## Abstract

This dissertation is concerned with two problems that lie at the interface of soft matter physics, geometry, and asymptotic analysis, but otherwise have no bearing on one another. In the first problem, I consider the equilibrium thermal fluctuations of deformable mechanical frameworks. These frameworks have served as highly idealized representations of mechanical structures that underlie a plethora of soft few-body systems at the submicron scale such as colloidal clusters and DNA origami. When the holonomic constraints in a framework cease to be linearly independent, singularities can appear in its configuration space, where it becomes energetically softer. Consequently, the framework's free-energy landscape becomes dominated by the neighborhoods of points corresponding to these singularities. In the second problem, I study the localization of elastic waves in thin elastic structures with spatially varying curvature profiles, using a curved rod and a uniaxially curved shell as concrete examples. Waves propagating on such structures have multiple components owing to the curvature-mediated coupling of the tangential and normal components of the displacement field. Here, using the semiclassical approximation, I show that these waves form localized, bound states around points where the absolute curvature of the structure has a minimum. Both these problems exemplify the subtle interplay between the mechanical properties of soft materials and their geometry, which further sets the stage for many interesting consequences.

# ASYMPTOTICS, GEOMETRY, AND SOFT MATTER

*by*

MANU MANNATTIL

M.Sc. (Integrated), Indian Institute of Technology Kanpur, 2015

Dissertation
Submitted in Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy in Physics


SYRACUSE UNIVERSITY

August 2023

Git commit 6db7f87; 2023-09-06 02:20:48 +0530

## Acknowledgments

I thank my adviser Chris Santangelo for taking me on as his Ph.D. student, for many useful conversations, and for giving me ample independence (among other things). I also thank Jen Schwarz, Joey Paulsen, and Will Wylie for serving on my dissertation committee. I am also grateful to Chris and Jen for helping me transition to greener pastures in soft matter physics after an unwise stint "doing" high-energy theory. Special thanks are due to my undergraduate adviser Sagar Chakraborty, with whom I did my first serious research work. Let me also collectively acknowledge all others (too numerous to list individually, yet too generous to forget) who helped me reach this stage of my academic career. Finally, I would like to thank my parents for their compassion and encouragement over the years.

## Contents

## Notation

Following the usual convention in physics literature, vectors in $\mathbb{R}^n$ are usually set in bold Latin letters, e.g., $\boldsymbol{v}$. Matrices are set in sans serif type, e.g., $\mathsf{M}$. Operators are distinguished with a hat, e.g., $\hat{a}$. Integrals without explicit limits are to be integrated over the entire range (usually $-\infty$ to $\infty$) of the integration variable(s). Unless explicitly indicated, repeated indices are to be summed over as usual. Other notational conventions are listed below.

| | |
|---|---|
| $\|\boldsymbol{v}\|$ | Euclidean norm $\sqrt{\boldsymbol{v}^\mathsf{T}\boldsymbol{v}}$ of a vector $\boldsymbol{v} \in \mathbb{R}^n$ |
| $\mathsf{I}_n$ | $n \times n$ identity matrix |
| $\nabla\phi$ | gradient of $\phi : \mathbb{R}^n \to \mathbb{R}$ considered as a row vector, or<br>the $m \times n$ Jacobian matrix of a map $\phi : \mathbb{R}^n \to \mathbb{R}^m$ |
| $(\nabla\phi)^\mathsf{T}$ | transpose gradient of $\phi : \mathbb{R}^n \to \mathbb{R}$ considered as a column vector, or<br>the $n \times m$ transpose of the Jacobian matrix of a map $\phi : \mathbb{R}^n \to \mathbb{R}^m$ |
| $\nabla\nabla\phi$ | $n \times n$ Hessian matrix of a scalar function $\phi : \mathbb{R}^n \to \mathbb{R}$ |
| $\det\mathsf{M}$ | determinant of a matrix $\mathsf{M}$ |
| $\ker\mathsf{M}$ | kernel (null space) of a matrix $\mathsf{M}$ |
| $\operatorname{tr}\mathsf{M}$ | trace of a matrix $\mathsf{M}$ |
| $\mathcal{O}(\cdot)$ | of the order of |

Chapter 1

# Introduction

This dissertation discusses two problems that are more or less unrelated apart from having a common origin in soft matter physics. During the course of the discussion, we shall rely on concepts from a potpourri of fields ranging from classical and quantum mechanics to statistical mechanics to structural engineering. Elementary notions from differential geometry and asymptotic analysis will also play a prominent role. This chapter provides a whirlwind tour of the dissertation, highlighting the key results, and concludes with an organizational summary.

Geometrical methods are now a mainstay of all branches of modern theoretical physics. Furthermore, vigorous research efforts in the second half of the 20th century led to the geometrization of the more classical fields of physics such as analytical mechanics [3, 156, 161] and elasticity theory [4, 112]. In some sense, the unreasonable effectiveness of geometrical methods in physics (among other mathematical sciences) should come as no surprise—after all, many physical problems are best formulated mathematically in terms of configuration (or parameter) spaces that are not Euclidean.

On a less abstract level, the intrinsic shape and structure of a physical system can also play a major role in dictating its physical properties. This is particularly true for soft, deformable materials, or "soft matter"—the stuff of everyday things. Indeed, mechanical pliability is often intimately connected to geometry, a fact that is appetizingly illustrated by a slice of pizza, which becomes stiff upon bending into a U shape, thanks to Gauss's remarkable theorem. Because soft materials are easily deformed, the presence of external perturbations such as thermal fluctuations or applied fields can have dramatic effects on their stability. For this reason, creating materials with designer microstructure to improve their strength has become a central research theme in many disciplines, with applications cutting across several length and energy scales.

Theoretical physics abounds with exactly solvable problems and "spherical cow" models involving conserved quantities, rigid bodies, ideal fluids, point particles, impenetrable walls, uniform fields, and elegant linear equations. Many of these models, however, stand in stark contrast to the fantastic haphazardness of the real world, which is messy and nonlinear, and continues to bewilder us as our experimental abilities evolve. For this reason, many physical problems, especially those in condensed matter and materials physics—archetypal examples of the sentiment that "more is different" [2]—can only be formulated approximately. Furthermore, such problems often require the employment of a range of asymptotic and perturbative methods for their solution. Such methods, perhaps fittingly, tend to be far

less rigorous in comparison to the elegant logical structure usually found in other fields of mathematics.

In this dissertation we discuss two problems, at two very different length and energy scales, that have a common origin in soft matter physics. Connecting the two problems is the basic notion that geometry, whether that of the abstract spaces describing a physical system, or its intrinsic shape and structure, plays a crucial role in dictating its properties. Both these problems, once formulated mathematically, have sufficient complexity that makes writing down exact solutions difficult. At the same time, both the problems are simple enough that asymptotic methods yield excellent approximate solutions, which means that we do not have to restrict ourselves to analyses of numerical experiments alone. In the following sections, we briefly summarize the main results of this dissertation, with subsequent chapters providing more detailed descriptions.

## 1.1 Singular frameworks and thermal fluctuations

In the first part of the dissertation, we study the effect of thermal fluctuations on submicron mechanical frameworks. Frameworks of similar nature have received extensive interest in recent years, with considerable attention paid to nanoscale frameworks made out of DNA origami [36, 114]. The problem is made fundamentally challenging due to the presence of singularities in the configuration spaces of these frameworks, which cause conventional approaches to break down [109].

### 1.1.1 Configuration spaces

Central to geometric mechanics is the idea that the configuration of a physical system, however complicated, can be fully described by a single point of a (usually high-dimensional) configuration space. To illustrate this point in more familiar settings, consider, for instance, the simple pendulum in Fig. 1.1(a). The pendulum's configuration at any given moment is fully specified by the angle $\theta_1$. Its configuration space is therefore equivalent to the circle $S^1$, which is a smooth manifold. To specify the configuration of the double pendulum in Fig. 1.1(b), on the other hand, requires two angles $\theta_1$ and $\theta_2$, and its configuration space is the torus $S^1 \times S^1$. Configuration spaces such as these, although abstract, have a one-to-one correspondence with the possible configurations that the system can be in.

To shed some more light on the above discussion and make the definitions more precise, consider a mechanical system with $n$ degrees of freedom (DOF), whose configuration at any given moment is fully described by a single configuration vector $\boldsymbol{r} \in \mathbb{R}^n$. Constraints in such a system are most clearly introduced by defining a constraint map $f : \mathbb{R}^n \to \mathbb{R}^m$ that vanishes when the constraints are satisfied, with $m$ being the number of constraints introduced. The constraint in the simple pendulum, for instance, is described by the constraint map

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(r_1, r_2) = r_1^2 + r_2^2 - \ell^2, \tag{1.1}$$

whereas the constraint map for the double pendulum would be of the form

$$f : \mathbb{R}^4 \to \mathbb{R}^2, \quad f(r_1, \ldots, r_4) = \begin{pmatrix} r_1^2 + r_2^2 - \ell_1^2 \\ (r_3 - r_1)^2 + (r_4 - r_2)^2 - \ell_2^2 \end{pmatrix}. \tag{1.2}$$

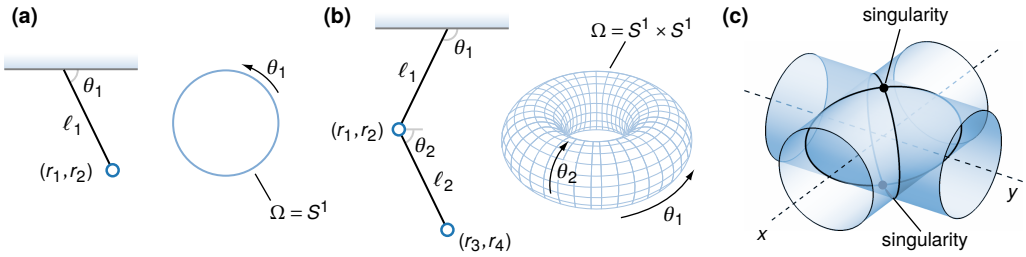


**Figure 1.1.** (a) A simple pendulum and its configuration space $\Omega = S^1$. (b) A double pendulum and its configuration space $\Omega = S^1 \times S^1$. (c) Configuration space of a particle constrained to move on two mutually perpendicular cylinders of equal radius is the cylinders' intersection curve, which is not a smooth manifold.

As we see from these examples, the constraint map $f$ is a general nonlinear map in the components of $\boldsymbol{r}$. The linear approximation of $f = [f_1(\boldsymbol{r}), \ldots, f_m(\boldsymbol{r})]$ is given by the $m \times n$ Jacobian matrix $\nabla f$ whose $ij$th entry is $\partial_j f_i$. With these definitions, the configuration space of a general mechanical system is the zero level set $\Omega = \{\boldsymbol{r} \in \mathbb{R}^n : f(\boldsymbol{r}) = 0\}$, which is the set of points where the constraints are satisfied exactly. Standard theorems[1] ensure that $\Omega$ is a smooth $(n-m)$-dimensional manifold if the Jacobian $\nabla f$ has full rank for all points in $\Omega$.

Configuration spaces in both the above examples arose as a result of holonomic constraints imposed on a physical system. However, the imposed holonomic constraints need not always be well-behaved. To illustrate this point, consider a particle constrained to move on two intersecting cylinders of equal radius with mutually perpendicular axes, as illustrated in Fig. 1.1(c). If the particle is to obey both constraints simultaneously, it can only move on the cylinders' intersection curve, which forms the configuration space of the particle. As we see from Fig. 1.1(c), the curve has two singularities where it self-intersects, which prevents the configuration space from being a smooth manifold. Mathematically, the constraints in the two-cylinder system are described by the map $f : \mathbb{R}^3 \to \mathbb{R}^2$, defined by $f(x, y, z) = (x^2 + z^2 - 1, y^2 + z^2 - 1)$. At singularities, such as the ones in Fig. 1.1(c), direct computation reveals that the Jacobian $\nabla f$ drops rank.[2] Such singularities, which arise when the constraints imposed on a system cease to be linearly independent, are not merely pathological irregularities, and they have been extensively studied in many fields, e.g., robotics and locomotion [123].

### 1.1.2 Frameworks

Soft few-body systems, composed of a small number of particles or units, interacting via short-ranged forces are commonplace in soft matter physics. A quintessential example of

[1] These theorems are almost never explicitly invoked in classical mechanics. However, they are implicit in the frequently used argument that a mechanical system with $n$ degrees of freedom and $m$ constraints has $(n-m)$ degrees of freedom, with the configuration space $\Omega$ parameterizable by $(n-m)$ generalized coordinates.

[2] Since the Jacobian is an $m \times n$ matrix, it drops rank whenever its rows—each representing a single linearized constraint—cease to become linearly independent.

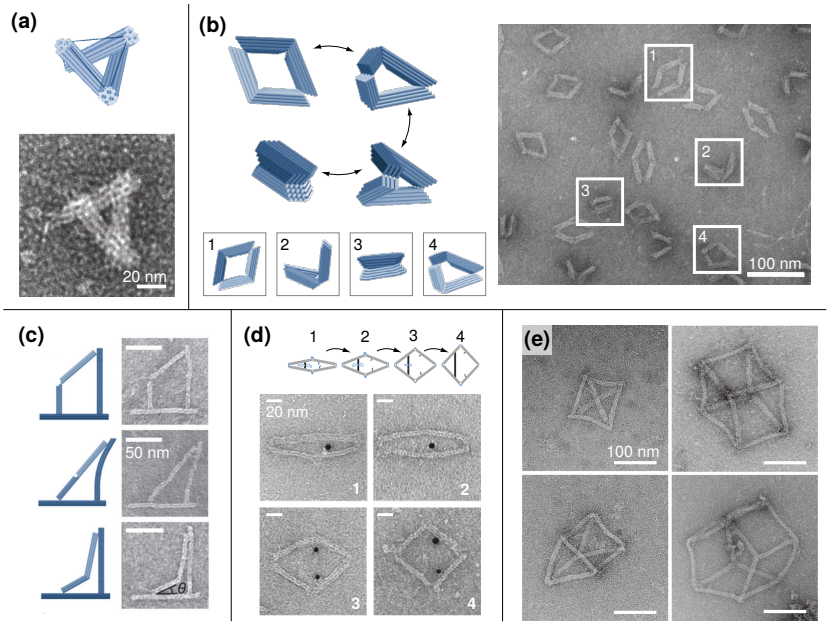


**Figure 1.2.** DNA origami has been widely used to self-assemble a variety of objects at the nanoscale. Depicted in the figure are (a) tensegrity structures [98]; (b), (c) linkage-based mechanisms [111, 181]; (d) a rhombus-shaped nanoactuator [84]; and (e) self-assembled polyhedra [70].

such a system is a colloidal cluster, which is composed of a small number of colloidal particles with sizes typically ranging from 1 nm to 0.1 $\mu$m. Although the particles in a colloidal cluster are held together by subtle effects of electrostatic and van der Waal forces, such details are often irrelevant if our goal is to describe the more macroscopic properties of the cluster. To a good approximation, therefore, the interactions between the different particles are effectively described using central-force bonds connecting their centers. Configurations of the resulting "bond skeleton", with each bond at its respective rest length, describe the different ground states of the cluster. Such colloidal skeletons are an example of a mechanical framework, which is an assembly of rigid bars that connect point-like joints. Frameworks are indeed holonomically constrained systems, not too different from the examples we have seen so far.

Apart from colloidal clusters, mechanical frameworks have been extensively used to understand a variety of mechanical structures found in viruses [67], crystals [140] and minerals [79], proteins [48], and of course, robots and machines [32, 41]. In recent years, nanoscale frameworks made out of multihelix DNA bundles (often dubbed "DNA origami") have received extensive interest, with applications ranging from drug delivery [180] to self-assembly [98]. As far as more generic descriptions of thermally driven frameworks are concerned, there has been long-standing interest in the effect of thermal fluctuations on the mechanical properties of

ordered and disordered lattices [172, 175, 179], and the folding of polymerized membranes [28, 125] and polyhedral nets [31, 117, 152]. There is, therefore, an arising need to understand how thermal excitations affect the physical properties of these frameworks, but only some attempts have been made so far [77, 147].

### 1.1.3 Free-energy landscapes

The effect of thermal fluctuations on a physical system is often represented by its free-energy landscape in terms of a set of collective variables that provide a coarse-grained description of its slowest dynamics. In theory [38, 50], one can obtain the free energy of a framework by integrating out the fast modes that are transverse to its shape space, i.e., the subset of its configuration space once rigid-body motions are removed. Doing this, however, becomes nontrivial when the framework has shape-space singularities [32, 103, 182]. For concreteness, consider the shape space of the planar four-bar linkage [55, 60, 153] with freely rotating joints [Figs. 1.3(a) and 1.3(c)]. Though this linkage has one degree of freedom up to Euclidean motions, it has two modes of deformation, one where the angle $\theta_1 = \theta_2$ and another where $\theta_1 \neq \theta_2$, meeting at two isolated singular points $(\theta_1, \theta_2) = (0, 0)$ and $(\pi, \pi)$. One generically expects the framework to be soft at these singularities, and indeed, as we see from the dashed, blue curves in Fig. 1.3(b), the free energy diverges in a harmonic approximation of the elastic energy [147]. These divergences must be cut off by higher-order nonlinear effects, yet how this happens and to what extent remains to be understood.

In Chapters 2 and 3, we develop a formalism to understand the thermal equilibration of common bar-joint frameworks that have isolated shape-space singularities. We show that the divergent contributions to the free energy arising in the harmonic approximation to the elastic energy are suppressed by anharmonic, quartic-order corrections. These findings show the existence of energetic free-energy barriers between configurations near the singularities and configurations farther from the singularities. Our results are consistent with a closely related work [68, 77] on singular colloidal clusters, but allow for isolated singularities of the shape space. We demonstrate our results using both the four-bar linkage as well as a flat, triangulated origami [18]. In addition to these two systems with one-dimensional configuration spaces, we also demonstrate the versatility of our methods using the five-bar linkage, which is a framework with a two-dimensional configuration space. The analyses presented in these chapters have direct consequences in the design and employment of nanoscale frameworks in applications ranging from self-assembly [98] to drug delivery [180], where relative thermodynamic stability of different configurations is of paramount importance.

## 1.2 Thin structures, elastic waves, and bound states

In the second part of this dissertation, we consider the trapping of elastic waves in thin elastic structures with spatially varying curvature profiles [108]. This problem was partly motivated by a fascinating acoustic phenomenon.

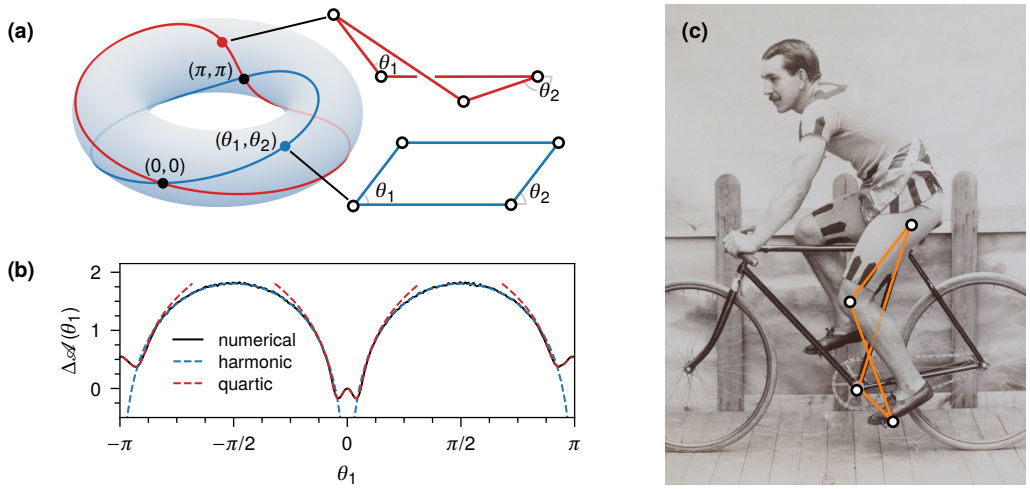


**Figure 1.3.** (a) The shape space of a four-bar linkage visualized as two intersecting curves on a torus. (b) Its free energy $\mathscr{A}(\theta_1)$ as a function of the angle $\theta_1$. (c) Four-bar linkage in the real world; photograph by J. Beau, *Photographie Sportive* (1898).

### 1.2.1 Can one hear the shape of a shell?

Take an ordinary hand saw used for cutting wood. Clamp down the handle end of the saw using your feet and bend its tip using your dominant hand so that its overall shape is similar to the letter S (see Fig. 1.4). The saw is now a "singing" saw: bowing it or striking it with a mallet produces a sharp, sustained sound. But how does the saw sing, and more importantly, why does it have to be shaped like an S? Although the sonorousness of the common saw has been known at least since the 19th century [160], the first scientific explanation for it was provided only in the early 90s by Scott and Woodhouse [150]. These authors modeled the saw as a thin elastic shell with a varying curvature profile and showed that the saw's sonority is due to vibrations that get trapped around its inflection point, where the local curvature vanishes. As trapped vibrations remain confined to the immediate vicinity of the inflection point, it reduces energy dissipation through the two ends of the saw, resulting in a sustained note. In a very recent work, Shankar *et al.* [151] reported that these vibrations have a topological origin on the basis of jumps in an integer-valued invariant as we move across the inflection point.

Despite the efforts of the above-mentioned authors, several critical questions remain unanswered: What kinds of waves on thin structures get trapped? Do we really need an inflection point to observe trapped waves? Is there a difference between the vibrational spectrum of one-dimensional and two-dimensional thin structures, i.e., rods vs. shells? Is it possible to compute, even approximately, the shapes and frequencies of the vibrational modes? In this dissertation, we explore these questions using a thin shell and a rod as concrete examples.

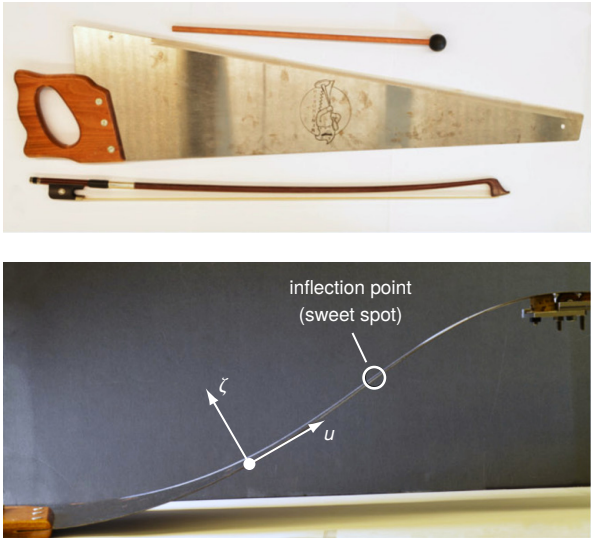


**Figure 1.4.** An ordinary hand saw, when bent into the shape of the letter S can be played like a musical instrument using a violin's bow or a mallet. A sustained note is produced on bowing or hitting the saw close to its inflection point, which is called a sweet spot by saw enthusiasts. In most models of thin elastic structures, such as the saw here, the displacement field is broken up into a normal (e.g., $\zeta$) and tangential (e.g., $u$) component, which remain coupled if the structure is curved. Photographs sourced from Ref. [151].

Central to understanding the trapping of waves in elastodynamic systems, such as the musical saw, is a commonplace eigenvalue problem of the form

$$\widehat{\mathsf{D}}\psi = \omega^2\psi. \tag{1.3}$$

Here, $\widehat{\mathsf{D}}$ is a self-adjoint linear differential operator, $\psi$ is the wave field, and $\omega$ is the frequency of vibration. A trapped wave is a solution $\psi$ to Eq. (1.3) satisfying the prescribed boundary conditions and decaying exponentially as we approach the physical boundaries. In physics, particularly in quantum-mechanical contexts, such solutions are called localized or bound states [143–145]—vocabulary that we will continue to use.

In elastodynamic problems involving thin structures, the wave field $\psi$ is almost always composed of displacements—broken up into tangential and normal components—from the neutral, undeformed configuration of the structure. The operator $\widehat{\mathsf{D}}$, the exact nature of which is model-dependent, is a square matrix of spatial derivatives that act on $\psi$. An uncurved elastic structure can support three basic wave types: extensional and shear waves, which respectively propagate by stretching and shearing the structure, and involve only the tangential components (e.g., $u$ in Fig. 1.4); and flexural waves that propagate by bending the structure and involve only the normal component (e.g., $\zeta$ in Fig. 1.4). By contrast, if the structure is curved, the normal and tangential components remain coupled. In this case, we

can only speak of waves that are predominantly flexural or extensional or shear-like, based on which component is more dominant in magnitude. It this curvature-mediated coupling between the various components that ultimately leads to the formation of bound states in curved structures.

### 1.2.2 Semiclassical approximation

The usual line of attack, when faced with an equation similar to Eq. (1.3), is to look for plane-wave solutions of the form $\psi \sim e^{\pm ikx}$. Such an endeavor, however, fails if the coefficients of the derivatives in $\widehat{\mathsf{D}}$ are not constants, which is the case for a thin structure with a varying curvature profile.[3] The semiclassical approximation or the Wentzel–Kramers–Brillouin (WKB) approximation is a widely used method to obtain approximate solutions to differential equations where these coefficients are slowly varying. In asymptotic analysis, it is usually introduced as an approximate method to find solutions to differential equations whose highest derivative is multiplied by a small parameter [6]. A representative example from physics is the time-independent Schrödinger equation for a particle of mass $m$ in a potential $V(x)$,

$$\left[-\frac{\hbar^2}{2m}\partial_x^2 + V(x)\right]\psi(x) = E\psi(x), \tag{1.4}$$

where the WKB method is frequently employed to find asymptotic solutions in the limit the (reduced) Planck's constant $\hbar \to 0$. At first glance, such a limit is perplexing and makes *no* physical sense as $\hbar$ is a fundamental constant, whose value is fixed by the units we choose to work in. In reality, the limit $\hbar \to 0$ represents the situation where the value of $\hbar$ is much smaller compared to the angular momentum scale, which is often the case with macroscopic systems described by classical physics. It is in this context that the WKB method gets the alternative moniker of semiclassical approximation.

Modern reformulations of the semiclassical method, based on the Weyl symbol calculus, provide unparalleled insights into the connection between classical and quantum mechanics. Weyl calculus provides an elegant way to set up a one-to-one association between differential operators (e.g., powers of $\partial_x$) defined on Hilbert spaces and ordinary functions defined on a position-momentum phase space. For instance, under the Weyl transform, the momentum operator $\hat{k} = -i\partial_x$ gets mapped to the momentum $k$. Likewise, the Schrödinger operator $\hat{H} = -\partial_x^2/2m + V(x)$, gets mapped to the classical Hamiltonian of a point particle $H(x,k) = k^2/2m + V(x)$. (For simplicity, we have suppressed factors of $\hbar$.) To make use of Weyl calculus, it is customary to express the derivatives of a given differential operator $D(x,\partial_x)$ in terms of $\hat{k}$ and its powers. Then, using an integral transform, the operator can be mapped to an ordinary function—called its Weyl symbol—defined on an $x$-$k$ phase space, i.e.,

$$\widehat{D}(x,\partial_x) = \widehat{D}\left(x, i\hat{k}\right) \quad \text{(Hilbert space)} \quad \rightleftharpoons \quad D(x,k) \quad \text{(phase space)}. \tag{1.5}$$

Once the operator is in its symbol form, it can be expanded and approximated just like any other ordinary function, providing a direct path towards obtaining asymptotic solutions to wave equations.

[3] To be clear, as the exact form of the operator $\widehat{\mathsf{D}}$ depends on the shell or rod model we choose to work with, even when the curvature is a constant, a plain-wave solution may not be applicable.

For matrix operators that act on multicomponent wave fields, such as the one in Eq. (1.3), the corresponding Weyl symbol $\mathsf{D}$ is called the dispersion matrix—an ordinary matrix, whose entries are functions defined on an $x$-$k$ phase space. The eigenvalues $\lambda(x,k)$ of the dispersion matrix serve as ray Hamiltonians of the different wave types represented by Eq. (1.3). This leads us to the phase space representation of waves as rays that satisfy the Hamilton's equations

$$\dot{x} = \partial_k \lambda(x,k) \quad \text{and} \quad \dot{k} = -\partial_x \lambda(x,k). \tag{1.6}$$

The advantage provided by the semiclassical approximation cannot be overstated: we have effectively reduced the wave equation to a Hamiltonian system describing point particles, a system that is much more easier to analyze.

### 1.2.3 Bound states in thin elastic structures

The semiclassical approximation also provides a direct route to extract the bound-state frequencies by "quantizing" the ray trajectories in the phase space. As bound waves are bounded in the phase space as well, the rays corresponding to these waves appear in the form of closed orbits when visualized [see Fig. 1.5]. Each bound orbit is associated with a wave of a specific frequency $\omega$. The single valuedness of the wave field as we traverse the orbit requires the overall phase to be a half-integral multiple of $2\pi$, which leads to the famous Bohr–Sommerfeld quantization rule. Additional complications and subtleties arise when the wave field has more than one component, which leads to a corrected quantization condition of the form

$$\oint \mathrm{d}x\, k(x) = 2\left(n + \tfrac{1}{2}\right)\pi - \left(\gamma_{\mathrm{G}} + \gamma_{\mathrm{NG}}\right). \tag{1.7}$$

Here $\gamma_{\mathrm{G}}$ and $\gamma_{\mathrm{NG}}$ are extra phases that arise because of the multicomponent nature of the problem. The phase $\gamma_{\mathrm{G}}$ is a geometric phase [10, 130], which was very recently shown to be responsible for the topological protection of equatorial oceanic waves on the Earth's surface [165]. That said, for the two example systems we study in this work, both the extra phases $\gamma_{\mathrm{G}}$ and $\gamma_{\mathrm{NG}}$ vanish. Although one might expect this observation to change based on the equations we work with, it seems to be a generic result, which we expect to hold in other models of curved structures as well.

Because the extra phases vanish, the analysis becomes less cumbersome, and ultimately leads to impressive agreement between the bound-state frequencies obtained through quantization and those seen in numerical experiments. The basic result from a combined semiclassical and numerical analysis is as follows: for the variably curved rod, which supports only extensional and flexural waves, only extensional waves form bound states; and in the case of the shell, all three wave types, i.e., flexural, extensional, and shear, exhibit bound states. For both the shell and the rod, the bound states develop around points where the absolute curvature has a minimum. The localization of flexural waves around the inflection point of a musical saw, which is what both Scott and Woodhouse [150] and Shankar *et al.* [151] restricted their attention to, is a special case of this more generic observation. Many practical applications of elastic waves, from acoustic cloaking to negative refraction, rely on delicate aspects of wave localization. The vast majority of these applications, however, require designer materials with intricate microstructures. Being able to induce localization by a mere

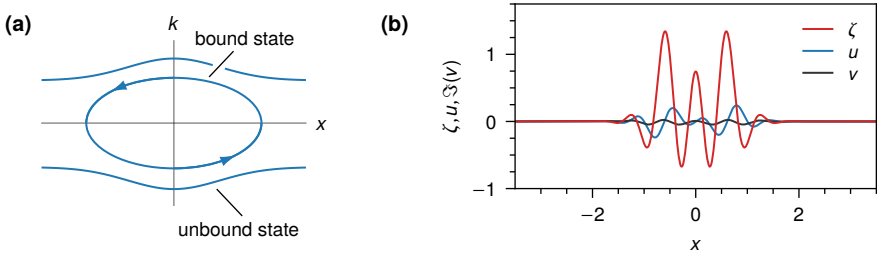


**Figure 1.5.** (a) Phase-space representation of waves as rays, showing bound and unbound waves. (b) A predominantly flexural bound state in a uniaxially curved shell with a curvature minimum at $x = 0$ obtained by solving Eq. (1.3) numerically. Here $x$ is an arc-length coordinate, $\zeta$ and $u, v$ are the normal and tangential components of the displacement field, respectively; see Fig. 1.4 and Chapter 5 for more details.

change of the material geometry is clearly advantageous. For this reason, the results presented here lay the groundwork for the design of even simpler instruments capable of inducing wave localization.

## 1.3 Organizational summary

In Chapter 2 we review basic facts about frameworks, their configuration spaces, and derive asymptotic expressions for their elastic energies. Chapter 3 deals with the effect that thermal fluctuations have on these frameworks, and presents a detailed calculation of the free-energy profiles for three example frameworks. Chapters 4 and 5 are concerned with the second problem discussed in this dissertation. The semiclassical approximation is reviewed in Chapter 4, with special attention paid to multicomponent wave equations. Chapter 5 makes use of the semiclassical approximation to understand the formation of bound states in thin elastic structures.

Chapter 2

# Constraints, frameworks, and singularities*

This chapter provides background material for the first part of this dissertation. First, we discuss holonomic constraints and how they can sculpt the configuration space of a classical mechanical system in a nontrivial manner. Next, we extend the general discussion on constraints to look at the zero-temperature behavior of mechanical frameworks and their configuration spaces. Finally, we derive two expressions for the energy of a mechanical framework to be used in the next chapter.

In classical mechanics, the concept of a configuration space makes it possible for us to abstract away a physical system as a single point, no matter how complicated the system is (see Fig. 2.1). The elegance behind the notion of a configuration space can be seen by considering the simple example of a rigid body, composed of an infinite number of particles. Yet, its configuration space is just a six-dimensional space whose points have a one-to-one correspondence with the location and the spatial orientation of the rigid body in Euclidean 3-space. In many example systems, such as the rigid body, the nature of a system's configuration space is entirely dictated by the (holonomic) constraints imposed on it, so we discuss them first.

## 2.1 Geometry of constraints

Consider a system of $N$ classical particles in $d$ dimensions. If the position of the $i$th particle is given by the position vector $\boldsymbol{r}_i \in \mathbb{R}^d$, then the entire configuration of the system can be fully described at any moment using a configuration vector $\boldsymbol{r} \in \mathscr{R} \subseteq \mathbb{R}^n$ defined by $\boldsymbol{r} = (\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$, with $n = Nd$. We shall call the set $\mathscr{R}$ as the ambient space.[1] Let us now impose $m \leq n$ holonomic constraints on the coordinates $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N$ by defining smooth functions $f_i : \mathscr{R} \subseteq \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, m$ that vanish when the constraints are satisfied. As an example, consider a system composed of a single particle with the coordinate $\boldsymbol{r} = (r_1, r_2, r_3)$

*This chapter and the next are expanded versions of M. Mannattil, J. M. Schwarz, C. D. Santangelo, Phys. Rev. Lett. **128**, 208005 (2022). The problem discussed in this paper emerged during conversations with my coauthors. I was responsible for all the analytical calculations and numerical experiments, and wrote the paper with inputs from my coauthors.

[1] Many authors [94, 102] call $\mathscr{R}$ as the configuration space. However, when constraints are imposed on the system, the actual configuration space is usually a lower-dimensional subset of $\mathscr{R}$. To make this distinction, we will call $\mathscr{R}$ as the ambient space.

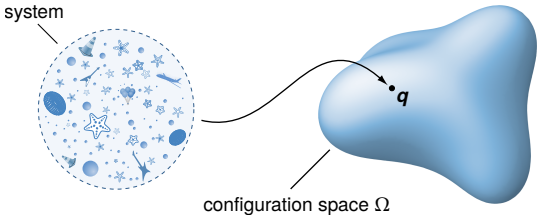


**Figure 2.1.** A classical system of point particles and rigid bodies can be represented by a single configuration vector $\boldsymbol{q}$ of a high-dimensional configuration space $\Omega$, which may or may not be a smooth manifold. (Inspired by Fig. 20.1 of Ref. [133].)

in $\mathbb{R}^3$. In this case $n=3$ and the ambient space is equivalent to the physical space the particle moves in, i.e., $\mathscr{R}=\mathbb{R}^3$. Now, consider the constraints

$$\begin{aligned} f_1(r_1,r_2,r_3) &= r_2^2+r_3^2-a^2, \\ f_2(r_1,r_2,r_3) &= r_1^2+r_3^2-b^2, \end{aligned} \tag{2.1}$$

where $a,b>0$ are constants. Geometrically, this corresponds to constraining a particle to move on two mutually perpendicular cylinders in $\mathscr{R}$.

For the general case, we will first look at the constraints individually. Each constraint function $f_i$ has an associated zero level set $\Omega_i = f_i^{-1}(0) = \{\boldsymbol{r}\in\mathscr{R} : f_i(\boldsymbol{r})=0\}$ whose elements satisfy the $i$th constraint exactly. If for all points $\boldsymbol{r}\in\Omega_i$, the gradient $\nabla f_i(\boldsymbol{r})$ is nonzero, then by the preimage theorem, $\Omega_i$ is an $(n-1)$-dimensional submanifold of the ambient space $\mathscr{R}$ [58, 92]. For the example system in Eq. (2.1), the level set $\Omega_1$ corresponding to the first constraint function $f_1$ is a cylinder of radius $a$ lying along the $r_1$ axis. Meanwhile, the level set $\Omega_2$ is a cylinder of radius $b$ lying along the $r_2$ axis. Although these constraints may seem somewhat abstract, they are in fact the ones seen in a three-slider mechanism (Fig. 2.2). Since the gradients $\nabla f_1$ and $\nabla f_2$ never vanish for any point in $\Omega_1$ and $\Omega_2$, they are both smooth submanifolds of $\mathscr{R}$. At each point $\boldsymbol{r}\in\Omega_i$ we can identify two vector spaces, namely the tangent space $T_{\boldsymbol{r}}\Omega_i = \{\boldsymbol{v}\in\mathbb{R}^n : \nabla f_i(\boldsymbol{r})\cdot\boldsymbol{v}=0\}$ and the normal space[2] $N_{\boldsymbol{r}}\Omega_i = (T_{\boldsymbol{r}}\Omega_i)^\perp$, the orthogonal complement of the tangent space in $\mathbb{R}^n$ under the standard Euclidean dot product. From these definitions, we see that the direct sum $T_{\boldsymbol{r}}\Omega_i \oplus N_{\boldsymbol{r}}\Omega_i = \mathbb{R}^n$.

The $m$ constraint functions $f_i$ can also be considered together as a single constraint map $f:\mathscr{R}\subseteq\mathbb{R}^n\to\mathbb{R}^m$ with $f(\boldsymbol{r}) = [f_1(\boldsymbol{r}), f_2(\boldsymbol{r}),\ldots,f_m(\boldsymbol{r})]$. The zero level set $\Omega$ of this map is a subset of the ambient space where all $m$ constraints functions $f_i$ vanish, i.e., it is formed by the intersection of the zero level sets corresponding to each constraint: $\Omega=\Omega_1\cap\Omega_2\cap\cdots\cap\Omega_m$. The constraint map $f$ has an $m\times n$ Jacobian matrix $\mathsf{J}=\nabla f$ and since $m\le n$, it has a maximum possible rank of $m$. A point $\bar{\boldsymbol{r}}$ on $\Omega$ is called *regular* if the Jacobian $\mathsf{J}$ has maximum rank at that point. As we have used here, from here on, an overbar on a point denotes that it belongs to $\Omega$. In the case of maps, the preimage theorem [92] states that if each $\bar{\boldsymbol{r}}\in\Omega$ is regular, then

[2] Here it is the rather trivial one-dimensional space spanned by $\{\nabla f_i\}$.

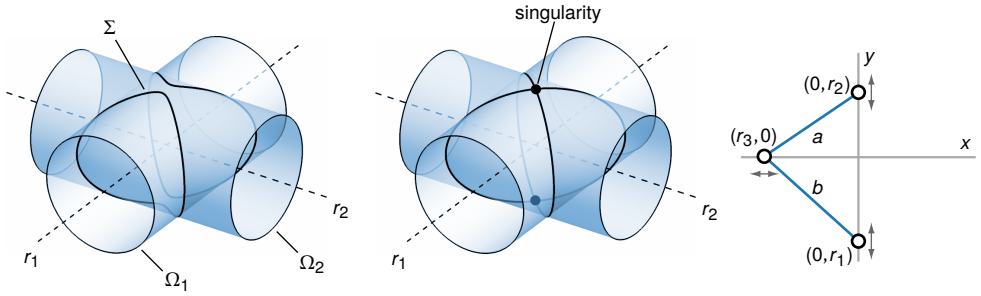


**Figure 2.2.** The intersection curves of two cylinders with mutually perpendicular axes can be (a) curves that are smooth manifolds; or (b) a curve with a self-intersection, i.e., a singularity. (c) A physical realization of the constraints in Eq. (2.1) is the three-slider mechanism [11], which is composed of three particles (the "sliders") that can freely move on two mutually perpendicular axes (the "rails"). The particles are connected by rigid bars of length $a$ and $b$, which gives rise to the constraints in Eq. (2.1).

$\Omega$ is a smooth submanifold of $\mathscr{R}$ of dimension $n-m$ and codimension $m$.[3] Since $\Omega$ is the set of points where all the constraints are obeyed, it forms the configuration space of the system.

Given a configuration space $\Omega$, we identify the tangent space $T_{\bar{\boldsymbol{r}}}\Omega = \ker \mathsf{J}(\bar{\boldsymbol{r}}) = \{\boldsymbol{v} \in \mathbb{R}^n : \mathsf{J}(\bar{\boldsymbol{r}})\boldsymbol{v} = 0\}$ and the normal space $N_{\bar{\boldsymbol{r}}}\Omega = (T_{\bar{\boldsymbol{r}}}\Omega)^{\perp}$, the orthogonal complement of the tangent space in $\mathbb{R}^n$ under the standard Euclidean dot product. Using the rank-nullity theorem, we see that $\dim T_{\bar{\boldsymbol{r}}}\Omega = n-m$. As $\Omega$ is an $(n-m)$-dimensional submanifold of $\mathbb{R}^n$, it is theoretically possible to find a smooth parameterization $\psi$ such that all points $\bar{\boldsymbol{r}} \in \Omega$ can be written as $\bar{\boldsymbol{r}} = \psi(\xi)$, where $\xi = (\xi_1, \xi_2, \ldots, \xi_{n-m}) \in \mathbb{R}^{n-m}$. We also remark that one generally needs multiple such parameterizations to cover all points in $\Omega$. The collective coordinates $\xi_i$ used to parameterize $\Omega$ could be any set of coordinates, although it is more convenient to choose coordinates physically relevant to the system.

It is also useful to analyze the intersection of the constraint level sets, say $\Omega_i$ and $\Omega_j$, locally at a point $\bar{\boldsymbol{r}} \in \Omega$. If the Jacobian $\mathsf{J}$ is full rank, then its rows, i.e., the gradients $\nabla f_i$ of the constraint functions $f_i$, are independent. Consequently, the tangent spaces $T_{\bar{\boldsymbol{r}}}\Omega_i$ and $T_{\bar{\boldsymbol{r}}}\Omega_j$—defined as sets of vectors perpendicular to $\nabla f_i$ and $\nabla f_j$—are different subspaces. Since both tangent spaces are $(n-1)$-dimensional subspaces of $\mathbb{R}^n$, the fact they are different implies (by dimension counting) $T_{\bar{\boldsymbol{r}}}\Omega_i + T_{\bar{\boldsymbol{r}}}\Omega_j = \mathbb{R}^n$. In other words, we say that the manifolds $\Omega_i$ and $\Omega_j$ are transverse to each other [58]. Hence, the Jacobian being full rank ensures that the level sets $\Omega_i$ of *all* constraints are pairwise transverse.

Going back to our example, the configuration space $\Omega$ in Fig. 2.2 is composed of the curves obtained by the intersection of the cylinders $\Omega_1$ and $\Omega_2$. When the cylinder radii satisfy $a \neq b$, then $\Omega$ is made up of two disconnected closed curves as illustrated in Fig. 2.2(a). Since $n = 3$ and $m = 2$, the configuration space $\Omega$ has a dimension equal to $n - m = 1$, consistent with

[3] This is equivalent to the familiar notion that imposing $m$ well-behaved constraints on a system with $n$ degrees of freedom reduces its number of degrees of freedom to $n-m$.

the fact that the curves are one-dimensional. The tangent spaces $T_{\bar{\boldsymbol{r}}}\Omega_1$ and $T_{\bar{\boldsymbol{r}}}\Omega_2$ for a point $\bar{\boldsymbol{r}} \in \Omega$ can be identified with the tangent planes of the two cylinders at that point. It is clear that there is no point in $\Omega$ where the two cylinders share a common tangent plane and it can be shown (by direct computation) that the Jacobian of the constraint map never drops rank. Hence, the intersection shown in Fig. 2.2(a) is indeed a transverse intersection.

More generally, there could be points $\bar{\boldsymbol{r}} \in \Omega$ where the Jacobian $J(\bar{\boldsymbol{r}})$ is not full rank. Such points are called *singular points*. If the Jacobian is rank deficient, at least two of its rows, say $\nabla f_i$ and $\nabla f_j$, are not independent and thus $T_{\bar{\boldsymbol{r}}}\Omega_i = T_{\bar{\boldsymbol{r}}}\Omega_j$. By the arguments in the preceding paragraphs, this would also imply that the corresponding level sets $\Omega_i$ and $\Omega_j$ cannot be mutually transverse since $T_{\bar{\boldsymbol{r}}}\Omega_i + T_{\bar{\boldsymbol{r}}}\Omega_j \neq \mathbb{R}^n$. Such a situation can be seen in Fig. 2.2(b), where the two cylinders have the same radius, in which case $\Omega$ reduces to a self-intersecting curve.[4] It is clear that the two cylinders share a common tangent plane at the point of self-intersection. Since the tangent spaces are not mutually independent at this point, their sum cannot possibly be equal to $\mathbb{R}^3$.

To summarize, when a set of constraints is imposed on a mechanical system, its configuration space can become nontrivial. Singularities may occur in the configuration space at points where the Jacobian of the constraint map drops rank. This means that, at these singular points, some of the constraints are equivalent to linear order. Geometrically, this corresponds to a nontransversal intersection of the zero level sets of the constraint functions at that point.

## 2.2 Frameworks

Let us now turn our attention to bar-joint frameworks made of $N \geq 3$ point-like joints in $d$ dimensions connected by $m < Nd - \frac{1}{2}d(d+1)$ freely rotating, massless bars. If the joints have position vectors $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N \in \mathbb{R}^d$ in the lab frame, the framework's configuration can be fully described at any given moment using a configuration vector $\boldsymbol{r} \in \mathbb{R}^{Nd}$ defined by $\boldsymbol{r} = (\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$. Let the $i$th bar in the framework have a length $\ell_i$. It is clear that the length $\ell(\boldsymbol{r}_i)$ can always be considered as a function of the configuration vector $\boldsymbol{r}$ describing the entire framework.

As before, the configuration space $\Omega$ of the framework is the set of configurations such that the length of the $i$th bar $\ell_i$ is equal to its rest length $\bar{\ell}_i$. Configurations in $\Omega$ are deformed states of the framework or the configurations obtained from such states after a rigid motion in $d$-dimensional space (i.e., after rotations, reflections, or translations). To separate the actual deformations of the framework from trivial rigid motions, it is more useful to consider the shape space $\Sigma$ [86, 87, 119], which is the set of all configurations with distinct shapes, and obtained after "filtering out" all the rigid motions from the configuration space.[5]

To practically identify $\Sigma$, we first switch to a Cartesian body frame attached to the framework so that all $\frac{1}{2}d(d+1)$ rigid motions are eliminated [38, 66]. We require $n = Nd - \frac{1}{2}d(d+1)$

[4] An analogous situation arises while considering the intersection of a sphere and a cylinder. If the sphere radius is equal to the cylinder diameter, the intersection curve has a double point, and is called Viviani's curve [159].

[5] More technically, the shape space is the quotient space $\Sigma = \Omega/SE(d)$, where $SE(d)$ is the special Euclidean group in $d$ dimensions [101].

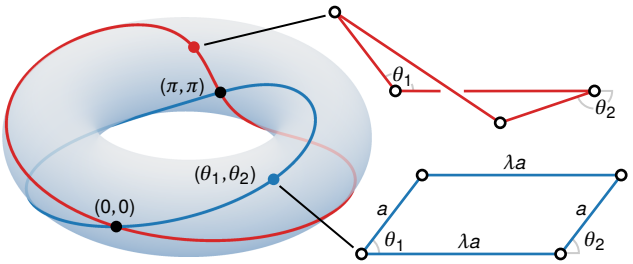


**Figure 2.3.** Shape space of the planar four-bar linkage visualized as two intersecting curves on a torus, each curve representing a "branch" of the shape space. The poloidal and toroidal angles along the branches correspond to the angles $\theta_1$ and $\theta_2$ of the linkage, which has two modes of deformation with $\theta_1 = \theta_2$ (blue curve) and $\theta_1 \neq \theta_2$ (red curve).

coordinates to specify the state of the framework in the body frame and let $\boldsymbol{q} \in \mathbb{R}^n$ be its configuration vector in this frame. Now consider $m$ holonomic constraint functions $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, m$, each associated with a single bar, and defined by

$$f_i(\boldsymbol{q}) = \frac{\ell_i^2(\boldsymbol{q}) - \bar{\ell}_i^2}{2\bar{\ell}_i}, \tag{2.2}$$

where the factor of $1/(2\bar{\ell}_i)$ is for later convenience. As before, the $m$ scalar constraint functions can also be considered together as a single constraint map $f : \mathbb{R}^n \to \mathbb{R}^m$ defined by $f(\boldsymbol{q}) = [f_1(\boldsymbol{q}), f_2(\boldsymbol{q}), \ldots, f_m(\boldsymbol{q})]$. Then, the shape space is the zero level set $\Sigma = \{\boldsymbol{q} \in \mathbb{R}^n : f(\boldsymbol{q}) = 0\}$. In the absence of external forces, each point in $\Sigma$ is a ground-state configuration of the framework with a distinct shape.

The compatibility matrix $\mathsf{C}(\bar{\boldsymbol{q}})$ [106, 132] at a configuration $\bar{\boldsymbol{q}} \in \Sigma$ is the $m \times n$ Jacobian matrix $\nabla f$ of the constraint map $f$.[6] If $\mathsf{C}$ has full rank for all points in $\Sigma$, then $\Sigma$ is an $(n-m)$-dimensional submanifold of $\mathbb{R}^n$ [92, 93]. But frameworks, like other holonomically constrained systems, can have shape spaces that are not smooth manifolds. Consider, for example, the four-bar linkage depicted in Fig. 2.3. The four-bar linkage is a simple framework made out of four bars with freely rotating joints [55, 60, 153]. Although the linkage has one DOF up to Euclidean motions, it has two modes of deformation: one where the angle $\theta_1 = \theta_2$ and another where $\theta_1 \neq \theta_2$. To visualize these two modes of deformation it is most natural to plot the angles $\theta_1$ and $\theta_2$ on a torus as in Fig. 2.3. From this figure we see that the deformation modes appear as two curves (branches) meeting at two isolated singular points $(\theta_1, \theta_2) = (0,0)$ and $(\pi, \pi)$. The two curves put together form the shape space $\Sigma$ of the four-bar linkage.

When the shape space of a framework has a "branched" structure, like in the case of the four-bar linkage, it is no longer a smooth manifold. In such a situation, the compatibility matrix $\mathsf{C}(\boldsymbol{q})$ drops rank at the singularity where the branches meet [104, 123], and the

[6] In Eq. (2.2) it might have been more natural to use the constraint function $f_i = \frac{1}{2}[\ell_i^2(\boldsymbol{q}) - \bar{\ell}_i^2]$, without the $\bar{\ell}_i^{-1}$ factor. This would have made $\nabla f$ the rigidity matrix instead of the compatibility matrix. For our purposes, however, it is more convenient to work with the compatibility matrix as its entries are dimensionless.

constraints cease to be linearly independent. Such singularities are the most common singularities [104, 105] found in a framework and here we consider the situation where they occur only at isolated points of $\Sigma$.[7] The branches of $\Sigma$, being $(n-m)$-dimensional submanifolds of $\mathbb{R}^n$, can be individually parameterized using a set of coordinates $\xi \in \mathbb{R}^{n-m}$, analogous to generalized coordinates, called shape coordinates [101] as they capture the shape changes of the framework as it moves on $\Sigma$. We also assume that $n$ is small enough that such parameterizations can be found without much difficulty and that the branches are linearly independent at the singularity [104]. Before moving on, we need to familiarize ourselves with a few other concepts that originate in rigidity theory and structural mechanics [22, 171].

### 2.2.1 Zero modes and self stresses

Irrespective of whether or not $\mathsf{C}$ is full rank, we define a *zero mode*[8] to be a perturbation $\boldsymbol{u} \in \mathbb{R}^n$ that preserves the constraints to linear order, i.e., $\mathsf{C}\boldsymbol{u} = 0$ or $\boldsymbol{u} \in \ker \mathsf{C}$. If $\mathsf{C}$ is full rank, then there is a well-defined tangent space $T_{\bar{\boldsymbol{q}}}\Sigma$ at $\bar{\boldsymbol{q}}$ and $\boldsymbol{u} \in T_{\bar{\boldsymbol{q}}}\Sigma$. In such a situation, the number $z$ of zero modes is

$$z = n - m\,, \tag{2.3}$$

which is equal to the dimension of tangent space.

However, as we have remarked, the compatibility matrix $\mathsf{C}$ drops rank at shape-space singularities. If the compatibility matrix is not full rank, then it implies that its rows are not linearly independent and there are $m$-dimensional nonzero vectors $\boldsymbol{\sigma} \in \mathbb{R}^m$ such that

$$\boldsymbol{\sigma}^{\mathsf{T}}\mathsf{C}(\bar{\boldsymbol{q}}) = 0\,. \tag{2.4}$$

Any vector $\boldsymbol{\sigma}$ that belongs to the left null space of the compatibility matrix has the physical interpretation of having components equal to the signed magnitude of spring tensions (up to normalization factors) that enforce static equilibrium of the framework. For this reason, $\boldsymbol{\sigma}$ is called a self stress. Forces involved in a state of self stress obey the strong form of Newton's third law and thus cannot possibly result in an unbalanced torque [51]. Thus, a framework under self stress is in a state of mechanical equilibrium. Note, however, that such an equilibrium may or may not be stable.

Applying the rank-nullity theorem to the transpose of Eq. (2.4) we find the number $s$ of (independent) self stresses to be

$$s = m - \operatorname{rank}\mathsf{C}^{\mathsf{T}}(\bar{\boldsymbol{q}}) = m - \operatorname{rank}\mathsf{C}(\bar{\boldsymbol{q}})\,. \tag{2.5}$$

If the compatibility matrix $\mathsf{C}$ is not full rank at a point $\bar{\boldsymbol{q}}$, there is no well-defined tangent space $T_{\bar{\boldsymbol{q}}}\Sigma$ at that point. However, one can still count the number $z$ of zero modes by applying the rank-nullity theorem on $\mathsf{C}$, which gives

$$z = n - \operatorname{rank}\mathsf{C}(\bar{\boldsymbol{q}})\,. \tag{2.6}$$

[7] In engineering literature they are usually called $c$-space singularities. For other, less common singularities that can occur in a framework, see Refs. [104, 122, 123], and references therein.

[8] Also called a soft mode, flex, etc. Sometimes it is also customary to explicitly differentiate between linear zero modes (those that preserve the constraints only to linear order) and nonlinear zero modes (those that preserve constraints to higher orders).

Eliminating rank$\mathsf{C}(\bar{\boldsymbol{q}})$ from Eqs. (2.5) and (2.6) we arrive at the Maxwell–Calladine [16, 115] relation

$$z - s = n - m, \tag{2.7}$$

which shows that the difference in the number of zero modes and the number of self stresses at a point $\bar{\boldsymbol{q}} \in \Sigma$ depends only on the number of DOF $n$ (i.e., the dimension of the ambient space) and the number of constraints $m$. Since we have assumed that $m \leq n$, we see that the presence of a self stress creates new zero modes. To illustrate these ideas, we shall now look at a simple example.[9]

**Example (A tethered particle in $\mathbb{R}^2$).** Consider a particle with position $\boldsymbol{q} = (q_1, q_2)$ tethered to two parallel walls using two bars as shown in Fig. 2.4. We will analyze the setup in Fig. 2.4(b) first and examine Fig. 2.4(a) as a special case. This framework clearly has zero DOF. And as the rigid motions have all been eliminated by tethering the particle to the walls, there is no need to transform to a body frame or distinguish between the shape space and the configuration space. In Fig. 2.4(b), the bars have a length of 1 unit and the inter-wall separation is $2(1-\epsilon)$ units. So following Eq. (2.2), we can write down the constraint map as

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad f(q_1, q_2) = \tfrac{1}{2}\left[(q_1 + 1 - \epsilon)^2 + q_2^2 - 1, (q_1 - 1 + \epsilon)^2 + q_2^2 - 1\right]. \tag{2.8}$$

When the parameter $\epsilon$, which controls the inter-wall separation, satisfies $0 < \epsilon < 1$, the configuration space $\Sigma$ is composed of two ground states, given by $\bar{\boldsymbol{q}} = (0, \pm\sqrt{2\epsilon - \epsilon^2})$. The compatibility matrices at the ground states are

$$\mathsf{C}(\bar{\boldsymbol{q}}) = \nabla f = \begin{pmatrix} \bar{q}_1 + 1 - \epsilon & \bar{q}_2 \\ \bar{q}_1 - 1 + \epsilon & \bar{q}_2 \end{pmatrix} = \begin{pmatrix} 1 - \epsilon & \pm\sqrt{2\epsilon - \epsilon^2} \\ -1 + \epsilon & \pm\sqrt{2\epsilon - \epsilon^2} \end{pmatrix}. \tag{2.9}$$

As we have assumed that $0 < \epsilon < 1$, the compatibility matrix has full rank at both ground states. Consequently, the framework does not support a state of self stress or have any zero modes. Clearly, in this case the configuration space is a zero-dimensional manifold. The framework failing to support a self stress is hardly surprising—if we were to put tensions in the bars, there would be a net vertical force that cannot be balanced. We can also analyze the situation geometrically. In Fig. 2.4(b), we have depicted the level sets $\Sigma_1$ and $\Sigma_2$ (circles of unit radius) of the two constraints individually. As we see from this figure, the level sets $\Sigma_1$ and $\Sigma_2$ have a transversal intersection at the ground state. In other words, the two constraints are linearly independent at the ground state, resulting in a compatibility matrix of full rank.

If the parameter $\epsilon = 0$, then there is only one ground state, i.e., $\bar{\boldsymbol{q}} = (0, 0)$. In this case the compatibility matrix becomes

$$\mathsf{C}(\bar{\boldsymbol{q}}) = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}. \tag{2.10}$$

Clearly, the above matrix does not have full rank and we see that there is a self stress $\boldsymbol{\sigma} \in \ker \mathsf{C}^{\mathsf{T}}$ given by $\boldsymbol{\sigma} = (1, 1)$. The physical interpretation of this self stress is that the framework will remain in equilibrium if we put tensions of equal magnitude in the two bars. Also, the constraints in this case are not linearly independent at the singularity—geometrically, we infer this by noticing that the level sets $\Sigma_1$ and $\Sigma_2$ in Fig. 2.4(a) have a nontransversal intersection and share a common tangent at the ground state. As the number of self stresses $s = 1$, the Maxwell–Calladine theorem tells us that there is a zero mode.

[9] Similar examples have been considered in Refs. [16, 96, 172, 179].

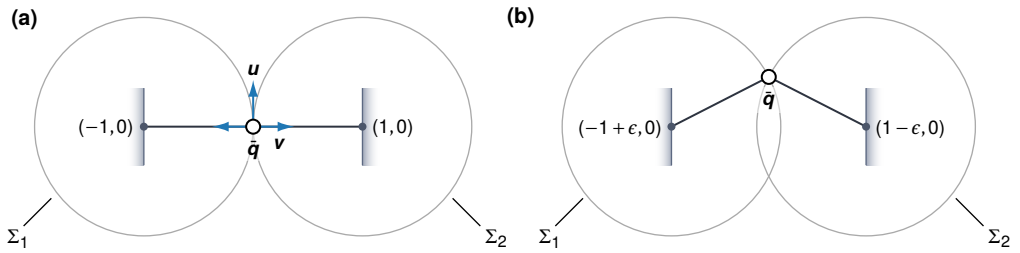


**Figure 2.4.** (a) Ground-state configuration of a particle with coordinate $\boldsymbol{q}$ tethered to two parallel walls using two bars—a simple framework that can support a state of self stress. Each bar has a length of 1 unit and the inter-wall separation is 2 units. The zero level sets $\Sigma_1$ and $\Sigma_2$ of the two constraint functions have a nontransversal intersection as they share a common tangent at the point of intersection. (b) When the wall separation is less than 2 units, the level sets $\Sigma_1$ and $\Sigma_2$ have a transversal intersection and the framework fails to support a state of self stress.

From $\ker \mathsf{C}(\bar{\boldsymbol{q}})$ we see that the zero mode $\boldsymbol{u} = (0,1)$ is the up/down displacement of the particle, which preserves the constraints to linear order.[10]

## 2.3 Elastic energy of a framework

In practice, there is no such thing as a framework with perfect constraints, and it is always possible to violate them by paying some sort of an energy cost. Multihelix bundles used in DNA origami, for instance, have measured stiffnesses in the range of 0.1–1 pN/nm [76]. For realistic frameworks, the shape space $\Sigma$ would then form a continuum of ground states (assuming that energy costs arise solely due to the constraints being violated). Keeping this in mind, we assume that the energy stored in the $i$th bar of the framework is a function $\phi_i(\ell_i)$ of its length $\ell_i$. We additionally assume that this function has a minimum when $\ell_i = \bar{\ell}_i$, the natural length of the bar, for all bars. The total energy of the framework is then

$$U(\boldsymbol{q}) = \sum_{i=1}^{m} \phi_i[\ell_i(\boldsymbol{q})]\,. \tag{2.11}$$

On the shape space $\Sigma$, the bars are at their natural lengths, i.e., for all points $\bar{\boldsymbol{q}} \in \Sigma$, we have $\ell_i(\bar{\boldsymbol{q}}) = \bar{\ell}_i$. The bars are also at their minimum potential energies $\phi_i(\bar{\ell}_i)$, which we assume to be zero for all bars without any loss of generality. We first look at the lowest-order approximation to the energy, namely, the harmonic approximation.

### 2.3.1 Harmonic approximation

The $i$th bar in the framework has a potential energy $\phi_i[\ell_i(\boldsymbol{q})]$. The constraint function associated with this bar is $f_i(\boldsymbol{q}) = [\ell_i^2(\boldsymbol{q}) - \bar{\ell}_i^2]/(2\bar{\ell}_i)$. Consider a point $\bar{\boldsymbol{q}}$ on the shape space $\Sigma$

[10] Of course, the constraints are not preserved at higher orders, and it can be shown that the framework is rigid at second order. Well-known results in rigidity theory then guarantee that the framework is rigid to all orders [22].

where the framework is in equilibrium. Dropping the index $i$ for now and noting that $\ell(\bar{\boldsymbol{q}}) = \bar{\ell}$, we first express the gradient and the Hessian of the length $\ell(\boldsymbol{q})$ at $\boldsymbol{q} = \bar{\boldsymbol{q}}$ in terms of the gradient and the Hessian of $f(\boldsymbol{q})$ as

$$\begin{aligned} \partial_j \ell &= \partial_j f, \\ \partial_j \partial_k \ell &= \partial_j \partial_k f - \bar{\ell}^{-2} \partial_j f \partial_k f, \end{aligned} \tag{2.12}$$

where all the partial derivatives are with respect to the components of $\boldsymbol{q}$ and evaluated at $\boldsymbol{q} = \bar{\boldsymbol{q}}$.

Since we have assumed the potential energy of all bars to have a minimum value (assumed to be zero) on the shape space, where $\ell = \bar{\ell}$, the derivative $\phi'(\bar{\ell})$ vanishes. We wish to find the lowest-order expansion of the total potential energy $U(\boldsymbol{q})$ near a point $\bar{\boldsymbol{q}} \in \Sigma$. First, let us consider the case where $\bar{\boldsymbol{q}}$ is not a point of self stress, in which case all zero modes are tangent to $\Sigma$ and can be extended to smooth deformations of the framework that do not cost energy to any order. In such a situation, after setting $\boldsymbol{q} \to \bar{\boldsymbol{q}} + \boldsymbol{q}$, the potential energy of a single bar to $\mathcal{O}(\|\boldsymbol{q}\|^2)$ is the familiar Hookean potential

$$\phi[\ell(\bar{\boldsymbol{q}} + \boldsymbol{q})] = \tfrac{1}{2}\phi''(\bar{\ell})(\partial_j \ell \, q_j)^2 + \mathcal{O}(\|\boldsymbol{q}\|^3). \tag{2.13}$$

Above, the partial derivatives of $\ell(\boldsymbol{q})$ and the derivatives of $\phi(\ell)$ have been evaluated at $\bar{\boldsymbol{q}} \in \Sigma$ and $\ell = \bar{\ell}$, respectively. The repeated indices are summed over as usual. Defining the stiffness of the $i$th bar to be $\kappa_i = \phi_i''(\ell_i)$ and noting that $\partial_j \ell_i \, q_j = \partial_j f_i \, q_j = \nabla f_i \cdot \boldsymbol{q}$, the total energy of the framework becomes (after reintroducing the free index $i$)

$$U = \sum_{i=1}^{m} \phi_i[\ell_i(\bar{\boldsymbol{q}} + \boldsymbol{q})] \approx \tfrac{1}{2} \sum_{i=1}^{m} \kappa_i (\nabla f_i \cdot \boldsymbol{q})^2 = \tfrac{1}{2} \boldsymbol{q}^{\mathsf{T}} \mathsf{C}^{\mathsf{T}} \mathsf{K} \mathsf{C} \boldsymbol{q} = \tfrac{1}{2} \boldsymbol{q}^{\mathsf{T}} \mathsf{D} \boldsymbol{q}. \tag{2.14}$$

Here $\mathsf{D} = \mathsf{C}^{\mathsf{T}}\mathsf{K}\mathsf{C}$ is the dynamical matrix evaluated at $\bar{\boldsymbol{q}}$ [106] (assuming joints of unit mass) and $\mathsf{K}$ is the $n \times n$ diagonal matrix of bar stiffnesses $\kappa_i$. Since the potential functions $\phi_i$ all have a minimum at $\ell_i = \bar{\ell}_i$, the stiffnesses $\kappa_i$ are positive nonzero numbers and the matrix $\mathsf{K}$ is positive definite. However, the dynamical matrix $\mathsf{D}$ is, in general, only positive semidefinite since $\ker \mathsf{C}$ need not be empty. Indeed, it will not be empty for the example frameworks we consider in the next chapter, as they all have internal degrees of freedom.

### 2.3.2 Higher-order approximations

Now let us analyze the case where there are additional zero modes due to the presence of a self stress. Dropping the index $i$ again, we first expand the energy of a bar to $\mathcal{O}(\|\boldsymbol{q}\|^4)$ and find

$$\begin{aligned} \phi[\ell(\bar{\boldsymbol{q}} + \boldsymbol{q})] = {} & \tfrac{1}{2}\phi''(\bar{\ell})(\partial_j \ell \, q_j)^2 + \tfrac{1}{6}\phi'''(\bar{\ell})(\partial_j \ell \, q_j)^3 + \tfrac{1}{2}\phi''(\bar{\ell})(\partial_j \ell \, q_j)\partial_k \partial_l \ell \, q_k q_l \\ & + \frac{1}{24}\phi''''(\bar{\ell})(\partial_j \ell \, q_j)^4 + \tfrac{1}{4}\phi'''(\bar{\ell})(\partial_j \ell \, q_j)^2 \partial_k \partial_l \ell \, q_k q_l \\ & + \tfrac{1}{8}\phi''(\bar{\ell})(\partial_j \partial_k \ell \, q_j q_k)^2 + \tfrac{1}{6}\phi''(\bar{\ell})(\partial_j \ell \, q_j)\partial_k \partial_l \partial_m \ell \, q_k q_l q_m + \mathcal{O}(\|\boldsymbol{q}\|^5). \end{aligned} \tag{2.15}$$

The subspace of zero modes is $\ker \mathsf{C}$ and its orthogonal complement in $\mathbb{R}^n$ is $(\ker \mathsf{C})^{\perp}$ and we write $\boldsymbol{q} = \boldsymbol{u} + \boldsymbol{v}$. Here $\boldsymbol{u} \in \ker \mathsf{C}$ is a zero mode and $\boldsymbol{v} \in (\ker \mathsf{C})^{\perp}$ is a fast vibrational mode of the system. First let us analyze the contribution of the zero mode $\boldsymbol{u}$ to the energy. Since $\boldsymbol{u}$ is a

zero mode, we have $\mathsf{C}\boldsymbol{u} = 0$. Now, as $\partial_j f$ is a row of C, this implies $\partial_j f\, u_j = 0$. Using this in Eq. (2.12) we get

$$\begin{aligned} \partial_j \ell\, u_j &= \partial_j f\, u_j = 0, \\ \partial_j \partial_k \ell\, u_j u_k &= \partial_j \partial_k f\, u_j u_k - \bar{\ell}^{-2} (\partial_j f\, u_j)(\partial_k f\, u_k) \\ &= \partial_j \partial_k f\, u_j u_k. \end{aligned} \tag{2.16}$$

Setting $\boldsymbol{v} = 0$ in the series expansion [Eq. (2.15)] and using the above simplifications, it is clear that the only nonvanishing contribution to the energy would come from the $\mathcal{O}(\|\boldsymbol{u}\|^4)$ term $\frac{1}{8}\phi''(\bar{\ell})(\partial_j \partial_k \ell\, u_j u_k)^2 = \frac{1}{8}\phi''(\bar{\ell})(\partial_j \partial_k f\, u_j u_k)^2$.

To understand how a fast mode $\boldsymbol{v}$ contributes to the energy, we similarly set $\boldsymbol{u} = 0$ in the energy expansion. Since $\boldsymbol{v}$ belongs to the orthogonal complement of the subspace of zero modes, by definition, $\partial_j \ell\, v_j = \partial_j f\, v_j \neq 0$. Hence, we see that the first nonvanishing contribution comes from the $\mathcal{O}(\|\boldsymbol{v}\|^2)$ term $\frac{1}{2}\phi''(\bar{\ell})(\partial_j \ell\, v_j)^2 = \frac{1}{2}\phi''(\bar{\ell})(\partial_j f\, v_j)^2$. This shows that the energy scales as $\mathcal{O}(\|\boldsymbol{u}\|^4) \sim \mathcal{O}(\|\boldsymbol{v}\|^2)$, i.e., zero and fast modes respectively make quartic- and harmonic-order contributions to the energy. Finally, we set $\boldsymbol{q} = \boldsymbol{u} + \boldsymbol{v}$ in Eq. (2.15), and find the energy of the $i$th bar to the lowest order in $\boldsymbol{u}$ and $\boldsymbol{v}$ as

$$\begin{aligned} \phi_i = {} & \tfrac{1}{2}\phi_i''(\bar{\ell}_i)(\partial_j f_i\, v_j)^2 + \tfrac{1}{2}\phi_i''(\bar{\ell}_i)(\partial_j f_i\, v_j)(\partial_k \partial_l f_i\, u_k u_l) + \tfrac{1}{8}\phi_i''(\bar{\ell}_i)(\partial_k \partial_l f_i\, u_k u_l)^2 \\ & + \mathcal{O}(\|\boldsymbol{u}\|^5) + \mathcal{O}(\|\boldsymbol{u}\|^3 \|\boldsymbol{v}\|) + \mathcal{O}(\|\boldsymbol{u}\| \|\boldsymbol{v}\|^2), \end{aligned} \tag{2.17}$$

where we have reintroduced the free index $i$. In the above equation, $\phi_i''(\bar{\ell}_i) = \kappa_i$, the stiffness of the $i$th bar, $\partial_j f_i v_j = \nabla f_i \cdot \boldsymbol{v}$, and $\partial_k \partial_l f_i\, u_k u_l = \boldsymbol{u}^{\mathsf{T}} (\nabla\nabla f_i)\boldsymbol{u}$, with $\nabla\nabla f_i$ being the Hessian matrix of the $i$th constraint function $f_i$. Then, the energy of the $i$th bar to the lowest order is

$$\phi_i \approx \tfrac{1}{2}\kappa_i \left[ \nabla f_i \cdot \boldsymbol{v} + \tfrac{1}{2}\boldsymbol{u}^{\mathsf{T}} (\nabla\nabla f_i)\boldsymbol{u} \right]^2. \tag{2.18}$$

For convenience, we define a column vector $\boldsymbol{w}(\boldsymbol{u}) \in \mathbb{R}^m$ such that its $i$th component $[\boldsymbol{w}(\boldsymbol{u})]_i = \frac{1}{2}\boldsymbol{u}^{\mathsf{T}} (\nabla\nabla f_i)\boldsymbol{u}$. With this definition, the total energy to the lowest order in $\boldsymbol{u}$ and $\boldsymbol{v}$ is

$$U = \sum_{i=1}^{m} \phi_i \approx \tfrac{1}{2} [\mathsf{C}\boldsymbol{v} + \boldsymbol{w}(\boldsymbol{u})]^{\mathsf{T}}\, \mathsf{K}\, [\mathsf{C}\boldsymbol{v} + \boldsymbol{w}(\boldsymbol{u})]. \tag{2.19}$$

This equation is a generalization of the energy of harmonic spring networks [172, 179] and colloidal clusters [77] that have zero modes, for an arbitrary number of zero modes and degrees of freedom, as well as a general interaction energy $\phi_i$ between the particles (i.e., the joints).

**Example (Tethered particle: energetics).** Let us consider again the tethered particle in Fig. 2.4(a). If the particle is now displaced from its ground state at $(0,0)$ to a point $(q_1, q_2)$, we can break up the displacement into a fast mode $\boldsymbol{v} = (q_1, 0)$ along the horizontal axis and a zero mode $\boldsymbol{u} = (0, q_2)$ along the vertical axis. The Hessian matrices of the two constraints are $\nabla\nabla f_1 = \mathsf{I}_2$ and $\nabla\nabla f_2 = \mathsf{I}_2$, so the vector $\boldsymbol{w}(\boldsymbol{u})$ is

$$\boldsymbol{w}(\boldsymbol{u}) = \left[ \tfrac{1}{2}\boldsymbol{u}\left(\nabla\nabla f_1\right)\boldsymbol{u}, \tfrac{1}{2}\boldsymbol{u}\left(\nabla\nabla f_1\right)\boldsymbol{u} \right] = \left( \tfrac{1}{2} q_2^2, \tfrac{1}{2} q_2^2 \right). \tag{2.20}$$

As $\mathsf{C}\boldsymbol{v} = (q_1, -q_1)$, and assuming that the bars have a constant stiffness $\kappa$, using Eq. (2.19), we find the energy[11]

$$U \approx \tfrac{1}{2}\kappa \|\mathsf{C}\boldsymbol{v} + \boldsymbol{w}(\boldsymbol{u})\|^2 = \kappa\left(q_1^2 + \tfrac{1}{4}q_2^4\right). \tag{2.21}$$

As we see from the energy expression, the zero mode $(0, q_2)$ contributes to the energy only at the quartic order—physically speaking, one would say that the framework has become softer due to the presence of the singularity.

We now briefly remark on an interesting comparison between Eq. (2.19) and the nonlinear strain that is considered while analyzing large deflections of a thin elastic plate. In our notation, this strain can be written as[12]

$$\epsilon_{ij} = \tfrac{1}{2}\left[\partial_i v_i + \partial_j v_i + (\partial_i u)(\partial_j u)\right]. \tag{2.22}$$

The above strain has been written down assuming that the plate, which is initially flat and described as a plane $\boldsymbol{r}(x, y)$ in $\mathbb{R}^3$, is undergoing a deformation $\boldsymbol{r}(x, y) \to \boldsymbol{r}(x, y) + v_1(x, y)\hat{\boldsymbol{x}} + v_2(x, y)\hat{\boldsymbol{y}} + u(x, y)\hat{\boldsymbol{z}}$. If we were to consider the plate as being composed of a set of discrete points connected by stiff bars, then the zero modes are the up/down out-of-plane displacements that the points can undergo without changing the constraints to linear order. Consequently, the fast vibrational modes would be composed of the in-plane displacements. In the discrete limit, the in-plane components of the displacement field, $v_1$ and $v_2$, would get replaced by a single vector $\boldsymbol{v} \in (\ker\mathsf{C})^\perp$. Meanwhile, the out-of-plane component $u$ of the displacement field would be replaced by a vector $\boldsymbol{u} \in \ker\mathsf{C}$. With this setup, the deformation energy can be obtained from Eq. (2.19). We then realize that the bracketed term in Eq. (2.19), i.e., $\mathsf{C}\boldsymbol{v} + \boldsymbol{w}(\boldsymbol{u})$ is a discrete version of the continuum strain in Eq. (2.22): the linear part $\mathsf{C}\boldsymbol{v}$ is roughly analogous to $(\partial_i v_j + \partial_j v_i)$, whereas the nonlinear part $\boldsymbol{w}(\boldsymbol{u})_i \sim \boldsymbol{u}^\mathsf{T}(\nabla\nabla f_i)\boldsymbol{u}$ is analogous to $(\partial_i u)(\partial_j u)$. Thus, the energy in Eq. (2.19) can be thought of as the square of a discrete, nonlinear strain.

[11] The curious reader may wonder if this energy is dimensionally correct—after all $q_1$ and $q_2$ have the same dimensions. The issue here is that we took the bars to have unit lengths, so the numerical factors in $\boldsymbol{w}(\boldsymbol{u})$ should be assumed to have units of inverse length.

[12] See, e.g., Section 2.14 of Ref. [90] or Section 6.2 of Ref. [125].

Chapter 3

# Thermal fluctuations of frameworks

In this chapter we study the effects that thermal fluctuations have on bar-joint frameworks at the submicron scale. In particular, we study the free-energy landscape of these frameworks, and show that the landscape is dominated by the neighborhoods of points that correspond to singularities in the framework's shape space.

Energy landscapes have been extensively used in all branches of science and engineering to provide an intuitive understanding of complex dynamical phenomena. They have been crucial in understanding a host of problems in chemical kinetics, protein folding, optimization, machine learning, etc. [166]. However, the energy $U(\boldsymbol{r})$ of a physical system is usually a scalar function of a possibly high-dimensional configuration vector $\boldsymbol{r}$ describing the system. This curse of dimensionality oftentimes makes it difficult to visualize energy landscapes directly. For this reason, particularly in the case of physical systems coupled to a heat bath, it is much more customary to consider the free-energy landscape obtained by projecting the free energy of the system down to a small number of physically relevant coordinates. Below, we provide a brief description of free-energy landscapes in a general setting.

## 3.1 Free-energy landscapes

Similar to the discussion in the previous chapter, consider a system of $N$ classical particles in $d$ dimensions. Let the position of the $i$th particle be $\boldsymbol{r}_i \in \mathbb{R}^d$. We can then use a configuration vector $\boldsymbol{r} \in \mathscr{R} \subseteq \mathbb{R}^{Nd}$ defined by $\boldsymbol{r} = (\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$ to describe all configurations of the system. Corresponding to $\boldsymbol{r}$ we define a momentum vector $\boldsymbol{p} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_N)$, where $\boldsymbol{p}_i$ is the momentum of the $i$th particle. The microscopic state of the system at a given moment is then described by the position-momentum pair $(\boldsymbol{r}, \boldsymbol{p})$ in phase space. Additionally, let the total energy of the system be given by the Hamiltonian

$$H(\boldsymbol{r}, \boldsymbol{p}) = E_{\text{kinetic}}(\boldsymbol{p}) + U(\boldsymbol{r}) = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m_i} + U(\boldsymbol{r}). \tag{3.1}$$

Here we have assumed that the Hamiltonian $H$ is separable, i.e., the energetic contributions of the particle momenta and positions are independent.

Physical systems that have been equilibrated with a thermal bath of temperature $T$ frequently arise in statistical physics. Although the energy of such a system undergoes fluctuations, the system itself is always at a fixed temperature. Then, the various microscopic

configurations of the systems, i.e., $(\boldsymbol{r}, \boldsymbol{p})$ pairs, are distributed according to the Boltzmann–Gibbs (or canonical) density

$$\mathscr{P}(\boldsymbol{r}, \boldsymbol{p}) = Z^{-1} \exp\left[-\beta H(\boldsymbol{r}, \boldsymbol{p})\right] \tag{3.2}$$

where $\beta = (k_B T)^{-1}$ is the inverse temperature, with $k_B$ being the Boltzmann constant. Also, the normalization factor $Z$ is the partition function defined by $Z = \int \mathrm{d}\boldsymbol{r}\, \mathrm{d}\boldsymbol{p} \exp[-\beta H(\boldsymbol{r}, \boldsymbol{p})]$. When the Hamiltonian $H$ is separable, we can perform the momentum integrals independently, so the effective, momentum-reduced probability density is

$$\mathscr{P}(\boldsymbol{r}) = \exp\left[-\beta U(\boldsymbol{q})\right]. \tag{3.3}$$

Above, we have purposefully omitted factors of normalization—something we will continue to do as much as possible.

As we remarked earlier, most physical systems, even those with a handful of particles, suffer from the curse of dimensionality. High-dimensional systems that are coupled to a heat bath are better analyzed by considering the marginal probability density of a lower-dimensional collective variable (CV), which is some physically relevant observable that (hopefully) tells us something useful about the system. Let $\xi \in \mathbb{R}^l$ with $l \geq 1$ be such an observable that depends only on particle positions and measurable by a sufficiently smooth map $\hat{\xi} : \mathbb{R}^n \to \mathbb{R}^l$, with $n = Nd$. The marginal probability density $\mathscr{P}(\xi)$ of the CV $\xi$ is defined as

$$\mathscr{P}(\xi) = \int \mathrm{d}\boldsymbol{r}\, \delta\left[\hat{\xi}(\boldsymbol{r}) - \xi\right] \exp\left[-\beta U(\boldsymbol{q})\right]. \tag{3.4}$$

Here $\delta[\cdot]$ is the $l$-dimensional Dirac delta function, which restricts the domain of integration to the $(n-l)$-dimensional CV level set $\hat{\xi}^{-1}(\xi) = \{\boldsymbol{q} \in \mathbb{R}^n : \hat{\xi}(\boldsymbol{q}) = \xi\}$ [61]. The usage of the term "collective variable" is now clear: the marginal density $\mathscr{P}(\xi)$ no longer refers to a single microstate—instead, it refers to a collection of an infinite number of microstates such that if we were to measure the value of the CV using any one of the states, we would get $\xi$.

To show that Eq. (3.4) gives the correct marginal density, we first note that the average $\langle\phi\rangle$ of an observable $\phi \equiv \phi(\xi) = \phi[\hat{\xi}(\boldsymbol{r})]$ that depends only on the CV $\xi$ can be computed using either the marginal density $\mathscr{P}(\xi)$ or the Boltzmann–Gibbs density $\mathscr{P}(\boldsymbol{r})$, i.e.,

$$\langle\phi\rangle = \int \mathrm{d}\xi'\, \phi(\xi')\mathscr{P}(\xi') = \int \mathrm{d}\boldsymbol{r}\, \phi[\hat{\xi}(\boldsymbol{r})]\mathscr{P}(\boldsymbol{r}). \tag{3.5}$$

Setting $\phi(\xi') = \delta(\xi' - \xi)$ yields

$$\int \mathrm{d}\xi'\, \delta(\xi' - \xi)\mathscr{P}(\xi') = \mathscr{P}(\xi) = \int \mathrm{d}\boldsymbol{r}\, \delta[\hat{\xi}(\boldsymbol{r}) - \xi]\mathscr{P}(\boldsymbol{r}), \tag{3.6}$$

consistent with the definition in Eq. (3.4).

The thermodynamic free energy associated with the CV $\xi$ is defined using the marginal density $\mathscr{P}(\xi)$ as [94]

$$\mathscr{A}(\xi) = -\beta^{-1} \log \mathscr{P}(\xi). \tag{3.7}$$

With this definition, we see that the free energy $\mathscr{A}(\xi)$ has an "inverse" relationship with $\mathscr{P}(\xi)$, i.e., a low value of $\mathscr{A}(\xi)$ indicates a higher probability of observing states with a CV value of $\xi$, and vice versa. Also, because of the arbitrariness of normalization factors, it makes

better sense to speak of free energy differences rather than absolute free energies, e.g., the free-energy difference between two CV values $\xi_1$ and $\xi_2$ is $\Delta\mathscr{A} = \mathscr{A}(\xi_2) - \mathscr{A}(\xi_1)$.

## 3.2 Free energy of frameworks

Recall that the total elastic energy of a framework is of the form $U(\boldsymbol{r}) = \sum_{i=1}^{m} \phi_i(\ell_i)$, where $\ell_i$ is the length of the $i$th bar with an energy $\phi_i(\ell_i)$, which is assumed to have a minimum value of zero at $\ell_i = \bar{\ell}_i$, the natural length of the $i$th bar. With the above form of the energy, all nontrivial ground states of a framework belong to its shape space $\Sigma$ [86, 87, 119], which is the set of all deformed configurations of the framework with the length of each bar equal to its natural length, once rotations and translations are removed. We remarked earlier that shape spaces of frameworks are parameterized by coordinates called shape coordinates [101]—usually internal angles—that capture the framework's shape changes (see the discussion in Section 2.2). As zero-energy shape changes constitute the most interesting dynamical feature of a framework, it follows that the shape coordinates are natural CVs for a low-dimensional description of a thermally excited framework.

Let us assume that the value of the chosen CV/shape coordinate for any configuration $\boldsymbol{q} \in \mathbb{R}^n$ of the framework in its body frame can be measured using the CV map $\hat{\xi}(\boldsymbol{q})$. In the case of the four-bar linkage, for example, if we choose $\theta_1$ as the CV, then $\hat{\xi}(\boldsymbol{q})$ is the map that computes $\theta_1$ for any $\boldsymbol{q}$, whether or not it lies on the branches of the linkage's shape space (see Fig. 2.3 on p. 15). For general frameworks, the marginal density $\mathscr{P}(\xi)$ of the CV, aside from factors of normalization, is

$$\mathscr{P}(\xi) = \int_{\mathbb{R}^n} \mathrm{d}\boldsymbol{q}\, I(\boldsymbol{q})\, \delta\left[\hat{\xi}(\boldsymbol{q}) - \xi\right] \exp\left[-\beta U(\boldsymbol{q})\right]. \tag{3.8}$$

In the above equation, except for the Jacobian factor $I(\boldsymbol{q})$ introduced by the change of coordinates from the lab frame to the body frame, all the other terms have the same meaning as in Eq. (3.4). Also, as the CV $\xi$ is an $(n-m)$-dimensional shape coordinate, the CV map is $\hat{\xi} : \mathbb{R}^n \to \mathbb{R}^{n-m}$. Furthermore, the CV level set $\hat{\xi}^{-1}(\xi)$ will be an $m$-dimensional manifold if $\nabla\hat{\xi}$ has full rank in $\hat{\xi}^{-1}(\xi)$. Assuming this, the marginal density $\mathscr{P}(\xi)$ can be written as an exact $m$-dimensional surface integral over $\hat{\xi}^{-1}(\xi)$ using the coarea formula [29, 61, 63, 94],

$$\mathscr{P}(\xi) = \int_{\hat{\xi}^{-1}(\xi)} \frac{\mathrm{d}\Theta(\boldsymbol{q})}{|\det \nabla\hat{\xi}(\nabla\hat{\xi})^{\mathsf{T}}|^{1/2}}\, I(\boldsymbol{q}) \exp\left[-\beta U(\boldsymbol{q})\right]. \tag{3.9}$$

Here $\mathrm{d}\Theta(\boldsymbol{q})$ is the surface measure on $\hat{\xi}^{-1}(\xi)$ and $\nabla\hat{\xi}$ is the $(n-m)\times n$ Jacobian matrix of $\hat{\xi}$ at $\boldsymbol{q}$. Sometimes, Eq. (3.9) is taken to be the definition of $\mathscr{P}(\xi)$ instead of Eq. (3.8). This only makes sense when $\nabla\hat{\xi}$ has full rank in $\hat{\xi}^{-1}(\xi)$ so that the determinant $\det \nabla\hat{\xi}(\nabla\hat{\xi})^{\mathsf{T}}$ does not vanish [94].[1] Furthermore, although we have in mind an $(n-m)$-dimensional CV $\xi$ that can be used to parameterize the branches of the shape space $\Sigma$, we do not require a parameterization at hand to use Eq. (3.9). For instance, in a framework the CV could be one of its internal angles, whose value can be directly computed from the coordinates of the framework's joints.

[1] Equation (3.9) is the arbitrary-dimensional analogue of the RHS in $\int_{\mathbb{R}} \mathrm{d}x\, f(x)\, \delta[g(x)] = \sum_i f(x_i)/|g'(x_i)|$, where $f$ and $g$ are scalar functions in $\mathbb{R}$, and $x_i$ are the roots of $g(x) = 0$. This is assuming $|g'(x_i)| \neq 0$, which is equivalent to the assumption that $\nabla\hat{\xi}$ has full rank in $\hat{\xi}^{-1}(\xi)$.

(The map that turns the coordinates $\boldsymbol{q}$ into the internal angle is the CV map $\hat{\xi}(\boldsymbol{q})$ for the framework.) Hence, in theory, using this equation only requires knowledge of the energy $U(\boldsymbol{q})$, the Jacobian factor $I(\boldsymbol{q})$, and the CV map $\hat{\xi}$. Most importantly, Eq. (3.9) makes no reference to the shape space $\Sigma$, or its branches, or whether or not it has singularities. However, in general, the CV level set $\hat{\xi}^{-1}(\xi)$ is bound to be a curved high-dimensional manifold. Hence, directly evaluating the integral in Eq. (3.9) becomes cumbersome, and we have to resort to asymptotic methods to evaluate it.

It is clear from both Eqs. (3.8) and (3.9) that in the large-$\beta$ limit, contributions to the marginal density would mainly come from the neighborhoods of the ground states in the CV level set $\hat{\xi}^{-1}(\xi)$ since the energy $U(\boldsymbol{q})$ vanishes at those points (see Fig. 3.1). Since all ground states belong to the shape space $\Sigma$, they could be regular (i.e., nonsingular) points or singularities of $\Sigma$. We call $\xi$ a regular value of the CV if $\hat{\xi}^{-1}(\xi)$ does not contain singularities of $\Sigma$ and vice-versa. After asymptotically evaluating the integral in Eq. (3.8) in the neighborhood of each ground state (e.g., using Laplace's method [13]), we can then sum the results to find the asymptotic expression of the marginal density. In the following, to simplify the presentation, we will only consider cases where the CV level set contains just one ground state, which is either a regular point or a singularity of $\Sigma$. More general cases can then be handled by using appropriate combinations of the results we derive.

### 3.2.1 Asymptotic marginal density: regular values

We first consider the case where $\xi$ is a regular value of the CV, i.e., when the CV level set $\hat{\xi}^{-1}(\xi)$ contains only regular points of $\Sigma$. As we remarked previously, using Eq. (3.9) to find the marginal density is difficult. Hence, we will use a Gaussian representation of the delta function to write the marginal density as [61, 62]

$$\mathscr{P}(\xi) = \lim_{\alpha\to\infty} \left(\frac{\alpha}{2\pi}\right)^{(n-m)/2} \int_{\mathbb{R}^n} \mathrm{d}\boldsymbol{q}\, I(\boldsymbol{q}) \exp\left[-\tfrac{1}{2}\alpha \|\hat{\xi}(\boldsymbol{q}) - \xi\|^2 - \beta U(\boldsymbol{q})\right], \tag{3.10}$$

where $\|\cdot\|$ is the $(n-m)$-dimensional Euclidean norm. For $\boldsymbol{q} \in \hat{\xi}^{-1}(\xi)$, the norm $\|\hat{\xi}(\boldsymbol{q}) - \xi\|$ vanishes. Similarly, for $\boldsymbol{q} \in \Sigma$, the energy $U(\boldsymbol{q})$ vanishes. This means that in the limit $\alpha, \beta \to \infty$, contributions to the above integral would mainly come from the neighborhood of the ground state $\bar{\boldsymbol{q}} = \Sigma \cap \hat{\xi}^{-1}(\xi)$. Hence, we can evaluate the above integral using Laplace's method after expanding the two terms in the exponent of Eq. (3.10) to the lowest order around $\bar{\boldsymbol{q}}$. To this end, we set $\boldsymbol{q} \to \bar{\boldsymbol{q}} + \boldsymbol{q}$, and expand the energy $U(\boldsymbol{q})$ using the harmonic approximation: $U \approx \frac{1}{2}\boldsymbol{q}^{\mathsf{T}}\mathsf{D}\boldsymbol{q}$ [see Eq. (2.14)], where $\mathsf{D}$ is the $n \times n$ dynamical matrix. Since $\bar{\boldsymbol{q}}$ is a regular point of $\Sigma$, the compatibility matrix $\mathsf{C}$ has full rank, and $\mathsf{D}$ has $n-m$ independent zero modes that belong to $\ker\mathsf{C}$ [106]. These zero modes are all tangent to $\Sigma$ and represent a degree of freedom [93]. Hence, to asymptotically evaluate Eq. (3.8) in the neighborhood of a regular point, we can safely use the harmonic approximation since any divergence [39, 147, 149] due to these zero modes is regularized by the delta function [or in the case of Eq. (3.10), its Gaussian approximation], which suppresses all contributions to the integral that are tangent to $\Sigma$ [142]. This gives us

$$\mathscr{P}(\xi) \sim \lim_{\alpha\to\infty} \frac{\alpha^{-m/2}}{(2\pi)^{(n-m)/2}} I(\bar{\boldsymbol{q}}) \int_{\mathbb{R}^n} \mathrm{d}\boldsymbol{q} \exp\left\{-\tfrac{1}{2}\boldsymbol{q}^{\mathsf{T}}\left[(\nabla\hat{\xi})^{\mathsf{T}}\nabla\hat{\xi} + \alpha^{-1}\beta\mathsf{D}\right]\boldsymbol{q}\right\}, \tag{3.11}$$

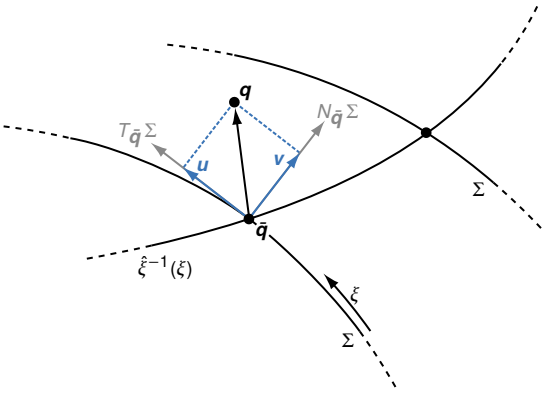


**Figure 3.1.** Cartoon illustrating the shape space $\Sigma$ parameterized by the CV $\xi$ and the CV level set $\hat{\xi}^{-1}(\xi)$. In general, the sets $\Sigma$ and $\hat{\xi}^{-1}(\xi)$ could intersect at multiple points, each of which is an energy ground state. Tangent and normal spaces of $\Sigma$ at such a point $\bar{\boldsymbol{q}} \in \Sigma \cap \hat{\xi}^{-1}(\xi)$ are $T_{\bar{\boldsymbol{q}}}\Sigma$ and $N_{\bar{\boldsymbol{q}}}\Sigma$.

where the dynamical matrix $\mathsf{D}$ and the Jacobian matrix $\nabla\hat{\xi}$ are evaluated at $\bar{\boldsymbol{q}}$. We have also rescaled $\boldsymbol{q} \to \alpha^{-1/2}\boldsymbol{q}$ for convenience.

We shall now evaluate the Gaussian integral in Eq. (3.11). Let us assume that the local parameterization of $\Sigma$ near $\bar{\boldsymbol{q}}$ in terms of the CV is $\psi : \mathbb{R}^{n-m} \to \mathbb{R}^n$, which means that $\bar{\boldsymbol{q}} = \psi(\xi)$. Now, the zero modes at $\bar{\boldsymbol{q}}$ belong to the tangent space $T_{\bar{\boldsymbol{q}}}\Sigma = \ker\mathsf{C}$ by definition [93]. This implies that the normal modes of the dynamical matrix, which are orthogonal to the zero modes, belong to the $m$-dimensional normal space $N_{\bar{\boldsymbol{q}}}\Sigma = (\ker\mathsf{C})^{\perp}$. Since $T_{\bar{\boldsymbol{q}}}\Sigma \oplus N_{\bar{\boldsymbol{q}}}\Sigma = \mathbb{R}^n$, we can always write $\boldsymbol{q} = \boldsymbol{u} + \boldsymbol{v}$ with $\boldsymbol{u} \in T_{\bar{\boldsymbol{q}}}\Sigma$ and $\boldsymbol{v} \in N_{\bar{\boldsymbol{q}}}\Sigma$, for any $\boldsymbol{q} \in \mathbb{R}^n$ (see Fig. 3.1). As an orthonormal basis for $N_{\bar{\boldsymbol{q}}}\Sigma$, choose the normal modes $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m \in \mathbb{R}^n$ and as the basis for $T_{\bar{\boldsymbol{q}}}\Sigma$ choose the columns of the $n \times (n-m)$ Jacobian matrix $\nabla\psi(\xi)$.[2] Therefore, we can write

$$\boldsymbol{q} = \boldsymbol{u} + \boldsymbol{v} = (\nabla\psi)\boldsymbol{x} + \mathsf{B}\boldsymbol{y}, \tag{3.12}$$

where $\mathsf{B} = \begin{pmatrix}\boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_m\end{pmatrix}$ is the change-of-basis matrix for $N_{\bar{\boldsymbol{q}}}\Sigma$. The vectors $\boldsymbol{x} \in \mathbb{R}^{n-m}$ and $\boldsymbol{y} \in \mathbb{R}^m$ represent the components of $\boldsymbol{u}$ and $\boldsymbol{v}$ in the chosen bases. Eq. (3.12) suggests the coordinate change $\boldsymbol{q} \to (\boldsymbol{x}, \boldsymbol{y})$. The quadratic form in the integral of Eq. (3.11) after such a coordinate change can be decomposed into blocks as

$$\tfrac{1}{2}\begin{pmatrix}\boldsymbol{x}^{\mathsf{T}} & \boldsymbol{y}^{\mathsf{T}}\end{pmatrix}\begin{pmatrix}(\nabla\psi)^{\mathsf{T}}(\nabla\hat{\xi})^{\mathsf{T}}\nabla\hat{\xi}\nabla\psi & (\nabla\psi)^{\mathsf{T}}(\nabla\hat{\xi})^{\mathsf{T}}\nabla\hat{\xi}\mathsf{B} \\ \mathsf{B}^{\mathsf{T}}(\nabla\hat{\xi})^{\mathsf{T}}\nabla\hat{\xi}\nabla\psi & \mathsf{B}^{\mathsf{T}}(\nabla\hat{\xi})^{\mathsf{T}}\nabla\hat{\xi}\mathsf{B} + \alpha^{-1}\beta\mathsf{D}^{\perp}\end{pmatrix}\begin{pmatrix}\boldsymbol{x} \\ \boldsymbol{y}\end{pmatrix}, \tag{3.13}$$

where $\mathsf{D}^{\perp}$ is the diagonal matrix of the $m$ nonzero eigenvalues of $\mathsf{D}$. Our assumption is that $\psi(\xi)$ is a valid parameterization of $\Sigma$ that is compatible with the CV map $\hat{\xi}$, i.e., $\bar{\boldsymbol{q}} = \psi(\xi)$

[2] The normal modes need not be orthonormal if there is degeneracy in the normal frequencies. But in such a situation, one can still pick a set of orthonormal vectors within the subspace spanned by degenerate normal modes. With some extra steps, the derivation can also be made to work for an arbitrary basis of $N_{\bar{\boldsymbol{q}}}\Sigma$ as well.

and $\hat{\xi}(\bar{\boldsymbol{q}}) = \xi$, which means that

$$(\hat{\xi} \circ \psi)(\xi) = \xi. \tag{3.14}$$

Taking derivatives with respect to $\xi$ on both sides of the above equation we see that

$$\nabla\hat{\xi}(\bar{\boldsymbol{q}})\nabla\psi(\xi) = \mathsf{I}_{n-m}, \tag{3.15}$$

where $\mathsf{I}_{n-m}$ is the $(n-m)\times(n-m)$ identity matrix. Using this, the $n\times n$ block matrix in Eq. (3.13) can be written as

$$\begin{pmatrix} \mathsf{I}_{n-m} & \nabla\hat{\xi}\mathsf{B} \\ \mathsf{B}^{\mathsf{T}}(\nabla\hat{\xi})^{\mathsf{T}} & \mathsf{B}^{\mathsf{T}}(\nabla\hat{\xi})^{\mathsf{T}}\nabla\hat{\xi}\mathsf{B} + \alpha^{-1}\beta\mathsf{D}^{\perp} \end{pmatrix}. \tag{3.16}$$

Since $\mathsf{I}_{n-m}$ is trivially invertible, the determinant of the above matrix is

$$\det \mathsf{I}_{n-m} \det\left[\mathsf{B}^{\mathsf{T}}(\nabla\hat{\xi})^{\mathsf{T}}\nabla\hat{\xi}\mathsf{B} + \alpha^{-1}\beta\mathsf{D}^{\perp} - \mathsf{B}^{\mathsf{T}}(\nabla\hat{\xi})^{\mathsf{T}}\mathsf{I}_{n-m}^{-1}\nabla\hat{\xi}\mathsf{B}\right] = \alpha^{-m}\beta^{m}\det\mathsf{D}^{\perp}. \tag{3.17}$$

Under the coordinate change $\boldsymbol{q} \to (\boldsymbol{x}, \boldsymbol{y})$, the volume element $\mathrm{d}\boldsymbol{q}$ in the integral in Eq. (3.11) acquires the factor $\sqrt{|\det \mathsf{J}^{\mathsf{T}}\mathsf{J}|}$, where $\mathsf{J}$ is the Jacobian of the transformation. The matrix $\mathsf{J}^{\mathsf{T}}\mathsf{J}$ can be readily cast into blocks as

$$\mathsf{J}^{\mathsf{T}}\mathsf{J} = \begin{pmatrix} (\nabla\psi)^{\mathsf{T}}\nabla\psi & 0 \\ 0 & \mathsf{B}^{\mathsf{T}}\mathsf{B} \end{pmatrix} = \begin{pmatrix} (\nabla\psi)^{\mathsf{T}}\nabla\psi & 0 \\ 0 & \mathsf{I}_m \end{pmatrix}. \tag{3.18}$$

Here $(\nabla\psi)^{\mathsf{T}}\nabla\psi$ is the metric induced by the embedding $\xi \mapsto \psi(\xi)$ and since the normal modes are orthonormal vectors, $\mathsf{B}^{\mathsf{T}}\mathsf{B} = \mathsf{I}_m$. Also, the off-diagonal blocks vanish since they involve inner products of the basis vectors of $T_{\bar{\boldsymbol{q}}}\Sigma$ and $N_{\bar{\boldsymbol{q}}}\Sigma$, which are orthogonal complements of each other. This gives $\sqrt{|\det \mathsf{J}^{\mathsf{T}}\mathsf{J}|} = \sqrt{|\det(\nabla\psi)^{\mathsf{T}}\nabla\psi|}$, which together with Eq. (3.17) lets us evaluate the Gaussian integral in Eq. (3.11) and write

$$\mathscr{P}(\xi) \sim I(\xi)\left(\frac{2\pi}{\beta}\right)^{m/2}\left|\frac{\det[\nabla\psi(\xi)]^{\mathsf{T}}\nabla\psi(\xi)}{\det\mathsf{D}^{\perp}(\xi)}\right|^{1/2}. \tag{3.19}$$

The form of the above equation suggests that the nonzero eigenvalues of $\mathsf{D}$ can be naturally interpreted as being inversely proportional to the effective widths of the fluctuations along the $m$ dimensions perpendicular to $\Sigma$.

### 3.2.2 Asymptotic marginal density: singular values

Equation (3.19) breaks down as we approach a singularity along $\Sigma$ since the lowest $s$ nonzero eigenvalues of $\mathsf{D}$ become very small and tend to zero, and we need an alternative method to find the marginal densities. At a singular point of $\Sigma$ with $s$ self stresses, the number of zero modes increases by $s$, which means that now there are only $m-s$ normal modes. The zero modes at a singularity are not all tangent to $\Sigma$, which means that the delta function in Eq. (3.8) fails to suppress the divergences due to these zero modes when the harmonic approximation is used. A cartoon of the situation is depicted in Fig. 3.2, which shows two branches of a one-dimensional shape space $\Sigma$ intersecting at a singularity $\bar{\boldsymbol{q}}^*$, where the CV has the value $\xi^*$. Because of the additional zero mode, the subspace of zero modes, $\ker\mathsf{C}$, is now a plane, which is tangent to both the branches at the singularity.

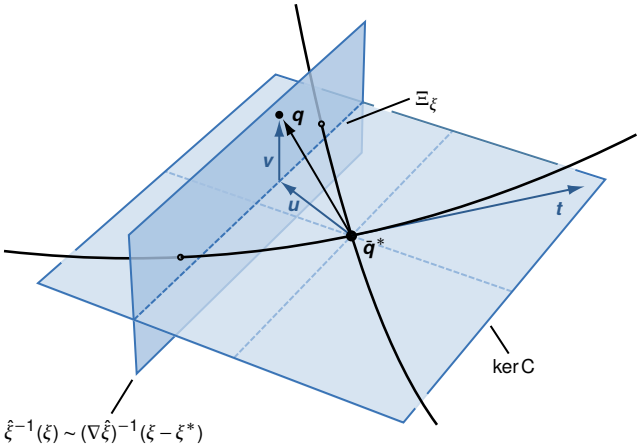


**Figure 3.2.** Cartoon illustrating the geometry of the branches and the level set $(\nabla\hat{\xi})^{-1}(\xi - \xi^*)$ of the linearized CV map near the singularity, which is at $\bar{\boldsymbol{q}}^*$. The two ground-state configurations in the CV level set are indicated by the intersection points of the branches with $(\nabla\hat{\xi})^{-1}(\xi - \xi^*)$. As $\xi \to \xi^*$, these ground states get infinitesimally close to each other and the framework becomes soft. The vector $\boldsymbol{u} \in \ker \mathsf{C}$ and the vector $\boldsymbol{v} \in (\ker \mathsf{C})^\perp$. Also indicated is a tangent vector $\boldsymbol{t}$ to a branch at the singularity. Here there is only one state of self stress and the delta function restricts the integral in Eq. (3.23) to the line $\Xi_\xi$ formed by the intersection of $(\nabla\hat{\xi})^{-1}(\xi - \xi^*)$ and $\ker \mathsf{C}$.

How should one proceed in such a situation? At first glance, one might think that all that is required is to replace the harmonic energy in the previous derivation with the quartic energy expansion, Eq. (2.19). This cannot be correct as Eq. (2.19) is only valid when the expansion is around the singularity, and our goal here is to find $\mathscr{P}(\xi)$ not just for the singular CV value $\xi^*$, but for all $\xi \to \xi^*$. When $\xi$ is close to $\xi^*$, the ground-state configurations in $\hat{\xi}^{-1}(\xi)$ are all regular points of $\Sigma$ that do not have any singular zero modes. However, at these configurations, the framework also becomes very soft in certain directions due to the presence of the nearby singularity, causing the harmonic approximation to break down. A possible strategy could then be to expand the energy to both harmonic and quartic order along these special soft modes, i.e., those that will ultimately become a zero mode at the singularity as we move along $\Sigma$. Apart from the complexity of such an expansion, this raises another issue: how should one perform Laplace asymptotics around the ground states in $\hat{\xi}^{-1}(\xi)$? When $\xi$ is close to $\xi^*$, the ground states in $\hat{\xi}^{-1}(\xi)$ get infinitesimally close to each other, as do the branches (see Fig. 3.2). Hence, on doing Laplace asymptotics around each of these ground states and on extending the integration domain to infinity, we would be adding contributions to the marginal density from the same region more than once. In this sense, the contributions from the branches are not separable in the vicinity of the singularity.

From the above discussion, it should be evident that the approach we used in deriving the harmonic marginal density, Eq. (3.19), will not work for finding an asymptotic expression

for $\mathscr{P}(\xi)$ for $\xi \to \xi^*$. The only possibility then is to rely on Eq. (3.9), which expresses $\mathscr{P}(\xi)$ as an exact integral over the CV level set $\hat{\xi}^{-1}(\xi)$. Note again that Eq. (3.9) makes no reference to the shape space $\Sigma$ as such, enabling us to sidestep the issues described above. This however, comes at the expense of having to do a higher-dimensional integral over $\hat{\xi}^{-1}(\xi)$. The next strategy is to make physically valid assumptions that can be used to dimensionally reduce this integral.

The first such assumption we make is that the CV map can be approximated by its Taylor expansion $\hat{\xi} = \xi^* + (\nabla\hat{\xi})\boldsymbol{q} + \mathcal{O}(\|\boldsymbol{q}\|^2)$ around the singularity $\bar{\boldsymbol{q}}^*$ for points close to $\bar{\boldsymbol{q}}^*$, with $\nabla\hat{\xi}$ being the Jacobian matrix of $\hat{\xi}$ at $\bar{\boldsymbol{q}}^*$. This linearizes the CV map and turns its level sets $(\nabla\hat{\xi})^{-1}(\xi - \xi^*) = \{\boldsymbol{u} \in \mathbb{R}^n : \xi^* + (\nabla\hat{\xi})\boldsymbol{u} = \xi\}$ near the singularity into hyperplanes (see Fig. 3.2). Since this hyperplane is considerably easier to parameterize in comparison to the actual CV level set $\hat{\xi}^{-1}(\xi)$, this makes the evaluation of $\mathscr{P}(\xi)$ as a surface integral using Eq. (3.9) less difficult. The second assumption is that the fast modes $\boldsymbol{v} \in (\ker \mathsf{C})^\perp$ are such that they do not change the value of the CV to linear order at the singularity, i.e., $(\nabla\hat{\xi})\boldsymbol{v} = 0$. This enables us to write $\boldsymbol{q} = \boldsymbol{u} + \boldsymbol{v}$, with $\boldsymbol{u} \in \ker \mathsf{C}$ and $\boldsymbol{v} \in (\ker \mathsf{C})^\perp$, and use the lowest-order approximation of the energy near the singularity [Eq. (2.19)] to get[3]

$$\begin{aligned}\mathscr{P}(\xi) &= \int \mathrm{d}\boldsymbol{q}\, I(\boldsymbol{q})\, \delta\left[\hat{\xi}(\boldsymbol{q}) - \xi\right] \exp\left[-\beta U(\boldsymbol{q})\right] \\ &\sim I(\xi^*) \int_{\ker \mathsf{C}} \mathrm{d}\boldsymbol{u}\, \delta\left[(\nabla\hat{\xi})\boldsymbol{u} - (\xi - \xi^*)\right] \int_{(\ker \mathsf{C})^\perp} \mathrm{d}\boldsymbol{v} \exp\left\{-\tfrac{1}{2}\beta\kappa[\mathsf{C}\boldsymbol{v} + \boldsymbol{w}(\boldsymbol{u})]^\mathsf{T}[\mathsf{C}\boldsymbol{v} + \boldsymbol{w}(\boldsymbol{u})]\right\},\end{aligned}$$

where we have used $(\nabla\hat{\xi})(\boldsymbol{u} + \boldsymbol{v}) = (\nabla\hat{\xi})\boldsymbol{u}$. As we can see from the above equation, the second assumption has allowed us to separate the contributions to $\mathscr{P}(\xi)$ from the fast vibrational modes in $(\ker \mathsf{C})^\perp$ and the zero modes in $\ker \mathsf{C}$. As before, we choose the $m - s$ normal modes as the basis for $(\ker \mathsf{C})^\perp$ and write $\boldsymbol{v} = \mathsf{B}\boldsymbol{y}$, with $\mathsf{B}$ being the change-of-basis matrix and $\boldsymbol{y} \in \mathbb{R}^{m-s}$ representing the components of $\boldsymbol{v}$ in the chosen basis. This turns the integral over $(\ker \mathsf{C})^\perp$ into an $(m - s)$-dimensional Gaussian integral over $\boldsymbol{y}$, which after a straightforward integration yields

$$\begin{aligned}\mathscr{P}(\xi) &\sim I(\xi^*) \int_{\ker \mathsf{C}} \mathrm{d}\boldsymbol{u}\, \delta\left[(\nabla\hat{\xi})\boldsymbol{u} - (\xi - \xi^*)\right] \int_{\mathbb{R}^{m-s}} \mathrm{d}\boldsymbol{y} \exp\left\{-\tfrac{1}{2}\beta\kappa[\mathsf{CB}\boldsymbol{y} + \boldsymbol{w}(\boldsymbol{u})]^\mathsf{T}[\mathsf{CB}\boldsymbol{y} + \boldsymbol{w}(\boldsymbol{u})]\right\} \\ &= \frac{I(\xi^*)}{\left|\det \mathsf{D}^\perp\right|^{1/2}} \left(\frac{2\pi}{\beta}\right)^{(m-s)/2} \int_{\ker \mathsf{C}} \mathrm{d}\boldsymbol{u}\, \delta\left[(\nabla\hat{\xi})\boldsymbol{u} - (\xi - \xi^*)\right] \exp\left[-\tfrac{1}{2}\beta\kappa\, \boldsymbol{w}(\boldsymbol{u})^\mathsf{T}\Pi\boldsymbol{w}(\boldsymbol{u})\right],\end{aligned} \tag{3.20}$$

where $\mathsf{D}^\perp$ is the diagonal matrix of the $m - s$ nonzero eigenvalues of $\mathsf{D}$ at the singularity and

$$\Pi = \mathsf{I}_m - \mathsf{CB}(\mathsf{B}^\mathsf{T}\mathsf{C}^\mathsf{T}\mathsf{CB})^{-1}\mathsf{B}^\mathsf{T}\mathsf{C}^\mathsf{T} \tag{3.21}$$

is a projection operator[4] that projects $\boldsymbol{w}(\boldsymbol{u})$ to the cokernel of $\mathsf{CB}$ (i.e., to $\ker \mathsf{B}^\mathsf{T}\mathsf{C}^\mathsf{T}$). Note that this operator would trivially have been the identity matrix if it were not for the cross-term $\boldsymbol{w}(\boldsymbol{u})^\mathsf{T}\mathsf{KC}\boldsymbol{v}$ in the energy expansion, Eq. (2.19), that couples the fast modes and the

[3] Changing $\boldsymbol{q} \to (\boldsymbol{u}, \boldsymbol{v})$ introduces a unit Jacobian factor when $\ker \mathsf{C}$ and $(\ker \mathsf{C})^\perp$ have orthonormal bases.

[4] A similar projection operator appears in the theory of polymerized membranes, once the in-plane components of the elastic free energy (i.e., the fast modes) are integrated out; see, e.g., Eqs. (6)–(8) of Ref. [124].

zero modes. Since the cokernel of CB is identical to the cokernel of C, it is spanned by an orthonormal basis of self stresses $\boldsymbol{\sigma}$, and one can write the action of $\Pi$ on $\boldsymbol{w}(\boldsymbol{u})$ as

$$\Pi \boldsymbol{w}(\boldsymbol{u}) = \sum_{\boldsymbol{\sigma}} \boldsymbol{\sigma} [\boldsymbol{\sigma} \cdot \boldsymbol{w}(\boldsymbol{u})]. \tag{3.22}$$

Using this in Eq. (3.20) and after a straightforward application of the coarea formula [Eq. (3.9)] we find to the lowest order, for $\xi \to \xi^*$,

$$\begin{aligned} \mathscr{P}(\xi) &\sim \frac{I(\xi^*)}{\left|\det \mathsf{D}^{\perp}\right|^{1/2}} \left(\frac{2\pi}{\beta}\right)^{(m-s)/2} \int_{\ker \mathsf{C}} \mathrm{d}\boldsymbol{u}\, \delta\left[(\nabla\hat{\xi})\boldsymbol{u} - (\xi - \xi^*)\right] \exp\left\{-\tfrac{1}{2}\beta\kappa \sum_{\boldsymbol{\sigma}} [\boldsymbol{\sigma}\cdot\boldsymbol{w}(\boldsymbol{u})]^2\right\} \\ &= \frac{I(\xi^*)}{\left|\det \mathsf{D}^{\perp} \det \nabla\hat{\xi} (\nabla\hat{\xi})^{\mathsf{T}}\right|^{1/2}} \left(\frac{2\pi}{\beta}\right)^{(m-s)/2} \int_{\Xi_\xi} \mathrm{d}\Theta(\boldsymbol{u}) \exp\left\{-\tfrac{1}{2}\beta\kappa \sum_{\boldsymbol{\sigma}} [\boldsymbol{\sigma}\cdot\boldsymbol{w}(\boldsymbol{u})]^2\right\}. \end{aligned} \tag{3.23}$$

Above, the integration domain $\Xi_\xi = (\nabla\hat{\xi})^{-1}(\xi - \xi^*) \cap \ker \mathsf{C}$, is a hyperplane of $(n-m+s)-(n-m) = s$ dimensions and $\mathrm{d}\Theta(\boldsymbol{u})$ is the surface measure on $\Xi_\xi$. After choosing a convenient parameterization for $\Xi_\xi$ in terms of the components of $\boldsymbol{u}$ in some basis of $\ker \mathsf{C}$, Eq. (3.23) becomes an $s$-dimensional integral involving the exponential of a quartic polynomial, which should converge if $\Sigma_{\boldsymbol{\sigma}} [\boldsymbol{\sigma}\cdot\boldsymbol{w}(\boldsymbol{u})]^2$ does not identically vanish in some region of $\Xi_\xi$ that extends to infinity (see below for an extended discussion on this). This completes the derivation of Eq. (3.23). In deriving Eq. (3.23), unlike in the derivation of Eq. (3.19), we have not looked at contributions to $\mathscr{P}(\xi)$ from the ground states on each branch of $\Sigma$ individually. Hence, Eq. (3.23) represents the collective contribution to the marginal density from all the branches that intersect to form the singularity at $\bar{\boldsymbol{q}}^*$.

*Convergence of the integral in Eq. (3.23)*

Before discussing the conditions that are required for the integral in Eq. (3.23) to converge, we digress slightly and discuss second-order rigidity, a frequently invoked notion in rigidity theory [23, 24]. In the notation that we have been using, a bar-joint framework is considered to be second-order rigid if there are no vector pairs $(\boldsymbol{u}, \boldsymbol{v})$ with $\boldsymbol{u} \in \ker \mathsf{C}$ and $\boldsymbol{v} \in (\ker \mathsf{C})^{\perp}$ such that[5]

$$\mathsf{C}\boldsymbol{v} + \boldsymbol{w}(\boldsymbol{u}) = 0. \tag{3.24}$$

For a singular configuration of the framework that supports nonzero self stresses $\boldsymbol{\sigma}$, second-order rigidity is equivalent to saying that there is no zero mode $\boldsymbol{u} \in \ker \mathsf{C}$ such that[6]

$$\boldsymbol{\sigma} \cdot \boldsymbol{w}(\boldsymbol{u}) = 0, \tag{3.25}$$

for all $\boldsymbol{\sigma} \in \ker \mathsf{C}^{\mathsf{T}}$. A well-known result is that a bar-joint framework is rigid to all orders if it is second-order rigid [24].

Clearly, a framework is not rigid (to any order) by definition and all tangents to the branches of the shape space $\Sigma$ at a singularity $\bar{\boldsymbol{q}}^*$ satisfy Eq. (3.25). Even though one can speak of tangent vectors to the branches of $\Sigma$ at $\bar{\boldsymbol{q}}^*$, the shape space $\Sigma$ itself ceases to be a

[5] Note that this equation is nothing but the Taylor expansion of the constraint map $f$ to the lowest order in $\boldsymbol{u}$ and $\boldsymbol{v}$.

[6] See, e.g., Corollary 5.2.2 of Ref. [24] or Corollaries 4.15–4.17 of Ref. [171].

smooth manifold at $\bar{\boldsymbol{q}}^*$. Hence, there is no well-defined tangent space at $\bar{\boldsymbol{q}}^*$ and it is common practice to consider instead the solution space of Eq. (3.25),

$$\mathscr{T} = \left\{ \boldsymbol{t} \in \ker \mathsf{C} : \boldsymbol{\sigma} \cdot \boldsymbol{w}(\boldsymbol{t}) = 0 \text{ for all } \boldsymbol{\sigma} \in \ker \mathsf{C}^\mathsf{T} \right\}, \tag{3.26}$$

which is the set of zero modes that preserve the constraints to second-order [18]. For this reason, $\mathscr{T}$ is often called the second-order tangent cone.[7] Even though the tangents to the branches of $\Sigma$ belong to $\mathscr{T}$, in general, there could be vectors in $\mathscr{T}$ that are not tangents. In such situations, the tangents cannot be resolved by considering the second-order tangent cone alone and higher-order analysis is necessary [104, 122]. To simplify things, from here on we assume that the tangents can be resolved at second order, i.e., every element of $\mathscr{T}$ is a tangent to the branches of the shape space $\Sigma$ at the singularity. For instance, in the cartoon in Fig. 3.2, $\mathscr{T}$ would be the union of the two tangents at the singularity.

Now, the integral that defines $\mathscr{P}(\xi)$ also includes a delta function $\delta[\hat{\xi}(\boldsymbol{q}) - \xi]$. After linearizing the CV map $\hat{\xi}$ around the singularity and integrating out the fast modes, the domain of integration in Eq. (3.23) becomes the hyperplane $\Xi_\xi = \ker \mathsf{C} \cap (\nabla\hat{\xi})^{-1}(\xi - \xi^*)$. At a singularity with $s$ self stresses $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \dots, \boldsymbol{\sigma}_s$, consider the map $g : \ker \mathsf{C} \to \mathbb{R}^{n-m+s}$ defined by $g(\boldsymbol{u}) = [(\nabla\hat{\xi})\boldsymbol{u} - (\xi - \xi^*), \boldsymbol{\sigma}_1 \cdot \boldsymbol{w}(\boldsymbol{u}), \boldsymbol{\sigma}_2 \cdot \boldsymbol{w}(\boldsymbol{u}), \dots, \boldsymbol{\sigma}_s \cdot \boldsymbol{w}(\boldsymbol{u})]$. In a basis of $\ker \mathsf{C}$, the equation $g(\boldsymbol{u}) = 0$ defines an exactly determined system of equations in $n - m + s$ unknowns, whose solutions (if they exist) are tangents that belong to $\Xi_\xi$. If this equation has isolated roots, then the term $\sum_{\boldsymbol{\sigma}} [\boldsymbol{\sigma} \cdot \boldsymbol{w}(\boldsymbol{u})]^2$ in the exponential of Eq. (3.23) is zero only for a discrete set of tangent vectors in $\Xi_\xi$ and is positive everywhere else, making the integral convergent. (In Fig. 3.2, this discrete set is composed of the tangents to the two branches, which when extended intersect the line $\Xi_\xi$, which is the domain of integration.) A necessary condition for the equation $g(\boldsymbol{u}) = 0$ to have isolated roots in $\Xi_\xi$ is that $(\nabla\hat{\xi})\boldsymbol{t} \neq 0$ for all $\boldsymbol{t} \in \mathscr{T}$. Given how $g(\boldsymbol{u})$ is a system of $n - m$ linear and $s$ quadratic equations, a stronger general condition that ensures this eludes us at present. On the other hand, if there is a tangent $\boldsymbol{t}$ such that $(\nabla\hat{\xi})\boldsymbol{t} = 0$, the integral diverges. Such cases are pathological and indicate a poor CV choice. After all, if $\boldsymbol{t}$ is tangent to $\Sigma$, then it is a slow mode that corresponds to a shape change in the framework, and one would definitely want the value of the CV to change along it.

To conclude, the convergence of Eq. (3.23) relies on the term $\sum_{\boldsymbol{\sigma}} [\boldsymbol{\sigma} \cdot \boldsymbol{w}(\boldsymbol{u})]^2$ in the exponential vanishing only for a finite number of isolated points in the integration domain $\Xi_\xi$. Two necessary conditions required for this are: (i) tangents to the branches of the shape space at the singularity can be resolved at second order and form the solution space $\mathscr{T}$ of Eq. (3.25), and (ii) the CV map is such that $(\nabla\hat{\xi})\boldsymbol{t} \neq 0$ for all $\boldsymbol{t} \in \mathscr{T}$.

#### *Scaling of the marginal density*

To see how the marginal density $\mathscr{P}(\xi^*)$ at a singular value $\xi^*$ scales with $\beta$ and $\kappa$, we first choose a basis for $\ker \mathsf{C}$ so that $\boldsymbol{u} = \mathsf{A}\boldsymbol{x}$, where $\boldsymbol{x} \in \mathbb{R}^{n-m+s}$ represents the components of $\boldsymbol{u}$ in the chosen basis and $\mathsf{A}$ is the associated change-of-basis matrix. Now, the $s$-dimensional hyperplane $\Xi_\xi$ formed by the intersection of the linearized CV level set and $\ker \mathsf{C}$ is defined

[7] See, e.g., Section 4.2 of Ref. [173]; also see the related discussions in Refs. [104, 122, 123]. Tangent cones themselves were originally introduced by Whitney [169] to study tangents to analytic varieties.

by $\nabla\hat{\xi}\mathsf{A}\boldsymbol{x}=0$, which is a set of $n-m$ homogeneous linear equations in $n-m+s$ variables. Without loss of generality, let us assume that we can solve these equations to obtain the last $n-m$ components of $\boldsymbol{x}$ in terms of its first $s$ components $\tilde{\boldsymbol{x}}=(x_1,x_2,\dots,x_s)$. This enables us to parameterize the hyperplane $\Xi_\xi$ using $\tilde{\boldsymbol{x}}$. As each component of the vector $\boldsymbol{w}(\boldsymbol{u})$ is a quadratic form in $\boldsymbol{x}$, after the elimination step, the term $\sum_{\boldsymbol{\sigma}}[\boldsymbol{\sigma}\cdot\boldsymbol{w}(\boldsymbol{u})]^2$ in the exponential of Eq. (3.23) becomes a homogeneous quartic polynomial $\widetilde{U}(\tilde{\boldsymbol{x}})$. A rescaling of the components $\tilde{\boldsymbol{x}}\to(\beta\kappa)^{-1/4}\tilde{\boldsymbol{x}}$, changes the surface measure from $\mathrm{d}\Theta(\tilde{\boldsymbol{x}})\to(\beta\kappa)^{-s/4}\mathrm{d}\Omega(\tilde{\boldsymbol{x}})$ and turns $\widetilde{U}(\tilde{\boldsymbol{x}})\to(\beta\kappa)^{-1}\widetilde{U}(\tilde{\boldsymbol{x}})$, yielding

$$\mathscr{P}(\xi^*)\sim\frac{I(\xi^*)}{\left|\kappa^{s/2}\det\mathsf{D}^\perp\det\nabla\hat{\xi}(\nabla\hat{\xi})^\mathsf{T}\right|^{1/2}}\left(\frac{2\pi}{\beta}\right)^{m/2-s/4}\int_{\Xi_\xi}\mathrm{d}\Theta(\tilde{\boldsymbol{x}})\exp\left[-\tfrac{1}{2}\widetilde{U}(\tilde{\boldsymbol{x}})\right]. \tag{3.27}$$

Clearly, the above integral is purely geometric in nature and all $\beta$ dependence has been extracted. Finally, noting that $\det\mathsf{D}^\perp\sim\kappa^{m-s}$ we see that the marginal density $\mathscr{P}(\xi^*)\sim(\beta\kappa)^{-m/2+s/4}$. It should be emphasized that we can only do this analysis for $\xi=\xi^*$. For other values of $\xi$ close to $\xi^*$, the hyperplane $\Xi_\xi$ is defined by the inhomogeneous equation $\nabla\hat{\xi}\mathsf{A}\boldsymbol{x}=\xi-\xi^*$, which makes $\widetilde{U}(\tilde{\boldsymbol{x}})$ a similarly inhomogeneous quartic polynomial, making a rescaling argument impossible.

The marginal density for regular values of the CV, which scales like $\mathscr{P}(\xi)\sim(\beta\kappa)^{-m/2}$ [Eq. (3.19)], is always subdominant to the marginal density at a singular value $\xi^*$, which scales like $\mathscr{P}(\xi^*)\sim(\beta\kappa)^{-m/2+s/4}$ for $s>0$. Hence, we see that the softening of the framework at a singularity causes an energetic free-energy barrier to develop between regular and singular values of the CV, with a temperature/stiffness dependence $\sim\ln\beta\kappa$. In comparison, the free-energy barriers between singular values of the CV are independent of $\beta$ and $\kappa$, and depend only on geometric parameters if the corresponding configurations have the same number of self stresses $s$. Although this might lead us to conclude that such barriers are entropic in origin, note that there would be energetic barriers along most realizable transition paths separating these configurations.

The stark difference in the asymptotic scaling of $\mathscr{P}(\xi)$ at a singular value $\xi=\xi^*$ and for values farther from it shows that the true scaling and behavior of $\mathscr{P}(\xi)$ for intermediate values of $\xi$ is nontrivial. A natural question is then: for what values of $\xi$ would the harmonic and quartic approximations capture the true behavior of $\mathscr{P}(\xi)$? Equation (3.19), derived using the harmonic approximation and a direct application of Laplace's method, is only accurate so long as the lowest nonzero eigenvalue $\omega_{\min}(\xi)$ of the dynamical matrix $\mathsf{D}$ is such that $\beta\omega_{\min}(\xi)$ is very large. This is also what causes it to break down as we approach a singularity, near which $\omega_{\min}(\xi)$ monotonically[8] decreases to zero as $\xi\to\xi^*$. This also implies that as $\beta$ becomes larger, the harmonic approximation starts capturing the true behavior of the marginal density for a larger range of $\xi$ values. For very large $\beta$, the range of validity of the quartic approximation is also bound to increase as the errors in the approximation become small. This means that, for large $\beta$, we expect to see some amount of overlap in the marginal density estimates using Eqs. (3.19) and (3.23). (For instance, see the example free-energy curves in Figs. 3.4 and 3.8(b),

[8] For one-dimensional shape spaces, using Rayleigh–Schrödinger perturbation theory [30] and considering the dynamical matrix at the singularity as the "unperturbed Hamiltonian", we can show that $\omega_{\min}\sim(\xi-\xi^*)^2$ for $\xi\to\xi^*$.

which we shall discuss in more detail later.) We expect the exact nature of the overlap to be problem-specific with a strong dependence on $\beta$ and we leave a more thorough analysis for future work.

We now use our formalism to find the free-energy profiles of two example frameworks with one-dimensional shape spaces with isolated singularities and compare them with results from Monte Carlo simulations. We will also use our formalism to analyze the five-bar linkage, a framework with a two-dimensional shape space. Motivated by typical DNA origami structures that have lengths in the range of a few hundred nanometers with stiffness in the range 0.1–1 pN/nm [76], we choose a nondimensional inverse temperature of $\beta = 10^4$ and use a potential of the form $\phi_i(\ell_i) = (\ell_i^2 - \bar{\ell}_i^2)^2/(8\bar{\ell}_i^2)$ so that $\phi_i''(\bar{\ell}_i) = \kappa = 1$. Further details on the simulations are given in Appendix 3.A.

## 3.3 Example: planar four-bar linkage

The four-bar linkage we consider (Fig. 3.3) is made out of two sets of bars of lengths $a$ and $\lambda a$, where $\lambda > 0$ is a dimensionless aspect ratio. For $\lambda \neq 1$, the linkage has shape-space singularities at $\theta_1 = 0$ and $\theta_1 = \pm\pi$ where the bars become collinear and support a state of self stress.[9] The shape space can be fully parameterized using the angle $\theta_1$, which we use as our CV.

### 3.3.1 Body frame

Rigid motions can be integrated out by transforming to a local Cartesian coordinate system (body frame) attached to the four-bar linkage with joint 1 at the origin and bar 1–2 lying along the horizontal axis as shown in Fig. 3.3. Let $(r_{i1}, r_{i2})$, $i = 1,2,3,4$ be the coordinates of the four joints in the lab frame. The configuration vector $\boldsymbol{q} \in \mathbb{R}^5$ of the linkage in the body frame is $\boldsymbol{q} = (q_1, q_2, \ldots, q_5)$. Also, two translational coordinates $x_1, x_2$ specify the position of joint 1, and an orientational coordinate $\eta$, which is the angle between the horizontal axes of the lab and body frames, gives the overall rotation of the linkage. The explicit coordinate transformation $\boldsymbol{r} \to (x_1, x_2, \eta, \boldsymbol{q})$ is given by

$$\begin{aligned} \begin{pmatrix} r_{11} \\ r_{12} \end{pmatrix} &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, & \begin{pmatrix} r_{21} \\ r_{22} \end{pmatrix} &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \mathsf{R}(\eta) \begin{pmatrix} q_1 \\ 0 \end{pmatrix}, \\ \begin{pmatrix} r_{31} \\ r_{32} \end{pmatrix} &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \mathsf{R}(\eta) \begin{pmatrix} q_2 \\ q_3 \end{pmatrix}, & \begin{pmatrix} r_{41} \\ r_{42} \end{pmatrix} &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \mathsf{R}(\eta) \begin{pmatrix} q_4 \\ q_5 \end{pmatrix}. \end{aligned} \tag{3.28}$$

Here $\mathsf{R}(\eta)$ is the rotation matrix in $\mathbb{R}^2$. Dropping the constant factor that one gets after integrating over $x_1, x_2$, and $\eta$, the overall Jacobian factor involved in the transformation given by Eq. (3.28) is

$$I(\boldsymbol{q}) = |q_1|. \tag{3.29}$$

[9] For simplicity, we do not discuss the square four-bar linkage with $\lambda = 1$ here as it has additional singularities at $(\theta_1, \theta_2) = (0, \pm\pi)$ [176].

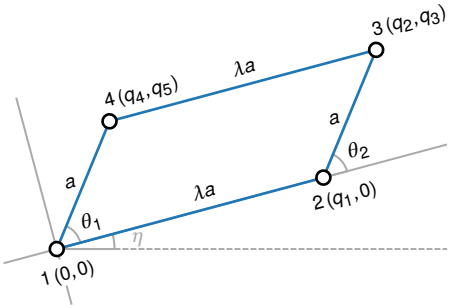


**Figure 3.3.** Body frame on the planar four-bar linkage with joint 1 at the origin and bar 1–2 lying along the horizontal axis. The angle between the horizontal axes of the body and the lab frame is $\eta$.

### 3.3.2 Branch parameterization

The four-bar linkage admits two modes of deformations, giving its shape space a branched appearance. On the parallel branch the angles $\theta_1$ and $\theta_2$ are equal, whereas on the twisted branch they have a nonlinear relationship with opposite signs. To find the exact relationship between $\theta_1$ and $\theta_2$ on the twisted branch, we first write down the constraint equation for bar 3–4 (see Fig. 3.3) of the four-bar linkage:

$$[\lambda a + a(\cos\theta_2 - \cos\theta_1)]^2 + (a\sin\theta_2 - a\sin\theta_1)^2 = \lambda^2 a^2. \tag{3.30}$$

Assuming that the aspect ratio $\lambda \neq 1$, this equation can be simplified and factorized in two different ways to get

$$[\cos\theta_2 - \cos\theta_1]\left[\cos\theta_2 - \frac{(1+\lambda^2)\cos\theta_1 - 2\lambda}{1 - 2\lambda\cos\theta_1 + \lambda^2}\right] = 0, \tag{3.31a}$$

$$[\sin\theta_2 - \sin\theta_1]\left[\sin\theta_2 - \frac{(1-\lambda^2)\sin\theta_1}{1 - 2\lambda\cos\theta_1 + \lambda^2}\right] = 0. \tag{3.31b}$$

The solutions to the above equations tell us the relationship between the "input" angle $\theta_1$ and the "output" angle $\theta_2$ on the two branches, e.g., on the parallel branch

$$\cos\theta_2 = \cos\theta_1, \qquad \sin\theta_2 = \sin\theta_1; \tag{3.32}$$

and on the twisted branch

$$\cos\theta_2 = \frac{(1+\lambda^2)\cos\theta_1 - 2\lambda}{1 - 2\lambda\cos\theta_1 + \lambda^2}, \qquad \sin\theta_2 = \frac{(1-\lambda^2)\sin\theta_1}{1 - 2\lambda\cos\theta_1 + \lambda^2}, \tag{3.33}$$

which is what we intended to find. We can express a point $\bar{\boldsymbol{q}}$ in the shape space $\Sigma \subset \mathbb{R}^5$ of the four-bar linkage in terms of the angles $\theta_1$ and $\theta_2$ as $\bar{\boldsymbol{q}} = [\lambda a, a(\lambda + \cos\theta_2), a\sin\theta_2, a\cos\theta_1, a\sin\theta_1]$. Using Eqs. (3.32) and (3.33) we find two parameterizations for $\Sigma$, namely $\psi_+ : \mathbb{R} \to \mathbb{R}^5$ (parallel branch) and $\psi_- : \mathbb{R} \to \mathbb{R}^5$ (twisted branch), defined by

$$\psi_+(\theta_1) = [\lambda a, a(\lambda + \cos\theta_1), a\sin\theta_1, a\cos\theta_1, a\sin\theta_1], \tag{3.34a}$$

$$\psi_-(\theta_1) = \left[\lambda a, a\frac{(1-\lambda^2)(\cos\theta_1 - 1)}{1-2\lambda\cos\theta_1+\lambda^2}, a\frac{(1-\lambda^2)\sin\theta_1}{1-2\lambda\cos\theta_1+\lambda^2}, a\cos\theta_1, a\sin\theta_1\right]. \tag{3.34b}$$

The above equations define two curves in $\mathbb{R}^5$ parameterized by the angle $\theta_1$. The induced metric on these two curves can be readily computed as

$$(\nabla\psi_+)^\mathsf{T}\nabla\psi_+ = \|\partial\psi_+/\partial\theta_1\|^2 = 2a^2, \tag{3.35a}$$

$$(\nabla\psi_-)^\mathsf{T}\nabla\psi_- = \|\partial\psi_-/\partial\theta_1\|^2 = \frac{2a^2[1-2\lambda(1-\lambda\cos\theta_1+\lambda^2)\cos\theta_1+\lambda^4]}{(1-2\lambda\cos\theta_1+\lambda^2)^2}. \tag{3.35b}$$

### 3.3.3 Marginal probability densities

*Regular values*

Now that we have explicit parameterizations of the two branches of the four-bar linkage in terms of the CV $\theta_1$, we can find the marginal density $\mathscr{P}(\theta_1)$ at regular values of $\theta_1$, far from the singular values, i.e., for $0 \ll |\theta_1| \ll \pi$. Note that for each value of $\theta_1$, there are two ground states—one on the parallel branch and one on the twisted branch. This means that we need to separately find the contributions of these ground states using Eq. (3.19) and then add them together to obtain $\mathscr{P}(\theta_1)$.

Starting with constraint function for bar 1–2 and going around the linkage in counterclockwise order, we find the constraint map $f:\mathbb{R}^5\to\mathbb{R}^4$ in the body frame to be

$$f(\boldsymbol{q}) = \left[\frac{q_1^2-\lambda^2a^2}{2\lambda a}, \frac{(q_2-q_1)^2+q_3^2-a^2}{2a}, \frac{(q_4-q_2)^2+(q_5-q_3)^2-\lambda^2a^2}{2\lambda a}, \frac{q_4^2+q_5^2-a^2}{2a}\right], \tag{3.36}$$

and the compatibility matrix $\mathsf{C}$ to be

$$\mathsf{C} = \nabla f = a^{-1}\begin{pmatrix} \lambda^{-1}q_1 & 0 & 0 & 0 & 0 \\ (q_1-q_2) & q_2-q_1 & q_3 & 0 & 0 \\ 0 & \lambda^{-1}(q_2-q_4) & \lambda^{-1}(q_3-q_5) & \lambda^{-1}(q_4-q_2) & \lambda^{-1}(q_5-q_3) \\ 0 & 0 & 0 & q_4 & q_5 \end{pmatrix}. \tag{3.37}$$

At regular points $\mathsf{C}$ has full rank, which implies that the dynamical matrix $\mathsf{D} = \mathsf{C}^\mathsf{T}\mathsf{K}\mathsf{C}$ and the matrix $\mathsf{K}\mathsf{C}\mathsf{C}^\mathsf{T}$ have the same nonzero eigenvalues. This gives $\det\mathsf{D}^\perp = \det\mathsf{K}\mathsf{C}\mathsf{C}^\mathsf{T} = \kappa^4\det\mathsf{C}\mathsf{C}^\mathsf{T}$. Inserting the parameterizations from Eqs. (3.34a) and (3.34b) into the compatibility matrix we compute $\det\mathsf{D}^\perp$ along the two branches as

$$\det\mathsf{D}_+^\perp = 2\kappa^4\sin^2\theta_1, \tag{3.38a}$$

$$\det\mathsf{D}_-^\perp = \frac{2\kappa^4\sin^2\theta_1[1-2\lambda(1-\lambda\cos\theta_1+\lambda^2)\cos\theta_1+\lambda^4]}{(1-2\lambda\cos\theta_1+\lambda^2)^2}. \tag{3.38b}$$

The asymptotic marginal density is then

$$\begin{aligned}\mathscr{P}(\theta_1) &\sim I(\theta_1)\left(\frac{2\pi}{\beta}\right)^2\left[\left|\frac{\det(\nabla\psi_+)^\mathsf{T}\nabla\psi_+}{\det\mathsf{D}_+^\perp}\right|^{1/2} + \left|\frac{\det(\nabla\psi_-)^\mathsf{T}\nabla\psi_-}{\det\mathsf{D}_-^\perp}\right|^{1/2}\right] \\ &= 2\lambda a^2\left(\frac{2\pi}{\beta\kappa}\right)^2|\sin\theta_1|^{-1},\end{aligned} \tag{3.39}$$

where we have used $I(\theta_1) = |q_1| = \lambda a$ [Eq. (3.29)] and the expressions for the induced metrics [Eqs. (3.35a) and (3.35b)] computed earlier.

*Singular values*

The four-bar linkage has shape-space singularities at $\theta_1 = 0$ and $\theta_1 = \pm\pi$ corresponding to configurations where the bars are collinear. Let us first look at the singularity at $\theta_1 = 0$, where the configuration vector $\bar{\boldsymbol{q}}^* = [\lambda a, a(\lambda+1), 0, a, 0]$. The compatibility and dynamical matrices at this point are

$$\mathsf{C} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \mathsf{D} = \mathsf{C}^\mathsf{T}\mathsf{K}\mathsf{C} = \kappa \begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.40}$$

The dynamical matrix has nonzero eigenvalues $(2+\sqrt{2})\kappa$, $2\kappa$, and $(2-\sqrt{2})\kappa$, which gives $\det \mathsf{D}^\perp = 4\kappa^3$. Also, the Hessian matrices of the four constraint functions $f_i$ at the singularity are

$$\begin{aligned} \nabla\nabla f_1 &= (\lambda a)^{-1} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, & \nabla\nabla f_2 &= a^{-1} \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \\ \nabla\nabla f_3 &= (\lambda a)^{-1} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \end{pmatrix}, & \nabla\nabla f_4 &= a^{-1} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{3.41}$$

As we remarked in Section 3.2.2, when using Eq. (3.23) to find the marginal density $\mathscr{P}(\theta_1)$, we will not make use of the parameterizations we derived earlier [Eqs. (3.34a) and (3.34b)]. Instead, we need to first linearize the CV map at the singularity. Since the CV we have chosen for the four-bar linkage is the angle $\theta_1$, the CV map that computes $\theta_1$ from the configuration vector $\boldsymbol{q}$ is $\hat{\theta}_1(\boldsymbol{q}) = \tan^{-1}(q_5/q_4)$.[10] This map has the Jacobian $\nabla\hat{\theta}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & a^{-1} \end{pmatrix}$ at the singularity. Direct inspection reveals that $(\nabla\hat{\theta}_1)\boldsymbol{v} = 0$ for all fast modes $\boldsymbol{v} \in (\ker \mathsf{C})^\perp$, justifying the usage of Eq. (3.23) to find $\mathscr{P}(\theta_1)$ when $\theta_1 \to 0$. This would not have been the case if, for instance, we had chosen the coordinate $q_4$ of joint 4 of the four-bar linkage (see Fig. 3.3), as our CV. In such a case $\nabla q_4 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \end{pmatrix}$ and the fast modes, all of which are along the collinear bars with a nonzero fourth component, do not satisfy $(\nabla q_4)\boldsymbol{v} = 0$. Incidentally, in this case, the tangent vectors $\boldsymbol{t}$ at the singularity are such that $(\nabla q_4)\boldsymbol{t} = 0$, which also makes

[10] To be more rigorous, we should be using the two-argument variant of the inverse tangent, sometimes denoted as $\mathrm{atan2}(q_5, q_4)$ in numerical software, so that $\theta_1$ is in $(-\pi, \pi]$ instead of $(-\pi/2, \pi/2)$. This is not an issue for the linearization since $\theta_1$ is small.

$q_4$ a poor choice as the CV since it does not capture the slow modes along $\boldsymbol{t}$. Continuing with $\theta_1$ as our CV, to evaluate the integral in Eq. (3.23), we choose the vector $\boldsymbol{u} \in \ker \mathsf{C}$ to be $\boldsymbol{u} = (0,0,q_3,0,q_5)$, so that the vector $\boldsymbol{w}(\boldsymbol{u})$ is

$$\begin{aligned}\boldsymbol{w}(\boldsymbol{u}) &= \left(\tfrac{1}{2}\boldsymbol{u}^\mathsf{T}\nabla\nabla f_1 \boldsymbol{u}, \tfrac{1}{2}\boldsymbol{u}^\mathsf{T}\nabla\nabla f_2 \boldsymbol{u}, \tfrac{1}{2}\boldsymbol{u}^\mathsf{T}\nabla\nabla f_3 \boldsymbol{u}, \tfrac{1}{2}\boldsymbol{u}^\mathsf{T}\nabla\nabla f_4 \boldsymbol{u}\right) \\ &= \left[0,\, q_3^2/(2a),\, (q_3-q_5)^2/(2\lambda a),\, q_5^2/(2a)\right].\end{aligned} \tag{3.42}$$

There is only one self stress $\boldsymbol{\sigma} \in \ker \mathsf{C}^\mathsf{T}$ at the singularity, and it is $\boldsymbol{\sigma} = (-1/2,-1/2,1/2,1/2)$. Using this in Eq. (3.23) along with $(\nabla\hat{\theta}_1)\boldsymbol{u} = a^{-1}q_5$ and the fact that $\det \mathsf{D}^\perp = 4\kappa^3$, we get[11]

$$\begin{aligned}\mathscr{P}(\theta_1) &\sim \frac{\lambda a}{2}\left(\frac{2\pi}{\beta\kappa}\right)^{3/2}\left(\int_{-\infty}^{\infty} \mathrm{d}q_3\,\mathrm{d}q_5\,\delta(a^{-1}q_5-\theta_1)\right. \\ &\qquad \left.\times \exp\left\{-\frac{\beta\kappa}{32\lambda^2a^2}\left[q_3-q_5\right]^2\left[(\lambda-1)q_3+(\lambda+1)q_5\right]^2\right\}\right) \\ &= \frac{\lambda a^2}{2}\left(\frac{2\pi}{\beta\kappa}\right)^{3/2}\int_{-\infty}^{\infty}\mathrm{d}q_3\,\exp\left\{-\frac{\beta\kappa}{32\lambda^2a^2}\left[q_3-a\theta_1\right]^2\left[(\lambda-1)q_3+(\lambda+1)a\theta_1\right]^2\right\}.\end{aligned} \tag{3.43}$$

At this point, it is useful to revisit the convergence criteria for Eq. (3.23), i.e., the requirement that the term $\sum_{\boldsymbol{\sigma}}[\boldsymbol{\sigma}\cdot\boldsymbol{w}(\boldsymbol{u})]^2$ in the exponential of Eq. (3.23) must only have isolated zeros. In the above equation, this term is $[q_3-a\theta_1]^2[(\lambda-1)q_3+(\lambda+1)a\theta_1]^2/(16\lambda^2a^2)$, which has isolated zeros $q_3 = a\theta_1$ and $q_3 = (1+\lambda)a\theta_1/(1-\lambda)$, consistent with the convergence requirement.

To evaluate the integral in Eq. (3.43), we symmetrize the expression in the exponential by changing variables $q_3 \to q_3 - (\lambda-1)^{-1}a\theta_1$, which yields

$$\begin{aligned}\mathscr{P}(\theta_1) &\sim \frac{\lambda a^2}{2}\left(\frac{2\pi}{\beta\kappa}\right)^{3/2}\int_{-\infty}^{\infty}\mathrm{d}q_3\,\exp\left\{-\frac{\beta\kappa(\lambda-1)^2}{32\lambda^2a^2}\left[q_3^2-\frac{\lambda^2a^2\theta_1^2}{(\lambda-1)^2}\right]^2\right\} \\ &= \lambda a^2\frac{(2\pi)^{3/2}}{(\beta\kappa)^{7/4}}\sqrt{\frac{\lambda a}{|\lambda-1|}}\exp\left[-\frac{\beta\kappa\lambda^2a^2\theta_1^4}{32(\lambda-1)^2}\right]\int_0^\infty \mathrm{d}x\,x^{-1/2}\exp\left(-\tfrac{1}{2}x^2+\frac{\sqrt{\beta\kappa}\lambda a\theta_1^2}{4|\lambda-1|}x\right) \\ &= \lambda a^2\frac{(2\pi)^{3/2}}{(\beta\kappa)^{7/4}}\sqrt{\frac{\pi\lambda a}{|\lambda-1|}}\exp\left[-\frac{\beta\kappa\lambda^2a^2\theta_1^4}{64(\lambda-1)^2}\right]D_{-1/2}\left(-\frac{\sqrt{\beta\kappa}\lambda a\theta_1^2}{4|\lambda-1|}\right),\end{aligned} \tag{3.44}$$

where $D_{-1/2}(\cdot)$ is the parabolic cylinder function [129, Eq. 12.5.1].

For the singularity at $\theta_1 = \pm\pi$, we have $\tilde{\boldsymbol{q}}^* = [\lambda a, (\lambda-1)a, 0, -a, 0]$ and we proceed with a similar calculation. The dynamical matrix and the Hessians of the constraint functions at this point are identical to those at $\theta_1 = 0$. Hence, we choose the vector $\boldsymbol{u}$ as before, yielding the same $\boldsymbol{w}(\boldsymbol{u})$ as in Eq. (3.42). However, the self stress at $\theta_1 = \pm\pi$ is different and it is $\boldsymbol{\sigma} = (-1/2,-1/2,-1/2,1/2)$. Using these results, we can evaluate the integral in Eq. (3.23) as before to get

[11] From the argument of the Dirac delta function, we see that the "hyperplane" $\Xi_\xi$ in Eq. (3.23) is just the line along $q_5 = a\theta_1$ in the $q_3$-$q_5$ plane.

$$\begin{aligned}\mathscr{P}(\theta_1) &\sim \frac{\lambda a^2}{2}\left(\frac{2\pi}{\beta\kappa}\right)^{3/2}\int_{-\infty}^{\infty} \mathrm{d}q_3 \exp\left\{-\frac{\beta\kappa(\lambda+1)^2}{32\lambda^2 a^2}\left[q_3^2 - \frac{\lambda^2 a^2(\pi-|\theta_1|)^2}{(\lambda+1)^2}\right]^2\right\} \\ &= \lambda a^2 \frac{(2\pi)^{3/2}}{(\beta\kappa)^{7/4}}\sqrt{\frac{\pi\lambda a}{\lambda+1}}\exp\left[-\frac{\beta\kappa\lambda^2 a^2(\pi-|\theta_1|)^4}{64(\lambda+1)^2}\right] D_{-1/2}\left[-\frac{\sqrt{\beta\kappa}\lambda a(\pi-|\theta_1|)^2}{4(\lambda+1)}\right].\end{aligned} \tag{3.45}$$

Note that the marginal densities at the singular values of $\theta_1$, i.e., $\mathscr{P}(0)$ and $\mathscr{P}(\pm\pi)$, scale as $(\beta\kappa)^{-7/4}$. Since $m = 4$ and the number of self stresses $s = 1$ at the singularities, this is consistent with the general scaling $\mathscr{P}(\xi^*) \sim (\beta\kappa)^{-(m/2-s/4)}$ for singular values $\xi^*$ of the CV [Eq. (3.27)].

### 3.3.4 Free energy

Using the marginal densities we have found for various regimes of $\theta_1$ [Eqs. (3.39), (3.44), and (3.45)] we see that the free-energy difference of the four-bar linkage $\Delta\mathscr{A}(\theta_1) = \mathscr{A}(\theta_1) - \mathscr{A}(0) = -\beta^{-1}\log\mathscr{P}(\theta_1) + \beta^{-1}\log\mathscr{P}(0)$ takes the form

$$\Delta\mathscr{A}(\theta_1) \sim \begin{cases} \beta^{-1}\left\{X^2\theta_1^4 - \ln\left[\dfrac{D_{-1/2}(-2X\theta_1^2)}{D_{-1/2}(0)}\right]\right\}, & \theta_1 \to 0 \\ \beta^{-1}\ln\left[X^{1/2}D_{-1/2}(0)|\sin\theta_1|\right], & 0 \ll |\theta_1| \ll \pi \\ \beta^{-1}\left(X^2Y^2(\pi-|\theta_1|)^4 - \ln\left\{Y^{1/2}\dfrac{D_{-1/2}\left[-2XY(\pi-|\theta_1|)^2\right]}{D_{-1/2}(0)}\right\}\right), & |\theta_1| \to \pi \end{cases} \tag{3.46}$$

where $X$ and $Y$ are positive dimensionless terms independent of $\theta_1$ and defined by

$$X = \frac{\sqrt{\beta\kappa}\lambda a}{8|\lambda-1|}, \qquad Y = \left|\frac{\lambda-1}{\lambda+1}\right|. \tag{3.47}$$

From Eq. (3.46), we also find the free-energy difference between the singular values $\theta_1 = 0$ and $\theta_1 = \pm\pi$ to be

$$\mathscr{A}(\pm\pi) - \mathscr{A}(0) \sim -\beta^{-1}\log Y^{1/2} = \tfrac{1}{2}\beta^{-1}\ln\left|\frac{\lambda+1}{\lambda-1}\right|, \tag{3.48}$$

which is a purely geometric quantity. This is exactly what we expect based on how the marginal density scales at a singular value [Eq. (3.27)], which shows that the free-energy difference must be a purely geometric quantity if the singular states support the same number of self stresses $s$. A comparison between the numerical results and asymptotic expressions in Eq. (3.46) shows excellent agreement for all values of $\theta_1$ (Fig. 3.4).

Clearly, the free-energy profile has minima for values of $\theta_1$ close to 0 or $\pm\pi$, corresponding to shape-space singularities. But a more assiduous reader may have noticed that the free-energy curves actually have double-well minima around $\theta_1 = 0$ and $\theta_1 = \pm\pi$. Also, there is an asymmetry in these curves—the free energies of the four-bar linkage for $\theta_1$ close to 0 is lower than $\theta_1$ close to $\pm\pi$. In the next two subsections, we will try to understand these two features.

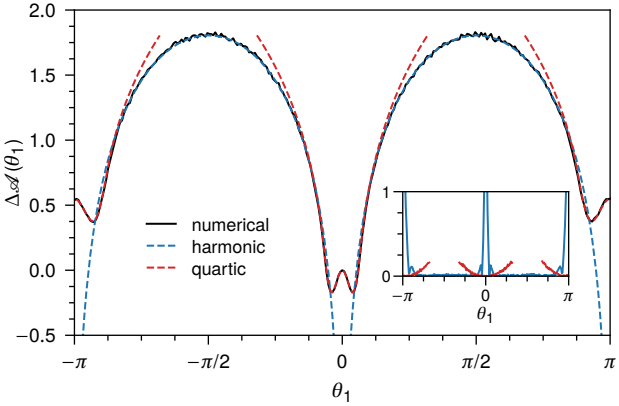


**Figure 3.4.** Free-energy difference $\Delta\mathscr{A}(\theta_1)$ of a four-bar linkage with parameters $a = 1$ and $\lambda = 2$ in units of $\beta^{-1}$ at $\beta = 10^4$. The inset shows the absolute errors between the numerical and asymptotic results using the harmonic and quartic approximations [Eq. (3.46)].

*Double-well structures*

To intuitively see why the free-energy landscape of the four-bar linkage has a double-well structure centered around the singular values, consider the projection of its total energy in the space of the angles $\theta_1$ and $\theta_2$. The projected energy can be found by setting the lengths of all bars equal to their natural lengths except for bar 3–4, which we assume to have an energy $\phi(\ell) = \kappa(\ell^2 - \bar{\ell}^2)^2/(8\bar{\ell}^2)$. (The exact form of the energy is irrelevant as long as it has a minimum at $\ell = \bar{\ell}$.) Writing $\ell$ in terms of the angles $\theta_1, \theta_2$ we get

$$\phi(\theta_1,\theta_2) = \frac{\kappa a^2}{8\lambda^2}\left[(\lambda + \cos\theta_2 - \cos\theta_1)^2 + (\sin\theta_2 - \sin\theta_1)^2 - \lambda^2\right]^2. \tag{3.49}$$

Figure 3.5 shows the level sets of $\phi(\theta_1,\theta_2)$ near $(\theta_1,\theta_2) = (0,0)$. For $|\theta_1| > 0$, there are two ground states in the CV level set $\hat{\theta}_1^{-1}(\theta_1)$, which is a straight line parallel to the $\theta_2$ axis in the $\theta_1$–$\theta_2$ space. Because of the extra softness of the linkage near the singularity, this means that for small nonzero values of $\theta_1$, there are more thermodynamically favorable states in $\hat{\theta}_1^{-1}(\theta_1)$ compared to $\hat{\theta}_1^{-1}(0)$, which is the CV level set at the singular value $\theta_1 = 0$. This is evidenced by the fact that for any given value $E > 0$, there are more points in the energy sublevel set $\{(\theta_1,\theta_2) : \phi(\theta_1,\theta_2) \le E\}$ along small nonzero values of $\theta_1$ than $\theta_1 = 0$ (see Fig. 3.5). The net increase in the number of thermodynamically favorable states near small nonzero values of $\theta_1$ lowers the free energy at those values, giving the landscape a double-well appearance.

We also remark that since an asymptotic expression for the free energy is available for the four-bar linkage, we can explicitly show the double-well nature of $\Delta\mathscr{A}(\theta_1)$ and find the locations of its minima. For example, expanding Eq. (3.44) in $\theta_1$ around $\theta_1 = 0$, we see that

$$\Delta\mathscr{A}(\theta_1) \sim \left[1 + 8\pi^2\Gamma^{-4}\left(\tfrac{1}{4}\right)\right] X^2\theta_1^4 - 4\pi\Gamma^{-2}\left(\tfrac{1}{4}\right) X\theta_1^2, \tag{3.50}$$

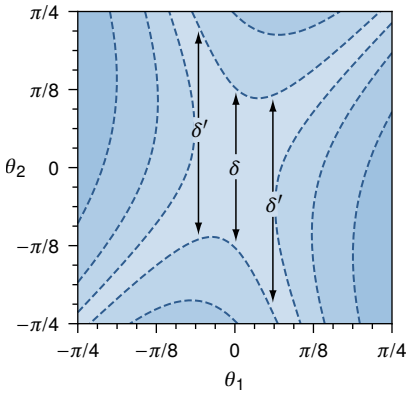


**Figure 3.5.** The level sets of the energy in Eq. (3.49) around $(\theta_1,\theta_2) = (0,0)$ for $a = \kappa = 1$ and $\lambda = 2$. The width of the energy sublevel sets is larger along small nonzero values of $\theta_1$ compared to $\theta_1 = 0$ (e.g., $\delta' > \delta$).

which is the equation for a double well, with $\Gamma(\cdot)$ being the gamma function. Expanding $\Delta\mathscr{A}(\theta_1)$ to $\mathcal{O}(|\theta_1|^6)$ we see that the two double-well minima are approximately at[12]

$$\theta_1^{\min} \approx \pm\sqrt{\frac{\Gamma^2\left(\frac{1}{4}\right)\left\{8\pi^2+\Gamma^4\left(\frac{1}{4}\right)-\sqrt{\left[8\pi^2+\Gamma^4\left(\frac{1}{4}\right)\right]^2-(16\pi^2)^2}\right\}}{64\pi^3 X}}. \tag{3.51}$$

Since $X \sim \sqrt{\beta}$, this also shows that as $\beta$ increases, $\theta_1^{\min}$ shifts closer to 0.

### *Asymmetries in the free-energy landscape*

We will now try to understand the origin of the asymmetry in the four-bar linkage's free-energy profile, i.e., the reason why the free-energy values close to $\theta_1 = \pm\pi$ are higher than the values at $\theta_1 = 0$. Consider Fig. 3.6 where we show a four-bar linkage in four different configurations, with the configurations on the left having an internal angle $\theta_1$ equal to the supplement of the value of $\theta_1$ for the configurations on the right. Geometrically, one might have expected the configurations on the left to be mirror images of the ones on the right. However, as Fig. 3.6 shows, that is only true when the linkage is being deformed along the parallel branch and when it is on the twisted branch, the two configurations are not mirror images. This is what leads to an asymmetric free-energy landscape for the four-bar linkage. More quantitative evidence for this asymmetry can be seen by expanding the angle relations along the twisted

[12] We expand $\Delta\mathscr{A}(\theta_1)$ to $\mathcal{O}(|\theta_1|^6)$ instead of $\mathcal{O}(|\theta_1|^4)$ since there are higher-order corrections to Eq. (3.50) that make an $\mathcal{O}(|\theta_1|^4)$ estimation less accurate.

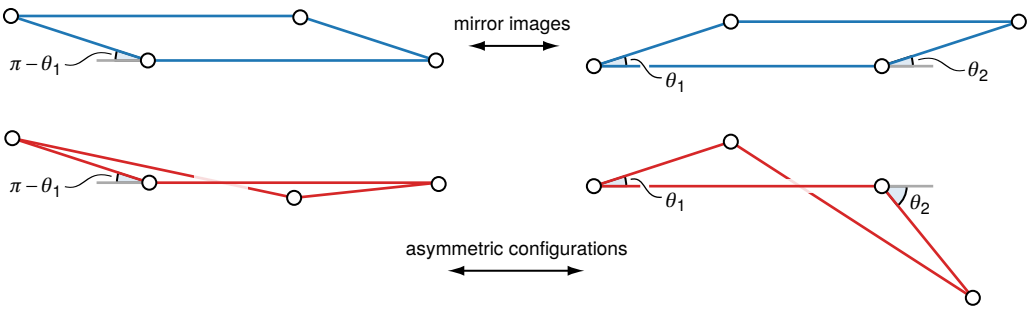


**Figure 3.6.** For values of the angle $\theta_1 \to 0$ or $\pi$, the configurations of the four-bar linkage display a certain asymmetry depending on the branch they are in. In case of the parallel branch, the two configurations are mirror images of each other (top figures). On the twisted branch, however, the two configurations are asymmetric and cannot be rotated/reflected into one another (bottom figures). The angle $\theta_1$ for the two configurations on the left is equal to the supplement of the same angle for the configurations on the right.

branch, Eq. (3.33), in $\theta_1$ around $\theta_1 = 0$ and $\theta_1 = \pm\pi$, whereby we find

$$\begin{aligned} \theta_2 &= \left(\frac{1+\lambda}{1-\lambda}\right)\theta_1 + \mathcal{O}(\theta_1^2) && \text{(expansion around } \theta_1 = 0\text{)}, \\ \pm\pi - \theta_2 &= \left(\frac{1-\lambda}{1+\lambda}\right)(\pm\pi - \theta_1) + \mathcal{O}(\theta_1^2) && \text{(expansion around } \theta_1 = \pm\pi\text{)}. \end{aligned} \tag{3.52}$$

The asymmetry in the free-energy profiles can also be seen if we use the angle $\theta_2$ as our CV instead of the angle $\theta_1$. Figure 3.7 shows the numerical free-energy profile as a function of the angle $\theta_2$, along with asymptotic results. Again, there is impressive agreement between the numerical and asymptotic results. However, now the free energy for $\theta_2$ values close to $\theta_2 = 0$ are higher than the values close to $\theta_2 = \pm\pi$. The asymptotic expressions for the free energy used in Fig. 3.7 have been obtained using the general procedure outlined in this section. But we do not write down the results explicitly as they can be read off from Eq. (3.46) after setting the nondimensional parameters $X \to \sqrt{\beta\kappa}\lambda a/(8|\lambda+1|)$ and $Y \to |(\lambda+1)/(\lambda-1)|$ [cf. Eq. (3.47)].

## 3.4 Example: triangulated origami

For further testing our methods, we consider an origami made by triangulating a unit square [18] and embedded in three dimensions [Fig. 3.8(a), lower-left corner]. To make the origami more realistic, in simulations, we avoid all configurations that result in face intersections. The one-dimensional shape space of this origami can be visualized as four intersecting branches in the space of the fold angles, i.e., the supplement of the dihedral angle at a fold. The intersection point is the singular flat state of the origami, where all the fold angles are zero.

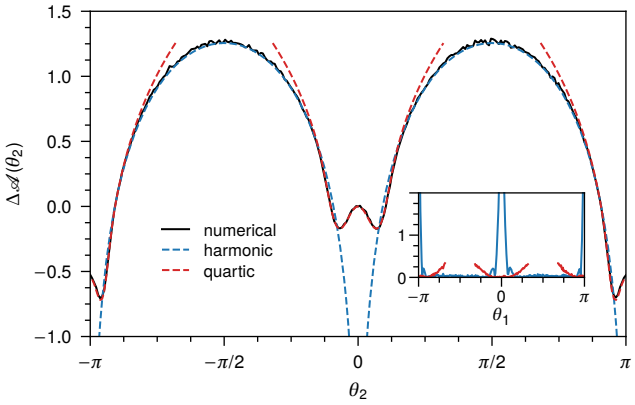


**Figure 3.7.** Free-energy difference $\Delta\mathscr{A}(\theta_2)$ of a four-bar linkage with parameters $a = 1$ and $\lambda = 2$ as a function of the angle $\theta_2$ [cf. Fig. 3.4]. The free energy is in units of $\beta^{-1}$ at $\beta = 10^4$. The inset shows the absolute errors between the numerical and asymptotic results.

### 3.4.1 Body frame

To remove the rigid motions, we transform to a body frame attached to the origami with joint 1 at the origin, bar 1–2 lying along the $x$ axis, and bar 2–6 constrained to move on the $xy$ plane [see Fig. 3.9]. The origami is made out of $N = 6$ joints and $m = 11$ bars and its configuration vector $\boldsymbol{q} \in \mathbb{R}^{12}$ in the body frame is $\boldsymbol{q} = (q_1, q_2, \ldots, q_{12})$. All orientations of the origami in the lab frame can be fully described by the two spherical polar angles that uniquely give the orientation of bar 1–2 and an azimuthal angle that gives the overall rotation of the origami about bar 1–2 [66]. After integrating over the coordinates of joint 1 (i.e., the translational coordinates) and the three angles, and dropping constant factors, the overall Jacobian factor involved in the transformation from the lab to the body frame is $I(\boldsymbol{q}) = |q_1^2 q_{12}|$.

### 3.4.2 Branch parameterization

Since the shape space of the origami (and other larger frameworks) is not amenable to analytical parameterization, we have to parameterize it numerically. We shall parameterize the branches of the shape space in terms of the fold angle $\rho_1$, which we will also use as our CV. We first express the tangent to the shape space at the flat state, $\boldsymbol{t}_0 \in \ker \mathsf{C}$, in terms of its unknown components in the basis of the $(N-3)$ out-of-plane displacement vectors of the joints.[13] An origami made by triangulating a square and having $N$ joints is expected to have

[13] Note that three of the $N$ joints are always constrained to move on a coordinate plane of the local Cartesian body frame. This means that only $(N-3)$ joints have out-of-plane displacements, and it is these displacement vectors that span $\ker \mathsf{C}$.

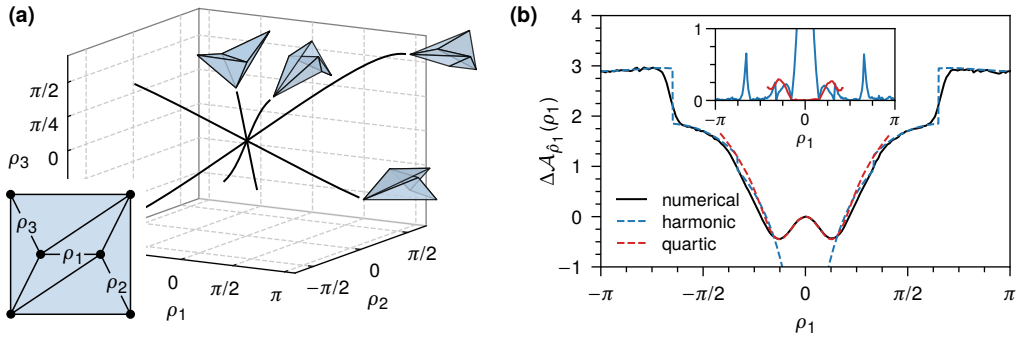


**Figure 3.8.** (a) A triangulated origami modeled as a bar-joint framework and (b) its shape space visualized in the space of fold angles $\rho_1, \rho_2$, and $\rho_3$. (c) Free-energy difference $\Delta\mathscr{A}(\rho_1)$ in units of $\beta^{-1}$ at $\beta = 10^4$. The inset shows the absolute errors between the numerical and asymptotic results using the harmonic and quartic approximations (blue and red curves, respectively).

at least $(N-4)$ states of self stress $\boldsymbol{\sigma} \in \ker \mathsf{C}^\mathsf{T}$ when it is flat [18]. To find the components of $\boldsymbol{t}_0$, we then solve the $(N-4)$ coupled quadratic equations [18, 162] [see Eq. (3.26)]

$$\boldsymbol{\sigma} \cdot \boldsymbol{w}(\boldsymbol{t}_0) = 0, \quad \text{for } \boldsymbol{\sigma} \in \ker \mathsf{C}^\mathsf{T} \tag{3.53}$$

along with the normalization condition $\|\boldsymbol{t}_0\| = 1$. Here the $i$th component of the vector $\boldsymbol{w}(\boldsymbol{t}_0) \in \mathbb{R}^m$ is $\frac{1}{2}\boldsymbol{t}_0^\mathsf{T} \nabla\nabla f_i \boldsymbol{t}_0$.

Numerical evidence [18] suggests that one would get $2^{N-4}$ unique tangent vectors—each corresponding to a particular branch—by solving the above quadratic equations. (Also see the related discussion regarding the solution space of these equations in Section 3.2.2.) Once a tangent vector to a branch at the flat state is selected, the rest of the branch can be parameterized by numerically solving the differential equation $\mathrm{d}\boldsymbol{q}/\mathrm{d}h = \boldsymbol{t}$, where $h$ is the arc length along a branch and $\boldsymbol{t}$ is the unit tangent vector to the branch at $\boldsymbol{q}$. If the $k$th point on the branch is $\boldsymbol{q}_k \in \mathbb{R}^n$, the next point $\boldsymbol{q}_{k+1}$ is then found by solving $f(\boldsymbol{q}_{k+1}) = 0$, e.g., using the Gauss–Newton method. As the initial guess we take $\boldsymbol{q}_{k+1} = \boldsymbol{q}_k + \Delta h\, \boldsymbol{t}_k$, where $\Delta h$ is the step length along a branch. The tangent vector $\boldsymbol{t}_k$ for $k \neq 0$ can be found by numerically computing $\ker \mathsf{C}$ at $\boldsymbol{q}_k$. Since the tangent vectors obtained this way do not preserve direction in general, we should also multiply each $\boldsymbol{t}_k$ that is found with the sign of its dot product with the previous tangent vector $\boldsymbol{t}_{k-1}$.

On successive repetition of the above steps starting at the flat state $(\boldsymbol{q}_0, \boldsymbol{t}_0)$, we get $\boldsymbol{q}$ as a function of the arc length $h$. To reparameterize the branch using the CV, which is the fold angle $\rho_1$ of fold 5–6, we first (linearly) interpolate between the arc-length parameterized points to obtain a set of points that are uniformly spaced in $\rho_1$. We then refine the interpolated points by solving $f(\boldsymbol{q}) = 0$ with the interpolated points as the initial guess. The interpolation/refinement steps can be repeated as many times as required to achieve the desired accuracy goal. Once the parameterization is complete, the induced metric along a branch can be computed by

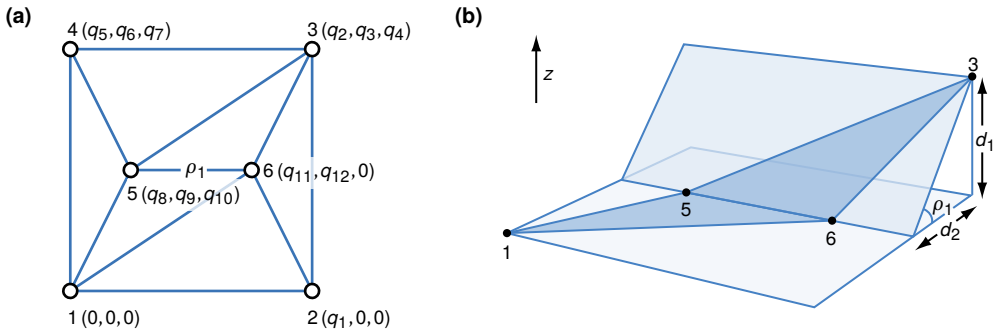


**Figure 3.9.** (a) A triangulated origami as a bar-joint framework illustrating the coordinates of the joints in the body frame. When the origami is flat, the external vertices lie on the corners of a unit square and the internal vertices 5 and 6 have coordinates $(1/4, 1/2, 0)$ and $(3/4, 1/2, 0)$. (b) Finding the CV (i.e., fold angle $\rho_1$) using simple geometry.

approximating the derivatives using difference quotients. Note that this parameterization is only used in conjunction with Eq. (3.19), the asymptotic expression for the marginal density under the harmonic approximation.

### 3.4.3 Marginal probability densities

*Regular values*

Similar to the calculation for the four-bar linkage, we first write down the constraint map $f : \mathbb{R}^{12} \to \mathbb{R}^{11}$ in the body frame and find the compatibility matrix $\mathsf{C} = \nabla f$. Once all four branches of the shape space have been numerically parameterized in terms of the CV $\rho_1$, the marginal probability density $\mathscr{P}(\rho_1)$ can be computed using Eq. (3.19). Here we remark that instead of computing $\det \mathsf{D}^\perp$ by finding the nonzero eigenvalues of $\mathsf{D}$, it is more convenient to calculate it using the fact that $\det \mathsf{D}^\perp = \det \mathsf{K}\mathsf{C}\mathsf{C}^\mathsf{T}$ at regular points. This gives $\mathscr{P}(\rho_1)$ for all $\rho_1$ far from the singular value (i.e., for $|\rho_1| \gg 0$).

*Singular value*

Our goal here is to use Eq. (3.23) to find the marginal density $\mathscr{P}(\rho_1)$ as $\rho_1 \to 0$. Since the calculation is similar in spirit to the case of the four-bar linkage, we only present the key steps here. As before, we numerically compute $\det \mathsf{D}^\perp$ from the nonzero eigenvalues of $\mathsf{D}$ at the singularity. The next step is to linearize the CV map so that the integral in Eq. (3.23) can be evaluated.

One can always numerically compute the fold angle $\rho_1$ as the angle between the normals to the faces that share fold 5–6. However, for linearizing the CV map, we need to find an expression that is more tractable analytically. A moment's thought [and Fig. 3.9(b)] shows that

the CV map can be written[14] as $\hat{\rho}_1(\boldsymbol{q}) = -\tan^{-1}[d_1(\boldsymbol{q})/d_2(\boldsymbol{q})]$. Here $d_1(\boldsymbol{q})$ is the perpendicular distance from joint 3 to the plane containing face 1–5–6, and $d_2(\boldsymbol{q})$ is the perpendicular distance from the projection of joint 3 on this plane to the line along fold 5–6. A straightforward calculation gives

$$d_1(\boldsymbol{q}) = \frac{q_{10}(q_2 q_{12} - q_3 q_{11}) + q_4(q_9 q_{11} - q_8 q_{12})}{\left|q_{10}^2 q_{11}^2 + q_{10}^2 q_{12}^2 + (q_9 q_{11} - q_8 q_{12})^2\right|^{1/2}}. \tag{3.54}$$

At the singular flat state $\bar{\boldsymbol{q}}^*$, we have $d_1(\bar{\boldsymbol{q}}^*) = 0$ and $d_2(\bar{\boldsymbol{q}}^*) = 0.5$, using which we find the Jacobian of the CV map to be

$$\begin{aligned}\nabla\hat{\rho}_1(\bar{\boldsymbol{q}}^*) = -\frac{d_2(\bar{\boldsymbol{q}}^*)\nabla d_1(\bar{\boldsymbol{q}}^*) - d_1(\bar{\boldsymbol{q}}^*)\nabla d_2(\bar{\boldsymbol{q}}^*)}{d_1^2(\bar{\boldsymbol{q}}^*) + d_2^2(\bar{\boldsymbol{q}}^*)} &= -\frac{\nabla d_1(\bar{\boldsymbol{q}}^*)}{d_2(\bar{\boldsymbol{q}}^*)} \\ &= \begin{pmatrix}0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0\end{pmatrix}.\end{aligned} \tag{3.55}$$

Here we see that one can linearize the CV map without needing an explicit expression for $d_2(\boldsymbol{q})$. Also, it is easy to verify that $(\nabla\hat{\rho}_1)\boldsymbol{v} = 0$ for all fast modes $\boldsymbol{v} \in (\ker \mathsf{C})^\perp$, enabling us to use Eq. (3.23) to find the marginal density as $\rho_1 \to 0$.

As the basis for $\ker\mathsf{C}$ we choose the out-of-plane displacements of the joints in the body frame and write the vector $\boldsymbol{u} \in \ker\mathsf{C}$ in terms of the $z$ coordinates $q_4$, $q_7$, and $q_{10}$ of the joints. This yields $\nabla\hat{\rho}_1 \cdot \boldsymbol{u} = 2(q_{10} - q_4)$. It also follows that the hyperplane $\Xi_\xi$ is the plane defined by $2(q_{10} - q_4) = \rho_1$ in the $q_4$-$q_7$-$q_{10}$ space and we can parameterize the points on $\Xi_\xi$ using either $(q_4, q_7)$ or $(q_7, q_{10})$. Once we write down the vector $\boldsymbol{w}(\boldsymbol{u})$ and find the self stresses $\boldsymbol{\sigma}$, we have all the ingredients to use Eq. (3.23). However, the integral that one is left with is not amenable to analytical integration, and so we have to resort to numerical quadrature. We do this by using a simple Monte Carlo integration scheme with $10^9$ sample points for each value of $\rho_1$ so that the maximum error is below $10^{-3}$. One also gets near-identical results with Mathematica's numerical integrator using a global adaptive method.[15]

### 3.4.4 Free energy

The free-energy difference $\Delta\mathscr{A}(\rho_1) = \mathscr{A}(\rho_1) - \mathscr{A}(0)$ for all values of $\rho_1$ can be easily computed from the corresponding probability densities using Eq. (3.7). As we see from Fig. 3.8(b), the numerical and the asymptotic results for the free-energy difference $\Delta\mathscr{A}(\rho_1)$ show good agreement in both regimes of $\rho_1$. Self-avoidance of the faces forces us to consider only a part of each branch of the shape space for our analysis. Since the extent of these parts (in $\rho_1$) varies for the four branches [Fig. 3.8(a)], it results in discontinuous jumps in the free-energy curves. Similar to the four-bar linkage, the free energy of the origami also has a double-well appearance around the singular value $\rho_1 = 0$. This is again because of the branched nature of its shape space and the increased softness near the singularity.

[14] To assign a sign to the fold angle, we first choose a unique normal to face 1–5–6 such that it coincides with the positive $z$ axis when the origami is flat. Signs are then chosen so that a mountain fold (as perceived by looking downwards along this normal) has a positive fold angle, e.g., the angle in Fig. 3.9(b) is negative since the fold is a valley fold.

[15] https://reference.wolfram.com/language/ref/NIntegrate.html

### 3.5 Example: planar five-bar linkage

The example frameworks we have considered so far have one-dimensional shape spaces. In this section, we illustrate the application of our formalism to a framework with a two-dimensional shape space, namely the planar five-bar linkage [26, 118] illustrated in Fig. 3.10(a). The five-bar linkage we consider is made out of four bars of equal length $a$ and a fifth bar of length $2a$. The shape space of the five-bar linkage [Fig. 3.10(b)] when visualized in the space of the angles $\zeta_1, \zeta_2$, and $\zeta_3$ appears as a two-dimensional surface with four isolated singularities,[16] all of which correspond to configurations where the bars become collinear and support a state of self stress. Since force balance in self-stressed states of a planar polygonal linkage requires its bars to be collinear [41], it is not surprising that these singularities are isolated, and in their neighborhoods, the shape space is very nearly a double cone [80, 118]. Similar isolated singularities are also seen in the shape spaces of origami that have self-stressed flat states [9].

For the sake of brevity and to avoid cluttering this SM with qualitatively similar results, we only discuss the nature of the free-energy landscape around the singularity at $(\zeta_1, \zeta_2, \zeta_3) = (0,0,0)$. Since the shape space of the five-bar linkage is two-dimensional, we choose a similarly two-dimensional CV, $\zeta = (\zeta_1, \zeta_2)$. Next, we employ Eq. (3.23) to find the asymptotic free-energy difference $\Delta\mathscr{A}(\zeta)$, choosing the singular value $\zeta^* = (0,0)$ as the point of zero free energy. Note that for doing this, we do not require an explicit parameterization of the linkage's shape space in terms of $\zeta$. Such a parameterization is only required if we want to use Eq. (3.19) to find the free energy far from $\zeta^* = (0,0)$ using the harmonic approximation. Since the other details are very similar to the previous calculations for the four-bar linkage and the triangulated origami, we just quote the final result:

$$\Delta\mathscr{A}(\zeta) \sim \beta^{-1}\left\{Z^2\zeta_1^2\zeta_2^2 - \ln\left[\frac{D_{-1/2}(-2Z\zeta_1\zeta_2)}{D_{-1/2}(0)}\right]\right\}, \tag{3.56}$$

where $Z$ is a positive dimensionless term defined by $Z = \sqrt{\beta\kappa}a/(2\sqrt{5})$ [cf. Eq. (3.44)]. A comparison between the theoretical and numerical results [Figs. 3.10(c) and 3.10(d)] shows very good agreement between the two. Similar to the four-bar linkage and the triangulated origami, the effective free energy minimum of the five-bar linkage is not at the singular value $\zeta^* = (0,0)$. Expanding $\Delta\mathscr{A}(\zeta)$ to $\mathcal{O}(\zeta_1^3\zeta_2^3)$ around $\zeta^*$ we see that the bottom of the free-energy valley near the singular value [white dashed curves in Figs. 3.10(c) and 3.10(d)] is approximately defined by the equation

$$\zeta_1\zeta_2 = \frac{\Gamma^2\left(\frac{1}{4}\right)\left\{8\pi^2 + \Gamma^4\left(\frac{1}{4}\right) - \sqrt{\left[8\pi^2 + \Gamma^4\left(\frac{1}{4}\right)\right]^2 - (16\pi^2)^2}\right\}}{64\pi^3 Z}, \tag{3.57}$$

which is that of a rectangular hyperbola [cf. Eq. (3.51)].

[16] At $(\zeta_1, \zeta_2, \zeta_3) = (0,0,0), (0, \pm\pi, \pm\pi), (\pm\pi, 0, \pm\pi)$, and $(\pm\pi, \pm\pi, 0)$.

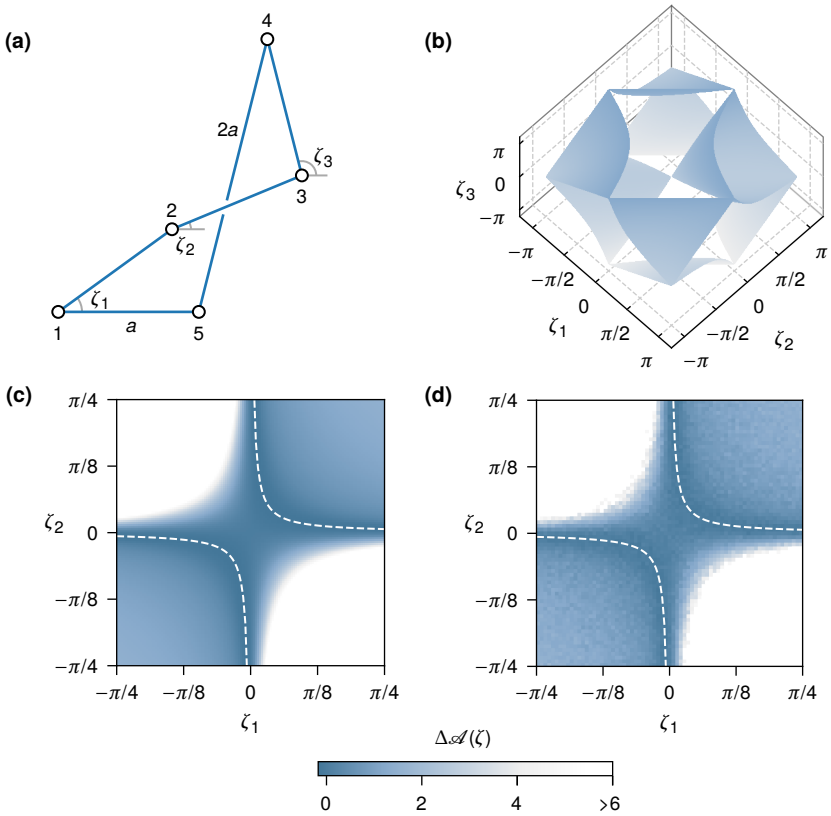


**Figure 3.10.** (a) The five-bar linkage and (b) its shape space visualized in terms of the angles $\zeta_1, \zeta_2$, and $\zeta_3$. These angles are measured counterclockwise from an axis parallel to bar 1–5. Free-energy profile (in units of $\beta^{-1}$) at $\beta = 10^4$ around the singular value $\zeta^* = (0,0)$ from (c) theory [Eq. (3.56)] and (d) simulations, with the dashed curves depicting the approximate bottom of the hyperbolic valley [Eq. (3.57)].

### 3.6 Permanently singular frameworks

Our discussion so far has been concerned with frameworks with isolated singularities. In such frameworks, the constraint map $f$ drops rank only at the singularities and has full rank everywhere else. Now consider the framework shown in Fig. 3.11(a). The bars connecting joints 1, 2, and 3 of this framework are in a permanent state of self stress, irrespective of the value of the angle $\theta$. By direct inspection we see that the one-dimensional shape space $\Sigma$ of this framework can be parameterized in terms of the internal angle $\theta$ using $\psi(\theta) = (a, \frac{1}{2}a, 0, a\cos\theta, a\sin\theta)$. This equation defines a smooth circle in $\mathbb{R}^5$, which does not have any "visible" singularities. However, on inserting this parameterization into the compatibility matrix $\mathsf{C} = \nabla f$ and the dynamical matrix $\mathsf{D} = \mathsf{C}^\mathsf{T}\mathsf{K}\mathsf{C}$ we find (assuming that all bars have stiffness $\kappa$)

$$\mathsf{C} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos\theta & \sin\theta \end{pmatrix}, \quad \mathsf{D} = \kappa \begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos^2\theta & \sin\theta\cos\theta \\ 0 & 0 & 0 & \sin\theta\cos\theta & \sin^2\theta \end{pmatrix}. \tag{3.58}$$

Clearly, $\mathsf{C}$ is rank deficient irrespective of the value of $\theta$ and $\mathsf{D}$ has two zero modes: $\boldsymbol{t} = \mathrm{d}\psi/\mathrm{d}\theta \in T_{\bar{\boldsymbol{q}}}\Sigma$ corresponding to tangential motion on $\Sigma$ and a singular zero mode $\boldsymbol{u} = (0,0,1,0,0)$ associated with the self stress [see Fig. 3.11(a)]. Hence, even though the shape space here is a smooth manifold, the presence of the singular zero mode causes the harmonic approximation to break down everywhere on $\Sigma$. This should be contrasted with the case of isolated singularities, where such singular modes appear only at the singularities of $\Sigma$. The results that we have derived so far (namely, Eqs. 3.19 and 3.23) will not let us analyze permanently singular frameworks [123, 173] such as the one in Fig. 3.11(a). However, such frameworks have been considered in the context of colloidal clusters [77]. It is thus instructive to rederive these results and compare them with our results for frameworks with isolated singularities.

Consider again a framework whose shape space $\Sigma$ is defined as the zero level set of a constraint map $f : \mathbb{R}^n \to \mathbb{R}^m$ with the compatibility matrix $\mathsf{C} = \nabla f$. When there are $s$ permanent states of self stress, out of the $n - m + s$ zero modes that belong to $\ker\mathsf{C}$, there are $n - m$ zero modes that belong to the tangent space $T_{\bar{\boldsymbol{q}}}\Sigma$, and the remaining $s$ zero modes are the singular zero modes. We can thus write $\ker\mathsf{C}(\bar{\boldsymbol{q}}) = T_{\bar{\boldsymbol{q}}}\Sigma \oplus \mathscr{S}$ for all $\bar{\boldsymbol{q}} \in \Sigma$. Here $\mathscr{S}$ is the subspace of the singular zero modes at $\bar{\boldsymbol{q}}$, defined as the orthogonal complement of $T_{\bar{\boldsymbol{q}}}\Sigma$ in $\ker\mathsf{C}$.[17] Note that such a decomposition of $\ker\mathsf{C}$ into two vector subspaces is not possible when $\Sigma$ has isolated singularities (where multiple branches cross) for two reasons: (i) singular zero modes exist only at the singularities of $\Sigma$ and (ii) at these singularities, there is no well-defined tangent space $T_{\bar{\boldsymbol{q}}}\Sigma$.

Since the subspace $\mathscr{S}$ of singular zero modes can be identified for all points in $\Sigma$, we can expand the energy to quartic order along $\boldsymbol{u} \in \mathscr{S}$ using Eq. (2.19). (This would not have been

[17] Since the zero modes are degenerate, the singular zero modes obtained by computing $\ker\mathsf{C}$ (or $\ker\mathsf{D}$) need not be orthogonal to $T_{\bar{\boldsymbol{q}}}\Sigma$. Hence, in writing $\ker\mathsf{C} = T_{\bar{\boldsymbol{q}}}\Sigma \oplus \mathscr{S}$, we are *defining* a singular mode to be one that belongs to $\ker\mathsf{C}$, but is orthogonal to $T_{\bar{\boldsymbol{q}}}\Sigma$. Clearly, such zero modes cannot be extended to a smooth deformation of the framework.

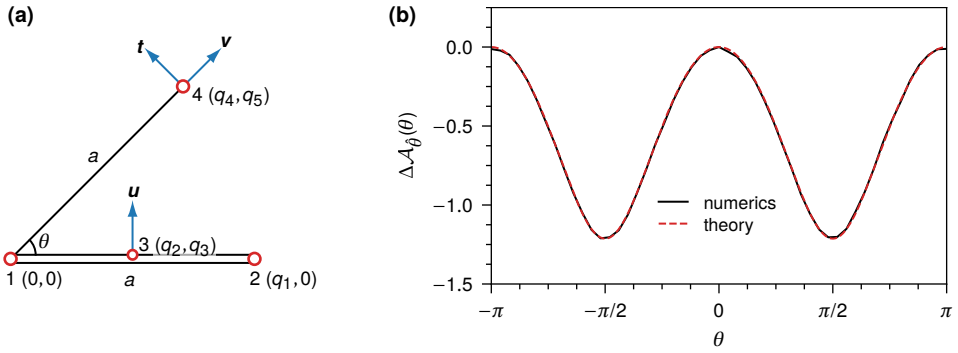


**Figure 3.11.** (a) A planar framework with four joints and four bars in a body frame attached to the framework. A bar of length $a$ connects joints 1 and 2, and two other bars of length $\frac{1}{2}a$ connect joint 3 with joints 1 and 2, which results in a state of self stress between joints 1, 2, and 3 [16] for all values of $\theta$. Here $\boldsymbol{t}$ is the zero mode corresponding to tangential motion on the shape space $\Sigma$, $\boldsymbol{u}$ is a singular zero mode associated with the self stress, and $\boldsymbol{v}$ is one of the three vibrational modes. (b) Free-energy difference $\Delta\mathscr{A}(\theta)$ [Eq. (3.60)] in units of $\beta^{-1}$ for a hypothetical $\theta$-dependent stiffness $\kappa(\theta) = \kappa_0(1+\cos^2\theta)$, chosen to verifying scaling. For constant stiffness, $\Delta\mathscr{A}(\theta)$ is zero and the free-energy landscape is flat (unlike in the case of isolated singularities).

possible for isolated singularities since Eq. (2.19) is valid only when the expansion is around a singularity.) To derive the asymptotic marginal density $\mathscr{P}(\theta)$, we can then proceed similar to the derivation of Eq. (3.19). As before, we pick coordinates associated with the framework's internal degrees of freedom as our CV $\xi$ and assume that $\Sigma$ can be parameterized using $\xi$. We expand both the terms in the exponent of Eq. (3.10) around each ground state $\bar{\boldsymbol{q}} \in \hat{\xi}^{-1}(\xi)$ after setting $\boldsymbol{q} \to \bar{\boldsymbol{q}} + \boldsymbol{t} + \boldsymbol{u} + \boldsymbol{v}$, where $\boldsymbol{t} \in T_{\bar{\boldsymbol{q}}}\Sigma$ is a zero mode that can be extended to a smooth deformation of the framework, $\boldsymbol{u} \in \mathscr{S}$ is a singular zero mode, and $\boldsymbol{v} \in (\ker \mathsf{C})^\perp$ is a fast mode. We also choose the columns of $\nabla\psi(\xi)$ as the basis for $T_{\bar{\boldsymbol{q}}}\Sigma$, with $\psi : \mathbb{R}^{n-m} \to \mathbb{R}^n$ being the parameterization of $\Sigma$ near $\bar{\boldsymbol{q}}$. Integrating over the components of $\boldsymbol{t}$ yields

$$\begin{aligned}\mathscr{P}(\xi) &\sim I(\xi)\left(\frac{2\pi}{\beta}\right)^{(m-s)/2}\left|\det[\nabla\psi(\xi)]^{\mathsf{T}}\nabla\psi(\xi)\right|^{1/2}\int_{\mathscr{S}}\mathrm{d}\boldsymbol{u}\int_{(\ker\mathsf{C})^\perp}\mathrm{d}\boldsymbol{v}\,\exp\left[-\tfrac{1}{2}\beta\kappa\|\mathsf{C}\boldsymbol{v}+\boldsymbol{w}(\boldsymbol{u})\|^2\right]\\ &= I(\xi)\left(\frac{2\pi}{\beta}\right)^{(m-s)/2}(\beta\kappa)^{-s/4}\left|\frac{\det[\nabla\psi(\xi)]^{\mathsf{T}}\nabla\psi(\xi)}{\det\mathsf{D}^\perp(\xi)}\right|^{1/2}\int_{\mathbb{R}^s}\mathrm{d}\boldsymbol{x}\,\exp\left\{-\tfrac{1}{2}\sum_{\boldsymbol{\sigma}}[\boldsymbol{\sigma}\cdot\boldsymbol{w}(\boldsymbol{x})]^2\right\}.\end{aligned} \tag{3.59}$$

In the last step, we have integrated over the fast modes in $(\ker\mathsf{C})^\perp$ and used the results in Eqs. (3.20)–(3.22), with $\det\mathsf{D}^\perp$ being the product of the $m-s$ nonzero eigenvalues of the dynamical matrix $\mathsf{D}$ at $\psi(\xi)$. Also, after picking an orthonormal basis for $\mathscr{S}$ and writing $\boldsymbol{u} = \mathsf{A}\boldsymbol{x}$, we can rescale the components $\boldsymbol{x} \to (\beta\kappa)^{-1/4}\boldsymbol{x}$ to extract all $\beta$ and $\kappa$ dependence as the term $\sum_{\boldsymbol{\sigma}}[\boldsymbol{\sigma}\cdot\boldsymbol{w}(\boldsymbol{x})]^2$ is a homogeneous quartic polynomial in the components of $\boldsymbol{x}$. Equation (3.59) is a rederivation of the integrand in Eq. (14) of Ref. [77] and although it looks

very similar to Eq. (3.23), it is fundamentally different. The similarities arise due to the fact that in deriving both Eqs. (3.59) and (3.23), we used the same quartic-order expansion of the energy. Nonetheless, it is worthwhile to compare the two equations. We can identify four major differences.

*Integration domain.* The integration domain in Eq. (3.59) is $\mathscr{S}$, the subspace of singular zero modes, which exists for all configurations of a permanently singular framework. In comparison, when the framework has isolated singularities, singular zero modes arise only at these singularities. Furthermore, the domain of integration in Eq. (3.23) is a CV-dependent hyperplane $\Xi_\xi = (\nabla\hat{\xi})^{-1}(\xi - \xi^*) \cap \ker \mathsf{C}$, which is not the subspace of singular zero modes, and such a vector subspace cannot be identified for a framework with isolated singularities.

*Relation to the shape space.* Using Eq. (3.59) requires a parameterization $\psi(\xi)$ of the shape-space $\Sigma$ and the factor $\det(\nabla\psi)^\mathsf{T}\nabla\psi$ is the determinant of the induced metric on $\Sigma$. This is similar to Eq. (3.19), which is the harmonic marginal density. In contrast, Eq. (3.23) does not make an explicit reference to $\Sigma$ and does not involve any parameterization of its branches. Equation (3.23) acquires the factor $|\det \nabla\hat{\xi}(\nabla\hat{\xi})^\mathsf{T}|^{-1}$ from Eq. (3.9), which a corollary of the coarea formula. Also, Eq. (3.59) is valid for all values of $\xi$, whereas Eq. (3.23) is only valid when $\xi$ is close to a singular value $\xi^*$ of the CV.

*Convergence.* If a permanently singular framework becomes second-order rigid (see Section 3.2.2) once the zero modes are restricted to the subspace $\mathscr{S}$, then $\boldsymbol{\sigma}\cdot\boldsymbol{w}(\boldsymbol{x}) > 0$ for all $\boldsymbol{\sigma}$ and the integral in Eq. (3.59) converges [77]. In particular, it will not converge for frameworks that are rigid, but not second-order rigid in $\mathscr{S}$ (e.g., see the example in Appendix A.3 of Ref. [24]). This should be contrasted with the case of Eq. (3.23) where there would always be vectors $\boldsymbol{t}$ in the integration domain $\Xi_\xi$ that satisfy $\boldsymbol{\sigma}\cdot\boldsymbol{w}(\boldsymbol{t}) = 0$ (each corresponding to a tangent to the branch at the singularity). Hence, convergence of Eq. (3.23) relies on the requirement that the number of such vectors is finite. Two necessary conditions required for this are (i) the tangents $\boldsymbol{t}$ are resolvable at second order and (ii) the CV map is such that $(\nabla\hat{\xi})\boldsymbol{t} \neq 0$ for all $\boldsymbol{t}$ (see Section 3.2.2).

*Scaling.* The scaling of Eq. (3.59) with respect to $\beta$ is consistent with Eq. (9) of Ref. [77]. Furthermore, since the scaling is independent of the value of the CV $\xi$, one would generically expect the free-energy barriers in a permanently singular framework to be dominated by entropic effects and the landscape would have the same appearance for all values of $\beta$ (provided it is large). Contrast this with Eq. (3.23), where the term in the exponent is an inhomogeneous quartic polynomial in the components of $\boldsymbol{u}$ for $\xi \neq \xi^*$ (see Section 3.2.2). This makes the scaling nontrivial and gives rise to temperature-dependent barriers in the free-energy landscape.

Returning to the example framework in Fig. 3.11(a), using Eq. (3.59), we find the marginal density $\mathscr{P}(\theta)$ and the free-energy difference $\Delta\mathscr{A}(\theta) = -\beta^{-1}\ln\mathscr{P}(\theta) + \beta^{-1}\ln\mathscr{P}(0)$ to be

$$\mathscr{P}(\theta) = \Gamma\left(\tfrac{1}{4}\right)\left|\frac{2\pi^6 a^{10}}{3(\beta\kappa)^7}\right|^{1/4}, \quad \Delta\mathscr{A}(\theta) = \tfrac{7}{4}\beta^{-1}\ln\left[\frac{\kappa(\theta)}{\kappa(0)}\right]. \tag{3.60}$$

Above, we have allowed for a hypothetical $\theta$-dependent stiffness $\kappa(\theta)$ so that the scaling factor 7/4 can be verified. The numerical results in Fig. 3.11(b) show excellent agreement with the analytical predictions. If the stiffness $\kappa$ is a constant, then $\Delta\mathscr{A}(\theta)$ vanishes for all values of $\theta$

and the landscape becomes flat. This should be contrasted with the examples for frameworks with isolated singularities, where temperature-dependent free-energy barriers exist even for constant stiffness.

## 3.7 Conclusion

In this chapter we have described a formalism to find the free-energy landscapes of common bar-joint frameworks with isolated singularities in their shape spaces. Our results indicate that configurations in the neighborhood of the singularities have relatively lower free energy compared to configurations farther from the singularities. More specifically, the free-energy landscapes of the four-bar linkage and the triangulated origami [Figs. 3.4 and 3.8(b)] demonstrate that the measured values of the CV tend to be closer to their values near the singularities. Yet, as free-energy landscapes (and even their extrema) do not always have a CV-agnostic interpretation [37, 44, 63], to draw conclusions we should also consider the physical meaning of the chosen CV (see Appendix 3.B for further discussion). The CVs we picked for both the example frameworks were internal angles whose values dictate the overall shape of the framework. Specifically, according to our results, we expect the bars of the four-bar linkage to tend to be collinear, as measured by the angle $\theta_1$ being close to 0 or $\pi$. Similarly, the origami will tend towards being flat, as measured by the fold angle $\rho_1$. This tendency increases at lower temperatures as the free-energy barriers become larger.

Our findings could help in programming the conformational dynamics of nanoframeworks [36] and highlight the interplay between the geometry of a framework's shape space and its thermodynamic properties. Since changes in configuration space topology is known to drive equilibrium phase transitions in certain physical systems [81], it would then be interesting to consider how shape-space singularities affect the physical properties of frameworks in the thermodynamic limit. Indeed, the affinity for the origami to be in nearly flat configurations is reminiscent of the well-known flat phase of a polymerized membrane [1, 12], which is the natural thermodynamic limit of the triangulated origami. Additionally, it would be interesting to explore the connection between recent mean-field theories [65] written to describe the anomalous elastic response of cellular tissues and the results presented in this work. The onset of rigidity in these tissues can be mapped to a critical configuration in a single cell's configuration space, rather similar to the singularities we have seen so far.[18] Other open questions include the behavior of these frameworks in the presence of active (nonthermal) noise [46, 47], which is known to preferentially actuate zero modes [172], and methods to bias their dynamics towards desired states [78], e.g., by introducing CV-dependent bias potentials [82].

## 3.A Numerical simulations

For all the frameworks we consider in this chapter, we perform our numerical simulations[19] in the lab frame at an inverse temperature $\beta = 10^4$ using a central-force potential $\phi_i[\ell_i(\boldsymbol{r})] =$

[18] Compare, for instance, Fig. 4 of Ref. [65] and Fig. 1.1(c) on page 3.

[19] The code we use for Monte Carlo simulations and numerical parameterization of the shape spaces is publicly available at https://github.com/manu-mannattil/thermmech.

$[\ell_i^2(\boldsymbol{r}) - \bar{\ell}_i^2]^2/(8\bar{\ell}_i^2)$, which has an absolute minimum at $\ell_i = \bar{\ell}_i$, for $\ell_i \geq 0$. With this potential, the stiffness $\kappa_i = \phi_i''(\bar{\ell}_i) = 1$ for all bars. Alternatively, any other potential $\phi_i(\ell_i)$ that depends only on the bar lengths and having a minimum at $\ell_i = \bar{\ell}_i$ can be used. We find the marginal probability densities of the CV [Eq. (3.8)] using histograms obtained from sampling the Boltzmann–Gibbs distribution using the classical Metropolis Monte Carlo algorithm (with an acceptance rate of about 50% and $\sim 10^9$ samples). The free-energy profile is then found using Eq. (3.7). For the triangulated origami, there is an additional need to reject all Monte Carlo moves that lead to face crossings:

(i) For faces that share an edge [e.g., faces 1–2–6 and 2–3–6 in Fig. 3.9(a)], a face crossing can be detected by looking for sign changes in the fold angle of the shared fold when it is close to $\pm\pi$.

(ii) For faces that do not share an edge, we use a triangle-triangle intersection test [164] to check if they intersect. Since there are eight such face pairs, to reduce computational costs, we only check this when the origami is sufficiently folded.

## 3.B Miscellaneous remarks

In this appendix, we collect some remarks regarding general free-energy landscapes and their interpretation.

### 3.B.1 Free energy under CV diffeomorphisms

The first issue we want to address is the lack of invariance of the free energy under a diffeomorphism (a differentiable transformation with a differentiable inverse) of the CV. Consider an $l$-dimensional CV $\xi \in \mathbb{R}^l$, measurable with a CV map $\hat{\xi} : \mathbb{R}^n \to \mathbb{R}^l$. Let $h : \mathbb{R}^l \to \mathbb{R}^l$ be a diffeomorphism such that $h(\xi) = \zeta$. The CV map for $\zeta$ is $\hat{\zeta} = h \circ \hat{\xi}$. On putting this in the coarea formula, Eq. (3.9), we find the new marginal density $\mathscr{P}(\zeta)$ to be

$$\mathscr{P}(\zeta) = \int_{\zeta^{-1}(\zeta)} \frac{\mathrm{d}\Theta(\boldsymbol{q})}{|\det \nabla\hat{\zeta}(\nabla\hat{\zeta})^{\mathsf{T}}|^{1/2}} \exp\left[-\beta U(\boldsymbol{q})\right]. \tag{3.61}$$

The CV level set $\hat{\xi}^{-1}(\xi)$ is identical to $\hat{\zeta}^{-1}(\zeta)$. To see this, note that for every point $\boldsymbol{q} \in \hat{\xi}^{-1}(\xi)$, we have $\hat{\zeta}(\boldsymbol{q}) = h(\hat{\xi}(\boldsymbol{q})) = h(\xi) = \zeta$, showing that $\boldsymbol{q} \in \hat{\zeta}^{-1}(\zeta)$ as well. Next, we use $\hat{\zeta} = h \circ \hat{\xi}$, to find $\nabla\hat{\zeta} = \nabla h \nabla\hat{\xi}$, and so $\det \nabla\hat{\zeta}(\nabla\hat{\zeta})^{\mathsf{T}} = \det \nabla\hat{\xi}(\nabla\hat{\xi})^{\mathsf{T}} \det(\nabla h)^{\mathsf{T}} \nabla h$. Using this in $\mathscr{P}(\zeta)$, we find

$$\mathscr{P}(\zeta) = \mathscr{P}(\xi)\, |\det (\nabla h)^{\mathsf{T}} \nabla h|^{-1/2}. \tag{3.62}$$

So the free energy in the new variable $\zeta$ is

$$\mathscr{A}(\zeta) = \mathscr{A}(\xi) + \tfrac{1}{2}\beta^{-1} \log|\det (\nabla h)^{\mathsf{T}} \nabla h|. \tag{3.63}$$

As the CV level sets $\hat{\xi}^{-1}(\xi)$ and $\hat{\zeta}^{-1}(\zeta)$ are identical, physically speaking, when working with both $\zeta$ and $\xi$, we are restricting our attention to the same set of microstates. Hence, the free energies $\mathscr{A}(\zeta)$ and $\mathscr{A}(\xi)$ not being equal is a bit perplexing. In reality, it is a simple consequence of probability density functions acquiring a Jacobian factor under a coordinate transformation. Because of the additional logarithmic term in Eq. (3.63) (often called a gauge term [61]) the functional forms of $\mathscr{A}(\xi)$ and $\mathscr{A}(\zeta)$ can be completely different. Consequently,

$\mathscr{A}(\xi)$ may have an entirely different set of extrema compared to $\mathscr{A}(\zeta)$ and for this reason, interpreting free-energy landscapes requires careful consideration of the CV used to describe them.

### 3.B.2 Choice of CV and free-energy extrema

This is best illustrated with an example. Consider a rotor of unit natural length in two dimensions centered around the origin with a rotationally symmetric potential energy[20] $U(x,y) = \frac{1}{2}\kappa(x^2+y^2-1)^2$, with $\kappa$ being a "spring constant" to set the units [see Fig. 3.12(b)]. This potential energy function achieves its minimum on the unit circle $\Omega$ defined by $x^2+y^2=1$, which we treat as the "configuration space" $\Omega$ of the rotor. If we choose the $x$ coordinate of the rotor as our CV, the CV map is $\hat{x}:(x,y)\mapsto x$. The marginal density $\mathscr{P}(x)$ of the CV is then

$$\begin{aligned}\mathscr{P}(x) &= \int_{-\infty}^{\infty} \mathrm{d}y \exp\left[-\tfrac{1}{2}\beta\kappa(x^2+y^2-1)^2\right] \\ &= (\beta\kappa)^{-1/4}\exp\left[-\tfrac{1}{2}\beta\kappa(x^2-1)^2\right]\int_0^{\infty}\mathrm{d}t\, t^{-1/2}\exp\left[-\tfrac{1}{2}t^2-\sqrt{\beta\kappa}(x^2-1)\,t\right] \\ &= \sqrt{\pi}(\beta\kappa)^{-1/4}\exp\left[-\tfrac{1}{4}\beta\kappa(x^2-1)^2\right]D_{-1/2}\left[\sqrt{\beta\kappa}(x^2-1)\right],\end{aligned} \tag{3.64}$$

where $D_{-1/2}[\cdot]$ is the parabolic cylinder function[21] [129, Eq. 12.5.1]. Using the above expression for $\mathscr{P}(x)$, and choosing $x=0$ as the point of zero free energy, we find the free-energy difference to be

$$\begin{aligned}\mathscr{A}(x)-\mathscr{A}(0) &= -\beta^{-1}\log\mathscr{P}(x)+\beta^{-1}\log\mathscr{P}(0) \\ &= \tfrac{1}{4}\kappa(x^4-2x^2)-\beta^{-1}\log\left\{\frac{D_{-1/2}\left[\sqrt{\beta\kappa}(x^2-1)\right]}{D_{-1/2}\left(-\sqrt{\beta\kappa}\right)}\right\}.\end{aligned} \tag{3.65}$$

This is an exact expression for the free-energy difference, represented by the black curve in Fig. 3.12(a). It is not readily obvious why a rotor with a rotationally symmetric potential would have a free-energy landscape that is not flat, particularly when it is clear that the rotor samples all orientations uniformly. To make things even more confusing, from Fig. 3.12(a), one might even think that the rotor spends more time near $x=\pm1$ compared to other points. But this is misleading. The only thing we can say from the rotor's free-energy landscape is that measured values of its $x$ coordinate have a propensity to be close to $x=\pm1$. Indeed, a much better CV for the thermalized rotor would be the polar angle $\hat{\theta}=\tan^{-1}(y/x)$. With this CV, we would find that $\mathscr{P}(\theta)$ is a constant, resulting in a flat free-energy landscape.

A similar issue arises when one evaluates the integral in Eq. (3.64) asymptotically for large $\beta$ using the harmonic approximation, Eq. (3.19). To do that, we need to parameterize $\Omega$ in

[20] We could have also chosen the more conventional energy function $U(x,y)=\kappa(\sqrt{x^2+y^2}-1)^2$, but that makes the integrals that follow difficult to evaluate exactly.

[21] We can also write $\mathscr{P}(x)$ in terms of the modified Bessel function of the second kind $K_{1/4}(\cdot)$ after making use of the relation $D_{-1/2}(z)=\sqrt{z/(2\pi)}K_{1/4}(z^2/4)$ [129, Eq. 12.7.10]. This is the result that computer algebra systems often produce.

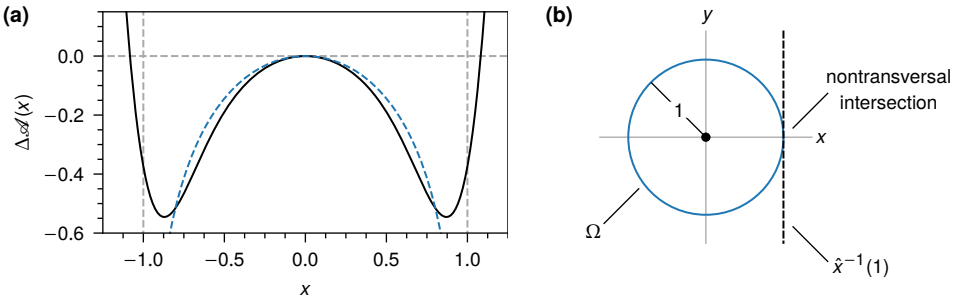


**Figure 3.12.** (a) Free energy $\Delta\mathscr{A}(x)$ of a stiff rotor at inverse temperature $\beta = 100$, from the exact expression [Eq. (3.65); black curve] and asymptotic expression [Eq. (3.70); dashed, blue curve]. Note how the asymptotic expression has a logarithmic divergence at $x = \pm 1$. Also note that the free-energy minimum is *not* at $x = \pm 1$ as one might guess (or how the asymptotic expression suggests). (b) Configuration space $\Omega$ of the rotor and the CV level set $\hat{x}^{-1}(x)$ have a nontransversal intersection at $x = \pm 1$.

terms of the coordinate $x$ and compute the nonzero eigenvalues of the dynamical matrix along $\Omega$. Two (standard) parameterizations in $x$ that cover a unit circle are

$$\psi_\pm(x) = \left(x, \pm\sqrt{1-x^2}\right). \tag{3.66}$$

The determinant of the corresponding induced metrics is

$$|\det(\nabla\psi_\pm)^\mathsf{T}\nabla\psi_\pm| = \left(1-x^2\right)^{-1}. \tag{3.67}$$

Meanwhile, the dynamical matrix is

$$\mathsf{D} = \nabla\nabla U(x, \pm\sqrt{1-x^2}) = 4\kappa \begin{pmatrix} x^2 & \pm x\sqrt{1-x^2} \\ \pm x\sqrt{1-x^2} & 1-x^2 \end{pmatrix}, \tag{3.68}$$

which has only one nonzero eigenvalue, $4\kappa$. Making use of Eq. (3.19), in the large-$\beta$ limit, we get the asymptotic marginal density[22]

$$\mathscr{P}(x) \sim \sqrt{\frac{2\pi}{\beta\kappa(1-x^2)}}. \tag{3.69}$$

Aside from normalization factors, $\mathscr{P}(x)$ above is the probability density of $x$ coordinates of points sampled uniformly (e.g., with respect to the polar angle) on a unit circle. This leads to a free-energy difference of the form

$$\mathscr{A}(x) - \mathscr{A}(0) \sim \beta^{-1}\log\sqrt{1-x^2}. \tag{3.70}$$

This free energy has a logarithmic divergence as $x \to \pm 1$. This can be traced back to the coordinate singularity [103] at $x = \pm 1$ when we parameterize a circle using $x$. Geometrically,

[22] Here we have added contributions from both $\psi_+$ and $\psi_-$.

as we show in the cartoon in Fig. 3.12(b), this corresponds to a nontransversal intersection between the CV level sets $\hat{x}^{-1}(\pm 1)$ and the configuration space $\Omega$.

In summary, the above discussions on CV diffeomorphisms and the example of the thermalized rotor shows that free-energy landscapes have to be carefully interpreted, taking into consideration the physical meaning of the chosen CV.

Chapter 4

# Semiclassical physics

This chapter presents a quick rundown of the semiclassical approximation as applied to multicomponent waves. Using a variational approach, we will derive the ray equations and describe the modifications necessary to extract the bound-state spectrum of a multicomponent wave operator through quantization. The general theory prescribed in this chapter will form the basis for our study of elastic waves in the next chapter.

Understanding wave propagation is vital to many disciplines across all branches of science and engineering. A widely employed asymptotic method to solve wave problems is the semiclassical approximation, which is sometimes referred to as the WKB approximation, the eikonal approximation, geometrical-optics limit, etc. Although the WKB approximation is discussed in almost all quantum mechanics books, several subtle issues, especially when multicomponent waves are involved, need to be addressed. Our primary goal in this chapter, therefore, is to review some existing results on the semiclassical approximation as applied to multicomponent wave fields. In particular, we will derive expressions for two extra phases that appear in the Bohr–Sommerfeld quantization rule for such wave fields. To keep the exposition simple, we will restrict ourselves to wave equations in one variable. For more detailed descriptions, we refer to the book by Tracy *et al.* [163] and references therein.

## 4.1 Introduction

We begin by considering a general wave equation of the form

$$\partial_t^2 \Psi(x,t) + \widehat{\mathsf{H}}\Psi(x,t) = 0, \tag{4.1}$$

where $\Psi(x,t)$ is an $N$-component wave field described by a one-dimensional coordinate $x$ and time $t$. As as example, in elastodynamics, $\Psi$ is usually composed of displacements from the neutral, undeformed state of some elastic structure, e.g., a rod or a shell. We also take $\widehat{\mathsf{H}}$ to be a Hermitian operator in the form of an $N \times N$ matrix, composed solely of spatial derivatives (i.e., powers of $\partial_x$) with time-independent coefficients. Assuming that the waves are time harmonic with frequency $\omega$, i.e., $\Psi(x,t) = \psi(x)e^{\pm i\omega t}$, where $\psi(x)$ is the time-independent part of the wave field, Eq. (4.1) can be recast as

$$\widehat{\mathsf{D}}\psi = 0, \quad \text{with} \quad \widehat{\mathsf{D}} = \widehat{\mathsf{H}} - \omega^2 \mathsf{I}_N, \tag{4.2}$$

where $\mathsf{I}_N$ is the $N \times N$ identity matrix.

If the coefficients of the spatial derivatives that appear in $\widehat{\mathsf{D}}$ are constants, then the eigenmodes $\psi$ are plain waves. In what follows we assume that these coefficients are slowly varying, with the variation controlled by a single positive parameter $\epsilon \ll 1$, called the eikonal parameter. It is useful to treat $\epsilon$ as an ordering parameter so that we can look for solutions at various orders of $\epsilon$. To this end, we rescale $x \to \epsilon^{-1}x$ so that a derivative $\partial_x$ becomes $\epsilon\partial_x$. With analogy to quantum mechanics, this allows us to recast the derivatives in $\widehat{\mathsf{D}}$ in terms of the wave number/momentum operator $\hat{k} = -i\epsilon\partial_x$, with $\epsilon$ playing the role of Planck's constant. Since we shall be considering $\widehat{\mathsf{D}}$ in the coordinate representation, the position operator $\hat{x} = x$.

### 4.1.1 Wave action

Central to the variational approach we will use in this chapter is the realization that a wave equation like Eq. (4.2) can be derived from a wave action of the form [83]

$$\begin{aligned}\mathscr{U}\left[\psi^*,\psi\right] &= \tfrac{1}{2}\langle\psi|\,\widehat{\mathsf{D}}\,|\psi\rangle = \tfrac{1}{2}\int \mathrm{d}x\,\mathrm{d}x'\,\langle\psi|x\rangle\,\langle x|\,\widehat{\mathsf{D}}\,|x'\rangle\,\langle x'|\psi\rangle \\ &= \tfrac{1}{2}\int \mathrm{d}x\mathrm{d}x'\,\psi^*_\mu(x)\,\widehat{\mathsf{D}}_{\mu\nu}(x,x')\,\psi_\nu(x').\end{aligned} \tag{4.3}$$

Above we have inserted the resolution of identity $\int \mathrm{d}x\,|x\rangle\langle x| = 1$ in appropriate places to express $\mathscr{U}$ in terms of $\psi$ and its conjugate $\psi^*$. Also, in the last step, we have made the matrix product between the matrix operator $\widehat{\mathsf{D}}$ and the wave vector $\psi$ explicit, with $\langle x|\widehat{\mathsf{D}}_{\mu\nu}|x'\rangle = \widehat{\mathsf{D}}_{\mu\nu}(x,x')$ being the matrix element of $\widehat{\mathsf{D}}_{\mu\nu}$ in the position basis. Varying $\mathscr{U}[\psi^*,\psi]$ with respect to $\psi^*$ gives the integral form of the eigenvalue problem, Eq. (4.2). For later use, we also note that $\mathscr{U}$ can also be written as

$$\begin{aligned}\mathscr{U}\left[\psi^*,\psi\right] &= \tfrac{1}{2}\int \mathrm{d}x'\,\langle\psi_\mu|x'\rangle\,\langle x'|\,\widehat{\mathsf{D}}_{\mu\nu}\,|\psi_\nu\rangle = \tfrac{1}{2}\int \mathrm{d}x'\,\langle x'|\,\widehat{\mathsf{D}}_{\mu\nu}\,|\psi_\nu\rangle\,\langle\psi_\mu|x'\rangle \\ &= \tfrac{1}{2}\int \mathrm{d}x'\,\langle x'|\,\widehat{\mathsf{D}}_{\mu\nu}\widehat{\mathsf{W}}_{\nu\mu}\,|x'\rangle = \tfrac{1}{2}\,\mathrm{tr}\left(\widehat{\mathsf{D}}_{\mu\nu}\widehat{\mathsf{W}}_{\nu\mu}\right),\end{aligned} \tag{4.4}$$

where the matrix operator $\widehat{\mathsf{W}}_{\nu\mu} = |\psi_\nu\rangle\langle\psi_\mu|$ is the density operator.

In order to solve Eq. (4.2) at various orders of $\epsilon$, it is convenient to make use of Weyl calculus, which allows one to map differential operators that are functions of $\hat{x}$ and $\hat{k}$ to ordinary functions, called Weyl symbols, defined on an $x$-$k$ phase space, and vice versa [17, 20]. A brief review of Weyl symbols is given in Appendix 4.A. Converting each entry of the matrix operator $\widehat{\mathsf{D}}$ into a Weyl symbol, we get the $N\times N$ dispersion matrix $\mathsf{D}$, which we express in various orders of $\epsilon$ as $\mathsf{D} = \mathsf{D}^{(0)} + \epsilon\mathsf{D}^{(1)} + \mathcal{O}(\epsilon^2)$.

The Wigner tensor $\mathsf{W}_{\nu\mu}$ is defined as the symbol of the density operator $\widehat{\mathsf{W}} = |\psi\rangle\langle\psi|$ with the matrix element $\widehat{\mathsf{W}}_{\nu\mu}(x',x) = \psi_\nu(x')\psi^*_\mu(x)$, and is given by

$$\mathsf{W}_{\nu\mu}(x,k) = \int \mathrm{d}r\, e^{-ikr/\epsilon}\,\psi_\nu\left(x+\tfrac{1}{2}r\right)\psi^*_\mu\left(x-\tfrac{1}{2}r\right). \tag{4.5}$$

By inverting the above expression, and setting $x+\frac{1}{2}r \to x'$ and $x-\frac{1}{2}r \to x$, we can write the kernel of the density operator in terms of its Weyl symbol as

$$\psi_\nu(x')\psi^*_\mu(x) = \frac{1}{2\pi\epsilon}\int \mathrm{d}k\, e^{ik(x'-x)/\epsilon}\,\mathsf{W}_{\nu\mu}\left[\tfrac{1}{2}(x'+x),k\right]. \tag{4.6}$$

In a similar fashion, we find

$$\widehat{\mathsf{D}}_{\mu\nu}(x,x') = \frac{1}{2\pi\epsilon}\int \mathrm{d}l\, e^{il(x-x')/\epsilon}\,\mathsf{D}_{\mu\nu}\left[\tfrac{1}{2}(x+x'),l\right]. \tag{4.7}$$

Putting Eq. (4.7) and (4.6) in Eq. (4.3), we arrive at

$$\mathscr{U} = \frac{1}{2(2\pi\epsilon)^2}\int \mathrm{d}x\,\mathrm{d}x'\,\mathrm{d}k\,\mathrm{d}l\, e^{i(l-k)(x-x')/\epsilon}\,\mathsf{D}_{\mu\nu}\left[\tfrac{1}{2}(x+x'),l\right]\mathsf{W}_{\nu\mu}\left[\tfrac{1}{2}(x'+x),k\right]. \tag{4.8}$$

Performing the change of variables $x \to \frac{1}{2}(x+x')$ and $x' \to x-x'$, which carries a Jacobian factor of unity, we arrive at

$$\begin{aligned}\mathscr{U} &= \frac{1}{2(2\pi\epsilon)^2}\int \mathrm{d}x\,\mathrm{d}x'\,\mathrm{d}k\,\mathrm{d}l\, e^{i(l-k)x'/\epsilon}\,\mathsf{D}_{\mu\nu}(x,l)\,\mathsf{W}_{\nu\mu}(x,k)\\ &= \frac{1}{4\pi\epsilon}\int \mathrm{d}x\,\mathrm{d}k\,\mathsf{D}_{\mu\nu}(x,k)\,\mathsf{W}_{\nu\mu}(x,k).\end{aligned} \tag{4.9}$$

So far there have been no approximations and the wave action in Eq. (4.9) is exact to all orders in the eikonal parameter. That said, the curious reader may wonder if there is an inconsistency here, in light of Eq. (4.4), where we have expressed $\mathscr{U}$ as the trace of the (scalar) operator $\widehat{O} = \widehat{\mathsf{D}}_{\mu\nu}\widehat{\mathsf{W}}_{\nu\mu}$. In terms of its symbol, the trace of an operator $\widehat{O}$ is

$$\operatorname{tr}\widehat{O} = \frac{1}{2\pi\epsilon}\int \mathrm{d}x\,\mathrm{d}k\,O(x,k), \tag{4.10}$$

so $\mathscr{U}$ must be

$$\mathscr{U} = \operatorname{tr}\widehat{O} = \frac{1}{4\pi\epsilon}\int \mathrm{d}x\,\mathrm{d}k\,\mathsf{D}_{\mu\nu}(x,k)\,e^{\frac{i\epsilon}{2}\widehat{\mathcal{L}}}\mathsf{W}_{\nu\mu}(x,k), \tag{4.11}$$

where we have used the Moyal formula (see Appendix 4.A for details) to write the symbol $O$ in terms of the symbols of $\mathsf{D}_{\mu\nu}$ and $\mathsf{W}_{\nu\mu}$. The issue here is a bit subtle—the resolution is that each correction term in the Moyal series can be expressed as divergence of a vector field in the $(x,k)$ phase space [163, Problem 3.16]. Hence, using the divergence theorem, the integrals over the correction terms vanish for all well-behaved wave fields. Thus, only the first term in the Moyal series remains, and we get Eq. (4.9) again.

Returning to our main problem, i.e., to employ the semiclassical approximation to solve Eq. (4.2), we will proceed as follows: (i) insert the eikonal ansatz into the wave action; (ii) expand the action to the lowest order in the eikonal parameter and form the *reduced action* $\mathscr{U}_{\mathrm{R}}$; (iii) extract the semiclassical equations of motions by varying the reduced action.

#### *Reduced wave action*

As we have assumed that the coefficients of the differential operators in $\widehat{\mathsf{D}}$ are slowly varying, it makes sense to look for an almost plain-wave-like solution of the form $\psi(x) = A(x)e^{iS(x)/\epsilon}$, where the amplitude $A(x)$ is an $N$-component spinor with complex components, and $S(x)/\epsilon$ is a rapidly varying phase, playing the role of an action. This is the *eikonal* ansatz and $\epsilon$ is the eikonal parameter, which controls the slowness of the variation. To employ the eikonal ansatz, we set $\psi_\mu = A_\mu e^{iS(x)/\epsilon}$ in Eq. (4.9) to obtain

$$\mathsf{W}_{\nu\mu}(x,k) = \int \mathrm{d}r\, e^{-ikr/\epsilon}\,A_\nu\left(x+\tfrac{1}{2}r\right)A^*_\mu\left(x-\tfrac{1}{2}r\right)\exp\left\{\frac{i}{\epsilon}\left[S\left(x+\tfrac{1}{2}r\right)-S\left(x-\tfrac{1}{2}r\right)\right]\right\}. \tag{4.12}$$

Next, we set $r \to \epsilon r$ and expand the amplitude $A_\nu(x+\frac{1}{2}\epsilon r) = A_\nu(x) + \frac{1}{2}(\partial_x A_\nu)\epsilon r + \mathcal{O}(\epsilon^2)$ and the phase $S(x+\frac{1}{2}\epsilon r) = S(x) + \frac{1}{2}[\partial_x S(x)]\epsilon r + \mathcal{O}(\epsilon^2)$ to obtain

$$\begin{aligned} \mathsf{W}_{\nu\mu}(x,k) &= \epsilon \int \mathrm{d}r\, A_\nu(x) A^*_\mu(x) \exp\{ir\,[\partial_x S(x) - k]\} + \mathcal{O}(\epsilon^2) \\ &= 2\pi\epsilon\, A_\nu(x) A^*_\mu(x) \delta\,[k - \partial_x S(x)] + \mathcal{O}(\epsilon^2). \end{aligned} \tag{4.13}$$

We have assumed that the dispersion matrix is ordered in the eikonal parameter $\epsilon$ so that we can write $\mathsf{D} = \mathsf{D}^{(0)} + \epsilon\mathsf{D}^{(1)} + \mathcal{O}(\epsilon^2)$. Hence, after putting the "reduced" Wigner tensor, Eq. (4.13), in Eq. (4.9), we find the action to the lowest order as

$$\mathscr{U} = \frac{1}{2}\int \mathrm{d}x\, \mathsf{D}^{(0)}_{\mu\nu}\,[x, k = \partial_x S(x)]\, A_\nu(x) A^*_\mu(x) + \mathcal{O}(\epsilon^2). \tag{4.14}$$

Next, we perform a spectral decomposition of the lowest-order dispersion matrix $\mathsf{D}^{(0)}_{\mu\nu}(x,k)$ and write it in terms of its eigenvectors $\tau_a$ and eigenvalues $\lambda_a$ as[1]

$$\mathsf{D}^{(0)}_{\mu\nu}(x,k) = \sum_a \lambda_a(x,k)\tau_{a,\mu}(x,k)\tau^*_{a,\nu}(x,k). \tag{4.15}$$

An eigenvector $\tau_a$ describes a specific wave type or "polarization" represented by the wave equation, Eq. (4.2). And by polarization, we mean a linear subspace of the total wave field that is usually of a distinct physical nature, e.g., flexural waves on a rod. We put Eq. (4.15) into Eq. (4.14), and define $B_a = \tau^*_{a,\nu} A_\nu = \langle \tau_a | A \rangle$, which is the projection of the amplitude $A_\nu$ along the direction of the polarization vector $\tau_a$. This results in a reduced action of the form

$$\mathscr{U}_\mathrm{R}\,[B_a, S] = \frac{1}{2}\sum_a \int \mathrm{d}x\, \lambda_a(x,k)|B_a(x,k)|^2, \quad \text{with} \quad k = \partial_x S(x). \tag{4.16}$$

### 4.1.2 Phase-space representation of waves

Variation of the reduced action, Eq. (4.16), with respect to the amplitude $B_a$ yields

$$\lambda_a[x, k(x)]\, B_a[x, k(x)] = 0 \quad \text{for all } a. \tag{4.17}$$

Assume that the amplitudes $B_a$ is nonzero only for one polarization, say, $a = I$. This means that the corresponding eigenvalue must vanish, which yields

$$\lambda_I\,[x, k(x)] = \lambda_I\,[x, \partial_x S(x)] = 0. \tag{4.18}$$

The above equation, which is known as the eikonal equation, is nothing but the Hamilton–Jacobi equation with $S(x)$ playing the role of the classical action. This leads us to the phase-space representation of waves as rays that satisfy the Hamilton's equations

$$\frac{\mathrm{d}x}{\mathrm{d}\sigma} = \partial_k \lambda_I(x,k) = \{x, \lambda_I\} \quad \text{and} \quad \frac{\mathrm{d}k}{\mathrm{d}\sigma} = -\partial_x \lambda_I(x,k) = \{k, \lambda_I\}, \tag{4.19}$$

where $\sigma$ is some parameter that parameterizes the rays $[x(\sigma), k(\sigma)]$ in the phase space, and $\{\cdot,\cdot\}$ is the $x$-$k$ Poisson bracket. Since the waves we consider propagate in a one-dimensional space, these rays are identical to the level curve defined by $\lambda_I(x,k) = 0$.

[1] In Eq. (4.15), the subscripts $\mu$ and $\nu$ indicate the components of $\tau_a$. We have also made the summation over the polarization index $a$ explicit—a convention that we will use throughout this dissertation.

In higher dimensions, solving the Hamilton-Jacobi equation can be nontrivial. However, in one dimension, it is always theoretically possible to solve for $k(x)$ from $\lambda_I(x,k)=0$. Variation of the action, Eq. (4.16), with respect to the phase $S(x)$ yields the amplitude transport equation[2]

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[|B_I(x)|^2\partial_k\lambda_I^{(0)}(x,k)\right]=0 \tag{4.20}$$

so $|B_I(x)|^2\partial_k\lambda_I^{(0)}(x,k)=C$ (constant). As the polarization vector $\tau_I$ and the wave amplitude $A$ are general complex vectors, $B_I=\langle\tau_I|A\rangle$ would generally be a complex number. However, by rephasing $\tau_I$ appropriately, we can always take $B_I$ to be real, and we find

$$B_I(x)=\frac{C}{\sqrt{\partial_k\lambda_I(x,k)}}. \tag{4.21}$$

Clearly, the amplitude projection $B_I$ becomes singular at points on the phase space where $\dot{x}=\partial_k\lambda_I(x,k)=0$, i.e., when the local tangent along a ray points along the $k$ axis. Such points are called caustics, and they arise because of ill-defined projections from the ray, which lives in the $x$-$k$ phase space, to the $x$ coordinate space. The issue can be sidestepped using the Keller–Maslov method [85, 113] where one Fourier transforms the wave field and uses the eikonal ansatz in the momentum space. Although we shall later rely on some results of this procedure, we will not discuss its technical details here. Interested readers are encouraged to look at the book by Tracy *et al.* [163, Chapter 6] (and references therein) for more information.

In writing down the eikonal equation, Eq. (4.18), we have assumed that the amplitude projection $B_a\neq 0$ only for $a=I$. Now, as $\mathsf{D}^{(0)}$ is a Hermitian matrix, its eigenvectors $\tau_a$ are mutually orthogonal, and we can write the wave amplitude as $A_\mu=\sum_a B_a\tau_{a,\mu}=B_I\tau_{I,\mu}$. This would clearly be wrong if more than one eigenvalue of the dispersion matrix $\mathsf{D}^{(0)}$ simultaneously vanish, because $A_\mu$ could then be proportional to any vector in the space spanned by the degenerate polarization vectors. For this reason, conventional semiclassical solutions also break down near points in the phase space where $\mathsf{D}^{(0)}$ is degenerate, leading to an exchange of energy and mode conversion between waves of different polarizations [163]. As the wave equations we study in this dissertation do not suffer from mode-conversion issues (at least in the parameter ranges we are interested in), we forgo a more detailed discussion of these problems here.

Assuming that we are far away from caustics and ignoring mode-conversion effects, one might write the final eikonal solution as

$$\psi_\mu(x)\sim B_I(x)\,\tau_{I,\mu}(x)e^{iS(x)/\epsilon}. \tag{4.22}$$

But there is a subtle problem. The polarization vector $\tau_I$ is only defined up to an overall phase. Thus, instead of the above equation, it is more appropriate to write the eikonal solution as

$$\psi_\mu(x)\sim B_I(x)\,\tau_{I,\mu}e^{i\gamma(x)}e^{iS(x)/\epsilon}, \tag{4.23}$$

[2]Here we have ignored the term $\partial_k B_I^2(x,k)\lambda_I$, which trivially vanishes as $\lambda_I(x,k)=0$ along a ray. Also, since $k$ is a function of $x$ on a ray, we can simply write the projected amplitude as $B(x)$ instead of $B(x,k)$. Note also that after evaluating the partial derivative of $\lambda_I$ with respect to $k$, we need to set $k\to k(x)$ for the full derivative with respect to $x$ to make sense.

where $\gamma(x) = \gamma[x, k = \partial_x S(x)]$ stands for an adiabatic phase that we have not determined so far. As $\tau_I$ is obtained from $\mathsf{D}^{(0)}$, whose entries are slowly varying functions of $x$, it makes sense to consider the phase $\gamma(x)$ to be slowly varying as well. For this reason, $\gamma$ cannot be absorbed into the eikonal phase $S(x)/\epsilon$, which is rapidly varying.

The phase $\gamma$ becomes important when quantizing orbits—a step that boils down to the single valuedness of $\psi_\mu$ as we move around a closed orbit, i.e., the sum of $\gamma$ and $S/\epsilon$ must be a half-integral multiple of $2\pi$ (see Section 4.4). Unfortunately, our lowest-order analysis is insufficient to determine this phase. We cannot also find an evolution equation for this phase from the reduced action, Eq. (4.16), as it is independent of $\gamma$. It can also be easily checked that the phase $\gamma$ cannot be introduced in the action by rephasing $\tau$ by the phase factor $e^{i\gamma(x,k)}$, as the action remains invariant. The message is clear: to find $\gamma$, we need to include higher-order corrections in the wave action. To this end, rather than repeat the original, but cumbersome derivations scattered across Refs. [7, 8, 83, 174], it is more sagacious to switch gears and follow a different approach due to Littlejohn and coworkers [99, 100] for (approximately) diagonalizing multicomponent differential operators.

## 4.2 Intermission: diagonalizing multicomponent operators

We start again, with a Hermitian linear differential operator $\widehat{\mathsf{D}}$ in the form of an $N \times N$ matrix satisfying a wave equation

$$\widehat{\mathsf{D}}\Psi = 0, \tag{4.24}$$

where $\Psi$ is an $N$-component wave field. We want to find a unitary operator $\widehat{\mathsf{U}}$ such that

$$\widehat{\mathsf{U}}^\dagger \widehat{\mathsf{D}} \widehat{\mathsf{U}} = \widehat{\Lambda} \tag{4.25}$$

is a diagonal operator. Clearly, this equivalent to demanding that $\widehat{\mathsf{D}}\widehat{\mathsf{U}} = \widehat{\mathsf{U}}\widehat{\Lambda}$. If we can find such an operator and solve the decoupled set of equations given by $\widehat{\Lambda}\Phi = 0$ somehow, the solutions to the original equation can then be recovered from $\Phi$ using $\Psi = \widehat{\mathsf{U}}\Phi$. However, Eq. (4.25) is an equation involving operator products and standard linear algebra methods do not (directly) help in finding the unitary operator $\widehat{\mathsf{U}}$. Instead, we start by finding the symbol form of the following two equations:

$$\begin{aligned} \widehat{\mathsf{D}}\widehat{\mathsf{U}} &= \widehat{\mathsf{U}}\widehat{\Lambda} \\ \widehat{\mathsf{U}}^\dagger\widehat{\mathsf{U}} &= \hat{\mathsf{I}}_N. \end{aligned} \tag{4.26}$$

Here we assume that the operators $\widehat{\mathsf{D}}$, $\widehat{\mathsf{U}}$, and $\widehat{\Lambda}$ have some problem-relevant ordering parameter $\epsilon$ so that we can expand[3] the operators (and their symbols) in terms of $\epsilon$ as follows:

$$\begin{aligned} \widehat{\mathsf{D}} &= \widehat{\mathsf{D}}^{(0)} + \epsilon\widehat{\mathsf{D}}^{(1)} + \epsilon^2\widehat{\mathsf{D}}^{(2)} + \cdots & & \mathsf{D} = \mathsf{D}^{(0)} + \epsilon\mathsf{D}^{(1)} + \epsilon^2\mathsf{D}^{(2)} + \cdots \\ \widehat{\mathsf{U}} &= \widehat{\mathsf{U}}^{(0)} + \epsilon\widehat{\mathsf{U}}^{(1)} + \epsilon^2\widehat{\mathsf{U}}^{(2)} + \cdots & \text{and} \quad & \mathsf{U} = \mathsf{U}^{(0)} + \epsilon\mathsf{U}^{(1)} + \epsilon^2\mathsf{U}^{(2)} + \cdots \\ \widehat{\Lambda} &= \widehat{\Lambda}^{(0)} + \epsilon\widehat{\Lambda}^{(1)} + \epsilon^2\widehat{\Lambda}^{(2)} + \cdots & & \Lambda = \Lambda^{(0)} + \epsilon\Lambda^{(1)} + \epsilon^2\Lambda^{(2)} + \cdots \end{aligned} \tag{4.27}$$

[3] In their papers, Littlejohn and coworkers [99, 168] assume that $\widehat{\mathsf{D}}$ is *not* ordered in $\epsilon$, i.e., $\widehat{\mathsf{D}} = \widehat{\mathsf{D}}_0$. This might seem confusing at first since the parameter $\epsilon$ must appear somewhere in the problem. The thing is that very often, $\epsilon$ appears as a factor to a spatial derivative operator, e.g., $\epsilon\partial_x$, which once expressed in terms of the momentum operator, does not have additional dependence on $\epsilon$. Of course, $\widehat{\mathsf{D}}$ could have further nontrivial dependence on $\epsilon$, which is a more general situation, and the one we want to consider here.

Because the Weyl transform preserves Hermiticity, $\mathsf{D}$ is a Hermitian matrix to all orders. Our requirement is that $\Lambda$ remains diagonal at all orders of $\epsilon$. To proceed, we use the Moyal star product to find the symbol form of Eq. (4.26) at different orders of $\epsilon$. At $\mathcal{O}(\epsilon^0)$ we find $\mathsf{D}^{(0)}\mathsf{U}^{(0)} = \mathsf{U}^{(0)}\Lambda^{(0)}$ and $\mathsf{U}^{(0)\dagger}\mathsf{U}^{(0)} = \mathsf{I}_n$, with the first equation being equivalent to

$$\mathsf{U}^{(0)\dagger}\mathsf{D}^{(0)}\mathsf{U}^{(0)} = \Lambda^{(0)}. \tag{4.28}$$

Since we demand $\Lambda^{(0)}$ to be diagonal, this is the usual linear algebra problem of diagonalizing the matrix $\mathsf{D}^{(0)}$. We thus deduce that the columns of the matrix $\mathsf{U}^{(0)}$ are composed of eigenvectors $\tau_a$ with eigenvalues $\lambda_a^{(0)}$ satisfying $\mathsf{D}^{(0)}\tau_a = \lambda_a^{(0)}\tau_a$, with $a = 1, 2, \ldots, N$. More explicitly,

$$\mathsf{U}^{(0)} = \begin{pmatrix} \tau_1 & \tau_2 & \cdots & \tau_n \end{pmatrix}. \tag{4.29}$$

The matrix $\Lambda^{(0)} = \mathrm{diag}[\lambda_1^{(0)}, \lambda_2^{(0)}, \ldots, \lambda_N^{(0)}]$ is composed of the eigenvalues $\lambda_a^{(0)}$.

At $\mathcal{O}(\epsilon^1)$, demanding $\mathsf{DU} = \mathsf{U}\Lambda$ gives us

$$\mathsf{D}^{(1)}\mathsf{U}^{(0)} + \mathsf{D}^{(0)}\mathsf{U}^{(1)} + (i/2)\left\{\mathsf{D}^{(0)}, \mathsf{U}^{(0)}\right\} = \mathsf{U}^{(1)}\Lambda^{(0)} + \mathsf{U}^{(0)}\Lambda^{(1)} + (i/2)\left\{\mathsf{U}^{(0)}, \Lambda^{(0)}\right\}. \tag{4.30}$$

Multiplying from the left by $\mathsf{U}^{(0)\dagger}$ and making use of $\mathsf{U}^{(0)\dagger}\mathsf{D}^{(0)} = \Lambda^{(0)}\mathsf{U}^{(0)\dagger}$, we get the $\mathcal{O}(\epsilon)$ correction[4]

$$\Lambda^{(1)} = \left(\mathsf{U}^{(0)\dagger}\mathsf{D}^{(1)}\mathsf{U}^{(0)} + (i/2)\mathsf{U}^{(0)\dagger}\left\{\mathsf{D}^{(0)}, \mathsf{U}^{(0)}\right\} - (i/2)\mathsf{U}^{(0)\dagger}\left\{\mathsf{U}^{(0)}, \Lambda^{(0)}\right\}\right) + \left[\Lambda^{(0)}, \mathsf{U}^{(0)\dagger}\mathsf{U}^{(1)}\right]. \tag{4.31}$$

In the last term above, $[\cdot,\cdot]$ denotes the matrix commutator. We can find $\mathsf{U}^{(0)}$ and $\Lambda^{(0)}$ by diagonalizing $\mathsf{D}^{(0)}$. The matrix $\mathsf{D}^{(1)}$ (if it is nonzero) can be found by expanding the symbol matrix $\mathsf{D}$. But that will still not let us find $\Lambda^{(1)}$ since we also need $\mathsf{U}^{(1)}$ to evaluate the commutator term $[\Lambda^{(0)}, \mathsf{U}^{(0)\dagger}\mathsf{U}^{(1)}]$. To proceed, we recall that we want $\Lambda$ to be diagonal at all orders, which means that $\Lambda^{(1)}$ must be a diagonal matrix as well. The $ab$th entry of the commutator term evalutes to

$$\left[\Lambda^{(0)}, \mathsf{U}^{(0)\dagger}\mathsf{U}^{(1)}\right]_{ab} = \left[\lambda_a^{(0)} - \lambda_b^{(0)}\right]\left(\mathsf{U}^{(0)\dagger}\mathsf{U}^{(1)}\right)_{ab}, \tag{4.32}$$

where we have made use of the diagonality of $\Lambda^{(0)}$, which implies $\Lambda_{ab}^{(0)} = \lambda_a^{(0)}\delta_{ab}$. And we see that the diagonal entries (with $a = b$) of the commutator vanish. So the commutator term does not contribute towards the diagonal entries of $\Lambda^{(1)}$ and we can find the $\mathcal{O}(\epsilon^1)$ correction $\lambda_a^{(1)}$ to the eigenvalues by carefully evaluating diagonal entries of the remaining terms in the RHS of Eq. (4.31). Before we do that, we need to discuss the role of the commutator term further. In fact, without this crucial term, the expansion will break down.

### 4.2.1 Role of the commutator term

None of our arguments so far guarantees that $\Lambda^{(1)}$ is diagonal. In fact, the off-diagonal entries of the four matrices in the RHS of Eq. (4.31) is generally not equal to zero. The only way for $\Lambda^{(1)}$ to be diagonal then is if these off-diagonal entries somehow cancel each other. We do not have the freedom to choose the off-diagonal entries in the first three terms since they only involve the matrices $\mathsf{U}^{(0)}$ and $\Lambda^{(0)}$, both of which are constrained by Eq. (4.28). However, the commutator term involves the matrix $\mathsf{U}^{(1)}$, which is something we have not found yet. At the

[4] For further higher-order corrections to $\Lambda$, see Eq. (19) of Ref. [168].

same time, $\mathsf{U}^{(1)}$ is not a completely arbitrary matrix, because on demanding unitarity of $\mathsf{U}$ at $\mathcal{O}(\epsilon)$ we get the following equation that puts constraints on $\mathsf{U}^{(1)}$:

$$\mathsf{U}^{(0)\dagger}\mathsf{U}^{(1)} + \mathsf{U}^{(1)\dagger}\mathsf{U}^{(0)} + (i/2)\left\{\mathsf{U}^{(0)\dagger}, \mathsf{U}^{(0)}\right\} = 0. \tag{4.33}$$

As we show below, this equation is not good enough to determine $\mathsf{U}^{(1)}$ completely, which is good for us since that gives us some freedom in choosing the off-diagonal elements of the commutator term in the way we want. To simplify the commutator term we first define $\mathsf{X} = \mathsf{U}^{(0)\dagger}\mathsf{U}^{(1)}$ so that the previous equation becomes

$$\mathsf{X} + \mathsf{X}^\dagger = -(i/2)\left\{\mathsf{U}^{(0)\dagger}, \mathsf{U}^{(0)}\right\}. \tag{4.34}$$

We see that Hermitian part of $\mathsf{X}$, given by $\mathsf{A} = (\mathsf{X} + \mathsf{X}^\dagger)/2$, is fixed by the above equation. Hence $\mathsf{A} = (-i/4)\{\mathsf{U}^{(0)\dagger}, \mathsf{U}^{(0)}\}$, which is clearly a Hermitian matrix. This still leaves us with the possibility of picking the anti-Hermitian part of $\mathsf{X}$, which we denote by $i\mathsf{B}$. Here $\mathsf{B}$ is a Hermitian matrix defined[5] by $\mathsf{B} = -i(\mathsf{X} - \mathsf{X}^\dagger)/2$. After replacing $\mathsf{X} = \mathsf{U}^{(0)\dagger}\mathsf{U}^{(1)}$ in the commutator term with $\mathsf{A} + i\mathsf{B}$ and using Eqs. (4.31) and (4.32), we find the off-diagonal entries of $\mathsf{B}$ to be

$$\mathsf{B}_{ab} = i\frac{\mathsf{Q}_{ab}}{\lambda_a^{(0)} - \lambda_b^{(0)}}, \quad \text{with } a \neq b, \tag{4.35}$$

where

$$\begin{aligned}\mathsf{Q} = {} & \mathsf{U}^{(0)\dagger}\mathsf{D}^{(1)}\mathsf{U}^{(0)} + (i/2)\mathsf{U}^{(0)\dagger}\left\{\mathsf{D}^{(0)}, \mathsf{U}^{(0)}\right\} - (i/2)\mathsf{U}^{(0)\dagger}\left\{\mathsf{U}^{(0)}, \Lambda^{(0)}\right\} \\ & - (i/4)\Lambda\left\{\mathsf{U}^{(0)\dagger}, \mathsf{U}^{(0)}\right\} + (i/4)\left\{\mathsf{U}^{(0)\dagger}, \mathsf{U}^{(0)}\right\}\Lambda.\end{aligned} \tag{4.36}$$

Although we have found an expression that gives the off-diagonal elements of $\mathsf{B}$, nothing in our arguments so far guaratees the Hermiticity of $\mathsf{B}$ and we have to explicitly check this.[6] From Eq. (4.35), we see that the matrix $\mathsf{Q}$ should be Hermitian if $\mathsf{B}$ is to be Hermitian. For two general matrices $\mathsf{F}$ and $\mathsf{G}$ we have $\{\mathsf{F}, \mathsf{G}\}^\dagger = -\{\mathsf{G}^\dagger, \mathsf{F}^\dagger\}$. Along with the assumption that $\mathsf{D}^{(1)}$ is Hermitian, we then find,

$$\begin{aligned}\mathsf{Q}^\dagger = {} & \mathsf{U}^{(0)\dagger}\mathsf{D}^{(1)}\mathsf{U}^{(0)} + (i/2)\left\{\mathsf{U}^{(0)\dagger}, \mathsf{D}^{(0)}\right\}\mathsf{U}^{(0)} - (i/2)\left\{\Lambda^{(0)}, \mathsf{U}^{(0)\dagger}\right\}\mathsf{U}^{(0)} \\ & - (i/4)\left\{\mathsf{U}^{(0)\dagger}, \mathsf{U}^{(0)}\right\}\Lambda + (i/4)\Lambda\left\{\mathsf{U}^{(0)\dagger}, \mathsf{U}^{(0)}\right\},\end{aligned} \tag{4.37}$$

which can be further simplified[7] to show that $\mathsf{Q}^\dagger = \mathsf{Q}$, from which the Hermiticity of $\mathsf{B}$ follows.

[5] Writing the anti-Hermitian part of $\mathsf{X}$ this way might look nonstandard and it is more natural to write it as $(\mathsf{X} - \mathsf{X}^\dagger)/2$. This is to ensure that the diagonal entries of matrix $\mathsf{B}$ are real numbers, for the sake of a future argument.

[6] Littlejohn and coworkers [99, 168] seem to not have stressed this subtle point in their papers and they do not prove the Hermiticity of $\mathsf{B}$ explicitly. At first glance, we might think that $\mathsf{B}$ would Hermitian by construction because we *took* $\mathsf{A}$ and $i\mathsf{B}$ to be the Hermitian and anti-Hermitian parts of $\mathsf{X}$. This is not true however, as we obtained $\mathsf{A}$ from requiring unitarity of $\mathsf{U}$ at $\mathcal{O}(\epsilon)$ and we are now attempting to find the off-diagonal entries of $\mathsf{B}$ from Eq. (4.31), which is an independent equation that guarantees the diagonalizability of $\mathsf{D}$ at $\mathcal{O}(\epsilon)$.

[7] The simplification proceeds (by explicitly computing matrix entries) by first showing that $\{\mathsf{U}^{(0)\dagger}, \mathsf{D}^{(0)}\}\mathsf{U}^{(0)} = \{\mathsf{U}^{(0)\dagger}, \mathsf{U}^{(0)}\}\Lambda - \mathsf{U}^{(0)\dagger}\{\mathsf{U}^{(0)}, \Lambda\} - \{{}^1\mathsf{U}^{(0)\dagger}, {}^3\mathsf{U}^{(0)}\}{}^2\mathsf{D}^{(0)}$ and $\{\Lambda^{(0)}, \mathsf{U}^{(0)\dagger}\}\mathsf{U}^{(0)} = \Lambda^{(0)}\{\mathsf{U}^{(0)\dagger}, \mathsf{U}^{(0)}\} - \mathsf{U}^{(0)\dagger}\{\mathsf{D}^{(0)}, \mathsf{U}^{(0)}\} - \{{}^1\mathsf{U}^{(0)\dagger}, {}^3\mathsf{U}^{(0)}\}{}^2\mathsf{D}^{(0)}$. Here the superscripts 1, 2, and 3 appearing before the matrices denote the order in which they are to be multiplied; see Ref. [100] for further details on this convention. Upon using these results in the RHS of Eq. (4.37), we see that $\mathsf{Q} = \mathsf{Q}^\dagger$.

Clearly, from Eq. (4.35) we see that our scheme would break down if the lowest-order symbol matrix $\mathsf{D}^{(0)}$ has an $\mathcal{O}(\epsilon^0)$ degeneracy, i.e., if $\mathsf{D}^{(0)}$ is degenerate with $\lambda_a^{(0)} = \lambda_b^{(0)}$ for some $a \neq b$. For similar reasons, we cannot also determine the diagonal entries $\mathsf{B}_{aa}$ from Eq. (4.35). However, as we show below, we can always take them to be zero, since they do not affect the unitarity of $\mathsf{U}$ to $\mathcal{O}(\epsilon)$. The symbol matrix $\mathsf{U}$ to $\mathcal{O}(\epsilon)$ is

$$\mathsf{U} = \mathsf{U}^{(0)} + \epsilon \mathsf{U}^{(1)} + \mathcal{O}(\epsilon^2) = \mathsf{U}^{(0)} \left[ \mathsf{I}_n + \epsilon \left( \mathsf{A} + i\mathsf{B}' + i\mathsf{B}'' \right) \right] + \mathcal{O}(\epsilon^2), \tag{4.38}$$

where we have made use of $\mathsf{U}^{(1)} = \mathsf{U}^{(0)}(\mathsf{A} + i\mathsf{B})$ and have written $\mathsf{B}$ as the sum of its diagonal part $\mathsf{B}'$ and off-diagonal part $\mathsf{B}''$.[8] Now, note that we have complete freedom in choosing phase factors for the columns of $\mathsf{U}^{(0)}$, namely the eigenvectors of $\mathsf{D}^{(0)}$. Rephasing the $a$th column by $e^{-i\epsilon \mathsf{B}'_{aa}}$ turns

$$\mathsf{U}^{(0)} \to \begin{pmatrix} e^{-i\epsilon \mathsf{B}'_{11}} \tau_1 & e^{-i\epsilon \mathsf{B}'_{22}} \tau_2 & \cdots & e^{-i\epsilon \mathsf{B}'_{nn}} \tau_n \end{pmatrix} = \mathsf{U}^{(0)} (\mathsf{I}_n - i\epsilon \mathsf{B}') + \mathcal{O}(\epsilon^2) \tag{4.39}$$

At the same time, the matrices $\mathsf{A} \to \mathsf{A} + \mathcal{O}(\epsilon)$ and $\mathsf{B}'' \to \mathsf{B}'' + \mathcal{O}(\epsilon)$. Thus, the RHS of Eq. (4.38) becomes

$$\left[ \mathsf{U}^{(0)} (\mathsf{I}_n - i\epsilon \mathsf{B}') + \mathcal{O}(\epsilon^2) \right] \left[ \mathsf{I}_n + \epsilon \left( \mathsf{A} + i\mathsf{B}' + i\mathsf{B}'' + \mathcal{O}(\epsilon) \right) \right] = \mathsf{U}^{(0)} \left[ \mathsf{I}_n + \epsilon \left( \mathsf{A} + i\mathsf{B}'' \right) \right] + \mathcal{O}(\epsilon^2). \tag{4.40}$$

In other words, we can absorb any nonzero diagonal entries of $\mathsf{B}$ by a suitable rephasing of the columns of $\mathsf{U}^{(0)}$. Thus, without loss of generality we take the diagonal part $\mathsf{B}'$ to be zero.

### 4.2.2 First-order correction

The matrix $\Lambda^{(1)}$, whose entries can be found from the parenthetical terms in the RHS of Eq. (4.31), is now diagonal to $\mathcal{O}(\epsilon)$. The diagonal entries of $\Lambda^{(1)}$ give the first-order correction $\lambda_a^{(1)}$ to the eigenvalue $\lambda_a$ for a specific polarization, i.e., $\lambda_a^{(1)} = \Lambda_{aa}^{(1)}$. We proceed by further simplifying the various parenthetical terms in Eq. (4.31) by noting that

$$\Lambda^{(0)} = \operatorname{diag}\left[ \lambda_1^{(0)}, \lambda_2^{(0)}, \ldots, \lambda_N^{(0)} \right], \quad \mathsf{U}_{\mu a}^{(0)} = \tau_{a,\mu}, \quad \text{and} \quad \mathsf{U}_{a\mu}^{(0)\dagger} = \tau_{a,\mu}^*. \tag{4.41}$$

Above, $\tau_{a,\mu}$ refers to the $\mu$th component of the polarization vector $\tau_a$ and $(\cdot)^*$ denotes complex conjugation. Here, we have also used mixed Latin/Greek indices to separate the polarization index from the component indices (both run from 1 to $N$, however). Using Eq. (4.41), the diagonal components of the first term in the parenthetical expression in Eq. (4.31) can be written as

$$\left( \mathsf{U}^{(0)\dagger} \mathsf{D}^{(1)} \mathsf{U}^{(0)} \right)_{aa} = \tau_{a,\mu}^* \mathsf{D}_{\mu\nu}^{(1)} \tau_{a,\nu}. \tag{4.42}$$

[8] Since $\mathsf{B}$ is Hermitian, its diagonal part $\mathsf{B}'$ is a real matrix.

Next, we simplify the diagonal entries of the second parenthetical term of Eq. (4.31) as

$$\begin{aligned}(i/2)\left(\mathsf{U}^{(0)\dagger}\left\{\mathsf{D}^{(0)},\mathsf{U}^{(0)}\right\}\right)_{aa} &= (i/2)\mathsf{U}^{(0)\dagger}_{a\mu}\left(\partial_x \mathsf{D}^{(0)}_{\mu\nu}\partial_k \mathsf{U}^{(0)}_{\nu a} - \partial_k \mathsf{D}^{(0)}_{\mu\nu}\partial_x \mathsf{U}^{(0)}_{\nu a}\right)\\ &= (i/2)\left[\partial_x\left(\Lambda^{(0)}_{a\mu}\mathsf{U}^{(0)\dagger}_{\mu\nu}\right) - \partial_x \mathsf{U}^{(0)\dagger}_{a\mu}\mathsf{D}^{(0)}_{\mu\nu}\right]\partial_k \mathsf{U}^{(0)}_{\nu a}\\ &\quad - (i/2)\left[\partial_k\left(\Lambda^{(0)}_{a\mu}\mathsf{U}^{(0)\dagger}_{\mu\nu}\right) - \partial_k \mathsf{U}^{(0)\dagger}_{a\mu}\mathsf{D}^{(0)}_{\mu\nu}\right]\partial_x \mathsf{U}^{(0)}_{\nu a}\\ &= (i/2)\lambda^{(0)}_a\left\{\tau^*_{a,\mu},\tau_{a,\mu}\right\} - (i/2)\tau^*_{a,\mu}\left\{\tau_{a,\mu},\lambda^{(0)}_a\right\}\\ &\quad - (i/2)\mathsf{D}^{(0)}_{\mu\nu}\left\{\tau^*_{a,\mu},\tau_{a,\nu}\right\}.\end{aligned} \tag{4.43}$$

In the second step above, we have also made use of $\mathsf{U}^{(0)\dagger}\mathsf{D}^{(0)} = \Lambda^{(0)}\mathsf{U}^{(0)\dagger}$. Finally, we can write the diagonal entries of the third parenthetical term in Eq. (4.31) as

$$-(i/2)\left(\mathsf{U}^{(0)\dagger}\left\{\mathsf{U}^{(0)},\Lambda^{(0)}\right\}\right)_{aa} = -(i/2)\tau^*_{a,\mu}\left\{\tau_{a,\mu},\lambda^{(0)}_a\right\}. \tag{4.44}$$

Putting the simplified expressions from Eqs. (4.42)–(4.44) in Eq. (4.31), we find the first-order correction in the eigenvalue $\lambda$ to be[9]

$$\lambda^{(1)}_a = \tau^*_{a,\mu}\mathsf{D}^{(1)}_{\mu\nu}\tau_{a,\nu} - i\tau^*_{a,\mu}\left\{\tau_{a,\mu},\lambda^{(0)}_a\right\} - (i/2)\left(\mathsf{D}^{(0)}_{\mu\nu} - \lambda^{(0)}_a\delta_{\mu\nu}\right)\left\{\tau^*_{a,\mu},\tau_{a,\nu}\right\} \tag{4.45}$$

Note that in the above equation there is no sum over the polarization index $a$, whereas the repeated Greek indices $\mu$ and $\nu$ indicate summation. Apart from Eq. (4.45), Littlejohn and coworkers [99, 100, 168] go on to derive a number of other results in their papers. For our purposes, however, Eq. (4.45), which provides the $\mathcal{O}(\epsilon)$ correction to the eigenvalue $\lambda^{(0)}_a$ is good enough to derive an evolution equation for the adiabatic phase.

## 4.3 Evolution of the polarization phase

To make the transition between the variational theory presented in Section 4.1 and the results in the previous section easier, we first note that we can write the wave operator $\widehat{\mathsf{D}}$ as

$$\widehat{\mathsf{D}} = \widehat{\mathsf{U}}\widehat{\Lambda}\widehat{\mathsf{U}}^\dagger. \tag{4.46}$$

Putting this in Eq. (4.4), we see that the wave action can also be written as

$$\mathscr{U} = \tfrac{1}{2}\operatorname{tr}\left(\widehat{\mathsf{D}}\widehat{\mathsf{W}}\right) = \tfrac{1}{2}\operatorname{tr}\left(\widehat{\mathsf{U}}\widehat{\Lambda}\widehat{\mathsf{U}}^\dagger\left|\psi\right\rangle\left\langle\psi\right|\right) = \tfrac{1}{2}\operatorname{tr}\left(\widehat{\Lambda}\widehat{\mathsf{U}}^\dagger\left|\psi\right\rangle\left\langle\psi\right|\widehat{\mathsf{U}}\right) = \tfrac{1}{2}\sum_a \operatorname{tr}\left(\hat{\lambda}_a\left|\widetilde{\psi}_a\right\rangle\left\langle\widetilde{\psi}_a\right|\right) \tag{4.47}$$

where we have used standard results pertaining to the trace of matrix products and have defined $\left|\widetilde{\psi}_a\right\rangle = \widehat{\mathsf{U}}_{a\mu}\left|\psi_\mu\right\rangle$. Proceeding by similar arguments as in Section 4.1, we express the wave action in terms of the Weyl symbol of the operator $\hat{\lambda}_a$, given by $\lambda_a$, and the Wigner function $W_a$ corresponding to the density matrix $\left|\psi_a\right\rangle\left\langle\psi_a\right|$ to find [cf. Eq. (4.9)]

$$\mathscr{U} = \frac{1}{4\pi\epsilon}\sum_a \int \mathrm{d}x\,\mathrm{d}k\,\lambda_a(x,k)W_a(x,k). \tag{4.48}$$

[9] This equation is identical to Eq. (3.21) of Ref. [100], except for the first term involving $\mathsf{D}^{(1)}$, as these authors assume that the operator $\widehat{\mathsf{D}}$ is not ordered in $\epsilon$.

The Wigner function $W_a$ to the lowest is

$$W_a = \mathsf{U}^{(0)\dagger}_{av}\mathsf{W}_{v\mu}\mathsf{U}^{(0)}_{\mu a} + \epsilon(\square) + \mathcal{O}(\epsilon^2) \tag{4.49}$$

where $\epsilon(\square)$ represents all $\mathcal{O}(\epsilon)$ correction terms, e.g., those that appear during the application of the Moyal formula to compute $W_a$, terms involving $\mathsf{U}^{(1)}$, etc. One might think that it is important to find these terms, which is of the same order as the phase correction term we are after. However, as we shall see, these terms drop out once we work out the equations of motion, and for this reason, we leave them unevaluated. Next, we insert the eikonal ansatz for the wave fields, $\psi_\mu = A_\mu e^{iS(x)/\epsilon}$, and using the steps we followed in p. 58, find the reduced Wigner function to be

$$\begin{aligned} W_{a,\mathrm{R}} &= 2\pi\epsilon\big[\tau^*_{a,\nu}A_\nu(x)\big]\big[A^*_\mu(x)\tau_{a,\mu}\big]\delta\big[k-\partial_x S(x)\big] + \mathcal{O}(\epsilon^2) \\ &= 2\pi\epsilon|B_a(x,k)|^2\delta\big[k-\partial_x S(x)\big] + \epsilon(\square) + \mathcal{O}(\epsilon^2). \end{aligned} \tag{4.50}$$

Above, the correction $\epsilon(\square)$ now includes previous correction terms as well as those arising from higher-order expansions of the eikonal phase $S(x)$ and amplitude $A_\mu$. To $\mathcal{O}(\epsilon^2)$, the eigenvalue $\lambda_a = \lambda^{(0)}_a + \epsilon\lambda^{(1)}_a + \mathcal{O}(\epsilon)$, so the wave action becomes

$$\begin{aligned} \mathscr{U}\big[B_a, S\big] &= \mathscr{U}^{(0)}\,[B_a, S] + \epsilon\mathscr{U}^{(1)}\,[B_a, S] + \mathcal{O}(\epsilon^2) \\ &= \tfrac{1}{2}\sum_a \int \mathrm{d}x\,\lambda^{(0)}_a(x)|B_a(x)|^2 + \epsilon\lambda^{(0)}_a(x)\times(\square) + \tfrac{1}{2}\epsilon\sum_a\int \mathrm{d}x\,\lambda^{(1)}_a(x)|B_a(x)|^2 + \mathcal{O}(\epsilon^2). \end{aligned} \tag{4.51}$$

The lowest-order action above $\mathscr{U}^{(0)}[B_a, S]$ is equivalent to the reduced action, Eq. (4.16), we saw previously. Varying $\mathscr{U}^{(0)}[B_a, S]$ with respect to $B_a$ and $S(x)$, we find, as before, the eikonal and amplitude transport equations

$$\begin{aligned} \lambda^{(0)}_I\big[x, k = \partial_x S(x)\big] &= 0, \\ \frac{\mathrm{d}}{\mathrm{d}x}\Big[|B_I(x)|^2\partial_k\lambda^{(0)}_I(x,k)\Big] &= 0, \end{aligned} \tag{4.52}$$

where we have assumed again that only one eigenvalue, with polarization index $a = I$, vanishes.

Varying the first-order action $\mathscr{U}^{(1)}[B_a, S]$ in Eq. (4.51) with respect to $B_a$, we find

$$\lambda^{(0)}_I\delta(\square) + \lambda^{(1)}_I\,[x, k = \partial_x S(x)] = 0, \tag{4.53}$$

where $\delta(\square)$ represents the variation of the correction terms $(\square)$ that we chose to not determine. However, from the eikonal equation we know that $\lambda^{(0)}_I = 0$, so once we insert Eq. (4.45) into the above equation, we get

$$\tau^*_{I,\mu}\mathsf{D}^{(1)}_{\mu\nu}\tau_{I,\nu} - (i/2)\mathsf{D}^{(0)}_{\mu\nu}\left\{\tau^*_{I,\mu}, \tau_{I,\nu}\right\} - i\tau_{I,\mu}\left\{\tau_{I,\mu}, \lambda^{(0)}_a\right\} = 0. \tag{4.54}$$

Upon explicitly introducing the adiabatic phase $\gamma$ by setting $\tau_I \to \tau_I e^{i\gamma}$, and after some straightforward algebra, we find that the phase $\gamma$ can be written as

$$\gamma = \gamma_{\mathrm{G}} + \gamma_{\mathrm{NG}}, \tag{4.55}$$

where the phases $\gamma_{\mathrm{G}}$ and $\gamma_{\mathrm{NG}}$ evolve according to

$$\dot{\gamma}_{\mathrm{G}} = i\tau_\mu^* \{\tau_\mu, \lambda^{(0)}\} \quad \text{and} \quad \dot{\gamma}_{\mathrm{NG}} = (i/2)\mathsf{D}^{(0)}_{\mu\nu}\left\{\tau_\mu^*, \tau_\nu\right\} - \tau_\mu^* \mathsf{D}^{(1)}_{\mu\nu}\tau_\nu. \tag{4.56}$$

Above, we have suppressed the polarization index $I$ for simplicity. The first of the extra phases $\gamma_{\mathrm{G}}$ has the form of a geometric phase upon treating the $x$-$k$ phase space as a parameter space. To see this, note that $\dot{\gamma}_{\mathrm{G}} = i\tau_\mu^*\{\tau_\mu, \lambda^{(0)}\} = i\tau^\dagger \cdot \dot{\tau}$. So, around a closed orbit $\mathcal{C}$ in phase space, the accumulated phase becomes

$$\gamma_{\mathrm{G}} = \oint_{\mathcal{C}} \dot{\gamma}_{\mathrm{G}}\, \mathrm{d}\sigma = i\oint_{\mathcal{C}} \tau^\dagger \cdot \mathrm{d}\tau = i\oint_{\mathcal{C}} \langle \tau | \nabla_\vartheta \tau\rangle \cdot \mathrm{d}\vartheta, \tag{4.57}$$

where $\vartheta = (x,k)$ denotes the "parameters" that are being varied along the orbit. The phase $\gamma_{\mathrm{G}}$, which is the integral of a differential 1-form, depends only on the path $\mathcal{C}$ and not how it is parameterized (like all geometric phases). The second phase $\gamma_{\mathrm{NG}}$ has no such interpretation and is not a geometric phase.

## 4.4 Bound waves in phase space

We expect the rays of bound waves to be bounded in phase space as well, with these rays being topologically equivalent to a circle [85, 116]. Such rays oscillate between two classical turning points where $k = 0$ and $\dot{x} = 0$ as shown in the cartoon in Fig. 4.1(a). From our previous discussion, we see that turning points are examples of caustics, i.e., points on the ray where $\dot{x} = 0$. Of course, on a bound orbit, apart from the classical turning points, there could be other caustics as well [see Fig. 4.1(b)].

The eikonal solution for the wave field is of the form $\psi(x) \sim B(x)\tau(x)e^{i\gamma(x)}e^{iS(x)/\epsilon}$, where $\gamma$ is the adiabatic phase that evolves according to Eq. (4.56). The eikonal phase $S(x)$ can be recovered by integrating $k(x)$ along a ray. As $\psi(x)$ must be single valued, its overall phase can only change by an integral multipe of $2\pi$ as we move around a bound orbit. This leads to a modified Bohr–Sommerfeld-type quantization condition of the form

$$\epsilon^{-1}\oint \mathrm{d}x\, k(x;\omega) = 2\left(n + \frac{\alpha}{4}\right)\pi - \gamma, \tag{4.58}$$

from which bound-state frequencies $\omega$ can be extracted. Above, the quantum number $n \in \mathbb{N}_0$ and $\alpha$ is the Keller–Maslov index [85, 113], which accounts for phase jumps at the turning points. Closed orbits in a two-dimensional phase space that can be smoothly deformed to a small circle always have $\alpha = 2$ [134].

## 4.A Weyl symbols

In physics, by quantization we mean an association of an operator, e.g., $\widehat{A}$, defined on some Hilbert space to an ordinary $c$-numbered function $A$ on the $x$-$k$ position-momentum phase space. In this context, the ordinary function is called the symbol of the operator [17, §2.3.1]. A particularly useful type of an operator symbol is the Weyl symbol, which is defined by the Weyl transform

$$A(x,k) = \int \mathrm{d}s\, e^{-iks/\epsilon}\left\langle x + \tfrac{1}{2}s\right|\widehat{A}\left|x - \tfrac{1}{2}s\right\rangle. \tag{4.59}$$

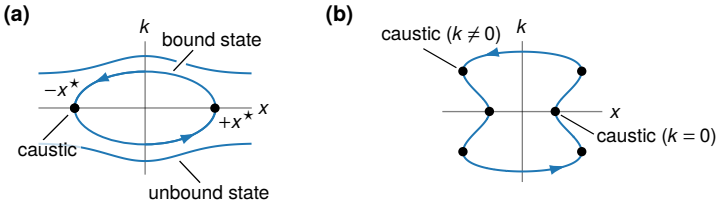


**Figure 4.1.** (a) In phase space, bound states are represented by rays in the form of closed orbits, which is analogous to that of a bound particle oscillating between two classical turning points ($\pm x^\star$ in the cartoon). Other trajectories represent unbound states. (b) A bound state represented by a "peanut"-shaped orbit has six caustics.

where $\langle x|\,\widehat{A}\,|x'\rangle = A(x,x')$ is the matrix element of the operator $\widehat{A}$. Weyl symbols are preferable over other symbols (e.g., the "common" $xk$ symbol [116]) because Weyl symbols of Hermitian operators are real functions of $x$ and $k$. To see this, consider a Hermitian operator $\widehat{A}$ with $\widehat{A}^\dagger = \widehat{A}$, and take the complex conjugate of Eq. (4.59) to find

$$A^*(x,k) = \int \mathrm{d}s\, e^{+iks/\epsilon} \left\langle x+\tfrac{1}{2}s\right|\widehat{A}\left|x-\tfrac{1}{2}s\right\rangle^* = \int \mathrm{d}s\, e^{+iks/\epsilon} \left\langle x-\tfrac{1}{2}s\right|\widehat{A}\left|x+\tfrac{1}{2}s\right\rangle = A(x,k), \quad (4.60)$$

which shows that the symbol $A(x,k)$ is real.

*Moyal formula*

In many situations, it is often necessary to find the symbol of the product of two operators, i.e., given two operators $\widehat{A}$ and $\widehat{C}$ what is the symbol of the operator $\widehat{C} = \widehat{A}\widehat{C}$? As $\widehat{A}$ and $\widehat{B}$ are general noncommuting operators, there is no reason to assume that $C = AB$, and indeed $C \neq AB$, in general. To derive a formula for the symbol product, we first express the matrix elements of $\widehat{A}$ and $\widehat{B}$ in terms of their symbols. Putting the resulting expressions in Eq. (4.59), and after a series of Taylor expansions and straightforward algebra, we find the Moyal star product, which can be used to compute the symbol of the product of two operators in terms of their respective symbols [17]:

$$\begin{aligned} C(x,k) = A(x,k) \star B(x,k) &= A(x,k)\, e^{\frac{i\epsilon}{2}\widehat{\mathcal{L}}} B(x,k) \\ &= A(x,k)\left[1 + \frac{i\epsilon}{2}\left(\overleftarrow{\partial_x}\overrightarrow{\partial_k} - \overleftarrow{\partial_k}\overrightarrow{\partial_x}\right) + \mathcal{O}(\epsilon^2)\right] B(x,k) \\ &= A(x,k)B(x,k) + \frac{i\epsilon}{2}\{A(x,k), B(x,k)\} + \mathcal{O}(\epsilon^2). \end{aligned} \quad (4.61)$$

Above, we have defined the Janus operator $\widehat{\mathcal{L}} = \overleftarrow{\partial_x}\overrightarrow{\partial_k} - \overleftarrow{\partial_k}\overrightarrow{\partial_x}$. Here the derivatives $\overleftarrow{\partial_x}$ and $\overleftarrow{\partial_k}$ act *only* on the terms to their left and the derivatives $\overrightarrow{\partial_x}$ and $\overrightarrow{\partial_k}$ act *only* on the terms to their right. In principle, it is possible to use the Weyl transform, Eq. (4.59), and the Moyal formula, Eq. (4.61), to compute the symbol of any given operator. For the purposes of this dissertation, however, the following simple rules suffice to convert operators to symbols:

$$f(x) \to f(x), \quad g(\hat{k}) \to g(k), \quad \text{and} \quad f(x)g(\hat{k}) \to f(x)g(k) + \frac{i\epsilon}{2} f'(x)g'(k) + \mathcal{O}(\epsilon^2). \quad (4.62)$$

Above, $f$ and $g$ are functions of $x$ and $\hat{k}$, with the primes denoting derivatives.

Chapter 5

# Wave localization in thin elastic structures*

In this final chapter we consider the localization of elastic waves in thin elastic structures with spatially varying curvature profiles, using a curved rod and a singly curved shell as concrete examples. Making use of the semiclassical/WKB approximation developed in the previous chapter, we show that elastic waves form localized, bound states around points where the absolute curvature of the structure has a minimum.

## 5.1 Introduction

Studying the propagation of elastic waves on thin structures is of crucial importance to a variety of problems in science and engineering, with applications ranging from acoustic cloaks to negative refraction [25, 42, 178]. Of particular relevance to many of these applications are localized waves, which are time-harmonic solutions to a wave equation that remain confined to a certain region of space without the presence of a confining potential or force. Indeed, such waves are observed in many physical systems and they are often caused by heterogeneities in the medium or the boundary. For instance, the Helmholtz equation admits bound states in arbitrary dimensions when solved on a tubular domain, provided that the tube is not everywhere straight [52]. Likewise, in waveguides in the form of an elastic plate, described again by coupled Helmholtz equations, waves localize around points of maximal curvature [57]. Bound waves of similar nature have also been predicted in waveguides in the form of rods [56], elastic strips with varying elastic moduli [43] and thickness [139], quantum waveguides [35], etc.

Localized waves can also arise in elastodynamic systems described by higher-order wave equations. In this context, Scott and Woodhouse [150] studied the localized vibrations of a musical saw—an ordinary hand saw bent into the shape of the letter S and playable like a musical instrument [97, 160]. More recently, Shankar *et al.* [151] revisited the musical saw using both experiments and theory. Forgoing an explicit analytical computation of the mode frequencies, they argued that the bound modes that appear at the inflection point of the saw are topologically protected. Localization of elastic waves on variably curved shells, such as the musical saw, is not entirely surprising as it is known that curvature acts as an effective refractive index for such waves [40, 128].

*This chapter is an extended version of M. Mannattil and C. D. Santangelo, arXiv:2306.07213 [cond-mat.soft]. The problem discussed in this work emerged during discussions with my coauthor. I was responsible for all the analytical and numerical calculations and wrote the paper taking into consideration my coauthor's comments.

Motivated by the above studies, in this chapter, we investigate further aspects of curvature-controlled localization of waves in thin elastic structures, choosing a singly curved shell and a curved rod as our examples. If the structure is uncurved, there are three basic types of waves that can propagate. Extensional waves propagate by stretching and compressing the structure, and involve only the tangential displacements ($u$ and $v$ in Fig. 5.1). Flexural waves, by contrast, propagate by bending the structure and involve only the normal displacement ($\zeta$ in Fig. 5.1). In flat plates, shear waves, which do not compress or expand the plate, and involve only the tangential displacements propagate as well [90].

The situation gets complicated when the structure is curved. First, curvature tends to couple the tangential and normal displacements, and therefore, we can only speak of waves that are predominantly flexural or extensional or shear like. Second, there are no universally accepted elastodynamic equations for curved structures, and in case of the rod and the shell, several choices exist [34, 88, 121, 135]. The simplest ones, however, are almost always a set of linear partial differential equations that couple the normal and tangential displacements, the independent variables being time and the coordinates that describe the undeformed configuration of structure ($x$ and $y$ in Fig. 5.1). The physical assumptions usually made (expressed here in terms of the wave number $k$ and the structure's curvature $m$) while writing down such a set of equations are [88, 128, 135]:

(i) The wavelength ($\sim k^{-1}$) is much larger that the thickness of the structure. If we work in length units such that the thickness is of order unity, we must then have $k \ll 1$. For bulk waves with wavelengths much smaller than the thickness (i.e., when $k \gg 1$), the structure should be treated as an infinite elastic medium [90].

(ii) The radius of curvature is much larger than the thickness, so $|m| \ll 1$ (in length units such that thickness is of order unity). This is a geometrical requirement as it is impossible to construct a structure whose local radius of curvature is smaller than its thickness.

(iii) The wavelength is smaller or of the same order of magnitude as the radius of curvature ($= m^{-1}$), so $k > m$ is a safe choice (in all length units).

Because the curvature couples the different displacement components, irrespective of the equations we use, to fully characterize wave propagation on curved structures, we have to consider multicomponent (i.e., vector) waves. Computing the exact spectrum of a multicomponent differential operator is often difficult, unless one resorts to numerical techniques. Indeed, for this reason, in their theoretical analyses of the musical saw, both Scott and Woodhouse [150], and Shankar *et al.* [151] chose to simplify matters by analyzing flexural vibrations alone.

As we discussed in Chapter 4, the semiclassical/WKB approximation is a widely used technique to obtain asymptotic solutions to wave problems, including those describing the elastodynamics of thin structures [120, 126, 137, 154]. We also saw that in wave equations supporting bound-state solutions, the semiclassical approximation allows one to extract the corresponding frequencies through quantization. As we discussed previously, subtleties can arise in multicomponent equations owing to the presence of an extra phase in the quantization rule [83, 99, 100, 174]. Recently, this phase has been shown [165] to be responsible for

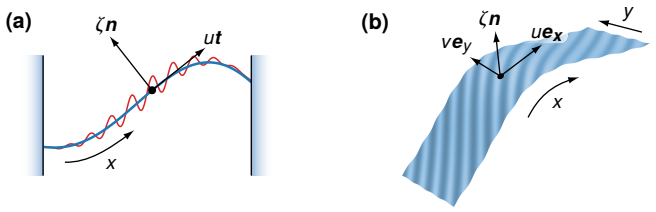


**Figure 5.1.** Waves can propagate on thin elastic structures such as (a) rods and (b) shells. The undeformed structure is parameterized by the coordinates $x$ and $y$. Curvature couples tangential displacements ($u$ and $v$) that stretch/shear the structure with normal displacements ($\zeta$) that bend it.

a spectral flow in the rotating shallow-water equations that describe oceanic waves on the Earth's surface, leading to the topological protection of equatorial waves [27].

In this chapter, we use the semiclassical approximation to study the bound-state spectrum of elastic waves in a curved rod and a singly curved shell with a varying curvature profile. To avoid losing the main results of this chapter in a thicket of details, we summarize them here:

(i) For both the rod and the shell, independent of the boundary conditions, waves exhibit robust localization around points where the absolute curvature has a minimum. Wave localization induced by the presence of an inflection point in an S-shaped musical saw [150, 151] is a special case of this more general observation. These findings also complement the prediction by Mohammed *et al.* [120] regarding the localization of flexural waves in a curved shell around points of maximal curvature.

(ii) In a curved rod, only extensional waves form bound states and flexural waves always form "unbound" states that are spread across the rod.

(iii) In a shell, waves of all three types can form states that are bound along the curved direction [$x$ in Fig. 5.1(b)]. In the frequency spectrum, flexural bound states appear first and have the lowest frequencies. They are then followed by shear and extensional bound states, in that order.

(iv) For both structures, flexural waves start propagating well below the frequency of the first bound state associated with an extensional wave. Hence, in very long rods and shells, these bound states coexist with a near-continuum of flexural waves, forming quasi-bound states in a continuum [69].

(v) Finally, both structures are described by equations for which the extra phase in the modified quantization rule vanishes—something that we expect to be generically true for equations of thin-walled structures (Appendix 5.A). This simplifies our analysis considerably and results in remarkable agreement between the quantization results and numerical experiments.

Our findings show that waves can be robustly trapped in thin elastic structures by a simple alteration of their geometry. This could help, for instance, in crafting better thin-plate acoustic cloaks [42], and aid the control of noise and vibration in thin structures [59, 107]. Curvature-

induced localization of waves could also be used to improve the acoustic black-hole effect in thin-walled structures [91, 131], which at the moment relies primarily on wave localization caused by a power-law tapered thickness profile [89]. Finally, a singly curved shell serves as a simple, yet effective single-mode waveguide that can steer flexural waves of specific frequencies in the uncurved direction.

## 5.2 Waves on a curved rod

As we remarked earlier, several rod theories [19, 167], with varying levels of sophistication, have been written down to describe wave propagation on rods—straight or curved. For our purposes, a simple linear model [88] of a curved rod would do. This model ignores higher-order effects like torsion, cross-sectional rotation, etc., and can be viewed as the lower-dimensional analog of the Donnell–Yu shell model [33, 177] that we shall later use to understand wave propagation on curved shells.

### 5.2.1 Equations of motion and semiclassical approximation

Let the undeformed state of the rod be in the form of a plane curve $\boldsymbol{\sigma} : \mathcal{X} \to \mathbb{R}^2$, which we take to be parameterized by its arclength $x \in \mathcal{X} \subset \mathbb{R}$. As a wave propagates along the rod, it undergoes a deformation $\boldsymbol{\sigma} \to \boldsymbol{\sigma} + \delta\boldsymbol{\sigma}$, where the displacement field $\delta\boldsymbol{\sigma}(x,t) = u(x,t)\boldsymbol{t}(x) + \zeta(x,t)\boldsymbol{n}(x)$. Here $\boldsymbol{t} = \mathrm{d}\boldsymbol{\sigma}/\mathrm{d}x$ is the unit tangent of the undeformed rod and $\boldsymbol{n}$ is its unit normal, obtained by rotating $\boldsymbol{t}$ counter-clockwise by $\pi/2$ [see Fig. 5.1(a)]. Using $\zeta$ and $u$ as the components of the wave field, and assuming the absence of external forces, the rod equations we are after are derived from the following energy functional [88]:

$$\mathscr{U}[\zeta, u] = \int \mathrm{d}t\,\mathrm{d}x\,\tfrac{1}{2}\left\{\rho\left(\dot{u}^2 + \dot{\zeta}^2\right) - E\left[\partial_x u - m(x)\zeta\right]^2 - B\left(\partial_x^2\zeta\right)^2\right\}. \tag{5.1}$$

Above, we have assumed that the rod is uniform with linear mass density $\rho$, with extensional stiffness $E$ and bending stiffness $B$. Also, the signed curvature of the rod is $m(x) = \boldsymbol{n}\cdot\mathrm{d}\boldsymbol{t}/\mathrm{d}x$, which we assume to vary with the arclength $x$. With these identifications, we can delineate the bending and stretching contributions to the energy.[1] Upon varying the energy functional $\mathscr{U}$, we get the dynamic rod equations

$$\rho\partial_t^2\zeta = -B\partial_x^4\zeta - Em(x)\left[m(x)\zeta - \partial_x u\right], \tag{5.2a}$$

$$\rho\partial_t^2 u = -E\left\{\partial_x\left[m(x)\zeta\right] - \partial_x^2 u\right\}. \tag{5.2b}$$

When the curvature $m = 0$, Eq. (5.2a) reduces to the dynamic Euler–Bernoulli equation representing purely transverse flexural waves that propagate by bending the rod. In the same limit, Eq. (5.2b) characterizes extensional waves that propagate longitudinally by stretching the rod. When the curvature is nonzero, which is the case we want to analyze, the components $\zeta$ and $u$ remain coupled.

[1] The stretching contribution to the energy is approximated by taking the usual one-dimensional strain $\partial_x u$ and adding to it the circumferential strain, which for a (convex) differential element of the rod with angular width $\delta\theta$ and radius $r = -m^{-1}$ is roughly $[(r+\zeta)\delta\theta - r\delta\theta]/(r\delta\theta) = -m\zeta$ [33]. The bending energy, on the other hand, is the usual Euler–Bernoulli bending energy, without any additional curvature-dependent corrections. See Section 5.4 for a rod model that includes additional corrections to the bending energy.

In terms of the rod's Young's modulus $Y$, cross-sectional area $A$, and its second moment of area $I$, the extensional and bending stiffnesses are $E = YA$ and $B = YI$, respectively. If the rod's cross-sectional "thickness" is $h$, then $A \sim h^2$ and $I \sim h^4$. This gives a natural length unit $\ell = \sqrt{B/E} \sim h$ and a time unit $\sqrt{B\rho}/E$ that can be used to conveniently nondimensionalize the rod equations. Noting that under a change of length units the curvature transforms as $m(x) \to \ell^{-1} m(x)$, we arrive at the nondimensional form of the rod equations, which in matrix form reads [cf. Eq. (4.1)]

$$\partial_t^2 \begin{pmatrix} \zeta \\ u \end{pmatrix} + \widehat{\mathsf{H}} \begin{pmatrix} \zeta \\ u \end{pmatrix} = 0, \text{ where } \widehat{\mathsf{H}} = \begin{pmatrix} \partial_x^4 + m^2(x) & -m(x)\partial_x \\ m(x)\partial_x + m'(x) & -\partial_x^2 \end{pmatrix}. \tag{5.3}$$

Although this is a rather simple set of coupled equations involving the curvature $m(x)$ as the only parameter, a nonuniform $m(x)$ makes obtaining a general solution difficult.

To employ the semiclassical approximation for solving Eq. (5.3), we assume that the curvature is a slowly varying function of the form $m(\epsilon x)$. Here $0 < \epsilon \ll 1$ is a small dimensionless parameter that controls the slowness of the variation. Because the length unit $\ell$ we chose for nondimensionalizing the arclength $x$ is proportional to the thickness, physically speaking, here we are assuming that the length scale over which the curvature varies significantly is much larger that the thickness of the rod. Next, we do a final change of variables $x \to \epsilon^{-1} x$ in Eq. (5.3) so that $m(\epsilon x) \to m(x)$ and all spatial derivatives get multiplied by $\epsilon$. For the eigenvalue problem with $\widehat{\mathsf{D}} = \widehat{\mathsf{H}} - \omega^2 \mathsf{I}_2$, this gives us

$$\widehat{\mathsf{D}} = \begin{pmatrix} \epsilon^4 \partial_x^4 + m^2(x) - \omega^2 & -m(x)\epsilon\partial_x \\ m(x)\epsilon\partial_x + \epsilon m'(x) & -\epsilon^2\partial_x^2 - \omega^2 \end{pmatrix} = \begin{pmatrix} \hat{k}^4 + m^2(x) - \omega^2 & -im(x)\hat{k} \\ im(x)\hat{k} + \epsilon m'(x) & \hat{k}^2 - \omega^2 \end{pmatrix}. \tag{5.4}$$

In the final step above, we have set $\epsilon\partial_x \to i\hat{k}$, with $\hat{k}$ being the momentum operator. Using the rules in Eq. (4.62) we can easily write down the Weyl symbol $\mathsf{D}$ for the operator in Eq. (5.4) as $\mathsf{D} = \mathsf{D}^{(0)} + \epsilon \mathsf{D}^{(1)}$, where

$$\mathsf{D}^{(0)} = \begin{pmatrix} k^4 + m^2(x) - \omega^2 & -im(x)k \\ im(x)k & k^2 - \omega^2 \end{pmatrix} \quad \text{and} \quad \mathsf{D}^{(1)} = \begin{pmatrix} 0 & \frac{1}{2}m'(x) \\ \frac{1}{2}m'(x) & 0 \end{pmatrix}. \tag{5.5}$$

The two eigenvalues of the dispersion matrix $\mathsf{D}^{(0)}$, representing waves of two different polarizations, are

$$\lambda_\pm = \tfrac{1}{2}\left\{k^2 + k^4 + m^2(x) \pm \sqrt{\left[k^2 - k^4 - m^2(x)\right]^2 + 4k^2 m^2(x)}\right\} - \omega^2. \tag{5.6}$$

For a given $\omega$, we have $\lambda_+(x,k;\omega) \geq \lambda_-(x,k;\omega)$ for all values of $x$ and $k$. Mode conversion between the two polarizations ensues near points where $\lambda_+(x,k;\omega) = \lambda_-(x,k;\omega)$, and the semiclassical approximation breaks down. But from Eq. (5.6), we see that $\lambda_+ = \lambda_-$ only when the discriminant in Eq. (5.6) vanishes, which happens only when $m(x)$ is zero, and $k = \pm 1$ or $k = 0$. As we discussed in Section 5.1, the rod equations are only applicable for short waves whose wavelength is much longer than the thickness, which translates to the requirement $0 \ll |k| \ll 1$. Hence, mode-conversion points with $k = \pm 1$ or $k = 0$, lie well beyond the range of applicability of these equations. For this reason, we ignore mode-conversion issues and assume that $\lambda_+ \neq \lambda_-$ throughout our analysis. Before moving on, it is useful to first analyze the propagation of waves on a rod of constant curvature.

### 5.2.2 Rods of constant curvature

First we analyze the case where the curvature is zero, i.e., when the rod is straight. In the limit of vanishing curvature $m$, we should recover the dispersion relations $\omega(k)$ for plane waves propagating on a straight rod from $\lambda_\pm$. On setting $m = 0$ and noting that $|k| \ll 1$, we see that $\lambda_+ = 0$ gives us the linear dispersion relation $\omega = k$, representing extensional waves propagating on a straight rod. Meanwhile, when $m = 0$, we find that $\lambda_- = 0$ gives us the quadratic dispersion relation $\omega = k^2$ of flexural waves on a straight rod.

For nonzero, but constant curvature, the rod forms part of a ring. If the curvature is sufficiently weak, we expect the eigenvalue $\lambda_+$ to continue to represent predominantly extensional waves and $\lambda_-$ to represent predominantly flexural waves. To verify this, we expand $\lambda_\pm$ in powers of $m^2$ and drop powers of $k$ in comparison to unity to find

$$\lambda_+ = k^2 + m^2 - \omega^2 + \mathcal{O}(m^4) \quad \text{and} \quad \lambda_- = k^4 - k^2 m^2 - \omega^2 + \mathcal{O}(m^4) \qquad (k \gg m). \tag{5.7}$$

Clearly, the above expansions can only be valid when the $\mathcal{O}(m^2)$ correction terms are less than the lowest-order terms, which is true only when $k \gg m$. In this limit we have $\lambda_+ \sim k^2 - \omega^2$ and $\lambda_- \sim k^4 - \omega^2$, confirming our expectation.

We next look at the case where both $k$ and $m$ are small,[2] but with $k \ll m$. To this end, we expand $\lambda_\pm$ in powers of $km^{-1}$ to find

$$\lambda_+ = k^2 + m^2 - \omega^2 + \mathcal{O}(k^4) \quad \text{and} \quad \lambda_- = \frac{k^6}{m^2}\left[1 + \mathcal{O}\left(\frac{k^2}{m^2}\right)\right] - \omega^2 \qquad (k \ll m). \tag{5.8}$$

Dispersion relations obtained from setting $\lambda_\pm = 0$ above show significant deviation from those of a straight rod and we expect the waves of both polarizations to result in bending and stretching of the rod. For the same reason, we expect both these waves to have both longitudinal and transverse characteristics. Despite this, for the sake of simplicity and identification, we shall continue to call waves represented by $\lambda_+$ as extensional waves and those represented by $\lambda_-$ as flexural waves.

The bending of the rod is associated with normal component $\zeta$ and stretching with tangential component $u$. To better understand how the two components contribute to the wave field in the presence of curvature, we define the amplitude ratio

$$\mathscr{R} = \frac{|\zeta|}{|\zeta| + |u|}. \tag{5.9}$$

With the above definition, for purely transverse and longitudinal waves $\mathscr{R} = 1$ and $\mathscr{R} = 0$, respectively. In the eikonal ansatz, at the lowest order, the wave field $\psi = (\zeta, u)$ is proportional to the eigenvectors $\tau_\pm$ of $\mathsf{D}^{(0)}$, so $\zeta \sim \tau_{\pm,1}$ and $u \sim \tau_{\pm,2}$. Making use of Eqs. (5.7) and (5.8), to the lowest order in $k$ and $m$, the eigenvectors $\tau_\pm$ are

$$\tau_+ \sim (m,\ ik) \quad \text{and} \quad \tau_- \sim (ik,\ m), \tag{5.10}$$

so the asymptotic amplitude ratios for the two wave polarizations become $\mathscr{R}_+ \sim |m|/(|m| + |k|)$ and $\mathscr{R}_- \sim |k|/(|m| + |k|)$. For $m = 0$, we know that extensional waves become entirely

[2] We consider this limit to analyze the behavior of the waves close to a classical turning point where $k = 0$. The wave decays beyond a turning point and it never gets a chance to complete a full-wavelength oscillation with $k \ll m$, so the short wavelength assumption is not violated.

longitudinal, and as expected the corresponding amplitude ratio $\mathscr{R}_+ = 0$. In same limit, flexural waves become entirely transverse as indicated by $\mathscr{R}_- = 1$. However, for small values of $k$ with $k \ll m$ we see that $\mathscr{R}_+ \to 1$ and $\mathscr{R}_- \to 0$. In other words, in the limit $k \ll m$, waves of the two polarizations would switch their nature from being predominantly longitudinal to being predominantly transverse, and vice versa.

### 5.2.3 Rods with varying curvature

For illustrative purposes, we consider rods of two curvature profiles $m_1(x)$ and $m_2(x)$, defined by

$$m_1(x) = b\tanh(\epsilon x) \quad \text{and} \quad m_2(x) = b - (b-a)\operatorname{sech}(\epsilon x), \tag{5.11}$$

which we will informally call tanh- and sech-type curvature profiles, respectively. Above, $b$ and $a$ are positive constants such that $a < b$ and the parameter $\epsilon$ controls the slowness of variation of the curvatures. A rod with a tanh-type curvature profile $m_1(x)$ possesses an inflection point at $x = 0$ where the curvature vanishes, and the curvature asymptotes to $\pm b$ as $x \to \pm\infty$ [see Figs. 5.3(a) and 5.3(b)]. On the other hand, a rod with a sech-type curvature profile $m_2(x)$ remains concave for all $x$, with the curvature acquiring its minimum value of $a$ at $x = 0$ and asymptoting to $b$ as $x \to \pm\infty$ [see Figs. 5.3(c) and 5.3(d)]. For the purpose of illustrating our results, we take $b = 0.1$ and $a = 0.01$, so the smallest ratio between the radius of curvature and thickness is $\mathcal{O}(10)$. Although such a ratio is comparatively smaller than typical experimental values, it would give us a clearer picture of the effects of a varying curvature profile. We also set $\epsilon = 0.01$, but as we rescale $x \to \epsilon^{-1}x$, the parameter $\epsilon$ does not always explicitly appear in our discussions. Note also that the numerical values of the local wave number $k$ is unaffected by this rescaling.

For a varying curvature profile, we find the rays for both polarizations by integrating the corresponding Hamilton's equations given by

$$\begin{aligned}
\dot{x} &= \frac{\partial \lambda_\pm}{\partial k} = k\left\{2k^2 + 1 \pm \frac{(2k^2+1)\left[k^2 + k^4 + m^2(x)\right] - 6k^4}{\sqrt{\left[k^2 - k^4 - m^2(x)\right]^2 + 4k^2 m^2(x)}}\right\}; \\
\dot{k} &= -\frac{\partial \lambda_\pm}{\partial x} = -\tfrac{1}{2}\left[m^2(x)\right]'\left\{1 \pm \frac{k^2 + k^4 + m^2(x)}{\sqrt{\left[k^2 - k^4 - m^2(x)\right]^2 + 4k^2 m^2(x)}}\right\}.
\end{aligned} \tag{5.12}$$

### 5.2.4 Extensional waves

We first discuss how a varying curvature profile affects extensional waves in a rod. In Section 5.3 we present an alternative approach, where we analyze the extensional limit of the rod equations and show that extensional waves form bound states. The purpose here, however, is to understand this using a semiclassical perspective, which will serve as a basis for understanding the variably curved shell.

As we discussed previously, extensional waves are to be associated with the ray Hamiltonian $\lambda_+$. Without loss of generality, let $x = 0$ be a point where $m^2(x)$ has a local extremum so that its derivative $[m^2(x)]'\big|_{x=0} = 0$ and the origin $(0,0)$ becomes a fixed point of the Hamilton's

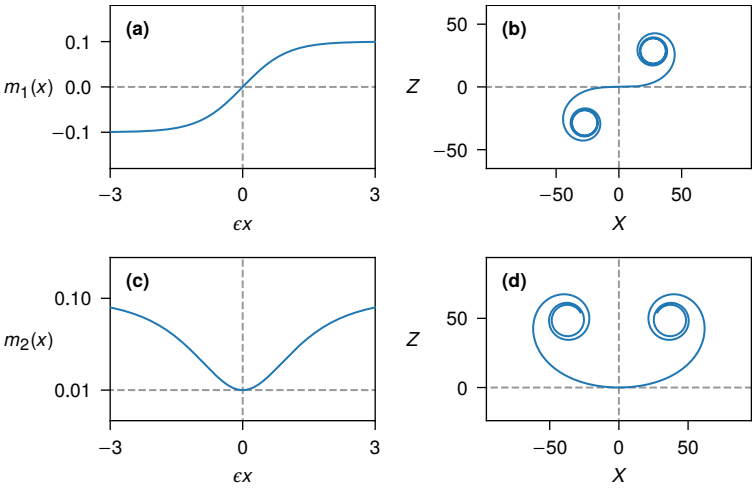


**Figure 5.2.** (a) Tanh-type curvature profile $m_1(x)$ with an inflection point at $x = 0$ and (b) the corresponding shape of the rod in Cartesian space with coordinates $X$ and $Z$. (c) Sech-type curvature profile $m_2(x)$ with no inflection point and (d) the corresponding shape of the rod. For both curvature types, as the arclength $x \to \pm\infty$, the rod becomes part of a circle. In all figures $a = 0.01$, $b = 0.1$, and $\epsilon = 0.01$.

equation in Eq. (5.12) for $\lambda_+$. Since Eq. (5.12) is a Hamiltonian system, the origin becomes a nonlinear center or a saddle depending on whether the Hamiltonian $\lambda_+(x,k)$, viewed as a function of $x$ and $k$, has an extremum or a saddle there [72, 158]. For the present problem, we find two situations[3] depending on the behavior of $m^2(x)$ at $x = 0$.

(i) *$m^2$ has a nonzero maximum at $x = 0$.* Then, for any point $x$ sufficiently close to 0, we have $m^2(x) < m^2(0) = \lambda_+(0,0)$. We also see that $\lambda_+(x,0) - \lambda_+(0,0) < 0$, whereas $\lambda_+(0,k) - \lambda_+(0,0) > 0$. In other words, if we move along the $x$ axis from the origin $(0,0)$, the value of $\lambda_+(x,k)$ decreases, whereas moving along the $k$ axis increases its value, showing that the origin becomes a saddle point.

(ii) *$m^2$ has a minimum at $x = 0$.* At any point $(x,k)$ sufficiently close to $(0,0)$, we have $m^2(x) > m^2(0)$. Then, using basic inequality arguments, $\lambda_+(x,k) - \lambda_+(0,0) > 0$, showing that $\lambda_+$ has a minimum at the origin, which means that it becomes a nonlinear center.

As the extrema of $m^2(x)$ are identical to the extrema of the absolute curvature $|m(x)|$, we can summarize as follows: points where the absolute curvature has a minimum become nonlinear centers of the ray equations associated with $\lambda_+$, and points where the absolute curvature has a maximum become saddles. Therefore, we expect extensional waves to get trapped only around points where the absolute curvature has a minimum.

[3] We can also infer this from the sign of the Hessian determinant of $\lambda_+$ at $(0,0)$ given by $\det \nabla\nabla\lambda_+|_{x=0,k=0} = 2(m^2)''|_{x=0}$, provided that $(m^2)'' \neq 0$ so that the origin is a nondegenerate fixed point with $\det \nabla\nabla\lambda_+ \neq 0$.

The black curves in Figs. 5.3(a) and 5.3(b) show the extensional rays for the two curvature profiles. Each ray is a level curve defined by $\lambda_+(x,k;\omega)=0$ for a specific $\omega$. Since $\lambda_+$ is invariant with respect to reflection about the $x$ and $k$ axes, the ray trajectories also share the same reflection symmetry. For both curvature types, there are three fixed points: the origin $(0,0)$ and $(\pm\infty,0)$. The fixed points at $(\pm\infty,0)$ are saddles since $m^2(x)$ achieves its maximum there. Two heteroclinic orbits, shown as black dashed curves, connect the fixed points at $(\pm\infty,0)$. As per our analysis, the origin is always a center as $m^2(x)$ achieves a minimum there for both curvature types. Close to the origin, the rays appear in form of closed orbits indicating bound states.

Before we present some example results, it is useful to analyze the parity of the bound states. If $\psi(x)=[\zeta(x),u(x)]$ is a bound state, then it is easy to check that for an odd curvature profile $m(x)$ with $m(-x)=m(x)$, the eigenmode $\psi(-x)=[\zeta(-x),u(-x)]$ is also a bound state satisfying the rod equations, Eq. (5.3), with the boundary conditions $\psi(\pm\infty)=0$. Assuming nondegeneracy in the eigenmodes, we must then have $[\zeta(x),u(x)]=\pm[\zeta(-x),u(-x)]$, which means that for an odd $m(x)$, the components $\zeta(x)$ and $u(x)$ must have the same parity, i.e., they are either both odd or both even.[4] For an even $m(x)$, however, we find that bound-state solutions must satisfy $[\zeta(x),u(x)]=\pm[\zeta(-x),-u(-x)]$, showing that the components $\zeta(x)$ and $u(x)$ always have different parity. In Figs. 5.3(c) and 5.3(d) we show two example bound states obtained by solving the rod equations numerically (see Appendix 5.B for details) and depicting both components of the wave field $\psi$. Consistent with our analysis, the example bound state of the tanh-type curvature profile has even $\zeta(x)$ and $u(x)$. Likewise, for the example bound state of the sech-type curvature profile, $\zeta(x)$ is odd and $u(x)$ is even.

Clearly, the displacement fields of the bound states of both curvature profiles show a significant presence of both normal and tangential components. To understand this better, consider again the phase portraits in Fig. 5.3(a) and 5.3(b), which have been color coded using the amplitude ratio $\mathscr{R}$ defined in Eq. (5.9), and computed[5] from the components of the polarization vector $\tau_+$. The closed orbits corresponding to the example bound states have been marked with white arrows in these phase portraits. Starting at the $k$ axis, if we move along a closed orbit towards one of its turning points, we see that $\mathscr{R}$ changes from 0 to 1, indicating that the wave acquires a strong normal component, which is what we see from the example bound states.

The existence of the bound states can also be intuitively understood from the dispersion relations for extensional waves. From Eq. (5.6), we see that at $k=0$, the dispersion curve given by $\lambda_+=0$ has a nonzero gap and a cut-on frequency[6] of $\omega^2_{\text{cut-on}}=m^2$. For waves with

[4] More quantum mechanically, this can be shown by considering the commutation of $\widehat{\mathsf{D}}$ with the operators $\widehat{\mathsf{P}}_\pm=\text{diag}(\hat{\pi},\pm\hat{\pi})$, where $\hat{\pi}$ is the usual parity operator [21]. Clearly, $\widehat{\mathsf{P}}_\pm\Psi(x)=[\zeta(-x),\pm u(-x)]$, so the eigenstates of $\widehat{\mathsf{P}}_+$ always have $\zeta(x)$ and $u(x)$ of the same parity, whereas the eigenstates of $\widehat{\mathsf{P}}_-$ always have $\zeta(x)$ and $u(x)$ of different parity. For odd and even $m(x)$, we can show that $\widehat{\mathsf{D}}$ commutes with $\widehat{\mathsf{P}}_+$ and $\widehat{\mathsf{P}}_-$, respectively. As commuting operators share the same eigenstates (assuming nondegeneracy), this proves the claim made above.

[5] Note that we are using the exact expression for $\tau_+$ and not the approximate expression in Eq. (5.10) to find $\mathscr{R}$. Using the approximate expression would also result in nearly similar results.

[6] Incidentally, this cut-on frequency is the ring frequency of extensional waves, i.e., when the wavelength is equal to the circumference of a ring of radius $m^{-1}$ [167].

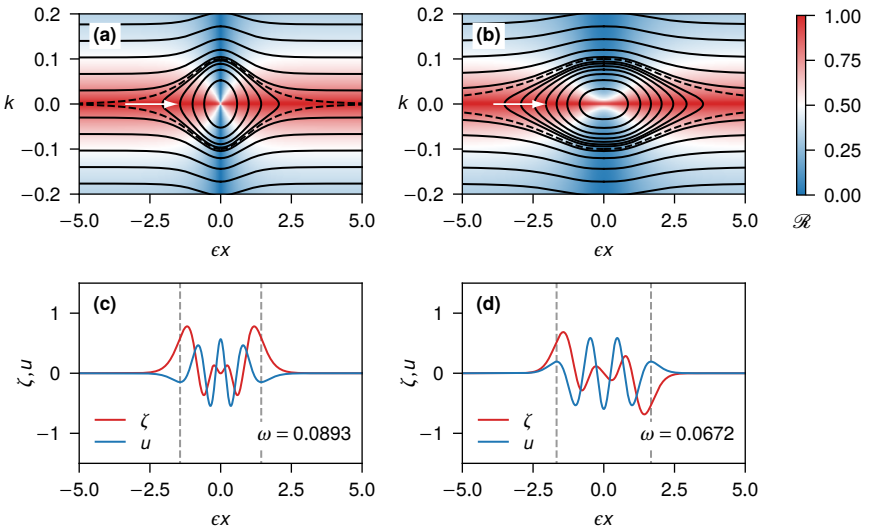


**Figure 5.3.** Ray trajectories for extensional waves on a rod with (a) tanh-type curvature profile $m_1(x)$ and (b) sech-type curvature profile $m_2(x)$, with the phase portraits color coded using the amplitude ratio $\mathscr{R}$ defined in Eq. (5.9). The white arrows in (a) and (b) indicate rays of the same frequency as the two bound states in (c) and (d). The grey vertical lines in (c) and (d) indicate the locations of the classical turning points.

$\omega < \omega_{\text{cut-on}}$, the wave number $k$ is always complex, preventing the wave from being able to propagate and the wave decays. Now, as an extensional wave enters a region of high curvature from a region of low curvature, the local cut-on frequency increases. This means that, at some point, the frequency of the wave would fall below the local cut-on frequency, and the wave gets reflected creating a bound state. As the cut-on frequency depends only on the magnitude of the curvature, and not its sign, bound states do not require the presence of an inflection point—an observation that was also made by Scott and Woodhouse [150] while analyzing the musical saw.

### *Quantization and bound states*

In order to evaluate the action in the quantization condition, Eq. (4.58), in principle, we should express $k$ in terms of $x$ from $\lambda_+(x,k;\omega)=0$. This proves to be difficult as $k$ would then have to be obtained as the root of a sixth-order polynomial. Instead, as described in Appendix 5.B, we compute the action by numerically integrating $k(x)$ between the classical turning points $\pm x^\star$. Setting $k=0$ in $\lambda_+(x,k;\omega)=0$, we see that these points are implicitly given by the equation $m^2(x^\star)=\omega^2$. To quantize the bound orbits, we also need the Keller–Maslov index, which is $\alpha=2$ as the orbits are topologically equivalent to a circle. Now, note that the off-diagonal operators of the matrix $\widehat{\mathsf{D}}$ in the rod equations, Eq. (5.3), are composed

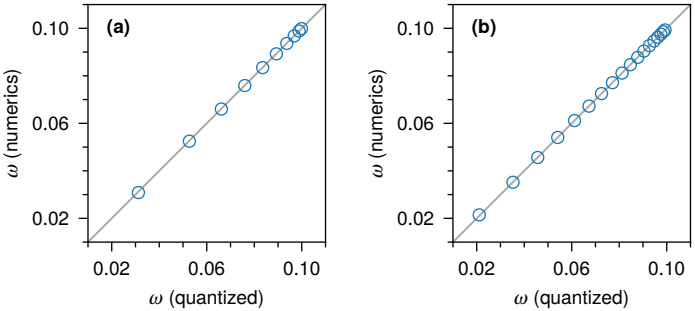


**Figure 5.4.** Bound-state frequencies obtained from numerics compared to that obtained through quantization for a rod with (a) tanh-type curvature profile and (b) sech-type curvature profile. In both plots, the gray guide line in the background represents $\omega$ (numerics) = $\omega$ (quantized).

only of odd derivatives. In Appendix 5.A, we show that the additional phase $\gamma$ in Eq. (4.58), vanishes for any operator of this form, enabling us to find the quantized frequencies without additional difficulty. A comparison between the numerically obtained bound-state frequencies and those obtained through quantization (Fig. 5.4) shows excellent agreement between the two. Furthermore, these frequencies are independent of the boundary conditions chosen to solve the rod equations, showing the robustness of the bound states.

Instead of computing the action by quadrature, we could have extracted $k(x)$ from the approximate expression for $\lambda_+$ [see Eqs. (5.7) and (5.8)], which gives $k(x) \approx \pm\sqrt{\omega^2 - m(x)^2}$. This expression can be used, for instance, to estimate the maximum number of bound states $n_{\max}$ for the two curvature profiles. The largest frequency for which we see an extensional bound state is $\omega = b$. As shown in Figs. 5.3(a) and 5.3(b), the rays corresponding to this frequency form two heteroclinic orbits connecting the classical turning points at $x = \pm\infty$. Computing the area inside these orbits, we find

$$n_{\max} \approx \frac{1}{\pi\epsilon}\int_{-\infty}^{+\infty} \mathrm{d}x\, k(x) = \begin{cases} b/\epsilon & \text{(tanh type)}, \\ \bar{b}/\epsilon + \mathcal{O}(a/\epsilon) & \text{(sech type)}. \end{cases} \tag{5.13}$$

Here, $\bar{b} = \pi^{-1} b \int_{-\infty}^{\infty} \mathrm{d}x\, \operatorname{sech} x\sqrt{2\cosh x - 1} \approx 2b$. Therefore, we expect an infinitely long rod with a sech-type curvature profile to support twice as many bound states as a rod with a tanh-type curvature profile. A similar expression for $n_{\max}$ is also derived in Appendix 5.3 where we directly consider the extensional limit of the rod equations, and find the quantized frequencies as well as the bound modes.

For the values $b = 0.1$ and $\epsilon = 0.01$ we use in our examples, Eq. (5.13) predicts a maximum of 10 bound states for a tanh-type rod and 20 bound states for a sech-type rod. This prediction is consistent with our numerical experiments with a finite-sized rod, where we find a total of 10 and 18 bound states for the two curvature profiles. Despite this, we only expect a small number of extensional bound states in actual experiments with curved rods, where we expect

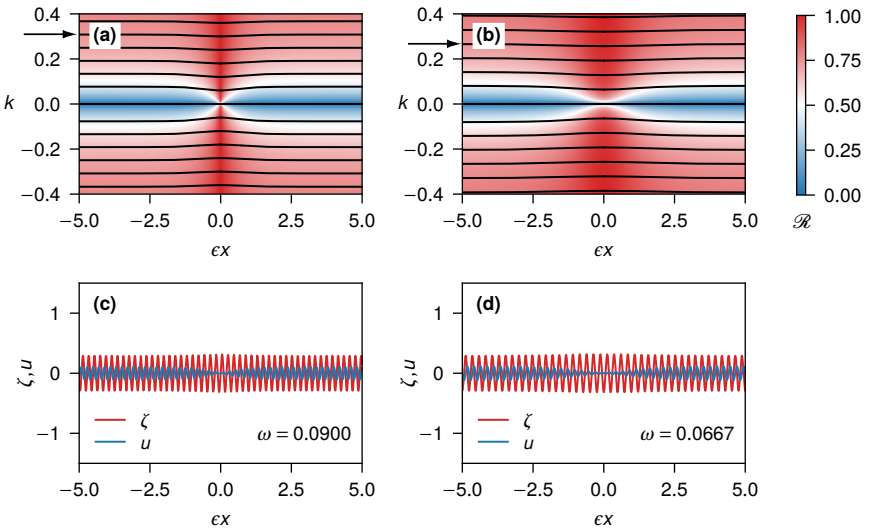


**Figure 5.5.** Ray trajectories for flexural waves on a rod with (a) tanh-type curvature profile $m_1(x)$ and (b) sech-type curvature profile $m_2(x)$, with the phase portraits color coded using the ratio $\mathscr{R}$ defined in Eq. (5.9). The black arrows in (a) and (b) indicate rays of the same frequency as the two eigenmodes in figures (c) and (d).

the finiteness of the rod to become more important. For one thing, as the arc length $x \to \pm\infty$, self-intersection is inevitable for both sech- and tanh-type rods having a nonzero thickness (see Fig. 5.2). (Self-intersection effects, however, can be ameliorated if we allow for small motions of the rod perpendicular to the plane containing it.) Furthermore, as the turning points of the higher-frequency bound states are far apart, they do not always appear to be spatially localized, even though they decay exponentially as $x \to \pm\infty$.

### 5.2.5 Flexural waves

Flexural waves are associated with the ray Hamiltonian $\lambda_-$ and from the corresponding Hamilton's equations in Eq. (5.12), we see that the derivatives $\dot{x}$ and $\dot{k}$ vanish everywhere on the $x$ axis. Since there are no isolated fixed points anywhere on the $x$ axis for these rays, we do not expect flexural waves to form bound states. This is confirmed by the phase portraits in Figs. 5.5(a) and 5.5(b). Like rays of extensional waves, the rays of flexural waves are also invariant under reflection about the $x$ and $k$ axes because $\lambda_-$ possesses the same symmetry. Two unbound flexural eigenmodes obtained numerically are displayed in Figs. 5.5(c) and 5.5(d). Rays corresponding to these modes have been marked by black arrows in Figs. 5.5(a) and 5.5(b). Although the normal component $\zeta$ dominates in these modes, they acquire a significant tangential component with increasing curvature, which we also see from the color coding of the phase portraits.

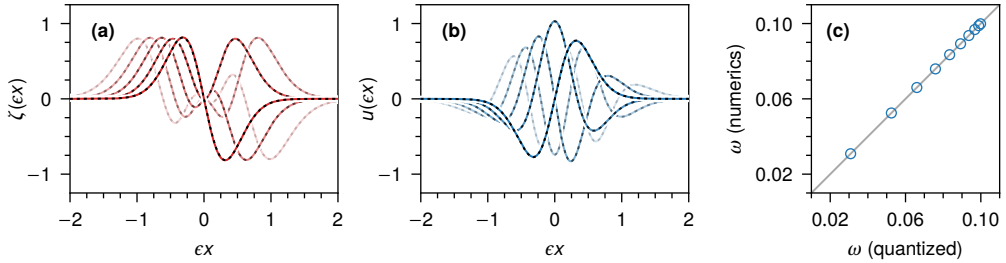


**Figure 5.6.** The first five extensional bound states of a rod with a tanh-type curvature profile showing (a) the normal component $\zeta(x)$ and (b) the tangential component $u(x)$, with lighter curves depicting higher-frequency states. The dashed black curves in (a) and (b) represent the solutions in Eq. (5.18). (c) Comparison between the bound-state frequencies $\omega$ obtained through quantization, and those from numerics [cf. Fig. 5.4(a)].

Although the frequencies of the extensional bound states in Figs. 5.3(c) and 5.3(d) and the flexural eigenmodes in 5.5(c) and 5.5(d) are rather close, the mode profiles differ significantly in appearance. Also, the frequencies of the flexural eigenmodes depend on the specific boundary conditions chosen and other physical parameters, e.g., the total length of the rod. When the rod length increases, we expect flexural waves to form a near continuum in the frequency spectrum, whereas the frequencies of the extensional bound states would continue to be determined by the quantization condition, Eq. (4.58). Setting $\lambda_- = 0$ and putting $k = 0$ in Eq. (5.6), we also see that flexural waves in a rod are gapless and start propagating well below the first bound extensional state. In this sense, in very long rods, extensional bound states appear as quasi-bound states in a continuum of flexural waves.

## 5.3 Extensional limit of the rod equations

We define the extensional limit of the rod equations as the limit in which the average bending energy of a deformed rod is considerably smaller than its stretching energy. The bending energy in the rod model we use is proportional to $(\partial_x^2\zeta)^2$ [see Eq. (5.1)], which gives rise to the fourth-order derivative $\partial_x^4\zeta$ in the rod equations, Eq. (5.3). Neglecting this term and Fourier transforming in time, we find the extensional limit of the rod equations, which take the form

$$m(x)\,[m(x)\zeta(x) - \partial_x u(x)] = \omega^2\zeta(x), \tag{5.14a}$$

$$\partial_x\,[m(x)\zeta(x) - \partial_x u(x)] = \omega^2 u(x). \tag{5.14b}$$

The above equations possess a soft-mode solution with $\omega = 0$ satisfying $m(x)\zeta(x) = \partial_x u(x)$, which corresponds to all linear isometries that do not stretch the rod to the lowest order.

We now look for bound-state solutions of Eqs. (5.14a) and (5.14b) satisfying $\zeta(\pm\infty) = u(\pm\infty) = 0$ with vanishing derivatives at $x = \pm\infty$. Let us additionally assume that the longitu-

dinal component $u(x) = \partial_x \phi(x)$, where $\phi(x)$ is an unknown differentiable function satisfying[7] $\phi(\pm\infty) = 0$. Equation (5.14b) then becomes a total derivative, which on integration yields

$$m(x)\zeta(x) = \partial_x^2 \phi(x) + \omega^2 \phi(x). \tag{5.15}$$

Putting the above equation in Eq. (5.14a) and ignoring the soft-mode solution with $\omega = 0$, we see that $\zeta(x) = m(x)\phi(x)$, from which we deduce that $\phi(x)$ satisfies a Schrödinger-like equation with the potential $m^2(x)$, and given by

$$-\partial_x^2 \phi(x) + m^2(x)\phi(x) = \omega^2 \phi(x). \tag{5.16}$$

As is well known from elementary quantum mechanics [14], Eq. (5.16) always admits a bound-state solution provided that potential $m^2(x)$ has a minimum.

For the tanh-type curvature profile with $m = m_1(x) = b\tanh(\epsilon x)$, Eq. (5.16) becomes the time-independent Schrödinger equation for a particle in a Pöschl–Teller potential [138, 148], whose solutions $\phi(x)$ can be written in terms of the associated Legendre polynomials $P_\nu^\mu(\cdot)$ [129]. On defining

$$\kappa(x) = \tanh(\epsilon x), \quad \nu = \tfrac{1}{2}\left[\sqrt{1 + 4b^2/\epsilon^2} - 1\right], \quad \text{and} \quad \mu = n - \nu \le 0 \quad (n \in \mathbb{N}_0), \tag{5.17}$$

and solving for $\phi(x)$, we find the (unnormalized) components $\zeta(x) = m(x)\phi(x)$, $u(x) = \partial_x \phi(x)$, and the quantized frequencies to be

$$\zeta(x) = b\kappa(x) P_\nu^\mu[\kappa(x)], \quad u(x) = \partial_x P_\nu^\mu[\kappa(x)], \quad \text{and} \quad \omega^2 = b^2 - \epsilon^2\mu^2. \tag{5.18}$$

Figure 5.6 presents a comparison of the bound states described by Eq. (5.18) and the numerical results for extensional bound states. The agreement is remarkable given the fact that the numerical results were obtained by solving the full wave equations, Eq. (5.3), without employing any additional approximations. From Eq. (5.17) we also see that maximum number of bound states is $n_{\max} \approx \nu = b/\epsilon + \mathcal{O}(b^2/\epsilon^2)$, which agrees with the semiclassical prediction in Eq. (5.13).

We are currently unaware of an exact solution of Eq. (5.16) for the potential $m_2^2(x) = [b - (b-a)\,\mathrm{sech}(\epsilon x)]^2$ corresponding to a sech-type curvature profile, even though exact solutions for similar potentials exist [71, 95, 127]. However, we can crudely approximate $m_2^2(x)$ as a deformed Pöschl–Teller potential of the form

$$m_\chi^2(x) = b^2 - (b^2 - a^2)\,\mathrm{sech}^2(\epsilon\chi x). \tag{5.19}$$

We fix the deformation parameter $\chi$ by demanding that the full widths at half minima of both $m_\chi^2(x)$ and $m_2^2(x)$ be equal, which gives

$$\chi = \mathrm{sech}^{-1}\sqrt{\frac{b^2 + a^2}{2(b^2 - a^2)}} \Bigg/ \mathrm{sech}^{-1}\left[\frac{b - \sqrt{(b^2 - a^2)/2}}{b - a}\right]. \tag{5.20}$$

When $b = 0.1$, $a = 0.01$, the deformation parameter $\chi \approx 0.49$ and $m_\chi(x)$ approximates $m(x)$ to around 97% accuracy when $|x| > \frac{1}{2}\Delta$, where $\Delta$ is the full width at half maximum. The

[7] Assuming $\phi(\pm\infty) = 0$ does not result in any loss in generality. Without this assumption, on integrating Eq. (5.14b) we see that $\phi(\pm\infty) = C$ (constant). We then obtain Eq. (5.16) in terms of $\tilde{\phi}(x) = \phi(x) - C$ with $\zeta = m\tilde{\phi}$ so that we can work in terms of $\tilde{\phi}$ alone, which is equivalent to setting $C = 0$.

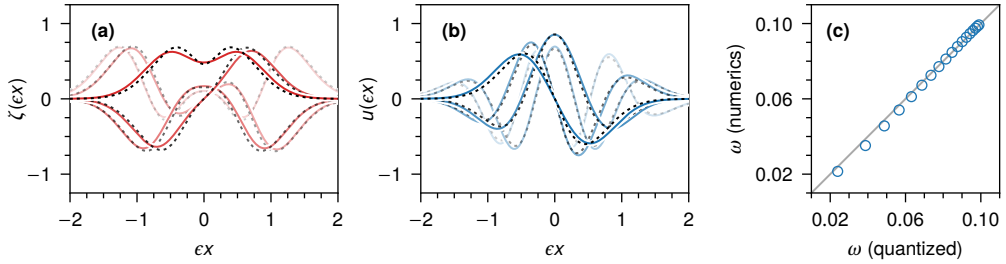


**Figure 5.7.** The first five extensional bound states of a rod with a sech-type curvature profile showing (a) the normal component $\zeta(x)$ and (b) the tangential component $u(x)$. Lighter curves in (a) and (b) depict higher-frequency states and the dashed black curves represent the solutions obtained analogously to those in Eq. (5.18) after setting $b^2 \to b^2 - a^2$ and $\epsilon \to \chi\epsilon$. (c) Comparison between the bound-state frequencies $\omega$ obtained through quantization, and those from numerics [cf. Fig. 5.4(b)].

approximation becomes significantly worse for $|x| < \frac{1}{2}\Delta$, with a maximum relative error of 68%. Therefore, we only expect the higher-frequency bound states of $m_\chi(x)$ to represent states of $m(x)$ reasonably well, which is exactly what we see from the results in Fig. 5.7.

At this juncture, let us remark that Eq. (5.16) can also be derived by diagonalizing the full wave equation in Eq. (5.3) using the asymptotic method outlined by Littlejohn and coworkers [99, 100, 168], without appealing to any physical arguments. In this method, aspects of which has been reviewed in Section 4.2, the operators of the diagonalized equations in symbol form are the two eigenvalues $\lambda_\pm$ of the dispersion matrix $\mathsf{D}^{(0)}$. To find the extensional bound states, we focus on $\lambda_+$, the ray Hamiltonian for extensional waves. Instead of using $\lambda_+$ from Eq. (5.6), which involves a radical expression, we use the approximate version $\lambda_+ \approx k^2 - m^2 - \omega^2$ [see Eqs. (5.7) and (5.8)]. Next, we promote $k \to \hat{k} = -i\partial_x$, which "quantizes" the ray Hamiltonian $\lambda_+$ and we obtain Eq. (5.16) for an unknown wave function $\phi(x)$. The full wave field can be recovered from $\phi(x)$ using $(\zeta, u) \sim \hat{\tau}_+\phi(x)$, where $\hat{\tau}_+$ is a two-component operator obtained by promoting $k \to \hat{k}$ in the symbol form of the polarization vector $\tau_+$. From Eq. (5.10), we see that $\tau_+ \approx [m(x), ik]$, so $\hat{\tau}_+ \approx [m(x), \partial_x]$. We then find $\zeta(x) \sim m(x)\phi(x)$ and $u(x) \sim \partial_x\phi(x)$, consistent with our analysis.

## 5.4 Higher-order rod equations

A slightly different set of rod equations, with higher-order curvature-dependent terms is sometimes considered instead of the equations that we have been using. Given the ubiquity of the higher-order equations,[8] it is useful to briefly discuss them here and contrast the results to those obtained so far. The higher-order rod model is derived from the following energy functional:

$$\mathscr{U}[\zeta, u] = \int \mathrm{d}t\,\mathrm{d}x\,\tfrac{1}{2}\left\{\rho\left(\dot{u}^2 + \dot{\zeta}^2\right) - E\left[\partial_x u - m(x)\zeta\right]^2 - B\left[\partial_x^2\zeta + m(x)\partial_x u\right]^2\right\}, \tag{5.21}$$

[8] See, e.g., Eq. (30) of Ref. [121], Eq. (3.5.24) of Ref. [54], and Eq. (3.34) of Ref. [34].

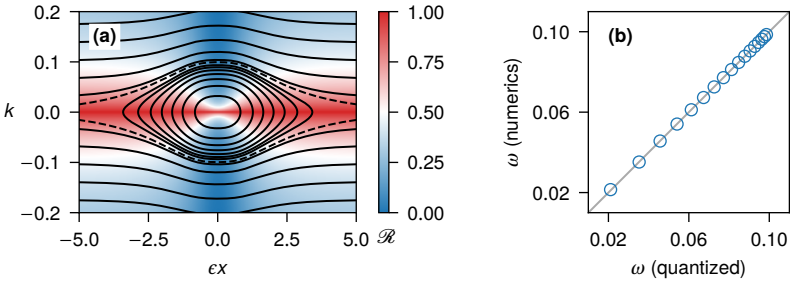


**Figure 5.8.** (a) Ray trajectories for extensional waves described by the higher-order rod equations, Eq. (5.22), for the sech-type curvature profile $m_2(x)$ [cf. Fig. 5.3(b)]. (b) Bound-state frequencies obtained from numerics compared to those obtained by quantization.

The only difference between Eq. (5.1) and Eq. (5.21) is the curvature-dependent term $m(x)\partial_x u$ in the bending energy density, with all other physical parameters having the same meaning. Using the same nondimensionalization as before, we find that Eq. (5.21) results in the following equations of motion:

$$\partial_t^2 \begin{pmatrix} \zeta \\ u \end{pmatrix} + \widehat{\mathsf{H}} \begin{pmatrix} \zeta \\ u \end{pmatrix} = 0, \tag{5.22}$$

where the matrix operator

$$\widehat{\mathsf{H}} = \begin{pmatrix} \partial_x^4 + m^2 & -m\partial_x(1-\partial_x^2) + 2m'\partial_x^2 + m''\partial_x \\ m\partial_x(1-\partial_x^2) + m'(1-\partial_x^2) & -(1+m^2)\partial_x^2 - 2mm'\partial_x \end{pmatrix}. \tag{5.23}$$

To see why Eq. (5.3) is a good approximation to Eq. (5.22), we write down the dispersion matrix by following the usual procedure of expressing the derivatives in terms of the momentum operator and converting operators to symbols, after which we find[9]

$$\mathsf{D}^{(0)} = \begin{pmatrix} k^4 + m^2 - \omega^2 & -imk(1+k^2) \\ imk(1+k^2) & (1+m^2)k^2 - \omega^2 \end{pmatrix} \quad \text{and} \quad \mathsf{D}^{(1)} = \tfrac{1}{2}\begin{pmatrix} 0 & m'(1-k^2) \\ m'(1-k^2) & 0 \end{pmatrix}. \tag{5.24}$$

Clearly, for small wave numbers ($k \ll 1$) and weak curvatures ($m \ll 1$), we can drop $k^2$ and $m^2$ in comparison to unity in the entries of $\mathsf{D}^{(0)}$ and $\mathsf{D}^{(1)}$. Doing so would lead us back to the dispersion matrices in Eq. (5.5).

The ray Hamiltonian for extensional and flexural waves ($\lambda_+$ and $\lambda_-$, respectively) are the two eigenvalues of $\mathsf{D}^{(0)}$, given by

$$\lambda_\pm(x,k) = \tfrac{1}{2}\left\{\left[1+k^2\right]\left[k^2+m^2(x)\right] + \sqrt{\left[k^2+1\right]^2\left[k^2+m^2(x)\right]^2 - 4\left[k^3 - km^2(x)\right]^2}\right\} - \omega^2. \tag{5.25}$$

A simple analysis, either by plotting the rays or writing down the corresponding Hamilton's equations, reveals that even when higher-order, curvature-dependent corrections are included,

[9] Note that unlike the dispersion matrices in Eq. (5.5), here we also have an $\mathcal{O}(\epsilon^2)$ correction $\mathsf{D}^{(2)}$, which we ignore in our analysis.

flexural waves fail to form bound states, and so we focus on extensional waves instead. Figure 5.8(a) shows the ray trajectories for extensional waves described by Eq. (5.22). Clearly, these rays are nearly identical to those seen previously in Fig. 5.3(b), showing again that the simpler rod equations are an excellent approximation to Eq. (5.22) in the weak curvature limit. For the same reason, the frequencies of the extensional bound states obtained through quantization and numerics also tend to agree rather well[10] with our previous results; compare, for example, Fig. 5.8(b) with Fig. 5.4(b).

Let us conclude the discussion on rods by emphasizing that all rod models are approximate to a certain degree. And although we have limited our analysis to just two models, we expect to see our basic result, i.e., extensional waves forming bound states around points where the absolute curvature has a minimum, in other models as well. We next turn to wave propagation on curved shells.

## 5.5 Waves on a curved shell

One of the simplest shell theories that involve a curvature-mediated coupling between the tangential and normal components of the displacement field is the widely used Donnell–Yu shell model [33, 177]. It is simple in the sense that it describes the undulations of a three-dimensional shell solely in terms of the deformation of the shell's two-dimensional midsurface, ignoring higher-order effects and retaining only the lowest-order derivatives. For an arbitrarily parameterized midsurface, it is easiest to extract the equations of motion from the covariant form of the Donnell–Yu shell equations derived by Pierce [136].

### 5.5.1 Equations of motion

We consider the middle surface of the shell to be a generalized cylinder [141] obtained by translating a plane curve $\boldsymbol{\sigma} : \mathcal{X} \to \mathbb{R}^3$, parameterized by $x \in \mathcal{X} \subset \mathbb{R}$, perpendicular to the plane containing it. Thus, the shell is defined by $\boldsymbol{\Sigma} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^3$, $\boldsymbol{\Sigma}(x, y) = \boldsymbol{\sigma}(x) + y\boldsymbol{e}_y$, where $\boldsymbol{e}_y = \partial_y \boldsymbol{\Sigma}$ is a constant unit vector perpendicular to the plane containing $\boldsymbol{\sigma}(x)$. Also, $y \in \mathcal{Y}$ is the coordinate along $\boldsymbol{e}_y$. For simplicity, and to make comparisons with the rod equations easier, we assume that $x$ is the arclength. So, $\boldsymbol{e}_x = \partial_x \boldsymbol{\Sigma} = \boldsymbol{t}$ is the unit tangent along the curve $\boldsymbol{\sigma}$. Furthermore, we orient $\boldsymbol{e}_y$ such that surface normal $\boldsymbol{e}_x \times \boldsymbol{e}_y$ coincides with the normal $\boldsymbol{n}$ to $\boldsymbol{\sigma}$. Then, the only nonvanishing principal curvature of the shell is equal to the curve's signed curvature $m(x)$. Propagating waves displace the shell from $\boldsymbol{\Sigma} \to \boldsymbol{\Sigma} + \delta\boldsymbol{\Sigma}$. We write the displacement field $\delta\boldsymbol{\Sigma}$ as $\delta\boldsymbol{\Sigma} = u\boldsymbol{e}_x + v\boldsymbol{e}_y + \zeta\boldsymbol{n}$. Assuming no external forces, the dynamic Donnell–Yu equations are

$$\varrho \frac{\partial^2 \zeta}{\partial t^2} = -\widetilde{B}\Delta^2\zeta - \widetilde{E}m^2(x)\zeta + \widetilde{E}m(x)\left(\frac{\partial u}{\partial x} + \eta\frac{\partial v}{\partial y}\right), \tag{5.26a}$$

$$\frac{\varrho}{\widetilde{E}}\frac{\partial^2 u}{\partial t^2} = -\frac{\partial [m(x)\zeta]}{\partial x} + \frac{\partial^2 u}{\partial x^2} + \frac{(1-\eta)}{2}\frac{\partial^2 u}{\partial y^2} + \frac{(1+\eta)}{2}\frac{\partial^2 v}{\partial x \partial y}, \tag{5.26b}$$

[10] Other aspects of quantization, e.g., the lack of the extra phase $\gamma$, location of the turning points, $m^2(x^\star) = \omega^2$, etc., also carry over.

$$\frac{\varrho}{\widetilde{E}}\frac{\partial^2 v}{\partial t^2} = -\eta m(x)\frac{\partial \zeta}{\partial y} + \frac{(1+\eta)}{2}\frac{\partial^2 u}{\partial x \partial y} + \frac{(1-\eta)}{2}\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}. \tag{5.26c}$$

Here $\Delta$ is the two-dimensional Laplacian, $\eta$ is the Poisson's ratio, and $\varrho$ is the density per unit area. Also, the extensional stiffness $\widetilde{E} = Yh/(1-\eta^2)$ and bending stiffness $\widetilde{B} = Yh^3/[12(1-\eta^2)]$, with $Y$ being the Young's modulus and $h$ being the thickness of the shell. On setting the length units to $\sqrt{\widetilde{B}/\widetilde{E}}$ and time units to $\sqrt{\widetilde{B}\varrho}/\widetilde{E}$, we arrive at the nondimensional form of these equations, which in matrix form reads

$$\partial_t^2 \begin{pmatrix} \zeta \\ u \\ v \end{pmatrix} + \widehat{\mathsf{H}} \begin{pmatrix} \zeta \\ u \\ v \end{pmatrix} = 0, \tag{5.27}$$

where the operator $\widehat{\mathsf{H}}$ is defined by

$$\widehat{\mathsf{H}} = \begin{pmatrix} \Delta^2 + m^2(x) & -m(x)\partial_x & -\eta m(x)\partial_y \\ m(x)\partial_x + m'(x) & -\partial_x^2 - \frac{1}{2}(1-\eta)\partial_y^2 & -\frac{1}{2}(1+\eta)\partial_x\partial_y \\ \eta m(x)\partial_y & -\frac{1}{2}(1+\eta)\partial_x\partial_y & -\frac{1}{2}(1-\eta)\partial_x^2 - \partial_y^2 \end{pmatrix}. \tag{5.28}$$

A set of equations analogous to the rod equations [see Eq. (5.3)] can be obtained by suppressing the $y$ derivatives in the submatrix obtained by deleting the third row and column of the operator $\widehat{\mathsf{H}}$ given above.

Translation invariance of $\widehat{\mathsf{H}}$ along $y$ lets us look for time-harmonic solutions with the common factor $e^{i(ly-\omega t)}$, where $l$ is the transverse wave number in the $y$ direction and $\omega$ is the frequency of oscillation. This makes the wave field depend only on the coordinate $x$, and makes the transverse wave number $l$ an additional parameter of the operator $\widehat{\mathsf{H}}$. But the operator $\widehat{\mathsf{H}}$ now has complex coefficients, and the components of its eigenmodes are complex functions. However, while discussing the numerical results, we use a phase convention such that $\zeta, u$ are real, and $v$ is imaginary.

Similar to what we did for the rod equations, we perform a change of variables $x \to \epsilon^{-1}x$ and recast the spatial derivatives in terms of the momentum operator $\hat{k}$. Finally, we find the dispersion matrix $\mathsf{D}^{(0)}$ as

$$\mathsf{D}^{(0)} = \begin{pmatrix} (k^2+l^2)^2 + m^2(x) - \omega^2 & -ikm(x) & -i\eta l m(x) \\ ikm(x) & k^2 + \frac{1}{2}(1-\eta)l^2 - \omega^2 & \frac{1}{2}(1+\eta)kl \\ i\eta l m(x) & \frac{1}{2}(1+\eta)kl & \frac{1}{2}(1-\eta)k^2 + l^2 - \omega^2 \end{pmatrix}, \tag{5.29}$$

and find its first-order correction $\mathsf{D}^{(1)}$ to be

$$\mathsf{D}^{(1)} = \frac{1}{2}\begin{pmatrix} 0 & m'(x) & 0 \\ m'(x) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.30}$$

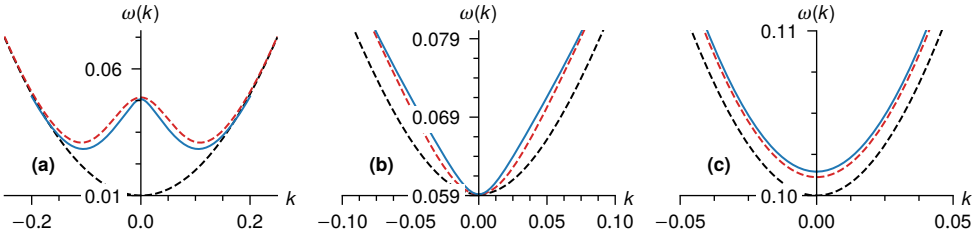


**Figure 5.9.** Dispersion curves for plain waves propagating on a cylinder of constant curvature $m = 0.05$ for (a) flexural, (b) shear, and (c) extensional waves. In each plot, the solid blue curves represent the actual dispersion curves obtained by finding $\omega$ using Eq. (5.35). The dashed red curves represent the approximate dispersion relation in Eq. (5.39). Also, the dashed black curves indicate the dispersion curves for an uncurved flat plate, i.e., when $m = 0$ [Eq. (5.32)]. The transverse wave number $l = 0.1$ and Poisson's ratio $\eta = 0.3$ for all curves.

For later analysis, it is also useful to note down the determinant[11] of $\mathsf{D}^{(0)}$, which is

$$\begin{aligned}\det \mathsf{D}^{(0)} = m^2(x)\left\{\left[\omega^2 - \tfrac{1}{2}(1-\eta)l^2\right]\left[\omega^2 - (1-\eta^2)l^2\right] - \tfrac{1}{2}(1-\eta)k^2\omega^2\right\} \\ - \left[\omega^2 - (k^2+l^2)^2\right]\left[\omega^2 - \tfrac{1}{2}(1-\eta)(k^2+l^2)\right]\left[\omega^2 - (k^2+l^2)\right].\end{aligned} \tag{5.31}$$

Before we continue, it is insightful to examine the dispersion relations for plain waves propagating on singly curved shells of constant curvature.

### 5.5.2 Shells of constant curvature

First we analyze the zero curvature limit, i.e., when the shell becomes a flat plate. In this limit, the wave equation, Eq. (5.28), decouples into two equations, the first of which involves only the normal component $\zeta$, and represents flexural waves. The second equation, representing extensional and shear waves, only involves the tangential components $u$ and $v$. Shear waves propagate transversely to the in-plane wave vector $k\boldsymbol{e}_x + l\boldsymbol{e}_y$, whereas extensional waves are longitudinal to it. Setting $m = 0$ in Eq. (5.31), we find the following flat-plate dispersion relations

$$\omega_0^2 = \left(k^2+l^2\right)^2, \quad \omega_0^2 = \tfrac{1}{2}\left(1-\eta\right)\left(k^2+l^2\right), \quad \text{and} \quad \omega_0^2 = \left(k^2+l^2\right), \tag{5.32}$$

which we recognize as the dispersion relations of flexural, shear, and extensional waves, respectively [45]. The above dispersion relations for a fixed transverse wave number $l$ are indicated by the dashed black curves in Fig. 5.9. Because $l$ is nonzero, there is now a gap in the dispersion curves and a corresponding nonzero cut-on frequency for each of these waves.

[11] The expression for $\det \mathsf{D}^{(0)}$ in Eq. (5.31) is sometimes referred to as the "full" dispersion relation [155]. In reality, $\det \mathsf{D}^{(0)}$ is the product of three dispersion relations—a fact that we will later exploit for finding the ray trajectories.

*Exact dispersion relations*

For nonzero but constant curvature, plain waves continue to propagate on the shell, which now becomes part of a thin cylinder. For simplicity, we shall continue to call these waves as being flexural, extensional, or shear in nature. However, as with the curved rod, we expect some amount of mixing of the tangential and normal displacements due to nonzero curvature. Also, to find the dispersion relations from Eq. (5.31) we now need to use the general cubic formula [129], which results in unwieldy analytical expressions. More specifically, these relations can be obtained by solving the following cubic in $\omega^2$:

$$\omega^6 - a\omega^4 - b\omega^2 - c = 0 \tag{5.33}$$

where the coefficients

$$\begin{aligned} a &= \tfrac{1}{2}q^2\left[(3-\eta)+2q^2\right], \\ b &= -\tfrac{1}{2}q^4\left[(3-\eta)q^2+(1-\eta)\right] - \tfrac{1}{2}m^2(1-\eta)\left[q^2+2(1+\eta)l^2\right], \\ c &= \tfrac{1}{2}(1-\eta)\left[q^8+m^2l^4\left(1-\eta^2\right)\right], \end{aligned} \tag{5.34}$$

and $q^2 = k^2 + l^2$ is the magnitude of the wave vector. The above cubic has the solution

$$\omega^2 = \begin{cases} \tfrac{1}{3}a + \dfrac{2^{1/3}\left(a^2+3b\right)}{3d} + \dfrac{d}{3\times 2^{1/3}} & \text{(extensional)}, \\ \tfrac{1}{3}a - \dfrac{\left(1-i\sqrt{3}\right)\left(a^2+3b\right)}{3\times 2^{2/3}d} - \dfrac{\left(1+i\sqrt{3}\right)d}{6\times 2^{1/3}} & \text{(shear)}, \\ \tfrac{1}{3}a - \dfrac{\left(1+i\sqrt{3}\right)\left(a^2+3b\right)}{3\times 2^{2/3}d} - \dfrac{\left(1-i\sqrt{3}\right)d}{6\times 2^{1/3}} & \text{(flexural)}. \end{cases} \tag{5.35}$$

where

$$d = \left(2a^3+9ab+27c+3\sqrt{3}\sqrt{4a^3c-a^2b^2+18abc-4b^3+27c^2}\right)^{1/3}. \tag{5.36}$$

It is not directly obvious how the roots in Eq. (5.35) can be associated with a specific wave polarization. To do so, we have to take the limit $m \to 0$ numerically and find which of the roots reduce to the flat-plate results in Eq. (5.32).

As an example, using Eq. (5.35), we find the dispersion curves at a curvature value of $m = 0.05$, which are indicated by the solid lines in Fig. 5.9. We make two observations on comparing these curves with the dispersion curves for a flat plate: (i) the gap in the dispersion curves for flexural and extensional waves [Figs. 5.9(a) and 5.9(c)] have increased, and the flexural dispersion curve now has a double-well appearance; (ii) although the dispersion curve for shear waves has changed in appearance [Fig. 5.9(b)], the gap remains the same and the cut-on frequency remains unchanged. The cut-on frequencies at nonzero $m$ can be computed from Eq. (5.31) after setting $k = 0$, and we find three roots

$$\omega^2_{\text{cut-on}} = \tfrac{1}{2}\left(l^2+l^4+m^2 \pm \sqrt{\left(l^2-l^4-m^2\right)^2+4\eta^2l^2m^2}\right) \quad \text{and} \quad \tfrac{1}{2}(1-\eta)l^2. \tag{5.37}$$

Our intuition and a series expansion in $m$ suggests that the lowest of the first two roots must be associated with flexural waves and the highest root must be associated with extensional

waves. The third root, which is independent of the curvature $m$, must then correspond to the cut-on frequency for shear waves. As we shall see, this association is only correct at very low curvatures.

*Approximate dispersion relations*

Although the exact dispersion relations for nonzero curvature are unwieldy, we can find an approximate expression for the dispersion relations in the very weak curvature limit. To this end, we write the roots $\omega^2$ as a regular perturbation series [6] in even powers of $m$, i.e.,

$$\omega^2(k,l) = \omega_0^2(k,l) + \sum_{n=1}^{\infty} m^{2n} Q_n(k,l), \tag{5.38}$$

where $\omega_0$ is one of the three roots in Eq. (5.32) and $Q_n(k,l)$ are coefficients to the correction terms that we have to determine.[12] Since the curvature is assumed to be very weak, the dispersion relations obtained this way can be associated with a wave type based on the choice we make for $\omega_0$. Putting Eq. (5.38) in Eq. (5.31), and dropping powers of $k$ and $l$ in comparison to unity, we find

$$\omega^2 = \begin{cases} \left(k^2+l^2\right)^2 + \left(1-\eta^2\right) m^2 \dfrac{l^4}{\left(k^2+l^2\right)^2} + \mathcal{O}(m^4) & \text{(flexural)}, \\ \frac{1}{2}\left(1-\eta\right)\left(k^2+l^2\right) + 2\left(1-\eta\right) m^2 \dfrac{k^2 l^2}{\left(k^2+l^2\right)^2} + \mathcal{O}(m^4) & \text{(shear)}, \\ \left(k^2+l^2\right) + m^2 \dfrac{\left(k^2+\eta l^2\right)^2}{\left(k^2+l^2\right)^2} + \mathcal{O}(m^4) & \text{(extensional)}. \end{cases} \tag{5.39}$$

The above approximate dispersion relations are identical to those found in the literature and derived using alternative approaches [49, 128, 135, 146]. Also, in their analytical characterization of bound waves in a musical saw, Shankar *et al.* [151] works exclusively with the above approximate dispersion relation for flexural waves, which has also been observed experimentally [170]. In Fig. 5.9, the approximate dispersion relations for $m = 0.05$ are indicated by the red dashed curves, from which we can see that Eq. (5.39) captures the true dispersion relations to a reasonably good accuracy.

Equation (5.39) must break down beyond a certain value of the curvature. Indeed, we only expect it to capture the true dispersion when the $\mathcal{O}(m^2)$ correction term in Eq. (5.38) is smaller than $\omega_0^2$, and more conservatively, only when $m \ll k^2 + l^2$. We would expect the dispersion relation to deviate significantly from Eq. (5.39) as $m$ increases. In fact, for small $k$ and $l$, Eq. (5.39) completely breaks down at a curvature at which the cut-on frequency for flexural and extensional waves become equal. For nonzero $l$, from Eq. (5.37), we see that this happens at a curvature value

$$m_{\ddagger}^2 = \frac{(1+\eta)\, l^2 \left[\frac{1}{2}(1-\eta) - l^2\right]}{\left(1+2\eta\right)\left(1-\eta\right)} = \frac{1}{2}\left(\frac{1+\eta}{1+2\eta}\right) l^2 + \mathcal{O}(l^4). \tag{5.40}$$

[12] Note that it is a futile task to expand the roots in Eq. (5.35) in powers of $m$ to find the approximate dispersion relationships.

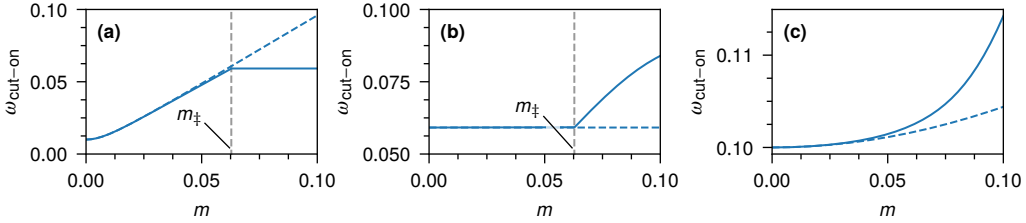


**Figure 5.10.** Cut-on frequencies for a singly curved shell as a function of the curvature $m$ for (a) flexural, (b) shear, and (c) extensional waves. The blue dashed curves represent the cut-on frequency predicted by Eq. (5.39), which holds only when $m$ is small. The solid curves represent the actual cut-on frequency obtained by finding $\omega$ from Eq. (5.31) using the general cubic formula, and taking the limit $k \to 0$ numerically. The transverse wave number $l = 0.1$ and Poisson's ratio $\eta = 0.3$ in all plots.

We graphically demonstrate this in Fig. 5.10, from which we can see that the expressions for the cut-on frequencies for shear and flexural waves get interchanged at large $m$. Hence, for $m > m_\ddagger$, the lowest of the first two cut-on frequencies in Eq. (5.37), should be associated with shear waves, whose dispersion curves would now have a curvature-dependent gap. The cut-on frequency for flexural waves, however, would now be equal to $\frac{1}{2}(1-\eta)l^2$. We remark that this switching of the cut-on frequencies does not happen if $l^2 > \frac{1}{2}(1-\eta)$ as Eq. (5.40) fails to have a real root. Also, as the cut-on frequencies of flexural and shear waves are equal for $m = m_\ddagger$, in principle, mode conversion can occur close to the entire $k$ axis on the phase plane. Despite such a possibility, we did not observe any discernible effects of mode conversion in our numerical experiments, and we shall it ignore in our analysis.

### 5.5.3 Shells with varying curvature

We shall now consider singly curved shells with varying curvature profiles. Given the complexity of the general dispersion relations and the myriad of subtleties, we shall, however, perform a less exhaustive analysis compared to what we did for the rod. We shall only look at a limited number of examples, and to simplify matters, we set the transverse wave number $l = 0.1$ and Poisson's ratio $\eta = 0.3$ (corresponding to that of steel) throughout. As for the rod, we will consider both tanh-type and sech-type curvature profiles, but we assume that the largest absolute curvature $b > m_\ddagger$. This would let us examine the problem beyond the range of validity of the approximate dispersion relation in Eq. (5.39). For the sech-type curvature profile, we additionally assume that the smallest curvature $a < m_\ddagger$. With $l = 0.1$ and $\eta = 0.3$, we have $m_\ddagger \approx 0.06$, and these assumptions are satisfied by the choices $a = 0.01$ and $b = 0.1$ that we made for the rod, and we use them for the shell as well.

From our earlier analysis, we saw that the spectral gap in the dispersion relation for all three wave polarizations grows with increasing curvature. Now, consider a wave traveling from a region of low curvature to one of high curvature. As the wave moves, at some point, the frequency of the wave would fall below the local cut-on frequency of waves, where it

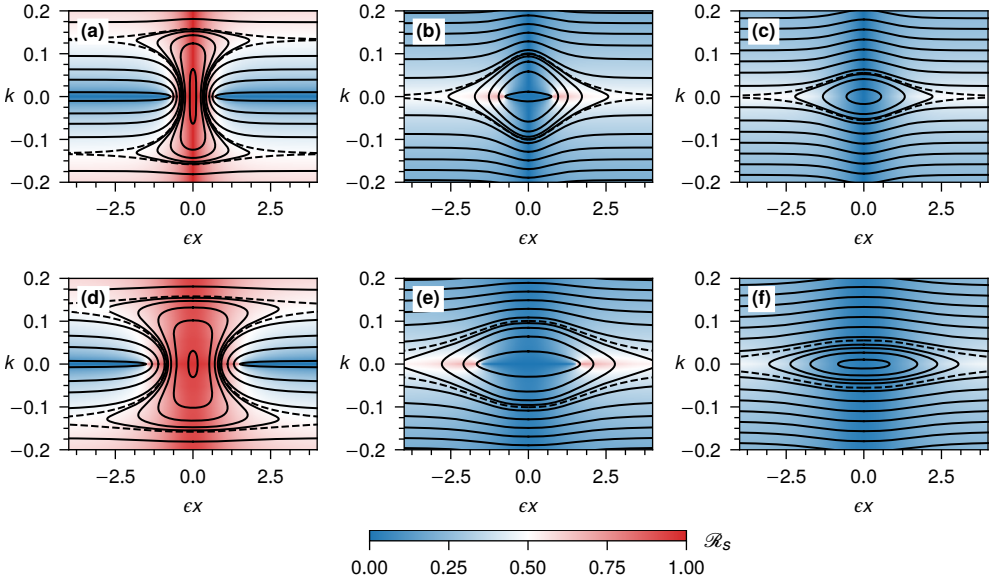


**Figure 5.11.** Rays trajectories for a curved shell with a tanh-type curvature profile (top panels) and a sech-type curvature profile (bottom panels) showcasing (a), (d) flexural waves, (b), (e) shear waves, and (c), (f) extensional waves. In all figures the closed curves represent rays associated with quantized bound states and the dashed black curves represent the highest frequency for which there is a bound state. The phase portraits have also been color coded with the ratio $\mathscr{R}_S$ defined in Eq. (5.41).

gets reflected back. Intuitively, we therefore expect bound states to occur for all three wave polarizations.

For ray analysis, we usually directly work with the eigenvalues of the dispersion matrix $\mathsf{D}^{(0)}$. This proves to be difficult for the shell as the expressions for the eigenvalues are unwieldy. In the absence of mode conversion, however, only one eigenvalue, say $\lambda$, will vanish at a given phase-space point $(x,k)$, causing the determinant $\det\mathsf{D}^{(0)}$ to vanish as well. Hence, the rays determined by $\lambda(x,k;\omega)=0$ and $\det\mathsf{D}^{(0)}=0$ are identical, allowing us to use $\det\mathsf{D}^{(0)}$ as the ray Hamiltonian.[13] Using $\det\mathsf{D}^{(0)}$ instead of $\lambda$ would amount to a trivial reparameterization of Hamilton's equations [163]. Judicious choices of the initial conditions, i.e., the coordinates $(x,k)$ and frequency $\omega$, obtained from the local dispersion curves at a given $(x,k)$, would determine the type of wave the ray represents. When analyzing a given wave type, it is also useful to compute an amplitude ratio, analogous to the one we used for the curved rod, and

[13] Let us remark that the ray trajectories can also be obtained by treating $\omega^2$ in Eq. (5.35) as the ray Hamiltonian. However, given the general awkwardness of the expressions, it is better to work with $\det\mathsf{D}^{(0)}$ instead.

defined by

$$\mathscr{R}_s = \frac{|\zeta|}{|\zeta| + |u| + |v|} \sim \frac{|\tau_1|}{|\tau_1| + |\tau_2| + |\tau_3|}. \tag{5.41}$$

Above, we have also made use of the fact that the wave field is asymptotic to the polarization vector $\tau$ to write $\mathscr{R}_s$ in terms of the components of $\tau$. With the above definition, flexural waves in a flat plate have $\mathscr{R}_s = 1$, whereas both shear and extensional waves have $\mathscr{R}_s = 0$. In a curved shell, because we expect both the normal and tangential components of the wave field to be significant, $\mathscr{R}_s$ for the three wave polarizations would deviate from their flat-plate counterparts. For all three wave types, a significant amount of normal and tangential contribution to the displacement field is indicated by values of $\mathscr{R}_s$ in the range $1/3 < \mathscr{R}_s < 1/2$. We discuss flexural waves first.

#### *Flexural waves*

Our intuitive expectation of flexural bound states is confirmed by the actual ray trajectories for flexural waves showcased in Figs. 5.11(a) and 5.11(d). For frequencies slightly above the cut-on frequency at $x = 0$, the rays appear in the form of closed, vertically elongated orbits that remain confined to a region where the curvature is very small. At larger frequencies, the rays begin to enter regions of higher curvature, and orbits change from being elliptical to highly eccentric, "peanut"-shaped curves. For these orbits, we have a total of six caustics where $\dot{x} = 0$. [Compare Figs. 5.11(a) and 5.11(d) with Fig. 4.1(b).] Two of these caustics, which are on the $x$ axis, are the usual classical turning points where $k = 0$. At the other four caustics $k \neq 0$, and they arise due to the two double-well minima in the (local) dispersion curves where $\mathrm{d}\omega/\mathrm{d}k = 0$.

For the peanut-shaped orbits, bound states do not occur beyond a frequency where the four caustics with $k \neq 0$ get pushed to $x = \pm\infty$. Since the absolute curvature $|m(\pm\infty)| = b$ for both curvature profiles, the largest frequency for which we see a bound state—represented by the dashed rays in Figs. 5.11(a) and 5.11(d)—must be the frequency of the double-well minimum in the dispersion curves for $m = b$. For small enough $b$, using Eq. (5.39), we find this minimum to be $\omega^2 \approx 2\sqrt{l^4 b^2 (1-\eta^2)}$. This expression, however, turns out to overestimate the actual minimum for larger $b$, and we must find it numerically. Also, this minimum exists only when $b \gtrsim l^2/\sqrt{1-\eta^2}$. For smaller $b$, rays of the bound states remain elliptical in nature, with the classical turning points being the only caustics.

An example profile of a low-frequency flexural bound state is shown in Fig. 5.12(a). This state has negligible tangential components $u$ and $v$, and remains confined to a region of similar extent as the classical turning points. At higher frequencies, flexural bound states grow beyond the classical turning points and enter regions of higher curvature. Here curvature effects become more prominent, and the states tend to have both tangential and normal components as seen from Fig. 5.12(b), and the color coding of the phase portraits. They, however, remain confined to a region of similar extent as the four caustics with $\dot{x} = 0$.

#### *Shear waves*

From the phase portraits in Fig. 5.11(b) and 5.11(e), we see that shear waves also form bound rays confined between two classical turning points. As the frequency of the orbits increase,

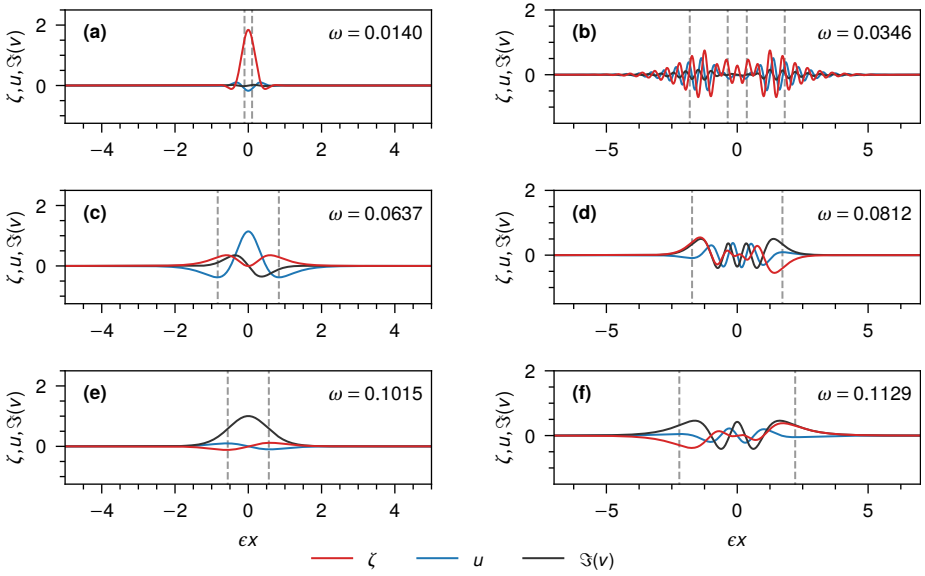


**Figure 5.12.** Numerical eigenmodes of a curved shell with a tanh-type curvature profile showing (a),(b) flexural, (c),(d) shear, and (e),(f) extensional bound states. In our phase convention, $\zeta$ and $u$ are always real, whereas $v$ is always complex, which is why we only show its imaginary component $\Im(v)$. Left panels depict low-frequency bound states, whereas the ones on the right panels have a higher frequency. The dashed vertical lines indicate the locations of the caustics.

the turning points move to $x = \pm\infty$, where the absolute curvature for both curvature types is $b$. Thus, the largest frequency for which we observe shear bound states is the shear-wave cut-on frequency for a cylindrical shell of curvature equal to $b$, obtained by setting $m = b$ in the second root of Eq. (5.37). Beyond this frequency, shear waves form unbound states. The rays of shear bound states are elongated along the $x$ direction as the spectral gap in the local dispersion curve does not begin increasing until $m(x) > m_\ddagger$. For the same reason, shear waves do not get localized if $b < m_\ddagger$. It is then natural to wonder if the localization of shear waves seen in the phase portraits is an artifact of having chosen a relatively large value of $b = 0.1$. But from Eq. (5.40) we see that $m_\ddagger$ can be made arbitrarily small by adjusting the value of transverse wave number $l$, so even for small $b$ we would expect shear bound states.

Example profiles of two shear bound states are shown in Figs. 5.12(c) and 5.12(d). The first one has a frequency that is only slightly above the shear-wave cut-on frequency and hence its local wave number $k$ is small. Therefore, its (local) wave vector is predominantly in the $y$ direction [see Fig. 5.1(b)]. Furthermore, as expected from the transverse nature of shear waves, the dominant tangential component is $u$, which is the displacement along $x$. The second shear bound state shown in Fig. 5.12(f) has a higher frequency, causing it to spread to regions of higher curvature, where curvature effects become more prominent. This can also

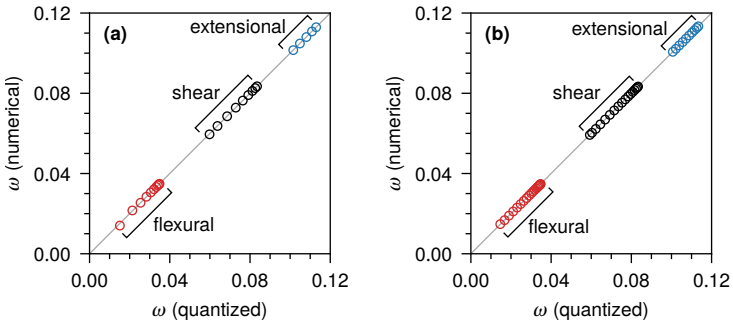


**Figure 5.13.** Bound-state frequencies of a curved shell obtained from numerics compared to that obtained through quantization for (a) tanh-type curvature profile $m_1(x)$ and (b) sech-type curvature profile $m_2(x)$. For both plots, the gray guide line in the background represents $\omega$ (quantized) = $\omega$ (numerical).

be inferred from the color coding of Figs. 5.11(b) and 5.11(e), which shows that shear bound states develop a significant normal component at higher frequencies.

*Extensional waves*

Color-coded phase portraits of extensional waves indicating bound states are shown in Figs. 5.11(c) and 5.11(f). The caustics for these bound states are the usual classical turning points with $k = 0$. For higher-frequency bound states, these points move to $\pm\infty$, where the absolute curvature is $b$. Hence, the largest frequency for which we observe extensional bound states must be the cut-on frequency for extensional waves in a shell having a curvature equal to $b$, obtained by putting $m = b$ in the first root in Eq. (5.37).

Example profiles of two extensional bound states are shown in Figs. 5.12(e) and 5.12(f). Low-frequency extensional bound states, such as the one in Fig. 5.12(e), are expected to displace the shell predominantly in the $y$ direction as seen from the comparatively large values of the tangential component $v$. The bound state in Fig. 5.12(f) has a slightly higher frequency, causing it to spread to regions of higher curvature, where it develops a significant normal component, which we also infer from the color coding of Figs. 5.11(c) and 5.11(f).

*Bound states and quantization*

To find the bound-state frequencies, we first set $k = 0$ in Eq. (5.31) and rearrange terms to find that that classical turning points $x^\star$ are given by the solutions to the implicit equation

$$m^2(x^\star) = \left[\frac{\omega^2 - l^2}{\omega^2 - (1-\eta^2)\, l^2}\right] \left(\omega^2 - l^4\right). \tag{5.42}$$

Depending on the value of $\omega$, the turning points found using the above equation could correspond to turning points on the bound rays of all three waves. We use the same quantization procedure as for the rod to determine the bound-state frequencies (Appendix 5.B). Furthermore, the extra phases $\gamma_G$ and $\gamma_{NG}$ in the quantization condition in Eq. (4.58) vanish for the

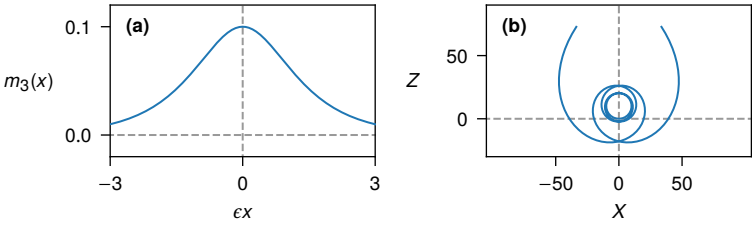


**Figure 5.14.** (a) Curvature profile $m_3(x) = b\,\mathrm{sech}\,\epsilon x$ and (b) a curve with this curvature profile in Cartesian space with coordinates $X$ and $Z$. As the curve has multiple self-intersections, only a physically unrealistic "phantom" shell can have this curvature profile. The parameters $b = 0.1$ and $\epsilon = 0.01$.

shell equations as well (Appendix 5.A). The Keller–Maslov index continues to be $\alpha = 2$ for all orbits—including the peanut-shaped orbits with six caustics—as they can be smoothly deformed into a circle centered around the origin [134]. From Fig. 5.13, we see that the bound-state frequencies obtained through quantization agree rather well with the numerical values for both curvature profiles.

Although waves of all three types form bound states, from our preceding analyses and Fig. 5.13, we see that flexural bound states appear first, followed by shear and extensional bound states. Thus, in very long shells, shear waves form bound states that lie in a quasi-continuum of flexural waves spread across the shell. Likewise, extensional bound states would lie in a quasi-continuum of unbound flexural and shear waves. Similar to the curved rod, the bound states of the curved shell are also of definite parity when the curvature profile $m(x)$ is odd or even. More specifically, when $m(x)$ is even, the components $\zeta(x)$ and $v(x)$ have the same parity, with $u(x)$ having the opposite parity. For odd $m(x)$, however, $\zeta(x)$ and $u(x)$ have the same parity, with $v(x)$ having the opposite parity, as can be seen from the example bound states in Fig. 5.12.

## 5.6 Shells with sharp bends

Flexural bound states have also been predicted to exist in curved shells around points of maximal curvature [120]. To illustrate this, we consider the curvature profile[14] $m_3(x) = b\,\mathrm{sech}(\epsilon x)$. For the parameter values $b = 0.1$ and $\epsilon = 0.01$, Fig. 5.14 shows the function $m_3(x)$ and the shape of a curve with this curvature profile. Clearly, the curve intersects itself at multiple points, and hence $m_3(x)$ can only be the curvature profile of a physically unrealizable "phantom" shell that allows self-intersections. Although self-intersections can be avoided by choosing a smaller value for $b$ (or a larger value for $\epsilon$), doing so would lead to the formation of very few or barely discernible bound states. For this reason, we will continue to work with the choice $b = 0.1$ so that the bound states appear more conspicuous.

[14] The authors of Refs. [120, 155] use a much more complicated curvature profile in their analysis. However, for our purposes, $m_3(x)$ has the basic features of the profiles that these authors consider.

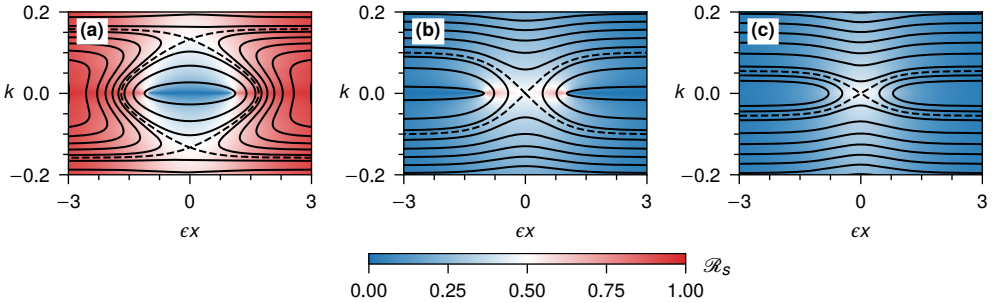


**Figure 5.15.** Rays trajectories for a curved shell with the curvature profile $m_3(x)$ showing (a) flexural waves, (b) shear waves, and (c) extensional waves. Only flexural waves form bound states. The phase portraits are color coded with the ratio $\mathscr{R}_s$ defined in Eq. (5.41).

Ray trajectories of all three wave types for the curvature profile $m_3(x)$ are shown in Fig. 5.15. As we see from this figure, only flexural waves form bound states. For this reason, we forgo a more detailed analysis of shear and extensional waves, and concentrate on flexural waves alone. In the phase portrait of Fig. 5.15(a), we see that in addition to the nonlinear center at the origin, there are two saddle points on the $k$ axis. The saddle points arise because of the double-well minimum in the flexural dispersion curves where $\mathrm{d}\omega/\mathrm{d}k = 0$. A more detailed analysis of the phase portraits and bound states of flexural waves reveals the following:

(i) The existence of bound states around regions of high curvature cannot simply be explained by considering the gap and/or cut-on frequency in the dispersion curves—after all, as we move away from regions of high curvature to regions of low curvature, this gap decreases. In reality, these bound states arise due to the double-well nature of the flexural dispersion curves at nonzero curvature. Indeed, if $l = 0$ or $b \lesssim l^2/\sqrt{1-\eta^2}$, the dispersion curves fail to have a double-well nature and bound states do not exist.[15]

(ii) Smaller orbits in the phase portraits represent bound states of *higher* frequency. Hence, higher-frequency bound states are represented by a smaller quantum number $n$. The smallest bound orbit has a frequency slightly above the cut-on frequency of flexural waves at curvature $m = b$. For $b > m_\ddagger$, as the frequency of the orbit approaches the cut-on frequency, the whole orbit flattens to a small line segment on the $x$ axis; for $b < m_\ddagger$, the orbit shrinks to the origin. [See Eq. (5.40) for the definition of $m_\ddagger$.]

(iii) As the frequency of the bound orbits gets smaller, they increase in size, but remain confined between the two heteroclinic orbits [shown as dashed curves in Fig. 5.15(a)] connecting the saddle points on the $k$ axis. The heteroclinic orbits have a frequency equal to the double-well minimum in the dispersion curves, which puts a limit on the *lowest* frequency for which we expect a bound mode.

[15] This should be contrasted with flexural bound states around points of minimal curvature, which continue to exist even for $b \lesssim l^2/\sqrt{1-\eta^2}$; see the discussion on p. 92.

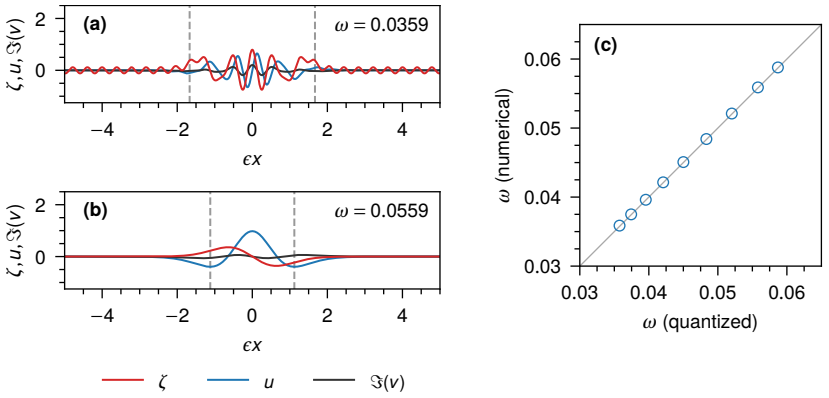


**Figure 5.16.** (a), (b) Example flexural bound states for a shell with curvature profile $m_3(x)$; the grey vertical lines indicate the locations of the classical turning points. Tunneling effects around the saddle points cause the lower-frequency modes, such as the one in (a), to leak out. (c) Comparison between the quantization results and numerics; the grey guideline in the background represents $\omega$ (quantized) = $\omega$ (numerical).

(iv) Incoming and outgoing rays close to the saddle point are nearly tangent to each other, leading to the formation of a tunneling region around the saddle point.[16] In this region, waves represented by one ray can transfer their energy to waves represented by the other ray [163], which causes the lower-frequency bound states to "leak out" as shown in Fig. 5.16(a). This, however, does not seem to have a considerable effect on the quantization results, and we recover the frequencies of the observed bound modes accurately [see Fig. 5.16(c)].

(v) Finally, bound states at higher frequencies, such as the one in Fig. 5.16(b), tend to have fewer nodes and appear to be more spatially confined.[17] They also have a more significant tangential contribution as evident from the color coding of Fig. 5.15(a).

In summary, flexural bound states in a shell can also develop around points of maximal curvature. Let us again remark, however, that for the parameter choices $b = 0.1$ and $\epsilon = 0.01$, the curvature profile $m_3(x)$ is a physical unrealizable one. On the other hand, numerical experiments with more realistic shells (with small $b$ and/or large $\epsilon$) revealed very few or practically nonexistent bound states—something that we expect to happen in real experiments as well.

## 5.7 Concluding remarks

In this chapter we have considered the localization of waves in thin elastic structures induced by variations in the structure's curvature profile. For both the example structures we

[16] Mohammed *et al.* [120] presents a detailed analysis of this tunneling phenomenon.

[17] This is in contrast to all the bound states, including those on the curved rod, that we have described so far.

considered, bound states develop around points where the structure's absolute curvature has a minimum. In case of the shell, flexural, shear, and extensional waves form bound states. Additionally, flexural bound states can also develop around points of a shell where the absolute curvature has a maximum. In contrast to shells, bound states in a curved rod (which are always extensional in nature) only exist around points where the absolute curvature has a minimum. These findings set the stage for the design of simple devices capable of inducing wave localization without relying on metamaterials with nontrivial microstructure.

Semiclassical approximation presents challenges of its own when used to study multi-component waves, particularly due to the presence of nontrivial phases in the quantization rule. The rod and shell equations we use in this chapter, however, have properties that cause these phases to vanish. Nevertheless, topologically protected waves in continuous media can arise when this phase is nonzero, especially when time-reversal symmetry is broken [165]. For this reason, it is worthwhile to explore the use of semiclassical methods in problems with broken time-reversal symmetry such as those in rotating elastic media [110], fluids with odd viscosity [157], and magnetoelastic waves [5], where one would generically expect this phase to be nonzero.

Ray tracing methods have continued to receive extensive attention in recent years. An interesting line of work involves the phenomenon of branched flows—the spatial branching of rays as a result of random, but weak inhomogeneities in the medium, the end result being the formation of tree-like structures and fluctuations of extreme intensity [64]. Branched flows have been observed in a host of systems, including tsunami waves in the ocean, light propagation in soap films, the flow of electrons in semiconductors, etc. In elastodynamics, branched flows have been shown to exist in thin elastic plates and cylinders with varying thickness profiles [73–75]. It would be interesting to extend these results to consider shells with random curvature profiles, akin to crumpled paper [53].

## 5.A Additional phases

In this appendix we will look at situations where the extra phase $\gamma$ that appears in the quantization condition in Eq. (4.58) vanishes. As we discussed during the derivation of Eq. (4.56), $\gamma = \gamma_{\mathrm{G}} + \gamma_{\mathrm{NG}}$, where $\gamma_{\mathrm{G}}$ is the term that gives rise to a nonzero geometric phase and $\gamma_{\mathrm{NG}}$ is the other (non-geometric) term. First, we shall analyze general $N$-component wave equations with an $N \times N$ dispersion matrix $\mathsf{D}^{(0)}$.

### 5.A.1 General wave equations

Consider a general $N$-component polarization vector $\tau$ of the dispersion matrix $\mathsf{D}^{(0)}$, given by

$$\tau = \left[ r_1(x,k)e^{i\varphi_1(x,k)},\ r_2(x,k)e^{i\varphi_2(x,k)},\ \ldots,\ r_N(x,k)e^{i\varphi_N(x,k)} \right]. \tag{5.43}$$

Above, we have expressed the $j$th component $\tau_j$ in terms of a real amplitude $r_j(x,k)$ and a phase $\varphi_j(x,k)$, both of which are functions of the phase-space coordinates $(x,k)$. Since $\tau$ is normalized, we have $\|\tau\|^2 = \sum_{j=1}^{n} r_j^2(x,k) = 1$ for all $(x,k)$. Putting Eq. (5.43) in Eq. (4.56), we

see that the rate of change of the first (geometric) phase $\gamma_G$ is

$$\begin{aligned}\dot{\gamma}_G = i\tau_j^* \{\tau_j, \lambda\} &= i r_j \{r_j, \lambda\} - r_j^2 \{\varphi_j, \lambda\} \\ &= (i/2)\{\|\tau\|^2, \lambda\} - r_j^2 \{\varphi_j, \lambda\} \\ &= -r_j^2 \{\varphi_j, \lambda\}.\end{aligned} \tag{5.44}$$

In the last step above, we have made use of the fact that $\|\tau\| = 1$ always, so $\{\|\tau\|^2, \lambda\}$ vanishes. Clearly, if the phases $\varphi_j(x,k)$ are constants, then $\dot{\gamma}_G$ vanishes. More generally, $\dot{\gamma}_G$ would vanish if all $(x,k)$ dependence in the phases $\varphi_j$ can be removed by an overall rephasing of $\tau$ (as such a rephasing does not affect the normalization of $\tau$). In other words, only the relative phases between the components of $\tau$ contribute to $\dot{\gamma}_G$. From here on we assume that the phases $\varphi_j$ are constants, so all Poisson brackets involving $\varphi_j$ can be set to zero. In that case $\dot{\gamma}_G = 0$ everywhere on the phase space, and the accumulated phase $\gamma_G$ as we move along an orbit can be taken to be zero.

But what about the second (non-geometric) phase $\gamma_{NG}$? From Eq. (4.56) we see that the rate of change of $\gamma_{NG}$ is given by

$$\begin{aligned}\dot{\gamma}_{NG} &= (i/2)\mathsf{D}^{(0)}_{jk}\left\{\tau_j^*, \tau_k\right\} - \tau_j^* \mathsf{D}^{(1)}_{jk}\tau_k \\ &= (i/2)\sum_{j<k}\left(\mathsf{D}^{(0)}_{jk} e^{-i\varphi_{jk}} - \mathsf{D}^{(0)^*}_{jk} e^{i\varphi_{jk}}\right)\{r_j, r_k\} - \tau_j^* \mathsf{D}^{(1)}_{jk}\tau_k.\end{aligned} \tag{5.45}$$

In the last step above, we have used Eq. (5.43) to simplify the first term on the RHS and have defined $\varphi_{jk} = \varphi_j - \varphi_k$. We have also made use of the Hermiticity of $\mathsf{D}^{(0)}$ to express the first term in terms of the off-diagonal entries of $\mathsf{D}^{(0)}$. From Eq. (5.45) we see that even when the phases $\varphi_j$ are constants, $\gamma_{NG}$ could be nonzero. However, in the following subsection we show that for wave equations of thin elastic structures with certain invariant properties, in addition to a vanishing $\gamma_G$, the phase $\gamma_{NG}$ vanishes as well.

### 5.A.2 Wave equations of thin elastic structures

We begin by noting that the rod equations, Eq. (5.3), as well as the shell equations, Eq. (5.28), remain invariant on simultaneously inverting the sign[18] of the spatial derivatives and the tangential components of the displacement field, i.e., under $(\partial_x, \partial_y) \to (-\partial_x, -\partial_y)$ and $(\zeta, u, v) \to (\zeta, -u, -v)$. This invariance can be traced back to the invariance of the strain expressions[19] used to derive these equations. The same invariance is also found in many higher-order theories of rods [19, 167] and shells [34]. For these reasons, it is useful to consider a general linear elastodynamic equation involving a 3-component wave field $\Psi = (\zeta, u, v)$ and possessing this invariance, and given by

$$\partial_t^2 \Psi(x,y,t) + \widehat{\mathsf{H}}\Psi(x,y,t) = 0 \quad \text{with} \quad \widehat{\mathsf{H}} = \begin{pmatrix} \widehat{Z} & \widehat{A} & \widehat{B} \\ \widehat{A}^\dagger & \widehat{U} & \widehat{C} \\ \widehat{B}^\dagger & \widehat{C}^\dagger & \widehat{V} \end{pmatrix}, \tag{5.46}$$

[18] Note that under $(\partial_x, \partial_y) \to (-\partial_x, -\partial_y)$ derivatives of the curvature transform as $m'(x) \to -m'(x)$.

[19] More specifically, this invariance arises when the linearized extensional and bending strains are comprised of terms involving only odd derivatives of $u$ and $v$, and even derivatives of $\zeta$, as in models based on the Kirchhoff–Love assumptions [151].

Above, the entries of $\widehat{\mathsf{H}}$ are linear differential operators comprised of powers of $\partial_x$ and $\partial_y$. Also, the diagonal entries $\widehat{Z}$, $\widehat{U}$, and $\widehat{V}$ are Hermitian operators and $\widehat{A}^\dagger$ is the Hermitian adjoint of $\widehat{A}$. Additionally, since Eq. (5.46) represents an elastodynamic system, we assume that the coefficients of all the derivatives in $\widehat{\mathsf{H}}$ are real so that $\Psi$ can be taken to be real as well.

Any invariance possessed by Eq. (5.46) must be shared by the (potential) energy density $\mathscr{J} = \frac{1}{2}\Psi^{\mathsf{T}}\widehat{\mathsf{H}}\Psi$ used to derive it from Hamilton's principle. If $\mathscr{J}$ is to be invariant under $(\partial_x, \partial_y) \to (-\partial_x, -\partial_y)$ and $(\zeta, u, v) \to (\zeta, -u, -v)$ for an arbitrary $\Psi$, the off-diagonal operators $\widehat{A}$ and $\widehat{B}$ must be odd under $(\partial_x, \partial_y) \to (-\partial_x, -\partial_y)$. In other words, they can only have terms involving exactly one odd power of $\partial_x$ (or $\partial_y$). Odd powers of $\partial_x, \partial_y$ acquire complex coefficients when expressed in terms of the momentum operator: $\partial_x^{2n+1} = (-1)^n i\hat{k}^{2n+1}$. Meanwhile, coefficients of even powers of $\partial_x, \partial_y$ remain real: $\partial_x^{2n} = (-1)^n \hat{k}^{2n}$. Using the rules in Eq. (4.62), we therefore conclude that the lowest-order symbols of the off-diagonal operators $\widehat{A}$ and $\widehat{B}$ must be purely complex (as $\widehat{\mathsf{H}}$ did not have complex coefficients to begin with). From Eq. (4.62) we also see that the $\mathcal{O}(\epsilon)$ corrections to these symbols must be real. Therefore, we can write down the symbols of the operators $\widehat{A}$ and $\widehat{B}$ as

$$\begin{aligned} A &= iA^{(0)} + \epsilon A^{(1)} + \mathcal{O}(\epsilon^2), \\ B &= iB^{(0)} + \epsilon B^{(1)} + \mathcal{O}(\epsilon^2). \end{aligned} \tag{5.47}$$

where $A^{(0)}, B^{(0)}$, etc., are real functions. A similar reasoning would reveal that the operator $\widehat{C}$ must be even under $(\partial_x, \partial_y) \to (-\partial_x, -\partial_y)$, and consequently, its symbol is of the form

$$C = C^{(0)} + i\epsilon C^{(1)} + \mathcal{O}(\epsilon^2), \tag{5.48}$$

where $C^{(0)}$ and $C^{(1)}$ are real. As $\widehat{\mathsf{H}}$ is a Hermitian operator, the symbols of the diagonal entries are all real and $\mathcal{O}(\epsilon)$ corrections to these symbols must vanish.

For finding the eigenmodes, after Fourier transforming in time, we define $\widehat{\mathsf{D}} = \widehat{\mathsf{H}} - \omega^2 \mathsf{I}_3$ and convert $\widehat{\mathsf{D}}$ to its symbol form $\mathsf{D} = \mathsf{D}^{(0)} + \epsilon \mathsf{D}^{(1)} + \mathcal{O}(\epsilon^2)$. From the above discussion, we see that most general dispersion matrix $\mathsf{D}^{(0)}$ and its $\mathcal{O}(\epsilon)$ correction $\mathsf{D}^{(1)}$ that can be written down is of the form

$$\mathsf{D}^{(0)} = \begin{pmatrix} Z^{(0)} - \omega^2 & iA^{(0)} & iB^{(0)} \\ -iA^{(0)} & U^{(0)} - \omega^2 & C^{(0)} \\ -iB^{(0)} & C^{(0)} & V^{(0)} - \omega^2 \end{pmatrix} \quad \text{and} \quad \mathsf{D}^{(1)} = \begin{pmatrix} 0 & A^{(1)} & B^{(1)} \\ A^{(1)} & 0 & iC^{(1)} \\ B^{(1)} & -iC^{(1)} & 0 \end{pmatrix}. \tag{5.49}$$

A polarization vector $\tau$ is defined up to an overall phase and normalization by $\mathsf{D}^{(0)}\tau = 0$. Direct inspection reveals that for $\mathsf{D}^{(0)}$ defined in Eq. (5.49), we can take $\tau$ to be of the form[20]

$$\tau = \begin{pmatrix} \tau_1 \\ i\tau_2 \\ i\tau_3 \end{pmatrix}, \tag{5.50}$$

where $\tau_1$, $\tau_2$, and $\tau_3$ are *real* functions defined on the phase space. Clearly, the relative phases $\varphi_{12}$ and $\varphi_{13}$ between the components of $\tau$ are either $\pm\pi/2$ or 0 (when $\tau_1$ or $\tau_2$ vanishes).

[20] To see this more explicitly, take $\tau = (r_1 e^{i\phi_1}, \tau_2, \tau_3)$ and solve for the components $\tau_2$ and $\tau_3$ from $\mathsf{D}^{(0)}\tau = 0$. Upon rephasing $\tau$ by $e^{-i\phi_1}$, we find that $\tau$ is of the general form in Eq. (5.50).

Likewise, the relative phase $\varphi_{23}$ is either 0 or $\pi$. Because the relative phases are constants, from our discussion in the previous subsection, it then follows that the geometric phase $\gamma_G = 0$. When the matrix $\mathsf{D}^{(1)}$ is of the form in Eq. (5.49), using the polarization vector $\tau$ in Eq. (5.50), a straightforward computation shows that second term in the expression for $\dot{\gamma}_{NG}$, Eq. (5.45), vanishes. Next, we note that $\mathsf{D}^{(0)}_{12} e^{-i\varphi_{12}} = \pm A^{(0)}$, $\mathsf{D}^{(0)}_{13} e^{-i\varphi_{13}} = \pm B^{(0)}$, and $\mathsf{D}^{(0)}_{23} e^{-i\varphi_{23}} = \pm C^{(0)}$, are all real. From Eq. (5.45) it then follows that $\dot{\gamma}_{NG}$ vanishes, and we can take $\gamma_{NG}$ to be zero as well.

The dispersion matrices for the thin shell we considered in Eq. (5.30) are of the form in Eq. (5.49), and hence the phase $\gamma = \gamma_G + \gamma_{NG}$ is zero for the shell. It can be verified that the dispersion matrices for many higher-order shell theories [34] would also be of this form. For the curved rod, the dispersion matrices in Eq. (5.5) are identical to the shell dispersion matrices once we delete the third row and column, and set $l = 0$. Proceeding by arguments similar to previous ones, we see that the extra phase $\gamma$ vanishes for the rod as well.

## 5.B Numerical details

To find the eigenmodes numerically, we solve the rod and shell equations, Eqs. (5.3) and (5.28), with $x \in \mathcal{X} = [-1000, 1000]$, using Dedalus [15] with a Chebyshev spectral decomposition and 2048 modes.[21] Bound states are identified by manual examination of the eigenmode profiles. To test the robustness of the bound states, we independently use clamped, simply supported, and mixed clamped–simply supported boundary conditions for both the rod and the shell. At the clamped end of a rod, the geometric boundary conditions are $\zeta(x) = \partial_x \zeta(x) = u(x) = 0$ [88]. For the shell, at the clamped end, we additionally have $v(x) = 0$ as well. At a simply supported end of a rod, we have the geometric boundary condition $\zeta(x) = u(x) = 0$ and the natural boundary condition $\partial_x^2 \zeta(x) = 0$ (no bending moment) [45]. In case of a shell, at a simply supported end, we have $\zeta(x) = u(x) = v(x) = 0$ and $\partial_x^2 \zeta(x) - \eta l^2 \zeta(x) = 0$ [177].

For finding the quantized frequencies numerically, for a given $n \in \mathbb{N}_0$, we start with an approximate guess for the frequency $\omega$ based on the numerical results. We then numerically integrate the ray equations starting at one of the classical turning points on the $x$ axis, e.g., the one at $x = -x^\star$, until the ray reaches the other turning point at $x = x^\star$ (see Fig. 4.1). Next, we compute

$$n(\omega) = (\pi\epsilon)^{-1} \int_{-x^\star}^{x^\star} \mathrm{d}x\, k(x) - \tfrac{1}{2} \tag{5.51}$$

using points $[x, k(x)]$ from the ray trajectory, with the integral evaluated by quadrature. For a general $\omega$, the estimated $n(\omega)$ will not be integer-valued. Quantized frequencies $\omega$ can be obtained by solving $n(\omega) = n$ using a numerical root finder. Alternatively, we could minimize the absolute "error" $|n - n(\omega)|$ using random values of $\omega$ spread around the initial guess, and take the quantized frequency to be $\mathrm{argmin}_\omega |n - n(\omega)|$. For the results reported in this chapter, this error is less than $10^{-10}$.

[21] Numerical code is publicly available at https://github.com/manu-mannattil/glwtes.

## Bibliography

*Note.*— All references that are publicly available on the Internet have archival versions at the Wayback Machine.*


[1] Abraham, F. F. and D. R. Nelson, "Fluctuations in the flat and collapsed phases of polymerized membranes," J. Phys. France **51**, 2653 (1990); cited on p. 51.

[2] Anderson, P. W., "More Is Different," Science **177**, 393 (1972); cited on p. 1.

[3] Arnold, V. I., *Mathematical Methods of Classical Mechanics*, trans. by Vogtmann, K. and Weinstein, A. (Springer, New York, NY, 1978); cited on p. 1.

[4] Audoly, B. and Y. Pomeau, *Elasticity and Geometry* (Oxford University Press, Oxford, 2010); cited on p. 1.

[5] Baños, A., "Normal Modes Characterizing Magnetoelastic Plane Waves," Phys. Rev. **104**, 300 (1956); cited on p. 98.

[6] Bender, C. M. and S. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw Hill, New York, NY, 1978); cited on pp. 8, 89.

[7] Berk, H. L. and D. Pfirsch, "KB method for systems of integral equations," J. Math. Phys. **21**, 2054 (1980); cited on p. 61.

[8] Bernstein, I. B., "Geometric optics in space- and time-varying plasmas," Phys. Fluids **18**, 320 (1975); cited on p. 61.

[9] Berry, M., M. E. Lee-Trimble, and C. D. Santangelo, "Topological transitions in the configuration space of non-Euclidean origami," Phys. Rev. E **101**, 043003 (2020), arXiv:1910.01008 [cond-mat.soft]; cited on p. 46.

[10] Berry, M. V., "Quantal phase factors accompanying adiabatic changes," Proc. R. Soc. A **392**, 45 (1984); cited on p. 9.

[11] Bohigas, O., "Singularity Set Computation: A Hands-On Session with the CUIK Suite," in *Singular Configurations of Mechanisms and Manipulators*, edited by Müller, A. and Zlatanov, D. (Springer International Publishing, Cham, 2019), pp. 39–65; cited on p. 13.

[12] Bowick, M. J., S. M. Catterall, M. Falcioni, G. Thorleifsson, and K. N. Anagnostopoulos, "The Flat Phase of Crystalline Membranes," J. Phys. I France **6**, 1321 (1996), arXiv:cond-mat/9603157; cited on p. 51.

[13] Breitung, K. W., *Asymptotic Approximations for Probability Integrals* (Springer, Berlin, 1994); cited on p. 25.

[14] Buell, W. F. and B. A. Shadwick, "Potentials and bound states," Am. J. Phys. **63**, 256 (1995); cited on p. 82.


*https:/web.archive.org

[15] Burns, K. J., G. M. Vasil, J. S. Oishi, D. Lecoanet, and B. P. Brown, "Dedalus: A flexible framework for numerical simulations with spectral methods," Phys. Rev. Research **2**, 023068 (2020), arXiv:1905.10388 [astro-ph.IM]; cited on p. 101.

[16] Calladine, C., "Buckminster Fuller's 'Tensegrity' structures and Clerk Maxwell's rules for the construction of stiff frames," Int. J. Solids Struct. **14**, 161 (1978); cited on pp. 17, 49.

[17] Chaichian, M. and A. Demichev, *Path Integrals in Physics*, Vol. 1: *Stochastic Processes and Quantum Mechanics* (IoP Publishing, Bristol, 2001); cited on pp. 57, 67, 68.

[18] Chen, B. G. and C. D. Santangelo, "Branches of Triangulated Origami Near the Unfolded State," Phys. Rev. X **8**, 011034 (2018), arXiv:1706.03869 [cond-mat.soft]; cited on pp. 5, 31, 41, 43.

[19] Chidamparam, P. and A. W. Leissa, "Vibrations of Planar Curved Beams, Rings, and Arches," Appl. Mech. Rev **46**, 467 (1993); cited on pp. 72, 99.

[20] Cohen, L., *The Weyl Operator and Its Generalization* (Birkhäuser, Basel, 2012); cited on p. 57.

[21] Cohen-Tannoudji, C., B. Diu, and F. Laloë, *Quantum Mechanics*, Vol. 1: *Basic Concepts, Tools, and Applications*, trans. by Hemley, S. R., Ostrowsky, N., and Ostrowsky, D., 2nd ed. (Wiley-VCH, Weinheim, 2019); cited on p. 77.

[22] Connelly, R. and S. D. Guest, *Frameworks, Tensegrities and Symmetry* (Cambridge University Press, Cambridge, 2022); cited on pp. 16, 18.

[23] Connelly, R. and H. Servatius, "Higher-order rigidity—What is the proper definition?" Discrete Comput. Geom. **11**, 193 (1994); cited on p. 30.

[24] Connelly, R. and W. Whiteley, "Second-Order Rigidity and Prestress Stability for Tensegrity Frameworks," SIAM J. Discrete Math. **9**, 453 (1996); cited on pp. 30, 50.

[25] Craster, R. V. and S. Guenneau, *Acoustic Metamaterials: Negative Refraction, Imaging, Lensing and Cloaking*, Vol. 166, Springer Series in Materials Science (Springer, Dordrecht, 2012); cited on p. 69.

[26] Curtis, R. and M. Steiner, "Configuration Spaces of Planar Pentagons," Am. Math. Mon. **114**, 183 (2007); cited on p. 46.

[27] Delplace, P., J. B. Marston, and A. Venaille, "Topological origin of equatorial waves," Science **358**, 1075 (2017), arXiv:1702.07583 [cond-mat.mes-hall]; cited on p. 71.

[28] Di Francesco, P., "Folding and coloring problems in mathematics and physics," Bull. Am. Math. Soc **37**, 251 (2000); cited on p. 5.

[29] Diaconis, P., S. Holmes, and M. Shahshahani, "Sampling from a Manifold," in *Advances in Modern Statistical Theory and Applications*, Vol. 10, Collections (Institute of Mathematical Statistics, Beachwood, OH, 2013), pp. 102–125; cited on p. 24.

[30] Dirac, P. A. M., *The Principles of Quantum Mechanics*, 4th ed. (Oxford University Press, London, 1958); cited on p. 32.

[31] Dodd, P. M., P. F. Damasceno, and S. C. Glotzer, "Universal folding pathways of polyhedron nets," Proc. Natl. Acad. Sci. U.S.A. **115**, E6690 (2018); cited on p. 5.

[32] Donelan, P. S., "Singularity-theoretic methods in robot kinematics," Robotica **25**, 641 (2007); cited on pp. 4, 5.

[33] Donnell, L. H., *Stability of thin-walled tubes under torsion*, tech. rep. 479 (National Advisory Committee for Aeronautics, Washington, D.C., 1933); cited on pp. 72, 85.

[34] Doyle, J. F., *Wave Propagation in Structures*, 3rd ed. (Springer, New York, NY, 2021); cited on pp. 70, 83, 99, 101.

[35] Duclos, P. and P. Exner, "Curvature-induced bound states in quantum waveguides in two and three dimensions," Rev. Math. Phys. **07**, 73 (1995); cited on p. 69.

[36] Dunn, K. E., F. Dannenberg, T. E. Ouldridge, M. Kwiatkowska, A. J. Turberfield, and J. Bath, "Guiding the folding pathway of DNA origami," Nature **525**, 82 (2015); cited on pp. 2, 51.

[37] E, W. and E. Vanden-Eijnden, "Metastability, conformation dynamics, and transition pathways in complex systems," in Multiscale Modelling and Simulation, edited by Attinger, S. and Koumoutsakos, P. (2004), pp. 35–68; cited on p. 51.

[38] Echenique, P., C. N. Cavasotto, and P. García-Risueño, "The canonical equilibrium of constrained molecular models," Eur. Phys. J. Spec. Top. **200**, 5 (2011), arXiv:1105.0374 [physics.chem-ph]; cited on pp. 5, 14.

[39] Ellis, R. S. and J. S. Rosen, "Asymptotic analysis of Gaussian integrals, II: Manifold of minimum points," Commun. Math. Phys. **82**, 153 (1981); cited on p. 25.

[40] Evans, A. A. and A. J. Levine, "Reflection and Refraction of Flexural Waves at Geometric Boundaries," Phys. Rev. Lett. **111**, 038101 (2013); cited on p. 69.

[41] Farber, M., *Invitation to Topological Robotics* (European Mathematical Society, Zurich, 2008); cited on pp. 4, 46.

[42] Farhat, M., S. Guenneau, S. Enoch, and A. B. Movchan, "Cloaking bending waves propagating in thin elastic plates," Phys. Rev. B **79**, 033102 (2009); cited on pp. 69, 71.

[43] Förster, C. and T. Weidl, "Trapped modes for an elastic strip with perturbation of the material properties," Q. J. Mech. Appl. Math. **59**, 399 (2006); cited on p. 69.

[44] Frenkel, D., "Simulations: The dark side," Eur. Phys. J. Plus **128**, 10 (2013), arXiv:1211.4440 [cond-mat.stat-mech]; cited on p. 51.

[45] Fung, Y. C., *Foundations of Solid Mechanics* (Prentice Hall, Englewood Cliffs, NJ, 1965); cited on pp. 87, 101.

[46] Gandikota, M. C. and A. Cacciuto, "The crumpling transition of active tethered membranes," Soft Matter **19**, 5328 (2023), arXiv:2306.13190 [cond-mat.soft]; cited on p. 51.

[47] Gandikota, M. C., S. Das, and A. Cacciuto, "Spontaneous crumpling of active spherical shells," preprint (2023), arXiv:2308.06416 [cond-mat.soft]; cited on p. 51.

[48] Gáspár, M. E. and P. Csermely, "Rigidity and flexibility of biological networks," Brief. Funct. Genom. **11**, 443 (2012), arXiv:1204.6389 [physics.bio-ph]; cited on p. 4.

[49] Germogenova, O. A., "Geometrical theory for flexure waves in shells," J. Acoust. Soc. Am. **53**, 535 (1973); cited on p. 89.

[50] Gō, N. and H. A. Scheraga, "On the Use of Classical Statistical Mechanics in the Treatment of Polymer Chain Conformation," Macromolecules **9**, 535 (1976); cited on p. 5.

[51] Goldstein, H., C. Poole, and J. Safko, *Classical Mechanics*, 3rd ed. (Addison-Wesley, New York, NY, 2002); cited on p. 16.

[52] Goldstone, J. and R. L. Jaffe, "Bound states in twisting tubes," Phys. Rev. B **45**, 14100 (1992); cited on p. 69.

[53] Gopinathan, A., T. A. Witten, and S. C. Venkataramani, "Trapping of vibrational energy in crumpled sheets," Phys. Rev. E **65**, 036613 (2002), arXiv:cond-mat/0109059 [cond-mat.mtrl-sci]; cited on p. 98.

[54] Graff, K. F., *Wave Motion in Elastic Solids* (Dover, Mineola, NY, 1991); cited on p. 83.

[55] Grashof, F., *Theoretische Maschinenlehre* (Voss, Leipzig, 1883); cited on pp. 5, 15.

[56] Gridin, D., A. I. Adamou, and R. Craster, "Trapped modes in bent elastic rods," Wave Motion **42**, 352 (2005); cited on p. 69.

[57] Gridin, D., R. V. Craster, and A. T. Adamou, "Trapped modes in curved elastic plates," Proc. R. Soc. A **461**, 1181 (2005); cited on p. 69.

[58] Guillemin, V. and A. Pollack, *Differential Topology* (Prentice Hall, Englewood Cliffs, NJ, 1974); cited on pp. 12, 13.
[59] Hansen, C., S. Snyder, X. Qiu, L. Brooks, and D. Moreau, *Active Control of Noise and Vibration*, 2nd ed. (CRC Press, Boca Raton, FL, 2012); cited on p. 71.
[60] Hartenberg, R. and J. Danavit, *Kinematic Synthesis of Linkages* (McGraw-Hill, New York, 1964); cited on pp. 5, 15.
[61] Hartmann, C., J. C. Latorre, and G. Ciccotti, "On two possible definitions of the free energy for collective variables," Eur. Phys. J. Spec. Top. **200**, 73 (2011); cited on pp. 23–25, 52.
[62] Hartmann, C., "Model Reduction in Classical Molecular Dynamics," PhD thesis (Freie Universität Berlin, 2007); cited on p. 25.
[63] Hartmann, C. and C. Schütte, "Comment on two distinct notions of free energy," Physica D **228**, 59 (2007); cited on pp. 24, 51.
[64] Heller, E. J., R. Fleischmann, and T. Kramer, "Branched flow," Phys. Today **74**, 44 (2021), arXiv:1910.07086 [cond-mat.soft]; cited on p. 98.
[65] Hernandez, A., M. F. Staddon, M. J. Bowick, M. C. Marchetti, and M. Moshe, "Anomalous elasticity of a cellular tissue vertex model," Phys. Rev. E **105**, 064611 (2022), arXiv:2109.10407 [cond-mat.soft]; cited on p. 51.
[66] Herschbach, D. R., H. S. Johnston, and D. Rapp, "Molecular Partition Functions in Terms of Local Properties," J. Chem. Phys. **31**, 1652 (1959); cited on pp. 14, 42.
[67] Hespenheide, B. M., D. J. Jacobs, and M. F. Thorpe, "Structural rigidity in the capsid assembly of cowpea chlorotic mottle virus," J. Phys.: Condens. Matter **16**, S5055 (2004); cited on p. 4.
[68] Holmes-Cerfon, M., "Sticky-Sphere Clusters," Annu. Rev. Condens. Matter Phys. **8**, 77 (2017), arXiv:1709.05138 [cond-mat.soft]; cited on p. 5.
[69] Hsu, C. W., B. Zhen, A. D. Stone, J. D. Joannopoulos, and M. Soljačić, "Bound states in the continuum," Nat. Rev. Mater. **1**, 1 (2016); cited on p. 71.
[70] Iinuma, R., Y. Ke, R. Jungmann, T. Schlichthaerle, J. B. Woehrstein, and P. Yin, "Polyhedra Self-Assembled from DNA Tripods and Characterized with 3D DNA-PAINT," Science **344**, 65 (2014); cited on p. 4.
[71] Ishkhanyan, A., "Schrödinger potentials solvable in terms of the general Heun functions," Ann. Phys. **388**, 456 (2018), arXiv:1601.03360 [quant-ph]; cited on p. 82.
[72] Jordan, D. W. and P. W. Smith, *Nonlinear Ordinary Differential Equations: An Introduction for Scientists and Engineers*, 4th ed. (Oxford University Press, Oxford, 2007); cited on p. 76.
[73] Jose, K., "Spatial variability of the dynamic response in periodic and non-homogeneous elastic media," PhD thesis (University of Southampton, 2023); cited on p. 98.
[74] Jose, K., N. Ferguson, and A. Bhaskar, "Branched flows of flexural waves in non-uniform elastic plates," Commun. Phys. **5**, 152 (2022); cited on p. 98.
[75] Jose, K., N. Ferguson, and A. Bhaskar, "Branched flows of flexural elastic waves in non-uniform cylindrical shells," PLOS ONE **18**, edited by Nawaz, R., e0286420 (2023); cited on p. 98.
[76] Jung, W.-H., E. Chen, R. Veneziano, S. Gaitanaros, and Y. Chen, "Stretching DNA origami: effect of nicks and Holliday junctions on the axial stiffness," Nucl. Acids Res. **48**, 12407 (2020); cited on pp. 18, 33.
[77] Kallus, Y. and M. Holmes-Cerfon, "Free energy of singular sticky-sphere clusters," Phys. Rev. E **95**, 022130 (2017), arXiv:1605.08678 [cond-mat.stat-mech]; cited on pp. 5, 20, 48–50.
[78] Kang, J.-H., H. Kim, C. D. Santangelo, and R. C. Hayward, "Enabling Robust Self-Folding Origami by Pre-Biasing Vertex Buckling Direction," Adv. Mater. **31**, 0193006 (2019); cited on p. 51.

[79] Kapko, V., C. Dawson, I. Rivin, and M. M. J. Treacy, "Density of Mechanisms within the Flexibility Window of Zeolites," Phys. Rev. Lett. **107**, 164304 (2011), arXiv:1109.5203 [cond-mat.mtrl-sci]; cited on p. 4.

[80] Kapovich, M. and J. Millson, "On the moduli space of polygons in the Euclidean plane," J. Differential Geom. **42**, 430 (1995); cited on p. 46.

[81] Kastner, M., "Phase transitions and configuration space topology," Rev. Mod. Phys. **80**, 167 (2008), arXiv:cond-mat.stat-mech/0703401; cited on p. 51.

[82] Kästner, J., "Umbrella sampling," WIREs Comput. Mol. Sci. **1**, 932 (2011); cited on p. 51.

[83] Kaufman, A. N., H. Ye, and Y. Hui, "Variational formulation of covariant eikonal theory for vector waves," Phys. Lett. A **120**, 327 (1987); cited on pp. 57, 61, 70.

[84] Ke, Y., T. Meyer, W. M. Shih, and G. Bellot, "Regulation at a distance of biomolecular interactions using a DNA origami nanoactuator," Nat. Commun. **7**, 10935 (2016); cited on p. 4.

[85] Keller, J. B., "Corrected Bohr–Sommerfeld quantum conditions for nonseparable systems," Ann. Phys. **4**, 180 (1958); cited on pp. 60, 67.

[86] Kendall, D. G., Barden, D., Carne, T. K., and Le, H., eds., *Shape & Shape Theory* (John Wiley & Sons, Ltd., Chichester, 1999); cited on pp. 14, 24.

[87] Kendall, D. G., "A Survey of the Statistical Theory of Shape," Stat. Sci. **4**, 87 (1989); cited on pp. 14, 24.

[88] Kernes, J. and A. J. Levine, "Effects of curvature on the propagation of undulatory waves in lower dimensional elastic materials," Phys. Rev. E **103**, 013002 (2021), arXiv:2006.13383 [cond-mat.soft]; cited on pp. 70, 72, 101.

[89] Krylov, V. V., "Overview of localised flexural waves in wedges of power-law profile and comments on their relationship with the acoustic black hole effect," J. Sound Vib. **468**, 115100 (2020); cited on p. 72.

[90] Landau, L. D. and E. M. Lifshitz, *Theory of Elasticity*, 3rd ed., Vol. 7, Course of Theoretical Physics (Butterworth–Heinemann, Oxford, 1986); cited on pp. 21, 70.

[91] Lee, J. Y. and W. Jeon, "Vibration damping using a spiral acoustic black hole," J. Acoust. Soc. Am. **141**, 1437 (2017); cited on p. 72.

[92] Lee, J. M., *Smooth Manifolds*, 2nd ed. (Springer, New York, NY, 2013); cited on pp. 12, 15.

[93] Leimkuhler, B. and S. Reich, *Simulating Hamiltonian Dynamics* (Cambridge University Press, Cambridge, 2005); cited on pp. 15, 25, 26.

[94] Lelièvre, T., M. Rousset, and G. Stoltz, *Free Energy Computations: A Mathematical Perspective* (Imperial College Press, London, 2010); cited on pp. 11, 23, 24.

[95] Lemieux, A. and A. Bose, "Construction de potentiels pour lesquels l'équation de Schrödinger est soluble," Ann. Inst. Henri Poincaré A **10**, 259 (1969); cited on p. 82.

[96] Lengyel, A., "Analogy between equilibrium of structures and compatibility of mechanisms," PhD thesis (Oxford University, 2002); cited on p. 17.

[97] Leonard, J. and J. E. Graebner, *Scratch My Back: A Pictorial History of the Musical Saw and How to Play It* (Kaleidoscope Press, Santa Ana, CA, 1989); cited on p. 69.

[98] Liedl, T., B. Högberg, J. Tytell, D. E. Ingber, and W. M. Shih, "Self-assembly of three-dimensional prestressed tensegrity structures from DNA," Nat. Nanotechnol. **5**, 520 (2010); cited on pp. 4, 5.

[99] Littlejohn, R. G. and W. G. Flynn, "Geometric phases and the Bohr–Sommerfeld quantization of multicomponent wave fields," Phys. Rev. Lett. **66**, 2839 (1991); cited on pp. 61, 63, 65, 70, 83.

[100] Littlejohn, R. G. and W. G. Flynn, "Geometric phases in the asymptotic theory of coupled wave equations," Phys. Rev. A **44**, 5239 (1991); cited on pp. 61, 63, 65, 70, 83.

[101] Littlejohn, R. G. and M. Reinsch, "Internal or shape coordinates in the $n$-body problem," Phys. Rev. A **52**, 2035 (1995); cited on pp. 14, 16, 24.
[102] Littlejohn, R. G. and M. Reinsch, "Gauge fields in the separation of rotations and internal motions in the $n$-body problem," Rev. Mod. Phys. **69**, 213 (1997); cited on p. 11.
[103] Liu, G., Y. Lou, and Z. Li, "Singularities of parallel manipulators: a geometric treatment," IEEE Trans. Robot. Autom. **19**, 579 (2003); cited on pp. 5, 54.
[104] López-Custodio, P., A. Müller, X. Kang, and J. Dai, "Tangential intersection of branches of motion," Mech. Mach. Theory **147**, 103730 (2020); cited on pp. 15, 16, 31.
[105] López-Custodio, P., A. Müller, J. Rico, and J. Dai, "A synthesis method for 1-DOF mechanisms with a cusp in the configuration space," Mech. Mach. Theory **132**, 154 (2019); cited on p. 16.
[106] Lubensky, T. C., C. L. Kane, X. Mao, A. Souslov, and K. Sun, "Phonons and elasticity in critically coordinated lattices," Rep. Prog. Phys. **78**, 073901 (2015), arXiv:1503.01324 [cond-mat.soft]; cited on pp. 15, 19, 25.
[107] Mace, B. R., "Active control of flexural vibrations," J. Sound Vib. **114**, 253 (1987); cited on p. 71.
[108] Mannattil, M. and C. D. Santangelo, *Geometric Localization of Waves on Thin Elastic Structures*, preprint, 2023, arXiv:2306.07213 [cond-mat.soft]; cited on p. 5.
[109] Mannattil, M., J. M. Schwarz, and C. D. Santangelo, "Thermal Fluctuations of Singular Bar-Joint Mechanisms," Phys. Rev. Lett. **128**, 208005 (2022), arXiv:2112.04279 [cond-mat.soft]; cited on p. 2.
[110] Marijanović, F., S. Moroz, and B. Jeevanesan, "Rayleigh waves and cyclotron surface modes of gyroscopic metamaterials," Phys. Rev. B **106**, 024308 (2022), arXiv:2111.11733 [cond-mat.soft]; cited on p. 98.
[111] Marras, A. E., L. Zhou, H.-J. Su, and C. E. Castro, "Programmable motion of DNA origami mechanisms," Proc. Natl. Acad. Sci. U.S.A. **112**, 713 (2015); cited on p. 4.
[112] Marsden, J. E. and T. J. R. Hughes, *Mathematical Foundations of Elasticity* (Dover, Mineola, NY, 1994); cited on p. 1.
[113] Maslov, V. P. and M. V. Fedoriuk, *Semiclassical Approximation in Quantum Mechanics*, trans. by Niederle, J. and Tolar, J. (Kluwer, Dordrecht, 1981); cited on pp. 60, 67.
[114] Matthies, M., N. P. Agarwal, and T. L. Schmidt, "Design and Synthesis of Triangulated DNA Origami Trusses," Nano Lett. **16**, 2108 (2016); cited on p. 2.
[115] Maxwell, J. C., "On the calculation of the equilibrium and stiffness of frames," Philos. Mag. **27**, 294 (1864); cited on p. 17.
[116] McDonald, S. W., "Phase-space representations of wave equations with applications to the eikonal approximation for short-wavelength waves," Phys. Rep. **158**, 337 (1988); cited on pp. 67, 68.
[117] Melo, H. P. M., C. S. Dias, and N. A. M. Araújo, "Optimal number of faces for fast self-folding kirigami," Commun. Phys. **3**, 154 (2020), arXiv:2005.08202 [cond-mat.soft]; cited on p. 5.
[118] Mermoud, O. and M. Steiner, "Visualisation of Configuration Spaces of Polygonal Linkages," J. Geom. Graph. **4**, 147 (2000); cited on p. 46.
[119] Mezey, P. G., *Shape in Chemistry* (Wiley-VCH, New York, NY, 1993); cited on pp. 14, 24.
[120] Mohammed, N. M., S. C. Creagh, and G. Tanner, "Tunnelling around bends—Wave scattering in curved shell structures," Wave Motion **101**, 102697 (2021); cited on pp. 70, 71, 95, 97.
[121] Morley, L. S. D., "Elastic Waves in a Naturally Curved Rod," Q. J. Mech. Appl. Math. **14**, 155 (1961); cited on pp. 70, 83.
[122] Müller, A., "Higher-order analysis of kinematic singularities of lower pair linkages and serial manipulators," J. Mech. Robot **10**, 011008 (2017); cited on pp. 16, 31.

[123] Müller, A. and Zlatanov, D., eds., *Singular Configurations of Mechanisms and Manipulators* (Springer, Cham, 2019); cited on pp. 3, 15, 16, 31, 48.
[124] Nelson, D. and L. Peliti, "Fluctuations in membranes with crystalline and hexatic order," J. Phys. France **48**, 1085 (1987); cited on p. 29.
[125] Nelson, D., T. Piran, and S. Weinberg, *Statistical Mechanics of Membranes and Surfaces*, 2nd ed. (World Scientific, Singapore, 2004); cited on pp. 5, 21.
[126] Nielsen, R. and S. Sorokin, "The WKB approximation for analysis of wave propagation in curved rods of slowly varying diameter," Proc. R. Soc. A **470**, 20130718 (2014); cited on p. 70.
[127] Nieto, M. M., "Exact wave-function normalization constants for the $B_0 \tanh z - U_0 \cosh^{-2} z$ and Pöschl–Teller potentials," Phys. Rev. A **17**, 1273 (1978); cited on p. 82.
[128] Norris, A. N. and D. A. Rebinsky, "Membrane and Flexural Waves on Thin Shells," J. Vib. Acoust. **116**, 457 (1994); cited on pp. 69, 70, 89.
[129] Olver, F. W. J., Lozier, D. W., Boisvert, R. F., and Clark, C. W., eds., *NIST Handbook of Mathematical Functions* (Cambridge University Press, New York, NY, 2010); cited on pp. 37, 53, 82, 88.
[130] Pancharatnam, S., "Generalized theory of interference, and its applications," Proc. Indian Acad. Sci. A **44**, 247 (1956); cited on p. 9.
[131] Pelat, A., F. Gautier, S. C. Conlon, and F. Semperlotti, "The acoustic black hole: A review of theory and applications," J. Sound Vib. **476**, 115316 (2020); cited on p. 72.
[132] Pellegrino, S. and C. Calladine, "Matrix analysis of statically and kinematically indeterminate frameworks," Int. J. Solids Struct. **22**, 409 (1986); cited on p. 15.
[133] Penrose, R., *The Road to Reality* (Jonathan Cape, London, 2004); cited on p. 12.
[134] Percival, I. C., "Semiclassical theory of Bound States," in *Adv. Chem. Phys.* Edited by Prigogine, I. and Rice, S. A. (John Wiley & Sons, Ltd, 1977), pp. 1–61; cited on pp. 67, 95.
[135] Pierce, A. D., "Waves on Fluid-Loaded Inhomogeneous Elastic Shells of Arbitrary Shape," J. Vib. Acoust. **115**, 384 (1993); cited on pp. 70, 89.
[136] Pierce, A. D., "Variational Formulations in Acoustic Radiation and Scattering," in *Physical Acoustics*, Vol. 22, edited by Pierce, A. D. and Thurston, R. N. (Academic Press, 1993), pp. 195–371; cited on p. 85.
[137] Pierce, A. D., "Physical Interpretation of the WKB or Eikonal Approximation for Waves and Vibrations in Inhomogeneous Beams and Plates," J. Acoust. Soc. Am. **48**, 275 (1970); cited on p. 70.
[138] Pöschl, G. and E. Teller, "Bemerkungen zur Quantenmechanik des anharmonischen Oszillators," Z. Phys. **83**, 143 (1933); cited on p. 82.
[139] Postnova, J. and R. V. Craster, "Trapped modes in elastic plates, ocean and quantum waveguides," Wave Motion **45**, 565 (2008); cited on p. 69.
[140] Power, S. C., "Polynomials for crystal frameworks and the rigid unit mode spectrum," Philos. Trans. R. Soc. A **372**, 20120030 (2014), arXiv:1102.2744 [math.CO]; cited on p. 4.
[141] Pressley, A., *Elementary Differential Geometry*, 2nd ed. (Springer, London, 2010); cited on p. 85.
[142] Ramond, P., *Field Theory: A Modern Primer* (Westview Press, Boulder, CO, 1997); cited on p. 25.
[143] Ravendranadhan, M. and M. Sabir, "Fermion-dyon bound states and fermion number fractionisation," J. Phys. G **15**, 741 (1989); cited on p. 7.
[144] Ravendranadhan, M. and M. Sabir, "Ground state charge of fermion soliton system in 1+1 and 3+1 dimensions," Int. J. Mod. Phys. A **08**, 705 (1993); cited on p. 7.
[145] Ravendranadhan, M. and M. Sabir, "Ground state charge of O(3) nonlinear $\sigma$ model solitons in 2+1 dimensions," Int. J. Mod. Phys. A **09**, 3657 (1994); cited on p. 7.

[146] Rebinsky, D. A. and A. N. Norris, "Dispersion of Flexural Waves on Shells," J. Vib. Acoust. **118**, 526 (1996); cited on p. 89.

[147] Rocklin, D. Z., V. Vitelli, and X. Mao, "Folding mechanisms at finite temperature," preprint (2018), arXiv:1802.02704 [cond-mat.soft]; cited on pp. 5, 25.

[148] Rosen, N. and P. M. Morse, "On the Vibrations of Polyatomic Molecules," Phys. Rev. **42**, 210 (1932); cited on p. 82.

[149] Schwarz, A. S., "Instantons and fermions in the field of instanton," Commun. Math. Phys. **64**, 233 (1979); cited on p. 25.

[150] Scott, J. F. M. and J. H. Woodhouse, "Vibration of an elastic strip with varying curvature," Philos. Trans. R. Soc. A **339**, 587 (1992); cited on pp. 6, 9, 69–71, 78.

[151] Shankar, S., P. Bryde, and L. Mahadevan, "Geometric control of topological dynamics in a singing saw," Proc. Natl. Acad. Sci. U.S.A. **119**, e2117241119 (2022), arXiv:arXiv:2108.10875 [cond-mat.soft]; cited on pp. 6, 7, 9, 69–71, 89, 99.

[152] Shenoy, V. B. and D. H. Gracias, "Self-folding thin-film materials: From nanopolyhedra to graphene origami," MRS Bull. **37**, 847 (2012); cited on p. 5.

[153] Shimamoto, D. and C. Vanderwaart, "Spaces of Polygons in the Plane and Morse Theory," Am. Math. Mon. **112**, 289 (2005); cited on pp. 5, 15.

[154] Søndergaard, N. and D. J. Chappell, "Ray and wave scattering in smoothly curved thin shell cylindrical ridges," J. Sound Vib. **377**, 155 (2016), arXiv:1608.04623 [physics.class-ph]; cited on p. 70.

[155] Sondergaard, N. and G. Tanner, "Wave chaos in the elastic disk," Phys. Rev. E **66**, 066211 (2002), arXiv:nlin/0209009; cited on pp. 87, 95.

[156] Souder, P. A., *Geometrical Mechanics*, Lecture Notes (Syracuse University, Syracuse, NY, 2017); cited on p. 1.

[157] Souslov, A., K. Dasbiswas, M. Fruchart, S. Vaikuntanathan, and V. Vitelli, "Topological Waves in Fluids with Odd Viscosity," Phys. Rev. Lett. **122**, 128001 (2019), arXiv:1802.09649 [cond-mat.soft]; cited on p. 98.

[158] Strogatz, S. H., *Nonlinear Dynamics and Chaos* (Westview Press, Cambridge, MA, 1994); cited on p. 76.

[159] Struik, D. J., *Lectures on Classical Differential Geometry*, 2nd ed. (Dover, Mineola, NY, 1988); cited on p. 14.

[160] Stuckenbruck, E. E., *The Singing Blade: The History, Acoustics, and Techniques of the Musical Saw*, Senior Project (Bard College, 2016); cited on pp. 6, 69.

[161] Sudarshan, E. C. G. and N. Mukunda, *Classical Dynamics: A Modern Perspective* (John Wiley & Sons, Ltd., New York, NY, 1974); cited on p. 1.

[162] Tarnai, T., "Kinematic Bifurcation," in *Deployable Structures*, edited by Pellegrino, S. (Springer, Vienna, 2001) Chap. 8, pp. 143–169; cited on p. 43.

[163] Tracy, E. R., A. J. Brizard, A. Richardson, and A. Kaufman, *Ray Tracing and Beyond: Phase Space Methods in Plasma Wave Theory* (Cambridge University Press, Cambridge, 2014); cited on pp. 56, 58, 60, 91, 97.

[164] Tropp, O., A. Tal, and I. Shimshoni, "A fast triangle to triangle intersection test for collision detection," Comp. Anim. Virtual Worlds **17**, 527 (2006); cited on p. 52.

[165] Venaille, A., Y. Onuki, N. Perez, and A. Leclerc, "From ray tracing to waves of topological origin in continuous media," SciPost Phys. **14**, 062 (2023), arXiv:2207.01479 [physics.flu-dyn]; cited on pp. 9, 70, 98.

[166] Wales, D. J., "Exploring Energy Landscapes," Annu. Rev. Phys. Chem **69**, 401 (2018); cited on p. 22.

[167] Walsh, S. J. and R. G. White, "Vibrational power transmission in curved beams," J. Sound Vib. **233**, 455 (2000); cited on pp. 72, 77, 99.

[168] Weigert, S. and R. G. Littlejohn, "Diagonalization of multicomponent wave equations with a Born–Oppenheimer example," Phys. Rev. A **47**, 3506 (1993); cited on pp. 61–63, 65, 83.

[169] Whitney, H., "Tangents to an Analytic Variety," Ann. Math. **81**, 496 (1965); cited on p. 31.

[170] Williams, E. G., B. H. Houston, and J. A. Bucaro, "Experimental investigation of the wave propagation on a point-driven, submerged capped cylinder using $K$-space analysis," J. Acoust. Soc. Am. **87**, 513 (1990); cited on p. 89.

[171] Williams, W. O., *A Primer on the Mechanics of Tensegrity Structures*, Technical Report (Center for Nonlinear Analysis, Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA, 2003); cited on pp. 16, 30.

[172] Woodhouse, F. G., H. Ronellenfitsch, and J. Dunkel, "Autonomous Actuation of Zero Modes in Mechanical Networks Far from Equilibrium," Phys. Rev. Lett. **121**, 178001 (2018), arXiv:1805.07728 [cond-mat.soft]; cited on pp. 5, 17, 20, 51.

[173] Wu, L., A. Müller, and J. S. Dai, "A matrix method to determine infinitesimally mobile linkages with only first-order infinitesimal mobility," Mech. Mach. Theory **148**, 103776 (2020); cited on pp. 31, 48.

[174] Yabana, K. and H. Horiuchi, "Adiabatic Viewpoint for the WKB Treatment of Coupled Channel System," Prog. Theor. Phys. **75**, 592 (1986); cited on pp. 61, 70.

[175] Yan, L., "Entropy favors heterogeneous structures of networks near the rigidity threshold," Nat. Commun. **9**, 1359 (2018), arXiv:1709.03015 [cond-mat.stat-mech]; cited on p. 5.

[176] Yang, R. and P. S. Krishnaprasad, "On the geometry and dynamics of floating four-bar linkages," Dynam. Stabil. Syst. **9**, 19 (1994); cited on p. 33.

[177] Yu, Y.-Y., "Free Vibrations of Thin Cylindrical Shells Having Finite Lengths With Freely Supported and Clamped Edges," J. Appl. Mech. **22**, 547 (1955); cited on pp. 72, 85, 101.

[178] Zangeneh-Nejad, F. and R. Fleury, "Active times for acoustic metamaterials," Rev. Phys. **4**, 100031 (2019); cited on p. 69.

[179] Zhang, L. and X. Mao, "Finite-temperature mechanical instability in disordered lattices," Phys. Rev. E **93**, 022110 (2016), arXiv:1503.05274 [cond-mat.soft]; cited on pp. 5, 17, 20.

[180] Zhao, S., F. Duan, S. Liu, T. Wu, Y. Shang, R. Tian, J. Liu, Z.-G. Wang, Q. Jiang, and B. Ding, "Efficient Intracellular Delivery of RNase A Using DNA Origami Carriers," ACS Appl. Mater. Interfaces **11**, 11112 (2019); cited on pp. 4, 5.

[181] Zhou, L., A. E. Marras, H.-J. Su, and C. E. Castro, "Direct Design of an Energy Landscape with Bistable DNA Origami Mechanisms," Nano Lett. **15**, 1815 (2015); cited on p. 4.

[182] Zlatanov, D., I. A. Bonev, and C. M. Gosselin, "Constraint Singularities as $C$-Space Singularities," in *Advances in Robot Kinematics*, edited by Lenarčič, J. and Thomas, F. (Springer, Dordrecht, 2002), pp. 183–192; cited on p. 5.

## Colophon

This dissertation was produced using LaTeX—a set of macros for the TeX typesetting system—and the *memoir* class. Figures and plots were created using a combination of Inkscape, Mathematica, and Matplotlib. The body text is set in 10 pt *Utopia* and the text in the figures is set in 8 pt (and occasionally 7 pt) *TeX Gyre Heros*, a free clone of *Helvetica*. Chapter and section headings are set in *Inter* and mathematics is set in Fourier-GUT*enberg*.

Much time was saved during the writing process thanks to several free software projects including BibTool, Debian GNU/Linux, Exuberant Ctags, Gimp, Git, ImageMagick, Inkscape, latexmk, Matplotlib, optipng, pdfsizeopt, pdfTeX, pdftk, Python, qpdfview, TeX Live, Vim, and VimTeX. Much time was lost yak shaving some of these tools to perfection.

*The lyf so short, the craft so longe to lerne.*