Elements of calculus of variations with applications in mechanics

Victor A. Eremeev[1]

[1] University of Cagliari, Italy

Computational Mechanics of Structures 2026, Cagliari , Italy

Università degli Studi di Cagliari
DIPARTIMENTO DI INGEGNERIA CIVILE, AMBIENTALE E ARCHITETTURA | DICAAR
Victor A. Eremeev
Professore Ordinario | ICAR/08 - Scienza delle Costruzioni

Contents

Historical Remarks

The history of mechanics, as well as the history of other branches of science, is a history of attempts to explain the world by means of *the smallest possible number of universal laws and general principles*. The most successful and fruitful attempts stem from the idea that the observable events are extreme in their character and that the general principles sought are variational, i.e. they assert that certain parameters obtain their maximum or minimum values in realizable physical processes.

This idea seems to endow Nature with some goal and appeared a long time ago. **Aristotle** (384–322 B.C.) claimed in his Physics, which served as the major source for natural philosophers for over 2000 years, that *in all its manifestations, Nature follows the easiest path that requires the least amount of effort*.

Historical Remarks

- Galileo (1564–1642),
- Descartes (1596–1650),
- Fermat (1601–1665),
- Newton (1643–1727),
- Leibnitz (1646–1716),
- R. Hooke (1635–1703),
- J. Bernoulli (1667–1748)...

For example, **Fermat** suggested that the experimentally observed refraction law corresponds to the principle of minimum time: *light moving from point A to point B chooses the trajectory for which the travel time is minimum.*

See, [V.L. Berdichevsky, Variational Principles of Continuum Mechanics, Volume 1, Springer, 2009].

Optimization in life and science

Optimization is a universal goal. Students would like to learn more, receive better grades, and have more free time; professors (at least some of them) would like to give better lectures, see students learn more, receive higher pay, and have more free time. These are the optimization problems of real life. In mathematics, optimization makes sense only when formulated in terms of a function $f(x)$ or other expression. One then seeks the minimum value of the expression. (It suffices to discuss minimization because maximizing f is equivalent to minimizing $-f$.)

1D Case

Let us recall some terminology for the one-variable case $y = f(x)$.

Definition

The function $f(x)$ has a *local minimum* at a point x_0 if there is a neighborhood $(x_0 - d, x_0 + d)$ in which $f(x) \geq f(x_0)$. We call x_0 the *global minimum* of $f(x)$ on $[a, b]$ if $f(x) \geq f(x_0)$ holds for all $x \in [a, b]$.

The necessary condition for a differentiable function $f(x)$ to have a local minimum at x_0 is

$$f'(x_0) = 0. \tag{1}$$

A simple and convenient sufficient condition is

$$f''(x_0) > 0. \tag{2}$$

Unfortunately, no available criterion for a local minimum is both sufficient and necessary. So the approach is to solve (1) for possible points of local minimum of $f(x)$ and then test these using an available sufficient condition.

A Function in n Variables

Consider the minimization of a function $y = f(\mathbf{x})$ with $\mathbf{x} = (x_1, \ldots, x_n)$. More cannot be expected from this theory than from the theory of functions in a single variable.

Definition

A function $f(\mathbf{x})$ has a *global minimum* at the point $\mathbf{x}^*$ if the inequality

$$f(\mathbf{x}^*) \leq f(\mathbf{x}^* + \mathbf{h}) \tag{3}$$

holds for all nonzero $\mathbf{h} = (h_1, \ldots, h_n) \in \mathbb{R}^n$. The point $\mathbf{x}^*$ is a *local minimum* if there exists $\rho > 0$ such that (3) holds whenever $\|\mathbf{h}\| = (h_1^2 + \cdots + h_n^2)^{1/2} < \rho$.

Necessary Conditions

Let $\mathbf{x}^*$ be a minimum point of a continuously differentiable function $f(\mathbf{x})$. Then $f(x_1, x_2^*, \ldots, x_n^*)$ is a function in one variable x_1 and takes its minimum at x_1^*. It follows that $\partial f / \partial x_1 = 0$ at $x_1 = x_1^*$. Similarly, the rest of the partial derivatives of f are zero at $\mathbf{x}^*$:

$$\left.\frac{\partial f}{\partial x_i}\right|_{\mathbf{x}=\mathbf{x}^*} = 0, \qquad i = 1, \ldots, n. \tag{4}$$

This is a necessary condition of minimum for a continuously differentiable function in n variables at the point $\mathbf{x}^*$.

Sufficient Conditions

To get sufficient conditions we must extend Taylor's formula. Let $f(\mathbf{x})$ possess all continuous derivatives up to order $m \geq 2$ in a ball centered at point $\mathbf{x}$, and suppose $\mathbf{x} + \mathbf{h}$ lies in this ball. Fixing these, we apply Taylor's formula in the variable t to $f(\mathbf{x} + t\mathbf{h})$:

$$f(\mathbf{x} + t\mathbf{h}) = f(\mathbf{x}) + \left.\frac{df(\mathbf{x} + t\mathbf{h})}{dt}\right|_{t=0} t + \frac{1}{2!}\left.\frac{d^2 f(\mathbf{x} + t\mathbf{h})}{dt^2}\right|_{t=0} t^2 + \cdots + \frac{1}{m!}\left.\frac{d^m f(\mathbf{x} + t\mathbf{h})}{dt^m}\right|_{t=0} t^m + o(t^m).$$

The remainder term is for the case when $t \to 0$. From this equality for sufficiently small t, the general Taylor formula can be derived.

Sufficient Conditions

The minimization problem for $f(\mathbf{x})$ is studied using only the first two terms of this formula:

$$f(\mathbf{x}+t\mathbf{h}) = f(\mathbf{x}) + \left.\frac{df(\mathbf{x}+t\mathbf{h})}{dt}\right|_{t=0} t + \frac{1}{2!}\left.\frac{d^2 f(\mathbf{x}+t\mathbf{h})}{dt^2}\right|_{t=0} t^2 + o(t^2). \qquad (5)$$

We calculate $df(\mathbf{x}+t\mathbf{h})/dt$ as a derivative of a composite function:

$$\left.\frac{df(\mathbf{x}+t\mathbf{h})}{dt}\right|_{t=0} = \frac{\partial f(\mathbf{x})}{\partial x_1}h_1 + \frac{\partial f(\mathbf{x})}{\partial x_2}h_2 + \cdots + \frac{\partial f(\mathbf{x})}{\partial x_n}h_n.$$

First and Second Differentials

The first differential is defined as

$$df = \frac{\partial f(\mathbf{x})}{\partial x_1}\, dx_1 + \frac{\partial f(\mathbf{x})}{\partial x_2}\, dx_2 + \cdots + \frac{\partial f(\mathbf{x})}{\partial x_n}\, dx_n. \tag{6}$$

The next term,

$$\left.\frac{d^2 f(\mathbf{x}+t\mathbf{h})}{dt^2}\right|_{t=0} = \sum_{i,j=1}^{n} \frac{\partial^2 f(\mathbf{x})}{\partial x_i\, \partial x_j} h_i h_j,$$

defines the second differential of f:

$$d^2 f = \sum_{i,j=1}^{n} \frac{\partial^2 f(\mathbf{x})}{\partial x_i\, \partial x_j}\, dx_i\, dx_j. \tag{7}$$

Taylor's formula of the second order becomes

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + \sum_{i=1}^{n} \frac{\partial f(\mathbf{x})}{\partial x_i} h_i + \frac{1}{2!} \sum_{i,j=1}^{n} \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} h_i h_j + o(\|\mathbf{h}\|^2). \tag{8}$$

Sufficient Conditions

The necessary condition for a minimum, $df = 0$, follows from (5) or (4). By (5), the condition

$$\left.\frac{d^2 f(\mathbf{x} + t\mathbf{h})}{dt^2}\right|_{t=0} > 0 \text{ for any sufficiently small } \|\mathbf{h}\|$$

suffices for $\mathbf{x}$ to minimize f. The corresponding quadratic form in the variables h_i is

$$\frac{1}{2!}\sum_{i,j=1}^{n} \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} h_i h_j = \frac{1}{2}\begin{pmatrix} h_1 & \cdots & h_n \end{pmatrix} \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}.$$

The $n \times n$ *Hessian matrix* is symmetric under our smoothness assumptions on f. Positive definiteness of the quadratic form can be verified via Sylvester's criterion.

On Global Minimum

The problem of global minimum for a function in many variables on a closed domain Ω is more complicated than the corresponding problem for a function in one variable. Indeed, the set of points satisfying (4) can be infinite for a multivariable function. Trouble also arises concerning the domain boundary $\partial\Omega$: since it is no longer a finite set (unlike $\{a, b\}$) we must also solve the problem of minimum on $\partial\Omega$, and the structure of such a set can be complicated. The algorithm for finding a point of global minimum of a function $f(\mathbf{x})$ cannot be described in several phrases; it depends on the structure of both the function and the domain.

Issues connected with the boundary can be avoided by considering the problem of global minimum of a function on an open domain. We will take this approach when treating the calculus of variations. Although analogous problems with closed domains arise in applications, the difficulties are so great that no general results are applicable to many problems. One must investigate each such problem separately.

Constraints

Constraints of the form

$$g_i(\mathbf{x}) = 0, \qquad i = 1, \dots, m, \tag{9}$$

permit reduction of constrained minimization to an unconstrained problem provided we can solve (9) and get

$$x_k = \psi_k(x_1, \dots, x_{n-m}), \qquad k = n - m + 1, \dots, n.$$

Substitution into $f(\mathbf{x})$ would yield an ordinary unconstrained minimization problem for a function in $n - m$ variables

$$f(x_1, \dots, x_{n-m}, \dots, \psi_n(x_1, \dots, x_{n-m})).$$

The resulting system of equations is nonlinear in general. This situation can be circumvented by the use of *Lagrange multipliers*.

Lagrange multipliers

The method proceeds with formation of the *Lagrangian function*

$$\mathcal{L}(x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_m) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j g_j(\mathbf{x}), \tag{10}$$

by which the constraints g_j are adjoined to f. Then the x_i and λ_i are all treated as independent, unconstrained variables. The resulting necessary conditions form a system of $n+m$ equations in the $n+m$ unknowns x_i, λ_j:

$$\begin{aligned} &\frac{\partial f(\mathbf{x})}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(\mathbf{x})}{\partial x_i} = 0, \qquad i = 1,\ldots,n, \\ &g_j(\mathbf{x}) = 0, \qquad j = 1,\ldots,m. \end{aligned} \tag{11}$$

Notion of Functional

The kind of dependence in which a real number corresponds to another (or to a finite set) is not enough to describe many natural processes.
For example the dependence described by quantities of energy type is such that *a numerical value* E *is uniquely defined by the distribution of fields of parameters characterizing the system.* We call this sort of dependence a *functional*. Of course, in mathematics we must also specify the classes to which the above fields may belong. The notion of functional generalizes that of function so that the minimization problem remains sensible. Hence we come to the object of investigation of our main subject: *the calculus of variations*.

Typical Examples: Membrane

A typical problem involves the total energy functional for an elastic membrane under load $F = F(x,y)$:

$$E(u) = \frac{1}{2}a \iint_S \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 \right] dx\,dy - \iint_S F u\,dx\,dy.$$

Here $u = u(x,y)$ is the deflection of a point (x,y) of the membrane, which occupies a domain S and has tension described by parameter a (we can put $a = 1$ without loss of generality). For a membrane with fixed edge, in equilibrium $E(u)$ takes its minimal value relative to all other *admissible* (or *virtual*) states. (An “admissible" function takes appointed boundary values and is sufficiently smooth, in this case having first and second continuous derivatives in S.)

Typical Examples: Membrane

The equilibrium state is described by Poisson's equation

$$\Delta u = -F. \tag{12}$$

Let us also supply the boundary condition

$$u\Big|_{\partial S} = \phi. \tag{13}$$

The problem of minimizing $E(u)$ over the set of smooth functions satisfying (13) is equivalent to the boundary value problem (12)–(13).

Typical Examples: Maximal Area

Other interesting problems come from geometry. Consider the following *isoperimetric problem*:

> *Of all possible smooth closed curves of unit length in the plane, find the equation of that curve L which encloses the greatest area.*

With $r = r(\phi)$ the polar equation of a curve, we seek to have

$$\int_0^{2\pi} \sqrt{r^2 + \left(\frac{dr}{d\phi}\right)^2}\, d\phi = 1, \qquad \frac{1}{2}\int_0^{2\pi} r^2\, d\phi \to \text{max}.$$

Notice how we denoted the problem of maximization.

Minimization of a simple functional using calculus

Consider a general functional of the form

$$F(y) = \int_a^b f(x, y, y')\, dx, \tag{14}$$

where $y = y(x)$ is smooth.
At this stage we do not stop to formulate strict conditions on the functions involved; we simply assume they have as many continuous derivatives as needed. Nor do we clearly specify the neighborhood of a function for which it is a local minimizer of a functional.

Example

The particular form

$$\int_a^b \sqrt{1 + (y')^2}\, dx$$

yields the length of the plane curve $y = y(x)$ from $(a, y(a))$ to $(b, y(b))$. The obvious minimizer is a straight line $y = kx + d$. Without boundary conditions (i.e., with $y(a)$ or $y(b)$ unspecified), k and d are arbitrary and the solution is not unique. We can impose no more than two restrictions on $y(x)$ at the ends a and b, because $y = kx + d$ has only two indefinite constants. However, the problem without boundary conditions also makes sense; its solution is the set of horizontal segments $y = d$ starting at the vertical line $x = a$ and ending at $x = b$.

Solution via Discretization

We begin by subdividing $[a, b]$ into n partitions each of length

$$h = \frac{b-a}{n}.$$

Denote $x_i = a + ih$ and $y_i = y(x_i)$, so $y_0 = y(a)$ and $y_n = y(b)$. Take an approximate value of $y'(x_i)$ as

$$y'(x_i) \approx \frac{y_{i+1} - y_i}{h}.$$

Approximation of (14) by the Riemann sum

$$\int_a^b f(x, y, y')\, dx \approx h \sum_{k=0}^{n-1} f(x_k, y_k, y'(x_k)) \tag{15}$$

gives

$$\int_a^b f(x, y, y')\, dx \approx h \sum_{k=0}^{n-1} f(x_k, y_k, (y_{k+1} - y_k)/h) = \Phi(y_0, \ldots, y_n).$$

Discretization 2

Since $\Phi(y_0, \ldots, y_n)$ is an ordinary function in $n+1$ independent variables, we set

$$\frac{\partial \Phi(y_0, y_1, \ldots, y_n)}{\partial y_i} = 0, \qquad i = 0, \ldots, n. \tag{16}$$

Again, any function f encountered is assumed to possess all needed derivatives. Henceforth we denote partial derivatives using

$$f_y = \frac{\partial f}{\partial y}, \qquad f_{y'} = \frac{\partial f}{\partial y'}, \qquad f_x = \frac{\partial f}{\partial x}, \tag{17}$$

and the total derivative using

$$\begin{aligned} \frac{df(x, y(x), y'(x))}{dx} &= f_x(x, y(x), y'(x)) + f_y(x, y(x), y'(x))\, y'(x) \\ &\quad + f_{y'}(x, y(x), y'(x))\, y''(x). \end{aligned} \tag{18}$$

Observe that in the notation $f_{y'}$ we regard y' as the name of a simple variable; we temporarily ignore its relation to y and even its status as a function in its own right.

Discretization 3

Consider the structure of (16). The variable y_i appears in the sum only once when $i = 0$ or $i = n$, twice otherwise. In the latter case (16) gives, using the chain rule and omitting the factor h,

$$\frac{f_{y'}\left(x_{i-1}, y_{i-1}, \dfrac{y_i - y_{i-1}}{h}\right)}{h} - \frac{f_{y'}\left(x_i, y_i, \dfrac{y_{i+1} - y_i}{h}\right)}{h} + f_y\left(x_i, y_i, \frac{y_{i+1} - y_i}{h}\right) = 0. \qquad (19)$$

For $i = 0$ the result is

$$h\left[f_y\left(x_0, y_0, \frac{y_1 - y_0}{h}\right) - \frac{f_{y'}\left(x_0, y_0, \dfrac{y_1 - y_0}{h}\right)}{h}\right] = 0$$

or

$$f_{y'}\left(x_0, y_0, \frac{y_1 - y_0}{h}\right) - h f_y\left(x_0, y_0, \frac{y_1 - y_0}{h}\right) = 0. \qquad (20)$$

Discretisations 4

For $i = n$ we obtain

$$f_{y'}\left(x_{n-1}, y_{n-1}, \frac{y_n - y_{n-1}}{h}\right) = 0. \tag{21}$$

In the limit as $h \to 0$, (20) and (21) give, respectively,

$$f_{y'}(x, y(x), y'(x))\big|_{x=a} = 0, \qquad f_{y'}(x, y(x), y'(x))\big|_{x=b} = 0.$$

Finally, considering the first two terms in (19) for $0 < i < n$,

$$-\frac{f_{y'}\left(x_i, y_i, \frac{y_{i+1} - y_i}{h}\right) - f_{y'}\left(x_{i-1}, y_{i-1}, \frac{y_i - y_{i-1}}{h}\right)}{h},$$

we recognize an approximation for the total derivative $-df_{y'}/dx$ at y_{i-1}.

Discretisations 5

Hence (19), after $h \to 0$ in such a way that x_{i-1} remains a fixed value c, reduces to

$$f_y - \frac{d}{dx} f_{y'} = 0 \qquad (22)$$

at $x = c$. A nonuniform partitioning will yield this equation similarly for any $x = c \in (a, b)$. In expanded form (22) is

$$f_y - f_{y'x} - f_{y'y} y' - f_{y'y'} y'' = 0, \qquad x \in (a, b). \qquad (23)$$

The limit passage has given us this second-order ordinary differential equation and two boundary conditions

$$f_{y'}\Big|_{x=a} = 0, \qquad f_{y'}\Big|_{x=b} = 0. \qquad (24)$$

Equations (22) and (24) play the same role for the functional (14) as equations (4) play for a function in many variables. In the absence of boundary conditions on $y(x)$, we get necessarily two boundary conditions for a function on which (14) attains a minimum.

Notation for various types of derivatives

It will be necessary to take derivatives of composite functions. When such functions are integrated by parts, we encounter "*total derivatives*" that must be distinguished from the usual partial derivatives. We denote total derivatives in the same way as ordinary derivatives, using the differential symbol d: therefore d/dx will denote a total derivative with respect to x.
We often denote partial derivatives by subscripts so that $\partial(\cdot)/\partial x$ will be denoted by $(\cdot)_x$ or sometimes $(\cdot)_1$. Let us consider two common cases.

Various types of derivatives: 1.

Suppose

$$f = f(x, y(x), y'(x))$$

so that f depends on x through (1) an independent variable x, and (2) the variables $p = y(x)$ and $q = y'(x)$ that are each functions of x as well. We will denote the partial derivative with respect to x as

$$f_x = \frac{\partial}{\partial x} f(x, p, q) \Big|_{p=y(x),\, q=y'(x)}$$

where, during differentiation, we regard p and q as independent variables. Other partial derivatives are

$$f_y = \frac{\partial}{\partial p} f(x, p, q) \Big|_{p=y(x),\, q=y'(x)}, \qquad f_{y'} = \frac{\partial}{\partial q} f(x, p, q) \Big|_{p=y(x),\, q=y'(x)}.$$

The total derivative with respect to x, denoted d/dx, arises when we differentiate while considering $y(x)$ and $y'(x)$ to be functions of x.

Total Derivative

The total derivative of the partial derivative $f_{y'}$ is, by the chain rule,

$$\frac{d}{dx}f_{y'} \equiv \frac{d}{dx}f_{y'}(x, y(x), y'(x)) = f_{y'x} + f_{y'y}y' + f_{y'y'}y'',$$

where, for example,

$$f_{y'y} = \frac{\partial}{\partial p}\frac{\partial}{\partial q}f(x, p, q)\Big|_{p=y(x),\, q=y'(x)}.$$

Various types of derivatives: 2.

Consider the composite function

$$f = f(x, y, u(x, y), u_x(x, y), u_y(x, y))$$

depending on independent variables x, y and on a function u and its derivatives, which depend on x, y as well. Now we denote

$$p = u(x, y), \qquad q = u_x(x, y), \qquad r = u_y(x, y),$$

where u_x and u_y are partial derivatives with respect to x and y, respectively. Introducing variables p, q, r, we get a function $f = f(x, y, p, q, r)$ in five independent variables.

Various types of derivatives: 2.

The following notations are used for partial derivatives:

$$f_x = \frac{\partial}{\partial x} f(x,y,p,q,r)\Big|_{p=u(x,y),\, q=u_x(x,y),\, r=u_y(x,y)},$$

$$f_y = \frac{\partial}{\partial y} f(x,y,p,q,r)\Big|_{p=u(x,y),\, q=u_x(x,y),\, r=u_y(x,y)},$$

$$f_u = \frac{\partial}{\partial p} f(x,y,p,q,r)\Big|_{p=u(x,y),\, q=u_x(x,y),\, r=u_y(x,y)},$$

$$f_{u_x} = \frac{\partial}{\partial q} f(x,y,p,q,r)\Big|_{p=u(x,y),\, q=u_x(x,y),\, r=u_y(x,y)},$$

and

$$f_{u_y} = \frac{\partial}{\partial r} f(x,y,p,q,r)\Big|_{p=u(x,y),\, q=u_x(x,y),\, r=u_y(x,y)}.$$

Total Derivative

Finally, let us display the notation for the total derivative d/dx of f_{u_x}, where f denotes $f = f(x, y, p, q, r)$:

$$\frac{d}{dx} f_{u_x} = \left(f_{qx} + f_{qp} u_x + f_{qq} u_{xx} + f_{qr} u_{yx} \right) \Big|_{p=u(x,y),\, q=u_x(x,y),\, r=u_y(x,y)},$$

and a similar formula for the total derivative with respect to y:

$$\frac{d}{dy} f_{u_x} = \left(f_{qy} + f_{qp} u_y + f_{qq} u_{xy} + f_{qr} u_{yy} \right) \Big|_{p=u(x,y),\, q=u_x(x,y),\, r=u_y(x,y)}.$$

The formulas for higher derivatives are denoted similarly.

Brief summary of important terms

A *functional* is a correspondence assigning a real number to each function in some class of functions. The calculus of variations is concerned with *variational problems*: i.e., those in which we seek the *extrema* (maxima or minima) of functionals.

An *admissible function* for a given variational problem is a function that satisfies all the constraints of that problem.

A function is *sufficiently smooth* for a particular development if all required actions (e.g., differentiation, integration by parts) are possible and yield results having the properties needed for that development.

Euler's Equation for the Simplest Problem. Functional

We begin with the problem of local minimum of the functional

$$F(y) = \int_a^b f(x, y, y')\,dx \qquad (25)$$

on the set of functions $y = y(x)$ that satisfy the boundary conditions

$$y(a) = c_0, \qquad y(b) = c_1. \qquad (26)$$

The existence of a solution can depend on the properties of this set. We must compare the values of $F(y)$ on all functions y satisfying (26). In view of (22) it is reasonable to seek minimizers that have continuous first and second derivatives on $[a, b]$. How should we specify a neighborhood of a function $y(x)$? Since all admissible functions must satisfy (26), we can consider the set of functions of the form $y(x) + \varphi(x)$ where

$$\varphi(a) = \varphi(b) = 0. \qquad (27)$$

Continuity of a Functional

With the intention of using tools close to those of classical calculus, we first introduce the idea of continuity of a functional with respect to an argument which, in turn, is a function on $[a, b]$. A suitably modified version of the classical definition of function continuity is as follows: given any small $\varepsilon > 0$, there exists a δ-neighborhood of $y(x)$ such that when $y(x) + \varphi(x)$ belongs to this neighborhood we have

$$|F(y + \varphi) - F(y)| < \varepsilon.$$

If the neighborhood of the zero function is specified by the inequality

$$\max_{x \in [a,b]} |\varphi(x)| + \max_{x \in [a,b]} |\varphi'(x)| < \delta, \tag{28}$$

the definition can become workable when $f(x, y, y')$ is continuous in the three independent variables x, y, y'.

Continuity

This is not the only possible definition of a neighborhood; later we shall discuss other possibilities. But one benefit is that the left side of (28) contains the expression usually used to define the norm on the set of all functions continuously differentiable on $[a,b]$:

$$\|\varphi(x)\|_{C^{(1)}(a,b)} = \max_{x\in[a,b]} |\varphi(x)| + \max_{x\in[a,b]} |\varphi'(x)|. \tag{29}$$

Definition

The space $C^{(1)}(a,b)$ is the normed space consisting of the set of all functions $\varphi(x)$ that are continuously differentiable on $[a,b]$, supplied with the norm (29). Its subspace of functions satisfying (27) is denoted $C_0^{(1)}(a,b)$. The set of all functions having k continuous derivatives on $[a,b]$ is denoted $C^{(k)}(a,b)$.

In many books these spaces are denoted by $C^{(k)}([a,b])$ to emphasize that $[a,b]$ is closed.

Norm

To keep our notation reasonable throughout the book, we introduce the convention

In cases where no ambiguity should arise, we typically abbreviate the space designation subscript on a norm symbol. □

For example, the notation $\|\cdot\|_{C^{(1)}(a,b)}$ (where the dot stands for the argument of the norm operation) is shortened to $\|\cdot\|$ in the present section. At times, only some aspect of the full label can be suppressed. For example, we may use the notation $\|\cdot\|_{C^{(1)}}$ if only the domain $[a, b]$ is understood.

Neighborhood and Strict Local Minimum

With this convention in mind let us proceed to

Definition

A δ-*neighborhood* of $y(x)$ of admissible functions is the set of all functions of the form $y(x) + \varphi(x)$ where $\varphi(x)$ is such that $\varphi(x) \in C_0^{(1)}(a, b)$ and $\|\varphi(x)\| < \delta$.

When no boundary conditions are imposed on y, then the definition of δ-neighborhood does not require φ to vanish at the endpoints.

Definition

A function $y(x)$ is a *point of local minimum* of $F(y)$ on the set satisfying (26) if there is a δ-neighborhood of $y(x)$, i.e., a set of functions $z(x)$ such that $z(x) - y(x) \in C_0^{(1)}(a, b)$ and $\|z(x) - y(x)\| < \delta$, in which $F(z) - F(y) \geq 0$. If in a δ-neighborhood the relation $F(z) - F(y) > 0$ holds for all $z(x) \neq y(x)$, then $y(x)$ is a *point of strict local minimum*.

Types of local minimum

We may speak of more than one type of local minimum. According to Definition 5, a function y is a minimum if there is a δ such that

$$F(y+\varphi) - F(y) \geq 0 \text{ whenever } \|\varphi\|_{C_0^{(1)}(a,b)} < \delta.$$

Historically this type of minimum is called “weak" and we shall use only this type and simply call it a minimum. Those who pioneered the calculus of variations also considered “strong" local minima, defining these as values of y for which there is a δ such that $F(y+\varphi) \geq F(y)$ whenever $\varphi(a) = \varphi(b) = 0$ and $\max|\varphi| < \delta$ on $[a,b]$. Here the modified condition on φ permits “strong variations" into consideration: i.e., functions φ for which φ' may be large even though φ itself is small. Note that when we “weaken" the condition on φ by changing the norm from the norm of $C_0^{(1)}(a,b)$ to the norm of $C_0(a,b)$ which contains only φ and not φ', we simultaneously strengthen the statement made regarding y when we assert the inequality $F(y+\varphi) \geq F(y)$.

Derivation of Euler's Equation

Let us turn to a rigorous justification of (22). We restrict the class of possible integrands $f(x, y, z)$ of (25) to the set of functions that are continuous in (x, y, z) when $x \in [a, b]$ and $|y - y(x)| + |z - y'(x)| < \delta$. Suppose the existence of a minimizer $y(x)$ for $F(y)$. Consider $F(y + t\varphi)$ for an arbitrary but fixed $\varphi(x) \in C_0^{(1)}(a, b)$. It is a function in the single variable t, taking its minimum at $t = 0$. If it is differentiable then

$$\left. \frac{dF(y + t\varphi)}{dt} \right|_{t=0} = 0. \tag{30}$$

To justify differentiation under the integral sign, let $f(x, y, y')$ be continuously differentiable in the variables y and y'. In addition let us assume $f(x, y, y')$ is twice continuously differentiable, in any combination of its arguments, in the domain of interest. By the chain rule, (30) yields

$$0 = \frac{d}{dt} \int_a^b f(x, y + t\varphi, y' + t\varphi')\, dx \bigg|_{t=0} = \int_a^b \left[f_y(x, y, y')\varphi + f_{y'}(x, y, y')\varphi' \right] dx. \tag{31}$$

First Variation

Definition

The right member of (31) is denoted $\delta F(y,\varphi)$ and called the *first variation* of the functional (25).

Integration by parts in the second term on the right in (31) gives

$$\int_a^b f_{y'}(x,y,y')\varphi'\,dx = -\int_a^b \varphi\frac{d}{dx}f_{y'}(x,y,y')\,dx$$

where the boundary terms vanish by (27). It follows that

$$\int_a^b \left[f_y(x,y,y') - \frac{d}{dx}f_{y'}(x,y,y')\right]\varphi\,dx = 0. \tag{32}$$

In the integrand we see the left side of (22). To deduce (22) from (32) we need the *fundamental lemma* of the calculus of variations.

Fundamental Lemma

Lemma

Let $g(x)$ *be continuous on* $[a, b]$*, and let*

$$\int_a^b g(x)\varphi(x)\,dx = 0 \tag{33}$$

hold for every function $\varphi(x)$ *that is differentiable on* $[a, b]$ *and vanishes in some neighborhoods of* a *and* b*. Then* $g(x) \equiv 0$*.*

Proof

Proof.

Suppose to the contrary that (33) holds while $g(x_0) \neq 0$ for some $x_0 \in (a, b)$. Without loss of generality we may assume $g(x_0) > 0$. By continuity we have $g(x) > 0$ in a neighborhood $[x_0 - \varepsilon, x_0 + \varepsilon] \subset (a, b)$. It is easy to construct a nonnegative bell-shaped function $\varphi_0(x)$ such that $\varphi_0(x)$ is differentiable, $\varphi_0(x_0) > 0$, and $\varphi_0(x) = 0$ outside $(x_0 - \varepsilon, x_0 + \varepsilon)$:

$$\varphi(x) = \begin{cases} \exp\left(\dfrac{\varepsilon^2}{(x - x_0)^2 - \varepsilon^2}\right), & |x - x_0| < \varepsilon, \\ 0, & |x - x_0| \geq \varepsilon. \end{cases}$$

See Fig. 1. The product $g(x)\varphi_0(x)$ is nonnegative everywhere and positive near x_0. Hence $\int_a^b g(x)\varphi(x)\, dx > 0$, a contradiction. □

Bell-shaped function

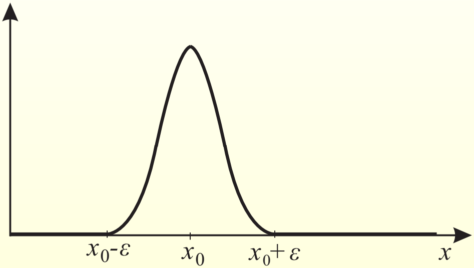

Figure: Bell-shaped function for the proof of Lemma 7.

Another formulation

It is possible to further restrict the class of functions $\varphi(x)$ in Lemma 7.

Lemma

Let $g(x)$ *be continuous on* $[a, b]$*, and let* (33) *hold for any function* $\varphi(x)$ *that is infinitely differentiable on* $[a, b]$ *and vanishes in some neighborhoods of* a *and* b*. Then* $g(x) \equiv 0$*.*

The proof is the same as that for Lemma 7: it is necessary to construct the same bell-shaped function $\varphi(x)$ that is infinitely differentiable. This form of the fundamental lemma provides a basis for the theory of generalized functions or distributions. These are linear functionals on the sets of infinitely differentiable functions, and arise as elements of the Sobolev spaces to be discussed later.

Euler's Equation

Now we can formulate the main result of this section.

Theorem

Suppose $y = y(x) \in C^{(2)}(a,b)$ *locally minimizes the functional* (25) *on the subset of* $C^{(1)}(a,b)$ *consisting of those functions satisfying* (26)*. Then* $y(x)$ *is a solution of the equation*

$$f_y - \frac{d}{dx} f_{y'} = 0. \tag{34}$$

Proof.

Under the assumptions of this section (including that $f(x,y,y')$ is twice continuously differentiable in its arguments), the bracketed term in (32) is continuous on $[a,b]$. Since (32) holds for any $\varphi(x) \in C_0^{(1)}(a,b)$, Lemma 7 applies. □

Euler's Equation

Definition

Equation (34) is known as the *Euler equation*, and a solution $y = y(x)$ is called an *extremal* of (25). A functional is *stationary* if its first variation vanishes.

Taken together,

$$f_y - \frac{d}{dx} f_{y'} = 0.$$

and

$$y(a) = c_0, \qquad y(b) = c_1.$$

constitute a boundary value problem for the unknown $y(x)$.

Example

Find a function $\bar{y} = \bar{y}(x)$ that minimizes the functional

$$F(y) = \int_0^1 [y^2 + (y')^2 - 2y]\, dx$$

subject to the conditions $y(0) = 1$ and $y(1) = 0$.
Here $f(x, y, y') = y^2 + (y')^2 - 2y$, so we obtain

$$f_y = 2y - 2, \qquad f_{y'} = 2y',$$

and the Euler equation is

$$y'' - y + 1 = 0.$$

Subject to the given boundary conditions, the solution is

$$\bar{y}(x) = 1 - \frac{e^x - e^{-x}}{e - e^{-1}}.$$

Example

We stress that this is an extremal: only supplementary investigation can determine whether it is an actual minimizer of $F(y)$. Consider the difference $F(\bar{y}+\varphi) - F(\bar{y})$ where $\varphi(x)$ vanishes at $x = 0, 1$. It is easily shown that

$$F(\bar{y}+\varphi) - F(\bar{y}) = \int_0^1 [\varphi^2 + (\varphi')^2]\, dx \geq 0,$$

so the global minimum of $F(y)$ really does occur at $\bar{y}(x)$. Although such direct verification is not always straightforward, a large class of important problems in mechanics (e.g., problems of equilibrium for linearly elastic structures under conservative loads) yield single extremals that minimize their corresponding total energy functionals. This happens because of the quadratic structure of the functional, as in the present example. □

Remark

We *assumed* the existence of a minimizer. This can lead to incorrect conclusions, and it is normally necessary to prove the existence of an object having needed properties.

Perron's paradox illustrates the trouble we may encounter by supposing the existence of a nonexistent object. Suppose there exists a greatest positive integer N. Since N^2 is also a positive integer we must have $N^2 \leq N$, from which it follows that $N = 1$. If we knew nothing about the integers we might believe this result and attempt to base an entire theory on it. □

Properties of Extremals

While attempting to seek a minimizer on a subset of $C^{(1)}(a,b)$, we imposed the illogical restriction that it must belong to $C^{(2)}(a,b)$ (note that f does not depend on y''). Let us consider how to circumvent this requirement.

Lemma

Let $g(x)$ be a continuous function on $[a,b]$ for which the following equality holds for every $\varphi(x) \in C_0^{(1)}(a,b)$:

$$\int_a^b g(x)\varphi'(x)\,dx = 0. \tag{35}$$

Then $g(x)$ is constant.

Proof.

For a constant c it is clear that $\int_a^b c\varphi'(x)\,dx = 0$ whenever $\varphi(x) \in C_0^{(1)}(a,b)$. So $g(x)$ can be an arbitrary constant. We show that there are no other forms for g. From (35) it follows that

$$\int_a^b [g(x) - c]\varphi'(x)\,dx = 0. \tag{36}$$

Take $c = c_0 = (b-a)^{-1}\int_a^b g(x)\,dx$. The function $\varphi(x) = \int_a^x [g(s) - c_0]\,ds$ is continuously differentiable and satisfies $\varphi(a) = \varphi(b) = 0$. Hence we can put it into (36) and obtain

$$\int_a^b [g(x) - c_0]^2\,dx = 0,$$

from which $g(x) \equiv c$. $\square$

Properties of Extremals

Lemma 11 provides a necessary condition for a relative minimum.

Theorem

Suppose $y = y(x) \in C^{(1)}(a,b)$ *locally minimizes* (25) *on the subset of functions in* $C^{(1)}(a,b)$ *satisfying* (26). *Then* $y(x)$ *is a solution of the following equation, where* c *is a constant:*

$$\int_0^x f_y(s, y(s), y'(s))\, ds - f_{y'}(x, y(x), y'(x)) = c. \tag{37}$$

Proof.

Let us return to the equality (31),

$$\int_a^b \left[f_y(x,y,y')\varphi + f_{y'}(x,y,y')\varphi'\right]\, dx = 0,$$

which is valid here as well. Integration by parts gives

$$\int_a^b f_y(x,y(x),y'(x))\varphi(x)\, dx = -\int_a^b \int_a^x f_y(s,y(s),y'(s))\, ds\, \varphi'(x)\, dx.$$

The boundary terms were zero by (27). It follows that

$$\int_a^b \left[-\int_a^x f_y(s,y(s),y'(s))\, ds + f_{y'}(x,y(x),y'(x))\right] \varphi'(x)\, dx = 0.$$

This holds for all $\varphi(x) \in C_0^{(1)}(a,b)$. So by Lemma 11 we have (37). □

Properties of Extremals

The integro-differential equation (37) has been called the *Euler equation in integrated form*.

Corollary

If

$$f_{y'y'}(x, y(x), y'(x)) \neq 0$$

along a minimizer $y = y(x) \in C^{(1)}(a, b)$ *of* (25), *then* $y(x) \in C^{(2)}(a, b)$.

Proof.

Rewrite (37) as

$$f_{y'}(x, y(x), y'(x)) = \int_0^x f_y(s, y(s), y'(s))\, ds - c.$$

The function on the right is continuously differentiable for any $y = y(x) \in C^{(1)}(a, b)$. Thus we can differentiate both sides of the last identity with respect to x and obtain

$$f_{y'x} + f_{y'y}y' + f_{y'y'}y'' = \text{a continuous function.}$$

Considering the term with $y''(x)$ on the left, we prove the claim. □

Ritz's Method

We now consider a numerical approach to minimizing the functional (25) with boundary conditions (26). In Ritz's method we seek a solution to the problem of minimization of the functional (25), with boundary conditions (26), in the form

$$y_n(x) = \varphi_0(x) + \sum_{k=1}^{n} c_k \varphi_k(x). \tag{38}$$

Here $\varphi_0(x)$ satisfies (26); a common choice is the linear function $\varphi_0(x) = \alpha x + \beta$ with

$$\alpha = \frac{d_1 - d_0}{b - a}, \qquad \beta = \frac{b d_0 - a d_1}{b - a}. \tag{39}$$

The remaining functions, called *basis functions*, satisfy the homogeneous conditions

$$\varphi_k(a) = \varphi_k(b) = 0, \qquad k = 1, \ldots, n.$$

The c_k are constants.

Ritz approximation

Definition

The function $y_n^*(x)$ that minimizes (25) on the set of all functions of the form (38) is called the n*th Ritz approximation*.

The Ritz approximations satisfy the boundary conditions (26) automatically. The above mentioned subspace is the space of functions of the form $\sum_{k=0}^{n} c_k\varphi_k(x)$. For a numerical solution it is necessary that the $\varphi_k(x)$ be linearly independent, which means that

$$\sum_{k=1}^{n} c_k\varphi_k(x) = 0 \quad \text{only if } c_k = 0 \text{ for } k = 1, \dots, n.$$

For manual calculation this was supplemented by the requirement that a small value of n — say 1, 2, or 3 at most — would suffice. The requirement could be met since the corresponding boundary value problems described real objects, such as bent beams, whose shapes under load were understood.

Complete System

Now, to provide a theoretical justification of the method, we require that the system $\{\varphi_k(x)\}_{k=1}^{\infty}$ be *complete*. This means that given any $y = g(x) \in C_0^{(1)}(a, b)$ and $\varepsilon > 0$ we can find a finite sum $\sum_{k=1}^{n} c_k \varphi_k(x)$ such that

$$\left\| g(x) - \sum_{k=1}^{n} c_k \varphi_k(x) \right\| < \varepsilon.$$

Here the norm is defined by (29). It is sometimes required that $\{\varphi_k(x)\}_{k=1}^{\infty}$ be a basis of the corresponding space, but this is not needed for either the justification of the method or its numerical realization.

Minimization

We therefore arrive at the problem of minimizing the functional

$$\int_a^b f(x, y_n, y_n')\, dx$$

where $y_n(x)$ is given by (38). The unknowns are the c_k, so the functional becomes a function in n real variables:

$$\Phi(c_1, \ldots, c_n) = \int_a^b f(x, y_n, y_n')\, dx.$$

To minimize this we solve the system

$$\frac{\partial \Phi(c_1, \ldots, c_n)}{\partial c_k} = 0, \qquad k = 1, \ldots, n. \tag{40}$$

Minimization

Denoting $c_0 = 1$, we have

$$
\begin{aligned}
\frac{\partial \Phi(c_1, \ldots, c_n)}{\partial c_k} &= \frac{\partial}{\partial c_k} \int_a^b f(x, y_n, y_n')\, dx \\
&= \frac{\partial}{\partial c_k} \int_a^b f\left(x, \sum_{i=0}^n c_i \varphi_i(x), \sum_{i=0}^n c_i \varphi_i'(x)\right) dx \\
&= \int_a^b f_y\left(x, \sum_{i=0}^n c_i \varphi_i(x), \sum_{i=0}^n c_i \varphi_i'(x)\right) \varphi_k(x)\, dx \\
&\quad + \int_a^b f_{y'}\left(x, \sum_{i=0}^n c_i \varphi_i(x), \sum_{i=0}^n c_i \varphi_i'(x)\right) \varphi_k'(x)\, dx,
\end{aligned}
$$

Minimization

hence (40) becomes

$$\int_a^b f_y\left(x, \sum_{i=0}^{n} c_i\varphi_i(x), \sum_{i=0}^{n} c_i\varphi_i'(x)\right)\varphi_k(x)\,dx + \int_a^b f_{y'}\left(x, \sum_{i=0}^{n} c_i\varphi_i(x), \sum_{i=0}^{n} c_i\varphi_i'(x)\right)\varphi_k'(x)\,dx = 0 \tag{41}$$

for $k = 1, \ldots, n$. This is a system of n simultaneous equations in the n variables $c_1, \ldots, c_n$. It is linear only if f is a quadratic form in c_k; i.e., only if the Euler equation is linear in $y(x)$. For methods of solving simultaneous equations, the reader is referred to books on numerical analysis.

Example

Consider the problem

$$\Psi(y) = \int_0^1 \{y'^2(x) + [1 + 0.1\sin(x)]y^2(x) - 2xy(x)\}\,dx \to \min$$

subject to $y(0) = 0$ and $y(1) = 10$. Find the Ritz approximations for $n = 1, 3, 5$ using $\varphi_0(x) = 10x$ and the following basis sets:

(a) $\varphi_k(x) = (1 - x)x^k$, $k \geq 1$,

(b) $\varphi_k(x) = \sin k\pi x$, $k \geq 1$.

Note that $\varphi_0(x)$ was chosen to satisfy the given boundary conditions.

Example

We find the expansion coefficients c_k by solving the system

$$\frac{\partial}{\partial c_k}\Psi\left(\varphi_0(x)+\sum_{i=1}^{n}c_i\varphi_i(x)\right)=0, \qquad i=1,\ldots,n.$$

For brevity let us denote

$$\langle y,z\rangle=\int_0^1\left\{y'(x)z'(x)+[1+0.1\sin(x)]y(x)z(x)\right\}dx$$

so that

$$\Psi(y)=\langle y,y\rangle-2\int_0^1 xy(x)\,dx.$$

Example

Using the symmetry of the form $\langle y, z\rangle$ we write out Ritz's equations:

$$
\begin{aligned}
c_1\langle\varphi_1,\varphi_1\rangle + c_2\langle\varphi_2,\varphi_1\rangle + \cdots + c_n\langle\varphi_n,\varphi_1\rangle &= -\langle\varphi_0,\varphi_1\rangle + \int_0^1 x\varphi_1(x)\,dx,\\
c_1\langle\varphi_1,\varphi_2\rangle + c_2\langle\varphi_2,\varphi_2\rangle + \cdots + c_n\langle\varphi_n,\varphi_2\rangle &= -\langle\varphi_0,\varphi_2\rangle + \int_0^1 x\varphi_2(x)\,dx,\\
&\vdots\\
c_1\langle\varphi_1,\varphi_n\rangle + c_2\langle\varphi_2,\varphi_n\rangle + \cdots + c_n\langle\varphi_n,\varphi_n\rangle &= -\langle\varphi_0,\varphi_n\rangle + \int_0^1 x\varphi_n(x)\,dx.
\end{aligned}
\tag{42}
$$

For small n this system can be solved by hand, otherwise computer solution is required.

Example

In the present case we find that for the first basis set the Ritz approximations are

$$y_1(x) = 10x - 2.162x(1-x),$$
$$y_3(x) = 10x + (-1.409x - 1.356x^2 - 0.246x^3)(1-x),$$
$$y_5(x) = 10x + (-1.404x - 1.404x^2 - 0.140x^3 - 0.063x^4 - 0.007x^5)(1-x).$$

For the second basis set we obtain the Ritz approximations

$$z_1(x) = 10x - 0.289\sin\pi x,$$
$$z_3(x) = 10x - 0.289\sin\pi x + 0.063\sin 2\pi x - 0.017\sin 3\pi x,$$
$$z_5(x) = 10x - 0.289\sin\pi x + 0.063\sin 2\pi x - 0.017\sin 3\pi x + 0.008\sin 4\pi x - 0.004\sin 5\pi x,$$

as required. $\square$

Notes on basis functions

For each type of approximation, if we appoint $\varepsilon = 0.01$ then we can stop at $k = 5$. Calculation out to $k = 10$ shows that the $k = 5$ approximations are both very good. However, they do differ from each other by a maximum of about 0.25. So which is “more" correct? We can answer this by substitution into the functional, which gives $\Psi(y_5) \approx 127.046$ and $\Psi(z_5) \approx 127.449$. This is evidence that polynomial approximation is preferable. It is not hard to see why: the true solution is not oscillatory, so the oscillatory behavior of the trigonometric polynomials is not helpful in this case. So the “practical" approach to terminating the numerical process may not work well for trigonometric approximation. In this particular example it can be shown that the trigonometric approximations do converge, but slowly.

Natural Boundary Conditions

Considering the problem of minimum of the functional (25) without boundary conditions (“with free boundary”) we obtain the Euler equation and some boundary conditions. They are known as *natural boundary conditions*.

Consider the minimization of (25) when there are no restrictions on the boundary for $y = y(x)$.

Theorem

Let $y = y(x) \in C^{(2)}(a,b)$ *be a minimizer of the functional* $\int_a^b f(x,y,y')\,dx$ *over the space* $C^{(1)}(a,b)$. *Then for* $y = y(x)$ *the Euler equation*

$$f_y - \frac{d}{dx} f_{y'} = 0 \qquad \text{for all } x \in (a,b) \tag{43}$$

holds along with the natural boundary conditions

$$f_{y'}\Big|_{x=a} = 0, \qquad f_{y'}\Big|_{x=b} = 0. \tag{44}$$

Proof.

We can repeat the initial steps. Namely, consider the values of the functional on the bundle of functions $y = y(x) + t\varphi(x)$ where $\varphi(x) \in C^{(1)}(a, b)$ is arbitrary but fixed. Here, however, there are no restrictions on $\varphi(x)$ at the endpoints of $[a, b]$.

For fixed $y(x)$ and $\varphi(x)$ the functional $\int_a^b f(x, y + t\varphi, y' + t\varphi')\, dx$ becomes a function of the real variable t, and attains its minimum at $t = 0$.

Differentiating with respect to t we get

$$\int_a^b \left[f_y(x, y, y')\varphi + f_{y'}(x, y, y')\varphi'\right]\, dx = 0.$$

□

Proof.

Integration by parts gives

$$\int_a^b \left[f_y(x,y,y') - \frac{d}{dx} f_{y'}(x,y,y')\right] \varphi\, dx + f_{y'}(x,y(x),y'(x))\varphi(x)\Big|_{x=a}^{x=b} = 0. \quad (45)$$

From this we shall derive the Euler equation for $y(x)$ and the natural boundary conditions. The procedure is as follows. We limit the set of all continuously differentiable functions $\varphi(x)$ to those satisfying $\varphi(a) = \varphi(b) = 0$. $\square$

Proof.

For these functions we have

$$\int_a^b \left[f_y(x,y,y') - \frac{d}{dx} f_{y'}(x,y,y') \right] \varphi \, dx = 0. \tag{46}$$

This equation holds for all functions $\varphi(x)$ that participate in the formulation of Lemma 7. Hence the continuous multiplier of $\varphi(x)$ in the integrand of (46) is zero, and the Euler equation (43) holds in (a,b). Now let us return to (45). The equality (46), because of the Euler equation, holds for all $\varphi(x)$. From (45) it follows that

$$f_{y'}(x, y(x), y'(x)) \varphi(x) \Big|_{x=a}^{x=b} = 0 \tag{47}$$

for any $\varphi(x)$. Taking $\varphi(x) = x - b$ we find that $f_{y'}|_{x=a} = 0$; taking $\varphi(x) = x - a$ we find that $f_{y'}|_{x=b} = 0$. $\square$

Remark

Natural boundary conditions are of great importance in mathematical physics. For some models of real bodies or processes it may be unclear which (and how many) boundary conditions are necessary for well-posedness of the problem. The variational approach usually clarifies the situation and provides natural boundary conditions dictated by the nature of the problem. The bending of a plate is a famous example. For her pioneering studies of this problem *Sophie Germain* received a prize from the French Academy of Sciences. She derived the biharmonic equation for the deflections of the midsurface of the plate, but with three boundary conditions as seemed to be in accordance with mechanical intuition; variational considerations later demonstrated that only two were independent.

Remark 2

Earlier we discussed the question of which boundary conditions can be imposed to get a well-posed boundary value problem for minimizing the functional (25). General considerations are nice; however, consider the minimization of

$$\int_0^1 y'^2\,dx \tag{48}$$

on the set of continuously differentiable functions. Its Euler equation is $y'' = 0$, thus all the extremals take the form

$$y = kx + b.$$

The natural boundary conditions are $y'(0) = 0$, $y'(1) = 0$. These imply $k = 0$. So the problem of minimum of (48) (with natural boundary conditions) has a family of solutions $y = b$ with arbitrary constant b. Thus we may impose an additional condition, say $y(0) = 2$. But in general, such a third condition for an ordinary differential equation of second order can yield a boundary value problem that has no solution.

Extensions to More General Functionals

Let us consider two extensions of the above results.

The functional $\int_a^b f(x, \mathbf{y}, \mathbf{y}') \, dx$

Let us replace $y(x)$ in (25) by a vector function

$$\mathbf{y}(x) = (y_1(x), \ldots, y_n(x)).$$

We denote the integrand of the functional as

$$f(x, \mathbf{y}(x), \mathbf{y}'(x)) \quad \text{or} \quad f(x, y_1(x), \ldots, y_n(x), y_1'(x), \ldots, y_n'(x))$$

interchangeably. The task is to treat functionals of the form

$$F(\mathbf{y}) = \int_a^b f(x, \mathbf{y}, \mathbf{y}') \, dx. \tag{49}$$

The functional $\int_a^b f(x, \mathbf{y}, \mathbf{y}')\, dx$

First consider the problem of minimizing (49) when $\mathbf{y}(x)$ takes boundary values

$$\mathbf{y}(a) = \mathbf{c}_0, \qquad \mathbf{y}(b) = \mathbf{c}_1, \tag{50}$$

with vector constants $\mathbf{c}_0 = (c_{01}, c_{02}, \ldots, c_{0n})$, $\mathbf{c}_1 = (c_{11}, c_{12}, \ldots, c_{1n})$. We take $\mathbf{y}(x) \in C^{(k)}(a, b)$ to mean that each coordinate function $y_i(x) \in C^{(k)}(a, b)$; that is, each $y_i(x)$ possesses all derivatives up to order k and these are all continuous on $[a, b]$. Imposing the norm

$$\|\mathbf{y}(x)\|_{C^{(k)}(a,b)} = \sum_{i=1}^{n} \|y_i(x)\|_{C^{(k)}(a,b)} \tag{51}$$

on $C^{(k)}(a, b)$, we can define ε-neighborhoods as needed to describe minimizers of (49).

We seek a minimizer $\mathbf{y}(x)$ of (49) from among all vector functions belonging to $C^{(1)}(a, b)$ and satisfying (50).

The functional $\int_a^b f(x, \mathbf{y}, \mathbf{y}')\, dx$

Theorem

Suppose $\mathbf{y}(x) \in C^{(2)}(a, b)$ *locally minimizes the functional* $\int_a^b f(x, \mathbf{y}, \mathbf{y}')\, dx$ *on the subset of vector functions of* $C^{(1)}(a, b)$ *satisfying* (50). *Then* $\mathbf{y}(x)$ *satisfies*

$$\nabla_{\mathbf{y}} f - \frac{d}{dx} \nabla_{\mathbf{y}'} f = 0. \tag{52}$$

Here we use the gradient notation

$$\nabla_{\mathbf{y}} = \left(\frac{\partial}{\partial y_1}, \ldots, \frac{\partial}{\partial y_n} \right), \qquad \nabla_{\mathbf{y}'} = \left(\frac{\partial}{\partial y'_1}, \ldots, \frac{\partial}{\partial y'_n} \right).$$

Eq. (52) *can be written as* n *scalar equations each having the form of the Euler equation*

$$f_{y_i} - \frac{d}{dx} f_{y'_i} = 0, \qquad i = 1, \ldots, n. \tag{53}$$

Proof.

Over the same construction of admissible functions, $\mathbf{y}(x) + t\varphi(x)$ where $\varphi(a) = \varphi(b) = 0$, we consider (49):

$$F(\mathbf{y}(x) + t\boldsymbol{\varphi}(x)) = \int_a^b f(x, \mathbf{y} + t\boldsymbol{\varphi}, \mathbf{y}' + t\boldsymbol{\varphi}')\, dx. \tag{54}$$

For fixed $\mathbf{y}(x)$ and $\varphi(x)$ this becomes a function of the real variable t and takes its minimum at $t = 0$ for any $\varphi(x)$. Take $\varphi(x)$ of the special form $\boldsymbol{\varphi}_1(x) = (\varphi(x), 0, \ldots, 0)$ where the only nonzero component stands in the first position. Then (54) becomes

$$F(\mathbf{y}(x) + t\boldsymbol{\varphi}_1(x)) = \int_a^b f(x, y_1(x) + t\varphi(x), y_2(x), \ldots, y_n(x), \\ y_1'(x) + t\varphi'(x), y_2'(x), \ldots, y_n'(x))\, dx. \tag{55}$$

□

Proof.

Now the function of t becomes a particular case of the function of the form $F(y(x) + t\varphi(x))$, with the evident notational change $y \mapsto y_1$. A consequence of the minimum of (55) at $t = 0$ is the corresponding Euler equation

$$f_{y_1} - \frac{d}{dx} f_{y_1'} = 0.$$

This is the first equation of (53). Similarly, the ith equation of (53) is derived by taking $\varphi(x)$ in the form $\varphi_1(x) = (0, \ldots, \varphi_i(x), \ldots, 0)$, where the only nonzero component stands in the ith position. □

Natural Boundary Conditions

Let us derive the natural boundary conditions for (49). Now we should not impose any conditions for $\mathbf{y}$ at points $x = a$ and $x = b$ in advance, and thus it is the same for φ at these points. For a moment consider all components of the minimizer $\mathbf{y}(x)$ other than $y_i(x)$ to be given. Then (49) can be formally considered as a particular case of (25) with respect to the ordinary function $y = y_i(x)$. Admissible vector functions differ from $\mathbf{y}(x)$ only in the ith component: $\boldsymbol{\varphi}(x) = \boldsymbol{\varphi}_i(x) = (0, \ldots, \varphi(x), \ldots, 0)$. Thus considering the problem of minimum of (49) without boundary restrictions, we get n pairs of boundary conditions:

$$f_{y_i'}\Big|_{x=a} = 0, \qquad f_{y_i'}\Big|_{x=b} = 0, \qquad\qquad i = 1, \ldots, n.$$

These are natural boundary conditions for a minimizer.

The functional $\int_a^b f(x, y, y', \ldots, y^{(n)})\, dx$

The functional

$$F_n(y) = \int_a^b f(x, y, y', \ldots, y^{(n)})\, dx \tag{56}$$

may be considered on the set of functions satisfying certain boundary conditions. Alternatively, we may impose no boundary conditions and seek natural boundary conditions.

First consider the problem with given boundary equations. The corresponding Euler equation will have order $2n$, hence we take n conditions at each endpoint:

$$\begin{aligned} y(a) &= c_0^*, & y(b) &= c_0^{**}, \\ y'(a) &= c_1^*, & y'(b) &= c_1^{**}, \\ &\vdots & &\vdots \\ y^{(n-1)}(a) &= c_{n-1}^*, & y^{(n-1)}(b) &= c_{n-1}^{**}. \end{aligned} \tag{57}$$

A sufficiently smooth integrand $f(x, y, y', \ldots, y^{(n)})$ belongs to $C^{(n)}$ on the domain of all of its variables.

The functional $\int_a^b f(x, y, y', \dots, y^{(n)})\, dx$

Theorem

Suppose $y(x) \in C^{(2n)}(a, b)$ *locally minimizes* $F_n(y)$ *in* (56) *on the subset of vector functions of* $C^{(n)}(a, b)$ *satisfying* (57). *Then* $y(x)$ *satisfies the* ***Euler–Lagrange equation***

$$f_y - \frac{d}{dx} f_{y'} + \frac{d^2}{dx^2} f_{y''} - \cdots + (-1)^n \frac{d^n}{dx^n} f_{y^{(n)}} = 0. \tag{58}$$

The functional $\int_a^b f(x, y, y', \ldots, y^{(n)})\, dx$

Proof.

Let us recall what it means for $y(x)$ to be a local minimizer of $F_n(y)$. Consider the bundle of functions $y(x) + \varphi(x)$ where $\varphi(x)$ is arbitrary and belongs to $C^{(n)}(a, b)$. Because the bundle must satisfy (57) for any $\varphi(x)$, we see that $\varphi(x)$ must satisfy the homogeneous conditions

$$\begin{aligned} \varphi(a) &= 0, & \varphi(b) &= 0, \\ \varphi'(a) &= 0, & \varphi'(b) &= 0, \\ &\vdots & &\vdots \\ \varphi^{(n-1)}(a) &= 0, & \varphi^{(n-1)}(b) &= 0. \end{aligned} \tag{59}$$

□

The functional $\int_a^b f(x, y, y', \ldots, y^{(n)})\, dx$

Proof.

Let $C_0^{(n)}(a,b)$ denote the subspace of $C^{(n)}(a,b)$ containing functions $\varphi(x)$ that satisfy (59). A function $y(x) \in C^{(n)}(a,b)$ satisfying (57) is a local minimizer of $F_n(y)$ if $F_n(y+\varphi) \geq F_n(y)$ for any $\varphi(x) \in C_0^{(n)}(a,b)$ such that $\|\varphi\|_{C^{(n)}(a,b)} < \varepsilon$ for some $\varepsilon > 0$.
As usual we introduce the parameter t and consider the values of $F_n(y)$ on the bundle $y(x) + t\varphi(x)$. Considering $F_n(y(x) + t\varphi(x))$ for a momentarily fixed $\varphi(x)$ as a function of t, we see that it takes its minimal value at $t = 0$ and thus

$$\left. \frac{dF_n(y(x) + t\varphi(x))}{dt} \right|_{t=0} = 0.$$

□

The functional $\int_a^b f(x, y, y', \dots, y^{(n)})\, dx$

Proof.

In detail,

$$\left.\frac{dF_n(y(x) + t\varphi(x))}{dt}\right|_{t=0}$$
$$= \left.\frac{d}{dt}\int_a^b f(x, y + t\varphi, y' + t\varphi', y'' + t\varphi'', \dots, y^{(n)} + t\varphi^{(n)})\, dx\right|_{t=0}$$
$$= \int_a^b \left(f_y\varphi + f_{y'}\varphi' + f_{y''}\varphi'' + \cdots + f_{y^{(n)}}\varphi^{(n)}\right) dx \qquad (60)$$

(in the last line of the formula the arguments are $f = f(x, y, y', \dots, y^{(n)})$). Now we apply (multiple) integration by parts to each term containing derivatives of φ so that on the last step the integrand contains only φ.
□

The functional $\int_a^b f(x, y, y', \dots, y^{(n)})\,dx$

Proof.

For the term $\int_a^b f_{y'}\varphi'\,dx$ we already have (45). For the term $\int_a^b f_{y''}\varphi''\,dx$ we produce

$$\begin{aligned}\int_a^b f_{y''}\varphi''\,dx &= -\int_a^b \varphi' \frac{d}{dx} f_{y''}\,dx + \varphi' f_{y''}\Big|_{x=a}^{x=b} \\ &= \int_a^b \varphi \frac{d^2}{dx^2} f_{y''}\,dx + \left(\varphi' f_{y''} - \varphi \frac{d}{dx} f_{y''}\right)\Big|_{x=a}^{x=b}.\end{aligned}$$

Similarly

$$\begin{aligned}\int_a^b f_{y'''}\varphi'''\,dx = &-\int_a^b \varphi \frac{d^3}{dx^3} f_{y'''}\,dx \\ &+ \left(\varphi'' f_{y'''} - \varphi' \frac{d}{dx} f_{y'''} + \varphi \frac{d^2}{dx^2} f_{y'''}\right)\Big|_{x=a}^{x=b}\end{aligned}$$

□

The functional $\int_a^b f(x, y, y', \ldots, y^{(n)})\,dx$

Proof.

and, in general,

$$\int_a^b f_{y^{(n)}} \varphi^{(n)}\,dx = (-1)^n \int_a^b \varphi \frac{d^n}{dx^n} f_{y^{(n)}}\,dx$$
$$+ \left(\varphi^{(n-1)} f_{y^{(n)}} - \varphi^{(n-2)} \frac{d}{dx} f_{y^{(n)}} + \cdots + (-1)^{n-1} \varphi \frac{d^{n-1}}{dx^{n-1}} f_{y^{(n)}} \right) \Bigg|_{x=a}^{x=b}.$$

By (58) the boundary terms vanish, and collecting results we have

$$\int_a^b \left(f_y - \frac{d}{dx} f_{y'} + \frac{d^2}{dx^2} f_{y''} - \cdots + (-1)^n \frac{d^n}{dx^n} f_{y^{(n)}} \right) \varphi\,dx = 0. \qquad (61)$$

Since this holds for any $\varphi(x) \in C_0^{(n)}(a, b)$, we can quote the fundamental lemma to complete the proof. □

Natural BCs for $\int_a^b f(x, y, y', \dots, y^{(n)})\,dx$

Let us investigate the natural boundary conditions for $F_n(y)$. Now $\varphi(x) \in C^{(n)}(a, b)$ with no boundary restrictions. The first steps of the previous discussion still apply; however, now there are the boundary terms in the expression for the first variation of $F_n(y)$ (the right side of (60)), so in obtaining the result analogous to (61) we should collect all terms including boundary terms.

Natural BCs for $\int_a^b f(x, y, y', \dots, y^{(n)})\, dx$

We rearrange the boundary terms, collecting coefficients of each $\varphi^{(i)}(x)$:

$$\begin{aligned}
&\int_a^b \left(f_y - \frac{d}{dx} f_{y'} + \frac{d^2}{dx^2} f_{y''} - \cdots + (-1)^n \frac{d^n}{dx^n} f_{y^{(n)}} \right) \varphi\, dx \\
&\quad + f_{y^{(n)}} \varphi^{(n-1)} \Big|_{x=a}^{x=b} \\
&\quad + \left(f_{y^{(n-1)}} - \frac{d}{dx} f_{y^{(n)}} \right) \varphi^{(n-2)} \Big|_{x=a}^{x=b} \\
&\quad + \left(f_{y^{(n-2)}} - \frac{d}{dx} f_{y^{(n-1)}} + \frac{d^2}{dx^2} f_{y^{(n)}} \right) \varphi^{(n-3)} \Big|_{x=a}^{x=b} \\
&\quad \vdots \\
&\quad + \left(f_{y'} - \frac{d}{dx} f_{y''} + \cdots + (-1)^{n-1} \frac{d^{n-1}}{dx^{n-1}} f_{y^{(n)}} \right) \varphi \Big|_{x=a}^{x=b} = 0.
\end{aligned} \tag{62}$$

Natural BCs for $\int_a^b f(x, y, y', \dots, y^{(n)})\, dx$

We now realize the common plan. First we consider (62) only on the subset $C_0^{(n)}(a,b)$ of all $\varphi(x) \in C^{(n)}(a,b)$. Then (62) reduces to (61), implying that (58) holds. Equation (62) becomes

$$
\begin{aligned}
& f_{y^{(n)}} \varphi^{(n-1)} \Big|_{x=a}^{x=b} \\
& \quad + \left(f_{y^{(n-1)}} - \frac{d}{dx} f_{y^{(n)}} \right) \varphi^{(n-2)} \Big|_{x=a}^{x=b} \\
& \quad + \left(f_{y^{(n-2)}} - \frac{d}{dx} f_{y^{(n-1)}} + \frac{d^2}{dx^2} f_{y^{(n)}} \right) \varphi^{(n-3)} \Big|_{x=a}^{x=b} \\
& \quad \vdots \\
& \quad + \left(f_{y'} - \frac{d}{dx} f_{y''} + \dots + (-1)^{n-1} \frac{d^{n-1}}{dx^{n-1}} f_{y^{(n)}} \right) \varphi \Big|_{x=a}^{x=b} = 0. \qquad (63)
\end{aligned}
$$

Natural BCs for $\int_a^b f(x, y, y', \dots, y^{(n)})\, dx$

It is easy to construct a set of polynomials $P_{ik}(x)$, for $k = 0, 1$ and $i = 0, \dots, n-1$, with the following properties:

$$\left.\frac{d^j P_{i0}}{dx^j}\right|_{x=a} = \delta_i^j, \qquad \left.\frac{d^j P_{i0}}{dx^j}\right|_{x=b} = 0, \qquad j = 0, 1, \dots, n-1,$$
$$\left.\frac{d^j P_{i1}}{dx^j}\right|_{x=a} = 0, \qquad \left.\frac{d^j P_{i1}}{dx^j}\right|_{x=b} = \delta_i^j, \qquad j = 0, 1, \dots, n-1,$$

where δ_i^j is the *Kronecker delta* symbol defined by $\delta_i^j = 1$ for $i = j$ and $\delta_i^j = 0$ otherwise.

Natural BCs for $\int_a^b f(x, y, y', \ldots, y^{(n)})\, dx$

Substituting these polynomials into (63), we get the natural boundary conditions for a minimizer $y(x)$:

$$f_{y^{(n)}}\Big|_{x=a} = 0,$$

$$f_{y^{(n)}}\Big|_{x=b} = 0,$$

$$\left(f_{y^{(n-1)}} - \frac{d}{dx} f_{y^{(n)}}\right)\Big|_{x=a} = 0,$$

$$\left(f_{y^{(n-1)}} - \frac{d}{dx} f_{y^{(n)}}\right)\Big|_{x=b} = 0,$$

$$\left(f_{y^{(n-2)}} - \frac{d}{dx} f_{y^{(n-1)}} + \frac{d^2}{dx^2} f_{y^{(n)}}\right)\Big|_{x=a} = 0,$$

$$\left(f_{y^{(n-2)}} - \frac{d}{dx} f_{y^{(n-1)}} + \frac{d^2}{dx^2} f_{y^{(n)}}\right)\Big|_{x=b} = 0,$$

$$\vdots$$

Natural BCs for $\int_a^b f(x, y, y', \dots, y^{(n)})\, dx$

$$\vdots$$

$$\left(f_{y'} - \frac{d}{dx} f_{y''} + \dots + (-1)^{n-1} \frac{d^{n-1}}{dx^{n-1}} f_{y^{(n)}} \right) \Big|_{x=a} = 0,$$

$$\left(f_{y'} - \frac{d}{dx} f_{y''} + \dots + (-1)^{n-1} \frac{d^{n-1}}{dx^{n-1}} f_{y^{(n)}} \right) \Big|_{x=b} = 0.$$

Note that the last two conditions contain $y^{(2n-1)}(x)$. In general, the natural boundary conditions contain higher derivatives than the equations (57).

Natural BCs for $\int_a^b f(x, y, y', \dots, y^{(n)})\, dx$

What if we appoint some of the boundary conditions (57)? For example, let $y(a) = c_1^*$ be the only boundary restriction for a minimizer. Then we need to require that $\varphi(a) = 0$, and we will get all the natural boundary conditions for $y(x)$ except the one whose expression is the multiplier of $\varphi(a)$ in the boundary sum (63). We must remove

$$\left(f_{y'} - \frac{d}{dx} f_{y''} + \cdots + (-1)^{n-1} \frac{d^{n-1}}{dx^{n-1}} f_{y^{(n)}} \right) \Big|_{x=a} = 0$$

from the list.

Example

Derive the Euler–Lagrange equation and natural boundary conditions for the energy functional whose minimizer defines the equilibrium of a bent cantilever beam described by parameters E, I. The beam is subjected to a distributed load $q(x)$, as well as a shear force Q^* and torque M^* applied to the end $x = l$:

$$E(y) = \frac{1}{2}\int_0^l EI(y'')^2\,dx - \int_0^l qy\,dx - Q^*y(l) - M^*y'(l),$$
$$y(0) = y'(0) = 0.$$

Note that the natural boundary conditions now have mechanical meaning: they account for the given torque and shear force at the “free” end $x = l$.

Solution of the Example

In this case the energy functional involves terms outside an integral, so it makes sense to repeat the derivation of the Euler–Lagrange equation for the functional $\int_a^b f(x, y, y', \dots, y^{(n)})\, dx$ to understand how M^* and Q^* enter the natural boundary conditions. Supposing y is a solution, we consider $E(y)$ on the bundle $y + t\varphi$ with arbitrary but fixed φ: that is, we consider $E(y + t\varphi)$ where $\varphi(0) = 0 = \varphi'(0)$. As a function of t this takes a minimum at $t = 0$, so its derivative at this point is zero:

$$\int_0^l EIy''\varphi''\, dx - \int_0^l q\varphi\, dx - Q^*\varphi(l) - M^*\varphi'(l) = 0.$$

Two integrations by parts in the first integral give

$$\int_0^l (EIy^{(4)} - q)\varphi\, dx + EIy''\varphi'\Big|_0^l - EIy'''\varphi\Big|_0^l - Q^*\varphi(l) - M^*\varphi'(l) = 0$$

and, because $\varphi(0) = 0 = \varphi'(0)$, we have

$$\int_0^l (EIy^{(4)} - q)\varphi\, dx + (EIy''(l) - M^*)\varphi'(l) - (EIy'''(l) + Q^*)\varphi(l) = 0.$$

Solution of the Example

Now we repeat the steps connected with the choice of φ. First we take those φ for which $\varphi(l) = 0 = \varphi'(l)$, which brings us to the equation

$$\int_0^l (EIy^{(4)} - q)\varphi \, dx = 0;$$

then, because of the arbitrariness of φ, we invoke the fundamental lemma to arrive at the Euler–Lagrange equation

$$EIy^{(4)} - q = 0 \quad \text{on } [0, l].$$

Hence for any φ that does not vanish at $x = l$ we have

$$(EIy''(l) - M^*)\varphi'(l) - (EIy'''(l) + Q^*)\varphi(l) = 0.$$

It follows that

$$EIy''(l) = M^*, \qquad EIy'''(l) = -Q^*,$$

which are the natural boundary conditions for the cantilever beam.

Functionals Depending on Functions in Many Variables

The two variable case is the simplest; extension to three or more independent variables is straightforward. Consider a functional of the form

$$F(u) = \iint_S f(x, y, u(x, y), u_x(x, y), u_y(x, y))\, dx\, dy. \tag{64}$$

Here u_x and u_y denote the partial derivatives $\partial u/\partial x$ and $\partial u/\partial y$, respectively. We confine ourselves to cases where S is simple; practical problems normally involve such domains and much complexity is thereby avoided. Let S be a closed domain in $\mathbb{R}^2$ with a piecewise smooth boundary ∂S. (We do not elaborate on the meaning of "smooth." Our attitude toward this issue is common among practitioners: we simply require everything needed in intermediate calculations.)

Boundary Conditions

We consider two main minimization problems for (64): the problem with the Dirichlet boundary condition

$$u(x,y)\Big|_{\partial S} = \psi(s), \tag{65}$$

and the problem "without" boundary conditions (i.e., the problem for which natural boundary conditions appear).
We first obtain the analogue to the Euler equation for (64). The general approach is to repeat the steps of previous proofs. Specifically we (1) introduce classes of functions over which we may consider the problem of minimum, (2) formulate the fundamental lemma for the two variable case, and (3) recall how to integrate by parts in the two variable case.

Functional Space

Let $C^{(k)}(S)$ denote the set of functions continuous on a compact domain S together with all their derivatives up to order k. The norm for defining a neighborhood of a function is

$$\|u\|_{C^{(k)}(S)} = \max_{\alpha+\beta\leq k} \max_{(x,y)\in S} \left| \frac{\partial^{\alpha+\beta} u(x,y)}{\partial x^{\alpha} \partial y^{\beta}} \right|. \tag{66}$$

$C_0^{(k)}(S)$ is the subset of $C^{(k)}(S)$ consisting of functions which, together with all their derivatives up to order $k-1$, vanish on the boundary ∂S. We shall use the corresponding notations $C^{(\infty)}(S)$ and $C_0^{(\infty)}(S)$ for sets of functions infinitely differentiable on S.

Basic Lemma

Lemma

Let $g(\mathbf{x})$ be continuous on S, and let

$$\iint_S g(\mathbf{x})\varphi(\mathbf{x})\,dx\,dy = 0 \tag{67}$$

hold for any function $\varphi(\mathbf{x}) \in C_0^{(\infty)}(S)$. Then $g(\mathbf{x}) \equiv 0$.

Proof.

We imitate the proof of Lemma 7. Suppose to the contrary that at some interior point $\mathbf{x}_0$ of S we have $g(\mathbf{x}_0) \neq 0$, say $g(\mathbf{x}_0) > 0$. Then $g(\mathbf{x}) > 0$ for all $\mathbf{x}$ in some disk C_ε having radius ε and center $\mathbf{x}_0$. It is easy to construct a bell-shaped surface of revolution centered at $\mathbf{x}_0$. The corresponding function $\varphi_0(\mathbf{x}) \in C_0^{(\infty)}(S)$ gives

$$\iint_S g(\mathbf{x})\varphi_0(\mathbf{x})\,dx\,dy = \iint_{C_\varepsilon} g(\mathbf{x})\varphi_0(\mathbf{x})\,dx\,dy > 0,$$

which contradicts (67). □

Integration by Parts

To integrate by parts we use

$$\iint_S u \frac{\partial v}{\partial x_i} \, dx \, dy = -\iint_S \frac{\partial u}{\partial x_i} v \, dx \, dy + \oint_{\partial S} uv \, n_i \, ds. \tag{68}$$

Here n_i is the cosine of the angle between the unit outward normal $\mathbf{n}$ and the unit vector along the x_i axis ($x_i = x, y$ for $i = 1, 2$, respectively). The length variable s parameterizes contour ∂S.

Euler equation

The main result now is the following. Let $f(x, y, u, p, q)$ be a continuous function having continuous first partial derivatives with respect to all of its arguments.

Theorem

Let $u = u(x, y) \in C^{(2)}(S)$ be a minimizer of the functional $\iint_S f(x, y, u, u_x, u_y)\, dx\, dy$ on the subset of $C^{(1)}(S)$ consisting of those functions satisfying (65). *Then the Euler equation*

$$f_u - \left(\frac{df_{u_x}}{dx} + \frac{df_{u_y}}{dy} \right) = 0 \tag{69}$$

holds in S. Here d/dx and d/dy are total partial derivatives, analogous to the total derivative in the one-dimensional case, when the function $u = u(x, y)$ as well as its partial derivatives u_x and u_y are considered as depending on x and y respectively.

Proof.

Consider the functional on the usual bundle $u = u(x,y) + t\varphi(x,y)$ where $\varphi(x,y)$ is a function from $C_0^{(1)}(S)$; that is, it has first derivatives continuous on S and satisfies

$$\varphi(x,y)\big|_{\partial S} = 0. \tag{70}$$

The functional $F(u+t\varphi)$ for a fixed $\varphi(x,y)$ becomes a function of the real variable t and takes its minimum at $t=0$. Thus

$$\begin{aligned}
0 &= \left.\frac{dF(u+t\varphi)}{dt}\right|_{t=0} \\
&= \frac{d}{dt}\left(\iint_S f(x,y,u+t\varphi,u_x+t\varphi_x,u_y+t\varphi_y)\,dx\,dy\right)\bigg|_{t=0} \\
&= \iint_S \left(f_u\varphi + f_{u_x}\varphi_x + f_{u_y}\varphi_y\right)dx\,dy.
\end{aligned}$$

Proof.

Integration by parts in the last two terms of the integrand gives

$$\iint_S \left[f_u - \left(\frac{df_{u_x}}{dx} + \frac{df_{u_y}}{dy} \right) \right] \varphi \, dx \, dy + \oint_{\partial S} \left(f_{u_x} n_x + f_{u_y} n_y \right) \varphi \, ds = 0. \qquad (71)$$

Remembering that $\varphi(x, y)$ satisfies (70), we get

$$\iint_S \left[f_u - \left(\frac{df_{u_x}}{dx} + \frac{df_{u_y}}{dy} \right) \right] \varphi \, dx \, dy = 0. \qquad (72)$$

Equation (69) follows from Lemma 18. □

Natural Boundary Conditions

Theorem

Let $u = u(x,y) \in C^{(2)}(S)$ *be a minimizer of the functional* $\iint_S f(x,y,u,u_x,u_y)\,dx\,dy$ *on* $C^{(1)}(S)$ *(without any boundary conditions). Then the Euler equation* (69) *holds in* S*, and* $u(x,y)$ *satisfies the natural boundary condition*

$$\left(f_{u_x} n_x + f_{u_y} n_y\right)\Big|_{\partial S} = 0. \tag{73}$$

Proof.

Consider $F(u+t\varphi)$ on the bundle $u+t\varphi$ where $\varphi(x,y) \in C^{(1)}(S)$ is arbitrary but momentarily fixed. For all such functions we establish (71) using the same reasoning as above. Restriction of $\varphi(x,y)$ to the set $C_0^{(1)}(S)$ then shows that (64) holds in S. So (72) holds whether φ belongs to $C_0^{(1)}(S)$ or $C^{(1)}(S)$. Hence

$$\oint_S \left(f_{u_x} n_x + f_{u_y} n_y\right) \varphi \, ds = 0. \tag{74}$$

Now we use the fact that on S, $\varphi = \varphi(s)$ is an arbitrary differentiable function. We do not prove the corresponding fundamental lemma for such an integral, but it is clear that a proof could be patterned after that of Lemma 7. Hence (73) follows from (74). □

Example

Demonstrate that for the functional

$$\Psi(u) = \frac{1}{2} \iint_S (u_x^2 + u_y^2)\, dx\, dy - \iint_S Fu\, dx\, dy \tag{75}$$

with $F = F(x, y)$ a given continuous function, the Euler equation and the natural boundary conditions are

$$\Delta u = -F \quad \text{in } S \tag{76}$$

and

$$\left.\frac{\partial u}{\partial n}\right|_{\partial S} = 0, \tag{77}$$

respectively. Show that on a solution u^* of the latter boundary value problem, if it exists, the functional $\Psi(u)$ attains a global minimum.

Solution

The derivation of (76) and (77) is straightforward. Denoting

$$f = \frac{1}{2}(u_x^2 + u_y^2) - Fu$$

we get

$$f_u - \left(\frac{df_{u_x}}{dx} + \frac{df_{u_y}}{dy} \right) = -F - \Delta u,$$

which leads to (76). The left-hand expression in (73) is

$$f_{u_x} n_x + f_{u_y} n_y = u_x n_x + u_y n_y,$$

which is $\partial u / \partial n$ on the boundary.

Solution 2

Before demonstrating the last statement in the example, we note that $\Psi(u)$ expresses the total energy of an elastic membrane. From physics we know that at points of minimum of a total energy functional for a mechanical system with conservative loads, the system is in equilibrium. In particular mechanics it is even shown that such an equilibrium state is stable at a point of strict minimum. Let us see what happens in this case of a spatially distributed object.
We suppose that a solution u^* of the boundary value problem (76)–(77) exists.

Solution 2

Consider the values of Ψ over the bundle $u^* + \varphi$, where φ is arbitrary:

$$
\begin{aligned}
\Psi(u^* + \varphi) = & \frac{1}{2} \iint_S \left((u_x^* + \varphi_x)^2 + (u_y^* + \varphi_y)^2 \right) dx\, dy \\
& - \iint_S F(u^* + \varphi)\, dx\, dy \\
= & \Psi(u^*) + \left[\iint_S \left(u_x^* \varphi_x + u_y^* \varphi_y \right) dx\, dy - \iint_S F \varphi\, dx\, dy \right] \\
& + \frac{1}{2} \iint_S \left(\varphi_x^2 + \varphi_y^2 \right) dx\, dy.
\end{aligned}
$$

Solution 3

Because of (76)–(77) (which, in the above theory, were derived as a direct consequence of the following equality and thus are equivalent to it when u^* is sufficiently smooth) we see that

$$\iint_S \left(u_x^* \varphi_x + u_y^* \varphi_y\right) dx\, dy - \iint_S F\varphi\, dx\, dy = 0.$$

So

$$\Psi(u^* + \varphi) - \Psi(u^*) = \frac{1}{2} \iint_S \left(\varphi_x^2 + \varphi_y^2\right) dx\, dy \geq 0,$$

which means that $\Psi(u)$ takes its global minimum at $u = u^*$. □

Thank you for your attention!!!

Further questions:

eremeyev.victor@gmail.com