

The Calculus of Variations

Peter J. Olver
School of Mathematics
University of Minnesota
Minneapolis, MN 55455
olver@umn.edu
<http://www.math.umn.edu/~olver>

Contents

1.	Introduction	2
2.	Examples of Variational Problems	2
	Minimal Curves, Optics, and Geodesics	3
	Minimal Surfaces	6
	Isoperimetric Problems and Constraints	8
3.	The Euler–Lagrange Equation	9
	The First Variation	9
	Curves of Shortest Length — Planar Geodesics	12
	Minimal Surface of Revolution	13
	The Brachistochrone Problem	16
	The duBois–Reymond Lemma	19
	A Cautionary Example	20
4.	Boundary Conditions	22
	Natural Boundary Conditions	22
	Null Lagrangians	25
	General Boundary Conditions	27
5.	The Second Variation	30
6.	Multi-dimensional Variational Problems	34
	The First Variation and the Euler–Lagrange Equations	35
	References	39

1. Introduction.

Minimization principles form one of the most wide-ranging means of formulating mathematical models governing the equilibrium configurations of physical systems. Moreover, many popular numerical integration schemes such as the powerful finite element method are also founded upon a minimization paradigm. In these notes, we will develop the basic mathematical analysis of nonlinear minimization principles on infinite-dimensional function spaces — a subject known as the “calculus of variations”, for reasons that will be explained as soon as we present the basic ideas. Classical solutions to minimization problems in the calculus of variations are prescribed by boundary value problems involving certain types of differential equations, known as the associated Euler–Lagrange equations. The mathematical techniques that have been developed to handle such optimization problems are fundamental in many areas of mathematics, physics, engineering, and other applications. In these notes, we will only have room to scratch the surface of this wide ranging and lively area of both classical and contemporary research.

The history of the calculus of variations is tightly interwoven with the history of mathematics, [12]. The field has drawn the attention of a remarkable range of mathematical luminaries, beginning with Newton and Leibniz, then initiated as a subject in its own right by the Bernoulli brothers Jakob and Johann. The first major developments appeared in the work of Euler, Lagrange, and Laplace. In the nineteenth century, Hamilton, Jacobi, Dirichlet, and Hilbert are but a few of the outstanding contributors. In modern times, the calculus of variations has continued to occupy center stage, witnessing major theoretical advances, along with wide-ranging applications in physics, engineering and all branches of mathematics.

Minimization problems that can be analyzed by the calculus of variations serve to characterize the equilibrium configurations of almost all continuous physical systems, ranging through elasticity, solid and fluid mechanics, electro-magnetism, gravitation, quantum mechanics, string theory, and many, many others. Many geometrical configurations, such as minimal surfaces, can be conveniently formulated as optimization problems. Moreover, numerical approximations to the equilibrium solutions of such boundary value problems are based on a nonlinear finite element approach that reduces the infinite-dimensional minimization problem to a finite-dimensional problem. See [21; Chapter 11] for full details.

Just as the vanishing of the gradient of a function of several variables singles out the critical points, among which are the minima, both local and global, so a similar “functional gradient” will distinguish the candidate functions that might be minimizers of the functional. The finite-dimensional calculus leads to a system of algebraic equations for the critical points; the infinite-dimensional functional analog results a boundary value problem for a nonlinear ordinary or partial differential equation whose solutions are the critical functions for the variational problem. So, the passage from finite to infinite dimensional nonlinear systems mirrors the transition from linear algebraic systems to boundary value problems.

2. Examples of Variational Problems.

The best way to appreciate the calculus of variations is by introducing a few concrete examples of both mathematical and practical importance. Some of these minimization

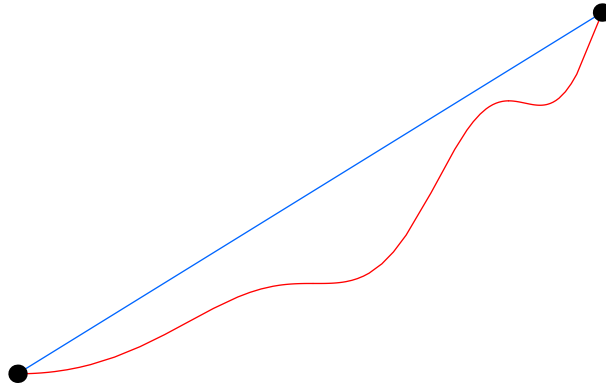


Figure 1. The Shortest Path is a Straight Line.

problems played a key role in the historical development of the subject. And they still serve as an excellent means of learning its basic constructions.

Minimal Curves, Optics, and Geodesics

The *minimal curve problem* is to find the shortest path between two specified locations. In its simplest manifestation, we are given two distinct points

$$\mathbf{a} = (a, \alpha) \quad \text{and} \quad \mathbf{b} = (b, \beta) \quad \text{in the plane } \mathbb{R}^2, \quad (2.1)$$

and our task is to find the curve of shortest length connecting them. “Obviously”, as you learn in childhood, the shortest route between two points is a straight line; see Figure 1. Mathematically, then, the minimizing curve should be the graph of the particular affine function[†]

$$y = cx + d = \frac{\beta - \alpha}{b - a} (x - a) + \alpha \quad (2.2)$$

that passes through or interpolates the two points. However, this commonly accepted “fact” — that (2.2) is the solution to the minimization problem — is, upon closer inspection, perhaps not so immediately obvious from a rigorous mathematical standpoint.

Let us see how we might formulate the minimal curve problem in a mathematically precise way. For simplicity, we assume that the minimal curve is given as the graph of a smooth function $y = u(x)$. Then, the length of the curve is given by the standard arc length integral

$$J[u] = \int_a^b \sqrt{1 + u'(x)^2} dx, \quad (2.3)$$

where we abbreviate $u' = du/dx$. The function $u(x)$ is required to satisfy the boundary conditions

$$u(a) = \alpha, \quad u(b) = \beta, \quad (2.4)$$

in order that its graph pass through the two prescribed points (2.1). The minimal curve problem asks us to find the function $y = u(x)$ that minimizes the arc length functional

[†] We assume that $a \neq b$, i.e., the points \mathbf{a}, \mathbf{b} do not lie on a common vertical line.

(2.3) among all “reasonable” functions satisfying the prescribed boundary conditions. The reader might pause to meditate on whether it is analytically obvious that the affine function (2.2) is the one that minimizes the arc length integral (2.3) subject to the given boundary conditions. One of the motivating tasks of the calculus of variations, then, is to rigorously prove that our everyday intuition is indeed correct.

Indeed, the word “reasonable” *is* important. For the arc length functional (2.3) to be defined, the function $u(x)$ should be at least piecewise C^1 , i.e., continuous with a piecewise continuous derivative. Indeed, if we were to allow discontinuous functions, then the straight line (2.2) does not, in most cases, give the minimizer. Moreover, continuous functions which are not piecewise C^1 need not have a well-defined arc length. The more seriously one thinks about these issues, the less evident the “obvious” solution becomes. But before you get too worried, rest assured that the straight line (2.2) is indeed the true minimizer. However, a fully rigorous proof of this fact requires a careful development of the mathematical machinery of the calculus of variations.

A closely related problem arises in geometrical optics. The underlying physical principle, first formulated by the seventeenth century French mathematician Pierre de Fermat, is that, when a light ray moves through an optical medium, it travels along a path that minimizes the travel time. As always, Nature seeks the most economical[†] solution. In an inhomogeneous planar optical medium, the speed of light, $c(x, y)$, varies from point to point, depending on the optical properties. Speed equals the time derivative of distance traveled, namely, the arc length of the curve $y = u(x)$ traced by the light ray. Thus,

$$c(x, u(x)) = \frac{ds}{dt} = \sqrt{1 + u'(x)^2} \frac{dx}{dt} .$$

Integrating from start to finish, we conclude that the total travel time along the curve is equal to

$$T[u] = \int_0^T dt = \int_a^b \frac{dt}{dx} dx = \int_a^b \frac{\sqrt{1 + u'(x)^2}}{c(x, u(x))} dx. \quad (2.5)$$

Fermat’s Principle states that, to get from point $\mathbf{a} = (a, \alpha)$ to point $\mathbf{b} = (b, \beta)$, the light ray follows the curve $y = u(x)$ that minimizes this functional subject to the boundary conditions

$$u(a) = \alpha, \quad u(b) = \beta,$$

If the medium is homogeneous, e.g., a vacuum[‡], then $c(x, y) \equiv c$ is constant, and $T[u]$ is a multiple of the arc length functional (2.3), whose minimizers are the “obvious” straight lines traced by the light rays. In an inhomogeneous medium, the path taken by the light ray is no longer evident, and we are in need of a systematic method for solving the minimization problem. Indeed, all of the known laws of geometric optics, lens design, focusing, refraction, aberrations, etc., will be consequences of the geometric and analytic properties of solutions to Fermat’s minimization principle, [3].

[†] Assuming time = money!

[‡] In the absence of gravitational effects due to general relativity.

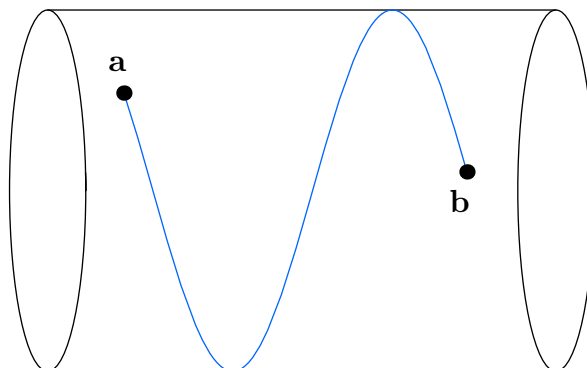


Figure 2. Geodesics on a Cylinder.

Another minimization problem of a similar ilk is to construct the *geodesics* on a curved surface, meaning the curves of minimal length. Given two points \mathbf{a}, \mathbf{b} lying on a surface $S \subset \mathbb{R}^3$, we seek the curve $C \subset S$ that joins them and has the minimal possible length. For example, if S is a circular cylinder, then there are three possible types of geodesic curves: straight line segments parallel to the center line; arcs of circles orthogonal to the center line; and spiral helices, the latter illustrated in Figure 2. Similarly, the geodesics on a sphere are arcs of great circles. In aeronautics, to minimize distance flown, airplanes follow geodesic circumpolar paths around the globe. However, both of these claims are in need of mathematical justification.

In order to mathematically formulate the geodesic minimization problem, we suppose, for simplicity, that our surface $S \subset \mathbb{R}^3$ is realized as the graph[†] of a function $z = F(x, y)$. We seek the geodesic curve $C \subset S$ that joins the given points

$$\mathbf{a} = (a, \alpha, F(a, \alpha)), \quad \text{and} \quad \mathbf{b} = (b, \beta, F(b, \beta)), \quad \text{lying on the surface } S.$$

Let us assume that C can be parametrized by the x coordinate, in the form

$$y = u(x), \quad z = v(x) = F(x, u(x)),$$

where the last equation ensures that it lies in the surface S . In particular, this requires $a \neq b$. The length of the curve is supplied by the standard three-dimensional arc length

[†] Cylinders are not graphs, but can be placed within this framework by passing to cylindrical coordinates. Similarly, spherical surfaces are best treated in spherical coordinates. In differential geometry, [5], one extends these constructions to arbitrary parametrized surfaces and higher dimensional manifolds.

integral. Thus, to find the geodesics, we must minimize the functional

$$\begin{aligned} J[u] &= \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2 + \left(\frac{dz}{dx}\right)^2} dx \\ &= \int_a^b \sqrt{1 + \left(\frac{du}{dx}\right)^2 + \left(\frac{\partial F}{\partial x}(x, u(x)) + \frac{\partial F}{\partial u}(x, u(x)) \frac{du}{dx}\right)^2} dx, \end{aligned} \tag{2.6}$$

subject to the boundary conditions $u(a) = \alpha$, $u(b) = \beta$. For example, geodesics on the paraboloid

$$z = \frac{1}{2}x^2 + \frac{1}{2}y^2 \tag{2.7}$$

can be found by minimizing the functional

$$J[u] = \int_a^b \sqrt{1 + u'^2 + (x + uu')^2} dx. \tag{2.8}$$

Minimal Surfaces

The minimal surface problem is a natural generalization of the minimal curve or geodesic problem. In its simplest manifestation, we are given a simple closed curve $C \subset \mathbb{R}^3$. The problem is to find the surface of least total area among all those whose boundary is the curve C . Thus, we seek to minimize the surface area integral

$$\text{area } S = \iint_S dS$$

over all possible surfaces $S \subset \mathbb{R}^3$ with the prescribed boundary curve $\partial S = C$. Such an area-minimizing surface is known as a *minimal surface* for short. For example, if C is a closed plane curve, e.g., a circle, then the minimal surface will just be the planar region it encloses. But, if the curve C twists into the third dimension, then the shape of the minimizing surface is by no means evident.

Physically, if we bend a wire in the shape of the curve C and then dip it into soapy water, the surface tension forces in the resulting soap film will cause it to minimize surface area, and hence be a minimal surface[†]. Soap films and bubbles have been the source of much fascination, physical, æsthetical and mathematical, over the centuries, [12]. The minimal surface problem is also known as *Plateau's Problem*, named after the nineteenth century French physicist Joseph Plateau who conducted systematic experiments on such soap films. A satisfactory mathematical solution to even the simplest version of the minimal surface problem was only achieved in the mid twentieth century, [18, 19]. Minimal surfaces and related variational problems remain an active area of contemporary research, and are of

[†] More accurately, the soap film will realize a local but not necessarily global minimum for the surface area functional. Non-uniqueness of local minimizers can be realized in the physical experiment — the same wire may support more than one stable soap film.

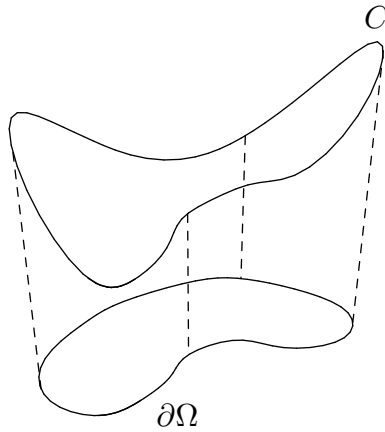


Figure 3. Minimal Surface.

importance in engineering design, architecture, and biology, including foams, domes, cell membranes, and so on.

Let us mathematically formulate the search for a minimal surface as a problem in the calculus of variations. For simplicity, we shall assume that the bounding curve C projects down to a simple closed curve $\Gamma = \partial\Omega$ that bounds an open domain $\Omega \subset \mathbb{R}^2$ in the (x, y) plane, as in Figure 3. The space curve $C \subset \mathbb{R}^3$ is then given by $z = g(x, y)$ for $(x, y) \in \Gamma = \partial\Omega$. For “reasonable” boundary curves C , we expect that the minimal surface S will be described as the graph of a function $z = u(x, y)$ parametrized by $(x, y) \in \Omega$. According to the basic calculus, the surface area of such a graph is given by the double integral

$$J[u] = \iint_{\Omega} \sqrt{1 + \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2} dx dy. \quad (2.9)$$

To find the minimal surface, then, we seek the function $z = u(x, y)$ that minimizes the surface area integral (2.9) when subject to the boundary conditions

$$u(x, y) = g(x, y) \quad \text{for} \quad (x, y) \in \partial\Omega, \quad (2.10)$$

that prescribe the boundary curve C . As we will see, (6.10), the solutions to this minimization problem satisfy a complicated nonlinear second order partial differential equation.

A simple version of the minimal surface problem, that still contains some interesting features, is to find minimal surfaces with rotational symmetry. A *surface of revolution* is obtained by revolving a plane curve about an axis, which, for definiteness, we take to be the x axis. Thus, given two points $\mathbf{a} = (a, \alpha)$, $\mathbf{b} = (b, \beta) \in \mathbb{R}^2$, the goal is to find the curve $y = u(x)$ joining them such that the surface of revolution obtained by revolving the curve around the x -axis has the least surface area. Each cross-section of the resulting surface is a circle centered on the x axis. The area of such a surface of revolution is given by

$$J[u] = \int_a^b 2\pi |u| \sqrt{1 + u'^2} dx. \quad (2.11)$$

We seek a minimizer of this integral among all functions $u(x)$ that satisfy the fixed boundary conditions $u(a) = \alpha$, $u(b) = \beta$. The minimal surface of revolution can be physically

realized by stretching a soap film between two circular wires, of respective radius α and β , that are held a distance $b - a$ apart. Symmetry considerations will require the minimizing surface to be rotationally symmetric. Interestingly, the revolutionary surface area functional (2.11) is exactly the same as the optical functional (2.5) when the light speed at a point is inversely proportional to its distance from the horizontal axis: $c(x, y) = 1/(2\pi|y|)$.

Isoperimetric Problems and Constraints

The simplest *isoperimetric problem* is to construct the simple closed plane curve of a fixed length ℓ that encloses the domain of largest area. In other words, we seek to maximize

$$\text{area } \Omega = \iint_{\Omega} dx dy \quad \text{subject to the constraint} \quad \text{length } \partial\Omega = \oint_{\partial\Omega} ds = \ell,$$

over all possible domains $\Omega \subset \mathbb{R}^2$. Of course, the “obvious” solution to this problem is that the curve must be a circle whose perimeter is ℓ , whence the name “isoperimetric”. Note that the problem, as stated, does not have a unique solution, since if Ω is a maximizing domain, any translated or rotated version of Ω will also maximize area subject to the length constraint.

To make progress on the isoperimetric problem, let us assume that the boundary curve is parametrized by its arc length, so $\mathbf{x}(s) = (x(s), y(s))$ with $0 \leq s \leq \ell$, subject to the requirement that

$$\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 = 1. \quad (2.12)$$

We can compute the area of the domain by a line integral around its boundary,

$$\text{area } \Omega = \oint_{\partial\Omega} x dy = \int_0^{\ell} x \frac{dy}{ds} ds, \quad (2.13)$$

and thus we seek to maximize the latter integral subject to the arc length constraint (2.12). We also impose periodic boundary conditions

$$x(0) = x(\ell), \quad y(0) = y(\ell), \quad (2.14)$$

that guarantee that the curve $\mathbf{x}(s)$ closes up. (Technically, we should also make sure that $\mathbf{x}(s) \neq \mathbf{x}(s')$ for any $0 \leq s < s' < \ell$, ensuring that the curve does not cross itself.)

A simpler isoperimetric problem, but one with a less evident solution, is the following. Among all curves of length ℓ in the upper half plane that connect two points $(-a, 0)$ and $(a, 0)$, find the one that, along with the interval $[-a, a]$, encloses the region having the largest area. Of course, we must take $\ell \geq 2a$, as otherwise the curve will be too short to connect the points. In this case, we assume the curve is represented by the graph of a non-negative function $y = u(x)$, and we seek to maximize the functional

$$\int_{-a}^a u dx \quad \text{subject to the constraint} \quad \int_{-a}^a \sqrt{1 + u'^2} dx = \ell. \quad (2.15)$$

In the previous formulation (2.12), the arc length constraint was imposed at every point, whereas here it is manifested as an integral constraint. Both types of constraints, pointwise

and integral, appear in a wide range of applied and geometrical problems. Such constrained variational problems can profitably be viewed as function space versions of constrained optimization problems. Thus, not surprisingly, their analytical solution will require the introduction of suitable Lagrange multipliers.

3. The Euler–Lagrange Equation.

Even the preceding limited collection of examples of variational problems should already convince the reader of the tremendous practical utility of the calculus of variations. Let us now discuss the most basic analytical techniques for solving such minimization problems. We will exclusively deal with classical techniques, leaving more modern direct methods — the function space equivalent of gradient descent and related methods — to a more in–depth treatment of the subject, [7].

Let us concentrate on the simplest class of variational problems, in which the unknown is a continuously differentiable scalar function, and the functional to be minimized depends upon at most its first derivative. The basic minimization problem, then, is to determine a suitable function $y = u(x) \in C^1[a, b]$ that minimizes the *objective functional*

$$J[u] = \int_a^b L(x, u, u') dx. \quad (3.1)$$

The integrand is known as the *Lagrangian* for the variational problem, in honor of Lagrange. We usually assume that the Lagrangian $L(x, u, p)$ is a reasonably smooth function of all three of its (scalar) arguments x, u , and p , which represents the derivative u' . For example, the arc length functional (2.3) has Lagrangian function $L(x, u, p) = \sqrt{1 + p^2}$, whereas in the surface of revolution problem (2.11), $L(x, u, p) = 2\pi |u| \sqrt{1 + p^2}$. (In the latter case, the points where $u = 0$ are slightly problematic, since L is not continuously differentiable there.)

In order to uniquely specify a minimizing function, we must impose suitable boundary conditions at the endpoints of the interval. To begin with, we concentrate on fixing the values of the function

$$u(a) = \alpha, \quad u(b) = \beta, \quad (3.2)$$

at the two endpoints. At the end of this section, we consider other possibilities.

The First Variation

The (local) minimizers of a (sufficiently nice) objective function defined on a finite-dimensional vector space are initially characterized as critical points, where the objective function’s gradient vanishes, [17]. An analogous construction applies in the infinite-dimensional context treated by the calculus of variations. Every sufficiently nice minimizer of a sufficiently nice functional $J[u]$ is a “critical function”. Of course, not every critical point turns out to be a minimum — maxima, saddles, and many degenerate points are also critical. The characterization of nondegenerate critical points as local minima or maxima relies on the second derivative test, whose functional version, known as the second variation, will be the topic of the following Section 5.

But we are getting ahead of ourselves. The first order of business is to learn how to compute the gradient of a functional defined on an infinite-dimensional function space. The general definition of the gradient requires that we first impose an inner product $\langle u, v \rangle$ on the underlying function space. The gradient $\nabla J[u]$ of the functional (3.1) will then be defined by the same basic directional derivative formula :

$$\langle \nabla J[u], v \rangle = \left. \frac{d}{d\varepsilon} J[u + \varepsilon v] \right|_{\varepsilon=0}. \quad (3.3)$$

Here $v(x)$ is a function that prescribes the “direction” in which the derivative is computed. Classically, v is known as a *variation* in the function u , sometimes written $v = \delta u$, whence the term “calculus of variations”. Similarly, the gradient operator on functionals is often referred to as the *variational derivative*, and often written δJ . The inner product used in (3.3) is usually taken (again for simplicity) to be the standard L^2 inner product

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx \quad (3.4)$$

on function space. Indeed, while the formula for the gradient will depend upon the underlying inner product, the characterization of critical points does not, and so the choice of inner product is not significant here.

Now, starting with (3.1), for each fixed u and v , we must compute the derivative of the function

$$h(\varepsilon) = J[u + \varepsilon v] = \int_a^b L(x, u + \varepsilon v, u' + \varepsilon v') dx. \quad (3.5)$$

Assuming sufficient smoothness of the integrand allows us to bring the derivative inside the integral and so, by the chain rule,

$$\begin{aligned} h'(\varepsilon) &= \frac{d}{d\varepsilon} J[u + \varepsilon v] = \int_a^b \frac{d}{d\varepsilon} L(x, u + \varepsilon v, u' + \varepsilon v') dx \\ &= \int_a^b \left[v \frac{\partial L}{\partial u}(x, u + \varepsilon v, u' + \varepsilon v') + v' \frac{\partial L}{\partial p}(x, u + \varepsilon v, u' + \varepsilon v') \right] dx. \end{aligned}$$

Therefore, setting $\varepsilon = 0$ in order to evaluate (3.3), we find

$$\langle \nabla J[u], v \rangle = \int_a^b \left[v \frac{\partial L}{\partial u}(x, u, u') + v' \frac{\partial L}{\partial p}(x, u, u') \right] dx. \quad (3.6)$$

The resulting integral often referred to as the *first variation* of the functional $J[u]$. The condition

$$\langle \nabla J[u], v \rangle = 0$$

for a minimizer is known as the *weak form* of the variational principle.

To obtain an explicit formula for $\nabla J[u]$, the right hand side of (3.6) needs to be written as an inner product,

$$\langle \nabla J[u], v \rangle = \int_a^b \nabla J[u] v dx = \int_a^b h v dx$$

between some function $h(x) = \nabla J[u]$ and the variation v . The first summand has this form, but the derivative v' appearing in the second summand is problematic. However, one can easily move derivatives around inside an integral through integration by parts. If we set

$$r(x) \equiv \frac{\partial L}{\partial p}(x, u(x), u'(x)), \quad (3.7)$$

we can rewrite the offending term as

$$\int_a^b r(x) v'(x) dx = [r(b)v(b) - r(a)v(a)] - \int_a^b r'(x)v(x) dx, \quad (3.8)$$

where, again by the chain rule,

$$r'(x) = \frac{d}{dx} \left(\frac{\partial L}{\partial p}(x, u, u') \right) = \frac{\partial^2 L}{\partial x \partial p}(x, u, u') + u' \frac{\partial^2 L}{\partial u \partial p}(x, u, u') + u'' \frac{\partial^2 L}{\partial p^2}(x, u, u'). \quad (3.9)$$

So far we have not imposed any conditions on our variation $v(x)$. We are only comparing the values of $J[u]$ among functions that satisfy the imposed boundary conditions (3.2). Therefore, we must make sure that the varied function

$$\widehat{u}(x) = u(x) + \varepsilon v(x)$$

remains within this set of functions, and so

$$\widehat{u}(a) = u(a) + \varepsilon v(a) = \alpha, \quad \widehat{u}(b) = u(b) + \varepsilon v(b) = \beta.$$

For this to hold, the variation $v(x)$ must satisfy the corresponding homogeneous boundary conditions

$$v(a) = 0, \quad v(b) = 0. \quad (3.10)$$

As a result, both boundary terms in our integration by parts formula (3.8) vanish, and we can write (3.6) as

$$\langle \nabla J[u], v \rangle = \int_a^b \nabla J[u] v dx = \int_a^b v \left[\frac{\partial L}{\partial u}(x, u, u') - \frac{d}{dx} \left(\frac{\partial L}{\partial p}(x, u, u') \right) \right] dx. \quad (3.11)$$

Since this holds for all variations $v(x)$, we conclude that[†]

$$\nabla J[u] = \frac{\partial L}{\partial u}(x, u, u') - \frac{d}{dx} \left(\frac{\partial L}{\partial p}(x, u, u') \right). \quad (3.12)$$

This is our explicit formula for the functional gradient or variational derivative of the functional (3.1) with Lagrangian $L(x, u, p)$. Observe that the gradient $\nabla J[u]$ of a functional is a *function*.

[†] See Lemma 3.4 and the ensuing discussion for a complete justification of this step.

The *critical functions* $u(x)$ are, by definition, those for which the functional gradient vanishes: $\nabla J[u] = 0$. Thus, $u(x)$ must satisfy

$$\nabla J[u] = \frac{\partial L}{\partial u}(x, u, u') - \frac{d}{dx} \frac{\partial L}{\partial p}(x, u, u') = 0. \quad (3.13)$$

In view of (3.9), the critical equation (3.13) is, in fact, a second order ordinary differential equation,

$$E(x, u, u', u'') = \frac{\partial L}{\partial u}(x, u, u') - \frac{\partial^2 L}{\partial x \partial p}(x, u, u') - u' \frac{\partial^2 L}{\partial u \partial p}(x, u, u') - u'' \frac{\partial^2 L}{\partial p^2}(x, u, u') = 0, \quad (3.14)$$

known as the *Euler–Lagrange equation* associated with the variational problem (3.1), in honor of two of the most important contributors to the subject. Any solution to the Euler–Lagrange equation that is subject to the assumed boundary conditions forms a critical point for the functional, and hence is a potential candidate for the desired minimizing function. And, in many cases, the Euler–Lagrange equation suffices to characterize the minimizer without further ado.

Theorem 3.1. *Suppose the Lagrangian function is at least twice continuously differentiable: $L(x, u, p) \in C^2$. Then any C^2 minimizer $u(x)$ to the corresponding functional $J[u] = \int_a^b L(x, u, u') dx$, subject to the selected boundary conditions, must satisfy the associated Euler–Lagrange equation (3.13).*

Let us now investigate what the Euler–Lagrange equation tells us about the examples of variational problems presented at the beginning of this section. One word of caution: there do exist seemingly reasonable functionals whose minimizers are not, in fact, C^2 , and hence do not solve the Euler–Lagrange equation in the classical sense; see [2] for examples. Fortunately, in most variational problems that arise in real-world applications, such pathologies do not appear.

Curves of Shortest Length — Planar Geodesics

Let us return to the most elementary problem in the calculus of variations: finding the curve of shortest length connecting two points $\mathbf{a} = (a, \alpha)$, $\mathbf{b} = (b, \beta) \in \mathbb{R}^2$ in the plane. As we noted in Section 3, such planar geodesics minimize the arc length integral

$$J[u] = \int_a^b \sqrt{1 + u'^2} dx \quad \text{with Lagrangian} \quad L(x, u, p) = \sqrt{1 + p^2}, \quad (3.15)$$

subject to the boundary conditions

$$u(a) = \alpha, \quad u(b) = \beta.$$

Since

$$\frac{\partial L}{\partial u} = 0, \quad \frac{\partial L}{\partial p} = \frac{p}{\sqrt{1 + p^2}}, \quad (3.16)$$

the Euler–Lagrange equation (3.13) in this case takes the form

$$0 = - \frac{d}{dx} \frac{u'}{\sqrt{1+u'^2}} = - \frac{u''}{(1+u'^2)^{3/2}}.$$

Since the denominator does not vanish, this is the same as the simplest second order ordinary differential equation

$$u'' = 0. \tag{3.17}$$

We deduce that the solutions to the Euler–Lagrange equation are all affine functions, $u = cx + d$, whose graphs are straight lines. Since our solution must also satisfy the boundary conditions, the only critical function — and hence the sole candidate for a minimizer — is the straight line

$$y = \frac{\beta - \alpha}{b - a} (x - a) + \alpha \tag{3.18}$$

passing through the two points. Thus, the Euler–Lagrange equation helps to reconfirm our intuition that straight lines minimize distance.

Be that as it may, the fact that a function satisfies the Euler–Lagrange equation and the boundary conditions merely confirms its status as a critical function, and does not guarantee that it is the minimizer. Indeed, any critical function is also a candidate for *maximizing* the variational problem, too. The nature of a critical function will be elucidated by the second derivative test, and requires some further work. Of course, for the minimum distance problem, we “know” that a straight line cannot maximize distance, and must be the minimizer. Nevertheless, the reader should have a small nagging doubt that we may not have completely solved the problem at hand . . .

The Brachistochrone Problem

The most famous classical variational principle is the so-called *brachistochrone problem*. The compound Greek word “brachistochrone” means “minimal time”. An experimenter lets a bead slide down a wire that connects two fixed points under the influence of gravity. The goal is to shape the wire in such a way that, starting from rest, the bead slides from one end to the other in minimal time. Naïve guesses for the wire’s optimal shape, including a straight line, a parabola, a circular arc, or even a catenary are wrong, and one can do better through a careful analysis of the associated variational problem. The brachistochrone problem was originally posed by the Swiss mathematician Johann Bernoulli in 1696, and served as an inspiration for much of the subsequent development of the subject.

We take, without loss of generality, the starting point of the bead to be at the origin: $\mathbf{a} = (0, 0)$. The wire will bend downwards, and so, to avoid distracting minus signs in the subsequent formulae, we take the vertical y axis to point downwards. The shape of the wire will be given by the graph of a function $y = u(x) \geq 0$. The end point $\mathbf{b} = (b, \beta)$ is assumed to lie below and to the right, and so $b > 0$ and $\beta > 0$. The set-up is sketched in Figure 4.

To mathematically formulate the problem, the first step is to find the formula for the transit time of the bead sliding along the wire. Arguing as in our derivation of the optics

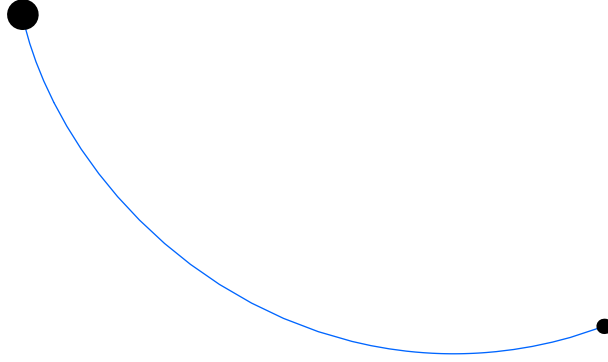


Figure 4. The Brachistochrone Problem.

functional (2.5), if $v(x)$ denotes the instantaneous speed of descent of the bead when it reaches position $(x, u(x))$, then the total travel time is

$$T[u] = \int_0^\ell \frac{ds}{v} = \int_0^b \frac{\sqrt{1+u'^2}}{v} dx, \quad (3.19)$$

where $ds = \sqrt{1+u'^2} dx$ is the usual arc length element, and ℓ is the overall length of the wire.

We shall use conservation of energy to determine a formula for the speed v as a function of the position along the wire. The kinetic energy of the bead is $\frac{1}{2}mv^2$, where m is its mass. On the other hand, due to our sign convention, the potential energy of the bead when it is at height $y = u(x)$ is $-mgu(x)$, where g the gravitational constant, and we take the initial height as the zero potential energy level. The bead is initially at rest, with 0 kinetic energy and 0 potential energy. Assuming that frictional forces are negligible, conservation of energy implies that the total energy must remain equal to 0, and hence

$$0 = \frac{1}{2}mv^2 - mgu.$$

We can solve this equation to determine the bead's speed as a function of its height:

$$v = \sqrt{2gu}. \quad (3.20)$$

Substituting this expression into (3.19), we conclude that the shape $y = u(x)$ of the wire is obtained by minimizing the functional

$$T[u] = \int_0^b \sqrt{\frac{1+u'^2}{2gu}} dx, \quad (3.21)$$

subject to the boundary conditions

$$u(0) = 0, \quad u(b) = \beta. \quad (3.22)$$

The associated Lagrangian is

$$L(x, u, p) = \sqrt{\frac{1+p^2}{u}}, \quad (3.23)$$

where we omit an irrelevant factor of $\sqrt{2g}$ (or adopt physical units in which $g = \frac{1}{2}$). We compute

$$\frac{\partial L}{\partial u} = -\frac{\sqrt{1+p^2}}{2u^{3/2}}, \quad \frac{\partial L}{\partial p} = \frac{p}{\sqrt{u(1+p^2)}}.$$

Therefore, the Euler–Lagrange equation for the brachistochrone functional is

$$-\frac{\sqrt{1+u'^2}}{2u^{3/2}} - \frac{d}{dx} \frac{u'}{\sqrt{u(1+u'^2)}} = -\frac{2uu'' + u'^2 + 1}{2\sqrt{u(1+u'^2)}} = 0. \quad (3.24)$$

Thus, the minimizing functions solve the nonlinear second order ordinary differential equation

$$2uu'' + u'^2 + 1 = 0. \quad (3.25)$$

Therefore, to find the critical functions, we need to solve a nonlinear second order ordinary differential equation — and not one in a familiar form. Rather than try to solve this differential equation directly, we note that the Lagrangian does not depend upon x , and therefore we can appeal to the following result.

Theorem 3.2. *Suppose the Lagrangian $L(x, u, p) = L(u, p)$ does not depend on x . Then the Hamiltonian function*

$$H(u, p) = L(u, p) - p \frac{\partial L}{\partial p}(u, p) \quad (3.26)$$

is a first integral for the Euler–Lagrange equation, meaning that it is constant on each solution:

$$H(u(x), u'(x)) = c \quad (3.27)$$

for some $c \in \mathbb{R}$, whose value can depend upon the solution $u(x)$.

Proof: Differentiating (3.26), we find

$$\frac{d}{dx} H(u, u') = \frac{d}{dx} \left[L(u, u') - u' \frac{\partial L}{\partial p}(u, u') \right] = u' \left(\frac{\partial L}{\partial u}(u, u') - \frac{d}{dx} \frac{\partial L}{\partial p}(u, u') \right) = 0,$$

which vanishes as a consequence of the Euler–Lagrange equation (3.13). This implies that the Hamiltonian function is constant, thereby establishing (3.27). *Q.E.D.*

Remark: This result is a special case of Noether’s powerful Theorem, [20; Chapter 4], that relates symmetries of variational problems — in this case translations in the x coordinate — with first integrals, a.k.a. conservation laws.

Equation (3.27) has the form of an implicitly defined first order ordinary differential equation which can, in fact, be integrated. Indeed, solving for

$$u' = h(u, c) \quad (3.28)$$

produces an autonomous first order differential equation, whose general solution can be obtained by integration:

$$\int \frac{du}{h(u, c)} = x + \delta, \quad (3.29)$$

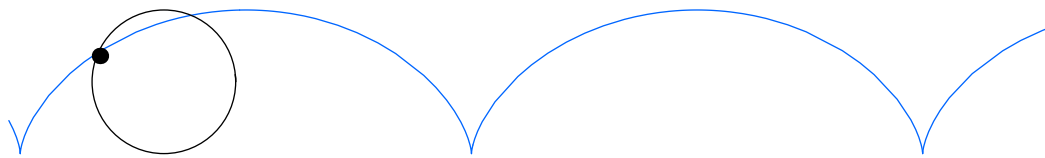


Figure 5. A Cycloid.

where δ is a second integration constant.

In our case (3.23), the Hamiltonian function is

$$H(u, p) = L - p \frac{\partial L}{\partial p} = \frac{1}{\sqrt{u(1+p^2)}}.$$

Thus, in view of (3.27),

$$H(x, u, u') = \frac{1}{\sqrt{u(1+u'^2)}} = c, \quad \text{which we rewrite as} \quad u(1+u'^2) = k,$$

where $k = 1/c^2$ is a constant. Solving for the derivative u' results in the first order autonomous ordinary differential equation

$$\frac{du}{dx} = \sqrt{\frac{k-u}{u}}.$$

This equation can be explicitly solved by separation of variables, and so

$$\int \sqrt{\frac{u}{k-u}} du = x + \delta$$

for some constant δ . The left hand integration relies on the trigonometric substitution

$$u = \frac{1}{2}k(1 - \cos \theta),$$

whereby

$$x + \delta = \frac{k}{2} \int \sqrt{\frac{1 - \cos \theta}{1 + \cos \theta}} \sin \theta d\theta = \frac{k}{2} \int (1 - \cos \theta) d\theta = \frac{1}{2}k(\theta - \sin \theta).$$

The left hand boundary condition implies $\delta = 0$, and so the solution to the Euler–Lagrange equation are curves parametrized by

$$x = r(\theta - \sin \theta), \quad u = r(1 - \cos \theta). \quad (3.30)$$

With a little more work, it can be proved that the parameter $r = \frac{1}{2}k$ is uniquely prescribed by the right hand boundary condition, and moreover, the resulting curve supplies the global minimizer of the brachistochrone functional, [9].

The parametrized curve (3.30) is known as a *cycloid*, which can be visualized as the curve traced by a point sitting on the edge of a rolling wheel of radius r , as plotted in Figure 5. (Checking this assertion is a useful exercise.) Interestingly, in certain configurations, namely if $\beta < 2b/\pi$, the cycloid that solves the brachistochrone problem dips below

the right hand endpoint $\mathbf{b} = (b, \beta)$, and so the bead is moving upwards when it reaches the end of the wire. Also note that the cycloid (3.30) has a vertical tangent at the initial point, so $u'(0) = \infty$ and the solution is not smooth there. This arises due to the singularity of the brachistochrone Lagrangian (3.23) when $u = 0$, or, equivalently, the fact that $u = 0$ is a singular point of the Euler–Lagrange equation (3.25).

Minimal Surface of Revolution

Consider next the problem of finding the curve connecting two points that generates a surface of revolution of minimal surface area. For simplicity, we assume that the curve is given by the graph of a *non-negative* function $y = u(x) \geq 0$. According to (2.11), the required curve will minimize the functional

$$J[u] = \int_a^b u \sqrt{1 + u'^2} \, dx, \quad \text{with Lagrangian} \quad L(x, u, p) = u \sqrt{1 + p^2}, \quad (3.31)$$

where we have omitted an irrelevant factor of 2π and used positivity to delete the absolute value on u in the integrand.

We have

$$\frac{\partial L}{\partial u} = \sqrt{1 + p^2}, \quad \frac{\partial L}{\partial p} = \frac{u p}{\sqrt{1 + p^2}},$$

and so the Euler–Lagrange equation (3.13) is

$$\sqrt{1 + u'^2} - \frac{d}{dx} \frac{u u'}{\sqrt{1 + u'^2}} = \frac{1 + u'^2 - u u''}{(1 + u'^2)^{3/2}} = 0. \quad (3.32)$$

Since the Lagrangian (3.23) does not depend upon x , we can again apply Theorem 3.2. The Hamiltonian function (3.26) is

$$H(u, p) = L - p \frac{\partial L}{\partial p} = \frac{u}{\sqrt{1 + p^2}},$$

and hence (3.27) implies

$$H(u, u') = \frac{u}{\sqrt{1 + u'^2}} = c \quad (3.33)$$

for some constant $c \in \mathbb{R}$. (This can be checked by directly calculating $dH/dx \equiv 0$.) Solving for[†]

$$\frac{du}{dx} = u' = \frac{\sqrt{u^2 - c^2}}{c}$$

results in an autonomous first order ordinary differential equation, which we can immediately solve:

$$\int \frac{c \, du}{\sqrt{u^2 - c^2}} = x + \delta,$$

[†] The square root is real since, by (3.33), $|u| \leq |c|$.

where δ is a constant of integration. The most useful form of the left hand integral is in terms of the inverse to the hyperbolic cosine function $\cosh z = \frac{1}{2}(e^z + e^{-z})$, whereby

$$\cosh^{-1} \frac{u}{c} = x + \delta, \quad \text{and hence} \quad u = c \cosh \left(\frac{x + \delta}{c} \right). \quad (3.34)$$

In this manner, we have produced the general solution to the Euler–Lagrange equation (3.32). Any solution that also satisfies the boundary conditions provides a critical function for the surface area functional (3.31), and hence is a candidate for the minimizer.

The curve prescribed by the graph of a hyperbolic cosine function (3.34) is known as a *catenary*. It is *not* a parabola, even though to the untrained eye it looks similar. Owing to their minimizing properties, catenaries are quite common in engineering design — for instance a hanging chain has the shape of a catenary, while the arch in St. Louis is an inverted catenary.

So far, we have not taken into account the boundary conditions. It turns out that there are three distinct possibilities, depending upon the configuration of the boundary points:

- (i) There is precisely one value of the two integration constants c, δ that satisfies the two boundary conditions.
- (ii) There are two different possible values of c, δ that satisfy the boundary conditions.
- (iii) There are *no* values of c, δ that allow (3.34) to satisfy the two boundary conditions.

This occurs when the two boundary points \mathbf{a}, \mathbf{b} are relatively far apart.

In the third configuration, the physical soap film spanning the two circular wires breaks apart into two circular disks, and this defines the minimizer for the problem; there is *no* surface of revolution that has a smaller surface area than the two disks. However, the “function”[†] that minimizes this configuration consists of two vertical lines from the boundary points to the x axis, along with that segment on the axis lying between them. More precisely, we can approximate this function by a sequence of genuine functions that give progressively smaller and smaller values to the surface area functional (2.11), but the actual minimum is not attained among the class of (smooth) functions.

There are further subtleties in the other two cases, which were completely resolved only relatively recently; see [15, 16], and further refinements in [14]. When only one catenary solution exists, it is *not* the minimizer, and the discontinuous solution constructed above has a smaller surface area. In the region where there are two possible catenaries, the one that is farther away from the axis, counterintuitively, gives a smaller surface area. However, this region splits into two subregions divided by a curve, called the *MacNeish frontier* where the farther away catenary has surface area exactly equal to the discontinuous solution, and either serves as the minimizer. The subregions on either side of the frontier correspond to when the catenary is the global minimizer or when the two disks have the smaller surface area. In the latter case, there are again nearby functions whose surface areas are arbitrarily close to the minimum offered by the disks.

[†] Here “function” must be taken in a *very* broad sense, as this one does not even correspond to a generalized function!



Figure 6. Bump Functions.

Thus, even in such a reasonably simple example, a number of the subtle complications can already be seen. (See also the cautionary example at the end of this section.) Lack of space precludes a more detailed development of the subject, and we refer the interested reader to more specialized books on the calculus of variations, including [6, 7, 8, 9, 10, 13].

The duBois–Reymond Lemma

The justification that a critical function satisfies the Euler–Lagrange equation relies on what is known as the *duBois–Reymond Lemma*. The simplest version of this key result can be stated as follows.

Lemma 3.3. *If $f(x)$ is continuous on the closed interval $[a, b]$, and*

$$\int_a^b f(x) v(x) dx = 0$$

for every continuous function $v(x)$, then $f(x) \equiv 0$ for all $a \leq x \leq b$.

Proof: A simple proof of the stated result proceeds by setting $v(x) = f(x)$. Then

$$0 = \int_a^b f(x) v(x) dx = \int_a^b [f(x)]^2 dx,$$

which, by continuity, is only possible when $f(x) \equiv 0$.

However, for many purposes, we would like to impose additional conditions on the functions $v(x)$, and the preceding proof will not work unless $f(x)$ satisfies the same conditions. A proof that has much wider applicability proceeds as follows. Suppose $f(x_0) \neq 0$ for some $a < x_0 < b$. By replacing $f(x)$ by $-f(x)$ if necessary, we can assume $f(x_0) > 0$. Then, by continuity, $f(x) > 0$ for all x lying in some small interval around x_0 , of the form $a < x_0 - \varepsilon < x < x_0 + \varepsilon < b$ for some $\varepsilon > 0$. Choose $v(x)$ to be a continuous function that is > 0 in this interval and $= 0$ outside. An example is the piecewise linear function

$$v(x) = \begin{cases} \varepsilon - |x - x_0|, & |x - x_0| < \varepsilon, \\ 0, & \text{otherwise,} \end{cases} \quad (3.35)$$

which is illustrated in the first plot in Figure 6. Then $f(x)v(x) \geq 0$ everywhere, hence $\int_a^b f(x)v(x) dx \geq 0$, and, by continuity, is equal to zero if and only if $f(x)v(x) \equiv 0$, which is a contradiction. *Q.E.D.*

Observe that one can replace the continuous function (3.35) by the smooth *bump function*

$$v(x) = \begin{cases} \varepsilon \exp [1 - (\varepsilon^{-2}(x - x_0)^2 - 1)^{-2}], & |x - x_0| < \varepsilon, \\ 0, & \text{otherwise,} \end{cases} \quad (3.36)$$

which, in addition to being positive on the required interval, is C^∞ everywhere, including at $x = x_0 \pm \varepsilon$, where all its derivatives vanish. The *support* of the bump function (3.36), meaning the closure of the set where it does not vanish, is the compact (closed and bounded) interval $\{|x - x_0| \leq \varepsilon\}$. This observation produces a useful strengthening of Lemma 3.3:

Lemma 3.4. *If $f(x)$ is continuous on $[a, b]$, and $\int_a^b f(x)v(x)dx = 0$ for every C^∞ function $v(x)$ with compact support in (a, b) , then $f(x) \equiv 0$ for all $a \leq x \leq b$.*

Note that the condition that $v(x)$ has compact support implies that $v(a) = v(b) = 0$, thus justifying our derivation of the Euler–Lagrange equation (3.13).

A Cautionary Example

While the Euler–Lagrange equation is fundamental in the analysis of variational problems, it does have limitations and analysis based solely thereon may miss important features. One complicating factor is that convergence in infinite-dimensional function space is considerably more subtle than in finite-dimensional Euclidean space. So let us conclude this section with one of a large variety of cautionary examples.

Consider the problem of minimizing the integral

$$J[u] = \int_0^1 \left[\frac{1}{2}(u'^2 - 1)^2 + \frac{1}{2}u^2 \right] dx \quad (3.37)$$

subject to the boundary conditions $u(0) = u(1) = 0$. The Euler–Lagrange equation is

$$0 = -\frac{d}{dx} [2u'(u'^2 - 1)] + u = (2 - 6u'^2)u'' + u = 0.$$

Since the Lagrangian is independent of x , Theorem 3.2 can be applied to reduce this to the first order ordinary differential equation given by the constancy of the Hamiltonian function

$$H(u, u') = L(u, u') - u' \frac{\partial L}{\partial p}(u, u') = -\frac{3}{2}u'^4 + u'^2 + \frac{1}{2} + \frac{1}{2}u^2 = c,$$

where $c \in \mathbb{R}$. As in (3.28, 29), the resulting quartic polynomial in u' can be solved for $u' = h(u, c)$, and the final autonomous first order ordinary differential equation can then be integrated. However, the resulting very complicated formula for the solution $u(x)$ is not particularly helpful, especially when one tries to impose the admitted boundary conditions to fix the values of the integration constants c, δ . A numerical approximation to the preceding two point boundary value problem would be considerably more enlightening.

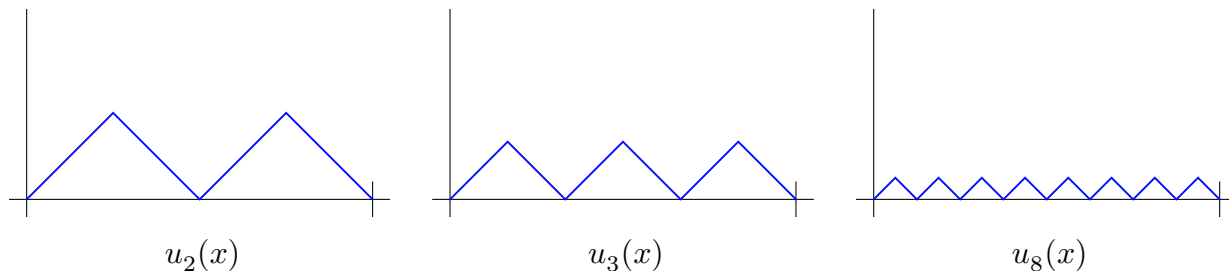


Figure 7. Sawtooth Functions.

Moreover, the solution to the Euler–Lagrange equation does not, in fact minimize[†] the functional, as the following direct analysis shows. Observe that the functional is clearly bounded below by $0 \leq J[u]$. However this minimum value cannot be achieved, since it would simultaneously require that $u'(x) = \pm 1$ and $u(x) = 0$. On the other hand, one can come arbitrarily close to the minimum as follows. Consider the “sawtooth” functions

$$u_n(x) = \begin{cases} x - \frac{k}{n}, & \frac{k}{n} \leq x \leq \frac{2k+1}{2n}, \\ \frac{k+1}{n} - x, & \frac{2k+1}{2n} \leq x \leq \frac{k+1}{n}, \end{cases} \quad \begin{matrix} k = 0, \dots, n-1, \\ n = 1, 2, 3, \dots, \end{matrix} \quad (3.38)$$

which are continuous, piecewise linear, with slope ± 1 , three of which are plotted in Figure 7. For these functions, the first term in the integrand vanishes, and hence

$$J[u_n] = \int_0^1 \frac{1}{2} u_n(x)^2 dx = \sum_{j=0}^{2n-1} \int_{j/(2n)}^{(j+1)/(2n)} \frac{1}{2} u_n(x)^2 dx = \frac{2n}{8n^2} = \frac{1}{4n}.$$

Thus, as $n \rightarrow \infty$, the value of $J[u_n] \rightarrow 0$. On the other hand, the functions (3.38) converge uniformly to the zero function: $u_n(x) \rightarrow u_*(x) \equiv 0$, and $J[u_*] = \frac{1}{2}$. In other words,

$$\lim_{n \rightarrow \infty} J[u_n] = 0 \neq \frac{1}{2} = J \left[\lim_{n \rightarrow \infty} u_n \right]. \quad (3.39)$$

We conclude that one can find functions $u(x)$ for which the value of the integral (3.37) is arbitrarily close to its overall minimum, but there is no function which achieves the value $J[u] = 0$.

Remark: The reader may object that the functions (3.38) used in this argument have corners at the points $x_j = j/(2n)$, and hence are not C^1 . But it is not hard to smooth out the corners, replacing each u_n by a nearby C^2 or even C^∞ function \tilde{u}_n , whose value $J[\tilde{u}_n]$ can be made arbitrarily close to that of u_n , and thus has the same limiting property (3.39).

[†] It may provide a local minimum — or maximum — but to ascertain this, one would need to invoke the second derivative test.

4. Boundary Conditions.

So far, we have only imposed fixed boundary conditions on the variational problem, in which the values of the minimizer are specified at the ends of the interval. Other types of boundary conditions can be handled, but there are some subtleties worth delving into.

Natural Boundary Conditions

A second common type of condition is that of a *free boundary*, in which no conditions are imposed on the minimizer at one or both ends of the interval. As an example of this type of problem, consider the planar geodesic problem in which one wishes to find the curve of shortest length that connects a given point to a vertical line. In other words, we seek to minimize the arc length functional (2.3) on the interval $[a, b]$ where the minimizing function is required to pass through the point (a, α) , so that $u(a) = \alpha$, but there are no conditions at the endpoint $x = b$. Thus we are seeking the shortest curve that goes from (a, α) to *any* point on the vertical line $L_b = \{x = b\}$. A moment's reflection will convince you that the minimizing curve is a horizontal straight line, so $u(x) \equiv \alpha$ for all $a \leq x \leq b$, that meets the vertical line L_b in a perpendicular direction. If we also impose no conditions at the initial endpoint, we then seek a curve of minimal length that goes from the vertical line $L_a = \{x = a\}$ to the vertical line $L_b = \{x = b\}$. In this case any horizontal line will serve as a minimizer, and so the minimizer is not unique.

Similarly, one could seek the solution to the brachistochrone problem with a free end, to determine the shape of a wire starting at the origin with the property that a bead sliding along it reaches the line L_b in the shortest possible time, without any need to specify which point on the line it ends up on. Thus, the functional to be minimized is again (3.21), but here one only imposes a condition, say $u(0) = 0$, at one end of the interval.

To analyze such problems, let us return to our general variational calculation. Thus, consider the usual *objective functional*

$$J[u] = \int_a^b L(x, u, u') dx. \quad (4.1)$$

Our goal is to minimize $J[u]$ among all sufficiently smooth function $u(x)$, defined for $a \leq x \leq b$, but now only subject to a single fixed boundary condition, namely

$$u(a) = \alpha, \quad (4.2)$$

while the value $u(b)$ at the other end of the interval remains arbitrary.

We can apply the same variational computation that leads to (3.7–8) since up to this point, the boundary conditions are not used. In order to preserve the remaining fixed boundary condition (4.2), the variation $v(x)$ must vanish at the left hand boundary point: $v(a) = 0$. However, there are no conditions imposed on $v(b)$ at the right hand end. Thus, only one of the boundary terms in (3.8) vanishes, and the result of the integration by parts computation, replacing (3.11), is

$$\begin{aligned} 0 &= \langle \nabla J[u], v \rangle \\ &= v(b) \frac{\partial L}{\partial p}(b, u(b), u'(b)) + \int_a^b v \left[\frac{\partial L}{\partial u}(x, u, u') - \frac{d}{dx} \left(\frac{\partial L}{\partial p}(x, u, u') \right) \right] dx. \end{aligned} \quad (4.3)$$

Thus, we can no longer identify the gradient of the functional $J[u]$ as the function given by the Euler–Lagrange expression, since there is an additional boundary component. Instead, let us work with (4.3) directly. Keep in mind that we have not imposed any conditions on $v(b)$. However, if we set $v(b) = 0$, then (4.3) reduces to the previous integral (4.3), whose vanishing, by the same argument based on Lemma 3.4, implies that $u(x)$ must still satisfy the Euler–Lagrange equation (3.13), and so the integral term in (4.3) vanishes. Thus for the entire right hand side to vanish for all such $v(x)$, the term multiplying $v(b)$ must also vanish, and we conclude that the minimizing function must satisfy

$$\frac{\partial L}{\partial p}(b, u(b), u'(b)) = 0. \quad (4.4)$$

This is known as a *natural boundary condition*, which imposes a constraint on the minimizer at the free boundary. We conclude that any minimizer for the boundary value problem that has only one endpoint fixed must be a solution to the Euler–Lagrange equation (3.13) supplemented by the two boundary conditions (4.2, 4) — one fixed and the other natural. Thus, the determination of the critical functions continues to require the solution to a two point boundary value problem for the second order Euler–Lagrange ordinary differential equation. If, instead of being fixed, the value of the minimizer at the other endpoint is also left free, then the same argument implies that the minimizer must also satisfy a natural boundary condition there:

$$\frac{\partial L}{\partial p}(a, u(a), u'(a)) = 0. \quad (4.5)$$

Thus, any minimizer of the free variational problem must solve the Euler–Lagrange equation along with the two natural boundary conditions (4.4–5). As always, the above are just necessary conditions for minimizers. Maximizers, if such exist, must also satisfy the same conditions, which serve to characterize the *critical functions* of the variational principle subject to the given boundary constraints. By this line of reasoning, we are led to the following conclusion.

At each endpoint of the interval, any critical function of a functional, including minimizers and maximizers, must satisfy either homogeneous or inhomogeneous fixed boundary conditions, or homogeneous natural boundary conditions.

Consequently, imposing any other type of boundary conditions at an endpoint, say $x = a$, will result in the critical function satisfying more than one boundary condition at the endpoint — the natural boundary condition plus the imposed boundary condition. This will effectively prescribe both $u(a)$ and $u'(a)$ which, by the basic existence and uniqueness theorem for the solution of second order ordinary differential equations, [4], would imply that they are satisfied by one and only one solution to the Euler–Lagrange equation, which may or may not satisfy the additional boundary condition(s) at the other endpoint; in the latter case there would be no critical functions associated with the variational principle, and hence no candidate minimizers

Example 4.1. Let us apply this method to the problem is to find the shortest path between a point and a straight line. By applying a rigid motion (translation and rotation), we can take the point to be $\mathbf{a} = (a, \alpha)$ and the line to be in the vertical direction, namely

$L_b = \{x = b\}$. Assuming the solution is given by the graph of a function $y = u(x)$, the length is given by the same arc length functional (3.15), but now there is only a single fixed boundary condition, namely $u(a) = \alpha$. To determine the second boundary condition at $x = b$, we use (4.4). In view of (3.16), this requires

$$\frac{u'(b)}{\sqrt{1 + u'(b)^2}} = 0, \quad \text{or, simply,} \quad u'(b) = 0.$$

This means that any critical function $u(x)$ must have horizontal tangent at the point $x = b$, or, equivalently, it must be perpendicular to the vertical line L_b .

The Euler–Lagrange equation is the same as before, reducing to $u'' = 0$, and hence the critical functions are affine: $u(x) = cx + d$. Substituting into the two boundary conditions $u(a) = \alpha$, $u'(b) = 0$, we conclude that the minimizer must be a horizontal straight line: $u(x) \equiv \alpha$ for all $a \leq x \leq b$, reconfirming our earlier observation. As before, this is not a complete proof that the horizontal straight line is a minimizer — based on what we know so far, it could be a maximizer — and resolving its status requires further analysis based on the second variation.

On the other hand, suppose we try to impose a different boundary condition at the endpoint — for example,

$$u'(b) = \beta u(b) + \gamma, \tag{4.6}$$

where β, γ are constants and $\beta \neq 0$, known as an inhomogeneous *Robin condition*. To maintain the boundary condition upon variation of $u \mapsto u + v$, the function $v(x)$ must satisfy the corresponding homogeneous Robin condition $v'(b) = \beta v(b)$. Since $v(b)$ can take on any value, according to the argument given above, any minimizer of the arc length functional must satisfy the homogeneous natural boundary conditions, namely $u'(b) = 0$. Combining this with the imposed Robin condition (4.6), there is only one possible solution to the Euler–Lagrange equation that satisfies both boundary conditions at $x = b$, namely the constant function $u(x) = -\gamma/\beta$, which does not satisfy the fixed boundary condition at $x = a$ unless $\alpha = -\gamma/\beta$. If this does not hold, the horizontal line $u(x) \equiv \alpha$ provides the shortest distance between the point (a, α) and the vertical line L_b , but there is no arc length minimizer that satisfies the Robin condition (4.6).

On the other hand, one can construct functions that satisfy the Robin boundary condition (4.6) whose arc length is arbitrarily close to that of the horizontal line, which is $b - a$. For example, we can slightly perturb the line by, say, setting

$$u_\varepsilon(x) = \begin{cases} \alpha, & a \leq x \leq b - \varepsilon, \\ \alpha + (\alpha\beta + \gamma)(x - b + \varepsilon)/(1 - \varepsilon\beta), & b - \varepsilon \leq x \leq b, \end{cases}$$

where $0 < \varepsilon < |\beta|$, which does satisfy (4.6); see Figure 8. Its arc length

$$b - a + \varepsilon \left(\sqrt{1 + \left(\frac{\alpha\beta + \gamma}{1 - \varepsilon\beta} \right)^2} - 1 \right) \longrightarrow b - a$$

as $\varepsilon \rightarrow 0$. One can even smooth off the corner, which has the effect of slightly decreasing its total arc length and thus does not affect the convergence as $\varepsilon \rightarrow 0$. Thus, while the

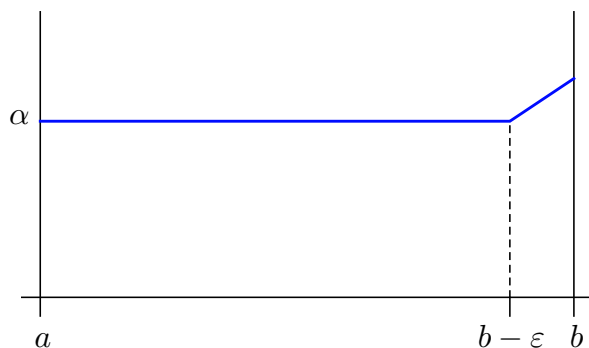


Figure 8. Perturbed Horizontal Line.

Robin boundary value problem has no minimizing solution, it does admit smooth functions that come arbitrarily close to the minimum possible value, which is $b - a$. This example exemplifies a common type of behavior for variational principles that are bounded from below, but have no bona fide function that achieves the minimum value.

Example 4.2. Let us return to the brachistochrone problem, but now let us seek the shape of the wire that enable the bead to get from the point $\mathbf{a} = (0, 0)$ to a point on the line $L_b = \{x = b\}$ where $b > 0$ in the shortest time. In this case, we no longer specify which vertical position on the line the bead ends up at, and so $x = b$ is a free end. Thus, we continue to minimize the same functional (3.21) subject to the single boundary condition $u(0) = 0$. The minimizer must satisfy the natural boundary condition at $x = b$. Since the Lagrangian is

$$L(x, u, p) = \sqrt{\frac{1 + p^2}{2gu}} \quad \text{with} \quad \frac{\partial L}{\partial p} = \frac{p}{\sqrt{2gu(1 + p^2)}},$$

the natural boundary condition (4.4) is simply $u'(b) = 0$. We conclude that the solution is provided by the cycloid (3.30) that starts out at the origin and has horizontal tangent at the point $x = b$. Note that, in view of (3.30), setting

$$\frac{du}{dx} = \frac{\sin \theta}{1 - \cos \theta} = 0 \quad \text{implies} \quad \theta = \pi,$$

which thereby serves to prescribe the parameter $r = b/\pi$. The resulting curve is the graph of a half of a cycloid arc.

Null Lagrangians

Despite the perhaps disappointing conclusion afforded by the inadmissibility of Robin conditions in Example 4.1, it turns out it is possible to impose other types of boundary conditions on a variational principle, by suitably modifying the functional. With this aim, we first introduce an important notion of independent interest in the calculus of variations.

Definition 4.3. A function $N(x, u, p)$ is called a *null Lagrangian* if the gradient of its associated variational functional $K[u] = \int_a^b N(x, u, u') dx$ vanishes identically: $\nabla K \equiv 0$.

Null Lagrangians can be viewed as the calculus of variations equivalent of constant functions, in that a function defined on a connected subset of \mathbb{R}^n has zero gradient if and only if it is constant, [17].

Example 4.4. Consider the variational problem

$$K[u] = \int_0^\infty [x^2 u' + 2xu] dx$$

for the Lagrangian $N = x^2 p + 2xu$. We compute its functional gradient using equation (3.12), and find

$$\nabla K[u] = \frac{\partial N}{\partial u} - \frac{d}{dx} \frac{\partial N}{\partial p} = 2x - \frac{d}{dx} (x^2) \equiv 0.$$

We conclude that N is a null Lagrangian.

Since the Euler–Lagrange equation of a null Lagrangian is simply $0 = 0$, every function satisfies it and thus, provided it also satisfies the imposed boundary conditions, is a critical function. We will shortly explain what this signifies. But first let us state an explicit characterization of all null Lagrangians.

Theorem 4.5. A function $N(x, u, p)$ defined for all[†] $(x, u, p) \in \mathbb{R}^3$ is a null Lagrangian if and only if it is a total derivative, so

$$N(x, u, u') = \frac{d}{dx} A(x, u) = \frac{\partial A}{\partial x} + u' \frac{\partial A}{\partial u} \quad (4.7)$$

for some function A that depends only on x, u .

Proof: Referring back to the explicit form of the Euler–Lagrange equation (3.14), the only term involving u'' is the last one, and so if the left hand side is to vanish for all possible functions $u(x)$, the coefficient of u'' must vanish, so

$$\frac{\partial^2 N}{\partial p^2} = 0, \quad \text{which implies} \quad N(x, u, p) = f(x, u) p + g(x, u)$$

for some functions f, g . Substituting the latter expression back into (3.14) yields

$$u' \frac{\partial f}{\partial u} + \frac{\partial g}{\partial u} - \frac{\partial f}{\partial x} - u' \frac{\partial f}{\partial u} = \frac{\partial g}{\partial u} - \frac{\partial f}{\partial x} = 0.$$

Since we are assuming N , and hence f, g are defined for all x, u , a standard result in multivariable calculus, [17, 21], implies $f = \frac{\partial A}{\partial u}$, $g = \frac{\partial A}{\partial x}$, for some function $A(x, u)$. *Q.E.D.*

Remark: Theorem 4.5 is a special case of a general theorem characterizing higher order and multidimensional null Lagrangians; see [20; Theorem 4.7].

[†] More generally, N can be defined on a subset of \mathbb{R}^3 with trivial topology.

With Theorem 4.5 in hand, we can apply the Fundamental Theorem of Calculus to write the functional associated with a null Lagrangian in the following form:

$$J[u] = \int_a^b N(x, u, u') dx = \int_a^b \frac{d}{dx} [A(x, u)] dx = A(b, u(b)) - A(a, u(a)). \quad (4.8)$$

In other words, the functional associated with a null Lagrangian depends only on the value of the function at the endpoints of the interval. This explains why every function solves the Euler–Lagrange equation and is in fact a minimizer. It is because the value of the functional depends only upon the values of $u(x)$ at the endpoints a, b , and not on its behavior for any $a < x < b$. Thus, $J[u]$ achieves the same value, meaning that it is *constant*, whenever $u(x)$ satisfies our usual fixed boundary conditions $u(a) = \alpha$, $u(b) = \beta$, and hence any such function is a minimizer (and maximizer).

General Boundary Conditions

So far we have treated fixed and free/natural boundary conditions, or combinations thereof. The flexibility afforded by null Lagrangians allows us to expand the range of boundary conditions that can be handled by the calculus of variations. Namely, we can modify the original variational problem by adding in a suitable null Lagrangian, which does not alter the Euler–Lagrange equations but does change the associated natural boundary conditions. In other words, if $N = D_x A$ is a null Lagrangian, then the *modified objective functional*

$$\begin{aligned} \widehat{J}[u] &= \int_a^b [L(x, u, u') + N(x, u, u')] dx = \int_a^b \left[L(x, u, u') + \frac{d}{dx} [A(x, u)] \right] dx \\ &= A(b, u(b)) - A(a, u(a)) + \int_a^b L(x, u, u') dx \end{aligned} \quad (4.9)$$

has the same Euler–Lagrange equations as $J[u] = \int_a^b L(x, u, u') dx$. Moreover, when subject to fixed boundary conditions, they have exactly the same critical functions, and hence the same minimizers, since their values differ only by the initial two boundary terms in the final expression, which depend only on the values of u at the endpoints a, b . On the other hand, as we will see, the two variational problems have different natural boundary conditions, and this flexibility allows us to admit any (reasonable) boundary conditions into the variational framework.

To prove this and explain how to construct the required null Lagrangian, let us assume that both boundary conditions explicitly involve the derivative of the function $u(x)$ at the endpoint. We can already treat fixed boundary conditions, and the remaining “mixed” case, in which one end is fixed and the other boundary condition involves the derivative, is left as an exercise for the reader. Under mild assumptions, we can solve the boundary conditions for the derivative:

$$u'(a) = B_1(u(a)), \quad u'(b) = B_2(u(b)), \quad (4.10)$$

for some functions B_1, B_2 depending only on the value of u at the endpoint in question. (They may, of course, depend on the respective endpoints a, b , but this is taken care of by

allowing different functions at each end.) Besides a free end, where $B_i(u) \equiv 0$, another case of interest in a variety of physical situations are the aforementioned Robin boundary conditions

$$u'(a) = \beta_1 u(a) + \gamma_1, \quad u'(b) = \beta_2 u(b) + \gamma_2, \quad (4.11)$$

in which $\beta_1, \beta_2, \gamma_1, \gamma_2$ are prescribed constants.

Given a variational problem with Euler–Lagrange equation supplemented by the prescribed boundary conditions (4.10), let us seek to modify the Lagrangian $L(x, u, u')$ by adding in a suitable null Lagrangian (4.7) in order that the natural boundary conditions associated with the modified Lagrangian

$$\tilde{L}(x, u, p) = L(x, u, p) + N(x, u, p) = L(x, u, p) + p \frac{\partial A}{\partial u}(x, u) + \frac{\partial A}{\partial x}(x, u) \quad (4.12)$$

are equivalent to the given boundary conditions (4.10). Thus, at the right hand endpoint, in view of (4.4), this means

$$0 = \frac{\partial \tilde{L}}{\partial p}(b, u(b), u'(b)) = \frac{\partial L}{\partial p}(b, u(b), u'(b)) + \frac{\partial A}{\partial u}(b, u(b))$$

is equivalent to the boundary equation $u'(b) = B_2(u(b))$. For this to occur, we must require

$$\frac{\partial A}{\partial u}(b, u(b)) = \frac{\partial L}{\partial p}(b, u(b), B_2(u(b))). \quad (4.13)$$

Similarly, at the left hand endpoint, we require

$$\frac{\partial A}{\partial u}(a, u(a)) = \frac{\partial L}{\partial p}(a, u(a), B_1(u(a))). \quad (4.14)$$

These two conditions suffice to prescribe (4.10) as the natural boundary conditions for the variational problem associated with the modified Lagrangian (4.12), thus justifying our claim that by a suitable choice of $A(x, u)$, or, equivalently, by adding in a suitable null Lagrangian $N = D_x A$ we can arrange for any boundary conditions of the above form to be the natural boundary conditions associated with the variational problem.

We can combine the conditions (4.13–14) into a simpler form as follows. Choose an “interpolating function” $B(x, u)$ such that

$$B(a, u(a)) = B_1(u(a)), \quad B(b, u(b)) = B_2(u(b)). \quad (4.15)$$

For example, we can use linear interpolation and set

$$B(x, u) = \frac{x-a}{b-a} B_2(u) - \frac{x-b}{b-a} B_1(u). \quad (4.16)$$

In particular, if $B_1(u) \equiv B_2(u)$, so that we have the same boundary condition at both ends, we can set $B(x, u) = B_1(u) = B_2(u)$. Then (4.13–14) are implied by the interpolated equation

$$\frac{\partial A}{\partial u}(x, u) = - \frac{\partial L}{\partial p}(x, u, B(x, u)), \quad \text{and thus} \quad A(x, u) = - \int \frac{\partial L}{\partial p}(x, u, B(x, u)) du \quad (4.17)$$

is any anti-derivative of the integrand. We have thus proved a general result about variational problems with specified boundary conditions.

Theorem 4.6. *Let $J[u] = \int_a^b L(x, u, u') dx$ be a variational problem whose minimizers are subject to the boundary conditions*

$$u'(a) = B(a, u(a)), \quad u'(b) = B(b, u(b)), \quad (4.18)$$

for some function $B(x, u)$. Let $A(x, u)$ be defined by (4.17). Then the modified variational problem

$$\tilde{J}[u] = \int_a^b [L(x, u, u') + D_x A(x, u)] dx = A(b, u(b)) - A(a, u(a)) + \int_a^b L(x, u, u') dx \quad (4.19)$$

has the same Euler–Lagrange equations and natural boundary conditions given by (4.18).

Observe that the modified variational problem (4.19) differs from the original only through the addition of certain boundary “corrections”. The point is that any solution to the resulting boundary value problem will be a candidate minimizer for the modified variational problem.

Example 4.7. Let us consider the problem of minimizing arc length subject to the Robin boundary conditions (4.11) on the interval $[a, b]$. In accordance with (4.16), set

$$B(x, u) = \beta(x)u + \gamma(x) \quad \text{where} \quad \beta(x) = \frac{x-a}{b-a}\beta_2 - \frac{x-b}{b-a}\beta_1, \quad \gamma(x) = \frac{x-a}{b-a}\gamma_2 - \frac{x-b}{b-a}\gamma_1.$$

Using the Lagrangian (3.15), we compute

$$\frac{\partial L}{\partial p} = \frac{p}{\sqrt{1+p^2}}.$$

Substituting into (4.17), the required function A is obtained by integration:

$$A(x, u) = - \int \frac{\beta(x)u + \gamma(x) du}{\sqrt{1 + [\beta(x)u + \gamma(x)]^2}} = - \frac{\sqrt{1 + [\beta(x)u + \gamma(x)]^2}}{\beta(x)},$$

provided $\beta(x) \neq 0$. The modified variational problem (4.19) can thus be written in the form[†]

$$\begin{aligned} \tilde{J}[u] &= - \frac{\sqrt{1 + [\beta_2 u(b) + \gamma_2]^2}}{\beta_2} + \frac{\sqrt{1 + [\beta_1 u(a) + \gamma_1]^2}}{\beta_1} + \int_a^b \sqrt{1 + u'(x)^2} dx \\ &= - \frac{\sqrt{1 + u'(b)^2}}{\beta_2} + \frac{\sqrt{1 + u'(a)^2}}{\beta_1} + \int_a^b \sqrt{1 + u'(x)^2} dx. \end{aligned} \quad (4.20)$$

[†] We avoid writing out the more complicated integral expression involving the corresponding null Lagrangian.

Thus, while the basic arc length functional does not, in general, admit a minimizer that satisfies the Robin boundary conditions, the modified arc length (4.20), which has the same Euler–Lagrange equation, (usually) does.

Let us solve the Robin boundary value problem for the Euler–Lagrange equation, which, as noted above, is merely $u'' = 0$, the solutions of which are straight lines $u = cx + d$. Substituting into the Robin boundary conditions (4.11) produces

$$c = \beta_1(ca + d) + \gamma_1 = \beta_2(cb + d) + \gamma_2.$$

Thus, if

$$(b - a)\beta_1\beta_2 + \beta_2 - \beta_1 \neq 0,$$

the problem admits a unique solution, while if the left hand side is zero, then there is either a one-parameter family of solutions that all give the same value to the modified variational problem (even though they have differing arc lengths), or there is no solution, depending on the values of γ_1, γ_2 .

5. The Second Variation.

The solutions to the Euler–Lagrange boundary value problem are the critical functions for the variational principle, meaning that they cause the functional gradient to vanish. For finite-dimensional optimization problems, being a critical point is only a necessary condition for minimality. One must impose additional conditions, based on the second derivative of the objective function at the critical point, in order to guarantee that it is a minimum and not a maximum or saddle point. Similarly, in the calculus of variations, the solutions to the Euler–Lagrange equation may also include (local) maxima, as well as other non-extremal critical functions. To distinguish between the possibilities, we need to formulate a second derivative test for the objective functional. In the calculus of variations, the second derivative of a functional is known as its *second variation*, and the goal of this section is to construct and analyze it in its simplest manifestation.

For a finite-dimensional objective function $F(u_1, \dots, u_n)$, the second derivative test was based on the positive definiteness of its Hessian matrix. The justification was based on the second order Taylor expansion of the objective function at the critical point. In an analogous fashion, we expand an objective functional $J[u]$ near the critical function. Consider the scalar function

$$h(\varepsilon) = J[u + \varepsilon v],$$

where the function $v(x)$ represents a variation. The second order Taylor expansion of $h(\varepsilon)$ takes the form

$$h(\varepsilon) = J[u + \varepsilon v] = J[u] + \varepsilon K[u; v] + \frac{1}{2} \varepsilon^2 Q[u; v] + \dots$$

The first order terms are linear in the variation v , and, according to our earlier calculation, given by the inner product

$$h'(0) = K[u; v] = \langle \nabla J[u], v \rangle$$

between the variation and the functional gradient. In particular, if u is a critical function, then the first order terms vanish,

$$K[u; v] = \langle \nabla J[u], v \rangle = 0,$$

for all allowable variations v , meaning those that satisfy the homogeneous boundary conditions. Therefore, the nature of the critical function u — minimum, maximum, or neither — will, in most cases, be determined by the second derivative terms

$$h''(0) = Q[u; v].$$

Now, if u is a minimizer, then $Q[u; v] \geq 0$. Conversely, if $Q[u; v] > 0$ for all $v \neq 0$, i.e., the second variation is positive definite, then the critical function u will be a strict local minimizer. This forms the crux of the second derivative test.

Let us explicitly evaluate the second variational for the simplest functional (3.1). Consider the scalar function

$$h(\varepsilon) = J[u + \varepsilon v] = \int_a^b L(x, u + \varepsilon v, u' + \varepsilon v') dx,$$

whose first derivative $h'(0)$ was already determined in (3.6); here we require the second variation

$$Q[u; v] = h''(0) = \int_a^b [A v^2 + 2B v v' + C (v')^2] dx, \quad (5.1)$$

where the coefficient functions

$$A(x) = \frac{\partial^2 L}{\partial u^2}(x, u, u'), \quad B(x) = \frac{\partial^2 L}{\partial u \partial p}(x, u, u'), \quad C(x) = \frac{\partial^2 L}{\partial p^2}(x, u, u'), \quad (5.2)$$

are found by evaluating certain second order derivatives of the Lagrangian at the critical function $u(x)$. In contrast to the first variation, integration by parts will not eliminate all of the derivatives on v in the quadratic functional (5.1), which causes significant complications in the ensuing analysis.

The second derivative test for a minimizer relies on the positivity of the second variation. So, in order to formulate conditions that the critical function be a minimizer for the functional, we need to establish criteria guaranteeing the positive definiteness of such a quadratic functional, meaning that $Q[u; v] > 0$ for all non-zero allowable variations $v(x) \neq 0$. Clearly, if the integrand is positive definite at each point, so

$$A(x)v^2 + 2B(x)vv' + C(x)(v')^2 > 0 \quad \text{whenever} \quad a < x < b, \quad \text{and} \quad v(x) \neq 0, \quad (5.3)$$

then $Q[u; v] > 0$ is also positive definite.

Example 5.1. For the arc length minimization functional (2.3), the Lagrangian is $L(x, u, p) = \sqrt{1 + p^2}$. To analyze the second variation, we first compute

$$\frac{\partial^2 L}{\partial u^2} = 0, \quad \frac{\partial^2 L}{\partial u \partial p} = 0, \quad \frac{\partial^2 L}{\partial p^2} = \frac{1}{(1 + p^2)^{3/2}}.$$

For the critical straight line function

$$u(x) = \frac{\beta - \alpha}{b - a} (x - a) + \alpha, \quad \text{with} \quad p = u'(x) = \frac{\beta - \alpha}{b - a},$$

we find

$$A(x) = \frac{\partial^2 L}{\partial u^2} = 0, \quad B(x) = \frac{\partial^2 L}{\partial u \partial p} = 0, \quad C(x) = \frac{\partial^2 L}{\partial p^2} = \frac{(b - a)^3}{[(b - a)^2 + (\beta - \alpha)^2]^{3/2}} \equiv c.$$

Therefore, the second variation functional (5.1) is

$$Q[u; v] = \int_a^b c (v')^2 dx,$$

where $c > 0$ is a positive constant. Thus, $Q[u; v] = 0$ vanishes if and only if $v(x)$ is a constant function. But the variation $v(x)$ is required to satisfy the homogeneous boundary conditions $v(a) = v(b) = 0$, and hence $Q[u; v] > 0$ for all allowable nonzero variations. Therefore, we conclude that the straight line is, indeed, a (local) minimizer for the arc length functional. We have at last rigorously justified our intuition that the shortest distance between two points is a straight line!

However, as the following example demonstrates, the pointwise positivity condition (5.3) is overly restrictive.

Example 5.2. Consider the quadratic functional

$$Q[v] = \int_0^1 [(v')^2 - v^2] dx. \tag{5.4}$$

We claim that $Q[v] > 0$ for all nonzero $v \not\equiv 0$ subject to homogeneous Dirichlet boundary conditions $v(0) = 0 = v(1)$. This result is not trivial! Indeed, the boundary conditions play an essential role, since choosing $v(x) \equiv c \neq 0$ to be any constant function will produce a negative value for the functional: $Q[v] = -c^2$.

To prove the claim, consider the quadratic functional

$$\tilde{Q}[v] = \int_0^1 (v' + v \tan x)^2 dx \geq 0,$$

which is clearly non-negative, since the integrand is everywhere ≥ 0 . Moreover, by continuity, the integral vanishes if and only if v satisfies the first order linear ordinary differential equation

$$v' + v \tan x = 0, \quad \text{for all} \quad 0 \leq x \leq 1.$$

The only solution that also satisfies boundary condition $v(0) = 0$ is the trivial one $v \equiv 0$. We conclude that $\tilde{Q}[v] = 0$ if and only if $v \equiv 0$, and hence $\tilde{Q}[v]$ is a positive definite quadratic functional on the space of allowable variations.

Let us expand the latter functional,

$$\begin{aligned}\tilde{Q}[v] &= \int_0^1 [(v')^2 + 2vv' \tan x + v^2 \tan^2 x] dx \\ &= \int_0^1 [(v')^2 - v^2 (\tan x)' + v^2 \tan^2 x] dx = \int_0^1 [(v')^2 - v^2] dx = Q[v].\end{aligned}$$

In the second equality, we integrated the middle term by parts, using $(v^2)' = 2vv'$, and noting that the boundary terms vanish owing to our imposed boundary conditions. Since $\tilde{Q}[v]$ is positive definite, so is $Q[v]$, justifying the previous claim.

To appreciate how delicate this result is, consider the almost identical quadratic functional

$$\hat{Q}[v] = \int_0^4 [(v')^2 - v^2] dx, \quad (5.5)$$

the only difference being the upper limit of the integral. A quick computation shows that the function $v(x) = x(4-x)$ satisfies the boundary conditions $v(0) = 0 = v(4)$, but

$$\hat{Q}[v] = \int_0^4 [(4-2x)^2 - x^2(4-x)^2] dx = -\frac{64}{5} < 0.$$

Therefore, $\hat{Q}[v]$ is *not* positive definite. Our preceding analysis does not apply because the function $\tan x$ becomes singular at $x = \frac{1}{2}\pi$, and so the auxiliary integral $\int_0^4 (v' + v \tan x)^2 dx$ does not converge.

The complete analysis of positive definiteness of quadratic functionals is quite subtle. Indeed, the strange appearance of $\tan x$ in the preceding example turns out to be an important clue! In the interests of brevity, let us just state without proof a fundamental theorem, and refer the interested reader to [9] for full details.

Theorem 5.3. *Let $A(x), B(x), C(x) \in C^0[a, b]$ be continuous functions. The quadratic functional*

$$Q[v] = \int_a^b [Av^2 + 2Bvv' + C(v')^2] dx$$

is positive definite, so $Q[v] > 0$ for all $v \not\equiv 0$ satisfying the homogeneous Dirichlet boundary conditions $v(a) = v(b) = 0$, provided

- (a) $C(x) > 0$ for all $a \leq x \leq b$, and
- (b) For any $a < c \leq b$, the only solution to its linear Euler–Lagrange boundary value problem

$$-(Cv')' + (A - B')v = 0, \quad v(a) = 0 = v(c), \quad (5.6)$$

is the trivial function $v(x) \equiv 0$.

Remark: A value c for which (5.6) has a nontrivial solution is known as a *conjugate point* to a . Thus, condition (b) can be restated that the variational problem has no conjugate points in the interval $(a, b]$.

Example 5.4. The quadratic functional

$$Q[v] = \int_0^b [(v')^2 - v^2] dx \quad (5.7)$$

has Euler–Lagrange equation

$$-v'' - v = 0.$$

The solutions $v(x) = k \sin x$ satisfy the boundary condition $v(0) = 0$. The first conjugate point occurs at $c = \pi$ where $v(\pi) = 0$. Therefore, Theorem 5.3 implies that the quadratic functional (5.7) is positive definite *provided* the upper integration limit $b < \pi$. This explains why the original quadratic functional (5.4) is positive definite, since there are no conjugate points on the interval $[0, 1]$, while the modified version (5.5) is *not*, because the first conjugate point π lies on the interval $(0, 4]$.

In the case when the quadratic functional arises as the second variation of a functional (3.1), the coefficient functions A, B, C are given in terms of the Lagrangian $L(x, u, p)$ by formulae (5.2). In this case, the first condition in Theorem 5.3 requires

$$\frac{\partial^2 L}{\partial p^2}(x, u, u') > 0 \quad (5.8)$$

for the candidate minimizer $u(x)$ that solves the Euler–Lagrange equation. This is known as the *Legendre condition*. The second, *conjugate point condition* requires that the so-called *linearized variational equation*

$$-\frac{d}{dx} \left[\frac{\partial^2 L}{\partial p^2}(x, u, u') \frac{dv}{dx} \right] + \left[\frac{\partial^2 L}{\partial u^2}(x, u, u') - \frac{d}{dx} \frac{\partial^2 L}{\partial u \partial p}(x, u, u') \right] v = 0, \quad (5.9)$$

where $u(x)$ is the candidate minimizer, has no nontrivial solutions $v(x) \neq 0$ that satisfy $v(a) = 0$ and $v(c) = 0$ for $a < c \leq b$. In this way, we have arrived at a rigorous form of the second derivative test for the simplest functional in the calculus of variations.

Theorem 5.5. *If the function $u(x)$ satisfies the Euler–Lagrange equation (3.13), and, in addition, the Legendre condition (5.8) holds and there are no conjugate points on the interval, then $u(x)$ is a strict local minimum for the functional.*

6. Multi-dimensional Variational Problems.

The calculus of variations encompasses a very broad range of mathematical applications. The methods of variational analysis can be applied to an enormous variety of physical systems, whose equilibrium configurations inevitably minimize a suitable functional, which, typically, represents the potential energy of the system. Minimizing configurations appear as critical functions at which the functional gradient vanishes. Following similar computational procedures as in the one-dimensional calculus of variations, we find that the critical functions are characterized as solutions to a system of partial differential equations, known as the Euler–Lagrange equations associated with the variational principle. Each solution to

the boundary value problem specified by the Euler–Lagrange equations subject to appropriate boundary conditions is, thus, a candidate minimizer for the variational problem. In many applications, the Euler–Lagrange boundary value problem suffices to single out the physically relevant solutions, and one need not press on to the considerably more difficult second variation.

Implementation of the variational calculus for functionals in higher dimensions will be illustrated by looking at a specific example — a first order variational problem involving a single scalar function of two variables. Once this is fully understood, generalizations and extensions to higher dimensions and higher order Lagrangians are readily apparent. Thus, we consider an objective functional

$$J[u] = \iint_{\Omega} L(x, y, u, u_x, u_y) \, dx \, dy, \quad (6.1)$$

having the form of a double integral over a prescribed domain $\Omega \subset \mathbb{R}^2$. The *Lagrangian* $L(x, y, u, p, q)$ is assumed to be a sufficiently smooth function of its five arguments. Our goal is to find the function(s) $u = f(x, y)$ that minimize the value of $J[u]$ when subject to a set of prescribed boundary conditions on $\partial\Omega$, the most important being our usual Dirichlet, Neumann, or mixed boundary conditions. For simplicity, we concentrate on the Dirichlet boundary value problem, and require that the minimizer satisfy

$$u(x, y) = g(x, y) \quad \text{for} \quad (x, y) \in \partial\Omega. \quad (6.2)$$

The First Variation and the Euler–Lagrange Equations

The basic necessary condition for an extremum (minimum or maximum) is obtained in precisely the same manner as in the one-dimensional framework. Consider the scalar function

$$h(\varepsilon) \equiv J[u + \varepsilon v] = \iint_{\Omega} L(x, y, u + \varepsilon v, u_x + \varepsilon v_x, u_y + \varepsilon v_y) \, dx \, dy$$

depending on $\varepsilon \in \mathbb{R}$. The *variation* $v(x, y)$ is assumed to satisfy homogeneous Dirichlet boundary conditions

$$v(x, y) = 0 \quad \text{for} \quad (x, y) \in \partial\Omega, \quad (6.3)$$

to ensure that $u + \varepsilon v$ satisfies the same boundary conditions (6.2) as u itself. Under these conditions, if u is a minimizer, then the scalar function $h(\varepsilon)$ will have a minimum at $\varepsilon = 0$, and hence

$$h'(0) = 0.$$

When computing $h'(\varepsilon)$, we assume that the functions involved are sufficiently smooth so as to allow us to bring the derivative inside the integral, and then apply the chain rule. At $\varepsilon = 0$, the result is

$$h'(0) = \frac{d}{d\varepsilon} J[u + \varepsilon v] \Big|_{\varepsilon=0} = \iint_{\Omega} \left(v \frac{\partial L}{\partial u} + v_x \frac{\partial L}{\partial p} + v_y \frac{\partial L}{\partial q} \right) dx \, dy, \quad (6.4)$$

where the derivatives of L are all evaluated at x, y, u, u_x, u_y . To identify the functional gradient, we need to rewrite this integral in the form of an inner product:

$$h'(0) = \langle \nabla J[u], v \rangle = \iint_{\Omega} h(x, y) v(x, y) dx dy, \quad \text{where} \quad h = \nabla J[u].$$

To convert (6.4) into this form, we need to remove the offending derivatives from v . In two dimensions, the requisite integration by parts formula is based on Green's Theorem:

$$\iint_{\Omega} \left(\frac{\partial v}{\partial x} w_1 + \frac{\partial v}{\partial y} w_2 \right) dx dy = \oint_{\partial\Omega} v (-w_2 dx + w_1 dy) - \iint_{\Omega} v \left(\frac{\partial w_1}{\partial x} + \frac{\partial w_2}{\partial y} \right) dx dy, \quad (6.5)$$

in which w_1, w_2 are arbitrary smooth functions. Setting $w_1 = \frac{\partial L}{\partial p}$, $w_2 = \frac{\partial L}{\partial q}$, we find

$$\iint_{\Omega} \left(v_x \frac{\partial L}{\partial p} + v_y \frac{\partial L}{\partial q} \right) dx dy = - \iint_{\Omega} v \left[\frac{\partial}{\partial x} \left(\frac{\partial L}{\partial p} \right) + \frac{\partial}{\partial y} \left(\frac{\partial L}{\partial q} \right) \right] dx dy,$$

where the boundary integral vanishes owing to the boundary conditions (6.3) that we impose on the allowed variations. Substituting this result back into (6.4), we conclude that

$$h'(0) = \iint_{\Omega} v \left[\frac{\partial L}{\partial u} - \frac{\partial}{\partial x} \left(\frac{\partial L}{\partial p} \right) - \frac{\partial}{\partial y} \left(\frac{\partial L}{\partial q} \right) \right] dx dy = \langle \nabla J[u], v \rangle, \quad (6.6)$$

where

$$\nabla J[u] = \frac{\partial L}{\partial u} - \frac{\partial}{\partial x} \left(\frac{\partial L}{\partial p} \right) - \frac{\partial}{\partial y} \left(\frac{\partial L}{\partial q} \right)$$

is the desired first variation or functional gradient. Since the gradient vanishes at a critical function, we conclude that the minimizer $u(x, y)$ must satisfy the *Euler-Lagrange equation*

$$\frac{\partial L}{\partial u}(x, y, u, u_x, u_y) - \frac{\partial}{\partial x} \left(\frac{\partial L}{\partial p}(x, y, u, u_x, u_y) \right) - \frac{\partial}{\partial y} \left(\frac{\partial L}{\partial q}(x, y, u, u_x, u_y) \right) = 0. \quad (6.7)$$

Once we explicitly evaluate the derivatives, the net result is a second order partial differential equation

$$L_u - L_{xp} - L_{yq} - u_x L_{up} - u_y L_{uq} - u_{xx} L_{pp} - 2u_{xy} L_{pq} - u_{yy} L_{qq} = 0, \quad (6.8)$$

where we use subscripts to indicate derivatives of both u and L , the latter being evaluated at x, y, u, u_x, u_y .

Example 6.1. As a first elementary example, consider the Dirichlet minimization problem

$$J[u] = \iint_{\Omega} \frac{1}{2} (u_x^2 + u_y^2) dx dy. \quad (6.9)$$

In this case, the associated Lagrangian is

$$L = \frac{1}{2}(p^2 + q^2), \quad \text{with} \quad \frac{\partial L}{\partial u} = 0, \quad \frac{\partial L}{\partial p} = p = u_x, \quad \frac{\partial L}{\partial q} = q = u_y.$$

Therefore, the Euler–Lagrange equation (6.7) becomes

$$-\frac{\partial}{\partial x}(u_x) - \frac{\partial}{\partial y}(u_y) = -u_{xx} - u_{yy} = -\Delta u = 0,$$

which is the two-dimensional Laplace equation. Subject to the selected boundary conditions, the solutions, i.e., the harmonic functions, are critical functions for the Dirichlet variational principle.

However, the calculus of variations approach, as developed so far, leads to a much weaker result since it only singles out the harmonic functions as *candidates* for minimizing the Dirichlet integral; they could just as easily be maximizing functions or saddle points. When dealing with a quadratic variational problem, the direct algebraic approach is, when applicable, the more powerful, since it assures us that the solutions to the Laplace equation really do minimize the integral among the space of functions satisfying the appropriate boundary conditions. However, the direct method is restricted to quadratic variational problems, whose Euler–Lagrange equations are linear partial differential equations. In nonlinear cases, one really does need to utilize the full power of the variational machinery.

Example 6.2. Let us derive the Euler–Lagrange equation for the minimal surface problem. From (2.9), the surface area integral

$$J[u] = \iint_{\Omega} \sqrt{1 + u_x^2 + u_y^2} \, dx \, dy \quad \text{has Lagrangian} \quad L = \sqrt{1 + p^2 + q^2}.$$

Note that

$$\frac{\partial L}{\partial u} = 0, \quad \frac{\partial L}{\partial p} = \frac{p}{\sqrt{1 + p^2 + q^2}}, \quad \frac{\partial L}{\partial q} = \frac{q}{\sqrt{1 + p^2 + q^2}}.$$

Therefore, replacing $p \rightarrow u_x$ and $q \rightarrow u_y$ and then evaluating the derivatives, the Euler–Lagrange equation (6.7) becomes

$$-\frac{\partial}{\partial x} \frac{u_x}{\sqrt{1 + u_x^2 + u_y^2}} - \frac{\partial}{\partial y} \frac{u_y}{\sqrt{1 + u_x^2 + u_y^2}} = \frac{-(1 + u_y^2)u_{xx} + 2u_x u_y u_{xy} - (1 + u_x^2)u_{yy}}{(1 + u_x^2 + u_y^2)^{3/2}} = 0.$$

Thus, a surface described by the graph of a function $u = f(x, y)$ is a critical function, and hence a candidate for minimizing surface area, provided it satisfies the *minimal surface equation*

$$(1 + u_y^2)u_{xx} - 2u_x u_y u_{xy} + (1 + u_x^2)u_{yy} = 0. \quad (6.10)$$

We are confronted with a complicated, nonlinear, second order partial differential equation, which has been the focus of some of the most sophisticated and deep analysis over the past two centuries, with significant progress on understanding its solution only within the past 70 years. We refer the interested reader to the advanced texts [18, 19] for further developments in this fascinating area of mathematical analysis.

Example 6.3. The deformations of an elastic body $\Omega \subset \mathbb{R}^n$ are described by the *displacement* field, $\mathbf{u}: \Omega \rightarrow \mathbb{R}^n$. Each material point $\mathbf{x} \in \Omega$ in the undeformed body will move to a new position $\mathbf{y} = \mathbf{x} + \mathbf{u}(\mathbf{x})$ in the deformed body

$$\tilde{\Omega} = \{ \mathbf{y} = \mathbf{x} + \mathbf{u}(\mathbf{x}) \mid \mathbf{x} \in \Omega \}.$$

The one-dimensional case governs bars, beams and rods, two-dimensional bodies include thin plates and shells, while $n = 3$ for fully three-dimensional solid bodies. See [1, 11] for details and physical derivations.

For small deformations, we can use a linear theory to approximate the much more complicated equations of nonlinear elasticity. The simplest case is that of a homogeneous and isotropic planar body $\Omega \subset \mathbb{R}^2$. The equilibrium configurations are described by the deformation function $\mathbf{u}(\mathbf{x}) = (u(x, y), v(x, y))$. A detailed physical analysis of the constitutive assumptions leads to a minimization principle based on the following functional:

$$\begin{aligned} J[u, v] &= \iint_{\Omega} \left[\frac{1}{2} \mu \|\nabla \mathbf{u}\|^2 + \frac{1}{2} (\lambda + \mu) (\nabla \cdot \mathbf{u})^2 \right] dx dy \\ &= \iint_{\Omega} \left[\left(\frac{1}{2} \lambda + \mu \right) (u_x^2 + v_y^2) + \frac{1}{2} \mu (u_y^2 + v_x^2) + (\lambda + \mu) u_x v_y \right] dx dy. \end{aligned} \quad (6.11)$$

The parameters λ, μ are known as the *Lamé moduli* of the material, and govern its intrinsic elastic properties. They are measured by performing suitable experiments on a sample of the material. Physically, (6.11) represents the stored (or potential) energy in the body under the prescribed displacement. Nature, as always, seeks the displacement that will minimize the total energy.

To compute the Euler–Lagrange equations, we consider the functional variation

$$h(\varepsilon) = J[u + \varepsilon f, v + \varepsilon g],$$

in which the individual variations f, g are arbitrary functions subject only to the given homogeneous boundary conditions. If u, v minimize J , then $h(\varepsilon)$ has a minimum at $\varepsilon = 0$, and so we are led to compute

$$h'(0) = \langle \nabla J, \mathbf{f} \rangle = \iint_{\Omega} (f \nabla_u J + g \nabla_v J) dx dy,$$

which we write as an inner product (using the standard L^2 inner product between vector fields) between the variation \mathbf{f} and the functional gradient $\nabla J = (\nabla_u J, \nabla_v J)$. For the particular functional (6.11), we find

$$h'(0) = \iint_{\Omega} \left[(\lambda + 2\mu) (u_x f_x + v_y g_y) + \mu (u_y f_y + v_x g_x) + (\lambda + \mu) (u_x g_y + v_y f_x) \right] dx dy.$$

We use the integration by parts formula (6.5) to remove the derivatives from the variations f, g . Discarding the boundary integrals, which are used to prescribe the allowable boundary conditions, we find

$$h'(0) = - \iint_{\Omega} \left(\begin{aligned} & [(\lambda + 2\mu) u_{xx} + \mu u_{yy} + (\lambda + \mu) v_{xy}] f \\ & + [(\lambda + \mu) u_{xy} + \mu v_{xx} + (\lambda + 2\mu) v_{yy}] g \end{aligned} \right) dx dy.$$

The two terms in brackets give the two components of the functional gradient. Setting them equal to zero, we derive the second order linear system of Euler–Lagrange equations

$$\begin{aligned} (\lambda + 2\mu) u_{xx} + \mu u_{yy} + (\lambda + \mu) v_{xy} &= 0, \\ (\lambda + \mu) u_{xy} + \mu v_{xx} + (\lambda + 2\mu) v_{yy} &= 0, \end{aligned} \quad (6.12)$$

known as *Navier's equations*, which can be compactly written as

$$\mu \Delta \mathbf{u} + (\mu + \lambda) \nabla(\nabla \cdot \mathbf{u}) = \mathbf{0} \quad (6.13)$$

for the displacement vector $\mathbf{u} = (u, v)$. The solutions to are the critical displacements that, under appropriate boundary conditions, minimize the potential energy functional.

Since we are dealing with a quadratic functional, a more detailed algebraic analysis will demonstrate that the solutions to Navier's equations are the minimizers for the variational principle (6.11). Although only valid in a limited range of physical and kinematical conditions, the solutions to the planar Navier's equations and its three-dimensional counterpart are successfully used to model a wide class of elastic materials.

In general, the solutions to the Euler–Lagrange boundary value problem are critical functions for the variational problem, and hence include all (smooth) local and global minimizers. Determination of which solutions are genuine minima requires a further analysis of the positivity properties of the second variation, which is beyond the scope of our introductory treatment. Indeed, a complete analysis of the positive definiteness of the second variation of multi-dimensional variational problems is quite complicated, and still awaits a completely satisfactory resolution!

Acknowledgments: It is a pleasure to thank Olivier de La Grandville, Aritra Lahiri, and Maria Clara Nucci for comments, corrections, and inspiration.

References

- [1] Antman, S.S., *Nonlinear Problems of Elasticity*, Appl. Math. Sci., vol. 107, Springer–Verlag, New York, 1995.
- [2] Ball, J.M., and Mizel, V.J., One-dimensional variational problem whose minimizers do not satisfy the Euler-Lagrange equation, *Arch. Rat. Mech. Anal.* **90** (1985), 325–388.
- [3] Born, M., and Wolf, E., *Principles of Optics*, Fourth Edition, Pergamon Press, New York, 1970.
- [4] Boyce, W.E., and DiPrima, R.C., *Elementary Differential Equations and Boundary Value Problems*, 7th ed., John Wiley, New York, 2001.
- [5] do Carmo, M.P., *Differential Geometry of Curves and Surfaces*, Prentice-Hall, Englewood Cliffs, N.J., 1976.
- [6] Courant, R., and Hilbert, D., *Methods of Mathematical Physics*, vol. I, Interscience Publ., New York, 1953.
- [7] Dacorogna, B., *Introduction to the Calculus of Variations*, Imperial College Press, London, 2004.
- [8] Forsyth, A.R., *Calculus of Variations*, Dover Publ., New York, 1960.
- [9] Gel'fand, I.M., and Fomin, S.V., *Calculus of Variations*, Prentice–Hall, Inc., Englewood Cliffs, N.J., 1963.

- [10] Giaquinta, M., and Hildebrandt, S., *Calculus of Variations I. The Lagrangian Formalism*, Springer–Verlag, New York, 1996.
- [11] Gurtin, M.E., *An Introduction to Continuum Mechanics*, Academic Press, New York, 1981.
- [12] Hildebrandt, S., and Tromba, A., *Mathematics and Optimal Form*, Scientific American Books, New York, 1985.
- [13] Kot, M., *A First Course in the Calculus of Variations*, American Mathematical Society, Providence, R.I., 2014.
- [14] de La Grandville, O., On a classic problem in the calculus of variations: setting straight key properties of the catenary, *Amer. Math. Monthly*, to appear.
- [15] MacNeish, H.F., On the determination of a catenary with given directrix and passing through two given points, *Ann. Math.* **7** (1906), 65–71.
- [16] MacNeish, H.F., Concerning the discontinuous solution in the problem of the minimum surface of revolution, *Ann. Math.* **7** (1906), 72–80.
- [17] Marsden, J.E., and Tromba, A.J., *Vector Calculus*, 6th ed., W.H. Freeman, New York, 2012.
- [18] Morgan, F., *Geometric Measure Theory: a Beginner’s Guide*, Academic Press, New York, 2000.
- [19] Nitsche, J.C.C., *Lectures on Minimal Surfaces*, Cambridge University Press, Cambridge, 1988.
- [20] Olver, P.J., *Applications of Lie Groups to Differential Equations*, 2nd ed., Graduate Texts in Mathematics, vol. 107, Springer–Verlag, New York, 1993.
- [21] Olver, P.J., *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics, Springer, New York, 2014.