ABSTRACT

Title of Dissertation: HAMILTONIAN FORMULATIONS OF THE EQUILIBRIUM CONDITIONS GOVERNING ELASTIC RODS: QUALITATIVE ANALYSIS AND EFFECTIVE PROPERTIES

Stefan Kehrbaum, Doctor of Philosophy, 1997

Dissertation directed by: Professor John H. Maddocks
Department of Mathematics and
Institute for Physical Science and Technology

A Hamiltonian viewpoint is adopted to study the qualitative behavior and effective properties of the equations governing the equilibrium conditions of elastic rods. The results fall naturally into two groups, the first concerning quite general rod models, and the second regarding effective properties of rods with rapidly varying material properties.

For general rods, a non-canonical Hamiltonian formulation governing the equilibria is developed, extending the well-known six-dimensional Poisson bracket for rigid body motion. It is then shown that at equilibrium the centerline of an infinite, uniform, hyperelastic rod is always bounded in a particular two-dimensional projection. Finally, a canonical Hamiltonian description using Euler parameters is exploited to show that the quadratures relating the frame evolution and the shape of the centerline of the rod can be entirely eliminated by using algebraic manipulations of an additional integral that appears only in the Euler parameter formulation.

In the second and major part of the dissertation, Hamiltonian systems governing the equilibria of elastic rods whose intrinsic shape depends on two different length scales are reformulated in such a way that the standard theory of averaging applies. The homogenized system is used to derive effective properties of rods whose intrinsic local structure is helical; in particular, for such rods the effective constitutive relation is isotropic, provided that all short-scale dependence arises through the unstressed shape. Specifically it is shown that in the asymptotic limit of high intrinsic twist, for a linearly elastic rod with diagonal strain-energy, there exists an effective isotropic bending stiffness given as the harmonic average of the two distinct bending moduli. This work is motivated by the application of continuum rod models as an approximation of large scale deformations of DNA molecules.

HAMILTONIAN FORMULATIONS OF THE EQUILIBRIUM CONDITIONS GOVERNING ELASTIC RODS: QUALITATIVE ANALYSIS AND EFFECTIVE PROPERTIES

by

Stefan Kehrbaum

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland at College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
1997

Advisory Committee:

Professor John H. Maddocks, Chairman/Advisor
Professor Stuart S. Antman
Professor Alex J. Dragt
Professor Royal B. Kellogg
Professor Garrett J. Stuck

DEDICATION

Für meine Familie. Danke ...

ACKNOWLEDGEMENTS

I am grateful to my advisor, Prof. John H. Maddocks, for guiding me through this project. He taught me the fine points of mathematical research and scientific writing; without his vision the completion of this project would certainly not have been possible. Professors Stuart Antman, Alex Dragt, Bruce Kellogg and Garrett Stuck generously agreed to serve on my dissertation committee. I wish to thank them for their valuable input and careful reading of the manuscript.

I am indebted to the former and current members of the Laboratory for Computation and Visualization in Mechanics at the University of Maryland. In particular, Randy Paffenroth and Dr. Tom Stone for invaluable support concerning computer issues (the graphics in this dissertation would not have been possible without their help); Dr. Kathleen Rogers, Dr. Donald Dichmann, Dr. Oscar Gonzalez and Dr. Robert Manning for proofreading the manuscript and various discussions pertaining to this subject. Altogether it has been an enjoyable experience to work with this group during the past two-and-a-half years.

There is a long list of people who have contributed to my academic progress during the past ten years. In particular, I would like to acknowledge PD Dr. Hans-Joachim Zwiesler from the Universität Ulm, and Prof. Donald Lutz from San Diego State University, who made my first stay in the United States, at San Diego State University, possible and provided me with a wonderful opportunity. My thesis advisor at San Diego State University, Prof. Joseph Mahaffy, supported me in the decision to continue my education towards a PhD degree; special thanks to him.

I would like to mention my old friends in Germany, who have not given up on me, while pursuing this goal so far away from home; and my new friends in the United States, who have shown me many new ways ahead.

Last but not least there is my family: my brother Jan, who is in many ways so far ahead of me; and my parents Irmgard and Peter Kehrbaum, who have never doubted that I was going to reach my goals. It is their support and love that made all this possible.

This work was supported by the AFOSR, NSF and ONR. A special thanks goes also to the Department of Mathematics at the University of Maryland for nominating me for a graduate school fellowship.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

In this dissertation we study several aspects of systems of ordinary differential equations governing the equilibrium conditions of elastic rods. A Hamiltonian viewpoint is adopted, emphasizing different canonical and non-canonical formulations of the equilibrium conditions, the corresponding integrals (or invariants) and completely integrable cases. One important application of the continuum rod model is the approximation of large-scale deformations of DNA molecules, which at the base-pair level have a fine scale structure. Motivated by this application, we consider rod models in which the intrinsic (or unstressed) shape has a rapid twist and periodic small scale structure. The main objective of this dissertation is to cast the corresponding Hamiltonian systems into a form that is suitable for an application of the theory of averaging to obtain effective (homogenized) rod models that can be used to approximate long scale deformations.

Several novel results and techniques concerning rods are presented in this dissertation: First, a non-canonical Hamiltonian formulation governing the equilibria of elastic rods is developed, extending the well-known six-dimensional Poisson bracket for rigid body motion. Second, it is shown that at equilibrium the centerline of an infinite, uniform, hyperelastic rod is always bounded in a particular two-dimensional projection. Third, a canonical, Euler parameter, Hamiltonian description is exploited to show that the quadratures relating the centerline of an elastic rod and the frame evolution can be entirely eliminated using algebraic manipulations of an additional integral that appears only in the Euler parameter formulation. Finally, Hamiltonian systems governing the equilibria of elastic rods whose intrinsic shape depends on two different length scales are reformulated in such a way that the standard theory of averaging applies. The homogenized system is used to derive effective properties of rods whose intrinsic local structure is helical.

The starting point of the analysis is the standard Cosserat theory of elastic rods, in which the rod is described by a curve $\mathbf{r}(s)$, representing the centerline, and a director frame $\{\mathbf{d}_i(s)\}_{i=1}^{3}$, locating the cross section of the rod. The main elements of this theory are reviewed in Chapter 2; as one example we discuss a rod theory based on the Kirchhoff Kinetic Analogy, which relates the governing equations for the statics of inextensible and unshearable, linearly elastic rods and the equations of motion of a heavy rigid body tumbling about a fixed point. We utilize this Analogy in several parts of our analysis.

We also exploit the fact that the equations governing the statics of rods possess a Hamiltonian structure, in which evolution is with respect to an arclength parameter. The parts of the theory of Hamiltonian systems relevant to this dissertation are reviewed in Chapter 3, and in Chapter 4 different non-canonical and canonical Hamiltonian systems governing the statics of elastic rods and the dynamics of rigid bodies are discussed. In particular, we extend a well-known six-dimensional non-canonical Hamiltonian formulation for heavy rigid body dynamics by introducing a twelve-dimensional Poisson bracket, written entirely in terms of body components (i.e. components relative to the frame $\{\mathbf{d}_i\}_{i=1}^3$). This extended system includes the components of the centerline $\mathbf{r}$ written with respect to the director frame, and fully describes the statics of elastic rods.

As an alternative to the non-canonical description, we also consider canonical Hamiltonian formulations for elastic rods and rigid bodies by choosing a parametrization of the director frame (or equivalently a parametrization of the rotation group $SO(3)$). Two typical choices for this parametrization are Euler angles and Euler parameters, both of which have been exploited to formulate a three (respectively four) degree-of-freedom system governing the motion of a heavy rigid body (see Holmes and Marsden [29] for a Hamiltonian formulation in terms of Euler angles, or Klein and Sommerfeld [33] for a Lagrangian formulation in terms of Euler parameters). Dichmann et al. [16] have included the spatial components of the centerline $\mathbf{r}$ to obtain a canonical, seven degree-of-freedom, Hamiltonian system in terms of Euler parameters describing the equilibria of elastic rods. Correspondingly, a six degree-of-freedom Euler angle formulation can readily be obtained. Integrals, and corresponding completely integrable cases, for the different Hamiltonian formulations are discussed in detail in Chapter 5.

In Chapter 6 we emphasize the formulation in terms of Euler parameters, and show that the existence of additional integrals due to the extra degree of freedom provides a different view on the Kirchhoff Kinetic Analogy. Since the centerline $\mathbf{r}$ is an ignorable coordinate in any of the Hamiltonian formulations, the full Hamiltonian system describing the statics of rods can always be reduced to a lower-dimensional system governing only the frame evolution. In this viewpoint the centerline can be recovered from the solution of the reduced system through three quadratures. The Kirchhoff Kinetic Analogy can be interpreted as a special case of this reduction, since the Analogy relates the centerline of an inextensible and unshearable rod to the frame evolution (or in this particular case the motion of a heavy rigid body) through the quadratures $\mathbf{r}' = \mathbf{d}_3$. However, many natural problems for rods involve two-point boundary conditions in which $\mathbf{r}$ appears explicitly. For such problems the reduction to the problem governing only the frame evolution is deceptive.

Ilyukhin [31] observed that in any formulation of the rod equilibrium conditions two of the three quadratures can always be eliminated algebraically, using certain integrals, independent of whether the problem is completely integrable or not. We exploit this fact to show that the projection of the centerline $\mathbf{r}$ onto the plane perpendicular to the (constant) force vector is always bounded, provided the rod is uniform. Independently, we observe that the extra degree of freedom in the Hamiltonian Euler parameter formulation yields an additional integral, through which even the third and last quadrature can be performed algebraically. Thus the Euler parameter Hamiltonian formulation yields in some sense a true Kinetic Analogy, in which the number of degrees of freedom of every two-point boundary value problem describing the statics of rods can be reduced from seven to four. Warner [65] exploits this result in computational benchmarks, and observes an increase in performance in numerical simulations by about a factor of three.

In Chapter 7 we turn our attention to the specific solution of the completely integrable case describing an inextensible and unshearable, uniform and isotropic rod. This case typically constitutes the starting point for an analysis of elastic rods, and in particular offers one route to an understanding of non-integrable systems via the computation of perturbation expansions. The main tool for a perturbation analysis of integrable Hamiltonian systems with bounded level sets is the construction of action-angle variables. The unboundedness of most level sets of the full system describing uniform and isotropic rods, makes it necessary to invoke again the Kirchhoff Kinetic Analogy, and construct the respective change of variables for the reduced system describing the motion of a Lagrange top. The formulation in terms of Euler angles seems to yield the only case where action-angle variables can be found explicitly. It is only here that the corresponding Hamilton-Jacobi equation is separable, so that a solution generating the appropriate change of variables can be calculated. As a byproduct of the action-angle variable description, we immediately compute closed-form solutions, although these can certainly be obtained through different means (cf. Ilyukhin [31]). Our treatment of action-angle variables and of closed-form solutions is motivated by two goals: first we would like to gain a better understanding of the integrable system describing inextensible and unshearable, uniform and isotropic rods, which is particularly useful for the interpretation of numerical simulations that complement our averaging analysis presented in Chapter 9. Second, action-angle variables allow us to discuss the homogenization result of Chapter 8 in the context of general Hamiltonian perturbation theory (cf. Appendix E), although I was not successful in carrying through any explicit computations in this direction.

Motivated by our consideration of the continuum rod model as an approximation of DNA, the main result of this dissertation, namely the construction of effective rod models, is derived in Chapter 8. The minimum energy (unstressed) shape of a rod is encapsulated in a configuration $\hat{\mathbf{r}}$ and $\{\hat{\mathbf{d}}_i\}_{i=1}^3$, and we assume that this configurations depends on two different length scales, a slow scale s and a fast scale $\tau = s/\epsilon$, where ϵ is a small parameter. Necessarily, the corresponding strain vector relative to the frame $\{\hat{\mathbf{d}}_i\}_{i=1}^3$ becomes unbounded for small values of ϵ, and a limiting problem cannot be obtained directly. We show how, under appropriate assumptions on the minimum energy shape of the rod, a new frame can be defined, so that the strain vector relative to this frame is bounded in the limit as ϵ tends to zero. The new frame generates a natural change of variables, resulting in a Hamiltonian system with rapidly varying coefficients, which is appropriate for an application of standard averaging theory. We obtain an effective (or homogenized) system, which in some sense can be regarded as an approximate rod model.

In Chapter 9, as a particular application of this homogenization result, we study rods whose centerline is highly coiled (or helical) in the minimum energy configuration. Assuming the radius and pitch of the helix are small and of the same order, we obtain an effective model, in which the helical centerline is replaced by a slowly varying centerline that captures the structure of the rod on longer length scales. It is shown that this model possesses a certain rotational invariance, which, under appropriate additional assumptions, can be seen to be equivalent to isotropy of the effective rod. As a special case of the helical rod model we obtain effective properties of rods that are highly twisted in their unstressed configuration. In particular, it is shown that for a linearly elastic rod, there exists an effective isotropic bending stiffness, computed as the harmonic average of the original distinct bending moduli. For the class of inextensible, unshearable rods with intrinsic helical structure, an analysis of the homogenized system yields a perhaps surprising observation, namely that the effective model does not allow for any extension or shearing effects, although the original helical structure suggests that these effects might exist. We investigate this phenomenon further in Chapter 10, together with different approaches to defining a helical rod model.

In Chapter 11 we summarize the main results and discuss an important aspect of the homogenization results that is not addressed in Chapters 8 and 9, namely, the approximation we formulate in Chapter 8 is only valid for initial value problems. We suggest several possible approaches for obtaining similar results for the two-point boundary value problems that arise in many interesting problems involving rod mechanics.

This dissertation concludes with several appendices. In particular, in Appendices A – C we present a summary of the notation used in this dissertation, properties of Legendre transforms and a definition of elliptic integrals. Furthermore, detailed proofs of the theorems stated in Chapter 7 are given in Appendix D, and several Mathematica scripts are listed in Appendix G. We also add some remarks about known results that relate to the homogenization theory developed in Chapters 8 and 9. First, the homogenization procedure can be interpreted as transforming the Hamiltonian system into a standard form, so that the techniques of Hamiltonian perturbation theory can be applied. In Appendix E, we show how this can be achieved using the action-angle variable description obtained in Chapter 7, and discuss our result in light of some of the main theorems of Hamiltonian perturbation theory. Second, the helical rod model studied in Chapter 9 can also be regarded as approximating a spring and in Appendix F we review some of the known results concerning effective spring theories.

Chapter 2

Cosserat Theory of Elastic Rods

In the following chapter the necessary background on the theory of elastic (Cosserat) rods is summarized. Most of the material in this chapter is standard and can be found for example in Antman [2, Ch. VIII]. In particular, in Section 2.1 we describe the director theory of rods and derive the relevant balance laws governing the equilibria of elastic rods. In Section 2.2 some aspects of the theory of constitutive laws are reviewed, and we list the relevant properties that are used throughout this dissertation. Perhaps one of the most important examples is a linearly elastic and isotropic rod, which typically constitutes the only case for which explicit calculations can be performed. We therefore provide a derivation of the most general form of a constitutive relation governing a linearly elastic, isotropic rod, which to my knowledge has not been recorded before. However, Antman [2] derives a representation for constitutive relations of isotropic, nonlinearly elastic rods, and the special case of a linearly elastic rod is certainly implied by this much more general result. Finally in Section 2.4 a special rod theory based on the Kirchhoff Kinetic Analogy is mentioned. This Analogy allows us to relate the equations governing the equilibria of linearly elastic rods to the equations describing the motion of a heavy rigid body tumbling about a fixed point.

2.1 Director Theory and Kinematic Equations

In (Cosserat) rod theory (cf. Antman [2, Ch. VIII]) the configuration of a rod is specified by three vector-valued functions

$$\mathbf{r}(s), \mathbf{d}_i(s) : [s_1, s_2] \to \mathbb{E}^3, \quad i = 1, 2,$$

where $\mathbb{E}^3$ denotes real, three-dimensional inner product space, and $\mathbf{d}_1(s)$ and $\mathbf{d}_2(s)$ are orthonormal. For equilibria these functions are parametrized by the single variable s. Then the curve $\mathbf{r}(s)$ can be interpreted as the centerline of the rod, and the vectors $\mathbf{d}_1(s)$ and $\mathbf{d}_2(s)$ serve to locate the orientation of a cross

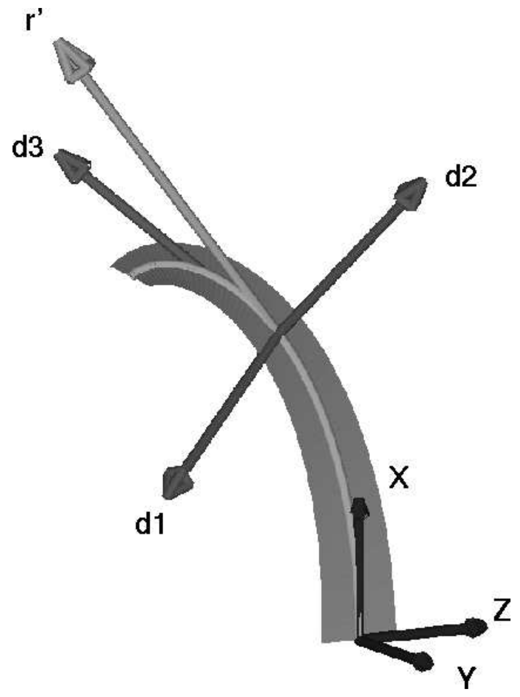

Figure 2.1: A configuration of an elastic rod. The curve $\mathbf{r}(s)$ represents the centerline. The triad $\{\mathbf{d}_i(s)\}$ indicates the orientation of the cross section, with $\mathbf{d}_1$ generating the ribbon that visualizes twisting.

section of the rod at s (cf. Figure 2.1). Define $\mathbf{d}_3(s) \equiv \mathbf{d}_1(s) \times \mathbf{d}_2(s)$ to obtain an orthonormal frame $\mathcal{D} = \{\mathbf{d}_i\}_{i=1}^3$, the *directors*.

The kinematics of the rod are described by the *strain vectors* $\mathbf{u}, \mathbf{v} \in \mathbb{E}^3$, which are defined through the relations

$$\mathbf{d}_i' = \mathbf{u} \times \mathbf{d}_i, \quad i = 1, 2, 3, \tag{2.1}$$

$$\mathbf{r}' = \mathbf{v}, \tag{2.2}$$

where $'$ denotes differentiation with respect to the independent variable s.

It will be natural to decompose vectors with respect to the director frame $\mathcal{D}$; for any vector $\mathbf{p} \in \mathbb{E}^3$, the triple of components $p_i = \mathbf{p} \cdot \mathbf{d}_i$ will be denoted with the sans-serif symbol p. The triple p will be referred to as the body components (or body variable) of $\mathbf{p}$. In contrast, we will denote the components $p^i = \mathbf{p} \cdot \mathbf{e}_i$ in a fixed (inertial), orthonormal frame $\mathcal{E} = \{\mathbf{e}_i\}_{i=1}^3$ by $\underline{\mathbf{p}}$.

Using the decompositions $\mathbf{u} = u_k \mathbf{d}_k$ and $\mathbf{v} = v_k \mathbf{d}_k$ we see that the kinematic equations become

$$\mathbf{d}_i' = u_k \mathbf{d}_k \times \mathbf{d}_i, \quad i = 1, 2, 3, \qquad \mathbf{r}' = v_k \mathbf{d}_k, \tag{2.3}$$

where we use (here and in the remainder of this dissertation) the *summation convention*, that twice repeated Latin indices are summed from 1 to 3, i.e. for example

$$u_i \mathbf{d}_i = \sum_{i=1}^{3} u_i \mathbf{d}_i.$$

The components $u_i = \mathbf{u} \cdot \mathbf{d}_i$ are the *bending* ($i = 1, 2$) and *twisting* ($i = 3$) strains. *Shear* and *dilatation* are measured by (v_1, v_2) and v_3, respectively.

In rod mechanics the forces and moments acting across a material cross section are averaged to yield a net *force* $\mathbf{n}(s) \in \mathbb{E}^3$ and a net *moment* $\mathbf{m}(s) \in \mathbb{E}^3$. Then balance of moments and forces imply the coordinate-free equilibrium equations

$$\mathbf{m}' + \mathbf{r}' \times \mathbf{n} = \mathbf{0}, \quad \mathbf{n}' = \mathbf{0}. \tag{2.4}$$

The components $m_i = \mathbf{m} \cdot \mathbf{d}_i$ are the *bending* ($i = 1, 2$) and *twisting moments* ($i = 3$) in the rod. We refer to n_1 and n_2 as the *shear forces* and n_3 as *tension*.

In some parts of our analysis we will impose material constraints, namely that the rod is *unshearable* and *inextensible*. In mathematical terms inextensibility is then expressed as the condition $|\mathbf{r}'| = 1$, and unshearability translates to $\mathbf{r}' \cdot \mathbf{d}_1 = \mathbf{r}' \cdot \mathbf{d}_2 = 0$. Consequently we may write

$$\mathbf{r}' = \mathbf{d}_3, \tag{2.5}$$

which replaces the kinematic equation (2.2). Note that this definition implies a restriction on the choice of the director frame $\mathcal{D}$ in the minimum energy (unstressed) configuration, as the director $\mathbf{d}_3$ is chosen to be parallel to $\mathbf{r}'$, which is reflected in the formulation of associated constitutive relations (see also the next section). For an inextensible, unshearable rod the parameter s can be interpreted as non-dimensionalized arclength in any configuration.

2.2 Constitutive Relations

The moments $\mathsf{m} = (m_1, m_2, m_3)$ and the shear forces and tension $\mathsf{n} = (n_1, n_2, n_3)$ must be related to the strains $\mathsf{u} = (u_1, u_2, u_3)$ and $\mathsf{v} = (v_1, v_2, v_3)$ via suitable *constitutive relations*. The material is called *hyperelastic* if there exists a *strain-energy function* $W(\mathsf{w}, \mathsf{z}, s)$ with $\mathsf{w} = (w_1, w_2, w_3)$, $\mathsf{z} = (z_1, z_2, z_3)$ such that

$$\frac{\partial W}{\partial \mathsf{w}}(\mathbf{0}, \mathbf{0}, s) = \frac{\partial W}{\partial \mathsf{z}}(\mathbf{0}, \mathbf{0}, s) = \mathbf{0},$$

and m and n are related to u and v through the constitutive relations

$$\mathsf{m} = \frac{\partial W}{\partial \mathsf{w}}(\mathsf{u} - \hat{\mathsf{u}}, \mathsf{v} - \hat{\mathsf{v}}, s), \quad \mathsf{n} = \frac{\partial W}{\partial \mathsf{z}}(\mathsf{u} - \hat{\mathsf{u}}, \mathsf{v} - \hat{\mathsf{v}}, s). \tag{2.6}$$

Here $\hat{\mathsf{u}}(s) = \big(\hat{u}_1(s), \hat{u}_2(s), \hat{u}_3(s)\big)$ and $\hat{\mathsf{v}}(s) = \big(\hat{v}_1(s), \hat{v}_2(s), \hat{v}_3(s)\big)$ are the components of (intrinsic) strain in the minimum energy unstressed configuration.

We assume that the function W is as often continuously differentiable as needed in the analysis. In addition we require that the strain-energy function is both convex and coercive. The function W is said to be *coercive* if

$$\frac{W(\mathsf{w}, \mathsf{z}, s)}{\sqrt{|\mathsf{w}|^2 + |\mathsf{z}|^2}} \to \infty \quad \text{as} \quad |\mathsf{w}|^2 + |\mathsf{z}|^2 \to \infty.$$

For a reasonable theory it is furthermore necessary to assume that any deformation of the rod is orientation-preserving, i.e. a rod of positive length cannot be compressed to zero length. A condition ensuring this requirement is that

$$v_3 = \mathbf{v} \cdot \mathbf{d}_3 > 0 \tag{2.7}$$

(see Antman [2, Sec. VIII.6] for a more general treatment of this property). Thus one must certainly enforce this condition in the minimum energy configuration, i.e. $\hat{v}_3 > 0$. It is important to realize that for a hyperelastic material there are solutions that violate (2.7), and which therefore do not correspond to physical rod configurations.

For an inextensible, unshearable rod the strain-energy function is a function of the strains $\mathsf{u} - \hat{\mathsf{u}}$ only, so that the constitutive relations (2.6) reduce to

$$\mathsf{m} = \frac{\partial W}{\partial \mathsf{w}}(\mathsf{u} - \hat{\mathsf{u}}, s). \tag{2.8}$$

A standard approximation in rod theory is to assume the rod to be linearly elastic, i.e. the strain-energy is a quadratic function of the strains

$$W(\mathsf{w}, \mathsf{z}, s) = \frac{1}{2}\begin{pmatrix}\mathsf{w}\\ \mathsf{z}\end{pmatrix} \cdot \mathbf{W}(s)\begin{pmatrix}\mathsf{w}\\ \mathsf{z}\end{pmatrix}, \tag{2.9}$$

where $\mathbf{W} \in \mathbb{R}^{6\times 6}$ is symmetric and positive-definite. For an inextensible, unshearable, linearly elastic rod the strain-energy function (2.9) reduces to

$$W(\mathsf{w}, s) = \frac{1}{2}\mathsf{w} \cdot \mathbf{K}(s)\mathsf{w}, \tag{2.10}$$

for a symmetric, positive-definite matrix $\mathbf{K}(s) \in \mathbb{R}^{3\times 3}$, which can in principle be chosen to be diagonal by choosing the frame $\mathcal{D}$ appropriately. However, recall that we defined an inextensible, unshearable rod to satisfy the constraint $\mathbf{r}' = \mathbf{d}_3$

(or equivalently $\mathsf{v} = (0,0,1)$) in any configuration, which provides a restriction on the possible choice of the director frame. Thus in general the frame $\mathcal{D}$ cannot be chosen to simultaneously diagonalize the matrix $\mathbf{K}(s)$ and satisfy $\hat{\mathsf{v}} = (0,0,1)$.

Nevertheless, there exist special rod theories based on the constraint $\mathbf{r}' = \mathbf{d}_3$ and constitutive relations of the form (2.8), where W is the diagonal quadratic form

$$W(\mathsf{w}, s) = \frac{1}{2}\sum_{i=1}^{3} K_i(s) w_i^2. \tag{2.11}$$

These theories typically start with the assumption that the rod is a solid body, whose diameter is small compared to its length. If one further assumes the material to be homogeneous, the coefficients K_i can be given the additional interpretation

$$K_1 = I_1 E, \quad K_2 = I_2 E, \quad K_3 = (I_1 + I_2)G,$$

where E is the *elastic (or Young's) modulus*, G the *shear modulus* and I_i the second moment of the area of cross section about the axis $\mathbf{d}_i$ (see Love [41]). The elastic and shear moduli are related through

$$G = \frac{E}{2(1+\nu)},$$

where ν is *Poisson's ratio* (for a typical material the range of this parameter is $\nu \in (0, \frac{1}{2})$).

The constitutive relation is *uniform* if the strain-energy function W does not explicitly depend on the parameter s. The rod is uniform if in addition the components of unstressed strain $\hat{\mathsf{u}}$ and $\hat{\mathsf{v}}$ are constant.

Under the assumption that the shear strains $\hat{\mathsf{v}}$ satisfy $\hat{\mathsf{v}} = (0,0,1)$ we will call the constitutive relation *isotropic* (or transversely isotropic) if the strain-energy function $W(\mathsf{w}, \mathsf{z}, s)$ is invariant under rotations about the $\mathbf{d}_3$ axis; equivalently

$$\frac{\partial}{\partial \alpha} W\big(\mathbf{Q}(\alpha)\mathsf{w}, \mathbf{Q}(\alpha)\mathsf{z}, s\big) = 0 \tag{2.12}$$

for all $0 \le \alpha < 2\pi$, $\mathsf{w}, \mathsf{z} \in \mathbb{R}^3$ with

$$\mathbf{Q}(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.13}$$

If the strains $\hat{\mathsf{v}}$ are of a more general form, this definition can be extended, and the strain-energy function should satisfy a rotational invariance about the axis defined by $\hat{v}_i \mathbf{d}_i$. However, for our purposes the above definition suffices. We will further describe the *rod* as isotropic if it has no preferred material direction for bending; provided that the constitutive relation is isotropic, the rod is isotropic if in addition $\hat{u}_i(s) = \hat{v}_i(s) = 0, i = 1, 2$. Antman [2, Theorem 9.19] shows that for an isotropic constitutive relation, the strain-energy function has the form

$$W(\mathsf{u}, \mathsf{v}, s) = W\big(J(\mathsf{u}, \mathsf{v}), s\big), \tag{2.14}$$

where

$$J(\mathsf{u}, \mathsf{v}) = (u_1^2 + u_2^2, v_1^2 + v_2^2, u_1 v_1 + u_2 v_2, u_3, v_3).$$

Consistent with this representation we show that for an isotropic, linearly elastic constitutive relation, the strain-energy function has the following form:

Lemma 2.1. *For an isotropic, linearly elastic constitutive relation, the strain-energy function is of the form* (2.9)*, where*

$$\mathbf{W} = \left(\begin{array}{ccc|ccc} a & 0 & 0 & e & g & 0 \\ 0 & a & 0 & -g & e & 0 \\ 0 & 0 & b & 0 & 0 & f \\ \hline e & -g & 0 & c & 0 & 0 \\ g & e & 0 & 0 & c & 0 \\ 0 & 0 & f & 0 & 0 & d \end{array}\right) \tag{2.15}$$

Proof. Rewrite the matrix $\mathbf{W} \in \mathbb{R}^{6\times 6}$ in (2.9) as

$$\mathbf{W} = \begin{pmatrix} \mathbf{W}_1 & \mathbf{W}_3 \\ \mathbf{W}_3^T & \mathbf{W}_2 \end{pmatrix}$$

with $\mathbf{W}_1$ and $\mathbf{W}_2$ symmetric. Then the linear constitutive relation is isotropic precisely if

$$\begin{pmatrix} \mathbf{Q}^T(\alpha) & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^T(\alpha) \end{pmatrix} \mathbf{W} \begin{pmatrix} \mathbf{Q}(\alpha) & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}(\alpha) \end{pmatrix}$$

is independent of α, or equivalently the matrices

$$\mathbf{Q}^T(\alpha)\mathbf{W}_i\mathbf{Q}(\alpha), \quad i = 1, 2, 3,$$

do not depend on α. We determine the general form of a matrix $\mathbf{M} = (m_{ij}) \in \mathbb{R}^{3\times 3}$, so that

$$\mathbf{Q}^T(\alpha)\mathbf{M}\mathbf{Q}(\alpha) = \mathbf{M} \quad \forall \alpha. \tag{2.16}$$

The entries $\bar{m}_{ij}$ of $\bar{\mathbf{M}} = \mathbf{Q}^T\mathbf{M}\mathbf{Q} - \mathbf{M}$ are given by

$$
\begin{aligned}
\bar{m}_{11} &= -\sin\alpha\big((m_{12}+m_{21})\cos\alpha + (m_{11}-m_{22})\sin\alpha\big),\\
\bar{m}_{12} &= \sin\alpha\big((m_{11}-m_{22})\cos\alpha - (m_{12}+m_{21})\sin\alpha\big),\\
\bar{m}_{13} &= m_{13}(\cos\alpha - 1) - m_{23}\sin\alpha,\\
\bar{m}_{21} &= \sin\alpha\big((m_{11}-m_{22})\cos\alpha - (m_{12}+m_{21})\sin\alpha\big),\\
\bar{m}_{22} &= \sin\alpha\big((m_{12}+m_{21})\cos\alpha + (m_{11}-m_{22})\sin\alpha\big),\\
\bar{m}_{23} &= m_{23}(\cos\alpha - 1) + m_{13}\sin\alpha,\\
\bar{m}_{31} &= m_{31}(\cos\alpha - 1) - m_{32}\sin\alpha,\\
\bar{m}_{32} &= m_{32}(\cos\alpha - 1) + m_{31}\sin\alpha,\\
\bar{m}_{33} &= 0.
\end{aligned}
$$

For (2.16) to hold we must have $\bar{m}_{ij} = 0 \ \ \forall\alpha$, so that

$$
m_{11} = m_{22}, \quad m_{12} = -m_{21}, \quad m_{13} = m_{23} = m_{31} = m_{32} = 0.
$$

This implies that if $\mathbf{M}$ is symmetric and (2.16) holds, then $\mathbf{M}$ is diagonal. Therefore we conclude that $\mathbf{W}_1$ and $\mathbf{W}_2$ are diagonal, $\mathbf{W}_3$ is skew-symmetric, and have the particular forms as indicated in (2.15). □

Note that the form of the matrix (2.15) conforms with the representation (2.14) of an isotropic strain-energy function, since

$$
\begin{aligned}
W(\mathsf{w},\mathsf{z},s) &= \frac{1}{2}\begin{pmatrix}\mathsf{w}\\ \mathsf{z}\end{pmatrix}\cdot\mathbf{W}(s)\begin{pmatrix}\mathsf{w}\\ \mathsf{z}\end{pmatrix}\\
&= \frac{1}{2}\Big[a(u_1^2+u_2^2) + bu_3^2 + c(v_1^2+v_2^2) + dv_3^2\\
&\quad +2e(u_1v_1+u_2v_2) + 2fu_3v_3\\
&\quad +2g(u_1v_2 - u_2v_1)\Big],
\end{aligned}
$$

where the last term can be rewritten as

$$
2g(u_1v_2 - u_2v_1) = 2g\sqrt{(u_1^2+u_2^2)(v_1^2+v_2^2) - (u_1v_1+u_2v_2)^2}.
$$

2.3 Parametrization of the Rotation Group

Specifying the orientation of the director frame $\mathcal{D}$ with respect to the inertial frame $\mathcal{E}$ is equivalent to choosing a parametrization of $SO(3)$, the group of proper orthogonal transformations in $\mathbb{E}^3$ (see Arnold [7, Ch. 6] or Goldstein [21, Ch. 4] for a treatment of the kinematics of rigid bodies; Shuster [58] provides a survey of different representations of $SO(3)$).

Direction cosines provide one elementary way to specify the orientation of the moving frame $\mathcal{D}$ relative to the inertial frame $\mathcal{E}$. However, these nine coordinates are not independent, since only three parameters are necessary to represent a rotation. Thus the constraints on these coordinates make direction cosines in general impractical for use in a variational formulation. However, it is possible to formulate the equations governing the motion of a tumbling rigid body (and even those describing elastic rods) entirely in terms of body components (i.e. components relative to the moving frame $\mathcal{D}$). Therefore the rotation matrix does not explicitly appear in the equations and is only needed to recover the spatial components of the variables (relative to the fixed frame $\mathcal{E}$), after a solution to the frame equations is computed. Non-canonical Hamiltonian formulations involving only body components are described in Sections 4.1 and 4.2.

A minimal representation of the rotation group is given by three *Euler angles*, which provide a local single-covering of $SO(3)$. However, a polar singularity arises, which can provide serious analytical and numerical difficulties. Additionally, the appearance of trigonometric functions is usually disadvantageous for both symbolic and numerical calculations. Nevertheless Euler angles certainly seem to be useful when considering integrable cases of elastic rods. In particular, the computation of action-angle variables for the Hamiltonian formulation of inextensible and unshearable, isotropic and uniform rods is possible only because the corresponding Hamilton-Jacobi equation in terms of Euler angles is separable (see the requisite calculations in Chapter 7).

The global double-covering of $SO(3)$ given by four *Euler parameters* (or quaternions) provides a considerable advantage for numerical computations. In addition, integrals arising in the Hamiltonian formulation of these variables seem to provide a different view on the Kirchhoff Kinetic Analogy, in that the quadratures inherent in (2.2), respectively (2.5), are reduced to simple algebraic manipulations, so that in some sense a true Kinetic Analogy is obtained (cf. Section 6.3).

We proceed to describe in more detail several parametrizations of $SO(3)$. The description of elastic rods and rigid bodies involves variables $\mathbf{p} \in \mathbb{E}^3$ that may be either expressed as components $\underline{\mathbf{p}}$ in the inertial frame $\mathcal{E}$ or as components p in the director frame $\mathcal{D}$. These variables are related via a rotation matrix $\mathbf{A} \in SO(3)$, i.e.

$$\mathsf{p} = \mathbf{A}\underline{\mathbf{p}},$$

and the different parametrizations of $SO(3)$ described in the following sections allow us to relate the components in the different frames in an explicit way.

2.3.1 Direction Cosines

Given two linearly independent, non-zero unit vectors $\mathbf{i}, \mathbf{j} \in \mathbb{E}^3$ one immediately constructs the rotation matrix $\mathbf{A}$ provided the components of $\mathbf{i}$ and $\mathbf{j}$ relative to both the frames $\mathcal{E}$ and $\mathcal{D}$ are known: Define a third vector $\mathbf{k}$ by $\mathbf{k} \equiv (\mathbf{i} \times \mathbf{j})/|\mathbf{i} \times \mathbf{j}|$. Then the components of $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ with respect to $\mathcal{E}$ and $\mathcal{D}$ are related through the rotation matrix $\mathbf{A}$ via

$$
\begin{aligned}
\mathsf{i} &= \mathbf{A}\underline{\mathbf{i}}, \\
\mathsf{j} &= \mathbf{A}\underline{\mathbf{j}}, \\
\mathsf{k} &= \mathbf{A}\underline{\mathbf{k}},
\end{aligned}
$$

which implies that $\mathbf{A}$ can be represented in terms of nine direction cosines (e.g. Goldstein [21, Ch. 4]) as

$$
\mathbf{A} = \begin{pmatrix} \mathsf{i} \cdot \underline{\mathbf{i}} & \mathsf{i} \cdot \underline{\mathbf{j}} & \mathsf{i} \cdot \underline{\mathbf{k}} \\ \mathsf{j} \cdot \underline{\mathbf{i}} & \mathsf{j} \cdot \underline{\mathbf{j}} & \mathsf{j} \cdot \underline{\mathbf{k}} \\ \mathsf{k} \cdot \underline{\mathbf{i}} & \mathsf{k} \cdot \underline{\mathbf{j}} & \mathsf{k} \cdot \underline{\mathbf{k}} \end{pmatrix},
$$

provided the frames $\mathcal{E}$ and $\mathcal{D}$ are orthonormal. This particular parametrization of the rotation group will be useful in connection with a twelve-dimensional, non-canonical Hamiltonian formulation for the equilibria of elastic rods (cf. Section 4.2).

2.3.2 Euler Angles

For the construction of Euler angles (see for example Antman [2, Sec. VIII.11]) the director $\mathbf{d}_3$ is represented in spherical coordinates

$$
\mathbf{d}_3 = \sin\theta\, \mathbf{h} + \cos\theta\, \mathbf{e}_3, \quad \mathbf{h} \equiv \cos\phi\, \mathbf{e}_1 + \sin\phi\, \mathbf{e}_2.
$$

Provided that $\sin\theta \neq 0$, an intermediate orthonormal frame $\{\mathbf{k}_i\}_{i=1}^3$ can be uniquely defined through

$$
\begin{aligned}
\mathbf{k}_2 &\equiv \frac{\mathbf{e}_3 \times \mathbf{d}_3}{|\mathbf{e}_3 \times \mathbf{d}_3|} = -\sin\phi\, \mathbf{e}_1 + \cos\phi\, \mathbf{e}_2, \\
\mathbf{k}_1 &\equiv \mathbf{k}_2 \times \mathbf{d}_3 = \cos\theta\, \mathbf{h} - \sin\theta\, \mathbf{e}_3, \\
\mathbf{k}_3 &= \mathbf{d}_3.
\end{aligned}
$$

Denoting by ψ the angle from $\mathbf{k}_1$ to $\mathbf{d}_1$, we can write the frame $\mathcal{D}$ in the coordinates $\boldsymbol{\phi} = (\phi, \theta, \psi)$ (the Euler angles) as

$$
\begin{aligned}
\mathbf{d}_1 &= \cos\psi\, \mathbf{k}_1 + \sin\psi \mathbf{k}_2, \\
\mathbf{d}_2 &= -\sin\psi\, \mathbf{k}_1 + \cos\psi\, \mathbf{k}_1,
\end{aligned}
$$

or in components relative to the frame $\mathcal{E}$

$$
\begin{aligned}
\underline{\mathbf{d}}_1 &= \begin{pmatrix} \cos\psi\cos\theta\cos\phi - \sin\psi\sin\phi \\ \cos\psi\cos\theta\sin\phi + \sin\psi\cos\phi \\ -\cos\psi\sin\theta \end{pmatrix}, \\
\underline{\mathbf{d}}_2 &= \begin{pmatrix} -\sin\psi\cos\theta\cos\phi - \cos\psi\sin\phi \\ -\sin\psi\cos\theta\sin\phi + \cos\psi\cos\phi \\ \sin\psi\sin\theta \end{pmatrix}, \\
\underline{\mathbf{d}}_3 &= \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix},
\end{aligned}
\tag{2.17}
$$

so that the rotation matrix $\mathbf{A}$ is

$$\mathbf{A}(\boldsymbol{\phi}) = (\underline{\mathbf{d}}_1, \underline{\mathbf{d}}_2, \underline{\mathbf{d}}_3).$$

Now define a basis $\{\mathbf{b}^i(\boldsymbol{\phi})\}_{i=1}^3$ for $\mathbb{R}^3$ in terms of these angles:

$$
\begin{aligned}
\mathbf{b}^1(\boldsymbol{\phi}) &= \begin{pmatrix} -\cos\psi\sin\theta \\ \sin\psi \\ 0 \end{pmatrix}, \\
\mathbf{b}^2(\boldsymbol{\phi}) &= \begin{pmatrix} \sin\psi\sin\theta \\ \cos\psi \\ 0 \end{pmatrix}, \\
\mathbf{b}^3(\boldsymbol{\phi}) &= \begin{pmatrix} \cos\theta \\ 0 \\ 1 \end{pmatrix},
\end{aligned}
\tag{2.18}
$$

so that the components of strain u_i can be written in the form

$$u_i = \mathbf{b}^i(\boldsymbol{\phi}) \cdot \boldsymbol{\phi}', \quad i = 1, 2, 3. \tag{2.19}$$

Euler angles have been employed in a Lagrangian description of elastic rods for example by Antman and Jordan [3] and Maddocks [43]. Holmes and Marsden [29] formulate a six degree-of-freedom, canonical Hamiltonian formulation involving these parameters, which is described in Section 4.4.

2.3.3 Euler Parameters

Euler parameters are a quadruple $\mathbf{q} = (q_1, q_2, q_3, q_4)$ with unit length, i.e.

$$\mathbf{q} \cdot \mathbf{q} = q_1^2 + q_2^2 + q_3^2 + q_4^2 = 1. \tag{2.20}$$

Euler's Theorem [21] asserts that every rotation $\mathbf{A} \in SO(3)$ is equivalent to a rotation about a unit vector $\mathbf{k}$ through an angle Φ. Then the Euler parameters are defined by

$$q_j = \mathbf{k} \cdot \mathbf{e}_j \sin(\Phi/2),\ j = 1,2,3, \quad q_4 = \cos(\Phi/2).$$

Euler parameters provide a double-covering of $SO(3)$ which is related to the fact that if $\mathbf{q}$ corresponds to the rotation $(\mathbf{k}, \Phi)$, then $-\mathbf{q}$ corresponds to the *coterminal* rotation $(\mathbf{k}, \Phi + 2\pi)$.

The rotation $\mathbf{A}(\mathbf{q}) = (\underline{\mathbf{d}}_1, \underline{\mathbf{d}}_2, \underline{\mathbf{d}}_3)$ can be written in terms of the Euler parameters as

$$\underline{\mathbf{d}}_1 = \begin{pmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 \\ 2(q_1q_2 + q_3q_4) \\ 2(q_1q_3 - q_2q_4) \end{pmatrix}, \quad \underline{\mathbf{d}}_2 = \begin{pmatrix} 2(q_1q_2 - q_3q_4) \\ -q_1^2 + q_2^2 - q_3^2 + q_4^2 \\ 2(q_1q_4 + q_2q_3) \end{pmatrix},$$

and

$$\underline{\mathbf{d}}_3 = \begin{pmatrix} 2(q_1q_3 + q_2q_4) \\ 2(-q_1q_4 + q_2q_3) \\ -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{pmatrix}.$$

Dichmann [15] introduces the notation

$$\mathbf{B}_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \mathbf{B}_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \tag{2.21}$$

and

$$\mathbf{B}_3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

to obtain the representation for the components of strain in the director frame $\mathcal{D}$

$$u_j = \frac{2}{|\mathbf{q}|^2}\, \mathbf{q}' \cdot \mathbf{B}_j \mathbf{q}, \quad j = 1,2,3.$$

2.4 Kirchhoff's Kinetic Analogy and the Dynamics of Tumbling Rigid Bodies

According to Love [41, Sec. 260] the classic form of the Kirchhoff Kinetic Analogy is based on a rod theory with the constraint $\mathbf{r}' = \mathbf{d}_3$ and linear constitutive relations of the form

$$\mathsf{m} = \mathbf{K}(s)\big(\mathsf{u} - \hat{\mathsf{u}}(s)\big), \tag{2.22}$$

where $\mathbf{K}$ is a positive-definite matrix. If the matrix $\mathbf{K}$ is constant and $\hat{\mathsf{u}} = \mathbf{0}$, equations (2.4) are equivalent to the classic equations of motion of a heavy rigid body in a uniform gravitational field tumbling about a fixed point; in this case the frame $\mathcal{D} = \{\mathbf{d}_i\}_{i=1}^3$ (referred to as the body frame) is fixed in the body in such a way that the center of mass lies on the $\mathbf{d}_3$ axis. The matrix $\mathbf{K}$ is the inertia tensor of the rigid body about the fixed point referred to the frame $\mathcal{D}$, and equations (2.4) are typically re-written as

$$\mathbf{m}' + \mathbf{d}_3 \times \mathbf{n} = \mathbf{0}, \quad \mathbf{n}' = \mathbf{0}, \tag{2.23}$$

with the independent parameter s now interpreted as time (and usually written t). We refer to this problem as describing the motion of a tumbling rigid body.

It is also often assumed that the body frame coincides with the principal axes of inertia so that $\mathbf{K}$ is a diagonal matrix with entries $K_1 \leq K_2 \leq K_3$, which is a restriction on the possible locations of the center of mass. Nevertheless in this special case, the equations describing the dynamics of a tumbling rigid body are identical in form to those describing the equilibria of a Kirchhoff rod with diagonal strain-energy function (2.11).

Two important special cases in which the problem is completely integrable (in the classic sense) have been studied extensively. A tumbling rigid body that is free to move with no external forces acting on it (or equivalently the fixed point is at the center of mass) is often described as the *free rigid body*. In the context of rods the free rigid body corresponds to special problems where the end-loadings are pure moments, so that $\mathbf{n} = \mathbf{0}$. The second case is that of a *Lagrange top*, which is a heavy rigid body with a diagonalized inertia tensor in which K_1 and K_2 are equal. The Lagrange top corresponds to the isotropic Kirchhoff rod.

A less widely known fact is that the Kirchhoff Kinetic Analogy persists for non-zero $\hat{\mathsf{u}}$. For example Wittenburg [66, Sec. 4.6] considers a gyrostat comprising a carrier rigid body in which is mounted a spinning sub-body (often called a momentum wheel). The function $\hat{\mathsf{u}}$ can then be interpreted as the angular velocity of the momentum wheel relative to the carrier, and the Analogy applies if one assumes that these relative velocities are prescribed functions of time. (The Analogy can also be made for more general systems, e.g. for gyrostats with multiple momentum wheels.)

Chapter 3

Hamiltonian Systems

This chapter contains a summary of the necessary background on finite-dimensional Hamiltonian systems as it pertains to this dissertation. After giving a Poisson description of the Hamiltonian system in Section 3.1, we discuss integrals and conditions for integrability in Section 3.2. For the understanding of integrable systems and their perturbations, coordinate transformations prove to be a useful tool (cf. Section 3.3). In particular for a canonical, integrable Hamiltonian system with compact level sets, action-angle variables exist, and we review their construction in Section 3.4.

3.1 Hamiltonian Structures

We give the Poisson description of a Hamiltonian system in terms of a Poisson bracket and specialize to the standard canonical formulation in local coordinates. General introductions to the theory of Hamiltonian systems are provided for example by Olver [51], Arnold [7] or Meyer and Hall [45]. For a succinct survey on Hamiltonian systems see also [6].

We define *phase space* to be a smooth m-dimensional manifold $\mathbf{X}$, where in applications often $\mathbf{X} = \mathbb{R}^{2n}$.

Definition 3.1. *A* Poisson bracket *on* $\mathbf{X}$ *is an operator that assigns to each pair of real-valued, smooth functions* F *and* G *on* $\mathbf{X}$ *a real-valued, smooth function* $\{F, G\}$ *on* $\mathbf{X}$*, satisfying the following properties:*

- *Bilinearity*

  $$\{aF + bG, H\} = a\{F, H\} + b\{G, H\}, \quad \{F, aG + bH\} = a\{F, G\} + b\{F, H\},$$

  for constants $a, b \in \mathbb{R}$*,*

- *Skew-Symmetry*

  $$\{F, G\} = -\{G, F\},$$

- *Jacobi Identity*

$$\{\{F,G\},H\} + \{\{H,F\},G\} + \{\{G,H\},F\} = 0,$$

- *Leibniz' Rule*

$$\{F,GH\} = \{F,G\}H + G\{F,H\}.$$

It can be shown that in local coordinates $\mathbf{z} = (z_1, \dots, z_m)$ this bracket can be written as

$$\{F,G\}(\mathbf{z}) = \nabla F(\mathbf{z}) \cdot \mathcal{J}(\mathbf{z}) \nabla G(\mathbf{z}),$$

where $\mathcal{J}(\mathbf{z}) = \big(\mathcal{J}_{ij}(\mathbf{z})\big)$ is the skew-symmetric *structure matrix* defined by $\mathcal{J}_{ij}(\mathbf{z}) = \{z_i, z_j\}, 1 \leq i,j \leq m$. The coordinates $\mathbf{z}$ on an even-dimensional manifold $\mathbf{X} = \mathbb{R}^{2n}$ will be called *canonical* if the structure matrix is the standard constant matrix

$$\mathbf{J}_{2n} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix}, \tag{3.1}$$

where $\mathbf{I}$ and $\mathbf{0}$ are the identity and null matrices of dimension n.

Definition 3.2. *Given a smooth function $\mathcal{H} : \mathbf{X} \to \mathbb{R}$ (the* Hamiltonian*) and a structure matrix $\mathcal{J}(\mathbf{z})$ in local coordinates $\mathbf{z}$, a* Hamiltonian system *is given by the system of m ordinary differential equations*

$$\frac{d}{dt}\mathbf{z} = \mathcal{J}(\mathbf{z}) \nabla_{\mathbf{z}} \mathcal{H}(\mathbf{z}, t). \tag{3.2}$$

If $\mathcal{J} = \mathbf{J}_{2n}$ then (3.2) *is called a canonical Hamiltonian system with n degrees of freedom. If the Hamiltonian $\mathcal{H}$ is independent of t, the system is called* autonomous.

3.2 Integrals and Complete Integrability

A function $I : \mathbf{X} \to \mathbb{R}$ is an *integral* (or invariant) of (3.2) if it is constant along solutions of (3.2). Equivalently I is an integral if

$$\{I, \mathcal{H}\}(\mathbf{z}) = \nabla I(\mathbf{z}) \cdot \mathcal{J}(\mathbf{z}) \nabla_{\mathbf{z}} \mathcal{H}(\mathbf{z}, t) = 0.$$

In particular, there may exist integrals, no matter what the Hamiltonian is, called *Casimirs*; the function $C : \mathbf{X} \to \mathbb{R}$ is a Casimir if ∇C is in the null-space of $\mathcal{J}$ (cf. [51]), i.e. $\{C, F\} = 0$ for all smooth functions F on $\mathbf{X}$. In a canonical formulation there are no Casimirs since the operator $\mathbf{J}_{2n}$ is non-singular.

We will call the Hamiltonian system (3.2) *completely integrable* if its solution can be reduced to quadratures, so that at least theoretically a solution in closed form can be found. In general an m-dimensional system of ordinary differential equations must have m linearly independent integrals in order to be completely integrable. For Hamiltonian systems the situation is different: fewer integrals are required provided that the integrals satisfy an additional commutativity condition. Two integrals I_1 and I_2 are said to be in *involution* or to commute if I_1 is an integral with respect to the flow generated by I_2, i.e.

$$\{I_1, I_2\}(\mathbf{z}) = \nabla I_1(\mathbf{z}) \cdot \mathcal{J}(\mathbf{z}) \nabla I_2(\mathbf{z}) = 0.$$

Suppose the Hamiltonian system (3.2) possesses k Casimirs and an additional l integrals in involution, all of which are linearly independent. Then the system is completely integrable if

$$2l + k = m.$$

3.3 Coordinate Transformations

Consider a change of variables $(\mathbf{z}, t) \to (\bar{\mathbf{z}}, t)$, where $\bar{\mathbf{z}} = \mathbf{f}(\mathbf{z}, t)$. Denote the inverse of $\mathbf{f}$ (regarded as a function of $\mathbf{z}$) by $\mathbf{g}$, i.e. $\mathbf{g}\big(\mathbf{f}(\mathbf{z}, t), t\big) = \mathbf{z}$. Then $\bar{\mathbf{z}}(t)$ satisfies the differential equation

$$\bar{\mathbf{z}}' = \left.\frac{\partial \mathbf{f}(\mathbf{z}, t)}{\partial t}\right|_{\mathbf{z}=\mathbf{g}(\bar{\mathbf{z}}, t)} + \mathbf{F}(\mathbf{z}, t) \mathcal{J}(\mathbf{z}) \mathbf{F}(\mathbf{z}, t)^T\big|_{\mathbf{z}=\mathbf{g}(\bar{\mathbf{z}}, t)} \nabla_{\bar{\mathbf{z}}} H(\mathbf{g}(\bar{\mathbf{z}}, t), t), \tag{3.3}$$

where $\mathbf{F}(\mathbf{z}, t) = \partial \mathbf{f}(\mathbf{z}, t)/\partial \mathbf{z}$ and $\partial \mathbf{f}(\mathbf{z}, t)/\partial \mathbf{z}$ denotes the matrix of partial derivatives of $\mathbf{f} = (f_1, \dots, f_m)$, i.e.

$$\begin{pmatrix} \frac{\partial f_1}{\partial z_1} & \cdots & \frac{\partial f_1}{\partial z_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial z_1} & \cdots & \frac{\partial f_m}{\partial z_m} \end{pmatrix}.$$

Now suppose the set of coordinates $\mathbf{z}$ is canonical, i.e. the structure matrix is the matrix $\mathcal{J} = \mathbf{J}_{2n}$ defined in (3.1). Then a change of variables $\bar{\mathbf{z}} = \mathbf{f}(\mathbf{z}, t)$ that satisfies

$$\frac{\partial \mathbf{f}}{\partial \mathbf{z}}(\mathbf{z}, t) \mathbf{J}_{2n} \left(\frac{\partial \mathbf{f}}{\partial \mathbf{z}}(\mathbf{z}, t) \right)^T = \mathbf{J}_{2n}$$

is called canonical or symplectic, and the transformed equations (3.3) are always Hamiltonian (see Meyer and Hall [45, Ch. IV.A]).

3.4 Action-Angle Variables and Hamilton-Jacobi Theory

Suppose the Hamiltonian system (3.2) on the manifold $\mathbf{X} = \mathbb{R}^{2n}$ is canonical and autonomous. Then setting $\mathbf{z} = (\mathbf{q}, \mathbf{p})$, $\mathbf{q}, \mathbf{p} \in \mathbb{R}^n$, we can write the equations in the form

$$\mathbf{q}' = \frac{\partial H(\mathbf{q}, \mathbf{p})}{\partial \mathbf{p}}, \quad \mathbf{p}' = -\frac{\partial H(\mathbf{q}, \mathbf{p})}{\partial \mathbf{q}}$$

with Hamiltonian $H(\mathbf{z}) = H(\mathbf{q}, \mathbf{p})$. Furthermore assume that the system is integrable, i.e. there exist n independent functions F_i in involution. Consider a level set of the functions F_i, i.e.

$$\mathbf{M_f} = \{\mathbf{z} : F_i(\mathbf{z}) = f_i,\ i = 1, \dots, n\}.$$

Then Liouville's Theorem gives a characterization of the Hamiltonian flow on this manifold (cf. Arnold [7, Ch. 10]); that is, if the manifold $\mathbf{M_f}$ is compact and connected, it is diffeomorphic to the n-dimensional torus

$$\mathbb{T}^n = \{(\varphi_1, \dots, \varphi_n) \mod 2\pi\}.$$

Moreover, the phase flow with Hamiltonian $F_1 = H$ determines a quasi-periodic motion on $\mathbf{M_f}$, i.e. in angular coordinates $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_n)$ we have

$$\frac{d\boldsymbol{\varphi}}{dt} = \boldsymbol{\omega}, \quad \boldsymbol{\omega} = \boldsymbol{\omega}(\mathbf{f}).$$

In a neighborhood of the torus $\mathbf{M_f}$ in phase space $\mathbb{R}^{2n}$ one can construct a particularly convenient set of variables $(\mathbf{I}, \boldsymbol{\varphi})$, the *action-angle variables*, so that in these variables the flow of the Hamiltonian system can be written as

$$\mathbf{I}' = \mathbf{0}, \quad \boldsymbol{\varphi}' = \boldsymbol{\omega}(\mathbf{I}).$$

These variables are canonical, and the Hamiltonian $H(\mathbf{I})$ is a function of the action variables $\mathbf{I}$ only, so that the angle variables $\boldsymbol{\varphi}$ vary with frequency $\boldsymbol{\omega}(\mathbf{I}) = \partial H(\mathbf{I})/\partial \mathbf{I}$.

Action-angle variables are constructed as follows (cf. Arnold [7, Ch. 10]): Let $\gamma_1(\mathbf{f}), \dots, \gamma_n(\mathbf{f})$ be a basis of one-dimensional cycles on the torus $\mathbf{M_f}$. Define the action variables

$$I_i(\mathbf{f}) = \frac{1}{2\pi} \oint_{\gamma_i} \mathbf{p} d\mathbf{q}.$$

In a neighborhood of a point $(\mathbf{q}_0, \mathbf{p}_0)$ this relation is invertible, so that we may define

$$S(\mathbf{q}, \mathbf{I}) = \int_{\mathbf{q}_0}^{\mathbf{q}} \mathbf{p}(\mathbf{I}, \mathbf{q}) d\mathbf{q},$$

where the path of integration lies on the torus $\mathbb{T}_{\mathbf{f}(\mathbf{I})}$. Then S is the *generating function* of the canonical transformation $(\mathbf{p}, \mathbf{q}) \to (\mathbf{I}, \boldsymbol{\varphi})$, i.e.

$$\mathbf{p} = \frac{\partial S}{\partial \mathbf{q}}, \quad \boldsymbol{\varphi} = \frac{\partial S}{\partial \mathbf{I}}.$$

In general it is difficult to compute a generating function for this transformation. However, it may be computed as the solution of a partial differential equation, the *Hamilton-Jacobi equation.* Specifically $S(\mathbf{q}, \boldsymbol{\alpha})$ is the solution to

$$H\left(\mathbf{q}, \frac{\partial S(\mathbf{q}, \boldsymbol{\alpha})}{\partial \mathbf{q}}\right) = h, \tag{3.4}$$

where $\boldsymbol{\alpha} \in \mathbb{R}^n$ are n integration constants. Suppose the Hamilton-Jacobi equation (3.4) is separable, i.e. the Hamiltonian has the form

$$H(\mathbf{q}, \mathbf{p}) = \phi_1\big(q_1, p_1, \phi_2(\mathbf{q}, \mathbf{p}), \dots, \phi_n(\mathbf{q}, \mathbf{p})\big),$$

where the functions ϕ_i are recursively defined by

$$\begin{aligned}
\phi_n(\mathbf{q}, \mathbf{p}) &= \hat{\phi}_n(q_n, p_n), \\
\phi_{n-1}(\mathbf{q}, \mathbf{p}) &= \hat{\phi}_{n-1}\big(q_{n-1}, p_{n-1}, \phi_n(\mathbf{q}, \mathbf{p})\big), \\
\phi_i(\mathbf{q}, \mathbf{p}) &= \hat{\phi}_i\big(q_i, p_i, \phi_{i+1}(\mathbf{q}, \mathbf{p}), \dots, \phi_n(\mathbf{q}, \mathbf{p})\big), \quad i = 2, \dots, n-2.
\end{aligned}$$

Then a solution of the form

$$S(\mathbf{q}, \boldsymbol{\alpha}) = S_1(q_1, \boldsymbol{\alpha}) + \cdots + S_n(q_n, \boldsymbol{\alpha})$$

can be easily found. The integration constants $\boldsymbol{\alpha}$ can then be expressed as functions of the action variables $\mathbf{I}$ through the relations

$$I_i = \frac{1}{2\pi} \oint_{\gamma_i} p_i dq_i = \frac{1}{2\pi} \oint_{\gamma_i} \frac{\partial S(\mathbf{q}, \boldsymbol{\alpha})}{\partial q_i} dq_i,$$

so that a generating function $S(\mathbf{q}, \mathbf{I})$ for the transformation is found.

This technique will be used in Chapter 7 to compute action-angle variables for the Hamiltonian formulation describing the motion of a Lagrange top.

Chapter 4

Hamiltonian Formulations for Elastic Rods and Rigid Bodies

We now present several Hamiltonian formulations governing the equilibria of elastic rods and the motion of a tumbling rigid body. In Section 4.1 we review a non-canonical Hamiltonian formulation for the dynamics of a tumbling body (cf. Novikov [49] or Holmes and Marsden [29]). This formulation has been used by Mielke and Holmes [46] in the context of elastic rods to determine the orientation of the director frame, and they then recover the spatial components of the centerline $\mathbf{r}$ by solving the additional differential equations (2.3), but from a non-Hamiltonian perspective (see also Simo et al. [59] for related Hamiltonian formulations in the context of rod dynamics). In Section 4.2 we present an extended non-canonical Hamiltonian formulation describing the equilibrium conditions of an extensible and shearable rod that includes the body components of both the centerline $\mathbf{r}$ and an additional fixed vector $\mathbf{i}$ in the phase space. To my knowledge this twelve-dimensional bracket has not been mentioned previously, although it is closely related to a nine-dimensional bracket, introduced in the context of satellite dynamics (see Wang et al. [63, 64]). A more classic approach is to obtain a canonical Hamiltonian formulation by combining the spatial components of $\mathbf{r}$ with some parametrization of the rotation group $SO(3)$. Dichmann et al. [16] derive a seven degree-of-freedom system using a parametrization in terms of Euler parameters. Alternatively one may use a parametrization in terms of Euler angles; we review these canonical formulations in Sections 4.3 and 4.4, respectively.

4.1 Non-canonical Formulation for the Dynamics of a Tumbling Rigid Body

In this formulation we rewrite the balance equations (2.23) in body coordinates

$$\mathsf{m}' = \mathsf{m} \times \mathsf{u} + \mathsf{n} \times \mathsf{e}_3, \tag{4.1}$$
$$\mathsf{n}' = \mathsf{n} \times \mathsf{u}, \tag{4.2}$$

where $\mathsf{e}_3 = (0,0,1)^T$ and u and m are related via (2.22). The Hamiltonian structure of this system has been investigated for example in [29, 49]. Specifically a Poisson bracket can be defined in local coordinates by the structure matrix

$$\mathbf{J}(\mathsf{m},\mathsf{n}) = \begin{pmatrix} \mathsf{m}\times & \mathsf{n}\times \\ \mathsf{n}\times & \mathbf{0} \end{pmatrix}, \tag{4.3}$$

where for any triple $\mathsf{p} = (p_1, p_2, p_3)$ we use the notation $\mathsf{p}\times$ to denote the matrix

$$\mathsf{p}\times = \begin{pmatrix} 0 & -p_3 & p_2 \\ p_3 & 0 & -p_1 \\ -p_2 & p_1 & 0 \end{pmatrix},$$

corresponding to the vector product. Then equations (4.1) and (4.2) are equivalent to the Hamiltonian system

$$\begin{pmatrix} \mathsf{m}' \\ \mathsf{n}' \end{pmatrix} = \mathbf{J}(\mathsf{m},\mathsf{n})\nabla H(\mathsf{m},\mathsf{n},s), \tag{4.4}$$

with Hamiltonian

$$H(\mathsf{m},\mathsf{n},s) = \frac{1}{2}\mathsf{m}\cdot\mathbf{K}^{-1}\mathsf{m} + \mathsf{m}\cdot\hat{\mathsf{u}}(s) + \mathsf{n}\cdot\mathsf{e}_3.$$

Equations (4.4) are the Hamiltonian form of the well-known Euler-Poisson equations (4.1) and (4.2) for the motion of a heavy rigid body about a fixed point.

4.2 Non-canonical Formulation for Elastic Rods

We now present a novel approach in which the full equations governing the equilibrium conditions of elastic rods are rewritten entirely in terms of body variables (i.e. components with respect to the director frame $\mathcal{D}$) to extend the non-canonical rigid body formulation of the previous section to a twelve-dimensional non-canonical Hamiltonian system. Assume that the force $\mathbf{n} \neq \mathbf{0}$, let $\mathbf{k} \equiv \mathbf{n}/|\mathbf{n}|$, let $\mathbf{i} \in \mathbb{E}^3$ be a fixed unit vector, and define $\mathbf{j} \equiv (\mathbf{k}\times\mathbf{i})/|\mathbf{k}\times\mathbf{i}|$. Then $\{\mathbf{i},\mathbf{j},\mathbf{k}\}$ define a basis of $\mathbb{E}^3$ provided that $\mathbf{i}$ and $\mathbf{n}$ are non-collinear.

We rewrite the differential equations (2.2) and (2.4) in terms of body coordinates to obtain

$$\mathsf{r}' = \mathsf{r}\times\mathsf{u} + \mathsf{v}, \tag{4.5}$$
$$\mathsf{m}' = \mathsf{m}\times\mathsf{u} + \mathsf{n}\times\mathsf{v}, \tag{4.6}$$
$$\mathsf{n}' = \mathsf{n}\times\mathsf{u}. \tag{4.7}$$

To complete the description we also rewrite the trivial equation $\mathbf{i}' = \mathbf{0}$ in body coordinates, i.e.

$$\mathsf{i}' = \mathsf{i}\times\mathsf{u}, \tag{4.8}$$

to obtain an additional Poisson equation. The constitutive relations (2.6) relate u and v to m and n. They can be inverted to yield

$$\begin{aligned} \mathsf{u} &= \frac{\partial W^*(\mathsf{m},\mathsf{n},s)}{\partial \mathsf{m}} + \hat{\mathsf{u}}, \\ \mathsf{v} &= \frac{\partial W^*(\mathsf{m},\mathsf{n},s)}{\partial \mathsf{n}} + \hat{\mathsf{v}}, \end{aligned}$$

where $W^*(\mathsf{m},\mathsf{n},s)$ is the Legendre transform of $W(\mathsf{u},\mathsf{v},s)$ (for a definition and properties of the Legendre transform see Appendix B). Equations (4.5)–(4.8) can then be seen to possess a Hamiltonian structure. Define the Hamiltonian

$$H(\mathsf{i},\mathsf{r},\mathsf{m},\mathsf{n},s) = W^*(\mathsf{m},\mathsf{n},s) + \mathsf{m}\cdot\hat{\mathsf{u}}(s) + \mathsf{n}\cdot\hat{\mathsf{v}}(s). \tag{4.9}$$

Then Hamilton's equations are of the form

$$\begin{pmatrix} \mathsf{i}' \\ \mathsf{r}' \\ \mathsf{m}' \\ \mathsf{n}' \end{pmatrix} = \mathbf{J}(\mathsf{i},\mathsf{r},\mathsf{m},\mathsf{n})\nabla H(\mathsf{i},\mathsf{r},\mathsf{m},\mathsf{n},s), \tag{4.10}$$

where the skew-symmetric operator $\mathbf{J}$ is defined by

$$\mathbf{J}(\mathsf{i},\mathsf{r},\mathsf{m},\mathsf{n}) = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathsf{i}\times & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathsf{r}\times & \mathbf{I} \\ \mathsf{i}\times & \mathsf{r}\times & \mathsf{m}\times & \mathsf{n}\times \\ \mathbf{0} & -\mathbf{I} & \mathsf{n}\times & \mathbf{0} \end{pmatrix}. \tag{4.11}$$

In order that $\mathbf{J}$ actually generate a Poisson bracket the Jacobi identity must be verified (cf. Olver [51, Proposition 6.8]). The necessary calculations are computationally tedious, but have been directly checked using the symbolic manipulation package Mathematica (cf. Appendix G).

For the non-canonical formulation to be meaningful it is important that the fixed frame components of all variables can be determined from the phase variables i, r, m and n. We now show that this is indeed the case. For convenience we assume that the fixed frame $\{\mathbf{i},\mathbf{j},\mathbf{k}\}$ is orthonormal. The rotation matrix can then be constructed in terms of the nine direction cosines relating the fixed frame $\{\mathbf{i},\mathbf{j},\mathbf{k}\}$ and the director frame $\{\mathbf{d}_i\}_{i=1}^3$. We write the direction cosines in the form

$$\mathbf{R} = \begin{pmatrix} \mathbf{d}_1\cdot\mathbf{i} & \mathbf{d}_1\cdot\mathbf{j} & \mathbf{d}_1\cdot\mathbf{k} \\ \mathbf{d}_2\cdot\mathbf{i} & \mathbf{d}_2\cdot\mathbf{j} & \mathbf{d}_2\cdot\mathbf{k} \\ \mathbf{d}_3\cdot\mathbf{i} & \mathbf{d}_3\cdot\mathbf{j} & \mathbf{d}_3\cdot\mathbf{k} \end{pmatrix} = (\mathsf{i},\mathsf{j},\mathsf{k}) = \left(\mathsf{i}, \frac{\mathsf{n}}{|\mathsf{n}|}\times\mathsf{i}, \frac{\mathsf{n}}{|\mathsf{n}|}\right).$$

Then the triple p of body components of any vector is related to the triple $\underline{\mathbf{p}}$ of components in the fixed frame $\{\mathbf{i},\mathbf{j},\mathbf{k}\}$ through the relation $\mathsf{p} = \mathbf{R}\underline{\mathbf{p}}$ or equivalently $\underline{\mathbf{p}} = \mathbf{R}^T\mathsf{p}$.

4.3 Canonical Euler Parameter Formulation for Elastic Rods

The four parameter description of $SO(3)$ has been exploited by Dichmann et al. [16] to formulate a seven degree-of-freedom Hamiltonian system governing the equilibria of elastic rods, which includes the spatial components of the centerline $\mathbf{r}$ and the force $\mathbf{n}$. The necessary derivation is not quite standard because the presence of four parameters means that the appropriate Lagrangian is not a strictly convex function of the generalized velocities.

The Hamiltonian system is derived via a constrained variational principle. The equilibria of hyperelastic rods are constrained critical points of the energy functional

$$\int_0^1 W(\mathsf{u}-\hat{\mathsf{u}}, \mathsf{v}-\hat{\mathsf{v}}, s)ds,$$

subject to the differentiated Euler parameter norm constraint (2.20):

$$\mathbf{q}\cdot\mathbf{q}' = 0. \tag{4.12}$$

Equivalently the equilibria are critical points of the Lagrangian

$$\mathcal{L}(\mathbf{r},\mathbf{r}',\mathbf{q},\mathbf{q}',s) = W(\mathsf{u}-\hat{\mathsf{u}}, \mathsf{v}-\hat{\mathsf{v}}, s) + \nu\ \mathbf{q}\cdot\mathbf{q}',$$

where the Lagrange multiplier $\nu \in \mathbb{R}$ enforces the constraint (4.12). We define conjugate variables $\mathbf{n}$ and $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)$ through

$$\begin{aligned}
\mathbf{n} &= \frac{\partial \mathcal{L}}{\partial \mathbf{r}'} = \frac{\partial W}{\partial z_j}(\mathsf{u}-\hat{\mathsf{u}}, \mathsf{v}-\hat{\mathsf{v}}, s)\mathbf{d}_j,\\
\boldsymbol{\mu} &= \frac{\partial \mathcal{L}}{\partial \mathbf{q}'} = \frac{\partial W}{\partial w_j}(\mathsf{u}-\hat{\mathsf{u}}, \mathsf{v}-\hat{\mathsf{v}}, s)\frac{2\mathbf{B}_j\mathbf{q}}{|\mathbf{q}|^2} + \nu\mathbf{q}
\end{aligned}$$

adopting the definition for the matrices $\mathbf{B}_j$ in (2.21). In particular, it can be seen that the conjugate variable $\mathbf{n}$ is merely the force vector introduced in Section 2.1.

In terms of these variables the moments are

$$m_i = \frac{1}{2}\boldsymbol{\mu}\cdot\mathbf{B}_i\mathbf{q}, \tag{4.13}$$

and the shear forces are

$$n_i = \mathbf{n}\cdot\mathbf{d}_i. \tag{4.14}$$

It can be verified that the full rod equilibrium conditions can be recast as the seven degree-of-freedom, canonical Hamiltonian system

$$\begin{pmatrix}\mathbf{r}'\\ \mathbf{q}'\\ \mathbf{n}'\\ \boldsymbol{\mu}'\end{pmatrix} = \mathbf{J}_{14}\nabla H(\mathbf{r},\mathbf{q},\mathbf{n},\boldsymbol{\mu},s), \tag{4.15}$$

where $\mathbf{J}_{14}$ is the standard structure matrix defined in (3.1). Furthermore the Hamiltonian is

$$H(\mathbf{r}, \mathbf{q}, \mathbf{n}, \boldsymbol{\mu}, s) = W^*(\mathsf{m}, \mathsf{n}, s) + \mathsf{m} \cdot \hat{\mathsf{u}}(s) + \mathsf{n} \cdot \hat{\mathsf{v}}(s),$$

where the variables m and n are expressed in terms of $\mathbf{q}$, $\boldsymbol{\mu}$ and $\mathbf{n}$ via (4.13) and (4.14).

Now suppose the rod is inextensible and unshearable. In this case the construction of the Hamiltonian system via a variational principle is more complicated; due to the constraint (2.5) additional multipliers arise in the Lagrangian formulation. A derivation of the corresponding Hamiltonian system is detailed in Li and Maddocks [39]. It is given by system (4.15), where the Hamiltonian is now

$$H(\mathbf{r}, \mathbf{q}, \mathbf{n}, \boldsymbol{\mu}, s) = W^*(\mathsf{m}, s) + \mathsf{m} \cdot \hat{\mathsf{u}} + n_3,$$

where the moments m are defined in (4.13) and $n_3 = \mathbf{n} \cdot \mathbf{d}_3$. Thus this formulation is obtained by simply replacing the Legendre transform of the general strain-energy function $W^*(\mathsf{m}, \mathsf{n}, s)$ by $W^*(\mathsf{m}, s)$ and $\hat{\mathsf{v}}$ by e_3

Alternatively, the correctness of this Hamiltonian formulation can be verified by checking that the system is equivalent to the kinematic equations (2.1), the balance laws (2.4), and the inextensibility and unshearability constraint (2.5), taken together with the constitutive relation (2.8).

4.4 Canonical Euler Angle Formulation for Elastic Rods

One may also adopt a parametrization of $SO(3)$ in terms of Euler angles to obtain a six degree-of-freedom canonical Hamiltonian system. Now the equilibria of elastic rods are unconstrained critical points of the energy functional

$$\int_0^1 W(\mathsf{u} - \hat{\mathsf{u}}, \mathsf{v} - \hat{\mathsf{v}}, s)\, ds,$$

where the strains u are expressed in terms of the Euler angles $\boldsymbol{\phi} = (\phi, \theta, \psi)$ and their derivatives according to (2.19).

Define conjugate variables $\mathbf{n}$ and $\mathbf{p} = (p_\phi, p_\theta, p_\psi)$ by

$$\begin{aligned}
\mathbf{n} &= \frac{\partial W}{\partial \mathbf{r}'} = \frac{\partial W}{\partial z_j}(\mathsf{u} - \hat{\mathsf{u}}, \mathsf{v} - \hat{\mathsf{v}}, s)\mathbf{d}_j, \\
\mathbf{p} &= \frac{\partial W}{\partial \boldsymbol{\phi}'} = \frac{\partial W}{\partial w_j}(\mathsf{u} - \hat{\mathsf{u}}, \mathsf{v} - \hat{\mathsf{v}}, s)\mathbf{b}^j(\boldsymbol{\phi})
\end{aligned}$$

with $\{\mathbf{b}^j(\boldsymbol{\phi})\}_{j=1}^3$ defined in (2.18). For example Holmes and Marsden [29] show how to obtain a three degree-of-freedom Hamiltonian formulation in terms of the variables $(\boldsymbol{\phi}, \mathbf{p})$ for the reduced problem describing rigid body dynamics, and this description is easily extended to obtain the six degree-of-freedom canonical Hamiltonian formulation in terms of the variables $(\mathbf{r}, \boldsymbol{\phi}, \mathbf{n}, \mathbf{p})$:

$$\begin{pmatrix} \mathbf{r}' \\ \boldsymbol{\phi}' \\ \mathbf{n}' \\ \mathbf{p}' \end{pmatrix} = \mathbf{J}_{12} \nabla H(\mathbf{r}, \boldsymbol{\phi}, \mathbf{n}, \mathbf{p}, s). \tag{4.16}$$

The Hamiltonian

$$H(\mathbf{r}, \boldsymbol{\phi}, \mathbf{n}, \mathbf{p}, s) = W^*(\mathsf{m}, \mathsf{n}, s) + \mathsf{m} \cdot \hat{\mathsf{u}}(s) + \mathsf{n} \cdot \hat{\mathsf{v}}(s),$$

is merely re-expressed in the new variables $(\mathbf{r}, \boldsymbol{\phi}, \mathbf{n}, \mathbf{p})$. In particular, the directors $\mathbf{d}_i$ have the form (2.17) and the moments and shear forces are

$$\begin{aligned} m_j &= \mathbf{p} \cdot \mathbf{b}_j(\boldsymbol{\phi}), && (4.17) \\ n_j &= \mathbf{n} \cdot \mathbf{d}_j, && (4.18) \end{aligned}$$

where $\{\mathbf{b}_j\}_{j=1}^3$ is the dual basis to the basis $\{\mathbf{b}^j\}_{j=1}^3$, i.e.

$$\begin{aligned} \mathbf{b}_1(\boldsymbol{\phi}) &= \begin{pmatrix} -\cos\psi/\sin\theta \\ \sin\psi \\ \cos\psi\cos\theta/\sin\theta \end{pmatrix}, \\ \mathbf{b}_2(\boldsymbol{\phi}) &= \begin{pmatrix} \sin\psi/\sin\theta \\ \cos\psi \\ -\sin\psi\cos\theta/\sin\theta \end{pmatrix}, \\ \mathbf{b}_3(\boldsymbol{\phi}) &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \end{aligned}$$

Consider the special case of a linearly elastic, inextensible, unshearable, isotropic rod with diagonal strain-energy function (2.11). If we assume that the rod is untwisted in its minimum energy configuration (i.e. $\hat{u}_3 = 0$), we have

$$H(\mathbf{r}, \boldsymbol{\phi}, \mathbf{n}, \mathbf{p}, s) = \frac{1}{2K_1 \sin^2\theta}\left[(p_\phi - p_\psi \cos\theta)^2 + p_\theta^2 \sin^2\theta\right] + \frac{1}{2K_3} p_\psi^2 + \mathbf{n} \cdot \mathbf{d}_3, \tag{4.19}$$

and this is the formulation that will be exploited to obtain a Hamiltonian system in terms of action-angle variables in Section 7.1.

In any of the Hamiltonian systems described in this chapter the centerline $\mathbf{r}$ represents three ignorable (or cyclic) variables, i.e. the variable $\mathbf{r}$ does not appear explicitly in the Hamiltonian. Accordingly the conjugate variables $\mathbf{n}$

automatically provide three commuting integrals that may therefore be used to reduce the dimension of the Hamiltonian system by six. Thus, for prescribed values of $\mathbf{n}$, the reduction to the frame variables inherent in the Kirchhoff Kinetic Analogy is entirely standard from the Hamiltonian perspective. Notice however that with the Euler parameter representation the reduced system has four degrees of freedom, whereas a parametrization in terms of Euler angles would lead to a reduced system with three degrees of freedom.

Chapter 5

Integrals and Completely Integrable Systems

The search for integrals (or invariants) of a system of ordinary differential equations is important, since the existence of integrals essentially reduces the order of the system of equations and the resulting lower-order system is typically easier to analyze. If for an m-dimensional system of differential equations, m linearly independent integrals are found, the system is completely integrable, so that at least in principle, a solution can be constructed through quadratures. For Hamiltonian systems the situations is special in the sense that a fewer number of integrals can guarantee complete integrability; it is necessary however, that these integrals satisfy an additional commutativity condition. In addition, for non-canonical Hamiltonian systems trivial integrals (or Casimirs) may exist, which arise independent of the Hamiltonian under study (recall Section 3.2).

The existence of integrals and associated integrable systems have been described for the equations of motion of a tumbling rigid body for example by Arnold [7]. Antman [2, Ch. IX] recognizes, from a non-Hamiltonian perspective, that the equations describing uniform and isotropic, hyperelastic rods are integrable. Detailed discussions about invariants for Hamiltonian formulations in rod mechanics can be found for example in [16, 39, 42], although the requisite calculations have been only carried out for inextensible and unshearable, linearly elastic rods. Most pertinently, Li and Maddocks [39] observed the existence of additional integrals in the Euler parameter Hamiltonian formulation, and we exploit this fact in our discussion of the last quadratures in Chapter 6.

We have carried out the calculations necessary to prove the existence of integrals and associated completely integrable systems for the non-canonical and Euler parameter description of general hyperelastic rods. Although these computations may be cumbersome, they have been efficiently implemented using the symbolic manipulation package Mathematica (see Appendix G). This software is capable of handling the calculations even for general constitutive functions of the form $W(\mathsf{u}, \mathsf{v})$. In particular, for an isotropic constitutive relation we use the representation (2.14) of the strain-energy function.

Considering the balance laws (2.4) it is apparent that the vector functions

$$\mathbf{m} + \mathbf{r} \times \mathbf{n} \tag{5.1}$$

and

$$\mathbf{n} \tag{5.2}$$

must be integrals of any of the Hamiltonian systems describing the equilibria of elastic rods. We will review how these (and additional) integrals arise in each of the formulations.

5.1 Integrals for the Non-Canonical Formulation

Consider the non-canonical Hamiltonian system (4.10). We need to distinguish trivial integrals that arise due to degeneracies of the operator $\mathbf{J}$ (i.e. Casimirs) from integrals that depend on the specific Hamiltonian under study. In our case the operator (4.11) has a two-dimensional null-space, spanned by the vectors

$$\nabla C_1 = \begin{pmatrix} \mathsf{m} + \mathsf{r} \times \mathsf{n} \\ \mathsf{n} \times \mathsf{i} \\ \mathsf{i} \\ \mathsf{i} \times \mathsf{r} \end{pmatrix}, \quad \nabla C_2 = \begin{pmatrix} \mathsf{i} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

The associated distinguished integrals are $C_1 = \mathsf{i} \cdot (\mathsf{m} + \mathsf{r} \times \mathsf{n})$ and $C_2 = \frac{1}{2}|\mathsf{i}|^2$; the integral C_2 arises since $\mathbf{i}$ is a fixed vector, the integral C_1 is the component of (5.1) with respect to $\mathbf{i}$ written in body coordinates.

Consider a general Hamiltonian $\mathcal{H}(\mathsf{m}, \mathsf{n}, s)$ dependent only on the combinations of variables m and n (and s). Then it is easy to verify that the components of $\mathbf{m} + \mathbf{r} \times \mathbf{n}$ with respect to $\mathbf{n}$ and $\mathbf{j}$ are integrals with respect to $\mathcal{H}$; in body coordinates these integrals are $\mathsf{m} \cdot \mathsf{n}$ and $(\mathsf{n} \times \mathsf{i}) \cdot (\mathsf{m} + \mathsf{r} \times \mathsf{n})$. Furthermore, the components of the constant force vector $\mathbf{n}$ in the basis $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ are trivially integrals of $\mathcal{H}$; in body coordinates $|\mathsf{n}|^2$ and $\mathsf{n} \cdot \mathsf{i}$ are the only independent integrals. In the absence of further symmetry properties of the constitutive function, and therefore of the Hamiltonian, the six quantities (5.1) and (5.2) constitute all known integrals. However if the rod is isotropic, the twisting moment

$$m_3 \tag{5.3}$$

is also an integral. Furthermore, if the rod is uniform the Hamiltonian becomes autonomous and so itself provides an additional integral.

5.2 Integrals for Canonical Formulations

The integrals discussed in the previous section arise independent of the particular variables adopted; they must also persist in the canonical Hamiltonian formulations (4.15) and (4.16) written in terms of either Euler parameters or Euler angles. The integrals (5.1) and (5.2) merely need to be re-expressed in terms of the Euler parameters $\mathbf{q}$ and their conjugate momenta $\boldsymbol{\mu}$ (or in terms of Euler angles $\boldsymbol{\phi}$ and the conjugate momenta $\mathbf{p}$). In particular, the components of the moment vector $\mathbf{m}$ relative to the inertial basis $\mathcal{E}$ are

$$\underline{m}_j = \frac{1}{2}\boldsymbol{\mu} \cdot \mathbf{F}_j \mathbf{q}, \tag{5.4}$$

where (see Dichmann [15])

$$\mathbf{F}_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{F}_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

and

$$\mathbf{F}_3 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

The point of interest here is that the parametrization in terms of Euler parameters gives rise to additional integrals (see Li and Maddocks [39] for more details). First the Euler parameter norm constraint (2.20) is manifested as the integral

$$\mathbf{q} \cdot \mathbf{q}, \tag{5.5}$$

associated with the gauge freedom generated by the absence of strict convexity of the corresponding Lagrangian with respect to the generalized velocities. Second, the function

$$\boldsymbol{\mu} \cdot \mathbf{q} + 2\,\mathbf{r} \cdot \mathbf{n} \tag{5.6}$$

is an integral. Thus in the Euler parameter description (4.15) there are in general always eight integrals (5.1), (5.2), (5.5) and (5.6), and isotropy and uniformity can yield two additional integrals.

5.3 Integrals for Rigid Body Equations

We now specialize to the equations of rigid body motion, and consider the force vector $\mathbf{n}$ as a parameter. Furthermore we assume a constitutive law of the particular form (2.22) as in the Kirchhoff Kinetic Analogy. In the reduced non-canonical formulation (4.4) the operator (4.3) has a two-dimensional null space spanned by the vectors

$$\nabla C_1 = \begin{pmatrix} \mathsf{n} \\ \mathsf{m} \end{pmatrix} \quad \nabla C_2 = \begin{pmatrix} \mathbf{0} \\ \mathsf{n} \end{pmatrix}$$

The associated Casimirs $C_1 = \mathsf{m} \cdot \mathsf{n}$ and $C_2 = \frac{1}{2}|\mathsf{n}|^2$ are the trivial integrals of (4.4) no matter what the Hamiltonian. In the non-canonical and the Euler angle representation, C_1 and C_2 constitute in general all known integrals. In the Euler parameter formulation the additional integral (5.5) remains (while the integral (5.6) does not persist because it involves the variable $\mathbf{r}$). In addition for a symmetric heavy rigid body (Lagrange top) the moment m_3 is an integral. The free rigid body has the integral $|\mathsf{m}|$. And finally, the Hamiltonian is an integral, provided that the system is autonomous (as is often the case).

5.4 Completely Integrable Systems

Finally, we describe the various integrable cases that arise in the above Hamiltonian formulations. One needs to verify the appropriate count on commuting integrals to obtain complete integrability (see Section 3.2). For the non-canonical Hamiltonian system (4.10) there always exist two Casimirs and four integrals, three of which are in involution. Isotropy implies the integral m_3, which is involution with all the others. Finally uniformity implies that the Hamiltonian is an integral, yielding a total of two Casimirs and five integrals in involution. Thus the equilibrium equations for a uniform and isotropic rod are integrable. A parametrization in terms of Euler parameters (respectively Euler angles) always yields eight (six) integrals, five (four) of which are in involution. Isotropy and uniformity each yield an additional integral, which raises the total number of integrals in involution to seven (six), and implies complete integrability for the system. A similar count arises for the reduced Hamiltonian formulations describing tumbling rigid body dynamics. The results are summarized in Tables 5.1 and 5.2.

	non-canonical	Euler angles	Euler parameters
no. of Casimirs	2	0	0
no. of integrals	4(6)	6(8)	8(10)
no. in involution	3(5)	4(6)	5(7)

Table 5.1: Numbers of integrals for Hamiltonian formulations describing elastic rods. The table indicates both the numbers of integrals that always exist and, in parentheses, the numbers that exist in the presence of the additional symmetries of isotropy and uniformity. The cases in parentheses all correspond to completely integrable systems.

	non-canonical	Euler angles	Euler parameters
no. of Casimirs	2	0	0
no. of integrals	0(2)	2(4)	3(5)
no. in involution	0(2)	2(3)	2(4)

Table 5.2: Numbers of integrals for Hamiltonian formulations describing a tumbling rigid body. The table indicates both the numbers of integrals that always exist and, in parentheses, the numbers that exist for the case of a Lagrange top. The parenthetical cases all correspond to completely integrable systems.

Chapter 6

The Last Quadrature

The main focus of this chapter is the investigation of the relationship between the frame evolution and the shape of the centerline of the rod. In its simplest case this relation is manifested in the Kirchhoff Kinetic Analogy (recall the discussion in Section 2.4). It states that for an inextensible, unshearable linearly elastic rod the equations governing the frame evolution are equivalent to the equations of motion of a tumbling rigid body. The centerline is then computed through the three quadratures inherent in the material constraint (2.5):

$$\mathbf{r}' = \mathbf{d}_3, \tag{6.1}$$

once a solution of the frame evolution is known. Equivalently this observation is manifested in the fact that the centerline $\mathbf{r}$ is an ignorable coordinate in any of the canonical Hamiltonian formulations discussed in Sections 4.3 and 4.4. Thus a version of Kinetic Analogy carries over to extensible and shearable hyperelastic rods, and the three quadratures are then described by the kinematic relations (2.2):

$$\mathbf{r}' = v_i \mathbf{d}_i, \tag{6.2}$$

where the strains v_i are related to the body variables m and n through the constitutive relations (2.6). However in this case, the interpretation of the reduced system as describing the motion of a tumbling rigid body no longer holds.

Given a solution to the reduced problem describing the frame evolution, the centerline $\mathbf{r}$ is certainly always computed via the quadratures (6.2). However, we will demonstrate that the use of various integrals allows the components of $\mathbf{r}$ in the fixed frame to be computed directly, without explicitly having to perform the three quadratures implicit in (6.2). In Section 6.2 we first review the manner in which two of the quadratures can always be eliminated. This is closely related to the fact that the centerline $\mathbf{r}$ can be conveniently written in terms of cylindrical coordinates. This calculation has been carried out by Ilyukhin [31] in the context of rods for which the material constraint (6.1) holds, but we demonstrate that the same computations hold for a general extensible and shearable rod.

We then make the observation that the extra information, and extra integrals, in the four degree-of-freedom Euler parameter Hamiltonian formulation of the frame evolution can be exploited to finesse the third and last quadrature (cf. Section 6.3).

In Section 6.4 we further restrict ourselves to study the class of uniform rods, for which some very specific qualitative features can be established.

6.1 Motivation

Our motivation for a careful investigation of the quadratures (6.2) is a thorough understanding of the relationship between the frame evolution and the shape of the centerline $\mathbf{r}$. We remark that, despite the Kinetic Analogy, considerable information concerning the shape of the centerline is lost if these quadratures are ignored.

Consider an inextensible and unshearable rod, so that the quadratures (6.2) reduce to the constraint (6.1). It can be observed numerically that these quadratures often seem to average over the potentially very complicated motion of the *tangent indicatrix* (i.e. the director $\mathbf{d}_3$) to yield a rather regular behavior of the centerline $\mathbf{r}$, at least on relatively long length scales. These striking differences are indicated in Figures 6.1, 6.2 and 6.3, which depict the evolution of both the tangent indicatrix and the corresponding shape of the centerline $\mathbf{r}$ in various integrable and nonintegrable cases. In Figure 6.1 a uniform, isotropic rod is shown. The corresponding Hamiltonian system is integrable and exhibits a typical helix-on-helix structure (see Chapter 7 for a more detailed discussion of uniform and isotropic rods). The corresponding motion of the Lagrange top is indicated by the tangent indicatrix oscillating between two parallel circles on the unit sphere. For a general anisotropic rod this regular behavior of the tangent indicatrix is destroyed, as indicated in Figures 6.2 and 6.3. The tangent indicatrix does not seem to follow any obvious structured pattern. However, despite the irregular local behavior, the corresponding shape of the rod centerline does average on a longer length scale to a helix-on-helix structure. Our consideration of the last quadratures is motivated by a desire to understand this behavior, some aspects of which will be explained in Section 6.4.

6.2 Cylindrical Coordinates for the Centerline

We will show how certain integrals that arise in the Hamiltonian formulation of elastic rods can be exploited to write the centerline of the rod with respect to cylindrical coordinates, independent of whether the problem is completely integrable or not. The analysis we follow here is adopted from Ilyukhin [31, Sec. 1.3].

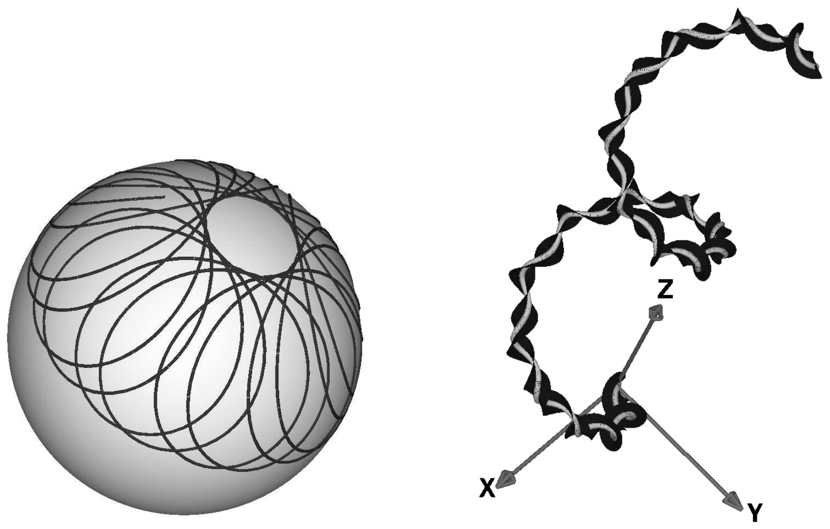

Figure 6.1: A configuration of an integrable rod, and the corresponding tangent indicatrix, tracing the director $\mathbf{d}_3$ on the unit sphere. Both representations are highly structured.

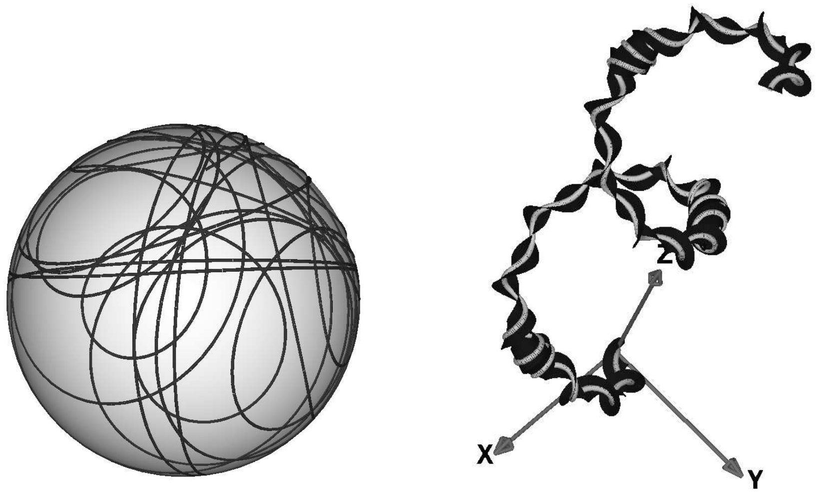

Figure 6.2: The full configuration of a nonintegrable rod exhibits a fairly regular structure, while the tangent indicatrix does not.

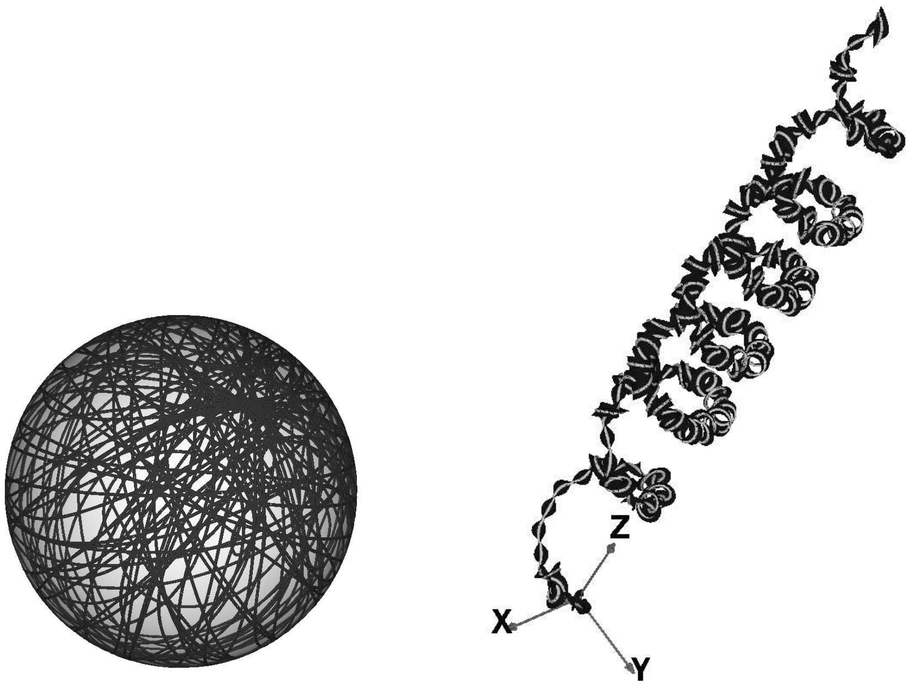

Figure 6.3: On a relatively long length scale, the rod exhibits a regular averaged structure, which is obscured in the tangent indicatrix.

We assume that the nonzero force vector $\mathbf{n}$ is parallel to the unit vector $\mathbf{e}_3$ of the inertial frame $\mathcal{E}$. In this case we can identify the frame $\mathcal{E}$ with the basis $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ introduced in Section 4.2, provided we take $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ to be orthonormal. Ilyukhin shows that the $\mathbf{i}$ and $\mathbf{j}$ components of $\mathbf{r}$ are readily computed using the integral (5.1) arising from the balance equations (2.4). Assume that the value of the integral is $\mathbf{c}$ and rewrite

$$\mathbf{m} + \mathbf{r} \times \mathbf{n} = \mathbf{c},$$

as

$$\mathbf{m} + \left(\mathbf{r} - \frac{\mathbf{n}}{|\mathbf{n}|^2} \times \mathbf{c}\right) \times \mathbf{n} = \frac{\mathbf{c} \cdot \mathbf{n}}{|\mathbf{n}|^2}\mathbf{n} = \frac{\mathbf{m} \cdot \mathbf{n}}{|\mathbf{n}|^2}\mathbf{n},$$

using the fact $\mathbf{m} \cdot \mathbf{n} = \mathbf{c} \cdot \mathbf{n}$. Without loss of generality, we may assume that the fixed coordinate system is translated by a shift $\mathbf{r}_0 = \mathbf{n}/|\mathbf{n}|^2 \times \mathbf{c}$, so that

$$\mathbf{m} + \mathbf{r} \times \mathbf{n} = \frac{\mathbf{m} \cdot \mathbf{n}}{|\mathbf{n}|^2}\mathbf{n}. \tag{6.3}$$

The components of $\mathbf{r}$ with respect to $\mathbf{i}$ and $\mathbf{j}$ are now readily computed in terms of the variables $\mathbf{m}$ and $\mathbf{n}$

$$\mathbf{r}\cdot\mathbf{i} = \frac{\mathbf{m}\cdot\mathbf{j}}{|\mathbf{n}|} = \frac{\mathsf{m}\cdot\mathsf{j}}{|\mathsf{n}|}, \tag{6.4}$$

$$\mathbf{r}\cdot\mathbf{j} = -\frac{\mathbf{m}\cdot\mathbf{i}}{|\mathbf{n}|} = -\frac{\mathsf{m}\cdot\mathsf{i}}{|\mathsf{n}|}. \tag{6.5}$$

This shows that the projection of $\mathbf{r}$ into the $\mathbf{ij}$-plane is merely a rotation and scaling of the moment vector $\mathbf{m}$ so that no integrations have to be performed to compute the two components of $\mathbf{r}$ in the $\mathbf{ij}$-plane.

We now explicitly introduce cylindrical coordinates (ρ, α, ξ) relative to $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, i.e.

$$\mathbf{r}\cdot\mathbf{i} = \rho\cos\alpha, \tag{6.6}$$

$$\mathbf{r}\cdot\mathbf{j} = \rho\sin\alpha, \tag{6.7}$$

$$\mathbf{r}\cdot\mathbf{k} = \xi.$$

Lemma 6.1. *Cylindrical coordinates (ρ, α, ξ) for the centerline $\mathbf{r}$ are defined through the relations*

$$\begin{aligned}
\rho^2 &= \frac{1}{|\mathbf{n}|^4}|\mathbf{m}\times\mathbf{n}|^2, \\
\alpha' &= \frac{1}{|\mathbf{n}|^3\rho^2}\{|\mathbf{n}|^2 m_i v_i - (\mathbf{m}\cdot\mathbf{n})n_i v_i\}, \\
\xi &= \mathbf{r}\cdot\frac{\mathbf{n}}{|\mathbf{n}|}.
\end{aligned}$$

In particular, for an inextensible, unshearable rod the angular component α is given by the identity

$$\alpha' = \frac{1}{|\mathbf{n}|^3\rho^2}\{|\mathbf{n}|^2 m_3 - (\mathbf{m}\cdot\mathbf{n})n_3\}.$$

Proof. The relations can be found in Ilyukhin [31] for inextensible, unshearable rods. The same computations hold for general hyperelastic rods.

We prove these relations utilizing (6.3). The radial component ρ of the projection is immediately computed as

$$\begin{aligned}
\rho^2 &= \frac{1}{|\mathbf{n}|^2}|\mathbf{r}\times\mathbf{n}|^2 \\
&= \frac{1}{|\mathbf{n}|^2}\left|\frac{\mathbf{m}\cdot\mathbf{n}}{|\mathbf{n}|^2}\mathbf{n} - \mathbf{m}\right|^2 \\
&= \frac{1}{|\mathbf{n}|^4}|\mathbf{m}\times\mathbf{n}|^2.
\end{aligned}$$

To obtain the expression for α', differentiate the defining relations (6.6) and (6.7) to obtain

$$\begin{aligned}\mathbf{r}'\cdot\mathbf{i} &= \rho'\cos\alpha - \rho\alpha'\sin\alpha,\\ \mathbf{r}'\cdot\mathbf{j} &= \rho'\sin\alpha + \rho\alpha'\cos\alpha,\end{aligned}$$

which yields after some manipulations

$$\begin{aligned}\alpha' &= \frac{1}{\rho^2}\Big((\mathbf{r}'\cdot\mathbf{j})(\mathbf{r}\cdot\mathbf{i}) - (\mathbf{r}'\cdot\mathbf{i})(\mathbf{r}\cdot\mathbf{j})\Big)\\ &= (\mathbf{r}'\times\mathbf{r})\cdot(\mathbf{j}\times\mathbf{i})\\ &= -\frac{1}{|\mathbf{n}|^2}(\mathbf{r}\times\mathbf{n})\cdot\mathbf{r}'.\end{aligned}$$

Using relation (6.3) and kinematic equation (6.2) yields the desired form for α'. Replacing (6.2) by the constraint (6.1) we verify the relation for inextensible and unshearable rods. □

6.3 The Last Quadrature Resolved

The goal of this section is to compute the components of the centerline in the fixed basis $\{\mathbf{i},\mathbf{j},\mathbf{k}\}$. In our treatment of cylindrical coordinates in Section 6.2 we have reviewed how the $\mathbf{i}$ and $\mathbf{j}$ components of $\mathbf{r}$ are readily computed using the integral (5.1) arising from the balance equations (2.4), in particular

$$\mathbf{r}\cdot\mathbf{i} = \frac{\mathbf{m}\cdot\mathbf{j}}{|\mathbf{n}|} = \frac{\mathsf{m}\cdot\mathsf{j}}{|\mathsf{n}|}, \tag{6.8}$$

$$\mathbf{r}\cdot\mathbf{j} = -\frac{\mathbf{m}\cdot\mathbf{i}}{|\mathbf{n}|} = -\frac{\mathsf{m}\cdot\mathsf{i}}{|\mathsf{n}|}. \tag{6.9}$$

This shows that the three quadratures inherent in (6.1) are reduced to the last quadrature

$$\mathbf{r}'\cdot\mathbf{k} = \mathbf{d}_3\cdot\mathbf{k} = \frac{n_3}{|\mathsf{n}|}.$$

Our observation is that if a solution to the reduced frame evolution is known in terms of Euler parameters, then, whether or not the rod is integrable, the last quadrature can be performed explicitly. The solution of the frame equations can be either obtained in a strictly Hamiltonian form involving Euler parameters or by solving the corresponding Euler-Lagrange equations and explicitly evaluating the Lagrange multiplier associated with the differentiated Euler parameter norm constraint (4.12).

Consider the specific representation of the Hamiltonian system (4.15) in terms of Euler parameters. It is only here that the integral (5.6) arises. Assume that the value of this integral is $\boldsymbol{\mu}\cdot\mathbf{q}+2\,\mathbf{r}\cdot\mathbf{n}=c$. Then the last quadrature is resolved as

$$\mathbf{r}\cdot\mathbf{k}=\frac{1}{2|\mathbf{n}|}(\boldsymbol{\mu}\cdot\mathbf{q}-c). \tag{6.10}$$

The value of c merely represents a shift in origin of $\mathbf{r}$ parallel to $\mathbf{n}$.

We may therefore conclude that the extra degree of freedom in an Euler parameter Hamiltonian description of the equilibria of elastic rods contains additional information that provides a strengthened Kinetic Analogy. Equations (6.8), (6.9) and (6.10) allow a complete reduction of two-point boundary value problems for rods to a two-point boundary value problem for the lower dimensional reduced Euler parameter Hamiltonian description of the frame evolution. Warner [65] has used this result in the numerical solution of the four degree-of freedom boundary value problem with the continuation software package AUTO, and observes an increase in performance by about a factor of three compared to a solution of the full seven degree-of-freedom boundary value problem.

6.4 Uniform Rods

For a uniform rod we can utilize the representation of the centerline in terms of cylindrical coordinates to show that the projection of the centerline into the plane perpendicular to the force is always bounded. We interpret this boundedness as the centerline of the rod evolving between two concentric cylinders. Figures 6.4, 6.5 and 6.6 illustrate this conclusion for various integrable and nonintegrable configurations of linearly elastic, inextensible and unshearable, uniform rods. In each Figure we show a skew projection of the rod together with a projection into the $\mathbf{ij}$-plane. For the integrable case of an isotropic rod, depicted in Figure 6.4 this conclusion was known, as it follows from the quasi-periodic behavior of the solution. However, the more general, nonintegrable case in Figure 6.5, illustrates that, despite the generally complicated behavior of the solution, the projection parallel to the $\mathbf{n}$ axis is bounded, and in fact may resemble that of an isotropic rod.

It is also possible that the radius of the inner cylinder can vanish. In fact this happens exactly if $\mathbf{m}(s)$ is parallel to $\mathbf{n}(s)$ at some value of s (cf. Figure 6.6). Again, in this configuration the skew projection appears rather complicated, while the projection along the $\mathbf{n}$ axis has an almost periodic behavior.

Lemma 6.2. *For a uniform rod the projection of the centerline* $\mathbf{r}$ *into the* $\mathbf{ij}$*-plane is bounded.*

Proof. The proof is a simple consequence of the observations in Section 6.2. The projection of $\mathbf{r}$ into the $\mathbf{ij}$-plane is simply a rotation of the projection of the

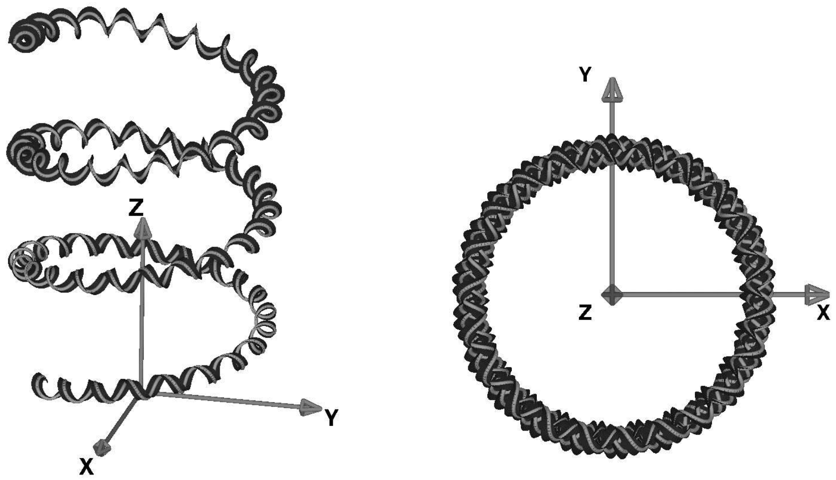

Figure 6.4: An integrable configuration. The centerline of the rod lies between two concentric circular cylinders (as indicated in the projection on the right).

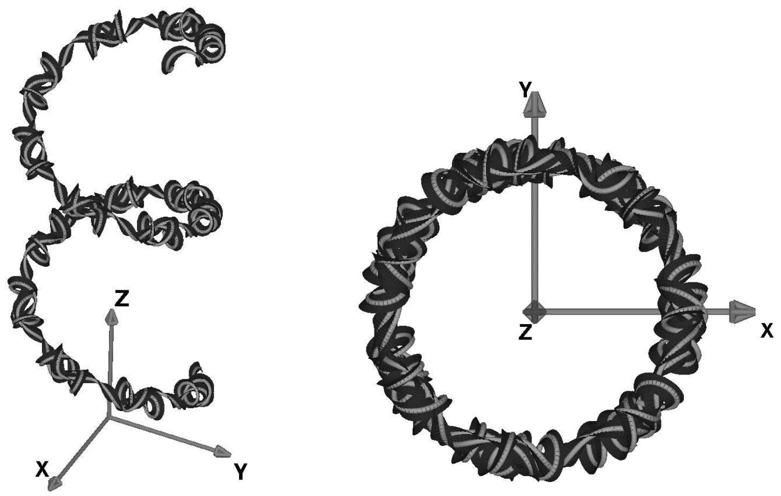

Figure 6.5: For a nonintegrable rod the overall structure becomes rather complicated. However, the existence of the two bounding cylinders persists.

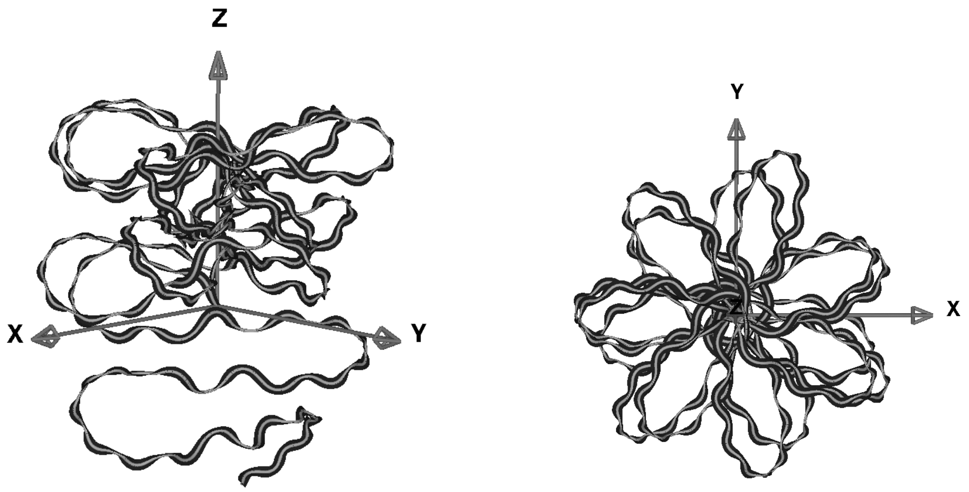

Figure 6.6: A configuration where the radius of the inner cylinder vanishes. The side view suggests a complicated structure, while the projection on the right seems almost periodic.

moment vector $\mathbf{m}$ into the same plane, thus we only need to show that $\mathbf{m}$ is bounded. The assumption of uniformity implies that the Hamiltonian is an integral, i.e. there exists a constant h such that

$$W^*(\mathsf{m},\mathsf{n}) + \mathsf{m}\cdot\hat{\mathsf{u}} + \mathsf{n}\cdot\hat{\mathsf{v}} = h,$$

or equivalently

$$\frac{W^*(\mathsf{m},\mathsf{n})}{|\mathsf{m}|} = \frac{h}{|\mathsf{m}|} - \frac{\mathsf{n}\cdot\hat{\mathsf{v}}}{|\mathsf{m}|} - \frac{\hat{\mathsf{u}}\cdot\mathsf{m}}{|\mathsf{m}|}. \tag{6.11}$$

Now assume that m is unbounded. The coercivity of W (as specified in Section 2.2) is equivalent to the coercivity of W^* (see Evans [19, Sec. 3.3.2]). Hence the left-hand-side of (6.11) becomes unbounded as m becomes unbounded. To the contrary the right-hand-side of (6.11) is bounded for all m, since $\mathbf{n}$ is an integral. With this contradiction we conclude that m must be bounded. □

Now suppose the rod is linearly elastic, inextensible and unshearable. If in the minimum energy configuration the rod is straight and untwisted (i.e. $\hat{\mathsf{u}} = \mathbf{0}$), we obtain more specific bounds on the radii of the two concentric cylinders. The radial component is given by

$$\rho^2 = \frac{1}{|\mathbf{n}|^2}\left(|\mathbf{m}|^2 - \frac{\mathbf{m}\cdot\mathbf{n}}{|\mathbf{n}|^2}\right).$$

Thus bounds on ρ are determined by bounds on $\mathbf{m}$, which we compute using the fact that the Hamiltonian is an integral, i.e.

$$\frac{1}{2}\mathsf{m} \cdot \mathbf{K}\mathsf{m} + \mathbf{n} \cdot \mathbf{d}_3 = h.$$

Denoting by K_{min} and K_{max} the minimum and maximum eigenvalue of $\mathbf{K}$, respectively, we obtain

$$2K_{max}^{-1}(h - |\mathbf{n}|) \leq |\mathbf{m}|^2 \leq 2K_{min}^{-1}(h + |\mathbf{n}|).$$

Chapter 7

Inextensible, Unshearable, Uniform, Isotropic Rods

In this chapter we review the construction of closed-form solutions for inextensible, unshearable, uniform, isotropic, linearly elastic rods by computing a transformation into action-angle variables for the reduced problem describing the frame evolution. Once these variables are constructed, closed-form solutions for the centerline are immediately obtained through quadratures. The treatment of closed-form solutions and action-angle variables will serve two purposes. First, explicit solutions provide insight into qualitative properties of the system that may otherwise be hard to obtain; and this knowledge will be be useful in our analysis of effective rod models in Chapter 8, in particular for the numerical simulations of helical models in Chapter 9. Second, action-angle variables are the main tool for computing perturbation expansions, and they allow us to relate the homogenization results in Chapter 8 to some of the main results in Hamiltonian perturbation theory (cf. Appendix E).

Klein and Sommerfeld [33], for example, obtain closed form solutions for the Euler angle formulation describing the Lagrange top. In addition they investigate the quadratures of these equations in terms of Cayley-Klein parameters, which are closely related to Euler parameters. Recently it has been shown (although from rather different perspectives, e.g. Benham [9], Ilyukhin [31], Langer and Singer [37], Li [38], Shi and Hearst [57]), that the quadratures (2.5) relating the centerline $\mathbf{r}$ of the rod to the motion of the Lagrange top can be carried out explicitly, provided a solution to the reduced integrable problem is known. It should be remarked however, that these closed form solutions do not represent truly explicit solutions of many two-point boundary value problems, because of the presence of undetermined constants. Tan and Witz [61] discuss certain rod boundary value problems that can be solved explicitly, but these cases can in some sense be regarded as exceptional.

A slightly more general approach to integrating Hamiltonian systems is the construction of action-angle variables, from which closed form solutions can readily be obtained. Action-angle variables for the Euler angle formulation governing the motion of a Lagrange top have been discussed by many authors; in particular, they are mentioned extensively in the Russian literature (see Arnold [4], Sergeev [56], Aksenenkova [1]). However, to our knowledge the derivation of this change of variables has not been published in the English language, so we provide a full derivation in Section 7.1.

Using the action-angle variable description of the Lagrange top and the quadratures (2.5), we determine an explicit solution for the centerline of the rod (2.5) in Section 7.2. Moreover, action-angle variables provide a different view on the qualitative behavior of the system; in particular, in Section 7.3 we investigate the helix-on-helix structure of the centerline, which is typical for this integrable system. Finally, in the last section we indicate how special solutions can be obtained, exploiting the knowledge of closed form solutions. In Section 6.4 we have shown that the centerline of the rod is always bounded in the projection into the plane perpendicular to the constant force vector. We show how to obtain conditions on the parameters of the problem to determine solutions for which the centerline is bounded in all directions. These solutions are special in the sense that the centerline of the rod must necessarily lie on a two-dimensional torus of revolution in $\mathbb{E}^3$.

7.1 Action-Angle Variables

The construction of action-angle variables for an integrable Hamiltonian system relies on the assumption that the level sets of the system are compact. This is in general not true for the full system describing rod mechanics, and we therefore focus our attention to the reduced system describing the motion of a Lagrange top. The standard method for the construction of action-angle variables is via the theory of Hamilton-Jacobi, which involves recognizing the system of ordinary differential equations as the characteristic system for a scalar partial differential equation (the Hamilton-Jacobi equation). Then a generating function for the change to action-angle variables is computed as the solution of this partial differential equation. In practice a solution can only be found if the Hamilton-Jacobi equation is separable. It is the formulation in terms of Euler angles that seems to yield the only separable case, provided the (constant) force vector is chosen to lie in the direction of the polar singularity.

Hence we seek action-angle variables for the reduced Hamiltonian system describing the motion of a Lagrange top using the parametrization in terms of Euler angles $\boldsymbol{\phi} = (\phi, \theta, \psi)$ and their conjugate momenta $\mathbf{p} = (p_\phi, p_\theta, p_\psi)$:

$$\begin{aligned} \boldsymbol{\phi}' &= \frac{\partial H(\boldsymbol{\phi}, \mathbf{p})}{\partial \mathbf{p}}, \\ \mathbf{p}' &= -\frac{\partial H(\boldsymbol{\phi}, \mathbf{p})}{\partial \boldsymbol{\phi}}. \end{aligned}$$

This system corresponds to (4.16), where the variables $\mathbf{r}$ and $\mathbf{n}$ are ignored. In particular, the Hamiltonian (4.19) becomes

$$H(\boldsymbol{\phi}, \mathbf{p}) = \frac{1}{2K_1 \sin^2\theta}\left((p_\phi - p_\psi \cos\theta)^2 + p_\theta^2 \sin^2\theta\right) + \frac{1}{2K_3} p_\psi^2 + \mathbf{n}\cdot\mathbf{d}_3,$$

where $\mathbf{n}$ is now treated as a parameter. Here and in the remainder of this chapter we assume that the frame $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ introduced in Section 4.2 is identified with the inertial frame $\mathcal{E}$, so that the potential term $\mathbf{n}\cdot\mathbf{d}_3$ becomes

$$\mathbf{n}\cdot\mathbf{d}_3 = N\mathbf{k}\cdot\mathbf{d}_3 = N\mathsf{e}_3 \cdot \underline{\mathbf{d}}_3 = N\cos\theta,$$

with $N = |\mathbf{n}|$.

Every level set of this system is compact, and is therefore, according to Liouville's Theorem, diffeomorphic to the three-dimensional torus $\mathbb{T}^3$ (compare also Section 3.4). Hence action-angle variables $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \varphi_3)$ and $\mathbf{J} = (J_1, J_2, J_3)$ exist, given by a transformation $(\boldsymbol{\phi}, \mathbf{p}) \to (\boldsymbol{\varphi}, \mathbf{J})$. In the new variables the Hamiltonian has the simple form

$$H(\boldsymbol{\phi}, \mathbf{p}) = \widetilde{H}(\mathbf{J}).$$

A generating function $S(\boldsymbol{\phi}, \mathbf{J})$ for this transformation can be found by solving the Hamilton-Jacobi equation

$$\frac{1}{2K_1 \sin^2\theta}\left[(S_\phi - S_\psi \cos\theta)^2 + S_\theta^2 \sin^2\theta\right] + \frac{1}{2K_3} S_\psi^2 + N\cos\theta = h, \tag{7.1}$$

where h is the value of the integral H. In practice the generating function is computed as a function $S(\boldsymbol{\phi}, \boldsymbol{\alpha})$ of the three integration constants $\boldsymbol{\alpha} \in \mathbb{R}^3$, which can in turn be related to the action variables $\mathbf{J}$ through

$$J_1 = \oint p_\phi d\phi = \oint \frac{\partial S(\boldsymbol{\phi}, \alpha)}{\partial \phi} d\phi, \quad \text{etc.},$$

where the line integral is taken over one cycle on the torus $\mathbb{T}^3$. Having computed the generating function as a function of the action variables $\mathbf{J}$, we obtain the transformation formulas for the angle variables $\boldsymbol{\varphi}$:

$$\boldsymbol{\varphi} = \frac{\partial S}{\partial \mathbf{J}}.$$

We summarize the calculation of action-angle variables in the following theorem (the notation and formulas are closely related to the ones used in Sergeev [56], but the derivation is not presented there).

Theorem 7.1 (Action Angle Variables). *Define*

$$\begin{aligned} f(u) &= K_1(2h - K_3^{-1}p_\psi^2 - 2Nu)(1-u^2) - (p_\phi - p_\psi u)^2 \\ &= (a - bu)(1-u^2) - (p_\phi - p_\psi u)^2 \\ &= b(u-u_1)(u-u_2)(u-u_3), \end{aligned}$$

i.e. $a = K_1(2h - K_3^{-1}p_\psi^2)$ *and* $b = 2K_1N$. *Let* u *and* v *be defined through*

$$u = \cos\theta = u_1 + (u_2 - u_1)\sin^2 v,$$

and denote

$$\epsilon_2^2 = u_2 - u_1, \quad \epsilon_3^2 = u_3 - u_1, \quad k^2 = \frac{\epsilon_2^2}{\epsilon_3^2}, \quad \alpha_\pm = \pm\frac{\epsilon_2^2}{1 \mp u_1}.$$

Under the assumption that $N > 0$ *the three roots* $u_i, i = 1, 2, 3$ *of* $f(u)$ *satisfy*

$$-1 < u_1 \leq u_2 < 1 < u_3.$$

In a neighborhood of the level set

$$\mathbf{M}_{\alpha_\phi,\alpha_\psi,h} = \Big\{(\boldsymbol{\phi}, \mathbf{p})\Big| p_\phi = \alpha_\phi, p_\psi = \alpha_\psi, H(\boldsymbol{\phi}, \mathbf{p}) = h\Big\}$$

the action variables $\mathbf{J}$ *and angle variables* $\boldsymbol{\varphi}$ *are given in terms of the Euler angles* $\boldsymbol{\phi}$ *and their conjugate variables* $\mathbf{p}$ *by*

$$\begin{aligned} J_1 &= p_\phi, \\ J_2 &= \frac{1}{\pi}\int_{u_1}^{u_2} \frac{\sqrt{f(u)}}{1-u^2}du, \\ J_3 &= p_\psi, \end{aligned} \tag{7.2}$$

and

$$\varphi_1 = \phi - \frac{1}{\sqrt{b}\epsilon_3 K}\Big(Q_+(v)K - Q_+(\frac{\pi}{2})F(v,k)\Big), \tag{7.3}$$

$$\varphi_2 = \frac{\pi}{K}F(v,k), \tag{7.4}$$

$$\varphi_3 = \psi + \frac{1}{\sqrt{b}\epsilon_3 K}\Big(Q_-(v)K - Q_-(\frac{\pi}{2})F(v,k)\Big). \tag{7.5}$$

Furthermore Equation (7.4) *can be inverted to yield*

$$u = \cos\theta = u_1 + \epsilon_2^2\mathrm{sn}^2\left(\frac{K\varphi_2}{\pi}, k\right). \tag{7.6}$$

Here $F(\varphi, k)$ and $\Pi(\varphi, \alpha^2, k)$ are elliptic integrals of the first and third kind, respectively. The complete elliptic integral of the first kind is denoted by K and $\mathrm{sn}(u, k)$ is an elliptic function defined as the inverse of the elliptic integral of the first kind (see Appendix C for more details). Furthermore we use the abbreviation

$$Q_{\pm}(v) = \frac{J_1 - J_3}{1 - u_1}\Pi(v, \alpha_+, k) \pm \frac{J_1 + J_3}{1 + u_1}\Pi(v, \alpha_-, k).$$

Proof. Most of the calculations will be carried out in Appendix D. We will here only indicate the main steps and add some remarks.

First, the variables p_ϕ and p_ψ are actually constant, since ϕ and ψ are ignorable coordinates, thus $f(u)$ can be regarded as a function of $u = \cos\theta$ only. Second, a solution to the Hamiltonian system only exists if there is a $u^* = \cos\theta^*$ such that $f(u^*) > 0$. Furthermore $f(\infty) > 0$, $f(-\infty) < 0$, $f(1) < 0$ and $f(-1) < 0$, so we conclude that there exist three real roots u_i of f satisfying

$$-1 < u_1 \leq u_2 \leq 1 < u_3.$$

We start by constructing a solution to the Hamilton-Jacobi equation (7.1). The fact that p_ϕ and p_ψ are cyclic variables immediately implies that the equation is separable and the generating function has the form

$$S(\boldsymbol{\phi}, \mathbf{J}) = p_\phi\phi + S_2(\theta) + p_\psi\psi,$$

where the constants p_ϕ, p_ψ and the function S_2 will be expressed in terms of the action variables $\mathbf{J}$. Using this form of the generating function in (7.1) yields

$$\left(\frac{\partial S_2(\theta)}{\partial\theta}\right)^2 = 2hK_1 - \frac{(p_\phi - p_\psi\cos\theta)^2}{\sin^2\theta} - \frac{K_1}{K_3}p_\psi^2 - 2NK_1\cos\theta.$$

With $u = \cos\theta$ we have

$$\left(\frac{\partial S_2}{\partial u}\right)^2 = \left(\frac{\partial S_2}{\partial\theta}\right)^2\frac{1}{1-u^2} = \frac{f(u)}{(1-u^2)^2}$$

and after integrating

$$S_2(\theta) = \int_{u_1}^{u}\frac{\sqrt{f(v)}}{1-v^2}dv.$$

The generating function S now defines the action variables $\mathbf{J}$:

$$\begin{aligned}
J_1 &= = \frac{1}{2\pi}\oint\frac{\partial S}{\partial\phi}d\phi = p_\phi,\\
J_2 &= \frac{1}{2\pi}\oint\frac{\partial S_2}{\partial\theta}d\theta,\\
J_3 &= \frac{1}{2\pi}\oint\frac{\partial S}{\partial\psi}d\psi = p_\psi.
\end{aligned}$$

We observe that $p_\theta = \frac{\partial S}{\partial \theta}$ evolves over a complete period, as u evolves from u_1 to u_2 and back; hence

$$J_2 = \frac{1}{\pi} \int_{u_1}^{u_2} \frac{\sqrt{f(u)}}{1-u^2} du.$$

Next the transformation formulas for the angle variables are computed through

$$\varphi_i = \frac{\partial S}{\partial J_i}, \quad i = 1, 2, 3.$$

The tedious computations involve elliptic integrals, and are carried out in detail in Appendix D.

Finally Equation (7.4) can be inverted to yield

$$\mathrm{sn}^2 \left(\frac{K\varphi_2}{\pi}, k \right) = \sin^2 v = \frac{u - u_1}{\epsilon_2^2},$$

and we obtain

$$\cos\theta \quad = \quad u_1 + \epsilon_2^2 \mathrm{sn}^2 \left(\frac{K\varphi_2}{\pi}, k \right).$$

□

The action variable J_2 defined in (7.2) is a function of J_1, J_3 and the integral $\widetilde{H} = h$, i.e.

$$J_2 = \widetilde{J}_2(J_1, J_3, h). \tag{7.7}$$

This equation cannot be explicitly solved for h; however, it is possible to determine the frequencies $\boldsymbol{\omega} = \partial\widetilde{H}/\partial\mathbf{I}$ describing the variation of the angle variables $\boldsymbol{\varphi}$ in an explicit form.

Corollary 7.1. *Adopt the notation and assumptions of Theorem 7.1. The Hamiltonian $\widetilde{H}$ is defined as the inverse of (7.7), provided $u_1 \neq u_2$. In this case the motion of a Lagrange top is given in action-angle variables by*

$$J_i' = 0, \qquad \varphi_i' = \omega_i, \qquad i = 1, 2, 3,$$

where the frequencies $\omega_i = \omega_i(\mathbf{J})$ are computed as

$$\begin{aligned}
\omega_1 \quad &= \quad \frac{1}{K_1 K} Q_+(\frac{\pi}{2}), \\
\omega_2 \quad &= \quad \frac{\pi\sqrt{b}\epsilon_3}{2K_1 K}, \\
\omega_3 \quad &= \quad \frac{K_1 - K_3}{K_1 K_3} J_3 - \frac{Q_-(\frac{\pi}{2})}{2K_1 K}.
\end{aligned}$$

Proof. Note that

$$\frac{\partial \widetilde{J}_2}{\partial h} = \frac{K_1}{\pi} \int_{u_1}^{u_2} \frac{1}{\sqrt{f(v)}} dv = 0,$$

if and only if $u_1 = u_2$. Hence by the Implicit Function Theorem [28, Sec. 169] the inverse $\widetilde{H}$ of $\widetilde{J}_2$ as a function of h exists for all h, provided $u_1 \neq u_2$. The frequencies $\boldsymbol{\omega}$ are computed by using the Inverse Function Theorem [28, Sec. 169]:

$$\begin{aligned}
\omega_1 = \frac{\partial \widetilde{H}}{\partial J_1} &= -\left(\frac{\partial \widetilde{J}_2}{\partial h}\right)^{-1} \frac{\partial \widetilde{J}_2}{\partial J_1}, \\
\omega_2 = \frac{\partial \widetilde{H}}{\partial J_2} &= \left(\frac{\partial \widetilde{J}_2}{\partial h}\right)^{-1}, \\
\omega_3 = \frac{\partial \widetilde{H}}{\partial J_3} &= -\left(\frac{\partial \widetilde{J}_2}{\partial h}\right)^{-1} \frac{\partial \widetilde{J}_2}{\partial J_3}.
\end{aligned}$$

The computation of these quantities is detailed in Appendix D. □

7.2 Closed-Form Solutions for Elastic Rods

A closed-form solution for the equations governing elastic rods is obtained by using the action-angle variable description of Theorem 7.1 and the representation of the centerline in cylindrical coordinates outlined in Lemma 6.1. We remark that closed-form solutions have been obtained before in several different contexts (see for example Langer and Singer [37], Li [38], Shi and Hearst [57])). In particular Benham [9] and Ilyukhin [31] study the Euler angle formulation and integrate the appropriate equations directly in light of the Kirchhoff Kinetic Analogy. Their result is equivalent to the following set of explicit solutions.

Corollary 7.2. *Adopt the notation of Theorem 7.1 and denote Jacobi's elliptic integral of the second kind by E (see Appendix C). Suppose the rod is uniform, isotropic and linearly elastic with strain energy (2.11). If the force $\mathbf{n}$ is non-vanishing and $k^2 \neq 0$, then cylindrical coordinates (ρ, α, ξ) of the centerline $\mathbf{r}$ can be expressed in terms of action-angle variables $\boldsymbol{\varphi}$ and $\mathbf{J}$ through*

$$\begin{aligned}
\xi(s) - \xi(0) &= u_1 s + \epsilon_2^2 \left(\frac{s}{k^2} - \frac{\pi}{k^2 K \omega_2} [E(\gamma(s), k) - E(\gamma(0), k)] \right), \\
\rho^2(s) &= \frac{c - bu}{N^2} = \frac{c - bu_1}{N^2} - \frac{b\epsilon_2^2}{N^2} \mathrm{sn}^2 \left(\frac{K \varphi_2(s)}{\pi}, k \right), \\
\alpha(s) - \alpha(0) &= \frac{N}{b} J_1 s + \frac{N}{b} \frac{J_3 b - J_1 c}{c - bu_1} \frac{\pi}{K\omega_2} \left[\Pi\Big(\gamma(s), \beta, k\Big) - \Pi\Big(\gamma(0), \beta, k\Big) \right],
\end{aligned}$$

where

$$c = a + J_3^2 - J_1^2, \quad \beta = \frac{b\epsilon_2^2}{c - bu_1}, \quad \gamma(s) = \mathrm{am}\left(\frac{K\varphi_2(s)}{\pi}, k\right).$$

Proof. Utilizing the representation of $\mathbf{r}$ in cylindrical coordinates (cf. Lemma 6.1) and the relation between Euler angles and action-angle variables we obtain the differential equations

$$\xi' = u_1 + \epsilon_2^2 \mathrm{sn}^2\left(\frac{K\varphi_2}{\pi}, k\right), \tag{7.8}$$

$$\rho^2 = \frac{c - bu}{N^2} = \frac{c - bu_1}{N^2} - \frac{b\epsilon_2^2}{N^2}\mathrm{sn}^2\left(\frac{K\varphi_2(s)}{\pi}, k\right), \tag{7.9}$$

$$\alpha' = \frac{N}{b}\left(J_1 + \frac{J_3 b - J_1 c}{c - bu}\right). \tag{7.10}$$

Equation (7.8) is straightforward to verify using the expression (7.6) for $u = \cos\theta$. In order to verify (7.9) we use the fact that the Hamiltonian is an integral, i.e.

$$h = \frac{1}{2}K_1^{-1}|\mathbf{m}|^2 + \frac{1}{2}(K_3^{-1} - K_1^{-1})m_3^2 + N\cos\theta,$$

to find an expression for $|\mathbf{m}|^2$:

$$|\mathbf{m}|^2 = a + J_3^2 - b\cos\theta,$$

using the fact that $m_3 = J_3$. Furthermore, the value of the integral $\mathbf{m}\cdot\mathbf{n}$ is

$$\mathbf{m}\cdot\mathbf{n} = Nm_i\mathbf{d}_i\cdot\mathbf{k} = N\mathbf{m}_i\underline{\mathbf{d}}_i\cdot\mathsf{e}_3 = Np_\phi = NJ_1,$$

using the representation of the directors $\{\underline{\mathbf{d}}_i\}$ and moments m_i in terms of Euler angles (cf. (2.17) and (4.17)). Hence we can express the radial component ρ as

$$\begin{aligned}\rho^2 &= \frac{1}{N^2}\left(|\mathbf{m}|^2 - \frac{(\mathbf{m}\cdot\mathbf{n})^2}{N^2}\right)\\ &= \frac{1}{N^2}(a + J_3^2 - J_1^2 - b\cos\theta)\\ &= \frac{1}{N^2}(c - b\cos\theta),\end{aligned}$$

and (7.9) is obtained by substituting (7.6) for $\cos\theta$. Finally, using previously computed quantities yields (7.10)

$$\begin{aligned}\alpha' &= \frac{1}{N^3\rho^2}\left(N^2 m_3 - (\mathbf{m}\cdot\mathbf{n})n_3\right)\\ &= \frac{N}{b}\left(J_1 + \frac{J_3 b - J_1 c}{c - bu}\right).\end{aligned}$$

Integrating equations (7.8) and (7.10) yields the final result. In particular, we compute

$$\xi(s) - \xi(0) = u_1 s + \epsilon_2^2 \int_0^s \mathrm{sn}^2\left(\frac{K\varphi_2(t)}{\pi}, k\right) dt.$$

We know the solution for the angle variable $\varphi_2(t) = \omega_2 t + \varphi_2(0)$, so using a linear change of variables we obtain

$$\begin{aligned}
\xi(s) - \xi(0) &= u_1 s + \epsilon_2^2 \int_0^s \mathrm{sn}^2\left(\frac{K\varphi_2(t)}{\pi}, k\right) dt \\
&= u_1 s + \epsilon_2^2 \frac{\pi}{K\omega_2} \int_{\frac{K\varphi_2(0)}{\pi}}^{\frac{K\varphi_2}{\pi}} \mathrm{sn}^2(u,k) du \\
&= u_1 s + \epsilon_2^2 \left(\frac{\pi}{K\omega_2} \frac{1}{k^2} \Big[u - E(\mathrm{am}(u,k),k) \Big]_{\frac{K\varphi_{20}}{\pi}}^{\frac{K\varphi_2}{\pi}} \right) \\
&= u_1 s + \epsilon_2^2 \left(\frac{s}{k^2} - \frac{\pi}{k^2 K\omega} \left[E(\gamma,k) - E(\gamma_0,k) \right] \right),
\end{aligned}$$

where the integral is evaluated using (D.5) in Appendix D. Finally we integrate (7.10) to be

$$\alpha(s) - \alpha(0) = \frac{N}{b} J_1 s + \frac{N}{b}(J_3 b - J_1 c) \int_0^s \frac{1}{c - bu} dt.$$

We calculate

$$\begin{aligned}
&\int_0^s \frac{1}{c - bu} dt \\
&\quad = \frac{1}{c - bu_1} \int_0^s \frac{1}{1 - \beta \mathrm{sn}^2(\frac{K\varphi_2}{\pi}, k)} dt \\
&\quad = \frac{1}{c - bu_1} \frac{\pi}{K\omega_2} \int_{\frac{K\varphi_{20}}{\pi}}^{\frac{K\varphi_2}{\pi}} \frac{1}{1 - \beta \mathrm{sn}^2(u,k)} du \\
&\quad = \frac{1}{c - bu_1} \frac{\pi}{K\omega_2} (\Pi(\gamma, \beta, k) - \Pi(\gamma_0, \beta, k)),
\end{aligned}$$

from which the expression for α immediately follows. □

In the following sections we will exploit this representation of the centerline to establish some interesting qualitative features of inextensible, unshearable uniform isotropic rods.

7.3 Helix-on-Helix Structure

We have established in Section 7.1 that the motion of a Lagrange top is diffeomorphic to a linear flow on a torus and therefore quasi-periodic. We wish to interpret the quasi-periodic behavior of the frame evolution for the centerline $\mathbf{r}$. In fact using the constraint (2.5) and the representation of $\mathbf{r}$ in cylindrical coordinates immediately gives us the structure of any solution of the equations governing inextensible, unshearable uniform isotropic rods, provided the force vector $\mathbf{n}$ is non-vanishing.

Corollary 7.3. *Under the assumptions of Corollary 7.2, the centerline* $\mathbf{r}$ *exhibits a helix-on-helix structure in the following sense. The cylindrical coordinates* (ρ, α, ξ) *can be written in the form*

$$\xi(s) - \xi(0) = \overline{\xi}(s) + \widetilde{\xi}(s), \quad \rho^2(s) = \overline{\rho}^2(s) + \widetilde{\rho}^2(s), \quad \alpha(s) - \alpha(0) = \overline{\alpha}(s) + \widetilde{\alpha}(s),$$

where $\overline{\xi}'' = \overline{\alpha}'' = (\overline{\rho}^2)' = 0$ *and* $\widetilde{\xi}$, $\widetilde{\rho}^2$ *and* $\widetilde{\alpha}$ *are periodic. Specifically we have*

$$\begin{aligned}
\overline{\xi} &= (u_1 + \epsilon_2^2 a^*)s, \\
\widetilde{\xi} &= \epsilon_2^2 \left[\frac{s}{k^2} - \frac{\pi}{k^2 K\omega_2}\{E(\gamma(s), k) - E(\gamma(0), k)\} - a^* s\right], \\
\overline{\rho}^2 &= \frac{1}{N^2}[c - b(u_1 + \epsilon_2^2 a^*)], \\
\widetilde{\rho}^2 &= -\frac{b}{N^2}\epsilon_2^2 \left[\mathrm{sn}^2(\frac{K\varphi_2}{\pi}, k) - a^*\right], \\
\overline{\alpha} &= \frac{N}{b}(J_1 + g^*)s, \\
\widetilde{\alpha} &= \frac{N}{b}\left[\frac{J_3 b - J_1 c}{c - bu_1}\frac{\pi}{K\omega_2}\{\Pi(\gamma(s), \beta, k) - \Pi(\gamma(0), \beta, k)\} - g^* s\right],
\end{aligned}$$

where

$$a^* = \frac{1}{Kk^2}(K - E), \quad g^* = \frac{1}{4K}\frac{J_3 b - J_1 c}{c - bu_1}\Pi(2\pi, \beta, k).$$

Proof. Since ξ', ρ^2 and α' are periodic functions in arclength s, it is possible to find, for each of these functions, decompositions of the form $g(s) = \bar{g} + \widetilde{g}(s)$, where $\bar{g}$ is constant and $\widetilde{g}$ is periodic with zero average. This is achieved by choosing (assuming the period of g is T)

$$\bar{g} = \frac{1}{T}\int_0^T g(s)ds,$$

so that $\widetilde{g} = g - \bar{g}$. Then ξ, ρ^2 and α have the desired form.

We use the representations (7.8)–(7.10) for ξ', ρ^2 and α' to compute the respective decompositions. First consider (7.8) and compute the average a^* of $\mathrm{sn}^2(u,k)$ over one period of length $4K$, i.e.

$$
\begin{aligned}
a^* &= \frac{1}{4K}\int_0^{4K} \mathrm{sn}^2(u,k)du \\
&= \frac{1}{4Kk^2}\Big[u - E(\mathrm{am}(u,k),k)\Big]_0^{4K} \\
&= \frac{1}{Kk^2}(K-E).
\end{aligned}
$$

This implies that

$$
\int_0^s \left[\mathrm{sn}^2\left(\frac{K\varphi_2}{\pi},k\right) - a^*\right] ds
$$

has zero average, so that

$$
\begin{aligned}
\overline{\xi} &= (u_1 + \epsilon_2^2 a^*)s, \\
\widetilde{\xi} &= \epsilon_2^2\left[\frac{s}{k^2} - \frac{\pi}{k^2K\omega_2}(E(\gamma,k) - E(\gamma_0,k)) - a^* s\right].
\end{aligned}
$$

Using the same calculations and (7.9) immediately yields the decomposition for ρ^2.

Finally, recall the form for α in (7.10) and compute the average g^* of

$$
\frac{J_3 b - J_1 c}{c - bu_1 - b\epsilon_2^2 \mathrm{sn}^2(u,k)}
$$

over one period of length $4K$, i.e.

$$
\begin{aligned}
g^* &= \frac{1}{4K}\int_0^{4K} \frac{J_3 b - J_1 c}{c - bu_1 - b\epsilon_2^2\mathrm{sn}^2(u,k)}du \\
&= \frac{1}{4K}\frac{J_3 b - J_1 c}{c - bu_1}\int_0^{4K}\frac{1}{1-\beta \mathrm{sn}^2(u,k)}du \\
&= \frac{1}{4K}\frac{J_3 b - J_1 c}{c - bu_1}[\Pi(\mathrm{am}(4K),\beta,k) - \Pi(\mathrm{am}(0),\beta,k)] \\
&= \frac{1}{4K}\frac{J_3 b - J_1 c}{c - bu_1}\Pi(2\pi,\beta,k),
\end{aligned}
$$

and thus $\alpha - \alpha(0) = \overline{\alpha} + \widetilde{\alpha}$ with

$$
\begin{aligned}
\overline{\alpha} &= \frac{N}{b}(J_1 + g^*)s, \\
\widetilde{\alpha} &= \frac{N}{b}\left(\frac{J_3 b - J_1 c}{c - bu_1}\frac{\pi}{K\omega_2}\left[\Pi(\gamma,\beta,k) - \Pi(\gamma_0,\beta,k)\right] - g^* s\right).
\end{aligned}
$$

□

The decomposition of the centerline, formulated in this theorem, may be interpreted as the centerline $\mathbf{r}$ of the rod having two different length scales, i.e. $\mathbf{r}(s,\tau) = \overline{\mathbf{r}}(s) + \widetilde{\mathbf{r}}(\tau)$, where in cylindrical coordinates $\overline{\mathbf{r}}$ is given by $(\overline{\rho}, \overline{\alpha}, \overline{\xi})$ and $\widetilde{\mathbf{r}}$ is given by $(\widetilde{\rho}, \widetilde{\alpha}, \widetilde{\xi})$. Note that $\overline{\mathbf{r}}$ describes a true helix, whereas $\widetilde{\mathbf{r}}$ should be interpreted in a local frame moving along $\overline{\mathbf{r}}$, and may not necessarily be helical.

Furthermore we remark that the decomposition of the centerline is not unique, since there is obviously some freedom in choosing the primary helix $\overline{\mathbf{r}}$. In particular the average value of $\tilde{\mathbf{r}}$ depends on this choice and a choice of the initial conditions $\boldsymbol{\varphi}(0)$. It seems natural however, to choose the average to be zero.

We conclude by investigating the structure of the projection of the centerline into the $\mathbf{ij}$-plane. As already shown in Section 6.4 this projection is always bounded for a uniform elastic rod. Under the additional assumption of isotropy the radial and angular components ρ and α are periodic functions. In particular, ρ^2 oscillates between the two values

$$\begin{aligned} \rho^2_{min} &= \frac{1}{N^2}(c - bu_2), \\ \rho^2_{max} &= \frac{1}{N^2}(c - bu_1). \end{aligned}$$

In the limit as u_2 approaches u_1 the projection becomes a circle and and we therefore conclude the centerline is a helix.

In the next section we will study the component ξ of the centerline in the direction of the force, to determine when solutions are bounded or closed.

7.4 Special Solutions

We have shown that the projection of the centerline $\mathbf{r}$ into the $\mathbf{ij}$–plane is bounded, so that in general $\mathbf{r}$ can only be unbounded in the direction of the force $\mathbf{n}$, given through the respective component ξ. This component depends on the initial condition of the angle variables $\boldsymbol{\varphi}$ and the values of the action variables $\mathbf{J}$. We will determine conditions on these values such that the centerline $\mathbf{r}$ is bounded in all directions.

Lemma 7.1. *Adopt the assumptions and notation of Corollary 7.2 and suppose the modulus for the elliptic integrals* $k^2 = \frac{u_2 - u_1}{u_3 - u_1}$ *is given. Define* $k^* = k^*(k) = \frac{K - E}{K}$ *and let*

$$c_1 = 3 - \frac{1 + k^2}{k^*}, \qquad c_2 = 3 - 2\frac{1 + k^2}{(k^*)^2} - \frac{k^2}{(k^*)^2}$$

and

$$c_3 = (1 - \frac{k^2}{k^*})(1 - \frac{1}{k^*}).$$

Then the centerline $\mathbf{r}$ *is bounded if and only if*

$$a_1^2 c_2 = a_2 c_1^2 \text{ and } a_1^3 c_3 = a_3 c_1^3,$$

where $a_1 = -\frac{a+J_3^2}{b}$, $a_2 = \frac{2J_1J_3-b}{b}$ *and* $a_3 = \frac{a-J_1^2}{b}$.

Proof. It is sufficient to show that the component ξ of $\mathbf{r}$ in the direction of the force $\mathbf{n}$ is bounded. From the representation of ξ this is the case if and only if

$$u_1 + \epsilon_2^2 a^* = 0,$$

which is equivalent to either of the following statements

$$u_2 = u_1\left(1 - \frac{k^2}{k^*}\right), \tag{7.11}$$

$$u_3 = u_1\left(1 - \frac{1}{k^*}\right). \tag{7.12}$$

Note that $k^* \neq 0$, since k^* vanishes if and only if $k = 0$, which we have excluded by assumption.

Let

$$\widetilde{f}(u) = \frac{1}{b} f(u) = u^3 + a_1 u^2 + a_2 u + a_3.$$

The roots of the polynomial $\widetilde{f}$ are related to the coefficients a_i through the following nonlinear system of equations

$$\begin{aligned} u_1 + u_2 + u_3 &= -a_1, \\ u_1u_2 + u_1u_3 + u_2u_3 &= a_2, \\ u_1u_2u_3 &= -a_3. \end{aligned}$$

These become by using representations (7.11) and (7.12) for u_2 and u_3

$$c_1 u_1 = -a_1, \quad c_2 u_1^2 = a_2, \quad c_3 u_1^2 = -a_3,$$

which imply the desired conditions. □

We have provided conditions on the parameters of the problem, such that the level sets of the Hamiltonian system are compact, and the general helix-on-helix structure described in Lemma 7.3 implies that the centerline $\mathbf{r}$ now lies on a two-dimensional torus of revolution in $\mathbb{E}^3$.

As a special case, one often considers solutions of the equilibrium conditions for which the centerline forms a smooth loop. In general these loops are found as solutions of two-point boundary value problems, i.e. for a rod of length $L = 1$ the boundary conditions

$$\mathbf{r}(0) = \mathbf{r}(1), \quad \mathbf{d}_3(0) = \mathbf{d}_3(1)$$

should be enforced. In addition we prescribe the angle α between the director $\mathbf{d}_1$ at $s = 0$ and $s = 1$, say

$$\mathbf{d}_1(0) \cdot \mathbf{d}_1(1) = \cos\alpha.$$

In practice however, one is mostly interested in solutions where the frame $\mathcal{D}$ closes upon itself so that $\alpha = 2n\pi$ (see for example the investigation of DNA mini-circles in Manning et al. [44]).

For the special integrable case of a rod equilibrium we have described in this chapter, conditions for ring closure can be formulated explicitly in terms of the parameters of the system, exploiting the knowledge of closed form solutions. Benham [10] formulates necessary conditions for ring closure. Starostin [60] determines explicit sufficient conditions for an elastic rod to form a smooth loop and exploits these for a bifurcation analysis. A bifurcation analysis is also obtained by Li and Maddocks [39] through a numerical treatment of the appropriate two-point boundary value problem for the Euler parameter formulation.

Chapter 8

Homogenization of Constitutive Relations Governing Elastic Rods with Rapidly Varying Intrinsic Shape

In this and the following chapters results concerning effective rod models will be developed and applied to some examples of practical importance. Motivated by the application of approximating large-scale deformations of DNA molecules by a continuum rod model, we assume the intrinsic shape of a rod to depend on two different length scales, a local (fast) scale and a global (slow) scale. For such rods it may be expected that averaging over the fast scale provides an effective (homogenized) constitutive relation, which captures the essential behavior of the rod on the global scale. In addition, the effective constitutive relations may be considerably simpler, and provide insight into properties that are not obvious when studying the original model. The elimination of the fast variation may furthermore be of advantage for numerical computations.

Introducing the two length scales s and $\tau = s/\epsilon$ in the minimum energy configuration of a rod yields a fast variation of the director frame, and generally the components of intrinsic strain relative to this frame will become unbounded in the limit as ϵ tends to zero. Consequently standard averaging techniques cannot be applied immediately. It will be shown that a new frame can be defined so that the associated strain vector remains bounded as ϵ tends to zero, and the coefficients of the system expressed in variables relative to the new frame inherit a rapid oscillation. The resulting system is in a form suitable for applying the theory of averaging for ordinary differential equations.

In Section 8.1 we will develop some necessary tools. In particular, we show that the Hamiltonian structure of the governing equations persists under the change of variables that is generated by introducing a new frame. These results are exploited in Section 8.2 in the context of elastic rods whose intrinsic shape depends on two different length scales. Averaging theory is applied in Section 8.3 to obtain an effective (homogenized) Hamiltonian system. In the final section of this chapter we discuss possible extensions of this model. Examples in which the homogenization procedure described below can be carried out explicitly are presented in Chapter 9.

8.1 Transformation of the Hamiltonian System Under a Change of Frames

The main tool for our analysis of effective rod models is the construction of a new frame $\mathcal{N}$ relative to the director frame $\mathcal{D}$ such that the associated strain vector in the reference configuration is bounded in the limit of fast oscillation. In this Section we will consider a general, arclength-dependent, change of frames, and examine how the strain vectors associated with the two frames are related. This change of frames induces an arclength-dependent change of variables in the governing Hamiltonian system. It will be shown that the Hamiltonian structure persists under this transformation, introducing an extra term in the Hamiltonian.

To clarify the relationship between different frames and associated strain vectors, we introduce some notation (see also Appendix A): In the following we use three different, orthonormal and right handed bases for $\mathbb{E}^3$: the fixed (inertial) frame $\mathcal{E} = \{\mathbf{e}_i\}_{i=1}^3$, the moving (director) frame $\mathcal{D} = \{\mathbf{d}_i(s)\}_{i=1}^3$, which have been defined in Section 2.1, and an additional moving frame $\mathcal{N} = \{\mathbf{D}_i(s)\}_{i=1}^3$. Recall that we denote the triples of components of any vector $\mathbf{p} \in \mathbb{E}^3$ with respect to $\mathcal{E}$ and $\mathcal{D}$ by $\underline{\mathbf{p}}$ and p, respectively. In addition the components of $\mathbf{p}$ with respect to $\mathcal{N}$ are denoted P. The three triples of components $\{\mathsf{e}_i\}_{i=1}^3$ denote a basis for $\mathbb{R}^3$ and are defined to be $\mathsf{e}_1 = (1,0,0)$, etc.

We now define a frame $\mathcal{N}$ relative to the director frame $\mathcal{D}$ as follows: Let

$$\mathbf{A}(s) = \left\{a_{ij}(s)\right\}_{0\le i,j\le 3} \tag{8.1}$$

be a general proper rotation matrix and define a frame $\mathcal{N} = \{\mathbf{D}_i\}_{i=1}^3$ through

$$\mathbf{D}_i = a_{ij}\mathbf{d}_j, \quad i = 1,2,3. \tag{8.2}$$

This frame has an associated strain vector $\mathbf{u}_{\mathcal{N}}$ given implicitly through the kinematic relation

$$\mathbf{D}_i' = \mathbf{u}_{\mathcal{N}} \times \mathbf{D}_i, \quad i = 1,2,3.$$

Each frame defines a rotation matrix containing the components of the frame relative to the fixed basis $\mathcal{E}$, i.e.

$$\mathbf{R}_{\mathcal{D}} = (\underline{\mathbf{d}}_1, \underline{\mathbf{d}}_2, \underline{\mathbf{d}}_3), \quad \mathbf{R}_{\mathcal{N}} = (\underline{\mathbf{D}}_1, \underline{\mathbf{D}}_2, \underline{\mathbf{D}}_3).$$

The matrices $\mathbf{R}_{\mathcal{D}}$ and $\mathbf{R}_{\mathcal{N}}$ can then be related through the rotation matrix (8.1), i.e.

$$\mathbf{R}_{\mathcal{N}}\mathsf{e}_3 = \underline{\mathbf{D}}_i = a_{ij}\underline{\mathbf{d}}_j = a_{ij}\mathbf{R}_{\mathcal{D}}\mathsf{e}_3 = \mathbf{R}_{\mathcal{D}}\mathbf{A}^T\mathsf{e}_3,$$

so that compactly we can write $\mathbf{R}_{\mathcal{N}} = \mathbf{R}_{\mathcal{D}}\mathbf{A}^T$ or

$$\mathbf{R}_{\mathcal{D}} = \mathbf{R}_{\mathcal{N}}\mathbf{A}. \tag{8.3}$$

We express the strain vectors $\mathbf{u}$ and $\mathbf{u}_{\mathcal{N}}$ in components relative to $\mathcal{D}$ and $\mathcal{N}$:

$$\mathbf{d}_i' = (\mathbf{u}\cdot\mathbf{d}_j)\mathbf{d}_j \times \mathbf{d}_i, \quad \mathbf{D}_i' = (\mathbf{u}_{\mathcal{N}}\cdot\mathbf{D}_j)\mathbf{D}_j \times \mathbf{D}_i.$$

Recalling the definition of the matrix $\mathsf{y}\times$ corresponding to the vector product, (cf. Appendix A),

$$\mathsf{y}\times = \begin{pmatrix} 0 & -y_3 & y_2 \\ y_3 & 0 & -y_1 \\ -y_2 & y_1 & 0 \end{pmatrix}, \quad \mathsf{y} = (y_1, y_2, y_3) \in \mathbb{R}^3,$$

we obtain the compact expressions

$$\mathbf{R}_{\mathcal{D}}' = \mathbf{R}_{\mathcal{D}}\mathsf{u}\times, \quad \mathbf{R}_{\mathcal{N}}' = \mathbf{R}_{\mathcal{N}}\mathsf{U}_{\mathcal{N}}\times, \tag{8.4}$$

where $\mathsf{u} = (\mathbf{u}\cdot\mathbf{d}_1, \mathbf{u}\cdot\mathbf{d}_2, \mathbf{u}\cdot\mathbf{d}_3)$ is the triple of components of the strain vector $\mathbf{u}$ relative to $\mathcal{D}$ and $\mathsf{U}_{\mathcal{N}} = (\mathbf{u}_{\mathcal{N}}\cdot\mathbf{D}_1, \mathbf{u}_{\mathcal{N}}\cdot\mathbf{D}_2, \mathbf{u}_{\mathcal{N}}\cdot\mathbf{D}_3)$ is the triple of components of the strain vector $\mathbf{u}_{\mathcal{N}}$ relative to $\mathcal{N}$. Utilizing this notation we express the relationship between the strain vectors $\mathbf{u}$ and $\mathbf{u}_{\mathcal{N}}$ as follows:

Lemma 8.1. *Assume the frames $\mathcal{D} = \{\mathbf{d}_i\}_{i=1}^3$ and $\mathcal{N} = \{\mathbf{D}_i\}_{i=1}^3$ are related through* (8.2)*, and let $\mathbf{u}$ and $\mathbf{u}_{\mathcal{N}}$ be the associated strain vectors. Then the triples of components u and $\mathsf{U}_{\mathcal{N}}$ of these vectors relative to $\mathcal{D}$ and $\mathcal{N}$, respectively, are related through*

$$\mathsf{U}_{\mathcal{N}}(s)\times = \big[\mathbf{A}(s)\mathsf{u}(s)\big]\times - \mathbf{A}'(s)\mathbf{A}^T(s).$$

Proof. Using the relations (8.3) and (8.4) we immediately compute

$$\begin{aligned} \mathbf{R}_{\mathcal{N}}\mathsf{U}_{\mathcal{N}}\times &= (\mathbf{R}_{\mathcal{N}})' \\ &= (\mathbf{R}_{\mathcal{D}}\mathbf{A}^T)' = \mathbf{R}_{\mathcal{D}}'\mathbf{A}^T + \mathbf{R}_{\mathcal{D}}(\mathbf{A}^T)' \\ &= \mathbf{R}_{\mathcal{D}}\mathsf{u}\times\mathbf{A}^T + \mathbf{R}_{\mathcal{D}}(\mathbf{A}^T)' \\ &= \mathbf{R}_{\mathcal{N}}\mathbf{A}\mathsf{u}\times\mathbf{A}^T + \mathbf{R}_{\mathcal{N}}\mathbf{A}(\mathbf{A}^T)' \\ &= \mathbf{R}_{\mathcal{N}}\Big([\mathbf{A}\mathsf{u}]\times - \mathbf{A}'\mathbf{A}^T\Big) \end{aligned}$$

so that

$$\mathsf{U}_{\mathcal{N}}\times = \big[\mathbf{A}\mathsf{u}\big]\times - \mathbf{A}'\mathbf{A}^T.$$

□

We will now show how the newly defined frame $\mathcal{N}$ can be used to define a change of variables in the Hamiltonian formulation governing elastic rods. Although the results in this chapter apply to any of the Hamiltonian formulations in Chapter 4, we specifically adopt the non-canonical formulation (4.10), since the Hamiltonian (4.9) seems to have the simplest structure. Recall that this system is defined by the structure matrix

$$\mathbf{J}(\mathbf{z}) = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathsf{i}\times & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathsf{r}\times & \mathbf{I} \\ \mathsf{i}\times & \mathsf{r}\times & \mathsf{m}\times & \mathsf{n}\times \\ \mathbf{0} & -\mathbf{I} & \mathsf{n}\times & \mathbf{0} \end{pmatrix},$$

which yields, together with the Hamiltonian

$$H(\mathbf{z}, s) = W^*(\mathsf{m}, \mathsf{n}, s) + \mathsf{m} \cdot \hat{\mathsf{u}}(s) + \mathsf{n} \cdot \hat{\mathsf{v}}(s),$$

the twelve-dimensional system

$$\mathbf{z}' = \mathbf{J}(\mathbf{z})\nabla H(\mathbf{z}, s), \tag{8.5}$$

where $\mathbf{z} = (\mathsf{i}, \mathsf{r}, \mathsf{m}, \mathsf{n})$ are the components of the centerline $\mathbf{r}$, the fixed vector $\mathbf{i}$ and the moment and force vectors $\mathbf{m}$ and $\mathbf{n}$ with respect to the director frame $\mathcal{D}$. We define a change of variables for the Hamiltonian system by expressing the vectors $\mathbf{r}$, $\mathbf{i}$, $\mathbf{m}$ and $\mathbf{n}$ in components with respect to the new frame $\mathcal{N}$, denoted $\mathbf{Z} = (\mathsf{I}, \mathsf{R}, \mathsf{M}, \mathsf{N})$. Explicitly this s-dependent change of variables is given by

$$\mathbf{Z} = \mathcal{A}(s)\mathbf{z} = \begin{pmatrix} \mathbf{A}(s) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}(s) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}(s) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}(s) \end{pmatrix} \mathbf{z}. \tag{8.6}$$

Generally, an s-dependent change of variables may not necessarily preserve the Hamiltonian structure. However, we next show that the linear transformation (8.6) is special in the sense that it preserves the Poisson bracket, but introduces an extra term in the Hamiltonian:

Lemma 8.2. *Under the s-dependent change of variables* (8.6) *the Hamiltonian system* (8.5) *becomes*

$$\mathbf{Z}' = \mathbf{J}(\mathbf{Z})\nabla \widetilde{H}(\mathbf{Z}, s)$$

with Hamiltonian

$$\widetilde{H}(\mathbf{Z}, s) = W^*\left(\mathbf{A}^T(s)\mathsf{M}, \mathbf{A}^T(s)\mathsf{N}, s\right) + \mathsf{M} \cdot \mathbf{A}(s)\left[\hat{\mathsf{u}}(s) - \mathsf{a}(s)\right] + \mathsf{N} \cdot \mathbf{A}(s)\hat{\mathsf{v}}(s),$$

where $\mathsf{a}(s)$ is defined through $\mathbf{A}'(s) = \mathbf{A}(s)\mathsf{a}(s)\times$, and the variables $\mathbf{Z} = (\mathsf{I}, \mathsf{R}, \mathsf{M}, \mathsf{N})$ are the components of $\mathbf{i}$, $\mathbf{r}$, $\mathbf{m}$ *and* $\mathbf{n}$ *with respect to the frame $\mathcal{N}$.*

Proof. According to (3.3) in Section 3.3, this Hamiltonian system transforms under the change of variables (8.6) to

$$\begin{aligned}\mathbf{Z}' &= \mathcal{A}'\mathcal{A}^T\mathbf{Z} + \mathcal{A}\,\mathbf{J}\big(\mathcal{A}^T\mathbf{Z}\big)\,\mathcal{A}^T\nabla_{\mathbf{Z}}H\left(\mathcal{A}^T\mathbf{Z}, s\right) \\ &= \mathcal{A}'\mathcal{A}^T\mathbf{Z} + \mathbf{J}(\mathbf{Z})\nabla_{\mathbf{Z}}H\left(\mathcal{A}^T\mathbf{Z}, s\right),\end{aligned}$$

where the identity $\mathcal{A}\,\mathbf{J}\big(\mathcal{A}^T\mathbf{Z}\big)\,\mathcal{A}^T = \mathbf{J}(\mathbf{Z})$ is verified using (A.1) in Appendix A. By definition $\mathbf{A}'(s) = \mathbf{A}(s)\mathsf{a}(s)\times$, so that

$$\begin{aligned}\mathcal{A}'\mathcal{A}^T\mathbf{Z} &= \mathcal{A}\begin{pmatrix}\mathsf{a}\times & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathsf{a}\times & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathsf{a}\times & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathsf{a}\times\end{pmatrix}\mathcal{A}^T\mathbf{Z} = \begin{pmatrix}(\mathbf{A}\mathsf{a})\times\mathsf{l}\\ (\mathbf{A}\mathsf{a})\times\mathsf{R}\\ (\mathbf{A}\mathsf{a})\times\mathsf{M}\\ (\mathbf{A}\mathsf{a})\times\mathsf{N}\end{pmatrix} \\ &= -\mathbf{J}(\mathbf{Z})\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{A}\mathsf{a}\\ \mathbf{0}\end{pmatrix} = \mathbf{J}(\mathbf{Z})\nabla_{\mathbf{Z}}(-\mathsf{M}\cdot\mathbf{A}\mathsf{a}).\end{aligned} \tag{8.7}$$

Finally, we express the Hamiltonian H in the new variables to get

$$H(\mathcal{A}^T\mathbf{Z}, s) = W^*\big(\mathbf{A}^T\mathsf{M}, \mathbf{A}^T\mathsf{N}, s\big) + \mathsf{M}\cdot\mathbf{A}\hat{\mathsf{u}} + \mathsf{N}\cdot\mathbf{A}\hat{\mathsf{v}}.$$

Thus with (8.7) we compute the Hamiltonian system in the variables $\mathbf{Z}$

$$\begin{aligned}\mathbf{Z}' &= \mathbf{J}(\mathbf{Z})\nabla_{\mathbf{Z}}\left[H(\mathcal{A}^T\mathbf{Z}, s) - \mathsf{M}\cdot\mathbf{A}\mathsf{a}\right] \\ &= \mathbf{J}(\mathbf{Z})\widetilde{H}(\mathbf{Z}, s).\end{aligned}$$

□

In the following sections we will apply this result to rods whose intrinsic shape depends on two different length scales.

We remark that the form of the Hamiltonian system derived in Lemma 8.2 is independent of which of the different Hamiltonian formulations in Chapter 4 is used. We have chosen the non-canonical formulation just for convenience. Alternatively one could use for example the canonical formulation in terms of Euler parameters. One parametrizes the director frame $\mathcal{D}$ by a quaternion $\mathbf{q}$, and obtains a Hamiltonian formulation in terms of $\mathbf{q}$ and their conjugate momenta $\boldsymbol{\mu}$. Next one parametrizes the new frame $\mathcal{N}$ by a different set of Euler parameters $\mathbf{Q}$ and introduces conjugate momenta $\boldsymbol{\Upsilon}$. It can then be shown that the transformation (8.6) expressed in the variables $(\mathbf{q}, \boldsymbol{\mu})$ and $(\mathbf{Q}, \boldsymbol{\Upsilon})$ is canonical.

8.2 Intrinsic Shape with Different Length Scales

In this section we consider rods with an intrinsic shape that depends on two different length scales. We take the specific form of the reference configuration to be quite general, but of a form that averaging theory can eventually be applied.

The *reference configuration* (or intrinsic shape) of a rod is given by four smooth vector-valued functions

$$\hat{\mathbf{r}}(s), \hat{\mathbf{d}}_i(s) : [s_1, s_2] \to \mathbb{E}^3, \quad i = 1, 2, 3,$$

describing the shape of the rod centerline and the director frame in the minimum energy configuration. Each frame $\hat{\mathcal{D}} = \{\hat{\mathbf{d}}_i\}_{i=1}^3$ defines a rotation matrix

$$\hat{\mathbf{R}}(s) = \left(\underline{\hat{\mathbf{d}}}_1(s), \underline{\hat{\mathbf{d}}}_2(s), \underline{\hat{\mathbf{d}}}_3(s)\right),$$

i.e. $\hat{\mathbf{R}}$ contains the components of the moving frame $\hat{\mathcal{D}}$ relative to the fixed frame $\mathcal{E}$. Thus equivalently we can assume the reference configuration to be defined by the pair $\{\hat{\mathbf{r}}, \hat{\mathbf{R}}\}$. The strain vectors $\hat{\mathbf{u}}(s)$ and $\hat{\mathbf{v}}(s)$ in the reference configuration are, analogous to (2.1) and (2.2), defined through the equations

$$\begin{aligned} \hat{\mathbf{d}}_i' &= \hat{\mathbf{u}} \times \hat{\mathbf{d}}_i, \quad i = 1, 2, 3, \\ \hat{\mathbf{r}}' &= \hat{\mathbf{v}}, \end{aligned}$$

and their components relative to the frame $\hat{\mathcal{D}}$ are denoted by

$$\hat{\mathsf{u}} = \begin{pmatrix} \hat{\mathbf{u}} \cdot \hat{\mathbf{d}}_1 \\ \hat{\mathbf{u}} \cdot \hat{\mathbf{d}}_2 \\ \hat{\mathbf{u}} \cdot \hat{\mathbf{d}}_3 \end{pmatrix}, \quad \hat{\mathsf{v}} = \begin{pmatrix} \hat{\mathbf{v}} \cdot \hat{\mathbf{d}}_1 \\ \hat{\mathbf{v}} \cdot \hat{\mathbf{d}}_2 \\ \hat{\mathbf{v}} \cdot \hat{\mathbf{d}}_3 \end{pmatrix}.$$

In particular, the triple of components $\hat{\mathsf{u}}$ can be conveniently written in terms of the rotation matrix $\hat{\mathbf{R}}$ through

$$\hat{\mathsf{u}}(s)\times = \hat{\mathbf{R}}(s)^T \frac{d}{ds} \hat{\mathbf{R}}(s). \tag{8.8}$$

(see Appendix A).

We now assume the reference configuration $\{\hat{\mathbf{r}}(s;\epsilon), \hat{\mathbf{R}}(s;\epsilon)\}$ of the rod to be dependent on a small parameter ϵ, and in particular upon two different length scales s and $\tau = s/\epsilon$. Equivalently, the director frame $\hat{\mathcal{D}} = \{\hat{\mathbf{d}}_i(s,\epsilon)\}_{i=1}^3$ in the reference configuration depends on ϵ, s and τ. For example the reference configuration of the rod may exhibit the helix-on-helix structure depicted in Figure 8.1 (locally helical rods will be studied in detail in Chapter 9). In Definition 8.1 we introduce a dependence of the configuration $\{\hat{\mathbf{r}}, \hat{\mathbf{R}}\}$ on the variables s, τ and ϵ that is motivated by the desire to cast the Hamiltonian system in a form that is suitable for an application of averaging theory:

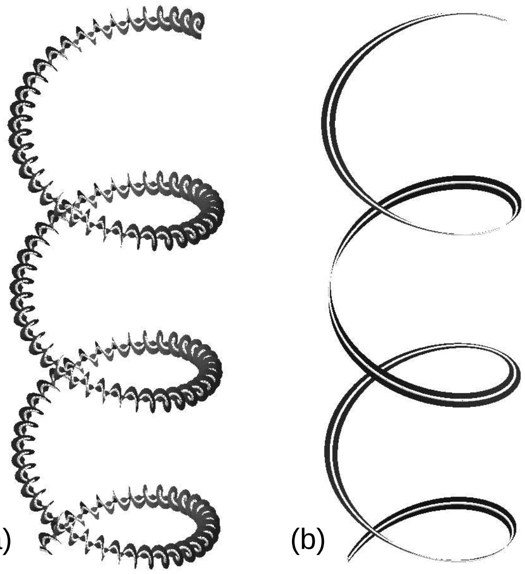

Figure 8.1: In (a) the unstressed configuration of a rod with two different length scales is shown. We expect averaging over the fast scale to yield the effective, unstressed approximation (b), which depends only on the slow length scale, but with a different constitutive relation.

Definition 8.1. *The reference configuration of the rod is said to depend on a fast scale $\tau = s/\epsilon$ superimposed on a slow scale s, if the pair $\{\hat{\mathbf{r}}, \hat{\mathbf{R}}\}$ has the form*

$$\begin{aligned} \hat{\mathbf{r}}(s;\epsilon) &= \mathbf{r}^\sharp(s,\tau;\epsilon), \\ \hat{\mathbf{R}}(s;\epsilon) &= \widetilde{\mathbf{R}}(s;\epsilon)\breve{\mathbf{R}}(\tau;\epsilon), \end{aligned} \tag{8.9}$$

where the functions $\widetilde{\mathbf{R}}(s;\epsilon)$, $\breve{\mathbf{R}}(\tau;\epsilon)$ and $\mathbf{r}^\sharp(s,\tau;\epsilon)$ are 2π-periodic in τ, and sufficiently smooth in the variables s, τ and ϵ. In particular, as $\epsilon \to 0$

$$\frac{d}{ds}\widetilde{\mathbf{R}}(s;\epsilon) = O(1), \quad \frac{d}{d\tau}\breve{\mathbf{R}}\,(\tau;\epsilon) = O(1),$$

$$\frac{\partial}{\partial s}\mathbf{r}^\sharp\,(s,\tau;\epsilon) = O(1), \quad \frac{\partial}{\partial \tau}\mathbf{r}^\sharp\,(s,\tau;\epsilon) = O(\epsilon),$$

for fixed s and τ.

We now compute the form of the components of intrinsic strain in the reference configuration that is implied by Definition 8.1. The assumptions on the functions

$\widetilde{\mathbf{R}}$, $\breve{\mathbf{R}}$ and $\mathbf{r}^\sharp$ imply in particular, that in the reference configuration the components of intrinsic strain $\hat{\mathsf{u}}$ and $\hat{\mathsf{v}}$ are of order $\hat{\mathsf{u}}(s;\epsilon) = O(\epsilon^{-1})$ and $\hat{\mathsf{v}}(s;\epsilon) = O(1)$ as $\epsilon \to 0$.

Lemma 8.3. *Assume the intrinsic shape of the rod* $\{\hat{\mathbf{r}}, \hat{\mathbf{R}}\}$ *is of the form*

$$\hat{\mathbf{r}}(s;\epsilon) = \mathbf{r}^\sharp(s,\tau;\epsilon), \quad \hat{\mathbf{R}}(s;\epsilon) = \widetilde{\mathbf{R}}(s;\epsilon)\breve{\mathbf{R}}(\tau;\epsilon)$$

according to Definition 8.1. Then the components of intrinsic strain $\hat{\mathsf{u}}$ *and* $\hat{\mathsf{v}}$ *are of the form*

$$\begin{aligned} \hat{\mathsf{u}}(s;\epsilon) &= \breve{\mathbf{R}}^T(\tau;\epsilon)\tilde{\mathsf{u}}(s;\epsilon) + \frac{1}{\epsilon}\breve{\mathsf{u}}(\tau;\epsilon), \\ \hat{\mathsf{v}}(s;\epsilon) &= \mathsf{v}^\sharp(s,\tau;\epsilon), \end{aligned} \tag{8.10}$$

where $\breve{\mathsf{u}}$ *and* $\tilde{\mathsf{u}}$ *are related to* $\breve{\mathbf{R}}$ *and* $\widetilde{\mathbf{R}}$ *through*

$$\frac{d}{d\tau}\breve{\mathbf{R}}(\tau;\epsilon) = \breve{\mathbf{R}}(\tau;\epsilon)\breve{\mathsf{u}}(\tau;\epsilon)\times, \quad \frac{d}{ds}\widetilde{\mathbf{R}}(s;\epsilon) = \widetilde{\mathbf{R}}(s;\epsilon)\tilde{\mathsf{u}}(s;\epsilon)\times,$$

and the components $\mathsf{v}^\sharp = (v_1^\sharp, v_2^\sharp, v_3^\sharp)$ *are defined through*

$$v_i^\sharp(s,\tau;\epsilon) = \hat{\mathbf{r}}(s;\epsilon)\cdot\hat{\mathbf{d}}_i(s;\epsilon) = \mathbf{r}^\sharp(s,\tau;\epsilon)\cdot\hat{\mathbf{d}}_i(s;\epsilon).$$

Consequently $\hat{\mathsf{u}}(s;\epsilon) = O(\epsilon^{-1})$ *and* $\hat{\mathsf{v}}(s;\epsilon) = O(1)$ *as* $\epsilon \to 0$.

Proof. The components of strain $\hat{\mathsf{u}}$ are defined through (8.8), so that we compute

$$\begin{aligned} &\hat{\mathsf{u}}(s;\epsilon)\times = \\ &= \hat{\mathbf{R}}^T(s;\epsilon)\frac{d}{ds}\hat{\mathbf{R}}(s;\epsilon) \\ &= \breve{\mathbf{R}}^T(\tau;\epsilon)\widetilde{\mathbf{R}}^T(s;\epsilon)\left(\left[\frac{d}{ds}\widetilde{\mathbf{R}}(s;\epsilon)\right]\breve{\mathbf{R}}(\tau;\epsilon) + \widetilde{\mathbf{R}}(s;\epsilon)\frac{1}{\epsilon}\left[\frac{d}{d\tau}\breve{\mathbf{R}}(\tau;\epsilon)\right]\right) \\ &= \breve{\mathbf{R}}^T(\tau;\epsilon)\widetilde{\mathbf{R}}^T(s;\epsilon)\widetilde{\mathbf{R}}(s;\epsilon)\left[\tilde{\mathsf{u}}(s;\epsilon)\times\right]\breve{\mathbf{R}}(\tau;\epsilon) \\ &\quad +\breve{\mathbf{R}}^T(\tau;\epsilon)\widetilde{\mathbf{R}}^T(s;\epsilon)\widetilde{\mathbf{R}}(s;\epsilon)\frac{1}{\epsilon}\breve{\mathbf{R}}(\tau;\epsilon)\left[\breve{\mathsf{u}}(\tau;\epsilon)\times\right] \\ &= \left[\breve{\mathbf{R}}^T(\tau;\epsilon)\tilde{\mathsf{u}}(s;\epsilon) + \frac{1}{\epsilon}\breve{\mathsf{u}}(\tau;\epsilon)\right]\times . \end{aligned}$$

We see that the components of intrinsic strain defined by

$$\tilde{\mathsf{u}}(s;\epsilon)\times = \widetilde{\mathbf{R}}^T(s;\epsilon)\frac{d}{ds}\widetilde{\mathbf{R}}(s;\epsilon), \quad \breve{\mathsf{u}}(\tau;\epsilon)\times = \breve{\mathbf{R}}^T\left(\frac{s}{\epsilon};\epsilon\right)\frac{d}{d\tau}\breve{\mathbf{R}}\left(\frac{s}{\epsilon};\epsilon\right)$$

are bounded as $\epsilon \to 0$, using the boundedness of the rotation matrices $\breve{\mathbf{R}}$ and $\widetilde{\mathbf{R}}$ together with the periodicity in the variable τ. Thus

$$\hat{\mathsf{u}}(s;\epsilon) = \breve{\mathbf{R}}^T(\tau;\epsilon)\tilde{\mathsf{u}}(s;\epsilon) + \frac{1}{\epsilon}\breve{\mathsf{u}}(\tau;\epsilon) = O\left(\frac{1}{\epsilon}\right),$$

as $\epsilon \to 0$.

Furthermore

$$\begin{aligned} v_i^\sharp(s,\tau;\epsilon) &= \left[\frac{d}{ds}\mathbf{r}^\sharp(s,\tau;\epsilon)\right] \cdot \hat{\mathbf{d}}_i(s;\epsilon) \\ &= \left[\frac{\partial}{\partial s} + \frac{1}{\epsilon}\frac{\partial}{\partial \tau}\right] \mathbf{r}^\sharp(s,\tau;\epsilon) \cdot \hat{\mathbf{d}}_i(s;\epsilon), \end{aligned}$$

so that $\mathsf{v}^\sharp(s,\tau,\epsilon) = O(1)$ using the boundedness and the periodicity of the partial derivatives $\partial\mathbf{r}^\sharp/\partial s$ and $\partial\mathbf{r}^\sharp/\partial\tau$. □

In rod mechanics it is customary to assume that the components of intrinsic strain $\hat{\mathsf{v}}$ have the particular form $\hat{\mathsf{v}} = \mathsf{e}_3 \equiv (0,0,1)$, so that $\hat{\mathbf{r}}'(s) = \hat{\mathbf{d}}_3(s)$. This can be achieved by rotating the frame $\hat{\mathcal{D}} = \{\hat{\mathbf{d}}_i\}_{i=1}^3$ to define a frame $\hat{\mathcal{D}}^\flat = \{\hat{\mathbf{d}}_i^\flat\}_{i=1}^3$ so that $\hat{\mathbf{r}}'$ is parallel to $\hat{\mathbf{d}}_3^\flat$, i.e.

$$\hat{\mathbf{r}}'(s) = \alpha(s)\hat{\mathbf{d}}_3^\flat(s)$$

for some scalar-valued function $\alpha(s)$. Then $\hat{\mathbf{r}}'(s)$ can be normalized by an appropriate change in the length scale of the independent variable s. Note that the definition of the frame $\hat{\mathcal{D}}^\flat$ along with the change in the length scale induces a change in the constitutive relations. However, for our considerations the assumption $\hat{\mathsf{v}} = \mathsf{e}_3$ is too restrictive since the frame $\hat{\mathcal{D}}$ (or equivalently the rotation matrix $\hat{\mathbf{R}}$) has already been chosen to have the form (8.9). In this case the transformation from the frame $\hat{\mathcal{D}}$ to the frame $\hat{\mathcal{D}}^\flat$ introduces a dependence on the fast variables τ in the strain-energy function, which may pose complications for the averaging analysis. We therefore adopt the general form for $\hat{\mathsf{v}}$ allowed in Definition 8.1. A special subcase is obtained by adopting the condition $\hat{\mathsf{v}} = \mathsf{e}_3$ and then assuming that the corresponding frame $\hat{\mathcal{D}}$ still has the special form (8.9). For inextensible and unshearable linearly elastic rods, one may also choose the frame $\hat{\mathcal{D}}^\flat$ to make the strain-energy function (2.9) a sum of squares. This implies that in general neither $\hat{\mathsf{v}} = \mathsf{e}_3$ nor the frame $\mathcal{D}^\flat$ can be assumed to have the form (8.9).

We have seen that the specific dependence of the reference configuration $\{\hat{\mathbf{r}}, \hat{\mathbf{R}}\}$ on the variables τ and ϵ given in Definition 8.1 implies that the components of strain $\hat{\mathsf{u}}$ become unbounded in the limit as ϵ tends to zero, which means that averaging theory cannot be applied immediately. We will exploit the results of Section 8.1 by explicitly defining a slowly varying frame $\mathcal{N}$, such that the strain vector associated with this frame is bounded. This immediately implies that the Hamiltonian system can be cast into a form that explicitly allows averaging to be applied.

We define the frame $\mathcal{N} = \{\mathbf{D}_i\}_{i=1}^3$ in (8.2) by identifying the rotation matrix $\mathbf{A}$ with the rapidly oscillating matrix $\breve{\mathbf{R}}$ defined in (8.9). Then we obtain the desired result, namely that in the reference configuration the components $\mathsf{U}_{\mathcal{N}}$ of the strain vector $\mathbf{u}_{\mathcal{N}}$ relative to $\mathcal{N}$ depend on the slow scale only.

Lemma 8.4. *Define the frame $\mathcal{N} = \{\mathbf{D}_i\}_{i=1}^3$ by*

$$\mathbf{D}_i(s;\epsilon) = \breve{r}_{ij}(\tau;\epsilon)\mathbf{d}_j(s;\epsilon), \quad i,j = 1,2,3,$$

with $\breve{\mathbf{R}}(\tau;\epsilon) = \{\breve{r}_{ij}(\tau;\epsilon)\}$, $0 \le i,j \le 3$ defined in (8.9), and assume the components of strain $\hat{\mathsf{u}}(s;\epsilon)$ are of the form (8.10), i.e.

$$\hat{\mathsf{u}}(s;\epsilon) = \breve{\mathbf{R}}^T(\tau;\epsilon)\tilde{\mathsf{u}}(s;\epsilon) + \frac{1}{\epsilon}\breve{\mathsf{u}}(\tau;\epsilon)$$

with $\frac{d}{d\tau}\breve{\mathbf{R}}(\tau;\epsilon) = \breve{\mathbf{R}}(\tau;\epsilon)\breve{\mathsf{u}}(\tau;\epsilon)\times$ and $\frac{d}{ds}\widetilde{\mathbf{R}}(s;\epsilon) = \widetilde{\mathbf{R}}(s;\epsilon)\tilde{\mathsf{u}}(s;\epsilon)\times$.

Then in the reference configuration the components $\hat{\mathsf{U}}_{\mathcal{N}}$ of the strain vector $\mathbf{U}$ relative to $\mathcal{N}$ satisfy $\hat{\mathsf{U}}_{\mathcal{N}}(s;\epsilon) = \tilde{\mathsf{u}}(s;\epsilon)$.

Proof. According to Lemma 8.1 the relationship between the triples $\mathsf{U}_{\mathcal{N}}$ and u can be explicitly written as

$$\begin{aligned}
\mathsf{U}_{\mathcal{N}}(s;\epsilon)\times &= -\left[\frac{d}{ds}\breve{\mathbf{R}}(\tau;\epsilon)\right]\breve{\mathbf{R}}^T(\tau;\epsilon) + \left[\breve{\mathbf{R}}(\tau;\epsilon)\mathsf{u}(s;\epsilon)\right]\times \\
&= -\frac{1}{\epsilon}\left[\frac{d}{d\tau}\breve{\mathbf{R}}(\tau;\epsilon)\right]\breve{\mathbf{R}}^T(\tau;\epsilon) + \left[\breve{\mathbf{R}}(\tau;\epsilon)\mathsf{u}(s;\epsilon)\right]\times \\
&= -\frac{1}{\epsilon}\breve{\mathbf{R}}(\tau;\epsilon)\ \breve{\mathsf{u}}(\tau;\epsilon)\times\ \breve{\mathbf{R}}^T(\tau;\epsilon) + \left[\breve{\mathbf{R}}(\tau;\epsilon)\mathsf{u}(s;\epsilon)\right]\times \\
&= \left[\breve{\mathbf{R}}(\tau;\epsilon)\left\{\mathsf{u}(s;\epsilon) - \frac{1}{\epsilon}\breve{\mathsf{u}}(\tau;\epsilon)\right\}\right]\times,
\end{aligned}$$

i.e. $\mathsf{U}_{\mathcal{N}}(s;\epsilon) = \breve{\mathbf{R}}(\tau;\epsilon)\{\mathsf{u}(s;\epsilon) - \frac{1}{\epsilon}\breve{\mathsf{u}}(\tau;\epsilon)\}$. Using (8.10) the strains $\hat{\mathsf{U}}_{\mathcal{N}}$ in the reference configuration become $\hat{\mathsf{U}}_{\mathcal{N}}(s;\epsilon) = \breve{\mathbf{R}}(\tau;\epsilon)\{\hat{\mathsf{u}}(s;\epsilon) - \frac{1}{\epsilon}\breve{\mathsf{u}}(\tau;\epsilon)\} = \tilde{\mathsf{u}}(s;\epsilon)$, in particular they do not depend on the fast variable τ. $\square$

Finally, we exploit Lemma 8.2 to compute the specific structure of the Hamiltonian system in the new frame $\mathcal{N}$. Now the change of variables (8.6) is explicitly dependent on the variables τ and ϵ, i.e.

$$\mathbf{Z} = \mathcal{A}(s;\epsilon)\mathbf{z} = \begin{pmatrix} \breve{\mathbf{R}}(\tau;\epsilon) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \breve{\mathbf{R}}(\tau;\epsilon) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \breve{\mathbf{R}}(\tau;\epsilon) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \breve{\mathbf{R}}(\tau;\epsilon) \end{pmatrix}\mathbf{z}. \tag{8.11}$$

Thus the matrix $\mathsf{a}(s;\epsilon)\times$ in Lemma 8.2 takes the explicit form

$$\mathsf{a}(s;\epsilon)\times = \breve{\mathbf{R}}^T(\tau;\epsilon)\frac{d}{ds}\breve{\mathbf{R}}(\tau;\epsilon) = \frac{1}{\epsilon}\breve{\mathbf{R}}^T(\tau;\epsilon)\frac{d}{d\tau}\breve{\mathbf{R}}(\tau;\epsilon) = \frac{1}{\epsilon}\breve{\mathsf{u}}(\tau;\epsilon)\times$$

adding the extra term $-\mathsf{M}\cdot\breve{\mathbf{R}}\mathsf{a} = -\mathsf{M}\cdot\breve{\mathbf{R}}\frac{1}{\epsilon}\breve{\mathsf{u}}$ in the Hamiltonian. Correspondingly, in the variables $\mathbf{Z} = (\mathsf{I},\mathsf{R},\mathsf{M},\mathsf{N})$ the Hamiltonian system governing the equilibria of elastic rods becomes

$$\mathbf{Z}' = \mathbf{J}(\mathbf{Z})\nabla\widetilde{H}(\mathbf{Z},s,\tau;\epsilon) \tag{8.12}$$

with Hamiltonian

$$\widetilde{H}(\mathbf{Z},s,\tau;\epsilon) = W^*\left(\breve{\mathbf{R}}^T(\tau;\epsilon)\mathsf{M},\breve{\mathbf{R}}^T(\tau;\epsilon)\mathsf{N},s\right) + \mathsf{M}\cdot\tilde{\mathsf{u}}(s;\epsilon) + \mathsf{N}\cdot\breve{\mathbf{R}}(\tau;\epsilon)\mathsf{v}^\sharp(s,\tau;\epsilon).$$

Thus the Hamiltonian system (8.12) is a system of differential equations with rapidly oscillating coefficients, but the Hamiltonian is smooth in the variables s, τ and ϵ, so that the right-hand-side of (8.12) is bounded for fixed τ, in the limit as ϵ tends to zero. The 2π-periodic dependence on the variable $\tau = s/\epsilon$ can now be conveniently exploited to define an averaged (or homogenized) system. A full description of the homogenized system along with estimates for the quality of the approximation will be given in Section 8.3.

8.3 Homogenization

In Section 8.2 we derived a Hamiltonian system with rapidly oscillating coefficients, and we will now show how averaging over the fast variable yields an effective (homogenized) system, which may under additional assumptions be interpreted as an effective rod model.

Consider the Hamiltonian system (8.12) formulated as an initial value problem

$$\mathbf{Z}' = \mathbf{J}(\mathbf{Z})\nabla\widetilde{H}\left(\mathbf{Z},s,\frac{s}{\epsilon},\epsilon\right), \quad \mathbf{Z}(0) = \mathbf{Z}_0, \tag{8.13}$$

where $\mathbf{Z}\in\mathbb{R}^{12}$ denotes the components of the variables $(\mathbf{i},\mathbf{r},\mathbf{m},\mathbf{n})$ with respect to the frame $\mathcal{N}$, i.e. $\mathbf{Z} = (\mathsf{I},\mathsf{R},\mathsf{M},\mathsf{N})$. This Hamiltonian system depends on two different length scales s and $\tau = s/\epsilon$, and is 2π-periodic in the fast variable τ. Our concern here is to apply the theory of averaging to determine the asymptotic behavior of the initial value problem for small ϵ.

Define an averaged Hamiltonian

$$\overline{H}(\mathbf{Z},s) = \frac{1}{2\pi}\int_0^{2\pi}\widetilde{H}(\mathbf{Z},s,\tau,0)d\tau \tag{8.14}$$

by integrating over one period with respect to the fast variable τ, where

$$H(\mathbf{Z}, s, \tau, 0) = \lim_{\epsilon \to 0} H(\mathbf{Z}, s, \tau, \epsilon)$$

for fixed τ. Then consider the corresponding averaged system

$$\mathbf{Z}' = \mathbf{J}(\mathbf{Z}) \nabla \overline{H}(\mathbf{Z}, s), \quad \mathbf{Z}(0) = \bar{\mathbf{Z}}_0. \tag{8.15}$$

We would like to understand the sense in which the solution of the averaged system (8.15) approximates the solution of (8.13). An answer can be found in the theory of averaging (cf. Guckenheimer and Holmes [23, Ch. 4], Hale [25, Sec. V.3] or Sanders and Verhulst [54, Ch. 3]). We summarize the results in the following theorem, which is a version of Theorem 4.1.1 in [23] adapted for our purposes.

This theorem requires a certain degree of smoothness in the variables $\mathbf{Z}$, s, τ and ϵ of the right-hand-side of the differential equations (8.13) to ensure in particular boundedness as ϵ tends to zero (specifically, the right-hand-side of (8.13) being at least two times continuously differentiable suffices). We satisfy this requirement by assuming the strain-energy function W to be sufficiently smooth. Note however, that the components of intrinsic strain (8.10) are unbounded in the limit as ϵ tends to zero. This motivated our consideration to construct a new frame relative to which the strains are bounded in the limit as ϵ tends to zero, so that the transformed Hamiltonian

$$\widetilde{H}(\mathbf{Z}, s, \tau, \epsilon) = W^* \left(\breve{\mathbf{R}}^T(\tau, \epsilon)\mathsf{M}, \breve{\mathbf{R}}^T(\tau, \epsilon)\mathsf{N}, s \right) + \mathsf{M} \cdot \tilde{\mathsf{u}}(s, \epsilon) + \mathsf{N} \cdot \breve{\mathbf{R}}(\tau)\mathsf{v}^\sharp(s, \tau, \epsilon)$$

and the Hamiltonian system (8.12) are as smooth as required.

Theorem 8.1. *Consider the solutions* $\mathbf{Z}_\epsilon, \bar{\mathbf{Z}} : [0, b] \to U$ *of* (8.13) *and* (8.15)*, respectively, for some bounded set* U*. If the initial data of problems* (8.13) *and* (8.15) *satisfy* $|\mathbf{Z}_\epsilon(0) - \bar{\mathbf{Z}}(0)| = O(\epsilon)$*, then* $|\mathbf{Z}_\epsilon(s) - \bar{\mathbf{Z}}(s)| = O(\epsilon)$ *on bounded intervals. In particular there exist constants* L*,* C_1 *and* C_2 *(dependent on* U *and* b*), so that*

$$|\mathbf{Z}_\epsilon(s) - \bar{\mathbf{Z}}(s)| \leq \left[|\mathbf{Z}_\epsilon(0) - \bar{\mathbf{Z}}(0)| + \epsilon \frac{C_1}{L} + \epsilon C_2 \right] e^{Ls} + \epsilon C_2.$$

Proof. To reach this conclusion we first cast the problem into a form suitable for standard averaging theory. Make the substitution $t = s/\epsilon$ in (8.13) to obtain the system

$$\frac{d}{dt}\mathbf{Z} = \epsilon \mathbf{J}(\mathbf{Z}) \nabla H(\mathbf{Z}, \epsilon t, t, \epsilon), \quad \mathbf{Z}(0) = \mathbf{Z}_0.$$

Introducing the new variable $\sigma = \epsilon t$ then yields the augmented system

$$\frac{d}{dt}\mathbf{x} = \epsilon \mathbf{f}(\mathbf{x}, t, \epsilon), \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{8.16}$$

where $\mathbf{x} = (\mathbf{Z}, \sigma)^T$, $\mathbf{x}_0 = (\mathbf{Z}_0, 0)^T$ and $\mathbf{f}(\mathbf{x}, t, \epsilon) = (\mathbf{J}(\mathbf{Z})\nabla H(\mathbf{Z}, \sigma, t, \epsilon), 1)^T$. This system is now exactly in a form suitable for applying Theorem 4.1.1 in [23], and we review the main steps of the proof of this theorem. In particular, there exists a smooth change of variables

$$\mathbf{x} = \mathbf{y} + \epsilon \mathbf{w}(\mathbf{y}, t, \epsilon), \tag{8.17}$$

which essentially defines the averaged system. Specifically, under the transformation (8.17) system (8.16) becomes

$$\frac{d}{dt}\mathbf{y} = \epsilon \bar{\mathbf{f}}(\mathbf{y}) + \epsilon^2 \mathbf{f}_1(\mathbf{y}, t, \epsilon), \tag{8.18}$$

with the corresponding averaged system

$$\frac{d}{dt}\mathbf{y} = \epsilon \bar{\mathbf{f}}(\mathbf{y}) = \frac{\epsilon}{2\pi}\int_0^{2\pi} \mathbf{f}(\mathbf{x}, t, 0)dt. \tag{8.19}$$

Denote the solution of (8.18) and (8.19) by $\mathbf{y}_\epsilon$ and $\bar{\mathbf{y}}$, respectively. Then Gronwall-type estimates imply that the distance between the solutions can be measured as

$$|\mathbf{y}_\epsilon(t) - \bar{\mathbf{y}}(t)| \leq \left[|\mathbf{y}_\epsilon(0) - \bar{\mathbf{y}}(0)| + \epsilon\frac{C_1}{L}\right] e^{\epsilon L t},$$

where L is a Lipschitz constant of $\bar{\mathbf{f}}$ on U, and C_1, C_2 are upper bounds of $\mathbf{f}_1$ and $\mathbf{w}$ on $U \times [0, b]$, respectively. The solution $\mathbf{x}_\epsilon$ of (8.16) can then be estimated via

$$\begin{aligned} |\mathbf{x}_\epsilon(t) - \bar{\mathbf{y}}(t)| &\leq |\mathbf{x}_\epsilon(t) - \mathbf{y}_\epsilon(t)| + |\mathbf{y}_\epsilon(t) - \bar{\mathbf{y}}(t)| \\ &\leq \left[|\mathbf{x}_\epsilon(0) - \bar{\mathbf{y}}(0)| + \epsilon\frac{C_1}{L} + \epsilon C_2\right] e^{\epsilon L t} + \epsilon C_2. \end{aligned}$$

Translating back to the original variables yields the desired estimates on $|\mathbf{Z}_\epsilon(s) - \bar{\mathbf{Z}}(s)|$. □

Fundamental to the proof of the theorem is the change of variables (8.17) that generates the transformed system (8.18). Under the imposed smoothness assumptions on the Hamiltonian this change of variables always exists. There is no guarantee however, that the transformation preserves the Hamiltonian structure of the system, i.e. that system (8.18) is Hamiltonian. For a canonical Hamiltonian system this change of variables can be chosen to be canonical, meaning that the transformed system is itself a (canonical) Hamiltonian system (cf. Goldstein [21, Sec. 9.1]). For the non-canonical structure we have used here, it is not clear whether the change of variables preserves the Hamiltonian structure. However, for the averaging result, as we have formulated above, this is not essential, since the lowest order approximation can be seen to be always Hamiltonian by inspection. Generally this may not be true, for example if the Poisson bracket depends on the fast variable τ.

We remark that the averaging result is independent of the particular parametrization chosen for the rotation group. It was just convenience that led us to use the non-canonical Hamiltonian description. In Appendix E we shall further discuss the implications of our averaging result in light of Hamiltonian perturbation theory.

It remains to show that the effective Hamiltonian system (8.15) actually defines a valid rod model. Consider the effective Hamiltonian (8.14) in the explicit form

$$\overline{H}(\mathsf{M},\mathsf{N},s) = \overline{W}^*(\mathsf{M},\mathsf{N},s) + \mathsf{M}\cdot\tilde{\mathsf{u}}(s,0) + \mathsf{N}\cdot\overline{\mathsf{V}}(s), \tag{8.20}$$

where

$$\overline{W}^*(\mathsf{M},\mathsf{N},s) = \frac{1}{2\pi}\int_0^{2\pi} W^*(\breve{\mathbf{R}}(\tau,0)\mathsf{M},\breve{\mathbf{R}}(\tau,0)\mathsf{N},s)d\tau,$$

and

$$\overline{\mathsf{V}}(s) = \frac{1}{2\pi}\int_0^{2\pi} \breve{\mathbf{R}}(\tau,0)\mathsf{v}^\sharp(s,\tau,0)d\tau.$$

Corollary 8.1. *If $\overline{V}_3(s) = \overline{\mathsf{V}}(s)\cdot\mathsf{e}_3 > 0$, $\forall s$ then the Hamiltonian system* (8.15) *defines a rod model. The configuration of the rod is given by a centerline $\mathbf{r}$ and a frame $\mathcal{N} = \{\mathbf{D}_i\}_{i=1}^3$, and the (effective) constitutive relations are*

$$\begin{aligned}
\mathsf{M} &= \frac{\partial\overline{W}}{\partial\mathsf{w}}\Big(\mathsf{U}_{\mathcal{N}} - \tilde{\mathsf{u}}(s,0), \mathsf{V} - \overline{\mathsf{V}}(s), s\Big),\\
\mathsf{N} &= \frac{\partial\overline{W}}{\partial\mathsf{z}}\Big(\mathsf{U}_{\mathcal{N}} - \tilde{\mathsf{u}}(s,0), \mathsf{V} - \overline{\mathsf{V}}(s), s\Big),
\end{aligned}$$

where the function $\overline{W}(\mathsf{w},\mathsf{z},s)$ is the Legendre transform of $\overline{W}^*(\mathsf{M},\mathsf{N},s)$. The function $\overline{W}$ has a minimum at $(\mathbf{0},\mathbf{0})$ and is both convex and coercive.

Here $\mathsf{U}_{\mathcal{N}} = (\mathbf{u}_{\mathcal{N}}\cdot\mathbf{D}_1, \mathbf{u}_{\mathcal{N}}\cdot\mathbf{D}_2, \mathbf{u}_{\mathcal{N}}\cdot\mathbf{D}_3)$ are the components of the strain vector $\mathbf{u}_{\mathcal{N}}$ relative to the slow frame $\mathcal{N}$. Furthermore, the triples V, M and N denote the components relative to $\mathcal{N}$ of the strain vector $\mathbf{v} = \mathbf{r}'$, the moment vector $\mathbf{m}$ and the force vector $\mathbf{n}$, respectively.

Proof. The averaged system (8.15) can be interpreted as a rod model, provided the effective Hamiltonian (8.20) satisfies several conditions. First, consistent with the definition in Section 2.2, the effective strain-energy function $\overline{W}$ must be convex, coercive and possess a minimum at $(\mathbf{0},\mathbf{0})$. Second, the effective component of intrinsic strain $\overline{V}_3(s)$ (measuring dilatation) must satisfy $\overline{V}_3(s) = \overline{\mathsf{V}}(s)\cdot\mathsf{e}_3 > 0$ for the constitutive relation to be orientation preserving.

To verify the first condition, define a function $g(\mathbf{x}) : \mathbb{R}^6 \to \mathbb{R}$ by

$$g(\mathbf{x}) = W^*(\mathsf{M}, \mathsf{N}, s),$$

and the corresponding average

$$\overline{g}(\mathbf{x}) = \int_0^{2\pi} g\big(\mathcal{A}(\tau)\mathbf{x}\big)\, d\tau = \int_0^{2\pi} W^*(\breve{\mathbf{R}}(\tau,0)\mathsf{M}, \breve{\mathbf{R}}(\tau,0)\mathsf{N}, s)\, d\tau,$$

where $\mathbf{x} = (\mathsf{M}, \mathsf{N})$ and

$$\mathcal{A}(\tau) = \begin{pmatrix} \breve{\mathbf{R}}^T(\tau;0) & \mathbf{0} \\ \mathbf{0} & \breve{\mathbf{R}}^T(\tau;0) \end{pmatrix}.$$

Since W is a strain-energy function, the function g is convex and coercive, and has a minimum at $\mathbf{0}$, since these properties are preserved under the Legendre transform. Thus provided that $\overline{g}$ is convex, coercive and is minimized at $\mathbf{0}$, the effective strain-energy function is defined as $\overline{W} = \big(\overline{W^*}\big)^*$ and has the required properties.

Let $\lambda \in [0,1], \mathbf{x}, \mathbf{y} \in \mathbb{R}^6$, and compute

$$\begin{aligned}
&\overline{g}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \\
&= \int_0^{2\pi} g\big(\lambda\mathcal{A}(\tau)\mathbf{x} + (1-\lambda)\mathcal{A}(\tau)\mathbf{y}\big)\, d\tau \\
&\le \int_0^{2\pi} \big[\lambda g\big(\mathcal{A}(\tau)\mathbf{x}\big) + (1-\lambda) g\big(\mathcal{A}(\tau)\mathbf{y}\big)\big]\, d\tau \\
&\le \lambda\overline{g}(\mathbf{x}) + (1-\lambda)\overline{g}(\mathbf{y})
\end{aligned}$$

using the convexity of g. Thus $\overline{g}$ is convex. To prove the coercivity of $\overline{g}$ note that

$$\frac{\overline{g}(\mathbf{x})}{|\mathbf{x}|} = \int_0^{2\pi} \frac{g\big(\mathcal{A}(\tau)\mathbf{x}\big)}{|\mathbf{x}|} d\tau = \int_0^{2\pi} \frac{g\big(\mathcal{A}(\tau)\mathbf{x}\big)}{|\mathcal{A}(\tau)\mathbf{x}|} d\tau \ge 2\pi \min_\tau \frac{g\big(\mathcal{A}(\tau)\mathbf{x}\big)}{|\mathcal{A}(\tau)\mathbf{x}|} \to \infty$$

as $|\mathbf{x}| \to \infty$, since g is coercive. It is apparent that $\overline{g}$ has a minimum at $\mathbf{0}$ if g has a minimum at $\mathbf{0}$.

$\square$

In Chapter 9 the homogenized system (8.15) will be exploited to obtain effective properties of elastic rods, whose intrinsic shape is highly coiled. For the special case of linearly elastic rods the averaged Hamiltonian can be computed in a fairly explicit form.

8.4 Extensions of the Model

In the previous sections we imposed several assumptions on the form of the reference configuration and the constitutive relations, some of which can easily be relaxed.

Recall the form of the Hamiltonian system (8.12) after the change of variables (8.11), in particular the Hamiltonian

$$\widetilde{H}(\mathbf{Z}, s, \tau; \epsilon) = W^*\big(\breve{\mathbf{R}}^T(\tau;\epsilon)\mathsf{M}, \breve{\mathbf{R}}^T(\tau;\epsilon)\mathsf{N}, s\big) + \mathsf{M}\cdot\tilde{\mathsf{u}}(s;\epsilon) + \mathsf{N}\cdot\breve{\mathbf{R}}(\tau;\epsilon)\hat{\mathsf{v}}(s;\epsilon).$$

Consider the following two generalizations of this model. First, assume that the strain-energy function W has an explicit dependence on the variables τ and ϵ and is sufficiently smooth in these variables, i.e. the Legendre transform takes the more general form

$$W^*(\mathsf{m}, \mathsf{n}, s, \tau, \epsilon). \tag{8.21}$$

Second, we relax the assumption that the Hamiltonian be periodic in the fast variable τ.

The more general form for the strain-energy function (8.21) still satisfies the assumptions of the homogenization theory of Section 8.3; in particular, Theorem 8.1 remains valid, i.e. the effective Hamiltonian is computed as the average of $\widetilde{H}(\mathbf{Z}, s, \tau; \epsilon)$ over one period in τ. However, the effective strain-energy function is now more generally calculated as

$$\overline{W}^*(\mathsf{M}, \mathsf{N}, s) = \frac{1}{2\pi}\int_0^{2\pi} W^*\big(\breve{\mathbf{R}}^T(\tau;\epsilon)\mathsf{M}, \breve{\mathbf{R}}^T(\tau;\epsilon)\mathsf{N}, s, \tau, 0\big)d\tau.$$

Furthermore Corollary 8.1 can be modified to apply in this case so that the system describes an effective rod model, provided the constitutive relation is orientation preserving.

Relaxing the periodicity assumption requires a more general averaging theory than the one presented in Section 8.3. Farkas [20] discusses averaging of differential equations for the case where the dependence on the fast variable is not periodic but rather quasi-periodic: The function $\widetilde{H}(\mathbf{Z}, s, \tau; \epsilon)$ is *quasi-periodic* with angular frequencies $\omega_1, \ldots, \omega_m$, $(\omega_k > 0, k = 1, \ldots, m)$ if

$$\widetilde{H}(\mathbf{Z}, s, \tau; \epsilon) = H^\sharp(\mathbf{Z}, s, \omega_1\tau, \ldots, \omega_m\tau; \epsilon),$$

and the function $H^\sharp(\mathbf{z}, s, \tau_1, \ldots, \tau_m; \epsilon)$ has period 2π in each of the variables $\tau_1, \ldots, \tau_m$. Suppose the limit

$$\overline{H}(\mathbf{Z}, s) = \lim_{T\to\infty}\frac{1}{T}\int_0^T \widetilde{H}(\mathbf{Z}, s, \tau, 0)d\tau$$

exists uniformly for $\mathbf{Z}$ and s. Then under appropriate smoothness assumptions on the Hamiltonian, the averaged system is given by

$$\mathbf{Z}' = \mathbf{J}(\mathbf{Z})\nabla \overline{H}(\mathbf{Z}, s),$$

and similar estimates as in Section 8.3 for the approximation of the true solution by the solution of the averaged system can be obtained (for details see [20] and the references therein). A quasi-periodic function is a special case of an *almost periodic* function, and averaging results can still be obtained in this case (see for example Hale [25]).

Finally we discuss an apparent generalization of Definition 8.1. Suppose that the frame $\hat{\mathcal{D}}$ (or equivalently the rotation matrix $\hat{\mathbf{R}}(s, \epsilon)$) takes the more general form

$$\hat{\mathbf{R}}(s;\epsilon) = \mathbf{R}_0(s, \tau; \epsilon),$$

rather than the factorization into a slow and fast rotation adopted in Definition 8.1. Furthermore assume that $\mathbf{R}_0$ is 2π-periodic in the variable τ, and as $\epsilon \to 0$

$$\frac{\partial}{\partial s}\mathbf{R}_0(s, \tau; \epsilon) = O(1), \quad \frac{\partial}{\partial \tau}\mathbf{R}_0(s, \tau; \epsilon) = O(1)$$

for fixed s and τ. Then the components of intrinsic strain $\hat{\mathsf{u}}$ are computed as

$$\hat{\mathsf{u}}(s;\epsilon) = \mathbf{R}_0^T(s, \tau; \epsilon)\left[\frac{\partial}{\partial s} + \frac{1}{\epsilon}\frac{\partial}{\partial \tau}\right]\mathbf{R}_0(s, \tau; \epsilon),$$

which implies that $\hat{\mathsf{u}}(s;\epsilon) = O(\epsilon^{-1})$ as $\epsilon \to 0$, (cf. also Lemma 8.3). The averaging theory in Section 8.2 still applies in this case. In particular, the frame $\mathcal{N}$ in Lemma (8.4) is defined through

$$\mathbf{D}_i(s;\epsilon) = r_{ij}^0(s, \tau; \epsilon)\mathbf{d}_j(s;\epsilon),$$

where $\mathbf{R}_0 = \{r_{ij}^0\}$, relative to which the components of the associated strain vector are bounded. Note that in the reference configuration the frame $\mathcal{N}$ is a parallel transport frame, i.e. it is constant. Both Theorem 8.1 and Corollary 8.1 apply in this case and the averaged Hamiltonian (8.20) takes the slightly more general form

$$\overline{H}(\mathsf{M}, \mathsf{N}, s) = \overline{W}^*(\mathsf{M}, \mathsf{N}, s) + \mathsf{N}\cdot\overline{\mathsf{V}}(s),$$

where

$$\overline{W}^*(\mathsf{M}, \mathsf{N}, s) = \frac{1}{2\pi}\int_0^{2\pi} W^*(\mathbf{R}_0(s, \tau, 0)\mathsf{M}, \mathbf{R}_0(s, \tau, 0)\mathsf{N}, s)d\tau,$$

and

$$\overline{\mathsf{V}}(s) = \frac{1}{2\pi}\int_0^{2\pi} \mathbf{R}_0(s, \tau, 0)\mathsf{v}^\sharp(s, \tau, 0)d\tau.$$

However, it is important to realize that the frame $\mathcal{N}$ should not necessarily be chosen to be a parallel transport frame. Since for this choice the components of intrinsic strain $\hat{\mathsf{U}}$ relative to $\mathcal{N}$ vanish, the entire intrinsic shape is introduced into the constitutive relation. This is in general not the best available choice as the examples in the next chapter will show. We therefore conclude that the above case may not truly represent a useful generalization of Definition 8.1.

Chapter 9

Application: Rods with Locally Helical Intrinsic Shape

The motivation for developing the homogenization theory derived in Chapter 8 was the use of the continuum rod model as an approximation for large scale deformations of DNA. The intrinsic structure of the DNA molecule suggests that the rod model should incorporate an intrinsic shape which is either highly coiled or as a limiting case highly twisted (cf. Section 9.1), and we wish to understand the effective long scale properties.

In Section 9.2 we define the intrinsic shape of the rod to depend on two different length scales: a fast scale that captures the local structure, which we assume to be helical, and a slow scale that describes the structure of the rod on a relatively long length scale in the reference configuration. The corresponding Hamiltonian system is analyzed in 9.3 with the homogenization result of Chapter 8 being applied to obtain an effective system. In Section 9.4 we prove that the effective system possesses an important property, namely that the constitutive relation is isotropic, provided that all short-scale dependence arises through the unstressed shape, (i.e. the strain-energy function does not depend on the fast length scale explicitly). In Section 9.5 we study the particular case of a linearly elastic rod and obtain a fairly explicit expression for the homogenized constitutive relation. Rods with a slowly varying unstressed shape but a highly twisted intrinsic frame provide an important special case that is discussed in Section 9.6. We conclude this chapter with some numerical examples to illustrate our results. Since the helical rod model discussed in this chapter can also be regarded as modeling a spring, we will review some related results about effective spring theories in Appendix F.

9.1 The Elastic Rod as a Model for DNA

The main application motivating our investigation of effective properties of elastic rods is to consider the rod as a continuum model of long length-scale deformations of DNA (see for example the recent survey articles by Schlick [55] and Olson [50] or the continuum computations of DNA mini-circles in Manning et al. [44]).

In these models it is typically assumed that the bending response is isotropic. It is however believed that on short, base pair level, length scales DNA has an anisotropic response to loads that induce bending. This anisotropic behavior is to be expected, since the two backbone strands define a chord in each circular cross-section of the double helix, and there is no reason to expect equal responses to bending about axes parallel and perpendicular to this chord. Moreover, the two backbones do not lie diametrically opposite each other, but instead have a significant angular offset, so that the chord creates what are called the major and minor grooves. Consequently there is no reason to expect equal response to two bending moments about a single axis parallel to the chord, but with opposite senses that correspond to deformations into either the major or minor groove.

On the other hand, the two helical backbone strands have a quite rapid intrinsic twist, so it has been conjectured that this rapid twist compensates for any asymmetry at the base pair level in such a way that on longer length scales there is a single isotropic bending stiffness.

We will prove a mathematical result along these lines. Specifically, we will show that in an asymptotic limit of high intrinsic twist, for a linearly elastic inextensible unshearable rod, there exists an effective isotropic bending stiffness, given as the harmonic average of the two original bending moduli. This result is actually a special case of our analysis of highly coiled rods, which may serve as a more general model of DNA.

9.2 Locally Helical Intrinsic Shape

Generally we could describe the rod as having an intrinsic helical shape whenever the centerline $\hat{\mathbf{r}}$ in the reference configuration is helical with no restriction on the frame $\hat{\mathcal{D}} = \{\hat{\mathbf{d}}_i\}_{i=1}^3$ (or equivalently on the rotation matrix $\hat{\mathbf{R}} = (\underline{\hat{\mathbf{d}}}_1, \underline{\hat{\mathbf{d}}}_2, \underline{\hat{\mathbf{d}}}_3)$). However, we shall focus on the case in which $\hat{\mathbf{r}}' = \hat{\mathbf{d}}_3$, so that the shape of the centerline is determined by the evolution of the director $\hat{\mathbf{d}}_3$. Since there is still some freedom in specifying the frame $\hat{\mathcal{D}}$, we further assume that the director $\hat{\mathbf{d}}_2$ traces the radius vector of the helix. Choosing the inertial frame $\mathcal{E}$ so that the basis vector $\mathbf{e}_3$ lies in the direction of the helix axis, we obtain the following parametrization of a helix with radius $\hat{\rho}$ and pitch $2\pi\hat{p}$:

$$\hat{\mathbf{r}}(s) = \frac{\hat{p}}{\sqrt{\hat{\rho}^2 + \hat{p}^2}} s\, \mathbf{e}_3 - \hat{\rho}\, \hat{\mathbf{d}}_2(s).$$

It is easily seen that this characterization of a helix is equivalent to the following definition:

Definition 9.1. *The reference configuration (or intrinsic shape)* $\{\hat{\mathbf{r}}(s), \hat{\mathbf{R}}(s)\}$ *of a rod is called helical with radius* $\hat{\rho}$ *and pitch* $2\pi\hat{p}$*, if* $\hat{\mathbf{r}}'(s) = \hat{\mathbf{d}}_3(s)$ *and* $\hat{\mathbf{R}}(s) = \breve{\mathbf{R}}(\hat{\alpha}s; \hat{\eta})$*, where*

$$\breve{\mathbf{R}}(\tau;\hat{\eta}) = \begin{pmatrix} \hat{\eta}_* \cos\tau & -\sin\tau & \hat{\eta}\cos\tau \\ \hat{\eta}_* \sin\tau & \cos\tau & \hat{\eta}\sin\tau \\ -\hat{\eta} & 0 & \hat{\eta}_* \end{pmatrix}, \tag{9.1}$$

with $\hat{\alpha} = \frac{1}{\sqrt{\hat{\rho}^2+\hat{p}^2}}$, $\hat{\eta} = \hat{\rho}\hat{\alpha}$ *and* $\hat{\eta}_* = \sqrt{1-\hat{\eta}} = \hat{p}\hat{\alpha}$.

The rotation matrix $\breve{\mathbf{R}}$ has the decomposition $\breve{\mathbf{R}}(\tau;\hat{\eta}) = \mathbf{Q}(\tau)\mathbf{S}_{\hat{\eta}}$, where

$$\mathbf{Q}(\tau) = \begin{pmatrix} \cos\tau & \sin\tau & 0 \\ -\sin\tau & \cos\tau & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{S}_{\hat{\eta}} = \begin{pmatrix} \hat{\eta}_* & 0 & \hat{\eta} \\ 0 & 1 & 0 \\ -\hat{\eta} & 0 & \hat{\eta}_* \end{pmatrix}, \tag{9.2}$$

so that the frame $\hat{\mathcal{D}}$ in the helical reference configuration is obtained by rotating the frame relative to the fixed frame $\mathcal{E}$ about the 3-axis through an angle $\hat{\alpha}s$, followed by a rotation about the resulting 2-axis through a fixed angle $\arcsin(\hat{\eta})$. Correspondingly the components of intrinsic strain $\hat{\mathsf{u}}$ are

$$\hat{\mathsf{u}}(s) = \breve{\mathbf{R}}^T(\hat{\alpha}s;\hat{\eta})\frac{d}{ds}\breve{\mathbf{R}}(\hat{\alpha}s;\hat{\eta}) = \hat{\alpha}\breve{\mathsf{u}},$$

where $\breve{\mathsf{u}} = (-\hat{\eta}, 0, \hat{\eta}_*)^T$.

Here we do not consider the case in which the director $\hat{\mathbf{d}}_2$ spins rapidly about the axis defined by $\hat{\mathbf{d}}_3$. However, the assumption that the director $\hat{\mathbf{d}}_2$ is the radius vector of the helix can be relaxed. For suppose that the director $\hat{\mathbf{d}}_2(s)$ maintains an angle $\beta(s)$ with the radius vector of the helix. Then the form of director frame $\hat{\mathcal{D}}$ in the reference configuration is given by the rotation matrix

$$\hat{\mathbf{R}}(s) = \breve{\mathbf{R}}(\hat{\alpha}s;\hat{\eta})\breve{\mathbf{R}}\big(\beta(s);0\big) = \mathbf{Q}(\hat{\alpha}s)\mathbf{S}_{\hat{\eta}}\mathbf{Q}\big(\beta(s)\big).$$

Assuming the director frame $\hat{\mathcal{D}}$ to be defined by the matrix $\breve{\mathbf{R}}(\hat{\alpha}s,\hat{\eta})$ only, we introduce a dependence on the angle $\beta(s)$ in constitutive relations (cf. also the remarks about the choice of the reference frame in Section 8.2).

In order to give the helical structure a local character relative to the total length of the rod we introduce a small parameter ϵ in the following way. Assume that the ratio of radius $\hat{\rho}$ to pitch $2\pi\hat{p}$ is fixed, say

$$\gamma = \frac{\hat{\rho}}{\hat{p}}, \tag{9.3}$$

i.e. $\hat{\eta} = \gamma/\sqrt{1+\gamma^2}$ is fixed. Now suppose $\hat{p}$ and $\hat{\rho}$ are small, i.e.

$$\epsilon = \sqrt{\hat{p}^2 + \hat{\rho}^2}.$$

Hence we have $\hat{\alpha} = 1/\epsilon$, which introduces the fast scale $\tau = \hat{\alpha}s = s/\epsilon$ in the rotation matrix (9.1) through $\mathbf{Q}(\hat{\alpha}s) = \mathbf{Q}(s/\epsilon)$.

Definition 9.1 can be generalized to account for an intrinsic shape in which the fast (helical) scale is superimposed on some arbitrary slow length scale (cf. Definition 8.1).

Definition 9.2. *The reference configuration* $\{\hat{\mathbf{r}}, \hat{\mathbf{R}}\}$ *of a rod is called locally helical if* $\hat{\mathbf{r}}'(s) = \hat{\mathbf{d}}_3(s)$ *and*

$$\hat{\mathbf{R}}(s;\epsilon) = \widetilde{\mathbf{R}}(s)\breve{\mathbf{R}}(\tau;\hat{\eta}),$$

where $\tau = s/\epsilon$ *and* $\hat{\eta} = \gamma/\sqrt{1+\gamma}$ *with* γ *being the local ratio of radius to pitch. In particular, the long scale evolution of the frame is determined by the rotation matrix* $\widetilde{\mathbf{R}}(s)$ *and the local, helical structure by the matrix* $\breve{\mathbf{R}}(\tau;\hat{\eta})$ *defined in* (9.1).

We have shown in Lemma 8.3 that the corresponding components of intrinsic strain satisfy

$$\hat{\mathsf{u}}(s,\tau) = \breve{\mathbf{R}}^T(\tau)\tilde{\mathsf{u}}(s) + \frac{1}{\epsilon}\breve{\mathsf{u}}, \tag{9.4}$$

where $\breve{\mathsf{u}} = (-\hat{\eta}, 0, \hat{\eta}_*)$ and $\tilde{\mathsf{u}}(s)\times = \widetilde{\mathbf{R}}^T(s)\frac{d}{ds}\widetilde{\mathbf{R}}(s)$.

The homogenization theory in Chapter 8 will be applied to this system to obtain explicitly an effective constitutive relation that is independent of the fast variable.

9.3 The Homogenized Hamiltonian System

We study the non-canonical Hamiltonian system (4.10) where the strains in the unstressed configuration have the form (9.4). In addition, we assume that the strain-energy function W does not explicitly depend on the fast scale τ. Following the analysis in Chapter 8 we define a new frame $\mathcal{N} = \{\mathbf{D}_i\}_{i=1}^3$ according to (8.2) with the matrix $\mathbf{A}$ taking the explicit form $\mathbf{A}(s;\epsilon) = \breve{\mathbf{R}}(s/\epsilon;\hat{\eta})$. Suppose for example the local structure of the rod is helical, and there is no additional slow length scale (cf. Figure 9.1(a)) so that the reference configuration of the rod is in fact exactly helical. Then the frame $\mathcal{N}$ is constant in the reference configuration and can be chosen to coincide with the inertial frame $\mathcal{E}$. In particular, the director $\mathbf{D}_3$ points in the direction of the helix axis $\mathbf{e}_3$ as shown in Figure 9.1(b).

The resulting Hamiltonian system has been derived in Section 8.2; in particular the Hamiltonian in the new variables is

$$\widetilde{H}(\mathsf{I},\mathsf{R},\mathsf{M},\mathsf{N},s,\tau) = W^*(\breve{\mathbf{R}}^T(\tau;\hat{\eta})\mathsf{M}, \breve{\mathbf{R}}^T(\tau;\hat{\eta})\mathsf{N}, s) + \mathsf{M}\cdot\tilde{\mathsf{u}}(s) + \mathsf{N}\cdot\breve{\mathbf{R}}(\tau;\hat{\eta})\mathsf{e}_3.$$

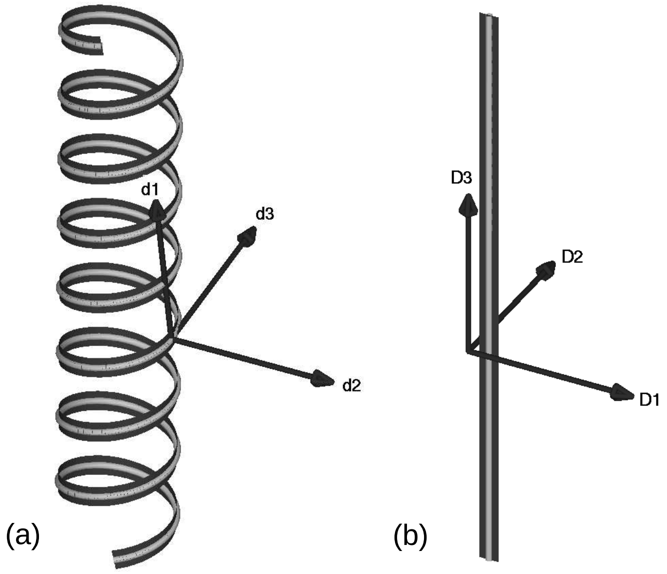

Figure 9.1: Fig. (a) shows a rod in its reference configuration that has a locally helical shape. The director frame $\mathcal{D} = \{\mathbf{d}_i\}_{i=1}^3$ is rapidly varying. In (b) we show the corresponding slow frame $\mathcal{N} = \{\mathbf{D}_i\}_{i=1}^3$. The rod configuration in (b) is obtained as the effective approximation.

The homogenization result in Theorem 8.1 implies that the effective Hamiltonian is computed as

$$\begin{aligned}\overline{H}(\mathsf{I},\mathsf{R},\mathsf{M},\mathsf{N},s) &= \frac{1}{2\pi}\int_0^{2\pi} \widetilde{H}(\mathsf{M},\mathsf{N},s,\tau)d\tau \qquad (9.5)\\ &= \frac{1}{2\pi}\int_0^{2\pi} \Big[W^*\big(\breve{\mathbf{R}}^T(\tau;\hat{\eta})\mathsf{M},\breve{\mathbf{R}}^T(\tau;\hat{\eta})\mathsf{N},s\big) + \mathsf{M}\cdot\tilde{\mathsf{u}}(s) + \mathsf{N}\cdot\breve{\mathbf{R}}(\tau;\hat{\eta})\mathsf{e}_3\Big]\,d\tau\\ &= \overline{W}^*(\mathsf{M},\mathsf{N},s) + \mathsf{M}\cdot\tilde{\mathsf{u}}(s) + \mathsf{N}\cdot\overline{\mathsf{V}},\end{aligned}$$

where

$$\overline{W}^*(\mathsf{M},\mathsf{N},s) = \frac{1}{2\pi}\int_0^{2\pi} W^*\big(\breve{\mathbf{R}}^T(\tau;\hat{\eta})\mathsf{M},\breve{\mathbf{R}}^T(\tau;\hat{\eta})\mathsf{N},s\big)d\tau \qquad (9.6)$$

and

$$\overline{\mathsf{V}} = \hat{\eta}_*\mathsf{e}_3.$$

with $\hat{\eta}_* = \sqrt{1-\hat{\eta}^2} = \hat{p}/\sqrt{\hat{\rho}^2+\hat{p}^2}$. According to Corollary 8.1 this system can be regarded as describing an effective rod model, since the component of strain

$\overline{V}_3 = \hat{\eta}_*$ is positive provided that the pitch $\hat{p}$ of the helix is positive. The fact that $\hat{\eta}_* < 1$ (whenever the radius of the helix $\hat{\rho}$ is strictly positive) implies that the parameter s does not parametrize arclength in the reference configuration. Thus we may introduce a new length scale $t = \hat{\eta}_* s$ which now plays the role of arclength in the reference configuration. The change in length scale induces a corresponding change in the constitutive relations.

As we have remarked before, in the effective approximation the helical (and rapidly varying) centerline of the rod is approximated by a slowly varying rod, and one may expect the extension and shearing effects, which are present in the helical structure, to be reflected in the approximation. However, the effective strain-energy function (9.6) shows that no additional extension and shearing effects are incorporated into the effective model. This becomes particularly obvious when considering an inextensible and unshearable rod so that the effective Hamiltonian reduces to

$$\overline{H}(\mathsf{I}, \mathsf{R}, \mathsf{M}, \mathsf{N}, s) = \overline{W}^*(\mathsf{M}, s) + \mathsf{M} \cdot \tilde{\mathsf{u}}(s) + \hat{\eta}_* \mathsf{N} \cdot \mathsf{e}_3,$$

with

$$\overline{W}^*(\mathsf{M}, s) = \frac{1}{2\pi} \int_0^{2\pi} W^*\big(\breve{\mathbf{R}}^T(\tau; \hat{\eta})\mathsf{M}, s\big)\, d\tau.$$

We see that the effective model does not include any extension and shearing effects whatsoever. We will investigate this conclusion in more detail in Section 10.4 within the context of closed form helical solutions. It will be seen that the extensional forces necessary to produce a helical deformation that extends the fast helix in the reference configuration become infinite as the small parameter ϵ tends to zero, which is an explanation of why extension effects do not arise in the homogenized model.

In the following sections we will derive important properties of the effective strain-energy function and obtain some fairly explicit representations if the rod is assumed to be linearly elastic.

9.4 Isotropy of the Constitutive Relation

For general non-quadratic strain-energy functions W, it does not seem feasible to explicitly compute a simple form for the effective strain-energy function $\overline{W}$. However, provided that the strain-energy function has no explicit dependence on the fast variable τ, it is possible to derive an important property of the averaged strain-energy function $\overline{W}$, namely $\overline{W}$ is invariant under rotations about the $\mathbf{D}_3$ axis. Equivalently the constitutive relation given by $\overline{W}$ is isotropic. Consequently the effective rod is isotropic precisely if the effective slowly varying curvature of the rod vanishes, i.e. $\tilde{u}_1(s) = \tilde{u}_2(s) = 0$.

Since $\overline{W}^*$, rather than $\overline{W}$ is given in the explicit form (9.6), it is convenient to first verify the rotational invariance of $\overline{W}^*$; the invariance of $\overline{W}$ then follows from the properties of the Legendre transform (see Appendix B). The discussion in this chapter is restricted to the case in which $\hat{\mathsf{v}} = \mathsf{e}_3$, or equivalently $\hat{\mathbf{r}}' = \mathbf{d}_3$, and thus we may adopt the characterization (2.12) of isotropy. The rotational invariance of $\overline{W}^*$ is characterized by the condition that

$$\overline{W}^*\big(\mathbf{Q}(\alpha)\mathsf{M}, \mathbf{Q}(\alpha)\mathsf{N}, s\big)$$

is independent of α for all $0 \le \alpha < 2\pi$, $\mathsf{M} = (M_1, M_2, M_3), \mathsf{N} = (N_1, N_2, N_3)$, where

$$\mathbf{Q}(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Using the decomposition $\breve{\mathbf{R}}(\tau;\hat{\eta}) = \mathbf{Q}(\tau)\mathbf{S}_{\hat{\eta}}$ for the rotation matrix (9.1) we calculate

$$\begin{aligned}
&2\pi\overline{W}^*\big(\mathbf{Q}(\alpha)\mathsf{M}, \mathbf{Q}(\alpha)\mathsf{N}, s\big) \\
&= \int_0^{2\pi} W^*\big(\breve{\mathbf{R}}^T(\tau;\hat{\eta})\mathbf{Q}(\alpha)\mathsf{M}, \breve{\mathbf{R}}^T(\tau;\hat{\eta})\mathbf{Q}(\alpha)\mathsf{N}, s\big)\,d\tau \\
&= \int_0^{2\pi} W^*\big(\mathbf{S}_{\hat{\eta}}\mathbf{Q}^T(\tau)\mathbf{Q}(\alpha)\mathsf{M}, \mathbf{S}_{\hat{\eta}}\mathbf{Q}^T(\tau)\mathbf{Q}(\alpha)\mathsf{N}, s\big)\,d\tau \\
&= \int_0^{2\pi} W^*\big(\mathbf{S}_{\hat{\eta}}\mathbf{Q}^T(\tau-\alpha)\mathsf{M}, \mathbf{S}_{\hat{\eta}}\mathbf{Q}^T(\tau-\alpha)\mathsf{N}, s\big)\,d\tau \\
&= \int_0^{2\pi} W^*\big(\breve{\mathbf{R}}^T(\tau-\alpha;\hat{\eta})\mathsf{M}, \breve{\mathbf{R}}^T(\tau-\alpha;\hat{\eta})\mathsf{N}, s\big)\,d\tau \\
&= \int_{-\alpha}^{2\pi-\alpha} W^*\big(\breve{\mathbf{R}}^T(\tau;\hat{\eta})\mathsf{M}, \breve{\mathbf{R}}^T(\tau;\hat{\eta})\mathsf{N}, s\big)\,d\tau \\
&= 2\pi\overline{W}^*(\mathsf{M}, \mathsf{N}, s),
\end{aligned}$$

since $W^*\big(\breve{\mathbf{R}}^T(\tau;\hat{\eta})\mathsf{M}, \breve{\mathbf{R}}^T(\tau;\hat{\eta})\mathsf{N}, s\big)$ has period 2π in τ. This demonstrates that the effective constitutive relation defined by $\overline{W}$ is isotropic.

9.5 Linearly Elastic Rods

For a linearly elastic rod the effective strain-energy function $\overline{W}$ can be computed in a fairly explicit form. In this case the strain-energy function is the quadratic form (2.9)

$$W(\mathsf{w}, \mathsf{z}, s) = \frac{1}{2}\begin{pmatrix}\mathsf{w}\\ \mathsf{z}\end{pmatrix}\cdot \mathbf{W}(s)\begin{pmatrix}\mathsf{w}\\ \mathsf{z}\end{pmatrix},$$

where $\mathbf{W} \in \mathbb{R}^{6\times 6}$ is a symmetric, positive-definite matrix (in particular, we do not assume $\mathbf{W}$ to be diagonal). The averaged strain-energy function is defined through (9.6), specifically

$$\overline{W}^*(\mathsf{M},\mathsf{N},s)=\frac{1}{2\pi}\int_0^{2\pi}W(\breve{\mathbf{R}}^T(\tau;\hat\eta)\mathsf{M},\breve{\mathbf{R}}^T(\tau;\hat\eta)\mathsf{N},s)d\tau=\frac{1}{2}\begin{pmatrix}\mathsf{M}\\ \mathsf{N}\end{pmatrix}\cdot\overline{\mathbf{W}}^{-1}(s)\begin{pmatrix}\mathsf{M}\\ \mathsf{N}\end{pmatrix},$$

with

$$\overline{\mathbf{W}}^{-1}(s)=\frac{1}{2\pi}\int_0^{2\pi}\begin{pmatrix}\breve{\mathbf{R}}(\tau;\hat\eta) & \mathbf{0}\\ \mathbf{0} & \breve{\mathbf{R}}(\tau;\hat\eta)\end{pmatrix}\mathbf{W}^{-1}(s)\begin{pmatrix}\breve{\mathbf{R}}^T(\tau;\hat\eta) & \mathbf{0}\\ \mathbf{0} & \breve{\mathbf{R}}^T(\tau;\hat\eta)\end{pmatrix}d\tau.$$

In particular, it involves the coefficients of the inverse matrix $\mathbf{W}^{-1}=(w_{ij})$, which we assume to be known. Then it can be calculated that the matrix $\overline{\mathbf{W}}^{-1}(s)$ has the following form

$$\overline{\mathbf{W}}^{-1}=\left(\begin{array}{ccc|ccc} a & 0 & 0 & e & g & 0\\ 0 & a & 0 & -g & e & 0\\ 0 & 0 & b & 0 & 0 & f\\ \hline e & -g & 0 & c & 0 & 0\\ g & e & 0 & 0 & c & 0\\ 0 & 0 & f & 0 & 0 & d \end{array}\right).$$

Recall the definition of the parameters $\hat\eta$ and $\hat\eta_*$:

$$\hat\eta=\frac{\gamma}{\sqrt{1+\gamma^2}},\qquad \hat\eta_*=\sqrt{1-\hat\eta^2}=\frac{1}{\sqrt{1+\gamma^2}},$$

with $\gamma=\frac{\hat\rho}{\hat p}$. Then the coefficients are

$$\begin{aligned}
a &= \frac{1}{2}\left(\hat\eta_*^2w_{11}+2\hat\eta\hat\eta_*w_{13}+w_{22}+\hat\eta^2w_{33}\right),\\
b &= \hat\eta^2w_{11}-2\hat\eta\hat\eta_*w_{13}+\hat\eta_*^2w_{33},\\
c &= \frac{1}{2}\left(\hat\eta_*^2w_{44}+2\hat\eta\hat\eta_*w_{46}+w_{55}+\hat\eta^2w_{66}\right),\\
d &= \hat\eta^2w_{44}-2\hat\eta\hat\eta_*w_{46}+\hat\eta_*^2w_{66},\\
e &= \frac{1}{2}\left(\hat\eta_*^2w_{14}+\hat\eta\hat\eta_*w_{16}+w_{25}+\hat\eta\hat\eta_*w_{34}+\hat\eta^2w_{36}\right),\\
f &= \hat\eta^2w_{14}-\hat\eta\hat\eta_*w_{16}-\hat\eta\hat\eta_*w_{34}+\hat\eta_*^2w_{36},\\
g &= \frac{1}{2}\left(\hat\eta_*w_{15}-\hat\eta_*w_{24}-\hat\eta w_{26}+\hat\eta w_{35}\right).
\end{aligned}$$

It can be shown that the matrix $\overline{\mathbf{W}}$ has the same sparsity pattern as $\overline{\mathbf{W}}^{-1}$, and we conclude that the structure of the matrix $\overline{\mathbf{W}}$ is exactly in the form required in Lemma 2.1. Therefore the effective constitutive equation defined by $\overline{W}$ is isotropic, in concordance with the result in Section 9.4.

The representation of the effective strain-energy function becomes even more explicit if the strain-energy is the diagonal quadratic form

$$W(\mathsf{w}, \mathsf{z}, s) = \frac{1}{2}\sum_{i=1}^{3}(K_i w_i^2 + A_i z_i^2).$$

In this case the effective strain-energy function $\overline{W}$ can be computed explicitly in terms of the coefficients K_i and A_i, namely

$$\overline{W}(\mathsf{w}, \mathsf{z}, s) = \frac{1}{2}\sum_{i=1}^{3}\bar{K}_i(s)w_i^2 + \frac{1}{2}\sum_{i=1}^{3}\bar{A}_i(s)z_i^2,$$

where

$$\begin{aligned}
\bar{K}_1^{-1} = \bar{K}_2^{-1} &= \frac{1}{2}\left(K_1^{-1}\hat{\eta}_*^2 + K_2^{-1} + K_3^{-1}\hat{\eta}^2\right), \\
\bar{K}_3^{-1} &= K_1^{-1}\hat{\eta}^2 + K_3^{-1}\hat{\eta}_*^2 \\
\bar{A}_1^{-1} = \bar{A}_2^{-1} &= \frac{1}{2}\left(A_1^{-1}\hat{\eta}_*^2 + A_2^{-1} + A_3^{-1}\hat{\eta}^2\right), \\
\bar{A}_3^{-1} &= A_1^{-1}\hat{\eta}^2 + A_3^{-1}\hat{\eta}_*^2.
\end{aligned}$$

Note that in this case it is immediately obvious that $\overline{W}$ defines an isotropic constitutive relation, since $\bar{K}_1 = \bar{K}_2$ and $\bar{A}_1 = \bar{A}_2$.

9.6 Elastic Rods with High Intrinsic Twist

We now study an example which arises as a limiting case of the model defined in Section 9.2. By considering a rod with locally helical intrinsic shape as the radius of the helix $\hat{\rho}$ vanishes, we obtain a rod with high intrinsic twist. This model is of particular importance for the continuum DNA model (see Manning et al. [44]).

Consistent with Definitions 9.1 and 9.2, we consider a rod that exhibits an intrinsic shape with two different length scales:

Definition 9.3. *A rod is said to have an intrinsic twist with rate β if the reference configuration of the rod is $\{\hat{\mathbf{r}}, \hat{\mathbf{R}}\}$ with $\hat{\mathbf{r}}' = \hat{\mathbf{d}}_3$ and $\hat{\mathbf{R}}(s) = \breve{\mathbf{R}}(\beta s, 0)$, where $\breve{\mathbf{R}}$ is defined in* (9.1)*, i.e.*

$$\breve{\mathbf{R}}(\tau, 0) = \begin{pmatrix} \cos\tau & -\sin\tau & 0 \\ \sin\tau & \cos\tau & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The rod is locally twisted, if a high intrinsic twist is superimposed on some arbitrary slow length scale, i.e. $\hat{\mathbf{R}}(s;\epsilon) = \widetilde{\mathbf{R}}(s)\breve{\mathbf{R}}(\tau, 0)$, where $\tau = s/\epsilon$.

Then the strains $\hat{\mathsf{u}}$ have the form

$$\hat{\mathsf{u}}(s,\tau,\epsilon) = \breve{\mathbf{R}}^T(\tau,0)\tilde{\mathsf{u}}(s) + \frac{1}{\epsilon}\breve{\mathsf{u}},$$

where $\breve{\mathsf{u}} = (0,0,1)^T$, and $\tilde{\mathsf{u}}$ is defined through

$$\frac{d}{ds}\widetilde{\mathbf{R}}(s) = \widetilde{\mathbf{R}}(s)\tilde{\mathsf{u}}(s)\times.$$

In this case the frame $\mathcal{N}$ introduced in Lemma 8.4 is often called the natural frame (see Langer and Singer [37]).

The effective rod model is immediately obtained by using the results of Section 9.3 and setting $\hat{\rho} = 0$. The effective Hamiltonian is given by (9.5) with $\hat{\eta} = 0$, i.e.

$$\widetilde{H}(\mathsf{I},\mathsf{R},\mathsf{M},\mathsf{N},s) = \overline{W}^*(\mathsf{M},\mathsf{N},s) + \mathsf{M}\cdot\tilde{\mathsf{u}}(s) + \mathsf{N}\cdot\mathsf{e}_3,$$

where

$$\overline{W}^*(\mathsf{M},\mathsf{N},s) = \frac{1}{2\pi}\int_0^{2\pi} W^*\big(\breve{\mathbf{R}}^T(\tau;0)\mathsf{M},\breve{\mathbf{R}}^T(\tau;0)\mathsf{N},s\big)d\tau.$$

We now consider some specific cases. Consider an inextensible, unshearable, linearly elastic rod with quadratic strain-energy function

$$W(\mathsf{w},s) = \frac{1}{2}\mathsf{w}\cdot\mathbf{K}(s)\mathsf{w}, \quad \mathbf{K} = \begin{pmatrix} K_1 & K_{12} & K_{13} \\ K_{12} & K_2 & K_{23} \\ K_{13} & K_{23} & K_3 \end{pmatrix}.$$

Following the computations in Section 9.5 we compute the effective strain-energy function

$$\overline{W}(\mathsf{w},s) = \frac{\bar{K}}{2}(w_1^2 + w_2^2) + \frac{\bar{K}_3}{2}w_3^2$$

with

$$\begin{aligned} \bar{K} &= 2\frac{-\left(K_{13}^2K_2\right) + 2K_{12}K_{13}K_{23} - K_1K_{23}^2 - K_{12}^2K_3 + K_1K_2K_3}{-K_{13}^2 - K_{23}^2 + K_1K_3 + K_2K_3}, \\ \bar{K}_3 &= \frac{-\left(K_{13}^2K_2\right) + 2K_{12}K_{13}K_{23} - K_1K_{23}^2 - K_{12}^2K_3 + K_1K_2K_3}{-K_{12}^2 + K_1K_2}. \end{aligned}$$

Thus for the diagonal strain-energy function (2.11),

$$W(\mathsf{w},s) = \frac{1}{2}\sum_{i=1}^{3} K_i w_i^2$$

we obtain the particularly simple effective coefficients

$$\bar{K} = \frac{2K_1K_2}{K_1+K_2}, \tag{9.7}$$

$$\bar{K}_3 = K_3, \tag{9.8}$$

i.e. the effective bending stiffness $\bar{K}$ is obtained as the harmonic average of the original distinct bending moduli. Suppose without loss of generality that $K_1 \leq K_2$, then

$$\frac{\bar{K}}{K_1} = \frac{2}{1 + \frac{K_1}{K_2}},$$

i.e. $1 \leq \bar{K}/K_2 \leq 2$, with $\bar{K}$ monotone increasing from K_1 to $2K_1$ as K_2 increases. Thus in particular if one bending stiffness is much larger than the other, say $K_2 \gg K_1$, then the effective bending stiffness is $\bar{K} \approx 2K_1$.

Finally we would like to comment on the implications of the averaging result on the material parameters. With regard to the twist stiffness recall (cf. Section 2.2) that for a linearly elastic material with diagonal energy (2.11), the bending and twisting moduli can be interpreted as

$$K_i = EI_i, \ i = 1, 2, \quad K_3 = G(I_1 + I_2), \quad G = \frac{E}{2(1+\nu)},$$

where ν is Poisson's ratio, E the elastic modulus, G the shear modulus and I_1 and I_2 are the geometric moments about the axes $\mathbf{d}_1$ and $\mathbf{d}_2$. For an isotropic rod the ratio of twist to bending stiffness is

$$\varrho = K_3/K_1 = \frac{1}{1+\nu}.$$

For a typical material one assumes Poisson's ratio to lie in the range $(0, \frac{1}{2})$, which implies the possible range for $\varrho \in (\frac{2}{3}, 1)$. In this section we have shown that for an anisotropic rod with high intrinsic twist, the material is effectively isotropic with the effective coefficients (9.7). This implies a ratio $\bar{\varrho}$ for the effective moduli

$$\bar{\varrho} = \frac{\bar{K}_3}{\bar{K}_1} = \frac{I_1 + I_2}{4(1+\nu)}\left(\frac{1}{I_1} + \frac{1}{I_2}\right) = \frac{(1+\delta)^2}{4\delta}\frac{1}{1+\nu},$$

where $\delta = I_1/I_2$. Regarding δ as the only free parameter, the possible range for $\bar{\varrho}$ is $(1/(1+\nu), \infty)$ or $\bar{\varrho} \in (\frac{2}{3}, \infty)$ for $\nu \in (0, \frac{1}{2})$. Thus for effectively isotropic rods higher ratios of twist to bending stiffness can be anticipated than for prismatic rods. Rogers [53] investigates the stability properties of closed loop solutions of isotropic, uniform, inextensible and unshearable, linearly elastic rods and finds that interesting stability exchanges occur for values of $\varrho = K_3/K_1 > 1$.

9.7 Numerical Experiments

We now illustrate our analysis of effective constitutive relations with some computations that compare the numerical solutions of averaged and unaveraged systems. We consider only the (diagonal) linearly elastic energy (2.11), and further restrict

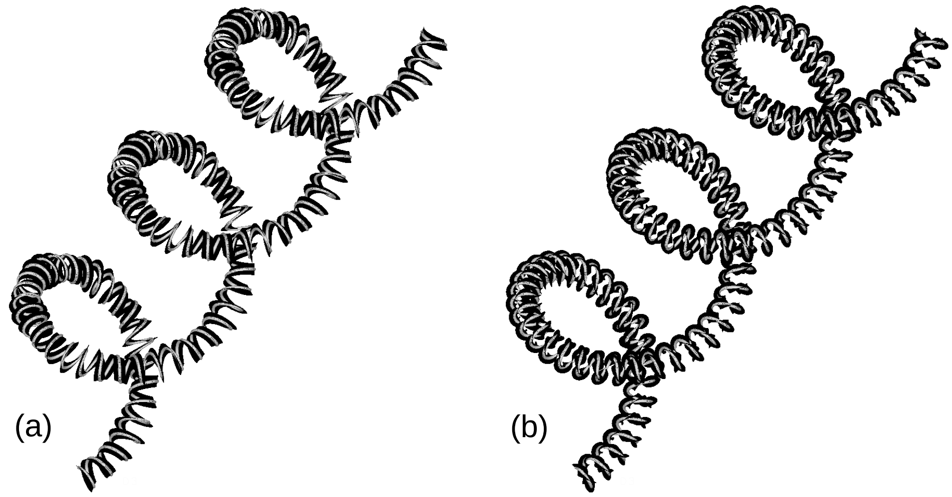

Figure 9.2: An integrable rod equilibrium on a relatively long length scale. Part (a) shows the typical helix-on-helix structure with no intrinsic twist. In (b) the intrinsic twist is increased, causing the frame to rotate more rapidly, but the configuration of the centerline is unchanged.

ourselves to uniform rods in which the coefficient functions reduce to constants. For the purpose of these calculations we use the canonical Hamiltonian description in terms of Euler parameters (cf. Section 4.3). The first set of numerical examples illustrates the homogenization result for rods with high intrinsic twist; we will then add an example in which the intrinsic shape is locally helical.

Consider first the model introduced in Section 9.6. We assume that the rod is straight ($\hat{u}_1 = \hat{u}_2 = 0$) in its reference configuration, but highly twisted with rate $\hat{u}_3$. In all cases we hold the initial conditions of the Hamiltonian system fixed, but vary the bending moduli K_1 and K_2, and the intrinsic twist rate $\hat{u}_3$. When interpreting these computational experiments it is useful to be aware of certain integrable cases and special solutions of the system with Hamiltonian (4.9) corresponding to the particular strain-energy (2.11). Recall (cf. Chapter 5) that the uniform rod with Hamiltonian (4.9) is completely integrable only when the rod is isotropic, i.e. the constitutive relation is isotropic ($K_1 = K_2$) and the rod is straight in the minimum energy configuration ($\hat{u}_1 = \hat{u}_2 = 0$). In the integrable case a general solution to the initial value problem can be found in a closed, but rather complicated, form involving elliptic functions (see Section 7.2), and the centerline of the general solution has the qualitative form of a helix-on-helix that is illustrated in Figure 9.2(a). One particularly simple solution to the integrable problem is the twisted loop, in which the centerline is a circle (cf. Figure 9.3(a)).

For the integrable rod the twist of the director frame about the tangent largely decouples from the shape of the centerline. In particular, as illustrated in Figure 9.2 and Figure 9.3, if the natural twist rate is varied for an integrable rod, the only change in the equilibrium configuration is a local rotation of the frame about the centerline, and the shape of the centerline is unaffected.

We now study the effects of anisotropy and intrinsic twist for the initial conditions giving rise to the twisted circle (cf. Figure 9.3(a); here $K_1 = K_2 = 0.1$). If no intrinsic twist is present, breaking isotropy, and therefore integrability, by taking $K_1 \neq K_2$ results in a configuration rather dissimilar to the twisted circle, see Figure 9.3(d). Here we have taken $K_1 = 0.2$ and $K_2 = 0.0666$. However, because we have chosen stiffnesses with harmonic average $\bar{K} = 0.1$, as the intrinsic twist is increased, the configuration better approximates the closed loop with one effective isotropic bending stiffness $K_1 = K_2 = 0.1$ (cf. Figure 9.3(e) and (f) and the effective approximations depicted in Figure 9.3(b) and (c)).

We can also compare the quality of the effective approximation by considering numerical configurations in which we hold the twist rate fixed, but vary the bending stiffnesses subject to the constraint that the harmonic average is prescribed (in our case $\bar{K} = 0.1$). The sequences Figure 9.3(b), Figure 9.3(e), and Figure 9.3(g), and Figure 9.3(c), Figure 9.3(f), and Figure 9.3(h), illustrate the results: as the bending stiffnesses are taken to be increasingly anisotropic, the approximation for fixed intrinsic twist worsens.

Furthermore we study the effects of anisotropy and high twist for the more general solution of the Hamiltonian system depicted in Figure 9.2(a). In Figure 9.4(a) we see that (as before) the integrable structure is destroyed if the bending stiffnesses are unequal and no intrinsic twist is present. However, the behaviors on short and long scales are very different. Despite nonintegrability, the long term structure is quite regular, and resembles the integrable case shown in Figure 9.2(a). However, the local structure, as shown in Figure 9.4(b), can be rather complicated. Finally, we increase the intrinsic twist to see that the resulting configuration (cf. Figure 9.4(c)) is very similar to its effective approximation illustrated in Figure 9.2(b).

As a last example we assume the intrinsic shape of the rod to be locally helical, so that the effective rod model must be a straight rod without twist. We choose initial conditions such that the effective rod is bent into a circular loop (see Figure 9.5(f)) . These initial conditions are then applied to the unhomogenized model, interpreting the initial conditions appropriately in the fast frame. We conclude that as ϵ is decreased we obtain a better and better approximation to the loop configuration (cf. Figure 9.5).

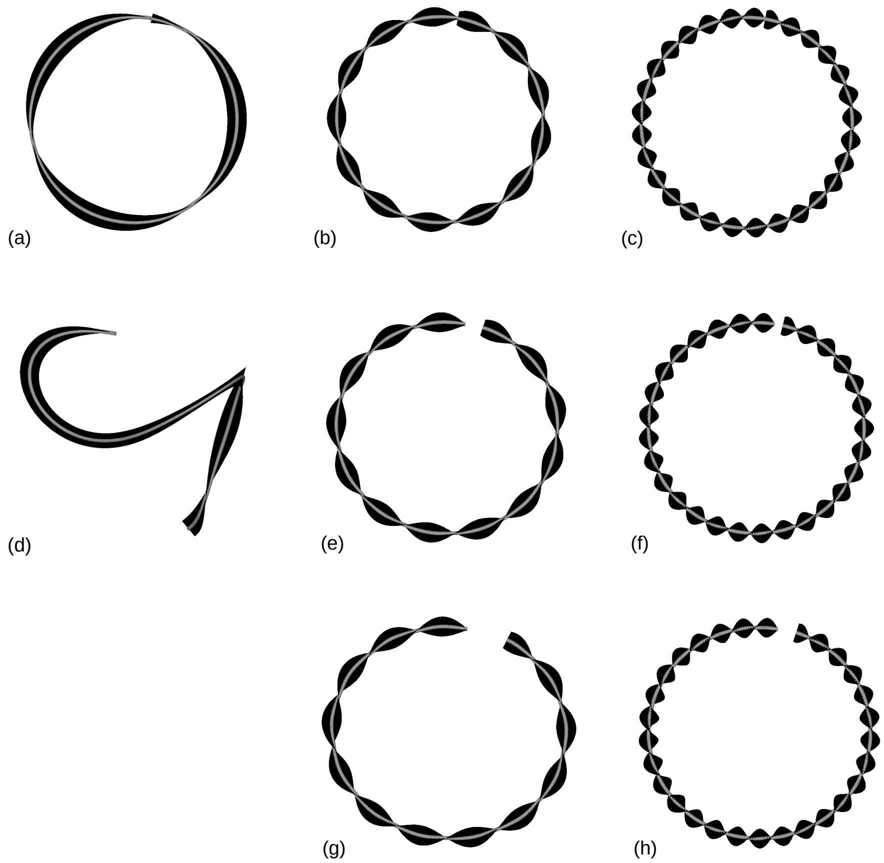

Figure 9.3: A closed twisted circular equilibrium (a) of an integrable rod (with $K_1 = K_2 = 0.1$). Integrability implies that the position of the centerline does not change as the intrinsic twist is increased: (a) $\hat{u}_3 = 0$, (b) $\hat{u}_3 = 50$ and (c) $\hat{u}_3 = 100$. For anisotropic bending stiffnesses $K_1 = 0.2$, and $K_2 = 0.0666$, the effective averaged bending stiffness is $\bar{K} = 0.1$. In the absence of intrinsic twist (cf. (d)) the solution is dissimilar to its counterpart (a). As the twist is increased ((e) $\hat{u}_3 = 50$ and (f) $\hat{u}_3 = 100$) the resemblance with the effective averaged solutions becomes apparent ((b) and (c)). For $K_1 = 1.0$, $K_2 = 0.0526$, $\bar{K} = 0.1$ with high twist ($\hat{u}_3 = 50$ in (g) and $\hat{u}_3 = 100$ in (h)), the approximation is worse than then in (e) and (f), since the anisotropy of the bending moduli is larger.

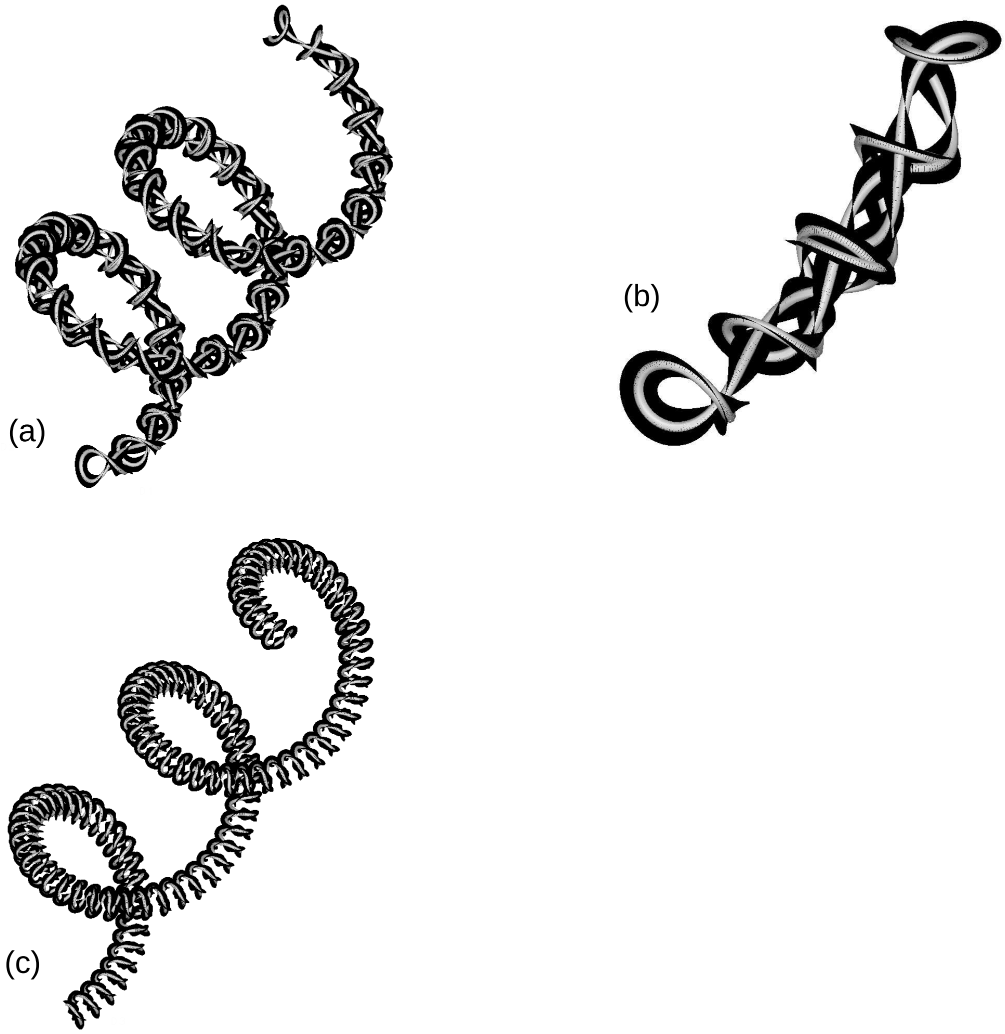

Figure 9.4: Equilibrium of a nonintegrable rod (straight in the minimum energy configuration, but with unequal bending stiffnesses). Despite nonintegrability the long term averaged behavior is still quite regular (cf. part (a)). However, the local structure, as indicated in part (b), may be rather complicated. Increased intrinsic twist leads to an approximately integrable structure (cf. part(c) and Figure 9.2).

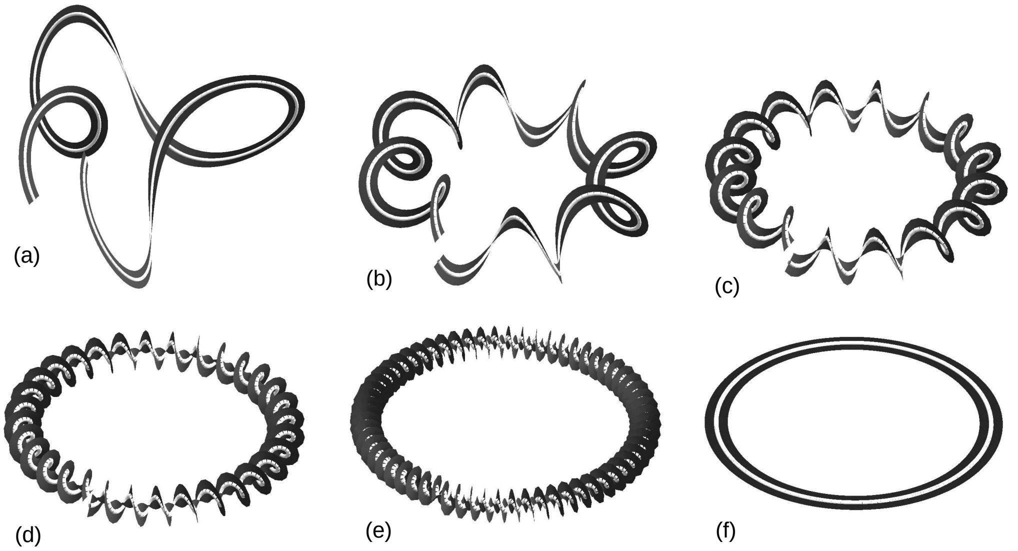

Figure 9.5: The linearly elastic, inextensible and unshearable rod of length 1 has a helical structure according to Definition 9.2 with $\hat{\rho}/\hat{p} = 0.5$, $K_1 = K_3 = 0.1$, $K_2 = 0.3$. We investigate the behavior as $\epsilon^2 = \hat{\rho}^2 + \hat{p}^2$ becomes small (in (a) $\epsilon = 0.05$, (b) $\epsilon = 0.02$, (c) $\epsilon = 0.01$, (d) $\epsilon = 0.005$, (e) $\epsilon = 0.0025$). The effective rod is shown is (f): it has length 0.447 and effective moduli $\bar{K}_1 = \bar{K}_2 = 0.15$, $\bar{K}_3 = 0.1$. In its intrinsic state it is a straight and untwisted rod. We apply initial conditions so that the rod is bent into a circular loop ($m_1 = -2.10744$, $m_2 = m_3 = 0$, $\mathbf{n} = \mathbf{0}$). The same initial conditions are then applied to the helical rod, but appropriately interpreted in the fast moving frame.

Chapter 10

More General Helical Rod Models

In the analysis of locally helical rod models we made several assumptions on the form of the helix and the relative magnitude of its parameters, and in this chapter we will examine how changes in these assumptions affect the averaging results obtained in Chapter 9. We will also consider our homogenization results in light of various families of closed-form helical solutions to the unaveraged system.

First, according to Definition 9.2 the locally helical rod model, and the analysis thereafter, relied on the assumption that the director $\hat{\mathbf{d}}_2$ in the reference configuration takes a fixed angle with the radius vector of the helix. One possible extension of this model is to allow the frame $\hat{\mathcal{D}}$ to have an additional (fast) rotation relative to the fast helical scale, and we show how to obtain effective systems for this model in Section 10.1.

Second, we obtained the locally helical character of the intrinsic shape by assuming the ratio of radius and pitch of the helix to be fixed, so that both are small of the same order in the parameter ϵ. This assumption can be relaxed by assuming radius and pitch to converge to zero with different orders as ϵ tends to zero. The implications of this generalization will be described in Section 10.2.

In Section 10.3 we show how to obtain a two-parameter family of closed-form solutions for helical deformations of linearly elastic rods with helical intrinsic shape. This particular example will be exploited in Section 10.4 to illustrate why our homogenization techniques do not lead to any extension and shearing effects. The closed form solutions also give some insight into the different behaviors of the limiting problems described in Section 10.2.

10.1 Locally Helical Models with High Intrinsic Twist

In this section the homogenization result for locally helical rods will be extended by incorporating an additional high intrinsic twist. Recall (cf. Definition 9.2) that the reference configuration of a locally helical rod is given by $\{\hat{\mathbf{r}}, \hat{\mathbf{R}}\}$, where

$\hat{\mathbf{r}}' = \hat{\mathbf{d}}_3$ and $\hat{\mathbf{R}}(s;\epsilon) = \widetilde{\mathbf{R}}(s)\breve{\mathbf{R}}(\tau;\hat{\eta})$ with $\tau = s/\epsilon$,

$$\breve{\mathbf{R}}(\tau;\hat{\eta}) = \mathbf{Q}(\tau)\mathbf{S}_{\hat{\eta}} = \begin{pmatrix} \cos\tau & \sin\tau & 0 \\ -\sin\tau & \cos\tau & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{\eta}_* & 0 & \hat{\eta} \\ 0 & 1 & 0 \\ -\hat{\eta} & 0 & \hat{\eta}_* \end{pmatrix},$$

$\hat{\eta} = \gamma/\sqrt{1+\gamma^2}$, and $\gamma = \hat{\rho}/\hat{p}$ is the fixed ratio of radius $\hat{\rho}$ to pitch $2\pi\hat{p}$.

Assume that an additional (fast) intrinsic twist is superimposed on the (fast) helical shape. Then the form of the director frame in the reference configuration is more generally defined by

$$\hat{\mathbf{R}}(s;\epsilon) = \widetilde{\mathbf{R}}(s)\breve{\mathbf{R}}(\tau;\hat{\eta})\breve{\mathbf{R}}(\beta\tau;0) = \widetilde{\mathbf{R}}(s)\mathbf{Q}(\tau)\mathbf{S}_{\hat{\eta}}\mathbf{Q}(\beta\tau).$$

The rotation matrix $\widetilde{\mathbf{R}}(s)$ describes the (arbitrary) slow length scale, and $\mathbf{T}(\tau) = \mathbf{Q}(\tau)\mathbf{S}_{\hat{\eta}}\mathbf{Q}(\beta\tau)$ the fast scale, where in particular $\mathbf{Q}(\beta\tau)$ corresponds to a fast rotation by an angle $\beta\tau$ about the director $\hat{\mathbf{d}}_3$.

Following the analysis in Chapter 8 we define a new frame $\mathcal{N}$ according to (8.2) via the matrix $\mathbf{T}(\tau)$, so that the associated strain vector is bounded as ϵ tends to zero. The change of variables (8.17) transforms the Hamiltonian in a form to which averaging can be applied:

$$\widetilde{H}(\mathbf{Z}, s, \tau) = W^*\left(\mathbf{T}^T(\tau)\mathsf{M}, \mathbf{T}^T(\tau)\mathsf{N}, s\right) + \mathsf{M}\cdot\tilde{\mathsf{u}}(s) + \mathsf{N}\cdot\mathbf{T}(\tau)\mathsf{e}_3.$$

However, contrary to the calculations in Chapter 9 the Hamiltonian is not necessarily a periodic function in the variable τ, but rather quasi-periodic with frequencies $\omega_1 = 1$ and $\omega_2 = \beta$, i.e.

$$\widetilde{H}(\mathbf{Z}, s, \tau) = H^\sharp(\mathbf{Z}, s, \tau, \beta\tau),$$

where $H^\sharp(\mathbf{Z}, s, \tau, \sigma)$ is periodic with period 2π in the variables τ and σ (recall the discussion in Section 8.4). Provided the limit

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \widetilde{H}(\mathbf{Z}, s, \tau)d\tau \tag{10.1}$$

exists, averaging results can still be obtained, although it is in general difficult to explicitly determine the limit (10.1).

Explicit computations can be performed if we assume in addition that the frequencies ω_1 and ω_2 are commensurable, i.e. there exist integers m and n, such that

$$\frac{\omega_2}{\omega_1} = \beta = \frac{m}{n}.$$

Then the function $\widetilde{H}(\mathbf{Z}, s, \tau)$ has period $2n\pi$, since

$$\begin{aligned}
&\widetilde{H}(\mathbf{Z}, s, \tau + 2n\pi) \\
&\quad = \; H^\sharp(\mathbf{Z}, s, \tau + 2n\pi, \beta(\tau + 2n\pi)) \\
&\quad = \; H^\sharp(\mathbf{Z}, s, \tau, \beta\tau + 2m\pi) \\
&\quad = \; H^\sharp(\mathbf{Z}, s, \tau, \beta\tau) \\
&\quad = \; \widetilde{H}(\mathbf{Z}, s, \tau).
\end{aligned}$$

In this case the averaged Hamiltonian is computed as

$$\overline{H}(\mathbf{Z}, s) = \frac{1}{2n\pi} \int_0^{2n\pi} \widetilde{H}(\mathbf{Z}, s, \tau) d\tau.$$

10.2 Locally Helical Models with Variable Ratio of Radius to Pitch

In Definition 9.2 we introduced the reference configuration $\{\hat{\mathbf{r}}, \hat{\mathbf{R}}\}$ of a locally helical rod, where $\hat{\mathbf{r}}' = \hat{\mathbf{d}}_3$ and

$$\hat{\mathbf{R}}(s; \epsilon) = \widetilde{\mathbf{R}}(s)\breve{\mathbf{R}}(\hat{\alpha}s; \hat{\eta}).$$

The rotation matrix $\widetilde{\mathbf{R}}(s)$ is arbitrary and $\breve{\mathbf{R}}(\tau; \hat{\eta})$ is defined by

$$\breve{\mathbf{R}}(\tau; \hat{\eta}) = \mathbf{Q}(\tau)\mathbf{S}_{\hat{\eta}} = \begin{pmatrix} \cos\tau & \sin\tau & 0 \\ -\sin\tau & \cos\tau & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{\eta}_* & 0 & \hat{\eta} \\ 0 & 1 & 0 \\ -\hat{\eta} & 0 & \hat{\eta}_* \end{pmatrix},$$

where $\hat{\alpha} = 1/\sqrt{\hat{\rho}^2 + \hat{p}^2}$, $\hat{\eta} = \hat{\rho}\hat{\alpha}$, $\hat{\eta}_* = \hat{p}\hat{\alpha}$. In particular, the local character of the helical structure is introduced by assuming the radius $\hat{\rho}$ and pitch $2\pi\hat{p}$ of the helix to be small. The ratio (9.3) of radius to pitch is assumed to be constant, i.e. $\gamma = \hat{\rho}/\hat{p}$ is constant, and the small parameter is introduced by setting

$$\frac{1}{\hat{\alpha}} = \sqrt{\hat{\rho}^2 + \hat{p}^2} = \epsilon.$$

As a consequence the rotation matrix $\breve{\mathbf{R}}(\tau; \hat{\eta})$ does not depend on the small parameter ϵ, and our analysis of effective locally helical rod models in Chapter 9 depends on this assumption. We will now assume the radius and pitch of the helix to be of different orders in ϵ, and study its implications.

More generally we may assume that the parameters of the helix are chosen to be

$$\hat{\rho} = \epsilon^{\kappa_1}, \quad \hat{p} = c\epsilon^{\kappa_2}, \quad \kappa_1, \kappa_2 \geq 0,$$

so that $\breve{\mathbf{R}}(\tau;\hat{\eta}) = \mathbf{Q}(\tau)\mathbf{S}_{\hat{\eta}(\epsilon)}$ now has an implicit ϵ-dependence through

$$\hat{\eta}(\epsilon) = \frac{\gamma(\epsilon)}{\sqrt{1+\gamma^2(\epsilon)}} = \frac{\epsilon^{\kappa_1}}{\sqrt{\epsilon^{2\kappa_1} + c^2\epsilon^{2\kappa_2}}}.$$

Then the strains (9.4) are of the form

$$\hat{\mathsf{u}}(s;\epsilon) = \breve{\mathbf{R}}^T\big(\tau;\hat{\eta}(\epsilon)\big)\tilde{\mathsf{u}}(s) + \hat{\alpha}(\epsilon)\breve{\mathsf{u}}(\tau;\epsilon),$$

where $\breve{\mathsf{u}}(\tau;\epsilon) = \big(-\hat{\eta}(\epsilon), 0, \hat{\eta}_*(\epsilon)\big)^T$ with

$$\hat{\alpha}(\epsilon) = \frac{1}{\sqrt{\epsilon^{2\kappa_1} + c^2\epsilon^{2\kappa_2}}}.$$

The order of $\hat{\alpha}(\epsilon)$ can be seen to be

$$\hat{\alpha}(\epsilon) = O\left(\frac{1}{\epsilon^{\min(\kappa_1,\kappa_2)}}\right).$$

Thus if $\kappa_1 > 0$ and $\kappa_2 > 0$ then $\tau = \hat{\alpha}s$ is a fast variable, and averaging over this variable yields a homogenized system. However, if $\kappa_1 = 0$ or $\kappa_2 = 0$, then $\hat{\alpha}$ is bounded as ϵ tends to zero and there is no fast variation in $\hat{\mathbf{R}}(\tau;\hat{\eta})$. In this case the Hamiltonian is a smooth function in ϵ and the behavior of the Hamiltonian system (4.10) can be determined directly without applying the theory of averaging.

Thus for the following analysis we assume that radius and pitch have the form

$$\hat{\rho} = \epsilon, \quad \hat{p} = c\epsilon^{\kappa},$$

where $\kappa > 0$. The effective Hamiltonian (9.5) is

$$\overline{H}(\mathsf{M},\mathsf{N},s) = \overline{W}^*(\mathsf{M},\mathsf{N},s) + \mathsf{M}\cdot\tilde{\mathsf{u}}(s) + \mathsf{N}\cdot\overline{\mathsf{V}}(s;\epsilon),$$

where

$$\overline{W}^*(\mathsf{M},\mathsf{N},s) = \frac{1}{2\pi}\int_0^{2\pi} W^*\Big(\breve{\mathbf{R}}^T\big(\tau,\hat{\eta}(0)\big)\mathsf{M}, \breve{\mathbf{R}}^T\big(\tau,\hat{\eta}(0)\big)\mathsf{N}, s\Big)d\tau$$

and

$$\begin{aligned}
\overline{\mathsf{V}}(s;\epsilon) &= \frac{1}{2\pi}\int_0^{2\pi} \breve{\mathbf{R}}(\tau,\hat{\eta}(0))d\tau\ \mathsf{e}_3 \\
&= \frac{1}{2\pi}\int_0^{2\pi} \mathbf{Q}(\tau)d\tau\ \mathbf{S}_{\hat{\eta}(\epsilon)}\mathsf{e}_3 \\
&= \mathbf{S}_{\hat{\eta}(\epsilon)}\mathsf{e}_3.
\end{aligned}$$

Thus depending on the size of κ the implications of this result are different:

- If $\kappa = 1$ the rotation matrix $\breve{\mathbf{R}}$ does not depend on ϵ, so that we recover the case discussed in Chapter 9.
- If $\kappa > 1$ then $\hat{\eta}(0) = 0$, so that the rotation matrix $\breve{\mathbf{R}}(\tau, 0)$ describes a rod with high intrinsic twist. A derivation of the corresponding homogenized system can be found in Section 9.6.
- The most interesting effect happens when $0 < \kappa < 1$. Now $\hat{\eta}(0) = 1$ and the effective components of strain become $\overline{\mathsf{V}} = (0, 0, 0)$. We recognize that this result violates the assumption that the constitutive relations should be formulated so that any deformation is orientation preserving (cf. Section 2.2 and Corollary 8.1). Thus the homogenized system does not constitute a valid rod model. The fact that in the reference configuration a rod of finite length degenerates into a rod of zero length as $\epsilon \to 0$ suggests that the problem should be formulated on a different length scale. One choice for a new length scale is $t = s/\hat{\eta}_*(\epsilon)$, so that t parametrizes the height of the helix rather than arclength. However, this scaling of the independent variable implies that the right-hand-side of the differential equations defined in Lemma 8.2 are multiplied by $\hat{\eta}_*(\epsilon)$, so that the problem is not in a form such that averaging can be applied. An effective model must therefore be obtained through different means.

10.3 Some Closed Form Helical Solutions for Rods with Intrinsic Helical Shape

In this section we construct closed-form solutions for helical equilibrium states of linearly elastic rods that are helical in their reference (unstressed) configuration. Specifically we assume the reference configuration to be defined according to Definition 9.1, i.e. in particular we assume that the radius vector $\hat{\mathbf{d}}_2$ traces the radius vector of the helix. Assuming the rod is inextensible and unshearable, and linearly elastic with diagonal strain-energy function (2.11) there exists a two-parameter family of helical deformations in which the director $\mathbf{d}_2$ in the stressed configuration necessarily traces the radius vector of the helix. More precisely, the stressed configuration $\{\mathbf{r}, \mathbf{R}\}$ (with $\mathbf{R} = (\underline{\mathbf{d}}_1, \underline{\mathbf{d}}_2, \underline{\mathbf{d}}_3)$) is helical if and only if

$$\mathbf{r}(s) = \frac{p}{\sqrt{\rho^2 + p^2}} s\mathbf{e}_3 - \rho \mathbf{d}_2(s).$$

With the rotation matrix (9.1)

$$\breve{\mathbf{R}}(\tau; \eta) = \begin{pmatrix} \eta_* \cos\tau & -\sin\tau & \eta\cos\tau \\ \eta_* \sin\tau & \cos\tau & \eta \sin\tau \\ -\eta & 0 & \eta_* \end{pmatrix}, \tag{10.2}$$

and $\eta_* = \sqrt{1-\eta^2}$, we can characterize any possible helical deformation of the rod:

Lemma 10.1. *Suppose the unstressed configuration of a rod is helical with radius $\hat\rho$ and pitch $2\pi\hat p$ according to Definition 9.1, i.e. the reference configuration is given by $\{\hat{\mathbf{r}}, \hat{\mathbf{R}}\}$, where $\hat{\mathbf{R}}(s) = \hat{\mathbf{R}}(\hat\alpha s; \hat\eta)$ and $\hat{\mathbf{r}}' = \mathbf{d}_3$ with $\hat\alpha = 1/\sqrt{\hat\rho^2 + \hat p^2}$, $\hat\eta = \hat\rho\hat\alpha$ and $\hat\eta_* = \sqrt{1-\hat\eta^2} = \hat p\hat\alpha$. Assume furthermore the rod is linearly elastic inextensible and unshearable, and the constitutive equation is given by the diagonal strain-energy function (2.11).*

Then the most general helical deformation is given by the two-parameter family $\{\mathbf{r}, \mathbf{R}\}$ with $\mathbf{R}(s) = \breve{\mathbf{R}}(\alpha s; \eta)$, which is parametrized by $\alpha = 1/\sqrt{\rho^2 + p^2}$ and $\eta = \rho\alpha$. This configuration is obtained by applying a force $\mathbf{n}(s)$ and a moment $\mathbf{m}(s)$ of the following form

$$\underline{\mathbf{m}}(s) = \underline{\mathbf{m}}^*(s) + \begin{pmatrix} 0 \\ 0 \\ M \end{pmatrix}, \quad \underline{\mathbf{n}}(s) = \begin{pmatrix} 0 \\ 0 \\ N \end{pmatrix}, \tag{10.3}$$

where

$$\underline{\mathbf{m}}^*(s) = N \begin{pmatrix} \frac{\eta}{\alpha} \cos \alpha s \\ \frac{\eta}{\alpha} \sin \alpha s \\ 0 \end{pmatrix},$$

and the axial components of moment and force $M = \mathbf{m} \cdot \mathbf{e}_3$ and $N = \mathbf{n} \cdot \mathbf{e}_3$ are

$$M = M(\eta, \alpha) = K_1 \eta \left(\alpha\eta - \hat\alpha\hat\eta \right) + K_3 \eta_* \left(\alpha\eta_* - \hat\alpha\hat\eta_* \right), \tag{10.4}$$

and

$$N = N(\eta, \alpha) = \hat\alpha^2 \hat\eta_* \left[K_1 \frac{\alpha}{\hat\alpha} \frac{\eta_*}{\hat\eta_*} \left(\frac{\hat\eta}{\eta} - \frac{\alpha}{\hat\alpha} \right) + K_3 \frac{\alpha}{\hat\alpha} \left(\frac{\alpha}{\hat\alpha} \frac{\eta_*}{\hat\eta_*} - 1 \right) \right], \tag{10.5}$$

where $\eta_* = \sqrt{1-\eta^2} = p\alpha$.

Furthermore there exists a one-parameter family of helical deformations for which the component of the moment vector $M = \mathbf{m} \cdot \mathbf{e}_3$ vanishes. In this configuration the parameters η and α satisfy

$$\frac{\alpha}{\hat\alpha} = \frac{K_1 \eta\hat\eta + K_3 \eta_* \hat\eta_*}{K_1 \eta^2 + K_3 {\eta_*}^2}, \tag{10.6}$$

and the force is given by

$$N = N(\eta) = \hat\alpha^2 \hat\eta_* \left[K_1 \frac{\alpha}{\hat\alpha} \frac{\eta_*}{\hat\eta_*} \left(\frac{\hat\eta}{\eta} - \frac{\alpha}{\hat\alpha} \right) + K_3 \frac{\alpha}{\hat\alpha} \left(\frac{\alpha}{\hat\alpha} \frac{\eta_*}{\hat\eta_*} - 1 \right) \right]. \tag{10.7}$$

Applying the force $\mathbf{n}$ and the moment $\mathbf{m}$ at the end of the rod is equivalent to applying a force $(0, 0, N)$ and a moment $(0, 0, M)$ at the centerline of the helix. Thus (10.6) and (10.7) correspond to the case where a pure force is applied to the centerline of the helix. Note that the ratio $h = \eta_*/\hat{\eta}_*$ is the ratio of the relative height of the helix in the stressed configuration to the height of the helix in the unstressed configuration.

Proof. It is easily verified that the moment vector $\mathbf{m}$ and the force vector $\mathbf{n}$ satisfy the balance equations (2.4) for any values of M and N. The appropriate form for M and N must be computed to satisfy the constitutive relations.

Assume the configuration $\{\mathbf{r}, \mathbf{R}\}$ of the rod has the following most general helical form $\mathbf{R}(s) = \breve{\mathbf{R}}(\alpha s; \eta)\breve{\mathbf{R}}(f(s), 0)$, so that the centerline defined by $\mathbf{r}' = \mathbf{d}_3$ is helical, and at s the frame is rotated around the director $\mathbf{d}_3$ by an angle $f(s)$. We will show that in order to satisfy the linear constitutive relations the angle $f(s)$ must vanish identically. The strains in this (stressed) configuration are

$$u_1 = -\eta\alpha \cos f(s), \quad u_2 = \eta\alpha \sin f(s), \quad u_3 = \eta_*\alpha + f'(s),$$

and similarly the strains in the unstressed configuration take the following simpler form

$$\hat{u}_1 = -\hat{\eta}\hat{\alpha}, \quad \hat{u}_2 = 0, \quad \hat{u}_3 = \hat{\eta}_*\hat{\alpha}.$$

These strains must satisfy the linear constitutive equations

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

where the moments m_i are computed as

$$m_1 = N\frac{\eta\eta_*}{\alpha} - M\eta, \quad m_2 = 0, \quad m_3 = N\frac{\eta^2}{\alpha} + M\eta_*.$$

Then M, N and $f(s)$ must satisfy the differential-algebraic equations

$$\begin{aligned} N\frac{\eta\eta_*}{\alpha} - M\eta &= K_1\left(-\eta\alpha\cos f(s) + \hat{\eta}\hat{\alpha}\right), && (10.8)\\ 0 &= K_2\eta\alpha\sin f(s), && (10.9)\\ N\frac{\eta^2}{\alpha} + M\eta_* &= K_3\left(\eta_*\alpha + f'(s) - \hat{\eta}_*\hat{\alpha}\right). && (10.10)\end{aligned}$$

Equation (10.9) implies that $f(s) \equiv 0$, and thus M and N must be chosen to satisfy (10.8) and (10.10) with $f(s) = 0$. The solution M and N to this system of linear equations is easily seen to be (10.4) and (10.5). $\square$

10.4 Inextensible Limit of Inextensible, Unshearable Rods

In Section 9.3 we developed an effective model for rods which are helical on a local length scale. In particular, for an inextensible and unshearable rod we obtained the effective Hamiltonian

$$\overline{H}(\mathsf{I}, \mathsf{R}, \mathsf{M}, \mathsf{N}, s) = \overline{W}^*(\mathsf{M}, s) + \mathsf{M} \cdot \tilde{\mathsf{u}}(s) + \hat{\eta}_* \mathsf{N} \cdot \mathsf{e}_3,$$

with effective strain-energy function

$$\overline{W}^*(\mathsf{M}, s) = \frac{1}{2\pi} \int_0^{2\pi} W^*\left(\breve{\mathbf{R}}^T(\tau; \hat{\eta})\mathsf{M}, s\right) d\tau.$$

In this system the centerline $\mathbf{r}$ satisfies (in the basis $\mathcal{N}$)

$$\mathsf{R}'(s) = \mathsf{R} \times \mathsf{U} + \hat{\eta}_* \mathsf{e}_3$$

or equivalently

$$\mathbf{r}' = \hat{\eta}_* \mathbf{D}_3,$$

and we conclude that the shearing and extension effects, which might be thought to be inherent in the helical structure, are negligible in the lowest-order approximation through averaging. In this section we investigate this phenomenon through a specific example.

Consider the helical deformation of a linearly elastic inextensible, unshearable rod computed in Lemma 10.1, in particular the one-parameter family of helices given by (10.6) and (10.7): Suppose the unstressed configuration of the rod is helical with radius $\hat{\rho}$ and pitch $2\pi\hat{p}$, and a force and moment of the size (10.3) is applied with $M = 0$ and

$$N = N(\eta) = \hat{\alpha}^2 \hat{\eta}_* \left[K_1 \frac{\alpha}{\hat{\alpha}} \frac{\eta_*}{\hat{\eta}_*} \left(\frac{\hat{\eta}}{\eta} - \frac{\alpha}{\hat{\alpha}} \right) + K_3 \frac{\alpha}{\hat{\alpha}} \left(\frac{\alpha}{\hat{\alpha}} \frac{\eta_*}{\hat{\eta}_*} - 1 \right) \right],$$

where

$$\hat{\alpha} = 1/\sqrt{\hat{\rho}^2 + \hat{p}^2}, \quad \hat{\eta} = \hat{\rho}\hat{\alpha}, \quad \hat{\eta}_* = \sqrt{1 - \hat{\eta}^2}$$

and

$$\alpha = 1/\sqrt{\rho^2 + p^2}, \quad \hat{\eta} = \rho\alpha, \quad \eta_* = \sqrt{1 - \eta^2}.$$

Then the rod maintains a helical configuration, where the parameters α and η satisfy

$$\frac{\alpha}{\hat{\alpha}} = \frac{K_1 \eta \hat{\eta} + K_3 \eta_* \hat{\eta}_*}{K_1 \eta^2 + K_3 {\eta_*}^2}.$$

In the above rod configuration the only free parameter is η, which we specify so that an extension of the unstressed helical shape of order $O(1)$ is maintained. Generally, in a helical configuration with parameters η and α the centerline $\mathbf{r}$ is

$$\mathbf{r}(s) = \begin{pmatrix} \frac{\eta}{\alpha}\sin\alpha s \\ -\frac{\eta}{\alpha}\cos\alpha s \\ \eta_* s \end{pmatrix}.$$

Thus the ratio of the height of the helix in the stressed configuration to the height of the helix in the unstressed configuration is $h = \eta_*/\hat{\eta}_*$, and we assume this ratio to be of order $O(1)$, i.e. h is constant. In addition we assume the size of the parameters $\hat{\rho}$ and $\hat{p}$ to have different order in the small parameter ϵ.

Lemma 10.2. *Assume the helical deformation is given by* (10.4) *and* (10.5) *where the parameters $\hat{\rho}$ and $\hat{p}$ have the form*

$$\hat{\rho} = \epsilon, \hat{p} = c\epsilon^{\kappa}, \quad \kappa > 0 \tag{10.11}$$

with the additional condition that

$$\frac{\eta_*}{\hat{\eta}_*} = h \tag{10.12}$$

for a constant $h > 0$. Then the order of the component of force N is

$$N = \begin{cases} O\left(\frac{1}{\epsilon^{3-\kappa}}\right) & \textit{if } \kappa \geq 1, \\ O\left(\frac{1}{\epsilon^{2\kappa}}\right) & \textit{if } 0 < \kappa < 1. \end{cases}$$

Proof. Choosing $\hat{\rho}$ and $\hat{p}$ of the particular form (10.11) we get

$$\hat{\alpha} = \frac{1}{\epsilon\sqrt{1 + c^2\epsilon^{2(\kappa-1)}}}, \quad \hat{\eta} = \frac{1}{\sqrt{1 + c^2\epsilon^{2(\kappa-1)}}}, \quad \hat{\eta}_* = \frac{\epsilon^{\kappa-1}}{\sqrt{1 + c^2\epsilon^{2(\kappa-1)}}}.$$

Hence

$$\hat{\alpha} = \begin{cases} O(\epsilon^{-1}) & \text{if } \kappa \geq 1, \\ O(\epsilon^{-\kappa}) & \text{if } 0 < \kappa < 1, \end{cases}$$

$$\hat{\eta} = \begin{cases} O(1) & \text{if } \kappa \geq 1, \\ O(\epsilon^{1-\kappa}) & \text{if } 0 < \kappa < 1, \end{cases}$$

which implies that

$$\hat{\eta}_* = \begin{cases} O(\epsilon^{\kappa-1}) & \text{if } \kappa \geq 1, \\ O(1) & \text{if } 0 < \kappa < 1. \end{cases}$$

Condition (10.12) implies

$$\eta_* = h\hat{\eta}_*, \quad \eta = \sqrt{1 - h^2\hat{\eta}_*^2},$$

so that $\eta_* = O(\hat{\eta}_*)$ and, if $h \neq 1$, $\eta = O(1)$. Thus it can be seen that for $\kappa \geq 0$, $\alpha/\hat{\alpha} = O(1)$. Hence the order of N is determined by

$$N = O(\hat{\alpha}^2 \hat{\eta}_*) = \begin{cases} O\left(\frac{1}{\epsilon^{3-\kappa}}\right) & \text{if } \kappa \geq 1, \\ O\left(\frac{1}{\epsilon^{2\kappa}}\right) & \text{if } 0 < \kappa < 1. \end{cases}$$

□

The case $\kappa = 1$ corresponds to the model discussed in Chapter 9, and the above lemma implies, that in order maintain any helical configuration of order $h = O(1)$, different from the reference configuration, a force of order $N = O(\epsilon^{-2})$ must be applied. Thus in the limiting (averaged) problem this may be interpreted as enforcing the inextensibility and unshearability constraint (2.5).

The cases $\kappa > 1$ and $0 < \kappa < 1$ correspond to the different models discussed in Section 10.2. There we have shown that if $0 < \kappa < 1$, the model converges to the high twist problem discussed in Section 9.6. Thus in this case it is expected that the averaged system should model an inextensible rod, and the order of magnitude of the force $N = O(\epsilon^{-2\kappa})$ required to maintain a finite extension is in concordance with this conclusion. It should be remarked that the homogenization theory formulated in Theorem 8.1 does not apply for this example (say for the case $\kappa = 1$), since in the helical deformation the force is of order $O(\epsilon^{-2})$. Therefore in particular the constant C_1 in Theorem 8.1 is of order $O(\epsilon^{-2})$, and the theorem becomes meaningless.

If $\kappa > 1$, we have seen that the averaged system does not constitute a rod model. The closed form helical solutions shows that in this case an effective system (possible determined with different methods) should include extension effects. The force necessary to maintain a finite extension is of order $O(\epsilon^{\kappa-3})$, and is therefore finite in the limit, provided $\kappa \geq 3$. However, the threshold $\kappa = 3$ is dependent on the particular family of solutions under study, and may be different for other deformations.

Chapter 11

Summary and Discussion

11.1 Summary

Several observations concerning the theory of elastic rods and the analysis of special cases have been made in this dissertation.

A new twelve-dimensional Poisson bracket describing the equilibrium conditions of elastic rods has been presented, filling a gap between the known six-dimensional non-canonical Euler-Poisson description of heavy rigid body dynamics and several canonical Hamiltonian descriptions for elastic rods and rigid bodies.

The fact that the centerline $\mathbf{r}$ is an ignorable coordinate in any of the Hamiltonian formulations describing elastic rods generally implies a reduction of the system by three degrees of freedom, and the reduced system governs the frame evolution only. In this viewpoint the shape of the centerline is computed from the frame evolution through three quadratures. In one instance, namely the description of linearly elastic, inextensible, unshearable rods, this reduction is known as the Kirchhoff Kinetic Analogy. In this case the reduced problem governing the frame evolution is equivalent to the equations of motion of a heavy rigid body tumbling about a fixed point. For boundary value problems, which typically arise in rod mechanics, this reduction is generally inappropriate since the variable $\mathbf{r}$ may explicitly appear in the boundary conditions. It was known that two of the three components of the centerline $\mathbf{r}$ can always be expressed in terms of variables of the reduced problem, and this fact has been exploited to prove that the centerline of an infinite, uniform rod is always bounded in a particular two-dimensional projection. Furthermore, it is shown that an additional integral appearing only in the Euler parameter Hamiltonian formulation may be used to algebraically express the last component of the centerline in terms of variables describing the frame evolution. As a consequence every two-point boundary value problem for the full seven degree-of-freedom Euler parameter formulation can be reduced to a boundary value problem for a four degree-of-freedom system.

The analysis of elastic rods, whose intrinsic shape exhibits a fine-scale periodic structure, constitutes the main part of this dissertation. Mathematically, these problems are formulated as Hamiltonian systems depending on two different length scales, and it has been investigated how, under suitable assumptions on the intrinsic shape of the rod, these systems can be recast into a form which is appropriate for standard averaging theory to be applied. The corresponding averaged (or homogenized) system can be regarded as describing an effective rod model. As a particular example, this theory has been applied to systems describing rods with a locally helical or, as a limiting case, highly twisted intrinsic shape. As an important consequence, it has been observed that in certain well-defined circumstances the effective system describes an isotropic rod. Several specific examples have been considered.

In the remainder of this conclusion we will discuss some important issues concerning the formulation of effective rod models that have not been resolved.

11.2 Homogenization of Boundary Value Problems

The homogenization theory in Chapter 8 has been formulated for initial value problems, and it is not obvious how to modify this result to hold for two-point boundary value problems, which typically arise in rod mechanics. In this discussion we will focus on a specific two-point boundary value problem, which arises in modeling DNA mini-circles. We will discuss how this problem can be properly formulated, incorporating a high intrinsic twist. Two methods to obtain effective models will be presented, both of which are open for further research.

In Section 7.4 we have discussed, in the context of integrable systems, the case of an inextensible and unshearable rod, straight in the unstressed configuration, that is bent to form a smooth closed loop. Specifically, one set of boundary conditions is

$$\left.\begin{array}{l} \mathbf{r}(0) = \mathbf{r}(1) = (0,0,0), \\ \mathbf{d}_3(0) = \mathbf{d}_3(1), \\ \mathbf{d}_i(0) = \mathbf{e}_i, \quad i = 1,2,3, \\ \mathbf{d}_1(0)\cdot\mathbf{d}_1(1) = \cos\alpha. \end{array}\right\} \tag{11.1}$$

As the underlying Hamiltonian system we adopt here the canonical Euler parameter description (4.15) with Hamiltonian

$$H(\mathbf{r},\mathbf{q},\mathbf{n},\boldsymbol{\mu},s) = W^*(\mathsf{m},s) + \mathsf{m}\cdot\hat{\mathsf{u}} + \mathsf{n}\cdot\mathsf{e}_3,$$

where $\hat{\mathsf{u}} = \mathbf{0}$ and the variables m and n are expressed in terms of the Euler parameters $\mathbf{q}$ and their conjugate momenta $\boldsymbol{\mu}$ (cf. (4.13) and (4.14) in Section

4.3). One appropriate set of boundary conditions in terms of the variables $\mathbf{q}$ and $\boldsymbol{\mu}$ have for example been described in Li and Maddocks [39]:

$$\left.\begin{array}{l} \mathbf{r}(0) = \mathbf{r}(1) = (0,0,0), \\ \big(q_1(0), q_2(0), q_3(0)\big) = (0,0,0), \\ \mathbf{q}(1) = \big(0,0,-\sin(\alpha/2), -\cos(\alpha/2)\big), \\ \mu_4(0) = 0. \end{array}\right\} \tag{11.2}$$

We now assume in addition that the rod is highly twisted in its unstressed configuration, so that the bending and twisting strains take the form $\hat{\mathsf{u}} = (0,0,1/\epsilon)$. As in Chapter 9 we are trying to determine the asymptotic behavior of the system as the twist rate becomes very large, or equivalently as ϵ tends to zero. We first observe that problem (11.2) is not well-defined in the limit, as the director frame rotates faster and faster around the axis $\mathbf{d}_3$, and therefore the specification of the angle α does not make sense. Hence we propose that the boundary value problem be formulated for system (8.12), i.e. in variables which are components relative to the slowly varying, natural frame $\mathcal{N} = \{\mathbf{D}_i\}_{i=1}^3$, according to Lemma 8.2. In these variables the Hamiltonian takes the form

$$\widetilde{H}(\mathbf{r}, \mathbf{Q}, \mathbf{n}, \boldsymbol{\Upsilon}, s, \tau) = W^*(\breve{\mathbf{R}}^T(\tau)\mathsf{M}, \breve{\mathbf{R}}^T(\tau)\mathsf{N}, s) + \mathsf{N}\cdot\breve{\mathbf{R}}(\tau)\mathsf{e}_3,$$

where $\tau = s/\epsilon$, the rotation matrix $\breve{\mathbf{R}}$ is defined in (9.1), and the variables M and N are components of the moment and force vectors $\mathbf{m}$ and $\mathbf{n}$ relative to the frame $\mathcal{N}$. Furthermore the frame $\mathcal{N}$ is parametrized in terms of Euler parameters $\mathbf{Q} = (Q_1, Q_2, Q_3, Q_4)$, and then the moments M and force N can be expressed in terms of $\mathbf{Q}$ and their conjugate momenta $\boldsymbol{\Upsilon} = (\Upsilon_1, \Upsilon_2, \Upsilon_3, \Upsilon_4)$. Instead of adopting (11.1), we now formulate the boundary conditions relative to the frame $\mathcal{N}$ in the following way:

$$\left.\begin{array}{l} \mathbf{r}(0) = \mathbf{r}(1) = (0,0,0), \\ \mathbf{D}_3(0) = \mathbf{D}_3(1), \\ \mathbf{D}_i(0) = \mathbf{e}_i, \quad i = 1,2,3, \\ \mathbf{D}_1(0)\cdot\mathbf{D}_1(1) = \cos\alpha. \end{array}\right\} \tag{11.3}$$

These boundary conditions are written in terms of the Euler parameters $\mathbf{Q}$ and their conjugate momenta $\boldsymbol{\Upsilon}$ as

$$\left.\begin{array}{l} \mathbf{r}(0) = \mathbf{r}(1) = (0,0,0), \\ (Q_1(0), Q_2(0), Q_3(0)) = (0,0,0), \\ \mathbf{Q}(1) = \big(0,0,-\sin(\alpha/2), -\cos(\alpha/2)\big), \\ \Upsilon_4(0) = 0. \end{array}\right\} \tag{11.4}$$

The techniques that have been employed in the general averaging theory (cf. Theorem 4.1.1 in [23]), and in this dissertation in Theorem 8.1, cannot, to my knowledge, be applied to prove similar averaging results for two-point boundary value problems. There exists however a vast literature concerning the homogenization of boundary value problems involving partial differential equations (see for example Bensoussan et al. [11]). Often these problems are formulated in a variational form, i.e. one searches for minimizers u_ϵ of a functional $\mathcal{F}(u, \nabla u, s, s/\epsilon, \epsilon)$ for u_ϵ in an appropriate function space, which incorporates the boundary conditions. Then one establishes the (typically weak) convergence of the sequence of minimizers. The homogenized problem is in general not defined as the average over the functional $\mathcal{F}$, but rather as the average over an appropriate dual functional $\mathcal{F}^*$.

The equations governing the equilibria of elastic rods can be cast into a variational form as we have already indicated in Sections 4.3 and 4.4. However, depending on the type of the problem and the choice for the parametrization of the rotation group the problem has to be formulated by using a constrained variational principle, which may complicate the analysis. A variational formulation involves minimizing a Lagrangian over an appropriate space, and in concordance with the theory for partial differential equations and our results for initial value problems, it is most likely that the homogenized Lagrangian is defined via the average over the Hamiltonian, which can be regarded as a dual functional of the Lagrangian via the Legendre transform.

Some of these methods can be found, for example, in Jikov et al. [32] (and references therein). However, problems involving partial differential equations are in some sense easier in that they typically involve one equation in several space variables, whereas our application actually involves more than one equation. On the other hand, for partial differential equations the form of the homogenized system depends on the type of the boundary conditions that are enforced. For ordinary differential equations one should probably not expect such a distinction, as in the homogenization of initial value problems the averaged system is independent of the initial conditions. Moreover, it seems likely, employing some type of shooting method, that the approximation through averaging is actually improved when dealing with boundary value problems.

For a special case of the closed-loop boundary value problem described above, a different approach to prove a homogenization result may be possible. Namely, we now require that the frame $\mathcal{N}$ at 0 closes up with itself at $s = 1$, i.e.

$$\mathbf{D}_i(0) = \mathbf{D}_i(1), \quad i = 1, 2, 3, \tag{11.5}$$

or equivalently $\alpha = 2m\pi$ for an integer m. Rather than using the Euler parameter Hamiltonian formulation we adopt the non-canonical formulation (8.12) developed in Lemma 8.2. Recall that the Hamiltonian has the form

$$H(\mathsf{I}, \mathsf{R}, \mathsf{M}, \mathsf{N}, s, \tau) = W^*(\breve{\mathbf{R}}^T(\tau)\mathsf{M}, \breve{\mathbf{R}}^T(\tau)\mathsf{N}, s) + \mathsf{N} \cdot \breve{\mathbf{R}}(\tau)\mathsf{e}_3,$$

where now the variables are the components of $\mathbf{i}$, $\mathbf{r}$, $\mathbf{m}$ and $\mathbf{n}$ relative to the frame $\mathcal{N}$. In this case the boundary conditions (11.3) in terms of the variables I, R, M, N take the particularly nice form

$$\mathsf{R}(0) = \mathsf{R}(1) = (0,0,0), \tag{11.6}$$
$$\mathsf{I}(0) = \mathsf{I}(1) = (1,0,0), \tag{11.7}$$
$$\mathsf{N}(0) = \mathsf{N}(1) = (0,0,1). \tag{11.8}$$

To see this, recall from the discussion in Section 4.2 that the configuration of the frame $\mathcal{N} = \{\mathbf{D}_i\}_{i=1}^3$ is equivalent to a specification of the components N and I of the constant vectors $\mathbf{n}$ and $\mathbf{i}$. Thus condition (11.5) is equivalent to (11.7) and (11.8). This implies in particular, that we may characterize the solution of this boundary value problem as a periodic orbit. The functions $\mathsf{R}(s)$, $\mathsf{I}(s)$ and $\mathsf{N}(s)$ are 1-periodic by assumption. To see that the function $\mathsf{M}(s)$ is periodic, note that $\mathbf{m} + \mathbf{r} \times \mathbf{n}$ is an integral (i.e. integral (5.1) in Chapter 5). This implies immediately that $\mathbf{m}(0) = \mathbf{m}(1)$ and thus with condition (11.5) the periodicity of $\mathsf{M}(s)$.

Thus rather than treating the above problem as a two-point boundary value problem, we study the full dynamical system and investigate the relationship between periodic orbits of the general system of differential equations and periodic orbits of the corresponding averaged system. Abstractly formulated, we are searching for 1-periodic solutions of

$$\mathbf{x}' = \mathbf{f}\left(\mathbf{x}, \frac{s}{\epsilon}\right) \tag{11.9}$$

close to 1-periodic solutions of the averaged system

$$\mathbf{x}' = \overline{\mathbf{f}}(\mathbf{x}) = \frac{1}{2\pi}\int_0^{2\pi} \mathbf{f}(\mathbf{x},\tau)d\tau, \tag{11.10}$$

where $\mathbf{f}(\mathbf{x},\tau)$ is 2π periodic in τ.

Averaging theory is often regarded as a method to locate periodic orbits of differential equations with a periodic time dependence (see e.g. Farkas [20, Ch. 6]). However, we should remark that the theory is generally formulated for problems depending on the time scale $t = s/\epsilon$, so that systems (11.9) and (11.10) become

$$\frac{d}{dt}\mathbf{x} = \epsilon\mathbf{f}(\mathbf{x},t), \tag{11.11}$$

and

$$\frac{d}{dt}\mathbf{x} = \epsilon\overline{\mathbf{f}}(\mathbf{x}). \tag{11.12}$$

The 1-periodic solution $\mathbf{x}(t)$ of system (11.9) is equivalent to a solution $\hat{\mathbf{x}}(t) = \hat{\mathbf{x}}(s/\epsilon) = \mathbf{x}(s)$ of (11.11) satisfying $\hat{\mathbf{x}}(0) = \hat{\mathbf{x}}(1/\epsilon)$. In the limit as ϵ tends to zero, these solutions may correspond to homoclinic orbits of system (11.12). Thus one investigates if there exist $1/\epsilon$-periodic solutions of system (11.11) close to homoclinic orbits of (11.12). For abstract results about the approximation of stable and unstable manifolds and fixed points see for example Guckenheimer and Holmes [23, Theorem 4.1.1].

Appendix A

Notation

We denote (Euclidean) real, three-dimensional inner product space by $\mathbb{E}^3$, and its elements $\mathbf{p} \in \mathbb{E}^3$ (denoted by boldface letters) are vectors. The inner product $\langle \mathbf{p}, \mathbf{q} \rangle$ between two vectors is denoted by $\mathbf{p} \cdot \mathbf{q}$ and the associated norm is

$$|\mathbf{p}| = \sqrt{\mathbf{p} \cdot \mathbf{p}}.$$

Consider three different bases for $\mathbb{E}^3$, a fixed (inertial) basis $\mathcal{E} = \{\mathbf{e}_i\}$, and two moving bases $\mathcal{D} = \{\mathbf{d}_i(s)\}$ and $\mathcal{N} = \{\mathbf{D}_i(s)\}$. All bases are assumed to be orthonormal and right-handed. Any vector $\mathbf{p} \in \mathbb{E}^3$ can be decomposed with respect to the above bases, e.g. $\mathbf{p} = (\mathbf{p} \cdot \mathbf{e}_i)\mathbf{e}_i$. Let $\underline{p}_i = \mathbf{p} \cdot \mathbf{e}_i$, $p_i = \mathbf{p} \cdot \mathbf{d}_i$ and $P_i = \mathbf{p} \cdot \mathbf{D}_i$, and denote triples of components with respect to the different bases by

$$\underline{\mathsf{p}} = \begin{pmatrix} \underline{p}_1 \\ \underline{p}_2 \\ \underline{p}_3 \end{pmatrix}, \quad \mathsf{p} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}, \quad \mathsf{P} = \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix}.$$

In order to simplify notation we adopt the summation convention that twice repeated Latin indices are summed from 1 to 3, e.g.

$$u_i \mathbf{d}_i = \sum_{i=1}^{3} u_i \mathbf{d}_i.$$

By a slight abuse of notation we denote the product of two triples of components by

$$\mathsf{p} \cdot \mathsf{q} = \mathsf{p}^T \mathsf{q} = p_i q_i.$$

For every skew linear transformation $\mathbf{A} : \mathbb{E}^3 \to \mathbb{E}^3$ there exists $\mathbf{a} \in \mathbb{E}^3$ such that $\mathbf{A}\mathbf{v} = \mathbf{a} \times \mathbf{v}$ for all $\mathbf{v} \in \mathbb{E}^3$. The symbolic notation $\mathbf{a}\times$ will be used to denote the skew linear transformation $\mathbf{A}$. Writing $\mathbf{y} = \mathbf{a} \times \mathbf{v}$ in components with respect to the basis $\{\mathbf{d}_i\}$ for example we get

$$\mathbf{y} \cdot \mathbf{d}_k = (\mathbf{A}\mathbf{v}) \cdot \mathbf{d}_k = (a_i \mathbf{d}_i \times v_j \mathbf{d}_j) \cdot \mathbf{d}_k = a_i v_j (\mathbf{d}_i \times \mathbf{d}_j) \cdot \mathbf{d}_k = \epsilon_{ijk} a_i v_j,$$

where ϵ_{ijk} is the alternating symbol, defined as

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } (i,j,k) = (1,2,3), (2,3,1), (3,1,2), \\ -1 & \text{if } (i,j,k) = (3,2,1), (1,3,2), (2,3,1), \\ 0 & \text{otherwise.} \end{cases}$$

Thus

$$\mathsf{y} = \mathsf{a}\times\mathsf{v},$$

where the matrix $\mathsf{a}\times$ is defined as

$$\mathsf{a}\times = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}$$

(the vector a is the axial vector of the matrix $\mathsf{a}\times$). Corresponding to the rotational invariance of the vector product, we have the property

$$\mathbf{R}\mathsf{a}\times\mathbf{R}^T = (\mathbf{R}\mathsf{a})\times. \tag{A.1}$$

Appendix B

Legendre Transforms

In this section we review the concept of a Legendre transform, and summarize certain simple properties. For a strictly convex function $\Phi : \mathbb{R}^n \to \mathbb{R}$ the Legendre transform is defined to be

$$\Phi^*(\mathbf{y}) = \sup_{\mathbf{x}}\{\mathbf{y}\cdot\mathbf{x} - \Phi(\mathbf{x})\}.$$

The Legendre transform is involutive, that is $\Psi = \Phi^*$ if and only if $\Psi^* = \Phi$.

The Legendre transform commutes with orthogonal changes of variables, that is if $\mathbf{R}$ is an orthogonal transformation and $\widetilde{\Phi}(\mathbf{x}) = \Phi(\mathbf{R}\mathbf{x})$, then $\widetilde{\Phi}^*(\mathbf{y}) = \Phi^*(\mathbf{R}\mathbf{y})$. To see this let $\mathbf{x}' = \mathbf{R}^T\mathbf{x}$ and compute

$$\begin{aligned}
\Phi^*(\mathbf{R}\mathbf{y}) &= \sup_{\mathbf{x}}\{\mathbf{R}\mathbf{y}\cdot\mathbf{x} - \Phi(\mathbf{x})\} \\
&= \sup_{\mathbf{x}'}\{\mathbf{y}\cdot\mathbf{x}' - \Phi(\mathbf{R}\mathbf{x}')\} \\
&= \sup_{\mathbf{x}'}\{\mathbf{y}\cdot\mathbf{x}' - \widetilde{\Phi}(\mathbf{x}')\} \\
&= \widetilde{\Phi}^*(\mathbf{y}).
\end{aligned}$$

This implies in particular that if Φ is invariant under an orthogonal transformation $\mathbf{R}$, i.e. $\Phi(\mathbf{R}\mathbf{x}) = \Phi(\mathbf{x})$ then its Legendre transform Φ^* is also invariant. For, if $\tilde{\Phi}(\mathbf{x}) = \Phi(\mathbf{x})$ for all $\mathbf{x}$, then

$$\Phi^*(\mathbf{y}) = \tilde{\Phi}^*(\mathbf{y}) = \Phi^*(\mathbf{R}\mathbf{y}).$$

Finally, the Legendre transform $\widetilde{\Phi}^*(\mathbf{y})$ of a shifted function $\tilde{\Phi}(\mathbf{x}) = \Phi(\mathbf{x} - \hat{\mathbf{x}})$ is $\tilde{\Phi}^*(\mathbf{y}) = \Phi^*(\mathbf{y}) + \hat{\mathbf{x}}\cdot\mathbf{y}$. For

$$\begin{aligned}
\tilde{\Phi}^*(\mathbf{y}) &= \sup_{\mathbf{x}}\{\mathbf{x}\cdot\mathbf{y} - \Phi(\mathbf{x} - \hat{\mathbf{x}})\} \\
&= \sup_{\mathbf{x}'}\{(\mathbf{x}' + \hat{\mathbf{x}})\cdot\mathbf{y} - \Phi(\mathbf{x}')\} \\
&= \sup_{\mathbf{x}'}\{\mathbf{x}'\cdot\mathbf{y} - \Phi(\mathbf{x}')\} + \hat{\mathbf{x}}\cdot\mathbf{y} \\
&= \Phi^*(\mathbf{y}) + \hat{\mathbf{x}}\cdot\mathbf{y}.
\end{aligned}$$

Appendix C

Elliptic Integrals and Elliptic Functions

The following definitions for elliptic integrals and functions can be found for example in Byrd [13].

Definition C.1. *For* $\varphi \in (0, \pi/2]$ *the* normal elliptic integrals *are defined by*

- *(elliptic integral of the first kind)*
$$F(\varphi, k) = \int_0^{\varphi} \frac{d\theta}{\sqrt{1 - k^2 \sin^2\theta}},$$
- *(elliptic integral of the second kind)*
$$E(\varphi, k) = \int_0^{\varphi} \sqrt{1 - k^2 \sin^2\theta}\, d\theta,$$
- *(elliptic integral of the third kind)*
$$\Pi(\varphi, \alpha^2, k) = \int_0^{\varphi} \frac{d\theta}{(1 - \alpha^2 \sin^2\theta)\sqrt{1 - k^2 \sin^2\theta}}, \quad \alpha^2 \in \mathbb{R}.$$

The complete *elliptic integrals are given by*

$$K = K(k) = F(\pi/2, k), \quad E = E(k) = E(\pi/2, k), \quad \Pi(\alpha^2, k) = \Pi(\pi/2, \alpha^2, k).$$

The number k is called the modulus *of the elliptic integrals, and for our purposes we assume $k \in (0, 1)$.*

Jacobi's elliptic functions are defined as the inverse functions of the elliptic integrals.

Definition C.2. *If $u = F(\varphi, k)$ then the* sine amplitude $\mathrm{sn}(u, k)$ *and the* amplitude $\mathrm{am}(u, k)$ *are defined through*

$$\sin\varphi = \mathrm{sn}(u, k),$$

and

$$\varphi = \mathrm{am}(u, k).$$

Appendix D

Proofs of Theorem 7.1 and Corollary 7.1

We will here carry out some of the calculations in the proofs of Theorem 7.1 and Corollary 7.1. Recall the notation used in these theorems: For $N > 0$ define

$$\begin{aligned} f(u) &= K_1(2h - K_3^{-1}p_\psi^2 - 2Nu)(1-u^2) - (p_\phi - p_\psi u)^2 \\ &= (a - bu)(1-u^2) - (p_\phi - p_\psi u)^2 \\ &= b(u-u_1)(u-u_2)(u-u_3), \end{aligned}$$

where $-1 < u_1 < u_2 < 1 < u_3$, $a = K_1(2h - K_3^{-1}p_\psi^2)$ and $b = 2K_1N$. Furthermore define v through $u = u_1 + (u_2 - u_1)\sin^2 v$ and let $\epsilon_2^2 = u_2 - u_1$, $\epsilon_3^2 = u_3 - u_1$, $k^2 = \epsilon_2^2/\epsilon_3^2$, $\alpha_\pm = \pm\epsilon_2^2/(1 \mp u_1)$ and

$$Q_\pm(v) = \frac{J_1 - J_3}{1 - u_1}\Pi(v, \alpha_+, k) \pm \frac{J_1 + J_3}{1 + u_1}\Pi(v, \alpha_-, k),$$

where $J_1 = p_\phi$ and $J_3 = p_\psi$.

The following proposition establishes the values of some specific integrals, which will be needed in the sequel.

Proposition D.1.

$$\int_{u_1}^{u} \frac{dw}{\sqrt{f(w)}} = \frac{2}{\sqrt{b}\epsilon_3}F(v,k), \tag{D.1}$$

$$\int_{u_1}^{u_2} \frac{dw}{\sqrt{f(w)}} = \frac{2K}{\sqrt{b}\epsilon_3}, \tag{D.2}$$

$$\int_{u_1}^{u} \frac{dw}{\sqrt{f(w)}(1-w)}dw = \frac{2}{\sqrt{b}\epsilon_3(1-u_1)}\Pi(v,\alpha_+,k), \tag{D.3}$$

$$\int_{u_1}^{u} \frac{dw}{\sqrt{f(w)}(1+w)}dw = \frac{2}{\sqrt{b}\epsilon_3(1+u_1)}\Pi(v,\alpha_-,k). \tag{D.4}$$

$$\int_{0}^{u} \mathrm{sn}^2 w\, dw = \frac{1}{k^2}\big[u - E\big(\mathrm{am}(u,k),k\big)\big]. \tag{D.5}$$

Proof. In order to evaluate integrals (D.1)–(D.4) we use the change of variables

$$w = u_1 + \epsilon_2^2 \sin^2 z,$$

so that

$$dw = 2\epsilon_2^2 \sin z \cos z dz$$

and

$$f(w) = b\epsilon_2^4 \epsilon_3^2 \sin^2 z \cos^2 z (1 - k^2 \sin^2 z).$$

If $w = u_1$ ($w = u$) then $z = 0$ ($z = u_1 + \epsilon_2^2 \sin u = v$), so that we immediately evaluate integral (D.1) to be

$$\int_{u_1}^{u} \frac{1}{\sqrt{f(w)}} dw = \frac{2}{\sqrt{b}\epsilon_3} \int_0^v \frac{dz}{\sqrt{1 - k^2 \sin^2 z}} = \frac{2}{\sqrt{b}\epsilon_3} F(v, k).$$

In particular if $w = u_2$ then $z = \pi/2$ and we obtain the expression for integral (D.2). Using the same change of variables we compute integral (D.3)

$$\begin{aligned}
&\int_{u_1}^{u} \frac{1}{\sqrt{f(w)}(1-w)} dw = \\
&= \frac{2}{\sqrt{b}\epsilon_3 (1 - u_1)} \int_0^v \frac{dz}{\sqrt{1 - k^2 \sin^2 z}(1 - \alpha_+ \sin^2 z)} \\
&= \frac{2}{\sqrt{b}\epsilon_3 (1 - u_1)} \Pi(v, \alpha_+, k),
\end{aligned}$$

and in the same manner (D.4). Finally the value of integral (D.5) can be found in Byrd [13]. □

We now turn to the proofs of Theorem 7.1 and Corollary 7.1. We have computed the generating function for the action-angle variable transformation $(\boldsymbol{\phi}, \mathbf{p}) \to (\boldsymbol{\varphi}, \mathbf{J})$ to be

$$S(\boldsymbol{\phi}, J_1, J_2, J_3) = J_1 \phi + S_2(\theta, J_1, J_2, J_3) + J_3 \psi,$$

where with $u = \cos\theta$

$$S_2(\theta, J_1, J_2, J_3) = \int_{u_1}^{u} \frac{\sqrt{f(v)}}{1 - v^2} dv.$$

The action-variable J_2 is computed as

$$J_2 = \widetilde{J}_2(J_1, J_3, h) = \frac{1}{\pi} \int_{u_1}^{u_2} \frac{\sqrt{f(u)}}{1 - u^2} du,$$

which implicitly defines the Hamiltonian $\widetilde{H} = h$, i.e.

$$\widetilde{J}_2\big(J_1, J_3, \widetilde{H}(J_1, J_2, J_3)\big) = J_2.$$

Furthermore note that the function S_2 is of the following form

$$S_2(\theta, J_1, J_2, J_3) = \widetilde{S}_2(\theta, J_1, J_3, \widetilde{H}(J_1, J_2, J_3))$$

with $\widetilde{S}_2 = \widetilde{S}_2(\theta, J_1, J_3, h)$ and h replaced by $h = \widetilde{H}(J_1, J_2, J_3)$. This is particularly important for the computation of the angle variables $\boldsymbol{\varphi}$. In particular

$$\begin{aligned}
\varphi_1 &= \frac{\partial S}{\partial J_1} = \phi + \frac{\partial S_2}{\partial J_1} \\
&= \phi + \frac{\partial \widetilde{S}_2}{\partial J_1} + \frac{\partial \widetilde{S}_2}{\partial h}\frac{\partial \widetilde{H}}{\partial J_1} \\
&= \phi + \frac{\partial \widetilde{S}_2}{\partial J_1} + \frac{\partial \widetilde{S}_2}{\partial h}\omega_1, \\
\varphi_2 &= \frac{\partial S_2}{\partial J_2} = \frac{\partial \widetilde{S}_2}{\partial h}\frac{\partial \widetilde{H}}{\partial J_2} = \frac{\partial \widetilde{S}_2}{\partial h}\omega_2, \\
\varphi_3 &= \frac{\partial S}{\partial J_3} = \psi + \frac{\partial S_2}{\partial J_3} \\
&= \psi + \frac{\partial \widetilde{S}_2}{\partial J_3} + \frac{\partial \widetilde{S}_2}{\partial h}\frac{\partial \widetilde{H}}{\partial J_3} \\
&= \psi + \frac{\partial \widetilde{S}_2}{\partial J_3} + \frac{\partial \widetilde{S}_2}{\partial h}\omega_3.
\end{aligned}$$

We compute the partial derivatives using (D.3) and (D.4)

$$\begin{aligned}
\frac{\partial \widetilde{S}_2}{\partial J_1} &= -\int_{u_1}^{u} \frac{J_1 - J_3 w}{\sqrt{f(w)}(1-w^2)} dw \\
&= -\frac{1}{2}\left(\int_{u_1}^{u} \frac{J_1 - J_3 w}{\sqrt{f(w)}(1-w)} dw + \int_{u_1}^{u} \frac{J_1 - J_3 w}{\sqrt{f(w)}(1+w)} dw\right) \\
&= -\frac{1}{2}\Bigg((J_1 - J_3)\int_{u_1}^{u} \frac{1}{\sqrt{f(w)}(1-w)} dw \\
&\quad + (J_1 + J_3)\int_{u_1}^{u} \frac{1}{\sqrt{f(w)}(1+w)} dw\Bigg) \\
&= -\frac{1}{\sqrt{b\epsilon_3}}\left(\frac{J_1 - J_3}{1-u_1}\Pi(v, \alpha_+, k) + \frac{J_1 + J_3}{1+u_1}\Pi(v, \alpha_-, k)\right) \\
&= -\frac{1}{\sqrt{b\epsilon_3}} Q_+(v)
\end{aligned}$$

and similarly

$$\frac{\partial \widetilde{J}_2}{\partial J_1} = -\frac{1}{\pi\sqrt{b}\epsilon_3} Q_+(\frac{\pi}{2}).$$

Furthermore using (D.1) and (D.2) we get

$$\frac{\partial \widetilde{S}_2}{\partial h} = K_1 \int_{u_1}^{u} \frac{1}{\sqrt{f(w)}} dw = \frac{2K_1}{\sqrt{b}\epsilon_3} F(v,k)$$

and

$$\frac{\partial \widetilde{J}_2}{\partial h} = \frac{K_1}{\pi} \int_{u_1}^{u_2} \frac{1}{\sqrt{f(w)}} dw = \frac{2K_1}{\pi\sqrt{b}\epsilon_3} K.$$

Finally using (D.3) and (D.4) we obtain

$$\begin{aligned}
\frac{\partial \widetilde{S}_2}{\partial J_3} &= -K_1 K_3^{-1} J_3 \int_{u_1}^{u} \frac{1}{\sqrt{f(w)}} dw + \int_{u_1}^{u} \frac{J_1 w - J_3 w^2}{\sqrt{f(w)}(1-w^2)} dw \\
&= J_3(1 - K_1 K_3^{-1}) \int_{u_1}^{u} \frac{1}{\sqrt{f(w)}} dw + \int_{u_1}^{u} \frac{J_1 w - J_3}{\sqrt{f(w)}(1-w^2)} dw \\
&= J_3(1 - K_1 K_3^{-1}) \frac{2}{\sqrt{b}\epsilon_3} F(v,k) \\
&\quad + \frac{1}{2} \left((J_1 - J_3) \int_{u_1}^{u} \frac{1}{\sqrt{f(w)}(1-w)} dw \right. \\
&\quad \left. -(J_1 + J_3) \int_{u_1}^{u} \frac{1}{\sqrt{f(w)}(1+w)} dw \right) \\
&= \frac{1}{\sqrt{b}\epsilon_3} \left(\frac{J_1 - J_3}{1 - u_1} \Pi(v, \alpha_+, k) - \frac{J_1 + J_3}{1 + u_1} \Pi(v, \alpha_-, k) \right) \\
&= J_3(1 - K_1 K_3^{-1}) \frac{2}{\sqrt{b}\epsilon_3} F(v,k) + \frac{1}{\sqrt{b}\epsilon_3} Q_-(v)
\end{aligned}$$

and

$$\frac{\partial \widetilde{J}_2}{\partial J_3} = J_3(1 - K_1 K_3^{-1}) \frac{2}{\pi\sqrt{b}\epsilon_3} K + \frac{1}{\pi\sqrt{b}\epsilon_3} Q_-(\frac{\pi}{2}).$$

Using these results we immediately compute the frequencies $\omega_i = \partial\widetilde{H}/\partial J_i$

$$
\begin{aligned}
\omega_1 = \frac{\partial\widetilde{H}}{\partial J_1} &= -\left(\frac{\partial\widetilde{J}_2}{\partial h}\right)^{-1}\frac{\partial\widetilde{J}_2}{\partial J_1} = \frac{1}{K_1K}Q_+(\frac{\pi}{2}),\\
\omega_2 = \frac{\partial\widetilde{H}}{\partial J_2} &= \left(\frac{\partial\widetilde{J}_2}{\partial h}\right)^{-1} = \frac{\pi\sqrt{b}\epsilon_3}{2K_1K},\\
\omega_3 = \frac{\partial\widetilde{H}}{\partial J_3} &= -\left(\frac{\partial\widetilde{J}_2}{\partial h}\right)^{-1}\frac{\partial\widetilde{J}_2}{\partial J_3} = \frac{K_1-K_3}{K_1K_3}J_3 - \frac{Q_-(\frac{\pi}{2})}{2K_1K},
\end{aligned}
$$

which completes the proof of Corollary 7.1. Using these and previously computed quantities we obtain the formulas for the angle variables

$$
\begin{aligned}
\varphi_1 &= \phi - \frac{1}{K\sqrt{b}\epsilon_3}\left(Q_+(v)K - Q_+(\frac{\pi}{2})F(v,k)\right),\\
\varphi_2 &= \frac{\pi}{K}F(v,k),\\
\varphi_3 &= \psi + \frac{1}{\sqrt{b}\epsilon_3K}\left(Q_-(v)K - Q_-(\frac{\pi}{2})F(v,k)\right)
\end{aligned}
$$

and the proof of Theorem 7.1 is complete.

Appendix E

Hamiltonian Perturbation Theory

The homogenization result in Chapter 8 neither relies on nor utilizes the fact that the system of differential equations has a Hamiltonian structure. However, this fact should not be ignored, since in particular for the computation of perturbation expansions it is relevant that higher-order terms are computed in such a way that the Hamiltonian structure persists in the approximation.

The averaging result in Chapter 8 relies on the fact that the change of variables (8.17) exists, under which the averaged system is generated, and estimates for the quality of the approximation can be obtained. In the Hamiltonian context this change of variables can be interpreted as transforming the system into a standard form, which provides the starting point for Hamiltonian perturbation theory. Assuming in addition that the averaged system is integrable, the action-angle variable description in Chapter 7 constitutes, at least theoretically, the tool for the computation of higher-order expansions and corresponding error estimates. In Section E.1 we will show in detail, how a standard form can be obtained. Then in Section E.2 we will discuss several techniques to obtain higher order expansions for perturbations of integrable Hamiltonian systems. The main results about the qualitative behavior of nearly integrable Hamiltonian systems are formulated in the Kolmogorov-Arnold-Moser (KAM) Theorem and Nekhoroshev's Theorem on exponential estimates. In Section E.3 we will briefly describe these results and discuss possible connections to our particular application.

E.1 Standard Form

The abstract setting of Hamiltonian perturbation theory is the autonomous and canonical, n degree-of-freedom Hamiltonian system

$$\begin{aligned} \mathbf{q}' &= \frac{\partial H(\mathbf{q},\mathbf{p},\epsilon)}{\partial \mathbf{p}}, \\ \mathbf{p}' &= -\frac{\partial H(\mathbf{q},\mathbf{p},\epsilon)}{\partial \mathbf{q}}, \end{aligned} \tag{E.1}$$

with Hamiltonian

$$H(\mathbf{q}, \mathbf{p}, \epsilon) = H_0(\mathbf{q}, \mathbf{p}) + \epsilon H_1(\mathbf{q}, \mathbf{p}, \epsilon).$$

It is assumed that the system corresponding to H_0 is completely integrable, so that in the neighborhood of a compact level set action-angle variables $\boldsymbol{\varphi}$ and $\mathbf{I}$ can be introduced (cf. Section 3.4). In these variables the Hamiltonian becomes

$$H(\mathbf{I}, \varphi) = H_0(\mathbf{I}) + \epsilon H_1(\boldsymbol{\varphi}, \mathbf{I}, \epsilon),$$

and thus

$$\begin{aligned} \boldsymbol{\varphi}' &= \frac{\partial H(\boldsymbol{\varphi}, \mathbf{I}, \epsilon)}{\partial \mathbf{I}}, \\ \mathbf{I}' &= -\frac{\partial H(\boldsymbol{\varphi}, \mathbf{I}, \epsilon)}{\partial \boldsymbol{\varphi}}, \end{aligned} \tag{E.2}$$

corresponding to the integrable system

$$\begin{aligned} \boldsymbol{\varphi}' &= \boldsymbol{\omega}(\mathbf{I}) = \frac{\partial H_0(\mathbf{I})}{\partial \mathbf{I}}, \\ \mathbf{I}' &= \mathbf{0}. \end{aligned} \tag{E.3}$$

The standard form (E.2) of a Hamiltonian system builds the basis for the perturbation theorems in Section E.3.

To cast the Hamiltonian system governing the equilibrium conditions of elastic rods into a similar form it is essential to use a canonical Hamiltonian description, rather than the non-canonical we have adopted in Chapter 8, since perturbation expansions can only be obtained in a systematic way for canonical systems. However, abstract perturbation results like the KAM theory and Nekhoroshev's Theorem are also valid for non-canonical Hamiltonian systems, since locally the non-canonical Poisson bracket can always be written in canonical coordinates. This requirement provides no restriction for us, since the homogenization result in Chapter 8 is independent of the particular parametrization of the rotation group and therefore independent of the particular Hamiltonian structure that is adopted. In the following discussion we will consider the reduced Hamiltonian system describing the frame evolution, parametrized in terms of Euler angles, since only in this case it is possible to construct an action-angle variable transformation (cf. Chapter 7), which is essential in Hamiltonian perturbation theory.

Consider the Hamiltonian system (8.12) in Lemma 8.2 reduced to describe the frame evolution only. Instead of using the canonical description adopted there, we parametrize the body variables M and N in terms of the Euler angles $\boldsymbol{\phi}$ and their conjugate momenta $\mathbf{p}$ to obtain the six degree-of-freedom system

$$\begin{pmatrix} \boldsymbol{\phi}' \\ \mathbf{p}' \end{pmatrix} = \mathbf{J}_6 \nabla H\left(\boldsymbol{\phi}, \mathbf{p}, s, \frac{s}{\epsilon}, \epsilon\right), \tag{E.4}$$

where $H(\boldsymbol{\phi}, \mathbf{p}, s, \tau, \epsilon)$ is the Hamiltonian $\widetilde{H}(\mathsf{M}, \mathsf{N}, s, \tau, \epsilon)$ in Lemma 8.2 expressed in terms of $\boldsymbol{\phi}$ and $\mathbf{p}$; in particular the Hamiltonian H is 2π-periodic in the variable τ.

Then we construct the change of variables (8.17), i.e.

$$\begin{pmatrix} \boldsymbol{\phi} \\ \mathbf{p} \end{pmatrix} = \mathbf{x} + \epsilon \mathbf{w}\left(\mathbf{x}, s, \frac{s}{\epsilon}, \epsilon\right),$$

that generates the averaged system in Theorem 8.1. This change of variables can be chosen to be canonical so that the system becomes

$$\mathbf{x}' = \mathbf{J}_6 \nabla_{\mathbf{x}} \left[\overline{H}(\mathbf{x}, s) + \epsilon \widetilde{H}\left(\mathbf{x}, s, \frac{s}{\epsilon}, \epsilon\right)\right],$$

where

$$\overline{H}(\mathbf{x}, s) = \frac{1}{2\pi} \int_0^{2\pi} H(\mathbf{x}, s, \tau, 0) d\tau.$$

Now assume the averaged system defined by $\overline{H}$ is integrable, i.e. describes the motion of a Lagrange top (and therefore does not depend on s). In Theorem 7.1 we have shown how to obtain an action angle variable description for the Lagrange top; correspondingly we can write the Hamiltonian in terms of the variables $\boldsymbol{\varphi}$ and $\mathbf{J}$ to get

$$\mathcal{H}(\boldsymbol{\varphi}, \mathbf{J}, s, \frac{s}{\epsilon}, \epsilon) = \overline{H}(\mathbf{J}) + \epsilon \widetilde{H}\left(\boldsymbol{\varphi}, \mathbf{J}, s, \frac{s}{\epsilon}, \epsilon\right).$$

Finally, we change to the length scale $t = s/\epsilon$, introduce the new variables $\xi = t$ and I to obtain an augmented, autonomous Hamiltonian

$$H'(\boldsymbol{\varphi}, \mathbf{J}, \xi, I, \epsilon) = I + \epsilon \mathcal{H}(\boldsymbol{\varphi}, \mathbf{J}, \epsilon\xi, \xi, \epsilon), \tag{E.5}$$

which defines the system

$$\begin{aligned} \frac{d}{dt}\boldsymbol{\varphi} &= \epsilon \frac{\partial \mathcal{H}(\boldsymbol{\varphi}, \mathbf{J}, \epsilon\xi, \xi, \epsilon)}{\partial \mathbf{J}}, \\ \frac{d}{dt}\mathbf{J} &= -\epsilon \frac{\partial \mathcal{H}(\boldsymbol{\varphi}, \mathbf{J}, \epsilon\xi, \xi, \epsilon)}{\partial \boldsymbol{\varphi}}, \\ \frac{d}{dt}\xi &= 1. \end{aligned} \tag{E.6}$$

The variable ξ introduces an extra half degree of freedom in the Hamiltonian formulation since the equation for the variable I is redundant. System (E.6) is in a form, similar to the standard form (E.2), with integrable part $I + \epsilon\overline{H}(\mathbf{J})$ and perturbation $\epsilon^2 \widetilde{H}(\boldsymbol{\varphi}, \mathbf{J}, \epsilon\xi, \xi, \epsilon)$.

It should be noted that, although it is theoretically possible to construct this standard form, it does not seem feasible to compute it in an explicit form, for two reasons. First, recall that the action-angle variable description of the Hamiltonian $\overline{H}$ is only given in the implicit form (7.2). Moreover the perturbation term $\widetilde{H}$ must be expressed in terms of action-angle variables, which does not in general seem to be possible. Second, it is not obvious whether the change of variables (8.17) can be formulated explicitly.

E.2 Perturbation Expansions and Hamiltonian Normal Forms

Consider the abstract Hamiltonian system (E.2) in standard form

$$\boldsymbol{\varphi}' = \frac{\partial H(\mathbf{I}, \boldsymbol{\varphi})}{\partial \mathbf{I}}, \quad \mathbf{I}' = -\frac{\partial H(\mathbf{I}, \boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}},$$

with $H(\mathbf{I}, \boldsymbol{\varphi}) = H_0(\mathbf{I}) + \epsilon H_1(\mathbf{I}, \boldsymbol{\varphi}, \epsilon)$. In order to obtain a perturbation expansion in the small parameter ϵ one applies successive, canonical changes of variables

$$(\mathbf{I}, \boldsymbol{\varphi}) \to (\mathbf{J}, \boldsymbol{\psi}), \tag{E.7}$$

under which the Hamiltonian is formally transformed into the N^{th} order *normal form*

$$\mathcal{H}(\mathbf{J}, \epsilon) = \sum_{j=0}^{N} \epsilon^j \mathcal{H}_j(\mathbf{J}) + \mathcal{R}(\mathbf{J}, \boldsymbol{\psi}, \epsilon), \tag{E.8}$$

where $\mathcal{R} = O(\epsilon^{N+1})$. The approximate system given by $\sum_{j=0}^{N} \epsilon^j \mathcal{H}_j(\mathbf{J})$ is integrable with the action-variables $\mathbf{J}$ as integrals. A systematic method for the construction of a normal form has first been proposed by Poincaré [52] in what is today known as Lindstedt's method. It has been extended to include systems with degeneracies and resonances in von Zeipel's method (see [6, Ch. 5] or [40, App. 8] for a general survey on Hamiltonian normal form theory).

In general, the methods of Lindtstedt and von Zeipel do not converge, as the transformation (E.7) have to be chosen carefully to generate a convergent series (E.8), if it exists at all. The KAM theory resolves this problem by constructing a super-convergent series that allows one to integrate most systems close to the unperturbed system (see also the remarks in Section E.3).

Even though the above methods have a significant theoretical value, they can typically not be used to compute an asymptotic expansion in practice, as seems to be the case in our application. As we have remarked above, an explicit action-angle variable description of the system cannot be obtained, and the appearance of elliptic functions and integrals suggests that it is infeasible to compute even the term of order ϵ in the expansion.

There are however cases, for which normal forms can be computed efficiently and systematically. Assume that the Hamiltonian system

$$\mathbf{q}' = \frac{\partial H(\mathbf{q},\mathbf{p})}{\partial \mathbf{p}}, \quad \mathbf{p}' = -\frac{\partial H(\mathbf{q},\mathbf{p})}{\partial \mathbf{q}}$$

is expanded about an equilibrium solution, say $(\mathbf{p},\mathbf{q}) = \mathbf{0}$, such that the Hamiltonian takes the form

$$H(\mathbf{q},\mathbf{p}) = \sum_{j=2}^{\infty} H_j(\mathbf{q},\mathbf{p}),$$

where the terms H_j are polynomials of degree j. This is a special case of the Hamiltonian system (E.1) as the system given by the lowest order term H_2 is linear, and therefore integrable. Assume the equilibrium is stable so that the quadratic approximation H_2 of the system can be written as

$$H_2(\mathbf{q},\mathbf{p}) = \frac{1}{2}\sum_{i=1}^{n} \omega_i(q_i^2 + p_i^2), \tag{E.9}$$

or in action-angle variables $(\mathbf{I},\boldsymbol{\varphi})$, $I_i = (q_i^2 + p_i^2)/2$

$$H_2(\mathbf{I}) = \sum_{i=0}^{n} \omega_i I_i.$$

If there exists an integer vector $\mathbf{k}$ with $|\mathbf{k}| = l$ such that $\mathbf{k}\cdot\boldsymbol{\omega} = 0$, the characteristic frequencies $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)$ are said to satisfy a resonance relation of order l. Birkhoff [12] shows that if the characteristic frequencies do not satisfy a resonance relation of order less than or equal to l, then there exists a canonical change of variables $(\mathbf{q},\mathbf{p}) \to (\mathbf{Q},\mathbf{P})$, such that the new Hamiltonian $\mathcal{H}$ is in *Birkhoff normal form*, i.e.

$$\mathcal{H}(\mathbf{P},\mathbf{Q}) = \mathcal{H}_0(\mathbf{J}) + \mathcal{R}, \quad \mathcal{R} = O(|\mathbf{P}| + |\mathbf{Q}|)^{l+1},$$

where $J_i = (P_i^2 + Q_i^2)/2$, and $\mathcal{H}_0$ is polynomial of degree $[l/2]$ in $\mathbf{J} = (J_1, \dots, J_n)$. If there are any resonances in the system, Birkhoff's transformation allows only a reduction up to finite order into normal form. Gustavson [24] extends Birkhoff's results to allow arbitrary commensurability between the nonzero frequencies. Dragt and Finn [18] also included zero frequencies. It should be remarked that normal forms can also be obtained if the system is expanded about a periodic orbit rather than an equilibrium solution (see e.g. Meyer and Hall [45]).

The classical approach for calculating the canonical transformations consists of computing a generating function to define the change of variables. A more practical approach involving Lie algebraic techniques has been proposed by Deprit [14] and modified in various contexts by Dragt and Finn [17, 18] (see also

Koseleff [35]). This technique allows the use of symbolic manipulation packages to compute canonical transformations and the associated normal form in an automatic and efficient way, in particular for the polynomial expansion (E.9).

Ilyukhin [31] considers the Euler angle formulation governing linearly elastic rods, and expands the Hamiltonian about an equilibrium solution that corresponds to a straight and untwisted rod. He computes, adopting the classical method using generating functions, a Birkhoff normal form of order four. Even in this example the computations are cumbersome, and the practical value of the very lengthy approximate expansion is not obvious.

E.3 Quality of Approximation

In this section we discuss two of the fundamental results of Hamiltonian perturbation theory, which attempt to answer the question in what sense an integrable system can be regarded as an approximation of the perturbed system; these are in particular the KAM Theorem and Nekhoroshev's Theorem on exponential estimates.

In Section E.1 we have shown how averaging transforms the reduced Hamiltonian formulation governing the frame evolution into the standard form (E.6) to which corresponds the integrable system

$$\begin{aligned} \frac{d}{dt}\boldsymbol{\varphi} &= \epsilon\frac{\partial \overline{H}(\mathbf{J})}{\partial \mathbf{J}}, \\ \frac{d}{dt}\mathbf{J} &= \mathbf{0}. \end{aligned} \tag{E.10}$$

One answer about the quality of the approximation (E.10) for system (E.6) has been given in Theorem 8.1, although this theorem is not restricted to integrable systems. We have shown that on the length scale $t = s/\epsilon$ the solutions of (E.6) and (E.10) are guaranteed to be close of order ϵ on intervals of length $O(1/\epsilon)$. Thus the interpretation of this result for the length scale of interest, namely arclength s, only provides an estimate on an interval of length $O(1)$. Moreover, this estimate has to be interpreted with care, since in general nothing can be said about the actual length of the interval, which may in fact be very short. System (E.10) cannot be regarded as a global approximation of (E.6), as for most initial conditions no statement can be made about the quality of the approximation on intervals of infinite length. For the special case where the averaged system is integrable we will exploit the Hamiltonian structure to show that for most initial conditions the approximation is expected to be much better than implied by Theorem 8.1.

Consider the general perturbation problem (E.2). The phase space of the unperturbed (integrable) system (E.3) is foliated by invariant tori determined by some constant value of $\mathbf{I}$. A torus with rationally independent frequencies $\omega(\mathbf{I}) = \partial H_0(\mathbf{I})/\partial\mathbf{I}$ is said to be *non-resonant* (for a *resonant* torus there exists a non-zero integer vector $\mathbf{k}$ so that $\mathbf{k}\cdot\boldsymbol{\omega}(\mathbf{I}) = \mathbf{0}$). The system is called *nondegenerate* if $\det(\partial\boldsymbol{\omega}/\partial\mathbf{I}) \neq 0$.

The celebrated KAM theorem (see Kolmogorov [34], Arnold [4, 5] and Moser [47]) analyzes the structure of the phase space under small Hamiltonian perturbations (for an introduction to this subject see for example [6, Ch. 5], [40, Ch. 7], [7, App. 8]). The conclusions of this theorem may be summarized as follows: If the unperturbed system is nondegenerate, then for a sufficiently small perturbation most non-resonant tori with frequencies $\boldsymbol{\omega}(\mathbf{I})$ are not destroyed but are only slightly deformed, so that in the phase space of the perturbed system there are invariant tori, on which the motion is quasi-periodic with a frequency $\bar{\boldsymbol{\omega}}$ close to $\boldsymbol{\omega}(\mathbf{I})$. The frequencies $\boldsymbol{\omega}(\mathbf{I})$ must be chosen to be sufficiently far away from resonances.

The KAM Theorem actually also holds if the unperturbed system is *isoenergetically nondegenerate*, that is

$$\det\begin{pmatrix} \frac{\partial^2 H_0}{\partial^2\mathbf{I}} & \frac{\partial H_0}{\partial\mathbf{I}} \\ \frac{\partial H_0}{\partial\mathbf{I}} & 0 \end{pmatrix} \neq 0,$$

and this is the situation that applies in our case. Recall that we obtained the autonomous Hamiltonian formulation (E.6) by introducing an extra half degree of freedom to eliminate the periodic s-dependency. Then the integrable part $I + \epsilon\overline{H}$ of the Hamiltonian (E.5) corresponds to an isoenergetically nondegenerate system, if $\overline{H}$ describes a nondegenerate system (see the remarks in [6, p. 185]).

Arnold [4] (see also Sergeev [56]) verifies the non-degeneracy condition for the Hamiltonian system describing a Lagrange top. Thus the KAM Theorem applies, i.e. for sufficiently small perturbations and for most initial conditions, the frame evolution of the perturbed system differs little from the frame evolution of the unperturbed system, and is in particular quasi-periodic. This shows that for most initial conditions Theorem 8.1 actually holds on infinite time intervals when applied to the reduced system (E.4).

It remains to interpret this result for the centerline $\mathbf{r}$. In Section 6.2 we have introduced the components of $\mathbf{r}$ relative to a frame $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, where $\mathbf{k}$ is parallel to the force vector $\mathbf{n}$. In particular, we have shown that the two components $\mathbf{r}\cdot\mathbf{i}$ and $\mathbf{r}\cdot\mathbf{j}$ can be algebraically expressed in terms of variables describing the frame evolution. Therefore the interpretation of the KAM result for the frame evolution also applies for these components of $\mathbf{r}$; for most initial conditions the components $\mathbf{r}\cdot\mathbf{i}$ and $\mathbf{r}\cdot\mathbf{j}$ are quasi-periodic functions, and suffer only a minimal change under the perturbation. It is not obvious if the result is true for the third component $\mathbf{r}\cdot\mathbf{k}$.

Finally, we mention Nekhoroshev's Theorem, which is considered the averaging theorem for Hamiltonian systems (see e.g. the articles by Nekhoroshev [48] and Benettin et al. [8], or the surveys in [6, Ch. 5] and [40, Ch. 7]): Consider the Hamiltonian system (E.2) and assume that the unperturbed Hamiltonian H_0 is convex. Then for sufficiently small perturbations, the variables $\mathbf{I}$ of the perturbed system vary little on exponentially long time intervals, i.e. there exist constants $a, b, c_1, c_2 > 0$ such that

$$|\mathbf{I}(t) - \mathbf{I}(0)| < c_1 \epsilon^b \quad \text{for} \quad 0 \leq t \leq c_2 \frac{1}{\epsilon} \exp \frac{1}{\epsilon^a}.$$

The theorem also holds if H_0 is a *steep* function (for the definition of this property see for example [40, App. 9]). However, for system (E.6) we were not able to verify either the convexity or the steepness condition, since the elliptic functions and integrals that are involved in the formulation of the action-angle variable description make specific computations very difficult.

Appendix F

The Elastic Rod as a Model for Springs

A coiled rod can be regarded as modeling a spring, and we discuss here the results of Chapter 9 as it relates to various theories of springs, in particular "effective" theories, which replace the spring by a straight (effective) rod (or an *equivalent column*, as it is often termed in the engineering literature).

For our purposes we say that the rod is a spring, if in its reference configuration, the centerline of the rod is helical (consistent with Definition 9.1). In the engineering literature the derivation of accurate spring theories is of interest in particular for a stability analysis. Suppose a spring, which is simply supported at both ends, is subject to a compressive force. Then at a critical load the compressed helical state may experience a (typically undesirable) *bifurcation*, where the spring buckles away from its helical state. A bifurcation (or *buckling*) analysis can be difficult for helical rod models, and therefore several attempts to produce "effective" rod models have been made. The helical rod is replaced by a straight rod with appropriate coefficients for compression, shear, bending and twisting (see for example the recent work by Krużelecki and Życzkowski [36] and references therein), however these approximations are only valid for helical deformations of the spring.

It seems in general appropriate that the "effective" model should incorporate shear and compression (or extension) effects in addition to bending and twisting. In the first papers on this subject by Hurlbrink [30] and Grammel [22] shear effects however were neglected, and their buckling analysis was accordingly inaccurate.

Haringx [26] derives an "effective" model for certain deformations of a spring composed of a linearly elastic, uniform and isotropic material. The additional assumption that the pitch angle is small compared to the length of the rod motivates the approximation of the spring by a straight (and isotropic), extensible and shearable rod. He then computes the appropriate bending, twisting, compression and shearing rigidities for the spring, when regarded as a rod, in its unstressed and stressed configuration, respectively. These rigidities are expressed in terms of the material parameters of the spring and build the basis for a buckling analysis.

In particular, he considers two different cases: First a rod that is simply supported at both ends and subjected to a compressive force at both ends, and second a rod that is fixed at one end and subjected to a combined force and moment at the other end. Through a linearized analysis, critical buckling loads are computed for the rod, which can then be related to the critical relative compression of the rod (and of the spring, for this matter) through the compression rigidity computed above.

In a later series of articles, Haringx [27] (see also Wahl [62, Ch. 24]) extends these results by allowing the spring to be composed of an anisotropic, linearly elastic material. Furthermore he then investigates the buckling behavior of the spring, when regarded as a helical rod, and he uses these results to justify the above approximation.

Recently Krużelecki and Życzkowski [36]) have reinvestigated Haringx's results by allowing the spring to be replaced by a rod, where the pitch of the helix is not necessarily small.

We conclude with some remarks as to how Haringx's result relate to our theory of effective constitutive relations. First, Haringx's approximation of the spring by a compressible and shearable rod is in no way rigorously justified, in particular the smallness of the pitch angle is not exploited for a rigorous asymptotic analysis. The only justification appears through a comparison of the critical buckling load of the spring and the rod. Second, the approximation is not universally valid in that it does not provide an effective constitutive relation valid for all deformations. It can only be regarded as an approximation to the spring in compressed helical states.

In our analysis of locally helical rods, or for that matter springs, we assume that both pitch and radius of the helix are small of the same order (and the model therefore differs from Haringx's). We apply an asymptotic analysis to obtain an effective constitutive relation describing an inextensible, unshearable rod. An asymptotic result for Haringx's model, where only the pitch angle p is considered small, is still an open problem.

Appendix G

Mathematica Code

The following two Mathematica scripts implement the non-canonical and the canonical Euler parameter Hamiltonian formulations governing the equilibrium conditions of elastic rods, we have described in Section 4.2. In particular, we verify that the bracket defined by the operator (4.11) satisfies the Jacobi identity using the characterization in [51, Prop. 6.8]. Furthermore, we establish the appropriate count on integrals in involution, which is needed to characterize completely integrable systems (cf. Section 5.4). The representation theorem in Antman [2, Theorem 9.16] is used to impose the isotropy condition on the strain-energy function W.

```
(*
Mathematica Script, SK 8/97, LCVM at UMCP
    1.Verifies that the extended 12-dimensional Poisson bracket
      governing the equilibria of elastic rods satisfies the
      Jacobi identity.
    2.Computes number of integrals in involution for various
      Hamiltonians
*)
z={0,0,0};Z={z,z,z};
Id=IdentityMatrix[3];
(*
    Variables r, i, m and n of the system
*)
r={x1,x2,x3};
i={x4,x5,x6};
m={x7,x8,x9};
n={x10,x11,x12};
(*
    Components of strain
*)
uhat={a,b,c};
```

```
vhat={d,e,f};
x=Flatten[{r,i,m,n}];
(*
    Construct matrix a x, corresponding to cross product,
    i.e. Cp[a].b = a x b
*)
Cp[x_]:={{0,-x[[3]],x[[2]]},{x[[3]],0,-x[[1]]},
    {-x[[2]],x[[1]],0}}
RR=Cp[r];
II=Cp[i];
MM=Cp[m];
NN=Cp[n];
(*
    Generate Poisson bracket, i.e. matrix J
*)
J={Flatten[{z,z,RR[[1]],Id[[1]]}],
   Flatten[{z,z,RR[[2]],Id[[2]]}],
   Flatten[{z,z,RR[[3]],Id[[3]]}],
   Flatten[{z,z,II[[1]],z}],
   Flatten[{z,z,II[[2]],z}],
   Flatten[{z,z,II[[3]],z}],
   Flatten[{RR[[1]],II[[1]],MM[[1]],NN[[1]]}],
   Flatten[{RR[[2]],II[[2]],MM[[2]],NN[[2]]}],
   Flatten[{RR[[3]],II[[3]],MM[[3]],NN[[3]]}],
   Flatten[{-Id[[1]],z,NN[[1]],z}],
   Flatten[{-Id[[2]],z,NN[[2]],z}],
   Flatten[{-Id[[3]],z,NN[[3]],z}]};
(*
    Implements test for Jacobi identity
    according to Proposition 6.8 in
    Olver 'Applications of Lie Groups to Differential Equations'
    test == 0 implies Jacobi identity satisfied.
*)
K[i_,j_,k_,l_]:=Part[J,i,l]*D[Part[J,j,k],x[[l]]]+
    Part[J,k,l]*D[Part[J,i,j],x[[l]]]+
    Part[J,j,l]*D[Part[J,k,i],x[[l]]];
L[i_,j_,k_]:=Sum[K[i,j,k,l],{l,1,12}];
test=0;
$RecursionLimit=500;
For[ii=1,ii<=2,ii++,
    For[jj=1,jj<=2,jj++,
        For[kk=1,kk<=2,kk++,
```

```
            test+=Abs[L[ii,jj,kk]]]]];
Print[test];
(*
    Grad[f] computes the gradient of f:R^12->R
    Poisson[f,g] evaluates the Poisson bracket
    Grad[f].J.Grad[g]
*)
Grad[f_]:=Table[D[f,x[[i]]],{i,12}];
Poisson[f_,g_]:=Grad[f].J.Grad[g];
Needs["LinearAlgebra`CrossProduct`"];
(*
    Several Hamiltonian for rod mechanics
    Hlin: linearly elastic, diagonal, inextensible, unshearable
    Hgen: general, extensible, shearable
    Hiso: isotropic, extensible, shearable
    for Hiso we use the Representation Theorem for
    isotropic functions (Theorem 9.16 in
    Antman `Nonlinear Problems in Elasticity')
*)
Hlin=1/2 (K1 m[[1]]^2 + K2 m[[2]]^2 + K3 m[[3]]^2)  + n[[3]];
Hgen=Wstar[m[[1]],m[[2]],m[[3]],n[[1]],n[[2]],n[[3]]]+
    m.uhat + n.vhat;
Hiso=Wstar[m[[1]]^2+m[[2]]^2,n[[1]]^2+n[[2]]^2+
    m[[1]]n[[1]]+m[[2]]n[[2]],m[[3]],n[[3]]]+n[[3]];
(*
Casimirs
*)
C1=1/2 i.i;
C2=i.(m+Cross[r,n]);
j=Cross[n,i];
(*
    Integrals, in particular I5 only is an integral if the rod
    is isotropic, and H is an integral if the system is
    autonomous (although in this script  it is always autonomous).
    I1, I2, I3 and I4 are always integrals
*)
I1=(m+Cross[r,n]).j;
I2= m.n;
I3= n.n;
I4= i.n;
I5=m[[3]];
```

```
(*
    Now compute the number of integrals in involution.
    For that choose a particular form of the Hamiltonian H,
    compute the involution matrix Invol, then
    no = (rank[Invol]+dim[Nullspace[Invol]])/2
    is the number of integrals in involution.
    As written below, the result should be
    no = 4, if H=Hlin or H=Hgen,
    no = 5, if H=Hlin (and K1=K2) or H=Hiso,
    no = 3, if I5 and H are not included in Inv (they are
    always in involution with all the other integrals !!!)
*)
H=Hlin;
Inv={I1,I2,I3,I4,I5,H};
Invol=Simplify[Transpose[Grad[Inv]].J.Grad[Inv]];
dimn=Dimensions[NullSpace[Invol]];
no=(dimn[[1]]+dimn[[2]])/2;
Print[no];
```

```
(*
    Mathematica Script, SK 8/97, LCVM at UMCP
    Implements the Euler parameter Hamiltonian formulation for
    elastic rods.
    We verify the count on integrals in involution under
    appropriate symmetry assumptions
*)
Needs["LinearAlgebra`CrossProduct`"];
(*
    The matrices B_i and F_i are used to represent the components
    of the moment vector m relative to the director frame,
    repectively the fixed frame.
*)
B1[q_]:={q[[4]],q[[3]],-q[[2]],-q[[1]]};
B2[q_]:={-q[[3]],q[[4]],q[[1]],-q[[2]]};
B3[q_]:={q[[2]],-q[[1]],q[[4]],-q[[3]]};
F1[q_]:={q[[4]],-q[[3]],q[[2]],-q[[1]]};
F2[q_]:={q[[3]],q[[4]],-q[[1]],-q[[2]]};
F3[q_]:={-q[[2]],q[[1]],q[[4]],-q[[3]]};
(*
    The variable rr (centerline), q (quaternion), nn force),
    mu (conjugate to q) for the Euler parameter forumlation.
*)
rr={x1,x2,x3};
q={x4,x5,x6,x7};
nn={x8,x9,x10};
mu={x11,x12,x13,x14};
uhat={u1,u2,u3};
x=Flatten[{rr,q,nn,mu}];
(*
    Directors in terms of variables q
*)
d1={q[[1]]^2-q[[2]]^2-q[[3]]^2+q[[4]]^2,
2*(q[[1]]q[[2]]+q[[3]]q[[4]]),
2*(q[[1]]q[[3]]-q[[2]]q[[4]])};
d2={2(q[[1]]q[[2]]-q[[3]]q[[4]]),
-q[[1]]^2+q[[2]]^2-q[[3]]^2+q[[4]]^2,
2(q[[2]]q[[3]]+q[[1]]q[[4]])};
d3={2(q[[1]]q[[3]]+q[[2]]q[[4]]),
2(q[[2]]q[[3]]-q[[1]]q[[4]]),
-q[[1]]^2-q[[2]]^2+q[[3]]^2+q[[4]]^2};
(*
```

```
    mm components of moment vector in fixed frame
    m  components of moment vector in moving frame
    n  components of force vector in moving frame
*)
mm=1/2{F1[q].mu,F2[q].mu,F3[q].mu};
m=1/2{B1[q].mu,B2[q].mu,B3[q].mu};
n={nn.d1,nn.d2,nn.d3};
Id=IdentityMatrix[7];
z={0,0,0,0,0,0,0};
(*
    canonical 14 by 14 J
*)
J={Flatten[{z,Id[[1]]}],
Flatten[{z,Id[[2]]}],
Flatten[{z,Id[[3]]}],
Flatten[{z,Id[[4]]}],
Flatten[{z,Id[[5]]}],
Flatten[{z,Id[[6]]}],
Flatten[{z,Id[[7]]}],
Flatten[{-Id[[1]],z}],
Flatten[{-Id[[2]],z}],
Flatten[{-Id[[3]],z}],
Flatten[{-Id[[4]],z}],
Flatten[{-Id[[5]],z}],
Flatten[{-Id[[6]],z}],
Flatten[{-Id[[7]],z}]};
(*
    Canonical Poisson Bracket
*)
Grad[f_]:=Table[D[f,x[[i]]],{i,14}];
Poisson[f_,g_]:=Grad[f].J.Grad[g];
(*
    Several Hamiltonians written in terms of Euler parameters
    Hgen: general extensible, shearable
    Hiso: isotropic
*)
Hgen=Wstar[m[[1]],m[[2]],m[[3]],n[[1]],n[[2]],n[[3]]]
+m.uhat+n[[3]];
Hiso=Wstar[m[[1]]^2+m[[2]]^2,n[[1]]^2+n[[2]]^2,
m[[1]]n[[1]]+m[[2]]n[[2]],m[[3]],n[[3]]]+n[[3]];
(*
    Integrals in Euler parameter formulation. Note that I7 and I8
```

```
    only appear in this formulation.
    I9 is an integral if the rod is isotropic
    and H is an integral if the system is autonomous
*)
I1=(mm+Cross[rr,nn])[[1]];
I2=(mm+Cross[rr,nn])[[2]];
I3=(mm+Cross[rr,nn])[[3]];
I4=nn[[1]];
I5=nn[[2]];
I6=nn[[3]];
I7=q.q;
I8=mu.q+2rr.nn;
I9=m[[3]];
(*
    Compute involution matrix Invol, then
    no = (rank(Invol)+dim(nullspace[Invol]))/2
    is the number of integrals in involution.
*)
Inv={I1,I2,I3,I4,I5,I6,I7,I8,I9};
Invol=Simplify[Transpose[Grad[Inv]].J.Grad[Inv]];
dimn=Dimensions[NullSpace[Invol]];
no=(dimn[[1]]+dimn[[2]])/2;
Print[no];
(*
    In the above example the result is no=6.
    If you include the Hamiltonian H=Hiso in Invol then the
    program runs forever. However, you may individually check
    that Hiso  commutes with all the other integrals,
    i.e.  Poisson[I_i, H]==0, which raises the total number of
    integrals in involution to 7.
*)
```

BIBLIOGRAPHY

[1] A.M. Aksenenkova. Canonical action-angle variables in the problem of Lagrange gyroscope. *Moscow State University, Ser. Matem. Mekh.*, 1, 1981.

[2] S.S. Antman. *Nonlinear Problems of Elasticity.* Springer-Verlag, New York, 1995.

[3] S.S. Antman and K.B. Jordan. Qualitative aspects of the spatial deformation of non-linearly elastic rods. *Proc. Roy. Soc. Edinburgh*, 73A(5):85–105, 1974.

[4] V.I. Arnold. Proof of a theorem of A.N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian. *Russian Math. Surv.*, 18(5):9–36, 1963.

[5] V.I. Arnold. Small denominators and problems of stability of motion in classical and celestical mechanics. *Russian Math. Surv.*, 18(6):85–192, 1963.

[6] V.I. Arnold. *Dynamical Systems III*, volume 3 of *Encyclopaedia of Mathematical Sciences.* Springer-Verlag, New York, 1988.

[7] V.I. Arnold. *Mathematical Methods of Classical Mechanics.* Springer-Verlag, 1989.

[8] G. Benettin, L. Galgani, and A. Giorgilli. A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems. *Celest. Mech.*, 37:1–25, 1985.

[9] C.J. Benham. An elastic model of the large-scale structure of duplex DNA. *Biopolymers*, 18:609–623, 1979.

[10] C.J. Benham. Geometry and mechanics of DNA superhelicity. *Biopolymers*, 22:2477–2495, 1983.

[11] A. Bensoussan, J.L. Lions, and G. Papanicolaou. *Asymptotic Analysis for Periodic Structures.* North-Holland Publishing Company, 1978.

[12] G.D. Birkhoff. *Dynamical Systems*, volume 9 of *American Mathematical Colloquium Publications.* American Mathematical Society, 1927.

[13] P.F. Byrd. *Handbook of elliptic integrals and functions.* Springer-Verlag, New York, 1971.

[14] A. Deprit. Canonical transformations depending on a small parameter. *Celest. Mech.*, 1:12–30, 1969.

[15] D.J. Dichmann. *Hamiltonian Dynamics of an Elastica and the Stability of Solitary Waves.* PhD thesis, University of Maryland, 1994.

[16] D.J. Dichmann, Y. Li, and J.H. Maddocks. Hamiltonian formulations and symmetries in rod mechanics. In J.P. Mesirov, K. Schulten, and D.W. Sumners, editors, *Mathematical Approaches to Biomolecular Structure and Dynamics*, volume 82 of *The IMA Volumes in Mathematics and Its Applications*, pages 71–113. 1996.

[17] A.J. Dragt and J.M. Finn. Lie series and invariant functions for analytic symplectic maps. *J. Math. Phys.*, 17(12):2215–2227, 1976.

[18] A.J. Dragt and J.M. Finn. Normal form for mirror machine Hamiltonians. *J. Math. Phys.*, 20(12):2649–2660, 1979.

[19] L.C. Evans. *Partial Differential Equations*, volume 3A of *Berkeley Mathematics Lecture Notes.* Center for Pure and Applied Mathematics, University of California, Berkeley, 1994.

[20] M. Farkas. *Periodic Motions.* Springer-Verlag, New York, 1994.

[21] H. Goldstein. *Classical Mechanics.* Addison-Wesley, 1980.

[22] R. Grammel. Die Knickung von Schraubenfedern. *Z. Angew. Mat. Mech.*, 4:384–389, 1924.

[23] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vectorfields.* Springer Verlag, New York, 1983.

[24] F.G. Gustavson. On constructing formal integrals of a Hamiltonian system near an equilibrium point. *Astron. J.*, 71(8):670–686, 1966.

[25] J. Hale. *Ordinary Differential Equations.* Robert E. Krieger Publishing Company, Inc., Malabar, Florida, 1990.

[26] J.A. Haringx. On the buckling and lateral rigidity of helical compression springs. *Proc. Nederl. Akad. Wet.*, 45:533–539, 650–654, 1942.

[27] J.A. Haringx. On highly compressible helical springs and rubber rods, and their applications for vibration-free mountings. *Philips Res. Reports*, 3:401–449; 4:49–80, 206–220 1948-1949.

[28] H. Heuser. *Lehrbuch der Analysis*, volume 2. Teubner, Stuttgart, 1986.

[29] P. Holmes and J. Marsden. Horseshoes and Arnold diffusion for Hamiltonian systems on Lie groups. *Indiana Univ. Math. J.*, 32(2):273–309, 1983.

[30] E. Hurlbrink. Berechnung zylindrischer Druckfedern auf Sicherheit gegen seitliches Ausknicken. *Z. Ver. Dt. Ing.*, 54:133–137, 181–184, 1910.

[31] A.A. Ilyukhin. *Spatial Problems of Nonlinear Theory of Elastic Rods.* Naukova Dumka, Kiev, 1979. in Russian.

[32] V.V. Jikov, S.M. Kozlov, and O.A. Oleinik. *Homogenization of Differential Operators and Integral Functionals.* Springer-Verlag, New York, 1994.

[33] F. Klein and A. Sommerfeld. *Über die Theorie des Kreisels.* Teubner, 1965.

[34] A.N. Kolmogorov. Preservation of conditionally periodic movements with small change in the Hamiltonian function. In G. Casati and J. Ford, editors, *Stochastic Behavior in Classical and Quantum Hamiltonian Systems*, volume 93 of *Lecture Notes in Physics*, pages 51–56. Springer-Verlag, 1979.

[35] P.-V. Koseleff. *Formal Calculus for Lie Methods in Hamiltonian Mechanics.* PhD thesis, École Polytechnique, 1993. translated by Lawrence Berkeley Laboratory, University of California.

[36] J. Krużelecki and M. Życzkowski. On the concept of an equivalent column in the stability problem of compressed helical springs. *Ingenieur-Archiv*, 60:367–377, 1990.

[37] J. Langer and D.A. Singer. Lagrangian aspects of the Kirchhoff elastic rod. *SIAM Review*, 38(4):605–618, 1996.

[38] H.-H. Li. *Geometrically Constrained Curves and the Calculus of Variations.* PhD thesis, Rice University, 1986.

[39] Y. Li and J.H. Maddocks. On the computation of equilibria of elastic rods, Part I: Integrals, symmetry and a Hamiltonian formulation. *J. Comp. Phys.*, 1996. to appear.

[40] P. Lochak and C. Meunier. *Multiphase Averaging for Classical Systems.* Springer-Verlag, 1988.

[41] A.E.H. Love. *A Treatise on the Mathematical Theory of Elasticity.* Dover Publishing, New York, 4th edn., 1944.

[42] J.H. Maddocks and D.J. Dichmann. Conservation laws in the dynamics of rods. *J. Elasticity*, 34:83–96, 1994.

[43] J.H. Maddocks. Stability of nonlinearly elastic rods. *Arch. Rational Mech. Anal.*, 85(4):180–198, 1984.

[44] R.S. Manning, J.H. Maddocks, and J.D. Kahn. A continuum rod model of sequence-dependent DNA structure. *J. Chem. Phys.*, 105(13):5626–5646, 1996.

[45] K.R. Meyer and G.R. Hall. *Introduction to Dynamical Systems and the N-Body Problem.* Springer-Verlag, 1992.

[46] A. Mielke and P. Holmes. Spatially complex equilibria of buckled rods. *Arch. Rational Mech. Anal.*, 101:318–348, 1988.

[47] J. Moser. On invariant curves of area preserving mappings of an annulus. *Nachr. Acad. Wiss. Göttingen Math. Phys. K1*, 11a(1):1–20, 1962.

[48] N.N. Nekhoroshev. An exponential estimate of the time of stability of nearly-integrable Hamiltonian systems. *Russian Math. Surv.*, 32(6):1–65, 1977.

[49] S.P. Novikov. The Hamiltonian formalism and a many-valued analogue of Morse theory. *Russian Math. Surv.*, 37(5):1–56, 1982.

[50] W.K. Olson. Simulating DNA at low resolution. *Current Opinion in Structural Biology*, 6:242–256, 1996.

[51] P.J. Olver. *Applications of Lie Groups to Differential Equations.* Springer-Verlag, New York, 1993.

[52] H. Poincaré. *Le méthodes nouvelles de la mécanique céleste*, volume 1. Gauthier-Villars, Paris, 1892.

[53] K.A. Rogers. *Stability Exchange in Parameter-Dependent Constrained Variational Principles with Applications to Elastic Rod Models of DNA Minicircles.* PhD thesis, University of Maryland, 1997.

[54] J.A. Sanders and F. Verhulst. *Averaging Methods in Nonlinear Dynamical Systems.* Springer-Verlag, New York, 1985.

[55] T. Schlick. Modeling superhelical DNA: Recent analytical and dynamical approaches. *Current Opinion in Structural Biology*, 5:245–262, 1995.

[56] V.S. Sergeev. Periodic motions of a heavy, rigid, dynamically almost symmetric body with a fixed point. *J. Appl. Math. Mech.*, 47(1):130–133, 1984.

[57] Y. Shi and J.E. Hearst. The Kirchhoff elastic rod, the nonlinear Schrödinger equation, and DNA supercoiling. *J. Chem. Phys.*, 101(6):5186–5200, 1994.

[58] M.D. Shuster. A survey of attitude representations. *J. Astronaut. Sci.*, 41:439–518, 1994.

[59] J.C. Simo, J.E. Marsden, and P.S. Krishnaprasad. The Hamiltonian structure of nonlinear elasticity: the material and convective representations of solids, rods, and plates. *Arch. Rat. Mech. Anal.*, 104:125–183, 1988.

[60] E.L. Starostin. Three-dimensional shapes of looped DNA. *Meccanica*, 31:235–271, 1996.

[61] Z. Tan and J.A. Witz. On the deflected configuration of a slender elastic rod subject to parallel terminal forces and moments. *Proc. Roy. Soc. Lond. A*, 449:337–349, 1995.

[62] A.M. Wahl. *Mechanical Springs.* McGraw-Hill, New York, 1963.

[63] L.S. Wang, P.S. Krishnaprasad, and J.H. Maddocks. Hamiltonian dynamics of a rigid body in a central gravitational field. *Celest. Mech. Dyn. Astron.*, 50:59–96, 1991.

[64] L.S. Wang, J.H. Maddocks, and P.S. Krishnaprasad. Steady rigid body motions in a central gravitational field. *J. Astron. Sci.*, 40(4):449–478, 1992.

[65] J.A. Warner. Numerical simulations of elastic rods. Master's thesis, University of Maryland, 1997.

[66] J. Wittenburg. *Dynamics of Systems of Rigid Bodies.* Teubner, Stuttgart, 1977.