Conjugate Points Revisited and Neumann-Neumann Problems

Robert S. Manning*

October 16, 2007

Submitted to SIAM Review, Education Section

Abstract

The theory of conjugate points in the calculus of variations is reconsidered with a perspective emphasizing the connection to finite-dimensional optimization. The object of central importance is the spectrum of the second-variation operator, analogous to the eigenvalues of the Hessian matrix in finite dimensions. With a few basic properties of this spectrum, one can gain a new perspective on the classic result that "stability requires the lack of conjugate points". Furthermore, we show how the spectral perspective allows the extension of the conjugate point approach to variants of the classic problems in the literature, such as problems with Neumann-Neumann boundary conditions.

Keywords: Conjugate points, calculus of variations, second variation

MSC Classification: 49K15

Abbreviated Title: Conjugate Points Revisited

1 Introduction

The calculus of variations is concerned with optimization problems in infinite dimensions, in which the domain is a set of functions rather than $\mathbb{R}^n$. The classic problem is to minimize a functional of the form

$$J[y] = \int_a^b F(s, y(s), y'(s))\, ds, \tag{1}$$

where y ranges over a given subset of a function space that includes boundary conditions pertinent to the physical problem. The central result in the calculus of variations is that any local minimizer of J must satisfy the Euler-Lagrange differential equation:

$$\frac{d}{ds}(F_{y'}) - F_y = 0 \tag{2}$$

(we subscript by y or y' to denote partial differentiation). Eq. (2) is the analogue of the familiar condition $\vec{\nabla} J = \mathbf{0}$ from finite-dimensional optimization, in which the goal is to minimize the function $J(\mathbf{x})$ where $\mathbf{x}$ ranges over $\mathbb{R}^n$. Indeed, the left-hand side of Eq. (2) is sometimes called

*Mathematics Department; Haverford College; Haverford, PA 19041

the variational derivative $\delta J/\delta y$ of J, and Eq. (2) can be derived [6] by discretizing the integral in Eq. (1) using intervals of width Δs and then taking $\Delta s \to 0$ (although this is not the usual proof of Eq. (2)).

In the finite-dimensional case, we next proceed to a "second-derivative test". Given a solution $\mathbf{x}_0$ to $\vec{\nabla} J = \mathbf{0}$, we Taylor-expand J about $\mathbf{x}_0$:

$$J(\mathbf{x}_0 + \Delta\mathbf{x}) = J(\mathbf{x}_0) + \frac{1}{2}(\Delta\mathbf{x})^T \mathbf{H_J}(\mathbf{x}_0)\Delta\mathbf{x} + \cdots, \tag{3}$$

where $\mathbf{H_J} \equiv d^2J/d\mathbf{x}^2$ is the Hessian matrix of second derivatives of J, and the first-order term of the Taylor expansion vanishes because $\vec{\nabla} J(\mathbf{x}_0) = \mathbf{0}$. The second-derivative test says that a sufficient condition for $\mathbf{x}_0$ to be a local minimizer of J is that $\mathbf{H_J}$ be positive-definite at $\mathbf{x}_0$, i.e., $(\Delta\mathbf{x})^T \mathbf{H_J}(\mathbf{x}_0)\Delta\mathbf{x} > 0$ for all $\Delta\mathbf{x} \in \mathbb{R}^n$. From a standard result in linear algebra, we can rephrase this: if the eigenvalues of $\mathbf{H}_J(\mathbf{x}_0)$ are all positive, then $\mathbf{x}_0$ is a local minimizer of J.

In the calculus of variations, too, the usual approach is also to consider the second-order term in the expansion of J about a solution y_0 of the Euler-Lagrange equations:

$$\delta^2 J[h] = \frac{1}{2}\int_a^b \left[F_{yy}(h(s))^2 + 2F_{yy'}h(s)h'(s) + F_{y'y'}(h'(s))^2\right] ds, \tag{4}$$

where F_{yy}, $F_{yy'}$ and $F_{y'y'}$ are evaluated at $(s, y_0(s), y_0'(s))$. The expression $\delta^2 J$ is usually called the second variation. Its argument h is the analogue of the variation $\Delta\mathbf{x}$ in the finite-dimensional case, and is usually called an "allowed variation" (see Sec. 2).

However, once the second variation has appeared on stage, the most common presentation of the calculus of variations seems to diverge from what we expect from our finite-dimensional experience. The typical sufficient condition for a local minimum involves two conditions: the (strengthened) Legendre condition ($F_{y'y'} > 0$) and the Jacobi condition (the absence of "conjugate points").

The role of the Legendre condition can be related fairly easily to the finite-dimensional theory. If $F_{y'y'} < 0$ at any point s, then by constructing a "bump" function h concentrated near s, and noting that $\delta^2 J[h]$ is then dominated by the $F_{y'y'}$ term, which is negative, we can conclude that $\delta^2 J$ is not positive definite and $y_0(s)$ is not a local minimizer of J (see, e.g., [3, p. 215]).

Surely, we feel, the idea of conjugate points must similarly be analogous to some piece of the finite-dimensional theory, but in the standard presentation, it is not clear how. Here is a typical definition, paraphrased from [11, §3.6], for the case that the boundary conditions on the original problem are Dirichlet ($y(a) = y_a$, $y(b) = y_b$): A conjugate point is a value σ between a and b so that the following equation has a solution h (other than $h \equiv 0$):

$$\begin{aligned} \frac{d}{ds}(F_{y'y'}h') &= (F_{yy} - \frac{d}{ds}F_{yy'})h, \quad a \le s \le \sigma \\ h(a) &= 0, \quad h(\sigma) = 0. \end{aligned} \tag{5}$$

It is not apparent how this definition relates to the eigenvalues that play a central role in the finite-dimensional theory. Our goal is present a formulation of conjugate points that makes

the connection to eigenvalues more apparent. Furthermore, we will show how this formulation indicates the appropriate definition of conjugate point when the problem is changed from the classic problem just presented.

Those interested in further exploring this connection between the calculus of variations and finite-dimensional optimization may be interested in the discussion in [6, §30], where $\delta^2 J$ is discretized using intervals of width Δs and a condition for its positive-definiteness is shown to converge to the Jacobi condition when $\Delta s \to 0$. In contrast, we will explore the analogy to the finite-dimensional problem while remaining in the infinite-dimensional setting (i.e., we will not discretize $\delta^2 J$), and we will also seek to extend the analogy from the special case of all positive eigenvalues to a more general connection between the spectrum of the Hessian matrix and the spectrum of an appropriate second-variation operator.

2 The operator view of the second variation

First, let us fill in some fundamental details of the theory sketched in Sec. 1. We seek a function y that will be a local minimum of the functional in Eq. (1). For technical reasons, we assume that the integrand F is C^3 in its arguments.

In order to define a local minimum, we must choose an underlying function space and associated norm. We will choose the function space $C^1([a,b])$ (the set of functions with continuous derivatives) with norm

$$\|y\| \equiv \sup_{a \le s \le b} |y(s)| + \sup_{a \le s \le b} |y'(s)|.$$

Many problems of interest involve boundary conditions imposed on the function y, so that the domain of J is a subset of $C^1([a,b])$. We will begin by assuming Dirichlet boundary conditions $y(a) = y_a$ and $y(b) = y_b$; in Sec. 6, we adapt the theory to Neumann boundary conditions.

In order to determine whether a particular function y is a local minimum, we investigate the behavior of J in a small neighborhood of y, by considering *allowed variations*. These are simply functions h for which $y + h$ obeys the same conditions required of y itself.[†] Thus, in terms of smoothness, we require $h \in C^1([a,b])$, and in order to respect the Dirichlet boundary conditions on y, we require $h(a) = h(b) = 0$.

To investigate a small neighborhood around y, we introduce a small parameter $\epsilon > 0$, and consider $J[y + \epsilon h]$, i.e., we replace y by $y + \epsilon h$ and y' by $y' + \epsilon h'$ in Eq. (1). We then perform a Taylor expansion, and collect terms of like order in ϵ, calling the first-order term $\delta J[h]$ and the second-order term $\delta^2 J[h]$. By straightforward analogy with the finite-dimensional problem, we would think that y will be a local minimum if $\delta J = 0$ and $\delta^2 J$ is positive definite. There is a little infinite-dimensional subtlety here, but this is basically true (we actually need a somewhat stronger condition than positive definiteness, namely $\delta^2 J[h] \ge k\|h\|^2$ for some $k > 0$ [6, §24], but as we will shall see, for many problems the spectrum of $\delta^2 J$ is discrete and the stronger condition follows from positive definiteness).

[†]Really, we are concerned with $y + \epsilon h$ for infinitesimal $\epsilon > 0$, but the two notions are equivalent in the current setting. In other settings, e.g., if we had an isoperimetric constraint, this distinction would be crucial.

The first-order term in the Taylor expansion takes the form

$$\delta J[h] = \int_a^b (F_y h + F_{y'} h')\, ds.$$

Integrating $F_{y'}h'$ by parts, we find that

$$\delta J[h] = \int_a^b h\left(F_y - \frac{d}{ds}(F_{y'})\right) ds,$$

and since this must be zero for all allowed h, we arrive at the Euler-Lagrange equation in Eq. (2), as can be seen in any calculus of variations text, e.g., [5, 6, 11].

Similarly, the second-order term in the Taylor expansion gives the form for $\delta^2 J$ found in Eq. (4). Here is where our story diverges from the traditional approach. We claim that the form of $\delta^2 J$ in Eq. (4) is not ideal for making the analogy to the finite-dimensional case. Instead, we rewrite $\delta^2 J$ as:

$$\delta^2 J[h] = \frac{1}{2}\int_a^b \left(F_{yy}h^2 + F_{yy'}hh' + F_{yy'}hh' + F_{y'y'}(h')^2\right) ds,$$

and integrate $F_{yy'}hh' + F_{y'y'}(h')^2 = (F_{yy'}h + F_{y'y'}h')h'$ by parts to find:

$$\begin{aligned}\delta^2 J[h] &= \frac{1}{2}\int_a^b \left(F_{yy}h^2 + F_{yy'}hh' - h\frac{d}{ds}(F_{yy'}h + F_{y'y'}h')\right) ds \\ &= \frac{1}{2}\int_a^b h\left(F_{yy}h + F_{yy'}h' - \frac{d}{ds}(F_{yy'}h + F_{y'y'}h')\right) ds = \frac{1}{2}\langle h, \mathcal{S}h\rangle,\end{aligned}$$

where $\langle\ ,\ \rangle$ is the L^2 inner product $\langle f, g\rangle = \int_a^b f(s)g(s)ds$ and $\mathcal{S}$ is the *second-variation operator*:

$$\mathcal{S}h \equiv F_{yy}h + F_{yy'}h' - \frac{d}{ds}(F_{yy'}h + F_{y'y'}h'). \tag{6}$$

This operator $\mathcal{S}$ is the natural analogue of the Hessian matrix $\mathbf{H_J}$: in the Taylor expansion of $J[y + \epsilon h]$, the second-order term is $\frac{1}{2}\langle h, \mathcal{S}h\rangle$, just as in the finite-dimensional case, the Taylor expansion in Eq. (3) has second-order term $\frac{1}{2}(\Delta\mathbf{x})^T\mathbf{H_J}\Delta\mathbf{x}$, which can be rewritten as $\frac{1}{2}\langle\Delta\mathbf{x}, \mathbf{H_J}\Delta\mathbf{x}\rangle$, using the standard inner product in $\mathbb{R}^n$, $\langle\mathbf{u}, \mathbf{v}\rangle = \mathbf{u}^T\mathbf{v}$.

Here we should inject a technical comment. In our initial setup, we chose an underlying function space that required one continuous derivative, and yet our second-variation operator $\mathcal{S}$ involves second derivatives. Thus, in our subsequent analysis, we will be considering variations h in the space

$$\mathcal{D}([a,b]) \equiv \{h \in C^2([a,b]) : h(a) = h(b) = 0\}.$$

Since this space is dense in the space of all allowed variations (which only require one derivative), it will be sufficient to show that $\delta^2 J$ is positive definite on $\mathcal{D}([a,b])$ in order to show that y is a local minimum.

Having defined $\mathcal{S}$, we reach our first payoff: a more natural definition of conjugate point in terms of the eigenvalues of $\mathcal{S}$. As usual, we say ρ is an eigenvalue of $\mathcal{S}$ on $\mathcal{D}$ if there exists a nonzero $h \in \mathcal{D}$ such that $\mathcal{S}h = \rho h$.

Definition 1. *A conjugate point is a number σ between a and b so that $\mathcal{S}$ has a zero eigenvalue on $\mathcal{D}([a,\sigma])$.*

More explicitly, σ is a conjugate point if we can find a solution to

$$\mathcal{S}h = 0, \quad h(a) = 0, \quad h(\sigma) = 0,$$

other than the obvious solution that h is the zero function. If we write out $\mathcal{S}h$ explicitly, we recover Eq. (5). Thus, we have developed a definition of conjugate point that is consistent with the classic theory, but whose significance is more apparent, and which will be more easily transferred to other problems.

Finally, let us draw another satisfying parallel between the Hessian matrix and the second-variation operator. We know that the Hessian matrix $\mathbf{H_J}$ is symmetric, i.e., $(\mathbf{H_J})^T = \mathbf{H_J}$. We can show that $\mathcal{S}$ is also symmetric, although what "symmetry" should mean in infinite dimensions is at first a bit mysterious, since the notion of the transpose of an operator is not obvious.

Take any pair of n-vectors $\mathbf{u}$ and $\mathbf{v}$ and an n-by-n square matrix $\mathbf{M}$ and consider $\langle \mathbf{u}, \mathbf{Mv} \rangle = \mathbf{u}^T\mathbf{Mv}$. This is a scalar, and hence it equals its own transpose:

$$\left(\mathbf{u}^T\mathbf{Mv}\right)^T = \mathbf{u}^T\mathbf{Mv}.$$

By the properties of transposes, this means $\mathbf{v}^T\mathbf{M}^T\mathbf{u} = \mathbf{u}^T\mathbf{Mv}$, or, in the language of inner products, $\langle \mathbf{v}, \mathbf{M}^T\mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{Mv} \rangle$.

Now, if $\mathbf{M}$ is symmetric, then $\mathbf{M}^T = \mathbf{M}$, and the derivation above shows that $\langle \mathbf{v}, \mathbf{Mu} \rangle = \langle \mathbf{u}, \mathbf{Mv} \rangle$ for all $\mathbf{u}$, $\mathbf{v}$.

On the other hand, if we had some matrix $\mathbf{M}$ so that $\langle \mathbf{v}, \mathbf{Mu} \rangle = \langle \mathbf{u}, \mathbf{Mv} \rangle$ for all $\mathbf{u}$, $\mathbf{v}$, then we could reverse the steps above to show that $\mathbf{u}^T\mathbf{M}^T\mathbf{v} = \mathbf{u}^T\mathbf{Mv}$ for all $\mathbf{u}$, $\mathbf{v}$, and this then would imply that $\mathbf{u}^T(\mathbf{M}^T - \mathbf{M})\mathbf{v} = 0$ for all $\mathbf{u}$, $\mathbf{v}$. From this, we can easily prove that $\mathbf{M}^T = \mathbf{M}$, by choosing $\mathbf{u} = \mathbf{e_i}$ and $\mathbf{v} = \mathbf{e_j}$ for arbitrary i, j, and thereby finding that $[\mathbf{M}^T - \mathbf{M}]_{ij} = 0$ for all i, j, thereby proving that $\mathbf{M}^T - \mathbf{M} = \mathbf{0}$. Thus, we have proven that $\mathbf{M}$ is a symmetric matrix if and only if $\langle \mathbf{v}, \mathbf{Mu} \rangle = \langle \mathbf{u}, \mathbf{Mv} \rangle$ for all $\mathbf{u}$, $\mathbf{v}$. This is the inspiration for the definition of symmetry of an operator:

Definition 2. *We say that an operator $\mathcal{S} : X \to X$ is symmetric on some inner product space X (with inner product $\langle\ ,\ \rangle$) if $\langle g, \mathcal{S}f \rangle = \langle f, \mathcal{S}g \rangle$ for all $f, g \in X$.*

We can verify that our second-variation operator $\mathcal{S}$ is symmetric by a few integrations by parts. Obtaining this symmetry was the reason for the artificial-looking split of the term $2F_{yy'}hh'$ into two equal pieces in our derivation of $\mathcal{S}$. Why do we care? As we will see in the next section, this symmetry gives a particularly nice structure to the spectrum of $\mathcal{S}$, since, just as in matrix theory, symmetric operators have special spectral properties. In the end, this spectrum is what determines whether our solution y is a local minimum.

3 The spectrum of $\mathcal{S}$

How does our definition of conjugate point help us determine whether a given solution y_0 to the Euler-Lagrange equation (plus boundary conditions) is a local minimum of J? As we have seen, we would like to determine whether $\delta^2 J$ is positive definite, i.e., whether $\langle h, \mathcal{S}h\rangle > 0$ for all $h \in \mathcal{D}([a,b]), h \not\equiv 0$.

In the finite-dimensional case, we know that $\langle \mathbf{\Delta x}, \mathbf{H_J \Delta x}\rangle > 0$ for all $\mathbf{\Delta x} \neq \mathbf{0}$ exactly when the eigenvalues of $\mathbf{H_J}$ are all positive. In the end, this will be our criterion as well: y_0 will be a local minimum of J if all the eigenvalues of $\mathcal{S}$ are positive. However, in infinite dimensions, there is a bit to worry about here.

To start, it's not even clear that $\mathcal{S}$ will have eigenvalues. In finite dimensions, we know that the matrix $\mathbf{H_J}$ is a linear operator from the vector space $\mathbb{R}^n$ to itself. The operator $\mathcal{S}$ is also a linear operator, but its domain and range do not match: we set its domain as $\mathcal{D}([a,b])$ in order to have the second derivatives in $\mathcal{S}$ make sense, but there is certainly no reason for a function in the range of $\mathcal{S}$ to have two continuous derivatives, nor for it to vanish at $s = a, b$. The natural range space for $\mathcal{S}$ is $L^2([a,b])$, the space of square-integrable functions on $[a,b]$.

Given this mismatch, is the eigenvalue problem even sensible? Fortunately, it is. The saving grace is that $\mathcal{S}$ is self-adjoint (a technical extension of the symmetry of $\mathcal{S}$ that we have already mentioned) with a domain $\mathcal{D}([a,b])$ that is *dense* in the Hilbert space $L^2([a,b])$. We will take advantage of a powerful spectral theory for operators of this type.

One last technical comment: this spectral theory involves the L^2-norm

$$\|y\|_2 \equiv \left(\int_a^b y^2 ds\right)^{1/2}.$$

This may be a little jarring, since we chose the C^1 norm $\|\cdot\|$ to define what constitutes a local minimum of J. There is not, in fact, a conflict here: the choice of the C^1 norm for defining the optimization problem led naturally to our need to understand the definiteness (or lack thereof) of $\mathcal{S}$, and now we use the L^2 norm merely as a means to answer that question, since we have at our disposal a powerful theory of operators defined on a dense subset of $L^2([a,b])$.

In general, the spectrum of an infinite-dimensional operator can contain more than just eigenvalues, and general results about positive definiteness are subtle. However, if $F_{y'y'} > 0$ (here is the role of the Legendre strengthened condition), then $\mathcal{S}$ is a Sturm-Liouville operator [10], and its spectrum consists of only a discrete increasing set of eigenvalues: $\rho_1 \leq \rho_2 \leq \rho_3 \leq \cdots$. Furthermore, the corresponding eigenfunctions of a Sturm-Liouville operator form a basis for the function space $\mathcal{D}([a,b])$, and this leads immediately to the fact that $\delta^2 J$ is positive definite if and only if all of its eigenvalues are positive (see the variational expression for eigenvalues in Sec. 5 for one justification of this fact). The existence of a discrete spectrum allows us to make the following definition (introduced by Morse [9]):

Definition 3. *The index of a solution y_0 to the Euler-Lagrange equation plus boundary conditions is the number of negative eigenvalues of $\mathcal{S}$ on $\mathcal{D}([a,b])$.*

Solutions with index zero are local minima. Higher-index solutions can be visualized by analogy with the finite-dimensional case, e.g., solutions with index one are "saddle points" where the graph of J has one downward-pointing direction.

So, our goal is simply to determine the index, the number of negative eigenvalues of $\mathcal{S}$ on $\mathcal{D}([a,b])$. How do conjugate points—values of σ for which zero is an eigenvalue of $\mathcal{S}$ on $\mathcal{D}([a,\sigma])$—relate to this question? The answer is clear once, for a given problem, we establish two additional facts:

Property 1. *For σ sufficiently close to a, the eigenvalues of $\mathcal{S}$ on $\mathcal{D}([a,\sigma])$ are all positive.*

It will be useful to give a name to this asymptotic notion of the eigenvalues of $\mathcal{S}$ on $\mathcal{D}([a,\sigma])$ for $\sigma \approx a$: we will call them the *inborn eigenvalues of $\mathcal{S}$*.

Property 2. *The eigenvalues of $\mathcal{S}$ on $\mathcal{D}([a,\sigma])$ are decreasing functions of σ.*

We emphasize that these two properties must be established for each particular problem in question. We will see in Ch. 5 that they do hold for the standard functional (1) with Dirichlet boundary conditions if the Legendre condition $L_{y'y'} > 0$ is true. However, for more general functionals, or different boundary conditions, the properties must be established anew (and, as we will see in Ch. 6, they may fail, necessitating a revised conjugate point theory).

Given Properties 1 and 2, we can show:

Proposition 1. *(Generalization of Jacobi's condition) The index of y_0 equals the number of conjugate points σ in (a,b).*

The reasoning behind Proposition 1 is as follows: if the inborn eigenvalues are all positive and then decrease with σ, and if conjugate points record when eigenvalues hit zero, then logically the number of negative eigenvalues when $\sigma = b$ must equal the number of conjugate points $\sigma \in (a,b)$. In particular, we can only have all positive eigenvalues at $\sigma = b$ if there are no conjugate points (the classic Jacobi condition).

To be sure, a full proof of Proposition 1 requires more careful analysis, e.g., on the question of the continuous dependence of the eigenvalues on σ, and worrying about the possibility of multiple zero eigenvalues, but this common-sense argument is a satisfying introduction to the idea of conjugate points.

4 A Fundamental Example

Before consider why Properties 1 and 2 are true for the Dirichlet problem, let us consider an example. Consider an elastic rod confined to a plane, with one end held at the origin, and with a mass m placed at the other end, as shown in Fig. 1. Assume the rod is inextensible (it can neither stretch nor compress), so it has a fixed arclength, and assume for simplicity that the total arclength is one. We impose "clamped" boundary conditions that the tangent vector to the rod must be vertical at each end. Parametrize the rod by arclength s, and let $\theta(s)$ be the (clockwise) angle from the positive y-axis to the rod's tangent vector at arclength s. The position $(x(s), y(s))$ of the rod can be determined from θ because inextensibility implies that $(x'(s), y'(s)) = (\sin\theta(s), \cos\theta(s))$, and so $(x(s), y(s)) = \int_0^s (\sin\theta(\tau), \cos\theta(\tau))d\tau$. Thus, θ describes a rod configuration completely.

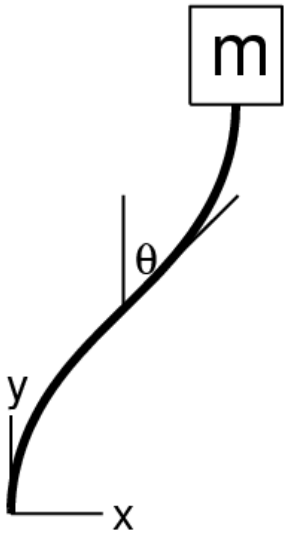

Figure 1: Elastic rod loaded at one end with a mass m and with clamped boundary conditions requiring a vertical tangent line at each end. At each point on the rod, θ is the (clockwise) angle from the positive y-axis to the rod's tangent vector.

We use a simple model for the total energy of this system:

$$E[\theta] = \int_0^1 \left(\frac{K(\theta')^2}{2} + mg\cos\theta \right) ds.$$

The first term models the bending energy of the rod: at each s, $\theta'(s)$ is the local rate of bending in the rod, and we assume a quadratic energy in this bending rate, weighted by a stiffness parameter K. The second term is the gravitational potential energy of the mass m, since mg may be pulled outside of the integral and $\int_0^1 \cos\theta ds = y(1)$. If we divide E by K (i.e., express energy in units of K), and define $\lambda \equiv mg/K$, we find:

$$J[\theta] = \int_0^1 \left(\frac{(\theta')^2}{2} + \lambda\cos\theta \right) ds. \tag{7}$$

Our goal is to study the equilibria of this functional as the parameter λ is varied (or, equivalently, as the mass m is varied).

If we assume that the amount of bending is small, and therefore replace $\cos\theta$ by the first two terms of its Taylor expansion, we arrive at an alternative small-angle energy functional:

$$J_s[\theta] = \int_0^1 \left(\frac{(\theta')^2}{2} + \lambda - \frac{\lambda\theta^2}{2} \right) ds, \tag{8}$$

which is useful for illustration since we can compute the eigenvalues of the associated operator $\mathcal{S}$ in closed form.

The Euler-Lagrange equation for J_s is $\theta'' + \lambda\theta = 0$ which must be solved subject to the imposed clamped boundary conditions $\theta(0) = \theta(1) = 0$. The general solution of the Euler-Lagrange equation is $\theta(s) = A\cos(s\sqrt{\lambda}) + B\sin(s\sqrt{\lambda})$. The condition $\theta(0) = 0$ implies that $A = 0$, and then the condition $\theta(1) = 0$ implies $B\sin(\sqrt{\lambda}) = 0$, which gives either $B = 0$ or $\lambda = j^2\pi^2$, for $j = 1, 2, \cdots$. These solutions are depicted in the "bifurcation diagram" in Fig. 2 in which we plot λ and $\max_{s\in[0,1]} |\theta(s)|$ for each solution. The horizontal line represents the

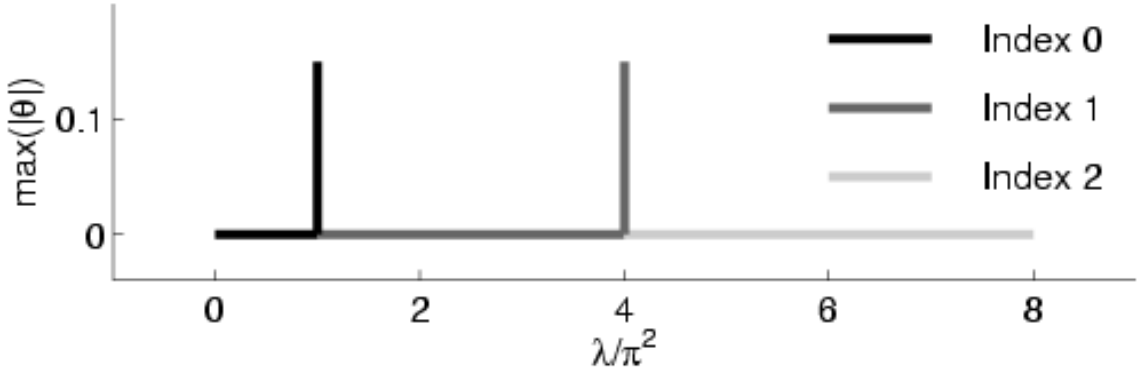

Figure 2: Bifurcation diagram for rod buckling using small-angle energy from Eq. (8) and clamped boundary conditions $\theta(0) = \theta(1) = 0$. For each solution, we plot $\max_{s\in[0,1]} |\theta(s)|$ against λ. The index of each solution is indicated by grayscale.

solutions $\theta(s) \equiv 0$ ($A = B = 0$) that exist for all λ, while the vertical lines at $\lambda = j^2\pi^2$ represent the solutions $B\sin(j\pi s)$ for $j = 1, 2, \cdots$ and any value of B. (Note that for these solutions $\max_{s\in[0,1]} |\theta(s)| = B$, and we have chosen to only plot small values of B consistent with our model assumption that we were considering only small amounts of buckling when using J_s).

The index of each solution in Fig. 2 is represented by grayscale, and comes from the following analysis. The second variation operator is:

$$\mathcal{S}h = -\frac{d}{ds}(h') - \lambda h = -h'' - \lambda h.$$

The eigenvalue equation $\mathcal{S}h = \rho h$ simplifies to $h'' = -(\lambda + \rho)h$ or $h'' + (\lambda + \rho)h = 0$, whose solutions are familiar from standard 2nd-order ODE theory:

$$h(s) = \begin{cases} A\exp\left(s\sqrt{-(\rho+\lambda)}\right) + B\exp\left(-s\sqrt{-(\rho+\lambda)}\right), & \text{if } \rho+\lambda < 0, \\ As + B, & \text{if } \rho+\lambda = 0, \\ A\cos\left(s\sqrt{\rho+\lambda}\right) + B\sin\left(s\sqrt{\rho+\lambda}\right), & \text{if } \rho+\lambda > 0. \end{cases} \tag{9}$$

In order for ρ to be an eigenvalue of $\mathcal{S}$ on $\mathcal{D}([0,\sigma])$, these solutions must satisfy $h(0) = h(\sigma) = 0$. It can be easily shown that neither the $\rho + \lambda < 0$ nor the $\rho + \lambda = 0$ solution can satisfy these two conditions (except in the trivial case $A = B = 0$). In the remaining case, the condition $h(0) = 0$ implies that $A = 0$, and then the condition $h(\sigma) = 0$ implies that $\sin(\sigma\sqrt{\rho+\lambda}) = 0$,

which implies that

$$\rho_n = -\lambda + \frac{n^2\pi^2}{\sigma^2}, \quad n = 1, 2, \cdots .$$

The validity of both Properties 1 and 2 are now clear for this example, since, for all λ, all eigenvalues $\rho > 0$ for σ sufficiently close to $a = 0$, and ρ is clearly decreasing in σ. We note further that, even though $\lambda \geq 0$ for our particular example, these two conclusions would still hold even if $\lambda < 0$.

Furthermore, setting $\sigma = 1$, these formulas allow us to find the index of a solution for each λ. If $\lambda \leq \pi^2$, then $\rho_n \geq 0$ for all n, so that the index is zero. If $\pi^2 < \lambda \leq 4\pi^2$, then $\rho_1 < 0$ and $\rho_n \geq 0$ for $n > 1$, so that the index is one. By similar arguments, if $j^2\pi^2 < \lambda \leq (j+1)^2\pi^2$ for $j \in \mathbb{N}$, then the index is j.

In terms of the buckling of the rod, we imagine we start with a small mass $m \approx 0$, and so the unbuckled solution $\theta \equiv 0$ has index zero, and hence is stable (it is a local minimizer of the energy). When the mass reaches a critical value $m = K\pi^2/g$, then the unbuckled solution becomes unstable (it has index one), and so the rod will instead switch to the first buckled branch of solutions, the vertical branch at $\lambda = \pi^2$.

We note that this problem is, in a sense, degenerate, in that $\mathcal{S}$ has a zero eigenvalue for all the buckled solutions, so the $\lambda = \pi^2$ solutions are only "marginally stable". In addition, it seems unphysical for all solutions $B\sin(\pi s)$ to be stable for $\lambda = \pi^2$. This awkwardness comes from our simplification of $\cos\theta$ to $1 - \theta^2/2$ in order to get closed-form solutions. If we use the full energy functional in Eq. (7), we get the diagram in Fig. 3, which is much like the diagram in Fig. 2 except the vertical branches now bend slightly and $\mathcal{S}$ only has zero eigenvalues at the "bifurcation points" where the curved branches meet the horizontal branch. (Fig. 3 includes

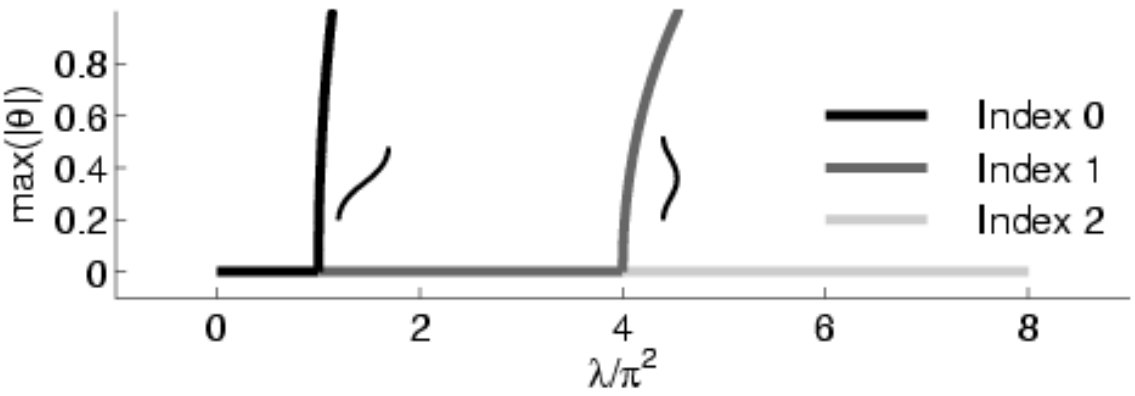

Figure 3: Bifurcation diagram for rod buckling using energy from Eq. (7) and clamped boundary conditions $\theta(0) = \theta(1) = 0$. Rod shapes are drawn next to the two branches of buckled configurations.

sample rod configurations on the first two branches of buckled solutions; this was not possible in Fig. 2 because replacing $\cos\theta$ by $1 - \theta^2/2$ implied the loss of the ability to determine x, y from θ).

Making Fig. 3 involved the computation of a numerical solution of the Euler-Lagrange equation for J (plus boundary conditions). This could be done by a variety of methods, including shooting or finite differences. Here, we used a parameter continuation package AUTO97 [4] (`http://indy.cs.concordia.ca/auto`) that is specialized for solving boundary value problems such as this one, including the ability to detect bifurcation points and follow bifurcating branches.

5 Analysis of the eigenvalues of $\mathcal{S}$

Why are Properties 1 and 2 true for the general Dirichlet problem? The reasoning for Property 1 is fairly straightforward. We adopt the shorthand

$$A(s) \equiv F_{y'y'}(s, y_0(s), y_0'(s)), \quad B(s) \equiv F_{yy'}(s, y_0(s), y_0'(s)), \quad C(s) \equiv F_{yy}(s, y_0(s), y_0'(s)).$$

Since we are considering σ asymptotically close to a, we expand A, B, C in Taylor series about a and consider just the zeroth order term, e.g., $A(s) \approx A(a)$. Then, the zeroth order terms in the eigenvalue equation $\mathcal{S}h = \rho h$ become:

$$C(a)h + B(a)h' - \frac{d}{ds}(B(a)h + A(a)h') = \rho h,$$

which simplifies to:

$$C(a)h - A(a)h'' = \rho h.$$

If we assume the Legendre strengthened condition holds, then $A(a) > 0$, and we divide both sides by it, and rearrange to find:

$$h'' + \left(-\frac{C(a)}{A(a)} + \frac{\rho}{A(a)}\right) h = 0.$$

This is the differential equation from our analysis of the second variation in Sec. 4, with $-C(a)/A(a)$ playing the role of λ, and $\rho/A(a)$ playing the role of ρ. In Sec. 4, we showed that, for any λ, the eigenvalues ρ were all positive for $\sigma \approx a$. Thus, in this case we have $\rho/A(a) > 0$ for all eigenvalues ρ, or, since $A(a) > 0$, $\rho > 0$ for all eigenvalues ρ.

The proof of Property 2 relies on some more machinery: a variational expression for the eigenvalues. From linear algebra, we know that the smallest eigenvalue ρ_1 of a symmetric matrix $\mathbf{M}$ is the minimal value of $\langle \mathbf{v}, \mathbf{M}\mathbf{v} \rangle$ over all unit vectors $\mathbf{v} \in \mathbb{R}^n$. Similarly, the second smallest eigenvalue ρ_2 is the minimal value of $\langle \mathbf{v}, \mathbf{M}\mathbf{v} \rangle$ over all unit vectors $\mathbf{v}$ orthogonal to the first eigenvector, etc. (see, e.g., [7], p. 468 for the "max" version of this property). The same characterization is available in infinite dimensions, at least when the spectrum consists of an increasing sequence of eigenvalues as we have here [12].

Specifically, let $\rho_i(\sigma)$ denote the ith smallest eigenvalue of $\mathcal{S}$ on $\mathcal{D}([a, \sigma])$, and $h_i(s; \sigma)$ the corresponding eigenfunction normalized to be unit-length ($\|h_i\|_2 = 1$). We define a sequence of spaces:

$$\begin{aligned}
\mathcal{B}_1(\sigma) &= \{h \in \mathcal{D}([a, \sigma]) : \|h\|_2 = 1\} \\
\mathcal{B}_2(\sigma) &= \{h \in \mathcal{D}([a, \sigma]) : \|h\|_2 = 1,\ \langle h, h_1 \rangle = 0\} \\
\mathcal{B}_3(\sigma) &= \{h \in \mathcal{D}([a, \sigma]) : \|h\|_2 = 1,\ \langle h, h_1 \rangle = \langle h, h_2 \rangle = 0\},
\end{aligned}$$

and so on, i.e., $\mathcal{B}_m(\sigma)$ is the set of unit-length allowed variations orthogonal to the first $m-1$ eigenfunctions. (In these definitions, the L^2-norms and inner products are all on the interval $[a,\sigma]$.) Then the eigenvalues have the following variational expression [12]:

Proposition 2.

$$\rho_m(\sigma) = \min_{h\in\mathcal{B}_m(\sigma)} \langle h, \mathcal{S}h\rangle.$$

The basic idea for the proof of Property 2 is then as follows. Consider $\sigma_1 < \sigma_2$ and define

$$j_i(s) = \begin{cases} h_i(s;\sigma_1), & \text{if } a \le s \le \sigma_1 \\ 0, & \text{if } \sigma_1 < s \le \sigma_2 \end{cases}, \quad i = 1, 2, \cdots, m, \tag{10}$$

i.e., j_i is constructed by appending a zero function to the eigenfunction h_i on $[a,\sigma_1]$. Now choose a_k not all zero so that $\sum_{k=1}^m a_k j_k$ is orthogonal to $h_1(s;\sigma_2)$, $\cdots$, $h_{m-1}(s;\sigma_2)$ on $[a,\sigma_2]$ (since we have $m-1$ orthogonality conditions and m coefficients at our disposal, this can always be done, by solving an $(m-1)$-by-m linear system). Rescale the set of a_k so that $\sum_{k=1}^m (a_k)^2 = 1$, which implies that $\sum_{k=1}^m a_k j_k \in \mathcal{B}_m(\sigma_2)$. Therefore, by Proposition 2,

$$\rho_m(\sigma_2) \le \int_a^{\sigma_2} \left[\sum_{i=1}^m a_i j_i(s)\right] \mathcal{S} \left[\sum_{k=1}^m a_k j_k(s)\right] ds.$$

The integrand vanishes for $\sigma_1 < s \le \sigma_2$, so we may change the upper integration limit to σ_1 and replace j_i by its restriction to (a,σ_1):

$$\rho_m(\sigma_2) \le \int_a^{\sigma_1} \left[\sum_{i=1}^m a_i h_i(s;\sigma_1)\right] \mathcal{S} \left[\sum_{k=1}^m a_k h_k(s;\sigma_1)\right] ds.$$

Now we use the fact that h_i are orthonormal eigenfunctions of $\mathcal{S}$ to find:

$$\rho_m(\sigma_2) \le \sum_{i,k=1}^m a_i a_k \rho_k(\sigma_1) \int_a^{\sigma_1} h_i h_k ds = \sum_{i=1}^m (a_i)^2 \rho_i(\sigma_1) \le \rho_m(\sigma_1).$$

The actual proof is somewhat more complicated because the functions j_k may not be in the space $\mathcal{D}([a,b])$ since they may not be C^2. Instead, we construct a sequence of functions $j_{kn} \in \mathcal{D}([a,b])$ that converge to j_k as $n \to \infty$ in a sufficiently nice fashion that we can use the above sequence of inequalities to prove

$$\rho_m(\sigma_2) \le \frac{\rho_m(\sigma_1) + O(1/n)}{1 + O(1/n^3)},$$

and then take $n \to \infty$. Details can be found in the Appendix.

6 The Neumann-Neumann problem

We now make a small change in our calculus of variations problem: changing to Neumann boundary conditions $y'(a) = \alpha$, $y'(b) = \beta$. It seems like a simple change that would have minimal effect on the theory: the Euler-Lagrange equation doesn't involve the boundary conditions at all, and there is a natural analogue to the conjugate point equation from Eq. (5):

$$\begin{aligned} \frac{d}{ds}(F_{y'y'}h') &= (F_{yy} - \frac{d}{ds}F_{yy'})h, \quad a \leq s \leq \sigma \\ h'(a) &= 0, \quad h'(\sigma) = 0. \end{aligned} \tag{11}$$

However, a statement of the Jacobi condition for the Neumann-Neumann problem is notably absent from all standard calculus of variations textbooks. Why is that? Because it isn't true! As we shall see, our alternate formulation of conjugate points clarifies why this natural extension of the theory to the Neumann case does not work, and also shows how to repair the theory.

For the purposes of concreteness, we will continue to work with our rod-buckling example, replacing the clamped boundary conditions $\theta(0) = \theta(1) = 0$ by "pinned" boundary conditions $\theta'(0) = \theta'(1) = 0$. Thus the rod is not forced to be vertical at each end, but rather is free to rotate, as if the end were attached to a frictionless pin. Said another way, the pinned boundary conditions mean that we apply no "moment" or torque at the ends of the rod (although we can apply forces) [1, p. 93]. In addition, to allow for a richer variety of examples, we will allow $\lambda < 0$, which represents pulling up on the rod rather than pushing down on it.

The allowed variations now live in the space

$$\mathcal{D}'([0,1]) \equiv \{h \in C^2([a,b]) : h'(0) = 0, h'(1) = 0\},$$

in order that $\theta + \epsilon h$ obey the pinned boundary conditions. The extraction of the first-order term from the Taylor expansion of $J[\theta + \epsilon h]$ yields the same Euler-Lagrange equation as before, and the second-order term yields the same second-variation operator $\mathcal{S}$. By our theory, the appropriate definition of conjugate point is a value of σ for which $\mathcal{S}$ has a zero eigenvalue on $\mathcal{D}'([0,\sigma])$, and it is readily verified that this yields the conjugate point equation given by Eq. (11).

So why is Jacobi's condition, or Proposition 1, not true? It relies on Properties 1 and 2, which we verified in Sec. 5 for Dirichlet boundary conditions. But these two properties are no longer true in the Neumann case. Let's investigate why for the rod-buckling example.

6.1 Number of negative inborn eigenvalues of $\mathcal{S}$

We have as before $(\mathcal{S}h)(s) = -h''(s) - \lambda\cos(\theta(s))h(s)$. If we take $\sigma = \epsilon$ for some small $\epsilon > 0$, and expand $\cos(\theta(s))$ about $s = 0$, we have:

$$(\mathcal{S}h)(s) = -h''(s) - Ch(s) + O(\epsilon).$$

where C denotes the constant $\lambda\cos(\theta(0))$. We seek the eigenvalues of $\mathcal{S}$ on $[0,\epsilon]$, i.e., solutions ρ_j to

$$-h''(s) - Ch(s) = \rho_j h(s), \quad h'(0) = 0, \quad h'(\epsilon) = 0,$$

or, equivalently,

$$h''(s) + (\rho_j + C)h(s) = 0, \quad h'(0) = 0, \quad h'(\epsilon) = 0,$$

The differential equation has the same array of possible solutions we previously derived in Eq. (9), with λ replaced by C. When we apply the Neumann boundary conditions, we find an eigenvalue $\rho = -C$ in addition to the infinite family $\rho = -C + \frac{n^2\pi^2}{\epsilon^2}$, $n = 1, 2, \cdots$. If $C < 0$, all of these eigenvalues are positive, so there are no negative inborn eigenvalues of $\mathcal{S}$ and Property 1 is true. If $C > 0$, however, then $\rho = -C$ is negative (and the rest of the eigenvalues are positive for ϵ small), so there is one negative inborn eigenvalue.

6.2 Direction of motion of zero eigenvalue at conjugate point

How about Property 2? It too can fail when we have Neumann boundary conditions, but the reason why is a bit subtle. Essentially, the problem is that the test function j_i defined by Eq. (10) as an eigenfunction h_i from $s = 0$ to $s = \sigma_1$ followed by the zero function from $s = \sigma_1$ to $s = \sigma_2$ is unacceptable. In the Dirichlet case, we had $h_i(\sigma_1) = 0$, so j_i was at least continuous, and therefore, as discussed in the Appendix, we could approximate it to within $1/n$ by C^2 functions with second derivatives that were only $O(n)$ and that was small enough to make the proof of Property 2 work. In the Neumann case, however, we may have $h_i(\sigma_1) \neq 0$ (we only know $h_i'(\sigma_1) = 0$), and so j_i may have a jump discontinuity. Then, if we try to approximate it to within $1/n$ by a C^2 function, the first derivative must be $O(n)$ somewhere, and therefore, the second derivative would be $O(n^2)$ and our analysis would break down.

Our solution to this quandary is to implement a numerical calculation that detects whether a zero eigenvalue corresponding to a conjugate point σ has positive or negative velocity, in the sense of determining whether the eigenvalue is positive or negative at $\sigma + \epsilon$ for small ϵ. We assume that a solution to the Euler-Lagrange equations has been determined numerically. The conjugate point equation from Eq. (11) can be solved numerically by solving the initial value problem (IVP)

$$h'' = (-\lambda \cos\theta)h, \quad h'(0) = 0, \quad h(0) = 1 \tag{12}$$

(we may set $h(0) = 1$, since the differential equation is linear, and so any solution with $h'(0) = 0$ is a multiple of the solution of Eq. (12)). Every value of σ between 0 and 1 for which $h'(\sigma) = 0$ in the numerical solution to Eq. (12) is a conjugate point. Now we wish to determine the direction of motion of the zero eigenvalue at each of these conjugate points, ideally in terms of quantities available from our IVP solver.

At the conjugate point just determined, we have $\rho_j(\sigma) = 0$ for some j, and $h(s)$ is a corresponding eigenfunction $h_j(s;\sigma)$. We introduce a small parameter $\epsilon > 0$ and determine the sign of $\rho_j(\sigma - \epsilon)$: if $\rho_j(\sigma - \epsilon) < 0$, the conjugate point is positive-velocity; otherwise it is negative-velocity.

We know that $\rho_j(\sigma-\epsilon) = \langle h_j(s;\sigma-\epsilon), \mathcal{S}h_j(s;\sigma-\epsilon)\rangle$, but $h_j(s;\sigma-\epsilon)$ is not available from our IVP solver. The restriction of $h_j(s;\sigma)$ to the interval $[0, \sigma-\epsilon]$ is a tempting choice, but it may not be in $\mathcal{D}([0,\sigma-\epsilon])$ because it may not satisfy the Neumann condition at $\sigma-\epsilon$. Still, if we make the natural assumption that when σ is changed by ϵ, both ρ_j and the corresponding eigenfunction h_j change by $O(\epsilon)$, the restriction should be within $O(\epsilon)$ of $h_j(s;\sigma-\epsilon)$. Furthermore, we can

tweak the restriction by only $O(\epsilon)$ and create an element of $\mathcal{D}([0, \sigma - \epsilon])$, e.g., by defining

$$\chi(s) \equiv h_j(s;\sigma) + As^2$$

(the function s^2 is chosen since it preserves the Neumann condition at $s = 0$), and determining the value of A so that $\chi'(\sigma - \epsilon) = 0$:

$$h_j'(\sigma - \epsilon;\sigma) + 2A(\sigma - \epsilon) = 0.$$

In order to express A in terms of values of $h_j(s;\sigma)$ at $s = \sigma$ rather than $\sigma - \epsilon$, we Taylor expand $h_j'(\sigma - \epsilon;\sigma)$ to find:

$$h_j'(\sigma;\sigma) - \epsilon h_j''(\sigma;\sigma) + 2A(\sigma - \epsilon) = 0.$$

By the Neumann condition, $h_j'(\sigma;\sigma) = 0$, and by $\mathcal{S}h = 0$, $h_j''(\sigma;\sigma) = -\lambda\cos(\theta(\sigma))h_j(\sigma;\sigma)$, so:

$$A = -\frac{\epsilon\lambda\cos(\theta(\sigma))h_j(\sigma;\sigma)}{2(\sigma - \epsilon)}.$$

Finally, we note that removing ϵ from the denominator would change A by $O(\epsilon^2)$, and $\chi(s)$ would thus still be within $O(\epsilon)$ of $h_j(s;\sigma - \epsilon)$. So, for convenience, we instead set

$$A = -\frac{\epsilon\lambda\cos(\theta(\sigma))h_j(\sigma;\sigma)}{2\sigma}.$$

Now we compute $\langle\chi, \mathcal{S}\chi\rangle$ (which we then argue has the same sign as $\rho_j(\sigma - \epsilon)$):

$$\mathcal{S}\chi = \mathcal{S}h_j + A\mathcal{S}(s^2) = 0 + A(-2 - \lambda\cos(\theta(s))s^2).$$

Thus,

$$\langle\chi, \mathcal{S}\chi\rangle = -A\int_0^{\sigma-\epsilon}(2\chi(s) + \lambda\cos(\theta(s))s^2\chi(s))ds$$

Now, changing the upper limit of the integral to σ changes the integral by $O(\epsilon)$, and since $A = O(\epsilon)$, changing $\chi(s)$ to $h_j(s;\sigma)$ also changes the integral by $O(\epsilon)$, so altogether

$$\langle\chi, \mathcal{S}\chi\rangle = -A\left[2\int_0^{\sigma} h_j(s;\sigma)ds + \lambda\int_0^{\sigma}\cos(\theta(s))s^2h_j(s;\sigma))ds\right] + O(\epsilon^2). \tag{13}$$

Thus, we adjoin two new differential equations to our IVP from Eq. (12):

$$\begin{aligned} v' &= h, \\ w' &= \cos(\theta)s^2h, \\ v(0) &= 0, \quad w(0) = 0, \end{aligned}$$

and at each conjugate point, the values of $v(\sigma) = \int_0^{\sigma} h(s)ds$ and $w(\sigma) = \int_0^{\sigma}\cos(\theta(s))s^2h(s)ds$ allow us to compute the leading-order term in $\langle\chi, \mathcal{S}\chi\rangle$ by Eq. (13). Note that the only remaining presence of ϵ is as a constant multiple of A, and hence it may be omitted altogether from the computations.

Claim: If $\langle \chi, \mathcal{S}\chi \rangle \neq 0$, then its sign is the same as the sign of $\rho_j(\sigma - \epsilon)$ for ϵ sufficiently small.

Proof: We may express χ as

$$\chi(s) = \sqrt{1 - b^2}\, h_j(s; \sigma - \epsilon) + b h_r(s),$$

where $b = O(\epsilon)$ and h_r is a unit vector orthogonal to $h_j(s; \sigma - \epsilon)$. Expanding h_r in the remaining eigenfunctions at $\sigma - \epsilon$:

$$\chi(s) = \sqrt{1 - b^2}\, h_j(s; \sigma - \epsilon) + b \sum_{i \neq j} c_i h_i(s; \sigma - \epsilon).$$

Since the h_i are orthonormal eigenfunctions of $\mathcal{S}$, a straightforward computation shows that:

$$\langle \chi, \mathcal{S}\chi \rangle = (1 - b^2)\rho_j(\sigma - \epsilon) + b^2 \sum_{i \neq j} c_i^2 \rho_i(\sigma - \epsilon).$$

Since $b = O(\epsilon)$ and $\rho_j(\sigma - \epsilon) = O(\epsilon)$, the first term dominates the second term, so that:

$$\langle \chi, \mathcal{S}\chi \rangle = \rho_j(\sigma - \epsilon) + O(\epsilon^2).$$

Since $\rho_j(\sigma - \epsilon) = O(\epsilon)$, we can see that $\langle \chi, \mathcal{S}\chi \rangle$ and $\rho_j(\sigma - \epsilon)$ agree in sign for ϵ sufficiently small.

Our computation finds the leading-order term of $\langle \chi, \mathcal{S}\chi \rangle$, and so, as long as it is non-zero, it too agrees in sign with $\rho_j(\sigma - \epsilon)$ for ϵ sufficiently small.

6.3 A Revised Theory

Once we know how to determine the number of negative inborn eigenvalues of $\mathcal{S}$, and the sign of the velocity of the zero-eigenvalue at each conjugate point, it is straightforward to adapt the classic Jacobi condition generalized in Proposition 1 to the following revised version:

Proposition 3.

$$\begin{aligned} \textit{Index} \;=\; & (\#\ \textit{of negative inborn eigenvalues of } \mathcal{S}) + (\#\ \textit{of negative-velocity conjugate points}) \\ & -(\#\ \textit{of positive-velocity conjugate points}) \end{aligned}$$

(Here, "positive-velocity conjugate point" is shorthand for "conjugate point at which the zero eigenvalue has positive velocity", and similarly for "negative-velocity conjugate point"). The logic of this proposition is the same as for the traditional Jacobi's condition using our spectrum-based perspective: the index, or number of negative eigenvalues on $[0, 1]$ is determined by tracking zero eigenvalues on $[0, \sigma]$ for values of σ between 0 and 1. In the traditional setting, we knew we started with no negative eigenvalues, and all zero eigenvalues were "becoming negative", so the index was just the number of conjugate points. With Proposition 3, we just have to remember that we may start with some negative eigenvalues, and that some zero eigenvalues are "becoming positive".

6.4 Example: Rod buckling with pinned ends

Fig. 4 shows a bifurcation diagram for the rod-buckling problem with Neumann-Neumann boundary conditions $\theta'(0) = \theta'(1) = 0$. As in the Dirichlet case, the horizontal branch contains

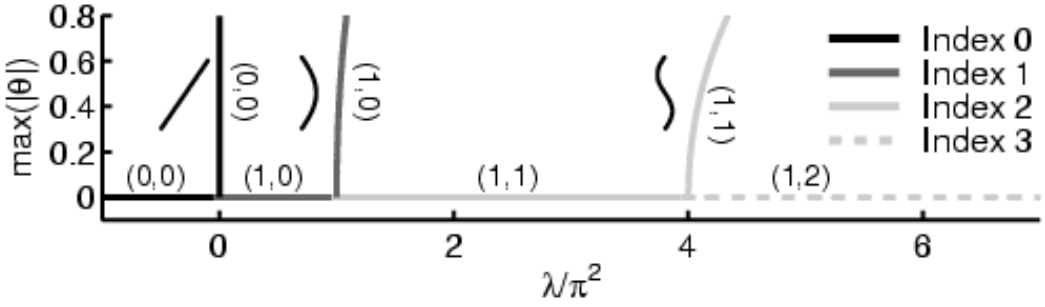

Figure 4: Bifurcation diagram for rod buckling using energy from Eq. (7) and pinned boundary conditions $\theta'(0) = \theta'(1) = 0$. Branches are labeled (n_i, n_c) where n_i is the number of negative inborn eigenvalues of $\mathcal{S}$ and n_c is the number of conjugate points (all negative-velocity).

the unbuckled straight-up configurations, and there are branches of buckled solutions bifurcating off the horizontal branch at $\lambda = \pi^2$ and $\lambda = 4\pi^2$. In this case, there is an additional vertical branch at $\lambda = 0$. Solutions on this branch are fully extended but tilted at some angle; the Neumann boundary conditions allow this.

The index suggests that the straight-up solution is stable (index zero) if $\lambda < 0$ (we are pulling up on the rod), but that for even a tiny downward force, that straight-up solution is unstable. Given that the rod is pinned at the origin and thus free to rotate freely, this makes sense. At zero force ($\lambda = 0$), the solutions that are fully extended and tilted are also stable, which also makes sense, since we are ignoring the gravitational force on the rod itself. (In fact, these solutions are marginally stable: $\mathcal{S}$ has a zero eigenvalue). If we extended the diagram to larger values of $\max|\theta|$, we would see a branch of straight-down solutions $\theta(s) = \pi$ that are stable (index zero) for $\lambda > 0$.

Notice the annotation of this diagram by (n_i, n_c), with n_i the number of negative inborn eigenvalues and n_c the number of conjugate points. If we had mistakenly used Jacobi's condition (ignoring the number of negative inborn eigenvalues), we would have concluded that the straight-up solution was stable for $0 < \lambda < \pi^2$, an erroneous conclusion. In addition, if we had continued either curve branch up to higher values of $\max|\theta|$, we would have reached a point at which $\theta(0) > \pi/2$, so that $\cos(\theta(0)) < 0$. At this point, we simultaneously gain a conjugate point and lose an negative inborn eigenvalue. This (correctly) leaves the index fixed, whereas Proposition 1 would erroneously predict that the index increases by one.

6.5 Example: Rod buckling with end torques

Finally, we consider an example that illustrates the importance of distinguishing between positive-velocity and negative-velocity conjugate points. We change the boundary conditions

to $\theta'(0) = 0.75$, $\theta'(1) = -1.5$, i.e., we impose torques in opposite directions at the two ends. The resulting bifurcation diagram in shown in Fig. 5.

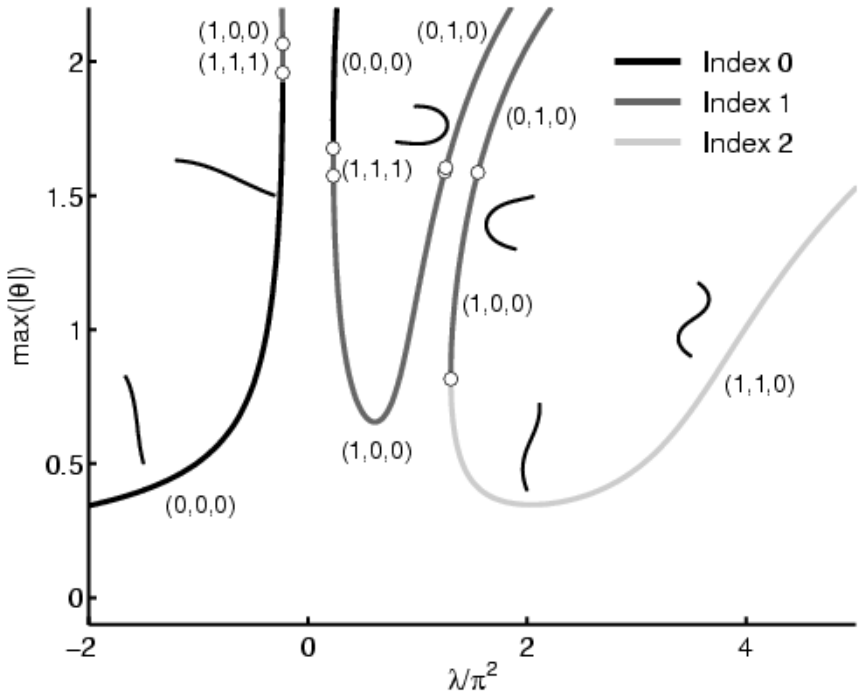

Figure 5: Bifurcation diagram for rod buckling using energy from Eq. (7) and torqued boundary conditions $\theta'(0) = 0.75, \theta'(1) = -1.5$. Branches are labeled (n_i, n_n, n_p) where n_i is the number of negative inborn eigenvalues of $\mathcal{S}$, n_n is the number of negative-velocity conjugate points, and n_p is the number of positive-velocity conjugate points. Open circles mark the points when any of these counts change. On the right side of the middle branch, there are two open circles very close to each other. The counts are $(1, 1, 1)$ between these circles.

We have chosen relatively large boundary torques, and thus we see no solutions with $\max|\theta| < 0.25$ radians. Still, we can understand this diagram as a perturbation of the diagram in Fig. 4, with the bifurcation points at $\lambda = 0, \pi^2, 4\pi^2$ "splitting" so that our new diagram consists of separated branches.

To clarify this connection between Fig. 5 and Fig. 4, we note that each branch in Fig. 4, except for the horizontal one, is in fact a pair of branches lying on top of each other, one for buckling left and one for buckling right (or, in the case of the $\lambda = 0$ branch, tilting left and tilting right). When we perturb the problem by adding end torques, the "children" of these branches become separately visible in Fig. 5.

The leftmost branch in Fig. 5 is the "child" of the $\lambda < 0$ portion of the horizontal branch from Fig. 4 plus one of the children of the $\lambda = 0$ branch. Each of these parent branches in Fig. 4 had the index equalling zero, and the same is true of the left branch in Fig. **??** (at least for $|\theta|$ small). Similarly, the rod configurations on this left branch imitate its parents in Fig. 4: we

have almost-straight-up configurations in the roughly-horizontal portion and mostly-extended configurations that are tilting left in the roughly-vertical portion.

The middle branch in Fig. 5 is the child of the "U" in Fig. 4 that is formed by the $\lambda = 0$ branch, the second branch of buckled configurations, and the horizontal branch for $0 < \lambda < \pi^2$. Once again, the rod configurations are similar (only one configuration is shown in Fig. 5 due to space considerations). The index of this middle branch in Fig. 5 is one (for $|\theta|$ small), matching the U from Fig. 4 except for the $\lambda = 0$ piece. This one discrepancy can be explained as follows. As previously noted, on the $\lambda = 0$ branch in Fig. 4, the operator $\mathcal{S}$ has a zero eigenvalue for each solution. In the child of this $\lambda = 0$ branch that is part of the middle branch in Fig. 5, the zero eigenvalue perturbs to a negative eigenvalue, yielding index one. (In the child that was part of the left branch, the zero eigenvalue perturbs to a positive eigenvalue, leaving the index zero).

Finally, the right branch in Fig. 5 is the child of the "U" in Fig. 4 that is formed by the second and third branches of buckled configurations and the horizontal branch for $\pi^2 < \lambda < 4\pi^2$: the index pattern and rod configurations match quite well.

We have annotated the diagram in Fig. 5 by the three counts that contribute to the index as per Proposition 3. Note in particular for larger values of $\max|\theta|$ several segments on which a positive-velocity conjugate point exists. In each case, if we counted all conjugate points as adding one to the index (as the traditional theory would dictate), our index prediction would be two larger than it currently is, yielding erroneous (and quite erratic) index patterns on the branches. Instead, we find index changes only at points in the diagram where λ folds (changes directions), matching the usual situation in bifurcation theory that the index only change at folds in the bifurcation parameter or at bifurcation points (not present in Fig. 5).

7 Discussion and Teaching Notes

Other extensions of basic conjugate point theory are possible in addition to accounting for Neumann-Neumann boundary conditions. It is relatively straightforward to handle the case in which J is a functional of a vector of unknown functions $\mathbf{y}(s) \in \mathbb{R}^n$ rather than a single real-valued function $y(s)$. The Euler-Lagrange equation becomes a vector differential equation; the differential equation $\mathcal{S}\mathbf{h} = 0$ has an n-dimensional basis of solutions to compute; and the conjugate point condition is that an appropriate determinant of this basis must cross zero [6, 8]. Although there is some matrix arithmetic required, this extension is significantly more straightforward than the Neumann-Neumann extension developed here.

In addition, "isoperimetric" constraints (of the form $\int_a^b g(y(s))ds = 0$) can be accounted for, using appropriate projection operators in the second-variation operator [8]. It is even possible to handle problems in which the Euler-Lagrange equations are integrodifferential equations [2].

The material in Sec. 2–4 should fit in any course featuring the calculus of variations, even if a full semester is not devoted to that topic. Indeed, this treatment should require minimally more time to present than the standard Jacobi condition conjugate point treatment. Ideally, this material would come after a treatment of numerical methods for initial and boundary value problems, so that an example such as the rod-buckling problem in Sec. 4, in which students

use a boundary-value problem method to solve the Euler-Lagrange equations and then solve the appropriate initial value problem to find conjugate points, can be treated.

8 Appendix: Proof of Property 2

Here we fill in the technical details of the proof of Property 2 for the case of Dirichlet boundary conditions, following the idea sketched at the end of Sec. 5. We are given σ_1, σ_2 with $a < \sigma_1 < \sigma_2 < b$, and a positive integer m, and we seek to prove that $\rho_m(\sigma_1) \leq \rho_m(\sigma_2)$ where $\rho_m(\sigma)$ is the eigenvalue of $\mathcal{S}$ on $[a, \sigma]$.

First, we note that given any positive integer n and real numbers a and b, there is a quintic polynomial p_n with $p_n(0) = p_n'(0) = p_n''(0) = 0$, $p_n(1/n) = 0$, $p_n'(1/n) = a$, and $p_n''(1/n) = b$ (there are six conditions to satisfy and six coefficients of the polynomial). It can be readily verified that $|p_n(s)| \leq \frac{a}{n} + \frac{b}{n^2}$, $|p_n'(s)| \leq a + \frac{b}{n}$, and $|p_n''(s)| \leq 5an + b$ for $0 < s < 1/n$.

Next, recall the functions

$$j_i(s) = \begin{cases} h_i(s; \sigma_1), & \text{if } a \leq s \leq \sigma_1 \\ 0, & \text{if } \sigma_1 < s \leq \sigma_2 \end{cases}, \quad i = 1, 2, \cdots, m$$

from Sec. 5. Consider any positive integer n such that $\sigma_1 + (1/n) < \sigma_2$. By translating and reflecting s from our construction of p_n above, we can find a polynomial q_n with $q_n(\sigma_1 + (1/n)) = q_n'(\sigma_1 + (1/n)) = q_n''(\sigma_1 + (1/n)) = 0$, $q_n(\sigma_1) = 0$, $q_n'(\sigma_1) = a$, and $q_n''(\sigma_1) = b$, where a is the left-derivative of j_i at $s = \sigma$ and b is the left-second-derivative of j_i at $s = \sigma$. Now we use q_n to smoothly connect j_i on $[0, \sigma_1]$ to j_i on $[\sigma_1 + (1/n), \sigma_2]$, giving us a function $j_{in}(s)$ that is C^2 on $[a, \sigma_2]$ with $j_{in}(s) = j_i(s)$ for $s \leq \sigma_1$ and $s \geq \sigma_1 + (1/n)$ and obeying the following:

$$\begin{aligned} |j_{in}(s)| &\leq A/n, \text{ for } \sigma_1 < s < \sigma_1 + (1/n) \\ |j_{in}'(s)| &\leq B, \text{ for } \sigma_1 < s < \sigma_1 + (1/n) \\ |j_{in}''(s)| &\leq Cn, \text{ for } \sigma_1 < s < \sigma_1 + (1/n), \end{aligned}$$

where A, B, C are constants that depend only on a and b.

Next, there are constants $a_{1n}, \cdots, a_{mn}$ so that $\sum_{i=1}^m a_{in} j_{in}$ is orthogonal to $h_1(s; \sigma_2), \cdots, h_{m-1}(s; \sigma_2)$ by the argument proposed in Sec. 5, and we scale the set of a_{in} so that $\sum_{i=1}^m (a_{in})^2 = 1$.

By construction, the (L^2-)length ℓ_n of $\sum_{i=1}^m a_{in} j_{in}$ is close to one for n large, and we can

quantify how close:

$$\begin{aligned}
(\ell_n)^2 &= \int_a^{\sigma_2} \left(\sum_{i=1}^m a_{in} j_{in}(s)\right)\left(\sum_{k=1}^m a_{kn} j_{kn}(s)\right) ds \\
&= \int_a^{\sigma_1} \left(\sum_{i=1}^m a_{in} h_i(s;\sigma_1)\right)\left(\sum_{k=1}^m a_{kn} h_k(s;\sigma_2)\right) ds + \int_{\sigma_1}^{\sigma_1+(1/n)} \left(\sum_{i=1}^m a_{in} j_{in}(s)\right)\left(\sum_{k=1}^m a_{kn} j_{kn}(s)\right) ds \\
&= \sum_{i,k=1}^m a_{in} a_{kn} \delta_{ik} + \int_{\sigma_1}^{\sigma_1+(1/n)} \left(\sum_{i=1}^m a_{in} j_{in}(s)\right)\left(\sum_{k=1}^m a_{kn} j_{kn}(s)\right) ds \\
&= \sum_{k=1}^m (a_{kn})^2 + \int_{\sigma_1}^{\sigma_1+(1/n)} \left(\sum_{i=1}^m a_{in} j_{in}(s)\right)\left(\sum_{k=1}^m a_{kn} j_{kn}(s)\right) ds \\
&= 1 + \int_{\sigma_1}^{\sigma_1+(1/n)} \left(\sum_{i=1}^m a_{in} j_{in}(s)\right)\left(\sum_{k=1}^m a_{kn} j_{kn}(s)\right) ds.
\end{aligned}$$

(The second equality holds because $j_{in}(s)$ vanishes for $s > \sigma_1 + (1/n)$, and the third equality holds because $\{h_i(s;\sigma_1)\}$ are orthonormal on $[a, \sigma_1]$.) So,

$$\begin{aligned}
|(\ell_n)^2 - 1| &= \left| \int_{\sigma_1}^{\sigma_1+(1/n)} \left(\sum_{i=1}^m a_{in} j_{in}(s)\right)\left(\sum_{k=1}^m a_{kn} j_{kn}(s)\right) ds \right| \\
&\leq \int_{\sigma_1}^{\sigma_1+(1/n)} \left(\sum_{i=1}^m |a_{in}||j_{in}(s)|\right)\left(\sum_{k=1}^m |a_{kn}||j_{kn}(s)|\right) ds \\
&\leq \frac{A^2}{n^2} \int_{\sigma_1}^{\sigma_1+(1/n)} \left(\sum_{i=1}^m |a_{in}|\right)\left(\sum_{k=1}^m |a_{kn}|\right) ds \\
&\leq \frac{A^2}{n^2} \int_{\sigma_1}^{\sigma_1+(1/n)} (m)(m) ds = \frac{A^2 m^2}{n^3}.
\end{aligned}$$

(The second inequality holds because of our bound on $|j_{in}|$ on $[\sigma_1, \sigma_1 + (1/n)]$, and the third inequality holds because $|a_{in}| \leq 1$.)

We have constructed a function $\frac{1}{\ell_n} \sum_{i=1}^m a_{in} j_{in}$ that is an element of $\mathcal{B}_m(\sigma_2)$, and so by Proposition 2,

$$\begin{aligned}
\rho_m(\sigma_2) &\leq \left\langle \frac{1}{\ell_n} \sum_{i=1}^m a_{in} j_{in}, \mathcal{S}\left(\frac{1}{\ell_n} \sum_{k=1}^m a_{kn} j_{kn}\right) \right\rangle. \\
&= \frac{1}{(\ell_n)^2} \left[\sum_{i,k=1}^m a_{in} a_{kn} \int_a^{\sigma_1} h_i(s;\sigma_1)(\mathcal{S} h_k(s;\sigma_1)) ds + \sum_{i,k=1}^m a_{in} a_{kn} \int_{\sigma_1}^{\sigma_1+(1/n)} j_{in}(s)(\mathcal{S} j_{kn}(s)) ds \right]
\end{aligned}$$

Since $h_i(s;\sigma)$ is an eigenfunction of $\mathcal{S}$, the first term in the brackets equals:

$$\sum_{i,k=1}^m a_{in} a_{kn} \int_a^{\sigma_1} h_i(s;\sigma_1) \rho_k(\sigma_1) h_k(s;\sigma_1) ds.$$

By the orthonormality of the h_i, this equals $\sum_{i,k=1}^{m} a_{in}a_{kn}\rho_k(\sigma_1)\delta_{ik}$, which equals $\sum_{k=1}^{m}(a_{kn})^2\rho_k(\sigma_1)$. By the ordering of the eigenvalues ρ_k, this is less than or equal to $\sum_{k=1}^{m}(a_{kn})^2\rho_m(\sigma_1)$, which, pulling ρ_m outside of the sum, equals $\rho_m(\sigma_1)$. So, we have shown:

$$\rho_m(\sigma_2) \leq \frac{1}{(\ell_n)^2}\left[\rho_m(\sigma_1) + \sum_{i,k=1}^{m} a_{in}a_{kn}\int_{\sigma_1}^{\sigma_1+(1/n)} j_{in}(s)(\mathcal{S}j_{kn}(s))ds\right]$$

Let e_n represent the second term in the bracket. Now, $\mathcal{S}$ is a second-order differential operator $(\mathcal{S}h)(s) = P(s)h''(s) + Q(s)h'(s) + C(s)h(s)$, and looking back at Eq. (6), we note that P, Q, C are continuous on $[a, b]$ because of our assumption that F is C^3. Therefore, there is a constant D such that $|P(s)| \leq D$, $|Q(s)| \leq D$, and $|R(s)| \leq D$ for all $s \in [a, b]$. Then, using the bounds on j_{kn} and its first two derivatives, we can say that $|\mathcal{S}j_{kn}(s)| \leq En$ for $\sigma_1 \leq s \leq \sigma_1 + (1/n)$, where E is a constant dependent only on a, b, and D. Thus,

$$|e_n| \leq \sum_{i,k=1}^{m} |a_{in}||a_{kn}| \int_{\sigma_1}^{\sigma_1+(1/n)} (A/n)(En)ds = \sum_{i,k=1}^{m} |a_{in}||a_{kn}|(AE/n) \leq AEm^2/n.$$

Thus, overall, we have shown that

$$\rho_m(\sigma_2) \leq \frac{\rho_m(\sigma_1) + e_n}{(\ell_n)^2},$$

where $|e_n| < AEm^2/n$ and $|(\ell_n)^2 - 1| \leq A^2m^2/n^3$. Taking $n \to \infty$, we then find, since $e_n \to 0$ and $\ell_n \to 1$, that $\rho_m(\sigma_2) \leq \rho_m(\sigma_1)$.

Acknowledgments: The author is indebted to John Maddocks, who got him started in this field, and to Kathleen Hoffman, a longtime collaborator on developing conjugate point theories in various settings. This work was supported by NSF grant DMS-0384739.

References

[1] S. S. Antman. *Nonlinear Problems of Elasticity*. Springer-Verlag, New York, 1995.

[2] G. Bulman and R. Manning. Stability of an elastic rod buckling into a soft wall. *Proc. R. Soc. Lon. Ser. A*, 461:2423–2450, 2005.

[3] R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume I. Interscience Publishers, Inc., New York, 1953.

[4] E. J. Doedel, H. B. Keller, and J. P. Kernévez. Numerical analysis and control of bifurcation problems, parts I and II. *Int. J. of Bif. and Chaos*, 3, 4:493–520, 745–772, 1991.

[5] C. Fox. *An introduction to the calculus of variations*. Oxford University Press, London, 1950.

[6] I. M. Gelfand and S. V. Fomin. *Calculus of Variations*. Dover, Mineola, NY, 1963.

[7] D. Lay. *Linear Algebra and Its Applications*. Addison-Wesley, Boston, 3rd edition, 2003.

[8] R. Manning, K. Rogers, and J. Maddocks. Isoperimetric conjugate points with application to the stability of dna minicircles. *Proc. R. Soc. Lond. A*, 454:3047–3074, 1998.

[9] M. Morse. *Introduction to Analysis in the Large*. Institute for Advanced Study, Princeton, New Jersey, 1951.

[10] A. Naylor and G. Sell. *Linear operator theory in engineering and science*. Springer-Verlag, New York, 1982.

[11] H. Sagan. *Introduction to the Calculus of Variations*. Dover Publications, New York, 1969.

[12] A. Weinstein and W. Stenger. *Methods of Intermediate Problems for Eigenvalues: Theory and Ramifications*. Academic Press, New York, 1972.