = \boldsymbol{m}(s)^T \boldsymbol{D}_i(s)$ of the moment vector in the new framing. We have

$$\begin{aligned} M_1(s) &= \boldsymbol{m}(s)^T\big(\cos\chi(s)\,\boldsymbol{d}_1(s) - \sin\chi(s)\,\boldsymbol{d}_2(s)\big) \\ &= \cos\chi(s)\,m_1(s) - \sin\chi(s)\,m_2(s), \end{aligned} \tag{4.27}$$

and, similarly,

$$M_2(s) = \sin\chi(s)\,m_1(s) + \cos\chi(s)\,m_2(s), \quad M_3(s) = m_3(s). \tag{4.28}$$

Combining these equalities yields

$$\mathbf{M}(s) = \mathbf{\Omega}(\chi(s))\,\mathbf{m}(s). \tag{4.29}$$

According to equations (4.24), (4.25) and (4.29), constitutive relation (4.10) can be expressed in the new variables as

$$\mathbf{M}(s) = \boldsymbol{L}(s)\left(\mathbf{U}(s) - \hat{\mathbf{U}}(s)\right), \tag{4.30}$$

where

$$\boldsymbol{L}(s) = \mathbf{\Omega}(\chi(s))\,\boldsymbol{K}(s)\,\mathbf{\Omega}^T(\chi(s)) \tag{4.31}$$

is the new stiffness matrix. It appears that the rod equilibrium equations in the new variables, namely equations (4.17), (4.18), (4.26) and (4.30), have the same form as the equilibrium equations in the original variables.

4.3 Boundary conditions

Many different sets of end conditions can be imposed on a rod (cf. Antman [6, sect. VIII.12]). Here we only consider three that occur in classic problems and are readily realizable in experiments. The boundary conditions are expressed with respect to the fixed frame $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$. We describe the $\boldsymbol{e}_3$-direction as *vertical*.

Clamped-free strut

A rod satisfies the *clamped-free strut* boundary conditions

$$\begin{gathered} \boldsymbol{r}(0) = \boldsymbol{0}, \quad \boldsymbol{d}_i(0) = \boldsymbol{e}_i, \quad i = 1, 2, 3, \\ \boldsymbol{n}(1) = (0, 0, -\lambda), \quad \boldsymbol{m}(1) = \boldsymbol{0} \end{gathered} \tag{4.32}$$

if it is vertically clamped at one end, and free to move under the influence of a prescribed vertical load of magnitude λ at the other end.

Clamped-sleeved strut

A rod subject to the *clamped-sleeved strut* boundary conditions

$$\begin{gathered} \boldsymbol{r}(0) = \boldsymbol{0}, \quad \boldsymbol{d}_i(0) = \boldsymbol{e}_i, \quad i = 1, 2, 3, \\ r_1(1) = r_2(1) = 0, \quad n_3(1) = -\lambda, \quad \boldsymbol{d}_3(1) = \boldsymbol{e}_3, \quad m_3(1) = 0 \end{gathered} \tag{4.33}$$

has one end vertically clamped at the origin, while the other end is held by a vertical sleeve on the $\boldsymbol{e}_3$-axis and loaded by a force with vertical component $-\lambda$. The end in the sleeve is free to rotate about its (vertical) axis.

Ring

The *ring* boundary conditions (cf. figure 4.2)

$$\begin{gathered} \boldsymbol{r}(0) = \boldsymbol{0}, \quad \boldsymbol{d}_i(0) = \boldsymbol{e}_i, \quad i = 1, 2, 3, \\ \boldsymbol{r}(1) = \boldsymbol{0}, \quad \boldsymbol{d}_1(1) = (\cos\alpha, \sin\alpha, 0), \\ \boldsymbol{d}_2(1) = (-\sin\alpha, \cos\alpha, 0), \quad \boldsymbol{d}_3(1) = \boldsymbol{e}_3, \end{gathered} \tag{4.34}$$

with $0 \leqslant \alpha < 2\pi$, force the rod to close upon itself at the origin tangentially to $\boldsymbol{e}_3$, with a prescribed angle α between $\boldsymbol{d}_1(0)$ and $\boldsymbol{d}_1(1)$.

Note that each of these three sets of end conditions eliminates the continuous symmetries of the equilibrium equations corresponding to rigid translations and rotations of the rod.

4.4 Isotropic rods

Isotropy (cf. Antman [6, sect. VIII.9]) is an important property in rod theory; it leads to a substantial simplification of the equilibrium equations. After giving a definition of isotropy, this section presents a continuous symmetry enjoyed by an isotropic rod subject to any of the three sets of boundary conditions introduced in section 4.3. A second part of the section is devoted to establishing various properties of the equilibria of an isotropic rod.

Definition 4.1 *We call a rod* constitutively isotropic *if its energy density function (4.8) is invariant under rotations of the directors $\boldsymbol{d}_1(s)$ and $\boldsymbol{d}_2(s)$ about $\boldsymbol{d}_3(s)$, i.e. if its stiffness matrix is of the special form (cf. Healey [25, remark 5.1])*

$$\boldsymbol{K}(s) = diag\big(K_1(s), K_1(s), K_3(s)\big), \tag{4.35}$$

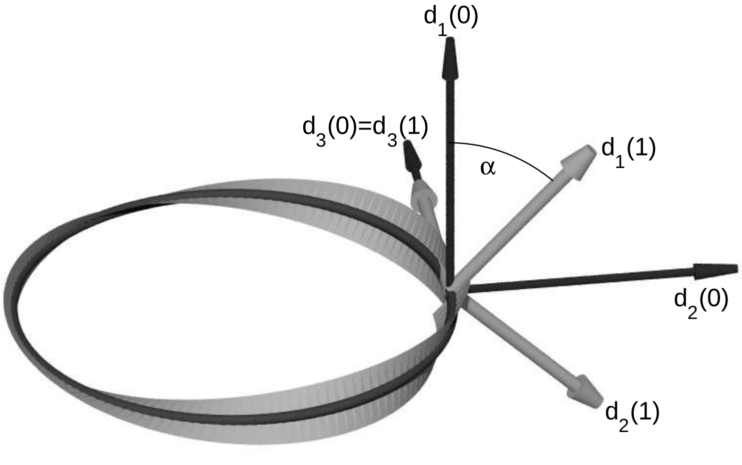


Figure 4.2: *Ring boundary conditions. At $s = 0$, the frame is oriented along the standard axes. At $s = 1$, the frame is rotated about $\mathbf{d}_3(0)$ by a prescribed angle α. The $s = 0$ frame is drawn darker and at a larger scale in order to distinguish it from the $s = 1$ frame.*

and constitutively anisotropic *otherwise. A rod is* isotropic *if it is constitutively isotropic and its unstressed configuration is straight, i.e.*

$$\hat{u}_1(s) = \hat{u}_2(s) = 0, \tag{4.36}$$

and anisotropic *otherwise.*

Note that the property of isotropy is independent of the choice of the framing, i.e. if a rod is isotropic (anisotropic) in one framing, then it is isotropic (anisotropic) in any framing (cf. equations (4.25) and (4.31)). For an isotropic rod, constitutive relation (4.10) simplifies to

$$m_1(s) = K_1(s)u_1(s), \tag{4.37}$$

$$m_2(s) = K_1(s)u_2(s), \tag{4.38}$$

$$m_3(s) = K_3(s)\big(u_3(s) - \hat{u}_3(s)\big). \tag{4.39}$$

The equilibrium equations for an isotropic rod are invariant under the continuous symmetry which consists of rotating the directors $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ about $\boldsymbol{d}_3(s)$ by a fixed angle for all $s \in [0, 1]$, leaving the centerline fixed.

Consider an isotropic rod subject to one of the sets of boundary conditions introduced in section 4.3. This problem is not invariant under the transformation associated to isotropy because of the boundary conditions. However, it is invariant under the transformation composed of (cf. Manning and Maddocks [37, sect. 3, transformation I])

step 1: a rigid rotation of the rod by an angle θ about $\boldsymbol{e}_3$,
step 2: a rotation of $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ by an angle $-\theta$ about $\boldsymbol{d}_3(s)$, leaving the centerline fixed. (4.40)

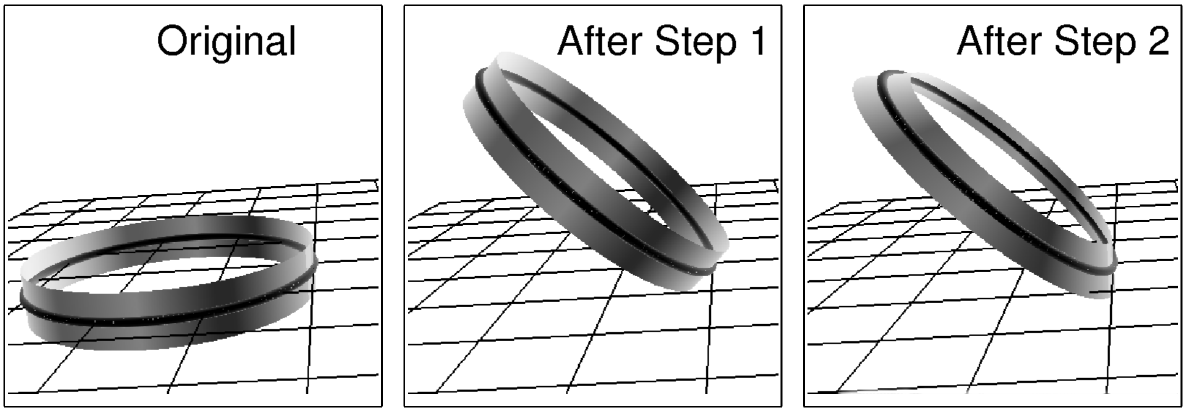


Figure 4.3: *Transformation (4.40): original configuration of the rod and its images after steps 1 and 2, for $\theta = 45°$. The $\boldsymbol{e}_3$-axis lies on the grid, pointing away from the viewer. (Reprinted from Manning and Maddocks [37, page 318], with permission from Elsevier Science.)*

This register symmetry is illustrated in figure 4.3. Performing transformation (4.40) on a solution $(\boldsymbol{r}^0, \boldsymbol{d}_i^0, \boldsymbol{n}^0, \mathbf{m}^0)$ produces a family of solutions expressed by

$$\begin{aligned} \boldsymbol{r}(s) = \boldsymbol{\Omega}(\theta)\boldsymbol{r}^0(s), \quad \boldsymbol{n} = \boldsymbol{\Omega}(\theta)\boldsymbol{n}^0, \quad \mathbf{m}(s) = \boldsymbol{\Omega}(\theta)\mathbf{m}^0(s), \\ \big(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s)\big) = \boldsymbol{\Omega}(\theta)\big(\boldsymbol{d}_1^0(s), \boldsymbol{d}_2^0(s), \boldsymbol{d}_3^0(s)\big)\boldsymbol{\Omega}^T(\theta), \end{aligned} \tag{4.41}$$

where $\boldsymbol{\Omega}(\theta)$ is the rotation matrix defined in (4.16). This family of equilibria is parametrized by the angle $\theta \in [0, 2\pi)$, and is diffeomorphic to a circle if $(\boldsymbol{r}^0, \boldsymbol{d}_i^0, \boldsymbol{n}^0, \mathbf{m}^0)$ is not a fixed point of the transformation.

The following lemma shows that, in the case of an isotropic rod, the twist moment constitutes an integral for the equilibrium equations.

Lemma 4.2 *Consider an equilibrium configuration of an isotropic rod. Then the twist moment is constant along the rod, and we write*

$$m_3(s) = m_3. \tag{4.42}$$

Proof By definition,

$$m_3'(s) = \big(\boldsymbol{m}(s)^T\boldsymbol{d}_3(s)\big)' = \boldsymbol{m}'(s)^T\boldsymbol{d}_3(s) + \boldsymbol{m}(s)^T\boldsymbol{d}_3'(s). \tag{4.43}$$

Combining (4.1) and (4.7) yields $\boldsymbol{m}'(s) = \boldsymbol{n}(s) \times \boldsymbol{d}_3(s)$, so that $\boldsymbol{m}'(s)^T\boldsymbol{d}_3(s) = 0$. Therefore, by (4.4),

$$m_3'(s) = \boldsymbol{m}(s)^T\big(\boldsymbol{u}(s) \times \boldsymbol{d}_3(s)\big). \tag{4.44}$$

Constitutive relations (4.37)-(4.39) can be written as

$$\boldsymbol{m}(s) = K_1(s)\boldsymbol{u}(s) + \Big(K_3(s)\big(u_3(s) - \hat{u}_3(s)\big) - K_1(s)u_3(s)\Big)\boldsymbol{d}_3(s). \tag{4.45}$$

Substituting this expression into (4.44) shows that $m_3'(s) = 0$, and the conclusion follows. ■

We proceed by establishing additional properties of the equilibria of an isotropic rod.

Proposition 4.3 *Consider an equilibrium configuration of an isotropic rod. Suppose that, at some point of the rod, the twist moment is zero and the moment is perpendicular to the force, i.e.*

$$m_3(s_0) = 0, \quad \boldsymbol{m}(s_0)^T \boldsymbol{n} = 0, \tag{4.46}$$

for some $s_0 \in [0,1]$. Then the centerline $\boldsymbol{r}(s)$ lies in a plane parallel to the direction of the force $\boldsymbol{n}$.

Proof By lemma 4.2 and hypothesis (4.46),

$$m_3(s) = 0, \tag{4.47}$$

and, by (4.39), $u_3(s) - \hat{u}_3(s) = 0$. It follows that constitutive relations (4.37)-(4.39) can be rewritten as

$$\boldsymbol{m}(s) = K_1(s)\big(\boldsymbol{u}(s) - \hat{\boldsymbol{u}}(s)\big), \tag{4.48}$$

and, by (4.4) and (4.36),

$$\boldsymbol{d}_3{}'(s) = \frac{1}{K_1(s)}\big(\boldsymbol{m}(s) \times \boldsymbol{d}_3(s)\big). \tag{4.49}$$

According to equation (4.7), $\boldsymbol{n}(s)$ is a constant function of s, and we write $\boldsymbol{n}(s) = \boldsymbol{n}$. Combining (4.1) and (4.7) leads to

$$\boldsymbol{m}'(s) = \boldsymbol{n} \times \boldsymbol{d}_3(s). \tag{4.50}$$

Differentiating this equation and using (4.49) shows that

$$\begin{aligned} \boldsymbol{m}''(s) &= \frac{1}{K_1(s)}\boldsymbol{n} \times \big(\boldsymbol{m}(s) \times \boldsymbol{d}_3(s)\big) \\ &= \frac{1}{K_1(s)}\Big(\big(\boldsymbol{n}^T \boldsymbol{d}_3(s)\big)\, \boldsymbol{m}(s) - \big(\boldsymbol{n}^T \boldsymbol{m}(s)\big)\, \boldsymbol{d}_3(s)\Big). \end{aligned} \tag{4.51}$$

By (4.7), the quantity $\boldsymbol{m}(s) + \boldsymbol{r}(s) \times \boldsymbol{n}$ is constant with respect to s, so that $\boldsymbol{m}(s)^T\boldsymbol{n}$ is also constant. Thus, by hypothesis (4.46),

$$\boldsymbol{m}(s)^T \boldsymbol{n} = 0, \tag{4.52}$$

and equation (4.51) simplifies to

$$\boldsymbol{m}''(s) = \frac{1}{K_1(s)}\, \big(\boldsymbol{n}^T \boldsymbol{d}_3(s)\big)\, \boldsymbol{m}(s). \tag{4.53}$$

To complete the proof, we consider two cases.

First we study the case where $\boldsymbol{m}(s) = \boldsymbol{0}$, for all $s \in [0,1]$. Then, by (4.49), $\boldsymbol{d}_3{}'(s) = \boldsymbol{0}$, so that $\boldsymbol{d}_3(s) = \boldsymbol{d}_3$. Therefore, by (4.1), $\boldsymbol{r}(s) = \boldsymbol{r}(0) + s\,\boldsymbol{d}_3$, i.e. the centerline of the rod is straight. Moreover, combining the equality $\boldsymbol{m}'(s) = \mathbf{0}$ with (4.50) shows that $\boldsymbol{n} \times \boldsymbol{d}_3(s) = \mathbf{0}$, i.e. the centerline is parallel to the direction of the force. This particular situation is consistent with the conclusion of the proposition.

We now study the case where there exists $\hat{s} \in [0,1]$ such that $\boldsymbol{m}(\hat{s}) \neq \mathbf{0}$. Consider the quantity

$$\boldsymbol{z}(s) = \boldsymbol{m}(s) \times \boldsymbol{m}(\hat{s}). \tag{4.54}$$

We have

$$\boldsymbol{z}(\hat{s}) = \mathbf{0}, \tag{4.55}$$

and, from (4.53),

$$\boldsymbol{z}''(s) = \frac{1}{K_1(s)} \left(\boldsymbol{n}^T \boldsymbol{d}_3(s)\right) \boldsymbol{z}(s). \tag{4.56}$$

By (4.50),

$$\begin{aligned} \boldsymbol{z}'(\hat{s}) &= \left(\boldsymbol{n} \times \boldsymbol{d}_3(\hat{s})\right) \times \boldsymbol{m}(\hat{s}) \\ &= \left(\boldsymbol{m}(\hat{s})^T \boldsymbol{n}\right) \boldsymbol{d}_3(\hat{s}) - \left(\boldsymbol{m}(\hat{s})^T \boldsymbol{d}_3(\hat{s})\right) \boldsymbol{n}, \end{aligned} \tag{4.57}$$

so that, by (4.47) and (4.52),

$$\boldsymbol{z}'(\hat{s}) = \mathbf{0}. \tag{4.58}$$

Consider the linear initial value problem defined by (4.56), (4.55) and (4.58) for $\boldsymbol{z}(s)$ on both the intervals $[0, \hat{s}]$ and $[\hat{s}, 1]$. Since the function multiplying $\boldsymbol{z}(s)$ in the right hand side of linear differential equation (4.56) is continuous, the initial value problem has a unique solution (cf. Hartman [24, lemma IV 1.1]). This solution is $\boldsymbol{z}(s) = \mathbf{0}$, so that, by (4.54), the direction of $\boldsymbol{m}(s)$ is constant with respect to s, i.e.

$$\boldsymbol{m}(s) = m(s)\boldsymbol{\eta}, \tag{4.59}$$

where $m(s)$ is the norm of $\boldsymbol{m}(s)$ and $\boldsymbol{\eta}$ is a fixed unit vector in $\mathbb{R}^3$. Then, by (4.47), $\boldsymbol{m}(s)^T \boldsymbol{d}_3(s) = 0$, and $\boldsymbol{d}_3(s)$ is perpendicular to $\boldsymbol{\eta}$. From (4.1),

$$\boldsymbol{r}(s) = \boldsymbol{r}(0) + \int_0^s \boldsymbol{d}_3(\sigma)\, d\sigma, \tag{4.60}$$

so that the centerline of the rod lies in a plane perpendicular to $\boldsymbol{\eta}$. Moreover, by (4.52), the force $\boldsymbol{n}$ is perpendicular to $\boldsymbol{\eta}$. The conclusion follows. ■

The following proposition is used in section 4.6 to obtain planar equilibria of a particular rod subject to different boundary conditions.

Proposition 4.4 *Suppose the centerline of an equilibrium configuration of an isotropic* *rod lies in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane. Then this configuration is characterized by*

$$\boldsymbol{r}(s) = \boldsymbol{r}(0) + \begin{pmatrix} \int_0^s \sin\phi(\sigma)\, d\sigma \\ 0 \\ \int_0^s \cos\phi(\sigma)\, d\sigma \end{pmatrix}, \quad \boldsymbol{n}(s) = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix},$$

$$\boldsymbol{d}_1(s) = \begin{pmatrix} \cos\psi(s)\cos\phi(s) \\ \sin\psi(s) \\ -\cos\psi(s)\sin\phi(s) \end{pmatrix}, \quad \boldsymbol{d}_2(s) = \begin{pmatrix} -\sin\psi(s)\cos\phi(s) \\ \cos\psi(s) \\ \sin\psi(s)\sin\phi(s) \end{pmatrix}, \tag{4.61}$$

$$\boldsymbol{d}_3(s) = \begin{pmatrix} \sin\phi(s) \\ 0 \\ \cos\phi(s) \end{pmatrix}, \quad \mathbf{m}(s) = \begin{pmatrix} K_1(s)\phi'(s)\sin\psi(s) \\ K_1(s)\phi'(s)\cos\psi(s) \\ m_3 \end{pmatrix},$$

where the constants n_1, n_2, n_3, m_3 and the scalar functions $\phi(s)$, $\psi(s)$ satisfy

$$K_3(s)\big(\psi'(s) - \hat{u}_3(s)\big) = m_3, \tag{4.62}$$

$$\big(K_1(s)\phi'(s)\big)' = n_3 \sin\phi(s) - n_1\cos\phi(s), \quad m_3\phi'(s) = n_2. \tag{4.63}$$

Proof By hypothesis, $\boldsymbol{r}(s)$ lies in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane, so that $\boldsymbol{d}_3(s) = \boldsymbol{r}'(s)$ also lies in that plane. It follows that $\boldsymbol{d}_3(s)$ can be expressed as in equation (4.61), where $\phi(s)$ is the angle between $\boldsymbol{d}_3(s)$ and $\boldsymbol{e}_3$. Since $\boldsymbol{d}_1(s)$ is perpendicular to $\boldsymbol{d}_3(s)$, it can be expressed as in (4.61), where $\psi(s)$ is the angle between $\boldsymbol{d}_1(s)$ and the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane. The expressions for $\boldsymbol{r}(s)$ and $\boldsymbol{d}_2(s)$ given in (4.61) follow from equations (4.1) and (4.2) respectively.

From (4.7) and (4.1), $\boldsymbol{n}(s) = \boldsymbol{n} = (n_1, n_2, n_3)$ and

$$\boldsymbol{m}'(s) + \boldsymbol{d}_3(s) \times \boldsymbol{n} = 0. \tag{4.64}$$

Equations (4.5), (4.37), (4.38), (4.39) and lemma 4.2 lead to the expression of $\mathbf{m}(s)$ given in (4.61) and equality (4.62). The moment vector is then

$$\begin{aligned} \boldsymbol{m}(s) &= m_1(s)\boldsymbol{d}_1(s) + m_2(s)\boldsymbol{d}_2(s) + m_3(s)\boldsymbol{d}_3(s) \\ &= \begin{pmatrix} m_3 \sin\phi(s) \\ K_1(s)\phi'(s) \\ m_3\cos\phi(s) \end{pmatrix}. \end{aligned} \tag{4.65}$$

Finally, inserting expressions (4.65) and (4.61) into (4.64) yields equalities (4.63). ∎

4.5 Uniform rods

In this section, we consider the case of a uniform rod (cf. Antman [6, sect. VIII.10]) subject to one of the three sets of boundary conditions introduced in section 4.3. The problem admits no continuous symmetry associated with the property of uniformity, except in three special cases. We describe the symmetry appearing in these special cases.

According to section 4.1, the intrinsic properties of a rod can be described by a stiffness matrix $\boldsymbol{K}(s)$ together with intrinsic strains $\hat{\mathbf{u}}(s)$. However, as shown in section 4.2, for a given rod, these quantities depend on the choice of the framing $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$.

Definition 4.5 *A rod is* uniform *if, for a suitable choice of framing* $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$, *the energy density function (4.8) does not explicitly depend on arclength s or, equivalently, the stiffness matrix $\boldsymbol{K}$ and the intrinsic strains $\hat{\mathbf{u}}$ are independent of s.*

Given a uniform rod, the framings such that $\boldsymbol{K}$ and $\hat{\mathbf{u}}$ do not depend on s are most natural and, in the following, we assume the framing denoted by $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ is chosen such that these conditions hold. The equilibrium equations of a uniform rod are invariant under the continuous symmetry of translation in arclength. However, in general, a uniform rod subject to one of the three sets of boundary conditions introduced in section 4.3 admits no symmetry associated with uniformity. The three following special cases constitute exceptions to this statement.

(a) An isotropic, uniform rod with the ring boundary conditions.

(b) A uniform rod under the ring boundary conditions with $\alpha = 0$, i.e. $\boldsymbol{d}_1(1) = \boldsymbol{d}_1(0)$.

(c) A uniform rod with a straight intrinsic configuration, i.e. $\hat{u}_1 = \hat{u}_2 = 0$, with a stiffness matrix of the block-diagonal form

$$\boldsymbol{K} = \begin{pmatrix} K_{11} & K_{12} & 0 \\ K_{12} & K_{22} & 0 \\ 0 & 0 & K_{33} \end{pmatrix}, \tag{4.66}$$

and subject to the ring boundary conditions with $\alpha = \pi$, i.e. $\boldsymbol{d}_1(1) = -\boldsymbol{d}_1(0)$.

The BVP corresponding to each of these three cases is invariant under the transformation defined by the following operations on an equilibrium $(\boldsymbol{r}^0, \boldsymbol{d}_i^0, \boldsymbol{n}^0, \mathbf{m}^0)$(cf. Manning and Maddocks [37, sect. 3, transformation II]):

step 1: a rigid translation of the rod by $-\boldsymbol{r}^0(\beta)$, with $0 \leqslant \beta < 1$, and a reparametrization so that $s = 0$ is at the origin,

step 2: for each $s > 1 - \beta$, a rotation of the framing by angle α about the tangent to the centerline, leaving the centerline fixed, (4.67)

step 3: a rigid rotation of the rod about the origin described by the rotation matrix $\big(\boldsymbol{d}_1^0(\beta), \boldsymbol{d}_2^0(\beta), \boldsymbol{d}_3^0(\beta)\big)^T$.

This transformation is illustrated in figure 4.4. It produces a family of solutions parametrized by $\beta \in [0, 1)$, and expressed by

$$\begin{aligned} \boldsymbol{r}(s) &= \big(\boldsymbol{d}_1^0(\beta), \boldsymbol{d}_2^0(\beta), \boldsymbol{d}_3^0(\beta)\big)^T \big(\boldsymbol{r}^0(t) - \boldsymbol{r}^0(\beta)\big), \\ \boldsymbol{n} &= \big(\boldsymbol{d}_1^0(\beta), \boldsymbol{d}_2^0(\beta), \boldsymbol{d}_3^0(\beta)\big)^T \boldsymbol{n}^0, \quad \mathbf{m}(s) = \boldsymbol{A}(s)^T \mathbf{m}^0(t), \\ \big(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s)\big) &= \big(\boldsymbol{d}_1^0(\beta), \boldsymbol{d}_2^0(\beta), \boldsymbol{d}_3^0(\beta)\big)^T \big(\boldsymbol{d}_1^0(t), \boldsymbol{d}_2^0(t), \boldsymbol{d}_3^0(t)\big) \boldsymbol{A}(s), \end{aligned} \tag{4.68}$$

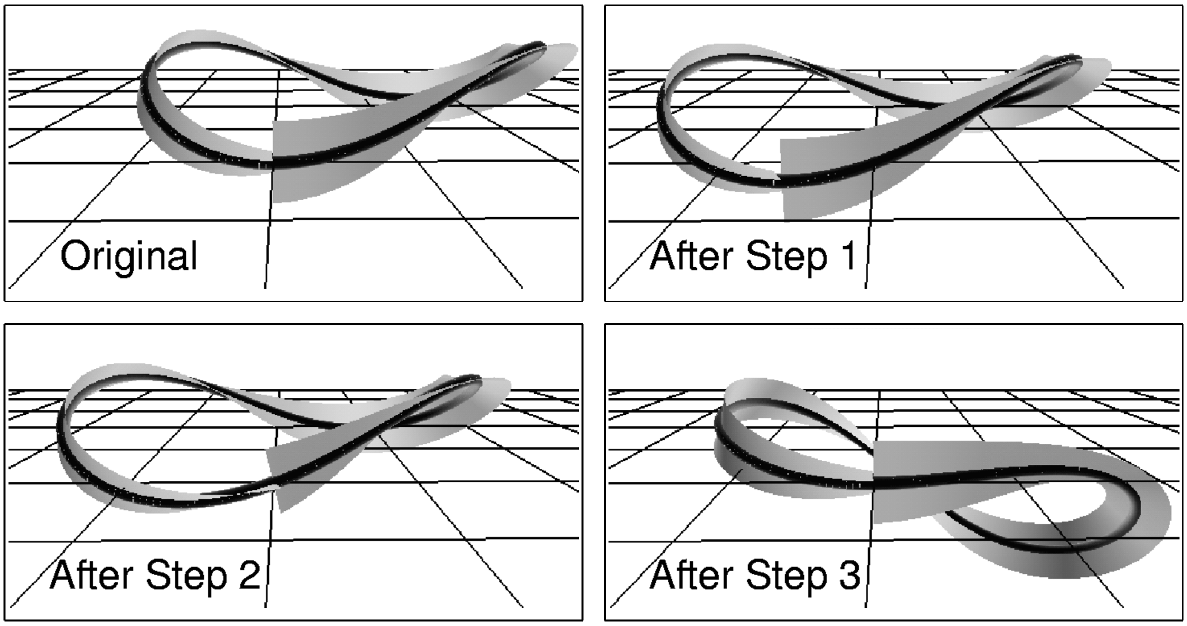


Figure 4.4: *Transformation (4.67): original configuration of the rod and its images after steps 1 through 3, for $\beta \simeq 0.06$. The $\boldsymbol{e}_3$-axis lies on the grid, pointing to the right. (Reprinted from Manning and Maddocks [37, page 318], with permission from Elsevier Science.)*

with $t = (s + \beta) \pmod 1$, and

$$\boldsymbol{A}(s) = \begin{cases} \boldsymbol{I} & \text{if } s < 1 - \beta, \\ \boldsymbol{\Omega}(\alpha) & \text{if } s > 1 - \beta, \end{cases} \tag{4.69}$$

where $\boldsymbol{I}$ denotes the 3×3 identity matrix and $\boldsymbol{\Omega}$ is defined in (4.16).

4.6 Equilibria of isotropic, uniform rods

In this section we study an isotropic, uniform rod under each of the three sets of boundary conditions introduced in section 4.3. We give rather explicit expressions for the planar equilibria of the rod under the clamped-free strut and the clamped-sleeved strut boundary conditions, as well as explicit expressions for the circular equilibria of the rod under the ring boundary conditions. The stiffness matrix and the intrinsic strains of the isotropic, uniform rod are denoted by

$$\boldsymbol{K}(s) = \operatorname{diag}(K_1, K_1, K_3), \quad \hat{\mathbf{u}}(s) = (0, 0, \omega). \tag{4.70}$$

Clamped-free strut

Consider an isotropic, uniform rod subject to the clamped-free strut boundary conditions (4.32). According to proposition 4.3, all equilibria of this strut lie in planes containing $\boldsymbol{e}_3$.

It follows from proposition 4.4 that the equilibria lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane take the form

$$\boldsymbol{r}^0(s) = \begin{pmatrix} \int_0^s \sin\phi(\sigma)\,d\sigma \\ 0 \\ \int_0^s \cos\phi(\sigma)\,d\sigma \end{pmatrix}, \quad \boldsymbol{n}^0(s) = \begin{pmatrix} 0 \\ 0 \\ -\lambda \end{pmatrix},$$

$$\boldsymbol{d}_1^0(s) = \begin{pmatrix} \cos(\omega s)\cos\phi(s) \\ \sin(\omega s) \\ -\cos(\omega s)\sin\phi(s) \end{pmatrix}, \quad \boldsymbol{d}_2^0(s) = \begin{pmatrix} -\sin(\omega s)\cos\phi(s) \\ \cos(\omega s) \\ \sin(\omega s)\sin\phi(s) \end{pmatrix}, \tag{4.71}$$

$$\boldsymbol{d}_3^0(s) = \begin{pmatrix} \sin\phi(s) \\ 0 \\ \cos\phi(s) \end{pmatrix}, \quad \mathbf{m}^0(s) = \begin{pmatrix} K_1\phi'(s)\sin(\omega s) \\ K_1\phi'(s)\cos(\omega s) \\ 0 \end{pmatrix},$$

where $\phi(s)$ satisfies

$$\phi''(s) = -\frac{\lambda}{K_1}\sin\phi(s), \quad \phi(0) = 0, \quad \phi'(1) = 0. \tag{4.72}$$

The solution set to this classic problem is described in appendix C.

The other equilibria of the clamped-free strut are obtained by applying transformation (4.40) on the equilibria lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane. Finally, note that performing transformation (4.40) with $\theta = \pi$ to an equilibrium of the form (4.71) with $\phi(s) = \hat{\phi}(s)$, yields again an equilibrium of the form (4.71), but with $\phi(s) = -\hat{\phi}(s)$.

Clamped-sleeved strut

Consider an isotropic, uniform rod subject to the clamped-sleeved strut boundary conditions (4.33). Although this strut possesses non-planar equilibria (cf. Maddocks [34], Neukirch and Henderson [42], for example), we only consider the planar equilibria here. According to boundary conditions $\boldsymbol{r}(0) = \mathbf{0}$ and $\boldsymbol{d}_3(0) = \boldsymbol{e}_3$, the planar equilibria all lie in planes containing $\boldsymbol{e}_3$.

It follows from proposition 4.4 that the equilibria lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane are expressed by

$$\boldsymbol{r}^0(s) = \begin{pmatrix} \int_0^s \sin\phi(\sigma)\,d\sigma \\ 0 \\ \int_0^s \cos\phi(\sigma)\,d\sigma \end{pmatrix}, \quad \boldsymbol{n}^0(s) = \begin{pmatrix} n_1 \\ 0 \\ -\lambda \end{pmatrix},$$

$$\boldsymbol{d}_1^0(s) = \begin{pmatrix} \cos(\omega s)\cos\phi(s) \\ \sin(\omega s) \\ -\cos(\omega s)\sin\phi(s) \end{pmatrix}, \quad \boldsymbol{d}_2^0(s) = \begin{pmatrix} -\sin(\omega s)\cos\phi(s) \\ \cos(\omega s) \\ \sin(\omega s)\sin\phi(s) \end{pmatrix}, \tag{4.73}$$

$$\boldsymbol{d}_3^0(s) = \begin{pmatrix} \sin\phi(s) \\ 0 \\ \cos\phi(s) \end{pmatrix}, \quad \mathbf{m}^0(s) = \begin{pmatrix} K_1\phi'(s)\sin(\omega s) \\ K_1\phi'(s)\cos(\omega s) \\ 0 \end{pmatrix},$$

where $\phi(s)$ satisfies

$$\begin{cases} \phi''(s) = -\dfrac{\lambda}{K_1}\sin\phi(s) - \dfrac{n_1}{K_1}\cos\phi(s), \\ \phi(0) = 0, \quad \phi(1) = 0 \pmod{2\pi}, \quad \displaystyle\int_0^1 \sin\phi(s)\,ds = 0. \end{cases} \tag{4.74}$$

The solutions to this problem are discussed in appendix C.

The other planar equilibria of the clamped-sleeved strut are obtained by applying transformation (4.40) on the equilibria lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane. Finally, note that performing transformation (4.40) with $\theta = \pi$ to an equilibrium of the form (4.73) with $\phi(s) = \hat{\phi}(s)$, yields again an equilibrium of the form (4.73), but with $\phi(s) = -\hat{\phi}(s)$.

Ring

Consider an isotropic, uniform rod satisfying the ring boundary conditions (4.34). This problem has been the subject of study by many authors, and we defer to Dichmann *et al.* [16] for a discussion of the literature. Here we restrict our attention to the circular equilibria, i.e. the equilibria with a circular centerline. According to boundary conditions $\boldsymbol{r}(0) = \boldsymbol{0}$ and $\boldsymbol{d}_3(0) = \boldsymbol{e}_3$, the circular equilibria all lie in planes containing $\boldsymbol{e}_3$.

The circular equilibria lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane can be found with the help of proposition 4.4; they have the form

$$\boldsymbol{r}^0(s) = \frac{1}{2\pi l}\begin{pmatrix} 1-\cos(2\pi ls) \\ 0 \\ \sin(2\pi ls) \end{pmatrix}, \quad \boldsymbol{n}^0(s) = \begin{pmatrix} 0 \\ 2\pi l K_3(\alpha-\omega+2\pi j) \\ 0 \end{pmatrix},$$

$$\boldsymbol{d}_1^0(s) = \begin{pmatrix} \cos(\alpha s+2\pi js)\cos(2\pi ls) \\ \sin(\alpha s+2\pi js) \\ -\cos(\alpha s+2\pi js)\sin(2\pi ls) \end{pmatrix}, \quad \boldsymbol{d}_2^0(s) = \begin{pmatrix} -\sin(\alpha s+2\pi js)\cos(2\pi ls) \\ \cos(\alpha s+2\pi js) \\ \sin(\alpha s+2\pi js)\sin(2\pi ls) \end{pmatrix},$$

$$\boldsymbol{d}_3^0(s) = \begin{pmatrix} \sin(2\pi ls) \\ 0 \\ \cos(2\pi ls) \end{pmatrix}, \quad \mathbf{m}^0(s) = \begin{pmatrix} 2\pi l K_1 \sin(\alpha s+2\pi js) \\ 2\pi l K_1 \cos(\alpha s+2\pi js) \\ K_3(\alpha-\omega+2\pi j) \end{pmatrix}, \tag{4.75}$$

with $l = \pm1, \pm2, \ldots$, and $j = 0, \pm1, \pm2, \ldots$. In these equilibria, the centerline covers itself $|l|$ times, and the directors $\boldsymbol{d}_1(s)$ and $\boldsymbol{d}_2(s)$ rotate about the tangent $\boldsymbol{d}_3(s)$ to the centerline with a constant spin rate $u_3^0 = \alpha + 2\pi j$.

The BVP describing the isotropic, uniform ring is invariant under transformations (4.40) and (4.67). However, a straightforward computation shows that, when applied to a circular solution of the form (4.75), transformation (4.67) produces the same result as transformation (4.40) with $\theta = -(\alpha + 2\pi j)\beta \pmod{2\pi}$. Hence these two continuous symmetries only give rise to one-parameter families of solutions. In other words, all the circular equilibria of the ring can be obtained by applying transformation (4.40) on the circular equilibria lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane. Note that performing transformation (4.40) with $\theta = \pi$ to a circular equilibrium of the form (4.75) with $l = \hat{l}$, yields the corresponding circular equilibrium with $l = -\hat{l}$.

Chapter 5

Elastic rods in a Hamiltonian formulation

Applying the averaging results of chapter 3 to rods with high intrinsic twist requires a canonical Hamiltonian formulation of the equilibrium equations. A classic method to obtain such a formulation involves the parametrization of the directors, i.e. a parametrization of the group of proper rotations $SO(3)$. There exist several common choices for such a parametrization (cf. Hughes [26, chap. 2], for example). Euler angles are often adopted because they involve the minimum number of variables, namely three. However, they give rise to a polar singularity which can be the source of serious difficulties. We prefer using quaternions which provide a four-parameter, singular-free, two-to-one parametrization of $SO(3)$. A canonical Hamiltonian formulation of the rod equilibrium equations based upon a parametrization of the directors by quaternions can be found in Li and Maddocks [33] and in Dichmann *et al.* [16]. The present chapter introduces this formulation of the equilibrium equations, and gives a translation of some of the results of chapter 4 into the new variables for further use.

The Hamiltonian formulation of the rod equilibrium equations is presented in section 5.1. Appropriate expressions for the clamped-free strut, the clamped-sleeved strut and the ring boundary conditions in the Hamiltonian variables are derived in section 5.2. An isotropic rod subject to one of these sets of boundary conditions is invariant under the register transformation (4.40). The explicit expression of this transfomation is translated into the new variables in section 5.3. Finally, the planar equilibria of an isotropic, uniform rod under the clamped-free strut and the clamped-sleeved strut boundary conditions, as well as the circular equilibria of an isotropic, uniform rod under the ring boundary conditions are expressed in the Hamiltonian variables in section 5.4.

5.1 Equilibrium equations

This section outlines the Hamiltonian formulation of the rod equilibrium equations presented by Li and Maddocks [33] and Dichmann *et al.* [16], and discusses two additional continuous symmetries of the equations due to the parametrization of the directors by quaternions.

We parametrize the directors by elements of $\mathbb{R}^4$ with unit euclidean norm, called *quaternions* or *Euler parameters*. By Euler's theorem on the motion of a rigid body (cf. Goldstein [22, sect. 4-6]), for each $s \in [0,1]$, there exist an angle $\varphi(s)$ and a unit vector $\boldsymbol{a}(s) \in \mathbb{R}^3$ such that the director frame $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ can be obtained from the reference frame $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ by a rotation of angle $\varphi(s)$ about the oriented axis $\boldsymbol{a}(s)$. The associated quaternion

$$\boldsymbol{\gamma}(s) = \big(\gamma_1(s), \gamma_2(s), \gamma_3(s), \gamma_4(s)\big) \tag{5.1}$$

is defined by

$$\gamma_i(s) = a_i(s)\, \sin\frac{\varphi(s)}{2}, \quad i = 1,2,3, \quad \gamma_4(s) = \cos\frac{\varphi(s)}{2}, \tag{5.2}$$

where $a_i(s)$ denote the components of $\boldsymbol{a}(s)$ with respect to the fixed frame $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$. Note that if $\boldsymbol{\gamma}(s)$ corresponds to a rotation $(\boldsymbol{a}(s), \varphi(s))$, then $-\boldsymbol{\gamma}(s)$ corresponds to the equivalent rotation $(\boldsymbol{a}(s), \varphi(s)+2\pi)$. Thus there is a two-to-one correspondence between quaternions and rotations. The directors are expressed in terms of the quaternions as

$$\begin{gathered}
\boldsymbol{d}_1 = \begin{pmatrix} \gamma_1^2 - \gamma_2^2 - \gamma_3^2 + \gamma_4^2 \\ 2\gamma_1\gamma_2 + 2\gamma_3\gamma_4 \\ 2\gamma_1\gamma_3 - 2\gamma_2\gamma_4 \end{pmatrix}, \quad \boldsymbol{d}_2 = \begin{pmatrix} 2\gamma_1\gamma_2 - 2\gamma_3\gamma_4 \\ -\gamma_1^2 + \gamma_2^2 - \gamma_3^2 + \gamma_4^2 \\ 2\gamma_2\gamma_3 + 2\gamma_1\gamma_4 \end{pmatrix}, \\
\boldsymbol{d}_3 = \begin{pmatrix} 2\gamma_1\gamma_3 + 2\gamma_2\gamma_4 \\ 2\gamma_2\gamma_3 - 2\gamma_1\gamma_4 \\ -\gamma_1^2 - \gamma_2^2 + \gamma_3^2 + \gamma_4^2 \end{pmatrix}.
\end{gathered} \tag{5.3}$$

We define a new variable, the *impetus* $\boldsymbol{\mu} : [0,1] \to \mathbb{R}^4$, by

$$\boldsymbol{\mu}(s) = \nu(s)\boldsymbol{\gamma}(s) + 2\sum_{i=1}^{3}\left(\sum_{j=1}^{3} K_{ij}(s)\Big(2\boldsymbol{\gamma}'(s)^T\boldsymbol{B}_j\boldsymbol{\gamma}(s) - \hat{u}_j(s)\Big)\right)\boldsymbol{B}_i\boldsymbol{\gamma}(s), \tag{5.4}$$

where the function $\nu(s) \in \mathbb{R}$, which can be interpreted as a Lagrange multiplier (cf. Li and Maddocks [33]), is expressed at the end of this section, $K_{ij}(s)$ denote the entries of the stiffness matrix $\boldsymbol{K}(s)$, and $\boldsymbol{B}_i$ are the 4×4 skew-symmetric matrices

$$\boldsymbol{B}_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{B}_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{B}_3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}. \tag{5.5}$$

The components of the moment with respect to the directors are expressed as $m_i = \frac{1}{2}\boldsymbol{\mu}^T\boldsymbol{B}_i\boldsymbol{\gamma}$, so that

$$\mathbf{m} = \tfrac{1}{2}\big(\boldsymbol{\mu}^T\boldsymbol{B}_1\boldsymbol{\gamma}, \boldsymbol{\mu}^T\boldsymbol{B}_2\boldsymbol{\gamma}, \boldsymbol{\mu}^T\boldsymbol{B}_3\boldsymbol{\gamma}\big). \tag{5.6}$$

Defining the phase variable $\boldsymbol{x} = (\boldsymbol{r}, \boldsymbol{\gamma}, \boldsymbol{n}, \boldsymbol{\mu}) : [0,1] \to \mathbb{R}^{14}$, where the force $\boldsymbol{n}$ is conjugate to the centerline $\boldsymbol{r}$ and the impetus $\boldsymbol{\mu}$ is conjugate to the quaternion $\boldsymbol{\gamma}$, the equilibrium equations for the rod can be formulated as the canonical Hamiltonian system

$$\boldsymbol{x}'(s) = \boldsymbol{J}H_{\boldsymbol{x}}(\boldsymbol{x}(s), s), \tag{5.7}$$

where $\boldsymbol{J}$ is the 14×14 canonical symplectic matrix, and the Hamiltonian is

$$H(\boldsymbol{x}(s), s) = \tfrac{1}{2}\mathbf{m}(s)^T \boldsymbol{K}^{-1}(s)\mathbf{m}(s) + \mathbf{m}(s)^T \hat{\mathbf{u}}(s) + \boldsymbol{n}(s)^T \boldsymbol{d}_3(s). \tag{5.8}$$

The right hand side of this equation can be written in terms of the phase variables by substituting the expressions given by (5.6) and (5.3) for $\mathbf{m}(s)$ and $\boldsymbol{d}_3(s)$. Note that the Hamiltonian is simply a polynomial of degree four in the phase variables.

Equilibrium equation (5.7) is invariant under rigid translations and rotations of the rod in space. Due to the use of quaternions $\boldsymbol{\gamma} \in \mathbb{R}^4$ to parametrize the three-dimensional group SO(3), additional symmetries arise, namely invariance under the two transformations

$$(\boldsymbol{r}, \boldsymbol{\gamma}, \boldsymbol{n}, \boldsymbol{\mu}) \to (\boldsymbol{r}, \boldsymbol{\gamma}, \boldsymbol{n}, \boldsymbol{\mu} + \delta\boldsymbol{\gamma}), \quad \delta \in \mathbb{R}, \tag{5.9}$$

$$(\boldsymbol{r}, \boldsymbol{\gamma}, \boldsymbol{n}, \boldsymbol{\mu}) \to (\rho^2 \boldsymbol{r}, \rho\boldsymbol{\gamma}, \rho^{-2}\boldsymbol{n}, \rho^{-1}\boldsymbol{\mu}), \quad \rho \in \mathbb{R}\backslash\{0\}. \tag{5.10}$$

These symmetries are associated to the following two integrals. If $\boldsymbol{x} = (\boldsymbol{r}, \boldsymbol{\gamma}, \boldsymbol{n}, \boldsymbol{\mu})$ is a solution to equation (5.7), then

$$\boldsymbol{\gamma}(s)^T \boldsymbol{\gamma}(s) = \eta, \quad \forall\, s \in [0, 1], \tag{5.11}$$

and

$$\boldsymbol{\gamma}(s)^T \boldsymbol{\mu}(s) + 2\boldsymbol{r}(s)^T \boldsymbol{n}(s) = c, \quad \forall s \in [0, 1], \tag{5.12}$$

where η and c are constants. The definition of quaternions requires the equality $\eta = 1$. The boundary conditions will be chosen in such a way to enforce this equality, and thereby to break symmetry (5.10). The value of the constant c has no influence on the description of the equilibria of the rod; it merely corresponds to a gauge freedom in the impetus variable. The boundary conditions will be chosen in such a way to eliminate this gauge freedom by fixing the value of c, and thereby to break symmetry (5.9).

From definition (5.4) and from the fact that $\{\boldsymbol{\gamma}(s), \boldsymbol{B}_1\boldsymbol{\gamma}(s), \boldsymbol{B}_2\boldsymbol{\gamma}(s), \boldsymbol{B}_3\boldsymbol{\gamma}(s)\}$ forms an orthonormal basis for $\mathbb{R}^4$, we find that $\boldsymbol{\gamma}(s)^T \boldsymbol{\mu}(s) = \nu(s)$. Comparing this equality with (5.12) shows that the function $\nu(s)$ appearing in the definition of the impetus can be expressed as

$$\nu(s) = -2\boldsymbol{r}(s)^T \boldsymbol{n}(s) + c. \tag{5.13}$$

Finally, a consequence of the two-to-one correspondence between quaternions and rotations is that if $(\boldsymbol{r}, \boldsymbol{\gamma}, \boldsymbol{n}, \boldsymbol{\mu})$ is a solution to equilibrium equation (5.7), then $(\boldsymbol{r}, -\boldsymbol{\gamma}, \boldsymbol{n}, -\boldsymbol{\mu})$ is also a solution. However, these two solutions are equivalent in the sense that they describe the same physical equilibrium configuration of the rod.

5.2 Boundary conditions

In this section, we use integrals (5.11) and (5.12) of the Hamiltonian equilibrium equations to derive appropriate formulations of the clamped-free strut, the clamped-sleeved strut and the ring boundary conditions in the Hamiltonian variables. These formulations of the boundary conditions define Lagrangian manifolds (cf. definition 2.4) at both end points. The formulations of the clamped-free strut and the clamped-sleeved strut boundary conditions are original, whereas the formulation of the ring boundary conditions already appears in Li and Maddocks [33].

Clamped-free strut

A direct translation of the clamped-free strut boundary conditions (4.32) into the Hamiltonian variables gives

$$\begin{aligned} \boldsymbol{r}(0) = \mathbf{0}, \quad \gamma_1(0) = \gamma_2(0) = \gamma_3(0) = 0, \quad \gamma_4(0) = \pm 1, \\ \boldsymbol{n}(1) = (0, 0, -\lambda), \quad \boldsymbol{\mu}(1)^T \boldsymbol{B}_i \boldsymbol{\gamma}(1) = 0, \quad i = 1, 2, 3. \end{aligned} \tag{5.14}$$

Because of the equivalence of solutions $(\boldsymbol{r}, \boldsymbol{\gamma}, \boldsymbol{n}, \boldsymbol{\mu})$ and $(\boldsymbol{r}, -\boldsymbol{\gamma}, \boldsymbol{n}, -\boldsymbol{\mu})$ to BVP (5.7)+(5.14), the boundary condition $\gamma_4(0) = \pm 1$ can be replaced by $\gamma_4(0) = 1$ without discarding any equilibrium configuration of the system. To eliminate the gauge freedom in the variable $\boldsymbol{\mu}$, we enforce the boundary condition

$$\boldsymbol{\mu}(1)^T \boldsymbol{\gamma}(1) = 0, \tag{5.15}$$

which fixes the value of the constant in (5.12) to

$$c = 2\boldsymbol{r}(1)^T \boldsymbol{n}(1). \tag{5.16}$$

Note that the constant depends on the solution considered. Since $\{\boldsymbol{\gamma}(1), \boldsymbol{B}_1\boldsymbol{\gamma}(1), \boldsymbol{B}_2\boldsymbol{\gamma}(1), \boldsymbol{B}_3\boldsymbol{\gamma}(1)\}$ forms a basis of $\mathbb{R}^4$, the last equalities in (5.14) together with (5.15) are equivalent to the condition $\boldsymbol{\mu}(1) = \mathbf{0}$. Accordingly, the clamped-free strut boundary conditions can be written as

$$\begin{aligned} \boldsymbol{r}(0) = \mathbf{0}, \quad \boldsymbol{\gamma}(0) = (0, 0, 0, 1), \\ \boldsymbol{n}(1) = (0, 0, -\lambda), \quad \boldsymbol{\mu}(1) = \mathbf{0}. \end{aligned} \tag{5.17}$$

This form of the clamped-free strut boundary conditions in the Hamiltonian variables is different from the formulation of Li and Maddocks [33, equation 4.1]; it has the advantage of satisfying property (**2.A5**), and therefore of defining Lagrangian manifolds at $s = 0$ and at $s = 1$, as is shown is appendix B.

Clamped-sleeved strut

Expressing the clamped-sleeved strut boundary conditions (4.33) into the Hamiltonian variables gives

$$\begin{aligned} \boldsymbol{r}(0) = \mathbf{0}, \quad \gamma_1(0) = \gamma_2(0) = \gamma_3(0) = 0, \quad \gamma_4(0) = \pm 1, \\ r_1(1) = r_2(1) = 0, \quad n_3(1) = -\lambda, \quad \gamma_1(1) = \gamma_2(1) = 0, \\ \gamma_3^2(1) + \gamma_4^2(1) = 1, \quad \mu_3(1)\gamma_4(1) - \mu_4(1)\gamma_3(1) = 0. \end{aligned} \tag{5.18}$$

The condition $\gamma_4(0) = \pm 1$ can be omitted from BVP (5.7)+(5.18), since it is implied by integral (5.11) of Hamilton's equation together with boundary conditions $\gamma_1(1) = \gamma_2(1) = 0$, $\gamma_3^2(1) + \gamma_4^2(1) = 1$ and $\gamma_1(0) = \gamma_2(0) = \gamma_3(0) = 0$. To eliminate the gauge freedom in the variable $\boldsymbol{\mu}$, we add the condition $\mu_4(0) = 0$, which fixes the value of the constant in (5.12) to

$$c = 0. \tag{5.19}$$

Accordingly, the clamped-sleeved strut boundary conditions can be formulated as

$$\begin{gathered}\boldsymbol{r}(0)=\mathbf{0},\quad \gamma_1(0)=\gamma_2(0)=\gamma_3(0)=0,\quad \mu_4(0)=0,\\ r_1(1)=r_2(1)=0,\quad n_3(1)=-\lambda,\quad \gamma_1(1)=\gamma_2(1)=0,\\ \gamma_3^2(1)+\gamma_4^2(1)=1,\quad \mu_3(1)\gamma_4(1)-\mu_4(1)\gamma_3(1)=0.\end{gathered} \tag{5.20}$$

This set of boundary conditions satisfies property (**2.A5**) and therefore defines Lagrangian manifolds at $s=0$ and at $s=1$, as is shown is appendix B. Note that if $(\boldsymbol{r},\boldsymbol{\gamma},\boldsymbol{n},\boldsymbol{\mu})$ is a solution to BVP (5.7)+(5.20), then $(\boldsymbol{r},-\boldsymbol{\gamma},\boldsymbol{n},-\boldsymbol{\mu})$ is also a solution, and the two solutions are equivalent.

Ring

A direct translation of the ring boundary conditions (4.34) into the Hamiltonian variables yields

$$\begin{gathered}\boldsymbol{r}(0)=\mathbf{0},\quad \gamma_1(0)=\gamma_2(0)=\gamma_3(0)=0,\quad \gamma_4(0)=\pm 1,\\ \boldsymbol{r}(1)=\mathbf{0},\quad \boldsymbol{\gamma}(1)=\pm\big(0,0,\sin\tfrac{\alpha}{2},\cos\tfrac{\alpha}{2}\big).\end{gathered} \tag{5.21}$$

The condition $\gamma_4(0)=\pm 1$ can be omitted since it is implied by integral (5.11) together with boundary conditions $\boldsymbol{\gamma}(1)=\pm\big(0,0,\sin\frac{\alpha}{2},\cos\frac{\alpha}{2}\big)$ and $\gamma_1(0)=\gamma_2(0)=\gamma_3(0)=0$. Because of the equivalence of solutions $(\boldsymbol{r},\boldsymbol{\gamma},\boldsymbol{n},\boldsymbol{\mu})$ and $(\boldsymbol{r},-\boldsymbol{\gamma},\boldsymbol{n},-\boldsymbol{\mu})$ to BVP (5.7)+(5.21), the boundary condition $\boldsymbol{\gamma}(1)=\pm\big(0,0,\sin\frac{\alpha}{2},\cos\frac{\alpha}{2}\big)$ can be replaced by $\boldsymbol{\gamma}(1)=\big(0,0,\sin\frac{\alpha}{2},\cos\frac{\alpha}{2}\big)$ without discarding any equilibrium configuration of the system. Finally, to eliminate the gauge freedom in the variable $\boldsymbol{\mu}$, we add the condition $\mu_4(0)=0$, which fixes the value of the constant in (5.12) to

$$c=0. \tag{5.22}$$

Accordingly, the ring boundary conditions can be expressed as

$$\begin{gathered}\boldsymbol{r}(0)=\mathbf{0},\quad \gamma_1(0)=\gamma_2(0)=\gamma_3(0)=0,\quad \mu_4(0)=0,\\ \boldsymbol{r}(1)=\mathbf{0},\quad \boldsymbol{\gamma}(1)=\big(0,0,\sin\tfrac{\alpha}{2},\cos\tfrac{\alpha}{2}\big).\end{gathered} \tag{5.23}$$

This set of boundary conditions satisfies property (**2.A5**) and therefore defines Lagrangian manifolds at $s=0$ and at $s=1$, as is shown in appendix B.

5.3 Isotropic rods

Consider an isotropic rod subject to one of the three sets of boundary conditions discussed in section 5.2. This problem is invariant under transformation (4.40). In this section it is shown that, when applied to a configuration $\boldsymbol{x}_0$, this transformation produces configurations $\boldsymbol{x}$ defined by

$$\boldsymbol{x}(s)=\boldsymbol{R}(\theta)\boldsymbol{x}_0(s), \tag{5.24}$$

with

$$\boldsymbol{R}(\theta)=\operatorname{diag}\big(\boldsymbol{\Omega}(\theta),\boldsymbol{\Upsilon}(\theta),\boldsymbol{\Omega}(\theta),\boldsymbol{\Upsilon}(\theta)\big), \tag{5.25}$$

and

$$\boldsymbol{\Omega}(\theta)=\begin{pmatrix}\cos\theta & -\sin\theta & 0\\ \sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix},\quad \boldsymbol{\Upsilon}(\theta)=\begin{pmatrix}\cos\theta & -\sin\theta & 0 & 0\\ \sin\theta & \cos\theta & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}. \tag{5.26}$$

This result was obtained by Manning and Maddocks [37] in the context of the rod problem with ring boundary conditions.

We will take advantage of the fact that if two rotations $\mathcal{A}$ and $\mathcal{B}$ are parametrized by quaternions $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\alpha_3,\alpha_4)$ and $\boldsymbol{\beta}=(\beta_1,\beta_2,\beta_3,\beta_4)$ respectively, then the composition of the two rotations (performing $\mathcal{A}$ first) is simply described by the *quaternion product*

$$\boldsymbol{\beta}\odot\boldsymbol{\alpha}=\begin{pmatrix}\beta_4 & -\beta_3 & \beta_2 & \beta_1\\ \beta_3 & \beta_4 & -\beta_1 & \beta_2\\ -\beta_2 & \beta_1 & \beta_4 & \beta_3\\ -\beta_1 & -\beta_2 & -\beta_3 & \beta_4\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\\ \alpha_4\end{pmatrix}. \tag{5.27}$$

Let $\boldsymbol{x}=(\boldsymbol{r},\boldsymbol{\gamma},\boldsymbol{n},\boldsymbol{\mu})$ be the result of transformation (4.40) on $\boldsymbol{x}_0=(\boldsymbol{r}^0,\boldsymbol{\gamma}^0,\boldsymbol{n}^0,\boldsymbol{\mu}^0)$. Then $\boldsymbol{r}(s)$ and $\boldsymbol{n}(s)$ are given by (4.41). The quaternion $\boldsymbol{\gamma}(s)$ corresponds to the framing $(\boldsymbol{d}_1(s),\boldsymbol{d}_2(s),\boldsymbol{d}_3(s))$, which is obtained from the framing $(\boldsymbol{d}_1^0(s),\boldsymbol{d}_2^0(s),\boldsymbol{d}_3^0(s))$ by a rotation of angle θ about $\boldsymbol{e}_3$, followed by a rotation of angle $-\theta$ about $\boldsymbol{\Omega}(\theta)\boldsymbol{d}_3^0(s)$. Consequently,

$$\boldsymbol{\gamma}(s)=\begin{pmatrix}-\boldsymbol{\Omega}(\theta)\boldsymbol{d}_3^0(s)\sin\frac{\theta}{2}\\ \cos\frac{\theta}{2}\end{pmatrix}\odot\begin{pmatrix}0\\ 0\\ \sin\frac{\theta}{2}\\ \cos\frac{\theta}{2}\end{pmatrix}\odot\boldsymbol{\gamma}^0(s). \tag{5.28}$$

We substitute the expression

$$\boldsymbol{\Omega}(\theta)\boldsymbol{d}_3^0(s)=\begin{pmatrix}\cos\theta & -\sin\theta & 0\\ \sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}2\gamma_1^0(s)\gamma_3^0(s)+2\gamma_2^0(s)\gamma_4^0(s)\\ 2\gamma_2^0(s)\gamma_3^0(s)-2\gamma_1^0(s)\gamma_4^0(s)\\ -(\gamma_1^0)^2(s)-(\gamma_2^0)^2(s)+(\gamma_3^0)^2(s)+(\gamma_4^0)^2(s)\end{pmatrix} \tag{5.29}$$

into (5.28), perform the quaternion products, simplify with the help of the identity $(\gamma_1^0)^2(s)+(\gamma_2^0)^2(s)+(\gamma_3^0)^2(s)+(\gamma_4^0)^2(s)=1$, and find

$$\boldsymbol{\gamma}(s)=\boldsymbol{\Upsilon}(\theta)\boldsymbol{\gamma}^0(s). \tag{5.30}$$

An expression for the impetus is given by substituting (5.13) into (5.4). The constant c appearing in this expression is invariant under transformation (4.40). Indeed, for the clamped-free strut boundary conditions,

$$c=2\boldsymbol{r}(1)^T\boldsymbol{n}(1)=2\boldsymbol{r}^0(1)^T\boldsymbol{\Omega}(\theta)^T\boldsymbol{\Omega}(\theta)\boldsymbol{n}^0(1)=2\boldsymbol{r}^0(1)^T\boldsymbol{n}^0(1)=c^0, \tag{5.31}$$

while, for the clamped-sleeved strut and ring boundary conditions,

$$c=0=c^0. \tag{5.32}$$

Denoting the stiffness matrix by

$$\boldsymbol{K}(s) = \operatorname{diag}\big(K_1(s), K_1(s), K_3(s)\big), \tag{5.33}$$

and expressing $\boldsymbol{r}(s)$, $\boldsymbol{n}(s)$, and $\boldsymbol{\gamma}(s)$ as in (4.41) and (5.30), the corresponding impetus takes the form

$$\boldsymbol{\mu}(s) = \boldsymbol{\mu}_{\mathrm{I}}(s) + 2K_1(s)\boldsymbol{\mu}_{\mathrm{II}}(s), \tag{5.34}$$

with

$$\begin{aligned}\boldsymbol{\mu}_{\mathrm{I}}(s) =& \Big(- 2\boldsymbol{r}^0(s)^T\boldsymbol{\Omega}(\theta)^T\boldsymbol{\Omega}(\theta)\boldsymbol{n}^0(s) + c^0\Big)\boldsymbol{\Upsilon}(\theta)\boldsymbol{\gamma}^0(s)\\ &+ 2K_3(s)\Big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{\Upsilon}(\theta)^T\boldsymbol{B}_3\boldsymbol{\Upsilon}(\theta)\boldsymbol{\gamma}^0(s) - \hat{u}_3(s)\Big)\boldsymbol{B}_3\boldsymbol{\Upsilon}(\theta)\boldsymbol{\gamma}^0(s),\end{aligned} \tag{5.35}$$

$$\begin{aligned}\boldsymbol{\mu}_{\mathrm{II}}(s) =& \Big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{\Upsilon}(\theta)^T\boldsymbol{B}_1\boldsymbol{\Upsilon}(\theta)\boldsymbol{\gamma}^0(s)\Big)\boldsymbol{B}_1\boldsymbol{\Upsilon}(\theta)\boldsymbol{\gamma}^0(s)\\ &+ \Big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{\Upsilon}(\theta)^T\boldsymbol{B}_2\boldsymbol{\Upsilon}(\theta)\boldsymbol{\gamma}^0(s)\Big)\boldsymbol{B}_2\boldsymbol{\Upsilon}(\theta)\boldsymbol{\gamma}^0(s).\end{aligned} \tag{5.36}$$

With the help of equalities

$$\boldsymbol{\Upsilon}(\theta)^T\boldsymbol{B}_3\boldsymbol{\Upsilon}(\theta) = \boldsymbol{B}_3, \quad \boldsymbol{B}_3\boldsymbol{\Upsilon}(\theta) = \boldsymbol{\Upsilon}(\theta)\boldsymbol{B}_3, \tag{5.37}$$

we find

$$\begin{aligned}\boldsymbol{\mu}_{\mathrm{I}}(s) =& \boldsymbol{\Upsilon}(\theta)\Big(- 2\boldsymbol{r}^0(s)^T\boldsymbol{n}^0(s) + c^0\Big)\boldsymbol{\gamma}^0(s)\\ &+ \boldsymbol{\Upsilon}(\theta)2K_3(s)\Big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{B}_3\boldsymbol{\gamma}^0(s) - \hat{u}_3(s)\Big)\boldsymbol{B}_3\boldsymbol{\gamma}^0(s).\end{aligned} \tag{5.38}$$

Using identities

$$\begin{aligned}\boldsymbol{\Upsilon}(\theta)^T\boldsymbol{B}_1\boldsymbol{\Upsilon}(\theta) &= \cos\theta\boldsymbol{B}_1 - \sin\theta\boldsymbol{B}_2,\\ \boldsymbol{\Upsilon}(\theta)^T\boldsymbol{B}_2\boldsymbol{\Upsilon}(\theta) &= \sin\theta\boldsymbol{B}_1 + \cos\theta\boldsymbol{B}_2,\end{aligned} \tag{5.39}$$

and

$$\begin{aligned}\cos\theta\boldsymbol{B}_1\boldsymbol{\Upsilon}(\theta) + \sin\theta\boldsymbol{B}_2\boldsymbol{\Upsilon}(\theta) &= \boldsymbol{\Upsilon}(\theta)\boldsymbol{B}_1,\\ -\sin\theta\boldsymbol{B}_1\boldsymbol{\Upsilon}(\theta) + \cos\theta\boldsymbol{B}_2\boldsymbol{\Upsilon}(\theta) &= \boldsymbol{\Upsilon}(\theta)\boldsymbol{B}_2,\end{aligned} \tag{5.40}$$

we obtain

$$\begin{aligned}\boldsymbol{\mu}_{\mathrm{II}}(s) \overset{(5.39)}{=}& \Big(\big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{B}_1\boldsymbol{\gamma}^0(s)\big)\cos\theta - \big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{B}_2\boldsymbol{\gamma}^0(s)\big)\sin\theta\Big)\boldsymbol{B}_1\boldsymbol{\Upsilon}(\theta)\boldsymbol{\gamma}^0(s)\\ &+ \Big(\big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{B}_1\boldsymbol{\gamma}^0(s)\big)\sin\theta - \big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{B}_2\boldsymbol{\gamma}^0(s)\big)\cos\theta\Big)\boldsymbol{B}_2\boldsymbol{\Upsilon}(\theta)\boldsymbol{\gamma}^0(s)\\ =& \Big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{B}_1\boldsymbol{\gamma}^0(s)\Big)\Big(\cos\theta\boldsymbol{B}_1\boldsymbol{\Upsilon}(\theta) + \sin\theta\boldsymbol{B}_2\boldsymbol{\Upsilon}(\theta)\Big)\boldsymbol{\gamma}^0(s)\\ &+ \Big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{B}_2\boldsymbol{\gamma}^0(s)\Big)\Big(- \sin\theta\boldsymbol{B}_1\boldsymbol{\Upsilon}(\theta) + \cos\theta\boldsymbol{B}_2\boldsymbol{\Upsilon}(\theta)\Big)\boldsymbol{\gamma}^0(s)\\ \overset{(5.40)}{=}& \boldsymbol{\Upsilon}(\theta)\Big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{B}_1\boldsymbol{\gamma}^0(s)\Big)\boldsymbol{B}_1\boldsymbol{\gamma}^0(s)\\ &+ \boldsymbol{\Upsilon}(\theta)\Big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{B}_2\boldsymbol{\gamma}^0(s)\Big)\boldsymbol{B}_2\boldsymbol{\gamma}^0(s).\end{aligned} \tag{5.41}$$

These calculations lead to the relation

$$\boldsymbol{\mu}(s) = \boldsymbol{\Upsilon}(\theta)\boldsymbol{\mu}^0(s). \tag{5.42}$$

Finally, gathering expressions (4.41), (5.30) and (5.42) for $\boldsymbol{r}(s)$, $\boldsymbol{n}(s)$, $\boldsymbol{\gamma}(s)$ and $\boldsymbol{\mu}(s)$ establishes result (5.24).

5.4 Equilibria of isotropic, uniform rods

Consider an isotropic, uniform rod. Section 4.6 presents expressions for the planar equilibria of the rod subject to the clamped-free strut and the clamped-sleeved strut boundary conditions, as well as for the circular equilibria of the rod under the ring boundary conditions. In the present section we rewrite these equilibria in the Hamiltonian variables. The stiffness matrix and the intrinsic strains of the isotropic, uniform rod are denoted by

$$\boldsymbol{K}(s) = \operatorname{diag}(K_1, K_1, K_3), \quad \hat{\mathbf{u}}(s) = (0, 0, \omega). \tag{5.43}$$

Clamped-free strut

Consider an isotropic, uniform rod subject to the clamped-free strut boundary conditions. The equilibria of the system are governed by BVP (5.7)+(5.17).

The equilibria lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane are expressed by $\boldsymbol{x}_0 = (\boldsymbol{r}^0, \boldsymbol{\gamma}^0, \boldsymbol{n}^0, \boldsymbol{\mu}^0)$, where $\boldsymbol{r}^0(s)$ and $\boldsymbol{n}^0(s)$ are given in (4.71). The quaternion $\boldsymbol{\gamma}^0(s)$ corresponds to the framing $(\boldsymbol{d}_1^0(s), \boldsymbol{d}_2^0(s), \boldsymbol{d}_3^0(s))$ expressed in (4.71). This framing is obtained from the frame $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ by a rotation of angle $\phi(s)$ about $\boldsymbol{e}_2$, followed by a rotation of angle ωs about $\boldsymbol{d}_3^0(s)$. Therefore,

$$\boldsymbol{\gamma}^0(s) = \begin{pmatrix} \sin\phi(s)\sin\frac{\omega s}{2} \\ 0 \\ \cos\phi(s)\sin\frac{\omega s}{2} \\ 0 \end{pmatrix} \odot \begin{pmatrix} 0 \\ \sin\frac{\phi(s)}{2} \\ 0 \\ \cos\frac{\phi(s)}{2} \end{pmatrix} = \begin{pmatrix} \sin\frac{\omega s}{2}\sin\frac{\phi(s)}{2} \\ \cos\frac{\omega s}{2}\sin\frac{\phi(s)}{2} \\ \sin\frac{\omega s}{2}\cos\frac{\phi(s)}{2} \\ \cos\frac{\omega s}{2}\cos\frac{\phi(s)}{2} \end{pmatrix}. \tag{5.44}$$

The corresponding impetus is given by

$$\begin{aligned} \boldsymbol{\mu}^0(s) =& \Big(-2\boldsymbol{r}^0(s)^T\boldsymbol{n}^0(s) + 2\boldsymbol{r}^0(1)^T\boldsymbol{n}^0(1)\Big)\boldsymbol{\gamma}^0(s) \\ &+ 2K_1\Big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{B}_1\boldsymbol{\gamma}^0(s)\Big)\boldsymbol{B}_1\boldsymbol{\gamma}^0(s) + 2K_1\Big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{B}_2\boldsymbol{\gamma}^0(s)\Big)\boldsymbol{B}_2\boldsymbol{\gamma}^0(s) \\ &+ 2K_3\Big(2\boldsymbol{\gamma}^{0\prime}(s)^T\boldsymbol{B}_3\boldsymbol{\gamma}^0(s) - \omega\Big)\boldsymbol{B}_3\boldsymbol{\gamma}^0(s). \end{aligned} \tag{5.45}$$

A straightforward computation then yields

$$\boldsymbol{\mu}^0(s) = -2\lambda\int_s^1 \cos\phi(\sigma)\,d\sigma \begin{pmatrix} \sin\frac{\omega s}{2}\sin\frac{\phi(s)}{2} \\ \cos\frac{\omega s}{2}\sin\frac{\phi(s)}{2} \\ \sin\frac{\omega s}{2}\cos\frac{\phi(s)}{2} \\ \cos\frac{\omega s}{2}\cos\frac{\phi(s)}{2} \end{pmatrix} + 2K_1\phi'(s) \begin{pmatrix} \sin\frac{\omega s}{2}\cos\frac{\phi(s)}{2} \\ \cos\frac{\omega s}{2}\cos\frac{\phi(s)}{2} \\ -\sin\frac{\omega s}{2}\sin\frac{\phi(s)}{2} \\ -\cos\frac{\omega s}{2}\sin\frac{\phi(s)}{2} \end{pmatrix}. \tag{5.46}$$

The other equilibria of the system are obtained by performing transformation (4.40) on the equilibria $\boldsymbol{x}_0$ lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane. According to section 5.3, they are of the form $\boldsymbol{x} = \boldsymbol{R}(\theta)\boldsymbol{x}_0$.

Clamped-sleeved strut

Consider an isotropic, uniform rod subject to the clamped-sleeved strut boundary conditions. The equilibria of the system satisfy BVP (5.7)+(5.20). As mentioned after equation (5.20), if $(\boldsymbol{r}, \boldsymbol{\gamma}, \boldsymbol{n}, \boldsymbol{\mu})$ is a solution to this BVP, then $(\boldsymbol{r}, -\boldsymbol{\gamma}, \boldsymbol{n}, -\boldsymbol{\mu})$ is also a solution, and the two solutions describe the same equilibrium configuration of the rod. To discard one of these equivalent solutions, we only consider the solutions satisfying $\gamma_4(0) = 1$ in what follows.

The equilibria lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane are expressed by $\boldsymbol{x}_0 = (\boldsymbol{r}^0, \boldsymbol{\gamma}^0, \boldsymbol{n}^0, \boldsymbol{\mu}^0)$, where $\boldsymbol{r}^0(s)$ and $\boldsymbol{n}^0(s)$ are given in (4.73). A computation similar to the one leading to equations (5.44) and (5.46) shows that

$$\boldsymbol{\gamma}^0(s) = \begin{pmatrix} \sin\frac{\omega s}{2}\sin\frac{\phi(s)}{2} \\ \cos\frac{\omega s}{2}\sin\frac{\phi(s)}{2} \\ \sin\frac{\omega s}{2}\cos\frac{\phi(s)}{2} \\ \cos\frac{\omega s}{2}\cos\frac{\phi(s)}{2} \end{pmatrix}, \tag{5.47}$$

and

$$\begin{aligned}\boldsymbol{\mu}^0(s) = &\left(2\lambda\int_0^s \cos\phi(\sigma)\,d\sigma - 2n_1\int_0^s \sin\phi(\sigma)\,d\sigma\right)\begin{pmatrix} \sin\frac{\omega s}{2}\sin\frac{\phi(s)}{2} \\ \cos\frac{\omega s}{2}\sin\frac{\phi(s)}{2} \\ \sin\frac{\omega s}{2}\cos\frac{\phi(s)}{2} \\ \cos\frac{\omega s}{2}\cos\frac{\phi(s)}{2} \end{pmatrix} \\ &+ 2K_1\phi'(s)\begin{pmatrix} \sin\frac{\omega s}{2}\cos\frac{\phi(s)}{2} \\ \cos\frac{\omega s}{2}\cos\frac{\phi(s)}{2} \\ -\sin\frac{\omega s}{2}\sin\frac{\phi(s)}{2} \\ -\cos\frac{\omega s}{2}\sin\frac{\phi(s)}{2} \end{pmatrix}.\end{aligned} \tag{5.48}$$

The other planar equilibria of the system are obtained by applying transformation (4.40) on the equilibria $\boldsymbol{x}_0$ lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane. According to section 5.3, they take the form $\boldsymbol{x} = \boldsymbol{R}(\theta)\boldsymbol{x}_0$.

Ring

Consider an isotropic, uniform rod subject to the ring boundary conditions. The equilibria of the system correspond to the solutions of BVP (5.7)+(5.23).

The equilibria lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane are expressed by $\boldsymbol{x}_0 = (\boldsymbol{r}^0, \boldsymbol{\gamma}^0, \boldsymbol{n}^0, \boldsymbol{\mu}^0)$, where $\boldsymbol{r}^0(s)$ and $\boldsymbol{n}^0(s)$ are given in (4.75). A computation similar to the one leading to equations (5.44) and (5.46) shows that

$$\boldsymbol{\gamma}^0(s) = (-1)^{l+j} \begin{pmatrix} \sin\left(\frac{\alpha s}{2} + \pi j s\right)\sin(\pi l s) \\ \cos\left(\frac{\alpha s}{2} + \pi j s\right)\sin(\pi l s) \\ \sin\left(\frac{\alpha s}{2} + \pi j s\right)\cos(\pi l s) \\ \cos\left(\frac{\alpha s}{2} + \pi j s\right)\cos(\pi l s) \end{pmatrix}, \tag{5.49}$$

and

$$\begin{aligned} \boldsymbol{\mu}^0(s) =& (-1)^{l+j} 4\pi l K_1 \begin{pmatrix} \sin\left(\frac{\alpha s}{2} + \pi j s\right)\cos(\pi l s) \\ \cos\left(\frac{\alpha s}{2} + \pi j s\right)\cos(\pi l s) \\ -\sin\left(\frac{\alpha s}{2} + \pi j s\right)\sin(\pi l s) \\ -\cos\left(\frac{\alpha s}{2} + \pi j s\right)\sin(\pi l s) \end{pmatrix} \\ &+ (-1)^{l+j} 2K_3(\alpha - \omega + 2\pi j) \begin{pmatrix} \cos\left(\frac{\alpha s}{2} + \pi j s\right)\cos(\pi l s) \\ -\sin\left(\frac{\alpha s}{2} + \pi j s\right)\sin(\pi l s) \\ \cos\left(\frac{\alpha s}{2} + \pi j s\right)\cos(\pi l s) \\ -\sin\left(\frac{\alpha s}{2} + \pi j s\right)\cos(\pi l s) \end{pmatrix}. \end{aligned} \tag{5.50}$$

The other circular equilibria of the system are obtained by performing transformation (4.40) on the circular equilibria $\boldsymbol{x}_0$ lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane. According to section 5.3, they take the form $\boldsymbol{x} = \boldsymbol{R}(\theta)\boldsymbol{x}_0$.

Chapter 6

Elastic rods with high intrinsic twist

Is it possible to approximate the equilibria of a constitutively anisotropic rod with a high, but finite, intrinsic twist subject to two-point boundary conditions by the equilibria of an effective, constitutively isotropic rod? We examine this question in light of the averaging results of chapter 3. It turns out that the magnitude of the intrinsic curvature of the rod is crucial for this question, and we distinguish the cases of an intrinsically curved rod with high intrinsic twist, of an intrinsically straight rod with high intrinsic twist, and of a rod with small intrinsic curvature and high intrinsic twist (for which the intrinsic curvature and the inverse of the intrinsic twist are of the same order of magnitude).

Section 6.1 gives a general definition of a rod with high intrinsic twist. In section 6.2, the problem governing the equilibria of a rod with high intrinsic twist subject to two-point boundary conditions is rewritten in a form suitable for averaging. Section 6.3 presents the effective rod, which is identical to the rod with high intrinsic twist, except for an effective stiffness matrix that makes it constitutively isotropic. The case of an intrinsically curved rod with high intrinsic twist, subject to clamped-free strut, clamped-sleeved strut or ring boundary conditions, is examined in section 6.4. In general, the equilibria of the associated effective rod are isolated and constitute approximations to the equilibria of the rod with high intrinsic twist.

The case of an intrinsically straight rod with high intrinsic twist, subject to clamped-free strut, clamped-sleeved strut or ring boundary conditions, is examined in section 6.5. The associated effective rod is isotropic and typically admits one-parameter families of equilibria. In general, only a finite number of members of such a family constitute approximations to equilibria of the rod with high intrinsic twist. These special members, which can be identified by means of a Melnikov function, are at least two in the case of a general stiffness matrix, and at least four in the case of a stiffness matrix with a certain block-diagonal form. In section 6.6, we explicitly determine the symmetry breaking in particular examples involving the clamped-free strut, the clamped-sleeved strut and the ring boundary conditions.

The case of a rod with small intrinsic curvature and high intrinsic twist, subject to clamped-free strut, clamped-sleeved strut or ring boundary conditions, is examined in section 6.7. A precise definition of such a rod is presented. The associated effective rod is isotropic, and typically admits one-parameter families of equilibria. In general, only a finite number, but at least two members of such a family constitute approximations to equilibria of the rod. The Melnikov function identifying these special members is the sum of two terms, one associated to intrinsic curvature and the other to intrinsic twist, which gives rise to a competition between the different effects in the determination of the symmetry breaking. In section 6.8, this competition is illustrated by explicit computation of the Melnikov function in a particular example involving the ring boundary conditions.

In section 6.9, we consider an intrinsically straight and twisted clamped-free strut with a diagonal stiffness matrix. By studying the linearization of the corresponding BVP about the trivial solution, we obtain and expression determining the buckling loads of the strut, for all values of the intrinsic twist. Then, for large values of the intrinsic twist, we estimate the error associated with the approximation of the first buckling load by the first buckling load of the corresponding effective strut. This error is proportional both to the inverse of the intrinsic twist and to a quantity measuring the anisotropy of the constitutive relation.

6.1 Definition

This section contains a precise definition of what we describe as a rod with high intrinsic twist. According to section 4.1, the intrinsic properties of a rod can be described by an intrinsic framing $(\hat{\boldsymbol{d}}_1(s), \hat{\boldsymbol{d}}_2(s), \hat{\boldsymbol{d}}_3(s))$ together with a stiffness matrix $\boldsymbol{K}(s)$. However, as shown in section 4.2, for a given rod, these quantities depend on the choice of the framing $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$.

Definition 6.1 *A rod is said to have a* high intrinsic twist *if it is constitutively anisotropic (cf. definition 4.1) and if, for a suitable choice of framing* $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$, *the following two properties hold.*

(1) The evolution of the intrinsic framing $\big(\hat{\boldsymbol{d}}_1(s), \hat{\boldsymbol{d}}_2(s), \hat{\boldsymbol{d}}_3(s)\big)$ *with respect to arclength consists in a rapid spin about* $\hat{\boldsymbol{d}}_3(s)$ *superimposed on a slow variation. More precisely,*

$$\big(\hat{\boldsymbol{d}}_1(s), \hat{\boldsymbol{d}}_2(s), \hat{\boldsymbol{d}}_3(s)\big) = \big(\hat{\boldsymbol{g}}_1(s), \hat{\boldsymbol{g}}_2(s), \hat{\boldsymbol{g}}_3(s)\big)\, \boldsymbol{\Omega}(\tau s), \tag{6.1}$$

where $\boldsymbol{\Omega}$ *is the rotation matrix defined in (4.16),* τ *is a large (positive or negative) parameter, and* $(\hat{\boldsymbol{g}}_1(s), \hat{\boldsymbol{g}}_2(s), \hat{\boldsymbol{g}}_3(s))$ *is a right-handed, orthonormal frame independent of* τ *(see figure 6.1).*

(2) The stiffness matrix $\boldsymbol{K}(s)$ *is independent of* τ.

This definition is essentially the same as the one introduced by Kehrbaum [27, sect. 9.6]; it only differs by the fact that it takes into account the possibility of different choices for the framing of a given rod. For a suitable choice of framing, the intrinsic properties of a rod with high intrinsic twist are described by quantities depending on two different length-scales: a short scale (with characteristic length $1/\tau$) over which the directors rotate about the centerline, and a long scale (with characteristic length

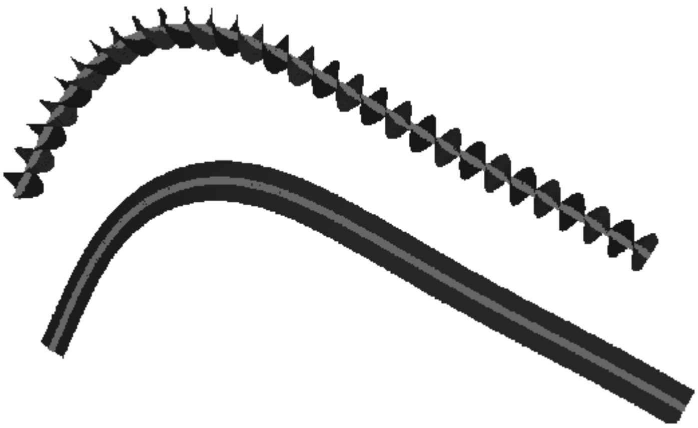


Figure 6.1: Top: *Intrinsic configuration of a rod with high intrinsic twist. The ribbon indicates the direction of* $\hat{\boldsymbol{d}}_1(s)$, *and* $\hat{\boldsymbol{d}}_3(s)$ *is tangent to the centerline.* Bottom: *Representation of the corresponding slowly varying frame* $(\hat{\boldsymbol{g}}_1(s), \hat{\boldsymbol{g}}_2(s), \hat{\boldsymbol{g}}_3(s))$. *The ribbon indicates the direction of* $\hat{\boldsymbol{g}}_1(s)$, *and* $\hat{\boldsymbol{g}}_3(s)$ *is tangent to the centerline.*

1) over which the tangent to the centerline and the stiffness matrix vary. Condition (2) imposes some correlation between the director $\boldsymbol{d}_1(s)$ and the orientation of the stiff plane for bending, at least locally, i.e. over the short scale. Thus, by condition (1), a rod with high intrinsic twist is characterized by a rapid rotation of the stiff plane about the centerline. Given a rod with high intrinsic twist, the framings such that conditions (1) and (2) hold are most natural and, in the following, we assume that the framing denoted by $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ is chosen such that these conditions hold.

Introduce the strains $\mathbf{u}(s) = (u_1(s), u_2(s), u_3(s))$ corresponding to a framing $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ of a rod with high intrinsic twist, and define the triple $\hat{\mathbf{v}}(s) = (\hat{v}_1(s), \hat{v}_2(s), \hat{v}_3(s))$ by

$$\hat{v}_1(s) = -\hat{\boldsymbol{g}}_3{}'(s)^T \hat{\boldsymbol{g}}_2(s), \quad \hat{v}_2(s) = \hat{\boldsymbol{g}}_3{}'(s)^T \hat{\boldsymbol{g}}_1(s), \quad \hat{v}_3(s) = \hat{\boldsymbol{g}}_1{}'(s)^T \hat{\boldsymbol{g}}_2(s), \tag{6.2}$$

where $(\hat{\boldsymbol{g}}_1(s), \hat{\boldsymbol{g}}_2(s), \hat{\boldsymbol{g}}_3(s))$ is the frame appearing in equation (6.1). In section 4.2 it is established that change of framing (4.15) leads to relation (4.24) between the associated strains. A similar computation based upon relation (6.1) shows that the intrinsic strains $\hat{\mathbf{u}}(s)$ are expressed by

$$\hat{\mathbf{u}}(s) = \boldsymbol{\Omega}^T(\tau s)\big(\hat{\mathbf{v}}(s) + \tau \mathbf{e}_3\big). \tag{6.3}$$

Note that, by definition, $\hat{\mathbf{v}}(s)$ is independent of τ. The name *rod with high intrinsic twist* is adopted since, for a choice of framing such that conditions (1) and (2) hold, the intrinsic twist $\hat{u}_3(s) = \hat{v}_3(s) + \tau$ is large.

6.2 Reformulation of the problem

Consider a rod with high intrinsic twist subject to one of the three sets of boundary conditions introduced in section 4.3. The purpose of this section is to rewrite the BVP governing the equilibria of the rod in a form suitable for applying the averaging results of chapter 3. This is achieved in two steps: a transformation of framing, and the introduction of a Hamiltonian formulation of the problem.

Given a rod with high intrinsic twist, define the number

$$\omega = \tau \pmod{2\pi} \in [0, 2\pi), \tag{6.4}$$

and the small parameter

$$\epsilon = \frac{1}{\tau - \omega} \in \mathbb{E}_{2\pi}, \tag{6.5}$$

where τ is the large parameter appearing in definition 6.1, and the set $\mathbb{E}_{2\pi}$ is defined in (3.6). Rotate the original framing $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ about $\boldsymbol{d}_3(s)$ through an angle $\frac{s}{\epsilon}$ to obtain the new framing

$$\big(\boldsymbol{D}_1(s), \boldsymbol{D}_2(s), \boldsymbol{D}_3(s)\big) = \big(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s)\big)\, \boldsymbol{\Omega}^T(\tfrac{s}{\epsilon}). \tag{6.6}$$

Note that this change of framing, which is slightly different from the change of framing consisting of a rotation of $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ about $\boldsymbol{d}_3(s)$ through an angle τs used by Kehrbaum [27], is more adapted for the treatment of two-point BVPs since it does not introduce any dependence on the value of τ in the boundary conditions (cf. last paragraph of this section). The equilibrium equations of the rod in terms of the new variables associated with the framing $(\boldsymbol{D}_1(s), \boldsymbol{D}_2(s), \boldsymbol{D}_3(s))$ are given in section 4.2. Denote the new strains by $\mathbf{U}(s) = (U_1(s), U_2(s), U_3(s))$, and the new triple of components of the moment by $\mathbf{M}(s) = (M_1(s), M_2(s), M_3(s))$. According to equation (4.31), the new stiffness matrix $\boldsymbol{L}(s, \frac{s}{\epsilon})$ is defined by

$$\boldsymbol{L}(s, \sigma) = \boldsymbol{\Omega}(\sigma)\, \boldsymbol{K}(s)\, \boldsymbol{\Omega}^T(\sigma), \tag{6.7}$$

where $\boldsymbol{K}(s)$ is the stiffness matrix associated with the original framing. Combining equations (6.1) and (6.6) shows that the new intrinsic framing is given by

$$\big(\hat{\boldsymbol{D}}_1(s), \hat{\boldsymbol{D}}_2(s), \hat{\boldsymbol{D}}_3(s)\big) = \big(\hat{\boldsymbol{g}}_1(s), \hat{\boldsymbol{g}}_2(s), \hat{\boldsymbol{g}}_3(s)\big)\, \boldsymbol{\Omega}(\omega s). \tag{6.8}$$

Then, by a computation similar to the one leading to equation (4.24), the new intrinsic strains are

$$\hat{\mathbf{U}}(s) = \boldsymbol{\Omega}^T(\omega s)\big(\hat{\mathbf{v}}(s) + \omega \mathbf{e}_3\big), \tag{6.9}$$

where $\hat{\mathbf{v}}(s)$ is defined in (6.2).

Following section 5.1, we now transform the equilibrium equations expressed in the new variables into a canonical Hamiltonian system. Introduce the quaternion variable $\boldsymbol{\gamma}$ parametrizing the new directors $(\boldsymbol{D}_1, \boldsymbol{D}_2, \boldsymbol{D}_3)$ and such that, in particular,

$$\boldsymbol{D}_3 = \begin{pmatrix} 2\gamma_1\gamma_3 + 2\gamma_2\gamma_4 \\ 2\gamma_2\gamma_3 - 2\gamma_1\gamma_4 \\ -\gamma_1^2 - \gamma_2^2 + \gamma_3^2 + \gamma_4^2 \end{pmatrix}. \tag{6.10}$$

The components of the moment relative to the new framing are expressed by

$$\mathbf{M} = \tfrac{1}{2}\big(\boldsymbol{\mu}^T \boldsymbol{B}_1 \boldsymbol{\gamma}, \boldsymbol{\mu}^T \boldsymbol{B}_2 \boldsymbol{\gamma}, \boldsymbol{\mu}^T \boldsymbol{B}_3 \boldsymbol{\gamma}\big), \tag{6.11}$$

where $\boldsymbol{\mu}$ is the impetus variable, conjugate to $\boldsymbol{\gamma}$. Introduce the phase variable $\boldsymbol{x} = (\boldsymbol{r}, \boldsymbol{\gamma}, \boldsymbol{n}, \boldsymbol{\mu})$, where $\boldsymbol{r}$ and $\boldsymbol{n}$ represent the centerline and the force respectively. The equilibrium equations for the rod with high intrinsic twist then become

$$\boldsymbol{x}'(s) = \boldsymbol{J} H_{\boldsymbol{x}}(\boldsymbol{x}(s), s, \tfrac{s}{\epsilon}), \tag{6.12}$$

with Hamiltonian

$$H(\boldsymbol{x}(s), s, \tfrac{s}{\epsilon}) = \tfrac{1}{2}\mathbf{M}(s)^T \boldsymbol{C}(s, \tfrac{s}{\epsilon})\mathbf{M}(s) + \mathbf{M}(s)^T \hat{\mathbf{U}}(s) + \boldsymbol{n}(s)^T \boldsymbol{D}_3(s), \tag{6.13}$$

where the flexibility matrix $\boldsymbol{C}(s, \frac{s}{\epsilon})$ is defined by

$$\boldsymbol{C}(s, \sigma) = \boldsymbol{L}^{-1}(s, \sigma) = \boldsymbol{\Omega}(\sigma)\boldsymbol{K}^{-1}(s)\boldsymbol{\Omega}^T(\sigma). \tag{6.14}$$

The quantities $\mathbf{M}(s)$ and $\boldsymbol{D}_3(s)$ appearing in the Hamiltonian are expressed in terms of the phase variables in equations (6.11) and (6.10).

Consider each of the three sets of boundary conditions introduced in section 4.3. According to equation (6.6), and since $\epsilon \in \mathbb{E}_{2\pi}$, the new framing $(\boldsymbol{D}_1(s), \boldsymbol{D}_2(s), \boldsymbol{D}_3(s))$ coincides with the original framing $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ both at $s = 0$ and at $s = 1$. It follows that each of these sets of boundary conditions takes an identical form in the two framings. In the Hamiltonian variables associated with the new framing, these sets of boundary conditions can therefore be expressed as in section 5.2.

6.3 Effective rods

Averaging the equilibrium equations for the rod with high intrinsic twist leads to the definition of an effective rod, with the property that the equilibria of the rod with high intrinsic twist can generally be approximated by equilibria of the effective rod. The description of the effective rod contained in this section is essentially the same as the description presented by Kehrbaum [27, sect. 9.6].

Hamiltonian function (6.13) depends on the intrinsic properties of the rod, i.e. on the intrinsic strains $\hat{\mathbf{U}}(s)$ and on the flexibility matrix $\boldsymbol{C}(s, \frac{s}{\epsilon})$ expressed in (6.9) and (6.14) respectively. The function $\hat{\mathbf{U}}(s)$ is independent of ϵ, while the function $\boldsymbol{C}(s, \frac{s}{\epsilon})$ has a 2π-periodic dependence on $\frac{s}{\epsilon}$. Then, as indicated in (6.13), the Hamiltonian is a function of both the long-scale variable s and the short-scale variable $\frac{s}{\epsilon}$, with a 2π-periodic dependence on $\frac{s}{\epsilon}$.

Hamilton's equation (6.12) governing the equilibria of the rod with high intrinsic twist is of the two-scale type studied in chapter 3. It follows that, if ϵ is small enough (i.e. if τ is large enough), the equilibria of the rod with high intrinsic twist can generally be approximated by solutions to the system

$$\boldsymbol{x}'(s) = \boldsymbol{J}\langle H\rangle_{\boldsymbol{x}}(\boldsymbol{x}(s), s) \tag{6.15}$$

with averaged Hamiltonian

$$\langle H\rangle(\boldsymbol{x}(s), s) = \tfrac{1}{2}\mathbf{M}(s)^T \langle \boldsymbol{C}\rangle(s)\mathbf{M}(s) + \mathbf{M}(s)^T \hat{\mathbf{U}}(s) + \boldsymbol{n}(s)^T \boldsymbol{D}_3(s). \tag{6.16}$$

It appears that averaged equation (6.15) governs the equilibria of an *effective rod* identical to the original rod, except that it possesses an averaged flexibility matrix $\langle \boldsymbol{C} \rangle(s)$ relative to the frame $(\boldsymbol{D}_1(s), \boldsymbol{D}_2(s), \boldsymbol{D}_3(s))$. Writing the stiffness matrix relative to the frame $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ as

$$\boldsymbol{K}(s) = \begin{pmatrix} K_{11}(s) & K_{12}(s) & K_{13}(s) \\ K_{12}(s) & K_{22}(s) & K_{23}(s) \\ K_{13}(s) & K_{23}(s) & K_{33}(s) \end{pmatrix}, \tag{6.17}$$

we have

$$\begin{aligned} \det \boldsymbol{K}(s) = {} & K_{11}(s)K_{22}(s)K_{33}(s) + 2K_{12}(s)K_{13}(s)K_{23}(s) \\ & - K_{11}(s)K_{23}^2(s) - K_{22}(s)K_{13}^2(s) - K_{33}(s)K_{12}^2(s), \end{aligned} \tag{6.18}$$

and

$$\langle \boldsymbol{C} \rangle(s) = \frac{1}{2\pi} \int_0^{2\pi} \boldsymbol{\Omega}(\sigma)\boldsymbol{K}^{-1}(s)\boldsymbol{\Omega}^T(\sigma)\, d\sigma = \operatorname{diag}\big(\bar{K}_1^{-1}(s), \bar{K}_1^{-1}(s), \bar{K}_3^{-1}(s)\big), \tag{6.19}$$

where

$$\begin{aligned} \bar{K}_1(s) &= \frac{2 \det \boldsymbol{K}(s)}{K_{11}(s)K_{33}(s) + K_{22}(s)K_{33}(s) - K_{13}^2(s) - K_{23}^2(s)}, \\ \bar{K}_3(s) &= \frac{\det \boldsymbol{K}(s)}{K_{11}(s)K_{22}(s) - K_{12}^2(s)}. \end{aligned} \tag{6.20}$$

From relation (6.14), the stiffness matrix $\bar{\boldsymbol{K}}(s)$ of the effective rod relative to the $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ frame is

$$\bar{\boldsymbol{K}}(s) = \boldsymbol{\Omega}^T(\tfrac{s}{\epsilon})\langle \boldsymbol{C} \rangle^{-1}(s)\boldsymbol{\Omega}(\tfrac{s}{\epsilon}) = \operatorname{diag}\left(\bar{K}_1(s), \bar{K}_1(s), \bar{K}_3(s)\right). \tag{6.21}$$

In the particular case where the original stiffness matrix is diagonal, i.e.

$$\boldsymbol{K}(s) = \operatorname{diag}\big(K_1(s), K_2(s), K_3(s)\big), \tag{6.22}$$

equation (6.20) simplifies to

$$\bar{K}_1(s) = \frac{2}{K_1^{-1}(s) + K_2^{-1}(s)}, \quad \bar{K}_3(s) = K_3(s), \tag{6.23}$$

i.e. the effective bending stiffness $\bar{K}_1(s)$ corresponds to the harmonic average of the original bending moduli, while the effective twist stiffness $\bar{K}_3(s)$ equals the original twist modulus.

To summarize, the effective rod is identical to the rod with high intrinsic twist, except that it has an effective stiffness matrix $\bar{\boldsymbol{K}}(s)$ expressed in (6.21). It appears that the effective rod is constitutively isotropic, even if the original rod is not. Note however that while averaging over high twist eliminates anisotropy in the constitutive relation, effects of anisotropy due to the slowly varying intrinsic curvatures $\hat{\mathbf{U}}(s)$ are unaffected and persist in the effective rod.

6.4 Intrinsically curved rods with high intrinsic twist

Consider a rod with high intrinsic twist and nonzero intrinsic curvature, i.e. a rod satisfying definition 6.1 with $\hat{\boldsymbol{g}}_3'(s) \neq \mathbf{0}$ or, equivalently, with $(\hat{v}_1(s), \hat{v}_2(s)) \neq \mathbf{0}$, for at least some $s \in [0,1]$. An example of such a rod is depicted in figure 6.1. Assume the rod is subject to one of the three sets of boundary conditions introduced in section 4.3.

The associated effective rod is anisotropic since, according to equation (6.9), its intrinsic strains $\hat{\mathbf{U}}(s)$ are such that $\big(\hat{U}_1(s), \hat{U}_2(s)\big) \neq \mathbf{0}$ for at least some $s \in [0,1]$. Thus the equilibrium equations for the effective rod, expressed in (6.15), admit only the following continuous symmetries: rigid translations and rotations of the rod in space, transformations (5.9) and (5.10) associated with the use of quaternions and, if the effective rod is uniform, translations in arclength. However, in general, the boundary conditions break all compositions of these symmetries, and the equilibria of the effective rod are isolated and satisfy the hypotheses of averaging theorem 3.4. Exceptions to this statement occur in the two following special cases, for example.

(1) The effective rod is uniform and subject to the ring boundary conditions with prescribed angle α such that $\boldsymbol{d}_1^*(1) = \boldsymbol{d}_1^*(0)$, where $(\boldsymbol{d}_1^*(s), \boldsymbol{d}_2^*(s), \boldsymbol{d}_3^*(s))$ is a framing for which the stiffness matrix $\boldsymbol{K}^*$ and the intrinsic strains $\mathbf{u}^*$ are independent of arclength. In this case, the effective rod is of type (b) defined in section 4.5 and possesses the continuous symmetry (4.67).

(2) An equilibrium of the effective rod corresponds to a bifurcation point.

In the general case where the hypotheses of theorem 3.4 hold, an equilibrium $\boldsymbol{x}_0$ of the effective rod generally constitutes an approximation to an equilibrium of the rod with high intrinsic twist if the parameter τ appearing in definition 6.1 is greater than some value (which usually depends on $\boldsymbol{x}_0$). To simplify, we say that the equilibria of the effective rod generally constitute approximations to the equilibria of the rod with high intrinsic twist.

6.5 Intrinsically straight rods with high intrinsic twist

Consider an intrinsically straight rod with a high intrinsic twist, i.e. a rod satisfying definition 6.1 with $\hat{\boldsymbol{g}}_3'(s) = \mathbf{0}$ or, equivalently, with $\hat{v}_1(s) = \hat{v}_2(s) = 0$, for all $s \in [0,1]$. Assume the rod is subject to one of the three sets of boundary conditions introduced in section 4.3.

The associated effective rod is characterized by intrinsic strains $\hat{\mathbf{U}}(s)$, with $\hat{U}_1(s) = \hat{U}_2(s) = 0$, and flexibility matrix of the form (6.19); it is therefore isotropic. Hence the effective problem composed of equilibrium equations (6.15) together with the boundary conditions is invariant under register transformation (4.40), and admits families of solutions of the form

$$\Gamma(\boldsymbol{x}_0) = \big\{\boldsymbol{R}(\theta)\boldsymbol{x}_0 : \theta \in [0, 2\pi)\big\}, \tag{6.24}$$

where $\boldsymbol{x}_0$ is a particular solution and $\boldsymbol{R}(\theta)$ is the transformation matrix defined in (5.25). In general, $\Gamma(\boldsymbol{x}_0)$ is diffeomorphic to a circle, and its members satisfy the hypotheses of averaging theorem 3.5. Exceptions to this statement occur in the three following special cases, for example.

(1) $\boldsymbol{x}_0$ is a fixed point of transformation (4.40). The only fixed points of this transformation are the trivial equilibria of the clamped-free strut and of the clamped-sleeved strut.

(2) Due to an additional continuous symmetry, $\Gamma(\boldsymbol{x}_0)$ is embedded in a two-dimensional manifold of solutions. This situation arises when $\boldsymbol{x}_0$ is a noncircular equilibrium of a uniform, isotropic rod under the ring boundary conditions, for example. In this case, the effective problem is of type (a) defined in section 4.5, and therefore possesses the additional symmetry (4.67).

(3) $\boldsymbol{x}_0$ corresponds to a bifurcation point.

In the general case where the hypotheses of theorem 3.5 hold, the equilibria in $\Gamma(\boldsymbol{x}_0)$ corresponding to simple zeroes of the associated Melnikov function give rise to equilibria of the rod with high intrinsic twist.

For any 2π-periodic function A of the fast variable $\sigma = \frac{s}{\epsilon}$, adopt the notation

$$\tilde{A}(\sigma) = \int_0^\sigma \Big(A(t) - \langle A \rangle\Big)\, dt. \tag{6.25}$$

Introduce the Poincaré function

$$\begin{aligned}\mathcal{P}(\theta) = \int_0^1 \Big(&\langle H_{\boldsymbol{q}}^T \tilde{H}_{\boldsymbol{p}} \rangle (\boldsymbol{R}(\theta)\boldsymbol{x}_0(s), s) - \langle H \rangle_{\boldsymbol{p}}^T \langle \tilde{H} \rangle_{\boldsymbol{q}} (\boldsymbol{R}(\theta)\boldsymbol{x}_0(s), s) \\ &- \langle \tilde{H} \rangle_s (\boldsymbol{R}(\theta)\boldsymbol{x}_0(s), s)\Big)\, ds,\end{aligned} \tag{6.26}$$

where H is given by (6.13), $\boldsymbol{q} = (\boldsymbol{r}, \boldsymbol{\gamma})$ and $\boldsymbol{p} = (\boldsymbol{n}, \boldsymbol{\mu})$. The Melnikov function can then be written as

$$\mathcal{M}(\theta) = \mathcal{P}_\theta(\theta), \tag{6.27}$$

and its simple zeroes coincide with the nondegenerate critical points of $\mathcal{P}(\theta)$. By definition, $\mathcal{P}(\theta)$ is 2π-periodic so that, in the general case, it possesses at least two nondegenerate critical points on $[0, 2\pi)$. Therefore, in general, among a family $\Gamma(\boldsymbol{x}_0)$ of equilibria of the effective rod, at least two members constitute approximations to equilibria of the rod with high intrinsic twist.

Consider now the case where the stiffness matrix of an intrinsically straight rod with high intrinsic twist is of the particular block-diagonal form

$$\boldsymbol{K}(s) = \begin{pmatrix} K_{11}(s) & K_{12}(s) & 0 \\ K_{12}(s) & K_{22}(s) & 0 \\ 0 & 0 & K_{33}(s) \end{pmatrix}. \tag{6.28}$$

Then, the Hamiltonian associated to the rod with high intrinsic twist is invariant under register transformation (4.40) with $\theta = \pi$. Indeed,

$$\begin{aligned} H(\boldsymbol{R}(\pi)\boldsymbol{z}, s, \tfrac{s}{\epsilon}) &= \tfrac{1}{2}\mathbf{M}^T\boldsymbol{\Omega}^T(\pi)\boldsymbol{\Omega}(\tfrac{s}{\epsilon})\boldsymbol{K}^{-1}(s)\boldsymbol{\Omega}^T(\tfrac{s}{\epsilon})\boldsymbol{\Omega}(\pi)\mathbf{M} \\ &\quad + M_3\hat{U}_3(s) + \boldsymbol{n}^T\boldsymbol{\Omega}^T(\pi)\boldsymbol{\Omega}(\pi)\boldsymbol{D}_3 \\ &= \tfrac{1}{2}\mathbf{M}^T\boldsymbol{\Omega}(\tfrac{s}{\epsilon})\boldsymbol{K}^{-1}(s)\boldsymbol{\Omega}^T(\tfrac{s}{\epsilon})\mathbf{M} + M_3\hat{U}_3(s) + \boldsymbol{n}^T\boldsymbol{D}_3 \\ &= H(\boldsymbol{z}, s, \tfrac{s}{\epsilon}). \end{aligned} \tag{6.29}$$

It follows that $\mathcal{P}(\theta)$ is π-periodic and generally admits at least four nondegenerate critical points on $[0, 2\pi)$. We conclude that, if the stiffness matrix is of the particular form (6.28), then a family $\Gamma(\boldsymbol{x}_0)$ of equilibria of the effective rod generally gives rise to at least four (instead of two) equilibria of the rod with high intrinsic twist.

6.6 Examples of intrinsically straight rods with high intrinsic twist

In this section, we consider a uniform, intrinsically straight rod with a diagonal stiffness matrix and a high intrinsic twist. By explicit computation of Melnikov functions, we resolve the symmetry breakings associated with the following one-parameter families of equilibria of the effective rod: the nontrivial equilibria in the case of the clamped-free strut boundary conditions, the nontrivial planar equilibria in the case of the clamped-sleeved boundary conditions, and the circular equilibria in the case of the ring boundary conditions with prescribed angle $\alpha \notin \{0, \pi\}$. In the case of the ring boundary conditions with prescribed angle $\alpha \in \{0, \pi\}$, we show that there is no symmetry breaking, i.e. the one-dimensional manifolds of circular equilibria of the effective rod generally give rise to one-dimensional manifolds of equilibria of the rod with high intrinsic twist.

Consider a rod with a high intrinsic twist satisfying definition 6.1 with $\hat{\boldsymbol{g}}_i'(s) = \mathbf{0}$, $i = 1, 2, 3$, or, equivalently, $\hat{\mathbf{v}}(s) = \mathbf{0}$. The intrinsic strains of the rod are then $\hat{\mathbf{u}}(s) = (0, 0, \tau)$. Suppose the stiffness matrix is of the form $\boldsymbol{K}(s) = \mathrm{diag}\,(K_1, K_2, K_3)$, with $K_1 \neq K_2$. According to (6.9) and (6.14), in the new frame $(\boldsymbol{D}_1(s), \boldsymbol{D}_2(s), \boldsymbol{D}_3(s))$, the rod is characterized by intrinsic strains

$$\hat{\mathbf{U}}(s) = (0, 0, \omega), \tag{6.30}$$

and flexibility matrix

$$\boldsymbol{C}(\tfrac{s}{\epsilon}) = \begin{pmatrix} \bar{K}^{-1} + \Delta\cos\left(\frac{2s}{\epsilon}\right) & \Delta\sin\left(\frac{2s}{\epsilon}\right) & 0 \\ \Delta\sin\left(\frac{2s}{\epsilon}\right) & \bar{K}^{-1} - \Delta\cos\left(\frac{2s}{\epsilon}\right) & 0 \\ 0 & 0 & K_3^{-1} \end{pmatrix}, \tag{6.31}$$

where

$$\bar{K} = \frac{2}{K_1^{-1} + K_2^{-1}}, \quad \Delta = \tfrac{1}{2}\left(K_1^{-1} - K_2^{-1}\right) \neq 0. \tag{6.32}$$

The equilibria of the rod then obey Hamilton's equation (6.12) with Hamiltonian

$$H(\boldsymbol{x}, \tfrac{s}{\epsilon}) = \frac{1}{2}\sum_{i,j=1}^{3} C_{ij}(\tfrac{s}{\epsilon}) M_i M_j + \omega M_3 + \boldsymbol{n}^T \boldsymbol{D}_3, \tag{6.33}$$

where C_{ij} denote the entries of the flexibility matrix $\boldsymbol{C}$.

The associated effective rod is characterized by intrinsic strains $\hat{\mathbf{U}}(s) = (0, 0, \omega)$ and flexibility matrix

$$\langle \boldsymbol{C} \rangle = \operatorname{diag}\big(\bar{K}^{-1}, \bar{K}^{-1}, K_3^{-1}\big), \tag{6.34}$$

or stiffness matrix $\bar{\boldsymbol{K}} = \operatorname{diag}(\bar{K}, \bar{K}, K_3)$. It is governed by equilibrium equation (6.15), with Hamiltonian

$$\langle H \rangle(\boldsymbol{x}) = \frac{M_1^2 + M_2^2}{2\bar{K}} + \frac{M_3^2}{2K_3} + \omega M_3 + \boldsymbol{n}^T \boldsymbol{D}_3. \tag{6.35}$$

Since it is isotropic, the effective rod admits families of equilibria of the form

$$\Gamma(\boldsymbol{x}_0) = \big\{\boldsymbol{R}(\theta)\boldsymbol{x}_0 : \theta \in [0, 2\pi)\big\}, \tag{6.36}$$

where $\boldsymbol{x}_0$ is a particular equilibrium and $\boldsymbol{R}(\theta)$ is the transformation matrix defined in (5.25).

In order to identify those members of $\Gamma(\boldsymbol{x}_0)$ giving rise to equilibria of the rod with high intrinsic twist, we compute the associated Melnikov function. Defining the quantity

$$\mathcal{I}(\boldsymbol{x}) = \langle H_{\boldsymbol{q}}^T \tilde{H}_{\boldsymbol{p}} \rangle(\boldsymbol{x}) - \langle H \rangle_{\boldsymbol{p}}^T \langle \tilde{H} \rangle_{\boldsymbol{q}}(\boldsymbol{x}), \tag{6.37}$$

where H, $\tilde{H}$ are defined by (6.33), (6.25), and $\boldsymbol{q} = (\boldsymbol{r}, \boldsymbol{\gamma})$, $\boldsymbol{p} = (\boldsymbol{n}, \boldsymbol{\mu})$, the Melnikov function can be written as

$$\mathcal{M}(\theta) = \frac{d}{d\theta}\int_0^1 \mathcal{I}(\boldsymbol{R}(\theta)\boldsymbol{x}_0)\, ds. \tag{6.38}$$

We have

$$\begin{aligned} H_{\boldsymbol{q}} &= \begin{pmatrix} \mathbf{0} \\ -\frac{1}{2}\sum_{i=1}^{3} U_i \boldsymbol{B}_i \boldsymbol{\mu} + \boldsymbol{E}^T \boldsymbol{n} \end{pmatrix}, \quad H_{\boldsymbol{p}} = \begin{pmatrix} \boldsymbol{D}_3 \\ \frac{1}{2}\sum_{j=1}^{3} U_j \boldsymbol{B}_j \boldsymbol{\gamma}, \end{pmatrix}, \\ \tilde{H}_{\boldsymbol{q}} &= \begin{pmatrix} \mathbf{0} \\ -\frac{1}{2}\sum_{i=1}^{3} \tilde{U}_i \boldsymbol{B}_i \boldsymbol{\mu} \end{pmatrix}, \quad \tilde{H}_{\boldsymbol{p}} = \begin{pmatrix} \mathbf{0} \\ \frac{1}{2}\sum_{j=1}^{3} \tilde{U}_j \boldsymbol{B}_j \boldsymbol{\gamma}, \end{pmatrix}, \end{aligned} \tag{6.39}$$

where

$$U_i = \hat{U}_i + \sum_{k=1}^{3} C_{ik} M_k, \tag{6.40}$$

and

$$\boldsymbol{E} = 2\begin{pmatrix} \gamma_3 & \gamma_4 & \gamma_1 & \gamma_2 \\ -\gamma_4 & \gamma_3 & \gamma_2 & -\gamma_1 \\ -\gamma_1 & -\gamma_2 & \gamma_3 & \gamma_4 \end{pmatrix}. \tag{6.41}$$

It follows that

$$\mathcal{I}(\boldsymbol{x}) = \frac{1}{2}\sum_{i,j=1}^{3}\alpha_{ij}\left(\frac{1}{2}\boldsymbol{\mu}^T\boldsymbol{B}_i\boldsymbol{B}_j\boldsymbol{\gamma}\right) + \sum_{j=1}^{3}\langle\tilde{U}_j\rangle\boldsymbol{n}^T\left(\frac{1}{2}\boldsymbol{E}\boldsymbol{B}_j\boldsymbol{\gamma}\right), \tag{6.42}$$

where

$$\begin{aligned}\alpha_{ij} &= \langle U_i\tilde{U}_j\rangle - \langle\tilde{U}_i\rangle\langle U_j\rangle \\ &= \hat{U}_i\sum_{l=1}^{3}\langle\tilde{C}\rangle_{jl}M_l + \sum_{k,l=1}^{3}\langle C_{ik}\tilde{C}_{jl}\rangle M_kM_l - \left(\sum_{k=1}^{3}\langle\tilde{C}\rangle_{ik}M_k\right)\left(\hat{U}_j + \sum_{k=1}^{3}\langle C\rangle_{jk}M_k\right).\end{aligned} \tag{6.43}$$

From (6.31), we have

$$\tilde{\boldsymbol{C}} = \frac{\Delta}{2}\begin{pmatrix}\sin\left(\frac{2s}{\epsilon}\right) & 1-\cos\left(\frac{2s}{\epsilon}\right) & 0\\ 1-\cos\left(\frac{2s}{\epsilon}\right) & -\sin\left(\frac{2s}{\epsilon}\right) & 0\\ 0 & 0 & 0\end{pmatrix}, \quad \langle\tilde{\boldsymbol{C}}\rangle = \frac{\Delta}{2}\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix}. \tag{6.44}$$

Inserting expressions (6.30), (6.31), (6.34) and (6.44) into equation (6.43), we obtain

$$\begin{aligned}&\alpha_{11} = \alpha_{22} = \alpha_{33} = 0,\\ &\alpha_{12} = -\alpha_{21} = -\frac{\Delta^2}{4}\left(M_1^2 + M_2^2\right) + \frac{\Delta}{2\bar{K}}\left(M_1^2 - M_2^2\right),\\ &\alpha_{13} = -\alpha_{31} = -\frac{\omega\Delta}{2}M_2 - \frac{\Delta}{2K_3}M_2M_3,\\ &\alpha_{23} = -\alpha_{32} = -\frac{\omega\Delta}{2}M_1 - \frac{\Delta}{2K_3}M_1M_3.\end{aligned} \tag{6.45}$$

Moreover,

$$\langle\tilde{U}_1\rangle = \frac{\Delta}{2}M_2, \quad \langle\tilde{U}_2\rangle = \frac{\Delta}{2}M_1, \quad \langle\tilde{U}_3\rangle = 0. \tag{6.46}$$

Substituting expressions (6.45) and (6.46) into equation (6.42), and using the equalities

$$\begin{aligned}\boldsymbol{B}_1\boldsymbol{B}_2 = -\boldsymbol{B}_2\boldsymbol{B}_1 = -\boldsymbol{B}_3, \quad \boldsymbol{B}_1\boldsymbol{B}_3 = -\boldsymbol{B}_3\boldsymbol{B}_1 = \boldsymbol{B}_2,\\ \boldsymbol{B}_2\boldsymbol{B}_3 = -\boldsymbol{B}_3\boldsymbol{B}_2 = -\boldsymbol{B}_1, \quad \tfrac{1}{2}\boldsymbol{E}\boldsymbol{B}_1\boldsymbol{\gamma} = -\boldsymbol{D}_2, \quad \tfrac{1}{2}\boldsymbol{E}\boldsymbol{B}_2\boldsymbol{\gamma} = \boldsymbol{D}_1,\end{aligned} \tag{6.47}$$

we find

$$\begin{aligned}\mathcal{I}(\boldsymbol{x}) = \frac{\Delta}{2}\left(\frac{M_3}{\bar{K}} - \frac{M_3}{K_3} - \omega\right)\left(M_2^2 - M_1^2\right) + \frac{\Delta}{2}\left(M_1\boldsymbol{n}^T\boldsymbol{D}_1 - M_2\boldsymbol{n}^T\boldsymbol{D}_2\right)\\ + \frac{\Delta^2}{4}\left(M_1^2 + M_2^2\right)M_3.\end{aligned} \tag{6.48}$$

Then, by equation (4.41),

$$\begin{aligned}\mathcal{I}(\boldsymbol{R}(\theta)\boldsymbol{x}_0) = \frac{\Delta}{2}\left(\frac{M_3^0}{\bar{K}} - \frac{M_3^0}{K_3} - \omega\right)\Big(\cos(2\theta)\left((M_2^0)^2 - (M_1^0)^2\right) + \sin(2\theta)2M_1^0M_2^0\Big)\\ + \frac{\Delta}{2}\Big(\cos(2\theta)\big(M_1^0\boldsymbol{n}^{0\,T}\boldsymbol{D}_1^0 - M_2^0\boldsymbol{n}^{0\,T}\boldsymbol{D}_2^0\big)\\ - \sin(2\theta)\big(M_1^0\boldsymbol{n}^{0\,T}\boldsymbol{D}_2^0 + M_2^0\boldsymbol{n}^{0\,T}\boldsymbol{D}_1^0\big)\Big) + \frac{\Delta^2}{4}\big((M_1^0)^2 + (M_2^0)^2\big)M_3^0,\end{aligned} \tag{6.49}$$

so that the Melnikov function (6.38) is expressed by

$$\begin{aligned}\mathcal{M}(\theta) = \Delta \int_0^1 \Big[\Big(\frac{M_3^0}{\bar{K}} - \frac{M_3^0}{K_3} - \omega \Big) \Big(&- \sin(2\theta) \big((M_2^0)^2 - (M_1^0)^2 \big) + \cos(2\theta) 2 M_1^0 M_2^0 \Big) \\ &- \sin(2\theta) \big(M_1^0 \boldsymbol{n}^{0\,T} \boldsymbol{D}_1^0 - M_2^0 \boldsymbol{n}^{0\,T} \boldsymbol{D}_2^0 \big) \\ &- \cos(2\theta) \big(M_1^0 \boldsymbol{n}^{0\,T} \boldsymbol{D}_2^0 + M_2^0 \boldsymbol{n}^{0\,T} \boldsymbol{D}_1^0 \big) \Big] \, ds,\end{aligned} \tag{6.50}$$

where the quantities carrying a superscript ^{0} are associated with the equilibrium $\boldsymbol{x}_0$.

Clamped-free strut

Consider the case where the uniform, intrinsically straight rod with a diagonal stiffness matrix and a high intrinsic twist defined above is subject to the clamped-free strut boundary conditions (4.32). In the Hamiltonian variables, both the rod and the effective rod obey version (5.17) of the boundary conditions.

The equilibria of the effective rod are discussed in sections 4.6 and 5.4. The trivial solution to the effective rod problem, corresponding to a straight equilibrium, is also an exact solution to the problem of the rod with intrinsic twist. The nontrivial equilibria of the effective rod occur in families of the form

$$\Gamma(\boldsymbol{x}_0) = \big\{ \boldsymbol{R}(\theta) \boldsymbol{x}_0 : \theta \in [0, 2\pi) \big\}, \tag{6.51}$$

where $\boldsymbol{x}_0$ is an equilibrium lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane and expressed (with slight differences in notation) by equations (4.71), (5.44) and (5.46). Substituting these expressions into equation (6.50) shows that the Melnikov function associated with $\Gamma(\boldsymbol{x}_0)$ is

$$\mathcal{M}(\theta) = \frac{\Delta \bar{K}^2}{2} \int_0^1 \Big(2\omega \sin(2\theta - 2\omega s) \phi'^2(s) - \frac{2\lambda}{\bar{K}} \cos(2\theta - 2\omega s) \sin(\phi(s)) \phi'(s) \Big) \, ds, \tag{6.52}$$

where $\phi(s)$ is a nontrivial solution to

$$\phi''(s) = -\frac{\lambda}{\bar{K}} \sin(\phi(s)), \quad \phi(0) = 0, \quad \phi'(1) = 0. \tag{6.53}$$

Finally, with the help of equation (6.53), we obtain

$$\mathcal{M}(\theta) = \frac{\Delta \bar{K}^2}{2} \cos(2\theta - 2\omega s) \phi'^2(s) \Big|_{s=0}^{s=1} = -\frac{\Delta \bar{K}^2}{2} \cos(2\theta) \phi'^2(0). \tag{6.54}$$

Since the trivial solution $\phi(s) = 0$ is excluded, (6.53) implies $\phi'(0) \neq 0$, and the Melnikov function admits simple zeroes for

$$\theta = \frac{\pi}{4}, \; \frac{3\pi}{4}, \; \frac{5\pi}{4}, \; \frac{7\pi}{4}. \tag{6.55}$$

These values of θ designate the equilibria of the effective rod contained in the two vertical planes making angles of $\frac{\pi}{4}$ and $\frac{3\pi}{4}$ with $\boldsymbol{d}_1(0)$ (cf. figure 6.2). We conclude that,

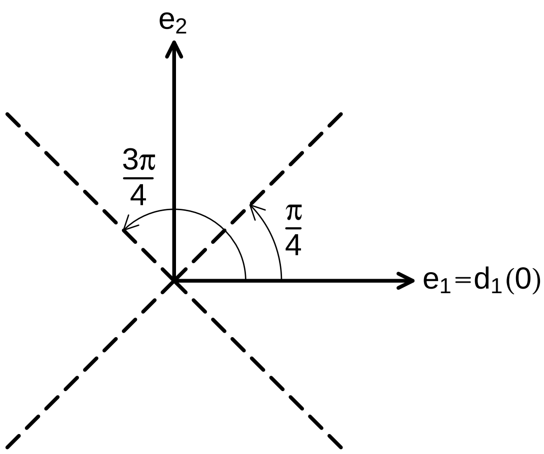


Figure 6.2: *Orientation of the two approximate vertical planes of buckling (dashed lines) for the clamped-free strut with high intrinsic twist.*

in general, only those equilibria of the effective strut lying in these two special planes constitute approximations to equilibria of the clamped-free strut with high intrinsic twist. Note that the angles defining the two special planes are fixed by $\boldsymbol{d}_1(0)$, which corresponds to a direction of extremal stiffness at the clamped end, but are independent of the values of the stiffness constants K_1, K_2, K_3, and of the load λ.

Clamped-sleeved strut

Consider the case where the uniform, intrinsically straight rod with a diagonal stiffness matrix and a high intrinsic twist defined above is subject to the clamped-sleeved strut boundary conditions (4.33). In the Hamiltonian variables, both the rod and the effective rod obey version (5.20) of the boundary conditions.

The planar equilibria of the effective rod are discussed in sections 4.6 and 5.4. The trivial solution to the effective rod problem, corresponding to a straight equilibrium, is also an exact solution to the problem of the rod with intrinsic twist. The nontrivial planar equilibria of the effective rod occur in families of the form

$$\Gamma(\boldsymbol{x}_0) = \big\{\boldsymbol{R}(\theta)\boldsymbol{x}_0 : \theta \in [0, 2\pi)\big\}, \tag{6.56}$$

where $\boldsymbol{x}_0$ is an equilibrium lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane and expressed (with slight differences in notation) by equations (4.73), (5.47) and (5.48). Substituting these expressions into equation (6.50) shows that the Melnikov function associated with $\Gamma(\boldsymbol{x}_0)$ is

$$\begin{aligned}\mathcal{M}(\theta) = \frac{\Delta\bar{K}^2}{2}\int_0^1 \Big[&2\omega\sin(2\theta - 2\omega s)\phi'^2(s) \\ &- 2\cos(2\theta - 2\omega s)\left(\frac{\lambda}{\bar{K}}\sin\phi(s) + \frac{n_1}{\bar{K}}\cos\phi(s)\right)\phi'(s)\Big]\,ds,\end{aligned} \tag{6.57}$$

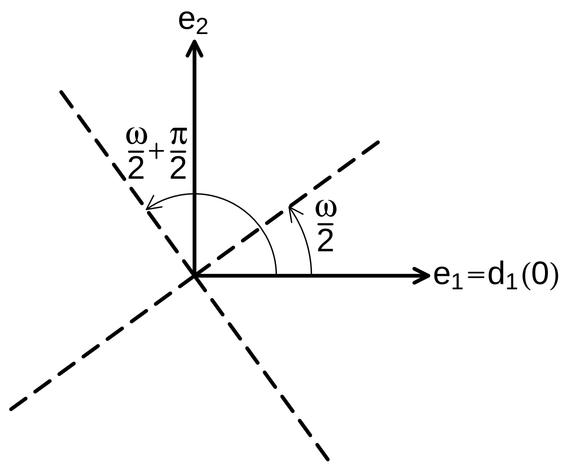


Figure 6.3: *Orientation of the two approximate vertical planes of buckling (dashed lines) for the clamped-sleeved strut with high intrinsic twist.*

where $\phi(s)$ is a nontrivial solution to

$$\begin{cases} \phi''(s) = -\dfrac{\lambda}{\bar{K}} \sin\phi(s) - \dfrac{n_1}{\bar{K}} \cos\phi(s), \\ \phi(0) = 0, \quad \phi(1) = 0 \pmod{2\pi}, \quad \displaystyle\int_0^1 \sin\phi(s)\, ds = 0. \end{cases} \tag{6.58}$$

Finally, with the help of equations (6.58) and (C.12), we obtain

$$\begin{aligned} \mathcal{M}(\theta) &= \frac{\Delta\bar{K}^2}{2} \cos(2\theta - 2\omega s)\phi'^2(s)\Big|_{s=0}^{s=1} \\ &= \frac{\Delta\bar{K}^2}{2}\Big(\cos(2\theta - 2\omega)\phi'^2(1) - \cos(2\theta)\phi'^2(0)\Big) \\ &= \Delta\bar{K}^2 \sin\omega \sin(2\theta - \omega)\phi'^2(0). \end{aligned} \tag{6.59}$$

As mentioned at the end of appendix C, the nontrivial solutions to (6.58) satisfying $\phi'(0) = 0$ are exceptional. Thus, in general, if the value of the high intrinsic twist τ of the rod is not a multiple of π, so that $\sin\omega \neq 0$, then the Melnikov function has simple zeroes for

$$\theta = \frac{\omega}{2}, \quad \frac{\omega}{2} + \frac{\pi}{2}, \quad \frac{\omega}{2} + \pi, \quad \left(\frac{\omega}{2} + \frac{3\pi}{2}\right) \pmod{2\pi}. \tag{6.60}$$

These values of θ designate the equilibria of the effective rod contained in the two vertical planes making angles of $\frac{\omega}{2}$ and $\frac{\omega}{2} + \frac{\pi}{2}$ with $\boldsymbol{d}_1(0)$ (cf. figure 6.3). We conclude that, if τ is not a multiple of π then, in general, among the planar equilibria of the effective strut, only those lying in these two special planes constitute approximations to equilibria of the clamped-sleeved strut with high intrinsic twist τ.

If τ is a multiple of π, then $\omega = 0$ or π, and theorem 3.5 does not say anything about the problem since the Melnikov function is identically zero. However, numerical experiments similar to the one presented in figure 1.2 seem to indicate that, in this case, the above conclusion remains valid, i.e. in general, among the planar equilibria of the effective strut, only those lying in the two planes making angles of 0 and $\frac{\pi}{2}$ with $\boldsymbol{d}_1(0)$ constitute approximations to equilibria of the strut with high intrinsic twist τ.

Finally note that the angles defining the two approximate planes of buckling of the strut are independent of the values of the stiffness constants K_1, K_2, K_3, and of the load λ. They depend only on the value of the high intrinsic twist τ in a periodic manner via the quantity $\omega = \tau \,(\mathrm{mod}\ 2\pi) \in [0, 2\pi)$. The quantity ω corresponds to the angle between the directions of extremal stiffness $\boldsymbol{d}_1(0)$ and $\boldsymbol{d}_1(1)$ of the effective rod, and is therefore an approximation of the angle between $\boldsymbol{d}_1(0)$ and $\boldsymbol{d}_1(1)$ of the rod with high intrinsic twist.

Ring

Consider the case where the uniform, intrinsically straight rod with a diagonal stiffness matrix and a high intrinsic twist defined above is subject to the ring boundary conditions (4.34). In the Hamiltonian variables, both the rod and the effective rod obey version (5.23) of the boundary conditions.

The circular equilibria of the effective rod, discussed in sections 4.6 and 5.4, occur in families of the form

$$\Gamma(\boldsymbol{x}_0) = \big\{\boldsymbol{R}(\theta)\boldsymbol{x}_0 : \theta \in [0, 2\pi)\big\}, \tag{6.61}$$

where $\boldsymbol{x}_0$ is a circular equilibrium lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane. The centerline and the force associated with $\boldsymbol{x}_0$ are of the form

$$\boldsymbol{r}^0(s) = \frac{1}{2\pi l}\begin{pmatrix} 1 - \cos(2\pi l s) \\ 0 \\ \sin(2\pi l s) \end{pmatrix}, \quad \boldsymbol{n}^0(s) = \begin{pmatrix} 0 \\ 2\pi l K_3(\alpha - \omega + 2\pi j) \\ 0 \end{pmatrix}, \tag{6.62}$$

with $l = \pm 1, \pm 2, \ldots$, and $j = 0, \pm 1, \pm 2, \ldots$. The components of the moment and the framing associated with $\boldsymbol{x}_0$ are given in the original variables by

$$\mathbf{m}^0(s) = \begin{pmatrix} 2\pi l \bar{K} \sin(\alpha s + 2\pi j s + \frac{s}{\epsilon}) \\ 2\pi l \bar{K} \cos(\alpha s + 2\pi j s + \frac{s}{\epsilon}) \\ K_3(\alpha - \omega + 2\pi j) \end{pmatrix}, \quad \boldsymbol{d}_1^0(s) = \begin{pmatrix} \cos(\alpha s + 2\pi j s + \frac{s}{\epsilon}) \cos(2\pi l s) \\ \sin(\alpha s + 2\pi j s + \frac{s}{\epsilon}) \\ -\cos(\alpha s + 2\pi j s + \frac{s}{\epsilon}) \sin(2\pi l s) \end{pmatrix},$$

$$\boldsymbol{d}_2^0(s) = \begin{pmatrix} -\sin(\alpha s + 2\pi j s + \frac{s}{\epsilon}) \cos(2\pi l s) \\ \cos(\alpha s + 2\pi j s + \frac{s}{\epsilon}) \\ \sin(\alpha s + 2\pi j s + \frac{s}{\epsilon}) \sin(2\pi l s) \end{pmatrix}, \quad \boldsymbol{d}_3^0(s) = \begin{pmatrix} \sin(2\pi l s) \\ 0 \\ \cos(2\pi l s) \end{pmatrix}, \tag{6.63}$$

and in the new variables by

$$\mathsf{M}^0(s) = \begin{pmatrix} 2\pi l\bar{K}\sin(\alpha s + 2\pi j s) \\ 2\pi l\bar{K}\cos(\alpha s + 2\pi j s) \\ K_3(\alpha - \omega + 2\pi j) \end{pmatrix}, \quad \boldsymbol{D}_1^0(s) = \begin{pmatrix} \cos(\alpha s + 2\pi j s)\cos(2\pi l s) \\ \sin(\alpha s + 2\pi j s) \\ -\cos(\alpha s + 2\pi j s)\sin(2\pi l s) \end{pmatrix},$$

$$\boldsymbol{D}_2^0(s) = \begin{pmatrix} -\sin(\alpha s + 2\pi j s)\cos(2\pi l s) \\ \cos(\alpha s + 2\pi j s) \\ \sin(\alpha s + 2\pi j s)\sin(2\pi l s) \end{pmatrix}, \quad \boldsymbol{D}_3^0(s) = \begin{pmatrix} \sin(2\pi l s) \\ 0 \\ \cos(2\pi l s) \end{pmatrix}. \tag{6.64}$$

In order to apply averaging results, the parameter ϵ must be small compared with all other quantities involved in the problem. Therefore, given a family of circular equilibria of the effective rod characterized by integers l and j, we assume $\frac{1}{\epsilon}$ is large compared with l and j. In particular, in the original framing, we suppose that the difference between the twist $u_3 = \alpha + 2\pi j + \frac{1}{\epsilon}$ and the intrinsic twist $\hat{u}_3 = \tau = \omega + \frac{1}{\epsilon}$ is small compared with the intrinsic twist.

The number of equilibria of the ring with high intrinsic twist crucially depends on the value of the prescribed angle α between $\boldsymbol{d}_1(0)$ and $\boldsymbol{d}_1(1)$: a continuous family $\Gamma(\boldsymbol{x}_0)$ of circular equilibria of the effective rod generally gives rise to four equilibria of the rod with high intrinsic twist if $\alpha \notin \{0, \pi\}$, and to a continuous family of equilibria if $\alpha \in \{0, \pi\}$. We examine each of these two cases.

Case $\alpha \notin \{0, \pi\}$

Substituting expressions given in (6.62) and (6.64) into equation (6.50) shows that the Melnikov function associated with $\Gamma(\boldsymbol{x}_0)$ is

$$\begin{aligned} \mathcal{M}(\theta) &= 4\pi^2 l^2 \bar{K}^2 \Delta \int_0^1 (\alpha + 2\pi j)\sin(2\theta - 2\alpha s - 4\pi j s)\, ds \\ &= 4\pi^2 l^2 \bar{K}^2 \Delta \sin\alpha \sin(2\theta - \alpha) \end{aligned} \tag{6.65}$$

and admits simple zeroes for

$$\theta = \frac{\alpha}{2}, \quad \frac{\alpha}{2} + \frac{\pi}{2}, \quad \frac{\alpha}{2} + \pi, \quad \left(\frac{\alpha}{2} + \frac{3\pi}{2}\right) \pmod{2\pi}. \tag{6.66}$$

These values of θ designate the circular equilibria of the effective rod contained in the two vertical planes making angles of $\frac{\alpha}{2}$ and $\frac{\alpha}{2} + \frac{\pi}{2}$ with $\boldsymbol{d}_1(0)$ (cf. figure 6.4). We conclude that, in general, among the circular equilibria of the effective ring, only those lying in these two special planes constitute approximations to equilibria of the ring with high intrinsic twist. Note that the angles defining the two special planes depend on the prescribed angle α between the directions of extremal stiffness $\boldsymbol{d}_1(0)$ and $\boldsymbol{d}_1(1)$, but not on the stiffness constants K_1, K_2, K_3.

Case $\alpha \in \{0, \pi\}$

In the formulation relative to the original framing $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$, the problem governing the equilibria of the ring with high intrinsic twist is of type (b) or (c) defined

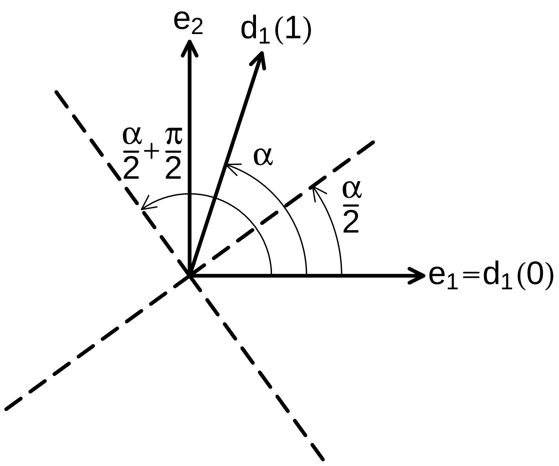


Figure 6.4: *Orientation of the two approximate vertical planes (dashed lines) close to which the equilibria of the ring with high intrinsic twist lie.*

in section 4.5, and it therefore invariant under transformation (4.67). Translating this transformation into the variables associated with the new framing $(\boldsymbol{D}_1(s), \boldsymbol{D}_2(s), \boldsymbol{D}_3(s))$ shows that if $(\boldsymbol{r}^0, \boldsymbol{D}_i^0, \boldsymbol{n}^0, \mathbf{M}^0)$ describes an equilibrium of the rod, then

$$
\begin{aligned}
\boldsymbol{r}(s) &= \boldsymbol{\Omega}^T\left(\tfrac{\beta}{\epsilon}\right)\left(\boldsymbol{D}_1^0(\beta), \boldsymbol{D}_2^0(\beta), \boldsymbol{D}_3^0(\beta)\right)^T\left(\boldsymbol{r}^0(t) - \boldsymbol{r}^0(\beta)\right), \\
\boldsymbol{n} = \boldsymbol{\Omega}^T\left(\tfrac{\beta}{\epsilon}\right)&\left(\boldsymbol{D}_1^0(\beta), \boldsymbol{D}_2^0(\beta), \boldsymbol{D}_3^0(\beta)\right)^T\boldsymbol{n}^0, \quad \mathbf{M}(s) = \boldsymbol{\Omega}\left(\tfrac{s}{\epsilon}\right)\boldsymbol{A}(s)^T\boldsymbol{\Omega}^T\left(\tfrac{t}{\epsilon}\right)\mathbf{M}^0(t), \\
\left(\boldsymbol{D}_1(s), \boldsymbol{D}_2(s), \boldsymbol{D}_3(s)\right) &= \boldsymbol{\Omega}^T\left(\tfrac{\beta}{\epsilon}\right)\left(\boldsymbol{D}_1^0(\beta), \boldsymbol{D}_2^0(\beta), \boldsymbol{D}_3^0(\beta)\right)^T \\
&\qquad \left(\boldsymbol{D}_1^0(t), \boldsymbol{D}_2^0(t), \boldsymbol{D}_3^0(t)\right)\boldsymbol{\Omega}\left(\tfrac{t}{\epsilon}\right)\boldsymbol{A}(s)\boldsymbol{\Omega}^T\left(\tfrac{s}{\epsilon}\right),
\end{aligned}
\tag{6.67}
$$

with $t = (s+\beta) \pmod 1$ and $\boldsymbol{A}(s)$ defined in (4.69), also corresponds to an equilibrium, for any $\beta \in [0, 1)$.

In the formulation relative to the original framing, the effective problem is invariant under transformations (4.40) and (4.67). However, when applied to a circular solution of the form (6.62), (6.63), transformation (4.67) produces the same result as transformation (4.40) with $\theta = -(\alpha + 2\pi j + \frac{1}{\epsilon})\beta \pmod{2\pi}$. This property follows from a comparison between expressions in (6.62), (6.63) and expressions in (4.75), together with the property mentioned in the last paragraph of section 4.6. Since $\alpha + 2\pi j + \frac{1}{\epsilon}$ is assumed to be large, and in particular larger than 2π, then, in the variables associated with the original framing, any cirular equilibrium of the effective rod is obtained by performing transformation (4.67), for some $\beta \in [0, 1)$, on a circular equilibria lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane. It follows that the circular equilibria of the effective rod occur in one-parameter families generated by transformation (4.67) in the original variables, and by transformation (6.67) in the new variables.

We have proven that, in the formulation relative to the new framing, both the equilibria of the rod with high intrinsic twist and the circular equilibria of the effective rod occur in one-parameter families generated by transformation (6.67). Moreover, any family of circular equilibria of the effective rod contains at least one representative lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane. The rest of this section is devoted to showing that, in general, a circular equilibrium of the effective rod lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane constitutes an approximation to an equilibrium of the rod with high intrinsic twist. It immediately follows that a family of circular equilibria of the effective rod typically constitutes an approximation to a family of equilibria of the rod with high intrinsic twist.

Consider the boundary conditions

$$\begin{gathered}\boldsymbol{r}(0) = \boldsymbol{0}, \quad \gamma_2(0) = \gamma_3(0) = 0, \quad \mu_1(0) = \mu_4(0) = 0, \\ \boldsymbol{r}(1) = \boldsymbol{0}, \quad \boldsymbol{\gamma}(1) = \left(0, 0, \sin\tfrac{\alpha}{2}, \cos\tfrac{\alpha}{2}\right),\end{gathered} \tag{6.68}$$

obtained from the ring boundary conditions (5.23) by replacing the constraint $\gamma_1(0) = 0$ by $\mu_1(0) = 0$. This modified boundary conditions imply $M_1(0) = 0$, and will be used to isolate the circular equilibria of the effective ring lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane. Although we do not use this property here, note that, as shown in appendix B, the modified boundary conditions satisfy property (**2.A5**) and therefore define Lagrangian manifolds at $s = 0$ and at $s = 1$.

Lemma 6.2 *Let $\boldsymbol{x}_0$ be the expression in the Hamiltonian variables of a circular equilibrium of the effective ring lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane. Then, there exists a $\mathcal{C}^0$-neighborhood V of $\boldsymbol{x}_0$ such that*

(i) any solution to the effective equilibrium equations (6.15) contained in V and satisfying the modified boundary conditions (6.68) also satisfies the original boundary conditions (5.23), and

(ii) any solution to the equilibrium equations (6.12) of the rod with high intrinsic twist contained in V and satisfying the modified boundary conditions (6.68) also satisfies the original boundary conditions (5.23).

Proof Consider a solution either to the effective equilibrium equations (6.15) or to the equilibrium equations (6.12) of the rod with high intrinsic twist, and satisfying modified boundary conditions (6.68). Combining the boundary conditions with integral $\boldsymbol{\gamma}^T\boldsymbol{\gamma}$ shows that $\gamma_1^2(0) + \gamma_4^2(0) = 1$, and we can write

$$\gamma_1(0) = \sin\frac{\zeta}{2}, \quad \gamma_4(0) = \cos\frac{\zeta}{2}, \quad \text{with } \zeta \in [-2\pi, 2\pi)\,. \tag{6.69}$$

It follows that

$$\boldsymbol{d}_1(0) = \boldsymbol{D}_1(0) = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \boldsymbol{d}_2(0) = \boldsymbol{D}_2(0) = \begin{pmatrix}0\\ \cos\zeta\\ \sin\zeta\end{pmatrix}, \quad \boldsymbol{d}_3(0) = \boldsymbol{D}_3(0) = \begin{pmatrix}0\\ -\sin\zeta\\ \cos\zeta\end{pmatrix}. \tag{6.70}$$

Since $\alpha \in \{0, \pi\}$, it follows from the boundary conditions that

$$\boldsymbol{d}_1(1) = \boldsymbol{D}_1(1) = \pm\boldsymbol{e}_1, \quad \boldsymbol{d}_2(1) = \boldsymbol{D}_2(1) = \pm\boldsymbol{e}_2, \quad \boldsymbol{d}_3(1) = \boldsymbol{D}_3(1) = \boldsymbol{e}_3. \tag{6.71}$$

Combining the boundary conditions with integral $\boldsymbol{m}+\boldsymbol{r}\times\boldsymbol{n}$ shows that $\boldsymbol{m}(0)=\boldsymbol{m}(1)$. Moreover,

$$\boldsymbol{m}(0)^T\boldsymbol{e}_1=\boldsymbol{m}(0)^T\boldsymbol{D}_1(0)=M_1(0)=0, \tag{6.72}$$

so that we can write

$$\boldsymbol{m}(0)=\boldsymbol{m}(1)=\begin{pmatrix}0\\ m_2^*\\ m_3^*\end{pmatrix}, \tag{6.73}$$

and

$$\begin{aligned}
&m_1(0)=M_1(0)=0, && m_1(1)=M_1(1)=0,\\
&m_2(0)=M_2(0)=m_2^*\cos\zeta+m_3^*\sin\zeta, && m_2(1)=M_2(1)=\pm m_2^*,\\
&m_3(0)=M_3(0)=-m_2^*\sin\zeta+m_3^*\cos\zeta, && m_3(1)=M_3(1)=m_3^*.
\end{aligned} \tag{6.74}$$

We have used the integrals $\boldsymbol{\gamma}^T\boldsymbol{\gamma}$ and $\boldsymbol{m}+\boldsymbol{r}\times\boldsymbol{n}$, and we need an additional integral to complete the proof. For part (i), the integrals associated with isotropy and uniformity constitute two options; we choose the former. For part (ii), the only option is the integral associated with uniformity.

We proceed with the proof of (i). Consider a solution to the effective equilibrium equations (6.15) satisfying the modified boundary conditions (6.68). Due to the isotropy of the effective rod, M_3 represents an integral (cf. lemma 4.2), and therefore

$$S(\zeta)=0, \tag{6.75}$$

where

$$S(\zeta)=M_3(0)-M_3(1)=-m_2^*\sin\zeta+m_3^*\cos\zeta-m_3^*. \tag{6.76}$$

A solution to (6.75) is $\zeta=0$, and we have $S'(0)=-m_2^*$. Suppose the solution to BVP (6.15)+(6.68) is close to a circular equilibrium $\boldsymbol{x}_0$ of the effective ring lying in the $(\boldsymbol{e}_1,\boldsymbol{e}_3)$-plane. Then ζ is close to 0, and $S'(0)$ is close to $-2\pi l\bar{K}$. In fact, if the solution is sufficiently close to $\boldsymbol{x}_0$, then $S'(0)\neq 0$, and it follows from equation (6.75) that $\zeta=0$. In this case, the condition $\gamma_1(0)=0$ holds, and the solution to BVP (6.15)+(6.68) also satisfies the original boundary conditions (5.23).

We complete this proof by establishing (ii). Following section 5.1, the original framing $(\boldsymbol{d}_1(s),\boldsymbol{d}_2(s),\boldsymbol{d}_3(s))$ can be parametrized by quaternions, and the equilibrium equations for the rod with high intrinsic twist can be expressed in Hamiltonian form, with Hamiltonian

$$H(\boldsymbol{z})=\frac{m_1^2}{2K_1}+\frac{m_2^2}{2K_2}+\frac{m_3^2}{2K_3}+m_3\tau+\boldsymbol{n}^T\boldsymbol{d}_3. \tag{6.77}$$

Due to the uniformity of the rod with high intrinsic twist, this Hamiltonian does not explicitly depend on the independent variable s, and therefore represents an integral. Consider a solution to the equilibrium equations (6.12) of the rod with high intrinsic twist, and satisfying the modified boundary conditions (6.68). We then have

$$T(\zeta)=0, \tag{6.78}$$

where

$$\begin{aligned} T(\zeta) &= H(\boldsymbol{z}(0)) - H(\boldsymbol{z}(1)) \\ &= \frac{1}{2K_2}\Big(\big(m_2^* \cos\zeta + m_3^* \sin\zeta\big)^2 - m_2^{*2}\Big) \\ &+ \frac{1}{2K_3}\Big(\big(-m_2^* \sin\zeta + m_3^* \cos\zeta\big)^2 - m_3^{*2}\Big) \\ &+ \big(-m_2^* \sin\zeta + m_3^* \cos\zeta - m_3^*\big)\tau - n_2 \sin\zeta + n_3 \cos\zeta - n_3. \end{aligned} \tag{6.79}$$

A solution to (6.78) is $\zeta = 0$, and

$$T'(0) = \left(\frac{1}{K_2} - \frac{1}{K_3}\right) m_2^* m_3^* - m_2^* \tau - n_2. \tag{6.80}$$

Suppose the solution to BVP (6.12)+(6.68) is close to a circular equilibrium $\boldsymbol{x}_0$ of the effective ring lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane. Then ζ is close to 0, and

$$m_2^* \approx 2\pi l \bar{K}, \quad m_3^* \approx K_3(\alpha - \omega + 2\pi j) \quad n_2 \approx 2\pi l K_3(\alpha - \omega + 2\pi j), \tag{6.81}$$

$$T'(0) \approx 2\pi l \left(\frac{\bar{K} K_3}{K_2} - K_3 - \bar{K}\right)(\alpha - \omega + 2\pi j) - 2\pi l \bar{K} \tau. \tag{6.82}$$

Since $\tau = \omega + \frac{1}{\epsilon}$ is assumed to be large compared with the other quantities involved in the problem, the right hand side of the latter equation is nonzero. In fact, if the solution is sufficiently close to $\boldsymbol{x}_0$, then $T'(0) \neq 0$, and it follows from equation (6.78) that $\zeta = 0$. In this case, the condition $\gamma_1(0) = 0$ holds, and the solution to (6.12)+(6.68) also satisfies the original boundary conditions (5.23). ■

In the following, we refer to equations (6.12)+(5.23) as the ring BVP, to equations (6.15)+(5.23) as the effective ring BVP, to equations (6.12)+(6.68) as the modified ring BVP, and to equations (6.15)+(6.68) as the modified effective ring BVP. An equilibrium $\boldsymbol{x}_0$ of the effective ring lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane enjoys the property $M_1(0) = 0$, and therefore satisfies the modified effective ring BVP. We now restrict our attention to some small $\mathcal{C}^0$-neighborhood of $\boldsymbol{x}_0$. The other members of the family of equilibria obtained by performing the transformation defined by (6.67) on $\boldsymbol{x}_0$ are such that $M_1(0) \neq 0$, and do not satisfy the modified effective ring BVP. Moreover, by lemma 6.2, any solution to the modified effective ring BVP also satisfies the effective ring BVP. Therefore, in general, $\boldsymbol{x}_0$ is an isolated solution to the modified effective ring BVP, and satisfies the hypotheses of averaging theorem 3.4. It follows that, for ϵ small, there exists a solution $\boldsymbol{\xi}(\epsilon)$ to the modified ring BVP lying in some small neighborhood of $\boldsymbol{x}_0$. Then, according to lemma 6.2, $\boldsymbol{\xi}(\epsilon)$ is also a solution to the ring BVP.

Having shown that, in general, a circular equilibrium of the effective ring lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane constitutes an approximation to an equilibrium of the ring with high intrinsic twist, we can conclude that, in general, the circular equilibria of the effective rod, which occur in one-parameter families generated by transformation (6.67), constitute approximations to equilibria of the ring with high intrinsic twist, which also occur in one-parameter families generated by transformation (6.67).

6.7 Rods with small intrinsic curvature and high intrinsic twist

In this section, we give a precise definition of what we describe as a rod with small intrinsic curvature and high intrinsic twist. We then study the equilibria of such a rod when it is subject to one of the sets of boundary conditions introduced in section 4.3.

To define a rod with small intrinsic curvature and high intrinsic twist, we slightly modify definition 6.1 of a rod with high intrinsic twist; instead of requiring the framing $(\hat{\boldsymbol{g}}_1(s), \hat{\boldsymbol{g}}_2(s), \hat{\boldsymbol{g}}_3(s))$ to be independent of the large parameter τ, we require its variation to be proportional to $\frac{1}{\tau}$.

Definition 6.3 *A rod is said to have* a small intrinsic curvature and a high intrinsic twist *if it is constitutively anisotropic and if, for a suitable choice of framing* $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$, *the following properties hold.*

(1) The intrinsic framing $\big(\hat{\boldsymbol{d}}_1(s), \hat{\boldsymbol{d}}_2(s), \hat{\boldsymbol{d}}_3(s)\big)$ *can be written as*

$$\big(\hat{\boldsymbol{d}}_1(s), \hat{\boldsymbol{d}}_2(s), \hat{\boldsymbol{d}}_3(s)\big) = \big(\hat{\boldsymbol{g}}_1(s), \hat{\boldsymbol{g}}_2(s), \hat{\boldsymbol{g}}_3(s)\big)\, \boldsymbol{\Omega}(\tau s), \tag{6.83}$$

where $\boldsymbol{\Omega}$ *is the rotation matrix defined in (4.16),* τ *is a large parameter, and* $(\hat{\boldsymbol{g}}_1(s), \hat{\boldsymbol{g}}_2(s), \hat{\boldsymbol{g}}_3(s))$ *is a right-handed, orthonormal frame with the property that the triple* $\hat{\mathbf{v}}(s)$ *defined by equation (6.2) is of the form*

$$\hat{\mathbf{v}}(s) = \frac{1}{\tau}\, \check{\mathbf{v}}(s), \tag{6.84}$$

with $\check{\mathbf{v}}(s)$ *independent of* τ.

(2) The stiffness matrix $\boldsymbol{K}(s)$ *is independent of* τ.

By equation (6.3), the intrinsic strains relative to the framing $(\boldsymbol{d}_1(s), \boldsymbol{d}_2(s), \boldsymbol{d}_3(s))$ of a rod with small intrinsic curvature and high intrinsic twist are given by

$$\hat{\mathbf{u}}(s) = \boldsymbol{\Omega}^T(\tau s) \left(\frac{1}{\tau}\, \check{\mathbf{v}}(s) + \tau \mathbf{e}_3 \right). \tag{6.85}$$

The intrinsic curvatures are of order of magnitude $\frac{1}{\tau}$, while the intrinsic twist is of order of magnitude τ.

Consider a rod with small intrinsic curvature and high intrinsic twist subject to one of the sets of boundary conditions introduced in section 4.3. Following section 6.2, we reformulate the problem in a form suitable for averaging. Consider the numbers ω and ϵ, and the new framing $(\boldsymbol{D}_1(s), \boldsymbol{D}_2(s), \boldsymbol{D}_3(s))$ defined by equations (6.5), (6.4) and (6.6) respectively. In this new framing, the flexibility matrix $\boldsymbol{C}(s, \frac{s}{\epsilon})$ is given by (6.14), and the intrinsic strains are expressed by

$$\hat{\mathbf{U}}(s, \epsilon) = \boldsymbol{\Omega}^T(\omega s) \left(\frac{\epsilon}{1 + \epsilon\omega}\, \check{\mathbf{v}}(s) + \omega \mathbf{e}_3 \right). \tag{6.86}$$

In the Hamiltonian variables relative to the new framing, the equilibrium equations are

$$\boldsymbol{x}'(s) = \boldsymbol{J} H_{\boldsymbol{x}}(\boldsymbol{x}(s), s, \tfrac{s}{\epsilon}, \epsilon), \tag{6.87}$$

with Hamiltonian

$$H(\boldsymbol{x}(s), s, \tfrac{s}{\epsilon}, \epsilon) = \tfrac{1}{2}\mathbf{M}(s)^T \boldsymbol{C}(s, \tfrac{s}{\epsilon})\mathbf{M}(s) + \mathbf{M}(s)^T \hat{\mathbf{U}}(s, \epsilon) + \boldsymbol{n}(s)^T \boldsymbol{D}_3(s). \tag{6.88}$$

Following section 6.3, we consider the averaged equilibrium equation

$$\boldsymbol{x}'(s) = \boldsymbol{J}\langle H\rangle_{\boldsymbol{x}}(\boldsymbol{x}(s), s), \tag{6.89}$$

with Hamiltonian

$$\langle H\rangle(\boldsymbol{x}(s), s) = \tfrac{1}{2}\mathbf{M}(s)^T \langle \boldsymbol{C}\rangle(s)\mathbf{M}(s) + \mathbf{M}(s)^T \hat{\mathbf{U}}(s, 0) + \boldsymbol{n}(s)^T \boldsymbol{D}_3(s). \tag{6.90}$$

This equation governs the equilibria of an effective rod characterized by flexibility matrix $\langle \boldsymbol{C}\rangle(s)$ given in (6.19), and intrinsic strains $\hat{\mathbf{U}}(s,0) = (0, 0, \omega)$. The effective rod is therefore isotropic and typically admits families of equilibria of the form

$$\Gamma(\boldsymbol{x}_0) = \big\{\boldsymbol{R}(\theta)\boldsymbol{x}_0 : \theta \in [0, 2\pi)\big\}, \tag{6.91}$$

where $\boldsymbol{x}_0$ is a particular solution and $\boldsymbol{R}(\theta)$ is the transformation matrix defined in (5.25).

The present situation is analogous to the one of the straight rods with high intrinsic twist discussed in section 6.5. If the hypotheses of theorem 3.5 hold, then the equilibria in $\Gamma(\boldsymbol{x}_0)$ corresponding to simple zeroes of the corresponding Melnikov function give rise to equilibria of the rod with small intrinsic curvature and high intrinsic twist. The Melnikov function can be split as

$$\mathcal{M}(\theta) = \mathcal{M}^t(\theta) + \mathcal{M}^c(\theta), \tag{6.92}$$

where the contribution $\mathcal{M}^t(\theta)$, associated with intrinsic twist, has the same expression as Melnikov function (6.27), and the contribution $\mathcal{M}^c(\theta)$, associated with intrinsic curvature, is defined by

$$\begin{aligned}\mathcal{M}^c(\theta) &= \frac{d}{d\theta}\int_0^1 \langle H\rangle_\epsilon(\boldsymbol{R}(\theta)\boldsymbol{x}_0(s), s)\, ds \\ &= \frac{d}{d\theta}\int_0^1 \mathbf{M}^0(s)^T \boldsymbol{\Omega}(\theta)^T \boldsymbol{\Omega}(\omega s)^T \check{\mathbf{v}}(s)\, ds.\end{aligned} \tag{6.93}$$

The second equality follows from (6.88), (6.86) and (4.41), and $\mathbf{M}^0(s) = \big(M_1^0(s), M_2^0(s), M_3^0(s)\big)$ represents the bending and twist moments associated with $\boldsymbol{x}_0$. Introducing the quantities

$$\begin{aligned} I_1 &= \int_0^1 \Big(M_1^0(s)\check{U}_2(s) - M_2^0(s)\check{U}_1(s)\Big)\, ds, \\ I_2 &= \int_0^1 \Big(-M_1^0(s)\check{U}_1(s) - M_2^0(s)\check{U}_2(s)\Big)\, ds, \end{aligned} \tag{6.94}$$

where $\check{U}_1(s)$ and $\check{U}_2(s)$ are defined by

$$\begin{pmatrix} \check{U}_1(s) \\ \check{U}_2(s) \end{pmatrix} = \begin{pmatrix} \cos(\omega s) & \sin(\omega s) \\ -\sin(\omega s) & \cos(\omega s) \end{pmatrix} \begin{pmatrix} \check{v}_1(s) \\ \check{v}_2(s) \end{pmatrix}, \tag{6.95}$$

equation (6.93) can be rewritten as

$$\mathcal{M}^c(\theta) = I_1 \cos\theta + I_2 \sin\theta. \tag{6.96}$$

Note that the condition given in Manning and Maddocks [37, equation (25)] for a circular equilibrium of a uniform, isotropic ring to persist under a small perturbation adding intrinsic curvature corresponds exactly to $\mathcal{M}^c(\theta) = 0$ in the present notation.

An argument similar to the one presented in section 6.5 shows that, in the general case, $\mathcal{M}(\theta)$ possesses at least two simple zeroes on $[0, 2\pi)$ and therefore, among a family $\Gamma(\boldsymbol{x}_0)$ of equilibria of the effective rod, at least two constitute approximations to equilibria of the rod with small intrinsic curvature and high intrinsic twist.

6.8 Example of a rod with small intrinsic curvature and high intrinsic twist

Consider a uniform rod with small intrinsic curvature and high intrinsic twist (cf. definition 6.3) characterized by $\check{\mathbf{v}}(s) = (c, 0, 0)$ and $\boldsymbol{K}(s) = \operatorname{diag}(K_1, K_2, K_3)$, with $K_1 \neq K_2$. Here c is a parameter controlling the value of the intrinsic curvature, and the centerline of the rod in the intrinsic configuration corresponds to an arc of a large circle (of radius $\frac{\varepsilon}{c}$). Assume the rod is subject to the ring boundary conditions (4.34). In this section, we explicitly compute the Melnikov function associated with the circular equilibria of the corresponding effective rod, and discuss the dependence of its zeroes on c.

By section 6.7, the effective rod is characterized by intrinsic strains $\hat{\mathbf{U}}(s, 0) = (0, 0, \omega)$ and stiffness matrix $\boldsymbol{K}(s) = \operatorname{diag}(\bar{K}, \bar{K}, K_3)$, where $\bar{K}$ is defined in (6.32). The circular equilibria of the effective rod, described in sections 4.6 and 5.4, occur in families of the form

$$\Gamma(\boldsymbol{x}_0) = \{\boldsymbol{R}(\theta)\boldsymbol{x}_0 : \theta \in [0, 2\pi)\}, \tag{6.97}$$

where $\boldsymbol{x}_0$ is a circular equilibrium lying in the $(\boldsymbol{e}_1, \boldsymbol{e}_3)$-plane.

According to equation (6.65), the part of the Melnikov function associated with intrinsic twist is

$$\mathcal{M}^t(\theta) = 4\pi^2 l^2 \bar{K}^2 \Delta \sin\alpha \sin(2\theta - \alpha), \tag{6.98}$$

where $\bar{K}$ and Δ are defined by (6.32) and l indicates the number of times the centerline of the solution $\boldsymbol{x}_0$ covers itself. If $\alpha \notin \{0, \pi\}$, $\mathcal{M}^t(\theta)$ admits four simple zeroes given by (6.66); if $\alpha \in \{0, \pi\}$, $\mathcal{M}^t(\theta) = 0$, $\forall\, \theta$.

Using expressions given in (6.64) for the bending moments of equilibrium $\boldsymbol{x}_0$, we find that the quantities I_1 and I_2 defined by equation (6.94) are

$$\begin{aligned} I_1 &= -2\pi l \bar{K} c \int_0^1 \cos(\alpha s - \omega s + 2\pi j s)\, ds, \\ I_2 &= -2\pi l \bar{K} c \int_0^1 \sin(\alpha s - \omega s + 2\pi j s)\, ds, \end{aligned} \tag{6.99}$$

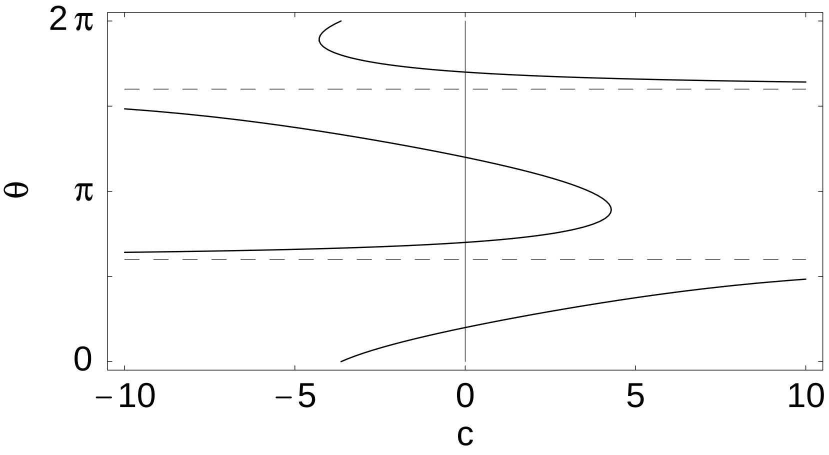


Figure 6.5: *Zeroes of the Melnikov function versus the intrinsic curvature parameter c, for $\alpha = \frac{2\pi}{5}$, $\omega = \frac{\pi}{5}$, $l = 1$, $j = 0$, $K_1 = 1$, $K_2 = 4$. The two horizontal dashed lines indicate the asymptotic values of the zeroes as $c \to \pm\infty$.*

so that the part of the Melnikov function associated with intrinsic curvature, expressed in (6.96), is

$$\begin{aligned}\mathcal{M}^c(\theta) &= -2\pi l\bar{K}c\int_0^1 \cos(\theta - \alpha s + \omega s - 2\pi j s)\,ds \\ &= \begin{cases} -\dfrac{4\pi l\bar{K}c}{\alpha-\omega+2\pi j}\sin\left(\dfrac{\alpha-\omega}{2}\right)\cos\left(\theta - \dfrac{\alpha-\omega}{2}\right) & \text{if } \alpha-\omega+2\pi j \neq 0, \\ -2\pi l\bar{K}c\cos\theta & \text{if } \alpha-\omega+2\pi j = 0. \end{cases}\end{aligned} \tag{6.100}$$

If $\frac{\alpha-\omega}{2\pi} \notin \mathbb{Z}\backslash\{-j\}$ and $c \neq 0$, $\mathcal{M}^c(\theta)$ admits 2 simple zeroes given by

$$\theta = \left(\frac{\pi}{2} + \frac{\alpha-\omega}{2}\right) (\mathrm{mod}\ \pi), \quad \left(\frac{\pi}{2} + \frac{\alpha-\omega}{2}\right) (\mathrm{mod}\ \pi) + \pi. \tag{6.101}$$

If $\frac{\alpha-\omega}{2\pi} \in \mathbb{Z}\backslash\{-j\}$ or $c = 0$, $\mathcal{M}^c(\theta) = 0, \forall\,\theta$.

It is instructive to consider the evolution of the zeroes of the Melnikov function $\mathcal{M}(\theta) = \mathcal{M}^t(\theta) + \mathcal{M}^c(\theta)$ as the intrinsic curvature parameter c varies. In the general case where $\alpha \notin \{0,\pi\}$ and $\frac{\alpha-\omega}{2\pi} \notin \mathbb{Z}\backslash\{-j\}$, the evolution, illustrated by the example of figure 6.5, has the following properties. When $c = 0$, the zeroes coincide with the four zeroes of $\mathcal{M}^t(\theta)$ given by (6.66). As c goes away from 0, four zeroes coexist, until two of them come together and vanish at some finite value of c. For larger values of c, two zeroes coexist and tend asymptotically to the zeroes of $\mathcal{M}^c(\theta)$ given by (6.101) as $|c| \to \infty$. In the case $\alpha \in \{0,\pi\}$ and $\frac{\alpha-\omega}{2\pi} \notin \mathbb{Z}\backslash\{-j\}$, the Melnikov function reduces to $\mathcal{M}^c(\theta)$. This function is identically zero when $c = 0$, and admits the two zeroes expressed by (6.101) when $c \neq 0$. In the case $\alpha \notin \{0,\pi\}$ and $\frac{\alpha-\omega}{2\pi} \in \mathbb{Z}\backslash\{-j\}$, the Melnikov function reduces to $\mathcal{M}^t(\theta)$. This function is independent of c and admits the four zeroes given by (6.66). Finally, in the case $\alpha \in \{0,\pi\}$ and $\frac{\alpha-\omega}{2\pi} \in \mathbb{Z}\backslash\{-j\}$, the Melnikov function vanishes for all values of θ and c.

6.9 Buckling loads of an intrinsically twisted strut

In this section, we obtain an expression determining the buckling loads of the intrinsically straight and twisted clamped-free strut discussed in chapter 1 (cf. figure 1.2), for all values of the intrinsic twist. Then, for large values of the intrinsic twist, we estimate the differences between the first of these buckling loads and the first buckling load of the corresponding effective strut.

For any value of the load, the problem governing the equilibria of the intrinsically straight and twisted clamped-free strut admits a trivial solution corresponding to a straight configuration. We call *buckling loads* the values of the load for which the linearization of the problem about the trivial solution admits nonzero solutions, since we expect branches of buckled equilibria to bifurcate off the trivial branch at these special values of the load.

Consider a rod characterized by intrinsic stains $\hat{\mathbf{u}}(s) = (0, 0, \tau)$, with $\tau \in \mathbb{R}$, and a stiffness matrix $\boldsymbol{K}(s) = \operatorname{diag}(K_1, K_2, K_3)$, with $K_1 < K_2$. According to section 4.1, and using the same notation, the equilibrium equations for this rod can be expressed as

$$\begin{gathered}
\boldsymbol{r}' = \boldsymbol{d}_3, \quad \boldsymbol{n}' = \mathbf{0}, \quad \sum_{i=1}^{3} (m_i \boldsymbol{d}_i)' + \boldsymbol{r}' \times \boldsymbol{n} = \mathbf{0}, \\
u_1 = -{\boldsymbol{d}_3'}^T \boldsymbol{d}_2, \quad u_2 = {\boldsymbol{d}_3'}^T \boldsymbol{d}_1, \quad u_3 = {\boldsymbol{d}_1'}^T \boldsymbol{d}_2, \\
m_1 = K_1 u_1, \quad m_2 = K_2 u_2, \quad m_3 = K_3 (u_3 - \tau),
\end{gathered} \tag{6.102}$$

where $(\boldsymbol{d}_1, \boldsymbol{d}_2, \boldsymbol{d}_3)$ forms a right-handed, orthonormal frame. Assume the rod is subject to the clamped-free strut boundary conditions

$$\boldsymbol{r}(0) = \mathbf{0}, \quad \boldsymbol{n}(1) = -\lambda \boldsymbol{e}_3, \quad \boldsymbol{d}_i(0) = \boldsymbol{e}_i, \quad m_i(1) = 0, \quad i = 1, 2, 3, \tag{6.103}$$

with load $\lambda > 0$. Substituting the expression $\boldsymbol{n}(s) = -\lambda\, \boldsymbol{e}_3$ and eliminating the variables $\boldsymbol{r}$ and u_i, problem (6.102)+(6.103) reduces to the system

$$\begin{gathered}
\sum_{i=1}^{3} (m_i \boldsymbol{d}_i)' - \lambda\, \boldsymbol{d}_3 \times \boldsymbol{e}_3 = \mathbf{0}, \\
m_1 = -K_1 {\boldsymbol{d}_3'}^T \boldsymbol{d}_2, \quad m_2 = K_2 {\boldsymbol{d}_3'}^T \boldsymbol{d}_1, \quad m_3 = K_3 \big({\boldsymbol{d}_1'}^T \boldsymbol{d}_2 - \tau \big),
\end{gathered} \tag{6.104}$$

with boundary conditions

$$\boldsymbol{d}_i(0) = \boldsymbol{e}_i, \quad m_i(1) = 0, \quad i = 1, 2, 3. \tag{6.105}$$

For any value of the load λ, problem (6.104)+(6.105) admits the trivial solution

$$\begin{gathered}
\boldsymbol{d}_1^0(s) = \begin{pmatrix} \cos \tau s \\ \sin \tau s \\ 0 \end{pmatrix}, \quad \boldsymbol{d}_2^0(s) = \begin{pmatrix} -\sin \tau s \\ \cos \tau s \\ 0 \end{pmatrix}, \quad \boldsymbol{d}_3^0(s) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \\
m_i(s) = 0, \quad i = 1, 2, 3,
\end{gathered} \tag{6.106}$$

corresponding to a straight configuration of the rod.

To obtain the linearized problem about this trivial solution, we consider expansions of the form

$$m_i = m_i^0 + \varepsilon z_i + O(\varepsilon^2), \quad i = 1,2,3, \tag{6.107}$$

and

$$\boldsymbol{d}_i = \boldsymbol{d}_i^0 + \varepsilon \boldsymbol{\xi}_i + O(\varepsilon^2), \quad i = 1,2,3, \tag{6.108}$$

where ε is a small parameter. Since $(\boldsymbol{d}_1, \boldsymbol{d}_2, \boldsymbol{d}_3)$ constitutes an orthonormal frame, there exists a Darboux vector $\boldsymbol{w}$ such that the latter expansion can be rewritten as

$$\boldsymbol{d}_i = \boldsymbol{d}_i^0 + \varepsilon \boldsymbol{w} \times \boldsymbol{d}_i^0 + O(\varepsilon^2), \quad i = 1,2,3. \tag{6.109}$$

Introduce the components $w_i = \boldsymbol{w}^T \boldsymbol{e}_i$ of the Darboux vector relative to the fixed frame, and the rotation matrix

$$\boldsymbol{P}(s) = \begin{pmatrix} \cos \tau s & -\sin \tau s \\ \sin \tau s & \cos \tau s \end{pmatrix}. \tag{6.110}$$

Substituting (6.107) and (6.109) into problem (6.104)+(6.105), collecting the coefficients of ε, and using expression (6.106), we obtain the linearized system

$$\begin{aligned} &\left[\boldsymbol{P} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right]' + \lambda \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \boldsymbol{0}, \quad z_3' = 0, \\ &\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} K_1 & 0 \\ 0 & K_2 \end{pmatrix} \boldsymbol{P}^T \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}', \quad z_3 = K_3 w_3', \end{aligned} \tag{6.111}$$

together with the boundary conditions

$$w_i(0) = 0, \quad z_i(1) = 0, \quad i = 1,2,3. \tag{6.112}$$

It immediately follows that

$$z_3(s) = 0, \quad w_3(s) = 0. \tag{6.113}$$

We proceed by transforming the equations for the remaining variables into a system that is easy to solve. Equation (6.111) leads to

$$\begin{aligned} &\boldsymbol{P}^T \left[\boldsymbol{P} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right]'' + \lambda \boldsymbol{P}^T \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}' = \boldsymbol{0}, \\ &\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}' = \boldsymbol{P} \begin{pmatrix} \frac{1}{K_1} & 0 \\ 0 & \frac{1}{K_2} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}. \end{aligned} \tag{6.114}$$

Substituting the second equality into the first one and simplifying, we find the system of differential equations with constant coefficients

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}'' + \begin{pmatrix} 0 & -2\tau \\ 2\tau & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}' + \begin{pmatrix} \frac{\lambda}{K_1} - \tau^2 & 0 \\ 0 & \frac{\lambda}{K_2} - \tau^2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \boldsymbol{0}. \tag{6.115}$$

According to equations (6.112) and (6.111), the associated boundary conditions are

$$z_1'(0) = \tau z_2(0), \quad z_2'(0) = -\tau z_1(0), \quad z_1(1) = 0, \quad z_2(1) = 0. \tag{6.116}$$

To determine the values of the load λ for which the linearized problem admits nonzero solutions, we distinguish five different cases defined in the following box.

case 0	:	$\tau = 0$
case 1	:	$\tau \neq 0, \quad 0 < \lambda < K_1\tau^2$ or $\lambda > K_2\tau^2$
case 2	:	$\tau \neq 0, \quad K_1\tau^2 < \lambda < K_2\tau^2$
case 3	:	$\tau \neq 0, \quad \lambda = K_1\tau^2$
case 4	:	$\tau \neq 0, \quad \lambda = K_2\tau^2$

Case 0: $\tau = 0$

The general solution to linearized equation (6.115) is

$$z_1(s) = a\cos\sqrt{\frac{\lambda}{K_1}}s + b\sin\sqrt{\frac{\lambda}{K_1}}s, \quad z_2(s) = c\cos\sqrt{\frac{\lambda}{K_2}}s + d\sin\sqrt{\frac{\lambda}{K_2}}s, \tag{6.117}$$

with $a, b, c, d \in \mathbb{R}$. From boundary conditions (6.116), this solution can be nonzero only if

$$\lambda = \left(j + \tfrac{1}{2}\right)^2 \pi^2 K_1 \quad \text{or} \quad \lambda = \left(j + \tfrac{1}{2}\right)^2 \pi^2 K_2, \quad j = 0, 1, 2, \ldots. \tag{6.118}$$

These values, corresponding to the buckling loads for a clamped-free strut with no intrinsic twist, are well-known.

Case 1: $\tau \neq 0$ **and** $(0 < \lambda < K_1\tau^2$ **or** $\lambda > K_2\tau^2)$

Introduce the quantities

$$\begin{gathered}\bar{K} = \frac{2}{K_1^{-1} + K_2^{-1}}, \quad \Delta = \frac{1}{2}\left(K_1^{-1} - K_2^{-1}\right), \quad R = \sqrt{\lambda^2\Delta^2 + 4\lambda\bar{K}^{-1}\tau^2}, \\ \rho_1, \rho_2 = \sqrt{\lambda\bar{K}^{-1} + \tau^2 \pm R}, \quad \Omega_1, \Omega_2 = -\lambda\Delta + 2\tau^2 \pm R.\end{gathered} \tag{6.119}$$

Note that, in this case, $\lambda\bar{K}^{-1} + \tau^2 - R > 0$, and ρ_2 is real. The general solution to equation (6.115) is

$$\begin{aligned} z_1(s) &= 2\tau\rho_1\big(a\sin\rho_1 s - b\cos\rho_1 s\big) + 2\tau\rho_2\big(c\sin\rho_2 s - d\cos\rho_2 s\big), \\ z_2(s) &= \Omega_1\big(a\cos\rho_1 s + b\sin\rho_1 s\big) + \Omega_2\big(c\cos\rho_2 s + d\sin\rho_2 s\big), \end{aligned} \tag{6.120}$$

with $a, b, c, d \in \mathbb{R}$. Boundary conditions (6.116) require the vector (a, b, c, d) to belong to the null space of the matrix

$$\boldsymbol{A}_1 = \begin{pmatrix} 2\rho_1^2 - \Omega_1 & 0 & 2\rho_2^2 - \Omega_2 & 0 \\ 0 & \rho_1(\Omega_1 - 2\tau^2) & 0 & \rho_2(\Omega_2 - 2\tau^2) \\ \rho_1\sin\rho_1 & -\rho_1\cos\rho_1 & \rho_2\sin\rho_2 & -\rho_2\cos\rho_2 \\ \Omega_1\cos\rho_1 & \Omega_1\sin\rho_1 & \Omega_2\cos\rho_2 & \Omega_2\sin\rho_2 \end{pmatrix}. \tag{6.121}$$

The solution to the linearized problem can be nonzero only if $\det \boldsymbol{A}_1 = 0$. After simplifications, we find $\det \boldsymbol{A}_1 = g f_1$, where

$$g = -8\lambda \left(\frac{\lambda}{K_1} - \tau^2 \right) \tag{6.122}$$

and

$$\begin{aligned} f_1 = 2\bar{K}^{-1}\tau^2 \sqrt{\left(\lambda\bar{K}^{-1} - \tau^2\right)^2 - \lambda^2\Delta^2} - \left(2\lambda\bar{K}^{-2} - \lambda\Delta^2 - 2\bar{K}^{-1}\tau^2\right)\tau^2 \sin\rho_1 \sin\rho_2 \\ + \left(2\bar{K}^{-1}\tau^2 + \lambda\Delta^2\right)\sqrt{\left(\lambda\bar{K}^{-1} - \tau^2\right)^2 - \lambda^2\Delta^2} \cos\rho_1 \cos\rho_2. \end{aligned} \tag{6.123}$$

Note that f_1 is ultimately a function of λ, τ, K_1 and K_2. In this case, $g \neq 0$, and the buckling loads are given implicitly by the values of λ for which $f_1 = 0$.

Case 2: $\tau \neq 0$ and $K_1\tau^2 < \lambda < K_2\tau^2$

Consider the quantities defined in (6.119). In the present case, $\lambda\bar{K}^{-1} + \tau^2 - R < 0$ and, as a substitute for ρ_2 which is imaginary, we introduce the real quantity

$$\nu = \sqrt{-\lambda\bar{K}^{-1} - \tau^2 + R}. \tag{6.124}$$

The general solution to equation (6.115) is

$$\begin{aligned} z_1(s) &= 2\tau\rho_1\left(a \sin\rho_1 s - b\cos\rho_1 s\right) - 2\tau\nu\left(c\sinh\nu s + d\cosh\nu s\right), \\ z_2(s) &= \Omega_1\left(a\cos\rho_1 s + b\sin\rho_1 s\right) + \Omega_2\left(c\cosh\nu s + d\sinh\nu s\right), \end{aligned} \tag{6.125}$$

with $a, b, c, d \in \mathbb{R}$. Boundary conditions (6.116) require the vector (a, b, c, d) to belong to the null space of the matrix

$$\boldsymbol{A}_2 = \begin{pmatrix} 2\rho_1^2 - \Omega_1 & 0 & -2\nu^2 - \Omega_2 & 0 \\ 0 & \rho_1(\Omega_1 - 2\tau^2) & 0 & \nu(\Omega_2 - 2\tau^2) \\ \rho_1 \sin\rho_1 & -\rho_1\cos\rho_1 & -\nu\sinh\nu & -\nu\cosh\nu \\ \Omega_1\cos\rho_1 & \Omega_1\sin\rho_1 & \Omega_2\cosh\nu & \Omega_2\sinh\nu \end{pmatrix}. \tag{6.126}$$

We have $\det \boldsymbol{A}_2 = g f_2$, where g is defined in (6.122) and

$$\begin{aligned} f_2 = 2\bar{K}^{-1}\tau^2 \sqrt{\lambda^2\Delta^2 - \left(\lambda\bar{K}^{-1} - \tau^2\right)^2} - \left(2\lambda\bar{K}^{-2} - \lambda\Delta^2 - 2\bar{K}^{-1}\tau^2\right)\tau^2 \sin\rho_1 \sinh\nu \\ + \left(2\bar{K}^{-1}\tau^2 + \lambda\Delta^2\right)\sqrt{\lambda^2\Delta^2 - \left(\lambda\bar{K}^{-1} - \tau^2\right)^2} \cos\rho_1 \cosh\nu. \end{aligned} \tag{6.127}$$

Note that f_2 is ultimately a function of λ, τ, K_1 and K_2. In this case, $g \neq 0$, and the buckling loads are given implicitly by the values of λ for which $f_2 = 0$.

Case 3: $\tau \neq 0$ and $\lambda = K_1\tau^2$

Define the numbers

$$\alpha = \frac{K_1}{K_2}, \quad \zeta = \tau\sqrt{3+\alpha}. \tag{6.128}$$

The general solution to equation (6.115), with λ replaced by $K_1\tau^2$, is

$$\begin{aligned} z_1(s) &= 2\big(a\sin\zeta s - b\cos\zeta s\big) + c\tau(1-\alpha) + d\tau(1-\alpha)s, \\ z_2(s) &= \sqrt{3+\alpha}\big(a\cos\zeta s + b\sin\zeta s\big) + 2d, \end{aligned} \tag{6.129}$$

where $a, b, c, d \in \mathbb{R}$. Boundary conditions (6.116) require the vector (a, b, c, d) to belong to the null space of the matrix

$$\boldsymbol{A}_3 = \begin{pmatrix} \sqrt{3+\alpha} & 0 & 0 & -1-\alpha \\ 0 & 1+\alpha & \tau(1-\alpha) & 0 \\ 2\sin\zeta & -2\cos\zeta & \tau(1-\alpha) & \tau(1-\alpha) \\ \sqrt{3+\alpha}\cos\zeta & \sqrt{3+\alpha}\sin\zeta & 0 & 2 \end{pmatrix}. \tag{6.130}$$

We have $\det \boldsymbol{A}_3 = \zeta(1-\alpha)f_3$, where

$$f_3 = 4(1+\alpha) + \big(\alpha^2 + 2\alpha + 5\big)\cos\zeta + (1-\alpha)\zeta\sin\zeta. \tag{6.131}$$

Note that f_3 is ultimately a function of τ, K_1 and K_2. Since $\zeta(1-\alpha) \neq 0$, the value $\lambda = K_1\tau^2$ is a buckling load if $f_3 = 0$.

Case 4: $\tau \neq 0$ and $\lambda = K_2\tau^2$

Define the numbers

$$\beta = \frac{K_2}{K_1}, \quad \eta = \tau\sqrt{3+\beta}. \tag{6.132}$$

The general solution to equation (6.115), with λ replaced by $K_2\tau^2$, is

$$\begin{aligned} z_1(s) &= \sqrt{3+\beta}\big(a\sin\eta s - b\cos\eta s\big) - 2d, \\ z_2(s) &= 2\big(a\cos\eta s + b\sin\eta s\big) + c\tau(1-\beta) + d\tau(1-\beta)s, \end{aligned} \tag{6.133}$$

where $a, b, c, d \in \mathbb{R}$. Boundary conditions (6.116) require the vector (a, b, c, d) to belong to the null space of the matrix

$$\boldsymbol{A}_4 = \begin{pmatrix} 1+\beta & 0 & -\tau(1-\beta) & 0 \\ 0 & \sqrt{3+\beta} & 0 & -1-\beta \\ \sqrt{3+\beta}\sin\eta & -\sqrt{3+\beta}\cos\eta & 0 & -2 \\ 2\cos\eta & 2\sin\eta & \tau(1-\beta) & \tau(1-\beta) \end{pmatrix}. \tag{6.134}$$

We have $\det \boldsymbol{A}_4 = \eta(1-\beta)f_4$, where

$$f_4 = 4(1+\beta) + \big(\beta^2 + 2\beta + 5\big)\cos\eta + (1-\beta)\eta\sin\eta. \tag{6.135}$$

Note that f_4 is ultimately a function of τ, K_1 and K_2. Since $\eta(1-\beta) \neq 0$, the value $\lambda = K_2\tau^2$ is a buckling load if $f_4 = 0$.

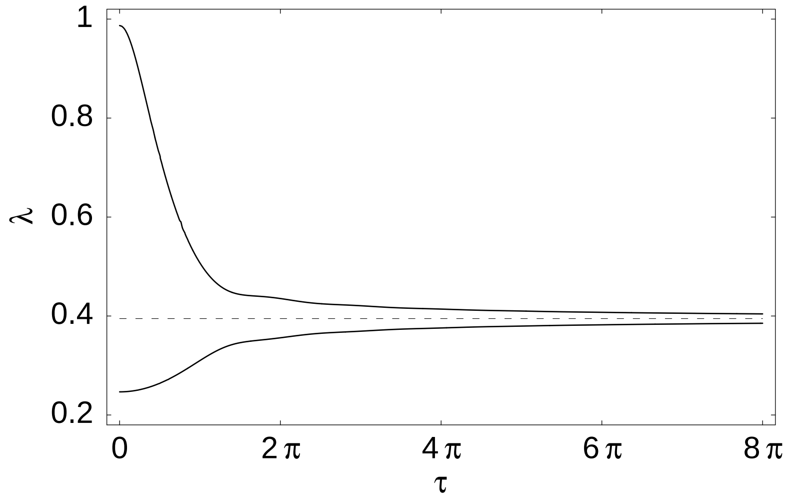


Figure 6.6: *First two buckling loads versus the intrinsic twist τ, for $K_1 = 0.1$, $K_2 = 0.4$. The horizontal dashed line indicates the value $\frac{1}{4}\pi^2\bar{K}$ corresponding to the first buckling load of the effective strut.*

The situation can be summarized as follows. For a clamped-free strut with no intrinsic twist, the buckling loads are expressed by equation (6.118). For a clamped-free strut with intrinsic twist $\tau \neq 0$, the buckling loads are the values of λ for which one of the following conditions holds:

$$
\begin{aligned}
&(1) \quad (0 < \lambda < K_1\tau^2 \text{ or } \lambda > K_2\tau^2) \text{ and } f_1 = 0,\\
&(2) \quad K_1\tau^2 < \lambda < K_2\tau^2 \text{ and } f_2 = 0,\\
&(3) \quad \lambda = K_1\tau^2 \text{ and } f_3 = 0,\\
&(4) \quad \lambda = K_2\tau^2 \text{ and } f_4 = 0.
\end{aligned}
$$

Note that the buckling loads depend on the absolute value of the intrinsic twist τ, but not on its sign, as well as on the bending stiffnesses K_1 and K_2, but not on the twist stiffness K_3. Figure 6.6 represents the first two buckling loads as functions of the intrinsic twist τ, for fixed values of the bending stiffnesses. At $\tau = 0$, the buckling loads take the values $\frac{1}{4}\pi^2 K_1$ and $\frac{1}{4}\pi^2 K_2$. As τ moves away from zero, they approach the first buckling load $\frac{1}{4}\pi^2\bar{K}$ of the associated effective strut fairly rapidly.

For large values of the intrinsic twist, an expansion of λ in powers of the small parameter $\frac{1}{\tau}$ in the equation $f_1 = 0$ shows that the first two buckling loads can be expressed as

$$
\frac{\pi^2\bar{K}}{4}\left(1 \pm \bar{K}\Delta\frac{1}{\tau} + O\left(\frac{1}{\tau^2}\right)\right). \tag{6.136}
$$

We conclude that the relative error between each of the two buckling loads of the strut with high intrinsic twist and the first buckling load of the effective strut is essentially given by $\bar{K}\Delta\left|\frac{1}{\tau}\right|$. This error is the product of the inverse of the intrinsic twist by the quantity

$$\bar{K}\Delta = \frac{K_2 - K_1}{K_2 + K_1} \in [0, 1), \tag{6.137}$$

which can be interpreted as a measure of the anisotropy of the constitutive relation of the strut.

Chapter 7

Conclusion

We conclude by presenting a summary of the main results of this dissertation, along with a discussion of alternative approaches and possible extensions. The first part of the thesis deals with perturbations in general Hamiltonian two-point BVPs, the second discusses averaging for Hamiltonian two-point BVPs depending on two scales, and the last concerns an application of the above techniques to the study of elastic rods with high intrinsic twist subject to two-point boundary conditions.

Perturbations in Hamiltonian BVPs

The Hamiltonian two-point BVP

$$\begin{cases} \boldsymbol{x}'(s) = \boldsymbol{J}H_{\boldsymbol{x}}(\boldsymbol{x}(s), s, \epsilon), \quad \forall\, s \in (0,1), \\ \boldsymbol{b}(\boldsymbol{x}(0)) = \boldsymbol{0}, \quad \boldsymbol{c}(\boldsymbol{x}(1)) = \boldsymbol{0}, \end{cases} \tag{7.1}$$

where H, $\boldsymbol{b}$, $\boldsymbol{c}$ are smooth functions and ϵ is a small parameter, is studied in chapter 2. The problem with $\epsilon = 0$ is called the unperturbed problem. First consider the case where the unperturbed problem possesses an isolated solution $\boldsymbol{x}_0$. Then, according to theorem 2.2, under nondegeneracy condition (**2.A3**) and for ϵ sufficiently small, problem (7.1) possesses a solution $\boldsymbol{\xi}(\epsilon)$, where $\boldsymbol{\xi}(0) = \boldsymbol{x}_0$ and $\boldsymbol{\xi}$ is a smooth function of ϵ. Hence $\boldsymbol{x}_0$ is the leading order approximation to $\boldsymbol{\xi}(\epsilon)$, and $\boldsymbol{\xi}(\epsilon)$ admits the expansion

$$\boldsymbol{\xi}(\epsilon) = \boldsymbol{x}_0 + \epsilon \boldsymbol{x}_1 + \epsilon^2 \boldsymbol{x}_2 + \dots, \tag{7.2}$$

where the corrections $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots$ can be determined by a perturbation expansion method. For example, the first correction $\boldsymbol{x}_1$ is the solution to problem (2.19).

Now consider the case where the unperturbed problem possesses a one-dimensional manifold of solutions described by $\boldsymbol{\zeta}(\theta)$, where $\theta \in I \subset \mathbb{R}$ (cf. assumption (**2.A6**)). Assume the boundary conditions in (7.1) enjoy property (**2.A5**) (and therefore define Lagrangian manifolds, cf. lemma 2.5), and define the Melnikov function

$$\mathcal{M}(\theta) = \frac{d}{d\theta} \int_0^1 H_\epsilon(\boldsymbol{\zeta}(\theta)(s), s, 0)\, ds. \tag{7.3}$$

Let $\boldsymbol{x}^* = \boldsymbol{\zeta}(\theta^*)$ denote a member of the manifold of solutions to the unperturbed problem. Then, according to theorem 2.10, under nondegeneracy condition (**2.A7**) and for ϵ sufficiently small, we have the following: if θ^* does not correspond to a zero of the Melnikov function, problem (7.1) possesses no solution in some neighborhood of $\boldsymbol{x}^*$; contrariwise, if θ^* corresponds to a simple zero of the Melnikov function, problem (7.1) possesses a solution $\boldsymbol{\xi}(\epsilon)$, where $\boldsymbol{\xi}(0) = \boldsymbol{x}^*$ and $\boldsymbol{\xi}$ is a smooth function of ϵ. In the latter case, $\boldsymbol{x}^*$ is the leading order approximation to $\boldsymbol{\xi}(\epsilon)$, and $\boldsymbol{\xi}(\epsilon)$ admits the expansion

$$\boldsymbol{\xi}(\epsilon) = \boldsymbol{x}^* + \epsilon \boldsymbol{x}_1 + \epsilon^2 \boldsymbol{x}_2 + \dots, \tag{7.4}$$

where the corrections $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots$ can be determined by a perturbation expansion method. For example, the first correction $\boldsymbol{x}_1$ is the unique solution to problem (2.121) making problem (2.122) solvable. It can be expressed as

$$\boldsymbol{x}_1 = \boldsymbol{v}_1^* - \frac{F_0}{\mathcal{M}_\theta(\theta^*)} \boldsymbol{x}_\theta^*, \tag{7.5}$$

where $\boldsymbol{x}_\theta^* = \boldsymbol{\zeta}_\theta(\theta^*)$, $\boldsymbol{v}_1^*$ is the unique solution to (2.121) satisfying $\int_0^1 \boldsymbol{v}_1^*(s)^T \boldsymbol{x}_\theta^*(s)\, ds = 0$, and F_0 is given by (2.127)

The above symmetry breaking results are related to the ones obtained by Ambrosetti and Badiale [3, 4] via a variational approach. Consider the functional $f_{(\epsilon)} : E \to \mathbb{R}$ of the form

$$f_{(\epsilon)}(\boldsymbol{u}) = f_{(0)}(\boldsymbol{u}) + \epsilon\, G(\boldsymbol{u}) + o(\epsilon), \tag{7.6}$$

where E is a Hilbert space and ϵ is a small parameter. Assume the unperturbed functional $f_{(0)}$ possesses a d-dimensional manifold

$$\Gamma = \left\{ \boldsymbol{\zeta}(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \mathbb{R}^d \right\} \tag{7.7}$$

of critical points. Define the Poincaré function $\mathcal{P} : \mathbb{R}^d \to \mathbb{R}$ by $\mathcal{P}(\boldsymbol{\theta}) = G(\boldsymbol{\zeta}(\boldsymbol{\theta}))$. Then, under some additional conditions (regularity, nondegeneracy, etc.), and for ϵ sufficiently small, we have the following.

(a) If $\boldsymbol{\theta}_1$ corresponds to a proper local minimum or maximum of $\mathcal{P}$, then the functional $f_{(\epsilon)}$ possesses a critical point $\boldsymbol{\xi}_{(\epsilon)}$, where $\boldsymbol{\xi}_{(\epsilon)} \to \boldsymbol{\zeta}(\boldsymbol{\theta}_1)$ as $\epsilon \to 0$.

(b) If $\boldsymbol{\theta}_2$ is a nondegenerate critical point of $\mathcal{P}$, i.e.

$$\mathcal{P}_{\boldsymbol{\theta}}(\boldsymbol{\theta}_2) = \boldsymbol{0} \quad \text{and} \quad \mathcal{P}_{\boldsymbol{\theta\theta}}(\boldsymbol{\theta}_2) \ \text{is invertible}, \tag{7.8}$$

then the functional $f_{(\epsilon)}$ possesses a critical point $\boldsymbol{\xi}_{(\epsilon)}$, where $\boldsymbol{\xi}_{(\epsilon)}$ is a smooth function of ϵ and $\boldsymbol{\xi}_{(0)} = \boldsymbol{\zeta}(\boldsymbol{\theta}_2)$.

Note that, in the particular case where the unperturbed functional possesses a one-dimensional manifold of critical points ($d = 1$), condition (7.8) is equivalent to the condition that $\boldsymbol{\theta}_2$ corresponds to a simple zero of the Melnikov function $\mathcal{M}(\boldsymbol{\theta}) = \mathcal{P}_{\boldsymbol{\theta}}(\boldsymbol{\theta})$.

Now consider a Hamiltonian BVP of the form (7.1), with the property that it is equivalent to a variational problem associated with a functional of the type (B.2) and imposed boundary conditions of the form (B.1). Note that, as shown in appendix B, each rod BVP considered in this dissertation is equivalent to such a variational problem.

Assume moreover that the imposed boundary conditions are *linear and homogeneous* (such as the clamped-free strut boundary conditions (B.27)). Then the problem under consideration is equivalent to the problem of finding critical points of a functional of the form (7.6), where the boundary conditions are included in the definition of the Hilbert space E. In this case, the symmetry breaking results of Ambrosetti and Badiale apply and constitute a generalization of the symmetry breaking results presented in this dissertation. However, in the case of more general imposed boundary conditions (such as the *nonlinear* clamped-sleeved strut boundary conditions (B.29)), we do not know how to treat the symmetry breaking of the problem via a variational approach. Note moreover that the bifurcation method used to obtain symmetry breaking results in this dissertation directly extends to the case of a non-Hamiltonian perturbation, whereas a variational approach does not.

Averaging in Hamiltonian BVPs

The two-scale Hamiltonian BVP

$$\begin{cases} \boldsymbol{x}'(s) = \boldsymbol{J} H_{\boldsymbol{x}}\big(\boldsymbol{x}(s), s, \tfrac{s}{\epsilon}, \epsilon\big), & \forall\, s \in (0,1), \\ \boldsymbol{b}(\boldsymbol{x}(0)) = \boldsymbol{0}, \quad \boldsymbol{c}(\boldsymbol{x}(1)) = \boldsymbol{0}, \end{cases} \tag{7.9}$$

where H, $\boldsymbol{b}$, $\boldsymbol{c}$ are smooth and $H(\boldsymbol{x}, s, \sigma, \epsilon)$ is T-periodic in σ, is studied in chapter 3. The associated averaged problem is

$$\begin{cases} \boldsymbol{y}'(s) = \boldsymbol{J} \langle H \rangle_{\boldsymbol{y}}(\boldsymbol{y}(s), s), \\ \boldsymbol{b}(\boldsymbol{y}(0)) = \boldsymbol{0}, \quad \boldsymbol{c}(\boldsymbol{y}(1)) = \boldsymbol{0}, \end{cases} \tag{7.10}$$

with Hamiltonian defined by

$$\langle H \rangle(\boldsymbol{y}, s) = \frac{1}{T} \int_0^T H(\boldsymbol{y}, s, \sigma, 0)\, d\sigma. \tag{7.11}$$

First consider the case where ϵ belongs to the set

$$\mathbb{E}_T = \left\{ \frac{1}{\nu T} : \nu \in \mathbb{Z} \backslash \{0\} \right\}. \tag{7.12}$$

Assume the averaged problem possesses an isolated solution $\boldsymbol{y}_0$. Then, according to theorem 3.4, under nondegeneracy condition (**3.A3**) and for $\epsilon \in \mathbb{E}_T$ sufficiently small, the two-scale problem possesses a solution $\boldsymbol{\xi}(\epsilon)$, where $\boldsymbol{\xi}(\epsilon)(s) \to \boldsymbol{y}_0(s)$ uniformly in $s \in [0,1]$ as $\epsilon \to 0$. The solution admits an expansion of the form

$$\boldsymbol{\xi}(\epsilon)(s) = \boldsymbol{y}_0(s) + \epsilon \boldsymbol{x}_1\big(s, \tfrac{s}{\epsilon}\big) + \epsilon^2 \boldsymbol{x}_2\big(s, \tfrac{s}{\epsilon}\big) + \ldots \tag{7.13}$$

about the leading order approximation $\boldsymbol{y}_0$, where, for $i = 1, 2, \ldots$, the function $\boldsymbol{x}_i\big(s, \sigma\big)$ is T-periodic in σ and $\boldsymbol{x}_i\big(s, \tfrac{s}{\epsilon}\big) = O(1)$ uniformly in $s \in [0,1]$ as $\epsilon \to 0$. A method to compute the corrections $\boldsymbol{x}_i$ is described in section 3.3. For example, the first correction is of the form

$$\boldsymbol{x}_1\big(s, \tfrac{s}{\epsilon}\big) = \boldsymbol{y}_1(s) + \boldsymbol{J} W^{(1)}_{\boldsymbol{y}}\big(\boldsymbol{y}_0(s), s, \tfrac{s}{\epsilon}\big), \tag{7.14}$$

where $\boldsymbol{y}_1(s)$ is the solution to problem (3.64) and $W^{(1)}$ is explicitly given by equation (3.39).

Now assume the averaged problem possesses a one-dimensional manifold of solutions described by $\boldsymbol{\zeta}(\theta)$, where $\theta \in I \subset \mathbb{R}$ (cf. assumption (**3.A5**)), and suppose the boundary conditions in (7.9) enjoy property (**3.A4**) (and therefore define Lagrangian manifolds). Define the Melnikov function

$$\mathcal{M}(\theta) = \frac{d}{d\theta} \int_0^1 \langle \mathcal{H}^{(1)} \rangle (\boldsymbol{\zeta}(\theta)(s), s) \, ds, \tag{7.15}$$

where $\mathcal{H}^{(1)}$ is given in (3.42). Let $\boldsymbol{y}^* = \boldsymbol{\zeta}(\theta^*)$ denote a member of the manifold of solutions to the averaged problem. Then, according to theorem 3.5, under nondegeneracy condition (**3.A6**) and for $\epsilon \in \mathbb{E}_T$ sufficiently small, we have the following: if θ^* does not correspond to a zero of the Melnikov function, the problem (7.9) possesses no solution in some neighborhood of $\boldsymbol{y}^*$; contrariwise, if θ^* corresponds to a simple zero of the Melnikov function, problem (7.9) possesses a solution $\boldsymbol{\xi}(\epsilon)$, where $\boldsymbol{\xi}(\epsilon)(s) \to \boldsymbol{y}^*(s)$ uniformly in $s \in [0,1]$ as $\epsilon \to 0$. The solution $\boldsymbol{\xi}(\epsilon)$ admits an expansion of the form

$$\boldsymbol{\xi}(\epsilon)(s) = \boldsymbol{y}^*(s) + \epsilon \boldsymbol{x}_1\big(s, \tfrac{s}{\epsilon}\big) + \epsilon^2 \boldsymbol{x}_2\big(s, \tfrac{s}{\epsilon}\big) + \dots \tag{7.16}$$

about the leading order approximation $\boldsymbol{y}^*$, where, for $i = 1, 2, \dots$, the function $\boldsymbol{x}_i\big(s, \sigma\big)$ is T-periodic in σ and $\boldsymbol{x}_i\big(s, \frac{s}{\epsilon}\big) = O(1)$ uniformly in $s \in [0,1]$ as $\epsilon \to 0$. A method to determine the functions $\boldsymbol{x}_i$ is described in section 3.4. For example, the first correction is

$$\boldsymbol{x}_1\big(s, \tfrac{s}{\epsilon}\big) = \boldsymbol{y}_1(s) + \boldsymbol{J} W^{(1)}_{\boldsymbol{y}}\big(\boldsymbol{y}^*(s), s, \tfrac{s}{\epsilon}\big), \tag{7.17}$$

where $\boldsymbol{y}_1$, which can be written as in equation (3.83), is the unique solution to problem (3.81) making problem (3.82) solvable, and $W^{(1)}$ is explicitly given by equation (3.39).

Now consider the general case where $\epsilon \notin \mathbb{E}_T$, which is discussed in section 3.5. Inserting the numbers

$$\omega = \frac{1}{\epsilon} \pmod{T} \in [0, T), \qquad \delta = \frac{\epsilon}{1 - \omega\epsilon} \in \mathbb{E}_T, \tag{7.18}$$

the two-scale problem (7.9) transforms into a problem of the same type, but with δ as small parameter and ω as an additional parameter (cf. equation (3.90)). The above averaging results apply to the new problem, and the corresponding averaged problem is (7.10), regardless of the value of the value of ω. In the case where the averaged problem possesses an isolated solution $\boldsymbol{y}_0$, then, under the usual nondegeneracy condition and for ϵ sufficiently small, problem (7.9) possesses a solution with $\boldsymbol{y}_0$ as leading order approximation. This solution can be expanded as in (3.91). In the case where the averaged problem possesses a one-dimensional manifold of solutions, then, under the usual hypotheses and for ϵ sufficiently small, problem (7.9) possesses a solution associated with each simple zero of the Melnikov function. Such a solution admits an expansion of the form (3.92). However, in general, the Melnikov function depends on the value of ω and, among the family of solutions to the effective problem, the set of members constituting approximations to solutions to (7.9) also depends on ω.

Elastic rods with high intrinsic twist

Elastic rods with high intrinsic twist subject to clamped-free strut, clamped-sleeved strut or ring boundary conditions are studied in chapter 6. In the case of an intrinsically curved rod with high intrinsic twist, the effective rod is also intrinsically curved (with the same intrinsic curvature) and has an effective stiffness matrix that makes it constitutively isotropic. In general, the equilibria of this effective rod are isolated and constitute approximations to the equilibria of the rod with high intrinsic twist.

In the case of an intrinsically straight rod with high intrinsic twist, the effective rod is also intrinsically straight and has an effective stiffness matrix that makes it isotropic. In general, this effective rod possesses one-parameter families of equilibria and, among such a family, only a finite number of members, which can be identified by means of a Melnikov function, constitute approximations to equilibria of the rod with high intrinsic twist. These special members are at least two in the case of a general stiffness matrix, and at least four in the case of a stiffness matrix of the block-diagonal form (6.28). For the particular example of a uniform, intrinsically straight rod with a diagonal stiffness matrix and a high intrinsic twist, we have the following explicit results.

(a) Under the clamped-free strut boundary conditions, only those equilibria of the effective rod lying in the two vertical planes making angles of $\frac{\pi}{4}$ and $\frac{3\pi}{4}$ with the soft plane at the clamped end constitute approximations to equilibria of the rod with high intrinsic twist.

(b) Under the clamped-sleeved boundary conditions, among the planar equilibria of the effective rod, only those lying in two special vertical planes constitute approximations to equilibria of the rod with high intrinsic twist. The orientation of these two planes actually depends on the exact value of the high intrinsic τ; they make angles of $\frac{\omega}{2}$ and $\frac{\omega}{2} + \frac{\pi}{2}$, where $\omega = \tau \pmod{2\pi}$, with the soft plane at the clamped end.

(c) Under the ring boundary conditions, the situation depends on the value of the prescribed angle α. In the case $\alpha \notin \{0, \pi\}$, among the circular equilibria of the effective rod, only those lying in the two vertical planes equidistant to the two soft planes at the ends (and making angles of $\frac{\alpha}{2}$ and $\frac{\alpha}{2} + \frac{\pi}{2}$ with the soft plane at one of the ends) constitute approximations to equilibria of the rod with high intrinsic twist. In the case $\alpha \in \{0, \pi\}$, due to an additional continuous symmetry of the problem, in general, all circular equilibria of the effective rod constitute approximations to equilibria of the rod with high intrinsic twist.

In the case of a rod with small intrinsic curvature and high intrinsic twist, the effective rod is intrinsically straight and has an effective stiffness matrix that makes it isotropic. In general, this effective rod possesses one-parameter families of equilibria and, among such a family, only a finite number, but at least two members constitute approximations to equilibria of the rod with small intrinsic curvature and high intrinsic twist. The Melnikov function identifying these special members is the sum of two terms, one associated to intrinsic curvature and the other to intrinsic twist, which gives rise to a competition between the different effects in the determination of the symmetry breaking. In section (6.8), this competition is illustrated by explicit results for a particular example involving the ring boundary conditions. In this case, among a family of circular equilibria of the effective rod, the number of members constituting approximations to equilibria of the rod changes from four to two as the relative importance of the intrinsic curvature increases.

The approximation which consists in replacing an equilibrium of a rod with high intrinsic twist by an equilibrium of an effective rod corresponds to keeping only the leading order term in an expansion of the form (7.13) or (7.16), where ϵ is of the order of magnitude of the inverse of the intrinsic twist. One way to find an estimate for the error associated with this approximation is to compute the first correction expressed in (7.14) or (7.17). The error generally decreases as the intrinsic twist increases. The error also depends on the difference between the Hamiltonian (6.13) and the averaged Hamiltonian (6.16), which is directly linked to the difference between the original stiffness matrix $\boldsymbol{K}(s)$ and the isotropic effective stiffness matrix $\bar{\boldsymbol{K}}(s)$ defined in (6.21). In some sense (which we do not make precise here), the more anisotropic the stiffness matrix $\boldsymbol{K}(s)$ is, the larger the difference between the two Hamiltonians is, and the bigger the error is.

The dependence of the error on the value of the intrinsic twist and on the anisotropy of the stiffness matrix has been illustrated in the particular case of a uniform, intrinsically straight and twisted rod with a diagonal stiffness matrix. One illustration is provided by the computation of equilibria of the rod subject to initial conditions by Kehrbaum [27, sect. 9.7]. Another illustration is given by the expression for the buckling loads of the rod subject to the clamped-free strut boundary conditions presented in section 6.9. It appears that, for high values of the intrinsic twist, the error associated with the approximation of the first buckling load by the first buckling load of the corresponding effective strut is proportional both to the inverse of the intrinsic twist and to the quantity $|K_1 - K_2|/(K_1 + K_2)$ measuring the anisotropy of the stiffness matrix $\boldsymbol{K} = \mathrm{diag}\,(K_1, K_2, K_3)$.

We have considered only three sets of rod boundary conditions in this dissertation, but rods with high intrinsic twist subject to other boundary conditions can easily be analyzed using the same techniques. In the study of rods with high intrinsic twist subject to initial conditions by Kehrbaum [27], the general case of an extensible, shearable rod with a general constitutive relation is considered. It is shown that the corresponding equilibirum equations can be written under the Hamiltonian two-scale form

$$\boldsymbol{x}'(s) = \boldsymbol{J} H_{\boldsymbol{x}}(\boldsymbol{x}(s), s, \tfrac{s}{\epsilon}), \tag{7.19}$$

and that the averaged equation defines an effective rod which is constitutively isotropic. Combining these results with the averaging method for Hamiltonian two-point BVPs presented in this dissertation, it is possible to analyze the equilibria of extensible, shearable elastic rods with high intrinsic twist and general constitutive relation, subject to two-point boundary conditions.

Appendix A

Some standard theorems

This appendix contains basic theorems of nonlinear analysis and differential geometry that are used in this thesis. The proofs of these theorems can be found in the associated references.

Banach fixed-point theorem

Theorem A.1 *Let U be a closed nonempty subset of a Banach space X and $\mathbf{\Phi} : U \to U$. Suppose $\mathbf{\Phi}$ is a contraction, i.e. there is a number $k \in [0, 1)$ such that*

$$\|\mathbf{\Phi}(\boldsymbol{x}) - \mathbf{\Phi}(\boldsymbol{y})\| \leqslant k\|\boldsymbol{x} - \boldsymbol{y}\|, \quad \forall\, \boldsymbol{x}, \boldsymbol{y} \in U. \tag{A.1}$$

Then the equation $\boldsymbol{x} = \mathbf{\Phi}(\boldsymbol{x})$ admits exactly one solution in U.

This theorem, sometimes called the *contraction mapping theorem*, can be found in Zeidler [53, sect. 1.1], for example.

Taylor's theorem

Theorem A.2 *Consider a mapping $\boldsymbol{\xi} \in C^k((-\epsilon_0, \epsilon_0), X)$, where $k \geqslant 1$, $\epsilon_0 > 0$, and X is a Banach space. Then, for $\epsilon \in (-\epsilon_0, \epsilon_0)$,*

$$\boldsymbol{\xi}(\epsilon) = \sum_{n=0}^{k-1} \frac{1}{n!} d^n \boldsymbol{\xi}(0)\epsilon^n + \boldsymbol{R}_k(\epsilon), \tag{A.2}$$

where $d^n\boldsymbol{\xi}$ denotes the n^{th} Fréchet derivative of $\boldsymbol{\xi}$ (with $d^0\boldsymbol{\xi} \equiv \boldsymbol{\xi}$), and

$$\boldsymbol{R}_k(\epsilon) = \frac{1}{(k-1)!} \int_0^1 (1-\tau)^{k-1} d^k \boldsymbol{\xi}(\tau\epsilon)\epsilon^k \, d\tau. \tag{A.3}$$

Note that

$$\|\boldsymbol{R}_k(\epsilon)\| \leqslant \frac{|\epsilon|^k}{k!} \max_{0 \leqslant t \leqslant \epsilon} \left\| d^k \boldsymbol{\xi}(t) \right\|, \tag{A.4}$$

so that, if we choose $\hat{\epsilon} \in (0, \epsilon_0)$ and define

$$M = \frac{1}{k!} \max_{-\hat{\epsilon} \leqslant t \leqslant \hat{\epsilon}} \left\| d^k \boldsymbol{\xi}(t) \right\|, \tag{A.5}$$

then, for ϵ satisfying $|\epsilon| < \hat{\epsilon}$,

$$\|\boldsymbol{R}_k(\epsilon)\| \leqslant M|\epsilon|^k. \tag{A.6}$$

This can be rewritten as

$$\|\boldsymbol{R}_k(\epsilon)\| = O(\epsilon^k), \quad \text{as } \epsilon \to 0. \tag{A.7}$$

The above version of Taylor's theorem is presented in Zeidler [53, sect. 3.2], for example.

Inverse mapping theorem

Theorem A.3 *Let Y, Z be Banach spaces, U an open set in Y and $\boldsymbol{\psi} \in C^k(U, Z)$, with $k \geqslant 1$. Let $\boldsymbol{y}_0 \in U$ and suppose the partial Fréchet derivative $\boldsymbol{\psi_y}(\boldsymbol{y}_0) : Y \to Z$ is bijective. Then there exists an open neighborhood $U_0 \times V_0$ of $(\boldsymbol{y}_0, \boldsymbol{\psi}(\boldsymbol{y}_0))$ in $U \times Z$ such that $\boldsymbol{\psi}$ defines a C^k-diffeomorphism between U_0 and V_0.*

This result is available in Zeidler [53, sect. 4.13] or in Abraham *et al.* [1], for example.

Implicit function theorem

Theorem A.4 *Let X, Λ, Y be Banach spaces, $U \times \Omega$ an open set in $X \times \Lambda$ and $\boldsymbol{\Phi} \in C^k(U \times \Omega, Y)$, with $k \geqslant 1$. Suppose $(\boldsymbol{x}_0, \boldsymbol{\lambda}_0) \in U \times \Omega$ is such that $\boldsymbol{\Phi}(\boldsymbol{x}_0, \boldsymbol{\lambda}_0) = \boldsymbol{0}$ and the partial Fréchet derivative $\boldsymbol{\Phi_x}(\boldsymbol{x}_0, \boldsymbol{\lambda}_0) : X \to Y$ is bijective. Then there exist a neighborhood $U_0 \times \Omega_0$ of $(\boldsymbol{x}_0, \boldsymbol{\lambda}_0)$ in $U \times \Omega$ and a function $\boldsymbol{\xi} \in C^k(\Omega_0, U_0)$ such that $\boldsymbol{\xi}(\boldsymbol{\lambda}_0) = \boldsymbol{x}_0$ and, for any $\boldsymbol{\lambda} \in \Omega_0$, $\boldsymbol{x} = \boldsymbol{\xi}(\boldsymbol{\lambda})$ is the unique solution to $\boldsymbol{\Phi}(\boldsymbol{x}, \boldsymbol{\lambda}) = \boldsymbol{0}$ in U_0.*

This theorem can be found in Zeidler [53, sect. 4.7] or in Abraham *et al.* [1], for example.

Crandall-Rabinowitz theorem

Theorem A.5 *Let X, Y be Banach spaces, $I \times U$ an open set containing $(\theta^*, \boldsymbol{0})$ in $\mathbb{R} \times X$, and $\boldsymbol{\Phi} \in C^k(I \times U, Y)$, with $k \geqslant 2$. Suppose $\boldsymbol{\Phi}(\theta, \boldsymbol{0}) = \boldsymbol{0}$ for any $\theta \in I$, and the operator $\boldsymbol{A} = \boldsymbol{\Phi_x}(\theta^*, \boldsymbol{0})$ is such that*

(i) Ker $\boldsymbol{A}$ = Span$\{\boldsymbol{u}^*\}$, with $\boldsymbol{u}^* \in X \backslash \{\boldsymbol{0}\}$,

(ii) Range $\boldsymbol{A}$ is closed and codim Range $\boldsymbol{A} = 1$,

(iii) $\boldsymbol{\Phi}_{\boldsymbol{x}\theta}(\theta^*, \boldsymbol{0})\boldsymbol{u}^* \notin$ Range $\boldsymbol{A}$.

If Z is any complement of Ker $\boldsymbol{A}$ in X, then there exist a neighborhood $I_0 \times U_0$ of $(\theta^*, \boldsymbol{0})$, a positive number ϵ_0, and functions

$$\phi \in C^{k-1}((-\epsilon_0, \epsilon_0), I_0), \quad \boldsymbol{\eta} \in C^{k-1}((-\epsilon_0, \epsilon_0), Z) \tag{A.8}$$

such that

$$\phi(0) = \theta^*, \quad \boldsymbol{\eta}(0) = \boldsymbol{0}, \quad \boldsymbol{\eta}_\epsilon(0) = \boldsymbol{0}, \tag{A.9}$$

and the set of solutions to $\boldsymbol{\Phi}(\theta, \boldsymbol{x}) = \boldsymbol{0}$ in $I_0 \times U_0$ is

$$\big\{(\theta, \boldsymbol{0}) : \theta \in I_0\big\} \cup \big\{\big(\phi(\epsilon), \epsilon\boldsymbol{u}^* + \boldsymbol{\eta}(\epsilon)\big) : |\epsilon| < \epsilon_0\big\}. \tag{A.10}$$

The first set in (A.10) is called a *trivial branch* of solutions, and the second set is called a *nontrivial branch* of solutions. The solution $(\theta^*, \mathbf{0})$, at the intersection of the two branches, is called a *bifurcation point.*

This theorem on bifurcation from a simple eigenvalue is due to Crandall and Rabinowitz [15]. Excellent introductions to this theorem can be found in Ambrosetti and Prodi [5, sect. 5.4, thm. 4.1] or in Buffoni and Toland [10].

Submersion theorem

Before stating this theorem of differential geometry, we give two definitions. Here $k \geqslant 1$.

Definition A.6 *Let $\boldsymbol{f} \in C^k(\mathbb{R}^m, \mathbb{R}^n)$. A point $\boldsymbol{y} \in \mathbb{R}^n$ is called a* regular value *of $\boldsymbol{f}$ if, for each $\boldsymbol{x}$ satisfying $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{y}$, the derivative $\boldsymbol{f}_{\boldsymbol{x}}(\boldsymbol{x})$ is surjective.*

Definition A.7 *A subset L of $\mathbb{R}^m$ is called a C^k-*submanifold of dimension *d if, for any $\boldsymbol{x} \in L$, there exists an open neighborhood U of $\boldsymbol{x}$ in $\mathbb{R}^m$, an open set $V \times Z \subset \mathbb{R}^d \times \mathbb{R}^{m-d}$ and a C^k-diffeomorphism $\boldsymbol{\chi} : U \to V \times Z$ such that $\boldsymbol{\chi}(L \cap U) = V \times \{\mathbf{0}\}$.*

Theorem A.8 *Let $\boldsymbol{f} \in C^k(\mathbb{R}^m, \mathbb{R}^n)$ and suppose $\boldsymbol{y} \in \mathbb{R}^n$ is a regular value of $\boldsymbol{f}$. Then the set*

$$L = \{\boldsymbol{x} \in \mathbb{R}^m : \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{y}\} \tag{A.11}$$

is a C^k-submanifold of dimension $m - n$ with tangent space given by $T_{\boldsymbol{x}}L = Ker\, \boldsymbol{f}_{\boldsymbol{x}}(\boldsymbol{x})$.

This theorem is presented in Abraham *et al.* [2] or in Berger and Gostiaux [7, thm. 2.1.2], for example.

Appendix B

Variational formulation

In a first part of this appendix, we consider a variational problem sometimes called the *modified Hamilton's principle* (cf. Goldstein [22, sect. 8-5]) with a certain type of imposed two-point boundary conditions. We prove that the associated Euler-Lagrange equation corresponds to a Hamiltonian BVP with the property that the boundary conditions at each end-point define a Lagrangian manifold. In a second part, we exhibit a variational formulation of the different rod BVPs composed of the equilibrium equations together with each of the sets of boundary conditions introduced in sections 5.2 and 6.6. We conclude that these sets of boundary conditions define Lagrangian manifolds.

General results

Consider the set Z of $(\boldsymbol{q}, \boldsymbol{p}) \in \mathcal{C}^1 \equiv C^1([0,1], \mathbb{R}^n \times \mathbb{R}^n)$ satisfying the separated boundary conditions

$$\boldsymbol{f}^\alpha(\boldsymbol{q}(\alpha)) = \boldsymbol{0}, \quad \alpha = 0, 1, \tag{B.1}$$

and the functional $\mathcal{J} : Z \to \mathbb{R}$ defined by

$$\mathcal{J}(\boldsymbol{q}, \boldsymbol{p}) = d^0(\boldsymbol{q}(0)) - d^1(\boldsymbol{q}(1)) + \int_0^1 \Big(\boldsymbol{p}(s)^T \boldsymbol{q}'(s) - H(\boldsymbol{q}(s), \boldsymbol{p}(s), s)\Big)\, ds. \tag{B.2}$$

We are interested in the variational problem of finding solutions $(\boldsymbol{q}, \boldsymbol{p}) \in Z$ which make $\mathcal{J}$ stationary.

In what follows it is implicitly meant that every statement containing the index α is true for $\alpha = 0, 1$. Assume

$$\begin{gathered} \boldsymbol{f}^\alpha \in C^2(\mathbb{R}^n, \mathbb{R}^{l^\alpha}), \quad \text{with } 0 \leqslant l^\alpha \leqslant n, \\ d^\alpha \in C^2(\mathbb{R}^n, \mathbb{R}), \quad H \in C^2(\mathbb{R}^{2n} \times [0,1], \mathbb{R}). \end{gathered} \tag{B.3}$$

Let

$$M^\alpha = \big\{\boldsymbol{\rho} \in \mathbb{R}^n : \boldsymbol{f}^\alpha(\boldsymbol{\rho}) = \boldsymbol{0}\big\}. \tag{B.4}$$

Suppose $\boldsymbol{0} \in \mathbb{R}^{l^\alpha}$ is a regular value of $\boldsymbol{f}^\alpha$, i.e., for any $\boldsymbol{\rho} \in M^\alpha$, the derivative $\boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})$ is surjective. Then, by the submersion theorem (cf. appendix A), M^α is a C^2-submanifold of dimension $n - l^\alpha$ in $\mathbb{R}^n$, with tangent space given by $T_{\boldsymbol{\rho}} M^\alpha = \operatorname{Ker} \boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})$. Assume moreover that there exist functions $\boldsymbol{g}^\alpha_1, \ldots, \boldsymbol{g}^\alpha_{n-l^\alpha} \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ such that

$$\operatorname{Span}\{\boldsymbol{g}^\alpha_1(\boldsymbol{\rho}), \ldots, \boldsymbol{g}^\alpha_{n-l^\alpha}(\boldsymbol{\rho})\} = T_{\boldsymbol{\rho}} M^\alpha = \operatorname{Ker} \boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho}), \quad \forall \boldsymbol{\rho} \in M^\alpha. \tag{B.5}$$

Proposition B.1 *Suppose the above assumptions are satisfied. Then $(\boldsymbol{q}, \boldsymbol{p}) \in \mathcal{C}^1$ makes $\mathcal{J}$ stationary if and only if it satisfies the Hamiltonian BVP*

$$\begin{cases} \boldsymbol{q}'(s) = H_{\boldsymbol{p}}(\boldsymbol{q}(s), \boldsymbol{p}(s), s), \quad \boldsymbol{p}'(s) = -H_{\boldsymbol{q}}(\boldsymbol{q}(s), \boldsymbol{p}(s), s), \\ \boldsymbol{b}^0(\boldsymbol{q}(0), \boldsymbol{p}(0)) = \boldsymbol{0}, \quad \boldsymbol{b}^1(\boldsymbol{q}(1), \boldsymbol{p}(1)) = \boldsymbol{0}, \end{cases} \tag{B.6}$$

where $\boldsymbol{b}^\alpha : \mathbb{R}^{2n} \to \mathbb{R}^n$ is defined by

$$\boldsymbol{b}^\alpha(\boldsymbol{\rho}, \boldsymbol{\pi}) = \begin{pmatrix} \boldsymbol{f}^\alpha(\boldsymbol{\rho}) \\ (\boldsymbol{\pi} - d^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho}))^T \boldsymbol{g}^\alpha_1(\boldsymbol{\rho}) \\ \vdots \\ (\boldsymbol{\pi} - d^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho}))^T \boldsymbol{g}^\alpha_{n-l^\alpha}(\boldsymbol{\rho}) \end{pmatrix}. \tag{B.7}$$

Moreover, defining the set

$$L^\alpha = \big\{(\boldsymbol{\rho}, \boldsymbol{\pi}) \in \mathbb{R}^{2n} : \boldsymbol{b}^\alpha(\boldsymbol{\rho}, \boldsymbol{\pi}) = \boldsymbol{0}\big\}, \tag{B.8}$$

we have, for any $(\boldsymbol{\rho}, \boldsymbol{\pi}) \in L^\alpha$,

$$(\tilde{\boldsymbol{\rho}}_1, \tilde{\boldsymbol{\pi}}_1), (\tilde{\boldsymbol{\rho}}_2, \tilde{\boldsymbol{\pi}}_2) \in \mathit{Ker}\, \boldsymbol{b}^\alpha_{(\boldsymbol{\rho},\boldsymbol{\pi})}(\boldsymbol{\rho}, \boldsymbol{\pi}) \Longrightarrow \tilde{\boldsymbol{\pi}}_1^T \tilde{\boldsymbol{\rho}}_2 - \tilde{\boldsymbol{\pi}}_2^T \tilde{\boldsymbol{\rho}}_1 = 0, \tag{B.9}$$

which implies L^α is a Lagrangian submanifold of $\mathbb{R}^{2n}$ (cf. definition 2.4).

Note that, at $s = \alpha$, the l^α scalar boundary conditions $\boldsymbol{f}^\alpha(\boldsymbol{q}(\alpha)) = \boldsymbol{0}$ *imposed* in the variational problem have been completed by $n - l^\alpha$ scalar *natural* boundary conditions $(\boldsymbol{p}(\alpha) - d^\alpha_{\boldsymbol{\rho}}(\boldsymbol{q}(\alpha)))^T \boldsymbol{g}^\alpha_i(\boldsymbol{q}(\alpha)) = 0$, so that Hamiltonian BVP (B.6) possesses n scalar boundary conditions both at $s = 0$ and at $s = 1$. Note also that property (B.9) is identical to property (**2.A5**). Therefore, if Hamiltonian (2.1) is equivalent to a variational problem of the form considered above, then property (**2.A5**) automatically holds.

Proof By definition of Z, a variation $(\boldsymbol{u}, \boldsymbol{v}) \in \mathcal{C}^1$ of $(\boldsymbol{q}, \boldsymbol{p}) \in Z$ is admissible if and only if

$$\boldsymbol{u}(\alpha) \in \operatorname{Ker} \boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{q}(\alpha)). \tag{B.10}$$

The functional $\mathcal{J}$ is stationary if and only if the first variation $\delta\mathcal{J}(\boldsymbol{u}, \boldsymbol{v})$ is zero for any admissible $(\boldsymbol{u}, \boldsymbol{v})$. We compute

$$\begin{aligned} \delta\mathcal{J}(\boldsymbol{u}, \boldsymbol{v}) &= d^0_{\boldsymbol{\rho}}(\boldsymbol{q}(0))^T \boldsymbol{u}(0) - d^1_{\boldsymbol{\rho}}(\boldsymbol{q}(1))^T \boldsymbol{u}(1) + \int_0^1 \boldsymbol{p}(s)^T \boldsymbol{u}'(s)\, ds \\ &\quad + \int_0^1 \Big(\boldsymbol{v}(s)^T \boldsymbol{q}'(s) - H_{\boldsymbol{q}}(\boldsymbol{q}(s), \boldsymbol{p}(s), s)^T \boldsymbol{u}(s) - H_{\boldsymbol{p}}(\boldsymbol{q}(s), \boldsymbol{p}(s), s)^T \boldsymbol{v}(s)\Big)\, ds \\ &= \big(\boldsymbol{p}(1) - d^1_{\boldsymbol{\rho}}(\boldsymbol{q}(1))\big)^T \boldsymbol{u}(1) - \big(\boldsymbol{p}(0) - d^0_{\boldsymbol{\rho}}(\boldsymbol{q}(1))\big)^T \boldsymbol{u}(0) \\ &\quad + \int_0^1 \Big(\boldsymbol{q}'(s) - H_{\boldsymbol{p}}(\boldsymbol{q}(s), \boldsymbol{p}(s), s)\Big)^T \boldsymbol{v}(s)\, ds \\ &\quad - \int_0^1 \Big(\boldsymbol{p}'(s) + H_{\boldsymbol{q}}(\boldsymbol{q}(s), \boldsymbol{p}(s), s)\Big)^T \boldsymbol{u}(s)\, ds, \end{aligned} \tag{B.11}$$

where the second equality follows from an integration by parts of the integral on the first line. We conclude that $(\boldsymbol{q},\boldsymbol{p}) \in Z$ makes $\mathcal{J}$ stationary if and only if

$$\boldsymbol{q}'(s) = H_{\boldsymbol{p}}(\boldsymbol{q}(s),\boldsymbol{p}(s),s), \quad \boldsymbol{p}'(s) = -H_{\boldsymbol{q}}(\boldsymbol{q}(s),\boldsymbol{p}(s),s) \tag{B.12}$$

and

$$\big(\boldsymbol{p}(\alpha) - d^\alpha_{\boldsymbol{\rho}}(\boldsymbol{q}(\alpha))\big) \in \big(\operatorname{Ker} \boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{q}(\alpha))\big)^\perp. \tag{B.13}$$

This last equation is equivalent to

$$\big(\boldsymbol{p}(\alpha) - d^\alpha_{\boldsymbol{\rho}}(\boldsymbol{q}(\alpha))\big)^T \boldsymbol{g}^\alpha_i(\boldsymbol{q}(\alpha)) = 0, \quad \forall\, i = 1,\dots,n-l^\alpha, \tag{B.14}$$

and the first part of the proposition is established.

We continue with the proof of property (B.9). We suppose $(\boldsymbol{\rho},\boldsymbol{\pi}) \in L^\alpha$ and

$$(\tilde{\boldsymbol{\rho}}_1,\tilde{\boldsymbol{\pi}}_1),(\tilde{\boldsymbol{\rho}}_2,\tilde{\boldsymbol{\pi}}_2) \in \operatorname{Ker} \boldsymbol{b}^\alpha_{(\boldsymbol{\rho},\boldsymbol{\pi})}(\boldsymbol{\rho},\boldsymbol{\pi}), \tag{B.15}$$

so that

$$\boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})\tilde{\boldsymbol{\rho}}_1 = \mathbf{0}, \quad \boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})\tilde{\boldsymbol{\rho}}_2 = \mathbf{0} \tag{B.16}$$

and, $\forall\, i = 1,\dots,n-l^\alpha$,

$$-\tilde{\boldsymbol{\rho}}_1^T d^\alpha_{\boldsymbol{\rho}\boldsymbol{\rho}}(\boldsymbol{\rho})\boldsymbol{g}^\alpha_i(\boldsymbol{\rho}) + \big(\boldsymbol{\pi} - d^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})\big)^T \boldsymbol{g}^\alpha_{i,\boldsymbol{\rho}}(\boldsymbol{\rho})\tilde{\boldsymbol{\rho}}_1 + \tilde{\boldsymbol{\pi}}_1^T \boldsymbol{g}^\alpha_i(\boldsymbol{\rho}) = 0. \tag{B.17}$$

By (B.16), $\tilde{\boldsymbol{\rho}}_2 \in \operatorname{Ker} \boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})$, so that there exist coefficients $\zeta_1,\dots,\zeta_{n-l^\alpha}$ such that

$$\tilde{\boldsymbol{\rho}}_2 = \sum_{i=1}^{n-l^\alpha} \zeta_i\, \boldsymbol{g}^\alpha_i(\boldsymbol{\rho}). \tag{B.18}$$

For $i = 1,\dots,n-l^\alpha$, we have

$$\boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})\, \boldsymbol{g}^\alpha_i(\boldsymbol{\rho}) = \mathbf{0}, \quad \forall\, \boldsymbol{\rho} \in M^\alpha, \tag{B.19}$$

and, by (B.16),

$$\tilde{\boldsymbol{\rho}}_1 \in \operatorname{Ker} \boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho}) = T_{\boldsymbol{\rho}} M^\alpha. \tag{B.20}$$

It follows that

$$\boldsymbol{f}^\alpha_{\boldsymbol{\rho}\boldsymbol{\rho}}(\boldsymbol{\rho})[\boldsymbol{g}^\alpha_i(\boldsymbol{\rho}),\tilde{\boldsymbol{\rho}}_1] + \boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})\, \boldsymbol{g}^\alpha_{i,\boldsymbol{\rho}}(\boldsymbol{\rho})\tilde{\boldsymbol{\rho}}_1 = \mathbf{0}. \tag{B.21}$$

By surjectivity of $\boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})$, the restriction $\boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})\Big|_{(\operatorname{Ker}\, \boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho}))^\perp}$ is bijective, and we can define the mapping

$$\boldsymbol{F}^\alpha(\boldsymbol{\rho}) = \left(\boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})\Big|_{(\operatorname{Ker}\, \boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho}))^\perp}\right)^{-1} : \mathbb{R}^{l^\alpha} \to \big(\operatorname{Ker} \boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})\big)^\perp. \tag{B.22}$$

Then, by (B.21),

$$\boldsymbol{g}^\alpha_{i,\boldsymbol{\rho}}(\boldsymbol{\rho})\tilde{\boldsymbol{\rho}}_1 = \boldsymbol{F}^\alpha(\boldsymbol{\rho})\big(\boldsymbol{f}^\alpha_{\boldsymbol{\rho}\boldsymbol{\rho}}(\boldsymbol{\rho})[\boldsymbol{g}^\alpha_i(\boldsymbol{\rho}),\tilde{\boldsymbol{\rho}}_1]\big) + \boldsymbol{k}_i, \tag{B.23}$$

with $\boldsymbol{k}_i \in \operatorname{Ker} \boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})$. We now compute

$$\begin{aligned}
\tilde{\boldsymbol{\pi}}_1^T \tilde{\boldsymbol{\rho}}_2 &\overset{(B.18)}{=} \sum_{i=1}^{n-l^\alpha} \zeta_i \Big(\tilde{\boldsymbol{\pi}}_1^T \boldsymbol{g}_i^\alpha(\boldsymbol{\rho}) \Big) \\
&\overset{(B.17)}{=} \sum_{i=1}^{n-l^\alpha} \zeta_i \Big(\tilde{\boldsymbol{\rho}}_1^T d^\alpha_{\boldsymbol{\rho\rho}}(\boldsymbol{\rho}) \boldsymbol{g}_i^\alpha(\boldsymbol{\rho}) - \big(\boldsymbol{\pi} - d^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})\big)^T \boldsymbol{g}^\alpha_{i,\boldsymbol{\rho}}(\boldsymbol{\rho}) \tilde{\boldsymbol{\rho}}_1 \Big) \\
&\overset{(B.23)}{=} \sum_{i=1}^{n-l^\alpha} \zeta_i \Big(\tilde{\boldsymbol{\rho}}_1^T d^\alpha_{\boldsymbol{\rho\rho}}(\boldsymbol{\rho}) \boldsymbol{g}_i^\alpha(\boldsymbol{\rho}) - \big(\boldsymbol{\pi} - d^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})\big)^T \big(\boldsymbol{F}^\alpha(\boldsymbol{\rho}) \big(\boldsymbol{f}^\alpha_{\boldsymbol{\rho\rho}}(\boldsymbol{\rho})[\boldsymbol{g}_i^\alpha(\boldsymbol{\rho}), \tilde{\boldsymbol{\rho}}_1] \big) \big) \\
&\qquad\qquad - \big(\boldsymbol{\pi} - d^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})\big)^T \boldsymbol{k}_i \Big) \\
&\overset{(B.18)}{=} \tilde{\boldsymbol{\rho}}_1^T d^\alpha_{\boldsymbol{\rho\rho}}(\boldsymbol{\rho}) \tilde{\boldsymbol{\rho}}_2 - \big(\boldsymbol{\pi} - d^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})\big)^T \big(\boldsymbol{F}^\alpha(\boldsymbol{\rho}) \big(\boldsymbol{f}^\alpha_{\boldsymbol{\rho\rho}}(\boldsymbol{\rho})[\tilde{\boldsymbol{\rho}}_2, \tilde{\boldsymbol{\rho}}_1] \big) \big).
\end{aligned} \tag{B.24}$$

The number of the equation used to obtain each equality is indicated above the corresponding equality sign. For the last equality we also used the fact that $(\boldsymbol{\pi} - d^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho}))^T \boldsymbol{k}_i = 0$, since $(\boldsymbol{\pi} - d^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})) \in (\operatorname{Ker} \boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho}))^\perp$ and $\boldsymbol{k}_i \in \operatorname{Ker} \boldsymbol{f}^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})$. In the same way, we find

$$\tilde{\boldsymbol{\pi}}_2^T \tilde{\boldsymbol{\rho}}_1 = \tilde{\boldsymbol{\rho}}_2^T d^\alpha_{\boldsymbol{\rho\rho}}(\boldsymbol{\rho}) \tilde{\boldsymbol{\rho}}_1 - \big(\boldsymbol{\pi} - d^\alpha_{\boldsymbol{\rho}}(\boldsymbol{\rho})\big)^T \big(\boldsymbol{F}^\alpha(\boldsymbol{\rho}) \big(\boldsymbol{f}^\alpha_{\boldsymbol{\rho\rho}}(\boldsymbol{\rho})[\tilde{\boldsymbol{\rho}}_1, \tilde{\boldsymbol{\rho}}_2] \big) \big). \tag{B.25}$$

By symmetry of $d^\alpha_{\boldsymbol{\rho\rho}}(\boldsymbol{\rho})$ and $\boldsymbol{f}^\alpha_{\boldsymbol{\rho\rho}}(\boldsymbol{\rho})$, expressions (B.24) and (B.25) are equal, so that property (B.9) is verified. Finally, property (B.9) is the same as condition (**2.A5**), so that the fact that L^α is a Lagrangian submanifold is a consequence of lemma 2.5. ■

Rod examples

We now present a variational formulation of the different rod BVPs composed of equilibrium equation (5.7) together with each of the sets of boundary conditions introduced in sections 5.2 and 6.6. The two conjugate variables are $\boldsymbol{q} = (\boldsymbol{r}, \boldsymbol{\gamma})$ and $\boldsymbol{p} = (\boldsymbol{n}, \boldsymbol{\mu})$.

Clamped-free strut BVP

Consider the variational problem defined by functional (B.2), with

$$d^0(\boldsymbol{r}(0), \boldsymbol{\gamma}(0)) = 0, \quad d^1(\boldsymbol{r}(1), \boldsymbol{\gamma}(1)) = -\lambda r_3(1), \tag{B.26}$$

and subject to the imposed boundary conditions

$$\boldsymbol{f}^0(\boldsymbol{r}(0), \boldsymbol{\gamma}(0)) = \begin{pmatrix} \boldsymbol{r}(0) \\ \gamma_1(0) \\ \gamma_2(0) \\ \gamma_3(0) \\ \gamma_4(0) - 1 \end{pmatrix} = \boldsymbol{0}. \tag{B.27}$$

It follows from proposition B.1 that this variational problem is equivalent to the clamped-free strut Hamiltonian BVP (5.7)+(5.17), and that the boundary conditions (5.17) satisfy property (**2.A5**) and therefore define Lagrangian submanifolds of $\mathbb{R}^{14}$ at $s = 0$ and at $s = 1$.

Clamped-sleeved strut BVP

Consider the variational problem defined by functional (B.2), with

$$d^0(\boldsymbol{r}(0), \boldsymbol{\gamma}(0)) = 0, \quad d^1(\boldsymbol{r}(1), \boldsymbol{\gamma}(1)) = -\lambda r_3(1), \tag{B.28}$$

and subject to the imposed boundary conditions

$$\boldsymbol{f}^0(\boldsymbol{r}(0), \boldsymbol{\gamma}(0)) = \begin{pmatrix} \boldsymbol{r}(0) \\ \gamma_1(0) \\ \gamma_2(0) \\ \gamma_3(0) \end{pmatrix} = \mathbf{0}, \quad \boldsymbol{f}^1(\boldsymbol{r}(1), \boldsymbol{\gamma}(1)) = \begin{pmatrix} r_1(1) \\ r_2(1) \\ \gamma_1(1) \\ \gamma_2(1) \\ \gamma_3^2(1) + \gamma_4^2(1) - 1 \end{pmatrix} = \mathbf{0}. \tag{B.29}$$

It follows from proposition B.1 that this variational problem is equivalent to the clamped-sleeved strut Hamiltonian BVP (5.7)+(5.20), and that the boundary conditions (5.20) satisfy property (**2.A5**) and therefore define Lagrangian submanifolds of $\mathbb{R}^{14}$ at $s = 0$ and at $s = 1$.

Ring BVP

Consider the variational problem defined by functional (B.2), with

$$d^0(\boldsymbol{r}(0), \boldsymbol{\gamma}(0)) = 0, \quad d^1(\boldsymbol{r}(1), \boldsymbol{\gamma}(1)) = 0, \tag{B.30}$$

and subject to the imposed boundary conditions

$$\boldsymbol{f}^0(\boldsymbol{r}(0), \boldsymbol{\gamma}(0)) = \begin{pmatrix} \boldsymbol{r}(0) \\ \gamma_1(0) \\ \gamma_2(0) \\ \gamma_3(0) \end{pmatrix} = \mathbf{0}, \quad \boldsymbol{f}^1(\boldsymbol{r}(1), \boldsymbol{\gamma}(1)) = \begin{pmatrix} \boldsymbol{r}(1) \\ \gamma_1(1) \\ \gamma_2(1) \\ \gamma_3 - \sin\frac{\alpha}{2} \\ \gamma_4 - \cos\frac{\alpha}{2} \end{pmatrix} = \mathbf{0}. \tag{B.31}$$

It follows from proposition B.1 that this variational problem is equivalent to the ring Hamiltonian BVP (5.7)+(5.23), and that the boundary conditions (5.23) satisfy property (**2.A5**) and therefore define Lagrangian submanifolds of $\mathbb{R}^{14}$ at $s = 0$ and at $s = 1$.

Modified ring BVP

Consider the variational problem defined by functional (B.2), with

$$d^0(\boldsymbol{r}(0), \boldsymbol{\gamma}(0)) = 0, \quad d^1(\boldsymbol{r}(1), \boldsymbol{\gamma}(1)) = 0, \tag{B.32}$$

and subject to the imposed boundary conditions

$$\boldsymbol{f}^0(\boldsymbol{r}(0), \boldsymbol{\gamma}(0)) = \begin{pmatrix} \boldsymbol{r}(0) \\ \gamma_2(0) \\ \gamma_3(0) \end{pmatrix} = \mathbf{0}, \quad \boldsymbol{f}^1(\boldsymbol{r}(1), \boldsymbol{\gamma}(1)) = \begin{pmatrix} \boldsymbol{r}(1) \\ \gamma_1(1) \\ \gamma_2(1) \\ \gamma_3 - \sin\frac{\alpha}{2} \\ \gamma_4 - \cos\frac{\alpha}{2} \end{pmatrix} = \mathbf{0}. \tag{B.33}$$

It follows from proposition B.1 that this variational problem is equivalent to the modified ring Hamiltonian BVP (5.7)+(6.68), and that the boundary conditions (6.68) satisfy property (**2.A5**) and therefore define Lagrangian submanifolds of $\mathbb{R}^{14}$ at $s = 0$ and at $s = 1$.

Appendix C

Planar equilibria of isotropic, uniform struts

According to section 4.6, the problem of finding the planar equilibria of an isotropic, uniform clamped-free or clamped-sleeved strut reduces to solving a second-order differential equation (with appropriate boundary conditions) for a single scalar variable. This equation is classic and has been extensively studied (cf. Love [32, chap. XIX], Antman [6, sect. IV.2] and references therein). The present appendix discusses the solutions to this reduced equation under the boundary conditions associated to the clamped-free strut and to the clamped-sleeved strut.

Clamped-free strut

The planar equilibria of an isotropic, uniform clamped-free strut are governed by equation (cf. section 4.6)

$$\phi''(s) = -\frac{\lambda}{K_1}\sin\phi(s), \quad \phi(0) = 0, \quad \phi'(1) = 0. \tag{C.1}$$

Although the solutions can be explicitly written in term of elliptic functions (cf. Reiss [43], for example), we only present a qualitative study of them here, following Chow and Hale [14, sect. 1.3].

If $\phi(s)$ is a solution to (C.1), then it satisfies the first integral equation

$$\phi'^2(s) - \frac{2\lambda}{K_1}\cos\phi(s) = -\frac{2\lambda}{K_1}\cos\phi_1, \tag{C.2}$$

with $\phi_1 = \phi(1)$. This equality provides the phase-plane diagram shown in figure C.1. Solutions to (C.1) correspond to trajectories in the phase-plane that begin at $s = 0$ on the ϕ'-axis and end at $s = 1$ on the ϕ-axis. The origin of the diagram corresponds to the *trivial* solution $\phi(s) = 0$. The other solutions, called *nontrivial*, correspond to periodic trajectories starting on the ϕ'-axis and encircling the origin $(\frac{1}{4} + \frac{1}{2}j)$-times, with $j \in \mathbb{N}$, before returning to the ϕ-axis. Finding these solutions reduces to determining which of the periodic orbits take exactly 1 unit of "time" s to complete $(\frac{1}{4} + \frac{1}{2}j)$ turns about the

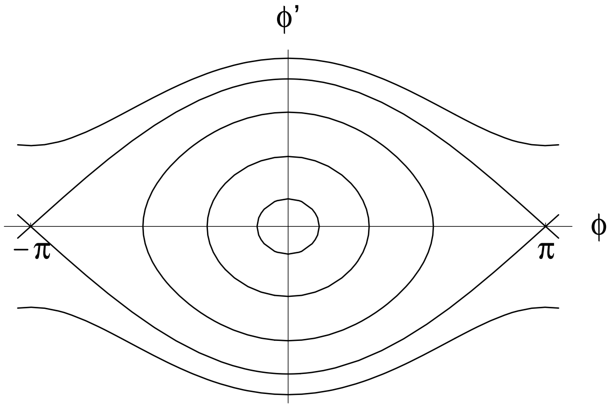


Figure C.1: *Phase-plane diagram. The diagram is symmetric about both axes, and repeats with period 2π along the ϕ-axis. The origin is a center, and the saddles $(\pm\pi, 0)$ are connected by heteroclinic orbits.*

origin. From (C.2), we have

$$\frac{d\phi}{ds} = \pm 2\sqrt{\frac{\lambda}{K_1}}\sqrt{\sin^2\frac{\phi_1}{2} - \sin^2\frac{\phi}{2}}, \tag{C.3}$$

and the condition can be written as

$$1 = \int_0^1 ds = \left(\frac{1}{2} + j\right)\sqrt{\frac{K_1}{\lambda}}\int_0^{|\phi_1|}\frac{d\phi}{\sqrt{\sin^2\frac{|\phi_1|}{2} - \sin^2\frac{\phi}{2}}}, \tag{C.4}$$

with $0 < |\phi_1| < \pi$. From the change of variable $\sin\frac{\phi}{2} = \sin\frac{|\phi_1|}{2}\sin\psi$, the condition becomes

$$\frac{\lambda}{K_1} = (1+2j)^2 K^2\left(\sin\tfrac{|\phi_1|}{2}\right), \tag{C.5}$$

where K denotes the complete elliptic integral of the first kind defined by (cf. Byrd and Friedman [11])

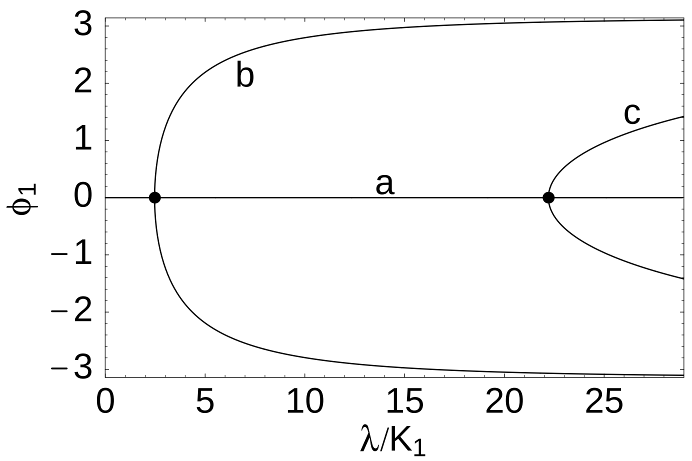
$$K(k) = \int_0^{\frac{\pi}{2}}\frac{d\nu}{\sqrt{1 - k^2\sin^2\nu}}. \tag{C.6}$$

This integral takes the value $\frac{\pi}{2}$ at $k = 0$, tends to ∞ as $k \to 1$, and is strictly increasing on the interval $[0, 1)$. Consequently, if $\lambda > \lambda_j$, where

$$\lambda_j = K_1\frac{(1+2j)^2\pi^2}{4}, \tag{C.7}$$

then equation (C.5) has two (symmetric) solutions for ϕ_1, and problem (C.1) has two (symmetric) solutions corresponding to $(\frac{1}{4} + \frac{1}{2}j)$ turns about the origin in phase plane.

The set of solutions to (C.1) (for λ in a limited interval) is represented by the bifurcation diagram in figure C.2. The trivial solution exists for all values of λ. Branches of nontrivial solutions bifurcate at $\lambda = \lambda_j$, for $j = 0, 1, 2, \ldots$, and the values λ_j are called *buckling loads*. The typical shapes of the planar centerlines $\boldsymbol{r}^0(s)$ (cf. equation (4.71)) of



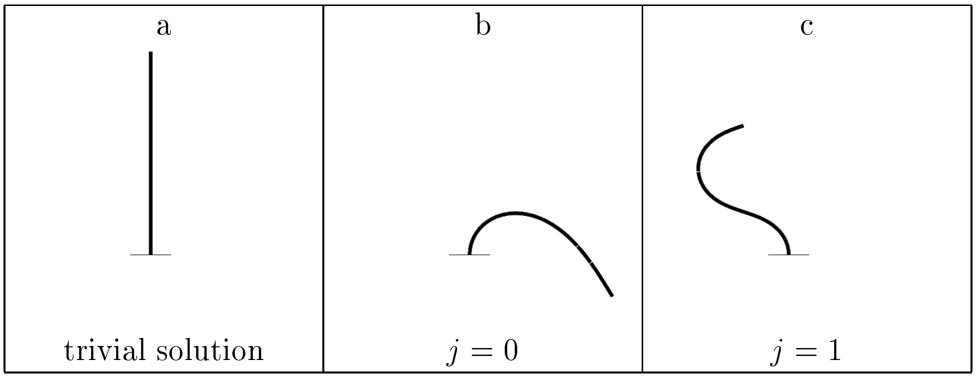


Figure C.2: Top: *Bifurcation diagram. The bifurcation points* $(\frac{\lambda_0}{K_1}, 0)$ *and* $(\frac{\lambda_1}{K_1}, 0)$ *are marked by dots.* Bottom: *Corresponding to branches a, b and c in bifurcation diagram, typical shapes of equilibria of the clamped-free strut.*

the rod associated to the different branches in the bifurcation diagram are also shown in figure C.2. The number of inflexion points of the centerline corresponds to the number of times the trajectory goes through the ϕ'-axis in phase-plane (excluding the starting point $s = 0$), and equals j.

Clamped-sleeved strut

The planar equilibria of an isotropic, uniform clamped-sleeved strut are governed by equation (cf. section 4.6)

$$\begin{cases} \phi''(s) = -\dfrac{\lambda}{K_1}\sin\phi(s) - \dfrac{n_1}{K_1}\cos\phi(s), \\ \phi(0) = 0, \quad \phi(1) = 0 \pmod{2\pi}, \quad \displaystyle\int_0^1 \sin\phi(s)\,ds = 0, \end{cases} \tag{C.8}$$

where $\phi(s)$ and n_1 are the unknowns, while λ and K_1 are parameters.

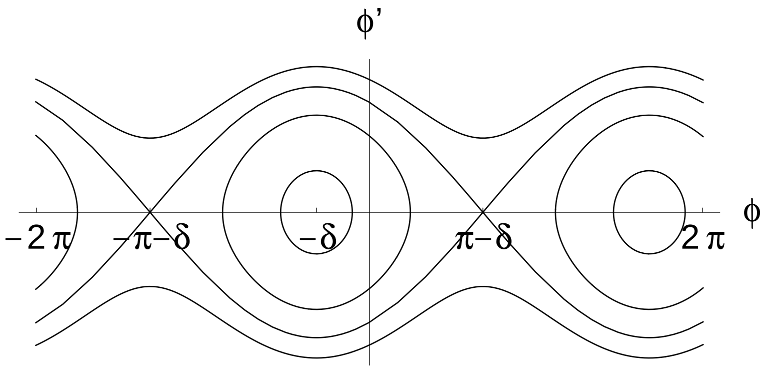


Figure C.3: *Phase-plane diagram. The diagram is symmetric about the ϕ-axis, and repeats with period 2π along the ϕ-axis. The point $(-\delta, 0)$ is a center, and the saddles $(\pm\pi - \delta, 0)$ are connected by heteroclinic orbits.*

If $\phi(s)$ is a solution to (C.8), then it satisfies the first integral equation

$$\phi'^2(s) - \frac{2\lambda}{K_1}\cos\phi(s) + \frac{2n_1}{K_1}\sin\phi(s) = \gamma, \tag{C.9}$$

where γ is a constant. Introducing the quantity

$$\delta = \begin{cases} \arctan\frac{n_1}{\lambda} & \text{if } \lambda \neq 0, \\ \frac{\pi}{2} & \text{if } \lambda = 0, \end{cases} \tag{C.10}$$

this first integral becomes

$$\phi'^2(s) - \frac{2\sqrt{\lambda^2 + n_1^2}}{K_1}\cos\big(\phi(s) + \delta\big) = \gamma. \tag{C.11}$$

A comparison with first integral (C.2) shows that the corresponding phase-plane diagrams are the same, except for a scaling along the ϕ'-direction and a shift by δ along the ϕ-direction (see figure C.3). Combining the boundary conditions in (C.8) with the above first integral, we obtain

$$\phi'(0) = \pm\phi'(1), \tag{C.12}$$

a property used in section 6.6.

Any solution to (C.8) corresponds to a trajectory in the phase-plane that begins at $s = 0$ on the ϕ'-axis, and ends at $s = 1$ on one of the lines $\phi = 2\pi l$, $l \in \mathbb{Z}$. Accordingly, problem (C.8) reduces to two scalar equations containing elliptic functions and relating n_1 and $\phi'(0)$ to λ. These equations express the condition that it takes exactly 1 unit of "time" s to go from the ϕ'-axis to one of the lines $\phi = 2\pi l$, $l \in \mathbb{Z}$, and the integral constraint $\int_0^1 \sin\phi(s)\, ds = 0$.

It is also possible to directly solve problem (C.8) numerically, as was done by Domokos [18], for example. Starting from special solutions, we perform a parameter continuation with the software AUTO (cf. Doedel *et al.* [17]) to obtain the portion of the bifurcation diagram shown in figure C.4. Other types of solutions can be found in

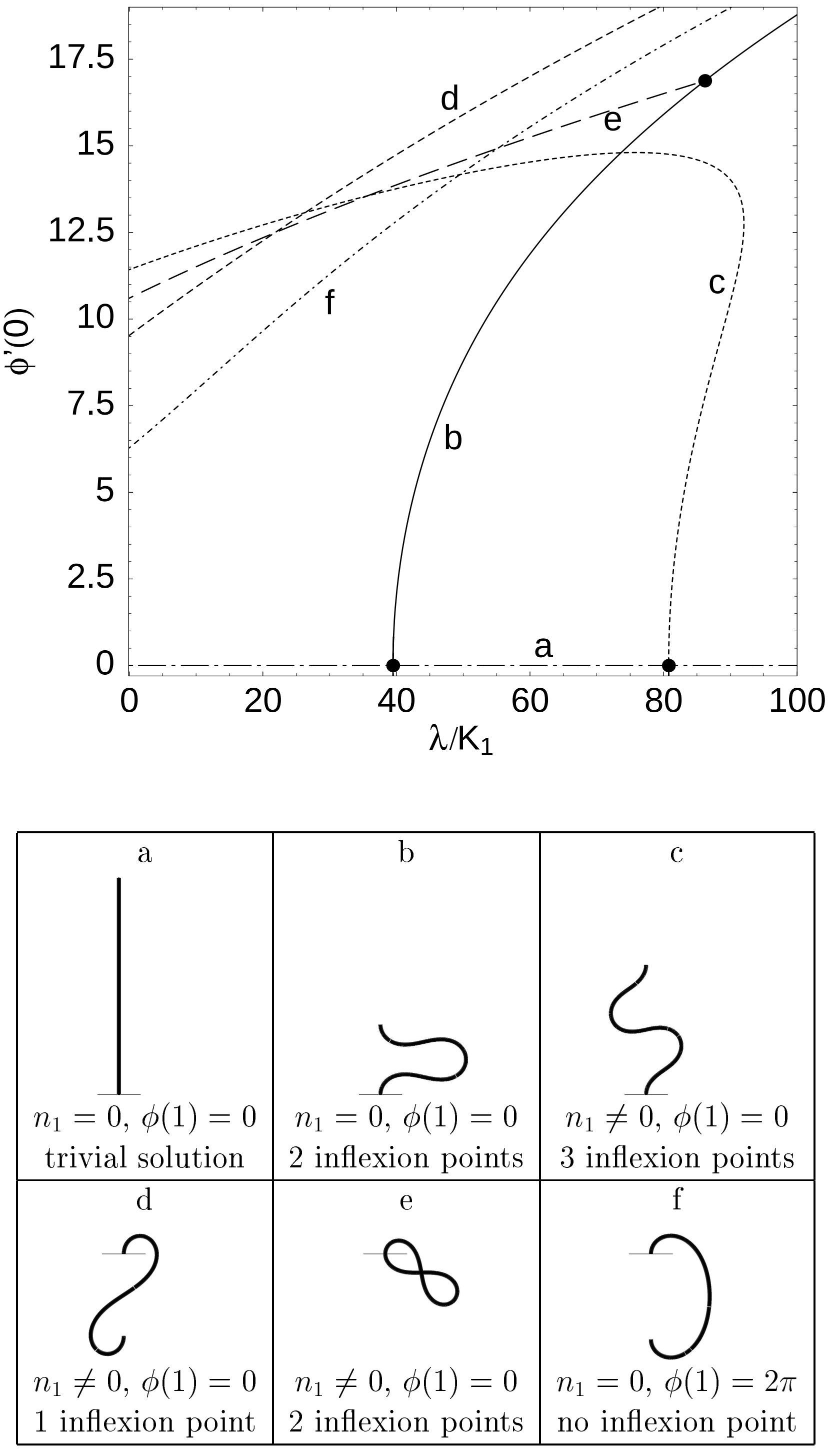


Figure C.4: Top: *Bifurcation diagram. Bifurcation points are marked by dots; all other apparent branch intersections are artifacts of the two-dimensional projection shown here. Branch e results from the superimposition of two branches with different values for* n_1. *The diagram is symmetric about the trivial branch a.* Bottom: *Typical shapes of planar equilibria of the clamped-sleeved strut, corresponding to the different branches in the bifurcation diagram. Each point on branch e corresponds to two closed, "8"-shaped equilibria with opposite values for* n_1.

van der Heijden *et al.* [52]. Note that the nontrivial solutions satisfying $\phi'(0) = 0$ are exceptional. The portion of the bifurcation diagram presented here does not contain any nontrivial solution with this property, but it is possible to find two such solutions by following branch e into the region of negative $\frac{\lambda}{K_1}$ (cf. Domokos [18, solutions 12 and 13 in figure 2]), for example.

Bibliography

[1] R. Abraham, J. E. Marsden, *Foundations of Mechanics*, second ed., Addison-Wesley, New York (1978)

[2] R. Abraham, J. E. Marsden, T. Ratiu, *Manifolds, Tensor Analysis, and Applications*, second ed., Springer-Verlag, New York (1988)

[3] A. Ambrosetti, M. Badiale, *Homoclinics: Poincaré-Melnikov type results via a variational approach*, Ann. Inst. Henri Poincaré **15** (1998) 233-252

[4] A. Ambrosetti, M. Badiale, *Variational perturbative methods and bifurcation of bound states from the essential spectrum*, Proc. Roy. Soc. Edinburgh **128A** (1998) 1131-1161

[5] A. Ambrosetti, G. Prodi, *A Primer of Nonlinear Analysis*, Cambridge University Press, Cambridge (1993)

[6] S. S. Antman, *Nonlinear Problems of Elasticity*, Springer-Verlag, New York (1995)

[7] M. Berger, B. Gostiaux, *Differential Geometry: Manifolds, Curves, and Surfaces*, Springer-Verlag, New York (1988)

[8] M. S. Berger, *Nonlinearity and Functional Analysis, Lectures on Nonlinear Problems in Mathematical Analysis*, Academic Press, New York (1977)

[9] H. Brezis, *Analyse fonctionnelle, théorie et applications*, Collection Mathématiques appliquées pour la maîtrise, Masson, Paris (1983)

[10] B. Buffoni, J. F. Toland, *Introduction à la théorie globale des bifurcations*, Presses polytechniques et universitaires romandes, Lausanne (2002)

[11] P. F. Byrd, M. D. Friedman, *Handbook of Elliptic Integrals for Engineers and Scientists*, second ed., Springer-Verlag, Berlin (1971)

[12] M. P. do Carmo, *Riemannian Geometry*, Birkhäuser, Boston (1992)

[13] P. Choquard, *Mécanique analytique*, Presses polytechniques et universitaires romandes, Lausanne (1992)

[14] S.-N. Chow, J. K. Hale, *Methods of Bifurcation Theory*, Springer-Verlag, New York (1982)

[15] M. G. Crandall, P. H. Rabinowitz, *Bifurcation from Simple Eigenvalues*, J. Funct. Anal. **8** (1971) 321-340

[16] D. J. Dichmann, Y. Li, J. H. Maddocks, *Hamiltonian formulations and symmetries in rod mechanics*, in *Mathematical Approaches to Biomolecular Structure and Dynamics*, The IMA Volumes in Mathematics and its Applications **82** (1996) 71-113

[17] E. J. Doedel, H. B. Keller, J. P. Kernévez, *Numerical analysis and control of bifurcation problems*, Int. J. Bifurcation and Chaos **1** (1991) 493-520 and 745-772

[18] G. Domokos, *Global Description of Elastic Bars*, Z. angew. Math. Mech. **74** (1994) T289-T291

[19] A. Friedman, *Foundations of Modern Analysis*, Holt, Rinehart and Winston Eds., New York (1970)

[20] I. M. Gelfand, S. V. Fomin, *Calculus of variations*, Prentice-Hall, Englewood Cliffs, New Jersey (1963)

[21] J. Guckenheimer, P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York (1983)

[22] H. Goldstein, *Classical Mechanics*, second ed., Addison-Wesley, Reading (1980)

[23] J. K. Hale, *Ordinary differential equations*, Wiley-Interscience, New York (1969)

[24] P. Hartman, *Ordinary Differential Equations*, second ed., Birkhäuser, Boston (1982)

[25] T. J. Healey, *Material Symmetry and Chirality in Nonlinearly Elastic Rods*, to appear in Math. Mech. Solids

[26] P. C. Hughes, *Spacecraft attitude dynamics*, John Wiley, New York (1986)

[27] S. Kehrbaum, *Hamiltonian Formulations of the Equilibrium Conditions Governing Elastic Rods: Qualitative Analysis and Effective Properties*, Ph. D. Thesis, University of Maryland, College Park (1997)

[28] S. Kehrbaum, J. H. Maddocks, *Elastic rods, rigid bodies, quaternions and the last quadrature*, Phil. Trans. R. Soc. Lond. A **355** (1997) 2117-2136

[29] S. Kehrbaum, J. H. Maddocks, *Effective properties of elastic rods with high intrinsic twist*, Proceedings of the 16th IMACS World Congress, Lausanne (2000)

[30] J. Kevorkian, J. D. Cole, *Multiple Scale and Singular Perturbation Methods*, Springer-Verlag, New York (1996)

[31] E. Kreyszig, *Introductory functional analysis with applications*, John Wiley, New York (1978)

[32] A. E. H. Love, *A treatise on the mathematical theory of elasticity*, fourth ed., Dover, New York (1944)

[33] Y. Li, J. H. Maddocks, *On the Computation of Equilibria of Elastic Rods, Part I: Integrals, Symmetry and a Hamiltonian Formulation*, unpublished (1996)

[34] J. H. Maddocks, *Stability of Nonlinearly Elastic Rods*, Arch. Rational Mech. Anal. **85** (1984) 311-354

[35] J. H. Maddocks, *Stability and Folds*, Arch. Rational Mech. Anal. **99** (1987) 301-328

[36] I. J. Maddox, *Elements of Functional Analysis*, Cambridge University Press, second ed., Cambridge (1988)

[37] R. S. Manning, J. H. Maddocks, *Symmetry breaking and the twisted elastic ring*, Comput. Methods Appl. Mech. Engrg. **170** (1999) 313-330

[38] R. S. Manning, J. H. Maddocks, J. D. Kahn, *A continuum rod model of sequence-dependent DNA structure*, J. Chem. Phys. **105** (1996) 5626-5646

[39] J. E. Marsden, T. S. Ratiu, *Introduction to Mechanics and Symmetry*, second ed., Springer-Verlag, New York (1999)

[40] A. Mielke, G. Schneider, A Ziegra, *Comparison of Inertial Manifolds and Application to Modulated Systems*, Math. Nachr. **214** (2000) 53-69

[41] J. A. Murdock, *Perturbations: Theory and Methods*, John Wiley, New York (1991)

[42] S. Neukirch, M. E. Henderson, *Classification of the spatial clamped elastica, part I: symmetries and zoology of solutions*, in preparation

[43] E. L. Reiss, *Column Buckling — An Elementary Example of Bifurcation*, in *Bifurcation Theory and Nonlinear Eigenvalue Problems*, J. B. Keller and S. S. Antman eds., Benjamin, New York (1969) 1-16

[44] W. K. Olson, *Simulating DNA at low resolution*, Curr. Opin. Struct. Biol. **6** (1996) 242-256

[45] S. Rey, J. H. Maddocks, *Buckling of an Elastic Rod with High Intrinsic Twist*, Proceedings of the 16th IMACS World Congress, Lausanne (2000)

[46] K. A. Rogers, *Stability Exchange in Parameter-Dependent Constrained Variational Principles with Applications to Elastic Rod Models of DNA Minicircles*, Ph. D. Thesis, University of Maryland, College Park (1997)

[47] J. A. Sanders, F. Verhulst, *Averaging Methods in Nonlinear Dynamical Systems*, Springer-Verlag, New York (1985)

[48] T. Schlick, *Modelling superhelical DNA: recent analytical and dynamic approaches*, Curr. Opin. Struct. Biol. **5** (1995) 245-262

[49] I. Stakgold, *Green's Functions and Boundary Value Problems*, second ed., John Wiley, New York (1998)

[50] C. A. Stuart, *An introduction to bifurcation theory based on differential calculus*, Nonlinear Mathematics and Mechanics, Heriot-Watt Symposium, Vol. IV, Res. Notes in Math. **39**, Pitman, (1979) 76-132

[51] C. A. Stuart, *Bifurcation of homoclinic orbits and bifurcation from the essential spectrum*, SIAM J. Math. Anal. **20** (1989) 1145-1171

[52] G. H. M. van der Heijden, S. Neukirch, V. G. A. Goss, J. M. T. Thompson, *Instability and self-contact phenomena in the writhing of clamped rods*, submitted to Int. J. Eng. Science

[53] E. Zeidler, *Nonlinear Functional Analysis and its Applications, I. Fixed-Point Theorems*, Springer-Verlag, New York (1986)

Curriculum vitae

Born in Sion, Switzerland, in 1971, I lived in Flanthey until I obtained a *Maturité cantonale et fédérale* at the *Collège des Creusets* in Sion, in 1990. I then attended the *École Polytechnique Fédérale* in Lausanne (EPFL), spent one year as exchange student at Carnegie Mellon University, Pittsburgh, and graduated from EPFL with a *Diplôme d'ingénieur physicien* in 1995. After spending one year in a Catholic Mission in N'dendé, Gabon, I attended the University of Tennessee, Knoxville, as graduate student in Mathematics and teaching associate; I obtained a Master of science degree in 1997. I began to work in the group of Prof. Maddocks at EPFL in 1998, preparing the present dissertation. I married Cathy Bagnoud in 2000.